

Mapecto Sistemático de las Herramientas de Soporte en la Capacitación de Minería de Datos

Systematic Mapping of Supporting Tools in Data Mining Training

Walter Garcia-Velez¹, Patricia Quiroz-Palma², y Alex Santamaria-Philco³

RESUMEN

Contexto: La importancia de la educación ha dado lugar a la búsqueda de diferentes técnicas de aprendizaje y al uso de herramientas informáticas. Este estudio realiza un mapeo sistemático de las herramientas utilizadas como soporte para la capacitación en minería de datos durante el periodo de 2020 a 2025. **Metodo:** Se siguió la metodología de mapeo sistemático adaptada de Kitchenham y Charters, analizando 14 estudios seleccionados de IEEE Xplore, Scopus y Google Scholar. **Resultados:** Se identificaron 15 herramientas de soporte, siendo las más destacadas Weka, RapidMiner y R Studio. El 42.86 % de los estudios utilizan aprendizaje automático como técnica principal. **Conclusiones:** Las herramientas identificadas permiten optimizar la capacitación en minería de datos, contribuyendo al desarrollo de talento en este ámbito crucial para el crecimiento tecnológico y económico.

Palabras clave: Capacitación, Aprendizaje, Herramientas, Minería de datos, Mapeo sistemático

ABSTRACT

Context: The importance of education has given rise to the search for different learning techniques and the use of computer tools. This study carries out a systematic mapping of the tools used to support training in data mining during the period from 2020 to 2025. **Method:** The systematic mapping methodology adapted from Kitchenham and Charters was followed, analyzing 14 selected studies from IEEE Xplore, Scopus and Google Scholar. **Results:** 15 support tools were identified, with Weka, RapidMiner and R Studio being the most prominent. 42.86% of the studies use machine learning as the main technique. **Conclusions:** The identified tools allow optimizing data mining training, contributing to the development of talent in this area crucial for technological and economic growth.

Keywords: Training, Learning, Tools, Data mining, Systematic mapping

Fecha de recepción: Diciembre 12, 2024

Fecha de aceptación: Marzo 30, 2025

Introducción

La capacitación adecuada permite a las personas obtener conocimientos en diferentes ámbitos. Para lograr una adecuada capacitación existen herramientas que apoyan el proceso de capacitación y entrenamiento en las instituciones de educación superior. El conocimiento puede ser transmitido a las personas mediante la capacitación y el entrenamiento, lo que denominaremos genéricamente capacitación. "La capacitación es toda actividad realizada en una organización, respondiendo a sus necesidades, que busca mejorar la actitud, conocimiento, habilidades o conductas de su personal" (Palma-Avellán, 2018). Con los conocimientos necesarios, toda persona, que denominaremos estudiante, puede ofrecer soluciones de aplicación de minería de datos. Por lo tanto, uno de los principales objetivos de las instituciones de educación es proporcionar o facilitar una capacitación eficaz, ya que aumenta la participación, la colaboración, la comunicación y el conocimiento, entre otros. Por otra parte, la evaluación continua a través de la capacitación y el aprendizaje influye directamente en el grado en que un estudiante puede responder satisfactoriamente a cualquier actividad.

Las capacidades adquiridas en la capacitación ayudan a las personas no sólo a mejorar sus conocimientos para realizar un trabajo, sino también para resolver problemas. De hecho, para evitar la obsolescencia de los conocimientos,

se debe llevar a cabo la capacitación periódica. El contenido de la capacitación debe contener información teórica, experiencias y prácticas y debe ser enfocado al campo de estudio.

En este caso, el campo de estudio es la capacitación de minería de datos. Según (Henry, 2018) el Data Mining (minería de datos) es un conjunto de técnicas y tecnologías que permiten explorar grandes bases de datos, de manera automática o semiautomática, con el objetivo de encontrar patrones repetitivos que expliquen el comportamiento de estos datos. (Pérez and Santín, 2014) mencionan que la minería de datos o Data Mining se basa en un proceso que permite el descubrimiento automático de relaciones o patrones que se obtienen al analizar gran volumen de datos. Estos autores también indican las técnicas de minería de datos. Entre las mejores técnicas de minería de datos tenemos: Seguimiento de Patrones, Asociación, Clasificación, Detección de valores atípicos, Agrupación,

¹ Universidad Laica Eloy Alfaro de Manabí. E-mail: walter.garcia@uleam.edu.ec

² Universidad Laica Eloy Alfaro de Manabí. E-mail: patricia.quiroz@uleam.edu.ec

³ Universidad Laica Eloy Alfaro de Manabí. E-mail: alex.santamaria@uleam.edu.ec



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International Licence.

Cómo citar: Garcia-Velez, W., Quiroz-Palma, P., & Santamaria-Philco, A. (2025). Mapeo Sistemático de las Herramientas de Soporte en la Capacitación de Minería de Datos. *Ecuadorian Science Journal*, 9(1), 31–38, Marzo–2025. DOI: <https://doi.org/10.46480/esj.9.1.223>

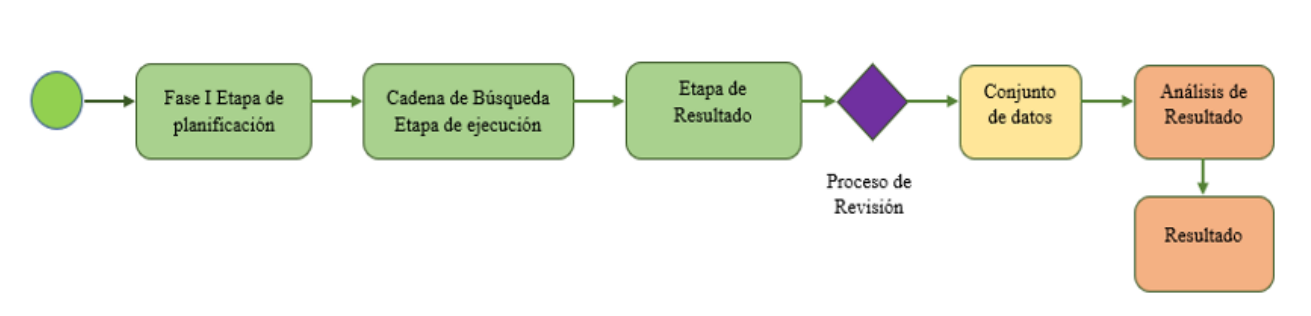


Figura 1. Proceso de mapeo sistemático basado en (Kitchenham and Charters, 2007).

Patrones secuenciales, Árboles de decisión, Análisis de Regresión, Procesamiento de la memoria a largo plazo y Redes Neuronales. Las TIC desempeñan un papel importante en la capacitación de minería de datos, a través de herramientas informáticas que proporcionan apoyo a la capacitación.

El objetivo principal de la investigación es llevar a cabo un mapeo sistemático de las herramientas de soporte utilizadas en la capacitación de minería de datos. Este estudio busca identificar y analizar las herramientas más empleadas, su efectividad, y las tendencias actuales en la educación de minería de datos, con el fin de proporcionar un panorama claro y actualizado de las prácticas y tecnologías empleadas en el proceso de enseñanza-aprendizaje. Este paper está estructurado de la siguiente forma: sección 2, materiales y métodos. Sección 3, los resultados y discusión. Finalmente, en la Sección 4 se establecen las conclusiones y trabajos futuros.

Las TIC desempeñan un papel importante en la capacitación de minería de datos, a través de herramientas informáticas que proporcionan apoyo a la capacitación.

El objetivo principal de la investigación es llevar a cabo un mapeo sistemático de las herramientas de soporte utilizadas en la capacitación de minería de datos. Este estudio busca identificar y analizar las herramientas más empleadas, su efectividad, y las tendencias actuales en la educación de minería de datos.

Metodología

Esta revisión bibliográfica se basa en la adaptación del modelo propuesto por (Kitchenham and Charters, 2007). Es para la obtención de las herramientas de soporte para la capacitación de minería de datos. Este método adaptado se divide en tres fases: Planificación de la revisión, Realización de la revisión y Resultado de la revisión. En la Figura 1 se muestra el proceso que se siguió en esta revisión bibliográfica.

Este método adaptado se divide en tres fases: Planificación de la revisión, Realización de la revisión (ejecución) y Resultado de la revisión. En el proceso de Planificación de la Revisión se definen las preguntas de investigación, el protocolo de revisión y se obtiene la cadena de búsqueda. En el proceso Ejecución de la Revisión, se seleccionan los estudios hasta obtener una síntesis de los datos. Finalmente, en el proceso Resultados de la Revisión, se obtienen los resultados de los estudios analizados y obtener un conjunto de datos para finalmente aplicar el metaanálisis y

obtener los gráficos estadísticos de los resultados del estudio bibliográfico. Cada proceso y subproceso se explicará en las siguientes subsecciones.

Fase 1: Planificación de la revisión

En esta fase, el primer paso se especificaron las preguntas de investigación (Tabla 1) se basaron sobre la información que queremos obtener sobre las herramientas para la enseñanza en minería de datos. El segundo paso fue desarrollar un protocolo de revisión aplicando los criterios de revisión, obteniendo a partir de esta etapa, la cadena de búsqueda (investigación).

RQ1. ¿Qué herramientas de TI se utilizan para apoyar el aprendizaje de la minería de datos?

RQ2. ¿Qué técnicas/algoritmos de minería de datos que aplican las herramientas de soporte?

Las búsquedas se basarán en los criterios de inclusión y exclusión.

Los criterios de inclusión referentes a la capacitación serán estudios en inglés referentes a herramientas de soporte a la capacitación en minería de datos, publicados en Journals y Conferencias desde 2020 a 2025, debido a que se consideraron los últimos 5 años para obtener herramientas actuales e innovadoras. Los términos de referencia a tomarse en cuenta serán: datamining, training, learning, teaching, education. Los campos de investigación incluidos son: Computer Science and Education.

Los criterios de exclusión: se excluirán los posters, trabajos en progreso, documentos, libros y workshops. Además, trabajos no relacionados a términos de referencia y diferentes los campos de investigación en Computación o Educación.

Evaluar el protocolo de revisión: Para definir la cadena de búsqueda se identificaron los términos claves de nuestro estudio: data mining, training, teaching, learning, education, tool. De esta forma, la cadena de búsqueda se estableció así: “(datamining OR data mining) AND (train* OR learn*) AND (educa* OR teach*) AND (tool)”. Las fuentes primarias fueron las librerías relevantes: IEEE Xlore, Scopus, y Google Scholar. El método es complementado con búsqueda manual en varias fuentes de información. Además de las referencias importantes citadas en los artículos relevantes.

Fase 2: Ejecución de la Revisión

El primer paso fue la selección de los estudios primarios. Las fuentes primarias fueron bibliotecas relevantes IEEE Xplore, Scopus y Google Scholar. El método se complementó con

Tabla 1. Cadena de búsqueda

Concepto	Subcadena	Conector	Razón
Minería de datos	Datamining Data mining	Y	Concepto central de la revisión
Educación	Educat*	O	Incluye educación, educar. . .
Aprendizaje	Learn*	O	Incluye aprender, learning. . .
Enseñanza	Teach*	O	Incluye enseñar, teaching. . .
Plataforma	Platform*	O	Incluye platform, platforms. . .
Herramientas	Tool	O	Incluye tool, tools. . .

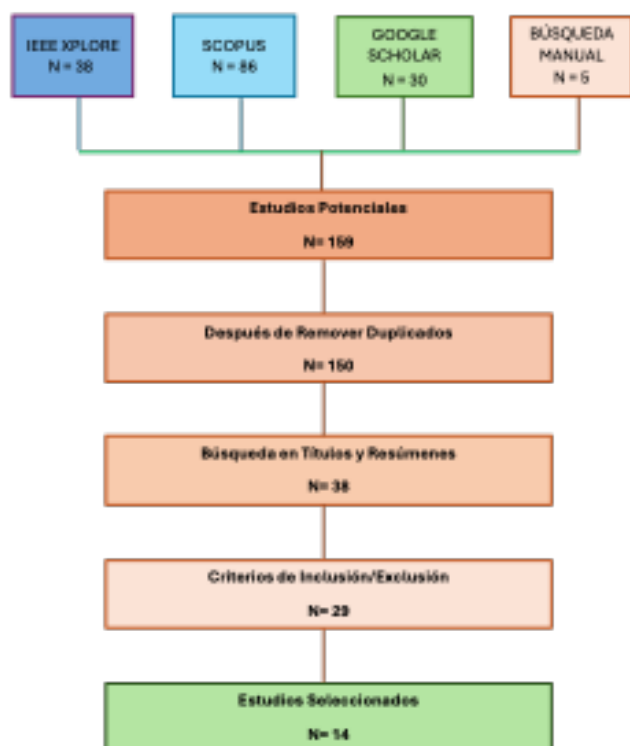
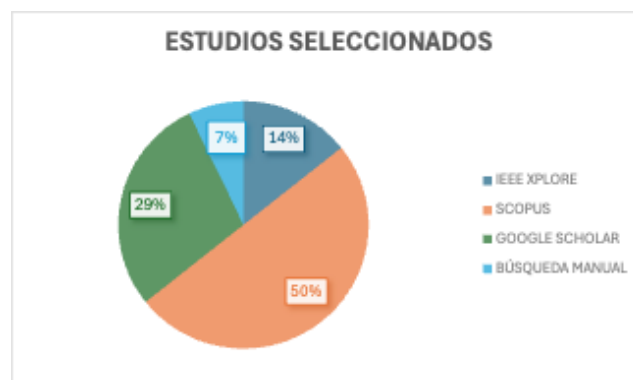
Cadena de búsqueda final: “(datamining OR data mining) AND (train* OR learn*) AND (educa* OR teach*) AND (tool)”

la búsqueda manual de diversas fuentes. Además de las importantes referencias citadas en los artículos relevantes. La búsqueda en cada una de las bases de datos indexadas se refinó mediante la selección de estudios primarios publicados en revistas, actas de congreso y se complementó con la búsqueda manual de estudios relacionados.

El tercer paso fue la extracción de datos y el seguimiento. Los datos extraídos se revisaron manualmente, aplicando los criterios de exclusión definidos que no cumplieron con la evaluación de calidad.

Los datos extraídos han sido revisados manualmente, aplicando los criterios de exclusión definidos y que no cumplen con la evaluación de la calidad. Se han seleccionado 159 estudios de las fuentes primarias, de los cuales se excluyen los estudios que no cumplen los criterios de calidad, quedando 14 estudios seleccionados, como muestra la Tabla 2 y Figuras 2 y 3.

Los estudios seleccionados se analizan uno por uno para garantizar la calidad y se clasifican según las preguntas de investigación planteadas.

**Figura 2.** Estudios seleccionados por búsquedas automáticas y manuales.**Figura 3.** Estudios seleccionados por búsquedas automáticas y manuales.

Fase 3: Resultado de la revisión

Esta fase consta de tres pasos. La primera etapa consiste en tabular los resultados obtenidos. El segundo paso es crear los informes de los resultados. El tercer paso es generar el conjunto de datos para realizar el análisis de los resultados. Esta fase se describe en la siguiente sección.

Resultados y Discusión

Una vez seleccionados los resultados, se han tabulado y organizado según la base de datos de búsqueda. En IEEE Xplore se han seleccionado 2 artículos, Scopus: 7 trabajos seleccionados y Google Scholar: 4 trabajos seleccionados. Además, se han seleccionado 1 estudio mediante búsqueda manual, con un total de 14 estudios seleccionados para esta revisión sistemática.

La Tabla 2, muestra los resultados del proceso de selección de estudios. Los resultados se han organizado en tablas. Para los resultados estadísticos gráficos, se procede a generar un conjunto de datos utilizando R Studio y el lenguaje R. Los resultados se han clasificado por años para determinar la frecuencia de los estudios, determinando que 2020 y 2023 son los años de mayor producción científica de la literatura sobre herramientas de soporte para la capacitación de personas en el ámbito de la minería de datos, como muestra la Tabla 3 y Figura 4.

Para organizar los estudios, se han clasificado según las preguntas de investigación planteadas en esta revisión bibliográfica. Se ha determinado que, los autores de 14 estudios han utilizado herramientas de soporte para la capacitación y preparación de los estudiantes de minería de datos, como se muestra en la Tabla 4. Estos estudios

Tabla 2. Proceso de selección de estudios

Fuente	Estudios Potenciales	Sin Duplicados	Criterios Aplicados	Seleccionados
IEEE Xplore	38	33	8	2
Scopus	86	84	12	7
Google Scholar	30	28	7	4
Búsqueda Manual	5	5	2	1
Total	159	150	29	14

Tabla 3. Estudios seleccionados por año

Año	IEEE Xplore	Scopus	Google Scholar	Manual	Total
2020	-	1	2	-	3
2021	1	-	1	-	2
2022	-	1	1	-	2
2023	1	3	-	1	5
2024	-	1	-	-	1
2025	-	1	-	-	1
Total	2	7	4	1	14

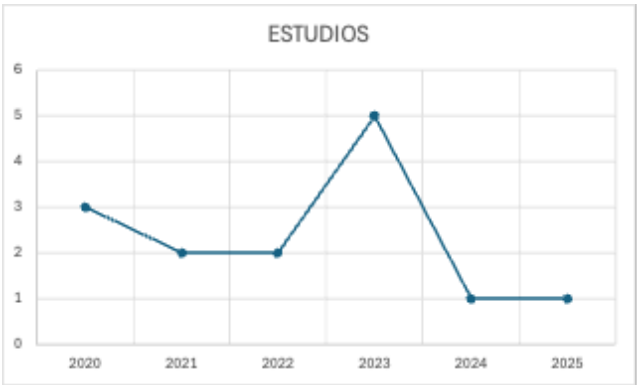


Figura 4. Estudios seleccionados por años.

fueron finalmente agrupados de acuerdo con las preguntas de investigación.

Los estudios seleccionados respecto a las herramientas de soporte en la capacitación de minería de datos se describen en la Tabla 5 a continuación. Siendo las herramientas relacionadas al aprendizaje automático las de mayor presencia en estos estudios.

En la Tabla 5, se presentan e identifican las herramientas de soporte, entre las cuales se destacan *Weka*, *RapidMiner*, *R Studio*, *Orange*, *TADA-ED*, *EID-Miner*, *Scikit-learn*, *TensorFlow*, *Jupyter Notebook*, *Power BI*, *e-Learning Toolbox*, *Course Prophet*, *Tweepy (Python)*, *MOOC* y *Palaute*. En la Tabla ??, se presentan las herramientas identificadas en los 14 estudios seleccionados.

Las herramientas identificadas en los 14 estudios seleccionados, presentadas en la Tabla 6, se describen a continuación:

Weka. Herramienta de software libre desarrollada por la Universidad de Waikato, en Nueva Zelanda. Se utiliza en el campo de la minería de datos y el aprendizaje automático. Esta herramienta se aplica en diversos ámbitos para determinar patrones en diferentes campos

como informática, medicina, educación y agricultura (Espinoza Mina, 2018; Gil, 2023; Shakil et al., 2015; Paguay Hurtado, 2021).

RapidMiner. Herramienta de gran relevancia en minería de datos, que facilita la preparación de datos, el análisis predictivo, el análisis de clústeres, la minería de textos, el análisis de series temporales, así como la validación y evaluación de modelos, todo sin necesidad de escribir código extenso. Esta herramienta permite optimizar procesos como la fijación de precios en una línea de productos de una empresa, mejorando la toma de decisiones empresariales (Álvarez Peralta and Recalde-Varela, 2020; López et al., 2016).

R y RStudio. *R* es una herramienta clave para el análisis de datos en diversas disciplinas, proporcionando una amplia gama de herramientas para la manipulación, análisis y visualización de datos. Su flexibilidad lo convierte en una opción popular tanto para la investigación académica como para la industria (Lortie, 2017; Lizana, 2020).

Orange. Herramienta de software de código abierto para minería de datos y aprendizaje automático, conocida por su interfaz gráfica interactiva. Es muy útil para explorar, analizar y visualizar datos sin necesidad de escribir código. *Orange* es compatible con una amplia variedad de análisis y técnicas en la ciencia de datos, tales como: análisis descriptivo y exploración de datos, preparación de datos, análisis predictivo, análisis de clúster o agrupamiento, minería de texto, análisis de series temporales, visualización de datos, evaluación y validación de modelos. Debido a su facilidad de manipulación y versatilidad, *Orange* está siendo ampliamente utilizada por analistas de datos en diversas áreas de conocimiento (Zupan and Pečar, 2018; Lemmerich and Voss, 2019; Babu and Raj, 2020; Kovar and Šojat, 2021).

Tabla 4. Resultado del Mapeo Sistemático

Pregunta de Investigación	Tipo de Propuesta	Estudios	Nº Estudios	Porcentaje %
RQ1. ¿Qué herramientas de TI se utilizan para apoyar el aprendizaje de la minería de datos?	Herramientas de soporte	E01, E02, E03, E04, E05, E06, E07, E08, E09, E10, E11, E12, E13, E14	14	100%
	Algoritmos predictivos	E01, E07, E13	3	21,43%
	Aprendizaje automático	E02, E04, E05, E07, E11, E14	6	42,86%
	Patrones secuenciales	E03	1	7,14%
	Lógica Difusa	E06	1	7,14%
	Algoritmo K-means	E09	1	7,14%
	Análisis de sentimientos	E10	1	7,14%
	Textmining	E12	1	7,14%
RQ2. ¿Qué técnicas/algoritmos de minería de datos que aplican las herramientas de soporte?	Algoritmos predictivos	E01, E07, E13	3	21,43%
	Aprendizaje automático	E02, E04, E05, E07, E11, E14	6	42,86%
	Patrones secuenciales	E03	1	7,14%
	Lógica Difusa	E06	1	7,14%
	Algoritmo K-means	E09	1	7,14%
	Análisis de sentimientos	E10	1	7,14%
	Textmining	E12	1	7,14%

TADA-ED (Tool for Analysis and Decision-making in Education). Herramienta avanzada diseñada para realizar análisis de datos educativos, enfocados en la toma de decisiones informadas que pueden mejorar los procesos de enseñanza y aprendizaje. Esta herramienta facilita el análisis de grandes volúmenes de datos relacionados con el rendimiento de los estudiantes, los métodos pedagógicos y los recursos educativos, con el objetivo de optimizar las estrategias educativas (Hertzberg et al., 2022).

EID-Miner. Conjunto de herramientas utilizadas para análisis de datos en educación, incluyendo rendimiento académico, estrategias pedagógicas y análisis de clústeres (Zhang et al., 2017).

Scikit-learn. Herramienta muy conocida para el análisis de datos en Python. Es de fácil acceso para realizar tareas de minería de datos y aprendizaje automático. *Scikit-learn* incluye una amplia gama de algoritmos y herramientas para el análisis de datos y la construcción de modelos predictivos (Pedregosa et al., 2011; Buitinck et al., 2016).

TensorFlow. Herramienta extremadamente potente para el desarrollo y la implementación de modelos de aprendizaje automático y *deep learning*, que permite a las máquinas aprender de manera autónoma patrones complejos en grandes volúmenes de datos (Abadi et al., 2016).

Jupyter Notebook. Herramienta de código abierto ampliamente utilizada en la ciencia de datos para análisis exploratorio, análisis estadístico, modelado predictivo, aprendizaje automático y visualización de datos. También permite la interacción directa con los datos (VanderPlas, 2016).

Power BI. Herramienta de visualización de datos desarrollada por Microsoft que permite transformar grandes volúmenes de datos para realizar análisis descriptivos,

predictivos, de tendencias, cohortes, geoespacial y de desempeño, integrando múltiples fuentes de datos en ámbitos empresariales y educativos. Por su variedad de funcionalidades, *Power BI* facilita la toma de decisiones mediante informes interactivos y visualizaciones (Mazurek and Wasilkowski, 2020).

e-Learning Toolbox. Herramienta basada en minería de datos con lógica difusa, utilizada en el ámbito educativo para enriquecer la experiencia de aprendizaje en entornos virtuales. Permite a educadores y estudiantes acceder a recursos digitales que facilitan la enseñanza (Perakovic et al., 2020).

Course Prophet. Herramienta impulsada por aprendizaje automático capaz de predecir si un estudiante está en riesgo de reprobar un curso determinado. Analiza el historial académico para identificar patrones y factores asociados al riesgo de fracaso (Caicedo-Castro, 2023).

Tweepy (Python). Biblioteca de Python que facilita la interacción con la API de Twitter, permitiendo a los desarrolladores acceder y manipular datos de Twitter de manera eficiente. Entre sus funcionalidades destacan la extracción de datos, análisis de sentimiento y visualización de información (Aleman Viteri, 2021; Darad and Krishnan, 2023).

MOOC (Massive Open Online Courses). Herramienta que ha permitido una innovación en el panorama educativo global, facilitando el acceso a contenidos de alta calidad desde cualquier región del mundo y eliminando barreras geográficas y económicas (Gómez et al., 2022).

Palaute. Herramienta en línea que facilita el análisis automático de comentarios escritos por estudiantes, permitiendo identificar temas y emociones relevantes para

Tabla 5. Estudios seleccionados

Estudio	Autor	Aplicación	Herramienta de Minería de Datos	Técnica de Minería de Datos	Tipo APP
E01	Quijije Quiroz and Maldonado Zuñiga (2023)	Estudiantes con necesidades especiales	Weka	Algoritmos predictivos	Web
E02	Romero and Ventura (2020)	Entornos de aprendizaje en línea	Weka Rapid Miner R y R Studio	Aprendizaje automático	Web Escritorio
E03	Fischer et al. (2020)	Minería de datos masivos	Orange TADA-ED EID-Miner	Patrones secuenciales	Web Escritorio
E04	Khan and Ghosh (2021)	Aprendizaje en el aula	Scikit-learn TensorFlow	Aprendizaje automático	Web
E05	Yağcı (2022)	Aplicación de algoritmos de aprendizaje automático	Weka Rapid Miner Jupyter Notebook	Aprendizaje automático	Web Escritorio
E06	Nebot et al. (2020)	Caja de herramientas e-learning	e-Learning Toolbox	Lógica difusa	Web
E07	Flores et al. (2022)	Uso de método SMOTE	Weka Power BI	Algoritmos predictivos	Web
E08	Caicedo-Castro (2023)	Minería de datos educativos	Course Prophet	Aprendizaje automático	Web
E09	Huatangari et al. (2023)	Uso del algoritmo K-means	Weka	Algoritmo K-means	Web
E10	Nuñez-Mora and Mendoza-Urdiales (2023)	Uso de Modelos GARCH	Tweepy (Python)	Análisis de Sentimientos	Web
E11	Lee et al. (2023)	Uso de IA y minería de datos	MOOC	Mapas de conocimientos y algoritmos predictivos	Web
E12	Gronberg et al. (2021)	Análisis automático de minería de texto	Palaute	Text mining	Web
E13	Okoye et al. (2024)	Aplicación KNN CRISP-DM	R y R Studio	Algoritmos predictivos	Web Escritorio
E14	Moraes et al. (2025)	Análisis de redes sociales de juegos	Jupyter Notebook	Aprendizaje automático	Web

mejorar la comprensión de contenidos específicos mediante minería de datos (Gronberg et al., 2021).

Discusión

Este estudio presenta un mapeo sistemático de las herramientas utilizadas en la capacitación en minería de datos, aplicando la metodología de revisión de literatura. A través de este análisis, se identificaron diversas herramientas que abarcan tanto opciones estándar como aquellas desarrolladas y propuestas por los autores, con el objetivo de optimizar el proceso de enseñanza-aprendizaje. Estas herramientas son fundamentales para enseñar las diferentes técnicas y aplicaciones de la minería de datos, mejorando la comprensión y la aplicación de los conceptos por parte de los estudiantes.

Entre las herramientas más destacadas se encuentran las que facilitan la visualización de datos, como Power BI, que son esenciales para identificar patrones y tendencias en los datos. Las herramientas de programación, como R y R Studio, Scikit-learn, TensorFlow, Jupyter Notebook y Tweepy (Python), permiten a los estudiantes aplicar

técnicas avanzadas de minería de datos, como clasificación, regresión y agrupación.

Las herramientas de interfaces gráficas, tales como Weka, Rapid Miner, TADA-ED y EID-Miner, proporcionan interfaces visuales que simplifican el aprendizaje sin la necesidad de conocimientos profundos en programación. Además, las herramientas de aprendizaje, como e-Learning Toolbox, e-learning, Course Prophet, MOOC y Palaute, optimizan el proceso educativo al ofrecer plataformas interactivas para medir el progreso y rendimiento de los estudiantes.

Conclusiones

Este mapeo sistemático identificó herramientas de soporte para la enseñanza de minería de datos, entre las que destacan Weka, Rapid Miner y otras herramientas especializadas. En relación a las técnicas/algoritmos, el mayor porcentaje corresponde al aprendizaje automático.

Las herramientas de Tecnologías de la Información permiten lograr la enseñanza de la minería de datos utilizando técnicas y algoritmos de este ámbito de estudio. Esta

Tabla 6. Herramientas de soporte en la capacitación de minería de datos.

Herramienta	Descripción
Weka	Herramienta de software libre con una amplia gama de algoritmos de minería de datos.
Rapid Miner	Herramienta de análisis predictivo.
R y R Studio	Herramientas potentes para análisis estadístico.
Orange	Herramienta con módulos de minería de datos y visualización.
TADA-ED	Herramienta específica para minería de datos en educación.
EID-Miner	Herramienta para identificar patrones de aprendizaje en educación.
Scikit-learn	Librería de Python con modelos de regresión, clasificación y predicción.
TensorFlow	Herramienta para modelos avanzados de aprendizaje profundo.
Jupyter Notebook	Herramienta para experimentación de modelos de aprendizaje automático.
Power BI	Herramienta de Microsoft para visualización de datos.
e-Learning Toolbox	Caja de herramientas de e-learning basada en técnicas de minería de datos con lógica difusa.
Course Prophet	Prototipo para predecir fracasos en cursos de aprendizaje automático.
Tweepy (Python)	Biblioteca Python de código abierto y fácil de usar para acceder a la API de Twitter.
MOOC	Herramienta de cursos masivos en línea, ofrece una plataforma flexible y accesible para el aprendizaje.
Palaute	Herramienta innovadora para analizar automáticamente la retroalimentación estudiantil mediante minería de texto.

investigación proporciona un catálogo de herramientas disponibles para que estudiantes y docentes accedan y utilicen estas herramientas de soporte.

Como trabajo futuro se pueden ampliar las preguntas de investigación para determinar detalles específicos de la capacitación, y proponer modelos y metodologías de enseñanza para la minería de datos.

Agradecimientos

Los autores agradecen a las instituciones participantes por el apoyo brindado para la realización de esta investigación.

Referencias

- Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., and Zheng, X. (2016). Tensorflow: aprendizaje automático a gran escala en sistemas distribuidos heterogéneos.
- Alemán Viteri, S. B. (2021). Análisis de sentimientos para twitter con vader y textblob. *Revista Odigos*, 2(3):9–25.
- Babu, B. and Raj, A. (2020). A review of data mining tools: Applications of orange in business intelligence. *International Journal of Data Science and Analytics*, 12(2):213–220.
- Buitinck, L. et al. (2016). Api design for machine learning software: Experiences from the scikit-learn project. *Data Mining and Knowledge Discovery*, 30(2):517–535.
- Caicedo-Castro, I. (2023). Course prophet: a system for predicting course failures with machine learning: a numerical methods case study. *Sustainability*, 15(18):13950.
- Darad, S. and Krishnan, S. (2023). Análisis de sentimiento de los datos de twitter de covid-19 utilizando modelos de aprendizaje profundo y aprendizaje máquina. *Ingenius. Revista de Ciencia y Tecnología*, 29:108–117.
- Espinoza Mina, M. A. (2018). Weka, áreas de aplicación y sus algoritmos: una revisión sistemática de literatura. *Revista Científica ECOCIENCIA*, 5:1–26.
- Fischer, C., Pardos, Z. A., Baker, R. S., Williams, J. J., Smyth, P., Yu, R., and Warschauer, M. (2020). Mining big data in education: Affordances and challenges. *Review of Research in Education*, 44(1):130–160.
- Flores, V., Heras, S., and Julian, V. (2022). Comparison of predictive models with balanced classes using the smote method for the forecast of student dropout in higher education. *Electronics*, 11(3):457.
- Gil, O. G. (2023). Predicción de la presión de burbujeo utilizando aprendizaje automático. *Revista Innovación y Software*, 4(1):204–218.
- Gronberg, N., Knutas, A., Hynninen, T., and Hujala, M. (2021). Palaute: An online text mining tool for analyzing written student course feedback. *IEEE Access*, 9:1–10.
- Gómez, M. J., Calderón, M., Sánchez, V., García Clemente, F. J., and Ruipérez-Valiente, J. A. (2022). Large scale analysis of open mooc reviews to support learners' course selection.
- Henry, F. (2018). Minería de datos. *ReciMundo*, pages 339–349.
- Hertzberg, J., Mundlos, S., Vingron, M., and Gallone, G. (2022). Tada|a machine learning tool for functional annotation-based prioritisation of pathogenic cnvs. *Genome Biology*, 23(1):67.

- Huatangari, L. Q., Ríos, M. A. C., and Camones, R. T. H. (2023). Segmenting the eating behaviour of university students using the k-means algorithm. *Bulletin of Electrical Engineering and Informatics*, 12(4):2363–2371.
- Khan, A. and Ghosh, S. K. (2021). Student performance analysis and prediction in classroom learning: A review of educational data mining studies. *Education and Information Technologies*, 26(1):205–240.
- Kitchenham, B. and Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering. In *Engineering*, volume 2, page 1051.
- Kovar, J. and Šojat, M. (2021). Applications of orange in big data analysis. *Applied Artificial Intelligence*, 35(4):422–433.
- Lee, C. A., Huang, N. F., Tzeng, J. W., and Tsai, P. H. (2023). Ai-based diagnostic assessment system: Integrated with knowledge map in moocs. *IEEE Transactions on Learning Technologies*, 16(5):873–886.
- Lemmerich, F. and Voss, A. (2019). Exploring data with orange: A user-friendly tool for data mining and machine learning.
- Lizana, M. F. (2020). Ventajas de r como herramienta para el análisis y visualización de datos en ciencias sociales. *Revista Científica de la UCSA*, 7(2):97–111.
- López, R. A. B., Guerrero Herrera, H. S., and Salazar González, J. D. (2016). Análisis de clústeres para la clasificación de datos económicos. *Revista Publicando*, 3(7):267–275.
- Lortie, C. J. (2017). Una revisión de r para la ciencia de datos: Elementos clave y un análisis crítico.
- Mazurek, G. and Wasilkowski, P. (2020). Using power bi for business decision support systems. *Journal of Business Intelligence and Analytics*, 14(2):56–67.
- Moraes, L. M., Ferreira, F. P., and Cordeiro, R. L. (2025). Modeling and analyzing social networks of games. *Expert Systems with Applications*, 261:125449.
- Nebot, A., Mugica, F., and Castro, F. (2020). An e-learning toolbox based on rule-based fuzzy approaches. *Applied Sciences*, 10(19):6804.
- Núñez-Mora, J. A. and Mendoza-Urdiales, R. A. (2023). Social sentiment and impact in us equity market: an automated approach. *Social Network Analysis and Mining*, 13(1).
- Okoye, K., Nganji, J. T., Escamilla, J., and Hosseini, S. (2024). Machine learning model (rg-dmml) and ensemble algorithm for prediction of students’ retention and graduation in education. *Computers and Education: Artificial Intelligence*, 6:100205.
- Paguay Hurtado, J. A. (2021). Desarrollo de un prototipo para la predicción de nuevos casos de covid-19 en el ecuador mediante el uso de inteligencia artificial.
- Palma-Avellán, A. M. (2018). Gestión del talento humano desde la perspectiva de capacitación del personal y rendimiento laboral. *Revista Científica FIPCAEC*, 3(7):52–69.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., and Duchesnay, E. (2011). Scikit-learn: Aprendizaje automático en python. *Revista de Investigación en Aprendizaje Automático*, 12:2825–2830.
- Perakovic, D., Remenar, V., and Grgurevic, I. (2020). Posibilidad de aplicar lógica difusa en sistemas de aprendizaje electrónico. In *Conferencia Centroeuropa sobre Información y Sistemas Inteligentes*.
- Pérez, C. and Santín, D. (2014). *Minería de Datos: Técnicas y Herramientas*. Thomson Ediciones Paraninfo, Madrid.
- Quijije Quiroz, H. B. and Maldonado Zuñiga, K. (2023). Técnica de minería de datos para procesos educativos en estudiantes con necesidades educativas especiales basado en un modelo predictivo. *Revista Científica Arbitrada Multidisciplinaria PENTACIENCIAS*, 5(5):205–217.
- Romero, C. and Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 10(3):e1355.
- Shakil, K. A., Anis, S., and Alam, M. (2015). Dengue disease prediction using weka data mining tool.
- VanderPlas, J. (2016). *Python Data Science Handbook: Essential Tools for Working with Data*. O’Reilly Media.
- Yağcı, M. (2022). Educational data mining: prediction of students’ academic performance using machine learning algorithms. *Smart Learning Environments*, 9(1):11.
- Zhang, X., Yao, L., Kanhere, S. S., Liu, Y., Gu, T., and Chen, K. (2017). Mindid: Person identification from brain waves through attention-based recurrent neural network. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence*.
- Zupan, B. and Pečar, S. (2018). Orange: Data mining and visualization tool. *Journal of Computational and Graphical Statistics*, 27(3):728–739.
- Álvarez Peralta, D. H. and Recalde-Varela, P. M. (2020). Optimización de precios en una empresa de retail utilizando la herramienta rapidminer. *Polo del Conocimiento*, 5(3):92–104.