



**UNIVERSIDAD POLITÉCNICA SALESIANA
SEDE GUAYAQUIL
CARRERA DE INGENIERÍA DE SISTEMAS**

**MAPEO SISTEMÁTICO SOBRE EL SEGUIMIENTO DEL APRENDIZAJE
DE ESTUDIANTES MEDIANTE EL USO DE LA MINERÍA DE DATOS
EDUCATIVOS**

Trabajo de titulación previo a la obtención del
Título de Ingeniero de Sistemas

AUTOR: Puga Paredes Jorge Luis

TUTOR: Ing. Joe Llerena Izquierdo MSig.

Guayaquil – Ecuador

2023

**CERTIFICADO DE RESPONSABILIDAD Y AUTORÍA DEL TRABAJO DE
TITULACIÓN**

Yo, Puga Paredes Jorge Luis con documento de identificación N° 0925635252 manifiesto que:
Soy el autor y responsable del presente trabajo; y, autorizo a que sin fines de lucro la Universidad Politécnica Salesiana pueda usar, difundir, reproducir o publicar de manera total o parcial el presente trabajo de titulación.

Guayaquil, 22 de julio del año 2023

Atentamente,



Puga Paredes Jorge Luis

0925635252

**CERTIFICADO DE CESIÓN DE DERECHOS DE AUTOR DEL TRABAJO DE
TITULACIÓN A LA UNIVERSIDAD POLITÉCNICA SALESIANA**

Yo, Puga Paredes Jorge Luis con documento de identificación No. 0925635252, expreso mi voluntad y por medio del presente documento cedo a la Universidad Politécnica Salesiana la titularidad sobre los derechos patrimoniales en virtud de que soy autor(a) del Artículo Académico: “**Mapeo sistemático sobre el seguimiento del aprendizaje de estudiantes mediante el uso de minería de datos educativos**”, el cual ha sido desarrollado para optar por el título de: Ingeniero de Sistemas, en la Universidad Politécnica Salesiana, quedando la Universidad facultada para ejercer plenamente los derechos cedidos anteriormente.

En concordancia con lo manifestado, suscribo este documento en el momento que hago la entrega del trabajo final en formato digital a la Biblioteca de la Universidad Politécnica Salesiana.

Guayaquil, 22 de julio del año 2023

Atentamente,



Puga Paredes Jorge Luis

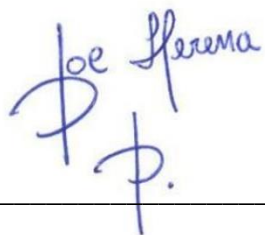
0925635252

CERTIFICADO DE DIRECCIÓN DEL TRABAJO DE TITULACIÓN

Yo, Ing. Joe Llerena Izquierdo MSig con documento de identificación N° 0914884879, docente de la Universidad Politécnica Salesiana, declaro que bajo mi tutoría fue desarrollado el trabajo de titulación: “**Mapeo sistemático sobre el seguimiento del aprendizaje de estudiantes mediante el uso de minería de datos educativos**”, realizado por Puga Paredes Jorge Luis con documento de identificación N° 0925635252, obteniendo como resultado final el trabajo de titulación bajo la opción Artículo Académico que cumple con todos los requisitos determinados por la Universidad Politécnica Salesiana.

Guayaquil, 22 de julio del año 2023

Atentamente,

A handwritten signature in blue ink that reads "Joe Llerena" above a stylized monogram "JL".

Ing. Joe Llerena Izquierdo MSig

0914884879

DEDICATORIA

Dedico este trabajo a todas las personas que buscan mejorar la calidad de la educación y promover el aprendizaje efectivo. Que este estudio pueda contribuir de alguna manera a la comprensión y aplicación de la MDE en beneficio de la educación y de aquellos que la buscan.

AGRADECIMIENTO

Agradezco a todas las personas que contribuyeron de alguna manera en la realización de este trabajo. Quiero expresar mi agradecimiento a mi tutor Ing. Joe Llerena Izquierdo MSig, cuyo conocimiento y guía fueron fundamentales en el desarrollo de este proyecto.

Además, quiero dedicar este trabajo a mis padres y hermanos, quienes siempre estuvieron a mi lado, brindándome su amor, comprensión y motivación. Su apoyo incondicional fue mi mayor fortaleza y me impulsó a dar lo mejor de mí en cada etapa de este proyecto.

A todos ellos, mi más sincero agradecimiento y dedicación. Sin su apoyo y contribución, este trabajo no habría sido posible.

RESUMEN

Este artículo presenta un mapeo sistemático de la literatura sobre el seguimiento del aprendizaje utilizando MD educativos, con un enfoque específico en la evaluación del rendimiento académico mediante árboles de decisión. La revisión de literatura se realizó utilizando una metodología PRISMA que incluyó la recopilación y análisis de datos de estudios relevantes en el campo.

En cuanto a la metodología, se utilizaron métodos y técnicas de recopilación y análisis de datos, como la identificación de palabras clave y una estrategia de búsqueda específica. Se establecieron criterios de inclusión y exclusión para la selección de estudios, y se aplicó una técnica de muestreo por bola de nieve para ampliar la búsqueda. Además, se evaluó la calidad de los estudios seleccionados.

En la implementación de árboles de decisión para la evaluación del rendimiento académico, se siguieron varios pasos. Primero, se realizó una selección de características relevantes, considerando variables y características específicas relacionadas con el rendimiento de los estudiantes. Luego, se llevó a cabo un preprocesamiento de datos para asegurar la calidad y consistencia de la información. A continuación, se construyó el modelo de árbol de decisión (MAD) utilizando los datos seleccionados y preprocesados. Se procedió a entrenar y evaluar el modelo. Por último, se interpretaron los resultados obtenidos a partir del árbol de decisión (AD), identificando patrones y relaciones relevantes. Los resultados del mapeo sistemático revelaron características generales de los estudios incluidos, así como los enfoques, técnicas, metodologías y herramientas utilizadas en la MDE y el seguimiento del aprendizaje. Se identificaron variables y características consideradas en los estudios, y se destacaron los principales hallazgos y resultados. En la discusión, se analizaron las tendencias y patrones identificados en la literatura, así como las limitaciones y desafíos encontrados en el campo. Además, se presentaron aplicaciones prácticas y recomendaciones basadas en los resultados obtenidos. El mapeo sistemático proporciona una visión general de la literatura existente sobre el seguimiento del aprendizaje utilizando MD educativos, con un enfoque en la evaluación del rendimiento académico mediante árboles de decisión. Se destacan las fortalezas, limitaciones y áreas de investigación futura en este campo.

Palabras claves: mapeo sistemático, minería de datos educativos, seguimiento del aprendizaje, rendimiento académico, árboles de decisión.

ABSTRACT

This article presents a systematic mapping of the literature on monitoring learning using educational data mining, with a specific focus on the assessment of academic achievement using decision trees. The literature review was conducted using a PRISMA methodology that included the collection and analysis of data from relevant studies in the field.

Regarding the methodology, data collection and analysis methods and techniques were used, such as the identification of keywords and a specific search strategy. Inclusion and exclusion criteria were found for study selection, and a snowball testing technique was applied to broaden the search. Also, see the quality assessment of selected studies.

In the implementation of decision trees for the evaluation of academic performance, several steps were followed. First, a selection of relevant characteristics was made, considering specific variables and characteristics related to student performance. Then, a data pre-processing was carried out to ensure the quality and consistency of the information. Next, the decision tree model was built using the selected and preprocessed data. The model was trained and evaluated. Finally, the results obtained from the decision tree were interpreted, identifying relevant patterns and relationships.

The results of the systematic mapping revealed general characteristics of the included studies, as well as the approaches, techniques, methodologies, and tools used in educational data mining and learning monitoring. Variables and characteristics considered in the studies were identified, and the main findings and results were highlighted.

In the discussion, the trends and patterns identified in the literature were analyzed, as well as the limitations and challenges found in the field. In addition, practical applications and recommendations based on the results obtained are presented.

Systematic mapping provides an overview of the existing literature on learning monitoring using educational data mining, with a focus on academic performance assessment using decision trees. Strengths, limitations, and areas for future research in this field are highlighted.

Keywords: systematic mapping, educational data mining, learning monitoring, academic performance, decision trees.

ÍNDICE DE CONTENIDO

1.	INTRODUCCIÓN	1
2.	REVISIÓN DE LITERATURA.....	1
3.	METODOLOGÍA	8
3.1.	Métodos y técnicas de Recopilación de datos empleadas	9
3.2.	Métodos y técnicas de Análisis de datos	10
3.3.	Diseño del mapeo sistemático	11
3.4.	Identificación de palabras clave y estrategia de búsqueda.....	11
3.5.	Criterios de inclusión y exclusión	11
3.6.	Proceso de selección de estudios.....	12
3.7.	Extracción de datos	12
3.8.	Evaluación de la calidad de los estudios	13
3.9.	Técnica de muestreo por bola de nieve	13
3.10.	Evaluación del rendimiento académico mediante árboles de decisión	14
3.10.1.	Selección de características relevantes.....	14
3.10.2.	Preprocesamiento de datos	14
3.10.3.	Construcción del modelo de árbol de decisión.....	14
3.10.4.	Entrenamiento y evaluación del modelo	15
3.10.5.	Interpretación de los resultados.....	15
4.	RESULTADOS	15
4.1.	Características generales de los estudios incluidos	15
4.2.	Enfoques y técnicas utilizadas en la MD educativos	16
4.3.	Metodologías y herramientas empleadas en el seguimiento del aprendizaje.....	16
4.4.	Variables y características consideradas en los estudios.....	18
4.5.	Resultados y hallazgos principales.....	18

5.	DISCUSIÓN.....	34
5.1.	Tendencias y patrones identificados en la literatura	34
5.2.	Limitaciones y desafíos encontrados.....	36
5.3.	Aplicaciones prácticas y recomendaciones	37
6.	CONCLUSIONES	38
	REFERENCIAS	39

1. INTRODUCCIÓN

El seguimiento del proceso de enseñanza de los estudiantes en el ámbito educativo es esencial para evaluar su progreso. La MDE ha surgido como un instrumento prometedor para analizar grandes volúmenes de datos generados en entornos educativos, brindando información valiosa para optimizar los procesos de enseñanza y aprendizaje (Echeverry et al., 2017). Esta disciplina interdisciplinaria combina técnicas de MD, aprendizaje automático e inteligencia artificial para descubrir patrones y tendencias en los datos educativos, como el desempeño académico y las interacciones de los estudiantes (Almasri et al., 2023).

En la era digital, los sistemas educativos están adoptando cada vez más herramientas digitales y plataformas en línea, lo que genera una gran cantidad de información (Atiaja Balseca, 2023). La MDE se muestra como una estrategia innovadora para aprovechar estos datos y optimizar la calidad de la educación (Albreiki et al., 2021). El mapeo sistemático del seguimiento del aprendizaje de los estudiantes a través de la MDE permite explorar y analizar la investigación existente, identificando fortalezas, debilidades y brechas en el conocimiento actual (Buenaño Fernández, 2020)(Salihoun, 2020). Además, la MDE fomenta la innovación en el diseño de estrategias pedagógicas y currículos adaptativos al proporcionar información sobre patrones de comportamiento y preferencias individuales de los estudiantes, promoviendo un aprendizaje más autónomo y significativo (Carrascal & Giraldo, 2019). Entender el potencial de la MDE permitirá identificar áreas de investigación futura y avanzar hacia un sistema educativo más eficiente y personalizado (Echeverry et al., 2017)(Wang et al., 2023).

2. REVISIÓN DE LITERATURA

En el estudio de la literatura, se destacará la importancia de la MDE como un instrumento para extraer información valiosa sobre el rendimiento de los estudiantes, los patrones de comportamiento y las características que contribuyen en el proceso de aprendizaje (Zhang et al., 2021)(Ferguson, 2012)(Raquel Ayala Carabajo, 2011). Se examinarán enfoques como la exploración de secuencias, la identificación de patrones y la predicción del rendimiento estudiantil (Parhizkar et al., 2023). Además, se explorarán las metodologías utilizadas, como el aprendizaje automático y el modelado estadístico, y las herramientas y software empleados en los estudios de MDE (Song et al., 2023). Al entender el estado existente de la investigación y

las tendencias emergentes, se podrán identificar áreas de oportunidad y brindar orientación para futuros trabajos en este campo en constante evolución (Yağcı, 2022).

2.1 Marco teórico

A continuación, se proporciona un marco teórico que brinda una base conceptual sólida para comprender y contextualizar la MDE y su relación con el seguimiento del aprendizaje de los estudiantes (Leitner et al., 2017)(Raquel Ayala Carabajo, 2008). Se explora conceptos clave como la analítica del aprendizaje (LA), la toma de decisiones pedagógicas apoyadas en datos y la importancia de la personalización educativa (Levine et al., 2023)(Raquel Ayala Carabajo, 2018).

MD web

La MD web se refiere al proceso de hallazgo de información de carácter valiosa y no trivial en considerables repositorios de datos, que previamente era desconocida y puede ser potencialmente útil (X. Chen et al., 2022). Esta área multidisciplinaria involucra diferentes enfoques computacionales, como la estructuración de árboles de decisión, inducción de reglas, redes neuronales artificiales, aprendizaje basado en instancias, aprendizaje bayesiano, programación lógica y algoritmos estadísticos (Morales et al., 2005)(Almasri et al., 2023). La MD web es un proceso que involucra diversas etapas, a partir de la preparación de los datos hasta la interpretación y aplicación del conocimiento descubierto (R Ayala Carabajo, 2009)(Joe Llerena-Izquierdo & Ayala-Carabajo, 2022b).

Las actividades y procedimientos principales en la MD incluyen la categorización, la asociación, la evaluación, técnica en la MD que busca establecer relaciones y conexiones entre diferentes variables o atributos en un conjunto de datos, la visualización y el descubrimiento de reglas. La MD es parte de los pasos dentro del proceso de develamiento de conocimiento, que comprende varias etapas:

- a. Preprocesamiento: incluye la obtención o extracción de datos, la depuración de estos, la transformación de variables, la selección de atributos relevantes para unificar información proveniente de diversas fuentes, con el fin de crear un conjunto de datos coherente y completo.
- b. MD: implica la elección de los algoritmos que se utilizan para MD apropiados y su aplicación a los datos.

c. Post procesamiento: se refiere al análisis los resultados obtenidos a través de las técnicas de extracción, procesamiento de datos y la medición de los resultados obtenidos, así como el uso del conocimiento descubierto.

Tipos de minería de Web

La MW se divide en tres tipos: minería de contenidos web, minería de estructura web y minería de utilización web. La minería de utilización web es la más utilizada en la enseñanza basada en web y se centra en analizar los datos de utilización de plataformas educativas en línea para identificar patrones de interacción de los estudiantes y comprender su comportamiento de aprendizaje (Dunajeva et al., 2022)(Baradwaj & Pal, 2012)(Raquel Ayala Carabajo, 2011). Esta técnica es fundamental para personalizar la enseñanza y optimizar la experiencia de aprendizaje en línea. En el marco teórico se abordan conceptos relacionados con la MD educativos, y también en relación con la LA, la toma de decisiones pedagógicas basadas en datos y la personalización educativa (Levine et al., 2023)(Raquel Ayala Carabajo, 2018). Estos elementos permiten optimizar la calidad de la enseñanza y proporcionar experiencias educativas más efectivas y adaptadas a las necesidades distintas de cada estudiante (Araka et al., 2022).

2.1.1 Analítica del aprendizaje

La analítica del aprendizaje se enfoca en recopilar, analizar y utilizar datos educativos para optimizar el proceso de aprendizaje. Según (Siemens & Long, 2011)(Tamada et al., 2022), esta disciplina utiliza técnicas de MD y análisis estadístico para identificar patrones y posibles relaciones en los datos educativos, con el fin de incrementar la eficacia de la enseñanza (R. Ayala Carabajo, 2016). (Romero & Ventura, 2020)(Llerena et al., 2021) también destacan el uso de técnicas de MD y análisis estadístico en LA para obtener información valiosa y tomar decisiones informadas en la enseñanza. Por su parte, (Papamitsiou & Economides, 2014)(Álava Morán, 2021)(Joe Llerena-Izquierdo & Ayala-Carabajo, 2022b) realizan una exploración sistemática de la literatura en la que analizan estudios de casos experimentales en el campo de LA y la MD educativos. Su objetivo es proporcionar una visión general de las evidencias empíricas y destacar la importancia de la gestión y análisis de datos como parte del procedimiento de toma de decisiones educativas. Además, resaltan la necesidad de un enfoque holístico que integre tanto en el análisis del aprendizaje, así como lo es la MDE para comprender completamente los procesos de aprendizaje. En conjunto, estos autores subrayan la importancia de utilizar datos de tipo educativo como forma de optimizar el proceso de

enseñanza y el aprendizaje, respaldando sus argumentos con evidencia empírica y destacando la relevancia de las diferentes metodologías que se utilizan para realizar la observación de datos en entornos educativos.

2.1.2 Proceso de elección de acciones o medidas educativas fundamentadas en datos

El proceso de realizar decisiones pedagógicas basadas en datos implica utilizar información derivada del proceso de MDE para decisiones informadas y personalizadas. Según (Aulakh et al., 2023), esto permite identificar las fortalezas, debilidades y necesidades específicas de los estudiantes. (Ordoñez-Avila et al., 2023) enfatizan en el uso de técnicas de MD y análisis estadístico en esta toma de decisiones (Moon et al., 2023). Por otro lado, destaca la importancia de utilizar los datos del aprendizaje de los estudiantes para identificar áreas de mejora y adaptar estrategias de enseñanza. (Švábensky et al., 2022) resaltan la relevancia de LA en la recopilación y uso de datos educativos. Además, (Papamitsiou & Economides, 2014)(Joe Llerena-Izquierdo & Ayala-Carabajo, 2022a) subrayan la necesidad de evidencia empírica y un enfoque holístico en el entendimiento de los procesos de aprendizaje. En conjunto, estos autores enfatizan cómo el análisis de datos educativos mejora la toma de decisiones pedagógicas y personalizadas (Lampropoulos, 2023).

2.1.3 Personalización educativa

La personalización educativa se centra en adaptar la enseñanza y los recursos educativos a las necesidades individuales de los estudiantes. (Taipalus, 2023) destacan que la MDE proporciona información detallada sobre el desempeño y las características de los estudiantes, lo que permite ofrecer experiencias educativas personalizadas, como la adaptación de la dificultad de las actividades y la recomendación de recursos adicionales. (Viberg & Grönlund, 2023) definen la personalización educativa como un conjunto diverso de estrategias y recursos que vinculan los intereses y experiencias del aprendiz con los objetivos educativos, a través de su participación en el control y elección de los procesos de aprendizaje. Destacan la importancia de las identidades en curso de los estudiantes y sugieren el uso de "artefactos identitarios" para proyectar sus intereses y aficiones. (Cardona et al., 2023) mencionan que la personalización se ha convertido en una tendencia creciente impulsada por las tecnologías de la información y comunicación (Issah et al., 2023). Destacan la importancia de la diferenciación de servicios educativos para permitir la personalización a nivel del estudiante. En resumen, la personalización educativa implica adaptar los contenidos y recursos educativos a las

necesidades individuales de los estudiantes, utilizando la información proporcionada por la MDE (Batoool et al., 2023).

2.1.4 Enfoques de MDE para el seguimiento del aprendizaje

Los enfoques de MDE se centran en el análisis de datos generados en entornos educativos para extraer información relevante y obtener conocimientos sobre el rendimiento de los estudiantes, patrones de comportamiento y características que puedan influir en el proceso de aprendizaje (Arnold et al., 2023)(J Llerena-Izquierdo, 2023). Estos enfoques utilizan técnicas de análisis de datos para descubrir relaciones, identificar tendencias y realizar predicciones sobre el rendimiento estudiantil (Arnold et al., 2023)(Alghamdi & Rahman, 2023). Los enfoques se basan en el análisis de masivas cantidades de datos generados en entornos educativos, como registros de actividades, resultados de evaluaciones, interacciones en plataformas de aprendizaje en línea (Pallathadka et al., 2023)(Albreiki et al., 2021). El objetivo principal es utilizar esta información para entender mejor el proceso de aprendizaje, personalizar la enseñanza y brindar apoyo individualizado a los estudiantes.

Enfoques utilizados en la MDE que se utilizan en el seguimiento del aprendizaje de los alumnos

En esta sección, se presentarán los diferentes enfoques utilizados en la MDE que se utiliza para el seguimiento del aprendizaje de los alumnos. Se discutirán técnicas como la exploración de secuencias, la identificación de patrones y la predicción del rendimiento estudiantil.

Enfoque análisis de secuencias

Los autores (Romero & Ventura, 2020) en su artículo mencionan que el análisis de secuencias es un procedimiento utilizado para analizar el orden temporal de los eventos de aprendizaje y las actividades de los estudiantes, permitiendo identificar patrones de comportamiento relevantes. En esta perspectiva, el estudio de secuencias permite detectar modelos de conducta recurrentes y establecer relaciones causales entre las acciones de los estudiantes y los resultados académicos. Al utilizar metodologías de MD, como algoritmos de secuencias frecuentes o procedimientos de minería de secuencias de patrones, se pueden descubrir secuencias de acciones que suelen estar asociadas con un mejor rendimiento académico.

El autor (de Baker & Inventado, 2014) en su trabajo resalta el análisis de secuencias como una técnica efectiva para descubrir patrones de comportamiento en el aprendizaje de los estudiantes,

incluyendo las secuencias de actividades que están asociadas con un mejor rendimiento académico. En este sentido, el enfoque de análisis de secuencias en la MDE ha demostrado ser un instrumento efectiva para comprender el orden temporal de los eventos de aprendizaje y las acciones de los estudiantes. Especialmente al examinar las secuencias de actividades y eventos en donde se pueden identificar patrones significativos que ofrecen conocimientos significativos acerca del proceso de adquisición de conocimiento y el rendimiento estudiantil.

El autor (Albreiki et al., 2021) en su estudio, destaca que el análisis de secuencias como una técnica utilizada en la MDE para identificar patrones secuenciales y establecer relaciones entre las acciones de los estudiantes y sus resultados académicos. Especialmente al analizar el orden temporal de los eventos de aprendizaje y las acciones realizadas por los estudiantes. Esta técnica permite identificar patrones de comportamiento, como las secuencias de actividades que suelen llevar a un mejor rendimiento académico o la detección de desviaciones en las secuencias de acciones que pueden indicar dificultades de aprendizaje.

Por ejemplo, en un estudio de los autores (Olaleye & Vincent, 2020) se utilizó el análisis de secuencias para analizar el comportamiento de los estudiantes en un entorno de aprendizaje en línea. Los investigadores descubrieron que los estudiantes que seguían una secuencia específica de actividades, como leer el material de estudio antes de participar en foros de discusión y luego completar ejercicios prácticos, tenían un mayor rendimiento académico en relación con aquellos que no seguían esa secuencia.

Asimismo, el análisis de secuencias también permitió detectar desviaciones en las secuencias de acciones que podrían indicar dificultades de aprendizaje o problemas específicos que enfrentan los estudiantes. Por ejemplo, si un estudiante generalmente sigue una secuencia de estudio establecida, pero de repente cambia su comportamiento y omite ciertas actividades importantes, esto podría indicar una falta de comprensión o una desconexión con el material.

El enfoque de análisis de secuencias en la MDE ofrece la capacidad de comprender el orden temporal en los eventos de aprendizaje y las prácticas de los estudiantes. Al identificar patrones de comportamiento y desviaciones, este enfoque proporciona información valiosa para optimizar el diseño de los entornos de aprendizaje, así como personalizar la instrucción y brindar apoyo individualizado a los estudiantes (Sanchez-Romero & Llerena-Izquierdo, 2023).

Enfoque sobre detección de patrones

El enfoque sobre la detección de patrones busca identificar relaciones entre diferentes variables y eventos en los datos educativos (Albreiki et al., 2021). Un ejemplo de este enfoque puede ser el análisis de las relaciones y dinámicas de intercambio con los estudiantes y el material de aprendizaje, las respuestas a las preguntas de evaluación y las actividades. La detección de patrones puede revelar relaciones ocultas y proporcionar información valiosa para comprender los factores que repercuten en el rendimiento estudiantil (Naseem et al., 2022).

La capacidad de identificar estos patrones de comportamiento y desviaciones en las secuencias de acciones proporciona a los educadores información valiosa para intervenir y brindar apoyo personalizado a los estudiantes (Weisi & Ahmadi, 2023)(López-Chila, Arteaga-Sotomayor, et al., 2023). Por ejemplo, si se detecta una desviación en la secuencia de acciones de un estudiante, el educador puede proporcionar retroalimentación específica o recursos adicionales para ayudar al estudiante a superar las dificultades (Naseem et al., 2022)(López-Chila, Sumba-Nacipucha, et al., 2023).

Enfoque sobre la MD educativos

El enfoque sobre la MDE describe en cómo se utiliza técnicas de aprendizaje automático y modelado estadístico, y se desarrollan modelos predictivos que estimen el rendimiento futuro de los estudiantes en función de variables y características específicas. Estos modelos pueden ayudar a reconocer a los alumnos que están en posibilidad de bajo rendimiento académico y permitir la intervención temprana para mejorar sus resultados.

Los autores (Romero & Ventura, 2020) en su artículo examina cómo la MDE puede ser utilizada para analizar el rendimiento de los estudiantes. Los autores aplican técnicas de clasificación y regresión para identificar factores influyentes en el rendimiento estudiantil y desarrollar modelos predictivos (Schelfhout et al., 2022).

Los autores (Albreiki et al., 2021) en su trabajo resaltan el uso de técnicas de MD para predecir el rendimiento académico de los estudiantes en un curso de ingeniería. Los autores emplean algoritmos de aprendizaje automático y variables relevantes para construir modelos predictivos y evaluar su precisión.

El artículo (Pallathadka et al., 2023) explora diversas técnicas de MD que se utilizan para predecir el rendimiento académico de los estudiantes. Los autores aplican algoritmos de

clasificación y regresión en conjuntos de datos educativos y comparan la eficacia de diferentes métodos de predicción.

El estudio de (Pallathadka et al., 2023) realizan una revisión sistemática de la literatura sobre la predicción del rendimiento estudiantil utilizando técnicas de MD. Los autores analizan varios enfoques y algoritmos utilizados en investigaciones anteriores y presentan un resumen de los métodos más efectivos y las variables más relevantes.

2.1.5 Metodologías y materiales empleadas en la MD educativos

Se describe las metodologías y materiales empleados en los estudios de MD educativos. Se explorarán técnicas como el adiestramiento automático, la representación gráfica de la información y la minería de texto, así como las plataformas y software utilizados para analizar los datos educativos.

Los autores (López Zambrano et al., 2021) en su artículo realizan una revisión exhaustiva de la literatura sobre la MDE y el análisis del aprendizaje. Se discuten las metodologías y técnicas utilizadas, como el adiestramiento automático, la explotación de texto y la visualización de datos, y se exploran las herramientas y plataformas empleadas en estos estudios.

Los autores (Silva & Fonseca, 2017) proporcionan una visión general de las herramientas utilizadas en la MD educativos. Los autores presentan una breve descripción de varias herramientas de código abierto y comerciales utilizadas para analizar datos educativos, como *WEKA*, *RapidMiner*, *Moodle*, también se discuten sus características y funcionalidades.

Según los autores (Namoun & Alshantqi, 2020) abordan herramientas de minería de texto y visualización utilizadas en la investigación educativa. Los autores presentan varias herramientas y técnicas, como *Topic Modeling*, *Social Network Analysis* y *Text Network Analysis*, que permiten analizar y visualizar datos de texto en contextos educativos.

3. METODOLOGÍA

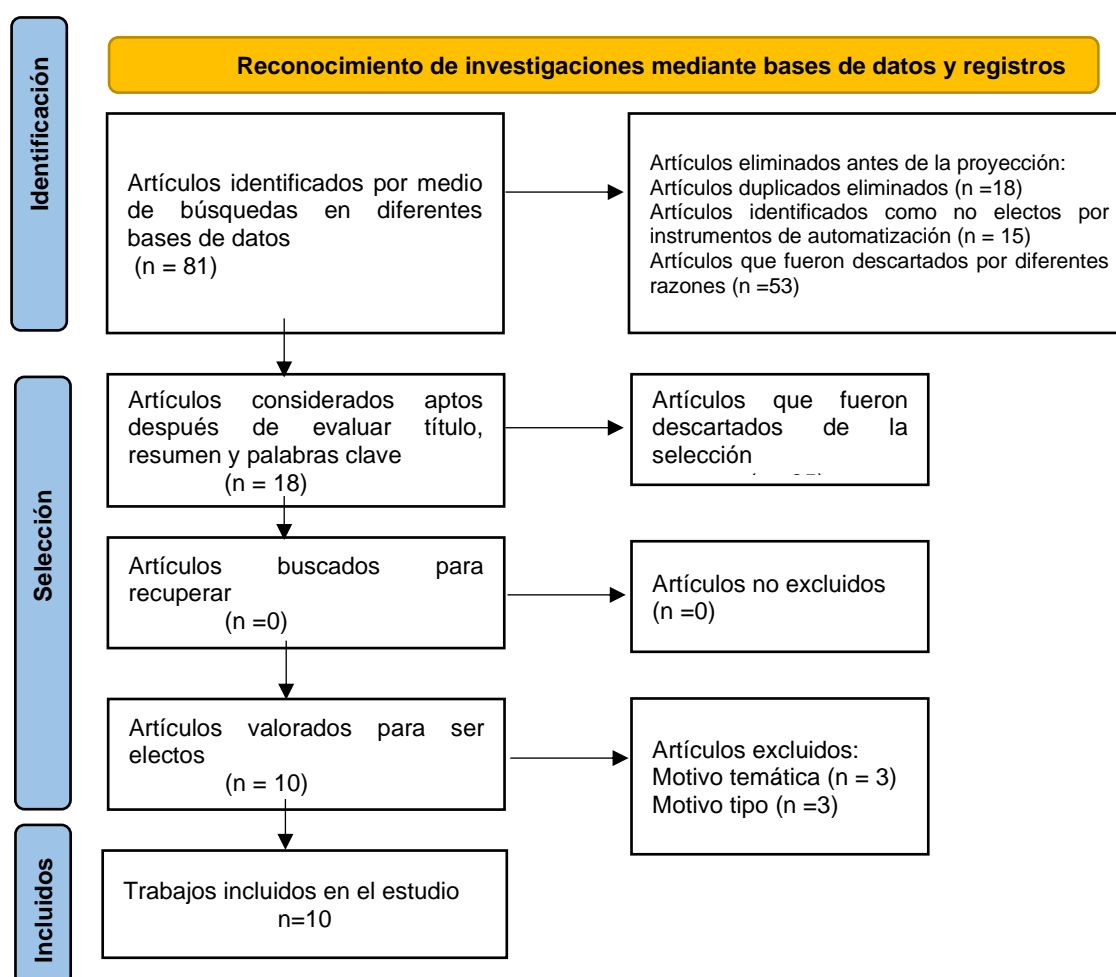
A continuación, se puntualiza la metodología que fue utilizada para realizar el proceso del mapeo ordenado para el seguimiento de la enseñanza de alumnos mediante el uso de MD

educativos. Se detallan los métodos y técnicas empleadas en la recolección de datos y para el estudio de los datos.

3.1. Métodos y técnicas de Recopilación de datos empleadas

Para recopilar los datos necesarios para el mapeo sistemático, se utilizaron diversas fuentes y estrategias. Primeramente, se hizo una exhaustiva indagación en bases de datos académicas, como Scopus, IEEEXplore y ACM Digital Library, utilizando palabras clave relacionadas con el tema de interés. También se consultaron actas de conferencias, revistas científicas y repositorios institucionales para identificar estudios relevantes, ver Fig. 1.

Figura 1. Reconocimiento de investigaciones mediante bases de datos y registros



Igualmente, de la búsqueda en bases de datos, se recurrió al procedimiento de selección de muestra mediante el método de bola de nieve, donde se revisó las referencias bibliográficas de los estudios escogidos y se identificaron otros artículos pertinentes. Esto permitió ampliar la

cobertura y asegurar la inclusión de estudios relevantes que podrían haber quedado fuera del proceso de la indagación precedente.

Asimismo, se utilizaron condiciones de admisión y exclusión para la selección los estudios adecuados. Se consideraron aquellos artículos publicados en los últimos cinco años, escritos en inglés o español, y que abordaran específicamente el tema del seguimiento del proceso de descubrimiento de conocimiento por parte de los alumnos mediante el uso de MD educativos. Se excluyeron estudios que no cumplieran con estos criterios o que no estuvieran disponibles en su versión completa.

3.2. Métodos y técnicas de Análisis de datos

Una vez recopilados los estudios relevantes, se realizó al examinar la información recopilada. En primer lugar, se realizó una lectura exhaustiva de los artículos seleccionados para comprender su contenido y extraer información relevante. Se registraron detalles como los objetivos de cada estudio, las metodologías utilizadas, las sistemáticas de MD empleadas y los efectos obtenidos. Después, se llevó a cabo una síntesis y organización de los hallazgos. Se identificaron patrones, tendencias y temas recurrentes en los estudios analizados. Se realizaron comparaciones y contrastes entre los diferentes enfoques y metodologías utilizadas, destacando similitudes y diferencias.

Además, se utilizó una matriz de análisis para clasificar los estudios según variables como el contexto educativo, las sistemáticas de MD empleadas, los efectos obtenidos y las limitaciones encontradas. Esto permitió tener una visión global y sistemática de la investigación en el campo del seguimiento del aprendizaje de estudiantes mediante MD educativos.

La metodología utilizada en este mapeo sistemático incluyó la recopilación exhaustiva de datos mediante búsqueda en bases de datos y muestreo por bola de nieve, así como el uso de discernimientos de inserción y eliminación para seleccionar los estudios pertinentes. Además, se realizó un análisis detallado de los datos recopilados, mediante la lectura, síntesis y organización de los hallazgos, así como la clasificación y comparación de los estudios seleccionados utilizando una matriz de análisis

3.3. Diseño del mapeo sistemático

En esta sección, se describe el diseño utilizado para realizar el mapeo sistemático sobre el seguimiento del proceso de adquisición de conocimiento por parte de los alumnos mediante el uso de MD educativos. El diseño seleccionado proporciona un marco estructurado y objetivo para realizar una exploración exhaustiva de la bibliografía existente. En este caso, se optó por un enfoque de mapeo sistemático, que se caracteriza por seguir un protocolo predefinido y transparente para la búsqueda, selección y análisis de estudios relevantes.

3.4. Identificación de palabras clave y estrategia de búsqueda

Se necesitó identificar los estudios relevantes, para lo cual se realizó un proceso de búsqueda exhaustiva utilizando palabras clave y términos relacionados con el seguimiento del aprendizaje de estudiantes y la MD educativos. Se utilizó una combinación de palabras clave como "seguimiento del aprendizaje", "MD educativos", "análisis de datos educativos", entre otros. Además, se realizaron realización de exploraciones en conjunto de datos académicas, lugares de almacenamiento digitales y revistas especializadas en educación y tecnología.

3.5. Criterios de admisión y rechazo

En esta etapa, se establecieron condiciones claras para la elección de estudios relevantes. Las condiciones de admisión se basaron en la temática específica del mapeo sistemático, centrándose en investigaciones relacionadas con el seguimiento del proceso de adquisición de conocimiento por parte de los alumnos mediante el uso de MD educativos. Por otro lado, se aplicaron criterios de exclusión para descartar estudios que no cumplieran con los objetivos y alcance del mapeo sistemático, como aquellos que trataban sobre temas no relacionados o que no estaban disponibles en formato completo.

Se determinaron diversas condiciones claras para la elección de estudios relevantes, considerando aquellos publicados a partir del año 2018 en adelante en español y otros idiomas considerados importantes para la investigación. Los criterios de inclusión se basaron en la temática específica del mapeo sistemático, centrándose en investigaciones relacionadas con el seguimiento del proceso de adquisición de conocimiento por parte de los alumnos mediante el uso de MD educativos.

Para la inclusión de estudios, se consideraron los siguientes criterios:

Fecha de publicación: los estudios debían haber sido publicados a partir del año 2018 hasta la fecha actual.

Idioma: se consideraron estudios en español y otros idiomas relevantes para la investigación.

Temática: los estudios debían abordar el seguimiento del proceso de adquisición de conocimiento por parte de los estudiantes utilizando sistemáticas de MD educativos.

Disponibilidad: se incluyeron solo aquellos estudios disponibles en formato completo y accesibles para su análisis.

Por otro lado, se aplicaron condiciones de exclusión para descartar estudios que no cumplían con los objetivos y alcance del mapeo sistemático, como aquellos que trataban sobre temas no relacionados con la MDE o aquellos que no estaban disponibles en formato completo.

La aplicación de estas condiciones de inclusión y condiciones de exclusión que garantizó la elección de estudios pertinentes y actualizados, suministrando un fundamento sólido para realizar el análisis y la obtención de conclusiones significativas en el contexto del mapeo sistemático sobre el seguimiento del proceso de adquisición de conocimiento mediante el uso de MD educativos.

3.6. Proceso de selección de estudios

En esta parte, se realizó un proceso riguroso de selección de estudios relevantes. Se hizo una exploración inicial de las etiquetas y síntesis de los escritos conseguidos a partir de la estrategia de búsqueda. Estos estudios que daban cumplimiento con las condiciones de entrada establecidos fueron seleccionados para una revisión más detallada. Luego, se hizo la revisión completa de los documentos elegidos para asegurar su relevancia y pertinencia con respecto al tema de investigación.

3.7. Extracción de datos

Una vez seleccionados los estudios pertinentes, se llevó a cabo la sustracción de datos relevantes de cada documento. Se diseñó una plantilla o formulario de extracción de datos que incluía variables y aspectos clave relacionados con el seguimiento del aprendizaje de estudiantes y la MD educativos. Se recopilaron datos como los objetivos de cada estudio, las

metodologías que fueron utilizadas, los resultados que fueron obtenidos y las conclusiones alcanzadas.

3.8. Valoración de la excelencia de las investigaciones" o "Análisis de la calidad de los estudios

En esta fase, se efectuó una valoración de la calidad de los documentos seleccionados para probar la confiabilidad y validez de los hallazgos. Se utilizaron criterios de calidad predefinidos, como la rigurosidad metodológica, la confianza en la exhibición de las consecuencias y la relevancia de las conclusiones. Se tuvieron en cuenta las fortalezas y limitaciones de cada estudio, así como la calidad de la evidencia proporcionada.

La metodología de investigación descrita en esta sección proporciona una estructura clara y sistemática para llevar a cabo un mapeo sistemático sobre el seguimiento del aprendizaje de estudiantes utilizando MD educativos. Cada paso del proceso se diseñó cuidadosamente para garantizar la objetividad y la exhaustividad en la recopilación y análisis de la literatura existente.

3.9. Enfoque de recolección de datos por bola de nieve

Además de la estrategia de búsqueda inicial, se hizo la utilización de la recolección de datos por bola de nieve para identificar otros artículos pertinentes. Esta técnica consiste en examinar las referencias bibliográficas de los estudios seleccionados y buscar estudios adicionales que podrían ser relevantes para la investigación. Al revisar las referencias bibliográficas, se identificaron otros artículos que no se encontraron inicialmente durante la búsqueda en bases de datos. Estos artículos adicionales fueron evaluados para determinar su relevancia y se incluyeron en el proceso de selección y análisis de estudios.

El enfoque de recolección de datos por bola de nieve permitió ampliar la búsqueda inicial y garantizar que se incluyeran estudios adicionales que podrían aportar información relevante para el mapeo sistemático sobre el alcance del proceso de adquisición de conocimiento de estudiantes a través del uso de MD educativos. De esta manera, se aseguró la exhaustividad en la recopilación de la literatura existente y se maximizó la cobertura de estudios pertinentes en el campo de investigación.

3.10. Evaluación del rendimiento académico mediante árboles de decisión

Se utiliza la metodología de estructuras de decisión en forma de árbol para evaluar el desempeño rendimiento académico de los alumnos. Los árboles de decisión como una técnica de aprendizaje automático que se basa en la realización de un modelo de árbol que representa las decisiones tomadas a partir de las características o atributos de los datos. A continuación, se detallan los pasos clave involucrados en esta metodología.

3.10.1. Selección de características relevantes

Anteriormente de construir el modelo de estructuras de decisión en forma de árbol, es fundamental realizar una selección de las características más relevantes que se utilizarán como entradas en el modelo. La selección de características tiene como objetivo identificar aquellas variables que poseen una marca más relevante en el cumplimiento académico de los alumnos. Para ello, se pueden aplicar técnicas como análisis de correlación, análisis de componentes principales u otras metodologías específicas de selección de características (Ramaswami et al., 2022).

3.10.2. Preprocesamiento de datos

Una vez seleccionadas las características relevantes, es primordial efectuar un preprocesamiento de los datos antes de crear el modelo de estructuras de decisión en forma de árbol. El preprocesamiento puede incluir la limpieza de los datos, la eliminación de valores diferentes, la normalización de las características o la codificación de los factores categóricos. Estas técnicas permiten asegurar la calidad de los datos y facilitar su procesamiento por parte del modelo de la estructura en forma de árbol (Y. Chen & Zhai, 2023).

3.10.3. Construcción del modelo la estructura de datos en forma de árboles (decisión)

La construcción de la estructura de datos en forma de árbol (decisión) es el paso central en la metodología utilizada para evaluar el rendimiento académico. En la fase, se utiliza el conjunto de datos preprocesados y se aplican algoritmos de aprendizaje automático para crear el AD . Los algoritmos más comunes utilizados son ID3, C4.5 o CART, los cuales buscan dividir los datos de manera óptima en cada nodo del árbol, maximizando la pureza o la ganancia de información (C.-H. Chen et al., 2021).

3.10.4. Entrenamiento y evaluación del modelo

Después de finalizar la creación del modelo de estructura de datos en forma de árbol, se procede a entrenarlo utilizando un grupo de datos de entrenamiento. Durante el entrenamiento, el modelo aprende a tomar decisiones basadas en las particularidades de los alumnos y los resultados académicos agrupados. Luego, se procede a evaluar el desempeño del modelo mediante la utilización de un conjunto de datos de prueba. Se analizan métricas como la exactitud o la precisión, la exhaustividad y la medida F1 se utilizan para medir la eficacia del modelo en la predicción del rendimiento académico (Schelfhout et al., 2022).

3.10.5. Traducción de los resultados

En la traducción de los resultados obtenidos a partir del modelo de AD es esencial para comprender las relaciones y elementos que impactan en el desempeño académico de los alumnos. A través de la visualización del AD, es posible identificar las reglas y condiciones que llevan a la toma de decisiones en cada nodo. Además, se pueden obtener medidas de importancia de características para determinar cuáles variables ejercen una influencia más significativa en el rendimiento académico (Yağcı, 2022).

La metodología de evaluación del rendimiento académico mediante árboles de decisión comprende la elección de factores notables, el preprocesamiento de datos, la construcción del modelo de AD, el entrenamiento y evaluación del modelo, además, se analiza la interpretación de los resultados obtenidos. Estos pasos permiten analizar y comprender las variables que tienen una impresión en el ejercicio académico de los alumnos. y tomar decisiones informadas para mejorar su desempeño.

4. RESULTADOS

4.1. Características generales de los estudios incluidos

En esta sección, se presentarán los aspectos comunes de los estudios incluidos en el mapeo sistemático sobre el seguimiento del proceso de adquirir conocimiento a través el uso de MD educativos. Se analizarán aspectos como el año que fue publicado, el idioma, el tipo de documento (empírico, teórico, etc.), la muestra utilizada y el contexto educativo en donde se llevó a cabo la investigación.

4.2. Enfoques y técnicas utilizadas en la MD educativos

Lista de enfoques y técnicas

- Aprendizaje automático: se utilizan algoritmos y modelos matemáticos para clasificar, predecir o agrupar estudiantes según su rendimiento académico o necesidades específicas. Algunos métodos comunes incluyen árboles de decisión, redes neuronales artificiales, algoritmos de vecinos más cercanos y clasificadores bayesianos.
- Modelado estadístico: se emplean métodos estadísticos para examinar y modelar las concordancias entre variables y características de los estudiantes. Por ejemplo, técnicas de regresión para determinar el predominio de varios valores independientes en el desempeño académico, análisis de varianza y pruebas de hipótesis.
- Algoritmos de clasificación: se utilizan para asignar a los alumnos en categorías o grupos específicos según sus factores y comportamientos. Algoritmos como Naïve Bayes, regresión logística o máquinas de direcciones de soporte se aplican para catalogar a los estudiantes en cualidades como cima, intermedio o inferior rendimiento académico.
- Agrupamiento: se utiliza para identificar patrones y similitudes entre los estudiantes, agrupándolos en clústeres con características similares. Métodos como k-means, agrupamiento jerárquico y agrupamiento basado en densidad se utilizan para este fin.
- Estimación: se emplean técnicas de modelado y estudio de datos para pronosticar el rendimiento futuro de los estudiantes en relación de las variables y características disponibles. Algoritmos de regresión, análisis de series temporales u otros modelos predictivos pueden ser utilizados para la estimación.

Estos enfoques y técnicas se aplican en el campo de la exploración de datos educativos con fines de minería. el seguimiento del aprendizaje de los estudiantes, permitiendo obtener información valiosa para mejorar la educación y brindar un apoyo más personalizado.

4.3. Metodologías y herramientas empleadas en el seguimiento del aprendizaje

En el seguimiento del aprendizaje en este proyecto, se utilizaron diversas metodologías y herramientas para llevar a cabo el monitoreo y análisis de los estudiantes. A continuación, se describen las principales metodologías y herramientas empleadas:

Metodologías:

- Estudios longitudinales: Se utilizó esta metodología para recopilar datos a lo largo del tiempo y realizar un seguimiento constante del desempeño de los estudiantes. Al estudiar los datos históricos, se pueden determinar modelos, tendencias y cambios en el desempeño a lo largo del tiempo.
- Estudios experimentales: Esta metodología se empleó para realizar experimentos controlados y valorar el hecho de que ciertos factores o intervenciones en el aprendizaje de los estudiantes. A través de pruebas y mediciones, se pudo determinar la eficacia de diferentes enfoques de enseñanza o estrategias de intervención.
- Estudios de observación: Esta metodología implicó la observación directa de los estudiantes en contextos de aprendizaje para recopilar datos cualitativos y cuantitativos sobre su comportamiento, participación y desempeño. La observación permitió obtener información ampliada y en tiempo real del aprendizaje de los estudiantes.

Herramientas tecnológicas:

- Plataformas de aprendizaje en línea: Se utilizaron plataformas en línea para recopilar y almacenar datos relacionados con el aprendizaje de los estudiantes. Estas plataformas proporcionaron un entorno virtual donde los estudiantes podían acceder a materiales de estudio, participar en actividades interactivas y realizar evaluaciones.
- Procedimientos de gestión de datos: Se emplearon sistemas de gestión de datos para almacenar y organizar los datos recopilados de los estudiantes. Estos sistemas facilitaron la gestión, el análisis y la extracción de información importante para el seguimiento del aprendizaje.
- Software de análisis de datos: Se utilizaron herramientas de software especializadas en estudio de reseñas para procesar, visualizar y recolectar información significativa de los datos recopilados. Estas herramientas permitieron realizar análisis estadísticos, identificar patrones y tendencias, y generar informes acerca del desempeño académico de los estudiantes.

En conjunto, estas metodologías y herramientas proporcionaron un enfoque integral con el motivo de monitorear el avance del aprendizaje de los estudiantes. Permitieron recopilar datos

relevantes, realizar análisis en profundidad y obtener información valiosa para optimizar la instrucción y el proceso de adquisición de conocimiento.

4.4. Variables y características consideradas en los estudios

En los estudios de seguimiento del aprendizaje mediante MD educativos, se consideran diversas variables y características. Algunas de las más comunes son el rendimiento académico, el tiempo destinado al estudio, la participación en actividades de aprendizaje, las interacciones en entornos virtuales, características socioeconómicas y demográficas, y motivación y actitudes hacia el aprendizaje.

4.5. Resultados y hallazgos principales

Los hallazgos y resultados obtenidos de los estudios de MDE muestran que se encuentra una relación significativa entre variables como el tiempo dedicado al estudio, la participación en actividades de aprendizaje y el rendimiento académico. También se han reconocido factores de riesgo y éxito, los hábitos de comportamiento de los estudiantes, y la detección anticipada de problemas de aprendizaje. Estos resultados han mejorado la toma de decisiones pedagógicas y brindado una visión más informada para mejorar el progreso de la enseñanza y del aprendizaje.

Resultados de la revisión de la Literatura Método Prisma

A continuación, se muestra los resultados del análisis de la literatura utilizando el método PRISMA, ver Fig. 2.

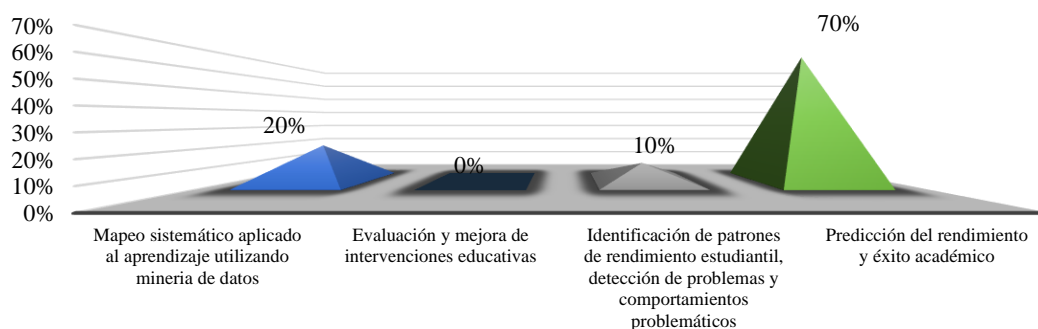
Figura 2. Resultados Método PRISMA

Item	Año de publicación	Artículos	Título del artículo	Número de citas	Temáticas y objetivos																
					Mapeo sistemático aplicado al aprendizaje utilizando minería de datos	Evaluación y mejora de intervenciones educativas	Identificación de patrones de rendimiento estudiantil, detección de problemas	Detección del rendimiento y éxito académico	Identificación de las limitaciones y desafíos asociados al uso de técnicas de minería de datos	Formulación del problema y datos implicados	Limitaciones encontradas	Propuestas y metodología aplicada	Solución encontrada, resultados y desafíos								
1	2018	Mapeo sistemático de la literatura sobre evaluación docente (2013-2017)		1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	2020	Minería de datos educativos: análisis del desempeño de estudiantes de ingeniería en las pruebas saber-pro		0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0
3	2022	Procedimiento de agrupación de estudiantes según riesgo de abandono para mejorar la gestión estudiantil en educación superior		0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0
4	2020	Comparación de técnicas de minería de datos para identificar indicios de deserción estudiantil, a partir del desempeño académico		0	0	0	1	0	0	1	0	0	0	0	1	0	0	0	0	0	0
5	2022	Minería predictiva aplicada al descubrimiento de factores asociados al desempeño en la competencia de lenguaje de los estudiantes de básica superior		0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0
6	2022	Modelo adaptativo de clasificación de profesiones en sistemas de orientación vocacional		0	0	0	1	0	0	0	0	1	0	1	0	0	0	0	0	0	0
7	2019	Caracterización de los estudiantes de una institución de educación superior mediante big data		0	0	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0
8	2021	Rendimiento académico de estudiantes en Educación Superior: predicciones de factores influyentes a partir de árboles de decisión		0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0
9	2023	Inteligencia artificial y aprendizaje colaborativo asistido por computadora en programación: un estudio de mapeo sistemático		0	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0
10	2020	Resolución de problemas algorítmicos y objetos de aprendizaje: una revisión de la literatura		1	0	0	0	1	0	0	0	0	0	1	0	0	1	0	0	0	0

Temáticas y objetivos

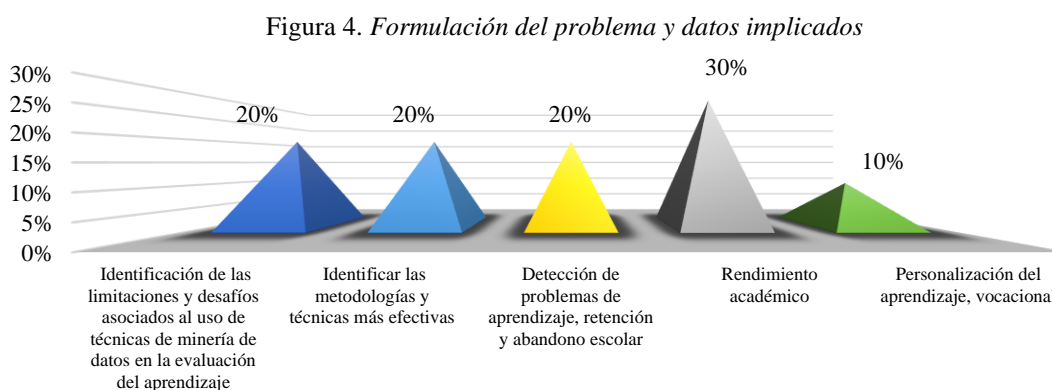
Con respecto a la temática se encontró que el 70% de artículos tenían la temática de Predicción del rendimiento y éxito académico, 10% sobre Identificación de patrones de rendimiento estudiantil, detección de problemas y comportamientos problemáticos. El 20% con respecto al Mapeo sistemático aplicado al aprendizaje utilizando MD y 0% la Evaluación y mejora de intervenciones educativas, ver Fig. 3.

Figura 3. Temáticas y objetivos



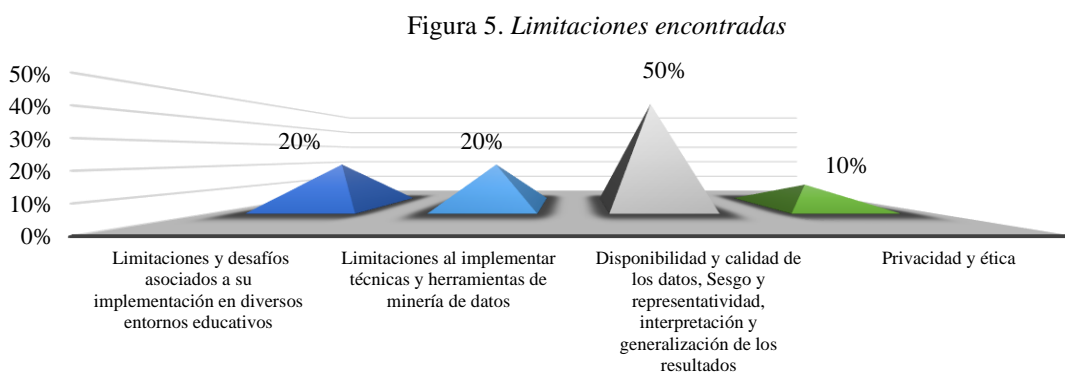
Formulación del problema y datos implicados

Con respecto a la formulación del problema y datos implicados, se encontró el 30% sobre Rendimiento académico, se encontró el 20% sobre Identificación de las limitaciones y desafíos asociados al empleo de sistemáticas de MD en la valoración del aprendizaje, Identificar las metodologías y técnicas más efectivas y Detección de problemas de aprendizaje, retención y abandono escolar. Se encontró 10% sobre Personalización del aprendizaje, vocacional, ver Fig. 4.



Limitaciones encontradas

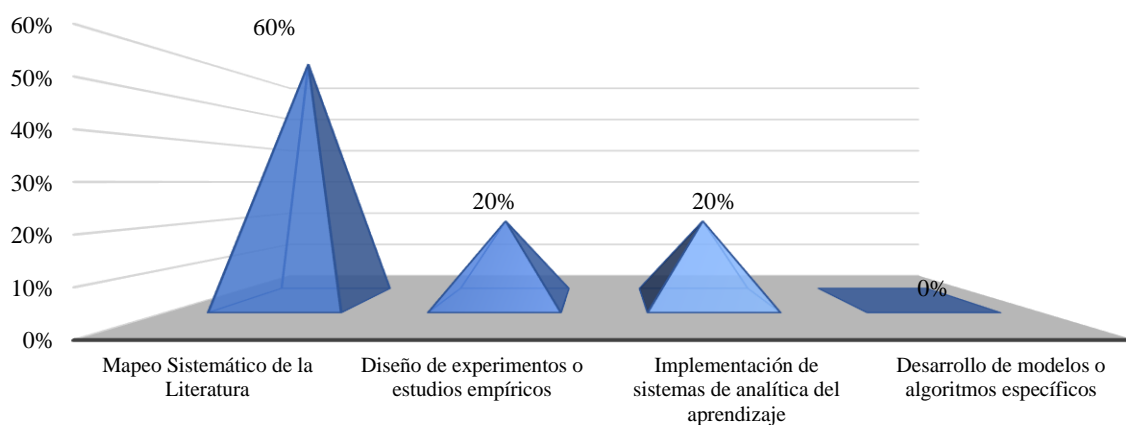
Se encontró el 50% sobre Disponibilidad y calidad de las identificaciones, Sesgo y representatividad, interpretación y generalización de las derivaciones. El 20% sobre Limitaciones y desafíos asociados a su implementación en diversos entornos educativos y Limitaciones al implementar instrumentales y metodologías de MD. Y el 10% sobre Privacidad y ética, ver Fig. 5.



Propuestas y metodología aplicada

Se encontró 60% sobre Mapeo Sistemático de la Literatura, el 20% sobre Diseño de experimentos o estudios empíricos e Implementación de sistemas de analítica del aprendizaje y el 0% sobre el Desarrollo de modelos o algoritmos específicos, ver Fig. 6.

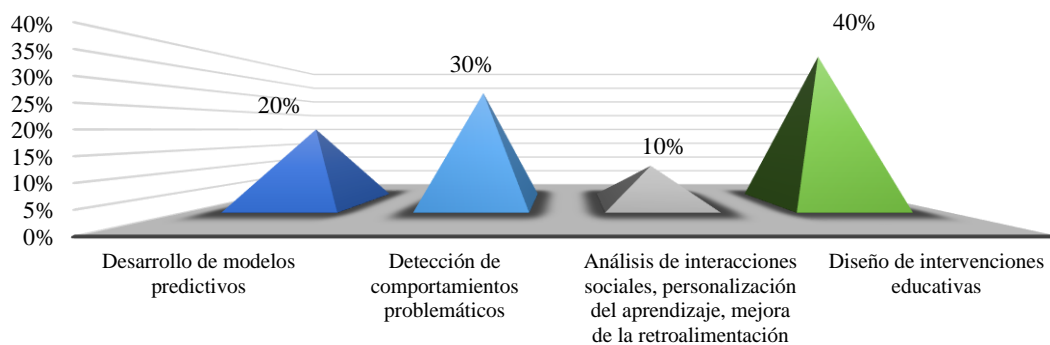
Figura 6. *Propuestas y metodología aplicada*



Solución encontrada, resultados y desafíos

Se encontró el 40% sobre Diseño de intervenciones educativas, el 30% sobre Detección de comportamientos problemáticos, el 20% sobre el Perfeccionamiento de patrones predictivos y el 10% sobre el Examen de interacciones sociales, personalización del aprendizaje, mejora de la retroalimentación, ver Fig. 7.

Figura 7. *Solución encontrada, resultados y desafíos*

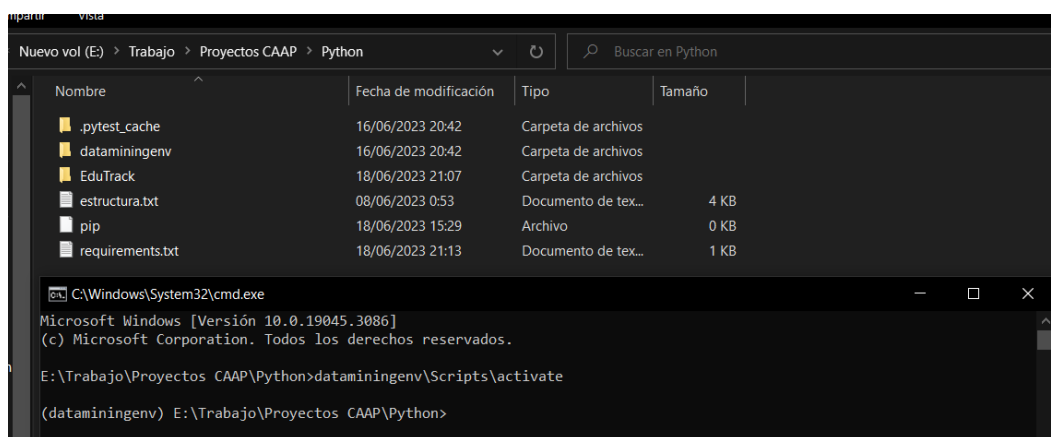


Resultados de la implementación

El proyecto involucra la configuración del entorno, la estructuración de las historias de usuario, la ejecución de interfaces, la MD basada en árboles de decisión, las pruebas unitarias y el análisis de datos provenientes de archivos Excel en un lugar donde se guarde los datos. En donde se utilizaron modelos entidad-relación para las bases de datos utilizadas.

- a. Para poder ejecutar la aplicación, pruebas o comprobación de cualquier archivo se debe gestionar el entorno virtual con el comando `dataminingenv\Scripts\activate` en el directorio raíz antes del env, ver Fig. 8.

Figura 8. Gestionar entorno virtual



- b. Una vez dentro del entorno se ejecuta el comando `pip install -r requirements.txt` para que se instalen todas las dependencias y bibliotecas que se usaron en el proyecto, ver Fig. 9.

Figura 9. Instalación dependencias

```
(dataminingenv) E:\Trabajo\Proyectos CAAP\Python>pip install -r requirements.txt
Requirement already satisfied: awsebcli==3.20.5 in c:\users\freddy bazante\appdata\roaming\python\python311\site-packages (from -r requirements.txt (line 1)) (3.20.5)
Requirement already satisfied: boto3==1.29.81 in c:\users\freddy bazante\appdata\roaming\python\python311\site-packages (from -r requirements.txt (line 2)) (1.29.81)
Requirement already satisfied: botores==2.8.2 in c:\users\freddy bazante\appdata\roaming\python\python311\site-packages (from -r requirements.txt (line 3)) (2.8.2)
Requirement already satisfied: certifi==2022.12.7 in c:\users\freddy bazante\appdata\roaming\python\python311\site-packages (from -r requirements.txt (line 4)) (2022.12.7)
Requirement already satisfied: charset-normalizer==2.0.12 in c:\users\freddy bazante\appdata\roaming\python\python311\site-packages (from -r requirements.txt (line 5)) (2.0.12)
Requirement already satisfied: colorama==0.4.3 in c:\users\freddy bazante\appdata\roaming\python\python311\site-packages (from -r requirements.txt (line 6)) (0.4.3)
Requirement already satisfied: contourpy==1.1.0 in c:\program files\python311\lib\site-packages (from -r requirements.txt (line 7)) (1.1.0)
```

- c. Una vez instalado pasaremos a comprobar el funcionamiento de cada parte del proyecto para esto tenemos una estructura de directorios, ver Fig. 10.

Figura 10. Estructura de directorios

```

Edutrack
| InterfazFinal.py
|
|
+---Historia_de_Usuario_1
| +---codigo_fuente
| | | main.py
| | | mineria_datos.py
| | | modelos.py
| |
| +---datos
| |   base_de_datos.sql
| |
| +---documentacion
| |   documentacion_adicional.md
| |   estructura_directorios.txt
| |   Historias_de_usuario.docx
| |   README.md
| |
| +---Interfaces
| |   maininterface.py
| |
| \---pruebas
| |   test_mineria_datos.py
| |   test_modelos.py
| |
+---Historia_de_Usuario_2
| +---codigo_fuente
| | | main.py
| | | mineria_datos.py
| | | modelos.py
| |
| +---datos
| |   base_de_datos.sql
| |
| +---documentacion
| |   documentacion_adicional.md
| |   estructura_directorios.txt
| |   Historias_de_usuario.docx
| |   README.md
| |
| +---Interfaces
| |   maininterface.py
| |
| \---pruebas
| |   test_mineria_datos.py

```

```

|   | test_modelos.py
|
+---Historia_de_Usuario_3
| +---codigo_fuente
| | | main.py
| | | mineria_datos.py
| | | modelos.py
| |
| +---datos
| |   base_de_datos.sql
| |
| +---documentacion
| |   documentacion_adicional.md
| |   estructura_directorios.txt
| |   Historias_de_usuario.docx
| |   README.md
| |
| +---Interfaces
| |   maininterface.py
| |
| \---pruebas
| |   test_mineria_datos.py
| |   test_modelos.py
|
+---Historia_de_Usuario_4
| +---codigo_fuente
| | | main.py
| | | mineria_datos.py
| | | modelos.py
| |
| +---datos
| |   base_de_datos.sql
| |
| +---documentacion
| |   documentacion_adicional.md
| |   estructura_directorios.txt
| |   Historias_de_usuario.docx
| |   README.md
| |
| +---Interfaces
| |   maininterface.py
| |
| \---pruebas
| |   test_mineria_datos.py
| |   test_modelos.py
|
\---Historia_de_Usuario_5
  +---codigo_fuente

```

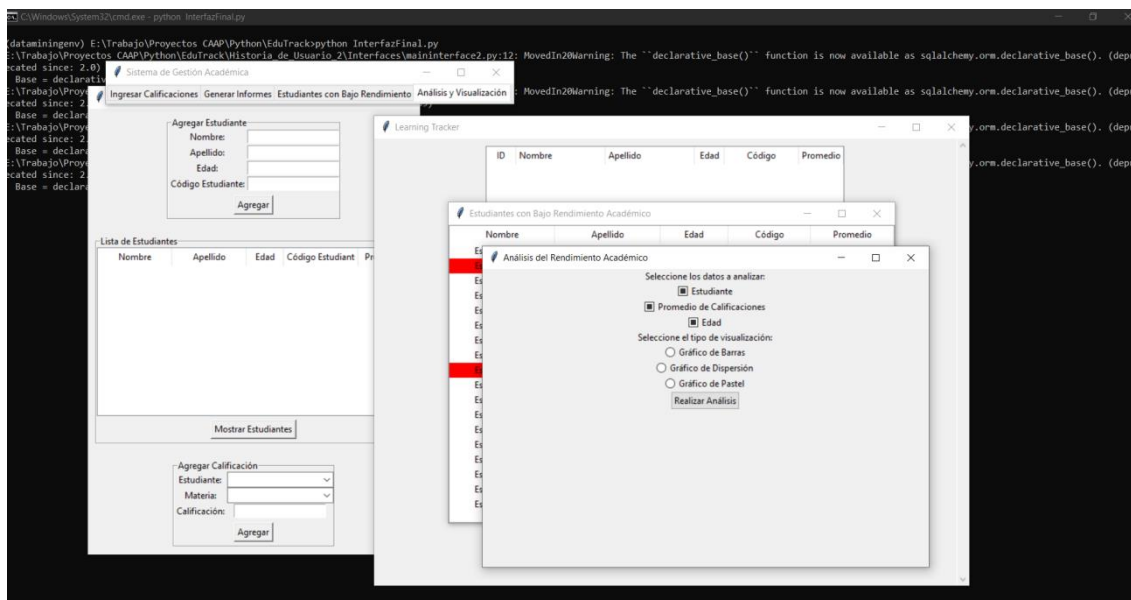
```

| | main.py
| | mineria_datos.py
| | modelos.py
|
+---datos
|   base_de_datos.sql
|
+---documentacion
|   documentacion_adicional.md
|   estructura_directorios.txt
|   Historias_de_usuario.docx
|   README.md
|
+---Interfaces
|   maininterface.py
|
\---pruebas
|   test_mineria_datos.py
|   test_modelos.py

```

- d. Comprobaremos la interfaz principal llamada InterfazFinal.py con el comando Python InterfazFinal.py, ver Fig. 11.

Figura 11. Interfaz principal



- e. Después de interactuar con cada módulo pasaremos a comprobar el modelo de MD los cuales se basan en árboles de decisión ingresando a cada directorio de historia de usuario y dentro de estas en el directorio codigo_fuente. Ejecutaremos el comando python main.py y con esto obtendremos las predicciones con respecto al rendimiento, ver Fig. 12.

Figura 12. Predicciones de rendimiento

```
(dataminingenv) E:\Trabajo\Proyectos CAAP\Python\EduTrack\Historia_de_Usuario_2\codigo_fuente>python main.py
Predicciones: [10. 10. 10. 10. 0. 10. 8. 10. 6.25 10. 10. 10.
10. 10. 10. 10. 0. 10. 10. 10. 10. 10. 10.
6.25 10. 10. 10. 10. 10. 10. 10. 10. 10. 10.
10. 0. 10. 10. 10. 10. 10. 10. 10. 10. 10.
10. 10. 10. 10. 10. 10. 8. 10. 10. 10. 6.63
10. 10. 10. 10. 10. 10. 10. 10. 10. 10.
10. 10. 10. 10. 10. 10. 10. 10. 10. 10.
10. 10. 10. ]
Puntaje de evaluación: 1.0

(dataminingenv) E:\Trabajo\Proyectos CAAP\Python\EduTrack\Historia_de_Usuario_2\codigo_fuente>
```

- f. Después de ver las predicciones del modelo pasaremos a realizar las pruebas unitarias para ello regresamos al directorio principal de cada historia de usuario y ejecutamos el comando `pytest -v pruebas/`. Esto nos indicará las pruebas realizadas y de estar correctas tendremos un estado PASSED, ver Fig. 13.

Figura 13. Pruebas unitarias

```
(dataminingenv) E:\Trabajo\Proyectos CAAP\Python\EduTrack\Historia_de_Usuario_2>pytest -v pruebas/
platform win32 -- Python 3.11.3, pytest-7.3.1, pluggy-1.0.0 -- C:\Program Files\Python311\python.exe
cachedir: .pytest_cache
rootdir: E:\Trabajo\Proyectos CAAP\Python\EduTrack\Historia_de_Usuario_2
collected 8 items

pruebas/test_mineria_datos.py::TestMineriaDatos::test_obtener_datos PASSED [ 12%]
pruebas/test_mineria_datos.py::TestMineriaDatos::test_obtener_datos_columnas_correctas PASSED [ 25%]
pruebas/test_mineria_datos.py::TestMineriaDatos::test_obtener_datos_tipos_correctos PASSED [ 37%]
pruebas/test_mineria_datos.py::TestMineriaDatos::test_obtener_datos_valores_no_nulos PASSED [ 50%]
pruebas/test_modelos.py::TestModelos::test_entrenar_modelo PASSED [ 62%]
pruebas/test_modelos.py::TestModelos::test_entrenar_modelo_no_datos_vacios PASSED [ 75%]
pruebas/test_modelos.py::TestModelos::test_evaluate_modelo_no_modelo_vacio PASSED [ 87%]
pruebas/test_modelos.py::TestModelos::test_predecir PASSED [100%]

===== 8 passed in 1.94s =====

(dataminingenv) E:\Trabajo\Proyectos CAAP\Python\EduTrack\Historia_de_Usuario_2>
```

- g. Hay que tomar algunas consideraciones en el código de la interfaz principal

```
def ingresar_calificaciones():
    # lógica para ingresar calificaciones
    subprocess.Popen(['python', E:\Trabajo\Proyectos
CAAP\Python\EduTrack\Historia_de_usuario2\Interfaces\maininterface2.py])
```

Aquí debemos modificar el directorio de acuerdo con la ubicación de la aplicación

- h. También debemos aclarar que se usó mysql como manejador de base de datos con el gestor Workbench. Para cargar la base de datos usaremos el script que se encuentra en cada historia de usuario. Para la historia de usuario 1 es un script diferente mientras que para las otras historias de usuario es la misma.
- i. En la parte final del proyecto se solicitó que se prueben algunos datos almacenados en archivos xlsx de Excel para lo cual se procedió a crear una base de datos llamada `learning_tracker_final`, ver Fig. 14.

Figura 14. Base de datos



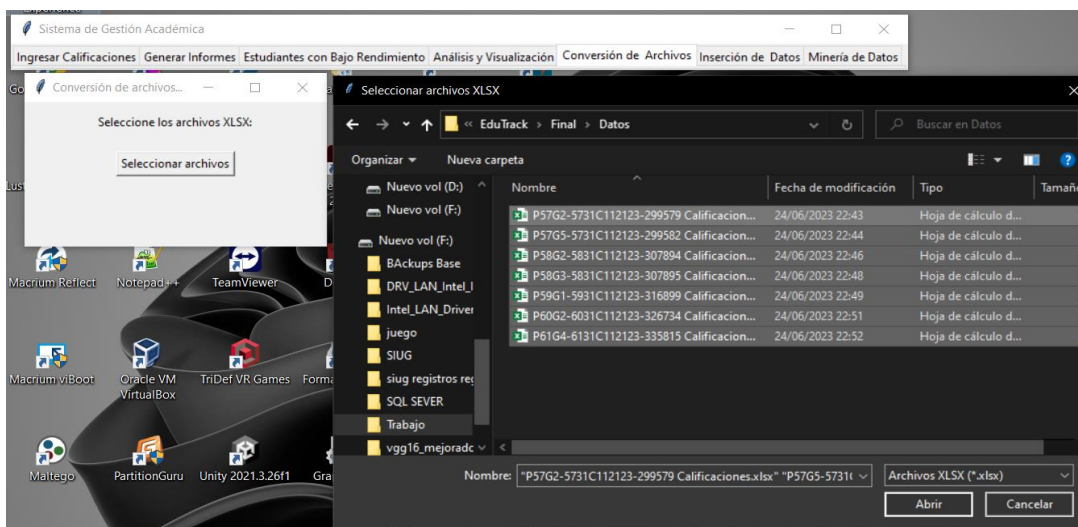
- j. Después de esto se procedió a analizar los archivos xlsx, para ello se abre cada archivo y se toma los datos esenciales a analizar y se los normaliza para poder interpretarlos con el código. Las columnas que se tomaron en cuenta son Total Parcial Final Aprovechamiento (Real), Cuestionario: Examen Parcial Final (Real), Cuestionario_Examen_Final, Total PF Examen (Real), Total Parcial Final (Real), Total del curso (Real), ver Fig.15.

Figura 15. Columnas seleccionadas para el análisis

Total Parcial Final Aprovechamiento (Real)	Cuestionario: Examen Parcial Final (Real)	Cuestionario_Examen_Final	Total PF Examen (Real)	Total Parcial Final (Real)	Total del curso (Real)
20	9	9	29	76	16

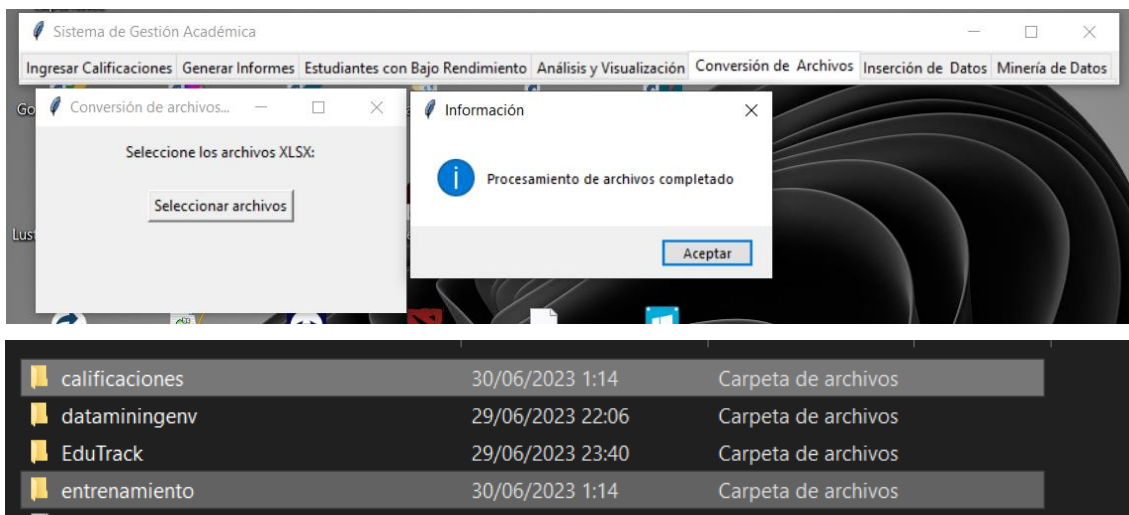
- k. Una vez normalizados los archivos se procede a convertirlos y almacenarlos en una carpeta para entrenamiento y otra para evaluar datos que se llama calificaciones, ver Fig. 16.

Figura 16 Preparación entrenamiento



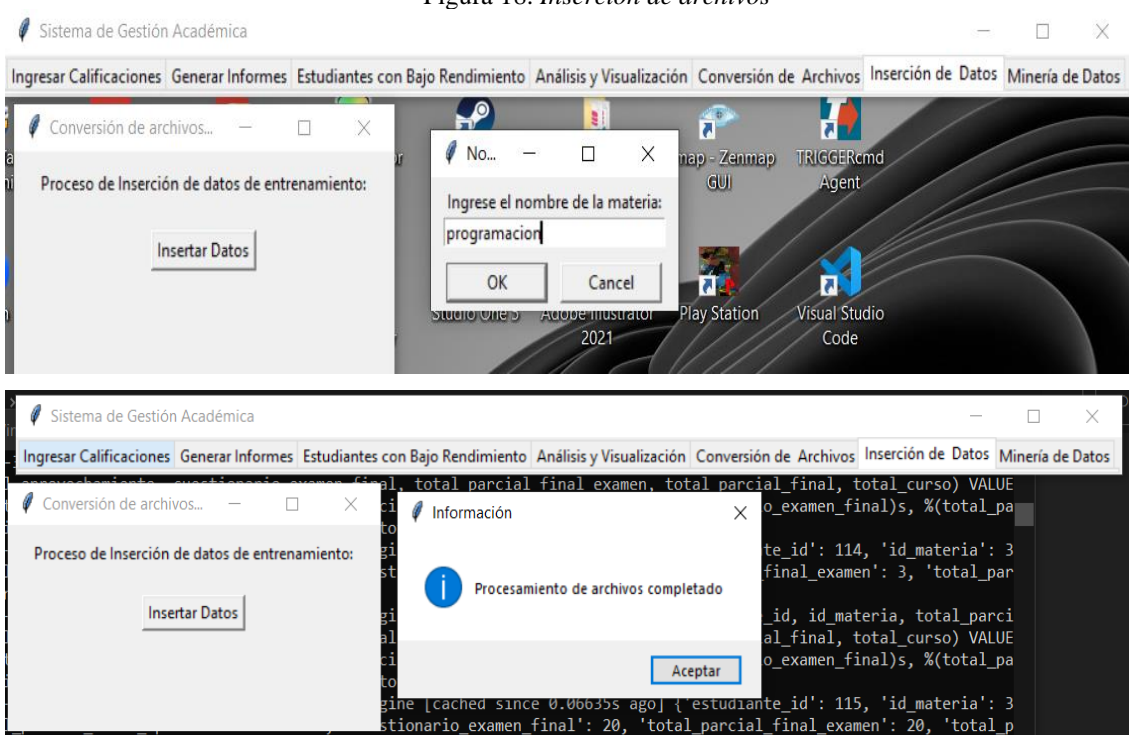
Una vez seleccionados los archivos para la preparación del entrenamiento. Se muestra en pantalla que el procesamiento de archivos fue completado, ver Fig. 17.

Figura 17. Preparación entrenamiento



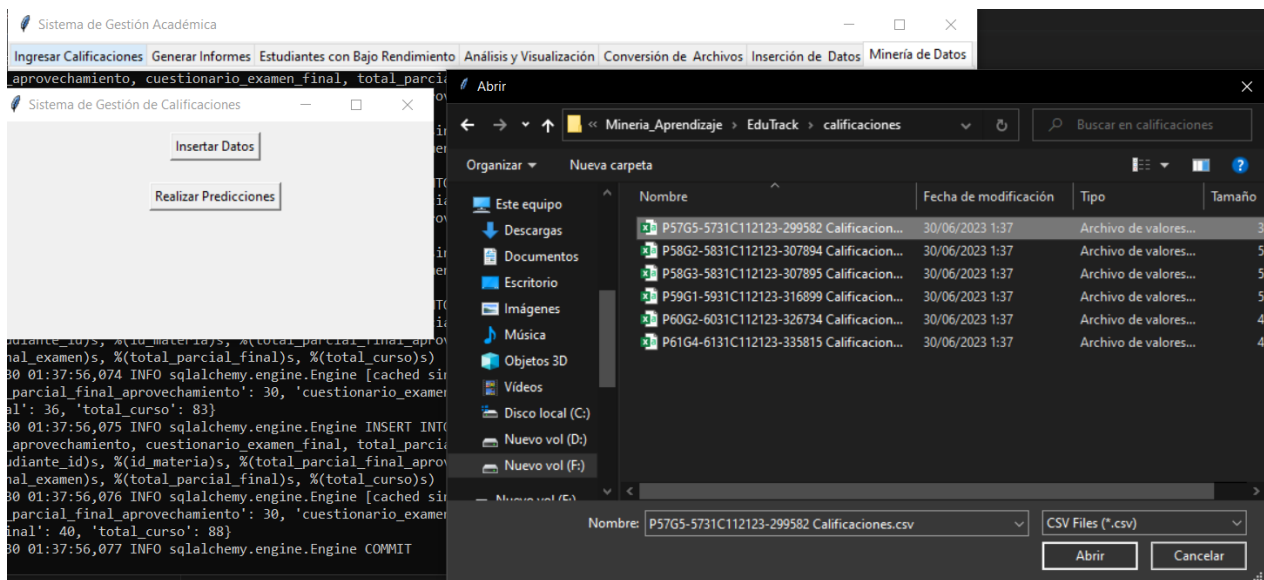
1. Luego se procede a insertar los datos en la base en la pestaña Inserción de archivos. Aquí se pedirá la materia, ver Fig. 18.

Figura 18. *Inserción de archivos*



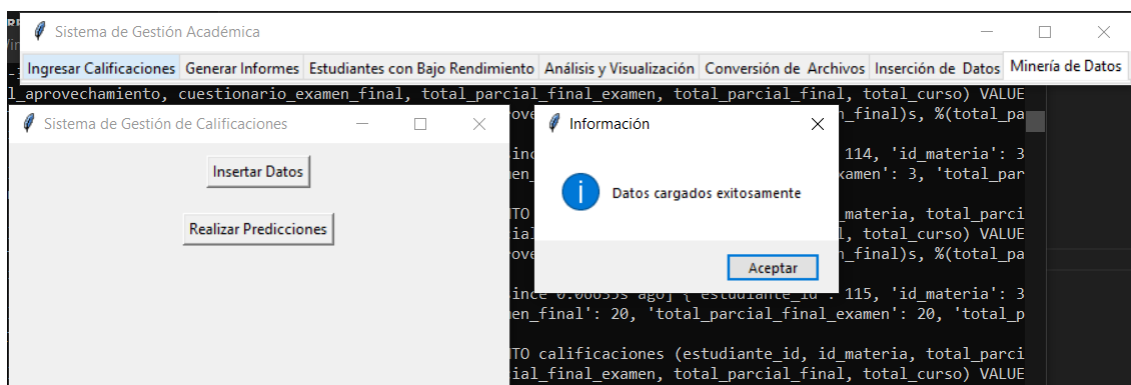
- m. Ya insertados los datos de entrenamiento se procede a ejecutar el modelo de minería en el que se evalúan datos de prueba y obtendremos datos como predicciones y el rendimiento del algoritmo mediante el error cuadrático medio, ver Fig. 19.

Figura 19. Ejecución de entrenamiento



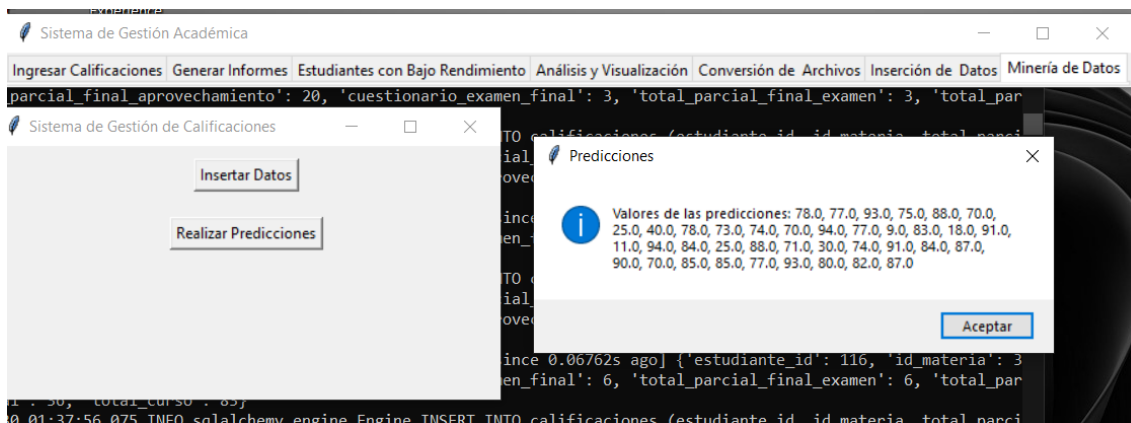
Después de cargar los archivos en Excel, se muestra en pantalla que los datos fueron cargados exitosamente, ver Fig. 20.

Figura 20. Datos cargados de entrenamiento



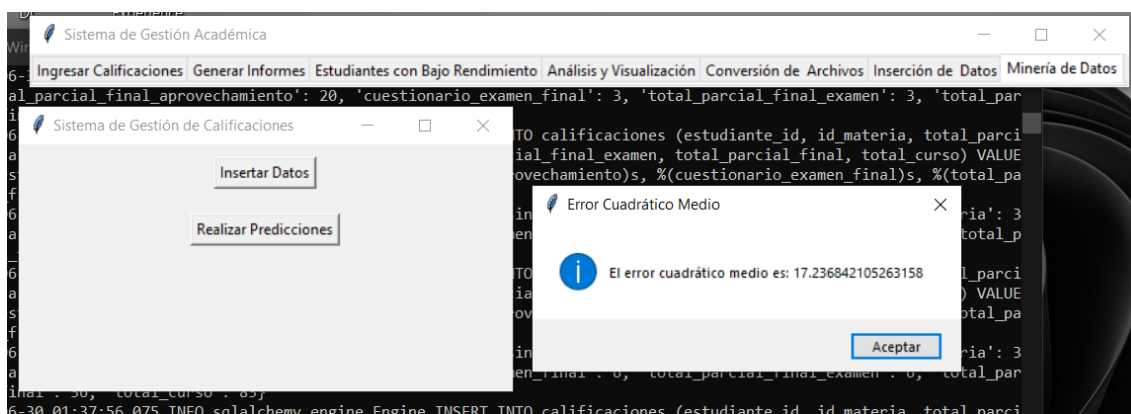
Después de que los datos fueron cargados exitosamente, se ejecuta el proceso del modelo y se muestran en pantalla los resultados del entrenamiento, ver Fig. 21.

Figura 21. Resultados de entrenamiento



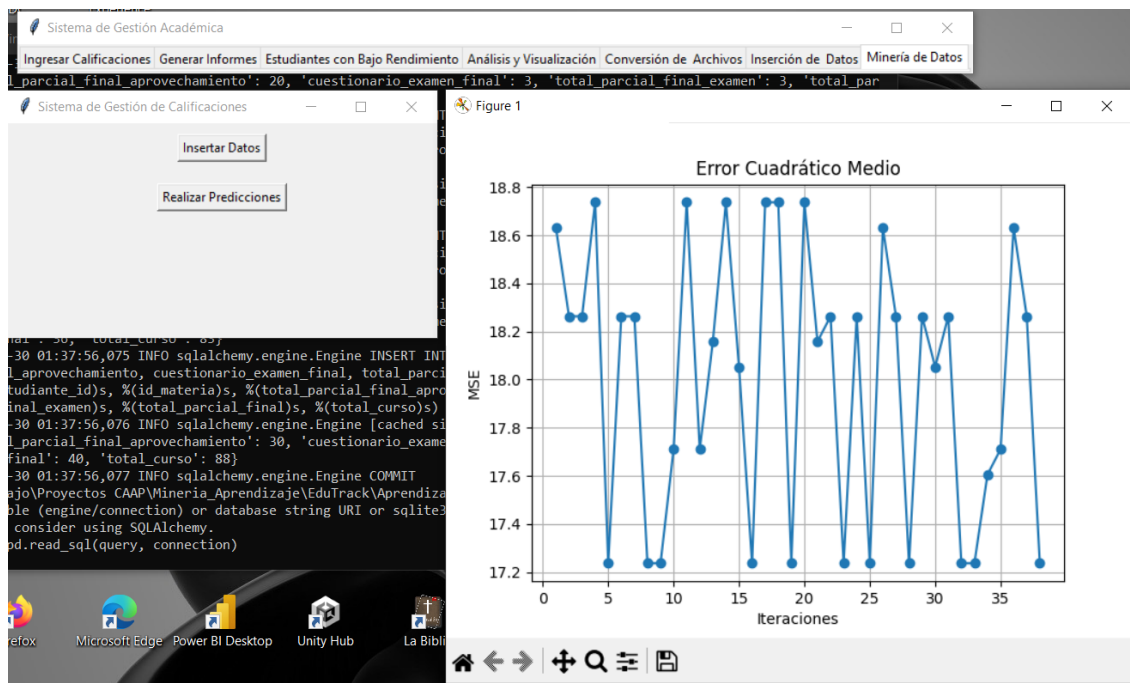
Después de mostrar en pantalla los resultados del entrenamiento (valores de las predicciones) se realiza el proceso para calcular el error cuadrático medio y se muestra en pantalla, ver Fig. 22.

Figura 22. Mensajes de error cuadrático medio



Luego de mostrar en pantalla el error cuadrático medio se presenta el gráfico de error cuadrático de error medio en donde se visualiza los datos en función de las iteraciones, en el eje x las iteraciones, y en el eje el valor del MSE, ver Fig. 23.

Figura 23. Gráfico de error cuadrático de error medio



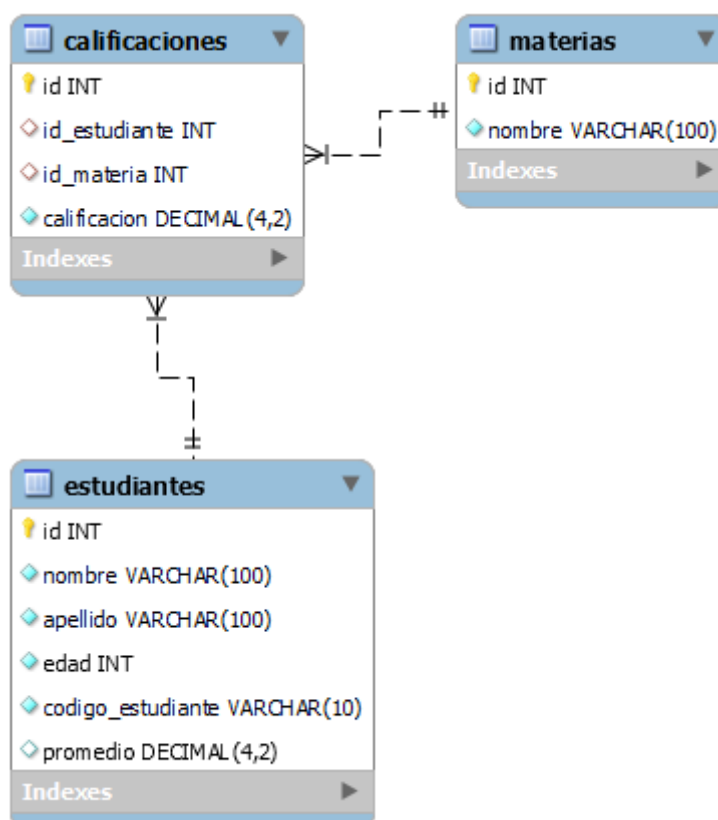
A continuación, se muestra los modelos entidad relación que fueron diseñadas para el modelo, primeramente, se muestra la tabla estudiante con los atributos, ver Fig. 24.

Figura 24. Tabla Estudiante de Base de datos learning tracker

students	
id	INT
name	VARCHAR(50)
age	INT
grade	INT
average_grade	FLOAT
Indexes	

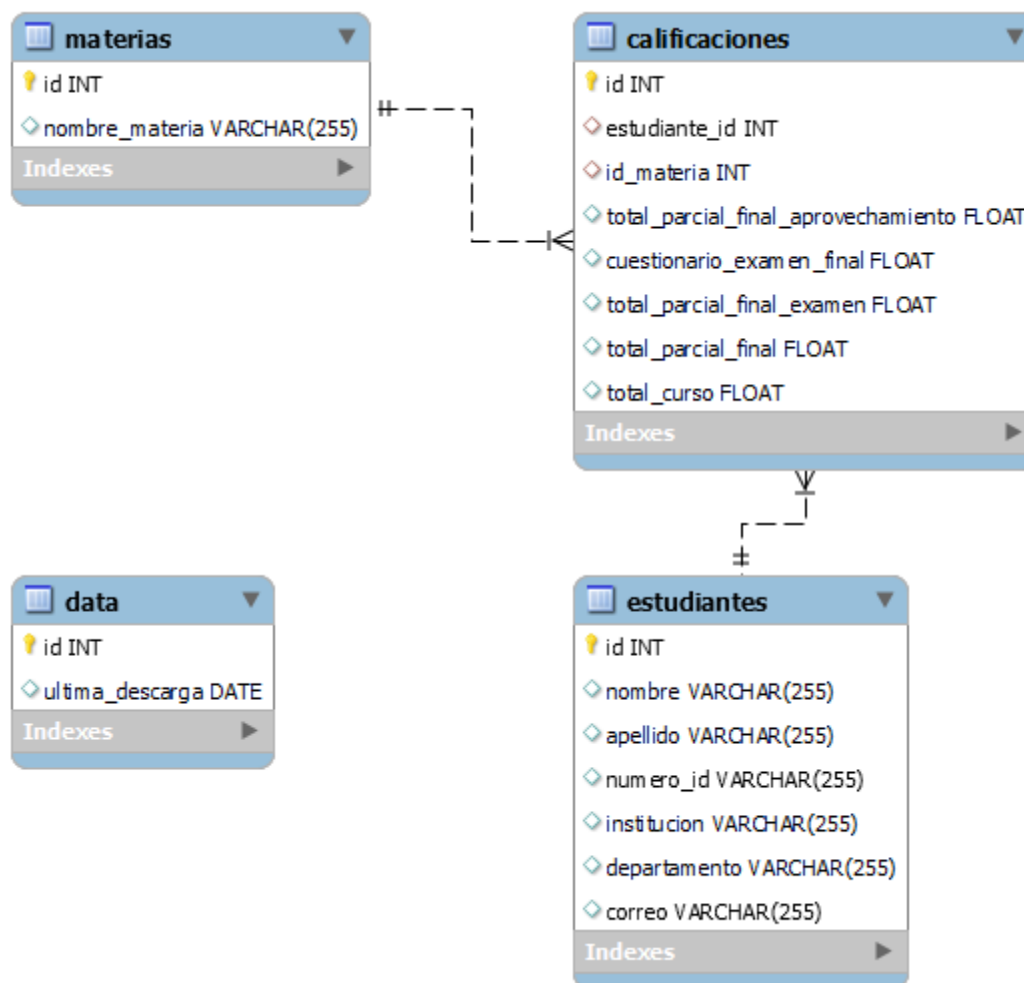
Posteriormente se procedió añadir las demás tablas y las relaciones en la versión inicial de la base de datos del modelo, ver Fig. 25.

Figura 25. Base de datos learning tracker h2



En la base de datos definitiva del modelo se añadió campos a las tablas, especialmente se añadió campos a la tabla calificaciones en donde se colocaron las notas del estudiante en los parciales, exámenes y nota final, así mismo se añadió campos en la tabla estudiantes como institución y departamento, ver Fig. 26.

Figura 26. Base de datos learning tracker definitivo



Repositorio en Github

A continuación, se incluye el enlace del repositorio en GitHub, <https://github.com/JorgeLuizP/Mineria>. Se incluye todo el código utilizado en la investigación.

5. DISCUSIÓN

5.1. Tendencias y patrones identificados en la literatura

Artículo "MD educativos: análisis del desempeño de estudiantes de ingeniería en las pruebas saber-pro", (Carrascal & Giraldo, 2019). Este artículo se relaciona con las tendencias y patrones identificados en la literatura, ya que utiliza técnicas de MD para analizar la habilidad de cada uno de los estudiantes de ingeniería en las pruebas saber-pro. Además, considera variables multidimensionales relacionadas con el proceso de adquisición de conocimiento de los estudiantes y utiliza herramientas tecnológicas para el análisis de datos. En donde se hizo uso generalizado de técnicas de proceso de adquisición automático para examinar grandes volúmenes de datos y personalizar la enseñanza. En donde se hizo el uso creciente de plataformas y herramientas tecnológicas en el seguimiento del aprendizaje. Y del énfasis en el diseño de estudios longitudinales para conseguir una comprensión más profunda del proceso educativo. El aporte de este artículo radica en utilizar técnicas de MD para analizar el desempeño de los estudiantes de ingeniería en las pruebas saber-pro. Su objetivo es proporcionar información y conocimientos sobre los elementos que actúan en el rendimiento de los alumnos en estas pruebas, lo cual puede ser útil para optimizar la calidad de la educación y el diseño de estrategias de enseñanza en el campo de la ingeniería.

Artículo "Mapeo sistemático de la literatura sobre evaluación docente (2013-2017)", (Corona & Montoya, 2018): En el documento que se realiza se hace un mapeo sistemático de la literatura existente sobre evaluación docente durante el período comprendido entre 2013 y 2017. Su aporte consiste en identificar y analizar las tendencias, enfoques y hallazgos en el campo de la evaluación docente, proporcionando una visión general de la investigación que se realizó en ese período.

Artículo "Procedimiento de agrupación de estudiantes según riesgo de abandono para mejorar la gestión estudiantil en educación superior", (Caroca et al., 2022): Este artículo presenta un procedimiento de agrupación de estudiantes de acuerdo con su riesgo de abandono en la educación superior. El aporte consiste en proponer una metodología que ayuda a identificar y gestionar de manera más efectiva a los estudiantes que enfrentan un mayor riesgo de abandonar sus estudios, con el propósito de mejorar la retención estudiantil y el éxito académico.

Artículo "Comparación de técnicas de MD para identificar indicios de deserción estudiantil, a partir del desempeño académico", (Pérez Gutierrez, 2020): El aporte de este artículo se centra

en comparar diferentes técnicas de MD para reconocer señales tempranas de deserción estudiantil basadas en el rendimiento escolar. Su contribución consiste en evaluar y comparar la eficacia de diversas sistemáticas de análisis de datos para predecir la deserción estudiantil, lo cual puede ser de gran utilidad para implementar intervenciones preventivas y mejorar las tasas de retención en instituciones educativas.

Artículo "Minería predictiva aplicada al descubrimiento de factores asociados al desempeño en la competencia de lenguaje de los estudiantes de básica primaria", (Pérez Gutierrez, 2020): El valor de este artículo radica en la implementación de la minería predictiva para descubrir los factores que están asociados al rendimiento de los estudiantes en la competencia de lenguaje en el nivel de educación básica primaria. Su objetivo es identificar qué variables y factores influyen en el progreso de mejorar la habilidad lingüística en los estudiantes de primaria, lo cual consigue ayudar a mejorar las estrategias de enseñanza y el diseño de programas educativos en este ámbito.

Artículo "Modelo adaptativo de clasificación de profesiones en sistemas de orientación vocacional", (Cruz-Eraso & González-Serrano, 2022): Este artículo propone un modelo adaptable de categorización de carreras en sistemas de orientación vocacional. Su aporte consiste en desarrollar un modelo que ayuda a los estudiantes a explorar y clasificar diferentes profesiones de acuerdo con sus intereses y aptitudes, brindando orientación y apoyo en la toma de decisiones vocacionales.

Artículo "Caracterización de los estudiantes de una institución de educación superior mediante big data", (Hoyos Pineda & Aponte-Novoa, 2019): El aporte de este artículo radica en la utilización de big data para caracterizar a los alumnos de una institución de educación superior. En donde el propósito es recolectar información detallada sobre el perfil de los estudiantes, sus características demográficas, académicas y socioeconómicas, lo cual puede ser de beneficio para la delineación de políticas educativas y el perfeccionamiento de la calidad de la educación superior.

Artículo "Rendimiento académico de estudiantes en Educación Superior: predicciones de factores influyentes a partir de árboles de decisión", (Landa et al., 2021): El aporte de este artículo se centra en predecir los elementos que impactan en el rendimiento escolar de los estudiantes en la educación de nivel superior utilizando árboles de decisión. Su contribución consiste en identificar qué variables y aquellos factores que ejercen una mayor influencia en el

desempeño académico., lo cual puede ayudar a implementar intervenciones y estrategias de apoyo para optimizar los efectos en los estudiantes en este nivel educativo.

Artículo "Inteligencia artificial y aprendizaje colaborativo asistido por computadora en programación: un estudio de mapeo sistemático", (Hidalgo et al., 2023): El aporte de este artículo se orienta en realizar un estudio de mapeo sistemático sobre el uso de la IA y la colaboración en el proceso de adquisición de conocimiento con la asistencia de la computadora en el campo de la programación. Su objetivo es identificar las tendencias, enfoques y aplicaciones de estas tecnologías en el saber de la programación, proporcionando una visión general de la investigación en este ámbito.

Artículo "Resolución de problemas algorítmicos y objetos de aprendizaje: una revisión de la literatura", (Velasco Ramírez, 2020): El aporte de este documento consiste en realizar una revisión de la literatura sobre la resolución de problemas algorítmicos y el uso de objetos de aprendizaje. Su objetivo es analizar y resumir la investigación existente en este campo, proporcionando una visión general de los enfoques, metodologías y resultados relacionados con la determinación de dificultades algorítmicos y la utilización de recursos educativos en el proceso de enseñanza-aprendizaje.

5.2. Limitaciones y desafíos encontrados

El mapeo sistemático sobre la MDE revela importantes tendencias y hallazgos en la investigación. Se destaca el uso generalizado de técnicas de aprendizaje automático, que permiten examinar grandes volúmenes de datos y personalizar la enseñanza de acuerdo con las necesidades de los estudiantes. Además, se ha observado una consideración creciente de variables multidimensionales vinculadas al proceso de aprendizaje de los estudiantes, lo que brinda una visión más completa y detallada de los elementos que tienen impacto o influencia en su rendimiento académico.

En cuanto a las herramientas y tecnologías utilizadas en el seguimiento del aprendizaje, se evidencia un aumento en el uso de plataformas y herramientas tecnológicas. Estas herramientas permiten recopilar y analizar datos de manera más eficiente, facilitando la detección de patrones y tendencias en el rendimiento estudiantil.

Sin embargo, también se han identificado limitaciones y desafíos en esta área de investigación. Entre ellos, se encuentran posibles sesgos en la selección de estudios y una disponibilidad

limitada de artículos en idioma español, lo que podría afectar la representatividad de los resultados. Además, la calidad y la accesibilidad de los datos utilizados en la MDE pueden ser un obstáculo, ya que es fundamental contar con información precisa y confiable para obtener conclusiones sólidas. Además, la interpretación de los resultados puede ser compleja debido a la naturaleza multifactorial y no lineal del aprendizaje de los estudiantes. Además, existen preocupaciones éticas y de privacidad relacionadas con el uso de información de tipo personal de los estudiantes, lo que requiere un enfoque cuidadoso y respetuoso en el manejo de la información.

5.3. Aplicaciones prácticas y recomendaciones

El análisis del mapeo sistemático sobre la MDE revela importantes tendencias y aspectos a considerar en la aplicación de esta técnica. Entre las tendencias identificadas se encuentran el uso de técnicas del proceso de adquisición de conocimiento automático, la consideración de variables multidimensionales, el empleo de tecnologías y el énfasis en estudios longitudinales. Sin embargo, también se reconocen ciertas limitaciones, como potenciales sesgos en la elección de los estudios y desafíos relacionados con la calidad de los datos, la comprensión de los hallazgos y la privacidad.

Las aplicaciones prácticas de la MDE son variadas y pueden tener un efecto significativo en la mejora de la educación. Algunas de estas aplicaciones incluyen la personalización del aprendizaje, la identificación temprana de dificultades en los estudiantes, en la mejora en el proceso de realizar decisiones pedagógicas y el desarrollo de políticas educativas basadas en datos. La personalización del aprendizaje permite adecuar la enseñanza a las necesidades particulares de cada estudiante, lo que puede desarrollar su motivación y desempeño académico. La identificación temprana de dificultades permite intervenir de manera oportuna para brindar apoyo y recursos adicionales a los estudiantes que lo necesiten. La mejora de la toma de decisiones pedagógicas se logra al utilizar los datos recopilados para comprender mejor el desarrollo y las necesidades de los estudiantes, lo que permite ajustar las estrategias y recursos educativos de manera más efectiva. Por último, el desarrollo de políticas educativas basadas en datos permite tomar decisiones informadas y fundamentadas en evidencia para apoyar la excelencia educativa.

En cuanto a las recomendaciones, se sugiere fomentar la colaboración entre investigadores, educadores y encargados de tomar decisiones, con el propósito de intercambiar saberes y vivencias en el ámbito de la MD educativos. También es importante establecer estándares éticos claros para el manejo y uso de los datos educativos, garantizando el resguardo de la privacidad y confidencialidad de los estudiantes. Además, se recomienda promover la formación y capacitación en el uso de la MD educativos, tanto para investigadores como para profesionales de la educación. Asimismo, se destaca la importancia de recopilar datos longitudinales que permitan un seguimiento a largo plazo del progreso de los estudiantes. Por último, se sugiere la realización de más investigaciones en este campo para seguir avanzando en el conocimiento y el empleo de la MD educativos.

6. CONCLUSIONES

En este estudio, se buscó evaluar a los estudiantes su aprendizaje con el uso de la MDE y diseñar un modelo óptimo que integre las mejores prácticas y se ajuste a las necesidades determinadas de la educación. Para lograrlo, se realizó un exhaustivo mapeo sistemático de artículos relevantes, lo que permitió obtener una visión completa de las investigaciones existentes en el campo.

Inicialmente, se hizo un análisis exhaustivo de la literatura existente sobre la MDE y su uso con el motivo de supervisar el avance educativo estudiantil. Esto permitió conseguir la comprensión amplia de los conceptos clave y los enfoques utilizados en este campo de estudio.

Además, se evaluaron las mejores prácticas y tendencias en el uso de la MDE con el objeto de monitorear el avance del aprendizaje de los alumnos. Se analizaron diferentes técnicas y metodologías utilizadas en investigaciones previas, identificando aquellas que demostraron ser más efectivas y relevantes para los objetivos del estudio.

Con relación a la exploración de la literatura y la evaluación de las mejores prácticas, se procedió al diseño y validación de un modelo de MD educativos. Este modelo buscó integrar las prácticas más efectivas identificadas y adaptarse a las necesidades específicas de la educación. Se utilizó el desempeño académico como uno de los mayores indicadores en el diseño del modelo, con el objetivo de conseguir datos destacados sobre el aprendizaje de los estudiantes.

Finalmente, se consideraron diferentes situaciones para proporcionar una visión más extensa y global de las posibles aplicaciones y beneficios del modelo de MDE diseñado. Esta aproximación permitió explorar diversas perspectivas y analizar cómo el modelo podría adaptarse y ser utilizado en distintos contextos educativos. Se planteó la aplicación del modelo de MDE en diferentes entornos de educación. En este caso, se analizó cómo el modelo podría utilizarse en vigilar el desempeño de los estudiantes y brindar retroalimentación instantánea y personalizada. Los resultados revelaron que el modelo de MD proporcionó datos valiosos sobre los patrones de interacción de los estudiantes en los distintos entornos, permitiendo el descubrimiento temprana de posibles problemas de aprendizaje y la implementación de estrategias de apoyo adecuadas.

En conjunto, este estudio se centró en evaluar, diseñar y validar un modelo de MDE para el monitoreo del aprendizaje estudiantil, incorporando las mejores prácticas y adaptándose a las necesidades específicas de la educación. Los resultados obtenidos contribuyen al campo de la educación al proporcionar herramientas y enfoques efectivos para optimizar la calidad de la enseñanza y personalizar la práctica del proceso de adquisición de conocimiento de los alumnos.

REFERENCIAS

- Álava Morán, N. S. (2021). *Metodologías y técnicas analíticas de aprendizaje en la educación superior: un mapeo sistemático*. <https://dspace.ups.edu.ec/handle/123456789/20737>
- Albreiki, B., Zaki, N., & Alashwal, H. (2021). A Systematic Literature Review of Student' Performance Prediction Using Machine Learning Techniques. In *Education Sciences* (Vol. 11, Issue 9). <https://doi.org/10.3390/educsci11090552>
- Alghamdi, A. S., & Rahman, A. (2023). Data mining approach to predict success of secondary school students: A Saudi Arabian case study. *Education Sciences*, 13(3), 293.
- Almasri, A., Obaid, T., Abumandil, M. S. S., Eneizan, B., Mahmoud, A. Y., & Abu-Naser, S. S. (2023). Mining Educational Data to Improve Teachers' Performance. In M. Al-Emran, M. A. Al-Sharafi, & K. Shaalan (Eds.), *International Conference on Information Systems and Intelligent Applications* (pp. 243–255). Springer International Publishing.
- Araka, E., Oboko, R., Maina, E., & Gitonga, R. (2022). Using educational data mining techniques to identify profiles in self-regulated learning: an empirical evaluation. *The International Review of Research in Open and Distributed Learning*, 23(1), 131–162.
- Arnold, J. C., Mühling, A., & Kremer, K. (2023). Exploring core ideas of procedural understanding in scientific inquiry using educational data mining. *Research in Science & Technological Education*, 41(1), 372–392.
- Atiaja Balseca, L. E. (2023). *Uso de la analítica del aprendizaje de los estudiantes para minimizar la pérdida escolar en las diferentes modalidades de estudio*. <http://dspace.ups.edu.ec/handle/123456789/25199>
- Aulakh, K., Roul, R. K., & Kaushal, M. (2023). E-learning enhancement through Educational Data Mining with Covid-19 outbreak period in backdrop: A review. *International Journal of*

- Educational Development*, 102814.
- Ayala Carabajo, R. (2016). Social science and humanities research training based on van manen's hermeneutic phenomenological approach within the hispanic american context. *Educacion XXI*, 19(2), 359–381. <https://doi.org/10.5944/educXXI.13945>
- Ayala Carabajo, R. (2009). La esperanza pedagógica en la experiencia vivida de los educadores: un estudio fenomenológico-hermenéutico. *Tomos I y II. 4ESIS DOCTORAL INÁDITA. Departamento de Pedagogía Sistemática y Social, Universidad Autónoma de Barcelona-España.*
- Ayala Carabajo, Raquel. (2008). La metodología fenomenológico-hermeneútica de M. Van Manen en el campo de la investigación educativa. Posibilidades y primeras experiencias. *Revista de Investigación Educativa, RIE*, 26(2), 409–430.
- Ayala Carabajo, Raquel. (2011). Pedagogical hope: A fresh and deep glance to educational experience from van Manen's approach. *Revista Espanola de Pedagogia*, 69(248), 119–143.
- Ayala Carabajo, Raquel. (2018). Pedagogical relationship: Max van manen's pedagogy in the sources of educational experience. *Revista Complutense de Educacion*, 29(1), 27–41. <https://doi.org/10.5209/RCED.51925>
- Baradwaj, B. K., & Pal, S. (2012). Mining educational data to analyze students' performance. *ArXiv Preprint ArXiv:1201.3417*.
- Batool, S., Rashid, J., Nisar, M. W., Kim, J., Kwon, H.-Y., & Hussain, A. (2023). Educational data mining to predict students' academic performance: A survey study. *Education and Information Technologies*, 28(1), 905–971.
- Buenaño Fernández, D. (2020). *Propuesta de un conjunto de herramientas de MD para evaluar el desempeño de los estudiantes y los procesos de enseñanza-aprendizaje en el ámbito de la educación en ingeniería.*
- Cardona, T., Cudney, E. A., Hoerl, R., & Snyder, J. (2023). Data mining and machine learning retention models in higher education. *Journal of College Student Retention: Research, Theory & Practice*, 25(1), 51–75.
- Caroca, A., Derpich, I., Gatica, G., Alfaro, M., Ruete, D., & Hinojosa, M. (2022). Procedimiento de agrupación de estudiantes según riesgo de abandono para mejorar la gestión estudiantil en educación superior. *Texto Livre: Linguagem e Tecnologia*, 15, 1–22.
- Carrascal, A. I. O., & Giraldo, J. J. (2019). MD educativos: Análisis del desempeño de estudiantes de ingeniería en las pruebas SABER-PRO. *Revista Politécnica*, 15(29), 128–139.
- Chen, C.-H., Yang, S. J. H., Weng, J.-X., Ogata, H., & Su, C.-Y. (2021). Predicting at-risk university students based on their e-book reading behaviours by using machine learning classifiers. *Australasian Journal of Educational Technology*, 37(4 SE-Articles), 130–144. <https://doi.org/10.14742/ajet.6116>
- Chen, X., Zou, D., Xie, H., Chen, G., Lin, J., & Cheng, G. (2022). Exploring contributors, collaborations, and research topics in educational technology: A joint analysis of mainstream conferences. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-022-11209-y>
- Chen, Y., & Zhai, L. (2023). A comparative study on student performance prediction using machine learning. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-023-11672-1>
- Corona, C. N., & Montoya, M. S. R. (2018). Mapeo sistemático de la literatura sobre evaluación docente (2013-2017). *Educação e Pesquisa*, 44.
- Cruz-Eraso, A.-F., & González-Serrano, C. (2022). Adaptive Model of Classification of Professions in Vocational Guidance Systems. *Revista Facultad de Ingeniería*, 31(61).
- de Baker, R. S. J., & Inventado, P. S. (2014). Chapter X: Educational Data Mining and Learning Analytics. *Comput. Sci*, 7, 1–16.
- Dunajeva, O., Pentel, A., & Maksimova, N. (2022). COVID-19's Impact on the Quality of Educational Process and the Academic Performance as Viewed by IT Students: A Case Study in Text Mining. In M. E. Auer, H. Hortsch, O. Michler, & T. Köhler (Eds.), *Mobility for Smart Cities and Regional Development - Challenges for Higher Education* (pp. 417–425). Springer International Publishing.
- Echeverry, C. E. M., Trujillo, M. L., & Salazar, M. H. M. (2017). MD en gestión del conocimiento de pymes de Colombia. *Revista Virtual Universidad Católica Del Norte*, 50, 224–237.
- Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. *International Journal*

- of *Technology Enhanced Learning*, 4(5–6), 304–317. <https://doi.org/10.1504/IJTEL.2012.051816>
- Hidalgo, C. G., Bucheli-Guerrero, V. A., & Ordóñez-Eraso, H. A. (2023). Artificial Intelligence and Computer-Supported Collaborative Learning in Programming: A Systematic Mapping Study. *Tecnura*, 27(75), 175–206.
- Hoyos Pineda, J. G., & Aponte-Novoa, F. A. (2019). Caracterización de los estudiantes de una institución de educación superior mediante big data. *Ingeniería y Desarrollo*, 37(2), 159–172.
- Issah, I., Appiah, O., Appiahene, P., & Inusah, F. (2023). A systematic review of the literature on machine learning application of determining the attributes influencing academic performance. *Decision Analytics Journal*, 100204.
- Lampropoulos, G. (2023). Educational Data Mining and Learning Analytics in the 21st Century. In *Encyclopedia of Data Science and Machine Learning* (pp. 1642–1651). IGI Global.
- Landa, B. D., Romero, R. M., & Rodriguez, W. J. M. (2021). Rendimiento académico de estudiantes en educación superior: predicciones de factores influyentes a partir de árboles de decisión. *Telos: Revista de Estudios Interdisciplinarios En Ciencias Sociales*, 23(3), 616–639.
- Leitner, P., Khalil, M., & Ebner, M. (2017). Learning analytics in higher education—a literature review. In A. Peña-Ayala (Ed.), *Studies in Systems, Decision and Control* (Vol. 94, pp. 1–23). Springer International Publishing. https://doi.org/10.1007/978-3-319-52977-6_1
- Levine, R. A., Rivera, P. E., He, L., Fan, J., & Bresciani Ludvick, M. J. (2023). A learning analytics case study: On class sizes in undergraduate writing courses. *Stat*, 12(1), e527. <https://doi.org/https://doi.org/10.1002/sta4.527>
- Llerena-Izquierdo, J. (2023). Adaptation of the curriculum in relation to student learning outcomes in initial programming courses. *2023 IEEE World Engineering Education Conference (EDUNINE)*, 1–6. <https://doi.org/10.1109/EDUNINE57531.2023.10102894>
- Llerena-Izquierdo, Joe, & Ayala-Carabajo, R. (2022a). Desarrollo de competencia investigativa de estudiantes universitarios de ingeniería en proceso de titulación. *Memoria Investigaciones En Ingeniería*, 23, 50–62.
- Llerena-Izquierdo, Joe, & Ayala-Carabajo, R. (2022b). Preparing undergraduate computer engineering students for publishable research. *2022 IEEE World Engineering Education Conference (EDUNINE)*, 1–6. <https://doi.org/10.1109/EDUNINE53672.2022.9782361>
- Llerena, J., Alava-Moran, N., & Zamora-Galindo, J. (2021). Learning analytics for student academic tracking, a comparison between Analytics Graphs and Edwiser Reports. *2021 Second International Conference on Information Systems and Software Technologies (ICI2ST)*, 101–107. <https://doi.org/10.1109/ICI2ST51859.2021.00022>
- López-Chila, R., Arteaga-Sotomayor, J., Caballero-Barros, E., & Llerena-Izquierdo, J. (2023). An e-learning evaluation method with the effectiveness of the questionnaire tool in the VLEs. *2023 IEEE World Engineering Education Conference (EDUNINE)*, 1–6. <https://doi.org/10.1109/EDUNINE57531.2023.10102811>
- López-Chila, R., Sumba-Nacipucha, N., & Llerena-Izquierdo, J. (2023). Effectiveness of digital resources for cloned Linear Algebra courses at the undergraduate level. *2023 IEEE World Engineering Education Conference (EDUNINE)*, 1–5. <https://doi.org/10.1109/EDUNINE57531.2023.10102901>
- López Zambrano, J., Lara Torralbo, J. A., Romero Morales, C., & others. (2021). Early prediction of student learning performance through data mining: A systematic review. *Psicothema*.
- Moon, J., Lee, D., Choi, G. W., Seo, J., Do, J., & Lim, T. (2023). Learning analytics in seamless learning environments: a systematic review. *Interactive Learning Environments*, 1–18.
- Morales, C. R., Soto, S. V., & Martínez, C. H. (2005). Estado actual de la aplicación de la MD a los sistemas de enseñanza basada en web. *Actas Del III Taller Nacional de Minería de Datos y Aprendizaje, TAMIDA2005*, 49, 56.
- Namoun, A., & Alshantqiti, A. (2020). Predicting student performance using data mining and learning analytics techniques: A systematic literature review. *Applied Sciences*, 11(1), 237.
- Naseem, M., Chaudhary, K., & Sharma, B. (2022). Predicting Freshmen Attrition in Computing Science using Data Mining. *Education and Information Technologies*, 27(7), 9587–9617. <https://doi.org/10.1007/s10639-022-11018-3>

- Olaleye, T. O., & Vincent, O. R. (2020). A Predictive Model for Students' Performance and Risk Level Indicators Using Machine Learning. *2020 International Conference in Mathematics, Computer Engineering and Computer Science, ICMCECS 2020*. <https://doi.org/10.1109/ICMCECS47690.2020.240897>
- Ordoñez-Avila, R., Reyes, N. S., Meza, J., & Ventura, S. (2023). Data mining techniques for predicting teacher evaluation in higher education. A SYSTEMATIC literature review. *Heliyon*.
- Pallathadka, H., Wenda, A., Ramirez-Asís, E., Asís-López, M., Flores-Albornoz, J., & Phasinam, K. (2023). Classification and prediction of student performance data using various machine learning algorithms. *Materials Today: Proceedings*, 80, 3782–3785.
- Papamitsiou, Z., & Economides, A. A. (2014). Learning analytics and educational data mining in practice: A systematic literature review of empirical evidence. *Journal of Educational Technology & Society*, 17(4), 49–64.
- Parhizkar, A., Tejeddin, G., & Khatibi, T. (2023). Student performance prediction using datamining classification algorithms: Evaluating generalizability of models from geographical aspect. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-022-11560-0>
- Pérez Gutierrez, B. R. (2020). Comparación de técnicas de minería de datos para identificar indicios de deserción estudiantil, a partir del desempeño académico. *Revista UIS Ingenierías*, 19(1 (2020)), 193–204.
- Ramaswami, G., Susnjak, T., Mathrani, A., & Umer, R. (2022). Use of Predictive Analytics within Learning Analytics Dashboards: A Review of Case Studies. *Technology, Knowledge and Learning*. <https://doi.org/10.1007/s10758-022-09613-x>
- Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. *WIREs Data Mining and Knowledge Discovery*, 10(3). <https://doi.org/10.1002/widm.1355>
- Salihoun, M. (2020). State of Art of Data Mining and Learning Analytics Tools in Higher Education. *International Journal of Emerging Technologies in Learning (IJET)*, 15(21), 58. <https://doi.org/10.3991/ijet.v15i21.16435>
- Sanchez-Romero, J., & Llerena-Izquierdo, J. (2023). Revisión de la literatura sobre el uso del aprendizaje profundo enfocado en sistemas de inspección ópticos automatizados para la detección de defectos superficiales en el sector de la manufactura. *Revista InGenio*, 6(2), 1–19. <https://doi.org/10.18779/ingenio.v6i2.680>
- Schelfhout, S., Wille, B., Fonteyne, L., Roels, E., Deros, E., De Fruyt, F., & Duyck, W. (2022). How accurately do program-specific basic skills predict study success in open access higher education? *International Journal of Educational Research*, 111, 101907. <https://doi.org/https://doi.org/10.1016/j.ijer.2021.101907>
- Siemens, G., & Long, P. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE Review*, 46(5), 30.
- Silva, C., & Fonseca, J. (2017). Educational data mining: a literature review. *Europe and MENA Cooperation Advances in Information and Communication Technologies*, 87–94.
- Song, Z., Sung, S.-H., Park, D.-M., & Park, B.-K. (2023). All-Year Dropout Prediction Modeling and Analysis for University Students. In *Applied Sciences* (Vol. 13, Issue 2). <https://doi.org/10.3390/app13021143>
- Švábensky, V., Vykopal, J., Čeleda, P., & Kraus, L. (2022). Applications of educational data mining and learning analytics on data from cybersecurity training. *Education and Information Technologies*, 27(9), 12179–12212.
- Taipalus, T. (2023). Systematic Mapping Study in Information Systems Research. *Journal of the Midwest Association for Information Systems (JMWAIS)*, 2023(1), 2.
- Tamada, M. M., Giusti, R., & Netto, J. F. (2022). Predicting Students at Risk of Dropout in Technical Course Using LMS Logs. In *Electronics* (Vol. 11, Issue 3). <https://doi.org/10.3390/electronics11030468>
- Velasco Ramírez, M. L. (2020). Resolución de problemas algorítmicos y objetos de aprendizaje: una revisión de la literatura. *RIDE. Revista Iberoamericana Para La Investigación y El Desarrollo Educativo*, 10(20).
- Viberg, O., & Grönlund, Å. (2023). *Practicable Learning Analytics*. Springer Nature.

- Wang, X., Zhao, Y., Li, C., & Ren, P. (2023). ProbSAP: A comprehensive and high-performance system for student academic performance prediction. *Pattern Recognition*, *137*, 109309. <https://doi.org/https://doi.org/10.1016/j.patcog.2023.109309>
- Weisi, H., & Ahmadi, R. (2023). Conceptualizing student voice in teacher-student curriculum partnerships within Iran's higher education. *Higher Education*. <https://doi.org/10.1007/s10734-023-01036-6>
- Yağcı, M. (2022). Educational data mining: prediction of students' academic performance using machine learning algorithms. *Smart Learning Environments*, *9*(1), 11. <https://doi.org/10.1186/s40561-022-00192-z>
- Zhang, Y., Yun, Y., An, R., Cui, J., Dai, H., & Shang, X. (2021). Educational Data Mining Techniques for Student Performance Prediction: Method Review and Comparison Analysis. *Frontiers in Psychology*, *12*. <https://doi.org/10.3389/fpsyg.2021.698490>