

ORIGINAL

Method with machine learning to carry out feasibility in data mining projects

Método con aprendizaje automático para realizar factibilidad en proyectos de minería de datos

Juan Camilo Giraldo Mejia¹  , Fabio Alberto Vargas Agudelo²  , Jorge Guadalupe Mendoza León²  

¹Tecnológico de Antioquia - Institución Universitaria, Antioquia. Medellín, Colombia.

²Instituto Tecnológico de Sonora- ITSON, Estado de Sonora. México.

Citar como: Giraldo Mejia JC, Vargas Agudelo FA, Mendoza León JG. Method with machine learning to carry out feasibility in data mining projects. Salud, Ciencia y Tecnología. 2026; 6:2446. <https://doi.org/10.56294/saludcyt20262446>

Enviado: 31-07-2025

Revisado: 03-10-2025

Aceptado: 16-11-2025

Publicado: 01-01-2026

Editor: Prof. Dr. William Castillo-González 

Autor para la correspondencia: Juan Camilo Giraldo Mejia 

ABSTRACT

Currently, there are Data Mining techniques aimed at increasing the accuracy of the information and the agility in the analysis. These are applied in the productive sector to characterize behaviors, based on the discovery of knowledge and, in this way, base decision-making in real and dynamic situations. Artificial intelligence (AI) drives research methods and data mining techniques for knowledge acquisition. For its use, the life cycle of data mining projects is followed, which involves stages of extraction, cleaning, preparation and transformation, modeling, and data evaluation. However, it is important to consider a feasibility study for data mining projects, with the objective of positively impacting organizations, by minimizing costly errors and guaranteeing an efficient distribution of resources, as well as the decision on the continuity of a project. This article presents a Machine Learning method to carry out feasibility in data mining projects, seeking to impact organizations by minimizing costly errors and guaranteeing an efficient distribution of resources.

Keywords: Data Mining; Machine Learning; Feasibility in Projects; Knowledge Discovery.

RESUMEN

En la actualidad, existen técnicas de Minería de Datos destinadas a incrementar la exactitud de la información y la agilidad en el análisis. Estas se aplican en el sector productivo para caracterizar comportamientos, apoyándose en el descubrimiento de conocimiento y, de esta manera, fundamentar la toma de decisiones en situaciones reales y dinámicas. La inteligencia artificial (IA) impulsa métodos de investigación y técnicas de minería de datos para la adquisición de conocimiento. Para su uso, se sigue el ciclo de vida de los proyectos de minería de datos, el cual involucra etapas de extracción, limpieza, preparación y transformación, modelado y evaluación de datos. Sin embargo, es importante considerar un estudio de factibilidad para los proyectos de minería de datos, con el objetivo de impactar positivamente en las organizaciones, mediante la minimización de errores costosos y la garantía de una distribución eficiente de recursos, así como de la decisión sobre la continuidad de un proyecto. Este artículo presenta un método de Aprendizaje Automático para realizar factibilidad en proyectos de minería de datos, buscando impactar en las organizaciones minimizando los errores costosos y garantizando una distribución eficiente de los recursos.

Palabras clave: Minería de Datos; Aprendizaje Automático; Factibilidad en Proyectos; Descubrimiento de Conocimiento.

INTRODUCCIÓN

La Minería de Datos tiene como propósito extraer patrones ocultos en grandes cantidades de información y obtener conclusiones relevantes que apoyen la toma de decisiones organizacional. A su vez, la inteligencia artificial impulsa métodos de investigación y técnicas de minería de datos para adquirir conocimiento.⁽¹⁾

En la actualidad, existen técnicas de Minería de Datos destinadas a incrementar la exactitud de la información y agilidad en el análisis.⁽²⁾ Estas se aplican en el sector productivo para el análisis de comportamientos, soportado en el descubrimiento de conocimiento, con el fin de fundamentar la toma de decisiones en situaciones reales y dinámicas. Permiten obtener información valiosa a partir de grandes cantidades de datos provenientes de diversas fuentes; mejorando la eficacia operacional y competitiva de las organizaciones. Entre las técnicas más comunes se encuentran la clasificación, la regresión, el agrupamiento y la extracción de reglas asociativas,⁽³⁾ las cuales permiten a las organizaciones anticipar tendencias, detectar errores, tomar decisiones de acuerdo con la información y mejorar la rentabilidad. Además, facilitan identificar datos ocultos, obtener conclusiones valiosas y guiar decisiones estratégicas en el contexto empresarial y científico.⁽⁴⁾ Asimismo, la minería de datos integra estadística, inteligencia artificial y bases de datos; y obtienen información de diversas fuentes. Su aplicación es amplia, cubriendo varios sectores como el educativo, industrial, finanzas, tecnología, mercados, salud, aportando en la predicción de comportamientos, optimización de recursos y prevención de riesgos ⁽⁵⁾. Para la implementación de estas técnicas se sigue el ciclo de vida para proyectos de minería de datos, el cual involucra etapas de extracción, limpieza, preparación y transformación, modelado, y evaluación de datos. Las fases de preparación y limpieza de datos, así como la selección de métodos y técnicas son fundamentales para dar respuesta a requerimientos de consultas específicas a partir de grandes colecciones de datos.⁽⁶⁾ Sin embargo, es importante considerar un estudio de factibilidad para este tipo de proyectos, ya que permite tomar decisiones sobre su continuidad, minimizando así los errores costosos y garantizando una distribución eficiente de los recursos.⁽⁷⁾ Cabe señalar que las metodologías para proyectos de Minería de Datos no siempre gestionan adecuadamente los elementos de implementación, debido a que es necesario especificar aspectos tecnológicos, organizativos y de datos de modo más efectivo.⁽⁸⁾ La minería de datos es útil para extraer patrones ocultos de grandes conjuntos de información, sin embargo, se presentan desafíos en la gestión y administración de los proyectos. No se logra un proyecto de minería de datos sin una estructura y un propósito, lo que implica realizar un estudio de viabilidad técnica, económica, y operativa. Es importante realizar un análisis de elementos financieros, técnicos, operativos, legales y de programación, con el objetivo de proponer un método que apoye a los analistas de decisiones para estimar y determinar la viabilidad de proyectos de minería de datos, así como en la identificación de riesgos o retos potenciales.^(9,10)

En las empresas, las herramientas tecnológicas juegan un papel importante para cumplir con los objetivos de un proyecto, por lo cual se requiere recursos técnicos, como sistemas de información que sean cruciales para ejecutar las actividades, además de modelar y analizar los datos de manera ágil. Así también, son necesarios recursos físicos, en cuanto a equipos con capacidad de procesamiento, velocidad y servidores que soporten grandes cantidades de información. Para que las organizaciones realicen estudios de viabilidad en proyectos de Minería de Datos; es necesario determinar el proceso de negocio, identificar debilidades y fuentes confiables, innovar e incluir nuevas prácticas, establecer equipo humano estratégico, incorporar modularidad, y definir criterios para toma de decisiones. La viabilidad operativa, determina si un proyecto puede sostenerse en el tiempo y sea exitoso en la organización. Para ello, es necesario involucrar un equipo humano capacitado para operar y analizar datos utilizando las diferentes herramientas tecnológicas implementadas.⁽¹⁰⁾

Asimismo, todo proyecto de minería de datos requiere una inversión financiera, por lo cual se debe validar los costos iniciales, los costos operativos, los beneficios económicos esperados y la optimización de procesos, ya que, si los gastos superan el retorno de la inversión, el proyecto no es factible. Lo anterior implica contar con información completa, suficiente y clara, lo que a su vez requiere un proceso de depuración para minimizar o preferentemente eliminar los datos innecesarios. Se sugiere mejorar los procesos y emplear técnicas específicas al momento de establecer los requisitos específicos del proyecto de minería de datos, teniendo en consideración el tamaño de los datos, el alcance la complejidad del proyecto, así como los recursos disponibles, para facilitar el diseño, la ejecución y el monitoreo de los flujos de trabajo.⁽¹¹⁾ La minería de datos aporta grandes beneficios a las organizaciones en el mejoramiento de procesos; sin embargo, se tienen retos con la calidad de los datos, la escalabilidad de los sistemas y cuestiones éticas. Las metodologías existentes para proyectos de minería de datos suelen aplicarse sin cambios detallados y extensiones con nuevas fases o etapas, que requieren las necesidades tecnológicas y organizacionales. Por ello, se sugiere optimizarlas mediante la incorporación de los elementos y recursos necesarios a nivel tecnológico, operativo, organizacional y de datos.⁽¹¹⁾

MÉTODO

Paso 1- revisión de características y etapas de metodologías para proyectos de Minería de Datos

Se revisaron las características de tres de las metodologías más usadas para desarrollar proyectos de Minería de Datos. Para conocer la dinámica que sigue cada una de ellas en sus diferentes fases o etapas.

Metodología SEMMA

Es una de las metodologías más usadas en el contexto de proyectos de minería de datos. Está estructurada a partir de cinco etapas: muestreo, consiste en seleccionar una muestra de datos a partir de una fuente específica. La segunda etapa conocida como exploración permite entender la estructura de los datos obtenidos previamente. Posteriormente en la etapa tres, se realiza la preparación y transformación de los datos. Una vez preparados los datos se pasa a la etapa de modelado, en esta se seleccionan y aplican algoritmos que permiten generar modelos predictivos o descriptivos. La etapa de evaluación de la funcionalidad y utilidad de los modelos de minería de datos.^(12,13) Asimismo, esta metodología se aplica en una variedad de contextos.^(14,15)

Metodología CRISP-DM

La metodología CRISP-DM surgió de la necesidad de establecer un modelo de proceso estándar.^(16,17) Consta de seis etapas, comprensión del negocio, en ella se establecen los propósitos de la organización, y el alcance de los indicadores del proceso. En la etapa dos se obtienen los datos de diferentes fuentes. En la etapa tres, preparación de los datos, se realiza limpieza, y transformación de datos. En la etapa cuatro se aplican técnicas de minería para generar modelos que exploren la articulación entre los datos. En la etapa cinco se evalúan los modelos obtenidos teniendo en cuenta su rendimiento. La última etapa se denomina despliegue; usan los modelos generados para beneficio de la organización, incorporándolos a la dinámica.

Metodología Catalyst

Fue validada por la comunidad científica y de la industria para el desarrollo de proyectos de minería de datos, habiendo sido creada alrededor del año 2000, al mismo tiempo que CRISP-DM.^(18,19) Se soporta en seis etapas: comprensión del negocio, en la cual se especifica el alcance del proyecto. La etapa dos se conoce como comprensión de los datos, en la cual se identifican las fuentes de datos. En la etapa tres se preparan los datos, realizando actividades para garantizar su calidad. En la fase cuatro se realiza el modelado, a través de herramientas y la técnica adecuada para dar solución a los propósitos de la minería de datos. La etapa cinco corresponde a la evaluación de los modelos obtenidos. La etapa seis denominada implementaciones, en esta se generan informes y reportes.^(18,19,20)

Paso 2- Análisis de características y actividades de las metodologías

Las tres metodologías revisadas son muy utilizadas para desarrollar proyectos de minería de datos, cada una con fases específicas que permiten guiar con buenas prácticas hasta generar modelos predictivos, o descriptivos que den respuesta a las necesidades organizacionales. Sin embargo, estas metodologías no presentan actividades relacionadas con estudios o análisis de factibilidad o viabilidad de los proyectos de Minería de Datos. No indican actividades que inviten a conocer y establecer los recursos técnicos necesarios para realizar el proyecto, tampoco aspectos relacionados con la factibilidad económica, y operativa del mismo. Véase la tabla 1.

Tabla 1. Comparativo de actividades que se realizan con las metodologías revisadas					
Metodología	Actividades				
	Conocer la organización y el proceso	Conocimiento y preparación de Datos	Modelado	Evaluación y Despliegue	Factibilidad técnica, económica, y operativa
SEMMA	No realiza	Se realiza	Se realiza	Se realiza	No realiza
CRISP-DM	Se realiza	Se realiza	Se realiza	Se realiza	No realiza
CATALYST	Se realiza	Se realiza	Se realiza	Se realiza	No realiza

Las metodologías son claras y detalladas respecto a la especificación del alcance de los proyectos, a partir de la selección y descripción de los procesos organizacionales que serán intervenidos con el análisis de minería de datos. Además, se establecen los propósitos del análisis a realizar, y se establecen los indicadores de proceso que marcan el alcance del proyecto.

Paso 3- Especificación de las fases para el método propuesto

El método debe permitir articularse con las fases de cualquiera de las metodologías usadas para proyectos de minería de datos, metodologías como CRISP-DM, SEMMA, Catalyst, Kdd, Kimball, entre otras. El método se estructura con tres fases o etapas. La primera se denomina especificación del proyecto, la segunda fase análisis de resultados, y la tercera visualización. Véase la figura 1.

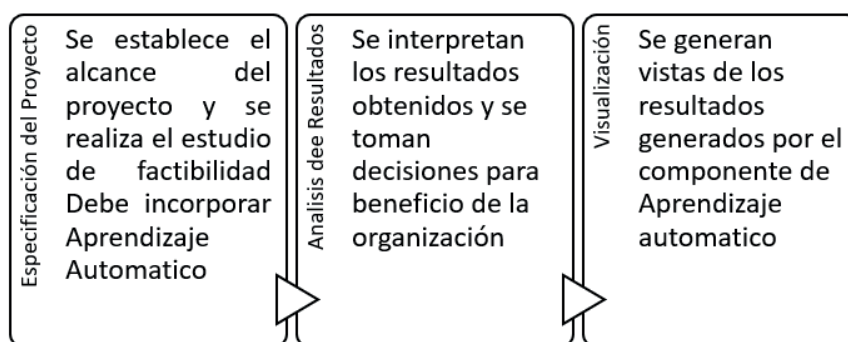


Figura 1. Fases del método para factibilidad de proyectos de minería de datos

Paso 4- Especificación de componentes del método

En el caso de las metodologías CRISP-DM, y SEMMA, en cada una, el método se articula con las fases de identificación del negocio, o conocimiento del negocio. Si la metodología no cuenta con una fase o componente en el cual se contextualiza el negocio, el método propuesto se integra como una nueva etapa o fase de la metodología, siendo la fase uno o inicial de la misma, como se indica en la figura 2.

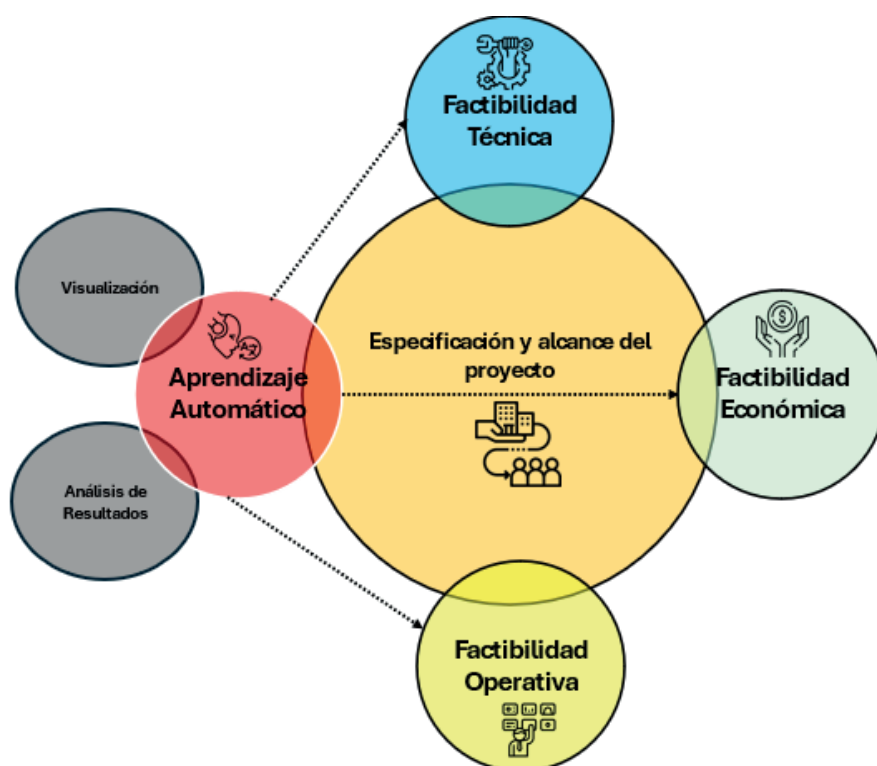


Figura 2. Método con aprendizaje automático -Factibilidad en proyectos de minería de datos

Fase 1 - Especificación y alcance del proyecto

En esta etapa se realiza la identificación y se describe con detalle el proceso a intervenir con Minería de Datos, conociendo las actividades que lo componen, y el flujo de datos. Además, se realizan actividades para determinar la factibilidad técnica, económica, y operativa del proyecto.

Factibilidad técnica

Se especifica la cantidad de datos y las fuentes de datos (tipos), los algoritmos que se utilizarán para la generación de modelos de minería, la infraestructura necesaria a nivel de hardware y software. Se realiza limpieza de datos. Asimismo, se identifica el tipo de algoritmo y la complejidad, de acuerdo con las características del proyecto de minería de Datos.

Para facilitar la dinámica de esta fase se articula Aprendizaje Automático, con el propósito de usar datos del proyecto, previamente realizados. Esta información se encuentra almacenada en una base de conocimiento.

Aprendizaje Automático: el modelo recoge los datos de proyectos de minería de datos ya finalizados. Con esta información se puede predecir resultados de factibilidad, a partir del entrenamiento de técnicas específicas de Aprendizaje Automático. El modelo indicará si el proyecto es técnicamente viable con un nivel de probabilidad.

Factibilidad Económica

Se determina el costo total del proyecto teniendo en cuenta los recursos de hardware, software, personal, retorno de inversión, y presupuesto inicial.

Aprendizaje automático: el modelo predice el retorno de inversión y el costo final del proyecto. Para esto, se puede usar un modelo de regresión como, por ejemplo, Regresión Lineal, o un modelo de decisiones basado en árboles de decisión, como, por ejemplo, Predicción de Precios de Viviendas, cuando el problema es que una empresa inmobiliaria quiere predecir el precio de venta de viviendas, basándose en características estructurales y de ubicación. Complementariamente, la inferencia proporciona una proyección del retorno de inversión e indica si el proyecto es económicamente viable con un determinado nivel de probabilidad.

Factibilidad Operativa

Se establece el equipo de trabajo, determinando la cantidad de recurso humano requerido, el tiempo necesario para la implementación, la puesta en marcha, y el monitoreo del proyecto.

Aprendizaje automático: el modelo predice el tiempo requerido para la implementación, puesta en marcha y monitoreo del proyecto, así como la cantidad de recursos humanos necesarios según sus características específicas. Al igual que en el modelo técnico, puede emplearse un modelo de clasificación o de regresión, en función de las variables dependientes e independientes definidas para el entrenamiento. Como resultado, el modelo indica la viabilidad operativa del proyecto con un nivel de probabilidad determinado.

Fase II - Análisis de Resultados para Toma de Decisiones

De acuerdo con las necesidades previamente establecidas para el desarrollo del proyecto de minería de datos, las predicciones obtenidas en cada una de las actividades de factibilidad se interpretan y se utilizan para la toma de decisiones.

Actividades

Obtención de Datos: los datos se extraen de diversas fuentes, específicamente de los modelos generados por el componente de aprendizaje automático en las tres fases de factibilidad. Los datos se obtienen en formato de valores separados por comas (CSV) y constituyen el insumo para su revisión y análisis. Para facilitar la interpretación de resultados, se emplean herramientas de visualización.

Fase III - Visualización

De forma gráfica se obtiene una vista de los datos generados por el componente de aprendizaje automático, y se cargan la información relacionada con el proyecto, respecto al alcance, permitiendo tomar decisiones más acertadas respecto a las necesidades del proceso organizacional.

RESULTADOS**Fase I - Especificación y Alcance del Proyecto**

Especificación del proceso Gestión de ventas: el proceso corresponde a la gestión de ventas; relacionado con las estrategias y buenas prácticas que aplica la organización para gestionar sus clientes.

Propósito del análisis: las ventas de la organización se han visto afectadas por la deserción de clientes. La organización planeó un análisis de datos que determine cuales son los clientes con riesgo de desertar. Esto permitirá a la organización establecer un conjunto de estrategias de retención.

Se implementó un modelo predictivo de análisis de datos, que determinará a partir de ciertas características demográficas, los clientes con alta probabilidad de desertar.

Indicadores para Análisis de datos

Se establecieron indicadores a el comportamiento de los clientes respecto a las ventas, determinando la articulación con nivel de satisfacción, gustos, inclinaciones o preferencias por productos.

Factibilidad Técnica

Fuente de datos: se identificaron datos en formato CSV, con aproximadamente 2 mil registros, estructurados con las características de ventas, las cuales muestran datos de clientes, productos, precios, tipos de productos, categorías, sucursales, empleados, valores facturados. Los datos se tomaron de las bases de datos transaccionales de ventas, pqrs, gestión de clientes.

Almacenamiento de Datos: los datos se consolidaron en una Bodega de Datos, conformada por tablas de dimensiones y tabla de hechos. Dimensiones (producto, sucursal, empleado, clientes, y preferencias), y una tabla de hechos que registra el movimiento de eventos históricos que se van generando desde la transaccionalidad de las ventas a clientes.

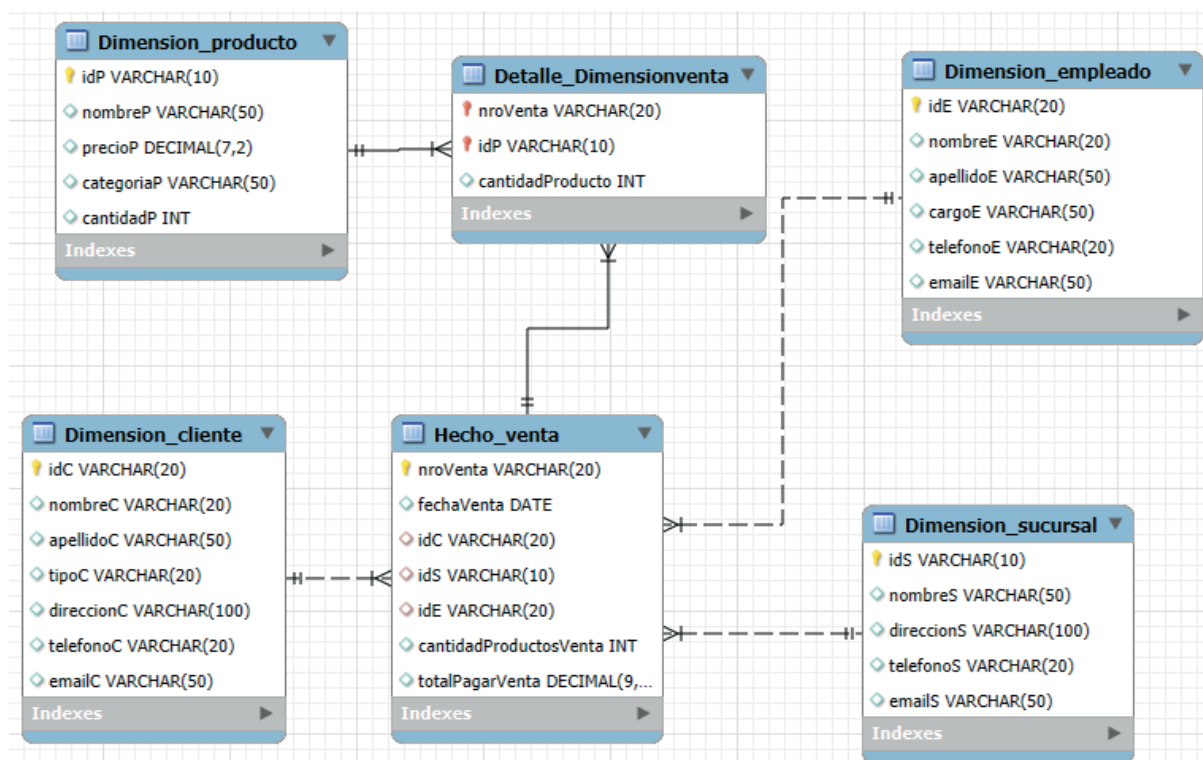


Figura 3. Bodega de Datos - eventos de ventas clientes

Base de Conocimiento Factibilidad Técnica: identificación del proyecto técnico, fecha de ejecución del proyecto, fecha fin del proyecto, propósito del proyecto, origen de datos, modelo de aprendizaje automático propuesto, costo estimado de factibilidad técnica, tiempo estimado, indicadores de resultado de factibilidad técnica.

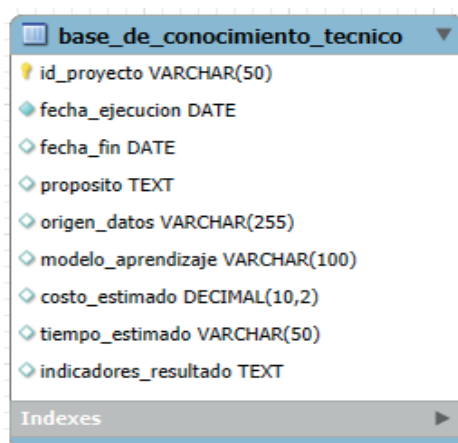


Figura 4. Base de Conocimiento - Factibilidad Técnica

Aprendizaje automático: se generó un modelo utilizando Python (librerías Pandas), entrenando un algoritmo de clasificación de Red Neuronal. La fuente que se utilizó para el entrenamiento es la base de conocimiento de factibilidad técnica. La inferencia presentó una probabilidad de viabilidad técnica del 82 %.

Factibilidad Económica

Se tiene en cuenta un tiempo de tres meses que dura el proyecto, el costo del Software y hardware necesarios para la implementación de modelos de aprendizaje automático, y para soportar las bases de conocimiento, y aplicaciones.

Costo de Software: licencias de código abierto para MySQL, Power BI, y Python.

Costo de recurso humano: científico de Datos, Analista de Datos. Treinta y seis millones de pesos (36 000 000).

Costo de Hardware: licencias para nube, y servidor dedicado para base de datos y aplicaciones. Cuarenta

millones de pesos (40 000 000).

Costo total del proyecto: setenta y seis millones de pesos (76 000 000).

Base de conocimiento económico

En este espacio se encuentran los registros históricos correspondientes a otros proyectos a los que se les calculó, los costos de personal, hardware, software, costo total del proyecto, el porcentaje de Retorno de Inversión. Todo esto teniendo en cuenta el tiempo de duración del proyecto, y el presupuesto disponible por la organización.

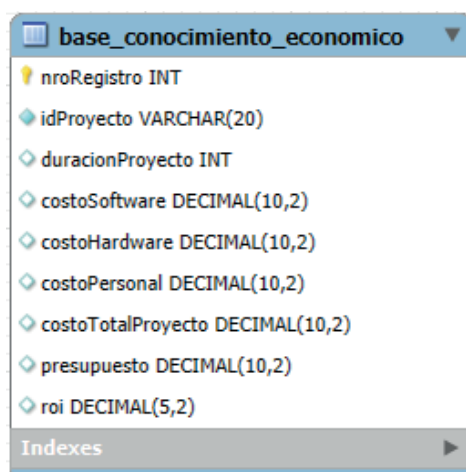


Figura 5. Base de Conocimiento - Factibilidad Económica

Aprendizaje automático para estimación económica: se generó un modelo utilizando Python (librerías Pandas), entrenando un algoritmo de Árboles de Decisión. La fuente que se utilizó para el entrenamiento es la base de conocimiento de factibilidad Económica. La conclusión del modelo indica que el proyecto es viable económicamente, teniendo en cuenta que presenta un Retorno de inversión alto, y el costo total del proyecto está dentro del presupuesto organizacional.

Factibilidad Operativa

Base de conocimiento operativa: este almacenamiento contiene eventos históricos relacionados con la factibilidad operativa determinada para otros proyectos, a partir de la duración, el total de recurso humano involucrado en el proyecto, y el rol y alcance que cada uno de ellos tiene para aportar al desarrollo, despliegue, y administración.

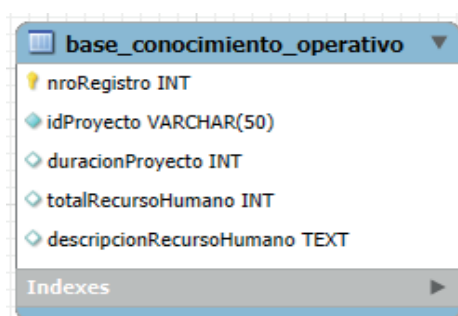


Figura 6. Base de Conocimiento - Factibilidad Operativa

Aprendizaje automático para estimación operativa: se utilizó un algoritmo de clasificación, y se generó un modelo que permitió determinar que el proyecto tiene viabilidad operativa, teniendo en cuenta los requerimientos de recursos, tiempo estimado, y esfuerzo requerido. Tiene una probabilidad del 79 %, a partir de otros proyectos establecidos y ejecutados. El modelo se entrenó con los eventos de la base de conocimiento operativa.

Fase II - Análisis de Resultados para la toma de decisiones

Viabilidad del proyecto: el proyecto es viable, se aprueba el desarrollo y despliegue del proyecto. Se cuenta con el presupuesto, el personal requerido, y está dentro de los tiempos estimados.

Fase III- Visualización

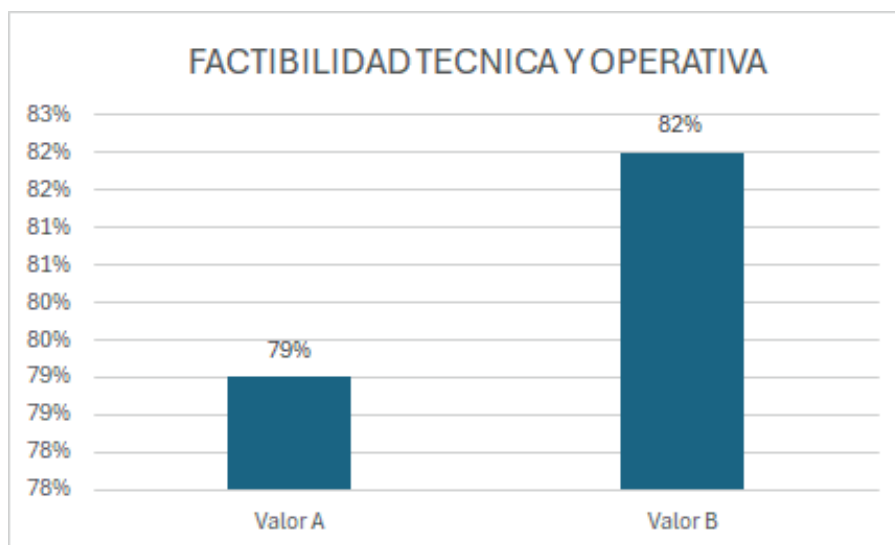


Figura 7. Comportamiento Factibilidad Técnica y Operativa

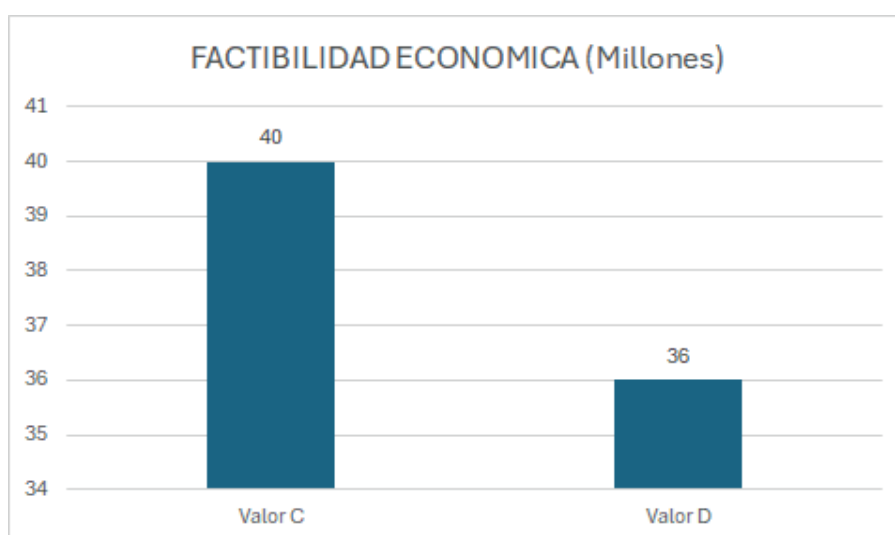


Figura 8. Comportamiento Factibilidad Económica

DISCUSIÓN

La implementación del método en las fases metodológicas de la minería de datos contribuye a la reducción de sobrecostos, el mejoramiento de la calidad proyectual y la distribución eficiente de los recursos humanos requeridos para las diversas actividades. Adicionalmente, permite confirmar la rentabilidad del proyecto posterior a su implementación y monitoreo, lo que justifica la inversión y asegura el cumplimiento de los objetivos estratégicos establecidos durante la fase de planificación.

CONCLUSIONES

La incorporación del método en las fases de las metodologías de minería de datos permite controlar sobrecostos, mejorar la calidad de los proyectos y optimizar la distribución de los recursos humanos necesarios para las distintas actividades. Asimismo, posibilita confirmar la rentabilidad del proyecto una vez implementado y monitoreado, lo que justifica la inversión y asegura el logro de los objetivos estratégicos establecidos durante la planificación.

La factibilidad de un proyecto de minería de datos requiere un análisis riguroso y equilibrado. La adopción de un enfoque integral que aborde las dimensiones técnica, económica y operativa incrementa las probabilidades de éxito de las organizaciones.

Esta propuesta busca fomentar la confianza entre los responsables de tales iniciativas, garantizando que resulten no solo tecnológicamente funcionales, sino también rentables, éticas y genuinamente transformadoras para la organización.

REFERENCIAS BIBLIOGRÁFICAS

1. Funes A, Dasso A. Methods and Techniques of Data Mining. En: Encyclopedia of Information Science and Technology, Fifth Edition. IGI Global; 2021. p. 749-767. Doi: 10.4018/978-1-7998-3479-3.ch045.
2. Zhu S. Analysis of the severity of vehicle-bicycle crashes with data mining techniques. Journal of Safety Research (J Saf Res). 2021. P. 76:218-227. Doi: \$10.1016/j.jsr.2020.11.011\$
3. P R, Nayak PP, Poojary P, Talekar PP. Data Mining: Concepts, Techniques, and Applications. Int J Adv Res Sci Commun Technol. 2024;120-128. doi:10.48175/ijarsct-2282.
4. Begum DPI, Banu DN. Data mining techniques. 2024. p. 188-206. Doi:10.58532/v3bfit2p6ch2.
5. Xie B, Zhang F. Design and Implementation of Data Mining in Information Management System. En: 2022 International Conference on Knowledge Engineering and Communication Systems (ICKES). IEEE; 2022. p. 1-5. Doi: 10.1109/ICKES58461.2022.9995817
6. Thakur A. A systematic review on data mining methods and applications. Int J Adv Res Comput Sci. 2022;13(2):28-31. doi:10.26483/ijarcs.v13i2.6807. Doi: 10.26483/ijarcs.v13i2.6807
7. Kumar M. FEASIBILITY STUDY & INPUT/ OUTPUT FORM DESIGN. GRF BOOKS; 2023. doi: 10.52458/9788196869434.2023.eb.grf.ch-12.
8. Plotnikova V, Dumas M, Milani F. Adaptations of data mining methodologies: a systematic literature review. PeerJ. 2020;6. doi:10.7717/PEERJ-CS.267.
9. Vasiliev AA, Goryachev AV. Models and Methods of Data Mining in Project Management. En: 2022 International Conference on Quality Management, Transport and Information Security, Information Technologies (IT&QM&IS). IEEE; 2022. p. 52-56. Doi: 10.1109/ElConRus54750.2022.9755501.
10. McLeod S. Feasibility studies for novel and complex projects: Principles synthesised through an integrative review. 2021; 2:100022. Doi: 10.1016/j.plas.2021.100022.
11. Renggli C, Rimanic L, Kolar L, Wu W, Zhang C. Automatic Feasibility Study via Data Quality Analysis for ML: A Case-Study on Label Noise. 2023. Doi: 10.1109/ICDE55515.2023.00024.
12. Gómez Palacios HJ, Jiménez Toledo RA, Hernández Pantoja GA, Martínez Navarro ÁA. A comparative between CRISP-DM and SEMMA through the construction of a MODIS repository for studies of land use and cover change. Adv Sci Technol Eng Syst J. 2017;2(3):598-604. Doi: 10.25046/aj020376.
13. Sangacha-Tapia L, González-Cañizalez Y, Rivas-Herrera J. Optimización de Criterios de Búsqueda avanzada para Nuevas Tendencias en la Académica mediante Machine Learning. Rev Cient Zambos. 2025;4(2):197-211. Doi:10.69484/rcz/v4/n2/114.
14. Alik A, Mirza H, Andri, Ferdiansyah A. Classification Of South Sumatra Songket Woven Fabric Motifs Using Deep Learning. Data J Inf Syst Manag. 2024;2(2):24-35. Doi:10.61978/data.v2i1.
15. Castro Coria EG, Ruiz Flores López P. Guía metodológica para el uso de minería de datos en la Plataforma Nacional de Transparencia. Estud Derecho Inf. 2024;(17). Doi:10.22201/ij.25940082e.2024.17.18782.
16. Chapman P, Clinton J, Kerber R, Khabaza T, Reinartz T, Shearer C, et al. CRISP-DM 1.0: Step-by-step Data Mining Guide. SPSS; 2000.
17. Espinosa-Zúñiga JJ. Aplicación de metodología CRISP-DM para segmentación geográfica de una base de datos pública. Ing Investig Tecnol. 2020; XXI (1):1-17. Doi:10.22201/fi.25940732e.2020.21n1.008.
18. Eckert KB. Modelo basado en la Toma Decisiones con Criterios Múltiples para la elección de Metodologías de Data Science. Universidad Nacional de Misiones; 2019.
19. Moine JM, Gordillo SE, Haedo AS. Análisis comparativo de metodologías para la gestión de proyectos de minería de datos. En: XVII Congreso Argentino de Ciencias de la Computación. 2011.

20. Rodríguez Montequín MT, Álvarez Cabal JV, Mesa Fernández JM, González Valdés A. Metodologías para la realización de proyectos de Data Mining. En: VII Congreso Internacional de Ingeniería de Proyectos. 2003. p. 257-265.

FINANCIACIÓN

Ninguna.

CONFLICTO DE INTERESES

Ninguno.

CONTRIBUCIÓN DE AUTORÍA

Conceptualización: Fabio Alberto Vargas, Juan Camilo Giraldo Mejía, Fabio Alberto Vargas Agudelo, Jorge Mendoza.

Análisis formal: Fabio Alberto Vargas, Juan Camilo Giraldo Mejía, Jorge Mendoza.

Investigación: Fabio Alberto Vargas, Juan Camilo Giraldo Mejía, Fabio Alberto Vargas Agudelo, Jorge Mendoza.

Metodología: Fabio Alberto Vargas, Juan Camilo Giraldo Mejía, Jorge Mendoza.

Administración del proyecto: Fabio Alberto Vargas, Juan Camilo Giraldo Mejía, Jorge Mendoza.

Recursos: Fabio Alberto Vargas, Juan Camilo Giraldo Mejía, Jorge Mendoza.

Arquitectura: Fabio Alberto Vargas, Juan Camilo Giraldo Mejía, Fabio Alberto Vargas Agudelo, Jorge Mendoza.

Supervisión: Fabio Alberto Vargas, Juan Camilo Giraldo Mejía, Jorge Mendoza.

Validación: Fabio Alberto Vargas, Juan Camilo Giraldo Mejía, Jorge Mendoza.

Redacción - borrador original: Fabio Alberto Vargas, Juan Camilo Giraldo Mejía, Jorge Mendoza.

Redacción - revisión y edición: Fabio Alberto Vargas, Juan Camilo Giraldo Mejía, Jorge Mendoza.