

Analítica de datos y rendimiento académico

Analítica de datos y rendimiento académico

Leonardo Emiro Contreras Bravo,
Giovanny Mauricio Tarazona Bermúdez,
Héctor Javier Fuentes López





UD
Editorial

ES
ESPACIOS

© Universidad Distrital Francisco José de Caldas
© Centro de Investigaciones y Desarrollo Científico
© Leonardo Emiro Contreras Bravo,
Giovanny Mauricio Tarazona Bermúdez, Héctor Javier Fuentes López
ISBN: 978-958-787-519-5
ISBN digital: 978-958-787-520-1
ISBN ePub: 978-958-787-521-8

Primera edición, mayo de 2023

Jefe Sección de Publicaciones
Rubén Eliécer Carvajalino C.

Gestión editorial
Edwin Pardo Salazar

Corrección de estilo
Proceditor

Diagramación
Catalina Osorio Mojica

Editorial UD
Universidad Distrital Francisco José de Caldas
Carrera 24 n°. 34-37
Teléfono: (601) 3239300 ext. 6202
Correo electrónico: publicaciones@udistrital.edu.co

*Sistema de Bibliotecas de la Universidad Distrital
Francisco José de Caldas
Catalogación en la publicación (CEP)*

Contreras Bravo, Leonardo Emiro

Analítica de datos y rendimiento académico / Leonardo Emiro
Contreras Bravo, Giovanny Mauricio Tarazona Bermúdez, Héctor Javier
Fuentes López. -- Primera edición. -- Bogotá: Universidad Distrital
Francisco José de Caldas, 2019.

300 páginas : gráficas ; 24 cm. -- (Colección Espacios)

ISBN: 978-958-787-519-5 ISBN digital: 978-958-787-520-1
ISBN ePub: 978-958-787-521-8

1. Aprendizaje supervisado (Aprendizaje automático) 2. Rendimiento
académico -- Mediciones 3. Enseñanza con ayuda de computadores --
Investigación bibliográfica 4. Educación superior -- Técnicas de
predicción. I. Tarazona Bermúdez, Giovanny Mauricio, autor.
II. Fuentes López, Héctor Javier, autor III. Serie.

006.31: CDD 21 edición.

Todos los derechos reservados.

Esta obra no puede ser reproducida sin el permiso previo escrito de la
Sección de Publicaciones de la Universidad Distrital Francisco José de Caldas.
Hecho en Colombia

Contenido

Capítulo 1. Introducción	17
Capítulo 2. Objetivos de la investigación	19
2.1. Objetivo general	19
2.2. Objetivos específicos	19
Capítulo 3. Metodología	21
3.1. Fase I. Fuentes de información	22
3.2. Fase II. Planificación de la revisión de literatura	22
3.3. Fase III. Criterios de selección de la información	24
3.4. Fase IV. Análisis e interpretación de datos relevantes	24
3.5. Fase V. Resultados	24
Capítulo 4. Tecnología y analítica en educación	27
4.1. Modalidad <i>E-Learning</i>	28
4.2. Modalidad <i>B-Learning</i>	29
4.3. Panorama de la analítica en educación	31
4.4. Analítica académica	36
4.5. Analítica del aprendizaje	40
Capítulo 5. Rendimiento académico	47
5.1. Campos de investigación	49
Capítulo 6. <i>Machine Learning</i>	51
6.1. Clasificación de <i>Machine Learning</i>	52
Capítulo 7. Tecnología y analítica del aprendizaje	57
7.1. Propósito	57
7.2. Planificación de la revisión de la literatura	57
7.3. Criterios de selección de la información	58

7.4. Análisis e interpretación de datos relevantes	59
7.5. Herramientas usadas para la analítica del aprendizaje en educación superior	61
7.6. Enfoques de aplicación de la analítica del aprendizaje en educación superior	67
7.7. Construcción social de la analítica del aprendizaje en educación superior	70
7.8. Conclusiones del capítulo	73
Capítulo 8. Tecnología y analítica académica	75
8.1. Propósito	75
8.2. Planificación de la revisión de literatura	75
8.3. Criterios de selección de la información	76
8.4. Análisis e interpretación de datos relevantes	77
8.5. Herramientas usadas para la analítica académica en educación superior	79
8.6. Enfoques de aplicación de la analítica académica en educación superior	84
8.7. Construcción social de la analítica académica en educación superior	86
8.8. Conclusiones del capítulo	88
Capítulo 9. Caracterización de metodologías para la evaluación del rendimiento académico en la educación superior	91
9.1. Propósito	92
9.2. Planificación de la revisión de literatura	92
9.3. Criterios de selección de la información	92
9.4. Análisis e interpretación de datos relevantes	93
9.5. Herramientas aplicadas al rendimiento académico	95
9.5.1. Descripción de algunas herramientas usadas para las investigaciones sobre rendimiento académico	108
9.5.2. Herramientas usadas para estudiar el rendimiento académico universitario	115
9.6. Construcción social del rendimiento académico en educación superior	118
9.7. Indicadores de medición del rendimiento académico	119
9.8. Relación de variables con el rendimiento académico	122
9.9. Conclusiones del capítulo	133

Capítulo 10. Resumen de investigaciones sobre analítica académica, analítica del aprendizaje y metodologías para determinar el rendimiento académico	135
10.1. Analítica del aprendizaje, periodo 2015-2020	135
10.2. Analítica académica periodo 2015-2020	172
10.3. Metodologías para evaluación del rendimiento académico	199
10.3.1. Psicología	200
10.3.2. Estadística - Matemática	249
10.3.3. Medicina	263
10.4. Conclusiones del capítulo	264
Referencias	267
Autores	299

Índice de figuras

Figura 1.	Metodología de la investigación (diagrama de fases)	22
Figura 2.	Modalidades de formación con apoyo <i>on-line</i> en educación (TEL)	30
Figura 3.	Panorama de la analítica en educación (TEL)	33
Figura 4.	Pasos de analítica de datos educativos	34
Figura 5.	Aportes de la analítica de datos a la educación	36
Figura 6.	Ciclo de vida del <i>learning analytics</i>	42
Figura 7.	Dimensiones de la analítica del aprendizaje	43
Figura 8.	Niveles y superposiciones de las diferentes formas de análisis	44
Figura 9.	Beneficiarios y enfoques por nivel para AA y LA	45
Figura 10.	Campos que han estudiado el rendimiento académico en educación superior	49
Figura 11.	Tendencia en herramientas de análisis de datos enfocadas en la educación	52
Figura 12.	Mapa del <i>Machine Learning</i>	53
Figura 13.	Cantidad de artículos por año según el enfoque de la analítica del aprendizaje	69
Figura 14.	Panorama de métodos en la analítica del aprendizaje	70
Figura 15.	Construcción social para la analítica de investigaciones	71
Figura 16.	Cantidad de artículos por año según el enfoque de la analítica del aprendizaje	72
Figura 17.	Panorama de países que realizan investigación en la temática	72
Figura 18.	Panorama de métodos en analítica académica	86
Figura 19.	Construcción social para la analítica académica	87
Figura 20.	Cantidad de artículos por año según el enfoque de analítica académica	87
Figura 21.	Panorama de países que realizan investigación en la temática	88
Figura 22.	Áreas que han estudiado la evaluación del rendimiento académico según la investigación referencial	96

Figura 23.	Principales herramientas para la recolección de información	116
Figura 24.	Herramientas usadas para el análisis de la información en la evaluación del rendimiento académico, según la revisión bibliográfica	117
Figura 25.	Trabajos que utilizaron la herramienta “encuestas/cuestionarios/test” en sus investigaciones a lo largo de los años. Trabajos que usaron la herramienta “Análisis descriptivo” en sus investigaciones a lo largo de los años	117
Figura 26.	Investigaciones sobre evaluación del rendimiento académico por países, según la investigación referencial	119
Figura 27.	Indicadores usados para evaluar el rendimiento académico, según la revisión bibliográfica	120
Figura 28.	Trabajos que utilizaron el indicador “promedio de calificaciones” en sus investigaciones a lo largo de los años	121
Figura 29.	Trabajos que utilizaron el indicador “calificaciones finales” en sus investigaciones a lo largo de los años. Trabajos que usaron el indicador “respuesta a una encuesta” en sus investigaciones a lo largo de los años	121
Figura 30.	Uso de los indicadores en los trabajos revisados a través del tiempo	122
Figura 31.	Relación de factores en aplicación <i>E-Learning</i>	138
Figura 32.	Correlación entre tiempo de navegación y carga cognitiva	140
Figura 33.	Efectos de las preferencias de dominio y retroalimentación sobre el rendimiento (coeficientes Beta estandarizados, $p < 0,01$)	144
Figura 34.	Efectividad de los algoritmos con y sin ajuste	177
Figura 35.	Rendimiento por semana de los algoritmos	182
Figura 36.	Correlación entre la variable asistencia frente a la prueba USMLE step 1	186
Figura 37.	Correlación entre el examen final del curso frente a la prueba USMLE step 1	186
Figura 38.	Matriz de comparación de los modelos	193
Figura 39.	Rendimiento Naive Bayes en diferentes <i>softwares</i>	198

Índice de tablas

Tabla 1.	Tipos de analítica de datos	33
Tabla 2.	Potencialidades de la analítica académica	39
Tabla 3.	Referencia y base de datos del estudio para la analítica del aprendizaje	59
Tabla 4.	Relación de investigaciones para la analítica del aprendizaje	62
Tabla 5.	Área de aplicación para la analítica del aprendizaje	67
Tabla 6.	Referencia y base de datos del estudio para la analítica académica	77
Tabla 7.	Relación de investigaciones para la analítica académica	80
Tabla 8.	Área de aplicación para la analítica académica	85
Tabla 9.	Enfoque, referencia y base de datos consultada	93
Tabla 10.	Resumen de características de los trabajos realizados en cada campo del conocimiento	96
Tabla 11.	Algunas herramientas para evaluar el rendimiento académico, según trabajos realizados en cada campo del conocimiento	109
Tabla 12.	Variables que afectan el rendimiento académico	123
Tabla 13.	Parte de la encuesta realizada en el estudio	136
Tabla 14.	Matriz de correlación de los indicadores analizados	140
Tabla 15.	Iteración con materiales el curso <i>on-line</i>	141
Tabla 16.	Puntuaciones de interpretación de rayos X (presentado como porcentaje con $\pm$ desviación estándar	145
Tabla 17.	Resultados de aprobación del curso	146
Tabla 18.	Características demográficas de los estudiantes del curso	148
Tabla 19.	Análisis multivariado de actitudes de <i>E-Learning</i>	149
Tabla 20.	Matriz DOFA	150
Tabla 21.	Puntuaciones del coeficiente Alfa para IMMS	153
Tabla 22.	Pruebas de normalidad de Shapiro-Wilk y el nivel de motivación del estudiante de pregrado, maestría y doctorado	153

Tabla 23.	Definiciones operativas de variables	154
Tabla 24.	Ítems para evaluar en la experiencia	155
Tabla 25.	Factores inhibidores de <i>E-Learning</i>	158
Tabla 26.	Prueba <i>t</i> de <i>student</i> de datos CEMEDEPU. Grupo completo de profesores	159
Tabla 27.	Estadística descriptiva sobre la relación de factores	162
Tabla 28.	Prueba de los efectos totales de las variables independientes sobre la preparación para el <i>E-Learning</i>	164
Tabla 29.	Resultados de agrupación de estilos	170
Tabla 30.	Precisión del algoritmo	170
Tabla 31.	Preguntas de la entrevista	171
Tabla 32.	Rendimiento de los algoritmos con dos conjuntos de variables	183
Tabla 33.	Coeficientes estandarizados <i>B</i> y tipificados <i>Beta</i>	185
Tabla 34.	Características de análisis	188
Tabla 35.	Resultados del nuevo enfoque	190
Tabla 36.	Resultados del nuevo modelo híbrido	190
Tabla 37.	Porcentaje de precisión según el conjunto de datos	191
Tabla 38.	Porcentaje de precisión de los algoritmos en dos instancias distintas	192
Tabla 39.	Resultados de los modelos utilizados	193
Tabla 40.	Métricas de evaluación en la interacción 8:2	197

Índice de siglas

DIMSI: Grupo de Investigación de Diseño, Modelamiento y Simulación

GICOECOL: Grupo de Investigación en Comercio Electrónico

TIC: Tecnologías de la información y la comunicación

***M-Learning*:** aprendizaje mediante dispositivos móviles

***E-Learning*:** aprendizaje electrónico

***B-Learning*:** aprendizaje bimodal

LMS (*Learning Management System*): sistemas de gestión de aprendizaje

TEL (*Technology Enhanced Learning*): aprendizaje mejorado por tecnología

CMS (*Content Management System*): sistema de gestión de contenidos

LA (*Learning Analytics*): analítica del aprendizaje

AA (*Academic Analytics*): analítica académica

ML (*Machine Learning*): aprendizaje automático

MLP: perceptrón, perceptrones multicapa

DGE (*Stochastic Gradient Descent*): descenso de gradiente estocástico

RNN (*Recurrent Neural Network*): redes neuronales recurrentes

CNN (*Convolutional Neural Networks*): redes neuronales convolucionales

DBM (*Deep Boltzmann Machine*): máquina profunda de Boltzmann

LSTM (*Long Short-Term Memory*): redes de memoria a corto y largo plazo

RBFN (*Radial Basis Function Network*): redes de funciones de base radial

KNN (*K-Nearest Neighbors*): K-vecino más cercano

SVC (*Support Vector Classification*): máquina de vectores de soporte para clasificación

SOM: mapas autoorganizados de Kohonen

Capítulo 1

Introducción

Hoy en día, están presentes en la comunidad científica diferentes técnicas, metodologías y herramientas tecnológicas que se están usando para el desarrollo y la posible solución de múltiples problemas en diversos campos del conocimiento. Este hecho abre la posibilidad de un mejor análisis de la información suministrada por las instituciones, con el fin de brindar soluciones que mejoren los problemas sociales, económicos o financieros, especialmente en el campo educativo. En este sentido, en el campo educativo colombiano se ha evidenciado, en la última década, un aumento en los niveles de los indicadores de este sistema, tal es el caso de los niveles de deserción de los estudiantes en la educación superior, así como el bajo nivel de rendimiento académico de los estudiantes (Guzmán y Durán, 2009). Las tecnologías de la información y la comunicación (TIC) y las herramientas computacionales han permeado diversos campos del conocimiento, facilitando la vida de los seres humanos, así como la forma de socializar e interactuar entre ellos. Tanto así, que ingresan al campo educativo generando diversos tipos de modalidades de formación que ayudan al proceso de aprendizaje de los estudiantes que utilizan técnicas, dispositivos y equipos tecnológicos.

El sector de las TIC, así como el sector de la inteligencia artificial, se muestran actualmente como líderes en la manipulación de datos provenientes de diferentes medios, como las plataformas tecnológicas, los dispositivos electrónicos de comunicación, las redes sociales y las aplicaciones que utilizan internet. Los gigantes tecnológicos (Amazon, Apple, Alphabet, Microsoft y Facebook, entre otros) requieren emplear la información

capturada o adquirida para ser transformada y suministrar informes estadísticos descriptivos o, mejor aún, informes sobre modelos predictivos aplicables a diversas áreas del saber. En este mismo sentido, Greller y Drachsler (2012) identificaron las ventajas de las tecnologías y la analítica en el campo de la educación, entre las cuales mencionan desarrollar procesos académicos que anteriormente eran difíciles de analizar con buenos resultados. De esta manera, la analítica como subárea que abarca conceptos de áreas, como la estadística, la matemática y el aprendizaje automático, entre otras, se está volviendo importante al querer predecir variables y factores, como el desempeño académico de los estudiantes y sus futuros campos de acción al finalizar su carrera, entre otros.

Por otro lado, el rendimiento académico es un tema que ha sido ampliamente estudiado y controvertido en el campo de la investigación educativa (Lamas, 2015; Montero *et al.*, 2007). Estas investigaciones varían desde su propio concepto hasta su medición y estimación. Algunos autores consideran que puede definirse como una medida de las capacidades adquiridas en el proceso formativo, como el cumplimiento de logros en un programa cursado o el nivel de conocimiento de un área respecto a una norma, entre otros (Lamas, 2015). Otros manifiestan que lo importante es cómo medirlo. Tal es el caso de los planteamientos sobre la medición con las calificaciones escolares (Lamas, 2015), o tal vez mediante la inclusión de otras variables que pueden incidir en su medición (Hernández, 2016; Rodríguez y Ruiz, 2009). En concordancia con lo anterior, este documento busca presentar algunas de las metodologías que diversos autores han desarrollado desde diferentes áreas de estudio (psicología, estadística, medicina, *Machine Learning*) y múltiples contextos, para la medición y evaluación del rendimiento académico en la educación superior. De esta manera, se podrán identificar algunas variables útiles que lo determinan y lo afectan.

Capítulo 2

Objetivos de la investigación

Este documento es producto del proyecto de investigación financiado por el Centro de Investigaciones y Desarrollo Científico (CIDC) de la Universidad Distrital Francisco José de Caldas, titulado: *Integración de datos del ámbito educativo y herramientas de análisis de datos para determinar el rendimiento académico en la educación superior* (código: 215593-1). Convocatoria 03-2019. Conformación de un banco de proyectos de investigación, desarrollo tecnológico e innovación en la Universidad Distrital Francisco José de Caldas.

2.1. Objetivo general

Proporcionar una visión general de cómo se ha estudiado el rendimiento académico por medio de la analítica y otras herramientas desde diversas áreas del conocimiento, con el fin de establecer cómo puede usarse el (aprendizaje automático) para identificar condiciones no satisfactorias y predecir el rendimiento académico de los estudiantes en educación superior.

2.2. Objetivos específicos

- Explorar la literatura científica concerniente al campo de la analítica del aprendizaje y analítica académica en educación superior durante los últimos cinco años, para identificar aspectos relacionados con sus campos de aplicación a diversas modalidades de aprendizaje, así como identificar las técnicas, los métodos y el *software* que se ha estado utilizando.

- Caracterizar las diversas metodologías utilizadas en los últimos años para evaluar el rendimiento académico en educación superior, con el propósito de identificar las variables que lo afectan y a partir de las cuales se puedan plantear modelos acordes con las condiciones de la Universidad Distrital Francisco José de Caldas.

Capítulo 3

Metodología

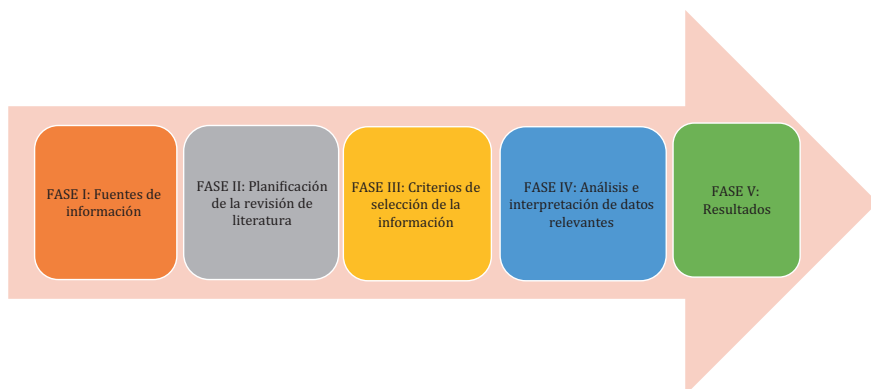
Con el fin de llevar a cabo el cumplimiento del objetivo general de esta parte de la investigación, enmarcada en el proyecto titulado *Integración de datos del ámbito educativo y herramientas de análisis de datos para determinar el rendimiento académico en la educación superior* (código 215593-1), financiado por el CIDC de la Universidad Distrital Francisco José de Caldas, se planteó la siguiente metodología. Esta se centró en dos aspectos fundamentales para la elaboración del presente documento de investigación: a) proporcionar una visión general de cómo la analítica de datos está permeando el campo educativo; y b) la búsqueda de metodologías o teorías que se han estado usando para determinar los indicadores de la educación superior, especialmente del rendimiento académico.

Para llevar a efecto la realización del estado del arte y debido a la gran cantidad de investigaciones o trabajos que están directamente relacionados con el objetivo de esta parte del proyecto, se utiliza el método de “revisiones sistemáticas de literatura” (RSL). Esta metodología tiene sus inicios en las revisiones bibliográficas elaboradas para las ciencias humanas, y también son utilizadas en el campo de la ingeniería (Kitchenham, 2004; Kitchenham y Charters, 2007; Kitchenham *et al.*, 2009).

En las cinco fases siguientes se expone la metodología empleada: 1) fuentes de información; 2) planificación de la revisión de literatura; 3) criterios de selección de la información; 4) análisis e interpretación de datos relevantes; y 5) resultados (figura 1). Cabe mencionar que para cada uno de los temas que se investigaron y que se describen en los capítulos siguientes,

se desglosa un poco más la metodología (palabras clave, fuente de datos, criterios, entre otros).

Figura 1. Metodología de la investigación (diagrama de fases)



Fuente: elaboración propia.

3.1. Fase I. Fuentes de información

Las fuentes de información registran las contribuciones sobre ciencia, tecnología o educación. Se incluyeron bases de datos académicas y motores de búsqueda académicos, en los cuales se encuentran: IEEE, Science Direct, Springer Link, Google Scholar.

3.2. Fase II. Planificación de la revisión de literatura

En esta fase, se revisan algunos aspectos básicos para tener en cuenta antes de avanzar en las fases que permitan seleccionar las investigaciones adecuadas para desarrollar una investigación referencial centrada en el tema en cuestión.

1. *Objetivos de la revisión de literatura.* Esta parte del proyecto pretende desarrollar dos objetivos clave dentro del macroproyecto al cual pertenece este trabajo de investigación: identificar la/s teoría/s disponible/s acerca del estudio de *Machine Learning* aplicadas al ámbito académico para la predicción del rendimiento académico; y establecer los últimos avances de carácter científico realizados sobre la aplicación de *Machine Learning* al tema en cuestión. Particularmente, este documento se centra en el

objetivo de proporcionar una visión general de cómo se ha estudiado el rendimiento académico por medio de la analítica y otras herramientas desde diversas áreas del conocimiento, con el fin de establecer cómo puede usarse el *Machine Learning* (aprendizaje automático) para identificar las condiciones no satisfactorias y predecir el rendimiento académico de los estudiantes en educación superior.

2. *Recursos disponibles.* El material físico sobre el tema, por lo general, no estuvo al alcance de los investigadores, ya que todos los recursos fueron electrónicos: revistas electrónicas, bases de datos científicas electrónicas y páginas de comunidades que realizan proyectos sobre el mismo tema (por lo regular, están ligadas a alguna universidad).
3. *Límites temporales.* La búsqueda de información contempló el periodo de cinco y diez años dependiendo del subtema investigado. La última revisión se hizo el 11 de diciembre de 2020.
4. *Definición del protocolo de búsqueda.* La búsqueda se llevó a cabo con palabras clave en inglés y español. Antes de mostrar los resultados en los capítulos 8, 9 y 10, se describen los criterios y palabras clave utilizadas para cada proceso de revisión referencial en cada tema.
5. *Consideraciones en la búsqueda.* Al hacer la búsqueda de información, todas las palabras clave de las ecuaciones de búsqueda se aplicaron a todas las fuentes que se consultaron; aunque, a veces, la ecuación cambió por el tipo de simbología que se usó en cada motor de búsqueda. Por ejemplo, en lugar de comillas se utilizaron paréntesis y símbolos adicionales. Para tener certeza de que la ecuación en el motor de búsqueda empleado quedó bien y la búsqueda fuera efectiva, se usaron las opciones de búsqueda avanzada.

Otro aspecto que vale la pena aclarar: aunque las variables que se están estudiando en la pregunta de investigación se centran en el rendimiento académico y no en la deserción académica, se aceptan en las revisiones sistemáticas de literatura (RSL) trabajos que estudien la deserción académica; pues estos indicadores (la deserción y el rendimiento) están íntimamente ligados. De acuerdo con la revisión de literatura inicial, que se hizo de manera exploratoria, la razón principal por la cual se busca predecir el rendimiento

académico es para evitar la deserción de un estudiante que esté en riesgo de hacerlo, ya que, a su vez, el rendimiento académico es la principal variable para determinar si un estudiante es un posible desertor.

3.3. Fase III. Criterios de selección de la información

Inicialmente, la búsqueda en las bases de datos sobre la temática principal arrojó una cantidad determinada de artículos o trabajos publicados en revistas científicas. Luego, se establecieron pasos tendientes a filtrar la información y las normas de exclusión para cada uno de los subtemas tratados en este documento (analítica académica, analítica del aprendizaje y metodologías aplicadas al rendimiento académico).

3.4. Fase IV. Análisis e interpretación de datos relevantes

A partir de los artículos ya filtrados, se procedió a examinar los tipos de variables, algoritmos, métodos y herramientas usadas para analizar la información recopilada por medio de las ecuaciones de búsqueda en las bases de datos. Por tanto, es posible que existan investigaciones no relacionadas en este documento, por cuanto no superaron los criterios establecidos o no cumplieron con las ecuaciones de búsqueda especificadas para abordar el tema.

3.5. Fase V. Resultados

Se pretende exponer, discutir y documentar resultados de la investigación y la aplicabilidad de los paradigmas de análisis de datos en el campo educativo y, especialmente, de la aplicación de estos en la determinación (predicción) del rendimiento académico. Todo ello, enmarcado en el panorama macro, que es la implementación de modelos de diversa índole que incluyen modelos contemporáneos de *Machine Learning*, para el caso de estudio “rendimiento académico en estudiantes de educación superior”.

La consecución de esta fase tiene por fin continuar con el desarrollo del proyecto *Integración de datos del ámbito educativo y herramientas de análisis de datos para determinar el rendimiento académico en la educación superior* (código 215593-1), que tiene como fin la implementación de modelos de *Machine Learning* mediante un *software* que permita tomar decisiones y acciones para

mejorar el rendimiento académico de los estudiantes de la Universidad Distrital Francisco José de Caldas, especialmente de la Facultad de Ingeniería (donde se desarrolla la investigación).

Capítulo 4

Tecnología y analítica en educación

Las TIC y las herramientas computacionales han penetrado diversos campos del conocimiento y, a su vez, han facilitado la vida de los seres humanos, así como la forma de socializar e interactuar entre ellos. Tanto así, que ingresaron progresivamente al campo educativo, generando diversos tipos de modalidades de formación que ayudan al proceso de aprendizaje de los estudiantes que utilizan técnicas, dispositivos y equipos tecnológicos. Uno de los pioneros de esta incursión es el denominado *E-Learning* (aprendizaje electrónico), que se lleva a cabo en la modalidad a distancia y permite a los estudiantes apropiarse de nuevos conocimientos mediante redes de computadores e internet (Contreras *et al.*, 2011).

Asimismo, la literatura ha sido enriquecida con los dispositivos móviles (tabletas y teléfonos) como medio de formación de los estudiantes; esta herramienta es conocida como *Mobile Learning* (*M-Learning*). De forma similar, se encuentran también los juegos como estrategia de enseñanza; es aquí donde aparece el *Game-Learning* o *G-Learning*, obtenido a partir de simulaciones gráficas que facilitan el proceso de aprendizaje de los estudiantes. Otra modalidad, no menos importante, es la denominada *B-Learning*, que utiliza las potencialidades del *E-Learning* y de la enseñanza tradicional presencial (Pineda y Tamayo, 2016). Esta última puede definirse como una forma de aprendizaje para los estudiantes, que combina lo mejor de la enseñanza por medio de actividades presenciales, con las mejoras tecnológicas y estrategias pedagógicas introducidas a la educación mediante la tecnología no presencial (*E-Learning*).

En el presente documento se abordará otro concepto interesante, la denominada “analítica académica”, la cual se diferencia de la “analítica del aprendizaje”. La analítica académica se enfoca en plantear soluciones a problemas relacionados con la academia, la gestión, la administración y los recursos financieros de las instituciones; mientras que la analítica del aprendizaje se centra en el estudio y la predicción de los aspectos relacionados con el proceso de enseñanza y aprendizaje de los estudiantes en un curso (Viberg *et al.*, 2018).

4.1. Modalidad *E-Learning*

La modalidad de aprendizaje electrónico conocida como *E-Learning* surge con la incursión de internet en diversos campos del conocimiento. Este último, al irrumpir en el campo educativo, buscaba el desarrollo del proceso de aprendizaje del estudiante de manera más autónoma y diferente a la modalidad tradicional presencial que se venía impartiendo en las instituciones. Los recursos que ofrecía para facilitar el aprendizaje autónomo eran de distintos tipos, como videos, textos, imágenes y diseños multimedia. Este tipo de recursos que ofrecía el *E-Learning* por medio de internet, según Boneu (2007), dio origen a la generación del *software* educativo. Este hecho facilitó aún más el desarrollo, la producción y el consumo de recursos para el aprendizaje en línea.

Vale destacar que hacia el 2004 apareció en escena la web social (web 2.0), que trajo consigo un apoyo a la naciente modalidad *E-Learning* por medio de los sistemas de gestión de aprendizaje (plataformas LMS), con las cuales los desarrolladores y docentes lograron generar diversos tipos de contenido didáctico apoyado en la web, que pueden ser fácilmente socializados y reutilizados por los estudiantes. Asimismo, los *Learning Management System* (LMS) o sistemas de manejo de aprendizaje, facilitaron el proceso de evaluación por medio de internet, haciendo que cada día se pudiese generar una alta cantidad de datos (información), gracias a la interacción de estudiantes y docentes sobre la misma. Esta información podía ser recopilada y almacenada, con el fin de analizarla posteriormente y, en cierta manera, “optimizar” de una forma incipiente el aprendizaje de los estudiantes.

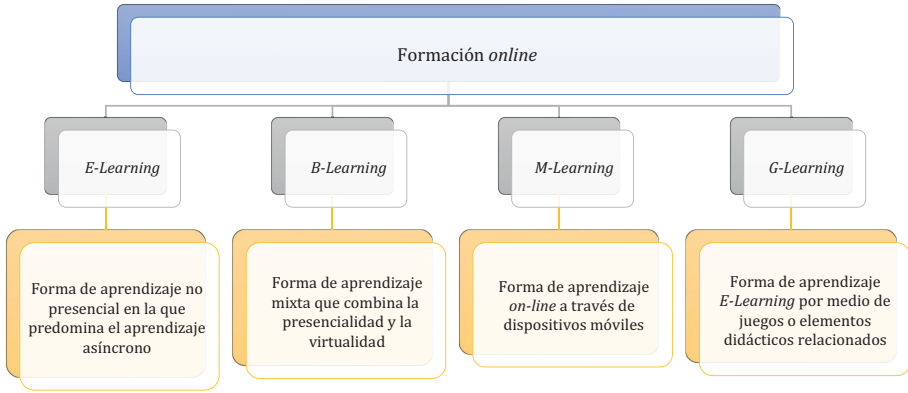
4.2. Modalidad *B-Learning*

Esta modalidad de enseñanza mediada por la tecnología, la definió Coaten (2003) como una forma de enseñar que facilita o ayuda a la modalidad tradicional mediante la combinación de actividades de tipo presencial con otras no presenciales (*on-line*). Por ende, el *Blended Learning* recibe también el nombre de “aprendizaje bimodal”, aprendizaje mezclado o mixto, ya que dicha metodología utiliza parte de la infraestructura digital disponible en la web; y otras que pueden crearse por los docentes de manera electrónica. Asimismo, usa estrategias didácticas adecuadas mediante la modalidad presencial, con el fin de despertar y fomentar la participación activa del estudiante en la clase y las actividades desarrolladas para facilitar su aprendizaje (Pascual, 2003).

Ahora bien, como sucedió con la incursión del *E-Learning* al campo educativo, el *B-Learning* también tuvo ciertos inconvenientes, como la resistencia al cambio de la enseñanza tradicional a un aprendizaje bimodal. Cambios que son necesarios, ya que implican la renovación, el progreso y el despertar de los estudiantes hacia un mundo cambiante. En este sentido, la educación requiere transformaciones que, aunque no sean radicales, merecen en sí una mejora de sus procesos y no solo de la enseñanza-aprendizaje, sino también de sus procesos administrativos y de gestión. En este aspecto, según Adell (1994), es donde el *B-Learning* puede dar su aporte, aprovechando la información disponible en internet, los recursos electrónicos y su infraestructura con fines académicos, unidos a unos métodos adecuados de participación activa del estudiante en su proceso de aprendizaje (Picciano, 2014).

La figura 2 presenta un resumen de diferentes modalidades de formación con apoyo *on-line* (Seoane, 2014), por medio del cual se puede apreciar cuáles son los cambios que ha experimentado el campo educativo con esta incursión de las TIC y que, tal vez, podrían ser considerados los antecedentes de la aplicación de la analítica en el sector educativo, por cuanto son unas fuentes generadoras de información que requieren ser analizadas y estudiadas para plantear soluciones al quehacer académico.

Figura 2. Modalidades de formación con apoyo *on-line* en educación (TEL)



Fuente: elaboración propia.

Como se mencionó, la aparición de los sistemas LMS revolucionó el sistema educativo, ya que los docentes querían conocer cómo hacerle seguimiento al aprendizaje y al rendimiento académico de sus estudiantes, así como tener claridad de su rendimiento frente a una tarea, una actividad específica, un cuestionario, entre otros recursos, que facilitaban los LMS. Los docentes se veían enfrentados a dificultades en su quehacer académico debido al carente uso de herramientas que les ayudaran a responder sus inquietudes sobre sus estudiantes. Los sistemas LMS y sus actividades pueden monitorear y generar evaluaciones numéricas sobre el progreso de una actividad de un curso; se pueden almacenar en una forma tabular estática, predefinida por los desarrolladores del sistema, pero no se pueden analizar en tiempo real para tomar las correspondientes acciones; se puede hacer análisis al culminar un curso y tomar acciones para el siguiente, pero ya no serán los mismos estudiantes que lo cursaron inicialmente (Park *et al.*, 2016). Ante estas y otras inquietudes planteadas por docentes, directivos e instituciones, es donde las herramientas de analítica de aprendizaje y analítica académica son útiles para recopilar la información, analizarla y presentarla a los interesados de forma automática y gráfica para facilitar las acciones pertinentes (Dyckhoff *et al.*, 2012). Es aquí donde aparece la analítica en educación, como medio que permite realizar un análisis más profundo de la información capturada sobre el progreso del proceso de enseñanza-aprendizaje de los estudiantes y, a su vez, hacer posible el análisis de los diversos indicadores de calidad

educativa, para que las instituciones puedan enfrentar los retos planteados por los gobiernos de diferentes países.

4.3. Panorama de la analítica en educación

Al usar el término *analítica*, se hace énfasis en establecer las relaciones entre diversas variables analizadas que, se supone, afectan a un fenómeno que se desea estudiar por medio de una variable de salida de un sistema de matrices. Las variables independientes que se analizan pueden ser de tipo biológico, tecnológico, demográfico, de gestión, financiero, entre otros (Oblinger *et al.*, 2007). Por tal razón, la analítica ha tenido cabida en muchos campos de la ciencia del conocimiento, como la biología, la ingeniería, la defensa, la salud, el comercio, el *marketing* y la educación, entre otros. Se debe tener en cuenta que cuando se habla de analítica de datos, se están utilizando diversas técnicas que permiten el análisis de un conjunto de datos, la estadística, la minería de datos y el aprendizaje automático con sus respectivas ramas afines (Burman y Som, 2019).

La interacción de docentes, estudiantes e instituciones ha generado, con el pasar de los años, alto volumen de información que requiere analizarse. Asimismo, se han desarrollado herramientas que facilitan el análisis de la información producto de diversas áreas. Los datos y las herramientas confluyen en lo que hoy se conoce como “analítica de datos” (Data y La, 2016), la cual, cuando analiza datos del área de la educación, recibe el nombre de “analítica de datos educativos” y tiene como fin recopilar, organizar, procesar y analizar la información contenida en las bases de datos y sistemas LMS usadas en las instituciones de educación para facilitar a los actores del proceso una toma de decisiones acertada, basados en la información mediante modelos explicativos o predictivos y no en suposiciones (Scholes, 2016).

Las tecnologías que incursionaron en el campo educativo dieron origen al término TEL (*Technology Enhanced Learning*) o el “aprendizaje mejorado por tecnologías”, que se utiliza para describir la aplicación de la tecnología al sector de la educación (enseñanza-aprendizaje), que permite la generación de conocimiento o solución a problemas a partir de datos del campo educativo. En este gran campo de la TEL, se encuentran dos conceptos

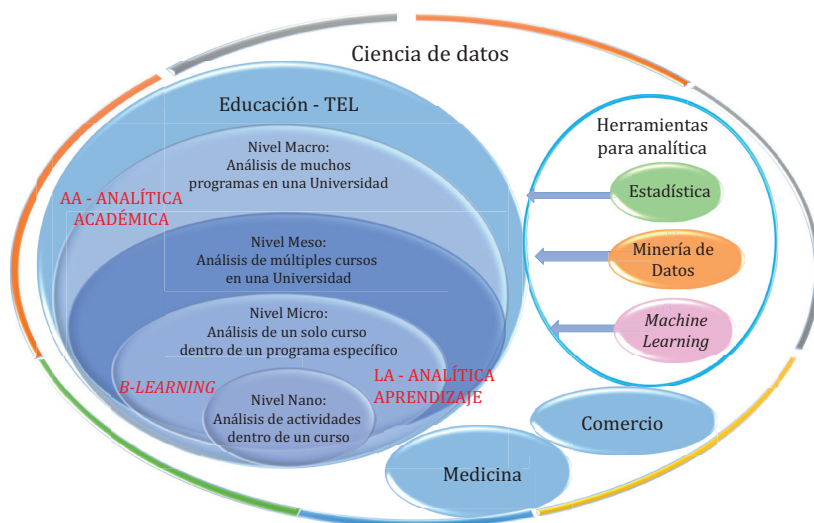
clave relacionados con la analítica de datos, que buscan generar estrategias y acciones futuras a partir de la información histórica contenida en los datos. El primero de ellos, la “analítica académica” (AA - *Academic Analytics*), abarca la escala institucional y permite predecir o estimar los indicadores que ejemplifican cómo se desarrolla la labor social de las instituciones a diversos niveles; por ejemplo, carreras, por facultades, por departamentos o por universidades (Daniel, 2015b; Ferguson, 2012). El segundo concepto, la “analítica del aprendizaje” (LA - *Learning Analytics*), abarca el proceso de enseñanza-aprendizaje y la interacción entre los diferentes actores, como los estudiantes y los profesores, en la que se busca, por medio de la recopilación y análisis de datos, comprender y optimizar el aprendizaje y los entornos en los que ocurre (Long, Siemens, Conole *et al.*, 2011). El primero, enfocado en procesos macro, y el segundo, en procesos meso del campo educativo.

Por otro lado, con el fin de identificar los problemas del campo educativo y su posible solución, aparece en la ciencia de datos una serie de herramientas, algoritmos y métodos provenientes de diferentes áreas del saber, que confluyen en una subárea conocida como “minería de datos” (*data mining*), que con la aparición de la inteligencia artificial facilitaron el desarrollo de algoritmos poderosos para analizar la información. Es aquí donde aparece en escena el concepto de “aprendizaje automático” o *Machine Learning*, que usará los datos generados por estudiantes, docentes e instituciones, convirtiéndolos en información útil que permita tomar acciones preventivas y acertadas sobre los métodos o actividades para fomentar la enseñanza y facilitar el aprendizaje (figura 3).

La analítica de datos es un concepto que ha tenido cabida en diversas áreas por capturar una serie de información analizada con el propósito de obtener elementos relevantes a partir de los datos. Por este motivo, ha incurrido en la educación, campo que se le ha facilitado en cierta medida a los investigadores. Existen diversos problemas ya identificados, que han tratado de modelarse con el pasar de los años, y aun así son fuente de investigación. Usar o implementar la analítica implica el conocimiento y manejo de una serie de herramientas matemáticas y estadísticas inicialmente, que junto con el desarrollo de las ciencias de la computación y de la inteligencia artificial

han permitido la toma de decisiones basadas en datos, con fines operativos en los diversos niveles de la educación (Contreras *et al.*, 2021).

Figura 3. Panorama de la analítica en educación (TEL)



Fuente: elaboración propia.

Según Murnion y Helfert (2013), la analítica de datos en educación ha surtido varias etapas con diferentes preguntas, cuestionamientos, herramientas de uso y tipo de información suministrada por los modelos implementados en cada una de las etapas. De esta manera, las diferentes etapas de la analítica pueden arrojar información importante sobre los datos analizados (tabla 1).

Tabla 1. Tipos de analítica de datos

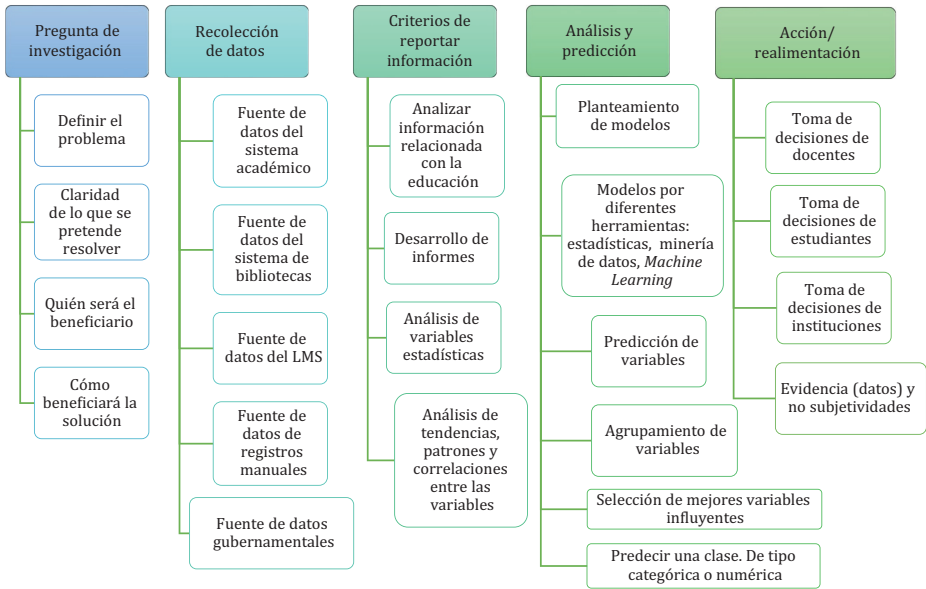
Analítica	Descriptiva: desempeño del total de las actividades.	Informes, cuadros, gráficas, estadísticas.
		Pregunta clave: ¿qué pasó?
	Diagnóstica: desempeño pasado para analizar la información.	Herramienta: <i>software</i> para análisis general, hojas de cálculo.
		Tendencias o relaciones causales entre las variables y los resultados.
		Pregunta clave: ¿por qué sucedió?
		Herramienta: estadística, matemáticas, ciencias computacionales.

Analítica	Predictiva: anticiparse a los comportamientos basándose en relaciones históricas entre las variables.	Predicciones, segmentaciones, optimizaciones, simulaciones.
		Pregunta clave: ¿qué podría pasar?
		Herramienta: minería de datos, aprendizaje automático.

Fuente: adaptada de Brownlee (2020).

Clow (2013) y Siemens (2013) plantearon una secuencia de pasos que permiten desarrollar la analítica de datos educativos a un conjunto de datos determinado. Estos se describen en la figura 4.

Figura 4. Pasos de analítica de datos educativos

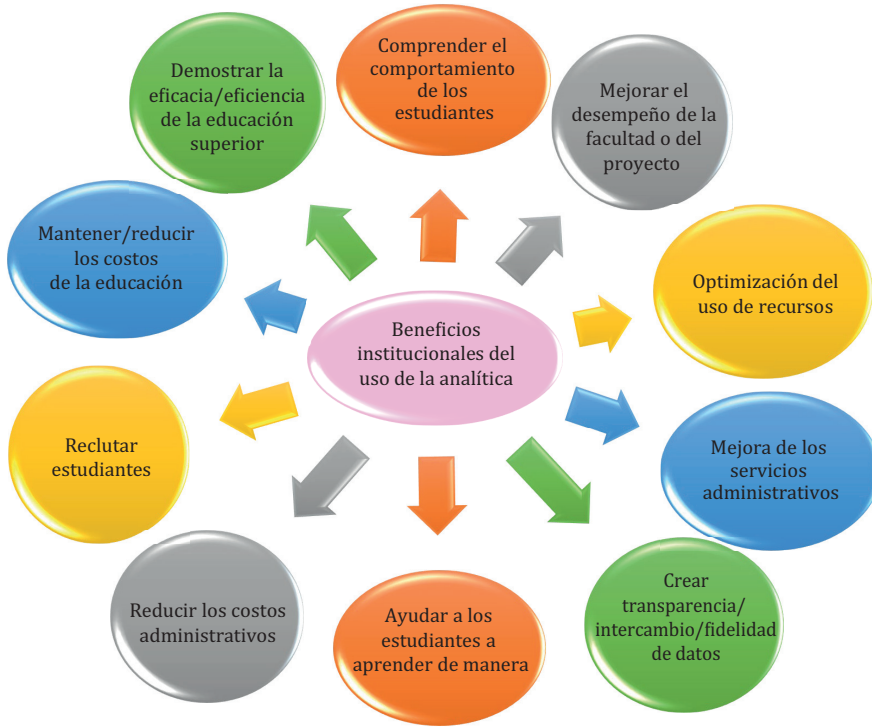


Fuente: adaptada de Clow (2013) y Siemens (2013).

El avance de las tecnologías y de la ciencia de datos ha visto en las instituciones de educación un campo por explorar. Tanto así, que las instituciones de educación superior han iniciado investigaciones en este campo, buscando identificar aspectos escondidos en los datos suministrados por las fuentes de información institucionales y gubernamentales. Este proceso se pudo llevar a cabo gracias a la analítica de datos, que abarca una serie de disciplinas que hacen su aporte, como el aprendizaje computacional, las técnicas de

visualización, los *software*, la ingeniería de características y *software* desarrollados por empresas del área de la computación. Por tanto, la analítica, en general, aporta diversos beneficios a la comunidad académica, como lo manifiestan Mahroeian *et al.* (2017), en el sentido de que esta robustece el cumplimiento de los indicadores de la calidad educativa por parte de las instituciones, permitiéndoles trabajar y reorganizar eficientemente sus recursos (figura 5). En concreto, en cuanto a Colombia, la analítica está incurriendo a paso lento en el campo educativo, a pesar de las convocatorias realizadas por el Ministerio de las TIC, que intenta que los investigadores aprendan herramientas para analizar datos y que planteen proyectos relacionados con esta temática. Esta incursión en la educación colombiana ha traído un cambio favorable en la forma de cómo visualizar la información almacenada por mucho tiempo en las instituciones. Antes era vista como un repositorio de información. Ahora son vistas las bases de datos educativas como materia prima para elaborar proyectos de desarrollo institucional y gubernamental, hecho que ha transformado la forma como se venían utilizando los datos almacenados, cuyo origen eran las iteraciones de docentes y estudiantes. Ahora, los datos no solo suministran información descriptiva, sino que también son usados para realizar predicciones de múltiples fenómenos y variables académicas, como la determinación del rendimiento académico, el progreso del aprendizaje de los estudiantes y la disminución de la tasa de deserción académica (Mahroeian *et al.*, 2017).

Figura 5. Aportes de la analítica de datos a la educación



Fuente: adaptada de Mahroeian *et al.* (2017).

4.4. Analítica académica

El término *analítica académica*, según Baepler y Murdoch (2010), fue descrito inicialmente por la compañía WebCT, con el fin de explicar la recolección de datos de sistemas CMS (*Content Management System* / sistema de gestión de contenidos). Con el pasar del tiempo, se ha ampliado esta definición, usada por las instituciones y el Gobierno como una forma de monitorear el progreso de los objetivos institucionales, propender por la disminución de los estudiantes que desertan de las instituciones, identificar grupos de estudiantes en riesgo de bajo rendimiento académico y sus cualidades, identificar estudiantes que requieren asistencia psicológica o vocacional, entre otros aspectos, cuyo fin es la optimización de los recursos físicos, humanos y académicos para mejorar los procesos educativos.

Con el pasar de los años, varios autores han aportado diversas definiciones sobre este concepto. Algunos, como Norris y Lefrere (2011), expresan que la analítica académica está centrada en los aspectos relativos a la academia; por ende, contempla diversos procesos de recopilación, de análisis y visualización de datos, por medio de los cuales se facilita medir, comparar y optimizar el desempeño de los estudiantes en un curso, un programa académico o una institución.

Por otro lado, una de las referencias más citadas en la literatura es la de Campbely Oblinger (2007), quienes afirman que la analítica académica consiste en la aplicación de herramientas y algoritmos que permiten explorar datos provenientes del campo educativo, con el fin de forjar nuevas estrategias y acciones que posibiliten predecir el progreso del estudiante, mejorar la enseñanza y las técnicas de aprendizaje, así como las variables y los indicadores relacionados para la toma de decisiones en la institución educativa (Dyckhoff *et al.*, 2012).

Tanto la analítica del aprendizaje como la analítica académica parten de la recopilación, limpieza, transformación y análisis de la información. Este último paso, inicialmente fue realizado por medio de la estadística y luego a través de un campo de las ciencias de la computación, denominado “minería de datos”, el cual posee diversos algoritmos que pueden aplicarse dependiendo la variable que se desea analizar (Iten *et al.*, 2008).

Goldstein y Katz (2005) consideran a la analítica académica similar a la “inteligencia de negocios”, ya que no solo apunta a la solución de problemas académicos, sino también a la solución de problemas del tipo de gestión y del área financiera. La analítica académica, conocida como *business intelligence*, aplicada al área educativa, permite a las instituciones y a los actores del proceso educativo, como estudiantes, docentes y recurso humano de las instituciones, hacer seguimiento al cumplimiento de los objetivos institucionales (García Tinisaray, 2015).

Al pertenecer la analítica académica al gran campo de la analítica de datos, su proceso es bastante similar. Estos pasos se pueden resumir así: recolección de la información, en cuyo caso las fuentes suelen ser diversas, pues el problema que se va a solucionar puede ser de tipo académico,

administrativo o financiero; también puede requerirse información interna a las instituciones como externas (bases de datos gubernamentales). Un segundo paso, lo constituye el análisis de la información estadística; esto es, búsqueda de tendencias de las variables de entrada, selección de variables e ingeniería de las características que puedan ayudar a los futuros modelos para que arrojen mejores resultados. Un tercer paso, es la aplicación de modelos, los cuales pueden variar si el problema es de clasificación de regresión o de *clustering*. Un cuarto paso, es la interpretación de los resultados entregados por los modelos implementados. Por último, está la etapa de acción, cuyo fin es implementar acciones a partir de la información obtenida de los modelos. Este paso de acción es llevado a cabo por docentes o instituciones o gobiernos que pueden estar interesados en la mejora del proceso educativo (García Tinisaray, 2015).

Establecer un equipo que desarrolle modelos dentro del concepto de analítica académica necesita múltiples variables, información diversa y un equipo interdisciplinario que, por un lado, recolecte, almacene y manipule la información básica de las fuentes de datos; y que posea conocimiento y entrenamiento en el manejo de las herramientas necesarias para proponer modelos predictivos; y, por otro, sus integrantes deben ser capaces de diseñar y ejecutar estrategias didácticas y académicas que permitan mejorar las experiencias de aprendizaje, los rendimientos académicos de los estudiantes y, en general, el quehacer universitario (Gómez, 2015). Un aspecto importante del equipo es el conocimiento sobre las herramientas claves para plantear análisis. Hoy en día, se utiliza el *software* o los lenguajes de programación como Weka, Matlab, Python y R., en los cuales existen diversas librerías o bibliotecas que permiten el análisis matricial de la información (datos). A partir de estos se aplican diversos algoritmos del tipo supervisados, no supervisados y de redes neuronales. Por supuesto, también incluyen bibliotecas para análisis.

El campo de la analítica académica ofrece múltiples beneficios al campo educativo, especialmente en la educación superior. Potencialidades que se han resumido atendiendo lo mencionado por Fournier *et al.* (2011) y Khalil *et al.* (2015), las cuales se resumen en la tabla 2.

Tabla 2. Potencialidades de la analítica académica

Grupo de interés	Objetivos	Ejemplo
Estudiantes	• Mejorar su rendimiento. • Personalizar el aprendizaje. • Inducir un proceso de aprendizaje autónomo y guiado de acuerdo con el perfil del alumno. • Buscar y recomendar actividades o cursos <i>on-line</i> que le faciliten el aprendizaje.	• Predicción de su estado de rendimiento futuro a partir de la información histórica. • Los estudiantes pueden comparar su rendimiento por <i>app</i> con otros estudiantes. • Los estudiantes pueden mejorar su rendimiento y aprendizaje a partir de tareas, actividades o cursos relacionados con la temática que se les dificulta.
Docentes/ investigadores	• Buscar patrones de comportamiento. • Mejorar sus métodos de enseñanza. • Realizar grupos de estudiantes dentro de los cursos o carreras de acuerdo con sus características. • Establecer modelos predictivos de los indicadores en términos numéricos o categóricos. • Tomar acciones acertadas basadas en la evidencia a tiempo. • Apoyar el aprendizaje para todos los estudiantes. • Evaluar los datos históricos del fenómeno o individuos a estudiar para mejorar la experiencia futura.	• Identificar características de estudiantes en riesgo de exclusión de la institución. • Seguimiento al proceso de enseñanza-aprendizaje de los estudiantes usando aplicaciones. • Proceso de recomendación hacia los estudiantes para evitar su deserción, bajo rendimiento. • Pueden comparar las técnicas de análisis de aprendizaje para recomendar acciones. • Recomienda un curso o libros relacionados con lo que los estudiantes anteriores consultaron. • Recomienda mejorar los métodos de enseñanza.
Instituciones educativas	• Organizar los recursos de orden institucional. • Optimizar o perfeccionar los indicadores educativos. • Apoyar los procesos de decisión.	• Mejorar la tasa de retención. • Mejorar el rendimiento académico. • Disminuir la tasa de deserción. • Tomar acciones acertadas basadas en la evidencia.

Fuente: adaptada de Fournier *et al.* (2011) y Khalil *et al.* (2015).

4.5. Analítica del aprendizaje

Hablar de analítica del aprendizaje o *Learning Analytics*, es referirse a una nueva incursión de la ciencia de datos al campo educativo, que ha tomado fuerza con el avance de grandes volúmenes de datos (*Big Data*), los cuales requieren ser analizados por algoritmos superiores a las metodologías tradicionales del campo estadístico. Estos modelos, que hoy se usan en la analítica del aprendizaje, no solo necesitan algoritmos poderosos de cálculo computacional, sino también de la suma de otros campos de las ciencias relacionadas con la computación que hacen su aporte por medio de aspectos, como el *software* con alta memoria de cálculo, componentes o bibliotecas que ayudan a una mejor visualización de la información para presentar los informes fruto de los análisis realizados y la mejora en la forma de cargue y limpieza de los formatos de los datos.

Según Stewart (2017), la analítica del aprendizaje ha tomado fuerza en el campo educativo gracias a los canales creados por la aparición de la web 2.0, internet y los sistemas gestores de aprendizaje que dieron origen también a los *software* académicos (Argonza, 2016), mediante los cuales se crearon diversas tareas, actividades y rutinas para los estudiantes. Estos han permitido la generación de datos de los que poco uso se ha hecho, desde el punto de vista de la analítica (Agudo-Peregrina *et al.*, 2014).

Banihashem *et al.* (2018) expresan que la analítica del aprendizaje (AA) es un campo joven de investigación que ha tenido impulso debido a la disponibilidad de un conjunto de datos proveniente de las iteraciones de docentes, estudiantes e instituciones, con el fin de mejorar la experiencia del aprendizaje, teniendo en cuenta que la procedencia de los datos, por lo general, es originada de las interacciones de los estudiantes con los sistemas de aprendizaje LMS durante el tiempo en que estos participan o interactúan con los contenidos presentados por los docentes o tutores.

Algunos autores han definido la analítica del aprendizaje como el aprovechamiento de un alto volumen de información generado por los educandos (minería de datos) durante su proceso de formación, con el objetivo de propender por una mejora en la calidad del aprendizaje de estos mediante estrategias de personalización en el proceso de aprendizaje (García Tinisaray,

2016). Otros añaden a esta definición, la importancia de la identificación y optimización de los ambientes en los cuales se presenta el aprendizaje de los estudiantes (Drachsler y Kalz, 2016).

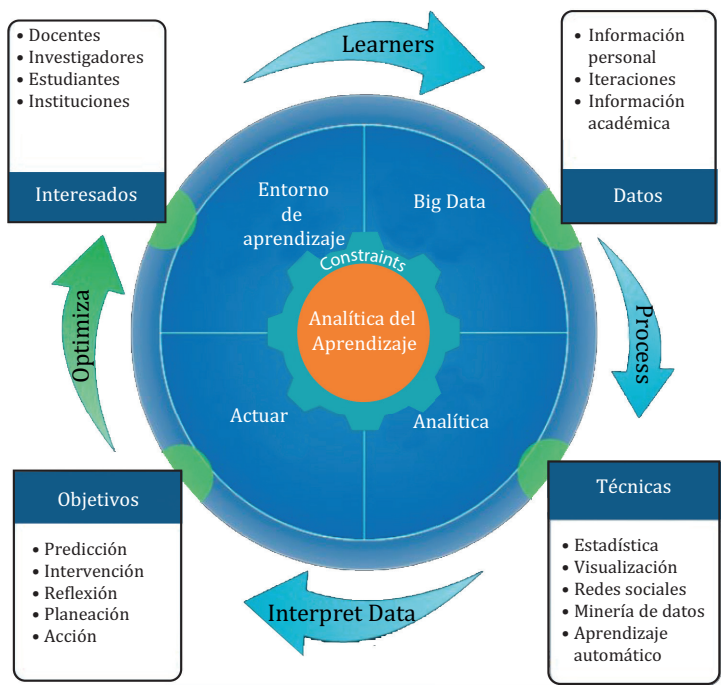
La analítica del aprendizaje considera para su análisis datos educativos históricos, por lo general provenientes de diversas fuentes que permiten estudiar el pasado, el presente y el futuro del comportamiento de los alumnos por medio de diferentes campos de acción, como la minería de datos, el aprendizaje automático, la analítica web y de la inteligencia comercial (Stewart, 2017).

Cabe mencionar que, según el *NMC Horizon Report: 2016 Higher Education Edition*, cuyo propósito es identificar y describir las tecnologías que tendrán gran impacto en la educación alrededor del mundo en los próximos cinco años, la analítica del aprendizaje es una tendencia a ser adoptada en la educación superior en un año, tal como lo hacen las compañías Amazon, Netflix y Google en el manejo del *marketing* y usuarios afiliados, en donde aquella puede ayudar a transformar la educación, pasando de un sistema de estándar de tamaño único a un sistema flexible, diseñado para satisfacer las necesidades y los intereses académicos de los estudiantes.

Partiendo del proceso analítico planteado por Elías (2011), autores como Khalil *et al.* (2015) generaron el ciclo de vida de *learning analytics* (figura 6), el cual consta de cuatro etapas principales. La primera, denominada “entorno de aprendizaje”, involucra varios sistemas LMS, como Moodle, la cual será utilizada para extraer algunos datos. La segunda fase, denominada “Big Data”, consiste en la información educativa de los diferentes grupos (interacción con LMS, datos personales, rastros o clic realizados en la LMS, entre otros) que pueden procesarse. La tercera fase de análisis, comprende diferentes métodos analíticos, como el análisis estadístico que busca descubrir patrones ocultos interesantes en el conjunto de datos educativos. En los métodos analíticos cuantitativos se encuentra: el análisis estadístico, que busca predecir con confianza lo que sucederá en el futuro para poder tomar decisiones más inteligentes, aplicando estadísticas y operaciones matemáticas. Otra opción, es el análisis cuantitativo de redes sociales, el cual permite identificar relaciones entre ellas (docentes y estudiantes). Algunas herramientas que ofrecen análisis de redes sociales enriquecidas

por visualizaciones son SNAPP y Cytoscape. La categoría de análisis cualitativo consiste en procesar datos, pero con una descripción más detallada. En este grupo, está la inteligencia emocional y el análisis cualitativo de redes sociales. La cuarta fase, el actuar, es donde se interpretan los resultados provenientes del análisis de los datos para conseguir los objetivos y optimizar el entorno de aprendizaje (predicción, recomendación, personalización, reflexión, retroalimentación).

Figura 6. Ciclo de vida del *learning analytics*



Fuente: adaptada de Khalil *et al.* (2015).

En el trabajo de Greller y Drachsler (2012) se describe la relación entre estrategias pedagógicas y actividades de aprendizaje, gracias a la representación y manifestación del conocimiento oculto tras los datos. Este modelo, que tal vez puede ser un poco más claro, considera seis dimensiones críticas de la analítica del aprendizaje (figura 7) que, por supuesto, pueden extrapolarse a la analítica académica; estas son: partes interesadas, métodos usados en los análisis, competencias, restricciones, objetivos y datos. Este modelo muestra

que al vincular los conjuntos de datos (públicos o privados) disponibles, se facilitaría el desarrollo de aplicaciones combinadas que pueden conducir a servicios más orientados al alumno y, por tanto, a una mejor personalización. Es evidente que se debe tener en cuenta una limitación importante para el desarrollo de modelos de analítica, el cual es la disponibilidad de datos, ya que muchas instituciones consideran la información personal de los estudiantes como un aspecto privado, hecho que es totalmente cierto. No obstante, para hacer analítica de datos en educación, según la investigación referencial, no es necesaria la información personal ni de estudiantes, ni de docentes, ni mucho menos del personal administrativo de las instituciones. Solo es necesaria la información de actividades del estudiante en plataformas como LMS, acceso de consultas en sistemas como el de bibliotecas o registros de aspectos académicos de forma anónima e información sociodemográfica, en el mejor de los casos.

Figura 7. Dimensiones de la analítica del aprendizaje

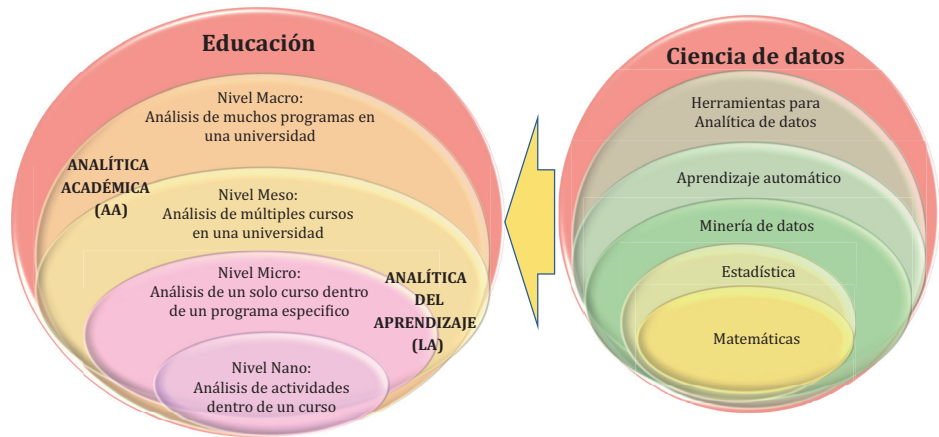


Fuente: adaptada de Greller y Drachsler (2012).

Los resultados de la analítica en el campo de la educación se pueden utilizar para elaborar informes, como el mapeo curricular; la personalización de la experiencia del estudiante (uso en tiempo real del proceso de aprendizaje de los estudiantes para mejorar su éxito); la predicción del comportamiento (como las tasas de abandono o deserción); el diseño de acciones para mejorar el aprendizaje y la determinación de las competencias en tiempo real; y no como se hace actualmente, en el que se recopilan datos de un semestre y las acciones de mejora se realizan en un tiempo posterior (Lonn *et al.*, 2015).

En resumen, la analítica académica contiene a la analítica del aprendizaje. Por tanto, al predecir el rendimiento académico como una clase (variable) de tipo numérico o categórico, se está inmerso en el campo de la analítica académica, y cuando se intenta predecir cuál será el resultado de un estudiante antes de que este curse una asignatura, se estaría incursionando en el campo de la analítica del aprendizaje (figura 8).

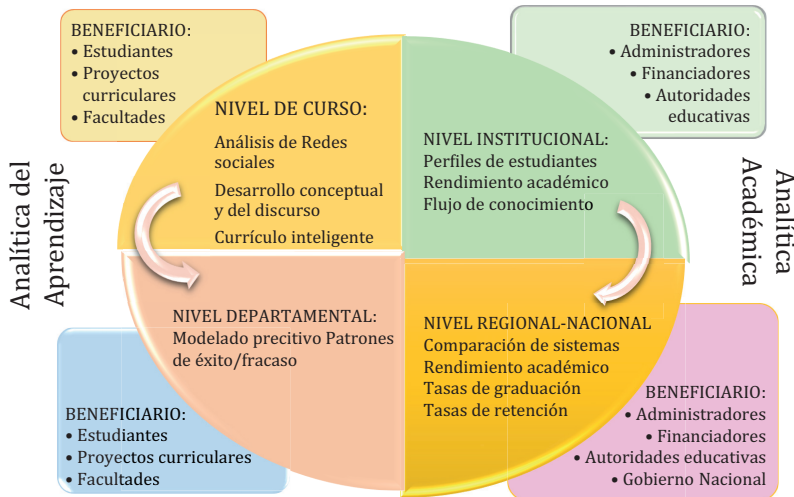
Figura 8. Niveles y superposiciones de las diferentes formas de análisis



Fuente: elaboración propia.

En cuanto a las diferencias entre los dos conceptos, se establecieron diversas definiciones, algunas de las cuales se expresaron con anterioridad y de las que se puede concluir dos enfoques. La AA, enfocada en la institución; y la LA, centrada en el estudiante, específicamente, en el proceso de aprendizaje (analiza la relación entre estudiante, docente y contenidos). Así se muestra en la figura 9.

Figura 9. Beneficiarios y enfoques por nivel para AA y LA



Fuente: adaptada de Long y Siemens (2011).

Así pues, algunas de las potencialidades de esta analítica del aprendizaje en el ámbito educativo son:

- Permite reconocer cuáles estudiantes pueden requerir intervenciones por parte del docente o de la institución, con el fin de mejorar su proceso y tratar de disminuir la deserción y mejorar la retención.
- Ofrece diversas recomendaciones de tipo académico a los estudiantes, como exhortar a leer un libro, tomar una lectura, un curso o hacer una actividad que facilite o mejore su aprendizaje.
- Evalúa los cursos ofrecidos, permitiendo evaluarlos y adaptarlos a las necesidades del entorno, brindando nuevas estrategias pedagógicas que faciliten la entrega exitosa de la información.
- A los docentes, no solo les permite tener acceso a la información de primera mano que permita identificar estudiantes en riesgo, sino que también ayuda tanto a ellos como a las instituciones a la identificación de profesores que poseen buen manejo de las herramientas y didácticas del campo educativo, así como la detección de aquellos con deficiencias en la didáctica y el uso de mejores técnicas de enseñanza.

Capítulo 5

Rendimiento académico

El rendimiento académico es un indicador que permite determinar si el estudiante, la institución y el docente tuvieron éxito o son eficientes durante el desarrollo del proceso de enseñanza-aprendizaje (García Tinisaray, 2015). A este indicador le han dado diversos nombres, como logro académico, rendimiento universitario, desempeño académico, éxito escolar, entre otros. Es considerado un indicador multidimensional, como lo establecen York *et al.* (2015), que lo definen teniendo en cuenta seis aspectos: el logro académico, la satisfacción, la adquisición de habilidades y competencias, la persistencia, el logro de objetivos de aprendizaje y el éxito profesional. De tal manera, que el análisis y el intento de una predicción temprana del desempeño de los estudiantes puede ayudar a instituciones y docentes a plantear acciones que permitan mejorar su tasa de éxito.

El rendimiento académico puede predecirse según diversos niveles, como lo expresan Alyahyan y Düştegör (2020):

- Nivel de titulación: predecir el éxito de los estudiantes en el momento de la obtención de la titulación.
- Nivel del año: predecir el éxito de los estudiantes al final del año.
- Nivel del curso: predecir el éxito de los estudiantes en un curso específico.
- Nivel de examen: predecir el éxito de los estudiantes en un examen de un curso específico.

El rendimiento académico, además, es un aspecto que posee múltiples elementos en interacción continua con el estudiante (componente importante del proceso de enseñanza y aprendizaje) y las instituciones de educación superior por medio de los docentes (Hernández, 2016). Entre dichos elementos de interacción, es posible mencionar el conocimiento de los estilos de aprendizaje para la creación de enfoques centrados en el estudiante, las emociones positivas y negativas, la transición entre sistemas educativos, la percepción académica sobre sus propias capacidades (Hernández, 2016), entre otros. En este sentido, la definición del rendimiento académico mencionada concuerda con la de Montero *et al.* (2007), la cual enuncia que el rendimiento académico es un resultado del aprendizaje que se ocasiona por el papel de los docentes, y se produce en el alumno, y es la derivación de una suma de elementos que interactúan y se agrupan en distintos factores, como institucionales, pedagógicos, psicosociales y sociodemográficos.

En un sentido más específico, las emociones, el deseo, la voluntad y otros fenómenos afectivos (conjugados con la capacidad cognoscitiva) implican una medida para la predicción del rendimiento académico (Luján *et al.*, 2019). Aunado a ello, en los factores contextuales que intervienen en el rendimiento académico, también se encuentra el apoyo social percibido, mediante el cual, un individuo se siente cuidado, valorado, querido y que pertenece a una red social en donde se comparten las responsabilidades (Fernández-Lasarte *et al.*, 2019). Incluso, en este ámbito emocional, el estrés tiene un papel importante y de bastante influencia en el rendimiento académico; pues, según algunos estudios, se ha reportado que altos niveles de estrés evidencian un rendimiento más bajo (Lin *et al.*, 2020).

El rendimiento académico también puede definirse como la expresión de las capacidades, los conocimientos y las características psicológicas que se desarrollan o se actualizan mediante el proceso formativo de enseñanza-aprendizaje, con el fin de que el estudiante alcance logros académicos, y que se evidencien al final de un semestre a través de una evaluación objetiva, que muchas veces es una nota numérica (Huamán, 2018; Montoya-Arenas *et al.*, 2021). Según Chilca (2017), las calificaciones finales de los estudiantes son un elemento clave del rendimiento académico en la imagen de una institución educativa. En adición, el rendimiento académico medido con

las calificaciones, representa un nivel de conocimientos de un área específica comparado con una norma establecida (según edad o nivel educativo) y las que, además, reflejan preferencias, dedicaciones y capacidades (Vargas *et al.*, 2019).

5.1. Campos de investigación

El rendimiento académico, al igual que otros indicadores de la calidad educativa, como el fenómeno de retención o el de graduación, se ha abordado desde diferentes áreas del conocimiento, como se muestra en la figura 10.

Figura 10. Campos que han estudiado el rendimiento académico en educación superior



Fuente: elaboración propia.

Para la primera de las áreas abordadas, se puede mencionar que la psicología es brevemente definida como una ciencia encargada de estudiar el comportamiento humano y los procesos mentales de las personas; en esta ciencia se encuentra inserto el estudio de la educación, con el fin de establecer principios, modelos, teorías, procedimientos y métodos para la enseñanza, instrucción y evaluación en el ámbito educativo (Sánchez *et al.*, 2011).

Por su parte, la estadística se ha convertido en una herramienta fundamental en múltiples aspectos cotidianos, por lo que se considera como

una ciencia transversal, incluida, además, en el ámbito educativo en el que las técnicas estadísticas funcionan de manera adecuada para la toma de decisiones; por ejemplo, en la investigación pedagógica y los procesos de investigación aplicada para el diseño de programas, comprensión de problemas de medición y evaluación, diagnóstico y orientación, entre muchos otros usos (Villanueva, 2012). Lo anterior viene de la mano con las matemáticas, pues a lo largo del tiempo, esta ciencia ha aportado a múltiples campos como la tecnología, los procesos de la naturaleza, la educación o la vida cotidiana, consolidándose como una herramienta útil en el ámbito de la investigación. Esto presenta varias de las razones por las cuales la estadística y la matemática se aplican directamente a la medición y evaluación del rendimiento académico en un contexto universitario.

En cuanto a la medicina, se reconoce que la salud es importante en la vida de los seres humanos, y que en su ausencia no es posible gozar de una buena educación, el logro de un trabajo o el desarrollo social y económico, es decir, de un progreso general para el individuo y la sociedad (Jaramillo, 2017). En este sentido, es valioso conocer cómo esta área tan amplia se relaciona con los procesos educativos llevados a cabo por los estudiantes de la educación superior.

El análisis del rendimiento es y será relevante para la mejora de la calidad educativa, razón por la cual se han dedicado numerosos estudios, tanto desde el punto de vista teórico como empírico, los cuales serán citados más adelante.

Capítulo 6

Machine Learning

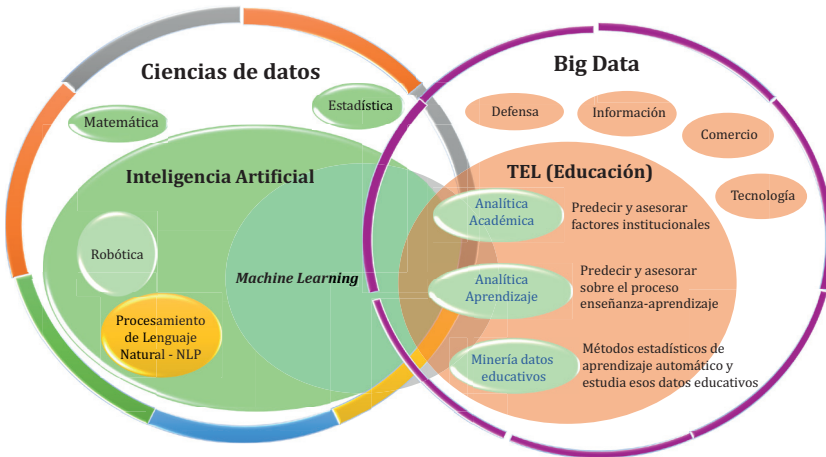
Machine Learning (ML) o aprendizaje automático, se puede definir como la ciencia que utiliza algoritmos para que puedan decir algo interesante con base en un conjunto de datos, sin tener que escribir ningún código específico para el problema. Es una subdisciplina de la inteligencia artificial que se basa en abordar y solucionar problemas a partir de disciplinas numéricas, como el razonamiento probabilístico, la investigación basada en la estadística, la recuperación de información y el reconocimiento de patrones; de esta forma, las máquinas mediante la ejecución de algoritmos hacen tareas que comúnmente realizan los humanos (Bagnato, s.f.). Otra definición bastante acertada por la comunidad científica es la de Hackeling (2014); el ML es el diseño y estudio de artefactos de programas computacionales que usan experiencia pasada para realizar decisiones futuras. Samuel (citado por Hackeling, 2014) dijo que el ML es el estudio que le da a los computadores la habilidad de aprender sin haber sido explícitamente programados.

El *Big Data* (gran volumen de datos) ha cogido mucho auge y, en parte, se debe a los métodos usados para analizar los datos en donde inicialmente los métodos eran estadísticos; posteriormente se usaron los algoritmos de *data mining* (minería de datos). Con el avance de la ciencia, de las ciencias de la computación y la aparición de la inteligencia artificial (figura 11), que consiste en buscar la manera como una máquina puede imitar o hacer acciones complejas como las que realizan los humanos, se han generado tres grandes campos: la robótica (movimientos humanos), el procesamiento por aprendizaje natural (texto y voz) y el *Machine Learning* (datos). Este último

utiliza algoritmos que predicen, previenen o actúan para mejorar el proceso académico de los estudiantes (Al-Barrak y Al-Razgan, 2016), y también para utilizarse en educación (Hackeling, 2014). De tal manera, que puedan emplearse diferentes algoritmos, como los de clasificación, las reglas de agrupación, regresión y asociación, entre otros, para predecir, por ejemplo, las calificaciones futuras de los estudiantes y, por ende, su rendimiento, partiendo de su rendimiento académico pasado y el de los estudiantes anteriores a él.

La inteligencia artificial también es considerada un área de la informática, pero tiene que ver con el desarrollo de programas que sean inteligentes o puedan hacer cosas inteligentes. Por tanto, el *Machine Learning* o aprendizaje automático es un subcampo de esta, que consiste básicamente en buscar un modelo o procedimiento que, por medio de la analítica, haga el mejor uso de los datos históricos compuestos por entradas y salidas para predecir hábilmente salidas nuevas a partir de entradas novedosas que no eran conocidas.

Figura 11. Tendencia en herramientas de análisis de datos enfocadas en la educación



Fuente: elaboración propia.

6.1. Clasificación de *Machine Learning*

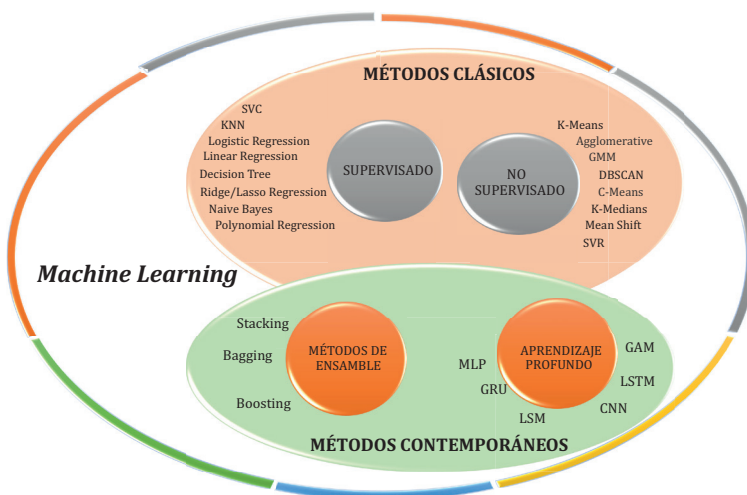
Inicialmente, se debe comprender cuáles son los posibles tipos de problemas en los que se puede utilizar el *Machine Learning* para después entender su clasificación y, luego, estar en capacidad de seleccionar los algoritmos

posibles dependiendo del problema que se va a solucionar. Los tipos de problemas en *Machine Learning* clásicos son:

- **Clasificación:** problemas en los que los datos están etiquetados, lo que significa que se le asignan una clase a la variable dependiente; por ejemplo: vive/no vive o aprueba/no aprueba. Con este tipo de tarea se busca que el modelo sea capaz de fijar a un dato nuevo (no conocido por el modelo) una etiqueta (por ejemplo: aprueba/no aprueba).
- **Regresión:** problemas en los que se trata de predecir un valor numérico. Un ejemplo de este tipo de problemas sería predecir el rendimiento académico de un estudiante, en donde la respuesta es un número entre 0 y 5 (para el caso de Colombia).
- **Agrupación:** este tipo de problemas se caracteriza por poseer datos que no están etiquetados, es decir, no tienen una variable dependiente definida. Por tanto, lo que se busca es dividir el conjunto de datos originales en grupos según la similitud y otras medidas de estructura natural en los datos.

La forma más común para empezar a comprender el campo del *Machine Learning* es por medio de la figura 12. En esta figura se explica de manera breve cuáles son los campos o subáreas.

Figura 12. Mapa del *Machine Learning*



Fuente: elaboración propia.

En el campo de aprendizaje clásico, *el Machine Learning* se encuentra dividido en aprendizaje supervisado y no supervisado. De acuerdo con Hastie *et al.* (2009), el aprendizaje supervisado consiste en predecir los valores de un conjunto de datos de salida a partir de un conjunto de datos de entrada. Se le llama “supervisado” porque, conforme al modelo, predice las salidas para datos de prueba, se calcula el error entre lo que predijo el algoritmo y el valor real. El objetivo es minimizar el error, ajustando la función de densidad de probabilidad que relaciona las entradas con las salidas. Es decir, son aquellos algoritmos donde el desarrollador actúa como un mediador para enseñar al algoritmo cuáles son las respuestas a las que debe llegar; esto lo hace con datos de entrenamiento.

En los métodos clásicos del aprendizaje automático o *Machine Learning*, aparece el llamado “aprendizaje no supervisado”; contrario a lo que ocurre en el aprendizaje supervisado, el algoritmo aprende de instancias o ejemplos que se le pasan para los cuales no se tiene una variable de respuesta conocida. Es decir, el algoritmo aprende sin ninguna respuesta de salida asociada; aprende por sí mismo, sin un humano (no se le enseña cuáles son las respuestas de salida). En esta categoría, los modelos de ML realizan, principalmente, agrupaciones o clasificaciones de los datos, según las variables de entrada y su diferenciación. Son modelos con métricas no precisas y que en ocasiones dependen de ciertas heurísticas para validar sus resultados (Hastie *et al.*, 2009).

Los métodos de ensamble o de conjunto son otra rama del *Machine Learning*, que se basan en establecer diferentes métodos que trabajan en conjunto, con el fin de reducir los errores que proporcionarían cada uno de sus algoritmos si trabajasen en solitario. Estos son métodos desarrollados por diferentes compañías dedicadas al análisis de datos, que buscan un mejor rendimiento de los algoritmos clásicos cuando trabajan en solitario. Algunos de los más usados en la literatura son: Stacking, Bagging, Boosting, Voting, Random Forest, Extra Tree, XGBoost.

- **Bagging:** este método, también conocido como *Bootstrap Aggregating*, tiene como característica usar un mismo algoritmo al que se le pasan diferentes subgrupos aleatorios de datos. Al final se promedian las respuestas.

- **Boosting:** este método se caracteriza porque cada uno de los algoritmos elegidos se entrenan secuencialmente. La entrada al primer algoritmo es un subconjunto de datos. El segundo algoritmo a usar tomará los datos que el algoritmo anterior no pudo procesar sumado a otra parte de los datos. Por tanto, lo que se busca es que cada nuevo algoritmo aprenda a corregir los errores del algoritmo anterior.
- **Voting:** básicamente consiste en tener diversos algoritmos (decididos por el analista de datos) a los que se les pasa el mismo grupo de datos. Los resultados de cada uno de los algoritmos se promedian para dar un resultado final. Según la revisión bibliográfica, este método no es tan popular, ya que los otros están dando una mayor precisión.

El aprendizaje profundo es un subcampo que a veces es llamado redes neuronales artificiales. Este subgrupo está compuesto por diferentes algoritmos, los cuales son inspirados en la estructura o función de las redes neuronales biológicas que buscan patrones (regresión y clasificación). Una red neuronal artificial, al inspirarse en una neurona biológica, se considera que tiene unos elementos que reciben señales y otros que emiten una respuesta de salida ante un estímulo. Por tanto, una red neuronal artificial es un cúmulo de neuronas organizadas en columnas (capas) que están interconectadas en cada una de sus capas y pueden transmitir información hacia adelante o hacia atrás a través de cada una de las capas.

Ahora bien, ¿por qué llamarlo “aprendizaje profundo”? La respuesta a este interrogante está en el que las redes neuronales contienen múltiples capas ocultas, de ahí el nombre de “profundo”, y el de llamar a este subgrupo “redes neuronales artificiales”. Se encuentran en este grupo: perceptrón, perceptrones multicapa (MLP), retropropagación (*back-propagation*), descenso de gradiente estocástico (*stochastic gradient descent*), redes neuronales recurrentes (RNN), redes Hopfield (*Hopfield network*), redes neuronales convolucionales (CNN), máquina profunda de Boltzmann (DBM), redes de memoria a corto y largo plazo (LSTM) y redes de funciones de base radial (*radial basis function network* —RBFN—). Las redes neuronales y los métodos de ensamble son dos grupos de métodos que están produciendo los resultados más precisos y son ampliamente utilizados en la inteligencia artificial.

Los algoritmos de *deep learning* o aprendizaje profundo son una actualización moderna de las redes neuronales artificiales básicas (perceptrón multicapa) que han explorado diferentes campos. Se caracterizan por la construcción de redes neuronales grandes y complejas aplicadas a conjuntos de grandes datos analógicos etiquetados, como imagen, texto, audio y video.

Capítulo 7

Tecnología y analítica del aprendizaje

Este capítulo se construye a partir del análisis realizado a diversas publicaciones relacionadas con la analítica del aprendizaje. El análisis incluye diversos aspectos, como la clasificación de las publicaciones por sus respectivos años de publicación, las variables propuestas y sus resultados soportados en gráficas. Se establecerán estadísticas descriptivas referentes a los métodos usados y la cantidad de artículos publicados por año.

La metodología para revisar la bibliografía del tema se puede sintetizar en los siguientes pasos: a) propósito de la búsqueda; b) planificación de la revisión de la literatura; c) establecer criterios de selección e idoneidad de la información; y d) analizar e interpretar datos relevantes.

7.1. Propósito

Se plantea la necesidad de explorar la literatura científica concerniente al campo de la analítica del aprendizaje aplicada en educación superior durante los últimos cinco años, identificando aspectos relacionados con sus campos de aplicación a diversas modalidades de aprendizaje, así como identificar las técnicas, los métodos y el *software* que se ha utilizado.

7.2. Planificación de la revisión de la literatura

Se consideró la revisión de artículos publicados en revistas y conferencias efectuadas por pares que describen sus métodos (cuantitativos, cualitativos, mixtos) y herramientas, así como el objetivo de aplicación. Se emplearon

combinaciones de palabras clave como *educational research*, *learning analytics*, *learning analytics research*, *E-Learning*, *M-Learning*, *G-Learning*.

7.3. Criterios de selección de la información

La búsqueda se hizo teniendo en cuenta los artículos científicos entre el 2015 y 2020. Inicialmente, la búsqueda en las bases de datos acerca de la temática principal “modelos de analítica del aprendizaje en educación superior” arrojó 200 investigaciones que contienen en sus palabras clave “analítica del aprendizaje”. Luego, se establecen pasos tendientes a filtrar la información y las normas de exclusión (se citan algunas):

- Se excluyen los estudios que no relacionen directamente la aplicación del paradigma de la analítica del aprendizaje.
- Se excluyen (para el estado del arte) los trabajos que no manejen una estructura o metodología secuencial propia de una investigación cuantitativa.
- Se excluyen los estudios que contienen una revisión bibliográfica de otros trabajos.
- Se excluyen los artículos que no arrojan resultados específicos del rendimiento de modelos utilizados para predecir mejoras del aprendizaje.
- Se excluyen los estudios que no incluyan el nombre o en el resumen las variables indicadas en la ecuación de búsqueda.

La recopilación se redujo aproximadamente a 90 artículos (muchos aplicados a la educación básica y media); sin embargo, dado nuestro interés en la aplicación de la analítica del aprendizaje en educación superior, solo se tuvieron en cuenta 60. Otro obstáculo que surgió en el proceso de recolección de información, fue porque algunos artículos que eran de nuestro interés no estaban disponibles en su totalidad debido a los permisos de autor, por lo cual se descartaron. En cuanto a la búsqueda de artículos referentes a la analítica del aprendizaje, se encontraron mediante el *G-Learning* (juegos en línea). Se descartó gran cantidad de artículos, puesto que estos se realizaban para estudiar el aprendizaje de estudiantes de secundaria (colegios). Esta parte del trabajo se enfocó en tratar de resolver las siguientes preguntas (Contreras *et al.*, 2021):

- ¿Qué herramientas y métodos se han utilizado para realizar la analítica del aprendizaje en educación superior?
- ¿Cuáles son las áreas de acción de la analítica del aprendizaje en la educación superior?
- ¿De qué manera se ha realizado la socialización de los trabajos efectuados en torno a la analítica del aprendizaje en la educación superior?
- ¿Hoy en día, sigue habiendo interés de la comunidad científica por este tema?

7.4. Análisis e interpretación de datos relevantes

Los 60 artículos que resultaron luego de la filtración se centraron en los siguientes campos de la analítica del aprendizaje: aplicación del aprendizaje *on-line* con plataformas LMS (*E-Learning*); aplicación del aprendizaje mediante dispositivos móviles, ML (*M-Learning*); combinación del aprendizaje *on-line* con clases presenciales (*B-Learning*), y aplicación del aprendizaje por medio de juegos (*G-Learning*).

La secuencia de pasos descrita brevemente permitió recopilar las investigaciones relacionadas con la analítica del aprendizaje. Cabe mencionar que en esta revisión pueden estar no incluidas algunas fuentes, puesto que no cumplieron con los criterios establecidos ni con las palabras clave especificadas. Los documentos consultados en esta revisión bibliográfica, discriminados para cada una de las bases de datos, se muestran en la tabla 3.

Tabla 3. Referencia y base de datos del estudio para la analítica del aprendizaje

Enfoque	Fuente	Base de datos
Analítica del aprendizaje	(Benavides y Pérez Montero, 2015)	Proquest
	(Back <i>et al.</i> , 2015)	Springer
	(Law, 2015)	Scopus
	(Giannakos <i>et al.</i> , 2015)	Scopus
	(Jo <i>et al.</i> , 2015)	Springer
	(Yen <i>et al.</i> , 2015)	Scopus

Enfoque	Fuente	Base de datos
Analítica del aprendizaje	(Martin y Whitmer, 2016)	Proquest
	(Kim <i>et al.</i> , 2016)	Scopus
	(Gibson y De Freitas, 2016)	Proquest
	(Nguyen <i>et al.</i> , 2016)	Proquest
	(Salajegheh <i>et al.</i> , 2016)	Springer
	(Alves <i>et al.</i> , 2017)	Scopus
	(Mamcenko y Kurilovas, 2017)	Springer
	(Ellis <i>et al.</i> , 2017)	Proquest
	(Luo <i>et al.</i> , 2017)	Springer
	(De Arco Paternina <i>et al.</i> , 2017)	Proquest
	(Gottipati y Shankararaman, 2018)	Springer
	(Vesin <i>et al.</i> , 2018)	Springer
	(Wang <i>et al.</i> , 2018)	Scopus
	(Wen <i>et al.</i> , 2018)	Scopus
	(Kew <i>et al.</i> , 2018)	Springer
	(Rajak <i>et al.</i> , 2018)	IEEE Xplorer
	(González-Yebra <i>et al.</i> , 2018)	Proquest
	(Eze <i>et al.</i> , 2018)	Springer
	(Núñez-Barriopedro <i>et al.</i> , 2019)	Proquest
	(Guillén-Guerrero <i>et al.</i> , 2019)	Proquest
	(Kim <i>et al.</i> , 2018)	Springer
	(Bote-Lorenzo <i>et al.</i> , 2019)	IEEE Xplorer
	(Chen <i>et al.</i> , 2019)	Springer
	(Klein <i>et al.</i> , 2019)	Springer
	(Azcona <i>et al.</i> , 2019)	Springer
	(Akçapınar <i>et al.</i> , 2019)	Springer

Enfoque	Fuente	Base de datos
Analítica del aprendizaje	(Al-araibi, Mahrin <i>et al.</i> , 2019)	IEEE Xplorer
	(Rajak <i>et al.</i> , 2018)	IEEE Xplorer
	(Al-araibi, Naz'ri bin Mahrin <i>et al.</i> , 2019)	Springer
	(Mirabolghasemi <i>et al.</i> , 2019)	Springer
	(Queiroga <i>et al.</i> , 2020)	Springer
	(Eze <i>et al.</i> , 2018)	Springer
	(Abdullah <i>et al.</i> , 2020)	IEEE Xplorer
	(Muljana y Luo, 2020)	IEEE Xplorer
	(Lerma y Silva, 2020)	Springer
	(Li <i>et al.</i> , 2020)	Springer
	(Olsen <i>et al.</i> , 2020)	Springer
	(Tahir y Wang, 2020)	IEEE Xplorer
	(Gizem <i>et al.</i> , 2020)	Springer
	(Chen, 2020)	Springer
	(Azzi <i>et al.</i> , 2020)	Springer
	(Almaiah <i>et al.</i> , 2020)	Springer
	(Teo <i>et al.</i> , 2020)	IEEE Xplorer
	(Kuliya y Usman, 2020)	Springer

Fuente: elaboración propia.

7.5. Herramientas usadas para la analítica del aprendizaje en educación superior

Como se ha mencionado al aplicar la analítica a un conjunto de datos, se utilizan diferentes métodos pertenecientes a variados campos, como el análisis estadístico, la minería de datos y el aprendizaje automático. Es importante destacar que la investigación sobre los métodos utilizados para analizar la aplicación de la analítica del aprendizaje en educación superior en los últimos años, es un insumo que interesa no solo a los expertos en

analítica, sino también a docentes de las áreas de educación e ingeniería interesados en abordar este tipo de temática para trabajos de maestría y doctorado, puesto que es un área en crecimiento en el campo educativo. De igual manera, proporciona información pertinente sobre los métodos analíticos, que orienta la tarea dispendiosa de una búsqueda bibliográfica.

La tabla 4 presenta los trabajos por años calendario del 2015 al 2020. Se pueden apreciar el(los) método(s) usado(s) en cada investigación: autor, campo de acción, método que proporcionó buen resultado y *software* o herramienta utilizada en el estudio. Los acrónimos presentados en la columna “estudio” refieren el enfoque de analítica del aprendizaje implementado en cada investigación; estos acrónimos específicamente se refieren a:

- EL: aplicación aprendizaje *on-line* con plataformas LMS (*E-Learning*).
- ML: presencial, dispositivos móviles (*M-Learning*)
- GL: por medio de juegos (*G-Learning*)
- BL: combinación del aprendizaje *on-line* con clases presenciales

Tabla 4. Relación de investigaciones para la analítica del aprendizaje

Año	Autor	Estudio	Algoritmos o método utilizado	Tamaño de la muestra	Mejor resultado	Software utilizado
2015	(Benavides y Pérez Montero, 2015)	BL	Estadística inferencial	77	Estadística inferencial	SPSS
2015	(Back <i>et al.</i> , 2015)	EL	Estadística descriptiva	48	Estadística descriptiva	Survey Monkey
2015	(Law, 2015)	EL	Logaritmo simple de PC y regresión lineal múltiple	365	Regresión lineal múltiple	No específica
2015	(Giannakos <i>et al.</i> , 2015)	ML	Serie de tiempo	11	Serie de tiempo	No específica
2015	(Jo <i>et al.</i> , 2015)	EL	Regresión múltiple	200	Regresión múltiple	No específica
2015	(Yen <i>et al.</i> , 2015)	EL	Correlación estadística	869	Correlación estadística	Line

Año	Autor	Estudio	Algoritmos o método utilizado	Tamaño de la muestra	Mejor resultado	Software utilizado
2016	(Martin y Whitmer, 2016)	EL	Correlación estadística	43	Correlación estadística	Stata
2016	(Kim <i>et al.</i> , 2016)	ML	Regresión múltiple	151	Regresión múltiple	No específica
2016	(Gibson y De Freitas, 2016)	GL	Análisis de Pareto	52 000	Análisis de Pareto	No específica
2016	(Nguyen <i>et al.</i> , 2016)	ML	Modelos lineales multivariantes	1062	Modelos lineales multivariantes	No específica
2016	(Salajegheh <i>et al.</i> , 2016)	EI	Estadística descriptiva	123	Estadística descriptiva	SPSS
2017	(Alves <i>et al.</i> , 2017)	EL	Estadística descriptiva e inferencial	2632	Estadística descriptiva e inferencial	SPSS
2017	(Mamcenko y Kurilovas, 2017)	EL	Sistema experto con implementación de árboles de decisión	6	Sistema experto con implementación de árboles de decisión	No específica
2017	(Ellis <i>et al.</i> , 2017)	EL	Regresión múltiple	145	Regresión múltiple	R FREE
2017	(Luo <i>et al.</i> , 2017)	BL	Modelo lineal mixto	119	Modelo lineal mixto	No específica
2017	(De Arco Paternina <i>et al.</i> , 2017)	BL	Análisis DOFA	132	Análisis DOFA	No aplica
2018	(Gottipati y Shankararaman, 2018)	EL	Correlación estadística	598	Correlación estadística	Visual SAP lumira
2018	(Vesin <i>et al.</i> , 2018)	EL	Diagrama de afinidad	66	Diagrama de afinidad	ProTus
2018	(Wang <i>et al.</i> , 2018)	ML	HMM (modelo oculto de Markov)	4399	HMM 0,849	Interfaz HTML basada en AJAX
2018	(Wen <i>et al.</i> , 2018)	GL	Simulación de juegos	25	Simulación de juegos	No específica

Año	Autor	Estudio	Algoritmos o método utilizado	Tamaño de la muestra	Mejor resultado	Software utilizado
2018	(Kew <i>et al.</i> , 2018)	EL	Correlación estadística y pruebas estadísticas	174	Correlación estadística y pruebas estadísticas	SPSS
2018	(Rajak <i>et al.</i> , 2018)	EL	Regresión múltiple	94	Regresión múltiple	SPSS 23.0
2018	(González-Yebra <i>et al.</i> , 2018)	BL	Modelo de Kano	40	Modelo de Kano	Second Life
2018	(Eze <i>et al.</i> , 2018)	EL	Estadística descriptiva	15	Entrevistas	NVivo
2019	(Núñez-Barriopedro <i>et al.</i> , 2019)	BL	K-means	508	K-means	SPSS
2019	(Guillén-Guerrero <i>et al.</i> , 2019)	BL	Método CEMEDEPU Estadística descriptiva	23	Método CEMEDEPU Estadística descriptiva	SPSS 22.0
2019	(Kim <i>et al.</i> , 2018)	EL	Modelo de mínimos cuadrados parciales	614	Modelo de mínimos cuadrados parciales	IBM SPSS 23
2019	(Bote-Lorenzo <i>et al.</i> , 2019)	EL	Regresión logística en modelos genéricos	1031	Regresión logística Modelo genérico	Canvas Network
2019	(Chen <i>et al.</i> , 2019)	ML	Estadística descriptiva	12	Estadística descriptiva	SPS
2019	(Klein <i>et al.</i> , 2019)	EL	Regresión lineal múltiple	27	Regresión lineal múltiple	Dedoose
2019	(Azcona <i>et al.</i> , 2019)	EL	Regresión logística, K-means, árboles de decisión, random forest	950	Random forest	No específica

Año	Autor	Estudio	Algoritmos o método utilizado	Tamaño de la muestra	Mejor resultado	Software utilizado
2019	(Akçapınar <i>et al.</i> , 2019)	ML	Algoritmo KNN	346	Algoritmo KNN	No específica
2019	(Al-araibi, Mahrin, <i>et al.</i> , 2019)	EL	Método Delphi	11	Método Delphi	No aplica
2019	m <i>et al.</i> , 2019)	EL	Análisis factorial exploratorio-confirmatorio y modelado de ecuaciones estructurales	1232	Modelado de ecuaciones estructurales	SPSS 25 y SmartPLS 3.0
2019	(Al-araibi, Naz'ri bin Mahrin <i>et al.</i> , 2019)	EL	Modelo de ecuaciones estructurales (SEM).	374	(SEM)	Amos
2019	(Mirabolghasemi <i>et al.</i> , 2019)	EL	Mínimos cuadrados parciales (PLS)	153	Mínimos cuadrados parciales (PLS)	Smart PLS
2020	(Queiroga <i>et al.</i> , 2020)	EL	Regresión logística, árbol de decisión, perceptrón, bosque aleatorio, AdaBoost	725	Algoritmo perceptrón con un AUC (área bajo la curva) de 0,8454	Python
2020	(Jivet <i>et al.</i> , 2020)	EL	MANCOVA (análisis multivalente de la varianza) con regresión múltiple	247	Regresión múltiple	No específica
2020	(Abdullah <i>et al.</i> , 2020)	GL	Correlación estadística	38	Correlación estadística	No específica
2020	(Muljana y Luo, 2020)	EL	Correlación estadística	18	Correlación estadística	R free
2020	(Lerma y Silva, 2020)	BL	Estadística descriptiva	40	Estadística descriptiva	Qda miner
2020	(Olsen <i>et al.</i> , 2020)	ML	Algoritmo Randomforest	28	Algoritmo Randomforest	Mplus

Año	Autor	Estudio	Algoritmos o método utilizado	Tamaño de la muestra	Mejor resultado	Software utilizado
2020	(Li <i>et al.</i> , 2020)	EL	J48 K-means	108	J48	Mplus 8,1
2020	(Tahir y Wang, 2020)	GL	No aplica	30	No aplica	No especifica
2020	(Gizem <i>et al.</i> , 2020)	EI	Estadística descriptiva	56	Encuestas	No especifica
2020	(Chen, 2020)	EI	Ecuaciones estructurales	241	Ecuaciones estructurales	AMOS 22.0
2020	Queiroga <i>et al.</i> , 2020)	EL	Algoritmo Fuzzy C-means (FCM)	95	Algoritmo Fuzzy C-means (FCM)	No especifica
2020	(Almaiah <i>et al.</i> , 2020)	EL	Estadística descriptiva	61	Estadística descriptiva	Nvivo
2020	(Teo <i>et al.</i> , 2020)	EL	Estadística descriptiva	18	Estadística descriptiva	No especifica
2020	(Kuliya y Usman, 2020)	EL	Correlación estadística	780	Correlación estadística	No especifica

Fuente: elaboración propia.

A partir de la información de la tabla 4 es posible afirmar que en el campo educativo y, especialmente, en la analítica del aprendizaje, se están usando herramientas clásicas para análisis de datos que tienen sus fundamentos en el área estadística (por ejemplo, correlación estadística y regresión lineal), aunque empiezan a presentarse diversos análisis que podrían considerarse en el campo de la minería de datos (por ejemplo, *K-Means* y árboles de decisión).

La industria está inundada con innumerables lenguajes de programación, con el objetivo de resolver las complejidades de los problemas y traer innovaciones tecnológicas. Pero parece no ser el común denominador cuando se pretende analizar los datos educativos relacionados con este tipo de analítica, en los que el *software* estadístico y las hojas de cálculo son las herramientas preferidas en los estudios revisados, dejando de lado las herramientas poderosas para la analítica como Matlab, Python y R, que ofrecen numerosas posibilidades y de bibliotecas para la aplicación de métodos o algoritmos en inteligencia artificial.

7.6. Enfoques de aplicación de la analítica del aprendizaje en educación superior

Las investigaciones consultadas de artículos se organizaron como se muestra en la tabla 5, teniendo en cuenta su área de investigación: aplicación aprendizaje *on-line* con plataformas LMS (*E-Learning*), dispositivos móviles (*M-Learning*), por medio de juegos (*G-Learning*) y aprendizaje mixto (*B-Learning*).

Tabla 5. Área de aplicación para la analítica del aprendizaje

<i>E-Learning</i>	<i>M-Learning</i>	<i>G-Learning</i>	<i>B-Learning</i>
(Back <i>et al.</i> , 2015)	(Giannakos <i>et al.</i> , 2015)	(Gibson y De Freitas, 2016)	(Benavides y Pérez Montero, 2015)
(Law, 2015)	(Nguyen <i>et al.</i> , 2016)	(Wen <i>et al.</i> , 2018)	(Salajegheh <i>et al.</i> , 2016)
(Jo <i>et al.</i> , 2015)	(Wang <i>et al.</i> , 2018)	(Abdullah <i>et al.</i> , 2020)	(Luo <i>et al.</i> , 2017)
(Yen <i>et al.</i> , 2015)	(Chen <i>et al.</i> , 2019)	(Tahir y Wang, 2020)	(De Arco Paternina <i>et al.</i> , 2017)
(Martin y Whitmer, 2016)	(Akçapınar <i>et al.</i> , 2019)		(González-Yebra <i>et al.</i> , 2018)
(Alves <i>et al.</i> , 2017)	(Olsen <i>et al.</i> , 2020)		(Núñez-Barriopedro <i>et al.</i> , 2019)
(Mamcenko y Kurilovas, 2017)			(Guillén-Guerrero <i>et al.</i> , 2019)
(Ellis <i>et al.</i> , 2017)			(Lerma y Silva, 2020)
(Gottipati y Shankararaman, 2018)			
(Vesin <i>et al.</i> , 2018)			
(Kew <i>et al.</i> , 2018)			
(Rajak <i>et al.</i> , 2018)			
(Eze <i>et al.</i> , 2018)			
(Kim <i>et al.</i> , 2018)			
(Bote-Lorenzo <i>et al.</i> , 2019)			

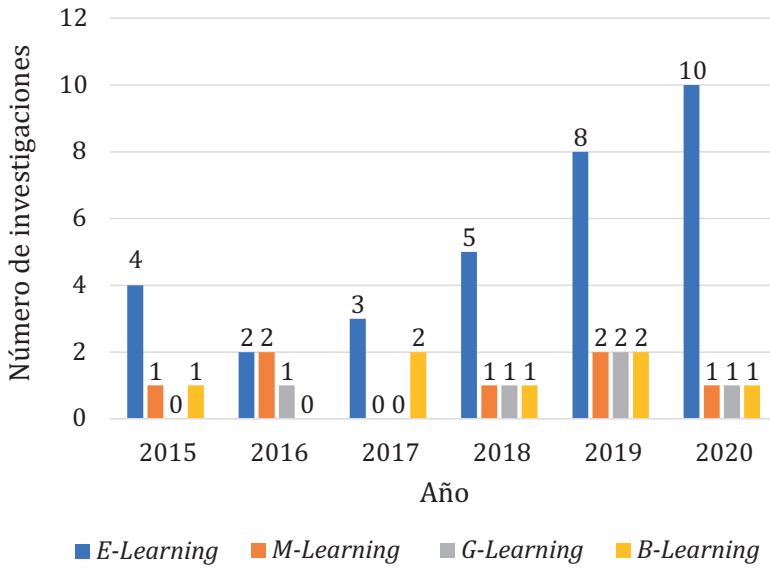
<i>E-Learning</i>	<i>M-Learning</i>	<i>G-Learning</i>	<i>B-Learning</i>
(Klein <i>et al.</i> , 2019)			
(Azcona <i>et al.</i> , 2019)			
(Al-araibi, Mahrin <i>et al.</i> , 2019)			
(Mirabolghasemi <i>et al.</i> , 2019)			
(Al-araibi, Naz'ri bin Mahrin, <i>et al.</i> , 2019)			
(Mirabolghasemi <i>et al.</i> , 2019)			
(Queiroga <i>et al.</i> , 2020)			
(Jivet <i>et al.</i> , 2020)			
(Muljana y Luo, 2020)			
(Li <i>et al.</i> , 2020)			
(Gizem <i>et al.</i> , 2020)			
(Chen, 2020)			
(Azzi <i>et al.</i> , 2020)			
(Almaiah <i>et al.</i> , 2020)			
(Teo <i>et al.</i> , 2020)			
(Kuliya y Usman, 2020)			

Fuente: elaboración propia.

La figura 13 presenta un panorama de las herramientas de análisis en la solución de problemas en este campo. Puede afirmarse que tiene fuerte tendencia a la aplicación de métodos de analítica de datos al campo educativo, especialmente al análisis del aprendizaje de estudiantes con metodología *E-Learning*. De esta manera, se corrobora que el problema es de interés actual. Cabe mencionar que en el estudio sobre el estado del arte de la analítica académica (Contreras *et al.*, 2021), los autores encontraron que esta tiene una fuerte tendencia en la aplicación hacia la búsqueda de modelos que

permitan explicar el rendimiento académico con herramientas de *Machine Learning*; un caso diferente a lo encontrado cuando se analiza la información relacionada con modelos de aprendizaje automático en educación superior, aplicados a la analítica del aprendizaje (mostrado posteriormente).

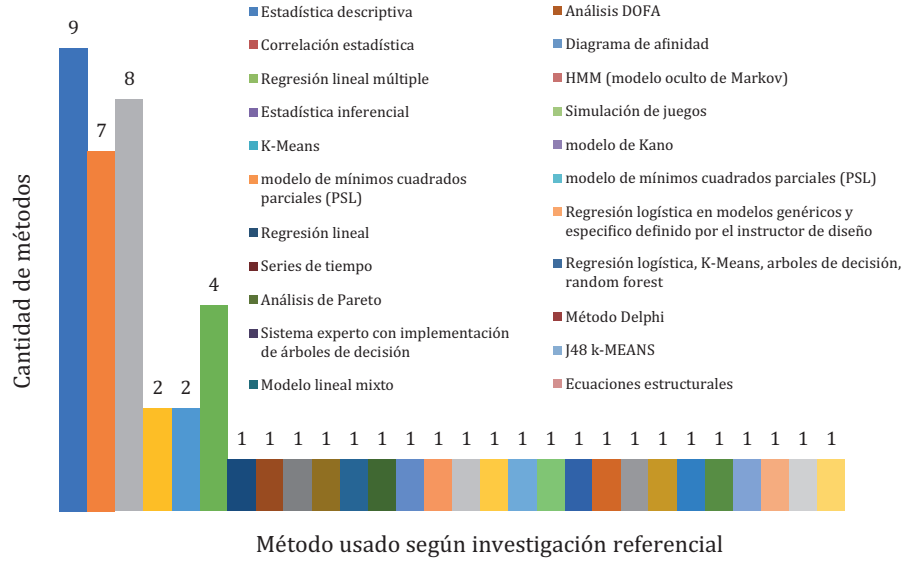
Figura 13. Cantidad de artículos por año según el enfoque de la analítica del aprendizaje



Fuente: elaboración propia.

Con el fin de identificar los métodos más usados en investigaciones relacionadas con el tema, se elaboró la figura 14, en la que se aprecia que se utilizan bastante las técnicas estadísticas, pero su enfoque es más hacia la analítica descriptiva y no la predictiva. Se muestran algunos métodos clásicos de predicción, como series de tiempo y regresión lineal. Los autores no están usando métodos que dan buenos resultados en el campo de la analítica académica para la predicción de indicadores; por ejemplo, el rendimiento académico, la deserción y la tasa de graduación como los algoritmos de aprendizaje supervisado clásicos y árboles de decisión, KNN, SVM (Adekitan y Noma-Osaghae, 2019; Contreras *et al.*, 2020; Tomasevic *et al.*, 2020; Wiyono y Tegal, 2019); o algoritmos más avanzados en el aprendizaje automático, como los métodos híbridos o de ensamble y las redes neuronales (Hasan *et al.*, 2020; Medina *et al.*, 2020; Zacharias y Athanasios, 2020).

Figura 14. Panorama de métodos en la analítica del aprendizaje



Fuente: elaboración propia.

7.7. Construcción social de la analítica del aprendizaje en educación superior

La socialización del conocimiento es un aspecto importante cuando se realizan investigaciones de alto impacto. En los últimos años, se nota una creciente tendencia de las investigaciones en aspectos relacionados con la analítica del aprendizaje, ya que es un área que está influenciando y permeando diversos campos como el educativo para tratar de solucionar los inconvenientes presentes (Lasarte *et al.*, 2020). Los resultados de las investigaciones son socializados por medio de revistas indexadas, revisadas por pares y en una menor proporción en congresos o simposios (figura 15). Esta fuerte tendencia a trabajar estos temas puede deberse, como lo manifiesta Oblinger *et al.* (2007), a la aparición de técnicas que facilitan el análisis de datos académicos que buscan no solo estudiar la influencia de factores tecnológicos y de gestión en las instituciones, sino también la forma cómo aprenden y cómo se puede mejorar el proceso de enseñanza-aprendizaje en las instituciones.



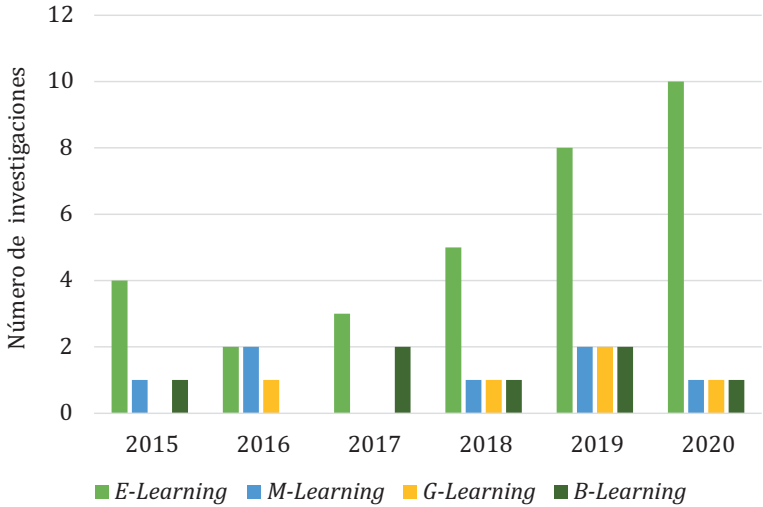
Fuente: elaboración propia.

La cantidad de publicaciones sobre analítica se redujo durante el periodo 2015-2018. Pero parece que es un campo en crecimiento en cantidad de estudios en los últimos años (reflejado en un aumento hacia el 2020), debido al auge de la minería de datos y el aprendizaje automático según el análisis referencial, así como al desarrollo tecnológico proveniente de la influencia del *big data* en diversos campos del conocimiento. Asimismo, en la figura 16 se observa el número de documentos reportados por año en el periodo de búsqueda, agrupados por campos de aplicación de la analítica del aprendizaje. A partir de los resultados, puede afirmarse que tiene fuerte tendencia a la aplicación del método de analítica de datos aplicados al campo educativo, especialmente al análisis del aprendizaje de estudiantes con metodología *E-Learning*, gracias al auge de las plataformas LMS y la web 2.0, según la información analizada en las investigaciones consultadas.

En la figura 17 se presentan las regiones o países que trabajan esta temática con mayor regularidad. Corea del Sur es el país con el que mayor frecuencia se comparte la autoría, seguido de China. De los 60 documentos obtenidos, el 55 % se desarrolla en Asia, seguido de Oceanía y Europa. Este hecho permite suponer que, en las universidades latinoamericanas, una gran cantidad de *software* o plataformas almacenan muchos datos provenientes

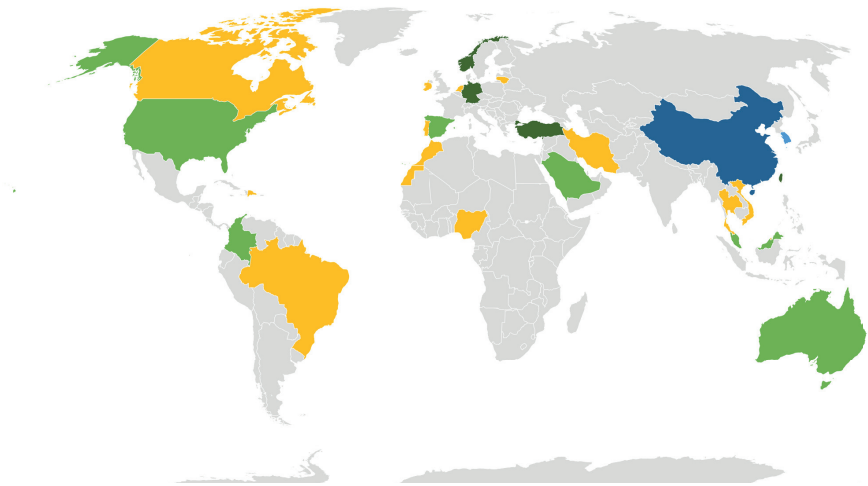
del estudiante, o almacenan datos generados durante su estadía en la universidad y estos no reciben ningún tratamiento en la búsqueda del mejoramiento de su proceso de aprendizaje. De tal manera, es necesaria la reutilización de la información, dando lugar a aplicaciones mucho más ricas que aporten mayor conocimiento o que complementen la información ya existente.

Figura 16. Cantidad de artículos por año según el enfoque de la analítica del aprendizaje



Fuente: elaboración propia.

Figura 17. Panorama de países que realizan investigación en la temática



Fuente: elaboración propia.

7.8. Conclusiones del capítulo

Con base en la información consultada y analizada, es posible plantear las siguientes conclusiones principales:

La analítica del aprendizaje es un campo que apenas está incursionando en la educación superior con algoritmos de minería de datos y *Machine Learning*. Este busca aprovechar el alto volumen de información generado por los estudiantes en sus actividades académicas, a la cual poco uso se le da, más allá de realizar estadísticas descriptivas y, en el mejor de los casos, una analítica diagnóstica (desempeño pasado para analizar la información).

Los resultados muestran que, en la educación superior, la analítica del aprendizaje se está encaminando hacia la solución o el desempeño de los estudiantes que utilizan la modalidad virtual (*E-Learning*). Este tipo de técnica puede enfocarse también hacia la aplicación en otro tipo de modalidad, pues lo que se requiere es la información que muchos dispositivos, plataformas y bases de datos de las instituciones tienen almacenada o capturan en el quehacer académico.

A pesar de que los algoritmos de son ampliamente usados en el campo del comercio y las comunicaciones (Amazon, Netflix y Google) para el manejo del *marketing* y usuarios afiliados, entre otros (Johnson *et al.*, 2015), no son usados para la búsqueda de problemas relacionados con el proceso enseñanza-aprendizaje para ayudar a transformar la educación, pasando de un sistema de estándar de tamaño único a un sistema flexible, diseñado para satisfacer necesidades e intereses académicos de los estudiantes.

El estudio del proceso de aprendizaje, según la revisión, no solo abarca los datos numéricos que se obtienen de las plataformas, analizados por medio de herramientas con el fin de ofrecer información relevante y útil para facilitar la toma de decisiones y acciones oportunas, sino que también es necesario tener en cuenta diferentes enfoques desde otras áreas como la psicología y la pedagogía. Por ende, no existe un camino definido en el estudio del proceso de aprendizaje de los estudiantes. Es un camino que conduce tanto a valores positivos como negativos (axiología).

Capítulo 8

Tecnología y analítica académica

Este capítulo se construye a partir del análisis realizado a diversas publicaciones relacionadas con la analítica académica. El análisis incluye diferentes aspectos, como clasificación de las publicaciones por sus respectivos años de publicación, variables propuestas y sus correspondientes resultados soportados en gráficas. Se establecerán estadísticas descriptivas referentes a los métodos usados, y cantidad de artículos publicados por año.

La metodología para hacer la revisión bibliográfica del tema se puede sintetizar en los siguientes pasos: a) propósito de la búsqueda; b) planificación de la revisión de la literatura; c) establecer criterios de selección e idoneidad de la información; y d) analizar e interpretar datos relevantes.

8.1. Propósito

Se plantea la necesidad de explorar la literatura científica concerniente al campo de la analítica académica aplicada en educación superior durante los últimos cinco años, identificando aspectos relacionados con sus campos de aplicación a diversas modalidades de aprendizaje, así como identificar las técnicas, los métodos y el *software* que se ha utilizado.

8.2. Planificación de la revisión de literatura

Se consideró la revisión de artículos publicados en revistas y conferencias revisadas por pares que describen sus métodos (cuantitativos, cualitativos, mixtos) y herramientas, así como el objetivo de aplicación utilizando combinaciones de palabras clave como: *educational research*, *academic analytics*,

academic analytics research, academic performance prediction, prediction of student dropout, academic graduation.

8.3. Criterios de selección de la información

El periodo de búsqueda de información corresponde a los últimos cinco años (2015-2020). Uno de los principales obstáculos presentados en su recopilación es el de encontrar datos referentes a casos de estudio de aplicación de métodos relacionados con la analítica académica en educación superior (105 artículos). Solo se tuvieron en cuenta 50 artículos para el análisis, pues se establecieron criterios tendientes a filtrar información y normas de exclusión (se citan algunas):

- Se excluyen los estudios que no relacionen directamente la aplicación del paradigma de la analítica académica.
- Se excluyen (para el estado del arte) los trabajos que no manejen una estructura o metodología secuencial propia de una investigación cuantitativa.
- Se excluyen los estudios que contienen una revisión bibliográfica de otros trabajos.
- Se excluyen los artículos que no arrojan resultados específicos del rendimiento de modelos utilizados para predecir aspectos relacionados con la analítica académica.
- Se excluyen los trabajos de aplicación de analítica académica a educación escolar, primaria o secundaria.
- Se excluyen los estudios que no incluyan el nombre o en el resumen las variables indicadas en la ecuación de búsqueda.

Esta parte del trabajo se enfoca en tratar de resolver las siguientes preguntas (Contreras *et al.*, 2021):

- ¿Qué herramientas y métodos se han utilizado para realizar la analítica académica en la educación superior?
- ¿Cuáles son los campos de acción de la analítica académica en la educación superior?

- ¿De qué manera se ha realizado la socialización de los trabajos elaborados en torno a la analítica académica en la educación superior?
- ¿Se continúa trabajando esta temática en los últimos años?

8.4. Análisis e interpretación de datos relevantes

La secuencia de pasos, descrita brevemente, permitió recopilar las investigaciones relacionadas con la analítica académica. Cabe mencionar que en esta revisión pueden estar no incluidas algunas fuentes, porque no cumplieron con los criterios establecidos y las palabras clave especificadas. Los documentos consultados en esta revisión bibliográfica discriminados para cada una de las bases de datos se muestran en la tabla 6.

Tabla 6. Referencia y base de datos del estudio para la analítica académica

Enfoque	Fuente	Base de datos
Analítica académica	(Echegaray-Calderón y Barrios-Aranibar, 2015)	IEEE Xplorer
	(Pradeep <i>et al.</i> , 2015)	IEEE Xplorer
	(Hernández, Meléndez <i>et al.</i> , 2016)	IEEE Xplorer
	(Dixson <i>et al.</i> , 2016)	Scopus
	(Ahamed <i>et al.</i> , 2016)	IEEE Xplorer
	(Hasbun <i>et al.</i> , 2016)	IEEE Xplorer
	(Devasia <i>et al.</i> , 2016)	IEEE Xplorer
	(Merchán-Rubiano <i>et al.</i> , 2017)	Proquest
	(Razaque <i>et al.</i> , 2017)	IEEE Xplorer
	(Qin <i>et al.</i> , 2017)	Scopus
	(Costa <i>et al.</i> , 2017)	IEEE Xplorer
	(Stanko <i>et al.</i> , 2017)	IEEE Xplorer
	(Khalil <i>et al.</i> , 2018)	Springer
	(Wheeler <i>et al.</i> , 2018)	Scopus
	(Liu <i>et al.</i> , 2018)	IEEE Xplorer
	(Limsathitwong <i>et al.</i> , 2018)	IEEE Xplorer

Enfoque	Fuente	Base de datos
Analítica académica	(Lima <i>et al.</i> , 2018)	Proquest
	(Hirokawa, 2018)	Springer
	(Wati <i>et al.</i> , 2018)	Springer
	(Miguéis <i>et al.</i> , 2018)	Springer
	(Burgos <i>et al.</i> , 2018)	Springer
	(Kaunang y Rotikan, 2018)	IEEE Xplorer
	(David y Anastasija, 2019)	Scopus
	(Ajibade, Ahmad <i>et al.</i> , 2019)	IEEE Xplorer
	(Nabizadeh <i>et al.</i> , 2019)	Springer
	(Martín, 2019)	Scopus
	(Kauffman <i>et al.</i> , 2019)	Scopus
	(Kwankajonwong <i>et al.</i> , 2019)	Proquest
	(Herodotou <i>et al.</i> , 2019)	Springer
	(Francis y Babu, 2019)	Springer
	(Vora y Rajamani, 2019)	Springer
	(Salal <i>et al.</i> , 2019)	Springer
	(Amazona y Hernández, 2019)	IEEE Xplorer
	(Qowidho <i>et al.</i> , 2019)	IEEE Xplorer
	(Ajibade, Bahiah Binti Ahmad <i>et al.</i> , 2019) Educational Data Mining (EDM)	IEEE Xplorer
	(Mohammadi <i>et al.</i> , 2019)	Scopus
	(Wiyono <i>et al.</i> , 2019)	IEEE Xplorer
	(Adekitan y Noma-Osaghae, 2019)	IEEE Xplorer
	(Nieto <i>et al.</i> , 2019)	Scopus
	(Francis y Babu, 2019)	Springer
	(Hussain <i>et al.</i> , 2019)	Scopus
	(Lasarte <i>et al.</i> , 2020)	Springer
	(Tan <i>et al.</i> , 2020)	Scopus

Enfoque	Fuente	Base de datos
Analítica académica	(Tenpipat y Akkarajitsakul, 2020)	IEEE Xplorer
	(Medina <i>et al.</i> , 2020)	IEEE Xplorer
	(Mavani, <i>et al.</i> , 2020)	IEEE Xplorer
	(Mengash, 2020)	IEEE Xplorer
	(Tomasevic <i>et al.</i> , 2020)	Scopus

Fuente: elaboración propia.

8.5. Herramientas usadas para la analítica académica en educación superior

Como se mencionó al aplicar la analítica a un conjunto de datos, se usan diferentes métodos pertenecientes a variados campos, como el análisis estadístico, la minería de datos y el aprendizaje automático. Es importante destacar que la investigación sobre los métodos que se han utilizado para analizar la aplicación de la analítica del aprendizaje en educación superior en los últimos años es un insumo que interesa no solo a los expertos en analítica, sino también a docentes de las áreas de educación e ingeniería interesados en abordar este tipo de temática para trabajos de maestría y doctorado, puesto que es un área en crecimiento en el campo educativo. De igual manera, proporciona información pertinente sobre los métodos analíticos, que orienta la tarea dispendiosa de una búsqueda bibliográfica.

La tabla 7 presenta los trabajos por años calendario del 2015 al 2020. Se pueden apreciar el(los) método(s) usado(s) en cada investigación: autor, campo de acción, método que proporcionó buen resultado y *software* o herramienta utilizada en el estudio. Los acrónimos presentados en la columna de “estudio” refieren el enfoque de analítica académica implementado en cada investigación; específicamente se refieren a:

- RA: rendimiento académico
- DA: deserción académica
- GR: graduación

Tabla 7. Relación de investigaciones para la analítica académica

Autor	Estudio	Algoritmos o métodos utilizados	Muestra	Mejor resultado	Software utilizado
(Echegaray-Calderón y Barrios-Araníbar, 2015)	RA	Algoritmos genéticos y redes neuronales	1271	Algoritmos genéticos y redes neuronales	No especifica
(Pradeep <i>et al.</i> , 2015)	RA	JRip, OneR, J48 y ADtree.	670	Algoritmo PRISM	Weka
(Hernández, Meléndez <i>et al.</i> , 2016)	DA	Regresión logística, árboles de decisión y red neuronal	134	Regresión logística	No especifica
(Dixon <i>et al.</i> , 2016)	RA	Regresión jerárquica	609	Regresión jerárquica	Mplus 7.3
(Ahamed <i>et al.</i> , 2016)	RA	Random Forest, Naive Bayes y redes neuronales	423	Redes neuronales	No especifica
(Hasbun <i>et al.</i> , 2016)	DA	Árboles de decisión	4840	Árboles de decisión	R free
(Devasia <i>et al.</i> , 2016)	DA	Naive Bayes	700	Naive Bayes	No especifica
(Merchán-Rubiano <i>et al.</i> , 2017)	DA	Árbol de decisión J48, Random Forest	7644	Árbol J48	WEKA
(Razaque <i>et al.</i> , 2017)	RA	Naive Bayes	660	Naive Bayes	No especifica
(Stanko <i>et al.</i> , 2017)	RA	Regresión lineal y logística	125	Ninguno	No especifica
(Qin <i>et al.</i> , 2017)	RA	Random forest	250	Random forest	No especifica
(Costa <i>et al.</i> , 2017)	RA	Árbol de decisión, SVM, redes neuronales, Naive Bayes	423	SVM	WEKA
(Khalil <i>et al.</i> , 2018)	GR	Correlación estadística	180	Correlación estadística	SPSS
(Wheeler <i>et al.</i> , 2018)	GR	Regresión logística multivariada	323	Regresión logística multivariada	No especifica
(Liu <i>et al.</i> , 2018)	DA	Serie de tiempo	325	Serie de tiempo	No especifica

Autor	Estudio	Algoritmos o métodos utilizados	Muestra	Mejor resultado	Software utilizado
(Limsathitwong <i>et al.</i> , 2018)	DA	Algoritmo de bosque aleatorio	2880	Algoritmo de bosque aleatorio	RapidMiner
(Lima <i>et al.</i> , 2018)	DA	K-means y C5.0	1844	K-means y C5.0	Excel - R free
(Hirokawa, 2018)	RA	SVM	480	SVM	WEKA
(Wati <i>et al.</i> , 2018)	RA	Árbol de decisión C4.5, Naive Bayes	279	Naive Bayes	RAPIDMINER
(Miguéis <i>et al.</i> , 2018)	RA	SVM, Naive Bayes, bosque aleatorio, Bagging, Bosting	2459	Bosque aleatorio	RAPIDMINER
(Burgos <i>et al.</i> , 2018)	RA	Redes neuronales SVM, LOGIT_ACT	104	LOGIT_ACT	-
(Kaunang y Rotikan, 2018)	RA	Árbol de decisión, bosque aleatorio	300	Árbol de decisión	WEKA
(David y Anastasija, 2019)	GR	Árbol de clasificación	582	Árbol de clasificación	R free
(Ajibade, Ahmad <i>et al.</i> , 2019)	RA	Naive Bayes, árbol de decisión, KNN y análisis discriminante	500	K-vecinos más cercano (KNN)	Kalboard 360
(Nabizadeh <i>et al.</i> , 2019)	RA	Modelado de ecuaciones estructurales (SEM)	380	(SEM)	AMOS
(Martín, 2019)	RA	Regresión lineal múltiple	508	Regresión lineal	No específica
(Kauffman <i>et al.</i> , 2019)	GR	Correlación estadística	77	Correlación estadística	No específica
(Kwankajonwong <i>et al.</i> , 2019)	RA	Correlación estadística	137	Correlación estadística	IBM SPSS Statistics
(Herodotou <i>et al.</i> , 2019)	RA	Naive Bayes, árbol de clasificación (CART) y KNN	1384	Análisis KNN y CART	No específica
(Francis y Babu, 2019)	RA	Algoritmo ensamble	-	Algoritmo ensamble	No específica

Autor	Estudio	Algoritmos o métodos utilizados	Muestra	Mejor resultado	Software utilizado
(Vora y Rajamani, 2019)	RA	Algoritmo de clasificación híbrido	456	Algoritmo de clasificación híbrido	Java
(Salal <i>et al.</i> , 2019)	GR	Naive Bayes, árbol J48, bosque aleatorio	21	Árbol de decisión J48	WEKA
(Amazona y Hernández, 2019)	RA	Naive Bayes, árbol de decisión, Deep Learning	300	Redes Deep Learning	RAPIDMINER
(Qowidho <i>et al.</i> , 2019)	RA	Árbol de decisión ID3, C4.5	230	C4,5	RAPIDMINER
(Ajibade, Bahiah Binti Ahmad <i>et al.</i> , 2019) Educational Data Mining (EDM	RA	Árbol de decisión, KNN, SVM, Bagging, Bosting	500	Árbol de decisión	-
(Mohammadi <i>et al.</i> , 2019)	RA	KNN, Naive Bayes, árbol de decisión	230	KNN	WEKA
(Wiyono y Tegal, 2019)	RA	KNN, árbol de decisión, SVM	382	SVM	WEKA
(Adekitan y Noma-Osaghae, 2019)	RA	Bosque aleatorio, árbol de decisión, Naive Bayes, regresión logística	1445	Regresión logística	KNIME, ORANGE
(Nieto <i>et al.</i> , 2019)	RA	Árbol de decisión, regresión logística, bosque aleatorio	6100	Bosque aleatorio	PYTHON
(Francis y Babu, 2019)	RA	Árbol de decisión, Naive Bayes, redes neuronales, SVM	-	Árbol de decisión	-
(Hussain <i>et al.</i> , 2019)	RA	Redes neuronales, regresión logística, Naive Bayes, SVM, árbol de decisión	100	SVM	MATLAB, SPSS
(Lasarte <i>et al.</i> , 2020)	DA	Bootstrap	1468	Bootstrap	AMOS 24
(Tan <i>et al.</i> , 2020)	RA	Regresión lineal	143	Regresión lineal 0,57	No especifica

Autor	Estudio	Algoritmos o métodos utilizados	Muestra	Mejor resultado	Software utilizado
(Tenpipat y Akkarajitsakul, 2020)	DA	Árbol de decisión, bosque aleatorio y aumento de gradiente	13 714	Árbol de decisión	No especifica
(Medina <i>et al.</i> , 2020)	DA	Redes neuronales Bayesianas y árboles de decisión	500	Redes Bayesianas	SPSS
(Mavani, <i>et al.</i> , 2020)	RA	Naive Bayes	200	Naive Bayes	WEKA, R, PYTHON
(Mengash, 2020)	RA	Árbol de decisión, SVM, Naive Bayes	2039	Redes neuronales	WEKA
(Tomasevic <i>et al.</i> , 2020)	RA	KNN, SVM, redes neuronales, árbol de decisión, Naive Bayes, regresión logística	32 593	Redes neuronales	MATLAB
(Bhutto <i>et al.</i> , 2020)		Regresión logística y SVM	480	SVM 78 %	WEKA
(Contreras <i>et al.</i> , 2020)	RA	Árbol de decisión, KNN, SVM, perceptrón	1620	Perceptrón	PYTHON
(Viloria <i>et al.</i> , 2020)	DA	ADTree, J48 RandomTree, REPTree SimpleCart	1200	AD tree 98,6%	WEKA
(Urteaga <i>et al.</i> , 2020)	DA	KNN, redes neuronales	-	Redes neuronales 97%	R
(Castrillón <i>et al.</i> , 2020)	RA	J48	460	91,67%	WEKA
(Mengash, 2020)	DA	Redes neuronales, SVM, árbol de decisión, Naive Bayes	2039	Redes neuronales 79,22%	WEKA
(L. E. Contreras <i>et al.</i> , 2020)	DA	K-vecino más cercano, SVC, perceptrón y árbol de decisión	2114	Árbol de decisión 85,80%	PYTHON

Autor	Estudio	Algoritmos o métodos utilizados	Muestra	Mejor resultado	Software utilizado
(Hasan <i>et al.</i> , 2020)	RA	Árbol de decisión, bosque aleatorio, Boosting, AdaBoost, Gradient Boosting (XGBoost)	4413	AdaBoost	PYTHON
(Hassan <i>et al.</i> , 2020)	RA	Bosque aleatorio, árbol de decisión, Naive Bayes, regresión logística, KNN, SVM	772	Bosque aleatorio	PYTHON
(Zacharias y Athanasios, 2020)	RA	Red neuronal	76	Red neuronal	WEKA

Fuente: elaboración propia.

8.6. Enfoques de aplicación de la analítica académica en educación superior

Se organizaron los artículos en tres grandes categorías (en cada uno de los campos se realizó una división atendiendo el área de investigación; su objetivo principal es predecir o estudiar indicadores institucionales). A partir de los resultados, se puede apreciar que la analítica ha entrado con firmeza al campo educativo en los últimos años, para intentar explicar diversos indicadores de tipo académico en la educación superior. Asimismo, de los tres indicadores que generalmente son usados para determinar la eficiencia del proceso educativo, la analítica académica muestra mayor tendencia a la explicación del rendimiento académico.

Las investigaciones consultadas de los artículos se organizaron como se muestra en la tabla 8, teniendo en cuenta su área de investigación: graduación de los estudiantes (GR), deserción académica (DA) y rendimiento académico (RA).

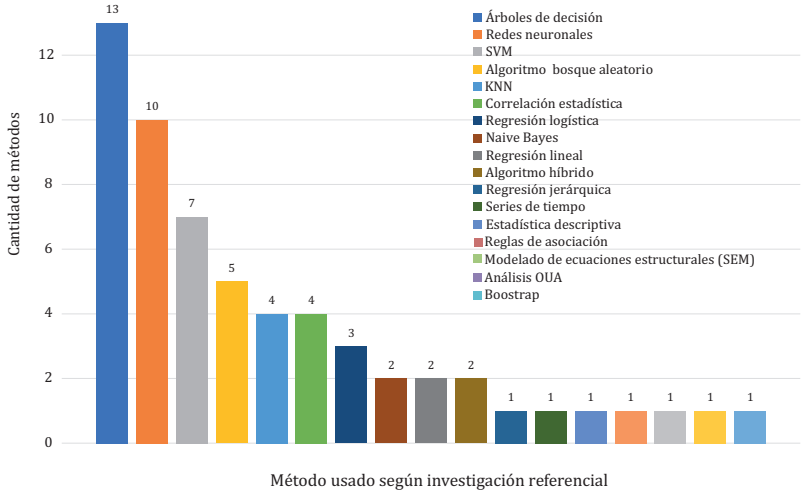
Tabla 8. Área de aplicación para la analítica académica

Graduación	(Khalil <i>et al.</i> , 2018); (Wheeler <i>et al.</i> , 2018); (David y Anastasija, 2019); (Vora y Rajamani, 2019); (Kauffman <i>et al.</i> , 2019)
Deserción	(Hernández, Meléndez <i>et al.</i> , 2016); (Hasbun <i>et al.</i> , 2016); (Devasia <i>et al.</i> , 2016); (Merchán-Rubiano <i>et al.</i> , 2017); (Liu <i>et al.</i> , 2018); (Limsathitwong <i>et al.</i> , 2018); (Lima <i>et al.</i> , 2018); (Lasarte <i>et al.</i> , 2020); (Tenpipat y Akkarajitsakul, 2020); (Medina <i>et al.</i> , 2020); (Contreras <i>et al.</i> , 2020)
Rendimiento académico	(Echegaray-Calderón y Barrios-Aranibar, 2015); (Pradeep <i>et al.</i> , 2015); (Dixon <i>et al.</i> , 2016); (Ahamed <i>et al.</i> , 2016); (Razaque <i>et al.</i> , 2017); (Stanko <i>et al.</i> , 2017); (Qin <i>et al.</i> , 2017); (Costa <i>et al.</i> , 2017); (Hirokawa, 2018); (Wati <i>et al.</i> , 2018); (Miguéis <i>et al.</i> , 2018); (Burgos <i>et al.</i> , 2018); (Kaunang y Rotikan, 2018); (Ajibade, Ahmad <i>et al.</i> , 2019); (Nabizadeh <i>et al.</i> , 2019); (Martín, 2019); (Kwankajonwong <i>et al.</i> , 2019); (Herodotou <i>et al.</i> , 2019); (Francis y Babu, 2019); (Vora y Rajamani, 2019); (Salal <i>et al.</i> , 2019); (Amazona y Hernández, 2019); (Qowidho <i>et al.</i> , 2019); (Ajibade, Bahiah Binti Ahmad <i>et al.</i> , 2019); (Mohammadi <i>et al.</i> , 2019); (Wiyono <i>et al.</i> , 2019); (Adekitan y Noma-Osaghae, 2019); (Nieto <i>et al.</i> , 2019); (Francis y Babu, 2019); (Hussain <i>et al.</i> , 2019); (Tan <i>et al.</i> , 2020); (Mavani <i>et al.</i> , 2020); (Mengash, 2020); (Tomasevic <i>et al.</i> , 2020); (Bhutto <i>et al.</i> , 2020); (Viloria <i>et al.</i> , 2020); (Urteaga <i>et al.</i> , 2020); (Castrillón <i>et al.</i> , 2020); (L. E. Contreras <i>et al.</i> , 2020); (Hasan <i>et al.</i> , 2020); (Hassan <i>et al.</i> , 2020); (Zacharias y Athanasios, 2020)

Fuente: elaboración propia.

La figura 18 presenta un panorama de las herramientas de análisis en la solución de problemas en este campo. Puede afirmarse que existe una fuerte tendencia a dejar de utilizar métodos tradicionalmente estadísticos. Se aprecia un fuerte crecimiento de algoritmos de minería de datos y aprendizaje automático de diversas ramas, como redes neuronales (aprendizaje profundo), bosque aleatorio (método de ensamble), árboles de decisión y SVM (aprendizaje supervisado clásico).

Figura 18. Panorama de métodos en analítica académica



Fuente: elaboración propia.

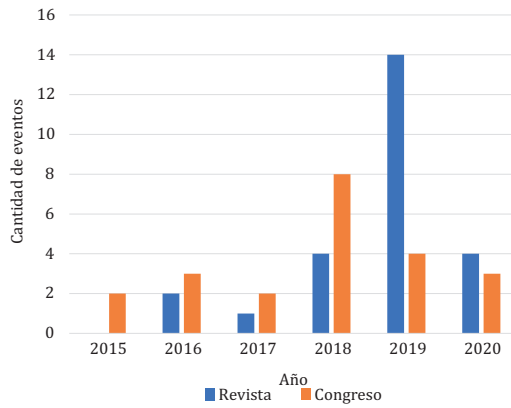
8.7. Construcción social de la analítica académica en educación superior

La socialización del conocimiento es un aspecto importante cuando se realizan investigaciones de alto impacto. En los últimos años, se nota una creciente tendencia de las investigaciones en aspectos relacionados con la analítica académica, pues es un área que está influenciando y permeando diversos campos como el educativo, para tratar de solucionar los inconvenientes presentes (Lasarte *et al.*, 2020). Los resultados de las investigaciones son socializados por medio de revistas indexadas, revisadas por pares y en una menor proporción en congresos o simposios (figura 19). Esta fuerte tendencia a trabajar estos temas puede deberse, como lo manifiestan Oblinger *et al.* (2007), a la aparición de técnicas que facilitan el análisis de datos académicos que buscan no solo estudiar la influencia de factores tecnológicos y de gestión en las instituciones, sino también la forma cómo aprenden y cómo se puede mejorar el proceso de enseñanza-aprendizaje en las instituciones.

La cantidad de publicaciones sobre analítica fue reducida durante el periodo 2015-2018. Pero parece que es un campo en crecimiento en cantidad de estudios en los últimos años (reflejado en un aumento hacia el 2020),

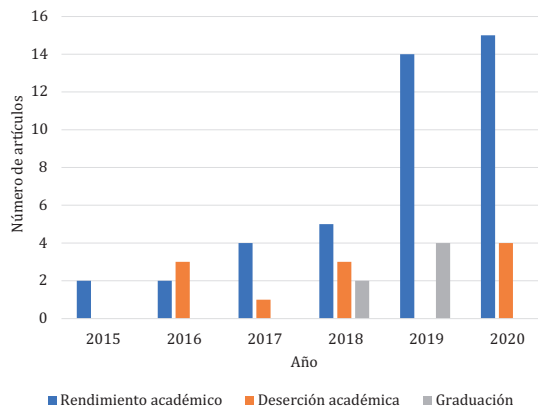
debido al auge de la minería de datos y el aprendizaje automático según el análisis referencial. Asimismo, en la figura 20 se observa el número de documentos reportados por año en el periodo de búsqueda, agrupados por campos de aplicación de la analítica académica. A partir de los resultados, puede afirmarse que tiene fuerte tendencia a la aplicación de métodos de analítica de datos aplicados al campo educativo, especialmente en lo concerniente a investigaciones relacionadas con la predicción del rendimiento académico de los estudiantes, seguido de lejos por las investigaciones sobre deserción académica.

Figura 19. Construcción social para la analítica académica



Fuente: elaboración propia.

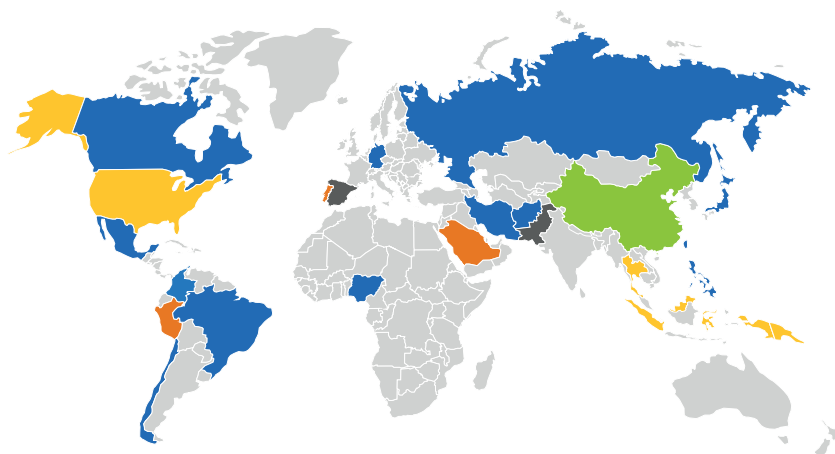
Figura 20. Cantidad de artículos por año según el enfoque de analítica académica



Fuente: elaboración propia.

En cuanto a las regiones o países que trabajan esta temática con mayor regularidad, se plantea que India es el país con el que mayor frecuencia se comparte autoría (figura 21). De los 55 documentos obtenidos, el 65 % se desarrolló en Asia y Europa, y en una baja proporción en los países de América. Este hecho permite suponer que, al igual que sucedió con la analítica del aprendizaje, en las universidades latinoamericanas un gran número de *software* o plataformas almacenan gran cantidad de datos provenientes del estudiante y estos parecen no recibir ningún tratamiento especial.

Figura 21. Panorama de países que realizan investigación en la temática



Fuente: elaboración propia.

8.8. Conclusiones del capítulo

Con base en la información consultada y analizada, es posible plantear las siguientes conclusiones principales (Contreras *et al.*, 2021):

Tanto la analítica académica como la del aprendizaje son campos emergentes que están apenas incursionando en la educación en los países latinoamericanos, a la luz de la revisión referencial consultada. Incursiona, hoy en día, por medio de las herramientas de aprendizaje automático que han solucionado múltiples problemas en otros campos de las ciencias.

Las investigaciones respecto a la analítica académica se han planteado en su gran mayoría para adelantar predicciones sobre el rendimiento académico, y lo han hecho desde diferentes áreas como la psicología y la estadística

principalmente, aunque se muestran indicios claros del uso de algoritmos supervisados y de ensamble correspondientes al *Machine Learning*.

Los tres indicadores de la calidad educativa que abordan la analítica académica se han visto fortalecidos con la aparición de mejores y potentes *software*, herramientas de captura de información y análisis, hecho que les ha permitido a los investigadores realizar aproximaciones para su determinación utilizando diferentes variables independientes en diversos trabajos.

Capítulo 9

Caracterización de metodologías para la evaluación del rendimiento académico en la educación superior

El rendimiento académico es un tema ampliamente estudiado y controvertido en el campo de la investigación educativa (Lamas, 2015; Montero *et al.*, 2007). Estas investigaciones varían desde su propio concepto hasta su medición y estimación.

Algunos autores consideran que puede definirse como una medida de las capacidades adquiridas en el proceso formativo, como el cumplimiento de logros en un programa cursado, el nivel de conocimiento de un área respecto a una norma, entre otros (Lamas, 2015). Otros manifiestan que lo importante es cómo medirlo. Tal es el caso de los planteamientos sobre la medición con las calificaciones escolares (Lamas, 2015), o tal vez mediante la inclusión de otras variables que pueden incidir en su medición (Hernández, 2016; Rodríguez y Ruiz, 2009).

En concordancia con lo anterior, este capítulo presenta algunas de las metodologías de diversos autores que han desarrollado, desde diferentes áreas de estudio, la medición y evaluación del rendimiento académico en la educación superior. De esta manera, se podrán identificar algunas variables útiles que lo determinan y lo afectan. La metodología para realizar esta parte de la revisión bibliográfica del tema es similar a la proporcionada anteriormente.

9.1. Propósito

El propósito de la revisión referencial consiste en caracterizar las metodologías existentes para la evaluación del rendimiento académico en educación superior, con el fin de identificar las variables útiles que permitan determinar el rendimiento académico de estudiantes universitarios.

9.2. Planificación de la revisión de literatura

Se consideró la revisión de artículos publicados utilizando combinaciones de palabras clave, como *rendimiento académico*, *rendimiento académico universidad*, *medición rendimiento académico*, *evaluación rendimiento académico*, *desempeño académico*, *academic performance*, *academic performance assessment*, *academic achievement* y *academic performance factors*.

9.3. Criterios de selección de la información

El periodo de búsqueda de información corresponde a los últimos cinco años (2015-2020). Es importante mencionar que se partió de un total de 118 documentos encontrados, se utilizaron efectivamente 50 de ellos, pues estos cumplieron con la característica de estar orientados hacia un estudio o investigación sobre la medición del rendimiento académico aplicado en múltiples contextos en educación superior. Los criterios establecidos para reducir la cantidad de artículos son similares a los descritos anteriormente, pero aplicados al rendimiento académico y no a la analítica de datos. Con el fin de ayudar a resolver el problema, las siguientes preguntas guiaron esta revisión:

- ¿Por medio de campos del conocimiento se ha abordado la temática de investigar el rendimiento académico en la educación superior?
- ¿Cuáles son los aspectos más relevantes de cada una de las metodologías analizadas, incluyendo las características propias del estudio?
- ¿Cómo ha sido la construcción social en torno al tema; por ejemplo, país, número de participantes, herramientas, resultados, entre otros, que permitan una adecuada caracterización y descripción de las mismas?

9.4. Análisis e interpretación de datos relevantes

Los investigadores agotaron la bibliografía utilizando el procedimiento, las palabras clave y las bases de datos descritas anteriormente. Por tanto, es posible que esta literatura no incluya fuentes no disponibles a través de los criterios y bases de datos no especificadas en este documento. Los documentos consultados en esta revisión bibliográfica se muestran en la tabla 9.

Tabla 9. Enfoque, referencia y base de datos consultada

Enfoque	Fuente	Base de datos
Psicología	(Hernández, 2016)	Scopus
	(Mushtaq y Khan, 2012)	Springer
	(Furnham <i>et al.</i> , 2013)	Springer
	(Zuniga-Jara <i>et al.</i> , 2018)	SciELO
	(Chilca, 2017)	Dialnet
	(Ortiz <i>et al.</i> , 2020)	SciELO
	(Hakyemez y Mardikyan, 2021)	ScienceDirect
	(Jain <i>et al.</i> , 2016)	Springer
	(Espinosa <i>et al.</i> , 2020)	ScienceDirect
	(Luján <i>et al.</i> , 2019)	Scopus
	(Fernández-Lasarte <i>et al.</i> , 2019)	Scopus
	(Guizado <i>et al.</i> , 2020)	SciELO
	(Zárate <i>et al.</i> , 2020)	SciELO
	(Lin <i>et al.</i> , 2020)	Biomed central
	(Deliens <i>et al.</i> , 2013)	Biomed central
	(Bahammam <i>et al.</i> , 2012)	Biomed central
	(Céspedes <i>et al.</i> , 2018)	SciELO
	García-Hernández y González-Ramírez, 2020)	Scopus
	(Rosales-Jaramillo <i>et al.</i> , 2018)	ERIC
	(Rivera <i>et al.</i> , 2020)	Open Journal Systems

Enfoque	Fuente	Base de datos
Psicología	(Castro y Closas, 2020)	1findr
	(Hayat <i>et al.</i> , 2020)	Biomed central
	(Mazumder <i>et al.</i> , 2020)	ERIC
	(Awadalla <i>et al.</i> , 2020)	Biomed central
	(Icekson <i>et al.</i> , 2020)	Scopus
	(Minichil <i>et al.</i> , 2020)	MEDLINE complete
	(Vargas <i>et al.</i> , 2011)	SciELO
	(Albela-García <i>et al.</i> , 2020)	1findr
	(Salmani <i>et al.</i> , 2020)637 university students participated in the study (response rate: 100%	Taylor y Francis <i>on-line</i>
	(Bayona, 2017)	Scopus
	(Bautista-Rodríguez y Gatica-Lara, 2020)	SciELO
	(Nicholson <i>et al.</i> , 2016)	Biomed central
	(Montoya-Arenas <i>et al.</i> , 2021)	Portal regional de la BVS
Estadística - matemática	(Muñoz-Villanueva <i>et al.</i> , 2020)	SciELO
	(Huamán, 2018)	Scopus
	(Shahrakipour <i>et al.</i> , 2016)	ResearchGate
	(Ñaupari, 2014)	Scopus
	(Vargas <i>et al.</i> , 2019)	Public Knowledge Project
	(Durán y Rosado, 2019)	Scopus
	(Bonsaksen, 2016)	SAGE Journals
	(Liang <i>et al.</i> , 2018)	ERIC
	(Pavelea y Moldovan, 2020)previous academic performance, in-class involvement and attention, class schedule, gender and other control variables	Sciendo
	(Farrokhi-Khajeh-Pasha <i>et al.</i> , 2012)	Scopus
	(Lenskiy <i>et al.</i> , 2020)	SAGE Journals
	(Grob <i>et al.</i> , 2015)	SciELO

Enfoque	Fuente	Base de datos
Medicina	(Sanguinetti <i>et al.</i> , 2013)	ScienceDirect
	(Muñoz-Comonfort <i>et al.</i> , 2014)	ScienceDirect
	(Buhagiar y Leo, 2018)	ERIC
	(Thomas <i>et al.</i> , 2020)SD = 2.76 years	Taylor y Francis <i>on-line</i>
	(Thomas y Viljoen, 2019)	Scopus

Fuente: elaboración propia.

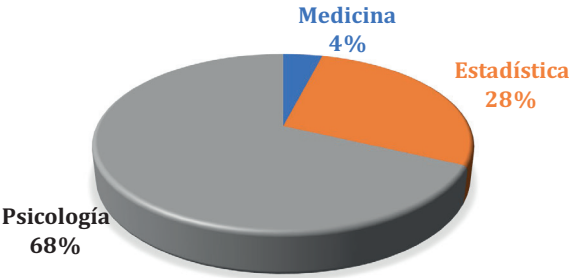
9.5. Herramientas aplicadas al rendimiento académico

De manera general, es posible observar que de los 50 estudios encontrados, la gran mayoría de estos (68 %) corresponden a estudios enfocados desde el área de la psicología, seguidos del área de la estadística-matemática con un porcentaje de 28 %, mientras que el área de medicina concierne al 4 % (figura 22). Esto demuestra la predilección de los investigadores por las herramientas relativas a cuestionarios sobre percepciones de los estudiantes respecto al rendimiento académico, para la recolección inicial de los datos; seguido principalmente de las herramientas estadísticas, las cuales tienen un enfoque mucho más marcado hacia los datos demográficos y su influencia sobre la variable de interés. Sin embargo, cabe aclarar que, en la mayoría de los estudios, independientemente de su categoría de enfoque, las herramientas para el análisis de la información son netamente estadísticas.

Cabe señalar que los autores vieron la necesidad, a partir de la información consultada, de examinar las herramientas de análisis de datos que utilizan algoritmos de minería de datos y *Machine Learning* (aprendizaje automático) que se aplican al campo educativo y, especialmente, al rendimiento académico. Por esta razón se está trabajando en este aspecto (no contenido en este documento).

El estudio del rendimiento académico en educación superior ha estado presente en investigaciones de diversos campos de las ciencias. A continuación, se presenta en la tabla 10 cada uno de los trabajos consultados, clasificados por año de publicación y algunos de sus aspectos relevantes para el periodo 2011-2021.

Figura 22. Áreas que han estudiado la evaluación del rendimiento académico según la investigación referencial



Fuente: elaboración propia.

Tabla 10. Resumen de características de los trabajos realizados en cada campo del conocimiento

Año	Autor	Estudio referente a:	Variables independientes	Variables dependientes	Tamaño de la muestra	País
2016	(Hernández, 2016)	Psicología	– Género – Carrera – Turno – Semestre – Número de asignaturas no aprobadas – Tiempo de trayecto – Trabajo – Tutor – Promedio	Rendimiento académico a través de las respuestas tipo Likert a una encuesta de 30 preguntas, obtenidas a partir de la revisión de la literatura	1651	México
2012	(Mushtaq y Khan, 2012)	Psicología	– Orientación adecuada – Habilidades de comunicación de los estudiantes – Instalaciones de aprendizaje – Estrés familiar – (todas las variables se derivan del cuestionario)	Rendimiento académico a través de los resultados obtenidos en el cuestionario, dentro de la sección para tal fin	155	Pakistán
2013	(Furnham <i>et al.</i> , 2013)	Psicología	Variables sobre la personalidad, resultado del test “BIG FIVE”, datos demográficos (sexo)	Rendimiento académico según dos indicadores: exámenes y trabajo de curso	1013	Gran Bretaña

Año	Autor	Estudio referente a:	Variables independientes	Variables dependientes	Tamaño de la muestra	País
2018	(Zuniga-Jara <i>et al.</i> , 2018)	Psicología	– Sensibilidad a la recompensa – Inadecuada calidad dietaria por sobrealimentación – Preferencia por alimentos grasos o dulces – (obtenidas a partir de cuestionarios específicos)	– Índice de masa corporal – Rendimiento académico según número de asignaturas reprobadas por un estudiante	173	Chile
2017	(Chilca, 2017)	Psicología	– Variables relacionadas con la autoestima y hábitos de estudio – (obtenidas de cuestionarios)	– Rendimiento académico a través del promedio de calificaciones de los estudiantes en la materia “Matemática básica I”	86	Perú
2020	(Ortiz <i>et al.</i> , 2020)	Psicología	– Variables relativas a la ansiedad matemática (ansiedad matemática hacia el aprendizaje, ansiedad matemática hacia la evaluación)	– Rendimiento académico según los resultados del semestre cursado, si los estudiantes aprobaron o desaprobaron	87	Colombia
2021	(Hakyemez y Mardikyan, 2021)	Psicología	– Datos demográficos (edad, género, ayuda financiera, tamaño pueblo natal, tamaño de la dependencia) – Variables sobre integración institucional y autoeficacia derivadas de los cuestionarios	– Rendimiento académico según el promedio de calificaciones	443	Turquía
2016	(Jain <i>et al.</i> , 2016)	Psicología	– Variables obtenidas de los cuestionarios de autoeficacia profesional y tutorías – Datos demográficos de los estudiantes	– Rendimiento académico medido a través del promedio porcentual acumulado	327	India

Año	Autor	Estudio referente a:	Variables independientes	Variables dependientes	Tamaño de la muestra	País
2020	(Espinosa <i>et al.</i> , 2020)	Psicología	- Variables obtenidas mediante la escala para la determinación del nivel de estrés - Factores demográficos (si un estudiante trabaja, personas con las que vive, sexo, edad)	Rendimiento académico mediante los promedios académicos de los estudiantes	115	Colombia
2019	(Luján <i>et al.</i> , 2019)	Psicología	- Variables referentes a la asertividad, obtenidas de un cuestionario - Datos demográficos (sexo, edad)	Rendimiento académico medido a través del promedio ponderado acumulado	69	Perú
2019	(Fernández-Lasarte <i>et al.</i> , 2019)	Psicología	- Apoyo social (apoyo del profesorado, apoyo familiar, apoyo de las amistades) - Inteligencia emocional (atención emocional, claridad emocional y reparación emocional)	Rendimiento académico obtenido a partir de los resultados de uno de los cuestionarios aplicados, a través de ítems con respuestas en escala Likert	419	España
2020	(Guizado <i>et al.</i> , 2020)	Psicología	Desempeño docente desde los ámbitos profesional, personal y social	Rendimiento académico medido a través del promedio de calificaciones	399	Perú
2020	(Zárate <i>et al.</i> , 2020)	Psicología	Competencia comunicativa intercultural (cognitiva, comportamental y afectiva)	Rendimiento académico medido a través de las calificaciones finales de los estudiantes en las asignaturas	280	Perú
2020	(Lin <i>et al.</i> , 2020)	Psicología	- Variables relativas a los resultados de los cuestionarios sobre el estrés - Datos demográficos (edad, sexo, año de estudio y estado civil)	Rendimiento académico a través del promedio de calificaciones	347	China

Año	Autor	Estudio referente a:	Variables independientes	Variables dependientes	Tamaño de la muestra	País
2013	(Deliens <i>et al.</i> , 2013)	Psicología	– Medidas antropométricas (peso, índice de masa corporal, porcentaje de grasa, contorno de cintura) – Datos sociodemográficos (género, edad, etnia, residencia, promedio de calificaciones del último año de secundaria) – Estatus socioeconómico (educación de los padres) – Hábitos dietarios – Hábitos de sueño – Hábitos de fumar – Hábitos de actividad física	– Rendimiento académico a través del promedio de calificaciones	101	Bélgica
2012	(Bahammam <i>et al.</i> , 2012)	Psicología	– Datos demográficos (sexo, índice de masa corporal, obesidad, hábito de fumar, uso de internet, trabajo) – Hábitos de sueños	– Rendimiento académico a través del promedio de calificaciones	410	Arabia Saudita
2018	(Céspedes <i>et al.</i> , 2018)	Psicología	– Ausentismo – Presentismo – (obtenidas de cuestionarios)	– Rendimiento académico a través de las respuestas al cuestionario, en las secciones de “Rendimiento académico disminuido”, “Cursos y ciclos desaprobados” y “Evolución del rendimiento académico”	8203	Perú

Año	Autor	Estudio referente a:	Variables independientes	Variables dependientes	Tamaño de la muestra	País
2020	García-Hernández y González-Ramírez, 2020)	Psicología	– Sexo – Provincia de procedencia – Puntaje en la prueba de ingreso de matemáticas – Opción en la que se solicitó la carrera – Materiales de estudio en la asignatura de matemáticas discretas	– Rendimiento académico a través de la nota final en matemáticas discretas	297	Cuba
2018	(Rosales-Jaramillo <i>et al.</i> , 2018)	Psicología	– Variables personales (edad, sexo, procedencia, nota bachillerato, hábito de leer, motivación intrínseca, entre otras) – Variables sociales (ciclo actual de estudio, tipo de bachillerato, motivación extrínseca, trabajo, acceso a internet, consumo de alcohol, entre otras) – Variables institucionales (horario actual, tipo de evaluación aplicada, aulas adecuadas, guía y motivación docente, entre otras)	– Rendimiento académico a través del promedio del puntaje de todas las materias (aprobadas y no aprobadas)	813	Ecuador
2020	(Rivera <i>et al.</i> , 2020)	Psicología	– Empatía (medida y evaluada a través de cuestionarios)	– Rendimiento académico a través del promedio acumulativo de todas las asignaturas en todos los semestres	128	Perú

Año	Autor	Estudio referente a:	Variables independientes	Variables dependientes	Tamaño de la muestra	País
2020	(Castro y Closas, 2020)	Psicología	– Constructos sobre enseñanza – Constructos sobre materiales – Constructos sobre aprendizaje	– Rendimiento académico a través de tres variables: promedio de calificaciones de pruebas escritas, cantidad de exámenes realizados durante la asignatura, situación académica (resultado general al final de semestre)	118	Argentina
2020	(Hayat <i>et al.</i> , 2020)	Psicología	– Variables relacionadas con emociones positivas del aprendizaje, autoeficacia, estrategias de aprendizaje metacognitivo (cada una recolectada de cuestionarios específicos)	– Rendimiento académico a través de los puntajes de exámenes finales del semestre, puntajes de exámenes parciales, trabajo semestral (quices, tareas, documentos y proyectos de investigación a lo largo del periodo)	279	Irán
2020	(Mazumder <i>et al.</i> , 2020)	Psicología	– Participación en clase medida a través de autorreporte, observaciones y gestos	– Rendimiento académico a través del promedio de calificaciones	44	Estados Unidos

Año	Autor	Estudio referente a:	Variables independientes	Variables dependientes	Tamaño de la muestra	País
2020	(Awadalla <i>et al.</i> , 2020)	Psicología	– Factores demográficos y socioeconómicos (edad, sexo, estado civil, especialización, año de estudio, nivel educativo de los padres, evaluación de los recursos materiales de la familia) – Depresión – Ansiedad – (estos dos últimos a través de cuestionarios específicos)	– Rendimiento académico a través del promedio de calificaciones más recientes, advertencias de asistencia, satisfacción con el curso mediante un cuestionario en escala Likert	404	Emiratos Árabes Unidos
2020	(Icekson <i>et al.</i> , 2020)	Psicología	– Optimismo – Conciencia – (cada uno de cuestionarios específicos)	– Rendimiento académico medido a través del promedio anual de calificaciones	175	Israel
2020	(Minichil <i>et al.</i> , 2020)	Psicología	– Factores sociodemográficos – Factores obstétricos y ginecológicos (en donde se incluía la percepción sobre el rendimiento académico) – Factores psicosociales – Consumo de sustancias	– Puntuación obtenida del cuestionario para el trastorno disfórico premenstrual (TDPM)	386	Etiopía
2011	(Vargas <i>et al.</i> , 2011)	Psicología	– Factores sociodemográficos – Trayectoria académica – Rasgos de personalidad – Pensamiento abstracto – Pensamiento creativo – Trastorno mental	– Rendimiento académico medido a través del promedio final de calificaciones de las asignaturas, así como uno general	94	México
2020	(Albela-García <i>et al.</i> , 2020)	Psicología	– Uso de técnicas de aprendizaje (tutorías y tutorías con seminarios)	– Rendimiento académico (en el estudio no se especifica su indicador)	80	Perú

Año	Autor	Estudio referente a:	Variables independientes	Variables dependientes	Tamaño de la muestra	País
2020	(Salmani <i>et al.</i> , 2020)	Psicología	– Características demográficas – Prevalencia de los trastornos del sueño	– Rendimiento académico medido a través del promedio de calificaciones	637	Omán
2017	(Bayona, 2017)	Psicología	– Hábitos de estudio medidos a través de las siguientes escalas: condiciones ambientales, planificación del estudio, utilización de materiales, asimilación de contenidos y sinceridad	– Rendimiento académico medido a través del promedio ponderado del semestre académico	50 y 21	Colombia
2020	(Bautista-Rodríguez y Gatica-Lara, 2020)	Psicología	– Características sociodemográficas (sexo, edad, lugar de residencia) – Características académicas (semestre cursado, asignaturas no acreditadas) – Características laborales (actividad laboral, antigüedad laboral) – Habilidades digitales (disposición de tecnología, habilidades digitales)	– Rendimiento académico medido a través de las calificaciones globales	126	México
2016	(Nicholson <i>et al.</i> , 2016)	Psicología	– Asistencia a un taller – Rendimiento académico previo (examen de mitad de semestre)	– Rendimiento académico medido a través de un examen final de anatomía	66	Australia
2021	(Montoya-Arenas <i>et al.</i> , 2021)	Psicología	– Variables relativas a la capacidad intelectual y a las funciones ejecutivas	– Rendimiento académico mediante el promedio aritmético ponderado de las calificaciones en los semestres cursados	42	Colombia

Año	Autor	Estudio referente a:	Variables independientes	Variables dependientes	Tamaño de la muestra	País
2020	(Muñoz-Villanueva <i>et al.</i> , 2020)	Psicología	– Variables relacionadas al cuestionario sobre hábitos alimenticios en el desayuno	– Rendimiento académico medido a través del promedio de calificaciones de tres asignaturas	39	Chile
2018	(Huamán, 2018)	Estadística Matemática	– Área formativa – Sexo – Curso/asignatura	– Rendimiento académico medido a través de las calificaciones obtenidas de cada uno de los cursos de la carrera	26	Perú
2016	(Shahrakipour <i>et al.</i> , 2016)	Estadística Matemática	– Datos demográficos (edad, sexo, tipo y nivel educativo de los padres/madres, tipo de facultad	– Rendimiento académico medido a través del promedio de calificaciones	200	Irán
2014	(Ñaupari, 2014)	Estadística Matemática	– Tipo de modalidad (presencial, virtual)	– Rendimiento académico medido a través de un indicador construido a partir del promedio o índice de aprobación	505	Perú
2019	(Vargas <i>et al.</i> , 2019)	Estadística Matemática	– Puntaje de examen de ingreso mediante dos modalidades distintas: puntaje de admisión y puntaje del centro preuniversitario	– Rendimiento académico a través del ranking académico de los estudiantes ingresantes a la universidad	1551	Perú

Año	Autor	Estudio referente a:	Variables independientes	Variables dependientes	Tamaño de la muestra	País
2019	(Durán y Rosado, 2019)	Estadística Matemática	– Variables relacionadas con la comprensión lectora de los estudiantes (didácticas docentes y prueba escrita aplicada)	– Rendimiento académico medido a través de dos indicadores: nivel de conocimiento en la prueba SABER 11 y otro es el reporte de calificaciones de cada corte semestral en las asignaturas de lenguaje	105	Colombia
2016	(Bonsaksen, 2016)	Estadística Matemática	– Factores sociodemográficos – Factores de relación – Factores educativos – Factores laborales	– Rendimiento académico a través del promedio de calificaciones autoevaluado, obtenido de un cuestionario – Variables relacionadas a la satisfacción de los estudiantes con su programa educativo, obtenidas a partir de un cuestionario	123	Noruega
2018	(Liang <i>et al.</i> , 2018)	Estadística Matemática	– Datos demográficos (género y etnia)	– Rendimiento académico medido a través del promedio numérico de calificaciones	1122	Estados Unidos

Año	Autor	Estudio referente a:	Variables independientes	Variables dependientes	Tamaño de la muestra	País
2020	(Pavelea y Moldovan, 2020)	Estadística Matemática	– Asistencia – Atención e intervención en clase – Características de los estudiantes (motivación y conciencia) – Otros factores (edad, género, lugar de residencia, distancia entre el lugar de estudio y hogar, tiempo del día para asistir al seminario)	– Rendimiento académico a través de una nota final de un <i>paper</i>	115	Rumanía
2012	(Farrokhi-Khajeh-Pasha <i>et al.</i> , 2012)	Estadística Matemática	– Puntaje total Konkoor y sus ocho subsecciones – Promedio de calificaciones de secundaria	– Rendimiento académico a través del puntaje CBSE (examen integral de ciencias básicas), puntaje CPIE (examen integral previo a la pasantía), promedio de calificaciones de la escuela de medicina	2126	Irán
2020	(Lenskiy <i>et al.</i> , 2020)	Estadística Matemática	– Receso académico durante la carrera (si se tomó o no se tomó, y el número de semestres que duró el receso)	– Rendimiento académico medido a través del promedio de calificaciones	584	Corea del Sur
2015	(Grob <i>et al.</i> , 2015)	Estadística Matemática	– Género – Promedio puntaje de notas de educación media – Puntaje PSU – Tipo de colegio de procedencia	– Rendimiento académico medido a través de las notas definitivas obtenidas en una asignatura	312	Chile

Año	Autor	Estudio referente a:	Variables independientes	Variables dependientes	Tamaño de la muestra	País
2013	(Sanguinetti <i>et al.</i> , 2013)	Estadística Matemática	– Datos demográficos (edad, sexo, provincia de origen) – Condición laboral – Computadora en el hogar – Uso del correo electrónico – Uso de redes sociales – Realización de deportes – Participación en las ONG – Satisfacción con la carrera – Libros leídos durante el último año – Aprobación del examen – Situación económica – Tipo de educación secundaria	– Rendimiento académico medido a través de la aprobación de exámenes y asistencia	188	Argentina
2014	(Muñoz-Comonfort <i>et al.</i> , 2014)	Estadística Matemática	– Evaluación diagnóstica – Ciclo de inicio de la carrera – Tipo de escuela de procedencia	– Rendimiento académico medido a través de calificación promedio de exámenes departamentales	6590	México
2018	(Buhagiar y Leo, 2018)	Estadística Matemática	– Uso de juegos para el aprendizaje	– Rendimiento académico medido a través de la calificación final en la asignatura “Programa integrado de empresas”	116	Estados Unidos

Año	Autor	Estudio referente a:	Variables independientes	Variables dependientes	Tamaño de la muestra	País
2020	(Thomas <i>et al.</i> , 2020)	Medicina	– Biomarcadores de niveles de inflamación – Datos demográficos (edad, sexo y etnia)	– Rendimiento académico mediante la media general estandarizada de calificaciones en un núcleo común de clases requeridas para la graduación y el desempeño en un módulo específico	180	Sudáfrica
2019	(Thomas y Viljoen, 2019)	Medicina	– Variabilidad de la frecuencia cardíaca a través de dos indicadores (PT y RMSSD)	– Rendimiento académico mediante la media general estandarizada de calificaciones en un núcleo común de clases requeridas para la graduación y el desempeño en un módulo específico, considerando cursos desaprobados con 50 puntos	76	Sudáfrica

Fuente: elaboración propia.

9.5.1. Descripción de algunas herramientas usadas para las investigaciones sobre rendimiento académico

Esta revisión bibliográfica abarca investigaciones de los tres campos mencionados con anterioridad, en los que se han trabajado teniendo en cuenta diversas variables y herramientas, algunas de las cuales se describen brevemente en la tabla 11.

Tabla 11. Algunas herramientas para evaluar el rendimiento académico, según trabajos realizados en cada campo del conocimiento

Grupo de herramientas	Herramienta	Descripción
Herramientas que miden la consistencia interna	Coeficiente Alfa de Cronbach (Oviedo y Campo-Arias, 2005, 2008):	Es un índice que se emplea para medir la confiabilidad y la consistencia interna de una escala o instrumento aplicado, evaluando el grado en que los ítems de este se correlacionan entre sí. Por lo general, un grupo de ítems que busca evaluar un factor común, muestra un elevado valor para el indicador, el cual es calculado a partir del promedio de las correlaciones entre los ítems del instrumento. El coeficiente Alfa de Cronbach es recomendado cuando el instrumento posee respuestas politómicas.
	Coeficiente de confiabilidad de Kuder-Richardson (Oviedo y Campo-Arias, 2005, 2008):	Al igual que el Alfa de Cronbach, el coeficiente de Kuder-Richardson se emplea para evaluar la consistencia interna de un instrumento compuesto por varios ítems; sin embargo, esta herramienta es recomendada para medir la consistencia interna cuando el instrumento posee respuestas dicotómicas. Matemáticamente, equivale al coeficiente de Cronbach con un desarrollo posterior.
Herramientas relacionadas con cuestionarios	Escala Likert (Maldonado, 2007):	Es un instrumento para la medición y recolección de datos de naturaleza cualitativa. Se constituye por una serie de ítems a modo de afirmaciones, mediante los cuales se evalúa la reacción del individuo ante estas. Las respuestas están dadas a través de grados de acuerdo o desacuerdo según la sentencia particular.
		Por lo general, estos ítems tienen cinco opciones de respuesta y a cada una se le asigna un valor numérico que al final consolida una puntuación total. Esta herramienta se emplea, en gran medida, para medir actitudes.
	Escala de capacidad percibida de Trapnell (Rodríguez, 2009):	Es un breve instrumento de 4 ítems valorados en una escala Likert de 9 puntos, para la medición indirecta de la capacidad intelectual autoatribuida. Las preguntas que plantea no se refieren a tareas particulares, sino a la capacidad intelectual general de un estudiante.

Grupo de herramientas	Herramienta	Descripción
Herramientas relacionadas con cuestionarios	Escala de estrés percibido (PSS) (Campo-Arias <i>et al.</i> , 2014):	Esta escala es una de las más conocidas para la cuantificación del estrés emocional a través de la medición de la respuesta psicológica de los individuos, especialmente aplicada en investigaciones clínicas y epidemiológicas. Este instrumento cuenta con varias versiones en la actualidad, entre las que se encuentran su versión original (de 14 ítems), y dos de las más conocidas a nivel mundial (una de 10 ítems y otra de 4 ítems), surgidas como un refinamiento de la primera, y que se encuentran disponibles en múltiples idiomas, incluyendo el español.
	Test “Big Five Inventory” (Depaula y Azzollini, 2013; Domínguez-Lara <i>et al.</i> , 2018):	Es un instrumento basado en el autoinforme, que a través de 44 ítems evalúa los cinco grandes factores de la personalidad: extraversión, agradabilidad, responsabilidad, neuroticismo y apertura a la experiencia. Cada uno de esos ítems se evalúa en una escala Likert según un grado de conformidad (desde Muy en desacuerdo hasta Muy de acuerdo). La puntuación total del test se determina con la suma de los ítems, lo que indica que una alta puntuación en alguno de los factores refiere a una mayor presencia de este atributo en el individuo.
	Binge Eating Scale (BES) (Cotter y Kelly, 2016):	La “escala de atracones” es un instrumento de autoinforme compuesto de 16 ítems, que busca capturar las características del comportamiento, emocionales y cognitivas de los “atracones” de comida, mediante preguntas que indagan sobre la cantidad de comida consumida, el control al comer, entre otros. Cada uno de los ítems se puede puntuar por medio de una escala de 3 o 4 puntos para ser sumados al final y obtener una puntuación total que oscila entre 0 y 46, y son las puntuaciones más altas, un indicador de problemas de atracones de gravedad.
	Escala abreviada de ansiedad matemática (Tejedor <i>et al.</i> , 2009):	Se diseñó este cuestionario para evaluar la ansiedad frente a las matemáticas. Se compone de nueve cuestiones para que los individuos decidan qué tanta ansiedad les generan algunas situaciones relacionadas con las matemáticas, a través de una valoración en una escala de 5 puntos que varía entre Muy baja ansiedad = 1, y Alta ansiedad = 5. El cuestionario presenta un resultado total por medio de la suma del número de puntos en cada una de las preguntas, oscilando así entre 9 y 45 puntos. Si dicha puntuación es alta, es indicativo de altos niveles de ansiedad respecto a las matemáticas.

Grupo de herramientas	Herramienta	Descripción
Herramientas relacionadas con cuestionarios	Escala breve de ajuste escolar (EBAE-10) (Moral <i>et al.</i> , 2010):	Es un instrumento que mide un constructo basado en indicadores sobre las buenas relaciones (con compañeros y docentes), aplicación (en las tareas y el aula), motivación para la asistencia a clases, y un autoconcepto positivo, expectativas de estudios posteriores, entre otros. Está compuesto por 10 ítems en formato Likert, que se evalúan en un rango de 6 puntos, que en total conforman una puntuación final que oscila entre 10 y 60 puntos, y es una mayor puntuación, un mayor ajuste escolar.
Herramientas del análisis factorial exploratorio	Análisis factorial exploratorio (Lloret-Segura <i>et al.</i> , 2014; López-Aguado y Gutiérrez-Provecho, 2019; Mavrou, 2015):	Es una de las herramientas más usadas para desarrollar, validar y adaptar instrumentos de medida psicológicos. Esta técnica estadística permite explorar, con gran precisión, tanto dimensiones subyacentes y constructos, como variables latentes de las variables analizadas (las que son observadas y medidas por el investigador) al comprobar qué tanto son estas, representadas por los ítems que constituyen el instrumento. En otras palabras, permite definir unas pocas dimensiones comunes que puedan explicar la mayor parte de la varianza observada de un conjunto mucho más grande de variables.
	Prueba Kaiser-Meyer-Olkin (Garmendia, 2007; López-Aguado y Gutiérrez-Provecho, 2019):	La prueba KMO permite comprobar el grado de relación entre un conjunto de variables; es decir, el grado en que cada una de ellas es predecible a partir de las otras. Esta medida toma valores entre 0 y 1, y es adecuado un valor mayor a 0,6, pues implica que las variables estarán más relacionadas entre sí (según algunos autores, incluso 0,5 es aún adecuado).
Herramientas para la evaluación de la normalidad y significancia estadística	Prueba de Kolmogorov-Smirnov (Romero-Saldaña, 2016):	La prueba de significación estadística K-S se emplea para verificar si los datos de una muestra determinada proceden de una distribución normal. Su uso se da cuando existen variables cuantitativas continuas, y cuando el tamaño de la muestra es mayor a 50.
	Corrección de Lilliefors (Romero-Saldaña, 2016):	En la práctica, la prueba K-S no es tan útil, dado que generalmente la media y la desviación estándar de la población son desconocidas y deben calcularse a partir de una distribución teórica de comparación, haciendo que se acepten hipótesis nulas la mayoría de las veces. Por ello, la corrección de Lilliefors es útil, pues está tabulada para el caso más habitual donde se desconoce la media y la desviación estándar.

Grupo de herramientas	Herramienta	Descripción
Herramientas para la evaluación de la normalidad y significancia estadística	Prueba Shapiro-Wilk (Pedrosa <i>et al.</i> , 2015; Segnini, 2008):	Es uno de los tantos métodos empleados para verificar la normalidad de una muestra de datos; además, es uno de los más sencillos y con mayor potencia estadística para aplicar. Para que se cumpla su condición de uso, la muestra debe ser igual o menor a 50.
	Prueba exacta de Fisher (Amat, 2016; Pértiga y Pita, 2004):	Es el test exacto que se emplea para verificar la existencia de una asociación entre dos variables cualitativas, al observar si las proporciones de una variable son diferentes según el valor que adquiera la otra variable. Esta prueba se emplea cuando la muestra de estudio es pequeña y no se pueden cumplir las condiciones para aplicar una prueba Chi-cuadrado.
	Prueba t-student (Díaz-Pinzón, 2020; Díaz y Ríos, 2018; Sánchez, 2015):	Esta prueba permite establecer si dos variables aleatorias normales y con igual varianza, tienen medias diferentes. De su resultado, se puede inferir que ambas muestras poseen distribuciones de probabilidad de media poblacional distinta, o bien, que la diferencia entre ellas se debe a oscilaciones estadísticas al azar. Para ello, en su metodología se propone la comprobación previa de la normalidad de ambas muestras, y con ello, proceder a calcular el p-valor (el cual se emplea para la determinación de la significación estadística en una prueba de hipótesis) de la prueba t.
Herramientas para la reducción de variables	Método de extracción de componentes principales (Peña, 2002):	El análisis de componentes principales busca representar de manera adecuada la información proveniente de n observaciones y p variables, a través de un número menor de variables construidas a partir de combinaciones lineales de las originales. Dentro de la utilidad de esta técnica, se encuentra que representa, de forma óptima, en un espacio dimensionalmente pequeño, las observaciones de un espacio general de mayores observaciones, pues es este el primer paso de la identificación de variables latentes o no observadas que son las que ocasionan la variabilidad de los datos; además, las variables originales correlacionadas, son transformadas en otras variables no correlacionadas, lo cual facilita la interpretación de los datos.
Herramientas para la bondad de ajuste	Rotación Varimax (F-Jardon y Martos, 2011; García y Media, 2015):	Corresponde a un método de rotación ortogonal empleado para minimizar el número de variables con cargas o saturaciones altas en cada factor, con el fin de mejorar y simplificar la interpretación de estos; asimismo, tiende a producir grupos de factores múltiples. Como resultado, en cada factor, las cargas altas se aumentan, mientras que las pequeñas se disminuyen, haciendo que cada factor resulte con un número pequeño de variables con grandes cargas.

Grupo de herramientas	Herramienta	Descripción
Herramientas para la bondad de ajuste	Estadístico chi-cuadrado (Mendivelso y Rodríguez, 2018):	Es una técnica o prueba estadística de libre distribución (no paramétrica) que se encarga de medir la discrepancia entre una distribución de frecuencias tanto observadas como esperadas. Los valores que puede tomar esta prueba oscilan de cero a infinito (no tiene valores negativos, ya que es la suma de valores elevados al cuadrado). Sus usos sirven para comprobar la bondad de ajuste de una variable, comprobar la independencia de dos variables y comprobar la homogeneidad de dos variables.
	Índice de ajuste comparativo (Lara, 2014; Martínez <i>et al.</i> , 2012):	Este índice relativo de amplio uso y buen comportamiento, se basa en la comparación de la discrepancia de la matriz de covarianzas predicha por un modelo, frente a una matriz de covarianzas observadas, con el fin de evaluar el grado de pérdida producido en el ajuste al cambiar del modelo planteado al modelo observado. Sus valores oscilan entre 0 y 1, y, por convención, requiere un valor mínimo de 0,9 para poder indicar que al menos el 90 % de la covarianza de los datos puede reproducirse por determinado modelo.
	Índice de bondad de ajuste de parsimonia (Salgado, 2009):	Se basa en el índice de bondad de ajuste (el cual indica la variabilidad explicada por un modelo, tomando valores entre 0 y 1, con 0,9 como un valor aceptable), pero considerando la parsimonia del modelo estimado, siendo esta referida al grado de ajuste para cada uno de los coeficientes o parámetros que allí se estiman. En este índice es adecuado obtener valores altos.
	El coeficiente de determinación R^2 (Laguna, s.f.):	Es una de las medidas más importantes en la bondad de ajuste, ya que este coeficiente indica el grado de ajuste de una recta de regresión, a los valores de una muestra; definiéndose como el porcentaje de variabilidad total de la variable dependiente (Y) que se puede explicar mediante la recta de la regresión. Este coeficiente es adimensional y puede tomar valores entre 0 y 1, acercándose cada vez más a este último valor cuando el ajuste es bueno, en caso contrario, el coeficiente se acercará a 0.
Herramientas sobre correlación	Coefficiente de correlación de Pearson (Marín, s.f.; Restrepo y González, 2007):Z	El objetivo de este coeficiente simétrico es medir el grado de asociación, fuerza o relación lineal que existe entre dos variables aleatorias de carácter cuantitativo y que, a su vez, poseen una distribución normal bivariada conjunta. Cuando el coeficiente obtenido es positivo, se establece que existe una relación directa entre las variables; en cambio, si el coeficiente es negativo, se observa una relación inversa, y si el coeficiente es igual a cero, las variables son independientes entre sí.

Grupo de herramientas	Herramienta	Descripción
Herramientas sobre correlación	Coeficiente Rho de Spearman (Marín, s.f.; Restrepo y González, 2007):	Es un coeficiente no paramétrico que se emplea de forma alternativa al coeficiente de correlación de Pearson, cuando este último no cumple algunos supuestos para su uso (por ejemplo, que las variables sean ordinales o si se incumple el supuesto de normalidad). Dentro de sus ventajas se encuentra que la técnica está libre de distribución probabilística al ser no paramétrica; además, los supuestos son menos estrictos. Este coeficiente es el coeficiente de Pearson con una transformación posterior de las puntuaciones originales en rangos.
	Estadístico de Durbin Watson (Catalán, s.f.; Ramírez, s.f.):	El estadístico de Durbin Watson se emplea en la identificación de autocorrelación de una serie de datos. Esta prueba depende del número de observaciones, número de parámetros, inclusión de variables rezagadas en el modelo, nivel de significancia, entre otros. Este estadístico puede tomar valores entre 0 y 4. Esta prueba es válida cuando las variables que se incluyen en la ecuación son exógenas.
Herramientas para la validez de constructos	Validez convergente y discriminante (Shuttleworth, s.f.):	Estos subgrupos de la validez de constructo, prueban que los constructos que se espera que se relacionen lo estén (validez convergente); o bien, que los constructos que no deben tener ninguna relación entre ellos, efectivamente no la tengan (validez discriminante). Este tipo de herramientas son empleadas especialmente en psicología, ciencias sociales, medicina y educación donde se investigan conceptos no cuantitativos o abstractos.
	Análisis factorial confirmatorio (Morata-Ramírez <i>et al.</i> , 2015):	Es un método que se emplea para la obtención de evidencias sobre la validez de un constructo, al brindar información acerca de la estructura interna de los instrumentos de medida. De forma general, este análisis comienza con la especificación e identificación del modelo; luego la recolección de datos, estimación de los parámetros y evaluación del ajuste al modelo. Con un ajuste adecuado del modelo, este puede utilizarse para la evaluación e interpretación de los parámetros; de otro modo, el modelo se debe modificar e iniciar el proceso de nuevo.
Otras herramientas	Regresión lineal múltiple (Montero-Granados, 2016):	Es una herramienta que permite ajustar modelos lineales o linealizables entre una variable dependiente (variable endógena) y varias variables independientes (variables exógenas), individualizando el fenómeno observado.

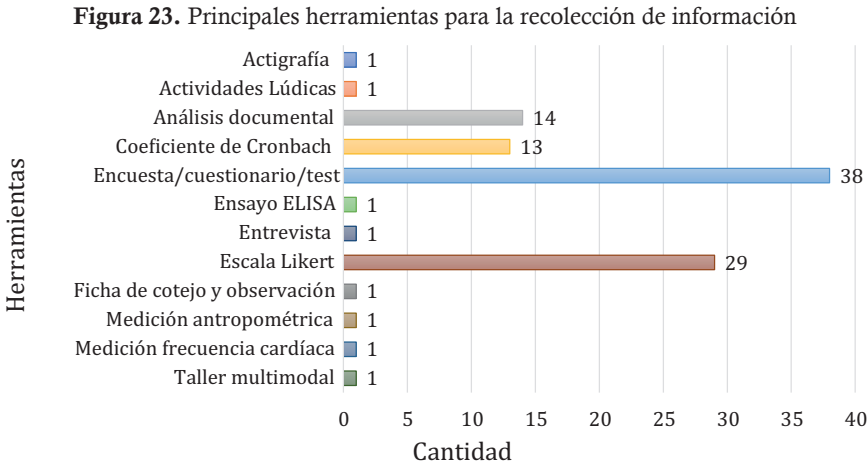
Grupo de herramientas	Herramienta	Descripción
	Modelamiento de ecuaciones estructurales (Manzano,2017):	Las ecuaciones estructurales son una herramienta estadística empleada para el estudio de las relaciones que existen entre las variables latentes (no se pueden medir de forma directa) y las observadas (medida de manera directa). Esta técnica permite determinar el efecto indirecto y total de una variable sobre otra. El planteamiento de estas ecuaciones estructurales inicia con una representación gráfica de las variables, asignando las variables observadas con cuadrados, y las latentes con un círculo o elipse, las cuales se asocian mediante una flecha unidireccional, o se correlacionan a través de una flecha bidireccional.
Otras herramientas	Análisis documental (Dulzaides y Molina, 2004):	Corresponde a una forma de investigación técnica a través de operaciones intelectuales, como describir, clasificar, indizar, anotar, extraer, traducir, y confeccionar una serie de documentos de forma sistemática para facilitar su recuperación.
	Diagrama de caja de bigotes:	Constituyen una representación visual que describe, resume y analiza múltiples características generales y particulares de variables cuantitativas de forma simultánea, por ejemplo, la dispersión y la simetría. En este tipo de gráficos se evidencian los datos centrales, datos adyacentes y datos atípicos o extremos; específicamente, se construye basado en cinco medidas estadísticas: valor mínimo, valor máximo, mediana, primer cuartil y tercer cuartil de los datos.

Fuente: elaboración propia.

9.5.2. Herramientas usadas para estudiar el rendimiento académico universitario

En cada uno de los artículos fue necesario emplear múltiples herramientas con el fin de lograr los objetivos propuestos de cada uno de los autores. Estas comprenden tanto las herramientas con las cuales se recolectó información, así como aquellas que permitieron su análisis, y, por ende, los hallazgos obtenidos. En la figura 23 se puede apreciar que la herramienta de mayor uso (en 38 artículos) corresponde a “encuestas/cuestionario/test” como forma de recolección de información (por ejemplo, conocimientos, habilidades, pensamientos y datos demográficos). Junto al uso de cuestionarios, se evidencia la “Escala Likert”, la cual, según los investigadores, resulta útil para captar información sobre sentimientos, pensamientos y percepciones de

cada uno de los individuos por medio de puntuaciones determinadas. Otra herramienta importante para la recolección de información la constituye el análisis documental, el cual consiste en una revisión exhaustiva de bases de datos y documentos. Por otra parte, el coeficiente de Cronbach también fue utilizado para la recolección de información, ya que permite validar la consistencia interna y confiabilidad de los instrumentos como cuestionarios.

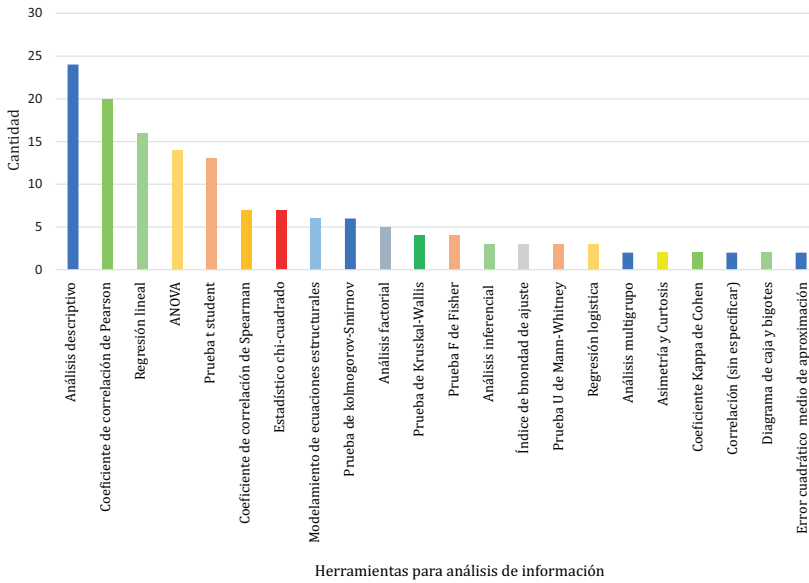


Fuente: elaboración propia.

En cuanto a las herramientas para el análisis de la información, en la figura 24 el “Análisis descriptivo” se constituye en el primer lugar de frecuencia de uso, por ser una herramienta estadística que engloba muchos indicadores y cálculos como el promedio, la desviación estándar, los porcentajes, las tablas de frecuencia, entre otros. Otra herramienta empleada en este ámbito fue el “Coeficiente de correlación de Pearson”, que permitió establecer relaciones y efectos entre las variables de estudio. En el análisis también fue posible encontrar el uso de regresiones lineales, el ANOVA (unidireccional o bidireccional) y la prueba t-student.

De cada una de las categorías en las que se dividieron las herramientas (recolección o análisis de información), se seleccionó la de mayor uso, para mostrar su comportamiento a través de los años. Lo anterior se presenta en la figura 25. Se observa que, en el 2020, se emplearon más herramientas, seguido del 2016.

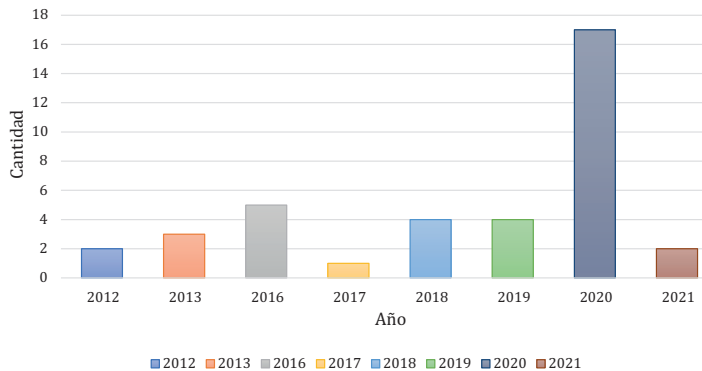
Figura 24. Herramientas usadas para el análisis de la información en la evaluación del rendimiento académico, según la revisión bibliográfica

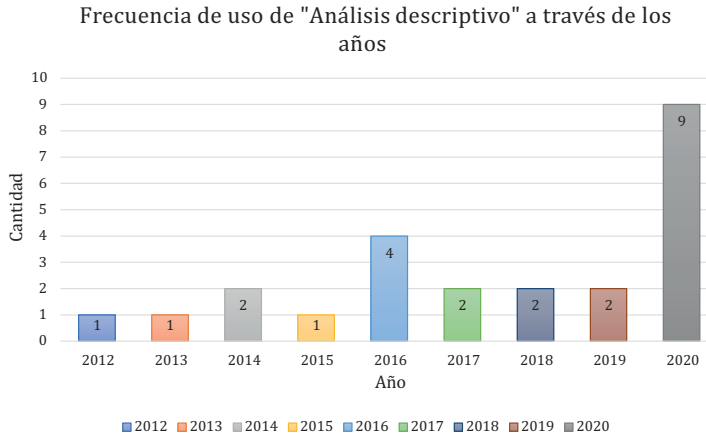


Fuente: elaboración propia.

Figura 25. (Arriba.) Trabajos que utilizaron la herramienta “encuestas/cuestionarios/test” en sus investigaciones a lo largo de los años. (Abajo.) Trabajos que usaron la herramienta “Análisis descriptivo” en sus investigaciones a lo largo de los años

Frecuencia de uso de "Encuesta/cuestionario/test" a través de los años





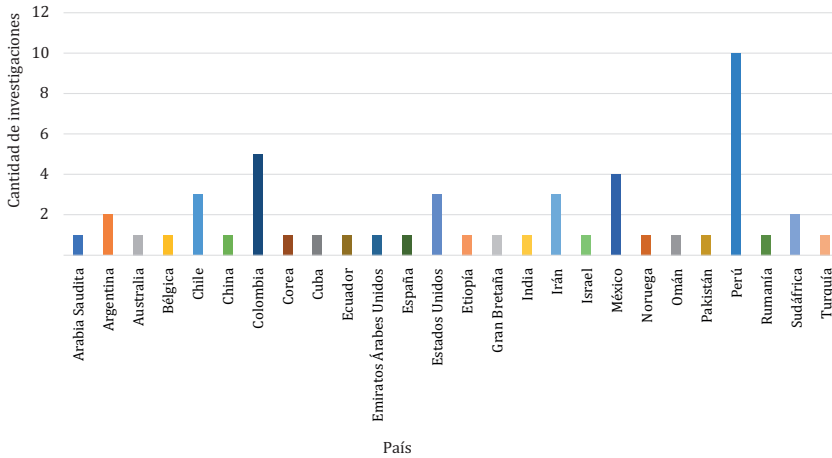
Fuente: elaboración propia.

9.6. Construcción social del rendimiento académico en educación superior

El análisis se puede centrar respecto a los países en donde más estudios se llevaron a cabo para la medición del rendimiento académico, así como los años de publicación de los mismos, tanto desde la forma general como de forma categorizada por las áreas de conocimiento planteadas. En este sentido, la figura 26 muestra la cantidad de trabajos socializados. Teniendo en cuenta las diversas metodologías suministradas por la literatura, se observa claramente que Perú es el lugar en el que más se hicieron este tipo de estudios para la medición del rendimiento académico. Le sigue Colombia, México, y posteriormente Chile, Estados Unidos e Irán, lo que permite inferir el amplio interés que existe en Latinoamérica (especialmente en los países de la Alianza del Pacífico) por investigar sobre este tema y sus posibles consecuencias en el mejoramiento del sistema de educación superior de la región.

En cuanto al comportamiento por categorías, a continuación se evidencian cuáles son los países que, según la investigación realizada en las bases de datos, han elaborado más trabajos publicados respecto a la evaluación del rendimiento académico para cada una de las tres grandes áreas establecidas para clasificar la información consultada.

Figura 26. Investigaciones sobre evaluación del rendimiento académico por países, según la investigación referencial



Fuente: elaboración propia.

En la categoría “Psicología”, teniendo en cuenta los países latinoamericanos en donde más se han desarrollado metodologías para la medición del rendimiento académico, el primer lugar lo ocupa Perú, seguido de Colombia, México y Chile en orden descendente; es decir, un comportamiento similar al registrado a nivel general. Con respecto a las fechas de publicación, el 2020 es aún el año con mayor cantidad de registros, seguido del 2016 y 2018.

De manera similar para los trabajos agrupados en la categoría “Estadística-Matemática”, es Irán el país donde se presenta con más frecuencia las metodologías para la medición del rendimiento académico, seguido de Estados Unidos. En cuanto a la categoría “Medicina”, es posible encontrar que Sudáfrica es el que más trabajos presenta sobre la metodología para la medición del rendimiento académico desde esta área de conocimiento.

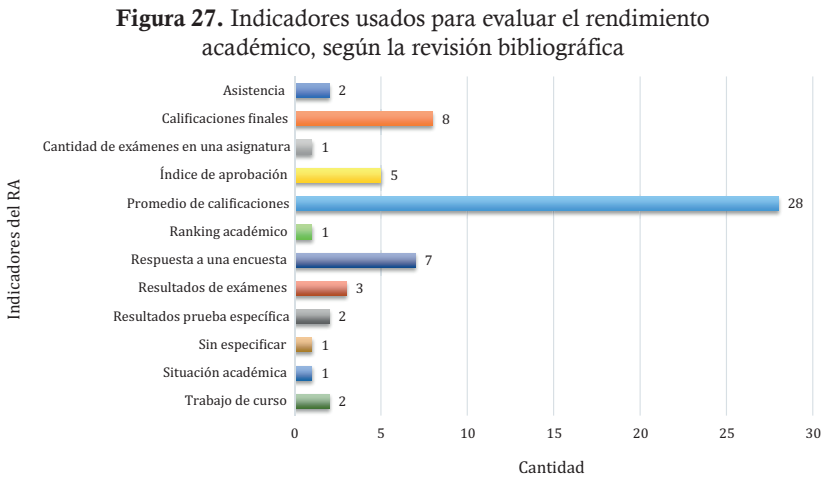
9.7. Indicadores de medición del rendimiento académico

En cada uno de los trabajos revisados y seleccionados, se emplearon distintos indicadores para la medición del rendimiento académico, según los objetivos que pretendía conseguir cada uno de los autores. En este sentido, es posible observar que las herramientas más usadas en orden decreciente son: “promedio de calificaciones” (con un uso predilecto de más del 50%), que varió entre las calificaciones de una sola materia, de varias materias o

cursos, promedios acumulados o ponderados, promedio de los exámenes, promedios en diversos periodos, entre otros. A esta herramienta, le siguen las “calificaciones finales”, bien sea de cada materia, por cortes en el semestre, o en una actividad específica. Posteriormente, le sigue “respuesta a una encuesta”, que implica una autopercepción del rendimiento por parte de los estudiantes, el índice de aprobación (número de materias aprobadas o no aprobadas sobre un total), resultados de exámenes, entre otros (figura 27). Cabe aclarar que se muestra una mayor cantidad de herramientas frente al número de artículos revisados; puesto que, en algunos de ellos, los autores usaban dos o tres de las herramientas para consolidar la variable “rendimiento académico”.

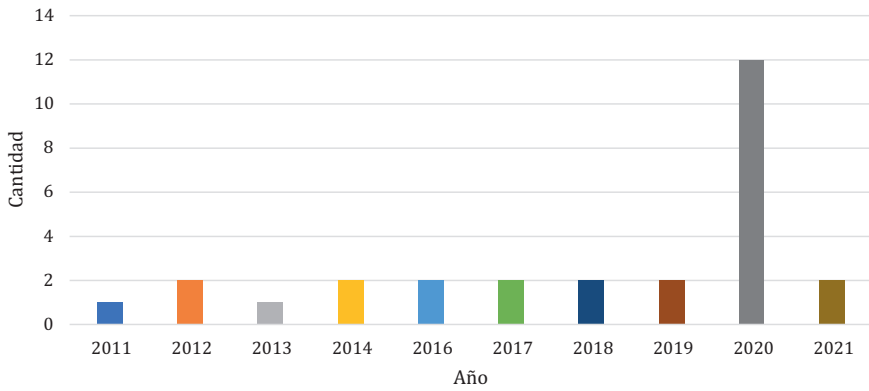
A partir de la figura 27, puede afirmarse que el “promedio de calificaciones” es la herramienta más utilizada en los campos mencionados para la evaluación del rendimiento académico a nivel universitario, hecho que se refleja en la figura 28, donde se observa que el 2020 es el año en el que más se ha empleado.

El segundo y tercer indicador de medición de rendimiento académico corresponde a las “calificaciones finales” y a la “respuesta a una encuesta”, respectivamente. El comportamiento de estos indicadores puede apreciarse en la figura 29. En ella, se observa que el 2020 fue uno de los años en donde más se emplearon los indicadores.



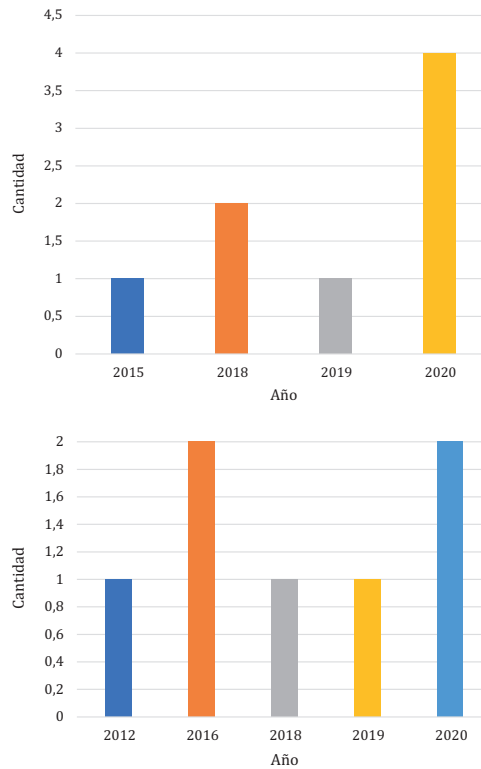
Fuente: elaboración propia.

Figura 28. Trabajos que utilizaron el indicador “promedio de calificaciones” en sus investigaciones a lo largo de los años



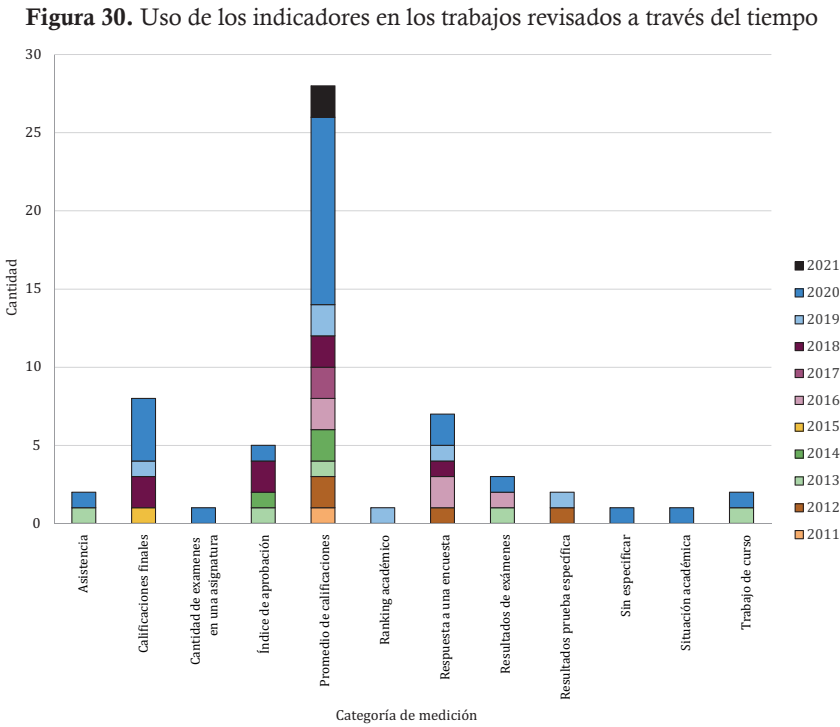
Fuente: elaboración propia.

Figura 29. (Arriba) Trabajos que utilizaron el indicador “calificaciones finales” en sus investigaciones a lo largo de los años. (Abajo) Trabajos que usaron el indicador “respuesta a una encuesta” en sus investigaciones a lo largo de los años



Fuente: elaboración propia.

También se emplearon los siguientes indicadores: avance de carrera, notas de exámenes, así como el promedio de los mismos, índices de estudio, las actividades desarrolladas durante todo un curso, número de asignaturas reprobadas o el índice de aprobación, promedio ponderado acumulado, notas de pruebas estatales, *ranking* de ingreso a la universidad, entre otros. En la figura 30, es posible evidenciar la totalidad de los indicadores encontrados; se relaciona la frecuencia de uso de cada uno de ellos, así como el año en el que se empleó.



Fuente: elaboración propia.

9.8. Relación de variables con el rendimiento académico

Diversas variables, según los autores consultados, pueden influir en el rendimiento académico de un estudiante y llevarlo al éxito o al fracaso académico; estas pueden agruparse en factores académicos, demográficos, socioeconómicos, de aprendizaje en línea, de gestión académica, psicosocial y de entorno académico. De ahí que se estén planteando diversas

investigaciones, con el fin de establecer cuáles podrían ser las variables más relevantes o importantes para una población objetivo y cuáles podrían ser las herramientas adecuadas para determinarlo. Los resultados de estas metodologías varían entre sí, aun teniendo en cuenta la teoría consultada por los autores. Sin embargo, es posible mencionar los hallazgos que relacionan algunos de estos estudios, los cuales se efectuaron a partir de poblaciones con diferentes características y tamaños, así como los múltiples planteamientos para las variables que incidían en el rendimiento académico. Algunas de estas se muestran en la tabla 12.

Tabla 12. Variables que afectan el rendimiento académico

Factor	Variables
Factores académicos	Puntaje global ICFES (Colombia). Año de presentación de prueba ICFES (< 2014 o > 2014). Puntaje ICFES para área de Matemáticas. Puntaje ICFES para área de Física. Puntaje ICFES para área de Química. Puntaje ICFES para área de Lenguaje. Puntaje ICFES para área de Biología. Puntaje ICFES para área de Ciencias Sociales. Puntaje ICFES para área de Filosofía. Puntaje ICFES para área de Ciencias Naturales. Puntaje ICFES para área de Lectura Crítica. Puntaje ICFES para área de Sociales y Ciudadanas. Puntaje ICFES para el área de Inglés.
Factores demográficos	Edad. Género. Estado civil. Tipo de colegio donde culminó la educación secundaria. Experiencia laboral. Tiempo transcurrido antes de ingresar a la universidad. Carácter del bachillerato.
Factores sociales	Nivel educativo de los padres. Número de miembros de la familia. Ocupación de los padres.
Factores socioeconómicos	Lugar de procedencia. Región de procedencia. Ingresos familiares. Número de autos. Casa propia o arrendada. Tipo de inscripción. Estrato.
Factores de gestión académica	Becas obtenidas. Año de ingreso a la universidad pénsum. Número de créditos. Número de créditos cursados. Número de créditos aprobados. Nota final para cada asignatura. Número de asignaturas cursadas. Número de asignaturas perdidas. Número de asignaturas ganadas. Número de veces que ha perdido una asignatura. Número de asignaturas repetidas. Número de asignaturas canceladas.
Factores tecnológicos	Posibilidades de acceso a internet. Número de accesos a LMS. Tiempo de uso total de LMS. Interacciones en línea. Número de recursos accedidos. Número de documentos descargados. Tiempo por actividad. Frecuencia de uso. Coordenadas x y y del mouse. Tiempo de clic en cada actividad. Cantidad de clics. Documentos accedidos. Participación en <i>quices</i> . Tiempo que el estudiante demora para resolver una pregunta. Participación en foros.
Factores de biblioteca	Frecuencia de acceso. Tipos de área de búsqueda. Número de libros en la biblioteca.

Factor	Variables
Factores institucionales	Tamaño del salón de clases. Organización y planificación del personal. Horarios de los cursos. Tipo y tamaño de la institución. Procesos de funcionamiento. Infraestructura. Políticas.
Factores pedagógicos	Número de estudiantes por docente. Número de estudiantes matriculados. Número de repetidores. Métodos y materiales didácticos utilizados. Motivación de los estudiantes. Tiempo dedicado por los profesores a la preparación de sus clases. Formación académica del profesor. Años de experiencia del profesor. Evaluación docente. Tipo de docente. Área académica.
Factores intelectuales y afectivos	Concentración. Memoria. Comprensión verbal. Razonamiento. Fluidez verbal. Sociabilidad. Seguridad. Integración. Relación padre-hijo. Relación profesor-alumno

Fuente: adaptada de Oladejo (2010); Tomás *et al.* (2014); Estrada y Quintero (2015); García (2015).

A continuación, se revisan algunas de las variables y se justifica su posible inclusión en esta investigación a partir de los resultados obtenidos de investigaciones previas que las han contemplado.

- **Edad:** para el caso de estudiantes universitarios, Page *et al.* (1990) señalan que la edad puede incidir en un rendimiento académico diferencial. En la categoría de la psicología, en el estudio de Espinosa *et al.* (2020) se encontró que la edad influye inversa e indirectamente en el rendimiento académico de los estudiantes de psicología y trabajo social colombianos, pues a mayor edad, se adquieren mayores responsabilidades; lo que ocasiona mayores niveles de estrés que finalmente se traduce en un rendimiento más bajo. Estos resultados concuerdan con la investigación de Awadalla *et al.* (2020), en la que se encontró que un menor rendimiento académico está relacionado con una mayor edad. De forma opuesta ocurre con lo planteado en la investigación de Bonsaksen (2016), desde el área de la estadística con estudiantes de terapia ocupacional de Noruega; pues, aunque se encontró que la edad también influye en el rendimiento académico, lo hace de forma directa: una mayor edad representa un mejor rendimiento académico. Otros trabajos que afirman la influencia positiva de esta variable en el rendimiento académico de los estudiantes son los de Hamoud *et al.* (2018) y Adekitan y Noma-Osaghae (2019). Lo anterior evidencia que tampoco hay un consenso claro sobre la influencia de la edad en el rendimiento académico, o de su relación con otros

factores, pues esta comparación se realizó entre estudiantes de diferentes nacionalidades y carreras.

- **Relación profesor-alumno:** según Hernández (2016), el papel de los profesores en el rendimiento académico de estudiantes de ingeniería y administración en México, es demasiado importante, incluso sobre los elementos de evaluación o la motivación extrínseca; lo que se correlaciona con los hallazgos de Mushtaq y Khan (2012), en donde se comprobó que la orientación adecuada por parte de padres y docentes tiene influencia en el rendimiento académico de los estudiantes de universidades privadas de Pakistán. De manera similar, Hakyemez y Mardikyan (2021) determinaron en sus hipótesis de investigación que, de forma indirecta, la integración de los estudiantes de negocios de Turquía con los docentes desempeña un papel importante sobre el rendimiento académico. Jain *et al.* (2016), en India, obtuvieron resultados similares en su estudio, para una muestra de estudiantes de negocio, en la cual encontraron un alto impacto de los programas de tutorías docentes en el rendimiento académico de los estudiantes, incluso una mayor influencia que la que tiene la autoeficacia profesional sobre la variable de interés. En la investigación de Castro y Closas (2020), el papel del docente también fue ampliamente valorado en los procesos de enseñanza-aprendizaje.
- **Género:** desde el ámbito de la psicología, Lin *et al.* (2020), al emplear una muestra de estudiantes de odontología de China, encontraron que la variable género posee un efecto sobre el rendimiento académico, pues las mujeres de dicha muestra tuvieron un mejor rendimiento académico; además, establecieron que las diferencias entre géneros son significativas en términos estadísticos. A esto mismo, llegaron Deliens *et al.* (2013) y Mazumder *et al.* (2020), pues en cuanto al género, las mujeres obtuvieron mejores calificaciones y una participación más activa. Esta posición se contrapone a los hallazgos de Furnham *et al.* (2013), puesto que en esta metodología se encuentra que las mujeres tuvieron un desempeño superior en el indicador de rendimiento académico; sin embargo, las diferencias no son significativas, lo que significa que sus efectos son invariantes sobre el rendimiento académico de los estudiantes de Gran Bretaña.

Por su parte, las investigaciones efectuadas desde la estadística, también evidencian que las mujeres muestran un mejor rendimiento académico, tal como lo enuncian Huamán (2018) en su investigación con estudiantes de estomatología del Perú y Shahrakipour *et al.* (2016) con estudiantes de carreras de medicina de Irán. Aunque en el estudio de Liang *et al.* (2018) con estudiantes de Estados Unidos, el rendimiento académico de las mujeres también fue mayor, se concluye que este efecto sobre la variable de interés es demasiado bajo como para afectarla realmente. Sin embargo, hay pocos estudios, como los de Almarabeh (2017) y Garg (2018), que muestran lo contrario, poca influencia de la variable género en el rendimiento académico.

- **Inteligencia emocional:** la investigación hecha por Fernández-Lasarte *et al.* (2019), midió el impacto del apoyo social percibido y la inteligencia emocional en el rendimiento académico. Frente al apoyo social, medido a través del apoyo del profesorado, del apoyo familiar y del apoyo de las amistades, fue el apoyo del profesorado la variable más significativa sobre el rendimiento académico de los estudiantes de muestra, provenientes de España. Lo mismo encontraron Guizado *et al.* (2020) en su investigación, la cual, por medio de varios ámbitos del desempeño docente, demostró una correlación positiva y significativa entre esta variable y el rendimiento académico de los estudiantes de Perú. La inteligencia emocional incide en el rendimiento académico por medio de algunos de sus factores (Fernández-Lasarte *et al.*, 2019). Como parte de la inteligencia emocional, Hayat *et al.* (2020) encontraron que tener y expresar emociones positivas frente a la academia también genera un efecto positivo en el rendimiento académico. Vargas *et al.* (2011) también hallaron que un pensamiento creativo puede predecir el rendimiento académico.
- **Comprensión verbal:** en cuanto al papel del docente desde la categoría “Estadística-Matemática”, Durán y Rosado (2019) investigaron una temática específica: la comprensión lectora y habilidades relacionadas con el lenguaje en estudiantes de ingeniería colombianos; allí se encontró que los docentes no implementan estrategias adecuadas para el desarrollo de

las clases (como incentivar la lectura o argumentación) y por ello se presentan resultados deficientes en el rendimiento académico.

- **Variables psicológicas:** el estrés también fue una variable que se examinó en varios de los estudios encontrados, específicamente en los de psicología. Mushtaq y Khan (2012), en su estudio, encontraron que el estrés familiar ocasiona parte de la variación negativa del rendimiento académico, pero que no es significativo en términos estadísticos. En contraparte, y como se mencionó anteriormente, Espinosa *et al.* (2020) indican que el estrés tiene una influencia no solo en el rendimiento académico, sino también en otros aspectos de la vida cotidiana; a estos resultados son los mismos a los que llegaron con el estudio de Lin *et al.* (2020), donde se enuncia que el estrés es estadísticamente significativo e inverso con el rendimiento académico. De manera similar ocurre con la ansiedad, que es un tema específico del área de la psicología: por ejemplo, una rama de la ansiedad (ansiedad matemática), estudiada por Ortiz *et al.* (2020) para el contexto colombiano de los estudiantes de ingeniería; se encontró que no existe relación entre esta y el rendimiento académico, lo que está acorde con los trabajos previos de otros autores. Awadalla *et al.* (2020) estudiaron tanto la depresión como la ansiedad, y concluyeron que ambas afectan de manera negativa el rendimiento académico de los estudiantes, mediado por el género (pues las mujeres tienden a sufrir más de estos trastornos). Otras variables que muestran influencia en el rendimiento, según algunos autores, son el interés del estudiante, la conducta hacia el estudio y la motivación (Garg, 2018; Hamoud *et al.*, 2018; Mueen *et al.*, 2016). Finalmente, en el estudio de Montoya-Arenas *et al.* (2021) se halló que la capacidad intelectual y de la función ejecutiva se relacionan fuertemente con el rendimiento académico.
- **Lugar de procedencia:** los estudios, como el de Cresswell y Underwood (2004) y Ortlieb y Cheek (2008), demuestran que existe una influencia del lugar de procedencia en la forma como aprenden los estudiantes; es decir, los estudiantes en diversas ubicaciones geográficas tienen conocimientos específicos, experiencias previas y modos de vida diferentes guiadas de distintas maneras por los docentes para satisfacer las necesidades educativas.

Sin embargo, existen investigaciones como las de Porto y Gresia (2005) y Garzón *et al.* (2010) que arrojan que no hay un efecto significativo.

- **Autoeficacia profesional:** como en otras variables, también se encuentran resultados contradictorios en las investigaciones de Jain *et al.* (2016), Hayat *et al.* (2020) y Hakyemez y Mardikyan (2021). En el primer estudio, aunque exista relación positiva entre las variables, aún sigue débil. En los siguientes dos estudios, se plantea que sí existe relación entre esta variable y el rendimiento académico.
- **Fluidez verbal:** según Mushtaq y Khan (2012), existe una relación entre las habilidades comunicativas de los estudiantes y el rendimiento académico. Mientras que, con la variable de asertividad, en el estudio de Luján *et al.* (2019) se concluyó que no hay relación entre esta y el rendimiento académico de los estudiantes de ciencias administrativas de Perú. En el estudio de Zárate *et al.* (2020) se demostró que la competencia comunicativa no interfiere de gran manera en el rendimiento académico, pues la relación entre estas es positiva pero débil.
- **Sociabilidad:** en el estudio de Furnham *et al.* (2013) se plantea que algunas características de la personalidad influyen tanto de forma positiva (conciencia, amabilidad y franqueza) como de manera negativa (la extraversión y el neuroticismo) en el rendimiento académico. Para el aspecto de la conciencia, Icekson *et al.* (2020) concuerdan con estos resultados, pues encontraron una relación positiva entre esta y el rendimiento académico, mediado por el género (favoreciendo a las mujeres).
- **Nivel educativo de los padres:** existe una relación positiva entre los niveles educativos de los padres y el rendimiento académico, según se plantea en el estudio de Shahrakipour *et al.* (2016). Esto también concuerda con lo encontrado por Awadalla *et al.* (2020).
- **Examen previo a la universidad:** según Vargas *et al.* (2019), existe una relación directa y significativa entre el puntaje de ingreso a la universidad y el rendimiento académico, por lo que la primera variable puede predecir el comportamiento de la otra mediante un modelo matemático lineal. Esto es afín con lo encontrado por García-Hernández y González-Ramírez

(2020), pues en su investigación se enuncia que la prueba de ingreso en matemáticas constituye un fuerte predictor. También vale la pena citar los trabajos de Grob *et al.* (2015), Ahmad *et al.* (2015) y Aluko *et al.* (2018), quienes concluyen que los datos de preadmisión, que son los resultados de las pruebas de acceso a la universidad, influyen en el rendimiento académico a nivel superior. Al contrario, Farrokhi-Khajeh-Pasha *et al.* (2012) indican que el examen de admisión (en este caso llamado Konkoor en Irán) es un predictor pobre del rendimiento académico de los estudiantes.

- **Datos preuniversitarios:** autores como Farrokhi-Khajeh-Pasha *et al.* (2012), Deliens *et al.* (2013), Grob *et al.* (2015), Anuradha y Velmurugan (2015), Mohamed y Waguih (2015), Asif *et al.* (2017), Garg (2018) y Pavelea y Moldovan (2020) concluyen que los datos preuniversitarios que incluyen resultados de secundaria ayudan a comprender el desempeño de los estudiantes.
- **Tipo de colegio de procedencia:** Muñoz-Comonfort *et al.* (2014) y Grob *et al.* (2015) concuerdan en sus resultados en los cuales indican que el rendimiento académico se ve afectado por el tipo de colegio en el que se estudió.
- **Rendimiento previo:** para Vargas *et al.* (2011), Yassein *et al.* (2017), Sivasakthi (2017) y Putpuek *et al.* (2018), la variable predictora más fuerte del rendimiento académico es el rendimiento previo (consiste en calificaciones ya obtenidas por los estudiantes desde que ingresaron a la universidad). De tal manera que permita determinar a las instituciones asignar una escala de valoración a los estudiantes, que facilite el cálculo del rendimiento académico como el “promedio de calificaciones” (GPA) o el “promedio de calificaciones acumulativo” (CGPA) (Parker *et al.*, 2004). De forma similar son los resultados obtenidos por Shahiri *et al.* (2015), quienes enfatizan que las calificaciones de la evaluación interna y CGPA son los factores más comunes utilizados para predecir el desempeño de los estudiantes utilizando la analítica. Para Sanguinetti *et al.* (2013), este rendimiento previo que afecta el rendimiento académico actual, implica la aprobación de una asignatura anterior, acorde con lo que encontraron Bautista-Rodríguez y Gatica-Lara (2020), en el cual el

rendimiento académico se ve afectado por aquellas materias que no se han aprobado.

- **Estado civil:** el estudio de Bonsaksen (2016) plantea que vivir con una pareja o compañero, y haber tenido experiencia previa en la educación superior, influye positivamente en el rendimiento académico.
- **Infraestructura institucional:** algunas variables relacionadas con el entorno del estudiante en la institución universitaria son importantes para la determinación del rendimiento académico, según diversos autores. El tipo de programa que se analiza muestra una influencia positiva, según Mohamed y Waguih (2015) y Hamoud *et al.* (2018); el tipo de modalidad de los estudios tiene influencia en el rendimiento académico, pues según la investigación de Ñaupari (2014), se evidencia que aquellos en modalidad presencial evidencian un mejor rendimiento académico que aquellos en modalidad virtual; el periodo académico semestral en el que ingresa y se analiza el rendimiento de los estudiantes influye también en el mismo (Mesarić y Šebalj, 2016). En cuanto al salón de clases, según Mushtaq y Khan (2012), las instalaciones del aprendizaje inciden positivamente en el rendimiento académico.
- **Variables tecnológicas:** en este grupo se encuentran las variables relacionadas con las actividades *on-line* propias de la metodología *E-Learning* o *B-Learning*, como la cantidad de horas de inicio de sesión, la cantidad de entradas en el tablero de discusión, la cantidad/tiempo total de material visto que, según Hamoud *et al.* (2018), tienen poca influencia en el rendimiento académico del estudiante. Sin embargo, para Castro y Closas (2020), los materiales (específicamente relacionados con las TIC y su uso en una modalidad semipresencial) sí son ampliamente apreciados por los estudiantes como un factor de incidencia sobre su rendimiento académico.
- **Conductas alimentarias:** en el estudio de Zuniga-Jara *et al.* (2018), se encontró que la sensibilidad por recompensa afecta de forma indirecta el rendimiento académico; es decir, esto influye en que un estudiante pueda estar alimentándose en exceso (alimentos dulces y grasos), lo cual afecta su índice de masa corporal (IMC), y por ende su rendimiento académico.

La investigación de Deliens *et al.* (2013) respalda lo anterior, indica que un mayor IMC, junto con un mayor peso y contorno de cintura, afectan de manera negativa el rendimiento académico de un estudiante.

- **Uso de alcohol, cigarrillo y otras sustancias:** Deliens *et al.* (2013) indican que un mayor consumo de alcohol afecta de manera negativa el rendimiento de los estudiantes. Bahammam *et al.* (2012), por su parte, indican que fumar determina y afecta de forma negativa el rendimiento académico.
- **Hábitos de sueño:** respecto a los hábitos de sueño, Bahammam *et al.* (2012), en su estudio, encontraron que aquellos estudiantes que tienen mejor sueño durante la noche, son capaces de evidenciar un mejor rendimiento académico. El estudio de Salmani *et al.* (2020) concuerda con lo anterior, pues los resultados señalan que un bajo rendimiento está relacionado con los trastornos del sueño, y mediado por el sexo (son las mujeres las más afectadas).
- **Técnicas y materiales de estudio:** García-Hernández y González-Ramírez (2020) encontraron que los materiales de estudio (especialmente los orientados a aumentar la motivación de los estudiantes) afectan de manera positiva el rendimiento académico. En el estudio de Albela-García *et al.* (2020), los resultados obtenidos indicaron que las técnicas de aprendizaje (como las tutorías y seminarios) también demuestran un efecto positivo sobre el rendimiento académico. En el estudio de Nicholson *et al.* (2016), se soportan los resultados de los estudios anteriores, ya que se encontró que emplear un taller interactivo y multimodal como preparación a un examen de anatomía, influyó de forma positiva en los resultados que los estudiantes pudieron lograr.
- **Asistencia:** Pavelea y Moldovan (2020) afirman que la asistencia a clase es un factor de relevancia en el rendimiento académico. Esto concuerda con lo encontrado en el estudio de Céspedes *et al.* (2018), donde el presentismo y ausentismo (junto con sus factores de medición) afectan tanto positiva como negativamente al rendimiento académico de los estudiantes.

- **Avance de carrera:** según Mazumder *et al.* (2020), Awadalla *et al.* (2020) y Bautista-Rodríguez y Gatica-Lara (2020), existe una relación entre el año de estudio (también llamado generación de estudio, año o semestre de estudio) y el rendimiento académico.
- **Hábito de lectura:** Sanguinetti *et al.* (2013) encontraron una relación positiva entre los hábitos de lectura y el rendimiento académico.

De los estudios revisados, también se pueden citar algunas variables no tan comunes en la mayoría de los trabajos y que, aunque pueden influir según la literatura, no lo hacen significativamente para el análisis de rendimiento académico a nivel universitario en la modalidad presencial. Se presentan los siguientes resultados:

- Para Chilca (2017), los hábitos de estudio se correlacionan moderadamente con el rendimiento académico, mientras que la autoestima no lo hace.
- En las hipótesis planteadas y comprobadas por Hakyemez y Mardikyan (2021), se encuentra que el desarrollo académico e intelectual incide directamente en el rendimiento académico.
- Hayat *et al.* (2020) hallaron una relación entre las estrategias de aprendizaje metacognitivo y el rendimiento académico.
- Para Icekson *et al.* (2020), el optimismo disposicional se relaciona con el rendimiento académico.
- En el estudio de Thomas y Viljoen (2019), la potencia total (una media de la variación de la frecuencia cardíaca) genera diferencias entre grupos de alto y bajo rendimiento, y se encuentra mediado por el género.
- Para Pavelea y Moldovan (2020), la intervención en clase está relacionada con el rendimiento académico.
- El uso de juegos entre semana afecta de manera negativa el rendimiento académico (Deliens *et al.*, 2013).
- Minichil *et al.* (2020) encontraron que aquellas estudiantes que perciben un rendimiento académico más bajo debido al dolor menstrual son más propensas a desarrollar el trastorno disfórico premenstrual.

9.9. Conclusiones del capítulo

Con base en la información consultada y analizada, es posible plantear las siguientes conclusiones principales:

En general, los anteriores hallazgos obtenidos por medio de la categoría “Psicología”, dejan en evidencia que los estudiantes ven en los docentes una figura importante en su propio rendimiento académico, no solo desde la parte de la enseñanza, sino también desde el apoyo que puedan recibir por un docente, lo cual es más notorio en sus propias percepciones, así como en los indicadores de rendimiento académico.

Las metodologías consideran la variable “docente” en su investigación, la encuentran relevante en el rendimiento académico, caso contrario con lo que ocurre con la variable “género”. En este sentido, aunque en todas las metodologías mencionadas anteriormente, el rendimiento académico de las mujeres fue superior, aún existe discrepancia de si hay o no diferencias significativas como para considerarla una variable de influencia.

El factor más influyente observado en la literatura, es el rendimiento académico previo. Los datos relacionados, es decir, los datos preuniversitarios o universitarios, se pueden recuperar fácilmente del Sistema de Información del Estudiante (SIS) de la universidad, tan ampliamente utilizado en la actualidad. El SIS también puede proporcionar algunos datos demográficos de los estudiantes (por ejemplo, edad, género, etnia), pero el estatus socioeconómico puede no estar disponible explícitamente.

La predicción temprana del desempeño de los estudiantes puede ayudar a las universidades, posibilitándoles brindar acciones oportunas, como planificar la capacitación adecuada para mejorar su tasa de éxito. La exploración de datos educativos ciertamente puede ayudar a lograr los objetivos educativos deseados. Al aplicar técnicas de minería de datos educativos (EDM), es posible desarrollar modelos de predicción para mejorar el éxito de los estudiantes. Sin embargo, usar técnicas de minería de datos puede resultar abrumador y desafiante para personas sin conocimientos técnicos.

Se ha investigado una variedad de factores en la literatura relacionados con su impacto en la predicción del éxito académico de los estudiantes, el cual se midió como rendimiento académico, ya que nuestra investigación mostró

que el rendimiento académico previo, la demografía de los estudiantes, la actividad de *E-Learning*, los atributos psicológicos, son los factores clave. En términos de técnicas de predicción, se han aplicado muchos algoritmos para predecir el éxito de los estudiantes bajo la técnica de clasificación.

En el área de medicina, no fue posible encontrar gran cantidad de literatura sobre el tema, lo que deja en evidencia que es un campo aún inexplorado con amplio potencial para futuras investigaciones.

Capítulo 10

Resumen de investigaciones sobre analítica académica, analítica del aprendizaje y metodologías para determinar el rendimiento académico

A continuación, se describe cada uno de los trabajos de los últimos años según la consulta realizada en las bases de datos mencionadas. Estos se encuentran divididos por periodo. Cada grupo de trabajos comprende: año, métodos utilizados, metodología y resultados, entre otros aspectos.

10.1. Analítica del aprendizaje, periodo 2015-2020

Benavides y Pérez Montero (2015) utilizan la metodología *B-Learning* como estrategia de mejora del proceso de aprendizaje de los estudiantes de primer semestre de Ingeniería Industrial de la Corporación Universitaria de Huila (CORHUILA), los cuales tienen notas inferiores o iguales a 3,5. Se tomó una muestra de 77 estudiantes, distribuidos en tres grupos de manera predeterminada por la Institución. Los grupos estudiados son los siguientes: a) ambiente *B-Learning* sin estrategia didáctica autorregulada y sin objetos para el aprendizaje; b) ambiente *B-Learning* con estrategia didáctica autorregulada y con objetos para el aprendizaje conceptual; y c) ambiente *B-Learning* con estrategia didáctica autorregulada y con objetos para el aprendizaje procedimental. La variable dependiente se refiere al logro académico alcanzado por los estudiantes, medido a través de los resultados obtenidos en una prueba de tipo texto de la temática del curso. Los resultados arrojaron que no existían

diferencias significativas en la construcción de estructuras de decisión, cuando se estudiaron los grupos de estudiantes en ambientes *B-Learning* con estrategia didáctica autorregulada y con objetos para el aprendizaje conceptual, en contraste con el grupo que trabajó objetos para el aprendizaje procedimental. Hubo un incremento de los valores de las medias, evidenciándose un aumento en el aprendizaje, situación que propicia y deja entrever la necesidad de continuar investigando sobre estrategias autorreguladoras y aplicaciones de las TIC en la formación.

Back *et al.* (2015) investigaron acerca de la implementación del *E-Learning* en escuelas de medicina. Se usó un cuestionario en línea (49 ítems) enviado a 48 facultades de medicina en Austria (n = 4), Alemania (n = 36) y Suiza (n = 8) utilizando SurveyMonkey® para el periodo febrero y septiembre de 2013. Se analizaron los datos de manera estadística. Las personas de contacto de la encuesta son miembros del grupo de trabajo Nuevos Medios de la Sociedad para la Educación Médica y miembros de la facultad en los decanatos de asuntos estudiantiles responsables del aprendizaje electrónico. Los resultados arrojaron que las plataformas de aprendizaje en línea son utilizadas por el 97% de las escuelas, usando el personal de *E-Learning* en 50%. El 81% de las escuelas ofrecieron programas de formación para sus profesores; el 76% otorgó beneficios orientados al desempeño; y el 17% otorgó tareas de *E-Learning*. El 97% utilizaría herramientas de aprendizaje electrónico compatibles con el plan de estudios producidas por otras facultades. La tabla 13 resume dichos resultados.

Tabla 13. Parte de la encuesta realizada en el estudio

Pregunta	Sí (%)
Información general sobre <i>E-Learning</i> en las facultades de medicina	
¿Existen herramientas de <i>E-Learning</i> para la educación de los estudiantes en su escuela de medicina?	100
- Predominantemente obligatorio	9,4
- Predominantemente opcional	90,6
¿Existe un conjunto específico de recomendaciones para la aplicación del aprendizaje electrónico por parte de su escuela de medicina, del Consejo Académico o del decanato?	43,8

Pregunta	Sí (%)
¿Estipulan el uso de un código de garantía de calidad para sus actividades de <i>E-Learning</i> ?	
- En caso afirmativo, ¿estos criterios de calidad están diseñados para cumplir con los estándares estipulados por la Asociación Médica Alemana o similar?	48,4
	20,0
¿Existe una estrategia para toda la facultad para aumentar el alcance y la calidad de las herramientas de aprendizaje electrónico en los próximos años en su escuela de medicina?	58,1
¿Tiene en cuenta criterios de género a la hora de desarrollar actividades de <i>E-Learning</i> ?	39,3
Desea más apoyo de sociedades disciplinarias relevantes (GMA/GMDS, etc.) e instituciones académicas o agencias públicas (DFG/BMBF/ Ministerios de Estado) para el desarrollo de actividades de E-Learning?	75,9

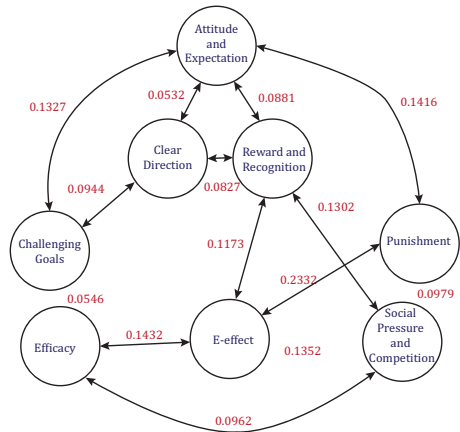
Fuente: adaptada de Back *et al.* (2015).

Law (2015) presenta una investigación exploratoria a través de la analítica de datos (regresión) que determina las influencias de los diversos factores de motivación del aprendizaje entre sí, así como sobre el aprendizaje en línea de un grupo de estudiantes del curso de programación de computadores. Se llevó a cabo una encuesta con 19 preguntas que permitieron identificar los factores que tienen un efecto motivador en el aprendizaje, y cinco adicionales para conocer el efecto electrónico y la eficacia percibidos por los estudiantes. En el estudio, se usaron las siguientes variables: factores intrínsecos (las actitudes, los objetivos y las emociones individuales) y factores extrínsecos (definición clara, recompensa, castigo, presión social). Las relaciones entre los factores y su influencia en el *E-Learning* se muestran mediante el gráfico de conectividad dirigida, aplicado al curso de programación informática. El número en cada borde representa la distancia semántica entre las construcciones asociadas directamente conectadas, obtenidas con el método de agrupamiento de gráficos jerárquicos (figura 31).

Giannakos *et al.* (2015) indagaron sobre los patrones e intereses de los estudiantes en un curso apoyado por videoconferencias. La muestra del estudio consistió en estudiantes de ciencias que se inscribieron en un curso de lectura de pregrado. Los instructores desarrollaron un programa de conferencias en video para ayudar a los estudiantes a prepararse antes de la clase. Las videoconferencias se integraron con las evaluaciones apropiadas,

que consistieron en preguntas que pudieron responderse en función del contenido del video. Las respectivas videoconferencias (con las evaluaciones integradas) se publicaron en la *wiki* del curso para su uso antes de la clase. Once estudiantes de primer año (18-20 años, seis mujeres y cinco hombres) participaron en el estudio. Los objetivos del aprendizaje del curso estaban relacionados con el uso de la información para la resolución de problemas, la investigación y la toma de decisiones. El curso duró cerca de diez semanas, mejorado con video-conferenciantes para ayudar a los estudiantes durante siete semanas (las primeras dos semanas y la última no requirieron videoconferencias).

Figura 31. Relación de factores en aplicación *E-Learning*



Fuente: tomada de Law (2015).

Se recopilaban tres tipos de datos; por tanto, se utilizó un análisis de datos apropiado para cada conjunto diferente de datos. De acuerdo con la navegación de video de los estudiantes, usaron la visualización de series de tiempo agregadas, con el fin de identificar los picos de las visualizaciones de video de los estudiantes (máximos globales). Posteriormente, investigaron cualquier relación potencial entre las opiniones de los estudiantes y el rendimiento del aprendizaje en las respectivas evaluaciones basadas en videos. Al final del análisis, observaron el pico global concerniente de cada video e intentaron dar una explicación de por qué esos segmentos son tan importantes para los estudiantes. Con la visualización de la actividad de los estudiantes mediante

gráficos, llegaron a la conclusión de que existe una relación positiva entre los puntajes de desempeño de aprendizaje de los estudiantes y sus opiniones sobre la respectiva evaluación de video.

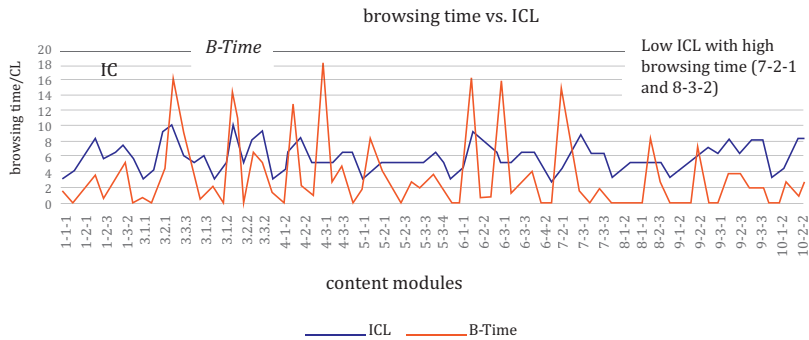
Jo *et al.* (2015) investigaron el proceso de construcción de variables *proxy* (variable aproximada usada en la predicción de fenómenos cuando la variable real no es medible o no es confiable) con un sistema de aprendizaje *on-line* para adultos mayores mediante una regresión lineal múltiple. Los 200 participantes trabajaban en el campo de los negocios financieros y para el estudio se reclutaron a partir de un curso de aprendizaje electrónico comercial titulado “crédito derivado” en Corea. El curso tenía como objetivo proporcionar a los estudiantes conocimientos sobre productos financieros (leyes, estrategia para una inversión segura, entre otros). Se esperaba que los alumnos que completaran el curso estuvieran preparados para la consulta con el cliente.

Se consideraron las siguientes variables: el tiempo total de inicio de sesión, la frecuencia de inicio de sesión, la regularidad del intervalo de inicio de sesión. La variable dependiente, el rendimiento en el aprendizaje (puntuación obtenida en la prueba final del curso). Los resultados mostraron una correlación de Pearson negativa entre la regularidad del intervalo de aprendizaje y la puntuación final de la prueba, definida como “rendimiento de aprendizaje”, así como también de esta última con la frecuencia de inicio de sesión. Por último, se analizó la regresión lineal múltiple, en la cual se encontró que las tres variables sugeridas explican el 34,7% de la varianza en el rendimiento del aprendizaje ($F = 36,267$, $p < 0,01$). De las tres variables, solo la regularidad del intervalo de aprendizaje resultó ser un predictor significativo del rendimiento del aprendizaje ($\beta = -0,590$, $t = -10,115$, $p < 0,01$).

Yen *et al.* (2015), en su trabajo, investigaron cómo se relacionan las características de los estudiantes en un Sistema de Manejo de Aprendizaje (LMS) con la carga cognitiva del curso “Aplicaciones y operaciones de teléfonos inteligentes” de la Universidad Nacional Abierta de Taiwán. Las redes sociales se utilizaron para mejorar la interacción y la presencia social de los estudiantes (adultos jóvenes). Se emplearon las siguientes variables: el tiempo que pasaron navegando por módulos de contenido específico, el tiempo de inicio de sesión y la cantidad de publicaciones que hicieron.

Entre los resultados importantes, está el de no tener una correlación significativa entre la carga cognitiva percibida por cada actividad. Esto quiere decir que los profesores y los estudiantes tienen una percepción diferente de la carga cognitiva de las tareas. En cuanto a la variable tiempo de navegación y la carga cognitiva, la prueba de correlación de Pearson indicó que sí existe una relación positiva y significativa (p - value = 0,268). En la figura 32 se muestran las curvas de tiempo de navegación promedio de los estudiantes por módulo y el índice de carga cognitiva determinada por los instructores para cada módulo; se puede notar que existen módulos, los cuales se calificaron con carga cognitiva baja, pero que requirieron más tiempo del pensado.

Figura 32. Correlación entre tiempo de navegación y carga cognitiva



Fuente: tomada de Yen *et al.* (2015).

La correlación entre las variables de densidad de publicaciones (DP) y la carga cognitiva (ICL) fue negativa y significativa ($p < 0,05$). La correlación entre DP y el número de visualizaciones (IV) fue significativa y positiva; y la correlación entre ICL y el IV no fue significativa (tabla 14).

Tabla 14. Matriz de correlación de los indicadores analizados

	DP	ICL	IV
DP	1		
ICL	-.261*	1	
IV	.276*	-.112	1

Nota: * $p < .05$.

Fuente: tomada de Yen *et al.* (2015).

Martin y Whitmer (2016) investigaron sobre el rendimiento de los estudiantes en cursos en línea y la inclusión de actividades cronometradas (cómo cambian los resultados de aprendizaje cuando se incorporan mecanismos de cronometraje en los cursos). Se contemplaron las siguientes variables: inicio de sesión (cantidad de veces que el estudiante ha iniciado sesión en el curso durante el semestre), tiempo total invertido (tiempo total que el estudiante ha estado en el curso durante el semestre), tiempo medio por sesión (número de elementos de contenido a los que accede el estudiante en el curso durante el semestre), diferencia de acceso al módulo (diferencia entre la fecha programada de ese elemento en el programa de estudios y la fecha en que el estudiante accedió por primera vez al elemento). La tabla 15 muestra datos descriptivos de los 39 estudiantes inscritos hasta el final del periodo. Los inicios de sesión del curso tienen un rango y una desviación estándar grandes, lo que muestra que algunos estudiantes inician sesión con mayor frecuencia que otros, con una distribución que no sigue una curva de campana normal. Esta variable tiene un rango relativamente bajo y una pequeña desviación estándar, lo que indica que la mayoría de los estudiantes accedieron a los materiales del curso (44,92 de un total de 52 ítems). La diferencia de acceso total tiene un rango alto, lo que se espera dada la condición experimental que afecta cuando los estudiantes pueden acceder a los materiales del curso.

Tabla 15. Iteración con materiales el curso *on-line*

Variable	Observations	Mean	SD	Min.	Max.
Total logins	39	203,85	77,66	75,00	460,00
Total items	39	44,92	4,09	28,00	52,00
Total student time (h)	39	97,23	34,49	36,16	171,86
Total access difference	39	38,59	30,14	-53,54	94,59

Fuente: tomada de Martin y Whitmer (2016).

Kim *et al.* (2016) plantearon su investigación sobre el panel de análisis de aprendizaje (LAD) y su efectividad sobre el aprendizaje en línea. Se dividió la población de 151 estudiantes que estaban tomando el curso en línea titulado “Estadísticas de gestión” en el 2014 en dos grupos. Las variables

estudiadas son las siguientes: frecuencia de uso del tablero, satisfacción del panel y logro de aprendizaje. Como resultado, se obtuvo, en primer lugar, que los estudiantes que recibieron el tratamiento del uso del tablero presentaron una puntuación final más alta que los que no lo recibieron. En segundo lugar, la frecuencia de uso del tablero, medida por la cantidad de veces que se abrió el tablero, no tuvo un impacto significativo en el aprendizaje (logro). Sin embargo, se observó una correlación ligeramente positiva entre la satisfacción con el LAD y el rendimiento del aprendizaje. Un análisis más detallado indicó que los alumnos que utilizaron el panel solo unas pocas veces mostraron una satisfacción relativamente alta con el LAD. Por otro lado, los estudiantes de alto rendimiento académico que abrieron el LAD con relativa frecuencia, mostraron una menor satisfacción con el tablero. Los resultados indican que el LAD debe revisarse de forma que motive a los alumnos de manera constante y apoye a los alumnos que tienen diferentes niveles de rendimiento académico.

Gibson y De Freitas (2016) estudiaron dos casos en su trabajo: el primero, se refiere al análisis de las acciones del alumno en una evaluación de rendimiento virtual basada en juegos creada en Harvard y analizada por uno de los autores. El segundo caso, se refiere a un estudio de retención longitudinal basado en cinco años de datos de desempeño de 52 000 estudiantes de una amplia gama de fuentes de información distribuidas. Estos dos casos destacan cuestiones clave que surgen al realizar estudios de analítica de aprendizaje en la educación superior y señalan la necesidad de una nueva forma de métodos mixtos (cualitativos y cuantitativos), específicamente soportados en la minería de datos y métodos de regresión (lineales y no lineales) aplicados a la analítica de aprendizaje.

De lo anterior, se extrajo un conjunto de datos analíticos (ADS) de 13 millones de elementos que sirvió como base para el modelo de aprendizaje automático semisupervisado de SOM (mapas autoorganizados de Kohonen), que se capacitó en 52 000 estudiantes con más de 250 atributos, incluidos algunos estudiantes en línea, a tiempo parcial y en el sitio. Los dos estudios de caso han demostrado que los datos de manera diferente se pueden utilizar para comprender mejor la experiencia de los alumnos. En el ejemplo del juego, se identificó cómo la experiencia de los individuos se puede modelar y

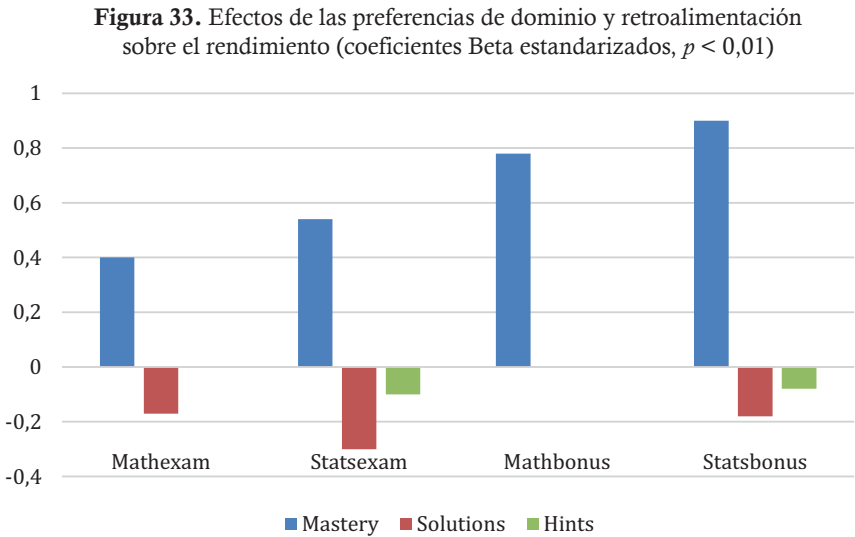
mapear de manera efectiva y se pueden proporcionar comentarios a lo largo de la experiencia. En el estudio de retención, la ventaja de utilizar datos de rendimiento, a lo largo del ciclo de vida del alumno, muestra cómo las intervenciones se pueden planificar e implementar mejor en los momentos adecuados durante el estudio.

Nguyen *et al.* (2016) investigaron cómo las disposiciones de aprendizaje y las preferencias de retroalimentación afectan el rendimiento de los estudiantes bajo un entorno combinado en línea (*M-Learning*). Consideraron 1069 estudiantes de un curso cuantitativo introductorio mixto (matemáticas y estadística) en una universidad pública en Países Bajos durante 2015 a 2016. En el curso investigado, los estudiantes trabajan en tutoriales electrónicos (un promedio de 10 horas en Sowiso y 25 horas en MyStatLab), lo que representa del 12,5% al 31% del tiempo disponible de 80 horas para aprender sobre ambos temas. Las variables corresponden a las medidas de desempeño en el examen, tanto de matemáticas (MathExam) como de estadística (StatsExam); a la bonificación agregada para ambos temas, que se basó en el rendimiento de los cuestionarios (MathBonus y StatsBonus); y al nivel de dominio alcanzado en los tutoriales de cada tema (MathMastery y StatsMastery).

La figura 33 resume la relación entre las preferencias de retroalimentación, según lo revelado por el comportamiento de los estudiantes. El número promedio de ejercicios terminados con éxito está fuertemente relacionado de manera positiva con todas las medidas de desempeño. La relación más fuerte resulta ser para el rendimiento en las pruebas, con MathBonus y StatsBonus con valores betas de 0,78 y 0,92, respectivamente, y algo menos fuerte para el rendimiento en los exámenes MathExam y StatsExam con valores betas de 0,40 y 0,53, respectivamente.

Salajegheh *et al.* (2016) investigaron sobre la efectividad del *E-Learning* para el desarrollo de habilidades de interpretación de rayos X en estudiantes de medicina preclínica de la Universidad Griffith (Brisbane, Australia). Las competencias en la interpretación clínica se evaluaron mediante la comparación de las puntuaciones anteriores y posteriores del curso denominado “intervención” durante un año. Los estudiantes estaban divididos en dos grupos: “grupo de intervención”, que recibió el curso de aprendizaje

electrónico, y el “grupo de control”, que no recibió *E-Learning*. A ambos grupos se les dio una charla introductoria sobre cómo interpretar una radiografía de tórax, seguida de un taller práctico en grupos pequeños, donde practicaron y desarrollaron sus habilidades mediante el estudio de diferentes radiografías de tórax. Los resultados de la tabla 16 muestran que los estudiantes del grupo de intervención mejoraron sus habilidades de interpretación de rayos X desde antes del curso de *E-Learning* (puntuación media del 57,9%) hasta después de la implementación (puntuación media del 70%); en términos estadísticos, esta mejora es significativa ($p < 0,05$). Estos resultados se sustentan en la tabla 16. Mientras que, para el otro grupo, no hubo una mejora significativa. De esta manera, una combinación de aprendizaje tradicional y aprendizaje electrónico puede ser más eficaz en las habilidades de interpretación radiológica de estudiantes de medicina.



Fuente: tomada de Nguyen *et al.* (2016).

Alves *et al.* (2017) realizaron un estudio cualitativo en el cual se buscaba comprobar si la frecuencia de acceso a un ambiente virtual de aprendizaje (AVA) tenía alguna relación con la media de las calificaciones, la asistencia a conferencias presenciales y las calificaciones aprobatorias de los estudiantes. La investigación tuvo los siguientes objetivos: valorar el grado de asociación entre las variables frecuencia de accesos al AVA y el número de unidades del

curso superadas, y analizar los resultados de aprendizaje de las unidades de curso superadas y la media de las clasificaciones según la frecuencia de accesos al AVA de la institución.

Tabla 16. Puntuaciones de interpretación de rayos X (presentado como porcentaje con +/- desviación estándar

	MCQ de preintervención	MCQ posterior a la intervención	Un año después de la intervención
Grupo de intervención	57,9 ± 9,7	70 ± 4,6	66,7 ± 5,8
Grupo de control	60,3 ± 9,6	N/A	N/A

Fuente: tomada de Salajegheh *et al.* (2016).

Se tomaron los datos de acceso de 2636 estudiantes de una institución (2015-2016), con los cuales se construyeron diez grupos independientes de estudiantes que pertenecen a las áreas de artes y humanidades, ciencias empresariales, deportes, educación, ingeniería y servicios sociales. Los resultados mostraron, en cuanto a la correlación de variables realizadas mediante el *software* SPSS, una correlación positiva; por ejemplo, se encontró una correlación de 0,709 entre las variables de cantidad de unidades de curso aprobadas y la asistencia a clases. También se halló que entre más alto sea el decil al cual pertenecen los datos (grupos 1 al 10), mayor es la media de las unidades aprobadas por los estudiantes. Para comparar si existían diferencias significativas entre las medias de cada grupo mostradas y teniendo en cuenta que no se conoce la distribución de los datos y la homogeneidad de la varianza, se utilizó la prueba Turkey, la cual comparó las medias de los grupos. Se puede inferir que los grupos compuestos por las asignaturas con mayor frecuencia de accesos al entorno virtual, obtienen mayores resultados en las unidades de curso que los grupos compuestos por las asignaturas con menor frecuencia de accesos al ambiente virtual.

Mamcenko y Kurilovas (2017) investigaron la analítica del aprendizaje como herramienta para promover la personalización del aprendizaje en entornos virtuales (Moodle). Tuvieron en cuenta una muestra de seis estudiantes. Se buscaba cuáles deberían ser las actividades de aprendizaje correctas para cada uno, teniendo en cuenta una clasificación previa que se realizó mediante el modelo de estilos de aprendizaje de Felder-Silverman

(FSLSM). Se consideraron las siguientes variables de la clasificación de los estudiantes en cuatro grupos objetivo:

- **Por tipo de información:** sensorial (SEN) frente al intuitivo (INT).
- **Por canal sensorial:** visual (VIS) frente al verbal (VER).
- **Por procesamiento de información:** activo (ACT) frente al reflexivo (REF).
- **Entendiendo:** secuencial (SEQ) frente al global (GLO).

El estudio también muestra evidencias de que los estudiantes aprenden de acuerdo con su estilo de aprendizaje; esto último, según el modelo de estilos de aprendizaje de Felder-Silverman (FSLSM). La tabla 17 sustenta la clasificación del estilo de cada uno de los seis estudiantes.

Tabla 17. Resultados de aprobación del curso

Nr.	Information Processing		Information Type		Sensorial Chanel		Understanding	
	Active	Reflective	Sensing	Intuitive	Visual	Verbal	Sequential	Global
1	72,7	27,3	54,5	45,5	72,7	27,3	54,5	45,5
2	45,5	54,5	27,3	72,7	100	0	18,2	81,8
3	72,7	27,3	36,4	63,6	36,4	63,6	45,5	54,5
4	72,7	27,3	63,6	36,4	27,3	72,7	45,5	54,5
5	63,6	36,4	72,7	27,3	54,5	45,5	45,5	54,5
6	27,3	72,2	72,2	27,3	90,9	9,1	72,7	27,3

Fuente: tomada de Mamcenko y Kurilovas (2017).

Ellis *et al.* (2017) dedicaron su estudio a determinar el valor de la analítica del aprendizaje para la calidad de la experiencia de aprendizaje del estudiante, y cómo los enfoques metodológicos combinados que utilizan evidencia observacional y de autoinforme pueden mejorar la variación cualitativa en el aprendizaje de los estudiantes. Se recopilaron 145 estudiantes en el curso de ingeniería (Australia). Las variables son: duración (tiempo que el alumno dedica a trabajar en una actividad), tablero (vistas de los estudiantes de un tablero con información sobre el compromiso con las actividades semanales), Col-Exp (vistas de los estudiantes de varias secciones de las notas del

curso), recurso (vistas de los estudiantes de cualquier página del curso), video (interacción con un video), MCQ (interacción con una pregunta de opción múltiple), VMCQ (interacción con preguntas de opción múltiple colocadas junto a un video), ejercicio (respuestas de los estudiantes a ejercicios de evaluación sumativa). Se realizaron dos modelos lineales de regresión múltiple: el primero incluyó solo el enfoque de aprendizaje superficial autoinformado, mientras que el segundo se obtuvo con la variable de autoinforme y las cinco variables observacionales que tienen correlaciones significativas con la puntuación final. Los resultados muestran que, en el primer modelo, la variable contribuyó significativamente al modelo de regresión. Sin embargo, el modelo representó solo el 9% de la variación en las notas finales. El segundo modelo explica el 34% de la variación. Esto significa que se explica un 25% adicional de variación en el rendimiento académico de los estudiantes cuando se incluyen las variables observacionales en el modelo.

Luo *et al.* (2017) aplicaron la metodología *B-Learning* a un curso de estadística médica de 119 estudiantes (China) para mejorar los resultados del aprendizaje y evaluar los factores de impacto del conocimiento, las actitudes y las prácticas (CAP) de los estudiantes al *E-Learning*. Los estudiantes experimentaron conferencias presenciales tradicionales, así como interacciones en línea basadas en la plataforma Moodle durante el curso. Se aplicó una encuesta que sumaba puntuaciones de diferentes aspectos (conocimiento, actitud y práctica). Se ajustó un modelo lineal mixto utilizando el procedimiento SAS PROC MIXED para analizar los factores de impacto. Las características demográficas de los estudiantes que participaron en el estudio se muestran en la tabla 18.

El carácter autopercebido, la residencia permanente registrada y las horas diarias en línea son factores de impacto significativos del conocimiento del *E-Learning*, aumentando la puntuación media en 6,45 puntos ($p < 0,0001$) durante la intervención. En cuanto a las actitudes de *E-Learning*, la reforma educativa, el número de comunidades, la edad de internet y las horas diarias en línea son factores de impacto significativos, aumentando la puntuación media en 5,10 puntos ($p = 0,0010$) durante la intervención. En cuanto al comportamiento del *E-Learning*, la reforma educativa y la alfabetización de

los padres son factores de impacto significativos, aumentando la puntuación media en 10,05 puntos ($p < 0,0001$) durante la intervención (tabla 19).

Tabla 18. Características demográficas de los estudiantes del curso

Variable	Clasificación	Número de personas (N = 116)
Género	Masculino	36 (31,03)
	Femenino	80 (68,97)
Número de comunidad	0	27 (23,28)
	1-2	77 (66,38)
	≥ 3	12 (10,34)
Carácter autopercebido	Tipo de introversión	16 (13,79)
	Tipo medio	79 (68,10)
	Tipo de extroversión	21 (18,10)
Residencia permanente registrada	Urbano	60 (51,72)
	Rural	56 (48,28)
Alfabetización de los padres	Universidad o superior	29 (25,00)
	Secundaria o preparatoria	70 (60,34)
	Primaria o inferior	17 (14,66)
Era de internet	Seis años o menos	55 (47,41)
	Siete a nueve años	33 (28,45)
	Diez años o más	28 (24,14)
Horas pasadas en línea por día	Menos de una hora	8 (6,90)
	De una a tres horas	78 (67,24)
	De tres a cinco horas	22 (18,97)
	Más de cinco horas	8 (6,90)

Fuente: tomada de Luo *et al.* (2017).

Tabla 19. Análisis multivariado de actitudes de *E-Learning*

Efecto	Clasificación	Estimar	t (P)
Interceptar		51,97	11,09 (<0,0001)
Reforma educativa	Antes de	-5,10	-3,4 (0,0010)
	Después	0,00	
Número de comunidad	0	-5,97	-1,63 (0,1046)
	1-2	-1,00	-0,31 (0,7554)
	≥ 3	0,00	
Era de internet	≤ 6 años	-7,00	-2,76 (0,0065)
	7-9 años	1,81	0,71 (0,4792)
	≥ 10 años	0,00	
Horas pasadas en línea por día	Menos de 1 hora	-2,39	-0,45 (0,6506)
	De 1 a 3 horas	5,83	1,49 (0,1386)
	De 3 a 5 horas	3,98	0,96 (0,3384)
	Más de 5 horas	0,00	

Fuente: tomada de Luo *et al.* (2017).

De Arco Paternina *et al.* (2017) estudiaron la factibilidad del *blended learning* en la especialización de Ingeniería de Procesos (Colombia). Para ello, se recolectó información por medio de encuestas virtuales realizadas a la población objetivo y se hizo el estudio de la viabilidad económica y financiera del proyecto. Se identificaron nueve agentes motivadores de decisión, al igual que la viabilidad financiera obtenida mediante indicadores como valor presente neto y tasa interna de retorno. Se realizó una investigación de mercado como soporte de decisión para implementar la plataforma básica de educación a distancia y, además, medir las variables involucradas que soportaría el sistema *B-Learning* propuesto. La muestra estuvo compuesta por 132 encuestados (nivel de confianza del 95 %). Para el análisis, se utilizó como método una matriz DOFA (debilidades-oportunidades-fortalezas y amenazas) con su respectiva escala (tabla 20).

Tabla 20. Matriz DOFA

Fortalezas	Oportunidades		Amenazas		
Única especialización semipresencial en Colombia en Ingeniería de Procesos de Refinación de Petróleo y Petroquímicos Básicos.	5	3	3	5	8
La Universidad San Buenaventura, seccional Cartagena, ya cuenta con la plataforma virtual en la cual se haría el montaje de los módulos de Especialización.	5	3	2	3	5
La Universidad San Buenaventura cuenta con la infraestructura física e informática para la realización de esta modalidad de Especialización.	5	3	2	3	5
La Universidad San Buenaventura cuenta con un sistema financiero sólido que posibilitaría diversas modalidades de pago de la Especialización.	4	4	2	3	5
El programa de Especialización ya cuenta en su versión presencial actual con prestigio en el mercado educativo de la región	4	3	1	2	3
La modalidad virtual posibilitaría la contratación de docentes prestigiosos y con alto nivel académico y mayor experiencia	1	1	1	1	2
	24	17	28	17	45

Fuente: tomada de De Arco Paternina *et al.* (2017).

Los hallazgos encontrados en la encuesta realizada se resumen en nueve aspectos importantes como base de evaluación financiera de *B-Learning* en la Especialización: competitividad laboral, deseos de superación, inversión en TIC para plataforma virtual, flexibilidad del método educativo, autodisciplina de estudio, confianza en el modelo educativo para asegurar el aprendizaje, frecuencia de sesiones presenciales, costo de la Especialización propuesta y plazo para iniciar el proyecto educativo. Concluyen que la oportunidad de negocio de implementación de un sistema de *B-Learning* en el programa de Especialización en Ingeniería de Procesos de Refinación de Petróleo y Petroquímicos básicos de la Universidad de San Buenaventura, tiene un alto potencial de éxito.

Gottipati y Shankararaman (2018) desarrollaron la herramienta de análisis de currículum (CAT) que genera puntuaciones sobre el currículum en dos dimensiones: niveles cognitivos y niveles de progresión, con el fin de mejorar el aprendizaje de los estudiantes. Se aplicó en un plan de estudios básico de pregrado: Licenciatura en Ciencias (Sistemas de Información) BSc (IS), ofrecida por la Escuela de Sistemas de Información (SIS), Singapore Management University (SMU). El estudio refleja que los cursos básicos se centran en “recordar” y “aplicar”. En contraste, los cursos avanzados del currículum se enfocan en el dominio mediante la “creación”. Asimismo, los cursos avanzados también enfatizan la comprensión y la aplicación. La precisión general del modelo fue de 74,69%.

Por su parte, Vesin *et al.* (2018) plantearon un estudio sobre el sistema de tutoría de programación (ProTuS) que proporciona contenido inteligente e interactivo, opciones de personalización, funciones adaptativas y análisis de aprendizaje como apoyo para los usuarios comprometidos en su aprendizaje. Los participantes docentes y estudiantes (66 que usaron ProTuS y cuestionario de usabilidad) están matriculados en el curso de tecnologías web. Los resultados arrojaron que el uso de ProTuS recibió la calificación general de 56,3 y 67,2 puntos por parte del grupo de docentes y estudiantes, respectivamente.

Wen *et al.* (2018) investigaron el proceso y las estrategias de 25 estudiantes de una universidad en Taiwán, en un juego de simulación para aprender física cinemática por medio del análisis de contenido. Los estudiantes contaban con un tiempo determinado para examinar el juego y construir un modelo sobre este. Luego, se comparó el modelo inicial, cuando el estudiante lo hacía de manera empírica con su modelo final al ya poseer conceptos de física. Se analizaron múltiples fuentes de datos, incluidos los modelos iniciales y revisados de los estudiantes, sus registros de actividad en el juego de simulación y el resultado de la prueba del modelo y la prueba anterior a la experiencia para lograr estos objetivos. Para entender cómo los estudiantes construyeron modelos para resolver el problema, los modelos inicial y revisado se categorizaron con base en los conceptos de física involucrados en el modelo. La categorización de los modelos se realizó por dos investigadores, y la concordancia

(Kappa) de la categorización fue de 0,83 para el modelo inicial y 0,73 para el modelo final.

Wang *et al.* (2018) crearon un modelo en su investigación en el cual segmentan el proceso de la clase en múltiples fases y construyen un modelo analítico de aprendizaje basado en HMM (modelo oculto de Markov), que puede analizar en detalle el estado de aprendizaje de los estudiantes, lo que admite el seguimiento continuo de cada estado de los estudiantes en un entorno de sala inteligente. En este estudio, el método de evaluación del aprendizaje de los estudiantes se realiza por medio de la interacción que tienen los estudiantes con sus tabletas, teniendo en cuenta unas probabilidades asociadas a la atención del estudiante en la clase, relacionadas con el grado de utilización de la tableta. Para dicho caso, se mostró una exactitud del 0,849, la cual es mayor a la del modelo de referencia (0,749). Por ello, se demuestra que las variables seleccionadas son adecuadas como métodos de medición del aprendizaje de los estudiantes. Se pudo comprobar que un estudiante que se encuentra en el estado TI (estudiantes que usan sus dispositivos Pad individuales para leer material didáctico o hacer ejercicios), cuando su maestro está en el estado NT (instructor que explica el conocimiento o discute con los estudiantes), es más probable que tenga un aprendizaje menor al resto.

Kew *et al.* (2018) buscaron mejorar el desempeño de los estudiantes (174) en su práctica de aprendizaje a través de *E-Learning*. En este sentido, la motivación es uno de los predictores importantes de la intención de utilizar el *E-Learning* que, a su vez, puede tener un impacto en la práctica de aprendizaje. Al conocer el nivel de motivación de los estudiantes, los instructores pueden proporcionar materiales didácticos adecuados en *E-Learning* para que se involucren más en las actividades. Se usaron dos tipos de instrumentos de recopilación de datos: cuestionario y entrevista. Al examinar el nivel de motivación, encontraron que el coeficiente Alfa (análisis) para los cuatro elementos sugeridos en el estudio es relativamente alto, lo que indica su influencia en el aprendizaje, según los autores (tabla 21).

Tabla 21. Puntuaciones del coeficiente Alfa para IMMS

Subescala	Número de elementos	Keller (1990)	Nivel Alfa en este estudio
Atención	12	.89	.80
Relevancia	9	.81	.84
Confianza	9	.90	.69
Satisfacción	6	.92	.83
Escala total	36	.96	.93

Fuente: tomada de Kew *et al.* (2018).

En la tabla 22 se refleja el resultado de la prueba de normalidad, que muestra que la significancia de estudiantes de pregrado, de maestría y de doctorado es 0,039, 0,814 y 0,682, respectivamente. Esto significa que no se distribuye normalmente. Por tanto, la prueba de Kruskal-Wallis se utiliza para examinar si existe una diferencia significativa entre estudiantes de pregrado, maestría y doctorado. También se comprueba que los estudiantes de doctorado (con una media de 4,18) tienen el mayor nivel de motivación. Luego le siguen los estudiantes de maestría (con una media de 3,87) y, por último, los de grado (con una media de 3,65).

Tabla 22. Pruebas de normalidad de Shapiro-Wilk y el nivel de motivación del estudiante de pregrado, maestría y doctorado

	Nivel de titulación	Shapiro-Wilk			Media
		Estadística	df	Sig.	
Total_IMMS	Licenciatura	,982	153	,039	3,65
	Maestro	,967	15	,814	3,87
	Doctor	,943	6	,682	4,18

Fuente: tomada de Kew *et al.* (2018).

Rajak *et al.* (2018) estudiaron la aceptación de los servicios de aprendizaje electrónico en las instituciones de educación superior en Brunéi (Isla de Borneo). Las variables independientes incluyen las características del profesor, el diseño de los contenidos de aprendizaje, los materiales didácticos y la

alegría; mientras que las variables intermedias son los beneficios percibidos y la facilidad de uso percibida. Por otro lado, la variable dependiente en esta investigación es la intención de utilizar el *E-Learning*. Se realizaron análisis de regresión múltiple para probar las hipótesis propuestas. Las variables definidas en esta investigación se describen en la tabla 23.

Tabla 23. Definiciones operativas de variables

Variable	Definición operacional
1. Características del profesor	El grado en que los profesores están dispuestos a echar una mano, se preocupan y cooperan con los estudiantes.
2. Material didáctico	El grado de adecuación de los materiales didácticos para los servicios de <i>E-Learning</i> .
3. Diseño de contenidos de aprendizaje	El grado de precisión y coherencia de los contenidos de aprendizaje transmitidos.
4. Alegría	El grado en que los estudiantes disfrutan los servicios de <i>E-Learning</i> .
5. Beneficios percibidos	El grado en que los estudiantes confían en los servicios de <i>E-Learning</i> impulsaría los resultados del aprendizaje.
6. Facilidad de uso percibida	El grado en que los estudiantes están seguros de que utilizar los servicios de <i>E-Learning</i> no supondrá ningún esfuerzo.
7. Intención de utilizar los servicios de aprendizaje electrónico	El grado en que los estudiantes anticipan participar en los servicios de aprendizaje electrónico.

Fuente: tomada de Rajak *et al.* (2018).

Los datos (94 encuestados) se analizaron con el SPSS 23.0. Se realizó un análisis de regresión múltiple para probar las hipótesis. Existe una relación positiva entre las características del profesor y los beneficios percibidos por los estudiantes de los servicios de *E-Learning*, $\beta = 0,270, p < 0,05$. Existe una relación positiva entre el diseño de contenidos de aprendizaje y la facilidad de uso percibida por los estudiantes, $\beta = 0,605, p < 0,05$. Existe una relación positiva entre la facilidad de uso percibida por los estudiantes y sus beneficios percibidos, $\beta = 0,451, p < 0,05$. Existe una relación positiva entre los beneficios percibidos por los estudiantes y su intención de utilizar los servicios de *E-Learning* $\beta = 0,274, p < 0,05$. También existe una relación positiva entre la facilidad de uso percibida por los estudiantes y su intención de utilizar los

servicios de *E-Learning*, $\beta = 0,244$, $p < 0,05$. Por último, no existe relación significativa entre las variables material didáctico y beneficios percibidos, y alegría e intención de utilizar los servicios de *E-Learning*, $\beta = 0,051$, $p > 0,05$ y $\beta = 0,181$, $p > 0,05$, respectivamente. La variable beneficios percibidos tiene el efecto más fuerte sobre la intención de utilizar los servicios de *E-Learning* como $\beta = 0,274$.

González-Yebra *et al.* (2018) en su proyecto tuvieron como objetivo el estudio del potencial de los entornos inmersivos 3D para la enseñanza-aprendizaje a distancia. En este sentido, se ha evaluado el uso de la plataforma *Second Life* para complementar y reforzar las clases presenciales de la asignatura diseño asistido por ordenador. El desarrollo del estudio se llevó a cabo durante tres fases diferenciadas: en la primera, fase preparatoria, se diseñó y planificó el procedimiento seguido en el desarrollo de la investigación experimental. En la segunda, fase exploratoria, se llevaron a cabo las diferentes sesiones programadas en el marco del proyecto y se procedió a la obtención de las valoraciones del alumnado. En la tercera y última, fase final, se aplicó y adaptó el conocido modelo de Kano, diseñado para clasificar tipos de atributos de calidad con diferente impacto en la satisfacción de las personas usuarias, con el fin de desarrollar y mejorar productos o servicios. Los estudiantes estaban divididos en tres grupos de trabajo. Los ítems para evaluar, se presentan en la tabla 24.

Tabla 24. Ítems para evaluar en la experiencia

Ítems	Atributos
Metodologías (M)	
M1	Ampliación de los recursos formativos (WebCT/Blackboard Learning System) de la asignatura DAO.
M2	La asignatura DAO se complementa con actividades interactivas en SL.
M3	Un tema teórico se imparte en SL, como alternativa a la docencia <i>face-to-face</i> .
M4	Un tema práctico se imparte en SL, como alternativa a la clase de prácticas.
M5	Desarrollo en SL de ejemplos de diseños CAD/CAE en tiempo real.
M6	Plus de innovación (desde el punto de vista docente) al desarrollo de la asignatura.

Ítems	Atributos
M7	Adquisición más fácil de los conocimientos técnicos, como alternativa a las clases magistrales.
M8	Refuerzo o ampliación de conocimientos sobre el Diseño Asistido por Ordenador.
M9	Estimulación para el desarrollo de habilidades relacionadas con la creatividad.
Interfaz (I)	
I1	Fácil navegación y exploración para seguir una clase de DAO.
I2	Buena calidad gráfica y realismo 3D de la Isla UAL-SEIS.
I3	Aspecto de comunicación agradable.
I4	Navegación y exploración en español.
I5	Requisito de un equipo informático y una buena conexión a internet.
I6	Parte de la asignatura se puede seguir desde casa, biblioteca, trabajo... (24h/24h).
Adaptibilidad (A)	
A1	Aumento del interés y la atención en el desarrollo de una asignatura en general.
A2	Aumento del interés por materias que emplean el paradigma <i>Lerning by doing</i> (asignaturas de alto contenido práctico, como DAO).
A3	Adaptación a su nivel de conocimientos informáticos.
A4	Necesidad de un buen nivel de percepción y visualización espacial.
A5	Estudiantes con mucha paciencia.
Propuestas (P)	
P1	Reuniones sin asistencia del profesorado para trabajar en grupos (por ejemplo, sesiones <i>brainstorming</i>).
P2	Mejora del realismo 3D de la Isla UAL-SEIS.

Fuente: tomada de González-Yebra *et al.* (2018).

Los atributos se clasificaron en seis categorías: atributos atractivos (A), atributos unidimensionales (O), atributos obligatorios (M), atributos indiferentes (i), atributos opuestos (R) y atributos cuestionables (Q). Basados en los resultados, se pudo establecer que la docencia, a través de entornos inmersivos 3D, supone un recurso docente útil, a la vez, innovador como complemento de un modelo de aprendizaje *B-Learning*. En conclusión, la

respuesta del alumnado participante con relación a la componente teórica de la asignatura ha sido ampliamente positiva, inclinándose por una docencia presencial en cuanto a la componente puramente práctica o aplicada.

Eze *et al.* (2018) estudiaron la adopción y utilización de las instalaciones de aprendizaje *E-Learning* por parte de profesores en una institución en Nigeria, utilizando 15 entrevistas semiestructuradas del personal académico (investigación descriptivo-cualitativa). La población de este estudio es el personal académico de tres facultades de la universidad. En general, los resultados indican que el 89% de los participantes están de acuerdo en que hay suficientes instalaciones de aprendizaje electrónico para su uso; las instalaciones son fáciles de usar y las facilidades ayudan a mejorar significativamente el aprendizaje. Sin embargo, el 72% de los participantes indicó que las actitudes de los usuarios, las instalaciones inadecuadas de internet y la formación inadecuada son los principales inhibidores. La universidad debería impartir formación con regularidad, especialmente cuando se emplean nuevos profesores, para asegurarse de que se adapten y la utilicen. Además, debería haber algún tipo de seguimiento semanal para garantizar que los profesores los utilicen en las clases en todo momento. También se identificaron los principales factores que inhiben el aprendizaje electrónico (véase tabla 25).

Según Núñez-Barriopedro *et al.* (2019), el principal objetivo de la investigación es analizar los condicionantes del proceso de enseñanza-aprendizaje de la universidad latinoamericana, para la adopción de los nuevos planteamientos educativos, como el B-Learning. El diseño de la investigación es mixto con triangulación recurrente. Se utilizó un cuestionario con escala de Likert aplicado a 50 profesores, 402 estudiantes y 6 directivos. Al realizar el clúster mediante el algoritmo K-means, para agrupar a los docentes a partir de sus paradigmas educativos, se encontró que la mayor cantidad de ellos se ubican en el grupo 2, profesores con perfil constructivista (33 sujetos, 68% del total); le siguen los profesores con tendencia cognitivista (grupo 3, con ocho individuos para el 16% de la muestra); luego continúan los docentes con tendencia conductista (grupo 4, cinco docentes, un 10% de los encuestados); en cuarto lugar están los profesores que emplean metodologías y modelos por competencias (grupo 1, tres profesores, para el 6%);

y en último lugar, con 2 %, pues solo un docente pertenece al grupo 5, está el modelo de enseñanza tradicional. Entre las principales conclusiones, se vislumbra la necesidad de fomentar el factor social del aprendizaje y darle mayor protagonismo al estudiante. Los recursos empleados son tradicionales y analógicos. Los docentes acusan cambios en su práctica docente, pero poco en la inclusión de las tecnologías de la información y la comunicación (TIC). Asimismo, los tres aspectos que más le cuesta cambiar a los docentes en su paso de la modalidad presencial a la B-Learning son: dejarle el protagonismo al estudiante, cambiar de la clase transmisiva al aprendizaje colaborativo, y de la evaluación sumativa a la formativa.

Tabla 25. Factores inhibidores de E-Learning

Códigos	Temas	Muestra en evidencias de apoyo
Factores inhibidores	Actitud de los usuarios	• “Actitud personal. Algunas personas no están acostumbradas y no quieren cambiar. La actitud de los estudiantes también. Les envías algunas cosas por correo o les das tareas a las que no acceden” (A1). • “La pereza, la actitud de ‘no me importa’ que muestran los estudiantes y algunos miembros del personal” (A7). • “No nos importa por qué la Universidad no busca nuestra opinión antes de comprar” (A8).
	Instalación de internet inadecuada	• “Fluctuación en internet” (A1). • “A veces el acceso a internet es deficiente, lo que dificulta el acceso a algunas plataformas de aprendizaje electrónico” (A2).
	Entrenamiento inadecuado	• “Falta de <i>know-how</i> y todavía somos nuevos en algunos” (A6). • “Falta de formación adecuada” (A7).

Fuente: tomada de Eze *et al.* (2018).

Guillén-Guerrero *et al.* (2019) aplicaron el cuestionario CEMEDEPU a una muestra de 23 profesores universitarios en su búsqueda de la influencia de la capacitación docente, en la aplicabilidad de cursos basados en B-Learning. Se utilizaron métodos mixtos (cuantitativos y cualitativos). Así, en el componente cuantitativo se usó un diseño preexperimental con comparación pretest y postest intragrupo. Para el componente cualitativo se utilizaron grupos focales como técnica de recogida de datos, con el fin de favorecer la explicación de los resultados. Para analizar si la capacitación docente en

metodologías centradas en el aprendizaje generó un cambio positivo en la metodología docente y evaluativa utilizada por los profesores, se comparó los puntajes del pretest y posttest de las escalas del cuestionario CEMEDEPU: “Modelo centrado en la enseñanza”, “Modelo centrado en el aprendizaje” y “Habilidades docentes”. Los resultados se pueden ver en la tabla 26.

Tabla 26. Prueba *t* de *student* de datos CEMEDEPU. Grupo completo de profesores

		M	D. T.	<i>t</i>	Signif.
Modelo centrado en la enseñanza	Pretest	2,7993	0,49999	-0,345	0,733
	Posttest	2,8294	0,49362		
Modelo centrado en el aprendizaje	Pretest	4,1111	0,45843	-2,632	0,015
	Posttest	4,3116	0,38224		
Habilidades docentes	Pretest	4,4261	0,37199	-3,411	0,003
	Posttest	4,6239	0,27298		

Fuente: tomada de Guillén-Guerrero *et al.* (2019).

Como se aprecia en los resultados, encontramos que las tres escalas aumentaron, pero el aumento en “Modelo centrado en la enseñanza” no es significativo a diferencia del “Modelo centrado en el aprendizaje” y “Habilidades docentes” que muestran cambios significativos en términos estadísticos. Estos resultados nos muestran que la capacitación que recibieron los docentes generó cambios, desde una orientación más centrada en la enseñanza hacia una centrada en el aprendizaje, así como una mejoría en las habilidades docentes.

Kim *et al.* (2018) se concentraron en analizar las percepciones de los estudiantes universitarios (muestra de 614 estudiantes) sobre el *E-Learning*, basándose en sus experiencias, y los roles mediadores del compromiso académico y la preparación digital en el contexto universitario de un entorno de *E-Learning* para el rendimiento académico (promedio de calificaciones). Se plantea un modelo de mínimos cuadrados parciales para desarrollar el estudio. Para analizar la estadística descriptiva (media, desviación estándar, análisis de correlaciones, prueba *t* independiente), se utilizó el *software* IBM SPSS 23. Para el modelo por mínimos cuadrados parciales, se empleó el

software SmartPLS 3.0. Los resultados son significativos, por la importancia del compromiso académico y la preparación digital de los estudiantes, como mediadores en sus percepciones del aprendizaje electrónico predichas por el rendimiento académico. Aunque los estudiantes percibieron positivamente las experiencias de *E-Learning* en el campus, deben tener fuertes habilidades digitales para realizar el trabajo académico y comprometerse con una participación activa en el contexto del aprendizaje académico en entornos universitarios de *E-Learning*.

Bote-Lorenzo *et al.* (2019) indagaron sobre la precisión de la predicción del resultado de aprendizaje de los estudiantes (muestra de 1031 alumnos), cuando estos se encuentran bajo un entorno MOOC (cursos en línea abiertos masivos). En este estudio, se buscó comparar dos modelos predictivos basados en regresión logística, uno de forma genérica y el otro diseñado por los instructores del curso, el cual es un enfoque conocido como “diseño del aprendizaje” (LD). Una universidad española ofreció el curso masivo, el cual enseñaba finanzas. El modelo impulsado por LD supera al modelo genérico, ya que incluye características más refinadas de la participación de los estudiantes en el curso. Como las intenciones de obtener el certificado final del curso, son los principales impulsores de los estudiantes MOOC activos que tienen la intención de completar tareas y asignaciones de aprendizaje. En ambos modelos, las características relacionadas con las actividades de las tareas de los estudiantes (envíos, visitas a las páginas de descripción de las tareas, descargas de los recursos PDF asociados con las tareas) son las más predictivas. De hecho, los autores sugieren que ciertas páginas de contenido eran más importantes para el aprendizaje de los estudiantes que otras. Estas diferencias no serían capturadas por una única característica genérica en el modelo genérico que acumule todas las páginas vistas de contenido.

Klein *et al.* (2019) estudiaron los incentivos tecnológicos y las barreras de adopción de la analítica del aprendizaje. Este trabajo de cohorte cualitativo tuvo en cuenta a 32 profesores de Estados Unidos. Los hallazgos indicaron que, si bien los profesores y los asesores conocen y usan las herramientas de la “analítica del aprendizaje” (LA), no lo hacen sin frustración y preocupación. Estos sentimientos están ligados a las barreras tecnológicas que existen en estas herramientas y en las infraestructuras tecnológicas que existen en

su campus, y han llevado a un uso desigual, desconfianza y métodos alternativos de práctica. Como se indicó en la discusión, el trabajo futuro debería centrarse en mejorar la integración de las herramientas de LA con otras tecnologías del campus y las necesidades y prácticas de los usuarios individuales; sobre la creación clara, precisa y visualizaciones de datos relevantes; e incluir las necesidades y perspectivas profesionales de los usuarios de profesores y asesores.

Azcona *et al.* (2019) investigaron cómo un conjunto de datos históricos, demográficos y sociales de los estudiantes permite a las instituciones comprender de mejor manera su aprendizaje. Muestra una nueva metodología para detectar automáticamente a los estudiantes “en riesgo” de reprobar una tarea en el curso de programación de computadoras. El modelo permitió realizar un análisis para clasificar a los estudiantes e identificar a aquellos que pueden necesitar ayuda para el aprendizaje y están en riesgo de reprobar su próxima evaluación formal. La variable final se dividió en dos clases: si su desempeño está por encima o por debajo del umbral de aprobación del módulo. Se utilizaron diferentes clasificadores para el modelo, como la regresión logística, las “máquinas de vectores de soporte” (SVM) con diferentes núcleos, K-vecinos, el árbol de decisión y el bosque aleatorio. Los mejores resultados se dieron con el clasificador K-vecinos ($K = 12$), el cual dio un alto desempeño con métricas para F1 y precisión.

Al-araibi, Mahrin *et al.* (2019) desarrollaron un modelo inicial para los aspectos tecnológicos (computadoras, internet, teléfonos móviles, audio-video, CD, DVD, videoconferencias, correos electrónicos, así como foros de discusión) de preparación para el *E-Learning* en instituciones de educación superior. La muestra la conformaron once expertos de diversas universidades del Medio Oriente con título de doctor (PhD), a los que se les analiza la información por la técnica denominada Delphi. En los resultados, se puede citar que el panel de expertos estuvo de acuerdo con los factores tecnológicos que se presentan en la tabla 27, que son los factores tecnológicos adecuados.

El estudio de Pham *et al.* (2019) examinó las relaciones entre los atributos de calidad, la satisfacción de los estudiantes y la lealtad de los estudiantes de *E-Learning* en Vietnam. Los datos de la encuesta de 60 aspectos que involucrados con confiabilidad, capacidad de respuesta, empatía, precisión,

facilidad de uso, seguridad, entre otros, se aplicaron a 1232 estudiantes universitarios. Se analizaron mediante el modelado de ecuaciones estructurales, utilizando SPSS 25 y SmartPLS 3.0. La calidad general del servicio de aprendizaje electrónico es un factor de segundo orden, con un coeficiente de regresión de la calidad del sistema de aprendizaje electrónico de 0,99; sobre el instructor de aprendizaje electrónico y la calidad de los materiales del curso 0,97; y sobre la calidad de los servicios administrativos y de apoyo del *E-Learning* 0,95. Todos los coeficientes de regresión en el modelo estructural son significativos en $p < 0,01$. La calidad general del servicio de aprendizaje electrónico explica el 98 % de la variación de la calidad del sistema de aprendizaje electrónico; 93,4 % de la varianza del instructor de *E-Learning* y la calidad de los materiales del curso; 89,5 % de la varianza de la calidad del servicio administrativo y de soporte de *E-Learning*.

Tabla 27. Estadística descriptiva sobre la relación de factores

No.	Factores tecnológicos	La relación entre factores tecnológicos	Media de R2	SD de R2	Media de R3	SD de R3
1	<i>Hardware</i>	<i>Hardware</i> (+) software	4,80	0,40	4,90	0,30
		Seguridad de <i>hardware</i> (+)	4,80	0,40	4,80	0,40
		<i>Hardware</i> (+) flexibilidad del sistema	3,90	0,94	3,90	0,94
		<i>Hardware</i> (+) habilidades técnicas y soporte	4,20	0,98	4,20	0,98
2	<i>Software</i>	<i>Software</i> (+) seguridad	4,60	0,49	4,90	0,30
		<i>Software</i> (+) habilidades técnicas y soporte	3,80	0,87	3,80	0,87
		<i>Software</i> (+) flexibilidad del sistema	4,00	0,89	4,00	0,89
		<i>Software</i> (+) computación en la nube	3,90	0,94	3,90	0,94
3	Conectividad	Conectividad (+) <i>hardware</i>	4,70	0,46	4,90	0,30
		<i>Software</i> de conectividad (+)	4,70	0,46	4,90	0,30
		Conectividad (+) seguridad	4,60	0,49	4,80	0,40
		Conectividad (+) habilidades técnicas y soporte	3,90	0,94	3,90	0,94

No.	Factores tecnológicos	La relación entre factores tecnológicos	Media de R2	SD de R2	Media de R3	SD de R3
4	Seguridad	Seguridad (+)habilidades técnicas y soporte	4,20	0,98	4,20	0,98
5	Centros de datos	Centro de datos (+) seguridad	4,70	0,46	4,70	0,46
		Centro de datos (+) computación en la nube	3,90	0,94	3,90	0,94

Fuente: tomada de Al-araibi, Mahrin *et al.* (2019).

Al-araibi, Naz'ri bin Mahrin *et al.* (2019) se concentraron en el aspecto tecnológico de la preparación para el *E-Learning*. Propusieron un modelo que incluye ocho factores tecnológicos: *software* (SA), *hardware* (HA), conectividad (CO), seguridad (SE), flexibilidad del sistema (FL), habilidades y soporte técnico (TS), computación en la nube (CC) y centro de datos (DC). Se realizó un estudio cuantitativo en seis universidades públicas de Malasia, con respuestas a la encuesta de 374 miembros del personal académico que utilizan el aprendizaje electrónico. Según los resultados de la tabla 28, solo siete variables independientes mostraron un efecto significativo en la preparación para el aprendizaje electrónico entre los encuestados. El *hardware* afectó de manera significativa la preparación para el *E-Learning* ($\beta = 0,261$, $p = 0,000$); mientras que el *software* tuvo un efecto sustancialmente positivo en la preparación para el *E-Learning* ($\beta = 0,258$, $p = 0,000$). La conectividad tuvo un efecto positivo y significativo en la preparación para el *E-Learning* ($\beta = 0,372$, $p < 0,027$). La seguridad también tuvo un efecto positivo y significativo en la preparación para el *E-Learning* ($\beta = 0,289$, $p < 0,021$). Además, la flexibilidad tuvo un efecto positivo y significativo en la preparación para el *E-Learning* ($\beta = 0,098$, $p < 0,049$). También, el soporte técnico tuvo un efecto positivo y significativo en la preparación para el aprendizaje electrónico ($\beta = 0,337$, $p = 0,000$). El centro de datos tuvo un efecto positivo y significativo en la preparación para el *E-Learning* ($\beta = 0,77$, $p < 0,015$). Sin embargo, la computación en la nube no tuvo un efecto positivo y significativo en la preparación para el *E-Learning* ($\beta = 0,034$, $p < 0,455$).

Tabla 28. Prueba de los efectos totales de las variables independientes sobre la preparación para el *E-Learning*

Camino	segundo	β	SE	CR	Valor p
HA -----> ER	0,423	0,261	0,088	4,833	0,000
SA -----> ER	0,387	0,285	0,080	4,868	0,000
CO -----> ER	-0,399	0,372	0,090	3,096	0,027
SE -----> ER	0,474	0,289	0,070	1,249	0,021
FL -----> ER	0,161	0,098	0,082	1,966	0,049
TS -----> ER	0,420	0,337	0,072	5,812	0,000
DC -----> ER	0,212	0,377	0,079	4,421	0,015
CC -----> ER	0,052	0,034	0,069	0,747	0,455

Fuente: tomada de Al-araibi, Mahrin *et al.* (2019).

Mirabolghasemi *et al.* (2019) realizaron 153 encuestas en una universidad privada en la región norte de Irán, con el fin de investigar el impacto de los factores humanos, organizacionales y tecnológicos en la preparación de los estudiantes para el aprendizaje electrónico, utilizando el método de “mínimos cuadrados parciales” (PLS), mientras que Smart PLS se utilizó para probar las hipótesis y validar el modelo propuesto. La confiabilidad del modelo se probó utilizando el Alfa de Cronbach y las pruebas de confiabilidad compuesta. Normalmente, se recomienda una puntuación superior a 0,60 para el primero, y en este caso osciló entre 0,717 y 0,916 para el Alfa de Cronbach. Mientras tanto, los resultados de la fiabilidad compuesta oscilaron entre 0,841 y 0,935, también por encima del valor mínimo recomendado. Por tanto, los ítems utilizados para representar los constructos son confiables. Los resultados de las pruebas de hipótesis mostraron que la autoeficacia, el apoyo administrativo, la ventaja relativa, la compatibilidad y la complejidad de la computadora tienen una relación significativa con la preparación de los estudiantes para el aprendizaje electrónico.

Queiroga *et al.* (2020) plantearon diferentes técnicas que tratan de predecir de la deserción de los estudiantes (muestra de 725 estudiantes) de un curso virtual del Instituto Federal Sul Rio-Grandense (IFSul) en Brasil, teniendo en cuenta los niveles de aprendizaje presentados por ellos, mediante la variable:

cantidad de horas de interacción con la plataforma Moodle. Se evaluaron las siguientes técnicas: algoritmo genético, regresión logística, árbol de decisión, perceptrón multicapa, bosque aleatorio, algoritmo AdaBoost. Se evidencia que el mejor modelo fue el que se implementó por medio de algoritmos genéticos con generación de hiperparámetros, con una media de 0,84; además, este algoritmo posee la menor desviación estándar.

Jivet *et al.* (2020) dedicaron su trabajo a la influencia que tienen las variables de transparencia del diseño (relacionada con el diseño de la interfaz), marcos de referencia (relacionada con indicadores que se presentan en pantalla) y apoyo a la acción (relacionada con recomendaciones que se le hace a cada estudiante con respecto a su desempeño en las pruebas presentadas en el tablero) con las habilidades que estos presentan en cuanto al “aprendizaje autorregulado” (SRL), entendiendo este como la autonomía en línea que presentan los estudiantes para desarrollar sus actividades. Los resultados arrojan que los objetivos del alumno tienen un efecto significativo solo en la relevancia percibida de los marcos de referencia. Saber qué factores influyen en la interpretación de los estudiantes conducirá a diseños más inclusivos y flexibles que atenderán las necesidades de los estudiantes novatos y expertos.

Abdullah *et al.* (2020) estudiaron la relación entre el compromiso y la participación percibidos por los estudiantes, visto desde tres puntos de vista: compromiso conductual, el cual se refiere a la participación de los estudiantes en su aprendizaje (actividades en clase); compromiso emocional, concerniente con las relaciones que desarrollan los estudiantes en el aula y en la escuela; y, por último, compromiso cognitivo, el cual se relaciona con la inversión de los estudiantes en su aprendizaje (motivación y autorregulación). Todo ello a través de la definición de unos emojis, que los estudiantes utilizan para clasificar cada clase que reciben, los cuales son: excelente, solo así (el estudiante se siente normal), no decir (referente a que el estudiante prefiere no informar sobre sus sentimientos) y tengo un mal día. Existe una correlación negativa débil con $r = 0,15$ entre el compromiso conductual y el logro cognitivo en la evaluación sumativa (examen final), lo que sugiere que la asistencia puede no ser significativa para cierto estudiante basado en dos argumentos: que hay estudiantes que obtuvieron una buena puntuación en

el examen final con un número bajo de asistencias y que hay estudiantes que obtuvieron una puntuación baja. A partir del análisis, son más los alumnos que no se sintieron positivos y emocionados durante la sesión presencial que logran sacar una buena puntuación (> 50) en el examen final. Sin embargo, también hay estudiantes que se sienten positivos y emocionados, pero no les fue bien durante el examen final (puntuación < 50).

Muljana y Luo (2020) realizaron un estudio sobre las características que se deben tener en cuenta en el diseño de entornos con aplicación de la analítica del aprendizaje, se tomó una muestra de 18 personas involucradas en este diseño específicamente en la educación superior. Se adoptó el “modelo de aceptación de tecnología” (TAM); este estudio se ha utilizado para explorar la aceptación de la tecnología por parte de los usuarios. Consta de dos componentes principales que afectan la aceptación del usuario, a saber: la utilidad percibida y la facilidad de uso percibida. Doce de los 18 diseñadores de entornos basados en la analítica del aprendizaje expresaron percepciones positivas preexistentes sobre los beneficios. Estuvieron de acuerdo en que la analítica del aprendizaje (LA) era útil para comprender a los estudiantes, su contexto y el progreso del aprendizaje, además de evaluar el compromiso y la eficacia del curso. Nueve de los diseñadores de entornos basados en LA mencionaron la importancia de tener pedagogía para impulsar el uso de la tecnología. Los diseñadores de entornos basados en LA parecían comprender el significado de LA en consonancia con su definición: la utilización de LA tiene como objetivo comprender a los alumnos y su contexto para mejorar el aprendizaje y su entorno. Este tipo de percepción resuena con los estudios existentes, que postulan que los datos de los estudiantes perderán el significado cuando se analicen e interpreten de manera incoherente con los objetivos pedagógicos y el contexto de aprendizaje. Pocos diseñadores de entornos basados en la analítica del aprendizaje creían que esta es simplemente una técnica de medición. En particular, dos diseñadores informaron que creían que la evaluación no analítica sigue siendo útil.

Olsen *et al.* (2020) estudiaron el análisis de múltiples flujos de datos. En este trabajo, investigaron cómo el uso multimodal de datos en un entorno colaborativo proporciona una mejor predicción de las ganancias del aprendizaje. Para abordar estas preguntas, analizaron datos multimodales utilizando

un sistema de tutoría inteligente (registro del tutor, el audio y los datos de diálogo con las ganancias de aprendizaje de los estudiantes). Encontraron que una combinación de modalidades, especialmente aquellas en una escala de tiempo más pequeña, como el video y el audio, proporcionan una predicción más precisa de las ganancias de aprendizaje que los modelos con una sola modalidad.

Li *et al.* (2020), en su estudio, examinaron el cambio de los perfiles de “aprendizaje autorregulado” (SRL) de los estudiantes a lo largo del tiempo en un entorno de “diseño asistido por computador” (CAD). Se identificaron tres perfiles de SRL: orientado a la reflexión, adaptativo y alumnos mínimamente autorregulados. Además, encontraron que los estudiantes con diferentes membresías de perfil diferían de manera significativa en su desempeño de diseño. En particular, los alumnos autorregulados adaptativos superaron a los alumnos mínimamente autorregulados en la complejidad del diseño, mientras que los alumnos autorregulados orientados a la reflexión demostraron una mayor eficiencia en el diseño que los alumnos mínimamente autorregulados. Este estudio proporciona a los investigadores conocimientos teóricos y metodológicos sobre la dinámica del SRL. Los resultados de esta investigación también informan a los profesionales sobre el diseño de intervenciones adaptativas.

Tahir y Wang (2020) indagan sobre una forma de establecer la metodología y terminología para abordar trabajos relacionados con el “aprendizaje basado en juegos” (GBL) debido al uso inconsistente de elementos. Este estudio cierra la brecha al delinear la terminología del GBL y el alcance con cuatro niveles conceptuales para luego categorizar sistemáticamente los elementos del GBL por alcance, definición y uso. La metodología utilizada es el análisis de contenido dirigido de la literatura del GBL recopilada mediante una revisión bibliográfica sistemática previa. La dimensionalización del GBL y su posterior descomposición en factores/subfactores basados en construcciones teóricas, ha dado como resultado un patrón consistente y claro que delinea la estructura del diseño, análisis y evaluación del juego educativo. La codificación adicional de métricas y el mapeo de la relación entre las dimensiones del GBL se deducen en un marco conceptual (llamado LEAGUÊ), que facilita una mayor comprensión del proceso de aprendizaje

con juegos educativos, dónde enfocarse y qué evaluar. El marco LEAGUÊ se puede aplicar para el análisis, diseño y evaluación de juegos de aprendizaje.

La multidimensionalidad del GBL requiere la evaluación de varios aspectos denominados dimensiones centrales (RQ1), incluyendo aprendizaje, factores del juego, reacciones afectivo-cognitivas, usabilidad, usuario y medio ambiente. Cada dimensión se centra en ciertos factores y subfactores que constituyen ese aspecto, y se requieren métricas para evaluarlos. El marco presenta 22 factores (que abordan RQ2), 74 subfactores (RQ3) y cinco categorías de métricas (RQ4). Las dimensiones del GBL están relacionadas entre sí, y es esencial evaluar las relaciones presentadas como una alta abstracción de LEAGUÊ para obtener información más significativa sobre los juegos educativos (RQ5). El marco proporciona una imagen detallada del GBL que guiará no solo a los investigadores y evaluadores, sino también a los diseñadores y desarrolladores de juegos educativos. El marco propuesto se basa en componentes basados en la teoría. Cada componente tiene una base sólida para la formación que está respaldada por constructos teóricos en la literatura del GBL y no meramente basada en sospechas.

La investigación de Gizem *et al.* (2020) tiene como objetivo examinar las opiniones de los profesores en formación sobre la recomendación personalizada y la retroalimentación de orientación basada en análisis de aprendizaje. Es un estudio de tipo cualitativo. El “sistema de manejo de aprendizaje” (LMS) semanal de los estudiantes obtuvo resultados de análisis de aprendizaje, a partir de los datos de registro relacionados con su comportamiento de uso. Luego, el investigador preparó mensajes personalizados de recomendación y orientación basados en los resultados de las analíticas de aprendizaje. Los resultados del análisis de aprendizaje y las recomendaciones relacionadas y los mensajes de orientación se enviaron por el LMS (desde la herramienta de mensajería) como comentarios.

Chen (2020) crea un modelo para comprender el sistema *E-Learning*. Se tienen en cuenta seis dimensiones: calidad de la información, calidad del sistema, uso del sistema, satisfacción del usuario, impacto individual e impacto organizacional. La calidad de la información y la calidad del sistema aumentan el uso del sistema y la satisfacción del usuario, lo que genera un impacto individual e influye en el impacto organizacional. Se

recopilaron y analizaron datos de 240 estudiantes universitarios que utilizaron sistemas de *E-Learning* en el aprendizaje de sus cursos. Se utilizó el “modelado de ecuaciones estructurales” (SEM) en el estudio, ya que era un método de análisis de datos comúnmente utilizado para la validación y extensión teóricas mediante AMOS 22.0 para el análisis de datos. Los resultados mostraron que la calidad de la información se asoció significativamente con el uso del sistema de los estudiantes universitarios y la satisfacción del usuario. Sin embargo, se demostró que la calidad del sistema se asoció únicamente con el empleo del sistema por parte de los estudiantes universitarios, pero no con la satisfacción del usuario. Finalmente, la sorpresa, que refleja la calidad del servicio, mostró una asociación significativa solo con la satisfacción del usuario, pero no con el uso del sistema.

Azzi *et al.* (2020), en su artículo, proponen un clasificador robusto que puede identificar el estilo de aprendizaje del alumno que usa un sistema *E-Learning*. El trabajo se centró en dos aspectos: mostrar la solidez del enfoque en cuanto al comportamiento del aprendizaje definido en categorías, según el modelo de estilo de aprendizaje de Felder-Silverman (FSLSM) y mostrar la precisión del algoritmo Fuzzy C Means (FCM) utilizado para agrupar los datos de comportamiento de aprendizaje capturados. Los cursos de análisis corresponden a las ciencias informáticas. En la primera prueba se tomaron datos recopilados de los comportamientos de los estudiantes que han seguido dos cursos: PHP / MySQL y análisis UML. En la segunda prueba, se consideran tres cursos de la materia informática de ciencias: PHP / MySQL, análisis UML y programación JAVA. Los resultados muestran la agrupación de los estilos de aprendizaje en la tabla 29.

También arrojaron que en la primera prueba la precisión del algoritmo FCM es del 93,41 % para 210 iteraciones y en la segunda prueba, la precisión obtenida es de 93,16 % para 220 iteraciones. Puesto que el número de valores de características en la segunda prueba es mayor que el de la primera prueba (véase tabla 30).

Tabla 29. Resultados de agrupación de estilos

Clusters	Número de secuencias	
	Primera prueba (2 cursos)	Segunda prueba (3 cursos)
Activo	138	141
Reflexivo	119	114
Detectando	107	105
Intuitivo	131	129
Visual	228	225
Verbal	99	104
Secuencial	115	112
Global	124	126

Fuente: tomada de Azzi *et al.* (2020).

Tabla 30. Precisión del algoritmo

Primera prueba (2 cursos)		Segunda prueba (3 cursos)	
Iteraciones	Exactitud (%)	Iteraciones	Exactitud (%)
70	85,27	80	84,91
110	87,38	120	87,12
210	93,41	220	93,16

Fuente: tomada de Azzi *et al.* (2020).

Almaiah *et al.* (2020) emplearon entrevistas utilizando análisis temático con el *software* NVivo. La entrevista se realizó con 30 estudiantes y 31 expertos en sistemas de *E-Learning* en seis universidades de Jordania y Arabia Saudita. Tuvieron como objetivo investigar los principales desafíos y factores que afectan el sistema de aprendizaje electrónico durante la pandemia de COVID-19. El propósito principal de este método, es capturar algo importante de los datos recopilados en relación con la pregunta de investigación. Se identificaron aspectos como la cultura, la calidad del aprendizaje *E-Learning*, las redes sociales, la cultura social y las fuentes financieras.

Teo *et al.* (2020) estudiaron la efectividad y la legitimidad en el *E-Learning* por medio de la aplicación de entrevistas. Es importante tener en cuenta que los autores definen la legitimidad como un juicio social de su aceptación, idoneidad o deseabilidad. Estas se realizaron en la sede de *Education Broadcasting Systems* (EBS) (Seúl), puesto que es pionero en el aprendizaje electrónico en Corea el Sur. Como resultado de este trabajo, se identificaron cuáles son las medidas clave para facilitar el éxito de la implementación de *E-Learning* en la educación superior de Corea del Sur: construir una infraestructura de aprendizaje electrónico eficiente para facilitar el acceso a los recursos de aprendizaje electrónico, estandarización de los sistemas de aprendizaje electrónico para mejorar la calidad de los recursos de aprendizaje electrónico, establecimiento de un sistema reglamentario sólido para la legitimidad reglamentaria, aumentar la conciencia pública sobre el aprendizaje electrónico para mejorar la legitimidad cognitiva y aplicar la pedagogía en línea socialmente apropiada para la legitimidad normativa.

Kuliya y Usman (2020), mediante un “modelo de aceptación de tecnología” extendida (TAM), intentaron predecir las intenciones de académicos y estudiantes de utilizar el aprendizaje electrónico en las instituciones de educación superior (IES) del noreste de Nigeria, por medio de un cuestionario para recopilar datos para el trabajo de investigación. En la tabla 31 se presenta las preguntas de la entrevista realizada. Se analizaron los datos extraídos y los resultados mostraron que, tanto para los académicos como para los estudiantes, la “facilidad de uso percibida” (PEOU) y la “utilidad percibida” (PU) del *E-Learning* predijeron fuertemente las intenciones conductuales de los participantes de utilizar el *E-Learning*. Además, los resultados confirmaron la confiabilidad de los componentes centrales del modelo de aceptación de tecnología.

Tabla 31. Preguntas de la entrevista

Preguntas
1. ¿Eres académico o estudiante?
2. ¿Cuál es su calificación más alta?
3. ¿Cuáles son los desafíos que experimentó en el sistema tradicional de aprendizaje?

Preguntas
4. ¿Tiene una computadora?
5. ¿Cómo se conecta a internet?
6. ¿Su institución cuenta con servicios de internet para académicos y estudiantes?
7. ¿Su institución ofrece servicios de aprendizaje electrónico?
8. ¿Ha utilizado alguna vez alguna plataforma <i>E-Learning</i> ?
9. Si respondió “sí” a la pregunta anterior, ¿cómo fue la experiencia?
10. ¿Está dispuesto a estudiar/enseñar en un entorno de <i>E-Learning</i> ?

Fuente: tomada de Kuliya y Usman (2020).

10.2. Analítica académica periodo 2015-2020

Echegaray-Calderón y Barrios-Araníbar (2015) trabajaron con el fin de identificar cuáles son los factores más influyentes en la predicción del éxito académico en entornos de educación superior, utilizando una red neuronal artificial como predictor de una universidad del sur del Perú. Se tiene en cuenta un total de 1271 estudiantes que ingresaron en el 2013. Los resultados de este trabajo muestran que no siempre el mejor estudiante no es el que tiene el mejor resultado de rendimiento académico. El tamaño de los factores se redujo de 39 a 8, a partir de una encuesta y de datos suministrados. Los factores más importantes son: edad, género, examen de admisión (prueba escrita), entrevista inicial, la práctica para aprender conocimientos, el conocer cómo sobrellevar situaciones difíciles en términos académicos y enfrentar situaciones adversas de la vida.

Pradeep *et al.* (2015) analizaron los factores que afectan el rendimiento académico de los estudiantes que contribuyen a la predicción del fracaso y abandono mediante el uso de la herramienta Weka, y 670 registros de estudiantes del periodo 2011-2013 de una institución de la India. Se parte de 57 atributos que incluyen variables demográficas, académicas y psicosociales. En esta investigación, se utilizaron los siguientes algoritmos, entre otros: JRip, OneR, J48, SimpleCart, Prism y ADtree. Este último dio los mejores resultados de precisión (97,2).

Hernández, Meléndez *et al.* (2016) compararon cuatro algoritmos: regresión logística, agrupamiento, árboles de decisión y red neuronal, con el fin de identificar los mejores resultados sobre la predicción de los riesgos de deserción de los 134 estudiantes en el ITSM-México (Instituto Tecnológico Superior de Misantla-México). Se observó que cuando el sistema utiliza el algoritmo de regresión logística, tiene el 100% en precisión, sensibilidad, exactitud y la medida-F; seguido por el algoritmo de red neuronal.

Dixon *et al.* (2016) examinaron las contribuciones de tres variables psicosociales (autoeficacia académica, esperanza y valor) al rendimiento académico (control de la capacidad percibida) a partir de 609 estudiantes. Los análisis de regresión indicaron que la esperanza y la autoeficacia académica eran predictores de la capacidad de percepción y que la autoeficacia académica era el predictor más sólido del logro académico. El valor no fue un predictor significativo de la capacidad percibida o del rendimiento académico. Se llegó a la conclusión de que las variables psicosociales pueden contribuir al rendimiento en estudiantes talentosos.

Ahamed *et al.* (2016) buscaron, en su estudio, la predicción del resultado de la HSC (examen de ingreso) en función de factores socioeconómicos, psicológicos y académicos. La información se obtuvo a partir de encuestas para lograr los atributos relevantes relacionados con los logros académicos de los estudiantes. Los 423 estudiantes tienen diferentes estatus económicos y distintas ubicaciones (sitio geográfico). Se cuenta con 33 atributos agrupados en los factores antes mencionados. En cuanto a los resultados, la precisión más alta se obtuvo utilizando redes neuronales artificiales (89,48 %) junto con la selección de funciones, SMOTE, MLD y normalización como método de preprocesamiento de los datos. A medida que se introduce el preprocesamiento, la precisión aumenta. Otro método con buenos resultados (82 %) es el uso de bosques aleatorios. No son buenos los resultados al usar Naive Bayes.

Hasbun *et al.* (2016) elaboran su trabajo con el propósito de detectar a los estudiantes propensos a la deserción y determinar cuáles son los factores que las instituciones deben observar al tratar de evitar esta situación, teniendo en cuenta las actividades extracurriculares de estudiantes de dos carreras: de ingeniería y negocios, de una universidad en Chile. Se recopilaban datos

de 4840 estudiantes del periodo 2008-2011 y se capacitaron y validaron dos modelos: uno que incluía todos los datos y otro que eliminaba créditos que valían cursos; lo que demuestra que las actividades extracurriculares son excelentes predictores de la deserción. Se obtuvieron cuatro conjuntos de datos de la base de datos: la primera tabla incluía información del estudiante: campus, estudio, grupo, sexo, fecha de nacimiento, título, estado (activo, abandonado, terminado), semestres inscritos, créditos aprobados, cursos aprobados como segundo idioma, puntajes de exámenes de selección universitaria (admisión) y “promedio de calificaciones acumuladas” (GPA). El segundo conjunto de datos contenía información detallada sobre las inscripciones de los estudiantes en los cursos y sus calificaciones finales. El tercer conjunto de datos contenía detalles sobre la información del curso, como los créditos otorgados (una vez aprobados) y los títulos en los que se incluyeron. El cuarto conjunto de datos incluyó la asistencia deportiva de los estudiantes, detallando cada fecha a la que asistieron y qué actividad física o deporte practicaron. Así como información sobre la escuela secundaria del estudiante, dependencia escolar (privada, pública) y la ubicación.

Se entrenaron dos modelos, uno que incluye todos los datos y otro que elimina las calificaciones y los créditos obligatorios de las actividades. Se utilizó el *software* R Studio usando el algoritmo CART. El primer árbol de decisión tuvo una precisión del 93,94% al predecir la deserción de los estudiantes. Son factores importantes para la predicción: la eficiencia de créditos regulares, el número de semestres aprobados, el promedio de calificaciones acumuladas, la escuela y el número de semestres inscritos. Para el segundo modelo, no se tuvieron en cuenta las siguientes variables: valor del crédito acumulado, eficiencia de los créditos regulares y los cursos regulares. El árbol de decisión CART tuvo una precisión de predicción del 79,29%. Son factores importantes para la predicción: eficiencia deportiva, eficiencia complementaria, ubicación de la escuela y dos asignaturas de matemáticas.

Devasia *et al.* (2016), por medio del método de Naive Bayes, clasificaron el rendimiento de 700 estudiantes del periodo 2013-2016. Se mostró el resultado mediante una aplicación web. Se clasificaron los estudiantes en cuatro clases o categorías de predicción. Se recopiló información como acciones de

grupo, pruebas de clase, seminarios y calificaciones de tareas previas de los estudiantes para predecir el desempeño en el semestre estudiado.

Merchán-Rubiano *et al.* (2017), por medio de un estudio de analítica de datos, buscaron predecir la deserción académica de estudiantes de ingeniería en la Universidad El Bosque, de Colombia. Los datos base son los resultados en las Pruebas de Estado ICFES o Saber 11 y datos demográficos de 7644 estudiantes. Los algoritmos de clasificación son: el árbol de decisión J48 y Random Forest. Los resultados mostraron que la deserción impacta más sobre los estudiantes con un puntaje general de la prueba ICFES por debajo de 46,8, puntajes de física y química por debajo de 44 y 54, respectivamente. Impacta más sobre los estudiantes que tienen un estrato social medio y no trabajan en el momento de la inscripción. Cuando el puntaje general de la prueba ICFES aumenta y el estado de si tienen o no trabajo no es relevante, el rendimiento es favorecido. También encontraron que, para predecir el rendimiento académico, las variables más influyentes son el puntaje general de la prueba ICFES, género, clases de día/noche y las puntuaciones de razonamiento cuantitativo obtenidas en la prueba. El bajo rendimiento es, principalmente, influenciado por el puntaje general de la prueba, el género masculino y el puntaje matemático por debajo de 56.

Razaque *et al.* (2017) utilizaron el algoritmo Naive Bayes para la evaluación del “desempeño del estudiante de posgrado” (GPA) al final del semestre. Para la evaluación general de los estudiantes, se tuvieron en cuenta los siguientes elementos: actividades educativas, asistenciales y extracurriculares. Se aplicó un método de preprocesamiento al conjunto de datos, dejando solo 500 registros para el estudio. También, congregaron a los estudiantes por medio del algoritmo de agrupación de K-means. Se recopilaron datos como la asistencia, la discusión y las calificaciones de las actividades. Permitió mejorar el resultado de los estudiantes que estaban en peligro de fracasar. Esta investigación también fue un esfuerzo para identificar a los estudiantes que requerían una consideración especial para disminuir el fracaso y tomar las medidas adecuadas para el próximo examen semestral.

Stanko *et al.* (2017) investigaron la predicción del rendimiento académico futuro de los estudiantes a partir del análisis de las pruebas estatales rusas. Demostraron, en términos estadísticos, que el examen estatal unificado

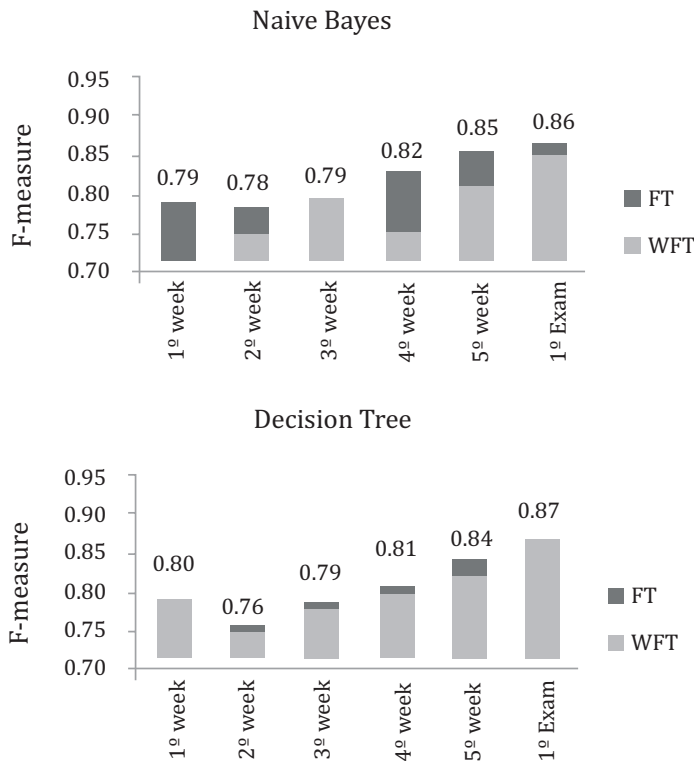
obligatorio de Rusia no proporciona datos suficientes para predecir el potencial académico de un solicitante. Para predecir el rendimiento académico potencial de un individuo, se tuvieron en cuenta cuatro cursos básicos: cálculo (C), matemáticas discretas (DM), arquitectura informática (CA), introducción a la programación (P), para los cuales se intentaba predecir la calificación de cada curso, así como el “desempeño” promedio. Ninguno de los modelos de regresión lineal pudo explicar más del 25 % de la varianza en los datos (por lo general, el modelo podría considerarse aceptable cuando describe al menos el 50 % de la varianza). Todos los modelos de regresión logística son constantes y predijeron el resultado más frecuente (aprobar en todos los casos, excepto en el curso de introducción a la programación). Por tanto, todos los modelos desarrollados, basados en el criterio de admisión numérico, no lograron predecir de manera confiable el rendimiento académico.

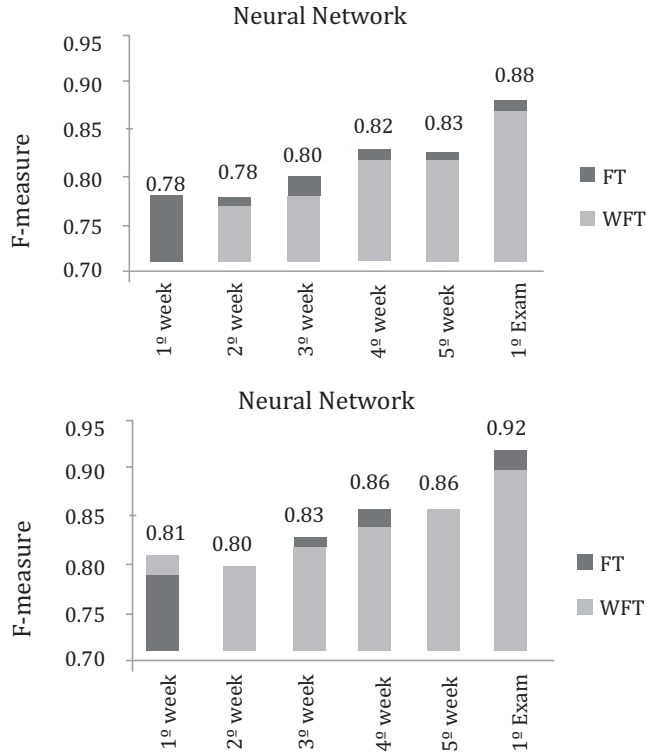
Qin *et al.* (2017) emplearon el índice de Gini para clasificar la importancia de los factores que afectan el rendimiento de los alumnos y el bosque aleatorio se utilizó para predecir el rendimiento. Los resultados del experimento muestran que la precisión de la predicción alcanza el 83,82 % para el bosque aleatorio. El trabajo contiene un análisis detallado para ordenar la importancia de las características de los estudiantes, que influyen en su rendimiento por el índice de Gini. Los resultados muestran que las puntuaciones de los cursos prácticos tienen un impacto mínimo en el aprendizaje posterior de los cursos avanzados.

Costa *et al.* (2017) presentaron un estudio comparativo sobre la efectividad de las técnicas de minería de datos educativos, para la detección temprana de aquellos estudiantes que probablemente reprobarían los cursos de programación introductoria. El estudio se llevó a cabo mediante cuatro técnicas de predicción, usando dos fuentes de datos diferentes e independientes en los cursos de programación introductoria, disponibles en una universidad pública de Brasil: uno proviene de la educación a distancia y el otro en el campus (véase figura 34). En la figura se muestran los resultados comparativos de la efectividad de los algoritmos utilizados sin ajuste (WFT) y luego con ajuste (FT) para el sistema de educación a distancia. Se observó una mejora significativa en el algoritmo de NB con el ajuste

para los datos en línea preprocesados. Para la primera, segunda, tercera, cuarta y quinta semana se obtuvo que la máquina de vectores de soporte (*support vector machine*) presentó la mayor efectividad con 81 %, 80 %, 83 %, 86 % y 86 % de efectividad, respectivamente, y 92 % para el primer examen. Mientras que árboles de decisión (*decision tree*) presentó la menor efectividad para la segunda y cuarta semana con 79 % y 81 %, respectivamente. Los algoritmos Naive Bayes y árboles de decisión (*decision tree*) presentaron la menor efectividad para la tercera semana con 79 %. Para la quinta semana, la red neuronal (*neural network*) presentó la menor efectividad con 83 %. Finalmente, en el primer examen, Naive Bayes presentó la menor efectividad con 86 %.

Figura 34. Efectividad de los algoritmos con y sin ajuste





Fuente: tomada de Costa *et al.* (2017).

Khalil *et al.* (2018) establecieron tres modelos para permitir la identificación e intervención tempranas para los estudiantes en riesgo de fracaso académico y de reprobación en el examen: a) el modelo de predicción previo a la matriculación incluye puntajes MCAT y GPA general de pregrado; b) el final del modelo M1 incluye los puntajes de las pruebas de progreso en el NBME administrado al final del primer año de la escuela de medicina (M1), los puntajes de la subescala LASSI y el desempeño ponderado en los exámenes internos M1, y c) el modelo final de M2 incluye los puntajes de las pruebas de progreso CBSE al final de M1 y el segundo año de la escuela de medicina (M2), los puntajes de la subescala LASSI y el desempeño ponderado en los exámenes internos M1 y M2. Se utilizaron los coeficientes de correlación del momento del producto de Pearson y el modelo de regresión multivariante para determinar combinaciones óptimas de variables independientes, que representan atributos de comportamiento y resultados de desempeño de los estudiantes (es decir, mínimos cuadrados ordinarios para optimizar la variación explicada y el error

estándar de la estimación). Los resultados mostraron que existe una asociación limitada, pero, en términos estadísticos, significativa entre el GPA de pregrado y el MCAT con el desempeño en el examen del USMLE. En resumen, el modelo de prematriculación explicó el 24 %, el final del modelo M1 explicó el 62 % y el final del modelo M2 explicó el 81 % de la variación en los puntajes del USMLE.

Liu *et al.* (2018) partieron de un conjunto de datos que contiene información sobre 22 cursos impartidos por la Universidad Abierta Beijing, China (OU) (32 593 estudiantes). Los resultados de sus evaluaciones y los registros de sus interacciones en línea están representados por resúmenes diarios de flujos de clics (10655280 entradas). Su objetivo consistía en estudiar la deserción en los cursos de aprendizaje a distancia abiertos y en línea por medio de un método de clasificación de series de tiempo (bosques de series de tiempo — TSF). Usaron la precisión como métrica de evaluación. El número de árboles en el modelo forestal de series de tiempo se establece en 500. A partir de estos resultados experimentales, se descubrió que la precisión general de la predicción es sorprendentemente buena. 14 de cada 15 presentaciones de módulos alcanzan una precisión superior al 80 %, dos de ellas alcanzan una precisión superior al 90 %. Significa que la metodología de construcción de datos y el modelo de clasificación de series de tiempo, pueden funcionar juntos como una herramienta eficaz para predecir la posibilidad de desvinculación de los estudiantes.

Limsathitwong *et al.* (2018) investigaron sobre la deserción escolar del Instituto de Tecnología Thai-Nichi (Tailandia). Desarrollaron una aplicación web para predecir el estado de los estudiantes a partir de su calificación de cada asignatura. Utilizaron el *software* RapidMiner Studio y algoritmos como árbol de decisión y bosque aleatorio. Algunas variables usadas en el estudio: año de ingreso, código del estudiante, promedio académico, facultad, cantidad de créditos mínimos, cantidad máxima de créditos, código del curso, código del estudiante, entre otras. Los resultados muestran que los algoritmos tienen valores de precisión, recuperación y de medida F1 como máximo de 81,20 %, 89,51 % y 85,15 %, respectivamente.

Lima *et al.* (2018) se centraron en la solución de problemas de deserción escolar y desempeño, tomando una muestra de 1844 estudiantes entre

egresados y desertores en el periodo 2007-2015, procesados mediante el *software* R, utilizando los algoritmos K-means y C5.0. La investigación se llevó a cabo en una universidad pública de Portugal. Se analizaron las siguientes variables: edad de acceso, número de unidades de curso aprobadas, número de intentos hasta éxito y la media de las clasificaciones obtenidas en las unidades de curso aprobadas. Como consecuencia de los análisis previos, se concluye que las principales características que pueden ayudar a distinguir entre un alumno con tendencia a completar el curso y un alumno que puede estar en riesgo de deserción son: la edad de acceso, el número de asignaturas aprobadas, el número de intentos de aprobación y la media de calificaciones. Cabe señalar que no se identificaron diferencias de género relevantes; sin embargo, es un factor que debe tenerse en cuenta. A partir de estas características, se utilizó el algoritmo K-means para identificar los grupos de estudiantes y las características inherentes a cada uno de ellos. Aplicando el método del codo, el número ideal de grupos para el conjunto de datos analizados es 4. El resultado del árbol muestra que, de los 223 casos de deserción, 150 son correctamente clasificados, y de los 275 casos de egresados, se clasificaron con precisión 256. Por tanto, se concluye que el modelo propuesto tiene una precisión del 67,3% para identificar características relacionadas con la deserción, mientras que el 93,1 % para datos sobre los egresados.

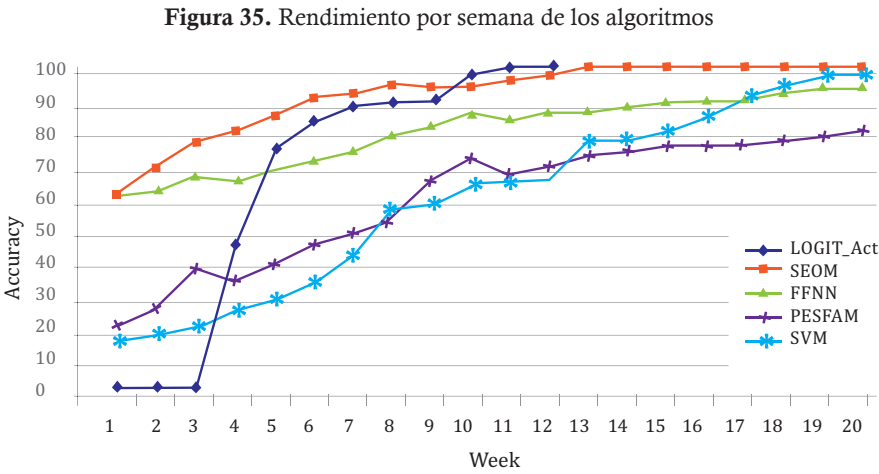
Hirokawa (2018) analizó la información de 480 estudiantes universitarios, utilizando las notas obtenidas en su curso, las cuales se clasifican en tres categorías: H (alto), M (medio) y L (bajo). Se utilizó este puntaje como objetivo de predicción. Las variables utilizadas se clasificaron en cuatro tipos de atributos: características demográficas, antecedentes académicos, participación de los padres y características de comportamiento. El modelo SVM se utilizó para el estudio. Su objetivo es el de identificar cuál combinación de los cuatro tipos de variables es efectiva para la predicción. “a” (antecedentes académicos), “b” (características de comportamiento), “c” (características demográficas) y “d” (información familiar). Los resultados muestran que al ejecutar el algoritmo SVM se obtuvo un porcentaje de precisión del 84,3%. Considerando a la variable de comportamiento como la clave para la predicción del rendimiento académico de los estudiantes.

Wati *et al.* (2018) decidieron adoptar los algoritmos Naive Bayes y C4.5 para predecir el resultado de un grupo de estudiantes de la Universidad de Mulawarman (Indonesia). Para comparar el rendimiento de los algoritmos, decidieron fraccionar los datos en nueve proporciones (10% al 90%) para cada entrenamiento y prueba de datos. Naive Bayes alcanzó una exactitud del 76,79% con una proporción del 80% de los datos y una precisión del 81,92%. Mientras que su desempeño más bajo es con una proporción del 10% de los datos, cuya exactitud es de 53,78% y su precisión de 42,38%. Por otro lado, el algoritmo C4.5 alcanzó una exactitud de 78,57%, con una proporción del 90% de los datos y una precisión del 76,72%.

Burgos *et al.* (2018), mediante el análisis de datos históricos acerca de las calificaciones de diferentes cursos de los estudiantes, buscaron predecir si un estudiante abandonaría o no un curso. Esto les permitió reducir la tasa de deserción en 14% con respecto a los años académicos anteriores, en los que no se realizó ningún estudio ni se aplicaron mecanismos de prevención que evitaran la deserción académica. En la figura 35 se observa la comparación de cada uno de los algoritmos en el transcurso de 20 semanas. A partir de la semana diez, la propuesta de los autores (LOGIT_Act) junto con el SEDM (sistema para minería de datos educativos) ofrecen la mayor exactitud. Mientras que PESFAM (mapeo de teoría de resonancia adaptativa) la menor. Desde la semana cinco hasta la semana nueve, la nueva propuesta muestra una tendencia positiva. En el punto medio decisivo (semanas 10-12), la nueva propuesta ofrece una mayor precisión superando otras propuestas.

Miguéis *et al.* (2018), para validar la metodología propuesta, utilizaron un conjunto de datos de 2459 estudiantes, que abarcan el 2003 al 2015, de la Escuela Europea de Ingeniería de una universidad pública. Se encuentra que los bosques aleatorios son superiores a las otras técnicas de clasificación que se consideraron (árboles de decisión, máquinas de vectores de soporte, Naive Bayes y árboles reforzados). El modelo de predicción se convierte en una herramienta útil para aplicar estrategias, con el fin de promover mayores niveles de rendimiento y mitigar el fracaso académico, aumentando en general la calidad de la experiencia académica proporcionada por una institución de educación superior. Los resultados demuestran la capacidad del

modelo propuesto para predecir el nivel de rendimiento de los estudiantes, el cual obtuvo una precisión superior al 95 %, en una etapa temprana de la trayectoria académica de los estudiantes. En resumen, las técnicas de clasificación utilizadas presentaron los siguientes porcentajes de precisión: bosque aleatorio con 96,1 %; Adaptive Boosting 95,7%; máquina de vectores de soporte (SVM) 93,9%; árboles de decisión (decision tree) 91,5%; Baggin 88,7%; Naive Bayes 75,9%.



Fuente: tomada de Burgos *et al.* (2018).

Kaunang y Rotikan (2018) trabajaron un modelo para predecir el rendimiento académico de los estudiantes del Departamento de Informática de la Universitas Klabat (Indonesia), utilizando dos algoritmos: árbol de decisión y bosque aleatorio. Con el fin de optimizar el parámetro de los clasificadores, se realizó una validación cruzada cinco y diez veces; además, permitió reducir el sesgo en la experimentación y aportar, en gran medida, en la construcción del modelo de predicción. Los resultados son los siguientes:

Árbol de decisión: para una validación cruzada de cinco veces, este algoritmo obtuvo una exactitud del 62,85% y una precisión en una escala de 0 a 1 de 0,543. Para la validación cruzada de diez veces, obtuvo una exactitud del 66,85% y una precisión en una escala de 0 a 1 de 0,622. De acuerdo con lo anterior, tuvo un mejor desempeño el algoritmo cuando se realizó una validación cruzada de diez veces.

Bosque aleatorio: cuando se realizó una validación cruzada de cinco veces, el algoritmo tuvo un porcentaje de exactitud del 61,14% y una precisión de 0,572 en una escala de 0 a 1. Con una validación cruzada de diez veces, el algoritmo tuvo un porcentaje de exactitud del 61,14% y una precisión de 0,573. De acuerdo con lo anterior, este algoritmo obtuvo el mismo rendimiento, sin importar la cantidad de veces de la validación cruzada.

Salal *et al.* (2019) compararon ocho algoritmos mediante el *software* WEKA, con el fin de determinar aquel con mayor porcentaje de precisión. Para ello, utilizaron 32 variables y los resultados mostraron que en la mayoría de los casos los algoritmos mejoran su precisión al reducir el número de variables a las más significativas. OneR, REPTree y Decision Tree (J48) obtuvieron un porcentaje de precisión superior al 76 %. La tabla 32 muestra el porcentaje de precisión para cada uno de los algoritmos.

Tabla 32. Rendimiento de los algoritmos con dos conjuntos de variables

Classification Algorithm taken for implementation	Classification Algorithm Accuracy with all attributes (%)	Classification Algorithm taken for implementation	Classification Algorithm Accuracy with eight attributes and with different parameter selection (%)
Naive Bayes	68,2589	Naive Bayes	73,1895
Decision Tree (J48)	67,7966	Decision Tree (J48)	76,2715
Random Tree	53,4669	Random Tree	67,9507
REPTree	75,1926	REPTree	76,7334
JRip	70,7242	JRip	74,114
OneR	76,7334	OneR	76,7334
SimpleLogistic	71,3405	SimpleLogistic	73,6518
ZeroR	30,9707 (Baseline accuracy)	ZeroR	30,9707 (Baseline accuracy)

Fuente: tomada de Salal *et al.* (2019).

David y Anastasija (2019) tuvieron en cuenta las variables como las familiares, el nivel educativo de la madre, el estrato socioeconómico de la vivienda, el nivel educativo del padre, la cantidad de libros en la casa y el poseer

computador en la vivienda, para predecir los resultados de los estudiantes del 2016, que presentaron la Prueba de Estado para acceder a la educación superior. Encontraron que las variables mencionadas en su orden, tienen relación con los resultados obtenidos. Como método, emplearon árboles de decisión.

Ajibade, Ahmad *et al.* (2019) evaluaron el desempeño de los métodos de selección de características y el uso de algoritmos de clasificación, como Naive Bayes (NB), Decision Tree (DT), K-vecino más cercano (KNN) y análisis discriminante (DISC). El algoritmo de evolución diferencial (DE) se propone como un mejor algoritmo de selección de características para evaluar el rendimiento académico de los estudiantes. El conjunto de datos utilizado se obtuvo de una LMS llamada Kalboard 360. En este estudio, el conjunto de datos asciende a 500 estudiantes con 16 características.

Nabizadeh *et al.* (2019) buscaron predecir el rendimiento académico con base en estrategias de motivación de aprendizaje y expectativas de resultados, a partir de un modelo teórico con la participación de 380 alumnos y alumnas de nueve facultades de ciencias médicas de la Universidad Shahid Beheshti de Teherán. Para la recolección de datos, se utilizó un muestreo en múltiples etapas, junto con el cuestionario de estrategias motivacionales para el aprendizaje y la escala de expectativas de resultados del estudiante. El “promedio de calificaciones universitarias” (CGPA) de las calificaciones anteriores de los estudiantes se consideró como la variable de desempeño académico. El análisis de datos se realizó utilizando el “modelado de ecuaciones estructurales” (SEM) en el *software* AMOS. La puntuación media de la estructura de las estrategias de aprendizaje, las estrategias de motivación, las expectativas de resultados y el “promedio de calificaciones” (GPA) de los estudiantes, no mostró diferencias estadísticas significativas en términos de género, estado civil, ubicación de la residencia, campo de estudio y nivel educativo. Hubo una relación directa y significativa entre las estructuras de las estrategias motivacionales ($R = 0,193, p < 0,001$), así como las estrategias de aprendizaje ($R = 0,243, p < 0,001$) y el CGPA; mientras que no hubo relación entre las expectativas de resultado y el CGPA. El análisis de la trayectoria reveló que las estrategias de aprendizaje autorreguladoras y las estrategias de motivación pueden predecir el rendimiento académico de estos estudiantes.

Martín (2019) realizó un estudio descriptivo e inferencial con una muestra de 508 estudiantes de ciencias de la salud pertenecientes a dos universidades privadas españolas. Por medio del *Motivated Strategies for Learning Questionnaire*, se evalúa la utilización de estrategias cognitivas, metacognitivas, de gestión de recursos y motivacionales; después, se realizó un modelo de regresión, tomando estas estrategias como variables independientes y el rendimiento medido mediante la revisión de las notas de los estudiantes como variable dependiente. Se establecieron cuatro variables independientes para el criterio nota media del primer cuatrimestre: la variable motivacional autoeficacia para el rendimiento (AR), la estrategia de aprendizaje de gestión del tiempo y lugar de estudio (TL), la estrategia de aprendizaje con otros compañeros y búsqueda de ayuda (AA) y la meta de autoensalzamiento (MAS). La autoeficacia para el rendimiento es la variable que mejor predijo el rendimiento académico en los estudiantes universitarios de la muestra. Otras variables que también se relacionaron con el rendimiento: la gestión del tiempo y el lugar, la búsqueda de ayuda y la meta de autoensalzamiento. Se obtuvieron los siguientes valores del coeficiente B en orden de definición de variables: 0,3, 0,21, 0,17, 0,07. El coeficiente estandarizado B, nos informa del número de unidades que aumenta la variable dependiente (media de calificación en el primer cuatrimestre) por cada unidad que aumenta la variable independiente (definidas en el modelo) correspondiente. En la tabla 33 se presentan dichos resultados de la investigación.

Tabla 33. Coeficientes estandarizados B y tipificados Beta

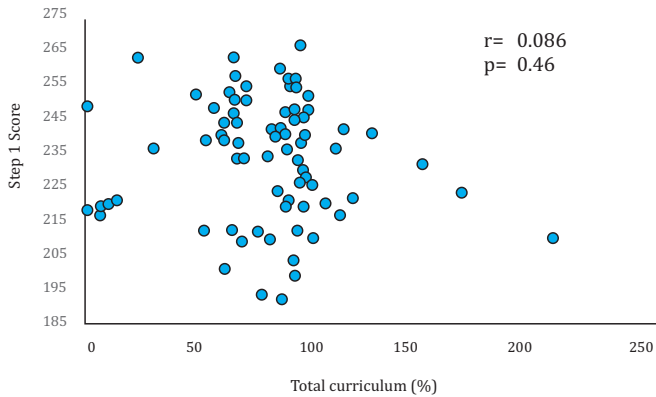
Variable	Coeficiente no estandarizado B	Coeficiente tipificado Beta	t	Sig T
AR	,30	,24	4,68	,00
TL	,21	,14	3,18	,00
AA	-,17	-,11	-2,32	,02
MAS	,07	,08	2,03	,04

Fuente: tomada de Martín (2019).

Kauffman *et al.* (2019), en su investigación, buscaron determinar si existía alguna relación entre el desempeño en la prueba USMLE step 1 (prueba que deben aprobar los estudiantes de medicina en Estados Unidos para optar

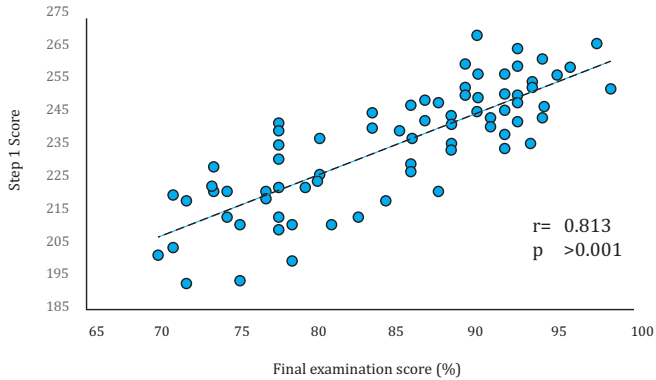
por su licencia médica) con las variables y asistencia a un curso de fisiopatología. Tomaron una muestra de 77 participantes. Se tuvieron en cuenta las siguientes variables: asistencia, acceso al plan de estudios formal, desempeño en evaluaciones internas de alto y bajo riesgo y acceso a materiales. En resumen, las variables asistencia y el acceso al plan de estudios formal no predicen el desempeño en la prueba USMLE step 1, mientras que el desempeño en evaluaciones internas de alto y bajo riesgo sí lo hacen. En las figuras 36 y 37 se presentan las gráficas de correlación, también conocidas como dispersión de datos, en las cuales se muestra el valor de correlación r , para el primer caso, este respecto a la asistencia; y en la segunda gráfica, en cuanto al examen final del curso.

Figura 36. Correlación entre la variable asistencia frente a la prueba USMLE step 1



Fuente: tomada de Kauffman *et al.* (2019).

Figura 37. Correlación entre el examen final del curso frente a la prueba USMLE step 1



Fuente: tomada de Kauffman *et al.* (2019).

Kwankajonwong *et al.* (2019) dedicaron su trabajo a investigar las relaciones entre el rendimiento académico de los estudiantes, su actitud, el conocimiento autoevaluado y la proporción de fisiología que se enseña en un bloque integrado de órganos y sistemas. El rendimiento académico se midió de la manera tradicional mediante evaluaciones formativas y sumativas. En este estudio participaron 130 estudiantes de fisiología en la medicina tailandesa. Se utilizó la correlación lineal para analizar las relaciones entre variables. Los resultados se resumen en virtud de que la variable tiempo de enseñanza dedicado a la fisiología se correlacionó de manera positiva con la actitud del estudiante ($r = 0,84$, $p = 0,001$) y se correlacionó de forma negativa con la calificación de autoevaluación ($r = -0,73$, $p = 0,007$). Los estudiantes de medicina tailandeses pueden beneficiarse de una actitud proactiva para estudiar fisiología, en contraste con las expectativas didácticas tradicionales de la educación tailandesa. La proporción del tiempo de enseñanza dedicado a la fisiología no influye en el rendimiento académico.

El objetivo del estudio de Herodotou *et al.* (2019) fue evaluar si al proporcionar a los profesores de una institución de educación superior a distancia, datos de los estudiantes, es posible predecir el rendimiento de los estudiantes y si permite a los profesores identificar y ayudar a los estudiantes en riesgo. Para tal fin, realizaron un análisis por medio de cuatro algoritmos: clasificador, Naive Bayes, árbol de clasificación y regresión (CART) y K-vecinos más cercanos (KNN). En los nueve cursos analizados, se presenta una variedad de actividades que incluyen la de asimilación, búsqueda y manejo de información, comunicación, actividades productivas, experimentales, interactivas y evaluación. Cada curso enfatiza algunas o la mayoría de estas actividades. Los mejores resultados mostraron que cuando un maestro accede al OUA (OUA es un sistema diseñado en la Universidad Abierta del Reino Unido, que utiliza una variedad de enfoques estadísticos y de aprendizaje automático avanzados para predecir los estudiantes en riesgo) entre un 1 % - 10 % más de la duración de las semanas del curso, los estudiantes tienen 7,4 veces más probabilidades de completar el curso. El éxito de la predicción, en general, es del 69,5 % (35,5 % por no aprobar un curso y 87,6 % por aprobar un curso). El criterio de Wald demostró que solo el uso semanal del tablero

OUA ($p = 0,001$) y la mejor puntuación del curso anterior de los estudiantes ($p = 0,001$) hizo una contribución significativa a la predicción.

Francis y Babu (2019) desarrollaron un modelo de predicción para evaluar el desempeño de los estudiantes en el ámbito académico, basado en técnicas de clasificación y agrupamiento y un conjunto de datos; para aplicarlo a los estudiantes de varias disciplinas académicas de instituciones de educación superior en Kerala (India). Se tuvieron en cuenta varias características de los estudiantes (véase tabla 34). Antes de plantear el nuevo algoritmo, experimentaron varios clasificadores: SVM, Naive Bayes, *decision tree* y *neural network*.

Se agruparon los estudiantes según las anteriores características en tres clases o grupos (alto, medio y bajo rendimiento). Agrupar los estudiantes para el desarrollo de modelos (modelo híbrido propuesto que combina agrupamiento y clasificación) da mejores resultados que experimentar sin métodos de agrupamientos (véase tabla 35). La precisión del modelo híbrido propuesto que combina agrupamiento y clasificación es de 0,7547, cuando se aplica a las características académicas, de comportamiento y adicionales del conjunto de datos de los estudiantes y se encuentra que es superior a la de los otros algoritmos existentes. Este modelo puede ayudar a los educadores a percibir que los alumnos reconocen a los alumnos débiles, para desarrollar el proceso de aprendizaje y reducir las tasas de fracaso de los académicos, y también ayuda a los administradores a gestionar mejor en función de los resultados del sistema de aprendizaje.

Tabla 34. Características de análisis

Categoría de característica	Subcaracterísticas	Descripción
Características demográficas	Género	Género de estudiantes
	Nacionalidad	Nacionalidad de los estudiantes
	Lugar de nacimiento	Lugar de nacimiento de los estudiantes
	Relación	Relación quién es responsable de los estudiantes

Categoría de característica	Subcaracterísticas	Descripción
Características académicas	ID de etapa	Etapa educativa de los estudiantes
	ID de grado	Nivel de grado de los estudiantes
	ID de sección	Sección del aula a la pertenece el alumno
	Tema	Tema del curso
	Semestre	Semestre del año escolar
Características de comportamiento	Manos levantadas	Comportamiento de los estudiantes durante el aprendizaje de la educación
	Recursos visitados	
	Vista de anuncio	
	Discusión	
Características adicionales	Encuesta de respuesta para padres	Encuesta respondida por padres y proporcionada por la escuela
	Satisfacción de los padres con la escuela	Satisfacción de los padres con respecto a las escuelas
	Días de ausencia del estudiante	Días de ausencia de los alumnos en las escuelas

Fuente: tomada de Francis y Babu (2019).

Vora y Rajamani (2019) pretenden predecir el desempeño educativo de los estudiantes a partir de la información sociodemográfica. Introducen una arquitectura de predicción de rendimiento que incluye dos módulos: uno es para manejar *Big Data* a través del marco MapReduce (incorpora el concepto de análisis de componentes principales), mientras que el otro es un módulo inteligente que predice el desempeño de los estudiantes usando etapas de procesamiento de datos. Para ello, este artículo presenta un nuevo clasificador híbrido, que hibrida aprendizaje profundo y la máquina de vectores de soporte. La implementación se llevó a cabo utilizando los detalles académicos de los estudiantes (17 atributos) y el detalle de los padres (13 atributos). El rendimiento del modelo de predicción propuesto, varía el porcentaje de entrenamiento al 50 %, 60 %, 70 %, 80 % y 90 % (véase tabla 36). Se observa que el modelo de predicción propuesto es superior a otros métodos con respecto a todas las medidas. En particular, la especificidad del modelo SVM propuesto con *deep learning* es 4,05 %, 58,57 % y 64,75 % mejor que DBN, SVM y NN.

Tabla 35. Resultados del nuevo enfoque

Algoritmo	Precisión	Recordar	Fscore	Exactitud
Agrupación	0,641509	0,641509	0,641509	0,641509
Árbol de decisión	0,622642	0,622642	0,622642	0,622642
Red neural	0,484277	0,484277	0,484277	0,484277

Fuente: tomada de Francis y Babu (2019).

Tabla 36. Resultados del nuevo modelo híbrido

Medidas	NN [34]	SVM [32]	DBN [33]	SVM con DBN
Especificidad	0,308411215	0,3625	0,842934372	0,875
Sensibilidad	0,029411765	0,126436782	0,791798107	0,796511628
Exactitud	0,37962963	0,669811321	0,680172414	0,714622642
Precisión	0,292929293	0,333333333	0,371428571	0,768292683
FPR	0,873563218	0,870588235	0,208201893	0,203488372
FNR	0,691588785	0,6375	0,159065628	0,025
VPN	0,029411765	0,126436782	0,791798107	0,796511628
FDR	0,707070707	0,666666667	0,628571429	0,231707317
F 1-puntaje	0,32038835	0,324022346	0,537931034	0,802973978
MCC	0,037979386	0,012964074	0,102889337	0,147059749

Fuente: tomada de Vora y Rajamani (2019).

Amazona y Hernández (2019) mostraron los resultados obtenidos por los algoritmos Naive Bayes, árboles de decisión, *deep learning* y *neural network* para analizar un conjunto de datos provenientes de 300 estudiantes de un Departamento de Tecnología de la Información. Variables como las historias académicas se utilizaron para entrenar a los modelos. Los resultados muestran que el clasificador de *deep learning* supera a otros dos clasificadores (árboles de decisión y Naive Bayes) al obtener un porcentaje de precisión general del 95 %. Los resultados para el algoritmo de árboles de decisión y Naive Bayes son del 93 % y 88 %, respectivamente.

Qowidho *et al.* (2019) compararon los algoritmos ID3 y C4.5, utilizando el *software* de RapidMiner. Se probaron los dos algoritmos con un conjunto de 230 datos, variando esta cantidad con el fin de obtener el mejor

porcentaje de precisión para cada uno. Al utilizar un conjunto de 150 datos, el algoritmo de C4.5 demostró tener un mayor porcentaje de precisión con 62,22 %; mientras que el ID3 obtuvo 51,11 % para la misma instancia. La tabla 37 muestra el porcentaje de precisión para los dos algoritmos con cada conjunto de datos utilizados.

Tabla 37. Porcentaje de precisión según el conjunto de datos

Test Data	Precisión	
	C4.5 (%)	ID3 (%)
25	28,57	28,57
50	33,33	33,00
100	43,33	50,00
150	62,22	51,11
200	43,33	40,00
230	34,78	42,03

Fuente: tomada de Qowidho *et al.* (2019).

Ajibade, Bahiah Binti Ahmad *et al.* (2019) utilizaron un conjunto de 500 datos provenientes de una plataforma de aprendizaje en línea conocida como KALBOARD 360. Variables asociadas con la interacción del estudiante en los módulos de la plataforma sirvieron de datos de entrada para la construcción de los algoritmos máquina de vectores de soporte (SVM), K-vecinos más cercanos (KNN) y árbol de decisión (DT). Además, con el fin de mejorar el rendimiento de los algoritmos, se aplicaron los métodos de conjunto, como Bagging, Boosting y bosque aleatorio. Los resultados obtenidos muestran que existe una fuerte relación entre las variables de comportamiento del estudiante como número de visitas al foro, asistencia a sesiones y recursos visitados, y su rendimiento académico. Se obtuvo una precisión del 91,5 % cuando los métodos de conjunto se aplicaron a los algoritmos para mejorar el rendimiento académico (véase tabla 38).

Tabla 38. Porcentaje de precisión de los algoritmos en dos instancias distintas

Algoritmo	Porcentaje de precisión	
	Resultados de prueba (%)	Resultados de validación (%)
DT	87,0	91,5
KNN	82,0	85,0
SVM	86,0	88,0

Fuente: tomada de Ajibade, Bahiah Binti Ahmad *et al.* (2019).

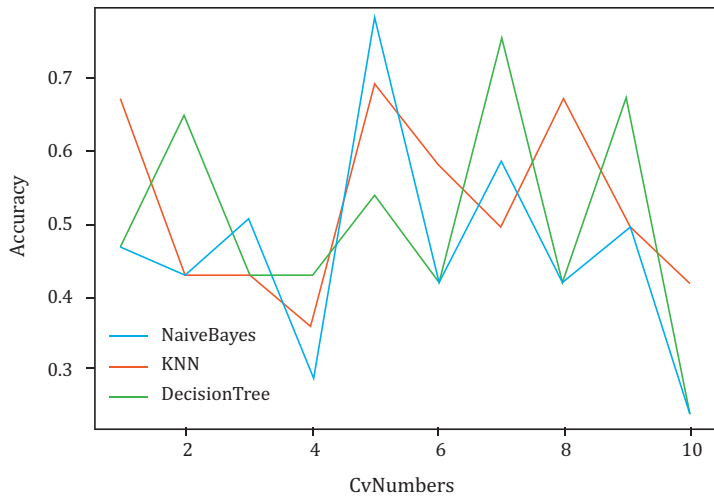
Mohammadi *et al.* (2019) utilizaron los clasificadores KNN, árbol de decisión y Naive Bayes en un conjunto de datos de 230 estudiantes de la Universidad de Kabul (Afganistán), para predecir su GPA como alto, medio y bajo. Los resultados muestran que el algoritmo de KNN obtuvo el mayor porcentaje de precisión con 54,64%, seguido del árbol de decisión con 53,25% y Naive Bayes con 46,16%. La figura 38 muestra la comparación del porcentaje de precisión de los modelos.

Wiyono *et al.* (2019) compararon los algoritmos KNN, SVM y árbol de decisión, con el fin de obtener un mejor porcentaje de precisión. El proceso de elaboración del modelo se hizo en cinco pasos: recolección de datos, preprocesamiento, construcción de modelos, comparación de modelos y evaluación. Los resultados muestran que el algoritmo SVM tiene la mejor precisión con un valor del 95%. El algoritmo del árbol de decisión tiene una precisión de predicción del 93% y el algoritmo KNN tiene un valor de predicción del 92%. El algoritmo KNN predijo que, de los 316 estudiantes activos, 308 eran activos y 8 inactivos y de un total de 66 estudiantes inactivos, predijo que 21 eran estudiantes activos y 45 inactivos. El algoritmo SVM predijo que, de los 316 estudiantes activos, 311 eran activos y 5 inactivos y de un total de 66 estudiantes inactivos, predijo que 13 eran estudiantes activos y 53 inactivos. El algoritmo de árbol de decisión predijo que, de los 312 estudiantes activos, 308 eran activos y 4 inactivos y de un total de 66 estudiantes inactivos, predijo que 18 eran estudiantes activos y 48 inactivos

Adekitan y Noma-Osaghae (2019) examinaron la relación entre los requisitos de admisión y el rendimiento académico de los estudiantes en su primer año, utilizando su CGPA y las calificaciones de sus cursos. Se construyeron

seis algoritmos en las plataformas KNIME y Orange. Se observó que el algoritmo de regresión logística presentó una precisión máxima de 50,23%. El estudio concluye que los requisitos de ingreso, son índices vitales para seleccionar aspirantes al ingreso en instituciones de educación superior. La tabla 39 presenta los resultados de cada uno de los modelos con cinco métricas de evaluación.

Figura 38. Matriz de comparación de los modelos



Fuente: tomada de Mohammadi *et al.* (2019).

Tabla 39. Resultados de los modelos utilizados

	Random Forest	Tree Ensemble	Decision Tree	Naive Bayes	Logistic Regression	Rprop. MLP
Correct Classified	210	205	172	186	218	213
Accuracy	48,38%	47,24%	38,631%	42,857%	50,23%	49,078%
Cohen's Kappa (k)	0,09	0,071	0,0041	0,106	0,081	0,084
Wrong Classified	224	229	262	248	216	221
Error	51,613%	52,765%	60,368%	57,143%	49,77%	50,922%

Fuente: tomada de Adekitan y Noma-Osaghae (2019).

Nieto *et al.* (2019) experimentaron con datos de la Facultad de Ingeniería de la Universidad Distrital (Bogotá). Entre los tres modelos de predicción, el bosque aleatorio tuvo la mayor precisión general (84,11 %), seguido de regresión logística (84,02 %), mientras que el árbol de decisión tuvo la tasa de precisión más baja (83,92 %). Los tres modelos tuvieron una tasa de precisión general relativamente alta que excedía el 80 %.

Francis y Babu (2019) utilizaron los algoritmos: máquina de vectores de soporte (SVM), Naive Bayes (NB), árboles de decisión (DT) y red neuronal (NN). Estos algoritmos evaluaron el rendimiento de los estudiantes basados en técnicas de clasificación y agrupación, y se alimentaron con un conjunto de datos de varias disciplinas académicas de instituciones de educación superior en Kerala (India). En la aplicación de los algoritmos SVM, NB, DT y NN usando características demográficas, SVM presentó la mayor exactitud con un porcentaje del 52,83 %, empleando características académicas arrojó que SVM logró el mayor porcentaje de precisión con 41,50 %, utilizando características de comportamiento SVM y DT obtuvieron el porcentaje más alto de exactitud con 54,71 %, aplicando características adicionales DT y NN presentaron la mayor exactitud con 62,26 %; mientras que NB la menor con 59,78 %, usando características adicionales sumadas con las de comportamiento, DT y NN presentaron la mayor exactitud con 62,26 %, mientras que NB la menor con 59,78 %, empleando características adicionales de comportamiento y académicos, DT presentó la mayor exactitud con 75,47 %, mientras que NB la menor con 52,83 %, utilizando características demográficas, académicas, de comportamiento y adicionales, DT presentó la mayor exactitud con 66,037 %, mientras NB la menor con 52,83 %.

Hussain *et al.* (2019) tomaron como objetivo predecir las dificultades a las que se podrían enfrentar los estudiantes en una sesión del curso de diseño digital posterior. Para ellos, se analizaron los datos registrados por un sistema de aprendizaje mejorado con tecnología (TEL, *technology enhanced learning*), llamado conjunto de educación y diseño de electrónica digital (DEEDS), utilizando algoritmos de aprendizaje automático. Se emplearon los siguientes algoritmos de aprendizaje automático: redes neuronales artificiales (ANN), máquinas de vectores de soporte (SVM), regresión logística, clasificadores Naive Bayes y árboles de decisión. Como resultado de la

investigación, se obtuvo que si el docente desea predecir los resultados de un individuo, así como la dificultad de una sesión completa de diseño digital al completar diferentes ejercicios y tareas durante el curso, las mejores alternativas serían los algoritmos SVM y ANN, ya que lograron una alta precisión (75 %), un valor estadístico Kappa (0,5), un alto valor ROC y bajos puntajes de RMSE (0,5 y 0,01, respectivamente), en comparación con los modelos Naive Bayes, regresión logística y árboles de decisión. Si el docente desea predecir la dificultad de un estudiante en la próxima sesión de un curso de diseño digital utilizando la menor cantidad de características, entonces se debe usar una máquina de vectores de soporte (SVM), seleccionando los atributos con la técnica Alpha-Investing. Se llega a esta conclusión, porque el algoritmo máquina de vectores de soporte (SVM) obtuvo la mayor precisión (80 %) en comparación con los modelos.

Lasarte *et al.* (2020) utilizaron un método de ensamble conocido como *bootstrap*, el cual es un método de remuestreo, que busca aproximar un determinado muestreo a una distribución de probabilidad. Esta investigación buscó determinar si el apoyo social por el profesorado y la familia, tiene influencia en la autoeficacia académica y en la implicación escolar, de una muestra de 1468 estudiantes. Se analizaron las siguientes variables: TCMS (subescala apoyo profesorado), AFA-R (subescalas apoyo familiar y apoyo amistades), SEM (escala implicación escolar) y EBAE-10 (subescala rendimiento académico percibido). Encontraron que las variables definidas tenían una influencia del 73 % en la autoeficacia académica y la implicación escolar de la muestra de estudiantes definida. En resumen, el profesorado y la familia deben ofrecer apoyo social al alumnado para reforzar la percepción de autoeficacia académica y la implicación escolar.

Tan *et al.* (2020) buscaron identificar la relación que pueden llegar a tener las variables de logro del estudiante, la motivación y las estrategias de aprendizaje con los tiempos de finalización, número de intentos y puntaje de unas determinadas pruebas. El estudio involucró 143 estudiantes de pregrado y se utilizó como método una regresión múltiple. El tiempo de finalización del examen fue la única variable de seguimiento del examen que mostró correlaciones de medianas a grandes con la motivación y el rendimiento del examen. El puntaje de la prueba y el número de intentos pueden

ser indicadores menos informativos, porque enredan la competencia con el uso ineficaz de la estrategia: los estudiantes competentes que emplean la práctica de recuperación pueden obtener puntajes altos en la prueba y menos intentos de prueba, así como los estudiantes que usan estrategias ineficaces (por ejemplo, buscar las respuestas).

Tenpipat y Akkarajitsakul (2020) analizaron los factores que afectan el estado educativo de los estudiantes universitarios, y crearon modelos de clasificación binaria para predecir su estado educativo si será abandonado u otro (es decir, otros estados que no son deserción) cuando estaban estudiando en la Universidad de Tecnología del Rey Mongkut Thonburi (KMUTT) (Bangkok, Tailandia). Los datos para el análisis se obtuvieron de la Oficina de Registro y Biblioteca de KMUTT. Los datos utilizados corresponden a 13 714 estudiantes de pregrado en siete cursos académicos, que van desde el curso 2012-2013 al 2019-2020. Se utilizaron algoritmos como el árbol de decisión, el bosque aleatorio y una clasificación de aumento de gradiente. Se agruparon las variables (30) en familiares, escolares, de admisión, de facultad y personal. Los resultados muestran que la precisión de predicción del aumento de gradiente, el árbol de decisión y el bosque aleatorio son del 93 %, 92 % y 92 %, respectivamente. Además, encontraron que las cinco características principales más importantes son el año académico del estudiante, el GPA de la escuela secundaria, los canales de admisión a la universidad, la facultad del estudiante y el género.

Medina *et al.* (2020) proponen un modelo de predicción que podría contribuir a reducir el índice de deserción de estudiantes universitarios en el Perú. Se compararon las técnicas de redes bayesianas con árboles de decisión, mediante una muestra de 500 estudiantes de pregrado. Estos datos se obtuvieron a partir de una aplicación web (encuestas virtuales). Las variables empleadas son demográficas y socioeconómicas identificadas. Para realizar la validación de la implementación del modelo se utilizó el método de validación cruzada. La validación cruzada permitirá evaluar los resultados de la eficiencia y tasa de error del modelo, por medio de una clasificación de los datos para las actividades de entrenamiento y pruebas. La evaluación de algoritmos se efectuó en tres iteraciones; en cada iteración se dividió la información recopilada en dos conjuntos de datos distintos: uno

para el entrenamiento del modelo y el otro para la validación o comprobación del modelo. En cada iteración, se manejó una razón distinta entre los datos de entrenamiento y los de validación: en la primera iteración, la razón fue de 6 a 4, en la segunda de 7 a 3 y en la tercera de 8 a 2 (véase tabla 40). Los resultados indican que las redes bayesianas se comportan mejor que los árboles de decisión con base en las métricas de precisión, exactitud, especificidad y tasa de error. Particularmente, la precisión de las redes bayesianas alcanza 67,10 %, mientras que para los árboles de decisión es de 61,92 % en la muestra de entrenamiento para la iteración con razón de 8:2. Por su parte, las variables persona deportista (0,29 %), vivienda propia (0,20 %) y calificaciones de preparatoria (0,15 %) son las que más contribuyen al modelo de predicción, tanto para las redes bayesianas como para los árboles de decisión.

Tabla 40. Métricas de evaluación en la interacción 8:2

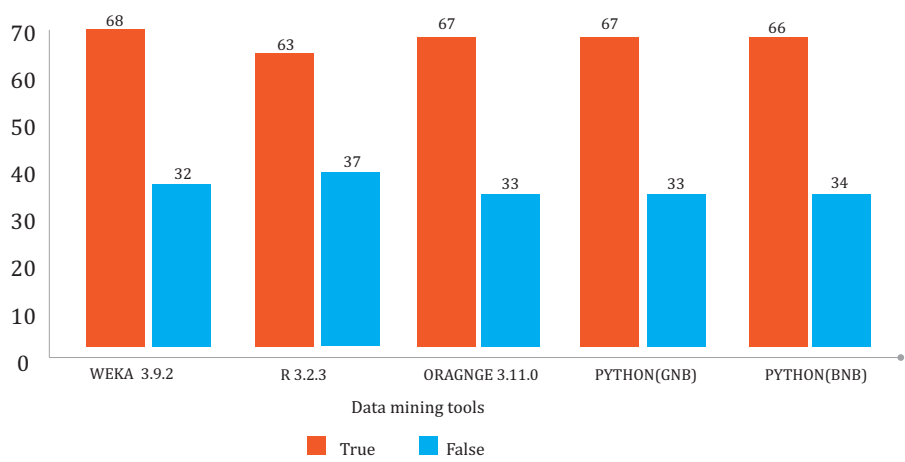
	Árbol decisión		Red Bayesiana	
	Entrenamiento (%)	Validación (%)	Entrenamiento (%)	Validación (%)
Exactitud	61,92	57,02	67,10	74,56
Especificidad	61,88	53,73	69,80	80,60
Precisión	59,69	48,33	65,92	70,45
Tasa de error	38,08	42,98	62,92	25,44

Fuente: tomada de Medina *et al.* (2020).

Mavani *et al.* (2020) realizaron un comparativo de los diferentes *softwares* utilizados para la construcción del algoritmo Naive Bayes; para ello, se utilizó un conjunto de 200 registros de estudiantes del Instituto de Tecnología de San Francisco (SFIT) (Mumbai, India). La figura 39 muestra la comparación entre el número de instancias clasificadas correcta e incorrectamente para el algoritmo Naive Bayes en cada una de las herramientas de *software*. De la figura, se puede inferir que todas las herramientas muestran resultados pronosticados similares. El más preciso es Weka, el cual tiene una precisión ligeramente mayor en comparación con la de otras tres herramientas, con 68 instancias correctas y 32 incorrectas. Para obtener resultados precisos, en segundo lugar, se encuentran Orange y Python (GNB), que muestran 67

instancias correctas y 33 incorrectas. La menor precisión se encontró en R con 63 instancias correctas y 37 incorrectas. Podemos ver que casi todas las herramientas proporcionan relativamente la misma predicción, pero si consideramos el resultado más preciso, Weka ofrece la predicción más precisa del 68 %. Esto significa que, de cada 100 estudiantes, el algoritmo trabajado en Weka predijo correctamente la colocación de 68 estudiantes en función de sus puntajes académicos. Orange y Anaconda (P-GNB) proporcionaron una precisión de 67%.

Figura 39. Rendimiento Naive Bayes en diferentes *softwares*



Fuente: tomada de Mavani *et al.* (2020).

Mengash (2020) se centró en tratar de predecir el rendimiento académico de los aspirantes a la universidad. Para su desarrollo, utilizó un conjunto de datos de 2039 estudiantes matriculados en una facultad de informática de una universidad pública (Saudi) de 2016 a 2019. Los resultados demuestran que el rendimiento universitario temprano de los solicitantes se puede predecir antes de la admisión en función de ciertos criterios (promedio de la escuela secundaria, puntaje de la Prueba de Admisión al Logro Académico y puntaje de la Prueba de Aptitud General). Los resultados también muestran que el puntaje de la Prueba de Admisión es una variable incidente en el rendimiento de los aspirantes. Por tanto, a este puntaje se le debe asignar más peso en los sistemas de admisión. Los resultados muestran que el algoritmo de red neuronal artificial tiene una tasa de precisión superior al 79%, por lo que es superior

a otras técnicas de clasificación consideradas (árboles de decisión, máquinas de vectores de soporte y Naive Bayes).

Tomasevic *et al.* (2020) tomaron como objetivo realizar un análisis detallado para comparar las técnicas de aprendizaje automático supervisadas de última generación, aplicadas a predecir el rendimiento de los estudiantes en los exámenes; es decir, descubrir a los estudiantes con un “alto riesgo” de abandonar el curso. Se planteó una combinación óptima de características relacionadas con el alumno (D - demografía, E - compromiso, P - datos de rendimiento). La precisión más alta, en general, se obtuvo con la red neuronal artificial (ANN), teniendo como variables de entrada los datos del compromiso y el rendimiento pasado. Seguida por la máquina de vectores de soporte (SVM) de núcleo lineal y radial. Los resultados más bajos se obtuvieron con el enfoque de Naive de Bayes. El promedio de error cuadrático medio (RMSE) representa la desviación estándar de la diferencia entre el valor predicho y el valor real para cada uno de los algoritmos usados. La red neuronal artificial (ANN) obtuvo el error más bajo.

10.3. Metodologías para evaluación del rendimiento académico

Las siguientes metodologías, las cuales se publicaron en un periodo comprendido entre 2011 y 2021 y en contextos diversos, están clasificadas y presentadas en tres grandes áreas de estudio: Psicología, Estadística-Matemática y Medicina. Para cada una de ellas, se mencionan sus principales características, los análisis y los resultados obtenidos.

El criterio de clasificación de las metodologías consistió en la revisión de la naturaleza de los instrumentos utilizados en ellas, así como los propósitos para las cuales se consolidan. Para las metodologías orientadas en la Psicología, se tuvieron en consideración los cuestionarios aplicados por los estudiantes (o docentes, en algunos casos), en donde demostraron su percepción sobre un tema en específico (generalmente en escalas Likert, que miden actitudes y grados de conformidad a través de categorías de respuesta, o mediante cuestionarios psicológicos que evalúan, entre otras cosas, niveles de ansiedad, habilidades sociales, autoestima). En cambio, en las metodologías orientadas en la Estadística-Matemática, se usaron datos específicos

de la actividad académica de los estudiantes, así como sus relaciones con los datos demográficos, sin los cuestionarios cualitativos. Finalmente, en la metodología centrada en el ámbito de la Medicina, utiliza procedimientos específicos del área mediante diversas mediciones para el diagnóstico de procesos fisiológicos que pueden incidir en el rendimiento académico. Cabe mencionar que pueden existir otros trabajos de investigación no contemplados en este documento, que no fueron descubiertos por medio de las palabras claves y criterios establecidos.

10.3.1. Psicología

A continuación, se presentan varios estudios en los cuales la fuente principal de información son las percepciones, opiniones y consideraciones de estudiantes y docentes, mediante encuestas y test para la medición y evaluación del rendimiento académico y las variables de influencia sobre este.

Hernández (2016) lleva a cabo una aproximación hacia la determinación de las variables que influyen en el rendimiento académico desarrollado en México durante el 2014, el cual tuvo como objetivo la realización de un diagnóstico de la percepción de los estudiantes, sobre aquellos factores que, según ellos, podrían llegar a influir en el rendimiento académico. Se aplicó una encuesta de 30 preguntas tipo Likert que representan variables sobre el rendimiento académico (obtenidas a partir de la revisión de la literatura), a 1651 estudiantes (43% mujeres y 57% hombres) de UPIICSA (Unidad Profesional Interdisciplinaria de Ingeniería y Ciencias Sociales y Administrativas) del IPN (Instituto Politécnico Nacional) de cinco carreras diferentes (Ingeniería industrial 61%, Licenciatura en administración industrial 24%, Ingeniería de transporte 12%, Ingeniería en informática 2%, Ciencias de la informática 1%), sobre una población total de 13 000 estudiantes. En este estudio diagnóstico, el indicador de medición del rendimiento académico está dado por los resultados obtenidos de la encuesta aplicada a los estudiantes. Las variables independientes en este estudio corresponden a algunos datos demográficos, así como información sobre asignaturas o tiempos de movilización: género, carrera, turno, semestre, número de asignaturas no aprobadas, tiempo de trayecto hacia la escuela,

trabajo, tutor, promedio. Las variables independientes es el rendimiento académico a partir de las respuestas de la encuesta.

De manera más específica, el cuestionario se realizó para obtener respuestas entre 1 y 5 (“totalmente en desacuerdo” y “totalmente de acuerdo”, respectivamente), y sus preguntas se dividieron en dos partes:

- **Primera parte:** variables independientes.
- **Segunda parte:** presenta los ítems sobre el rendimiento académico para ser evaluados en la escala tipo Likert.

Considerando que el análisis factorial exploratorio es una técnica usada en el desarrollo y validación de instrumentos de medida psicológicos, al explorar cómo un conjunto de variables latentes caracterizan a una persona según sus respuestas a los ítems (Lloret-Segura *et al.*, 2014). En el presente diagnóstico se realizó un AFE por medio de la reducción de dimensiones en el *software* SPSS (Statistical Package for the Social Sciences), obteniendo resultados para la prueba de Kaiser Meyer Olkin de $0,936 > 0,50$, lo cual indica la existencia de un mayor grado de correlación entre los ítems, y, por tanto, da la posibilidad de reducir los datos (variables estudiadas) a factores. A partir de la extracción de componentes principales y rotación Varimax, se determinó que los nuevos factores serían: rol de los profesores, elementos de evaluación y motivación extrínseca del estudiante. Como resultado de este diagnóstico, se evidencia la importancia de los profesores en el rendimiento académico, incluso sobre los otros factores determinados; pues a partir de ellos, se ejecuta una correcta evaluación de las temáticas, así como su potencial influencia en la motivación del estudiante. Los profesores, además, son vistos como acompañamiento y respaldo respecto a las dificultades y metas académicas. Sin embargo, este estudio diagnóstico se vio limitado por la participación de otras carreras para determinar sus propias percepciones y si existe diferencia respecto a las estudiadas. Además, el límite también está dado por las opciones cerradas de las preguntas que pudieran no representar correctamente el sentir de los estudiantes.

Mushtaq y Khan (2012) buscaron conocer cuáles son los factores que afectan el rendimiento académico; se centraron en las universidades privadas de Pakistán de las ciudades Rawalpindi e Islamabad. Esta investigación es,

además, innovadora, pues incluye la variable “orientación adecuada” y su efecto sobre el rendimiento académico. Esta y otras variables se midieron en una escala Likert de 20 ítems, cada uno de ellos calificado en una escala de 5 puntos (que oscilan entre “muy en desacuerdo” y “muy de acuerdo”, de 1 a 5, respectivamente) en un total de 175 cuestionarios distribuidos; de estos, 155 se consideraron como válidos para la muestra. La forma de medición del rendimiento académico se realiza mediante el análisis de los resultados de los test aplicados. Este estudio también se caracteriza por plantear una serie de hipótesis con respecto a la revisión previa de la literatura, una por cada variable que puede afectar el rendimiento de los estudiantes, y de la comprobación de las mismas dependen los resultados de la investigación:

- **H1:** “existe una relación positiva entre la comunicación y el desempeño del estudiante”.
- **H2:** “existe una relación positiva entre las instalaciones de aprendizaje y el desempeño de los estudiantes”.
- **H3:** “existe una relación positiva entre la orientación adecuada y el desempeño del estudiante”.
- **H4:** “existe una relación negativa entre el estrés familiar y el desempeño del estudiante”.

Este estudio toma como indicador de medición del rendimiento académico los resultados obtenidos en el cuestionario, de la sección de “rendimiento académico”. La variable dependiente en este estudio corresponde al desempeño del estudiante, mientras que las variables independientes son la “orientación adecuada”, “habilidades de comunicación de los estudiantes”, “instalaciones de aprendizaje” y “estrés familiar”. Todas las variables se obtuvieron mediante un cuestionario. Se empleó el Alpha de Cronbach para determinar la fiabilidad de todos los ítems del instrumento, con el que se obtuvo un resultado de 0,710. De igual forma, se aplicó esta misma herramienta para la confiabilidad de los ítems individuales, clasificados por las variables “orientación adecuada”, “habilidades de comunicación de los estudiantes”, “instalaciones de aprendizaje” y “estrés familiar”, así como el “desempeño

del estudiante”, obteniendo así confiabilidades de 0,806, 0,497, 0,735, 0,258 y 0,716, respectivamente.

Las correlaciones entre variables se realizaron con el coeficiente de correlación de Pearson, con el cual se determinó un grado de asociación entre las habilidades de comunicación y el rendimiento de los estudiantes de alrededor del 13 %, así como también la variable dependiente tiene relación con las instalaciones de aprendizaje, orientación adecuada y estrés familiar (13 %, 20 % y 2 %, respectivamente). Respecto al análisis de regresión, se hizo un resumen del modelo, así como el análisis de varianza (ANOVA). En el primer caso, R^2 es igual a 0,553, lo que indica que el 55 % de la variación en el desempeño del estudiante es resultado de las variables independientes. Para el segundo caso, y al determinar el estadístico F, se obtuvo un valor de 20,083 que sugiere que el modelo es altamente significativo. Finalmente, el cálculo de los coeficientes estandarizados y no estandarizados de la regresión permite la confirmación o anulación de las hipótesis planteadas.

Como se planteó en un principio, con la confirmación de las hipótesis planteadas, se podría determinar el efecto de las variables predictoras sobre la variable dependiente. De acuerdo con lo anterior, los coeficientes Beta del modelo indican que:

- Las habilidades comunicativas del estudiante causan el 20 % de variación positiva en el rendimiento académico. Aunado a lo anterior, el valor t también es significativo, por lo que se acepta la primera hipótesis. Esto se refiere a las habilidades comunicativas del estudiante, así como su nivel de competencia en inglés.
- Las instalaciones del aprendizaje causan el 16 % de la variación positiva en el rendimiento académico del estudiante. El valor t , al igual que en el primer caso, es también significativo, por lo que se acepta la segunda hipótesis. En este caso, se considera el entorno en el que el estudiante se desenvuelve y los elementos de este, como el acceso a bibliotecas y computadores.
- La orientación adecuada, es la variable innovadora y especial de este estudio, causa el 17 % de la variación positiva en el rendimiento académico del estudiante. El valor t es significativo, y esta hipótesis también es

aceptada. Esto significa que la orientación dada por los padres y docentes es importante en el rendimiento académico.

- El estrés familiar causa el 13 % de la variación negativa en el desempeño académico; además, el valor t es insignificante, por lo que esta es la única hipótesis rechazada. En esta variable se encuentran los siguientes contenidos: el ingreso familiar, la educación de los padres y otros factores socioeconómicos.

De este modo, tanto las habilidades comunicativas del estudiante (las cuales incluyen las habilidades en el idioma inglés), como las instalaciones del aprendizaje y la orientación adecuada, afectan de manera positiva el desempeño académico del estudiante. Por otro lado, el estrés familiar incide de forma negativa en el rendimiento académico. Las limitaciones en este estudio pueden estar dadas, en principio, por el tamaño y procedencia de la muestra, el cual fue pequeño y reducido únicamente a dos ciudades. También pudieron quedar excluidos otros factores que afectan o pueden afectar el rendimiento académico.

Furnham *et al.* (2013) buscan en su estudio determinar los efectos de los rasgos de la personalidad de los estudiantes sobre el rendimiento académico. Este estudio establece la relación entre estas variables al examinar la consistencia entre los resultados autorreportados por los estudiantes (mediante cuestionarios) y los resultados específicos de su rendimiento académico. Se empleó una muestra de 1013 estudiantes (662 mujeres y 351 hombres) de cuatro universidades británicas en distintas facultades (12,5 % artes/humanidades, 4 % arquitectura, 11,6 % leyes, 21 % ciencias sociales e históricas, 27,3 % ciencias de la vida y medicina, 3 % biomédica, 3,3 % ingeniería, 7,2 % matemáticas/física, 7 % desconocido). Las edades de estos estudiantes oscilaban entre 18 a 51, con una edad promedio de 21,33 años y una desviación estándar igual a 3,76. El indicador de medición del rendimiento académico consiste en los resultados de dos métodos de evaluación distintos: exámenes y trabajo de curso. Mientras que los exámenes hacen referencia a la evaluación de conocimiento y habilidades en periodos específicos de tiempo, el trabajo de curso reúne diversas actividades, desde proyectos a gran escala hasta cortas tareas escritas. Las variables exógenas están relacionadas con

la personalidad, provenientes como resultado del test “BIG FIVE”, así como el dato demográfico del género, mientras que la variable endógena es el rendimiento académico.

A los participantes del estudio se les indicó que completaran una versión resumen del test “Big Five Inventory”, consistente de tres ítems, orientados hacia la evaluación de la personalidad de cada uno de ellos, así como también se les requirió una medida autorreportada sobre su propio rendimiento académico. Se recolectaron estos datos mediante una página web construida para tal fin, que permitía proveer la información de manera “anónima, no intrusiva y pragmática”.

Posteriormente, se utilizó el modelamiento de ecuaciones estructurales (SEM) para construir un modelo con diferencias individuales para las variables predictoras del rendimiento académico. La bondad de ajuste del modelo se evaluó mediante el estadístico chi-cuadrado (X^2), el índice de bondad de ajuste y su versión ajustada, así como la raíz cuadrada media residual y el índice de bondad de ajuste de parsimonia. En total, en esta investigación se construyeron dos modelos: el primero incluía los resultados del test “BIG FIVE” como variables exógenas, y el factor de RA como variable endógena. El segundo modelo, se añadió el género como otra variable exógena.

En el primer modelo (considerando únicamente los factores del “BIG FIVE” como exógenos) se ejecutó un análisis multigrupo para medir la invarianza entre un género y otro; considerando, además, la previa evaluación de un modelo anterior para cada grupo (M-F) de forma individual. Este modelo original se hizo sin restricciones entre grupos, mientras que el modelo posterior se realizó con restricciones. Se obtuvo que el modelo de línea de base era invariante para ambos géneros. Este mismo análisis también se aplicó para determinar la invarianza entre grupos según el tipo de estudios (carreras), por lo que el modelo de línea de base se modificó al agregar una nueva variable llamada “amabilidad-neuroticismo”, con el fin de explicar los datos de la muestra de estudiantes de artes y humanidades (al igual que con los géneros, también se estableció un modelo de línea de base sin restricciones y otro con restricciones, dando como resultado que el modelo era invariante entre tipos de grado).

A partir de los resultados de este estudio, es posible afirmar que las variables de la personalidad son “estables, robustas y predecibles” en la correlación y determinación del rendimiento académico. De manera más específica, los rasgos de personalidad más fuertes en la predicción positiva del rendimiento académico son la conciencia, amabilidad y franqueza (mientras que la conciencia y la amabilidad se relacionan con los dos métodos de evaluación del rendimiento académico, la franqueza solo estaba relacionada con las notas de trabajo de curso); pues las personas amables tienen la capacidad de brindar, solicitar y recibir ayuda de sus profesores y compañeros. Al contrario, la extraversión y el neuroticismo se relacionan a un desempeño mucho más bajo, lo que se explica al existir mayores distracciones y una preferencia por socializar, así como una memoria a largo plazo más débil.

Por otra parte, el papel del género se demostró por medio de la correlación negativa y modesta con ambos métodos de evaluación, con una pequeña diferencia en donde el género femenino evidenció un desempeño superior. Sin embargo, en el modelo de línea de base, los parámetros son invariantes entre “femenino” o “masculino”, pues los coeficientes Beta son idénticos en ambas muestras; este comportamiento también fue similar en el test “Big Five”, ya que los efectos eran similares entre géneros. El género y la personalidad tienen efectos invariantes en las diferentes áreas de estudio abordadas en este estudio (humanidades, ciencias sociales, ciencias de la vida, ciencias duras). Cabe aclarar que las variables de personalidad evidenciaron mayor varianza en el trabajo de curso que en los exámenes, lo que explica que las diferencias individuales entre estudiantes son más notorias en situaciones menos construidas y controladas (como ocurre normalmente en un examen). Este estudio se ve limitado por una medida corta del rasgo de personalidad y las notas autorreportadas por los estudiantes, ya que no se conoce a ciencia cierta su historia académica o sus actitudes reales hacia las mediciones.

Zuniga-Jara *et al.* (2018) analizaron la relación entre conductas alimentarias de riesgo y rendimiento académico en universitarios. Los autores se basan en una hipótesis central (con consecuencias en cadena), la cual enuncia que cuando los estudiantes poseen un mayor SaR (sensibilidad a recompensa), se incide en el tipo de alimentación que consumen, conduciendo

a un desorden alimenticio, que se traduce en sobrealimentación y preferencia por alimentos dulces y grasos, que se refleja en sobrepeso (IMC), y finalmente a un menor desempeño académico. La muestra se constituyó de 173 estudiantes de la Universidad Católica del Norte (Coquimbo, Chile) de diversas carreras, a quienes se les aplicaron una serie de encuestas, así como la medición de su masa corporal. La media de edad de estos estudiantes es de 21-40 años. En este estudio, el indicador de medición del rendimiento académico corresponde al número de asignaturas reprobadas por un estudiante; es decir, que un menor número de asignaturas reprobadas es equivalente a un buen rendimiento, contrario a lo que sucedería cuando el número de materias reprobadas es mayor. Las variables independientes en esta metodología corresponden a aquellas relacionadas con los constructos establecidos (sensibilidad a la recompensa, inadecuada calidad dietaria por sobrealimentación y preferencia por alimentos grasos o dulces), mientras que las variables dependientes son el índice de masa corporal de los estudiantes (IMC) y eventualmente el rendimiento académico.

Se utilizó una especificación del modelamiento de ecuaciones estructurales, en la cual se postula que la SaR se relaciona con una inadecuada calidad dietaria y preferencia por alimentos grasos o dulces. Ambos afectan el IMC de los estudiantes, y esto finalmente se vincula con el rendimiento académico. Con esta técnica de análisis estadístico multivariado se analizaron las relaciones estructurales latentes, y para la cual se usó, además, la función de ajuste de máxima verosimilitud. Estos constructos (sensibilidad a la recompensa, inadecuada calidad dietaria por sobrealimentación y preferencia por alimentos grasos o dulces) se midieron y se observaron mediante variables, obtenidas con cuestionarios específicos:

- **Vulnerabilidad por SaR:** se empleó el cuestionario BAS (Behavioural Activation System), en donde se mide “el sistema de aproximación conductual ante el placer y la recompensa” por medio de tres subescalas (pulsión; búsqueda de diversión, placer o sensaciones; y responsividad a la recompensa). En este cuestionario, que había demostrado confiabilidades satisfactorias en la teoría consultada por los autores mediante coeficientes de Cronbach, se hicieron preguntas como:

- “Cuando estoy haciendo algo bien, me encanta seguir haciéndolo”.
 - “Siempre estoy dispuesto(a) a probar algo nuevo, sí creo que es divertido”.
 - “Si veo la oportunidad de obtener algo que quiero, me muevo de inmediato hacia ello”.
- **Sobrealimentación:** este constructo se midió por medio de dos cuestionarios (los cuales han mostrado confiabilidades satisfactorias con el Alpha de Cronbach) para observar “la tendencia a sobrealimentarse en respuesta a ciertos desencadenantes”. El primer cuestionario es “The Dutch Eating Behaviour Questionnaire (DEBQ)”, donde se aplicaron subescalas de alimentación emocional y alimentación externa. Algunas de las preguntas en este cuestionario son:
 - “¿Deseas seguir comiendo cuando es delicioso?”.
 - “¿Comes algo inmediatamente cuando es delicioso?”.
 - **En cuanto al segundo cuestionario, “Binge Eating Scale (BES)”**, este evaluó el comportamiento de la persona frente a la comida.
 - **Preferencias alimentarias:** se usó el “Cuestionario de Preferencias Alimentarias (CPA)” con adaptaciones para incluir 42 alimentos representativos para la población chilena. Se utilizó una escala de Likert de cinco ítems, desde alto a bajo contenido de grasas, y desde alto en azúcares simples, hidratos de carbono de alta complejidad hasta bajo en carbohidratos/alta proteína. Algunos de los alimentos en este cuestionario son:
 - “Helados de crema y chocolate”.
 - “Chocolates”.
 - “Alfajor relleno con manjar cubierto de chocolate”.
 - “Leche entera fluida o en polvo”.

Al finalizar la aplicación de los cuestionarios y la tabulación de sus resultados, se realizó un análisis factorial confirmatorio (CFA) para examinar la validez de los constructos a través de un SEM reducido (con covarianzas entre las variables independientes, pero sin las dependencias entre ellas). A partir de estos resultados, sí fue posible estimar el modelo estructural completo bajo la selección de los constructos que proveían los mejores resultados de acuerdo con la validez convergente y discriminante, por lo que se consideraron válidos. Los coeficientes obtenidos en el modelo de ecuaciones estructurales evidencian que todas las relaciones estructurales planteadas son positivas, tal como se planteó en la hipótesis inicial: SaR está relacionado con la sobrealimentación y la preferencia por alimentos dulces y grasos. Estos dos, a su vez, evidencian un efecto positivo sobre el IMC, y este finalmente evidencia que altos IMC influyen de manera positiva con un mayor número de asignaturas reprobadas (menor rendimiento académico). Lo anterior plantea que las características psicológicas frente a los comportamientos alimenticios de los estudiantes sí tienen una influencia indirecta sobre el rendimiento académico (es indirecta, pues estas respuestas psicológicas influyen primeramente en el IMC y este en el rendimiento académico).

Chilca (2017) plantea su hipótesis de la existencia de una relación entre la autoestima y los hábitos de estudio con el rendimiento académico de estudiantes universitarios, considerando que la población de estudio correspondía a 196 estudiantes de la Facultad de Ingeniería de la Universidad Tecnológica del Perú (UTP), matriculados en el curso de matemática básica durante el semestre académico 2016 - III en seis aulas en el turno de la mañana. En este estudio se empleó una muestra de 86 estudiantes bajo la fórmula del muestreo aleatorio simple. En esta metodología, el promedio de calificaciones de los estudiantes en la materia “matemática básica I”, corresponde al indicador de medición del rendimiento académico. Las variables independientes están relacionadas con la autoestima y hábitos de estudio, obtenidas de cuestionarios, mientras que la variable dependiente es el rendimiento académico.

La obtención de datos para las variables independientes (autoestima y hábitos de estudio) estuvo dada a partir de la aplicación de dos instrumentos: el inventario de autoestima de Coopernsmith y el inventario de hábitos de

estudio de Vicuña. El primero presenta un coeficiente de confiabilidad de Kuder-Richardson (0,64) para su validez y confiabilidad, y está constituido por 58 ítems con respuestas dicotómicas, de aplicación individual y colectiva, que a su vez posee cinco subtest (“Test de mentira”, “Sí mismo”, “Social pares”, “Hogar padres”, “Universidad”). En este test es posible obtener un puntaje de hasta 100 puntos cuando se suman todos los ítems que responden correctamente (con posible invalidación de la prueba si en el test de mentiras se obtenía un puntaje superior a 4) y luego multiplicado por 2 (sin contar el puntaje de mentiras). Esta sumatoria permite ubicar al estudiante en una categoría, según el baremo propuesto por el autor.

El segundo instrumento evidencia una forma de evaluación similar al anterior, se compone de 53 ítems con respuestas dicotómicas, con aplicación individual y colectiva, y también presenta cinco subtest (“¿Cómo estudia usted?”, “¿Cómo hace sus tareas?”, “¿Cómo se prepara para sus exámenes?”, “¿Cómo escucha clases?”, “¿Qué acompaña sus momentos de estudio?”). En este test se puede obtener un puntaje máximo de 53 puntos mediante una calificación de 1 o 0 a cada ítem. Luego de la sumatoria total, se ubica al estudiante en una categoría, según el baremo planteado por su autor. Para este cuestionario, la validez y confiabilidad están dadas a partir del coeficiente de confiabilidad de Kuder Richardson (0,83).

Para el proceso de prueba de hipótesis, fue necesario, primero, evaluar la normalidad de los datos, seguido del proceso de prueba de hipótesis mediante análisis de regresión lineal múltiple (con una previa comprobación de linealidad, independencia de errores, homocedasticidad de residuos y pronósticos, normalidad de los errores, multicolinealidad). También fue necesario llevar a cabo la evaluación de bondad de ajuste al usar el análisis de varianza ANOVA, para finalmente realizar el cálculo del coeficiente de correlación de Pearson, así como los coeficientes del modelo y su significancia estadística (para todos los casos, la significancia estadística tiene un nivel de 0,05). Este análisis estadístico se hizo con la ayuda del *software* SPSS v.22.0.

El porcentaje de la varianza del rendimiento académico que se explica por la autoestima y hábitos de estudio es de 24,5 %, lo que indica que la correlación entre las variables es significativa pero moderadamente baja. En

este mismo sentido, el análisis de varianza demostró que existe una correlación significativa entre las dos variables independientes y el rendimiento académico. Sin embargo, al calcular los coeficientes del modelo de regresión lineal múltiple, se observa que con la prueba *t* de student, existe significancia estadística únicamente para el coeficiente de la variable hábitos de estudio —pues su valor *p* (0,000) es menor a la significancia estadística (establecida como 0,05), pero no para la variable autoestima (la cual tiene un valor *p* igual a 0,401)—.

Los anteriores resultados también se apoyan en que cerca del 41,9% de los estudiantes tienen bajo rendimiento académico, y, sin embargo, casi todos los estudiantes tienen un nivel de autoestima entre “medio alto” y “alto” (94,2%); es decir, que independientemente de su rendimiento académico, el nivel de autoestima de los estudiantes es bueno. En términos específicos, para este estudio, esto puede estar dado porque la muestra se constituye por estudiantes con estudios técnicos previos, o que es su segunda carrera profesional con estudios de noche. Con los resultados obtenidos, se descarta parcialmente la hipótesis inicial, pues si bien los hábitos de estudio sí influyen en el rendimiento académico, este no es el caso para la autoestima.

Ortiz *et al.* (2020) pretenden analizar la relación existente entre la ansiedad matemática percibida y el rendimiento académico en estudiantes de ingeniería de una universidad privada de Barranquilla (Colombia), con una muestra de 87 estudiantes (ingeniería industrial, 43,2%; ingeniería de mercados, 20,7%; ingeniería de sistemas, 36,4%), de I a II semestre durante el periodo académico 2017-2, en donde predominó el género masculino (65,95% de participación). El indicador de medición del rendimiento académico en este estudio corresponde a los resultados del semestre cursado, según si los estudiantes aprobaron o desaprobaron. Las variables independientes son las relativas a la ansiedad matemática (hacia el aprendizaje o hacia la evaluación), edad, sexo e ingeniería, obtenidas por medio de un cuestionario específico y los datos demográficos. La variable dependiente del modelo es el rendimiento académico.

Se utilizó el instrumento “Escala Abreviada de Ansiedad Matemática” (AMAS) de Hopko en su adaptación al español, para evaluar la ansiedad a las matemáticas en dos dimensiones (ansiedad hacia el aprendizaje

y la ansiedad hacia la evaluación). Este se compone de nueve ítems, con los cuales los estudiantes pueden determinar el nivel de ansiedad que les genera diversas situaciones relacionadas con las matemáticas. El Alpha de Cronbach del instrumento en su totalidad fue de 0,809, para la subescala de ansiedad en el aprendizaje de las matemáticas fue de 0,793 y para la subescala de la ansiedad en la evaluación de las matemáticas de 0,704. A los resultados obtenidos mediante el instrumento anterior, se les realizó un análisis descriptivo (representación gráfica de las variables mediante tablas de distribución de frecuencia y porcentajes) e inferencial (para contrastar las hipótesis planteadas) mediante el *software* SPSS v 20. La normalidad de los datos se comprobó con la prueba de Kolmogorov-Smirnov (pues la muestra es mayor a 50 individuos), con la cual se determinó que la muestra no se comporta de forma normal ni en el total del cuestionario, ni en la subescala de ansiedad matemática y tampoco en la subescala de ansiedad en la evaluación. También, se llevó a cabo un análisis correlacional a través del coeficiente de Spearman para comprobar la forma en que se relacionan las distintas variables entre sí. Finalmente, también se realizó un análisis univariado para examinar posibles relaciones existentes de las variables independientes y la variable dependiente del modelo, utilizando el estadístico de Kruskal-Wallis.

Los resultados del coeficiente de Spearman indican la existencia de una correlación moderada, pero negativa entre la edad y la ingeniería (a mayor edad, los estudiantes prefieren no estudiar ingeniería, pues estas carreras requieren de mucha dedicación, y en general, las personas mayores no disponen de tanto tiempo porque tienen otras obligaciones); una correlación positiva entre la ingeniería y la ansiedad matemática global (se explicará más adelante con los resultados del análisis univariado), y una correlación moderada y negativa entre la edad y la ansiedad matemática global (con el avance de la edad, se observa que la percepción de la ansiedad —en situaciones de aprendizaje y evaluación en matemáticas— tiende a disminuir, y de igual forma, la proporción de estudiantes en ingeniería son jóvenes, por lo que serán estos quienes presenten mayores niveles de ansiedad). Del análisis univariado se encuentra una varianza significativa en la ansiedad matemática global en función de la ingeniería, ansiedad en el aprendizaje y

ansiedad en la evaluación (existen diferencias significativas en la varianza de la ansiedad según la ingeniería elegida y del pénsum de las mismas; es decir, esta ansiedad matemática interviene al seleccionar una carrera). A pesar de la existencia de las correlaciones anteriormente mencionadas, el mayor resultado obtenido en esta metodología es que no existe correlación entre el rendimiento académico y la ansiedad matemática (la cual se esperaba, según el trabajo de otros autores). Esto puede explicarse debido a la variable temporal; pues, por ejemplo, en un examen existe un límite temporal que puede aumentar la ansiedad, haciendo que los estudiantes no puedan desarrollar y plasmar todas sus habilidades y competencias matemáticas. Ello permite concluir que el rendimiento académico y su relación con la ansiedad matemática, no están definidos con claridad, por lo que se podría pensar que el desempeño puede estar más influenciado por otros factores.

Hakyemez y Mardikyan (2021) buscan revelar cómo la interacción entre la integración institucional y la autoeficacia, tienen un efecto sobre el rendimiento académico. Los datos provienen de 520 estudiantes de primer año de la escuela de negocios de una universidad pública en Turquía. Los datos se recolectaron por medio de una encuesta durante la quinta semana del semestre de primavera del año académico 2014-2015. Por medio de la revisión de la literatura, se planteó una serie de hipótesis (con las cuales, los autores construyeron un modelo con sus respectivas relaciones):

- **H1:** “la integración académica afecta positivamente el desarrollo académico e intelectual”.
- **H2:** “el desarrollo académico e intelectual afecta positivamente el rendimiento académico”.
- **H3:** “la integración académica incide positivamente en el compromiso institucional”.
- **H4:** “la integración social incide positivamente en el compromiso institucional”.
- **H5:** “el compromiso institucional incide positivamente en el rendimiento académico”.

- **H6:** “la autoeficacia general afecta positivamente la autoeficacia académica”.
- **H7:** “la autoeficacia general afecta positivamente la integración social y la integración académica”.
- **H8:** “la integración académica afecta positivamente la autoeficacia académica”.
- **H9:** “la autoeficacia académica afecta positivamente el desarrollo académico e intelectual”.
- **H10:** “la autoeficacia académica afecta positivamente el rendimiento académico”.

El indicador de medición del rendimiento académico en esta metodología, consiste en el promedio de calificaciones de los estudiantes, proporcionado por la oficina de registro. Las variables modeladoras corresponden a algunos datos demográficos, como edad, género, ayuda financiera, tamaño del pueblo natal, tamaño de la dependencia, así como las variables derivadas de los cuestionarios sobre integración institucional y autoeficacia. Por otro lado, la variable dependiente corresponde al rendimiento académico. En esta encuesta aplicada, en principio se les preguntó por los siguientes datos, con el fin de determinar las variables modeladoras, como edad, género, ayuda financiera, tamaño del pueblo natal, tamaño de la dependencia:

- **Edad:** este dato permitió clasificar a los estudiantes en las categorías tradicionales (entre 18 y 21 años) y no tradicionales.
- **Género:** se tuvo un código específico para cada uno de los géneros (femenino = 1, masculino = 2).
- **Estado de ayuda financiera:** se respondía de forma binaria (sí = 1, no = 0), considerando que la respuesta positiva podría ser a través de una donación, un préstamo o ambas (donación y préstamo).
- **El pueblo natal:** también se midió por una variable binaria (una metrópolis con una población de 1 000 000 o más = 1, o todas las demás = 0).

- **Tamaño de la dependencia:** según el radio profesorado-alumnos ($0,1 = 1$, y los departamentos más pequeños = 2).

En otra parte de la encuesta se utilizó “La Escala de Integración Institucional” de Pascarella y Terenzini, la cual, mediante sus 34 ítems, captura la integración percibida de los estudiantes en cinco áreas diferentes: interacciones de pares-grupo (1), interacciones con profesores (2), preocupación del profesorado por el desarrollo y la enseñanza del estudiante (3), el desarrollo académico e intelectual (4) y el compromiso institucional y de metas (5). En este sentido, la integración académica se midió con las áreas (2) y (3), y la integración social por el área (1). Por otra parte, las subescalas (5) y (4) midieron el compromiso institucional y el desarrollo académico e intelectual, respectivamente.

Para las creencias de autoeficacia generales y específicas de dominio, se usaron dos escalas separadas: “La Escala de Autoeficacia General” de Chen (la cual posee ocho ítems) y “La Escala de Autoeficacia Académica” de Yılmaz (con siete ítems). Respondieron todos los ítems en una escala Likert de cinco puntos. Las relaciones entre autoeficacia, integración institucional y rendimiento académico se examinaron usando los paquetes estadísticos SPSS 20 y SPSS AMOS 21.

Con los resultados del modelo estructural, se evidenció que todas las relaciones propuestas son soportadas, excepto la relación entre el compromiso institucional y el desempeño académico, pues no se demostró una asociación positiva en ninguno de los modelos desarrollados. En este sentido, se encontró que los factores institucionales son fuertes predictores del rendimiento académico, excepto el mencionado anteriormente (compromiso institucional, lo que implicaría que la hipótesis H5 es rechazada), y también se sugiere que la autoeficacia académica y el desarrollo académico e intelectual están vinculados positivamente al rendimiento académico.

Por una parte, el análisis de múltiples grupos indicó que las relaciones propuestas están moderadas por la ayuda financiera. Las limitaciones en este estudio están dadas por la naturaleza del mismo; es decir, con la adopción de un diseño transversal, se asumió que los estudiantes no cambiarían de actitud desde la aplicación de la encuesta hasta la finalización del semestre, lo que

podría afectar la causalidad de los factores. Por otra, este estudio está dado para una sola institución con una muestra pequeña, por lo que no es posible generalizar el modelo. Finalmente, la gama de variables consideradas (a pesar de ser amplia) pudo haber dejado de lado algunos factores asociados.

Jain *et al.* (2016) examinaron si los programas de tutorías docentes y la autoeficacia profesional tienen un impacto en el rendimiento académico de los estudiantes. Estudio realizado en la Universidad Guru Gobind Singh Indraprastha (India). Se recolectó la información de un cuestionario de varias secciones, aplicado a 327 estudiantes (187 hombres y 140 mujeres) de posgrado en negocios, seleccionados aleatoriamente, y quienes, en promedio, tenían una edad de 20,5 años. En esta investigación, el indicador de medición del rendimiento académico corresponde al promedio porcentual acumulado. Las variables que inciden sobre el rendimiento académico (variable dependiente) se obtienen de los cuestionarios sobre la autoeficacia profesional y las tutorías, así como los datos demográficos de los estudiantes. Para esta investigación, se plantearon, en principio, las siguientes 2 hipótesis:

- **H1:** “no hay relación significativa entre las tutorías y la autoeficacia profesional”.
- **H2:** “no hay impacto significativo de las tutorías y la autoeficacia profesional en el rendimiento académico”.

Como se mencionó previamente, se usó una metodología de encuesta para probar las hipótesis de investigación. Dicho instrumento consistía en tres partes diferentes:

- En la primera parte se preguntó sobre los datos demográficos y sobre la historia académica de los estudiantes. Entre estas preguntas se encuentran, por ejemplo, la edad y los años de estudio, entre otras.
- La segunda parte consistió en la medición de las variables sobre la autoeficacia profesional mediante “la escala de autoeficacia ocupacional específica de la tarea”, en una adaptación propia (originalmente consistía de 60 ítems que medían cuatro habilidades: habilidades interpersonales verbales; habilidades cuantitativas, lógicas y de negocios; fuerza física y agilidad, y habilidades estéticas. Sin embargo, para este estudio solo se

consideraron las preguntas para las dos primeras habilidades, en un total de 30 ítems, y la puntuación para cada uno de los ítems estaba dada por una escala Likert de 10 puntos). Para este cuestionario, se calculó un Alpha de Cronbach de 0,943, lo cual corrobora su confiabilidad.

- La tercera parte consistió en la medición de las variables sobre las tutorías, para lo cual se usó “la escala de tutoría para estudiantes universitarios”. Este instrumento es medido desde cuatro factores: apoyo psicológico y emocional (8 ítems), grado y apoyo profesional (6 ítems), conocimiento de la materia académica (5 ítems), existencia de un modelo a seguir (6 ítems), y cada ítem se midió con una escala Likert de cinco puntos. El cuestionario, en general, presenta un Alpha de Cronbach de 0,920, corroborando así su confiabilidad.

El análisis estadístico de los resultados obtenidos se hizo con el *software* SPSS 22.0. Además, se usaron herramientas estadísticas como la media, Alpha de Cronbach, regresión, correlación y la prueba *t*. De forma más específica, se utilizó la prueba *t* para verificar la significancia estadística de las relaciones planteadas entre las variables. Por otro lado, se llevó a cabo la regresión para examinar el impacto de las tutorías en la autoeficacia. Además, se usó el estadístico de Durbin Watson para analizar el problema de la autocorrelación en los residuos del análisis de la regresión, el cual mostró un resultado de 1,820 lo que indica que las variables no están autocorrelacionadas.

Mediante los resultados de la prueba *t*, se encontró que la correlación entre el rendimiento académico y la autoeficacia es baja y positiva (0,213), mientras la correlación entre el rendimiento académico y las tutorías es alta y positiva (0,477). También se encontró que las tutorías y la autoeficacia están positivamente conectadas (0,319). En general, los valores encontrados para las relaciones, al ser menores a 0,5, son débiles pero positivas. La regresión demostró que la tutoría tiene resultados Beta positivos sobre la autoeficacia (0,287), lo que indica que la hipótesis H1 es rechazada. El valor de R² ajustado (6,6%) muestra la varianza explicada sobre la autoeficacia, lo que significa que las variables independientes en esta regresión (las tutorías) explican un porcentaje bajo en la varianza de los puntajes de los estudiantes sobre la autoeficacia profesional.

Por otro lado, la correlación muestra una asociación moderada y positiva entre las tutorías y el rendimiento académico. El resultado de R^2 muestra que el 20,3 % del rendimiento académico es explicado por las variables de las tutorías y la autoeficacia profesional. Esto indica que la hipótesis H2 también es rechazada, puesto que hay un impacto estadísticamente significativo de las tutorías y la autoeficacia profesional en el rendimiento académico de los estudiantes de negocios; pues las tutorías son las que tienen mayor impacto en el rendimiento académico que la autoeficacia profesional. Las limitaciones en este estudio pueden estar dadas a los resultados que no pueden generalizarse a otras áreas, pues el área de estudio fue solo una (estudiantes de negocios). Por otro lado, en este estudio tampoco se tuvieron en cuenta otros factores, como las características personales de los estudiantes (motivación al logro, competencias en las habilidades de estudio, características sociales y demográficas).

Espinosa *et al.* (2020) describieron la influencia del estrés en el rendimiento académico de estudiantes a nivel universitario. La muestra consta de 115 estudiantes (de ambos sexos entre las edades de 18 a 24 años) de psicología y trabajo social de la Universidad Simón Bolívar (sede Cúcuta, Colombia), que asistían a la jornada nocturna del segundo y cuarto semestre académico, y con promedios de 3,5 a 5 (en una escala de 0 a 5). El indicador de medición del rendimiento académico, en este estudio, está constituido por los promedios académicos de los estudiantes. Las variables de entrada se obtuvieron por medio de la escala para la determinación del nivel de estrés, junto con los factores demográficos que incluían, por ejemplo, si un estudiante trabajaba, las personas con las que vive, sexo y edad.

La recolección de datos se efectuó con el fin de probar la hipótesis central (la relación del estrés con el rendimiento académico), con base en la medición numérica y el análisis estadístico. De este modo, para la medición del nivel de estrés, se empleó la escala creada por Borges y Melgosa, la cual consistía en 30 preguntas cerradas (de tipo dicotómicas o politómicas), y de nueve preguntas más, relacionadas con la motivación, autoestima, depresión en razones familiares, laborales y personales. Entre de estas preguntas se encuentran:

- “¿Fumas?”
- “¿Usas tranquilizantes habitualmente?”
- “¿Vives en un vecindario ruidoso?”
- “¿Te irritas con facilidad?”
- “¿Te llevas trabajo a casa?”
- “¿Te irritas cuanto tienes que esperar en una fila?”

Los resultados obtenidos del cuestionario se analizaron con el paquete estadístico SPSS 25 y Excel. En cuanto a la motivación y la autoestima, resulta que son factores primordiales en lo personal, familiar y laboral, mientras que la depresión es un factor que los estudiantes minimizan porque existe un interés en la superación. De los resultados del estrés, se encontró que el 22 % de los estudiantes tiene estrés bajo, 31 % estrés medio y el 46 % tiene estrés alto. Al indagar estos resultados un poco más, y relacionarlos con otras variables como la edad o el género, se infiere que, en los estudiantes con edades inferiores al promedio, la mayoría presenta niveles de estrés bajos; pero a medida que se incrementa la edad de los estudiantes, esta tendencia se invierte debido a un mayor número de responsabilidades adquiridas (son las mujeres las que presentan la mayor frecuencia de estrés). Cabe mencionar que el estrés de los estudiantes que trabajan también se ve influido por una variable diferente, reflejada en la migración de venezolanos que constituyen una amenaza para el empleo. Por el contrario, quienes no trabajan, y no experimentan cambios (como moverse de una ciudad a otra para llegar a la universidad, no vivir solos, están en una zona de confort), son aquellos que presentan niveles moderados de estrés. Finalmente, y como respuesta a la hipótesis central que depende de la relación del estrés con el rendimiento académico, se observa que existe una relación no directa entre estos; pues, a un rendimiento académico menor, se corresponde una mayor puntuación referida al estrés. En síntesis, a partir del cuestionario planteado (en el cual se interrogó sobre diversos temas como el trabajo y la familia, entre otros), se determina que el estrés efectivamente es capaz de tener efectos sobre múltiples aspectos de la vida, en especial con el rendimiento académico.

Luján *et al.* (2019) pretenden determinar la relación entre la asertividad (es decir, la capacidad de comunicación asertiva) y el rendimiento académico entre los estudiantes de la Facultad de Ciencias Administrativas de la Universidad Nacional de Piura (Perú), según edad y sexo. La muestra empleada fue de 69 estudiantes en el 2017-1, a quienes se les aplicó un cuestionario para medir la asertividad. El indicador de medición del rendimiento académico en esta investigación, se representó por el promedio ponderado acumulado. La metodología empleada en este estudio consistió, inicialmente, en la conformación de tres distintos grupos, clasificados por el promedio. Específicamente, se conformó un primer grupo con los estudiantes que tenían promedios generales más altos (tercio superior), el segundo grupo con los promedios ubicados en el tercio medio, y el tercer grupo con los alumnos del tercio inferior. A estos estudiantes también se les aplicó el cuestionario de asertividad adaptado de Rathus, con el fin de medir la asertividad de cada uno de ellos. Dicho cuestionario se ejecutó por medio de 30 ítems, y donde fue posible clasificarlos en cuatro escalas diferentes (inasertivo, asertividad regular, asertivo, muy asertivo). Para el análisis estadístico principal, se utilizó el método del Momento del Producto de Pearson. De este modo, se correlacionaron las variables referentes a la asertividad con el rendimiento académico, con pruebas de asociación para variables nominales, de intervalo y ordinales. Se observa que los estudiantes hombres con mayor rendimiento académico presentan una asertividad un poco más alta; sin embargo, no existe una correlación significativa entre estas variables ($\text{Sig.} > 0,05$ [0.976]), por lo que no es posible afirmar que sea la asertividad un predictor del rendimiento académico masculino (pues el nivel de error obtenido es casi del 100%).

En cuanto a los resultados relativos a la edad, los estudiantes que tienen 24 años o más (61,5%), son los que muestran mayor asertividad. La asertividad regular se presenta entre los estudiantes de 20 a 21 años (65,45%) y de 22 a 23 años (80%). En este caso, la prueba exacta de Fisher evidencia una relación significativa ($\text{Sig.} = 0,027$) entre la asertividad y la edad de los estudiantes. Para la relación del rendimiento académico con la asertividad, se observa que los estudiantes con promedios acumulados de 15 o más, son asertivos o muy asertivos en el 20% de cada categoría; los estudiantes con promedios acumulados de 11 a 12, son asertivos en el 20%; mientras que para los estudiantes

con promedios de 13 y 14 se presenta el 26,5 % de asertividad y 2,0 % de muy asertivos. En general, la mayoría de los estudiantes evidencian asertividad regular (70 %, 67,3 % y 60 % para los rendimientos de 11 a 12, 13 a 14, y de 15 a más, respectivamente). Sin embargo, al analizar los resultados, es notable que no existe una correlación significativa ($\text{Sig.} > 0,05$ [0,093]) entre la asertividad y el rendimiento académico de los estudiantes; pues aunque la asociación entre asertividad y rendimiento es positiva débil ($r = 0,204$), no existe sustento para asegurar que esta relación presenta el 95 % de confianza para ser aceptada. Finalmente, se concluye que el rendimiento académico no se asocia de manera significativa con la asertividad de los estudiantes. Este mismo comportamiento ocurre con el género. Por otro lado, la relación significativa entre asertividad y edad, se da en estudiantes de 22 y 23 años.

Fernández-Lasarte *et al.* (2019) sugieren evidenciar la forma en que el apoyo social percibido influye en el rendimiento académico, así como la influencia positiva de la inteligencia emocional sobre el rendimiento académico. Además de ello, se busca analizar la capacidad predictiva del apoyo social percibido sobre la inteligencia emocional, y ambas variables sobre el RA. Para lograr el objetivo, se usó una muestra de 419 estudiantes de la Universidad del País Vasco (España), con edades entre 18 y 40 años. Las edades de los estudiantes se pueden clasificar como adolescencia tardía (estudiantes entre 18 a 20 años —66 %—), y como juventud o adultez temprana (estudiante entre 21 y 40 años —34 %—). Su participación fue voluntaria y anónima. El indicador de medición del rendimiento académico en esta investigación está dado por los resultados obtenidos a partir de uno de los cuestionarios aplicados (adaptación de EBAE-10), en el cual se mide el rendimiento académico por medio de un número de ítems con respuestas en escala Likert, en donde los estudiantes autorreportan su percepción sobre su propio desempeño. Las variables predictoras en este estudio, específicamente corresponden al apoyo social recibido (apoyo del profesorado, apoyo familiar y apoyo de las amistades), a la “inteligencia emocional” (atención emocional, claridad emocional y reparación emocional), mientras que la variable dependiente es el rendimiento académico. Como instrumentos para la recolección de datos, se emplearon las siguientes escalas:

- **Escala de Apoyo Social Percibido (AFA: apoyo familiar y amistades):** se compone de 14 ítems con cinco rangos de respuesta en escala Likert (1 = nunca a 5 = siempre), y dos factores: apoyo familiar (7 ítems) y apoyo de amistades (7 ítems). Por medio de estos ítems, se analizan la satisfacción y la disponibilidad percibida por los estudiantes ante la necesidad de hablar, así como recibir ayuda y afecto de su familia y amistades. La consistencia interna de la escala en esta investigación con el Alpha de Cronbach es de 0,83 y 0,82 para el apoyo familiar y el apoyo de amistades, respectivamente.
- **Teacher and Classmate Support Scale (TCMS):** presenta 4 ítems con escala Likert de cinco puntos (1 = totalmente en desacuerdo a 5 = totalmente de acuerdo), en donde se evalúa el trato justo, la ayuda, el interés y la amabilidad de los profesores hacia los estudiantes. La fiabilidad de esta escala mediante el Alpha de Cronbach para el presente estudio es de 0,73.
- **Trait Meta Mood Scale (TMMS-24):** cuenta con 24 ítems y escala Likert de cinco opciones de respuesta (1 = nada de acuerdo a 5 = totalmente de acuerdo), para medir sus tres dimensiones (atención emocional, claridad emocional y reparación emocional; se refiere a la percepción, la comprensión y el empleo de estrategias de regulación de las propias emociones). La consistencia interna es de 0,91, 0,90 y 0,85 para las dimensiones de atención emocional, claridad emocional y reparación emocional, respectivamente.
- **Escala Breve de Ajuste Escolar (EBAE-10):** la subescala de rendimiento académico consta de 3 ítems y seis grados de respuesta (1 = completamente en desacuerdo a 6 = completamente de acuerdo), con el fin de examinar la percepción de autoeficacia, motivación y buenas calificaciones. El Alfa de Cronbach de la subescala es de 0,70.

Las correlaciones bivariadas que se analizan con el coeficiente de correlación de Pearson (con el cual se mide su dirección e intensidad), parten de la influencia del apoyo social percibido sobre la inteligencia emocional, y ambas sobre el rendimiento académico.

Para los análisis estadísticos, se utilizó el *software* SPSS. Los resultados permiten observar que todas las relaciones planteadas son positivas y significativas, salvo en las tres variables de la “inteligencia emocional” y el apoyo

del profesorado que no alcanzaron significación estadística. Entre las asociaciones más altas se encuentran: el apoyo del profesorado con el rendimiento académico, el apoyo familiar con la reparación emocional y la reparación emocional con el rendimiento académico (pues esta es la relación más fuerte entre “inteligencia emocional y rendimiento” académico). Las variables del apoyo social no predicen la atención emocional de forma significativa. Para la predicción de las otras dimensiones de la “inteligencia emocional” (claridad emocional y reparación emocional), la única variable predictora corresponde al apoyo familiar. A pesar de esto, el R^2 de estos modelos es pequeño, demostrando una baja capacidad explicativa ($R^2 = 0,057$ y $R^2 = 0,037$ para la claridad emocional y la reparación emocional, respectivamente). En cuanto al rendimiento académico, esa variable es explicada por las variables “apoyo del profesorado” y “apoyo de las amistades”, teniendo un mayor peso el apoyo del profesorado ($\beta = 0,282$ frente a $\beta = 0,175$). El rendimiento académico también se explica por la reparación y atención emocional ($\beta = 0,133$ y $\beta = 0,117$, respectivamente).

Guizado *et al.* (2020) trataron de determinar la relación entre el desempeño docente y el rendimiento académico de los estudiantes (2015). La población estuvo conformada por 399 alumnos de las promociones 2011-2012 de la Universidad Nacional de Educación “Enrique Guzmán y Valle” (Perú), matriculados en las 14 especialidades de dicha universidad. Las variables que inciden sobre el rendimiento académico de los estudiantes (variable dependiente), están relacionadas con el desempeño docente (desde el ámbito profesional, personal y social). Se plantearon cuatro hipótesis (tres específicas y una general):

- **H1:** “relación entre el desempeño docente (ámbito profesional) y el rendimiento académico”.
- **H2:** “relación entre el desempeño docente (ámbito personal) y el rendimiento académico”.
- **H3:** “relación entre el desempeño docente (ámbito social) y el rendimiento académico”.
- **Hipótesis general:** “relación entre el desempeño docente y el rendimiento académico”.

Para recolectar la información sobre el desempeño docente, se elaboró un cuestionario estructurado de 32 ítems con respuestas en una escala Likert de cinco puntos. Para el análisis de la información, se utilizó tanto la estadística descriptiva como la inferencial (prueba r de Pearson para el cálculo de los niveles de correlación existente entre las diferentes variables de las hipótesis planteadas). A partir de la prueba r se determinó que existen correlaciones significativas en las hipótesis específicas y la general. De forma específica:

- **H1:** “existe una fuerte relación entre las variables desempeño docente (ámbito profesional) y el rendimiento académico del 75,7% con un nivel de confianza del 95 % y 5 % de probabilidad de error, y significancia estadística bilateral de 0,000”.
- **H2:** “existe una fuerte relación entre las variables desempeño docente (ámbito personal) y el rendimiento académico del 77,9% con un nivel de confianza del 95 % y 5 % de probabilidad de error, y significancia estadística bilateral de 0,000”.
- **H3:** “existe una fuerte relación entre las variables desempeño docente (ámbito social) y el rendimiento académico del 77,1 % con un nivel de confianza del 95 % y 5 % de probabilidad de error, y significancia estadística bilateral de 0,000”.
- **Hipótesis general:** “existe una fuerte relación entre las variables desempeño docente y el rendimiento académico del 76,8% con un nivel de confianza del 95 % y 5 % de probabilidad de error, y significancia estadística bilateral de 0,000”.

De este modo, con un nivel de 76,8% se afirma que hay relación significativa entre el desempeño docente y el rendimiento académico de los estudiantes de las promociones 2011 y 2012 del ciclo académico 2015-I en la Facultad de Tecnología de la Universidad Nacional de Educación.

Zárate *et al.* (2020) buscan establecer el grado de relación entre la competencia comunicativa intercultural y el rendimiento académico en estudiantes de lenguas extranjeras. La muestra estuvo conformada por 280 estudiantes de Lenguas Extranjeras de la Universidad Nacional de Educación “Enrique Guzmán y Valle” (Perú), de las especialidades de inglés, francés, italiano y

español como lengua extranjera. Esta investigación propone como indicador de medición del rendimiento académico, las calificaciones finales de los estudiantes en las asignaturas de la especialidad (que oscilan de 1 a 20) en distintos niveles: rendimiento académico bajo (calificaciones menores o iguales a 10), rendimiento académico regular (calificaciones de 11, 12 y 13), rendimiento académico bueno (14 y 15), rendimiento académico muy bueno (16, 17, 18), rendimiento académico excelente (19 y 20). La variable independiente en este estudio es la “competencia comunicativa intercultural” (clasificada entre CCI cognitiva, CCI comportamental y CCI afectiva), mientras la variable dependiente es el rendimiento académico. Para lograr el objetivo de esta investigación, se plantearon cuatro hipótesis (una general y tres específicas):

- **Hipótesis general:** “correlación entre ‘competencia comunicativa intercultural’ y el rendimiento académico”.
- **H1:** “correlación entre ‘competencia comunicativa intercultural cognitiva’ y el rendimiento académico”.
- **H2:** “correlación entre ‘competencia comunicativa intercultural comportamental’ y el rendimiento académico”.
- **H3:** “correlación entre ‘competencia comunicativa afectiva’ y el rendimiento académico”.

La recolección de datos de la primera variable (“competencia comunicativa intercultural”) se efectuó por medio de un cuestionario de tipo Likert, con una confiabilidad de 0,826 (calculada con el Alpha de Cronbach). Por otra parte, para obtener las notas finales (segunda variable), se llevó a cabo un análisis documental de las actas de notas de las asignaturas de especialidad en lenguas extranjeras. Después, se tabularon y se analizaron los datos mediante el paquete estadístico SPSS. De toda la muestra empleada, 7 estudiantes presentaron un rendimiento académico deficiente (lo que significa que no pudieron comprender frases y expresiones de uso frecuente y que se relacionan con las experiencias personales, información sobre sí mismos, familia, lugares de interés, ocupaciones, entre otros); 37 estudiantes presentaron un nivel regular (sí lograron comprender frases y expresiones de uso

frecuente relacionadas con su experiencia personal en un nivel muy básico); 132 estudiantes evidenciaron un nivel académico bueno (son capaces de entender textos y producir textos sencillos); 101 sujetos presentaron un nivel muy bueno (desarrollaron habilidades más complejas, como entender textos y producir textos orales y escritos de mayor complejidad, conversar con hablantes nativos, entre otros); y 3 estudiantes evidenciaron un rendimiento académico excelente (lograron de manera eficiente el dominio de las cuatro habilidades: escuchar, hablar, leer y escribir).

Con respecto a las hipótesis planteadas, se observa que:

- **Hipótesis general:** “los resultados evidencian una relación entre las variables, de grado positivo débil según el indicador Rho de Spearman con el valor de 0,172. El valor de $p = 0,04$ es menor a 0,05, por lo que la relación entre dichas variables es significativa”.
- **H1:** “el grado de relación entre la CCI cognitiva y el RA es no significativo ($p = 0,054$). Mediante la aplicación del coeficiente Rho de Spearman, se obtuvo un coeficiente de 0,115, demostrando que, además, el grado de relación entre ambas variables es débil y positiva. Sin embargo, a pesar de no ser significativa, es necesario tenerla en cuenta, pues viene a ser un instrumento que incrementa el nivel de complejidad cognitiva”.
- **H2:** “el grado de relación entre la CCI comportamental y el RA es significativo ($p = 0,013$), y con el coeficiente Rho de Spearman (0,149) se evidenció que el grado de relación entre ambas variables es débil y el sentido (dirección) de esta asociación es positivo”.
- **H3:** “el grado de relación entre la CCI afectiva y el RA no es significativo ($p = 0,086$). El coeficiente Rho de Spearman (0,103) señaló que el grado de relación, entre ambas variables, es débil y positivo. A pesar de no ser significativo, no se debe descuidar, ya que es importante el control de las emociones y la ansiedad, desarrollo de empatía, actitudes positivas, entre otros”.

De este modo, y sintetizando, la CCI (cognitiva, comportamental o afectiva) no interfiere de gran manera en el aprendizaje de lenguas extranjeras y su

rendimiento académico, puesto que los resultados obtenidos en cada dimensión dan cuenta que son relaciones positivas pero débiles.

Lin *et al.* (2020) investigaron la cantidad y fuentes de estrés en estudiantes universitarios de odontología en Fujian (China), así como los factores asociados con el estrés. El indicador de medición del rendimiento académico corresponde al promedio de calificaciones en una escala de 0 a 4. Las variables en este estudio son: rendimiento académico (variable dependiente) y la relativa a los cuestionarios sobre el estrés (variable independiente). También se incluyen las variables asociadas como edad, sexo, año de estudio, estado civil.

En principio, se aplicó el “Cuestionario de estrés ambiental dental” (DES) y “La escala de estrés percibido” (PSS) usando un sistema de encuesta *on-line*. También se debió recolectar la información demográfica (sexo, edad, año de estudio) de cada uno de los estudiantes, así como el promedio de calificaciones. El primer cuestionario (que tenía una consistencia interna de 0,86 por el Alpha de Cronbach) estaba adaptado del original, y resultó con una muestra con 32 ítems agrupados en cinco categorías de posibles factores de estrés: “factores de estrés social” (ítems 1 a 10), “facultad y administración” (ítems 11 a 16), “carga de trabajo” (ítems 17 a 22), “creencias de autoeficacia” (ítems 23 a 27) y “presión de desempeño” (ítems 28 a 32). Cada punto tenía opción de respuesta en una escala Likert de cuatro puntos, desde 1 (no estresante) hasta 4 (muy estresante), junto con una quinta respuesta “no aplicable”, si ese era el caso. El segundo cuestionario (con Alpha de Cronbach igual a 0,81), consistía en 10 ítems mediante los cuales se indagaba sobre los sentimientos y pensamientos de los estudiantes durante el último mes, con el fin de evaluar el estrés general percibido. Estas respuestas estaban dadas en escala Likert de cinco puntos (desde 0 = nunca hasta 4 = a menudo).

Para efectuar los análisis estadísticos, se utilizó PASW Statistics. El puntaje promedio obtenido de uno de los cuestionarios (PSS) fue de 18,06 (en un rango de 4 a 33). A partir de estos, no se hallaron diferencias significativas en los puntajes de esta escala por año, estudio o sexo. Respecto a las fuentes de estrés, los participantes informaron tres causas principales: “a menudo piensan en cosas que tienen que lograr”, “a menudo se sienten nerviosos y estresados” y “a menudo sentían que las cosas iban bien”. Por un lado, el puntaje promedio de DES de todos los participantes fue de 52,21

(dentro del rango: 0-128). Tampoco se encontraron diferencias significativas en las puntuaciones totales de la DES por año de estudio, ni por sexo. Por otro, se encontraron diferencias significativas en las puntuaciones de “carga de trabajo” ($p = 0,015$) y “creencias de autoeficacia” ($p = 0,020$) del DES entre los estudiantes en los diferentes años de estudio. De forma específica, los estudiantes de tercer y cuarto año mostraron puntajes de “carga de trabajo” de DES, significativamente más altos que los estudiantes de primer y segundo año. También, los estudiantes de segundo, tercer y cuarto año tenían creencias sobre la autoeficacia más fuertes que los estudiantes de primer año. Respecto a la relación del estrés con el rendimiento académico a través de la regresión lineal múltiple, se observa que las puntuaciones de estrés están correlacionadas significativa e inversamente con el promedio de calificaciones ($r = -0,119$, $p = 0,029$ con el cuestionario DES; $r = -0,116$, $p = 0,037$ con el cuestionario PSS); ello significa que una puntuación de estrés más alta puede predecir un rendimiento académico más bajo en el promedio de calificaciones. Al contrario, no hubo correlación significativa entre las puntuaciones de estrés y la edad, las puntuaciones de estrés y el sexo, o las puntuaciones de estrés y el año de estudio (todos $p > 0,05$). El sexo sí se correlaciona de manera significativa y positiva con las puntuaciones promedio, pues las estudiantes presentaron un mejor rendimiento académico. No se encontró correlación significativa entre datos demográficos y puntuaciones de estrés.

Deliens *et al.* (2013) examinaron las diferencias entre las características sociodemográficas y las conductas alimentarias de los estudiantes, con el fin de identificar cómo estas últimas, junto con el peso, pueden estar correlacionadas con el rendimiento académico. La muestra estuvo compuesta por 101 estudiantes de primer año de la Universidad Libre de Bruselas (Bélgica); en dicha muestra se encuentran los estudiantes que asistieron a todos los exámenes finales del curso, así como los que no lo hicieron. El indicador de medición del rendimiento académico se midió por el promedio de calificaciones obtenido al final del año académico. Las variables independientes en este estudio corresponden a los datos recolectados sobre las medidas antropométricas (peso, índice de masa corporal, porcentaje de grasa, contorno de cintura), acerca de las características sociodemográficas (género, edad, etnia, residencia, promedio de calificaciones del último año de secundaria),

respecto al estatus socioeconómico (educación de los padres), también de los hábitos dietarios, de sueño, de fumar, de actividad física, entre otros, que provienen de los cuestionarios respondidos por los estudiantes. Por su parte, la variable dependiente corresponde al rendimiento académico.

Para analizar los datos recolectados, se empleó la herramienta SPSS. El promedio de calificaciones del último año de secundaria, fue más alto para aquellos que asistieron a todos los exámenes (un promedio de calificaciones de 68,6); en contraste con los que no presentaron todos los exámenes (un promedio de calificaciones de 63,6). Asimismo, estos estudiantes que no asistieron a todos los exámenes mostraron un incremento en su contorno de cintura (+1,2 cm) (comparado con quienes pasaron sus exámenes [-0,5 cm] y con quienes presentaron sus exámenes, pero fallaron [1,0 cm]). En los hábitos alimenticios, el consumo de papas fritas fue ligero, pero significativamente mayor en los estudiantes que no atendieron todos los exámenes del curso. De igual manera, el RA más bajo se presentó cuando los estudiantes reportaron estar haciendo una dieta, que cuando comían con más frecuencia en el restaurante, o cuando consumían más soda o papas fritas, o cuando reportaron un mayor uso de alcohol. Del análisis de regresión univariado se evidenció que las estudiantes con promedios de calificaciones altos durante su último año de secundaria, también mostraron un promedio de calificaciones alto en su primer año de universidad. Por otra parte, la ganancia de peso, incremento en el índice de masa corporal y el contorno de cintura estuvieron negativamente correlacionados con el RA. De los hábitos de juego entre semana, también se evidenció que representan calificaciones más bajas para el estudiante. Asimismo, la correlación más fuerte con el RA estaba dada con el contorno de cintura.

Respecto al modelo multivariado (en el que no se incluyó el promedio de calificaciones de secundaria, pues la muestra era limitada), se encontró que el 30% de la varianza en el rendimiento académico estaba explicado por dicho modelo. En sus resultados, únicamente el estado de la dieta y la frecuencia del consumo de alcohol permanecieron significativos; esta última es la que más se correlaciona con el promedio de calificaciones. Como conclusión, se infiere que ser hombre, tener bajas calificaciones en secundaria, un incremento en el peso, IMC y contorno de cintura, el juego entre semana, estar

en una dieta, comer en el restaurante estudiantil con mayor frecuencia, un mayor consumo de soda y papas fritas, y mayor frecuencia de consumo de alcohol, puede predecir un rendimiento académico más bajo.

Bahammam *et al.* (2012) plantearon su trabajo teniendo como principal objetivo la determinación de la relación existente entre los hábitos y la duración de sueño, con el rendimiento académico. Se llevó a cabo entre 2009 y 2010 en la Facultad de Medicina de la Universidad King Saud (Arabia Saudita), con una muestra de 410 estudiantes de los primeros tres niveles académicos. El indicador de medición del rendimiento académico empleado en esta investigación corresponde al promedio de calificaciones, el cual se clasificó en dos grandes grupos: “excelente” (con un puntaje entre 3,5 y 5) y “promedio” (con un puntaje menor a 3,5). En este estudio se emplearon las siguientes variables independientes: datos demográficos (sexo, índice de masa corporal, obesidad, uso de internet, trabajo), y respuestas sobre hábitos de sueño derivadas de los cuestionarios, como horarios para dormir y despertar, duración, quejas del sueño, entre otros; hábitos sobre la ingesta de café o fumar, tiempo asignado para actividades y hábitos diarios, entre otros. El análisis se realizó con el *software* SPSS. De la muestra estudiada, el 28% de ellos (115 estudiantes) presentaron un RA excelente, mientras que la gran mayoría (72% = 295 estudiantes) presentaron un RA promedio. De este último grupo, se observa que estos tuvieron mayores calificaciones en el ESS (escala de somnolencia de Epworth), así como también mostraron un mayor porcentaje de estudiantes que se sentían somnolientos en las clases (68,8% en esta categoría, frente al 55,3% de la categoría “excelente”).

Los grupos con RA “excelente” presentaron una hora de dormir más temprano, así como un mayor tiempo total de sueño entre los días de la semana. También se menciona que la prevalencia de obesidad, el hábito de fumar y el uso diario de internet de más de dos horas al día, fue mayor en el grupo “promedio”. Asimismo, las diferencias entre grupos están dadas por la variable de “obtener suficiente sueño”, pues la proporción de los estudiantes del grupo excelente que obtuvo suficiente sueño (50%) fue mayor que la proporción en el grupo promedio (30,6%). Por otro lado, el análisis demostró que el sexo o el nivel educativo, no influyó en las diferencias en el rendimiento académico, así como consumir bebidas con cafeína. Al final

del análisis, se concluyó que el sentimiento subjetivo de obtener suficiente sueño y no fumar, son las únicas variables consideradas como predictoras de un rendimiento académico excelente.

Céspedes *et al.* (2018), en su investigación, tuvieron como objetivo la determinación del efecto tanto del ausentismo como del presentismo sobre el rendimiento académico, mediante el estudio de una muestra de 8203 estudiantes de pregrado que se encontraban desde el primer al décimo ciclo, en las escuelas de negocios e ingenierías de las universidades peruanas (34 universidades públicas y privadas). Los indicadores de medición del rendimiento académico en la presente investigación corresponden a los ítems asociados a las categorías de rendimiento académico disminuido, cursos y ciclos desaprobados y la evolución del rendimiento académico.

La información se recolectó mediante una encuesta de autor reporte, compuesta por diferentes escalas que permitieron el análisis del ausentismo, el presentismo y el rendimiento académico, así como una serie de preguntas en donde se requirió información demográfica, académica y laboral. De manera específica, la escala de ausentismo con 5 ítems, se dividió en dos dimensiones (cantidad de días ausentes debido a la salud y por otras razones). Sus opciones de respuesta son: nunca (0), 1 día (1), 2 días (2), de 3 a 5 días (4), de 6 a 10 días (8), más de 10 días (15). Por su parte, la escala de presentismo, la cual indicaba la disminución de la productividad académica por la distracción y el agotamiento, tuvo las siguientes opciones de respuesta: nunca (0), 1 día (1), entre 2 y 5 días (4), entre 6 y 10 días (8), casi todos los días (15).

Los estudiantes que presentaron presentismo, el 95,7% reportó haber tenido problemas para concentrarse durante las clases, el 82,3% afirmó haber estado distraído, el 87,2% informó haber estado cansado, otros estudiaron más lento (77,45%), así como tuvieron preocupaciones externas al estudio (76,3%). Respecto a la disminución del rendimiento académico, el 82,5% expresó que, durante el último mes, este estuvo por debajo de sus capacidades. En cuanto a la desaprobación de ciclos y cursos, el 53,6% indicó que ha desaprobado al menos un curso, y el 23% ha repetido al menos un ciclo académico. Referente al género, el 51,8% de las mujeres ha desaprobado al menos un curso y el 18,2% ha repetido un ciclo académico; mientras que los

hombres, el porcentaje de desaprobación fue de 55,4% y el de repetición de un ciclo académico de 23,8%.

La relación entre el ausentismo y el presentismo evidencia un efecto directo en el rendimiento académico disminuido, ya que lo explica en 43,3%, así como sobre la evolución del RA en 9,7%. Asimismo, el ausentismo es capaz de explicar el 5,3% del presentismo académico, así como también lo hace con los cursos y ciclos desaprobados, en un porcentaje de 3,7%.

Con los coeficientes Beta obtenidos en el modelo estructural, se observa que la relación más fuerte y significativa es la del presentismo sobre la disminución del RA ($B = 0,628$ y $p = 0,000$), seguida del ausentismo sobre el presentismo ($B = 0,231$ y $p = 0,000$) y del ausentismo sobre los cursos y ciclos desaprobados ($B = 0,163$ y $p = 0,000$). Otras de las relaciones a un nivel de significancia igual a 0,000, son las presentadas por el ausentismo sobre la disminución del RA ($B = 0,101$) y el presentismo sobre la evolución del RA ($B = -0,145$).

García-Hernández y González-Ramírez (2020) proyectaron su trabajo en identificar las variables predictivas asociadas al rendimiento académico en matemáticas, en la carrera de ingeniería en ciencias informáticas (Cuba), durante el curso académico 2017-2018. La muestra consta de 807 estudiantes. La variable dependiente es el rendimiento académico en matemáticas discretas, mientras que las variables independientes son: sexo, provincia de procedencia, puntaje en la prueba de ingreso de matemáticas, opción en la que se solicitó la carrera y los materiales de estudio en la asignatura de matemáticas discretas. Como primera fase, se desarrolló un análisis descriptivo de las variables independientes para poder inferir las características de la población total. También, y mediante un estudio correlacional (Pearson), se llevó a cabo un análisis preliminar sobre las correlaciones entre las variables. A partir del estudio correlacional, se determinó que aquellas variables que más se relacionan con el RA son: la prueba de ingreso de matemática ($r = 0,441$ a un nivel de significancia de 0,01) y los materiales de estudio (en especial, las subvariables sobre capacidad de motivación — $r = 0,535$ a un nivel de significancia de 0,01—, adecuación didáctica — $r = 0,401$ a un nivel de significancia de 0,01— y calidad tecnológica — $r = 0,498$ a un nivel de significancia de 0,01—). El resto de pares de variables presentaron

correlaciones bajas y poco representativas. Sin embargo, como ninguna de las correlaciones son superiores a 0,9, esto indica que no existe multicolinealidad entre las variables de estudio. Del modelo de regresión logística, se determinó que los factores seleccionados demuestran una contribución significativa en la explicación de la probabilidad de tener un bajo RA, pues alcanzó un porcentaje de predicción de 85,6%. De forma específica, estos factores representativos son el sexo (con una significancia estadística de 0,000 y un valor odds-ratio de 2,091, lo que significa que un hombre tiene 2,091 menos posibilidades de suspender que las mujeres). Por su parte, la variable de prueba de ingreso de matemáticas (con 0,000 de significancia y valor odds-ratio de 0,338) indica que un estudiante que logre altas calificaciones de ingreso, puede reducir la probabilidad de suspender la materia “matemáticas discretas” en 66%. Respecto a los materiales de estudio, para la subvariable de capacidad de motivación (con significancia de 0,005 y odds-ratio de 0,257), se encuentra que los estudiantes altamente motivados tienen cuatro veces más de probabilidad de alcanzar un buen RA; en las subvariables adecuación didáctica y calidad tecnológica (significancia de 0,000 y odds-ratio de 0,758, y significancia de 0,002 y odds-ratio de 0,809) se determina que los estudiantes que se encuentran más satisfechos respecto a estas, tienen 20% menos de probabilidad de alcanzar un bajo RA. En resumen, los materiales de estudio y las calificaciones obtenidas en la prueba de ingreso de matemáticas para la educación superior cubana, son las dos variables que mejor predicen el RA para la materia matemáticas discretas.

Rosales-Jaramillo *et al.* (2018) buscan caracterizar a los estudiantes de las carreras de medicina, veterinaria y zootecnia (MVZ) e ingeniería agronómica (IAG), con el fin de establecer los determinantes de su rendimiento académico. Se realizó en la Facultad de Ciencias Agropecuarias de la Universidad de Cuenca (Ecuador), para una muestra compuesta de 813 estudiantes. Las variables determinantes en este estudio están clasificadas en las categorías de variables personales, sociales e institucionales. Por su parte, la variable independiente corresponde al rendimiento académico. La recolección de información primaria se llevó a cabo mediante una encuesta estructurada y con el uso de una escala Likert, la cual contenía aspectos sociodemográficos y de satisfacción sobre temas educativos.

La información recolectada se analizó con el SPSS. Del análisis de todas las variables determinantes, se evidenció que 18 de estas poseen una relación significativa con el rendimiento académico; sin embargo, la mayoría de estas son débiles. Cabe resaltar que, de estas variables, la “carrera” es aquella que presenta un mayor grado de relación ($r = 0,323$ con significancia de 0,000), lo que permite deducir que los estudiantes de las carreras investigadas presentan diferencias en su RA. Al contrario, algunas de las variables planteadas, no presentan una relación con el RA. Entre estas se encuentran: motivación extrínseca, acceso a internet, lugar de estudio en el domicilio, frecuencia de lectura, consumo de alcohol, tabaco, niveles educativos de los padres, entre otras.

Rivera *et al.* (2020), en la Escuela Profesional de Educación Primaria de la Universidad Nacional Hermilio Valdizán en Huánuco (Perú) plantearon su trabajo, cuyo objetivo principal es determinar y describir la empatía y su relación con el rendimiento académico de los estudiantes. La muestra correspondió al total de la población de 185 estudiantes. El indicador de medición del rendimiento académico corresponde al promedio acumulativo del total de asignaturas cursadas en todos los semestres. La información sobre la variable empatía se midió por medio de una encuesta, basada en un instrumento tipo Likert, llamado “índice de reactividad personal”, de 28 ítems divididos en cuatro subescalas: toma de perspectiva (TP), fantasía (FS), preocupación empática (EC) y malestar personal (PD), y con una confiabilidad establecida en un rango de 0,71 a 0,77 para dichas subescalas. Cada uno de los ítems presentados se midió a través de cinco opciones, con puntuaciones de 0 a 4 con las siguientes respuestas: “no me describe bien” (0), “me describe un poco” (1), “me describe bien” (2), “me describe bastante bien” (3), “me describe muy bien” (4). Las puntuaciones mayores implican una mayor presencia en el estudiante del constructo medido. Las correlaciones del RA con cada una de las dimensiones de la empatía no son significativas ni fuertes (se presentó un $r = 0,025$ y significancia de 0,778). Aunque con anterioridad, se indicó que los estudiantes con mejor RA, también son aquellos que, según el instrumento, presentan mayores niveles de empatía, se demostró que no existen diferencias estadísticamente significativas en la variable empatía, frente a aquellos que no presentaron un buen rendimiento académico.

Castro y Closas (2020), en su estudio, evaluaron la eficacia de un proceso educativo que se caracteriza por el uso de las TIC en la Facultad de Ciencias Económicas de la Universidad Nacional del Nordeste (Argentina), para la asignatura de contabilidad básica a partir de una muestra de 118 estudiantes. El indicador de medición del rendimiento académico se constituye mediante tres variables: el promedio de calificaciones de las pruebas escritas parciales acumulativas teórico-prácticas; la cantidad de exámenes escritos realizados durante la asignatura, y la situación académica. Los datos obtenidos se procesaron con el SPSS. Del modelo planteado, las relaciones que se presentaron son significativas en términos estadísticos (algunos con $p < 0,05$ y $p < 0,01$), y con un coeficiente de regresión considerable:

- Enseñanza y materiales (0,466)
- Enseñanza y aprendizaje (0,513)
- Enseñanza y rendimiento (0,237)
- Materiales y aprendizaje (0,347)
- Materiales y rendimiento (0,616)
- Aprendizaje y rendimiento (0,293)

La variable latente que mejor explica el rendimiento académico es la enseñanza, pues no solo incide en él de forma directa, sino que también lo hace de manera indirecta a través de los factores materiales y aprendizaje. Las limitaciones del estudio se encuentran en la falta de nuevas variables que puedan incidir sobre el rendimiento académico, o bien, las mismas variables, pero con una forma de medición distinta. Además, al ser la muestra específica, no es conveniente extender los resultados de la investigación a otros ámbitos, por lo que no es posible hacer inferencias sobre las poblaciones que no se ven representadas en la muestra.

Hayat *et al.* (2020) indagaron acerca de los efectos mediadores que tienen tanto las estrategias de aprendizaje metacognitivo, así como las emociones relacionadas con el aprendizaje entre la relación de la autoeficacia académica con el rendimiento académico. El tamaño de la muestra incluye 279 estudiantes de medicina de la Universidad de Ciencias Médicas de Shiraz

(Irán). Se analizaron las siguientes variables: puntajes de los exámenes finales en el semestre de estudio, puntajes de curso por exámenes parciales y componente de trabajo semestral (*quices*, tareas-presentaciones en clase individuales y en grupo, revisión de un libro o trabajo, discusiones grupales y proyectos de investigación a lo largo del periodo). Para el desarrollo del objetivo, se plantearon las siguientes nueve hipótesis:

- **H1:** “la autoeficacia académica tiene un efecto directo sobre el rendimiento académico”.
- **H2:** “las emociones académicas positivas tienen un efecto directo en el rendimiento académico”.
- **H3:** “las estrategias de aprendizaje metacognitivo tienen un efecto directo sobre el rendimiento académico”.
- **H4:** “la autoeficacia académica tiene un efecto directo sobre las emociones académicas positivas”.
- **H5:** “la autoeficacia académica tiene un efecto directo sobre las estrategias de aprendizaje metacognitivas”.
- **H6:** “las emociones académicas positivas tienen un efecto directo sobre las estrategias de aprendizaje metacognitivas”.
- **H7:** “las estrategias de aprendizaje metacognitivo median la relación entre la autoeficacia académica y el rendimiento académico”.
- **H8:** “las estrategias de aprendizaje metacognitivo median la relación entre las emociones académicas positivas y el rendimiento académico”.
- **H9:** “las emociones académicas positivas median la relación entre la autoeficacia académica y el rendimiento académico”.

La matriz de correlación permite identificar que la autoeficacia presenta relación tanto significativa como positiva con el RA ($r = 0,46$, $p \leq 0,01$). Igualmente, para la relación entre las estrategias de aprendizaje metacognitivo y el RA, el comportamiento es similar a lo anteriormente descrito ($r = 0,45$, $p \leq 0,01$), así como la relación entre las emociones positivas relacionadas con el aprendizaje y el RA ($r = 0,48$, $p \leq 0,01$); pues esta relación es la más fuerte. Para las variables independientes, también se evidencia

relación positiva y significativa entre el aprendizaje metacognitivo y la autoeficacia ($r = 0,59$, $p \leq 0,01$), así como las emociones positivas con la autoeficacia ($r = 0,65$, $p \leq 0,01$). Finalmente, la última relación es entre las emociones positivas y las estrategias de aprendizaje metacognitivo que poseen el mismo comportamiento de las relaciones anteriores ($r = 0,55$, $p \leq 0,01$). A partir del modelado de ecuaciones estructurales, fue posible determinar que la autoeficacia de los estudiantes presenta un impacto tanto en sus emociones (relativas al aprendizaje) así como en las estrategias de aprendizaje metacognitivas, las cuales, a su vez, afectan el RA. Asimismo, las emociones relativas al aprendizaje influyen en las estrategias de aprendizaje metacognitivas, que median el efecto de las emociones sobre el RA.

Mazumder *et al.* (2020) investigaron la participación de los estudiantes durante la clase y su efecto sobre su rendimiento académico en tres diferentes clases de ingeniería mecánica de la Universidad de Michigan-Flint (Estados Unidos) para una muestra de 44 estudiantes. La participación de los estudiantes en clase se determinó mediante tres métodos: una encuesta de autorreporte, las observaciones hechas por los estudiantes durante la clase, así como la captura de los gestos de los mismos. Estos datos se recolectaron en dos momentos diferentes (los primeros 15 minutos de la clase y 45 minutos después de iniciada la clase). Para el cumplimiento del objetivo del estudio, se plantearon dos hipótesis:

- **H1:** “no hay diferencia significativa entre la participación y RA con respecto a hombres y mujeres en el salón”.
- **H2:** “no hay diferencia significativa en el nivel de participación durante el periodo de clase”.

Los datos recolectados se analizaron con el *software* SPSS. Existe una diferencia significativa en términos estadísticos entre ambos géneros ($p = 0,04$), lo que conduce a rechazar también la primera hipótesis. Esto va acorde con lo encontrado sobre las estudiantes, ya que ellas parecen ser más activas que los estudiantes. En cuanto al comportamiento en los intervalos de tiempo, se determinó que no hubo diferencia estadísticamente significativa durante ambos momentos ($p = 0,06$). Sin embargo, al comparar el comportamiento y el nivel de participación, se evidenció que los estudiantes son más participativos

durante el primer intervalo de medición ($p = 0,004$), lo que permite que la segunda hipótesis sea rechazada. Finalmente, la diferencia entre la participación y el RA no fue significativa en términos estadísticos ($p = 0,53$), lo que permite concluir que la participación no es un buen predictor del rendimiento académico. Los autores señalan que los estudiantes pueden complementar su aprendizaje con actividades externas a la clase, y no exclusivamente en esta.

Awadalla *et al.* (2020) estudiaron la relación entre la ansiedad y la depresión respecto al rendimiento académico (el cual varía de 0 a 4, pues tiene un puntaje alto un indicador de mejor desempeño). La muestra está compuesta por 404 estudiantes con edades que oscilaban entre 17 y 25 años, quienes asistieron a una universidad de los Emiratos Árabes Unidos. Se planteó la hipótesis de que los estudiantes que posean un posible trastorno depresivo, tendrán promedios de calificación más bajos, y que los niveles más altos de los síntomas depresivos al inicio del estudio, predicen promedios más bajos en el seguimiento posterior. Las variables independientes corresponden a los puntajes del cuestionario PHQ-9, GAD-7, edad, sexo, estado civil, año de estudio, facultad, educación materna-paterna, escala de riqueza familiar, corresponden a las variables independientes. A los estudiantes se les pidió completar una encuesta escrita (al inicio del estudio y seis meses después), en donde se les preguntó por los factores socioeconómicos y su rendimiento académico. Por otro lado, para evaluar los síntomas depresivos y de ansiedad se emplearon las escalas PHQ-9 y GAD-7. De manera más específica, se recolectó la siguiente información:

- Factores demográficos y socioeconómicos: se consideraron la edad, sexo, estado civil, especialización, año de estudio, nivel educativo de los padres. También se empleó la escala de afluencia familiar (FAS), con el fin de evaluar los recursos materiales familiares (propiedad de un automóvil, dormitorio propio, computador en el hogar, número de vacaciones familiares por año); sus puntuaciones oscilan de 0 a 9, donde a mayor puntaje, se indica mayor afluencia (abundancia), mientras que las puntuaciones más bajas pueden significar mayores riesgos de depresión.
- El Cuestionario de salud del paciente (PHQ-9) es una medida de detección para la evaluación de síntomas de depresión. Posee 9 ítems, cada

uno con una escala de 0 a 3, para obtener una puntuación total de 27. Los estudiantes con 10 puntos o más, son clasificados en el grupo de posible trastorno depresivo mayor, y por debajo de este puntaje, sin TDM.

- La escala de 7 ítems de trastorno de ansiedad general (GAD-7) sirve para detectar la gravedad del trastorno de ansiedad generalizada (TAG). Se reporta a partir del autoinforme con 7 ítems, y con un puntaje total de 0 a 21. Las personas con 10 o más puntos se asignaron al grupo de posible TAG, y debajo de dicho puntaje, sin TAG.
- Rendimiento académico: corresponde al promedio de calificaciones más reciente, las advertencias académicas por mala asistencia en el último semestre y el nivel de satisfacción con el curso.

De los factores socioeconómicos preguntados, se determinó que respecto a la educación de los padres, se encontró que el 56,7% de los padres de los estudiantes (229) y el 48,5% de las madres de los estudiantes (196) eran graduados de una universidad. En cuanto al estado civil, la mayoría de los participantes eran solteros (94,3%). También se determinó que los hombres son mayores que las mujeres, vienen de familias más ricas y tienen padres con menor educación que las mujeres. En cuanto al rendimiento académico:

- El promedio de calificaciones de los estudiantes osciló entre 0,66 y 4 (con un promedio de 2,8).
- La mitad de la muestra (47,8%) había recibido una advertencia de asistencia (los estudiantes con promedios más bajos recibieron más advertencias) ($r = 0,24, p < 0,01$).
- Los mismos estudiantes con promedios más bajos también se sintieron menos satisfechos con sus estudios ($r = -0,182, p < 0,01$).
- Un menor RA se asoció con mayor edad ($r = -0,176, p < 0,01$), mayor año de estudio ($r = -0,17, p < 0,01$), menor nivel de educación paterna ($r = 0,127, p = 0,011$) y menor nivel de educación materna ($r = 0,1, p = 0,028$).

Respecto a ambos trastornos, fue posible encontrar que el 34,2% (138) de los estudiantes se evaluaron con posible TDM, mientras que el 22,3% (89) se evaluaron para un posible TAG. Además, las mujeres tuvieron puntuaciones más

altas para el cuestionario PHQ-9 con respecto a los hombres ($p < 0,001$); así como también, ellas tuvieron mayores puntuaciones en la escala de ansiedad ($p < 0,001$), y con mayores posibilidades de estar en el grupo TAG; sin embargo, para un posible TDM, los porcentajes son similares entre ambos géneros (hombres 31,6 % y mujeres 35,3 %). Las puntuaciones del cuestionario PHQ-9 y del GAD-7 no tuvieron relación con el año de estudio, la edad o el nivel educativo paterno (pues no son significativos en términos estadísticos).

Con respecto al RA, los puntajes más altos de TAG se asociaron con promedios académicos más bajos ($r = -0,176$, $p > 0,001$), pero no se asociaron con las advertencias de asistencia o el nivel de satisfacción con los estudios. Esta relación descrita está moderada por los efectos del género; es decir, que aquellas mujeres con este trastorno presentaron un RA más bajo que el de las del grupo sin TAG. Para el caso de los hombres, no hubo diferencia significativa entre ambos grupos. Para el TDM, el grupo que tenía este trastorno, también fue el que mostró el promedio de calificaciones más bajo ($p = 0,003$) y también se sentían menos satisfechos con los estudios ($p = 0,015$).

Para el seguimiento posterior a seis meses, se encontró que el grupo TDM tuvo promedios de calificaciones más bajos durante esta medición, contrario a lo que ocurrió con el grupo TAG, en donde el promedio de calificaciones fue más bajo al inicio del estudio y no difirió en la segunda medición. A partir de la hipótesis planteada, el grupo TDM tuvo un promedio más bajo comparado con aquellos que no se clasificaron de esta forma; también estuvieron menos satisfechos con los estudios y no hubo relación con las advertencias de asistencia.

Icekson *et al.* (2020) investigaron la determinación de los efectos moderadores de la conciencia y del género en la relación existente entre el optimismo disposicional y el rendimiento académico. Partieron de una muestra de 175 estudiantes de la Universidad Ben Gurion (Israel). El indicador de medición del rendimiento académico es el promedio anual de calificaciones, evaluado en una escala de 0 a 100, y el cual se recopiló al final del año. Las variables predictoras son el optimismo, la conciencia y el género (estas dos últimas también variables moderadoras). Los datos se analizaron con SPSS Windows 24.0.

De la matriz de correlación, observaron que la conciencia estuvo asociada positivamente tanto con el RA ($r = 0,21$, $p < 0,01$) como con el optimismo disposicional ($r = 0,20$, $p < 0,01$), lo que da a entender que los estudiantes con mayores niveles de conciencia tenían más probabilidades de tener perspectivas optimistas y así un mayor RA, con respecto a los estudiantes con niveles menores de conciencia. No hubo correlación significativa entre el optimismo y el RA ($r = 0,08$). Por su parte, de la prueba t se determinó que no existieron diferencias de género en el optimismo, conciencia o en el RA.

Del análisis de regresión múltiple jerárquica, se evidenció que cuando se ingresó el género de forma aislada, no se mostraron efectos importantes ni significativos sobre el RA ($\beta = 0,13$, $p = 0,08$). Cuando se ingresó la conciencia como factor moderador, esta variable predijo de forma significativa el RA ($\beta = 0,17$, $p < 0,05$). Finalmente, cuando se ingresó el optimismo, no se presentó ningún incremento significativo en el modelo ($\beta = 0,05$, $p = 0,48$). Para los otros conjuntos de interacciones, se tuvo que la interacción optimismo y conciencia, y la interacción optimismo disposicional y género, resultaron ser significativos, lo que confirma el papel moderador de la conciencia y el género.

En conclusión, un alto optimismo aumentó el RA de las mujeres con alta conciencia, y tuvo efectos adversos sobre el RA de los hombres con poca conciencia. Los resultados indican que la gran conciencia puede regular los efectos negativos de un alto optimismo, como expectativas poco realistas y exceso de confianza.

Minichil *et al.* (2020), en su estudio, trataron de determinar la magnitud del trastorno disfórico premenstrual (TDPM) y los factores asociados a este, en los cuales se incluye el rendimiento académico. Este estudio se llevó a cabo en la Facultad de Medicina y Ciencias de la Salud de la Universidad de Gondar (Etiopía). El tamaño de la muestra incluye 386 estudiantes (mujeres) con edades entre 18 y 26 años (20,9 años en promedio). Las siguientes son las variables independientes: factores sociodemográficos (edad, religión, etnia, estado civil, residencia, nivel educativo de la madre), factores obstétricos y ginecológicos (duración de la menstruación en un ciclo y su regularidad, dolor premenstrual y su nivel, edad de la menarquia, antecedentes familiares, problemas relacionados con la menstruación y el impacto en

el rendimiento académico), factores psicosociales (como el apoyo social y estrés académico), factores relacionados con las sustancias (alcohol, khat y cigarrillos). Como variable de salida, los autores plantearon el TDPM (trastorno disfórico premenstrual).

La comprobación y análisis de los datos se realizó con el EpiData v.3.1 y SPSS v.21.0. Mediante el análisis de regresión logística, se determinó que ser soltera, tener antecedentes familiares con problemas de menstruación, dolor menstrual severo, percibir un impacto del dolor menstrual en el RA, sangrado menstrual abundante, estrés percibido (clasificado como moderado o alto), un sangrado prolongado durante seis días, y el uso de khat y cigarrillos, estuvo asociado con el TDPM, bajo una significancia estadística de $p < 0,05$. Posteriormente, y luego de la regresión logística ajustada, se estableció que el dolor menstrual severo, la percepción de un impacto en el rendimiento académico debido al dolor menstrual y el alto estrés percibido, se asociaron significativamente con el TDPM.

Finalmente, se determinó que aquellas estudiantes que enunciaron tener un impacto negativo en el RA debido al dolor menstrual, tienen 2,31 veces más posibilidades de desarrollar el TDPM, en contraste con las estudiantes que no lo percibieron así. El desarrollo del TDPM también fue 2,82 veces mayor cuando las estudiantes percibieron un dolor menstrual mayor. Además, se determinó que existía 3,52 veces más posibilidades de desarrollar TDPM en las estudiantes con alto estrés percibido.

Vargas *et al.* (2011) determinaron la influencia de algunas características propias de los estudiantes sobre el rendimiento académico. El indicador de medición del rendimiento académico se consolidó a través de los promedios finales del primer año de los estudiantes de la carrera de medicina; es decir, por cada asignatura vista (anatomía, bioquímica, biología del desarrollo, biología celular y tisular, psicología médica, salud pública), así como un promedio general de las mismas. Esta información se calculó mediante los expedientes académicos. Se empleó una muestra compuesta de 94 estudiantes (48 hombres y 46 mujeres) del Programa de Alta Exigencia Académica (PAEA) de primer año de medicina, de la Universidad Nacional Autónoma de México (UNAM).

Se estableció que el tipo de bachillerato de procedencia no representó un factor que pudiera generar diferencias entre las calificaciones de los estudiantes. Sin embargo, el examen general de conocimientos sí representó diferencias significativas en términos estadísticos entre los grupos de bachillerato ($p < 0,001$). A partir de la aplicación de los modelos lineales generales univariados, se encontró que, para los promedios de las asignaturas evaluadas, las variables predictoras significativas son:

- **Anatomía:** examen general de conocimientos (EGC), fluidez, psicastenia y esquizofrenia, con significancias estadísticas que oscilaron entre 0,001 y 0,27.
- **Biología del desarrollo:** EGC ($p = 0,018$) (único predictor).
- **Bioquímica:** EGC, fluidez, puntaje global de creatividad, esquizofrenia y manía con significancias estadísticas entre 0,000 y 0,036.
- **Biología celular y tisular:** EGC, fluidez, originalidad, índice de creatividad, esquizofrenia, manía, con significancias estadísticas que variaron entre 0,000 y $p = 0,033$.
- **Salud pública:** EGC, hipocondriasis, paranoia, esquizofrenia y manía, con un valor p para significancia estadística que osciló entre 0,001 y 0,028.
- **Psicología médica:** EGC, fluidez, originalidad, hipocondriasis, paranoia, psicastenia, esquizofrenia y manía, con valores para la significancia estadística entre 0,000 y 0,040.
- **Promedio general:** EGC ($p = 0,000$).

Por otro lado, se detectó la presencia de psicopatología en los estudiantes finalizando el año de medición. Específicamente, el 80,21 % (77 estudiantes) no presentaron datos que se pudieran asociar a este trastorno, mientras que 19 de ellos sí los presentaron (19,79%): 5 hombres y 14 mujeres. Sin embargo, en términos estadísticos, no hubo diferencias significativas entre estos grupos para el promedio final general, biología del desarrollo, bioquímica, biología celular y tisular y psicología médica (pero sí para las materias de anatomía y salud pública). Como conclusión, el Examen General de Conocimientos (EGC) fue el único predictor significativamente estadístico

del promedio final de las asignaturas en separado, y para el promedio final general. Por su parte, el pensamiento creativo y los rasgos de la personalidad son predictores significativos del promedio final de las materias, aunque no se presentó constante para todas. Además, se determinó que ni los aspectos sociodemográficos, ni antecedentes académicos (preseas, dominio de idiomas, experiencia en investigación, docencia) se relacionaron con el RA.

Albela-García *et al.* (2020) determinan el rendimiento académico mediante técnicas de aprendizaje, para una asignatura llamada radiología, en la Facultad de Odontología de la Universidad de San Martín de Porres (Perú). Se considera una muestra de 80 estudiantes. El indicador de medición del rendimiento académico no se especifica para este estudio, simplemente en el estudio se menciona el rendimiento académico de los estudiantes en la asignatura de radiología de la carrera de odontología. Las variables independientes corresponden al uso de técnicas de aprendizaje.

10.3.1.1. Análisis

Como primer paso del estudio, se recolectó la información sobre el rendimiento académico de los estudiantes antes de alguna técnica de aprendizaje (pre-test). Luego, se empleó alguna de las técnicas propuestas de aprendizaje (tutorías, y tutorías más seminarios). Por último, se requirió el RA final de los estudiantes (post-test) luego de la aplicación de la técnica de aprendizaje, con el fin de realizar comparaciones. Con la prueba de rangos de Wilcoxon, se comprobó que ambas técnicas (tutoría, y tutoría con seminario) muestran una diferencia estadísticamente significativa ($p < 0,001$) entre los promedios iniciales, y los promedios después de emplear una técnica de aprendizaje. De acuerdo con esto, y como conclusión del estudio, se encontró que posterior a la aplicación de técnicas de aprendizaje, se generan distintas puntuaciones en los promedios el RA de los estudiantes (es decir, comparativamente entre el momento antes y después de haberlos empleado); la diferencia se evidenció específicamente en una mejoría de las notas.

Salmani *et al.* (2020) plantearon determinar tanto la prevalencia como los efectos en la salud y en el rendimiento académico debido a los trastornos de sueño, diferenciados por género. Se parte de una muestra de 637 estudiantes de la Universidad Sultán (Omán). El indicador de medición del rendimiento

académico en la presente investigación fue el promedio de calificaciones. Con el fin de realizar comprobaciones y comparaciones, este indicador se agrupó en dos categorías (los promedios de calificaciones menores a 2, y los promedios de calificaciones mayores o iguales a 2). Las variables independientes corresponden a las características demográficas, así como a la prevalencia de los trastornos del sueño. Los análisis estadísticos se llevaron a cabo con el SPSS v.25.0.

Respecto al rendimiento académico, los estudiantes con un promedio de calificaciones superior a 3 fue el 28,9%, el 68,0% presentó promedios de calificaciones que oscilaron entre 2 y 3, mientras que el 3,1% (20) presentó un promedio de calificaciones menor a 2. En los trastornos de sueño más presentados en los estudiantes, se encuentran: narcolepsia (con el 88,2%), síndrome de piernas inquietas (41,1%), insomnio (36,4%), apnea obstructiva del sueño (21,5%), trastornos afectivos (14,1%). Los siguientes trastornos son los menos frecuentes: trastorno del ritmo circadiano (9,1%), pesadillas (7,4%) y sonambulismo (4,2%). En cuanto al género, de los resultados de las mujeres encuestadas, se encontró que la narcolepsia y el insomnio son significativamente mayores entre ellas ($p = 0,000$ y $p = 0,001$, respectivamente). Respecto a la edad, los trastornos más prevalentes son el insomnio y la apnea obstructiva del sueño en estudiantes con edades iguales o mayores a 20 años ($p = 0,032$ y $0,007$, respectivamente). Para el resto de los trastornos estudiados, no se encontraron relaciones significativas con el género o la edad.

Como conclusión, resultó significativa la asociación entre los trastornos de sueño y un bajo promedio de calificaciones entre las estudiantes: considerando que los universitarios se encuentran más expuestos a sufrir de trastornos de sueño, se puede afirmar que estos se correlacionan con la disminución en el RA de las mujeres (en concreto, con los trastornos de apnea obstructiva del sueño, sonambulismo y trastornos afectivos, con una significancia estadística de $p = 0,002$, $p = 0,003$ y $p = 0,021$, respectivamente). Para los hombres, no existió asociación significativa con ningún trastorno.

Bayona (2017) desarrolló su trabajo para determinar si existe alguna correlación entre los hábitos de estudio y rendimiento académico de los estudiantes. Se empleó una muestra de 50 estudiantes en una primera medición, y 21 estudiantes para la segunda de la carrera de ingeniería de sistemas de la

Universidad Francisco de Paula Santander (Colombia). El indicador de medición del rendimiento académico se asumió como el promedio ponderado del estudiante en el semestre académico. Las siguientes son las variables independientes objeto de estudio: las condiciones ambientales para el estudio, planificación del estudio, utilización de materiales, asimilación de contenidos y sinceridad, las cuales conforman una variable superior, que corresponde a los hábitos de estudio. Se plantearon dos hipótesis de acuerdo con el objetivo del estudio para ser confirmadas o descartadas, mediante los resultados obtenidos:

- **H0:** “el rendimiento académico es independiente de los hábitos de estudio”.
- **Ha:** “el rendimiento académico depende de los hábitos de estudio”.

En el análisis de los resultados, se empleó el coeficiente de correlación R entre las variables de estudio, con el fin de confirmar si el RA dependía de los hábitos de estudio. A partir de la primera medición, se encontró que la única escala en la que los estudiantes presentaron buenos resultados, fue en las condiciones ambientales; sin embargo, para las escalas de planificación del estudio, utilización de materiales, asimilación de contenidos y grado de sinceridad, la tendencia se registró como normal. Concluyó que los hábitos de estudio no se correlacionan con el rendimiento académico, lo que significa que para tener buen rendimiento no necesariamente se deben tener buenos hábitos de estudio.

Bautista-Rodríguez y Gatica-Lara (2020) determinan la relación existente entre las características sociodemográficas (sexo, edad, lugar de residencia, entre otras), académicas (semestre cursado, asignaturas no acreditadas, entre otras), laborales (actividad laboral, antigüedad laboral, entre otras) y habilidades digitales (disposición de tecnología, habilidades digitales, entre otras), con el rendimiento académico de la carrera de técnico superior universitario en urgencias médicas (TSU en UM) en la modalidad en línea, de la Universidad Abierta y a Distancia de México (UnADM). La muestra de trabajo fue de 126 estudiantes. El indicador de medición del rendimiento académico, propuesto por los autores, corresponde a las calificaciones globales, las cuales son otorgadas por la coordinación de carrera.

Para el análisis de los resultados, se empleó el *software* SPSS. De forma específica, se empleó el coeficiente de Spearman y Pearson, con el fin de identificar el grado y la dirección de la correlación entre las características evaluadas. Se determinó que algunas de las variables independientes (personales, laborales y habilidades digitales) frente al RA no presentaron una correlación significativa en términos estadísticos entre ellas (pues el valor p superó el límite de 0,05). Entre estas se encuentran: el sexo, la edad y el lugar de residencia (con coeficientes “ r ” iguales a 0,085, 0,073 y -0,040, respectivamente); la actividad laboral ($r = -0,144$) y la antigüedad laboral ($r = -0,132$); la disposición de tecnología ($r = 0,055$) y las habilidades digitales ($r = 0,080$). De las variables que sí presentaron poder predictivo sobre el RA en términos estadísticos son significativas ($p < 0,005$), como las variables académicas “semestre de la carrera que cursa” ($r = 0,450$) y “asignaturas no acreditadas” ($r = -0,294$). Esto significa que un estudiante activo puede reprobar menos asignaturas, así como cuando se avanza en la carrera, el RA tiende a mejorar. Se determinó también que el porcentaje en que estas variables predicen el RA es igual al 23 % (mediante el cálculo de R^2 ajustada).

Nicholson *et al.* (2016) investigan para determinar si una herramienta con un enfoque profundo para el aprendizaje (como lo es un taller interactivo de anatomía), podría afectar el RA de los estudiantes, a diferencia de los enfoques que se adoptan con las conferencias o prácticas de laboratorio. El tamaño de la muestra consta de 66 estudiantes pertenecientes a la asignatura de anatomía musculoesquelética funcional (Sídney). En esta investigación, el indicador de medición del rendimiento académico se midió con un examen de anatomía antes y después del taller, con el fin de realizar comparaciones entre ellos. El análisis de los datos se hizo con la herramienta IBM SPSS 20.0, en donde se calcularon estadísticas descriptivas (para los datos demográficos y los resultados de los exámenes), un ANOVA de un factor y un análisis de regresión lineal.

Los resultados de los asistentes y no asistentes al taller en el examen de mitad de semestre son, en promedio, de 73,3 y 69,7 puntos, respectivamente. En cuanto al examen final, 74,9 para los asistentes y de 67,1 para los no asistentes. Este comportamiento corresponde a una diferencia significativa en términos estadísticos, confirmada mediante el ANOVA de un factor, para

los resultados de ambos exámenes (examen de mitad de semestre $p = 0,05$ y examen final $p < 0,001$), observando así que los promedios de los asistentes a los talleres eran mejores antes y después del taller que los del otro grupo.

Lo anterior denota un cambio positivo para los asistentes al taller, y un cambio negativo para los no asistentes. Esto también se evidencia con la regresión lineal (donde se relacionaban los resultados de ambos exámenes), pues el mejor ajuste se dio en $R^2 = 0,711$, con una pendiente de 0,832 para los asistentes, y un $R^2 = 0,569$ con una pendiente de 0,695 para quienes no asistieron. Exactamente, se presentó un aumento de 1,6% frente a una disminución de 2,3%. En conclusión, el taller interactivo sí fue efectivo para la mejora en el RA de los estudiantes, por lo que, a través de esta modalidad de aprendizaje, se promovió tanto la investigación, así como la profundidad en el aprendizaje.

Montoya-Arenas *et al.* (2021) presentan como objetivo el análisis de los factores de la capacidad intelectual (CI), como cociente intelectual verbal (CIV) y cociente intelectual ejecutivo (CIE) y las funciones ejecutivas (FE), como memoria operativa verbal, memoria operativa visual, rastreo y monitoreo visual, velocidad de procesamiento, organización verbal, categorización visual, flexibilidad cognitiva y control inhibitorio; sobre el rendimiento académico de los estudiantes. La muestra consta de 42 estudiantes de la Fundación Universitaria Luis Amigó (Colombia). El indicador de medición del rendimiento académico se estableció según el promedio aritmético ponderado de las calificaciones de los semestres cursados. Este se clasificó como bajo rendimiento académico si el valor obtenido era menor a 2,9, mientras que un alto rendimiento académico correspondía a un valor de 4,0 sin materias perdidas. Con respecto al análisis de las variables CI y FE, los estudiantes con alto rendimiento mostraron una mejor ejecución de las tareas, salvo en la memoria operativa visual, la velocidad de procesamiento perceptual visual y errores persistentes en la flexibilidad cognitiva. En el caso contrario, la memoria operativa visual fue más rápida en el grupo de bajo rendimiento (estadísticamente significativo con $p \leq 0,05$), además de demostrar un efecto completo ($d=1,93$) para dicho grupo. Lo anterior puede identificarse como una acción compensatoria.

Muñoz-Villanueva *et al.* (2020), en su investigación, plantearon el estudio de la relación entre el desayuno y el rendimiento académico del técnico de nivel superior en enfermería de la Universidad Adventista de Chile. El indicador de medición del rendimiento académico se midió según el promedio de las calificaciones de las materias de anatomía y fisiología, enfermería en urgencias y principios de farmacología. Se empleó una encuesta autoadministrada para la recolección de información, en la cual se indagaron aspectos sobre hábitos alimenticios. Específicamente: hábitos de desayuno, alimentos consumidos en el desayuno, frecuencia de consumo, y aspectos demográficos como el trabajo, entorno familiar, diagnóstico de problemas de aprendizaje y el rendimiento académico. Al realizar la prueba de Fisher, se demostró que no existe relación entre las variables “desayuno regular” y “consideras que tienes problemas de atención o memoria”, así como las variables “desayuno regular” y “sentirse despierto” en las clases de las primeras horas, pues se obtuvo un nivel de significancia estadística de $p = 1,0$ y $p = 0,516$, respectivamente. Con la prueba Kruskal Wallis, se concluyó que no se presentó una relación significativa en términos estadísticos, entre la calidad del desayuno y el RA, pero no se descarta su influencia del todo.

10.3.2. Estadística - Matemática

Huamán (2018) planteó su investigación con el fin de demostrar la importancia y relevancia social que tiene el rendimiento académico en la calidad de la preparación del estudiante. Trabajo desarrollado en la carrera de estomatología de la Universidad Cayetano Heredia (Perú). Esta investigación tuvo como indicador de medición del rendimiento académico las calificaciones de los estudiantes, obtenidas en cada uno de los cursos tomados en los cinco años de carrera. Las variables independientes empleadas en este estudio son el área formativa, sexo, curso/asignatura; mientras la variable dependiente corresponde al rendimiento académico. En síntesis, en las áreas básicas, formativas y profesionales, el comportamiento discriminado por género evidencia un mejor rendimiento académico para las mujeres. Respecto a la comparación entre áreas, se aprecia que la más crítica es el área formativa, pues en esta se incluyen asignaturas como estadística, microbiología o patología, las cuales representan la entrada al enfoque de la carrera, que se

relacionan directamente con temas de salud, y pueden resultarles difíciles a los estudiantes que no cuentan con conocimientos previos en estos temas. Por el contrario, los estudiantes demuestran un mejor desempeño en el área profesional, que permite indagar que se realiza un mayor esfuerzo por las materias que no se imparten únicamente por medio de la teoría, sino de la práctica, aplicando todo lo aprendido previamente.

Shahrakipour *et al.* (2016) buscaron la relación entre variables independientes que corresponden a datos demográficos, como edad, sexo, tipo y nivel educativo de los padres/madres, así como el tipo de facultad y la variable dependiente corresponde al rendimiento académico. Trabajo desarrollado en una universidad de Irán de las carreras de enfermería, rehabilitación, ciencias médicas, odontología, paramédicos, durante el 2015. El coeficiente de correlación sugiere que existe correlación significativa entre el género y el rendimiento académico, pues las mujeres obtuvieron logros superiores a la media, lo cual es consistente con estudios previos. También hay una correlación significativa entre la educación parental/maternal y el rendimiento académico, que puede estar dada por la motivación que, desde pequeños, los estudiantes observan en sus familias, comportamientos, interacciones, estilos de vida, entre otros. Por otro lado, la relación entre edad y facultad, no fue significativa con el rendimiento académico.

Ñaupari (2014) evaluó el impacto de la modalidad presencial y virtual en el rendimiento académico de los estudiantes de la Universidad Continental (Perú). La muestra consta de 338 estudiantes en modalidad presencial y 167 en modalidad virtual. El indicador de medición del rendimiento académico se construyó en función de dos indicadores básicos: promedio (en una escala de 20 puntos) e índice de aprobación (porcentaje de asignaturas cursadas sobre el total de asignaturas inscritas). Desde la escala unificada del rendimiento académico, y al comparar las dos modalidades, se observa que un mayor porcentaje de los estudiantes en modalidad virtual, presenta un rendimiento académico bajo (69,5%) comparado con el 45,8% registrado en la modalidad presencial. En cuanto a un rendimiento alto, no hay diferencia significativa (9,6% modalidad presencial frente al 9,5% modalidad virtual). Estos resultados indican que siete de cada diez estudiantes en modalidad virtual desaprobaban con notas menores a 10,5 y logran aprobar menos de

ocho asignaturas de diez que cursan. En contraste, en la modalidad presencial, desaprueban de esta manera el 45,8 %. En general, para ambos tipos de modalidades, y como se menciona en el estudio, el rendimiento académico es bajo y no corresponde a un nivel deseado, especialmente para los estudiantes en modalidad virtual. Estos resultados son similares a la teoría consultada por el autor, en donde se destacaron los grupos en modalidades presenciales en niveles de conocimiento suficiente, con la diferencia de que se observaba que, en ambos tipos de modalidades, la mayoría de los estudiantes se encontraba en un desempeño sobresaliente (regular), difiriendo de los resultados del presente estudio.

Vargas *et al.* (2019) plantearon determinar la medida en que un modelo matemático es capaz de explicar la relación entre el puntaje del examen de ingreso y el rendimiento académico de los estudiantes que ingresan a la Universidad Nacional Jorge Basadre Grohmann (Perú). Parte de una muestra de 1551 estudiantes. El indicador de medición del rendimiento académico para este modelo matemático, consiste en el *ranking* académico de los estudiantes ingresantes a la universidad, que, para los datos empleados en el estudio, oscilan entre 0,091 y 17,591, con un promedio de 10,674. Las variables predictoras de este estudio corresponden al puntaje de examen de ingreso mediante dos modalidades distintas (puntaje de admisión - puntaje del centro preuniversitario), mientras la variable dependiente es el rendimiento académico. Para este estudio, se consideraron tres hipótesis (una general y dos específicas):

- **Hipótesis general:** “relación entre el puntaje de examen de ingreso y el rendimiento académico”.
- **H1:** “relación entre el puntaje de examen de admisión y el rendimiento académico”.
- **H2:** “relación entre el puntaje de examen del CEPU (Centro preuniversitario) y el rendimiento académico”.

Del análisis inferencial a través de la correlación de Pearson, se evidencia que existe una relación directa entre las variables “puntaje de examen de ingreso” y “rendimiento académico”, pues hay relación directa y significativa entre estas variables con un $p = 0,372$. Asimismo, esto sucede con el puntaje

de ingreso y el rendimiento académico. De esta manera, se confirma la hipótesis general (obtenida con el primer resultado) y las hipótesis específicas del estudio (segundo y tercer resultado). De lo anterior, se observa que la diferenciación entre las modalidades parece no ser buen predictor del RA, pues, según la teoría, los estudiantes obtienen resultados similares independientemente de la modalidad de ingreso.

Durán y Rosado (2019) indagan sobre las habilidades en producción de textos y comprensión de los estudiantes de ingeniería y su relación sobre el rendimiento académico. La muestra tomada fue de 105 estudiantes del primer semestre académico de ingeniería de la Universidad Francisco de Paula Santander (Colombia). En esta investigación, se consideran dos indicadores de medición del rendimiento académico: el nivel de conocimiento de los estudiantes en la prueba SABER 11 y el reporte de calificaciones de cada corte semestral en las asignaturas de lenguaje. Las variables independientes corresponden a aquellas relacionadas con la comprensión lectora de los estudiantes (didácticas docentes y prueba escrita aplicada), mientras la variable dependiente corresponde al rendimiento académico. Como resultado del cruce de datos, se identificó que aquellos estudiantes con deficientes resultados en la prueba SABER y las calificaciones promedio en cada corte académico (indicadores del rendimiento académico), también presentan bajos resultados en la prueba diagnóstica. Esto permite inferir que los estudiantes con este comportamiento pueden reprobar el semestre, incluso corren el riesgo de desertar.

Bonsaksen (2016) llevó a cabo su trabajo para determinar los factores de importancia para el desempeño académico y la satisfacción de los estudiantes de terapia ocupacional. En esta investigación, se plantea el indicador de medición de rendimiento académico como el promedio de calificaciones autoevaluado con base en los exámenes completados hasta el momento en el programa educativo, clasificadas como excelentes (1), muy buenas (2), buenas (3), satisfactorias (4), suficientes (5) y falla (6). En este estudio, al momento de la recolección, esta medida se basó en un solo examen para los estudiantes de primer año, mientras que los de segundo y tercer año estaba basada en 6 y 10 exámenes. La muestra se compone de 123 estudiantes.

En cuanto al rendimiento académico, se encontró que los participantes de mayor edad, los que vivían con un cónyuge/pareja y los que tenían experiencia previa en educación superior, presentaron un mejor rendimiento académico frente a los participantes que eran más jóvenes, solteros y sin experiencia en educación superior antes de ingresar a la educación en terapia ocupacional. La satisfacción, por otro lado, fue mayor entre los estudiantes que no tenían la terapia ocupacional como su opción educativa de mayor prioridad y entre aquellos que dedicaron más tiempo a los autoestudios, en comparación con sus contrapartes. De manera específica, las variables independientes explican solo una cantidad moderada de varianza en las medidas de resultado, pues las relacionadas con la educación son las más importantes para explicar el rendimiento académico y la satisfacción con el programa educativo.

Liang *et al.* (2018) investigaron los efectos de la etnia y el género en el promedio de calificaciones de fin de año de los estudiantes universitarios en Texas (Estados Unidos). De manera específica, se investigaron los estigmas étnicos y de género que podrían llegar a contribuir al rendimiento académico. El indicador de medición del rendimiento académico se midió con el promedio numérico de calificaciones. Las variables independientes en este estudio son género y etnia, mientras la variable dependiente es el rendimiento académico. La muestra se compone de 1122 estudiantes. Se empleó un ANOVA bidireccional para el análisis de los datos. Según los resultados del estudio, el rendimiento académico se vio afectado por el género, ya que las estudiantes universitarias obtuvieron mejores promedios de calificación en comparación con los hombres. Asimismo, el rendimiento académico no estuvo influenciado por el origen étnico o la interacción entre el género y el origen étnico, pues no hubo significancia estadística que demostrara esas relaciones.

Pavelea y Moldovan (2020) establecieron la influencia de múltiples y variados factores (como la asistencia a clase, el rendimiento académico previo, la intervención y atención en clase, el horario de clase, el género, entre otros) sobre el rendimiento académico de estudiantes de la Universidad Babeş-Bolyai (Rumanía). En este estudio, el indicador de medición del rendimiento académico corresponde a la nota final obtenida por el *paper* enviado como parte del examen del curso “escritura académica”. Para dicho examen, los

estudiantes redactaron y enviaron una propuesta de investigación de un tema de su preferencia con tres partes (introducción, revisión de la literatura y metodología). Las variables independientes corresponden a la asistencia, atención e intervención en clase, características de los estudiantes (como motivación y conciencia); otros factores (edad, género, lugar de residencia, distancia entre el lugar de estudio y el hogar, tiempo del día en que se asistió a un seminario). Se evidenció que la asistencia y el RA están positivamente correlacionados (con $R^2 = 0,278$ y $p = 0,000$), lo que quiere decir que un incremento de una unidad en la asistencia, predice un incremento en la nota final de los estudiantes. Al contrario, la atención no se correlacionó significativamente con el RA, difiriendo con lo esperado por los autores. En resumen, los principales factores que pueden influenciar el rendimiento académico de esta muestra de estudiantes, son tanto la asistencia a los seminarios, como la intervención en la clase, por lo que esta condición podría ser la limitación del presente estudio.

Farrokhi-Khajeh-Pasha *et al.* (2012) buscan determinar la validez predictiva que tienen los resultados del “Konkoor” (un examen nacional altamente competitivo, que es el único medio de ingreso a las escuelas médicas) en combinación con los promedios de calificaciones de secundaria, sobre el rendimiento académico de los estudiantes. La muestra consta de 2126 estudiantes de escuelas médicas públicas en Irán. El indicador de medición de rendimiento académico está basado en tres diferentes puntajes:

- **Puntaje CBSE (“examen integral de ciencias básicas”):** este examen nacional, que se lleva a cabo al final del quinto semestre, se centra en el entendimiento y aplicación de las áreas de ciencias básicas. Este es un requisito para que los estudiantes pasen a la etapa clínica.
- **Puntaje CPIE (“examen integral previo a la pasantía”):** este examen, también de carácter nacional, se lleva a cabo al final de la etapa clínica, y evalúa el conocimiento clínico de los estudiantes. Este es un requisito para empezar una pasantía.
- **Promedio de calificaciones de la escuela de medicina:** es un promedio calculado a partir de todos los cursos completados antes de la pasantía, y se ponderaron de acuerdo con el número de horas de cada curso.

Las variables predictoras se constituyen por el puntaje total del Konkoor y de sus ocho subsecciones (idioma persa, idioma árabe, estudios religiosos, idioma extranjero, matemáticas, biología, física y química), así como el promedio de calificaciones de secundaria (que específicamente corresponde al promedio de notas de los últimos tres años de los cursos de secundaria, que tiene un puntaje máximo de 20). Las correlaciones entre las predictoras (promedio de calificaciones de secundaria, idioma persa, idioma árabe, estudios religiosos, idioma extranjero, el total de la categoría sin ciencia, matemáticas, biología, física, química, el total de la categoría con ciencia y el puntaje total del Konkoor) y las variables de salida (CBSE, CPIE y promedio de calificaciones de la escuela de medicina) son estadísticamente significativas para $p < 0,001$. Observaron que la mayor de todas estas correlaciones se presenta entre el puntaje total del Konkoor y el CBSE con un coeficiente de correlación de 0,473, el cual se puede considerar mediano, igual que las relaciones entre el puntaje Konkoor y el promedio de calificaciones de la escuela de medicina y el CPIE. El promedio de calificaciones de secundaria, por su parte, tiene un efecto mediando en el promedio de calificaciones de la escuela de medicina y en el CBSE. La correlación más pequeña se dio respecto a los estudios religiosos y el CPIE. Al comparar los coeficientes de correlación de las categorías del Konkoor que incluían ciencia y los que no la incluían, respecto a las variables de salida, el CPIE presentó una mejor correlación con los temas que incluían ciencia, mientras que el promedio de calificaciones de la escuela de medicina se correlacionó mejor con los temas que no incluían ciencia.

Lenskiy *et al.* (2020) se centraron en la determinación de las consecuencias que tiene al tomar un receso académico durante la carrera y por un número determinado de semestres, sobre el rendimiento académico de los estudiantes, contrastado con quienes no tomaron ningún tipo de receso (grupo de control). Trabajo planteado en las carreras de ingeniería de la información y las comunicaciones e ingeniería de la Universidad de Tecnología y Educación (Corea). Se parte de una muestra de 653 estudiantes. La variable independiente en este estudio corresponde al receso académico (si este se tomó o no durante el ciclo total académico y el número de semestres que

duró el receso), mientras que la variable dependiente es el rendimiento académico. De forma específica, la muestra se dividió de la siguiente manera:

- 24 estudiantes tomaron un receso de 1 semestre (4%)
- 45 estudiantes tomaron un receso por 2 semestres (7%)
- Ningún estudiante tomó un receso por 3 semestres (0%)
- 263 estudiantes tomaron un receso por 4 semestres (40%)
- 159 estudiantes tomaron un receso por 5 o más semestres (24%)
- 93 estudiantes estudiaron de forma continua sin interrupciones (14%)
- Estudiantes excluidos de la muestra (11%)

De forma inicial, se observa un incremento en el RA de los estudiantes luego de haber tomado un receso académico, pero al realizar un análisis estadístico de correlación en la comparación de distribuciones del RA de los estudiantes con 1, 2, 4, 5 o más descansos, el valor p superó el límite 0,05. Quiere decir que tomar recesos académicos durante la carrera, no influye en el RA de los estudiantes. De aquellos quienes sí tomaron descansos, se encontró que las razones son múltiples para esta decisión (servicio militar, viajes, estudios en el exterior, enfermedad, frustración, entre otras), y según sus percepciones, el 41 % percibió un efecto positivo en el RA, el 29,5 % percibió un efecto negativo, mientras que el 29,5 % no percibió ningún cambio. Quienes afirmaron tener un efecto positivo en el RA, una mayoría enunció que había tenido un receso largo (66 %).

Grob *et al.* (2015) evaluaron la relación entre un puntaje de selección, puntaje previo durante la enseñanza media y el rendimiento académico de estudiantes de primer año de odontología de la Universidad de los Andes (Chile), en la asignatura de morfología. El tamaño de la muestra consta de 312 estudiantes. El indicador de medición del rendimiento académico, corresponde a las notas definitivas obtenidas en la asignatura de morfología al final del primer año de estudio. Las variables independientes corresponden al género, al promedio puntaje de notas de educación media, al puntaje prueba de selección universitaria (PSU) y al tipo de colegio de procedencia (según el financiamiento del mismo: municipal, particular subvencionado y

particular). Se determinó que existe una relación significativa en términos estadísticos, entre la nota de enseñanza media y el RA ($r = 0,35$ y $p < 0,0001$), así como una relación significativa en términos estadísticos, entre el puntaje obtenido en el PSU y el RA ($r = 0,39$ y $p < 0,0001$). Lo anterior indica que sí existe correlación positiva entre el RA, con las notas de enseñanza media, y el PSU, lo que se traduce en que, a mayores puntajes en las variables predictoras, se refleja en un mejor RA en la asignatura de morfología. Respecto a la variable de género, se evidenció igualmente una correlación positiva y estadísticamente significativa entre las variables de estudio, nota de enseñanza media y PSU sobre el RA de las mujeres. En cuanto al tipo de colegio, se presentó una correlación también positiva y significativa para el puntaje de enseñanza media y el PSU sobre el RA. Las demás relaciones no se mostraron significativas (salvo en la relación entre los colegios municipales y la nota PSU). Como conclusión, se evidencia que el RA se encuentra correlacionado con el género, el tipo de colegio (si es particular) y en menor grado con los años de ingreso, lo que finalmente deja en entredicho que las variables seleccionadas como independientes, sí son predictoras del RA bajo algunos escenarios.

Sanguinetti *et al.* (2013) determinaron los factores relacionados con el rendimiento académico en alumnos de fisiopatología. Para ello, se consolidó una muestra de 188 estudiantes (91 % de mujeres) de la carrera de nutrición de la Universidad Nacional de Salta (Argentina), quienes estaban tomando la cátedra de fisiopatología. El indicador de medición del rendimiento académico en el presente estudio fue dado por “regularizar la asignatura”, lo cual significa, aprobar los exámenes parciales, y de este modo, poder tener acceso al examen final. Igualmente, los estudiantes debieron presentar una asistencia a las clases del 70 %. Las variables independientes del estudio son: datos demográficos (edad, sexo, provincia de origen), condición laboral de los estudiantes, si tenían computadora en el hogar, el uso del correo electrónico, uso de redes sociales, realización de deportes, participación en las ONG, satisfacción con la carrera, libros leídos durante el último año (que no se relacionaran con el estudio), aprobación del examen final de anatomía (la cual era una condición para cursar la asignatura de estudio), situación económica y tipo de educación secundaria. La variable independiente

correspondió al rendimiento académico. En cuanto a las relaciones entre el RA y las variables independientes propuestas, mediante el test de Fisher, se encontró que no existe diferencia significativa en términos estadísticos, entre los estudiantes de una provincia de origen y otra ($p = 0,74$), así como tampoco existe diferencia significativa sobre la actividad laboral de los estudiantes ($p = 0,48$). Otras de las variables que no se correlacionaron con el RA: participación en las ONG ($p = 0,61$), recursante ($p = 0,138$), tipo de educación media ($p = 0,75$) y situación económica ($p = 0,138$). Para el caso de las variables que sí presentaron diferencias significativas en términos estadísticos sobre el RA de los estudiantes, se encuentran los que aprobaron anatomía ($p = 0,0002$), así como tener hábitos de lectura ($p = 0,019$).

Muñoz-Comonfort *et al.* (2014) buscaron establecer la correlación existente entre una evaluación diagnóstica y el rendimiento académico de los estudiantes universitarios de la licenciatura de médico cirujano, de la Facultad de Medicina de la UNAM (México). La muestra es de 6590 estudiantes. El indicador de medición del rendimiento académico corresponde a la calificación promedio de los exámenes departamentales de cada una de las seis asignaturas tomadas durante el primer año de carrera (anatomía, embriología, bioquímica, biología celular, salud pública, salud mental). La variable independiente es la evaluación diagnóstica realizada a los mismos (calificación promedio de los 8 temas que lo componen), tomando en consideración el ciclo de inicio de la carrera, el tipo de escuela de procedencia. En cada uno de los ciclos escolares, se observó que aquellos estudiantes que no provenían del sistema de la UNAM, presentaron mayores calificaciones. Se encontró que las diferencias son significativas en términos estadísticos por pares de medias para los tres tipos de bachillerato (a un nivel de 0,05).

Buhagiar y Leo (2018), en su proyecto, investigaron la existencia de una relación entre el rendimiento académico y el acceso a material complementario mediante procesos de ludificación, es decir, empleando juegos que apoyan el aprendizaje. Se planteó para 116 estudiantes de la Facultad de Administración de la Universidad de Florida Central (Orlando, Florida). El indicador de medición del rendimiento académico correspondió a la calificación final de la clase estudiada. La variable independiente está dada por los juegos para el aprendizaje.

La parte del grupo que ejecutó el proceso de ludificación, empleó siete diferentes juegos como medio para reforzar los conocimientos del curso (unir, tarjetas flash, búsqueda de palabras, crucigrama, ahorcado, esta o aquella, *last chance cadet*), mientras que el grupo de control no accedió a ningún juego. Al finalizar la asignatura, cada estudiante obtuvo su calificación final (independientemente de si había empleado los juegos o no durante su aprendizaje), con el fin de hacer las comparaciones entre estas. Para dicha comparación, se empleó una prueba t de dos colas (para varianzas desiguales), y la prueba F (para la igualdad de varianzas) mediante el *software* SAS, con un nivel de significancia estadística de 0,05. En comparación, se encontró que la calificación promedio final del grupo que empleó los juegos fue menor a la calificación promedio del grupo de control (80,7 y 82,4, respectivamente). A pesar de lo anterior, no se determinó una diferencia significativa en términos estadísticos, entre las puntuaciones de los dos grupos del curso. Específicamente, con la prueba t (varianzas desiguales) se encontró un $p < 0,09$, mientras que para la prueba F (igualdad de varianzas) se encontró un $p < 0,000001$. Se concluye que los juegos en el aprendizaje no determinan un mejor rendimiento académico.

Gómez *et al.* (2011) intentan determinar el rendimiento de los estudiantes de la unidad académica multidisciplinaria zona media (UAMZM) de la Universidad Autónoma de San Luis Potosí (México). Plantearon dos hipótesis de estudio:

- **H1:** “existe relación entre el rendimiento del estudiante y el género, carrera y semestre al que está inscrito el alumno”.
- **H2:** “existe relación entre el rendimiento de los estudiantes, el promedio de calificaciones y la satisfacción con las carreras”.

De un total de 837 estudiantes de las carreras de licenciatura en administración, contaduría pública, licenciatura en enfermería, licenciatura en mercadotecnia e ingeniería civil, tomaron una cantidad de 238 estudiantes. De esta muestra, el 40 % son hombres y el 60 % mujeres. Se realizó un instrumento de tipo cuestionario elaborado por profesores de la institución, conformado por 26 ítems, de los cuales, once son de carácter sociodemográfico de tipo nominal, ocho para medir la satisfacción de los estudiantes con su carrera

y siete para medir el rendimiento percibido, ambos de tipo ordinal en una escala de seis puntos. Se utilizaron las variables: género, carrera y semestre. Respecto al género, se empleó la prueba T, para la carrera se empleó la prueba ANOVA y para la variable “semestre” se utilizó el coeficiente de Spearman. Entre los hallazgos, los autores encontraron que existe evidencia estadística de la asociación entre el género y el rendimiento percibido ($t = -2,751$; significancia = 0,007). El género femenino cuenta con un rendimiento académico superior con respecto al género masculino. A su vez, hallaron una evidencia débil entre el semestre y el rendimiento percibido ($\rho = -0,190$; significancia = 0,009) y una nula relación entre la variable carrera y el rendimiento del estudiante ($F = 0,821$; significancia = 0,513).

Siddiqi y Memon (2017) tienen como propósito determinar si la adicción a internet impactaba en la gestión del tiempo y si desencadenaba en un bajo rendimiento académico. Se efectuó un muestreo aleatorio sistemático para elegir 32 participantes (es decir, ocho participantes para cada uno de los cuatro grupos focales de la población total). Los resultados revelaron que solo el 30 % de los estudiantes estaban afectados por la adicción a internet y menos del 15 % de ellos enfrentaban dificultades en la gestión del tiempo. El coeficiente de correlación mostró una relación moderada (0,42) entre la adicción a internet y la gestión del tiempo. El coeficiente de correlación describe que cuanto mayor sea el nivel de adicción a internet, mayor será la dificultad que enfrentará un estudiante en la gestión del tiempo. El rendimiento del coeficiente de correlación puede mejorarse aumentando el tamaño de la muestra. El estudio también reveló que la gestión del tiempo y el rendimiento académico son directamente proporcionales entre sí. Un análisis de varianza en las respuestas de los estudiantes masculinos y femeninos reveló que el género no tiene impacto en los resultados generales del estudio.

Bajpai y Maneesha (2018) se centraron en hallar el impacto que tienen las redes sociales en el rendimiento académico de un grupo de estudiantes universitarios graduados. Como metodología de recopilación de datos, llevaron a cabo una encuesta a los estudiantes de ingeniería de una universidad de Dubái y usaron el *software* SPSS para el posterior análisis de la información. Se envió el cuestionario por correo electrónico a los estudiantes,

luego se recibieron las respuestas de 191 estudiantes de programas de ingeniería informática, mecánica, electrónica e instrumentación, electrónica y comunicación, civil, química y biotecnología. Del total de estudiantes, 46 respondieron que usan las redes sociales para socializarse, seis respondieron que le dan uso para fines académicos, 122 que las usan para socializarse y fines académicos y siete no aplicaron. En cuanto a la pregunta: si las redes sociales tienen un efecto negativo o positivo sobre el rendimiento académico, 129 respondieron que afectan de manera positiva y 52 que afectan de forma negativa. En cuanto a la pregunta: si las redes sociales afectan el ejercicio físico, 81 estudiantes respondieron que sí afectan y 100 que no afectan. En cuanto a las preferencias sobre la comunicación con los demás, 155 estudiantes prefirieron una comunicación cara a cara y 26 a través de redes sociales. El 99% están activos en alguna plataforma de redes sociales. Las tres plataformas de redes sociales más visitadas son WhatsApp (59,6%), Instagram (13,8%) y YouTube (9,3%). Cerca del 70% de los estudiantes utilizan WhatsApp y Facebook para establecer contactos con exalumnos. LinkedIn (64%) se utiliza más eficazmente para encontrar las perspectivas de empleo. Realizaron un conjunto de histogramas que les permitieron conocer la frecuencia con la que se repiten algunos resultados de las variables estudiadas. Luego hicieron la prueba ANOVA para llevar a cabo una comparación de medias, pues el promedio de calificaciones es la variable dependiente y la constante el tiempo utilizado. Al ser el p -value menor a 0,05, los autores encontraron que el modelo de regresión predice de manera efectiva el tiempo que gastan los estudiantes en redes sociales. Los autores encontraron que a medida que aumenta el tiempo en redes sociales, el promedio de calificaciones disminuye.

Nabizadeh *et al.* (2019) emprendieron un estudio con la participación de 380 estudiantes varones y mujeres de nueve facultades de ciencias médicas de la Universidad Shahid Beheshti (Teherán). A cada estudiante se le proporcionó un cuestionario de estrategias motivantes para que el estudio y las expectativas den resultados. Se tuvieron en cuenta las siguientes variables: género, calificaciones, campo de estudio, estado civil, residencia, segunda vocación. La comparación del GPA (*grade point average*, o promedio de calificaciones) de estudiantes, no mostró ninguna diferencia significativa

en términos de las variables de género, campo de estudio, nivel educativo, estado matrimonial, estado de empleo y lugar de residencia. El cálculo de los coeficientes de correlación de Pearson entre las variables de estudio y el GPA de los estudiantes ($p < 0,05$) reveló una correlación lineal directa y positiva entre las estrategias de motivación y las estrategias de aprendizaje. Las personas con alta autoeficacia tienden a estar más inclinadas a visualizar resultados positivos sobre sus tareas.

Vijayalakshmi *et al.* (2018) utilizaron el enfoque de aprendizaje basado en proyectos (PBL), en el cual se emplearon diferentes estrategias de formación de grupos basados en la agrupación natural, calificaciones mixtas, el género y las habilidades complementarias. En las actividades complementarias, se encuentran las habilidades profesionales y técnicas. Como habilidades profesionales, se destacan: capacidad de explorar e innovar, capacidad para trabajar en equipo, capacidad de liderazgo. Como habilidades técnicas, se encuentran: habilidades de programación, habilidades de prueba, habilidades de comunicación. Como parte del proceso experimental, se les pidió a los estudiantes que se autocalificaran en dichas habilidades. Luego, se les pidió que formaran equipos en los cuales todos sus miembros tuvieran habilidades diferentes y escogieran a un líder. Cada uno de estos equipos se formó para desarrollar y ejecutar miniproyectos y proyectos de ingeniería. Una vez se finalizaron los proyectos, se llevó a cabo la evaluación por pares de los estudiantes tenidos en cuenta y, posteriormente, se le dio la calificación final a cada uno. Luego de hacer una limpieza de datos, con el fin de eliminar aquellos atípicos, los autores obtuvieron alrededor de 30 equipos que se podrían formar según las pautas dadas inicialmente. Se llevó a cabo la prueba T para el conjunto de datos, siguiendo dos estrategias de formación de equipos: basado en el promedio de calificaciones y basado en la agrupación natural. Dicha prueba se realizó para analizar el efecto de las estrategias de formación de equipos en el rendimiento de los estudiantes. A partir de la comparación de medias, se pudo observar el aumento en los puntajes de los mismos estudiantes en proyectos menores, comparados con aquellos proyectos en los que no se tenían estrategias de agrupación. Se efectuó una prueba T para los datos de proyectos menores a dos años (2013-2014 y 2014-2015), para analizar el efecto que tienen las estrategias

de formación de equipos en el rendimiento de los estudiantes. Los autores asumen un valor de significancia de 0,5; como resultado, la prueba arroja un p -value de 0,0055. Por tanto, concluyen que la formación de un equipo impacta en el aprendizaje de los estudiantes, en el contexto del aprendizaje basado en proyectos.

10.3.3. Medicina

Thomas *et al.* (2020) plantean la evaluación entre los niveles de biomarcadores proinflamatorios registrados en los estudiantes universitarios del primer año y el rendimiento académico, partiendo de una muestra de 180 estudiantes de la Facultad de Ciencias de la Salud de la Universidad de Johannesburgo (Sudáfrica). El indicador de medición del rendimiento académico, corresponde a la media general estandarizada de calificaciones en un núcleo común de clases requeridas para la graduación y el desempeño en un módulo específico. La variable independiente corresponde a los biomarcadores que indican los niveles de inflamación de los estudiantes, así como algunos datos demográficos, como edad, sexo y etnia. Para la medición de la actividad inflamatoria, se evaluó la relación en proporción de neopterina-creatinina (μmol neopterina/ mol creatinina) mediante una sola recolección de muestra de orina. La actividad inflamatoria también se midió a través de la interleucina-6 plasmática, neopterina plasmática y neopterina plasmática-creatinina, en un subgrupo pequeño de 37 estudiantes. Ambas mediciones se hicieron mediante un ensayo inmunoabsorbente ligado a enzimas (ELISA) en fase sólida. Por otro lado, el rendimiento académico se representó a través de la media general estandarizada de calificaciones en un núcleo común de clases requeridas para la graduación, así como la de un módulo específico (anatomía y fisiología). Para el análisis de los datos, se utilizaron los análisis de varianza unidireccional (ANOVA) y los análisis de regresión mediante el STATA 14. El ANOVA de una vía, indicó que ninguno de los biomarcadores inflamatorios utilizados en este estudio difirió en función de la etnia (negro, blanco, indio). Similarmente, ninguno de los biomarcadores inflamatorios mostró una diferencia de ANOVA en función del sexo. Por otro lado, el análisis de regresión indicó que la edad no fue un predictor significativo de ninguno de los biomarcadores inflamatorios. En

cuanto al rendimiento académico, ninguno de los biomarcadores inflamatorios medidos se correlacionó significativamente con ninguna medida de este.

Thomas y Viljoen (2019) buscan determinar si la variabilidad de la frecuencia cardíaca (VFC) de una persona en reposo o en respuesta a un desafío cognitivo y ortostático, se relaciona de forma significativa con el rendimiento académico de los estudiantes. La muestra correspondió a 76 estudiantes de una universidad de Sudáfrica. El indicador de medición del rendimiento académico para el presente estudio, correspondió al promedio de todas las calificaciones obtenidas durante un año, en un núcleo común de las clases que se requerían para la graduación, mientras que los cursos reprobados y abandonados tuvieron una puntuación de 50. Los puntajes del rendimiento académico se clasificaron en dos grupos: bajo y alto. La variable independiente corresponde a la variabilidad de la frecuencia cardíaca, que se midió con dos indicadores: potencia total (PT) como medida de la VFC de forma general y la raíz cuadrada de las diferencias sucesivas (RMSSD) para realizar mediciones de la influencia parasimpática en el corazón. Para la recolección de los datos, se les equipó de un monitor de frecuencia cardíaca. A partir de los análisis preliminares, se determinó que no había interacciones significativas entre la etnia o la edad con los indicadores de la VFC (pues los valores de p encontrados son mayores a 0,1). En cuanto al género, este sí influyó en los valores de potencia total, es decir, que para las mujeres este marcador del VFC es más alto que para los hombres.

El indicador de PT en posición supina (global) y PT en posición supina para mujeres, son las únicas dos medidas que demostraron diferencia significativa en términos estadísticos, entre los grupos de alto y bajo rendimiento. Es decir, cuando el VFC es mayor en supino, medido a través de la PT, el rendimiento académico es significativamente mayor (especialmente en las mujeres).

10.4. Conclusiones del capítulo

De acuerdo con el trabajo presentado y con los resultados obtenidos, se pueden plantear las siguientes conclusiones principales:

Los investigadores se han preocupado en los últimos años en trabajar en el desarrollo de modelos que permitan comprender aspectos de la vida académica del estudiante, docentes e instituciones que permitan la elaboración y toma de decisiones acertadas.

La analítica es un campo en crecimiento usado en el ámbito educativo que tiene por objetivo, estudiar cómo influyen los factores tecnológicos y de gestión en las instituciones; por ejemplo, aspectos relacionados con la organización y su área financiera (Oblinger *et al.*, 2007). Busca la aplicación de los principios y herramientas de la minería de datos y aprendizaje automático a la academia.

El rendimiento académico, la tasa de deserción y la tasa de graduación en su respectivo orden, constituyen factores imprescindibles en el abordaje del tema de la calidad de la educación superior; el primero, el más influyente, según las investigaciones analizadas.

Las herramientas tecnológicas avanzadas en las ciencias de la computación, como lo es el aprendizaje automático y sus algoritmos, están apoyando y fortaleciendo la toma de decisiones desde el punto de vista operacional, académico y financiero.

De acuerdo con lo encontrado en el análisis de la investigación referencial, es posible utilizar los algoritmos de aprendizaje automático en un enfoque para estimar la cantidad de estudiantes graduados por cohorte y los mejores campos de acción de un grupo de estudiantes al terminar la carrera universitaria, así como a los indicadores anteriormente mencionados. Pero tal como lo menciona Daniel (2015a), la tecnología para entregar este potencial es aún joven y la investigación en la comprensión de la utilidad pedagógica está todavía en su infancia.

La analítica del aprendizaje es un campo que apenas está incursionando en la educación superior, que busca aprovechar el alto volumen de información generado por los estudiantes en sus actividades académicas, a los cuales poco uso se les da más allá de realizar estadísticas descriptivas y, en el mejor de los casos, una analítica diagnóstica (desempeño pasado para analizar la información).

Los resultados muestran que, en la educación superior, la analítica del aprendizaje se está centrando en la solución o desempeño de los estudiantes que estudian mediante la modalidad virtual. Este tipo de técnica puede centrarse también en la aplicación en otro tipo de modalidad, pues lo que se requiere es la información de muchos dispositivos, plataformas y bases de datos que las instituciones tienen almacenados o capturan en el quehacer académico.

El estudio del proceso de aprendizaje, según la revisión, no solo abarca los datos numéricos que se obtienen de las plataformas, analizados por medio de herramientas, con el fin de ofrecer información relevante y útil para facilitar la toma de decisiones y acciones oportunas. Sino que es necesario tener en cuenta diferentes enfoques desde otras áreas, como la psicología y la pedagogía. Por ende, no existe un camino definido en el estudio del proceso de aprendizaje de los estudiantes. Es un camino que conduce a valores positivos y a negativos (axiología).

Las metodologías trabajadas en el área de la psicología, consideraron los cuestionarios aplicados por los estudiantes (o docentes, en algunos casos), en donde demostraron su percepción sobre un tema en específico (generalmente en escalas Likert, que miden actitudes y grados de conformidad por medio de categorías de respuesta, o mediante cuestionarios psicológicos que evalúan, entre otras cosas, niveles de ansiedad, habilidades sociales, autoestima, entre otros).

Las metodologías investigativas usadas en el campo estadístico-matemático, emplearon datos específicos de la actividad académica de los estudiantes, así como sus relaciones con los datos demográficos, sin el uso de cuestionarios cualitativos.

Finalmente, en la metodología centrada en el ámbito de la medicina, utiliza procedimientos específicos del área a través de diversas mediciones para el diagnóstico de procesos fisiológicos, que pueden incidir en el rendimiento académico.

Referencias

- Abdullah, A., Asar, A. K. y Alnisa'Anis Alanna Binti Ruzelan, N. (2020). A novel learning engagement data model (LEDLM) for online attendance system. *IOP Conference Series: Materials Science and Engineering*, 769(1), 012026. <https://doi.org/10.1088/1757-899X/769/1/012026>
- Adekitan, A. I. y Noma-Osaghae, E. (2019). Data mining approach to predicting the performance of first year student in a university using the admission requirements. *Education and Information Technologies*, 24(2), 1527-1543. <https://doi.org/10.1007/s10639-018-9839-7>
- Adell, J. (1994). World Wide Web: un sistema hipermedia distribuido para la docencia universitaria. *Nuevas Tecnologías de la Información y la Comunicación para la Educación*. Ediciones Alfar, 114-121. <http://nti.uji.es/docs/nti/badajoz.html>
- Agudo-Peregrina, Á. F., Iglesias-Pradas, S., Conde-González, M. Á. y Hernández-García, Á. (2014). Can we predict success from log data in VLEs? Classification of interactions for learning analytics and their relation with performance in VLE-supported F2F and online learning. *Computers in Human Behavior*, 31, 542-550. <https://doi.org/10.1016/j.chb.2013.05.031>
- Ahamed, A. T. M. S., Mahmood, N. T. y Rahman, R. M. (2016). Prediction of HSC examination performance using socioeconomic, psychological and academic factors. *International Conference on Electrical and Computer Engineering*, 263-266.
- Ahmad, F., Ismail, N. H. y Aziz, A. A. (2015). The prediction of students' academic performance using classification data mining techniques. *Applied Mathematical Sciences*, 9(129), 6415-6426. <https://doi.org/10.12988/ams.2015.53289>

- Ajibade, S. M., Ahmad, N. B. y Shamsuddin, S. M. (2019). An heuristic feature selection algorithm to evaluate academic performance of students. *10th Control and System Graduate Research Colloquium (ICSGRC)*, 2-3.
- Ajibade, S. S. M., Bahiah Binti Ahmad, N. y Mariyam Shamsuddin, S. (2019). Educational data mining: Enhancement of student performance model using ensemble methods. *IOP Conference Series: Materials Science and Engineering*, 551(1). <https://doi.org/10.1088/1757-899X/551/1/012061>
- Akçapınar, G., Altun, A. y Aşkar, P. (2019). Using learning analytics to develop early-warning system for at-risk students. *International Journal of Educational Technology in Higher Education*, 16(1). <https://doi.org/10.1186/s41239-019-0172-z>
- Al-araibi, A. A. M., Mahrin, M. N. bin y Yusoff, R. C. M. (2019). Technological aspect factors of *E-Learning* readiness in higher education institutions: Delphi technique. *Education and Information Technologies*, 24(1), 567-590. <https://doi.org/10.1007/s10639-018-9780-9>
- Al-araibi, A. A. M., Naz'ri bin Mahrin, M., Yusoff, R. C. M. y Chuprat, S. B. (2019). A model for technological aspect of *E-Learning* readiness in higher education. *Education and Information Technologies*, 24(2), 1395-1431. <https://doi.org/10.1007/s10639-018-9837-9>
- Al-Barrak, Mashael A. y Al-Razgan, Muna. (2016). Predicting students final GPA using decision trees: A case study. *International Journal of Information and Education Technology*, 6-7, 528-533. <https://doi.org/10.7763/IJiet.2016.V6.745>
- Albela-García, G. P., Luna-Yáñez, C., Orezzaoli, H. M., Guevara-Canales, J.-O. y García, H. (2020). Relación del rendimiento académico con técnicas de aprendizaje en estudiantes de odontología de la asignatura de radiología. *Kiru*, 17(4), 2410-2717. <https://doi.org/10.24265/kiru.2020.v17n4.01>
- Almaiah, M. A., Al-Khasawneh, A. y Althunibat, A. (2020). Exploring the critical challenges and factors influencing the *E-Learning* system usage during COVID-19 pandemic. *Education and Information Technologies*, 25, 5261-5280. <https://doi.org/10.1007/s10639-020-10219-y>
- Almarabeh, H. (2017). Analysis of students' performance by using different data mining classifiers. *International Journal of Modern Education and Computer Science*, 8, 9-15. <https://doi.org/10.5815/ijmecs.2017.08.02>

- Aluko, R. O., Daniel, E. I., Oshodi, O. S., Aigbavboa, C. O. y Abisuga, A. O. (2018). Towards reliable prediction of academic performance of architecture students: Using data mining techniques. *Journal of Engineering, Design and Technology*, 16(3), 385-397. https://wlv.openrepository.com/bitstream/handle/2436/623424/Daniel_Towards_Reliable_Prediction_2020.pdf
- Alves, P., Morais, C. y Miranda, L. (2017). Learning analytics in higher education: Assessing learning outcomes. *Conferencia Europea sobre E-Learning*, 25-32.
- Alyahyan, E. y Düşteğör, D. (2020). Predicting academic success in higher education: Literature review and best practices. *International Journal of Educational Technology in Higher Education*, 17(3). <https://doi.org/10.1186/s41239-020-0177-7>
- Amat, J. (2016). Test exacto de Fisher, chi-cuadrado de Pearson, McNemar y Q-Cochran. En *Test estadísticos para variables cualitativas*. https://rpubs.com/Joaquin_AR/220579
- Amazona, M. V. y Hernández, A. A. (2019). Modelling student performance using data mining techniques. *Proceedings of the 2019. 5th International Conference on Computing and Data Engineering (ICCDE)*, 19, 36-40. <https://doi.org/10.1145/3330530.3330544>
- Anuradha, C. y Velmurugan, T. (2015). A comparative analysis on the evaluation of classification algorithms in the prediction of students performance. *Indian Journal of Science and Technology*, 8(15).
- Argonza, J. (2016). *Big Data* en la educación. *Revista Digital Universitaria*, 17(1), 1-16. <http://www.revista.unam.mx/vol.17/num1/art06/>
- Asif, R., Merceron, A., Ali, S. A. y Haider, N. G. (2017). Analyzing undergraduate students' performance using educational data mining. *Computers and Education*, 113, 177-194. <https://doi.org/10.1016/j.compedu.2017.05.007>
- Avella, J. T., Kebritchi, M., Nunn, S. G. y Kanai, T. (2016). Learning analytics methods, benefits, and challenges in higher education: A systematic literature review. *Journal of Asynchronous Learning Network*, 20(2). <https://doi.org/10.24059/olj.v20i2.790>
- Awadalla, S., Davies, E. B. y Glazebrook, C. (2020). A longitudinal cohort study to explore the relationship between depression, anxiety and academic performance among Emirati university students. *BMC*

- Psychiatry*, 20(448). <https://bmcpsy psychiatry.biomedcentral.com/track/pdf/10.1186/s12888-020-02854-z.pdf>
- Ayesha, Sh., Mustafá, T., Sattar, A. R. y Khan, M. I. (2010). Data mining model for higher education system. *European Journal of Scientific Research*, 43, 24-32. <http://connection.ebscohost.com/c/articles/52621007/data-mining-model-higher-education-system>
- Azcona, D., Hsiao, I. H. y Smeaton, A. F. (2019). Detecting students-at-risk in computer programming classes with learning analytics from students' digital footprints. *User Modeling and User-Adapted Interaction*, 29(4), 759-788. <https://doi.org/10.1007/s11257-019-09234-7>
- Azzi, I., Jeghal, A., Radouane, A., Yahyaouy, A. y Tairi, H. (2020). A robust classification to predict learning styles in adaptive *E-Learning* systems. *Education and Information Technologies*, 25(1), 437-448. <https://doi.org/10.1007/s10639-019-09956-6>
- Back, D. A., Behringer, F., Harms, T., Plener, J., Sostmann, K. y Peters, H. (2015). Survey of *E-Learning* implementation and faculty support strategies in a cluster of mid-European medical schools. *BMC Medical Education*, 15(145), 1-9. <https://doi.org/10.1186/s12909-015-0420-4>
- Baepler, P. y Murdoch, C. J. (2010). Academic analytics and data mining in higher education. *International Journal for the Scholarship of Teaching & Learning*, 4(2). <https://doi.org/10.20429/ijstol.2010.040217>
- Bagnato, J. (s.f.). *Aprende machine learning*. <https://www.aprendemachine-learning.com>
- Bahammam, A. S., Alaseem, A. M., Alzakri, A. A., Almeneessier, A. S. y Sharif, M. M. (2012). The relationship between sleep and wake habits and academic performance in medical students: A cross-sectional study. *BMC Medical Education*, 12(61). <https://doi.org/10.1186/1472-6920-12-61>
- Bajpai, P. y Maneesha. (2018). Analyzing effect of social media on academic performance of university graduates. *ACM International Conference Proceeding Series*, 40-44. <https://doi.org/10.1145/3234825.3234830>
- Banihashem, S. K., Aliabadi, K., Pourroostaei Ardakani, S., Delaver, A. y Nili Ahmadabadi, M. (2018). Learning analytics: A systematic literature review. *Interdisciplinary Journal of Virtual Learning in Medical Sciences*, 9(2), 63024. <https://doi.org/10.5812/ijvlms.63024>

- Bautista-Rodríguez, G. y Gatica-Lara, F. (2020). Factores relacionados con el rendimiento académico en una carrera técnica en salud impartida en línea. *Investigación en Educación Médica*, 9(33), 89-97. <http://www.scielo.org.mx/pdf/iem/v9n33/2007-5057-iem-9-33-89.pdf>
- Bayona, E. (2017). Hábitos de estudio y rendimiento académico en ingeniería de sistemas. *Revista Colombiana de Computación*, 18(1), 61-69. <https://doi.org/10.29375/25392115.3198>
- Benavides, E. S. y Pérez Montero, E. L. (2015). Estructuras de decisión a partir del aprendizaje autorregulado en ambientes *B-Learning*. *Revista Tecnura*, 15-24. <https://doi.org/10.14483/udistrital.jour.tecnura.2015.SE1.a01>
- Bhutto, S., Siddiqui, I. F., Arain, Q. A. y Anwar, M. (2020). Predicting students' academic performance through supervised (ICISCT). *2nd International Conference on Information Science and Communication Technology*. <https://doi.org/10.1109/ICISCT49550.2020.9080033>
- Boneu, J. (2007). Plataformas abiertas de *E-Learning* para el soporte de contenidos educativos abiertos. *Revista de Universidad y Sociedad del Conocimiento*, 4(1), 36-47. https://www.redalyc.org/pdf/780/Resumenes/Abstract_78040109_2.pdf
- Bonsaksen, T. (2016). Predictors of academic performance and education programme satisfaction in occupational therapy students. *British Journal of Occupational Therapy*, 79(6). <https://doi.org/10.1177/0308022615627174>
- Bote-Lorenzo, M. L., Gómez-Sánchez, E. y Asensio-Pérez, J. I. (2019). Aligning learning design and learning analytics through instructor involvement: A MOOC case study. *Interactive Learning Environments*, 1, 1-15.
- Brownlee, J. (2020). *Mastery*. <https://machinelearningmastery.com/>
- Buhagiar, T. y Leo, C. (2018). Does gamification improve academic performance? *Journal of Instructional Pedagogies*, 20, 1-6. <https://files.eric.ed.gov/fulltext/EJ1178731.pdf>
- Burgos, C., Campanario, M. L., Peña, D. de la, Lara, J. A., Lizcano, D. y Martínez, M. A. (2018). Data mining for modeling students' performance: A tutoring action plan to prevent academic dropout. *Computers and Electrical Engineering*, 66, 541-556. <https://doi.org/10.1016/j.compeleceng.2017.03.005>

- Burman, I. y Som, S. (2019). Predicting students academic performance using support vector machine. *ProceedingsK. Amity International Conference on Artificial Intelligence (AICAI) 2019*, 756-759. <https://doi.org/10.1109/AICAI.2019.8701260>
- Campbell, J. P. y Oblinger, D. (2007). Academic analytics: A new tool for a new era. *EDUCAUSE Review*, 42(4), 40-57. <https://er.educause.edu/articles/2007/7/academic-analytics-a-new-tool-for-a-new-era>
- Campo-Arias, A., Oviedo, H. C. y Herazo, E. (2014). Escala de estrés percibido-10: desempeño psicométrico en estudiantes de medicina de Bucaramanga, Colombia. *Rev. Fac. Med.*, 62(3), 407-413. <https://doi.org/10.15446/revfacmed.v62n3.43735>
- Castrillón, O. D., Sarache, W. y Ruiz-Herrera, S. (2020). Prediction of academic performance using artificial intelligence techniques. *Formacion Universitaria*, 13(1), 93-102. <https://doi.org/10.4067/S0718-50062020000100093>
- Castro, I. de y Closas, A. (2020). Modelo estructural de las relaciones entre enseñanza en contextos tecnológicos y el rendimiento académico en contabilidad. *Cuaderno de Pedagogía*, 17(34), 31-47. <https://cuaderno.pucmm.edu.do/index.php/cuadernodepedagogia/article/view/394/409>
- Catalán, H. (s.f.). *Curso internacional: construcción de escenarios económicos y econometría avanzada*. Universidad Nacional de Colombia. https://www.cepal.org/sites/default/files/courses/files/04_pruebas_especificacion.pdf
- Céspedes, R. C., Vara-Horna, A., López-Odar, D., Díaz-Rosillo, A. y Asencios-González, Z. (2018). Ausentismo, presentismo y rendimiento académico en estudiantes de universidades peruanas. *Propósitos y Representaciones*, 6(1), 83-133. <https://doi.org/10.20511/pyr2018.v6n1.177>
- Chen, H. J. (2020). Clarifying the impact of surprise in *E-Learning* system design based on university students with multiple learning goals orientation. *Education and Information Technologies*, 25, 5873-5892. <https://doi.org/10.1007/s10639-020-10249-6>
- Chen, L., Yoshimatsu, N., Goda, Y., Okubo, F., Taniguchi, Y., Oi, M., Konomi, S., Shimada, A., Ogata, H. y Yamada, M. (2019). Direction of collaborative problem solving-based STEM learning by learning

- analytics approach. *Research and Practice in Technology Enhanced Learning*, 14(24). <https://doi.org/10.1186/s41039-019-0119-y>
- Chilca, M. (2017). Autoestima, hábitos de estudio y rendimiento académico en estudiantes universitarios. *Propósitos y Representaciones*, 5(1), 71-127. <https://dialnet.unirioja.es/servlet/articulo?codigo=5904759>
- Clow, D. (2013). An overview of learning analytics. *Teaching in Higher Education*, 18(6), 683-695. <https://doi.org/10.1080/13562517.2013.827653>
- Coaten, N. (2003). *Educaweb*. Suplemento del boletín de educaweb | ISSN: 1578-5793. <https://www.educaweb.com/esp/servicios/monografico/formacionvirtual/1181076-a.html>
- Contreras, L. E., Fuentes, H. J. y Rodríguez, I. (2020). Academic interruption model using automatic learning algorithms. *International Journal of Mechanical and Production Engineering Research and Development (IJMPERD)*, 10(3), 16075-16086. www.tjprc.org
- Contreras, L., González, K. y Fuentes, H. (2011). Use of ICT and especially of blended learning in higher education. *Revista Educación y Desarrollo Social*, 5(1), 151-160. <https://doi.org/10.18359/REDS.898>
- Contreras, L., Fuentes, H. y Molano, J. (2021). Analítica académica: nuevas herramientas aplicadas a la educación. *Revista Boletín Redipe*, 10(3), 137-158.
- Contreras, L., Fuentes, H. y Rodríguez, J. (2020). Academic performance prediction by as a success/failure indicator for engineering students. *Formación Universitaria*, 13(5), 233-246. <https://doi.org/10.4067/S0718-50062020000500233>
- Costa, E. B., Fonseca, B., Santana, M. A., de Araújo, F. F. y Rego, J. (2017). Evaluating the effectiveness of educational data mining techniques for early prediction of students' academic failure in introductory programming courses. *Computers in Human Behavior*, 73, 247-256. <https://doi.org/10.1016/j.chb.2017.01.047>
- Cotter, E. W. y Kelly, N. R. (2016). Binge eating scale (BES). In *Encyclopedia of Feeding and Eating Disorders* (pp. 1-5). Springer. https://doi.org/10.1007/978-981-287-087-2_9-2
- Cresswell, J. y Underwood, C. (2004). *Location, location, location: Implications of geographic situation on Australian student performance in PISA 2000*. Australian Council for Educational Research.

- Daniel, B. (2015a). *Big Data* and analytics in higher education: Opportunities and challenges. *British Journal of Educational Technology*, 46-5(0007-1013), 904-920. <https://www.learntechlib.org/p/151804/>
- Daniel, B. (2015b). *Big Data* and analytics in higher education: Opportunities and challenges. *British Journal of Educational Technology*, 46(5), 904-920. <https://doi.org/10.1111/bjet.12230>
- David, J. y Anastasija, G. (2019). Predicting academic performance based on students' family environment: Evidence for Colombia using classification trees. *Psychology, Society and Education*, 11(3), 299-311. <https://www.uco.es/ucopress/ojs/index.php/psye/article/view/13947>
- De Arco Paternina, L. K., Barrios Barceló, H. de J. y Parra Ortega, G. M. (2017). Análisis de factibilidad de un sistema de educación *B-Learning*: caso de estudio. Universidad de San Buenaventura. *Revista Academia y Virtualidad*, 10(2), 7-19. <https://doi.org/10.18359/ravi.2706>
- Deliens, T., Clarys, P., Bourdeaudhuij, I. de y Deforche, B. (2013). Weight, socio-demographics, and health behaviour related correlates of academic performance in first year university students. *Nutrition Journal*, 12, 162. <https://doi.org/10.1186/1475-2891-12-162>
- Depaula, P. D. y Azzollini, S. C. (2013). Analysis of the Big five model of personality as a predictor of cultural intelligence. *PSIENCIA Revista Latinoamericana de Ciencia Psicológica*, 5(1), 35-43. <https://doi.org/10.5872/psiencia/5.1.24>
- Devasia, T., Vinushree, T. P. y Hegde, V. (2016). Prediction of students performance using educational data mining. *International Conference on Data Mining and Advanced Computing (SAPIENCE)*, 91-95.
- Díaz-Pinzón, J. E. (2020). Estudio comparativo entre el contagio durante la cuarentena obligada por el COVID-19 y el contagio durante la apertura gradual y controlada para algunos sectores de la economía en Colombia. *Repertorio de Medicina y Cirugía*, 29(1), 52-58. <https://revistas.fucsulud.edu.co/index.php/repertorio/article/view/1073/1245>
- Díaz, L. y Ríos, F. (2018). El valor *p*. Interpretación, orígenes y su utilización actual. *Revista Argentina de Terapia Intensiva*, 35(3), 55-59.
- Dixson, D. D., Worrell, F. C., Olszewski-Kubilius, P. y Subotnik, R. F. (2016). Beyond perceived ability: The contribution of psychosocial factors to academic performance. *Annals of the New York Academy of Sciences*, 1377(1), 67-77. <https://doi.org/10.1111/nyas.13210>

- Domínguez-Lara, S., Merino-Soto, C. y Guevara-Cordero, C. (2018). *Big five inventory* en universitarios peruanos: resultados preliminares de su validación. *PSYKHE*, 27(2), 1-12. <https://doi.org/10.7764/psykhe.27.2.1052>
- Drachsler, H. y Kalz, M. (2016). The MOOC and learning analytics innovation cycle (MOLAC): A reflective summary of ongoing research and its challenges. *Journal of Computer Assisted Learning*, 32, 281-290. <https://onlinelibrary.wiley.com/doi/abs/10.1111/jcal.12135>
- Dulzaides, M. y Molina, A. (2004). Análisis documental y de información: dos componentes de un mismo proceso. *ACIMED*, 12(2).
- Durán, C. y Rosado, A. (2019). La comprensión lectora y el rendimiento académico en estudiantes de ingeniería. *Revista Colombiana de Tecnologías de Avanzada (RCTA)*, 1(33), 9-15. <https://doi.org/10.24054/16927257.v33.n33.2019.3317>
- Dyckhoff, A. L., Zielke, D., Bültmann, M., Chatti, M. A. y Schroeder, U. (2012). Design and implementation of a learning analytics toolkit for teachers. *Educational Technology and Society*, 15(3), 58-76.
- Echegaray-Calderón, O. A. y Barrios-Aranibar, D. (2015). Optimal selection of factors using genetic algorithms and neural networks for the prediction of students' academic performance. *Congreso Latinoamericano de Inteligencia Computacional*, 1-6.
- Elías, T. (2011). *Learning analytics: Definitions, processes and potential*. <http://learninganalytics.net/LearningAnalyticsDefinitionsProcessesPotential.pdf>
- Ellis, R. A., Han, F. y Pardo, A. (2017). Improving learning analytics - Combining observational and self-report data on student learning. *Educational Technology and Society*, 20(3), 158-169.
- Espejo, L. M. D. (s.f.). *Diagrama de caja y bigotes*. Liceo Marta Donoso Espejo, Talca, Chile. http://inst-mat.ugal.cl/tem/sitiolmde/primeroguias-liceo/recuperacion/Diagrama_de_Caja_y_Bigotes-2.pdf
- Espinosa, J., Hernández, J., Rodríguez, J., Chacín, M. y Bermúdez, V. (2020). Influencia del estrés sobre el rendimiento académico. *AVFT-Archivos Venezolanos de Farmacología y Terapéutica*, 39(1). <https://bonga.unisimon.edu.co/bitstream/handle/20.500.12442/6322/PDF.pdf;jsessionid=BCAE649A5ED0968F81C3D9B047C8039E?sequence=1>

- Estrada, J. y Quintero, R. (2015). *Bajo rendimiento académico en la Universidad Distrital Francisco José de Caldas*. 1ª ed., Editorial Universidad Distrital.
- Eze, S. C., Chinedu-Eze, V. C. y Bello, A. O. (2018). The utilisation of *E-Learning* facilities in the educational delivery system of Nigeria: A study of M-University. *International Journal of Educational Technology in Higher Education*, 15(34). <https://doi.org/10.1186/s41239-018-0116-z>
- F-Jardon, C. M. y Martos, M. S. (2011). Un método para determinar competencias distintivas en pequeñas y medianas empresas. *Revista de Administração da Universidade*, 4(2), 195-214. <https://www.redalyc.org/pdf/2734/273419420001.pdf>
- Farrokhi-Khajeh-Pasha, Y., Nedjat, S., Mohammadi, A., Rad, E. M., Majdzadeh, R., Monajemi, F., Jamali, E. y Yazdani, S. (2012). The validity of Iran's national university entrance examination (Konkoor) for predicting medical students academic performance. *BMC Medical Education*, 12(60). <https://doi.org/10.1186/1472-6920-12-60>
- Ferguson, R. (2012). Learning analytics: Drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, 4(x), 304-317. <https://doi.org/10.1504/IJTEL.2012.051816>
- Fernández-Lasarte, O., Ramos-Díaz, E. y Sáez, I. A. (2019). Rendimiento académico, apoyo social percibido e inteligencia emocional en la universidad. *European Journal of Investigation in Health, Psychology and Education*, 9(1), 39. <https://www.sciencegate.app/document/10.30552/ejihpe.v9i1.315>
- Fournier, H., Kop, R. y Sitlia, H. (2011). The value of learning analytics to networked learning on a personal learning environment. *LAK'11 Proceedings of the 1st International Conference on Learning Analytics and Knowledge*, 104-109. <https://doi.org/10.1145/2090116.2090131>
- Francis, B. K. y Babu, S. S. (2019). Predicting academic performance of students using a hybrid data mining approach. *Journal of Medical Systems*, 43(6). <https://doi.org/10.1007/s10916-019-1295-4>
- Furnham, A., Nuygards, S. y Chamorro-Premuzic, T. (2013). Personality, assessment methods and academic performance. *Instructional Science*, 41(5), 975-987. <https://www.jstor.org/stable/43575408?seq=1>
- García-Hernández, A. y González-Ramírez, T. (2020). El rendimiento académico en matemáticas discretas: un estudio predictivo. *Atenas Revista Científico Pedagógica*, 49(1), 118-134. <https://cache.1science.com/3b/18/3b18065b3a38adb07839c6bf07bd00c87e01147b.pdf>

- García, A. y Media, H. (2015). *Evaluación docente institución universitaria colegios de Colombia - UNICOC: análisis factorial para la consolidación de indicadores confiables internamente bajo la percepción de estudiantes, docentes y jefe inmediato*. Fundación Universitaria Los Libertadores. <https://repository.libertadores.edu.co/bitstream/handle/11371/246/AstridLilianaGarciaCordoba.pdf?sequence=2&isAllowed=y>
- García Tinisaray, D. K. (2015). *Construcción de un modelo para determinar el rendimiento académico de los estudiantes basado en learning analytics (análisis del aprendizaje), mediante el uso de técnicas multivariantes*. Universidad de Sevilla. <https://idus.us.es/bitstream/handle/11441/40436/Tesis%20realizada%20por%20Daysi%20K.%20Garc%3%ADa.pdf?sequence=1&isAllowed=y>
- Garg, R. (2018). Predicting student performance of different regions of punjab using classification techniques. *International Journal of Advanced Research in Computer Science*, 9(1), 236-240.
- Garmendia, M. L. (2007). Análisis factorial: una aplicación en el cuestionario de salud general de Goldberg, versión de 12 preguntas. *Revista Chilena de Salud Pública*, 11(2), 57-65. <https://doi.org/10.5354/0717-3652.2007.3095>
- Garzón, R., Rojas, M. O., Riesgo, L. del, Pinzón, M. y Salamanca, A. L. (2010). Factores que pueden influir en el rendimiento académico de estudiantes de bioquímica que ingresan en el programa de medicina de la Universidad del Rosario (Colombia). *Educación Médica*, 13(2), 85-96. <https://scielo.isciii.es/pdf/edu/v13n2/original1.pdf>
- Giannakos, M. Nm., Chorianopoulos, K. y Chrisochoides, N. (2015). Making sense of video analytics: Lessons learned from clickstream interactions, attitudes, and learning outcome in a video-assisted course. *International Review of Research in Open and Distributed Learning*, 16(1), 260-284.
- Gibson, D. y de Freitas, S. (2016). Exploratory analysis in learning analytics. *Technology, Knowledge and Learning*, 21(1), 5-19. <https://doi.org/10.1007/s10758-015-9249-5>
- Gizem, F., Yilmaz, K. y Yilmaz, R. (2020). Student opinions about personalized recommendation and feedback based on learning analytics. *Technology, Knowledge and Learning*, 25(4), 753-768. <https://doi.org/10.1007/s10758-020-09460-8>
- Goldstein, P. J. y Katz, R. N. (2005). Academic analytics: The uses of management information and technology in higher education. *EDUCAUSE*

- Quarterly*, 8(December), 113. <https://library.educause.edu/resources/2005/12/academic-analytics-the-uses-of-management-information-and-technology-in-higher-education>
- Gómez, D., Oviedo, R. y Martínez, E. (2011). Factores que influyen en el rendimiento académico del estudiante universitario. *Tecnociencia, Educación y Humanidades*, 5(2), 90-97. https://www.researchgate.net/publication/305335495_Factores_que_influyen_en_el_rendimiento_academico_del_estudiante_universitario
- Gómez, D. (2015). *Analítica visual en E-Learning*. Universidad de Salamanca. <https://gredos.usal.es/handle/10366/125950>
- González-Yebra, Ó., Aguilar, M. A., Aguilar, F. J. y Lucas, M. (2018). Evaluación de entornos inmersivos 3D como herramienta de aprendizaje *E-Learning. Educación XXI*, 21(2), 417-440. <https://doi.org/10.5944/educXXI.16204>
- Gottipati, S. y Shankararaman, V. (2018). Competency analytics tool: Analyzing curriculum using course competencies. *Education and Information Technologies*, 23(1), 41-60. <https://doi.org/10.1007/s10639-017-9584-3>
- Greller, W. y Drachsler, H. (2012). Translating learning into numbers: A generic framework for learning analytics. *Journal of Educational Technology & Society*, 15(3), 42-57. https://www.researchgate.net/publication/234057371_Translating_Learning_into_Numbers_A_Generic_Framework_for_Learning_Analytics
- Grob, M., Becerra, D., Rodríguez, A., Cristiane, J., Ramírez, V. y Sabag, N. (2015). Relación entre puntaje de prueba de selección universitaria y nota enseñanza media, y el rendimiento académico de la asignatura de morfología en alumnos de primer año de odontología de la Universidad de los Andes. *International Journal of Morphology*, 33(2), 527-531.
- Guillén-Guerrero, G., Antonio, O. y Molina, M. (2019). Impacto de la capacitación docente con formato *B-Learning* sobre la metodología docente y evaluativa de los profesores. *Iberian Journal of Information Systems and Technologies*, 21(1), 166-176.
- Guizado, G., Valenzuela, M. y Vallejo, P. (2020). Desempeño docente y el rendimiento académico de los estudiantes de la Facultad de Tecnología en la Universidad Nacional de Educación de Perú. *Revista Conrado*, 16(72), 200-203. <http://scielo.sld.cu/pdf/rc/v16n72/1990-8644-rc-16-72-200.pdf>

- Guzmán, C. y Durán, D. (2009). *Deserción estudiantil en la educación superior colombiana*. Ministerio de Educación Nacional - Republica de Colombia. http://www.mineducacion.gov.co/sistemasdeinformacion/1735/articulos-254702_libro_desercion.pdf
- Hackeling, G. (2014). *Mastering with scikit-learn: Learn to implement and evaluate solutions with scikit-learn*. 1. Packt Publishing Ltd.
- Hakyemez, T. C. y Mardikyan, S. (2021). The interplay between institutional integration and self-efficacy in the academic performance of first-year university students: A multigroup approach. *International Journal of Management Education*, 19(1). <https://doi.org/10.1016/j.ijme.2020.100430>
- Hamoud, A. K., Hashim, A. S. y Awadh, W. A. (2018). Predicting student performance in higher education institutions using decision tree analysis. *International Journal of Interactive Multimedia and Artificial Intelligence*, 5(2), 26-31. <https://doi.org/10.9781/ijimai.2018.02.004>
- Harshil, M. y Agrawal, H. (2015). Student performance prediction using machine learning. *International Journal of Engineering Research And Technology (IJERT)*, 4(03), 110-113. <https://doi.org/10.17577/ijertv4is030127>
- Hasan, R., Palaniappan, S., Mahmood, S., Abbas, A., Sarker, K. U. y Sattar, M. U. (2020). Predicting student performance in higher educational institutions using video learning analytics and data mining techniques. *Applied Sciences*, 10(11), 3894. <https://doi.org/10.3390/app10113894>
- Hasbun, T., Araya, A. y Villalon, J. (2016). Extracurricular activities as dropout prediction factors in higher education using decision trees. *16th International Conference on Advanced Learning Technologies Extracurricular*, 3-5. <https://doi.org/10.1109/ICALT.2016.66>
- Hassan, H., Ahmad, N. B. y Anuar, S. (2020). Improved students' performance prediction for multi-class imbalanced problems using hybrid and ensemble approach in educational data mining. *Journal of Physics*, 1529(5). <https://doi.org/10.1088/1742-6596/1529/5/052041>
- Hastie, T., Tibshirani, R. y Friedman, J. (2009). *The elements of statistical learning. Data mining, inference, and prediction*. Second edition, Springer.
- Hayat, A. A., Shateri, K., Amini, M. y Shokrpour, N. (2020). Relationships between academic self-efficacy, learning-related emotions, and metacognitive learning strategies with academic performance in medical

- students: A structural equation model. *BMC Medical Education*, 20(76). <https://doi.org/10.1186/s12909-020-01995-9>
- Hernández, A. G., Meléndez, R. A., Morales, L. A., García, A., Tecpanecatl, J. L. y Algreto, I. (2016). Comparative study of algorithms to predict the desertion in the students at the ITSM-Mexico. *Latin America Transactions*, 14(11), 4573-4578.
- Hernández, C. (2016). Diagnóstico del rendimiento académico de estudiantes de una escuela de educación superior en México. *Revista Complutense de Educación*, 27(3), 1369-1388. <https://revistas.ucm.es/index.php/RCED/article/view/48551/48839>
- Herodotou, C., Rienties, B., Boroowa, A. y Zdrahal, Z. (2019). A large-scale implementation of predictive learning analytics in higher education: The teachers' role and perspective. *Educational Technology Research and Development*, 67, 1273-1306. <https://doi.org/10.1007/s11423-019-09685-0>
- Hirokawa, S. (2018). Key attribute for predicting student academic performance. *ACM International Conference Proceeding Series*, 308-313. <https://doi.org/10.1145/3290511.3290576>
- Huamán, J. (2018). *Evaluación del rendimiento académico estudiantil de la cohorte 2011-2015, según áreas de la carrera de estomatología, Universidad Peruana Cayetano Heredia*. http://repositorio.upch.edu.pe/bitstream/handle/upch/1429/Evaluacion_HuamanSanchez_Jeanette.pdf?sequence=1&isAllowed=y
- Hussain, M., Zhu, W., Zhang, W., Abidi, S. M. R. y Ali, S. (2019). Using to predict student difficulties from learning session data. *Artificial Intelligence Review*, 52(1), 381-407. <https://doi.org/10.1007/s10462-018-9620-8>
- Icekson, T., Kaplan, O. y Slobodin, O. (2020). Does optimism predict academic performance? Exploring the moderating roles of conscientiousness and gender. *Studies in Higher Education*, 45(3), 635-647. <https://doi.org/10.1080/03075079.2018.1564257>
- Iten, L., Arnold, K. y Pistilli, M. (2008). Mining real-time data to improve student success in a gateway course. *Eleventh Annual TLT Conference*.
- Jain, R., Chaudhary, B. y Jain, N. (2016). Impact of mentoring on academic performance & career self-efficacy of business students. *Indian Journal of Industrial Relations*, 51(4), 684-693. https://www-jstor-org.ezproxy.unal.edu.co/stable/43974591?seq=1#metadata_info_tab_contents

- Janani, P., Verma, S., Natarajan, S. y Sinha, A. K. (2019). Communication using IoT. In *Information and communication technology for sustainable development, advances in intelligent systems and computing*. Springer. <https://doi.org/10.1007/978-981-13-7166-0>
- Jaramillo, J. (2017). La importancia de ser médico. *Acta Médica Costarricense*, 59(2). https://www.scielo.sa.cr/scielo.php?script=sci_arttext&pid=S0001-60022017000200048
- Jivet, I., Scheffel, M., Schmitz, M., Robbers, S., Specht, M. y Drachsler, H. (2020). From students with love: An empirical study on learner goals, self-regulated learning and sense-making of learning analytics in higher education. *Internet and Higher Education*, 47(May), 100758. <https://doi.org/10.1016/j.iheduc.2020.100758>
- Jo, I., Kim, D. y Yoon, M. (2015). Constructing proxy variables to measure adult learners' time management strategies in LMS. *Educational Technology & Society*, 18, 214-225.
- Johnson, L., Adams, S., Estrada, V. y Freeman, A. (2015). *NMC horizon report: 2015 higher education edición*. <https://library.educause.edu/resources/2015/2/2015-horizon-report>
- Kauffman, C., Derazin, M., Asmar, A. y Kibble, J. (2019). Patterns of medical student engagement in a second-year pathophysiology course: Relationship to USMLE Step 1 performance. *Advances in Physiology Education*, 43(01), 512-518. <https://doi.org/10.1152/advan.00082.2019>
- Kaunang, F. J. y Rotikan, R. (2018). Students' academic performance prediction using data mining. *Proceedings of the 3rd International Conference on Informatics and Computing (ICIC), 2018*, 1-5. <https://doi.org/10.1109/IAC.2018.8780547>
- Kew, S. N., Petsangsri, S., Ratanaolarn, T. y Tasir, Z. (2018). Examining the motivation level of students in *E-Learning* in higher education institution in Thailand: A case study. *Education and Information Technologies*, 23(6), 2947-2967. <https://doi.org/10.1007/s10639-018-9753-z>
- Khalil, M. K., Hawkins, H. G., Crespo, L. M. y Buggy, J. (2018). The design and development of prediction models for maximizing students' academic achievement. *Medical Science Educator*, 111-117.
- Khalil, M., Khalil, M. y Ebner, M. (2015). Learning analytics: Principles and constraints. *EdMedia: World Conference on Educational Media and Technology*, 2015(1), 1789-1799. <https://www.learntechlib.org/p/151455/>

- Kim, H. J., Hong, A. J. y Song, H. (2019). The roles of academic engagement and digital readiness in students' achievements in university *E-Learning environments*. *International Journal of Educational Technology in Higher Education*, 16(21), 4-18.
- Kim, J., Jo, I. H. y Park, Y. (2016). Effects of learning analytics dashboard: Analyzing the relations among dashboard utilization, satisfaction, and learning achievement. *Asia Pacific Education Review*, 17(1), 13-24. <https://doi.org/10.1007/s12564-015-9403-8>
- Kitchenham, B. (2004). *Procedures for performing systematic reviews*. <http://www.inf.ufsc.br/~aldo.vw/kitchenham.pdf>
- Kitchenham, B. A. y Charters, S. (2007). *Guidelines for performing systematic literature reviews in software engineering*. https://www.researchgate.net/publication/302924724_Guidelines_for_performing_Systematic_Literature_Reviews_in_Software_Engineering
- Kitchenham, B., Pearl Brereton, O., Budgen, D., Turner, M., Bailey, J. y Linkman, S. (2009). Systematic literature reviews in *software engineering* - A systematic literature review. *Information and Software Technology*, 51(1), 7-15. <https://doi.org/10.1016/j.infsof.2008.09.009>
- Klein, C., Lester, J., Rangwala, H. y Johri, A. (2019). Technological barriers and incentives to learning analytics adoption in higher education: Insights from users. *Journal of Computing in Higher Education*, 31(3), 604-625. <https://doi.org/10.1007/s12528-019-09210-5>
- Kuliya, M. y Usman, S. (2020). Perceptions of *E-Learning* among undergraduates and academic staff of higher educational institutions in north-eastern Nigeria. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-020-10325-x>
- Kwankajonwong, N., Ongprakobkul, C., Qureshi, S. P., Watanatada, P., Thanprasertsuk, S. y Bongsebandhu-phubhakdi, S. (2019). Attitude, but not self-evaluated knowledge, correlates with academic performance in physiology in Thai medical students. *Advances in Physiology Education*, 43(3), 324-331.
- Laguna, C. (s.f.). *Correlación y regresión lineal*. https://www.academia.edu/39203766/CORRELACION_Y_REGRESION_LINEAL_Autor_Clara_Laguna_4_1_INTRODUCCION
- Lamas, H. (2015). Sobre el rendimiento escolar. *Propósitos y Representaciones: Revista de Psicología Educativa*, 3(1), 313-386.

- Lasarte, O. F., Díaz, E. R., Palacios, E. G. y Fernández, A. R. (2020). The role of social support in school adjustment during secondary education. *Psicothema*, 32(1), 100-107. https://addi.ehu.es/bitstream/handle/10810/55974/Fernandez-Lasarte%2c%20O_The%20role%20of.pdf?sequence=1
- Law, S. N. (2015). Exploratory network analysis of learning motivation factors in *E-Learning* facilitated computer programming courses. *The Asia-Pacific Education Researcher*, 24(4), 705-717. <https://doi.org/10.1007/s40299-014-0223-0>
- Lenskiy, A., Shariat, R. y Seol, S. (2020). The effect of academic breaks on undergraduate academic performance. *The International Journal of Electrical Engineering & Education*. <https://doi.org/10.1177/0020720920922518>
- Lerma, L. G. y Silva, C. L. (2020). Fortalecimiento de la práctica docente con *Learning Analytics*: estudio de caso. *Praxis & Saber*, 11(2020), 221-245.
- Li, S., Chen, G., Xing, W., Zheng, J. y Xie, C. (2020). Longitudinal clustering of students' self-regulated learning behaviors in engineering design. *Computers & Education*, 153. <https://psycnet.apa.org/record/2020-33762-001>
- Liang, Y.-W., Jones, D. y Robles-Pina, R. A. (2018). Ethnic and gender stereotypes on college students' academic performance. *Research in Higher Education Journal*, 35. <https://files.eric.ed.gov/fulltext/EJ1194415.pdf>
- Lima, J., Alves, P., Pereira, M. y Almeida, S. (2018). Using academic analytics to predict dropout risk in engineering courses. *European Conference on E-Learning*.
- Limsathitwong, K., Tiwatthanont, K. y Yatsungnoen, T. (2018). Dropout prediction system to reduce discontinue study rate of information technology students. *International Conference on Business and Industrial Research (ICBIR)*, 110-114.
- Lin, X. J., Zhang, C. Y., Yang, S., Hsu, M. L., Cheng, H., Chen, J. y Yu, H. (2020). Stress and its association with academic performance among dental undergraduate students in Fujian, China: A cross-sectional online questionnaire survey. *BMC Medical Education*, 20(181), 181. <https://doi.org/10.1186/s12909-020-02095-4>
- Liu, H., Wang, Z., Benachour, P. y Tubman, P. (2018). A time series classification method for behaviour-based dropout prediction. *18th International*

- Conference on Advanced Learning Technologies A*, 191-195. <https://doi.org/10.1109/ICALT.2018.00052>
- Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A. y Tomás-Marco, I. (2014). El análisis factorial exploratorio de los ítems: una guía práctica, revisada y actualizada. *Anales de Psicología*, 30(3), 1151-1169. <https://doi.org/10.6018/analesps.30.3.199361>
- Long, P., Siemens, G., Conole, G. y Gašević, D. (2011). *LAK'11: Proceedings of the 1st International Conference on Learning Analytics and Knowledge*. <https://iucat.iu.edu/iub/13609333>
- Long, P. y Siemens, G. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE Review*, 46(5), 30-32. <https://er.educause.edu/articles/2011/9/penetrating-the-fog-analytics-in-learning-and-education>
- Lonn, S., Aguilar, S. J. y Teasley, S. D. (2015). Investigating student motivation in the context of a learning analytics intervention during a summer bridge program. *Computers in Human Behavior*, 47, 90-97. <https://doi.org/10.1016/J.CHB.2014.07.013>
- López-Aguado, M. y Gutiérrez-Provecho, L. (2019). Cómo realizar e interpretar un análisis factorial exploratorio utilizando SPSS. *REIRE Revista d'Innovació i Recerca en Educació*, 12(2), 1-14. <https://doi.org/10.1344/reire2019.12.227057>
- Luján, P., Trelles, L. y Mogollón, M. (2019). Asertividad y rendimiento académico en estudiantes de la Facultad de Ciencias Administrativas de la Universidad Nacional de Piura. *UCV - Scientia*, 11(1). <http://revistas.ucv.edu.pe/index.php/UCV-SCIENTIA/article/view/2397/1971>
- Luo, L., Cheng, X., Wang, S., Zhang, J., Zhu, W., Yang, J. y Liu, P. (2017). Blended learning with Moodle in medical statistics: An assessment of knowledge, attitudes and practices relating to E-Learning. *BMC Medical Education*, 17(1), 1-8. <https://doi.org/10.1186/s12909-017-1009-x>
- Mahroeian, H., Daniel, B. y Butson, R. (2017). The perceptions of the meaning and value of analytics in New Zealand higher education institutions. *International Journal of Educational Technology in Higher Education*, 14(1), 1-17. <https://doi.org/10.1186/s41239-017-0073-y>
- Maldonado, S. (2007). *Manual práctico para el diseño de la Escala Likert*.
- Mamcenko, J. y Kurilovas, E. (2017). On using learning analytics to personalise learning in virtual learning environments. *European Conference on E-Learning*, 1, 353-361.

- Manzano, A. (2017). Introducción a los modelos de ecuaciones estructurales. *Investigación en Educación Médica*, 7(25), 67-72. <http://www.scielo.org.mx/pdf/iem/v7n25/2007-5057-iem-7-25-67.pdf>
- Marín, J. (s.f.). Análisis de correlación lineal. Los procedimientos. Correlaciones bivariadas y correlaciones parciales. En *SPSS 10 Guía para el análisis de datos* (pp. 323-335). <http://halweb.uc3m.es/esp/Personal/personas/jmmarin/esp/GuiaSPSS/17corlin.pdf>
- Martín, A. N. (2019). Variables motivacionales y cognitivas predictivas del rendimiento en estudiantes universitarios de ciencias de la salud. *Educación Médica Superior*, 33(1), 1-29. <http://scielo.sld.cu/pdf/ems/v33n1/1561-2902-ems-33-01-e1397.pdf>
- Martin, F. y Whitmer, J. C. (2016). Applying learning analytics to investigate timed release in online learning. *Technology, Knowledge and Learning*, 21(1), 59-74. <https://doi.org/10.1007/s10758-015-9261-9>
- Martínez, E., García-Alandete, J., Sellés, P., Bernabé, G. y Soucase, B. (2012). Análisis factorial confirmatorio de los principales modelos propuestos para el *purpose-in-life test* en una muestra de universitarios españoles. *Acta Colombiana de Psicología*, 15(1), 67-76. <http://www.scielo.org.co/pdf/acp/v15n1/v15n1a07.pdf>
- Mavani, U. y Brian, V. (2018). *Information and communication technology for sustainable development. Proceedings of ICT4SD 2018*, 933(2194-5357), 362-372.
- Mavrou, I. (2015). Análisis factorial exploratorio: cuestiones conceptuales y metodológicas. *Revista Nebrija*, 19, 71-80. <https://www.nebrija.com/revista-linguistica/analisis-factorial-exploratorio.html>
- Mazumder, Q. H., Sultana, S. y Mazumder, F. (2020). Correlation between classroom engagement and academic performance of engineering students. *International Journal of Higher Education*, 9(3), 240-247. <https://doi.org/10.5430/ijhe.v9n3p240>
- Medina, E. C., Chunga, C. B., Armas-Aguirre, J., Grandón, E. E. y Bío-bío, U. (2020). Modelo predictivo para reducir el índice de deserción de estudiantes universitarios en el Perú: redes bayesianas vs. árboles de decisión. *15th Iberian Conference on Information Systems and Technologies (CISTI)*, June, 24-27.
- Mendivelso, F. y Rodríguez, M. (2018). Prueba CHI-cuadrado de independencia aplicada a tablas 2xN. *Revista Médica Sanitas*, 21(2), 92-95. <https://revistas.unisanitas.edu.co/index.php/rms/article/view/342>

- Mengash, H. A. (2020). Using data mining techniques to predict student performance to support decision making in university admission systems. *IEEE Access*, 8, 55462-55470. <https://doi.org/10.1109/ACCESS.2020.2981905>
- Merchán-Rubiano, S., Beltrán-Gómez, A. y Duarte-García, J. (2017). Engineering students' academic performance prediction using ICFES test scores and demographic data. *Ingeniería Solidaria*, 13(21). <https://doi.org/10.16925/in.v13i21.1729>
- Mesarić, J. y Šebalj, D. (2016). Decision trees for predicting the academic success of students. *Croatian Operational Research Review*, 7, 367-388. <https://doi.org/10.17535/corr.2016.0025>
- Miguéis, V. L., Freitas, A., García, P. J. V. y Silva, A. (2018). Early segmentation of students according to their academic performance: A predictive modelling approach. *Decision Support Systems*, 115, 36-51. <https://doi.org/10.1016/j.dss.2018.09.001>
- Minichil, W., Eskindir, E., Demilew, D. y Mirkena, Y. (2020). Magnitude of premenstrual dysphoric disorder and its correlation with academic performance among female medical and health science students at University of Gondar, Ethiopia, 2019: A cross-sectional study. *BMJ Open*, 10(e034166). <https://doi.org/10.1136/bmjopen-2019-034166>
- Mirabolghasemi, M., Choshaly, S. H. y Iahad, N. A. (2019). Using the HOT-fit model to predict the determinants of *E-Learning* readiness in higher education: A developing country's perspective. *Education and Information Technologies*, 24(6), 3555-3576. <https://doi.org/10.1007/s10639-019-09945-9>
- Mohamed, M. H. y Waguih, H. M. (2015). Early prediction of student success using a data mining classification technique. *International Journal of Science and Research*, 6(10), 126-131. <https://doi.org/10.21275/ART20177029>
- Mohammadi, M., Dawodi, M., Tomohisa, W. y Ahmadi, N. (2019). Comparative study of supervised learning algorithms for student performance prediction. *1st International Conference on Artificial Intelligence in Information and Communication (ICAIIIC)*, 2019, 124-127. <https://doi.org/10.1109/ICAIIIC.2019.8669085>
- Montero-Granados, R. (2016). *Modelos de regresión lineal múltiple*. Universidad de Granada. https://www.ugr.es/~montero/matematicas/regresion_lineal.pdf

- Montero, E., Villalobos, J. y Valverde, A. (2007). Factores institucionales, pedagógicos, psicosociales y sociodemográficos asociados al rendimiento académico en la Universidad de Costa Rica: un análisis multinivel. *RELIEVE - Revista Electrónica de Investigación y Evaluación Educativa*, 13(2), 215-234. www.uv.es/RELIEVE/v13n2/RELIEVEv13n2_5.htm
mwww.uv.es/RELIEVE]pag.215
- Montoya-Arenas, D. A., Bustamante-Zapata, E. M., Díaz-Soto, C. M. y Pineda, D. (2021). Factores de la capacidad intelectual y de la función ejecutiva relacionados con el rendimiento académico en estudiantes universitarios. *Revista de la Escuela de Ciencias de la Salud de la Universidad Pontificia Bolivariana*, 40(1), 10-18. <https://revistas.upb.edu.co/index.php/medicina/article/view/7014/6548>
- Moral, J., Sánchez, J. y Villarreal, M. (2010). Desarrollo de una escala multidimensional breve de ajuste escolar. *Revista Electrónica de Metodología Aplicada*, 15(1), 1-11.
- Morata-Ramírez, M., Holgado-Tello, F., Barbero-García, I. y Méndez, G. (2015). Análisis factorial confirmatorio. Recomendaciones sobre mínimos cuadrados no ponderados en función del error tipo I de Ji-cuadrado y RMSEA. *Acción Psicológica*, 12(1), 79-90. <https://doi.org/10.5944/ap.12.1.14362>
- Mueen, A., Zafar, B. y Manzoor, U. (2016). Modeling and predicting students' academic performance using data mining techniques. *International Journal of Modern Education and Computer Science*, 8(11), 36-42. <https://doi.org/10.5815/ijmecs.2016.11.05>
- Muljana, P. S. y Luo, T. (2020). Utilizing learning analytics in course design: Voices from instructional designers in higher education. *Journal of Computing in Higher Education*, 206-234. <https://doi.org/10.1007/s12528-020-09262-y>
- Muñoz-Comonfort, A., Leenen, I. y der Goes, T. I. F. (2014). Correlación entre la evaluación diagnóstica y el rendimiento académico de los estudiantes de medicina. *Investigación en Educación Médica*, 3(10), 85-91. <https://www.sciencedirect.com/science/article/pii/S2007505714727310>
- Muñoz-Villanueva, P., Alarcón-Araneda, M., Jarpa-Parra, M. y Morales-Ojeda, I. A. (2020). Hábito y calidad del desayuno; su efecto en el rendimiento académico de estudiantes de técnico de nivel superior en enfermería. *Horizonte Sanitario*, 19(3), 405-414. <https://doi.org/10.19136/hs.a19n3.3706>

- Murnion, P. y Helfert, M. (2013). Academic analytics in quality assurance using organisational analytical capabilities. In University Oxford (ed.), *Annual Conference of the UK Academy of Information Systems (UKAIS)*. <https://doi.org/10.13140/2.1.3368.1600>
- Mushtaq, I. y Khan, S. N. (2012). Factors affecting students' academic performance. *Global Journal of Management and Business Research*, 12(9). https://globaljournals.org/GJMBR_Volume12/3-Factors-Affecting-Students-Academic.pdf
- Nabizadeh, S., Hajian, S., Sheikhan, Z. y Rafiei, F. (2019). Prediction of academic achievement based on learning strategies and outcome expectations among medical students. *BMC Medical Education*, 19(1), 1-11. <https://doi.org/10.1186/s12909-019-1527-9>
- Nguyen, Q., Tempelaar, D. T., Rienties, B. y Bas, G. (2016). What learning analytics-based prediction models tell us about feedback preferences of students. *Quarterly Review of Distance Education*, 17(3), 13-33.
- Nicholson, L. L., Reed, D. y Chan, C. (2016). An interactive, multi-modal anatomy workshop improves academic performance in the health sciences: A cohort study. *BMC Medical Education*, 16(7). <https://doi.org/10.1186/s12909-016-0541-4>
- Nieto, Y., García, V., Montenegro, C., González, C. y González, R. (2019). Usage of for strategic decision making at higher educational institutions. *IEEE Access*, 7, 75007-75017. <https://doi.org/10.1109/ACCESS.2019.2919343>
- Norris, D. M. y Lefrere, P. (2011). Transformation through expeditionary change using online learning and competence-building technologies. *Research in Learning Technology*, 19(1). <https://doi.org/10.3402/rlt.v19i1.17105>
- Núñez-Barriopedro, E., Monclúz, I. M. y Ravina-Ripoll, R. (2019). El impacto de la utilización de la modalidad *B-Learning* en la educación superior. *Alteridad Revista de Educación*, 14(1), 26-39.
- Ñaupari, F. (2014). Evaluación del rendimiento académico de estudiantes universitarios en la modalidad presencial y virtual. *Apuntes. Revista de Ciencias Sociales*, 4(1). <https://dialnet.unirioja.es/servlet/articulo?codigo=5042944>
- Oblinger, G., Campbell, J. P., Deblois, P. B. y Oblinger, D. G. (2007). Academic analytics: A new tool for a new era. *Research in Higher Education*, 1(2),

- 727-742. <http://er.educause.edu/~media/files/article-downloads/erm0742.pdf>
- Oladejo, M. (2010). A path-analytic study of socio-psychological variables and academic performance of distance learners in nigerian universities. *Thesis*, 197. <https://doi.org/DOI: 10.13140/RG.2.2.19443.73762>
- Olsen, J., Sharma, K., Rummel, N. y Aleven, V. (2020). Temporal analysis of multimodal data to predict collaborative learning outcomes. *British Journal of Educational Technology*, 51(5), 1527-1547.
- Ortiz, M., Paredes, M., Soto, R. y Aldana, E. (2020). Ansiedad matemática y desempeño académico en estudiantes en la formación básica de ingeniería. *Formación Universitaria*, 13(4), 93-100. <https://doi.org/10.4067/S0718-50062020000400093>
- Ortlieb, E. T. y Cheek, E. H. (2008). How geographic location plays a role within instruction: Venturing into both rural and urban elementary schools. *Educational Research Quarterly*, 31(2), 48-64. https://www.researchgate.net/publication/234748194_How_Geographic_Location_Plays_a_Role_within_Instruction_Venturing_into_both_Rural_and_Urban_Elementary_Schools
- Oviedo, H. C. y Campo-Arias, A. (2005). Aproximación al uso del coeficiente Alfa de Cronbach. *Revista Colombiana de Psiquiatría*, 34(4), 572-580. <http://www.scielo.org.co/pdf/rcp/v34n4/v34n4a09.pdf>
- Oviedo, H. C. y Campo-Arias, A. (2008). Propiedades psicométricas de una escala: la consistencia interna. *Revista Salud Pública*, 10(5), 831-839. <https://www.redalyc.org/pdf/422/42210515.pdf>
- Page, M., Gaviria, J.-L. y Gómez Bueno, C. (1990). *Hacia un modelo causal del rendimiento académico* (M. de Educación Centro de Publicaciones - Secretaría General - Técnica —ed.—). Gráficas JUMA. <https://www.researchgate.net/publication/39127951>
- Park, Y., Yu, J. H. y Jo, I.-H. (2016). Clustering blended learning courses by online behavior data: A case study in a Korean higher education institute. *The Internet and Higher Education*, 29(29), 1-11. <https://doi.org/10.1016/j.iheduc.2015.11.001>
- Parker, J. D. A., Summerfeldt, L. J., Hogan, M. J. y Majeski, S. A. (2004). Emotional intelligence and academic success: Examining the transition from high school to university. *Personality and Individual Differences*, 36, 163-172. [https://doi.org/10.1016/S0191-8869\(03\)00076-X](https://doi.org/10.1016/S0191-8869(03)00076-X)

- Pascual, M. (2003). El *Blended learning* reduce el ahorro de la formación *on-line* pero gana en calidad. *Educaweb - Formación virtual*. <https://www.educaweb.com/esp/servicios/monografico/formacionvirtual/1181108-a.html>
- Pavelea, A. M. y Moldovan, O. (2020). Why some fail and others succeed: Explaining the academic performance of pa undergraduate students. *NISPAcee Journal of Public Administration and Policy*, 13(1), 109-132. <https://doi.org/10.2478/nispa-2020-0005>
- Pedrosa, I., Juarros-Basterretxea, J., Robles-Fernández, A., Basteiro, J. y García-Cueto, E. (2015). Pruebas de bondad de ajuste en distribuciones simétricas, ¿qué estadístico utilizar? *Universitas Psychologica*, 14(1), 245-254. <https://doi.org/10.11144/Javeriana.upsy13-5.pbad>
- Peña, D. (2002). *Análisis de datos multivariantes*. https://www.researchgate.net/publication/40944325_Analisis_de_Datos_Multivariantes
- Pértega, S. y Pita, S. (2004). Asociación de variables cualitativas: el test exacto de Fisher y el test de McNemar. *Cuadernos de Atención Primaria*, 11(5), 304-308.
- Pham, L., Limbu, Y. B., Bui, T. K., Nguyen, H. T. y Pham, H. T. (2019). Does *E-Learning* service quality influence E-Learning student satisfaction and loyalty? Evidence from Vietnam. *International Journal of Educational Technology in Higher Education*, 16(1). <https://doi.org/10.1186/s41239-019-0136-3>
- Picciano, A. (2014). *Big Data* and learning analytics in blended learning environments: Benefits and concerns. *International Journal of Interactive Multimedia and Artificial Intelligence*, 2(7), 35. <https://doi.org/10.9781/ijimai.2014.275>
- Pineda, J. y Tamayo, L. (2016). E-moderating and E-tivities: The implementation of a workshop to develop online teaching skills in in-service teachers. *Profile Issues in Teachers' Professional Development*, 18(1), 97-114. <https://doi.org/10.15446/profile.v18n1.44269>
- Porto, A. y Gresia, L. Di. (2005). *Performance of university students and their determinants*. http://sedici.unlp.edu.ar/bitstream/handle/10915/54674/Documento_completo_.pdf-PDFA.pdf?sequence=1
- Pradeep, A., Das, S. y Kizhekkethottam, J. J. (2015). Students dropout factor prediction using EDM techniques. *International Conference on Soft-Computing and Network Security*.

- Putpuek, N., Rojanaprasert, N., Atchariyachanvanich, K. y Thamrongthanyawong, T. (2018). Comparative study of prediction models for final GPA score: A case study of Rajabhat Rajanagarindra University. *2018 IEEE/ACIS 17th International Conference on Computer and Information Science (ICIS)*, 92-97. <https://ieeexplore-ieee-org.ezproxy.unal.edu.co/stamp/stamp.jsp?tp=&arnumber=8466475>
- Qin, F., Zhu, L., Cheng, Z. y Zhang, Q. (2017). Research on student performance evaluation based on random forest. *DEStech Transactions on Engineering and Technology Research*, 387-393. <https://doi.org/10.12783/dtetr/eeta2017/7762>
- Qowidho, T., Zarlis, M., Nababan, E. B., Agusnady, A. y Sembiring, B. S. (2019). Analysis of c4.5 and ID3 methods in determining student graduation. *Journal of Physics: Conference Series*, 1255(1). <https://doi.org/10.1088/1742-6596/1255/1/012025>
- Queiroga, E. M., Lopes, J. L., Kappel, K., Aguiar, M., Araújo, R. M., Munoz, R., Villarroel, R. y Cechinel, C. (2020). A learning analytics approach to identify students at risk of dropout: A case study with a technical distance education course. *Applied Sciences*, 1, 1-20. <https://doi.org/10.3390/app10113998>
- Rajak, A. N. H. A., Nurul, D., Pg, N. y Bakar, A. (2018). *E-Learning services acceptance in higher educational institutes: A case study in Brunei. Educ Inf Technol (2018)*, 2341-2361.
- Ramírez, D. (s.f.). *Autocorrelación*. http://webdelprofesor.ula.ve/economia/dramirez/MICRO/FORMATO_PDF/Materialeconometria/Autocorrelacion.pdf
- Razaque, F., Soomro, N., Ahmed, S., Soomro, S., Samo, J. A. y Dharejo, H. (2017). Using Naive Bayes algorithm to students' bachelor academic performances analysis. *Conferencia Internacional IEEE sobre Tecnologías de Ingeniería y Ciencias Aplicadas (ICETAS)*, 1-5.
- Restrepo, L. F. y González, J. (2007). De Pearson a Spearman. *Revista Colombiana de Ciencias Pecuarias*, 20, 183-192. <https://www.redalyc.org/pdf/2950/295023034010.pdf>
- Rivera, E. E., Becerra, S. C., Cotrina, A. R. y Acero, A. C. (2020). Empatía y rendimiento académico en estudiantes universitarios. *Educare*, 24(2), 26. <https://revistas.investigacion-upelipb.com/index.php/educare/article/view/1319/1289>

- Rodríguez, M. (2009). Análisis factorial confirmatorio de la versión uruguaya de la escala Smart de Trapnell para medir capacidad intelectual percibida. *RIDEP*, 1(27), 85-105. <https://www.aidep.org/sites/default/files/2017-07/R275.pdf>
- Romero-Saldaña, M. (2016). Pruebas de bondad de ajuste a una distribución normal. *Enfermería del Trabajo*, 6(3), 105-114.
- Rosales-Jaramillo, C. A., Aguilar-Valladares, Y. M., Vásquez-Zamora, A. L. y Guzmán-Espinoza, C. K. (2018). Caracterización y determinantes del rendimiento académico de estudiantes de ciencias agropecuarias, Universidad de Cuenca. *Polo del Conocimiento*, 3(10), 285-299. <https://doi.org/10.23857/casedelpo.2018.3.10.octubre.285-299>
- Salajegheh, A., Jahangiri, A., Dolan-Evans, E. y Pakneshan, S. (2016). A combination of traditional learning and *E-Learning* can be more effective on radiological interpretation skills in medical students: A pre- and post-intervention study. Approaches to teaching and learning. *BMC Medical Education*, 16(1), 1-7. <https://doi.org/10.1186/s12909-016-0569-5>
- Salal, Y. K., Abdullaev, S. M. y Kumar, M. (2019). Educational data mining: Student performance prediction in academic. *International Journal of Engineering and Advanced Technology*, 8(4C), 54-59.
- Salgado, L. (2009). *Instrumentos de marketing aplicados a la compra de productos ecológicos: un caso de estudio entre Barcelona, España y La Paz, México*. https://www.tesisenred.net/bitstream/handle/10803/1485/08.LSB_Capítulo_8.pdf
- Salmani, A. A. Al, Shidhani, A. Al, Qassabi, S. S. Al, Yaaribi, S. A. Al y Musharfi, A. M. Al. (2020). Prevalence of sleep disorders among university students and its impact on academic performance. *International Journal of Adolescence and Youth*, 25(1), 974-981. <https://doi.org/10.1080/02673843.2020.1815550>
- Sánchez, M., Árraga, M. y Maldonado, L. (2011). *Psicología y educación*. 1ª ed., Universidad de Zulia. https://www.researchgate.net/publication/298344051_Psicologia_y_Educacion
- Sánchez, R. (2015). t-Student: usos y abusos. *Revista Mexicana de Cardiología*, 26(1). http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0188-21982015000100009

- Sanguinetti, J. M., López, R., Vieta, M. E., Berruezo, S. y Chagra, C. (2013). Factores relacionados con el rendimiento académico en alumnos de fisiopatología. *Investigación en Educación Médica*, 2(8), 177-182. <https://www.sciencedirect.com/science/article/pii/S2007505713727108>
- Scholes, V. (2016). The ethics of using learning analytics to categorize students on risk. *Educational Technology Research and Development*, 64(5), 939-955. <https://doi.org/10.1007/s11423-016-9458-1>
- Segnini, S. (2008). Apéndice A: Prueba de normalidad de Shapiro Wilk. En *Fundamentos de Bioestadística* (pp. A1-A4).
- Seoane, A. (2014). *Formalización de un modelo de formación online basado en el factor humano y la presencia docente mediante un lenguaje de patrón*. Tesis doctoral, Universidad de Salamanca. <https://doi.org/10.14201/gredos.123342>
- Shahiri, A. M., Husain, W. y Rashid, N. A. (2015). A review on predicting student's performance using data mining techniques | Elsevier enhanced reader. *Procedia Computer Science*, 72, 414-422. <https://www.sciencedirect.com/science/article/pii/S1877050915036182>
- Shahrakipour, M., Arbabisarjou, A., Zare, S., Ghoreishinia, G., Nazemzadeh, M. S., Ansari, H., Rigi, S., Jahadpanah, Y., Danesh, A. y Jafari, H. (2016). Evaluation of academic performance among students of medical sciences. *International Journal of Pharmacy & Technology*, 8(2), 12241-12247. https://www.researchgate.net/publication/306215731_Evaluation_of_academic_performance_among_students_of_medical_sciences
- Shuttleworth, M. (s.f.). *Validez convergente y discriminante*. <https://explorable.com/es/validez-convergente-y-discriminante>
- Siddiqi, S. y Memon, Z. A. (2017). Internet addiction impacts on time management that results in poor academic performance. *Proceedings - 14th International Conference on Frontiers of Information Technology, FIT 2016*, 63-68. <https://doi.org/10.1109/FIT.2016.020>
- Siemens, G. (2013). Learning analytics. *American Behavioral Scientist*, 57(10), 1380-1400. <https://doi.org/10.1177/0002764213498851>
- Sivasakthi, M. (2017). Classification and prediction based data mining algorithms to predict students' introductory programming performance. *International Conference on Inventive Computing and Informatics (ICICI 2017)*, 346-350.

- Stanko, T., Zhirosh, O., Johnston, D. y Gartsev, S. (2017). On possibility of prediction of academic performance and potential improvements of admission campaign at IT university. *IEEE Global Engineering Education Conference*, 862-866.
- Stewart, C. (2017). Learning analytics: Shifting from theory to practice. *Journal on Empowering Teaching Excellence*, 1(1), 10. <https://doi.org/10.15142/T3G63W>
- Tahir, R. y Wang, A. I. (2020). Codifying game-based learning: Development and application of LEAGUE framework for learning games. *The Electronic Journal of E-Learning*, 18(1 Jan 2020), 69-87. <https://doi.org/10.34190/ejel.20.18.1.006>
- Tejedor, B., Santos, M., García-Orza, J., Carratalà, P. y Navas, M. (2009). Variables explicativas de la ansiedad frente a las matemáticas: un estudio de una muestra de 6° de primaria. *Anuario de Psicología*, 40(3), 345-355.
- Tenpipat, W. y Akkarajitsakul, K. (2020). Student dropout prediction: A KMUTT case study. *International Conference on Big Data Analytics and Practices (IBDAP) Student*.
- Teo, T. S. H., Kim, S. L. y Jiang, L. (2020). *E-Learning* implementation in South Korea: Integrating effectiveness and legitimacy perspectives. *Information Systems Frontiers*, 22(2), 511-528. <https://doi.org/10.1007/s10796-018-9874-3>
- Thomas, B. L., Bipath, P. y Viljoen, M. (2020). Inflammatory activity and academic performance in university students. *Journal of Psychology in Africa*, 30(1), 53-57. <https://doi.org/10.1080/14330237.2020.1712802>
- Thomas, B. L. y Viljoen, M. (2019). Heart rate variability and academic performance of first-year university students. *Neuropsychobiology*, 78(4), 175-181. <https://doi.org/10.1159/000500613>
- Tomás, J., Expósito, M. y Sempere, S. (2014). Determinantes del rendimiento académico en los estudiantes de grado. Un estudio en administración y dirección de empresas. *Revista de Investigación Educativa*, 32(2), 379-392. <https://doi.org/dx.doi.org/10.6018/rie.32.2.177581>
- Tomasevic, N., Gvozdenovic, N. y Vranes, S. (2020). An overview and comparison of supervised data mining techniques for student exam performance prediction. *Computers and Education*, 143(August 2019), 103676. <https://doi.org/10.1016/j.compedu.2019.103676>

- Urteaga, I., Siri, L. y Garófalo, G. (2020). Predicción temprana de deserción mediante aprendizaje automático en cursos profesionales en línea. *RIED. Revista Iberoamericana de Educación a Distancia*, 23(2), 147-167. <https://doi.org/10.5944/ried.23.2.26356>
- Vargas, H., Solórzano, L. y Chanini, W. (2019). Modelo matemático entre el puntaje de examen de ingreso y el rendimiento académico de los estudiantes ingresantes a la Universidad Nacional Jorge Basadre Grohmann, año académico 2018. *Ciencias*, 3(3). <http://revistas.unjbg.edu.pe/index.php/cs/article/view/949/1069>
- Vargas, I., Ramírez, C., Cortés, J., Farfán, A. y Heinze, G. (2011). Factores asociados al rendimiento académico en alumnos de la Facultad de Medicina: estudio de seguimiento a un año. *Salud Mental*, 34, 301-308. <http://www.scielo.org.mx/pdf/sm/v34n4/v34n4a2.pdf>
- Vesin, B., Mangaroska, K. y Giannakos, M. (2018). Learning in smart environments: User-centered design and analytics of an adaptive learning system. *Smart Learning Environments*, 5(24). <https://doi.org/10.1186/s40561-018-0071-0>
- Viberg, O., Hatakka, M., Bälter, O. y Mavroudi, A. (2018). The current landscape of learning analytics in higher education. *Computers in Human Behavior*, 89, 98-110. <https://doi.org/10.1016/j.chb.2018.07.027>
- Vijayalakshmi, M., Salimath, S., Shettar, A. S. y Bhadri, G. (2018). A study of team formation strategies and their impact on individual student learning using educational data mining (EDM). *Proceedings - IEEE 9th International Conference on Technology for Education, T4E 2018*, 182-185. <https://doi.org/10.1109/T4E.2018.00047>
- Villanueva, A. B. (2012). El progreso de la estadística y su utilidad en la evaluación del desarrollo. *Papeles de Población*, 18(73), 1-31. <https://www.redalyc.org/pdf/112/11224638010.pdf>
- Viloria, A., García Guliany, J., Niebles Núñez, W., Palma, H. H. y Niebles Núñez, L. (2020). Data mining applied in school dropout prediction. *Journal of Physics*, 1432, 12092. <https://doi.org/10.1088/1742-6596/1432/1/012092>
- Vora, D. R. y Rajamani, K. (2019). A hybrid classification model for prediction of academic performance of students: A *Big Data* application. *Evolutionary Intelligence*, 15, 1083-1096. <https://doi.org/10.1007/s12065-019-00303-9>

- Wang, Q., Wu, W. y Qi, Y. (2018). A learning analytic model for smart classroom. In *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)*. Vol. 11268 LNCS. Springer International Publishing. https://doi.org/10.1007/978-3-030-01298-4_19
- Wati, M., Indrawan, W., Widiars, J. A. y Puspitasari, N. (2018). Data mining for predicting students' learning result. *Proceedings of the 2017 4th International Conference on Computer Applications and Information Processing Technology, CAIPT 2017, 2018-Janua*, 1-4. <https://doi.org/10.1109/CAIPT.2017.8320666>
- Wen, C. T., Chang, C. J., Chang, M. H., Fan Chiang, S. H., Liu, C. C., Hwang, F. K. y Tsai, C. C. (2018). The learning analytics of model-based learning facilitated by a problem-solving simulation game. *Instructional Science*, 46(6), 847-867. <https://doi.org/10.1007/s11251-018-9461-5>
- Wheeler, E., Finucane, S., Soper, S. y Mayhew, T. (2018). Perceived value of preadmission physical therapy observation hours and their ability to predict academic performance. *Journal of Physical Therapy Education*, 32(1), 17-25.
- Wiyono, S., Abidin, T., Bersama, P. H., Lor, P. y Tegal, K. (2019). *Comparative study of KNN, SVM, and decision tree algorithm to predict student's performance*. 7(Iss. 1) (January), 190-196. <https://doi.org/10.5281/zenodo.2550651>
- Yassein, N. A., Gaffer, R., Helali, M. y Mohomad, S. B. (2017). Predicting student academic performance in KSA using data mining techniques. *Journal of Information Technology & Software Engineering*, 7(5), 1-5. <https://doi.org/10.4172/2165-7866.1000213>
- Yen, C., Chen, I., Lai, S. y Chuang, Y. (2015). An analytics-based approach to managing cognitive load by using log data of learning management systems and footprints of social media. *Journal of Educational Technology & Society*, 18, 141-158.
- York, T. T., Gibson, C. y Rankin, S. (2015). Defining and measuring academic success. *Practical Assessment, Research, and Evaluation*, 20(5). <https://doi.org/10.7275/hz5x-tx03>
- Zacharias, D. y Athanasios, A. (2020). Monitoring of the results through a survey concerning the socio-economic characteristics of the elderly using geographic information systems (GIS): A case study in Greece.

- International Journal of Innovation and Economic Development*, 6(3), 36-45. <https://doi.org/10.18775/ijied.1849-7551-7020.2015.64.2004>
- Zárate, E., Lavado, B. y Pomahuacre, W. (2020). Competencia comunicativa intercultural y rendimiento académico en lenguas extranjeras. *Revista Conrado*, 16(74). <https://orcid.org/0000-0002-2924-6771>
- Zuniga-Jara, S., Pizarro-Díaz, R., Escudero-López, M. E. y Honores-Marín, G. (2018). Rendimiento académico universitario y conductas alimentarias. *Formación Universitaria*, 11(2), 57-64. <https://doi.org/10.4067/S0718-50062018000200057>

Autores

Leonardo Emiro Contreras Bravo

Ingeniero mecánico, máster en Ingeniería (materiales y procesos de manufactura), PhD(c) en Ingeniería. Docente de planta, Universidad Distrital Francisco José de Caldas. Director, grupo de investigación en Diseño, Modelamiento y Simulación, DIMSI. Áreas de conocimiento: diseño, manufactura, análisis de datos, aprendizaje automático.

Giovanny Mauricio Tarazona Bermúdez

Ingeniero industrial, máster en Diseño y Gestión de proyectos tecnológicos. Doctor en Sistemas y servicios informáticos para internet. Docente de planta, Universidad Distrital Francisco José de Caldas. Director, grupo de investigación en Comercio Electrónico Colombiano, GICOECOL. Áreas de conocimiento: innovación estratégica, comercio electrónico.

Héctor Javier Fuentes López

Economista. Máster en Economía. PhD(c) en Geografía. Docente de planta, Universidad Distrital Francisco José de Caldas. Grupo de Investigación en Diseño, Modelamiento y Simulación DIMSI. Áreas del conocimiento: economía, econometría, geografía económica, estudios sociales, analítica.

Este libro se terminó
de imprimir en mayo
de 2023 en
DGP Editores
Bogotá, Colombia