

Análisis predictivo y gestión optimizada de la cartera de clientes morosos en CNEL EP, Unidad de Negocio Bolívar

Predictive Analysis and Optimized Management of the Delinquent Customer Portfolio at CNEL EP Bolívar Business Unit

Marco Vinicio Carrillo Trujillo¹ <https://orcid.org/0009-0003-7699-1728>,
Pedro Stalyn Aguilar Encarnación¹ <https://orcid.org/0009-0005-1664-2280>

¹Escuela Superior Politécnica de Chimborazo, ESPOCH Sede Orellana, Ecuador
marcov.carrillo@epoch.edu.ec, pedro.aguilar@epoch.edu.ec



Esta obra está bajo una licencia internacional
Creative Commons Atribución-NoComercial 4.0.

Enviado: 2025/04/15
Aceptado: 2025/05/15
Publicado: 2025/06/30

Resumen

La morosidad representa un obstáculo estructural para la sostenibilidad financiera de las empresas públicas de distribución eléctrica; sin embargo, a diferencia del sector financiero, aún no se han adaptado modelos predictivos basados en aprendizaje automático a este contexto. En particular, CNEL EP enfrenta dificultades para anticipar el incumplimiento de sus clientes, lo cual limita la eficacia de su gestión comercial. Ante esta problemática, el presente estudio tiene como objetivo determinar el modelo de aprendizaje automático más preciso para predecir el riesgo de morosidad en la Unidad de Negocio Bolívar. Para ello, se adoptó un enfoque metodológico basado en Design Science Research y CRISP-DM, el cual permitió integrar una revisión sistemática de literatura (PRISMA), el análisis de un conjunto de 72.483 registros históricos y la aplicación de técnicas como PCA, SMOTE y modelos ensemble (RandomForest, Gradient Boosting, AdaBoost y VotingClassifier). Gradient Boosting y VotingClassifier alcanzaron métricas casi perfectas (Accuracy: 0.9982; F1 Macro: 0.9957; AUC ROC: 1.000) y (Accuracy: 0.9983; F1 Macro: 0.9959; AUC ROC: 1.000), incluso en escenarios de estrés con ruido, desbalance y pérdida de datos; además, la incorporación de SHAP y LIME facilitó la interpretación de las predicciones, garantizando transparencia para usuarios no técnicos. Los hallazgos evidencian que la solución es robusta, replicable y aplicable en la práctica. Este estudio constituye un aporte significativo al demostrar que los modelos de aprendizaje automático pueden fortalecer la gestión de cartera en el sector eléctrico público ecuatoriano.

Sumario: Introducción, Metodología, Resultados, Discusión, Conclusiones.

Como citar: Carrillo, M. & Aguilar, P. (2025). Análisis predictivo y gestión optimizada de la cartera de clientes morosos en CNEL EP, unidad de negocio bolívar. *Revista Tecnológica - Espol*, 37(1), 285-308. <https://rte.espol.edu.ec/index.php/tecnologica/article/view/1312>

Palabras clave: Aprendizaje Automático, Ciencia del Diseño, Evaluación de Modelos, Modelos Ensemble, Minería de Datos, Inteligencia Artificial Explicable, Riesgo de Morosidad.

Abstract

Payment delinquency represents a structural barrier to the financial sustainability of public electricity distribution companies; however, unlike the financial sector, predictive models based on machine learning have not yet been adapted to this context. In particular, CNEP faces challenges in anticipating customer defaults, which limits the effectiveness of its commercial management. This study aims to identify the most accurate machine learning model to predict delinquency risk in the Bolívar Business Unit. A methodological approach based on Design Science Research and CRISP-DM was adopted, incorporating a systematic literature review (PRISMA), the analysis of 72,483 historical records, and the application of techniques such as PCA, SMOTE, and ensemble models (RandomForest, Gradient Boosting, AdaBoost, and VotingClassifier). Gradient Boosting and VotingClassifier achieved near-perfect performance metrics (Accuracy: 0.9982; F1 Macro: 0.9957; AUC ROC: 1.000) and (Accuracy: 0.9983; F1 Macro: 0.9959; AUC ROC: 1.000), even under stress scenarios involving noise, imbalance, and data loss. Furthermore, the integration of SHAP and LIME enabled transparent interpretation of predictions for non-technical users. The findings show that the proposed solution is robust, replicable, and applicable in practice. This study makes a significant contribution by demonstrating that machine learning models can enhance portfolio management in Ecuador's public electricity sector.

Keywords: Data Mining, Default Risk, Design Science, Ensemble Models, Machine Learning, Model Evaluation, Explainable Artificial Intelligence.

Introducción

La morosidad en la cartera de clientes constituye un desafío estructural para la sostenibilidad financiera de las empresas públicas de distribución eléctrica; en contraste, en el sector financiero se han desarrollado modelos de aprendizaje automático altamente efectivos para predecir el riesgo de incumplimiento (Liu et al., 2024; Xu, 2024); no obstante, estos avances aún no han sido adaptados ni validados en contextos operativos del sector eléctrico, donde las particularidades institucionales y regulatorias exigen soluciones especializadas para anticipar el riesgo y optimizar la recuperación de cartera (Akinjole et al., 2024; Ha et al., 2024; Rios & Arbeláez, 2023).

En la literatura científica se han documentado múltiples enfoques exitosos de clasificación del riesgo crediticio, destacando el uso de modelos ensemble como Random Forest, Gradient Boosting y Voting Classifier, junto con técnicas de balanceo de clases como SMOTE y métodos de reducción de dimensionalidad como PCA (Dastile et al., 2020; Kim et al., 2020; Ren, 2025); sin embargo, la mayoría de estas aplicaciones se han desarrollado en el sector bancario, bajo condiciones de mercado muy distintas a las del sector eléctrico público.

No se han documentado estudios empíricos aplicados al sector eléctrico público que implementen y validen modelos de aprendizaje automático para la predicción del riesgo de morosidad, considerando la complejidad operativa, los esquemas tarifarios diferenciados y los factores sociales que influyen en el comportamiento de pago de los usuarios (Ha et al., 2024; Rios & Arbeláez, 2023). Esta brecha representa un obstáculo para incorporar soluciones tecnológicas en la gestión comercial de las empresas públicas de energía. (Mujo et al., 2025; Wang et al., 2025).

El objetivo del estudio es, determinar el modelo de aprendizaje automático más preciso mediante un proceso de minería de datos para gestionar de forma efectiva a los clientes en riesgo de morosidad en CNEL EP Unidad de Negocio Bolívar.

La investigación busca proporcionar una solución replicable, robusta y basada en evidencia que permita automatizar la clasificación de clientes morosos, reducir las pérdidas por incobrabilidad y fortalecer la gestión comercial mediante la incorporación de inteligencia artificial (Kim et al., 2020; Ren, 2025; Xu, 2024). Asimismo, aporta al desarrollo del conocimiento académico al extender la aplicación de modelos de aprendizaje automático hacia dominios no financieros, demostrando su utilidad en contextos operativos complejos como la gestión de cartera en empresas públicas del sector eléctrico (Akinjole et al., 2024; Ha et al., 2024).

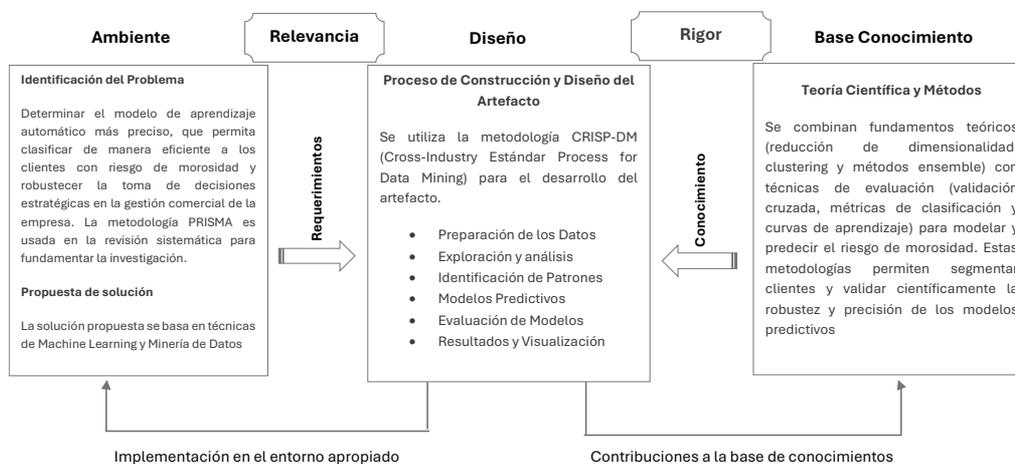
El artículo se organiza en cinco secciones: Sección 1, “Introducción, presentación del tema, relevancia, brecha de conocimiento, objetivo y justificación”; Sección 2, “Metodología, describe el enfoque Design Science Research y CRISP-DM para el desarrollo del artefacto, detalla las fases de Entendimiento del Negocio, Comprensión y Preparación de datos, Modelado y Evaluación”; Sección 3, “Resultados de los modelos implementados, su rendimiento comparativo, pruebas de estrés”; Sección 4, “Discusión, contraste entre los resultados alcanzados y los reportados en investigaciones previas”; Sección 5, “Conclusiones, se exponen las principales conclusiones y se sugieren futuras líneas de investigación para optimizar y ampliar la solución desarrollada.

Metodología

El estudio se desarrolló bajo el marco metodológico de la Ciencia del Diseño (Design Science Research, DSR), un enfoque que permite abordar problemas complejos mediante la construcción y validación de artefactos tecnológicos en contextos reales (Akoka et al., 2023; Auna & Richard, 2023; Goecks et al., 2021). Esta metodología ha demostrado ser particularmente eficaz en el desarrollo de modelos predictivos aplicables a contextos organizacionales, al permitir una articulación efectiva entre teoría y práctica. Su enfoque se sustenta en la interacción dinámica de tres ciclos fundamentales: el ciclo de rigor, el ciclo de relevancia y el ciclo de diseño (Delpont et al., 2024; Goecks et al., 2021; Montevechi et al., 2024). La Figura 1 detalla las actividades desarrolladas en cada ciclo.

Figura 1

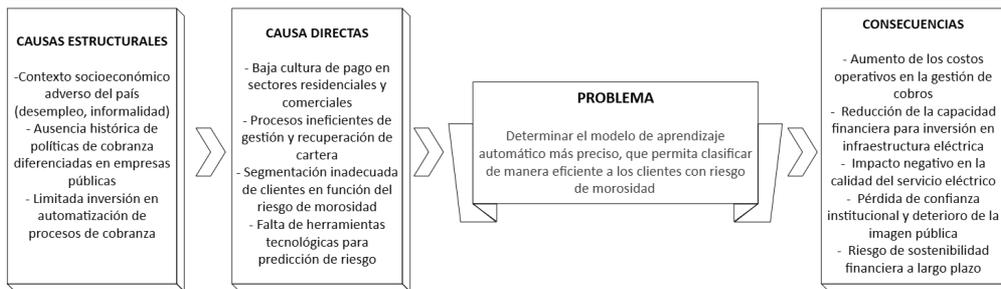
Metodología de la Ciencia del Diseño Aplicada (DSR)



El Ciclo de Relevancia vinculó la solución propuesta con una necesidad institucional crítica: anticipar el riesgo de morosidad para optimizar la gestión de cartera. Esta necesidad se asocia con el problema formulado en la presente investigación, el cual muestra sus causas estructurales y efectos en la Figura 2. En consecuencia, la investigación inició con un proceso de revisión de la literatura a fin de determinar los modelos de aprendizaje automático más utilizados en la predicción del riesgo de morosidad y la recuperación de cartera en contextos generales.

Figura 2

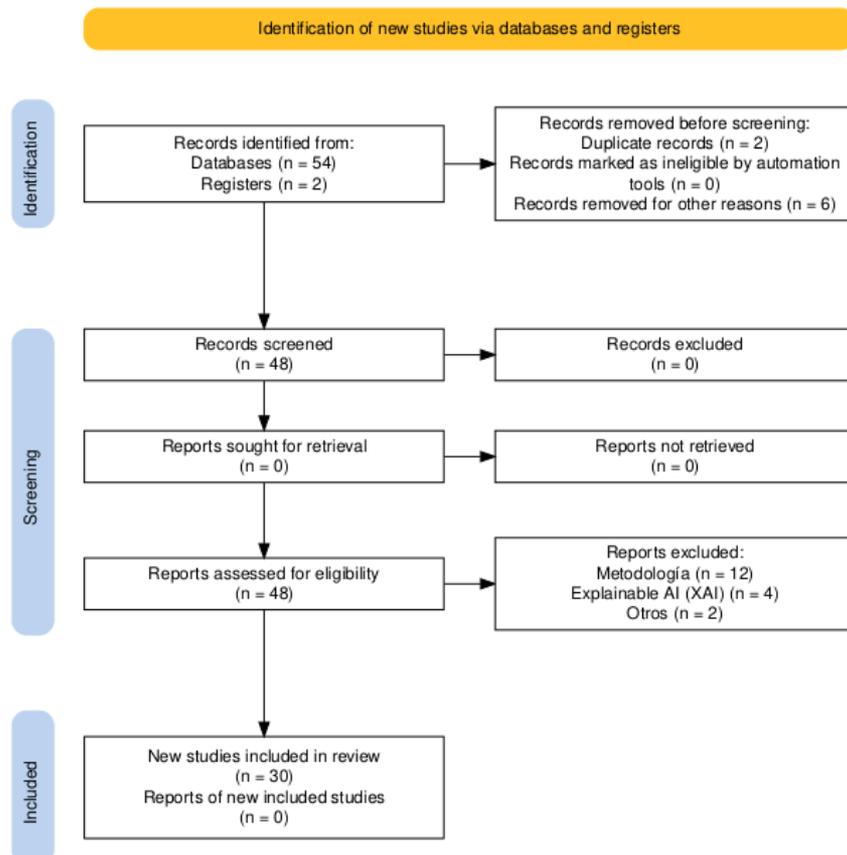
Análisis de Causas Estructurales, Causas Directas y Consecuencias



El proceso de revisión sistemática de literatura se realizó siguiendo el protocolo PRISMA, el diagrama de flujo de este proceso en esta investigación se observa en la Figura 3. Este protocolo permitió establecer criterios claros para la búsqueda, selección y evaluación de fuentes (Alvi et al., 2024; Kim et al., 2020).

Figura 3

Diagrama de Flujo de la Metodología PRISMA para la Revisión de Literatura



Identificación

Se recopilaron un total de 56 registros iniciales, provenientes de bases de datos científicas reconocidas por su rigor académico, tales como IEEE Xplore, SpringerLink, Elsevier y ScienceDirect. En esta fase, se eliminaron 2 documentos duplicados y otros 6 artículos que no cumplían con los siguientes criterios temáticos, contextuales:

- Artículos que traten sobre modelos de aprendizaje automático.
- Artículos cuyo enfoque sea la predicción de riesgo de morosidad o default crediticio.
- Artículos que empleen técnicas de minería de datos.
- Artículos cuyo contexto del estudio sea operativo, no solo teórico o simulado.

Screening (Cribado)

Se procedió al análisis preliminar de los 48 estudios restantes, sin que se registraran exclusiones adicionales en esta etapa, ya que todos los estudios cumplían con los requisitos básicos para pasar a la fase de evaluación completa.

Elegibilidad

Se realizó un análisis en profundidad de los 48 estudios filtrados, de los cuales se excluyeron 18 trabajos por las siguientes razones:

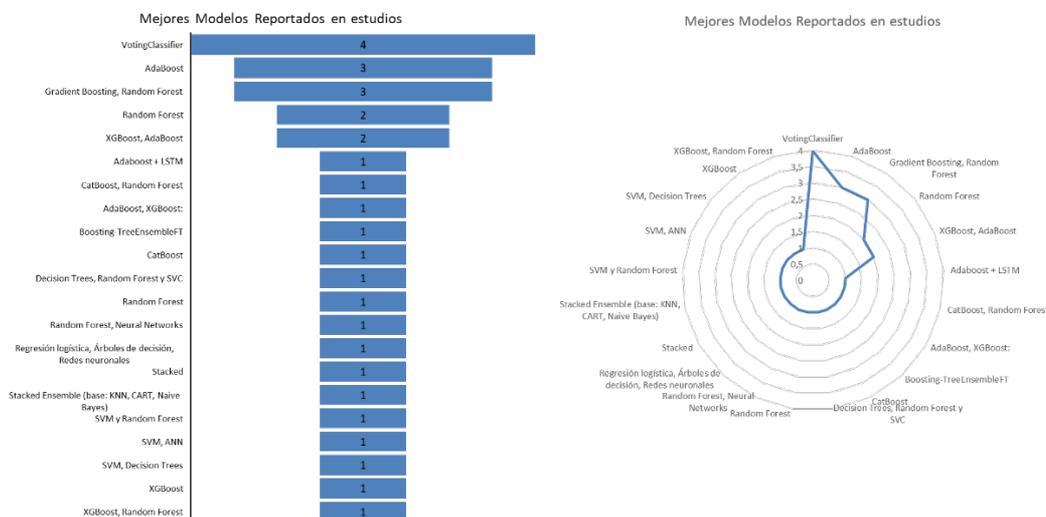
- 12 estudios se centraban exclusivamente en aspectos metodológicos sin aplicación directa a problemas de morosidad o riesgo crediticio.
- 4 estudios estaban orientados al desarrollo de técnicas de interpretabilidad (Explainable AI - XAI) sin abordar directamente la predicción de morosidad.
- 2 estudios fueron descartados por pertenecer a otras temáticas que no se alineaban con los objetivos de la presente investigación.

Inclusión

Finalmente, se obtuvieron 30 estudios en la revisión final, los cuales aportaron evidencia empírica y mejores prácticas sobre la implementación de modelos de aprendizaje automático para la predicción de morosidad y riesgo crediticio en entornos organizacionales, conforme el detalle de la Figura 4.

Figura 4

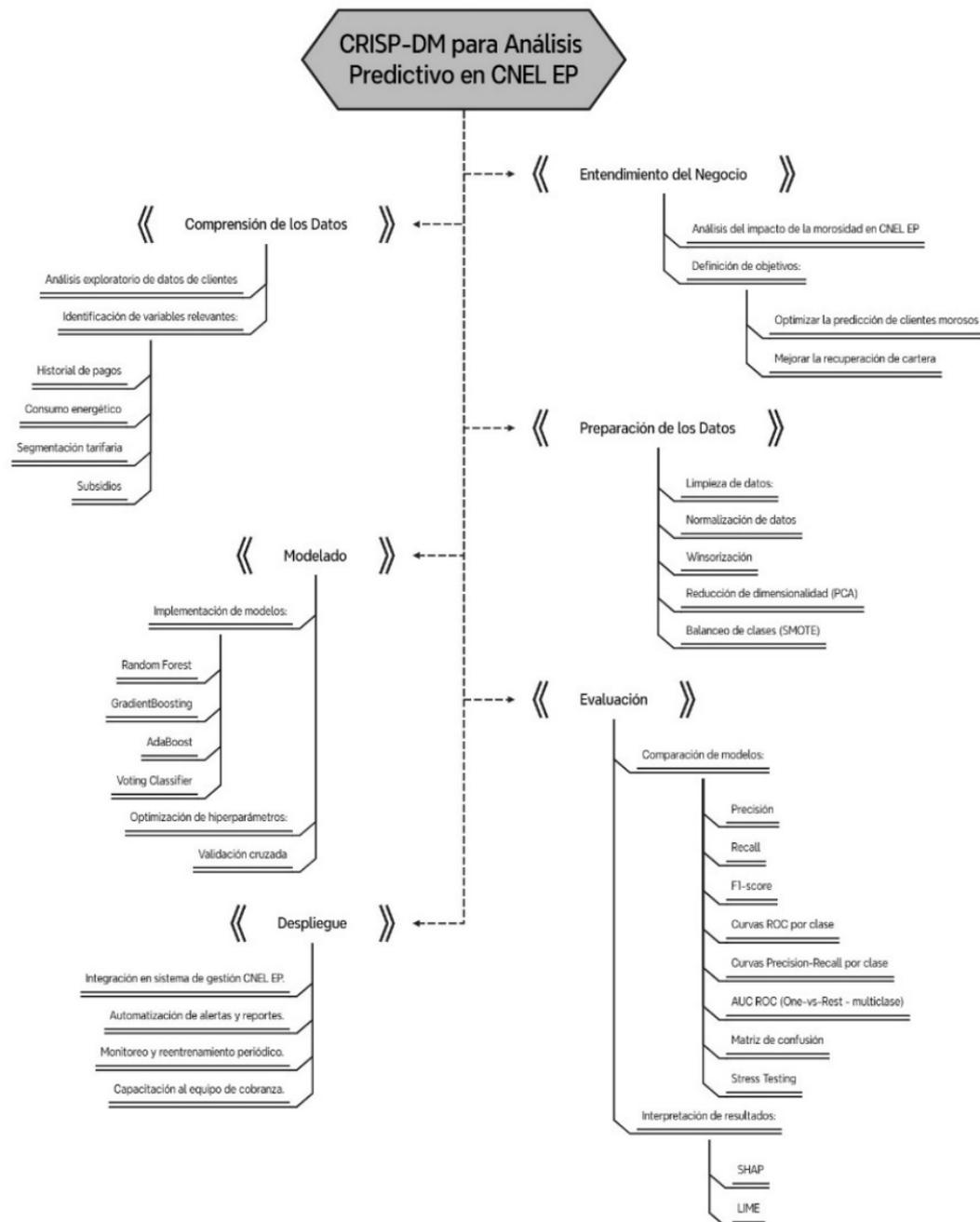
Cantidad de modelos utilizados por estudio en los trabajos analizados



El Ciclo del Diseño permitió la construcción del artefacto, facilitando la validación empírica de las soluciones propuestas. En esta etapa, se describe la construcción de los modelos predictivos propuestos orientados a resolver el problema identificado en el entorno organizacional. Para estructurar este proceso de manera sistemática y efectiva, se adoptó la metodología CRISP-DM (Cross Industry Standard Process for Data Mining), la cual se observa en la Figura 5.

Figura 5

Metodología CRISP-DM para el análisis predictivo de CNEL EP



Entendimiento del Negocio

En una primera fase se analizó el impacto y estado de la cartera vencida en CNEL EP Unidad de Negocio Bolívar en la gestión financiera, identificándose un porcentaje de cartera vencida del 13,6% a diciembre de 2024, porcentaje considerado alto según el Banco Mundial

que recomienda para empresas de servicios públicos mantener este indicador por debajo del 5% para garantizar su sostenibilidad financiera. Así mismo, se realizó un análisis integral de las variables objeto del estudio.

Comprensión de los Datos

La información analizada estuvo compuesta por 72.483 registros y 43 variables relacionadas con la gestión de cartera de CNEL EP Unidad de Negocio Bolívar. Esta incluye datos de identificación del cliente, ubicación geográfica, tarifas, aplicación de subsidios, nivel de morosidad y un desglose detallado de la deuda, como meses de mora, valor total de la deuda, fechas de última factura y suspensión, así como categorizaciones por rangos de antigüedad y valor. También se contemplan componentes específicos de valores vencidos por concepto de energía, alumbrado, intereses, entre otros.

Análisis Exploratorio

Se realizó una depuración de columnas irrelevantes y se clasificaron las variables en numéricas y categóricas. Se identificaron valores atípicos utilizando el rango intercuartílico (IQR), detectándose anomalías especialmente en variables como "MESES_DEUDA", "VALOR_DEUDA", y "V_VENCIDO_ENERGIA".

Se implementaron histogramas, *boxplots*, gráficos de barras para categorías con menos de 20 niveles, mapas de calor de correlación entre variables numéricas, y diagramas de dispersión, como se muestra en las Figuras Figura 6 y Figura 7. Se identificó una alta dispersión y presencia de valores extremos, como deudas que alcanzan hasta los 27.854 USD y períodos de mora de hasta 193 meses. Este análisis permitió identificar patrones de comportamiento de la morosidad y sentó las bases para el preprocesamiento y modelado predictivo posterior (Alonso Robisco & Carbó Martínez, 2022; Kim et al., 2020; Shahid et al., 2023).

Figura 6

Gráficas Descriptivas

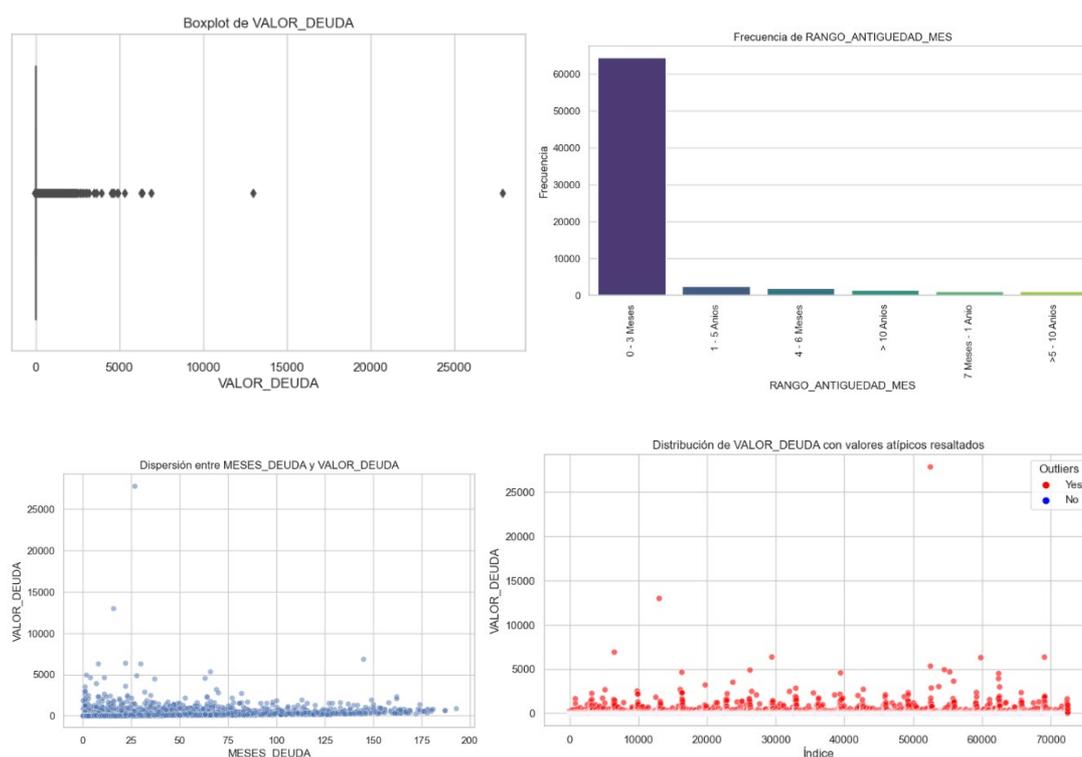
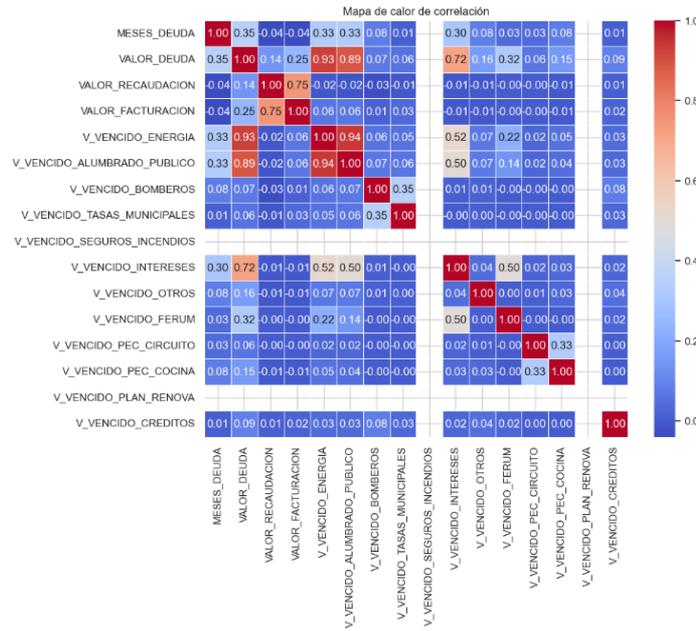


Figura 7

Mapa de calor de Correlación de variables

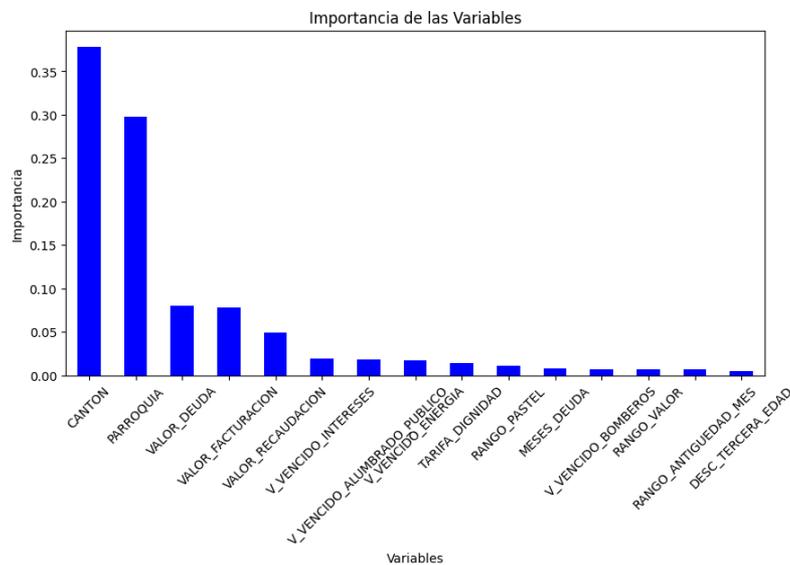


Preparación de Datos

Se realizó una depuración de registros. Se excluyeron clientes con estado “PASIVO” por representar casos en etapa coactiva, no aptos para predicción temprana, así como registros con deudas negativas y duplicados. Posteriormente, se eliminaron variables irrelevantes o redundantes, y se estandarizaron los nombres de las columnas. Se manejaron valores nulos, principalmente imputando la moda o eliminando registros según el caso. Se identificaron y trataron valores atípicos mediante IQR, y se aplicó winsorización para reducir su impacto sin eliminar datos. Finalmente, se implementó un modelo Random Forest con selección de características para reducir la dimensionalidad, mejorando la eficiencia del modelo y su capacidad de generalización sin comprometer la precisión (Akinjole et al., 2024; Alonso Robisco & Carbó Martínez, 2022), el resultado se muestra en la Figura 8.

Figura 8

Importancia de las Variables (Random Forest con selección de características)



Implementación de Modelos

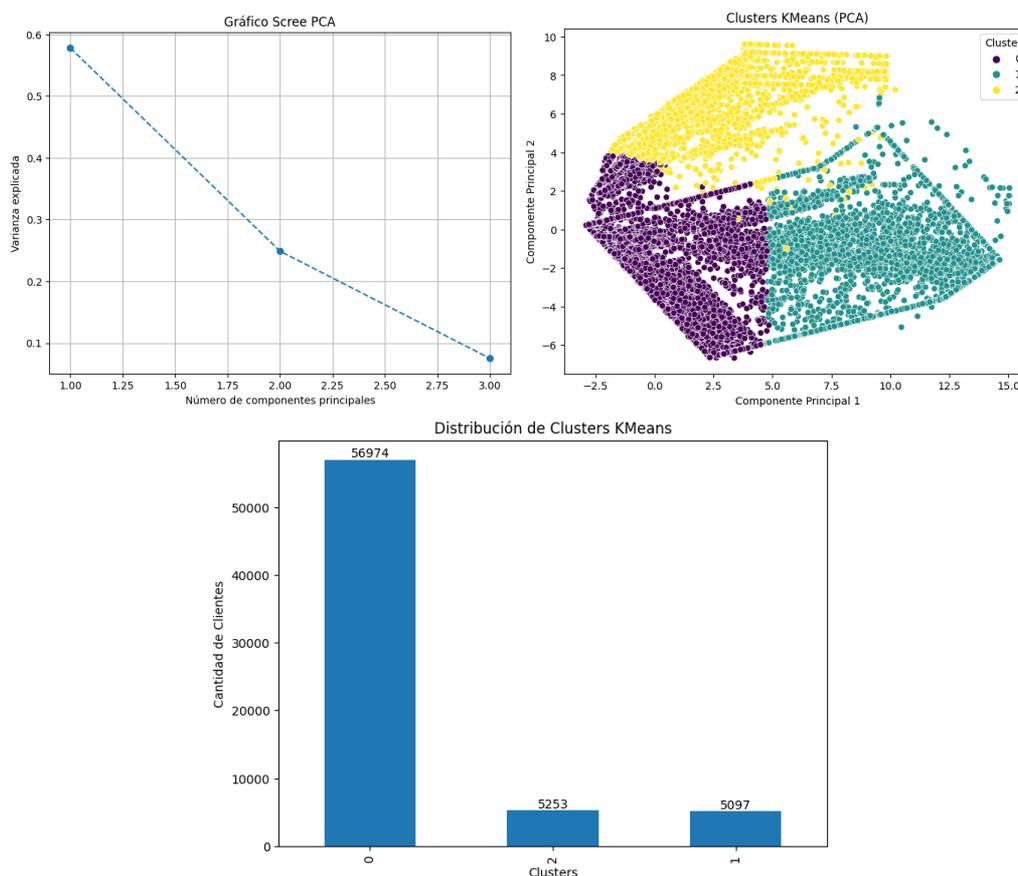
Para identificar el riesgo de morosidad, se desarrolló un enfoque en dos fases: una fase no supervisada, orientada a la segmentación de clientes mediante técnicas de agrupamiento y una fase supervisada, centrada en la construcción de modelos predictivos basados en aprendizaje automático.

Fase No Supervisada

Se seleccionaron las 10 variables más relevantes y se aplicó RobustScaler para escalar los datos, mitigando el efecto de outliers. Posteriormente, se utilizó Análisis de Componentes Principales (PCA) para reducir la dimensionalidad, conservando más del 90% de la varianza. Sobre esta representación reducida se aplicó el algoritmo K-Means, para determinar el número óptimo de clústeres se empleó el método del codo, el cual indicó que tres clústeres ofrecían un buen equilibrio de agrupamiento, representados gráficamente en la Figura 9. El índice de Silhouette del modelo fue de 0.681, clasificando a los clientes en riesgo bajo (0), medio (1) y alto (2). Este enfoque permitió identificar patrones de morosidad sin necesidad de etiquetas predefinidas (Liu et al., 2024; Sheikh et al., 2020; Yemmanuru et al., 2024).

Figura 9

Clasificación con KMeans



Fase Supervisada

A partir de un conjunto de datos previamente segmentado mediante técnicas no supervisadas, se procedió con la codificación de las variables categóricas mediante la técnica de Label Encoding y la selección de variables explicativas relevantes, entre ellas: VALOR_DEUDA, VALOR_FACTURACION, VALOR_RECAUDACION, MESES_DEUDA y los componentes del valor vencido, en función de su contribución al estudio. La variable objetivo fue RIESGO_DE_MOROSIDAD, categorizada en tres niveles

(bajo, medio y alto). Se dividió el conjunto de datos en subconjuntos de entrenamiento y prueba bajo una proporción 80/20. Dado que se identificó un desbalance entre las categorías de la variable objetivo, se aplicó la técnica SMOTE (Synthetic Minority Over-sampling Technique) exclusivamente al conjunto de entrenamiento, con el fin de equilibrar las clases mediante la generación de instancias sintéticas de las categorías minoritarias, mejorando así la representatividad de los datos y la capacidad del modelo para generalizar (Akinjole et al., 2024; Shahid et al., 2023).

Se implementaron tres algoritmos de aprendizaje ensemble: *Random Forest*, *Gradient Boosting* y *AdaBoost*, además de un *Voting Classifier* que integra las predicciones de los modelos base mediante votación suave.

Se realizó un ajuste de hiperparámetros clave en el caso de Random Forest, se aumentó el número de árboles ($n_estimators=150$) para incrementar la estabilidad del modelo, se limitó la profundidad máxima ($max_depth=12$) y se establecieron umbrales mínimos para divisiones y hojas ($min_samples_split=5$, $min_samples_leaf=3$) para evitar el sobreajuste, mientras que la selección aleatoria de características ($max_features="sqrt"$) y el ajuste de pesos por clase ($class_weight="balanced"$) permitieron mejorar la diversidad y compensar el desbalance de clases. Para Gradient Boosting, se incrementó el número de árboles ($n_estimators=200$) con una tasa de aprendizaje reducida ($learning_rate=0.05$) para estabilizar el aprendizaje progresivo, se limitaron los árboles base en profundidad y tamaño de partición, así mismo, se aplicó una técnica de muestreo parcial ($subsample=0.8$) para aumentar la robustez frente al ruido. En el caso de AdaBoost, se configuraron 150 clasificadores débiles con una tasa de aprendizaje moderada ($learning_rate=0.1$) que permite mejorar la estabilidad del modelo sin comprometer su capacidad de ajuste. Estos hiperparámetros, para alcanzar altos niveles de precisión y reducir la tendencia al sobreajuste mejorando la generalización (scikit-learn.org, n.d.-b).

Se desarrollaron curvas de aprendizaje completas con el objetivo de identificar signos de sobreajuste.

Se aplicó validación cruzada estratificada con $k=5$ asegurando que cada fold conserve la proporción original de clases (riesgo bajo, medio y alto), lo que es fundamental en problemas multiclase con datos desbalanceados (scikit-learn.org, n.d.-a).

Las métricas de evaluación empleadas incluyeron Accuracy, Balanced Accuracy, F1 Score Macro, Precision Macro, Recall Macro, Cohen's Kappa y Matthews Correlation Coefficient, lo que proporcionó una valoración robusta del desempeño de los modelos en contextos con clases desbalanceadas. Adicionalmente, se aplicaron métricas basadas en probabilidades, como AUC ROC (One-vs-Rest) y Average Precision Score (PR AUC), lo que permitió analizar su capacidad discriminativa entre clases y su rendimiento frente a clases minoritarias. También se elaboraron matrices de confusión, curvas ROC y curvas Precision-Recall por clase, lo cual enriqueció la interpretación visual de los resultados y fortaleció la validación de los modelos predictivos.

Se construyó un entorno de pruebas sistemático en el que, para cada modelo, se midió su capacidad de generalización y su sensibilidad frente a clases con menor representación. Los escenarios definidos incluyeron: (1) evaluación con el conjunto base original, (2) adición de ruido moderado ($\sigma=0.3$) para simular datos con perturbaciones, (3) reducción del 50% de los datos correspondientes a la clase minoritaria (clase 2), y (4) eliminación aleatoria del 40% del total de muestras. Estos escenarios permitieron analizar la robustez y estabilidad de los modelos

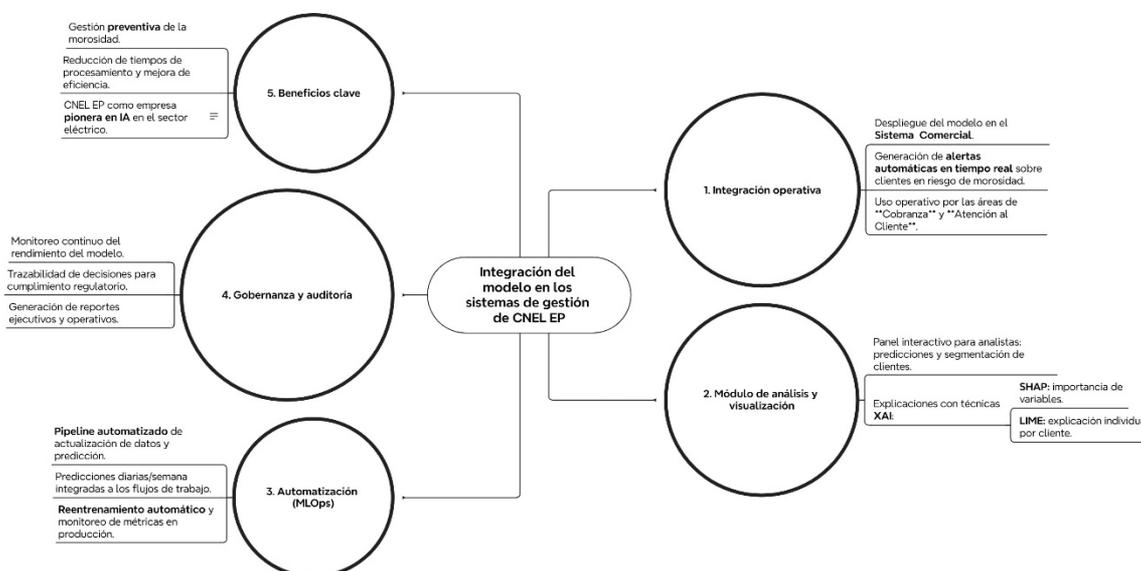
ante condiciones de incertidumbre, pérdida de información o desbalance inducido, consolidando una metodología orientada a validar el comportamiento de los algoritmos en situaciones cercanas a un entorno real de aplicación (Akinjole et al., 2024; Ren, 2025; Shahid et al., 2023).

También se implementaron técnicas de interpretabilidad en inteligencia artificial (*Explainable AI – XAI*), específicamente SHAP (Shapley Additive Explanations) y LIME (Local Interpretable Model-Agnostic Explanations). Estas herramientas permitieron analizar individualmente las decisiones del modelo, identificando el aporte específico de cada variable en la clasificación de riesgo de morosidad.

Finalmente, se realiza una propuesta de integración del modelo en los sistemas de gestión de CNEL EP, tal como se detalla en la Figura 10. La arquitectura propuesta incluye un módulo de alertas en tiempo real y un sistema de visualización interactiva que permite segmentar y priorizar clientes según su nivel de riesgo estimado. El sistema incorpora técnicas de interpretabilidad como SHAP y LIME, lo que garantiza la transparencia de las predicciones y facilita su evaluación por parte de usuarios no expertos; además, se plantea un entorno de actualización continua (MLOps) con monitoreo de desempeño y generación de reportes ejecutivos, alineado con principios de trazabilidad y auditoría en contextos regulados. Estos elementos permiten vincular los resultados obtenidos con aplicaciones prácticas en la gestión de cartera, conforme a lo sugerido en estudios previos (Qin et al., 2021; Dinh & Thanh, 2022; Zanke, 2023).

Figura 10

*Propuesta de integración del modelo en los sistemas de gestión comercial de CNEL EP
Unidad de Negocio Bolívar*



Ciclo de Rigor

La construcción del modelo se fundamentó en principios teóricos ampliamente validados en la literatura científica, los cuales fueron identificados de manera sistemática mediante el protocolo PRISMA. Esta revisión rigurosa permitió explorar y sintetizar críticamente los enfoques más relevantes en segmentación de clientes, técnicas de clasificación en contextos multiclase, y métodos para el tratamiento de desbalance de clases, como SMOTE (Synthetic Minority Over-sampling Technique). A partir de esta base teórica consolidada, se

justificó la selección de los algoritmos Random Forest, Gradient Boosting, AdaBoost y un Voting Classifier, cuya eficacia ha sido ampliamente demostrada en investigaciones aplicadas a problemáticas similares.

Asimismo, se adoptó el modelo metodológico CRISP-DM para estructurar el proceso de análisis en seis fases, desde el entendimiento del negocio hasta la evaluación de los modelos. El uso de esta metodología permitió implementar una limpieza exhaustiva de datos, selección de variables relevantes mediante Random Forest, reducción de dimensionalidad con PCA y una fase exploratoria que identificó valores atípicos críticos como deudas de hasta 27.854 USD o 193 meses de mora acumulada.

La robustez del diseño se reforzó a través de la validación cruzada estratificada ($k=5$), y de un conjunto amplio de métricas de evaluación: Accuracy, F1 Score Macro, Balanced Accuracy, Matthews Correlation Coefficient, AUC ROC OVR y Average Precision Score (PR AUC). Las métricas fueron seleccionadas por su idoneidad en la evaluación de modelos predictivos multiclase con desbalance. Por ejemplo, el modelo Gradient Boosting alcanzó una accuracy del 0.9982 y una F1 Macro de 0.9957, mostrando desempeño casi perfecto incluso ante perturbaciones simuladas.

Se implementaron escenarios de stress testing, para evaluar la estabilidad de los modelos frente a la introducción de ruido ($\sigma=0.3$), eliminación aleatoria del 40% de los registros, y reducción del 50% de la clase minoritaria. El modelo de Gradient Boosting y el Voting Classifier mantuvieron su capacidad de generalización con AUC ROC de 1.000 y PR AUC de 1.000, validando empíricamente su idoneidad para aplicaciones reales. De esta forma, el estudio no solo aplica conocimientos existentes, sino que contribuye al cuerpo científico con evidencia replicable sobre la aplicación de algoritmos ensemble en entornos públicos con alto grado de incertidumbre y desbalance estructural en los datos.

Resultados

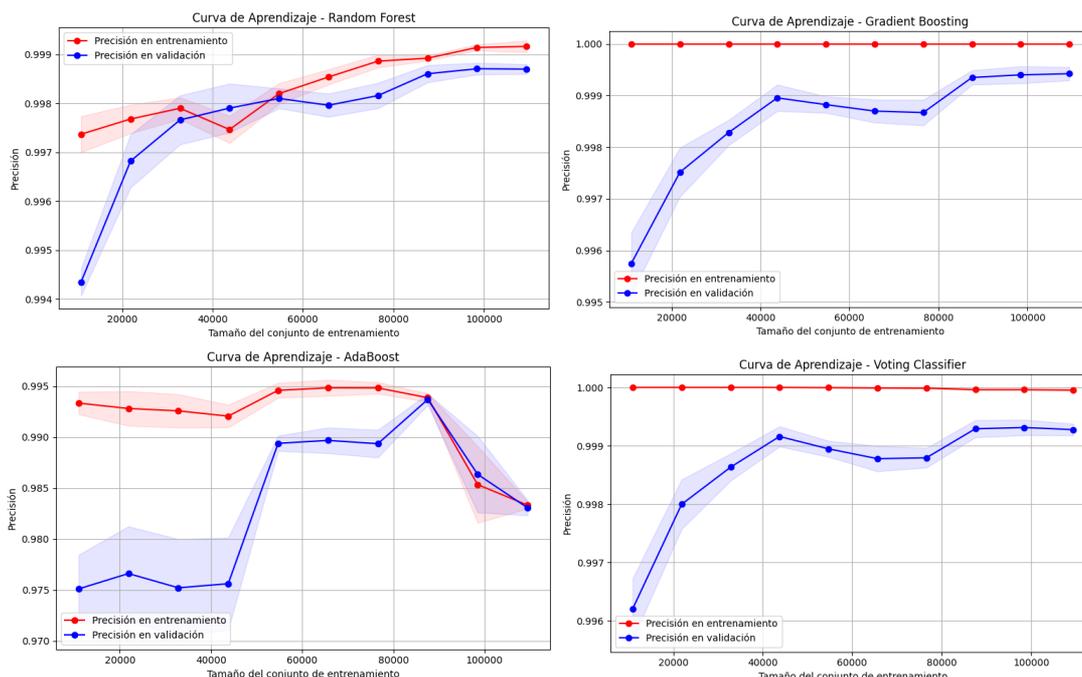
En esta sección, se presentan los resultados derivados de la implementación de los modelos de aprendizaje automático Random Forest, Gradient Boosting y AdaBoost, además de un Voting Classifier; se implementan curvas de aprendizaje para identificar tendencias de convergencia y posibles indicios de sobreajuste; se comparan las métricas de desempeño utilizando validación cruzada estratificada y escenarios simulados representativos de condiciones operativas reales; se evalúa la robustez de los modelos a través de análisis de sensibilidad y pruebas de estrés, complementadas con técnicas de inteligencia artificial explicable (SHAP y LIME) para interpretar la influencia de las variables en las predicciones y finalmente se propone un modelo matemático que aborde el problema de clasificación multiclase en la predicción del riesgo de morosidad.

Curvas de Aprendizaje

Los modelos Gradient Boosting y Voting Classifier presentaron un comportamiento altamente estable, con precisiones cercanas al 100%, lo que evidenció una sólida capacidad de generalización. Random Forest mantuvo también un desempeño consistente, con una brecha controlada entre ambas curvas. En contraste, el modelo AdaBoost evidenció una mayor variabilidad en la precisión de validación y una separación más pronunciada respecto al entrenamiento, lo que indicó una menor robustez frente a incrementos en el tamaño del conjunto de datos y una posible tendencia al sobreajuste, tal como se observa en la Figura 11.

Figura 11

Clasificación con KMeans



Validación Cruzada

Los modelos fueron evaluados mediante validación cruzada estratificada (k=5). Todos los modelos, excepto AdaBoost, presentan un rendimiento sobresaliente con una concordancia cercana entre el accuracy puntual y el promedio de validación cruzada, lo que sugiere una alta capacidad de generalización y robustez, los resultados se presentan en la Tabla 1. GradientBoosting y VotingClassifier emergen como las opciones más sólidas para entornos críticos, mientras que AdaBoost, aunque funcional, podría no ser óptimo en contextos con alta variabilidad o donde las clases minoritarias son relevantes.

Tabla 1

Resultados Validación Cruzada (k=5)

Modelo	Accuracy	Cross-Validation	Análisis
RandomForest	0.997	0.999	Excelente capacidad de ajuste, con mínima diferencia entre entrenamiento y validación, lo que indica buena generalización y bajo riesgo de sobreajuste.
GradientBoosting	0.998	0.999	Ligeramente superior en precisión puntual. Su alto desempeño en validación cruzada refuerza su confiabilidad, incluso en entornos variados.
AdaBoost	0.984	0.983	Notablemente menor que los otros modelos. Puede tener mayor sensibilidad al ruido o menor capacidad de adaptación a patrones complejos, pese al balanceo con SMOTE.
VotingClassifier	0.998	0.999	Iguala el rendimiento del GradientBoosting y lo supera en robustez al combinar varios algoritmos, lo que le permite mantener consistencia entre precisión individual y validación cruzada.

Evaluación de modelos

Los resultados mostrados en la Tabla 2 indican que el modelo GradientBoosting obtuvo el mejor desempeño global con métricas casi perfectas en todas las clases. VotingClassifier también mostró resultados sobresalientes, siendo robusto y estable, ideal para escenarios operativos. RandomForest alcanzó una precisión muy alta y balance entre clases. En contraste, AdaBoost, aunque competitivo, presentó menor rendimiento en clases minoritarias. La

evaluación incluyó métricas tradicionales y probabilísticas, así como análisis gráfico (ROC y Precision-Recall), confirmando la efectividad de SMOTE en mejorar el aprendizaje de clases minoritarias.

Tabla 2

Evaluación de Modelos

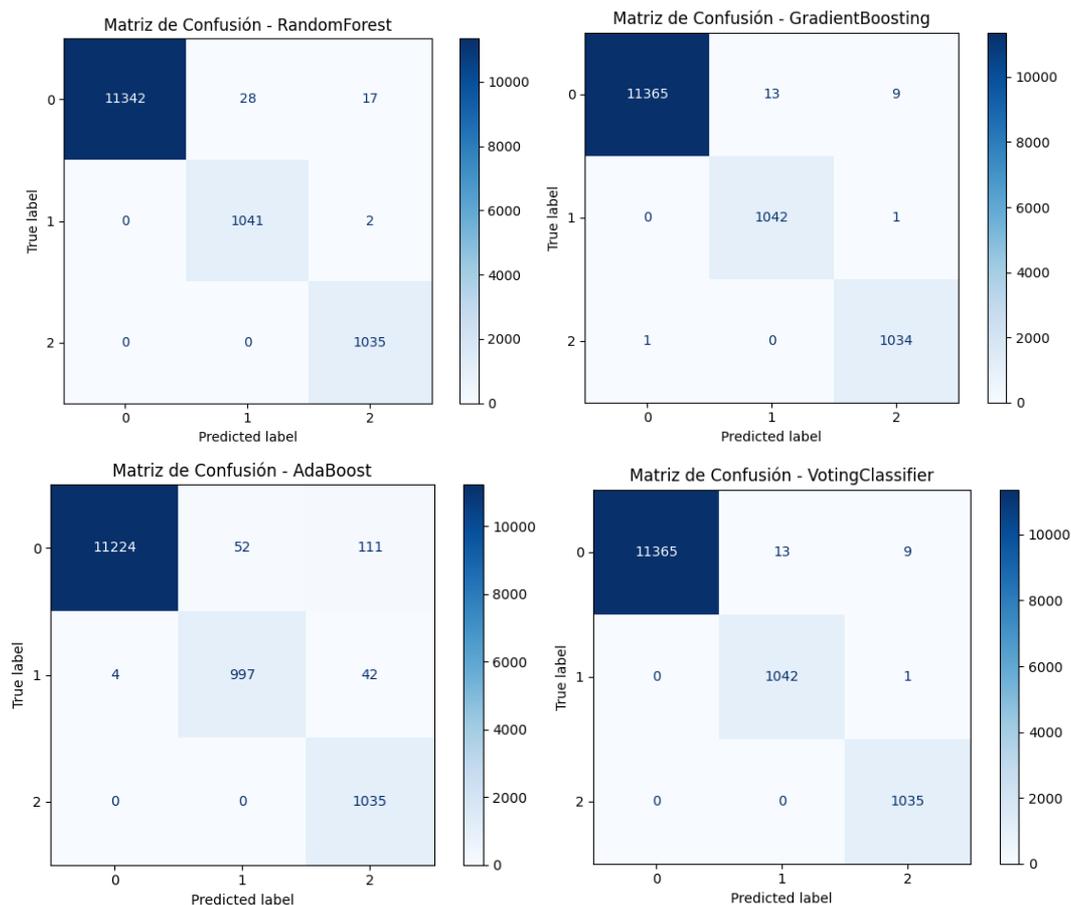
Modelo	Accuracy	Balanced Accuracy	F1 Macro	Precision Macro	Recall Macro	Cohen's Kappa	Matthews Corrcoeff	AUC ROC (OVR)	Avg. Precision (PR AUC)
RandomForest	0.9965	0.9980	0.9916	0.9853	0.9980	0.9873	0.9874	1.0	0.9998
GradientBoosting	0.9982	0.9987	0.9957	0.9927	0.9987	0.9935	0.9935	1.0	1.0
AdaBoost	0.9845	0.9805	0.9590	0.9404	0.9805	0.9450	0.9459	0.9997	0.9981
VotingClassifier	0.9983	0.9990	0.9959	0.9927	0.9990	0.9938	0.9938	1.0	0.9999

Matrices de Confusión

Las matrices de confusión muestran que GradientBoosting y VotingClassifier son los modelos con mejor desempeño general, presentando apenas 23 errores cada uno y una excelente capacidad para discriminar entre las tres clases de riesgo de morosidad (bajo, medio y alto). RandomForest también demuestra alta precisión, aunque con un leve incremento en errores de clasificación, especialmente en la clase 0. En contraste, AdaBoost es el modelo con menor desempeño relativo, acumulando más errores en las clases 0 y 1, lo que evidencia una menor sensibilidad frente a clases minoritarias, a pesar de clasificar correctamente todos los casos de la clase 2, tal como se observa en la Figura 12.

Figura 12

Matrices de Confusión modelos implementados

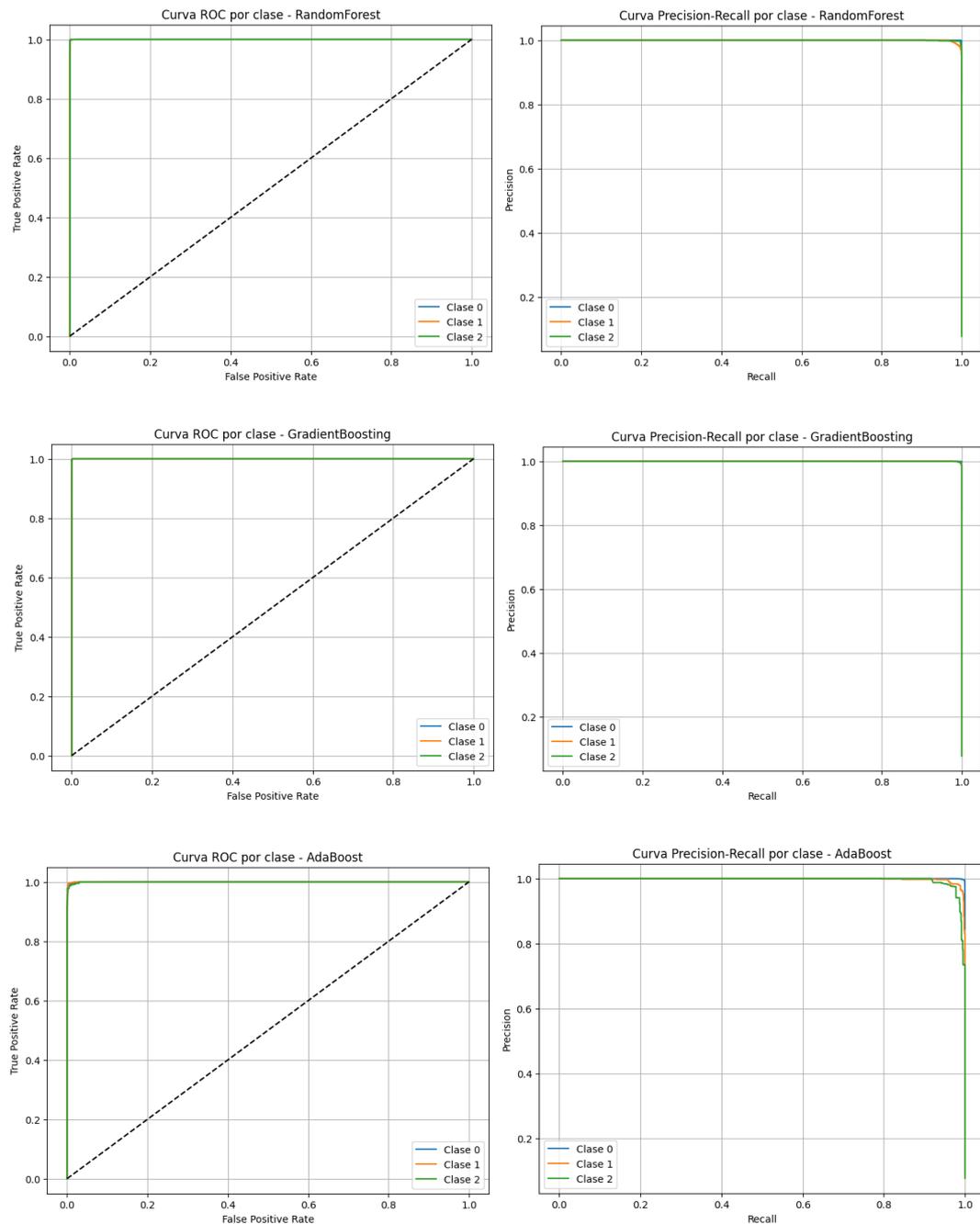


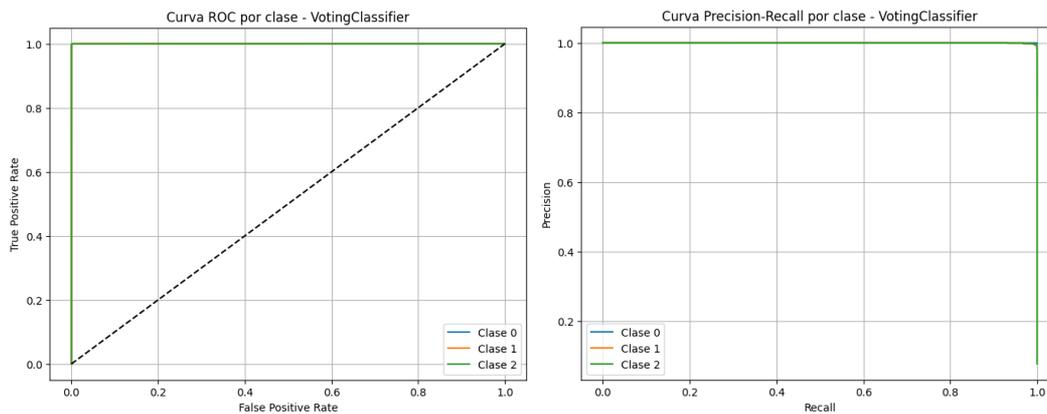
Curvas Precision-Recall y ROC

Las curvas ROC y Precision-Recall por clase reflejaron un desempeño sobresaliente en todos los modelos evaluados. GradientBoosting y VotingClassifier, cuyos perfiles se mantuvieron próximos al punto ideal, evidenciando una excelente capacidad de discriminación y precisión en la clasificación de las tres clases. RandomForest mostró un comportamiento igualmente sólido, aunque con leves desviaciones en la clase 1. Por su parte, AdaBoost presentó una ligera disminución en la precisión y sensibilidad, particularmente en las clases minoritarias, lo que se tradujo en curvas menos ajustadas al óptimo, como se ilustra en la Figura 13.

Figura 13

Curva ROC y Curva de Precisión-Recall de los modelos implementados





Stress Testing

Los resultados del stress testing demostraron que todos los modelos evaluados mantuvieron un alto rendimiento en el escenario base, destacándose GradientBoosting y VotingClassifier con métricas cercanas a la perfección en precisión, sensibilidad, F1, AUC y PR AUC. Ante escenarios adversos como ruido, reducción de clases y eliminación de datos, GradientBoosting mostró la mayor robustez y consistencia, con caídas mínimas en las métricas clave, mientras que VotingClassifier también mantuvo un desempeño muy alto, aunque con una leve sensibilidad al balance de clases. En contraste, AdaBoost evidenció mayor vulnerabilidad, especialmente en escenarios de reducción de datos o desbalance. En función de estos resultados, GradientBoosting se perfila como el modelo más adecuado para entornos operativos exigentes, donde la precisión y la resiliencia frente a condiciones cambiantes son fundamentales, como se detalla en la Tabla 3.

Tabla 3

Stress Testing modelos implementados

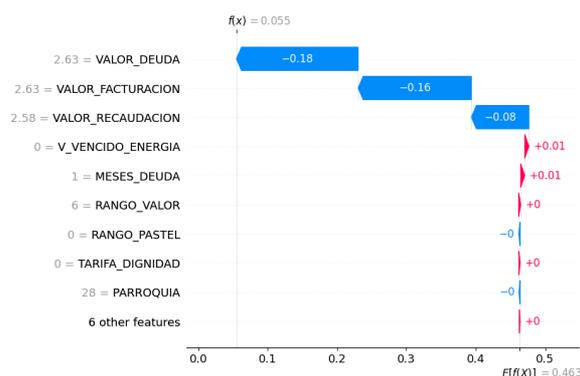
Modelo	Escenario	Accuracy	Balanced Accuracy	F1 Macro	Precision Macro	Recall Macro	AUC ROC (OVR)	Avg. Precision (PR AUC)	Observaciones
RandomForest	Base (original)	0,9965	0,998	0,9916	0,9853	0,998	1	0,9998	Desempeño sólido y estable.
GradientBoosting	Base (original)	0,9982	0,9987	0,9957	0,9927	0,9987	1	1	Máxima precisión y robustez.
AdaBoost	Base (original)	0,9845	0,9805	0,959	0,9404	0,9805	0,9997	0,9981	Alta precisión pero menor balance.
VotingClassifier	Base (original)	0,9983	0,999	0,9959	0,9927	0,999	1	0,9999	Desempeño sólido y estable.
RandomForest	Ruido moderado ($\sigma=0.3$)	0,9926	0,9939	0,9823	0,9711	0,9939	0,9997	0,9982	Ligeramente afectado por ruido.
GradientBoosting	Ruido moderado ($\sigma=0.3$)	0,9952	0,9903	0,9865	0,983	0,9903	0,9999	0,9994	Buena tolerancia al ruido.
AdaBoost	Ruido moderado ($\sigma=0.3$)	0,9811	0,9783	0,9517	0,9302	0,9783	0,9995	0,9962	Disminución significativa por ruido.
VotingClassifier	Ruido moderado ($\sigma=0.3$)	0,9954	0,9941	0,9885	0,9831	0,9941	0,9999	0,9993	Resistente a ruido con leve pérdida.
RandomForest	Reducción clase 2 al 50%	0,9964	0,998	0,9886	0,9795	0,998	1	0,9997	Pérdida de sensibilidad clase 2.
GradientBoosting	Reducción clase 2 al 50%	0,9982	0,999	0,9943	0,9896	0,999	1	1	Reducción marginal con buena generalización.
AdaBoost	Reducción clase 2 al 50%	0,9839	0,9805	0,939	0,9074	0,9805	0,9997	0,9973	Menor precisión en clase minoritaria.
VotingClassifier	Reducción clase 2 al 50%	0,9982	0,999	0,9943	0,9896	0,999	1	0,9999	Mantuvo desempeño elevado.
RandomForest	Eliminación aleatoria del 40%	0,9965	0,9982	0,9917	0,9854	0,9982	1	0,9998	Afectado por reducción de muestra.
GradientBoosting	Eliminación aleatoria del 40%	0,9984	0,9989	0,9962	0,9936	0,9989	1	1	Estable bajo eliminación aleatoria.
AdaBoost	Eliminación aleatoria del 40%	0,9853	0,9821	0,961	0,9427	0,9821	0,9998	0,9983	Alto impacto por pérdida de datos.
VotingClassifier	Eliminación aleatoria del 40%	0,9984	0,9994	0,9962	0,9931	0,9994	1	0,9999	Alto rendimiento con buena resiliencia.

Aplicación de técnicas de interpretabilidad de modelos (Explainable AI - XAI)

El análisis SHAP aplicado al modelo VotingClassifier mostró que la predicción individual se ajustó desde un valor esperado de 0.463 a una probabilidad final de 0.055 de morosidad. Las variables que más contribuyeron a esta reducción fueron VALOR_DEUDA = 2.63 (-0.18), VALOR_FACTURACION = 2.63 (-0.16) y VALOR_RECAUDACION = 2.58 (-0.08), todas asociadas a bajos montos que disminuyeron significativamente el riesgo estimado. Otras variables como MESES_DEUDA = 1, V_VENCIDO_ENERGIA = 0 y RANGO_VALOR = 6 tuvieron un impacto marginal positivo (+0.01 o menor), sin alterar el resultado final. En conjunto, el modelo clasificó correctamente al cliente como No Moroso, basándose en características económicas clave, como se visualiza en la Figura 14.

Figura 14

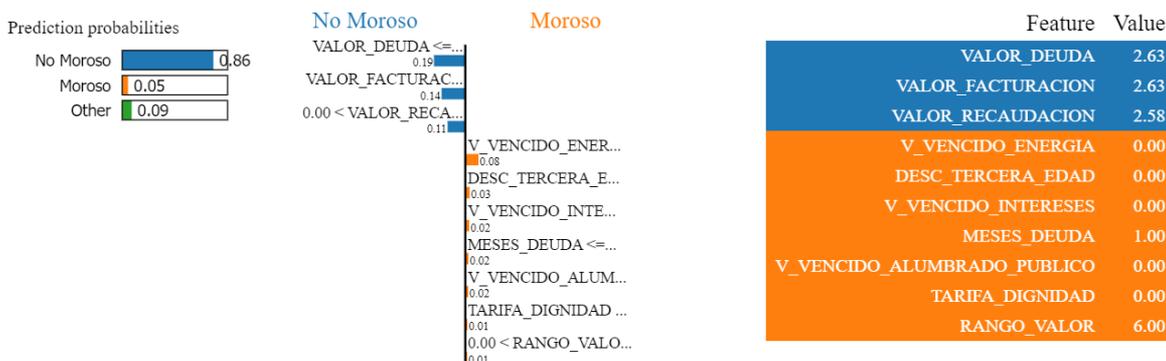
SHAP aplicado al modelo VotingClassifier



Así mismo, LIME de VotingClassifier mostró que el cliente fue clasificado como No Moroso con una probabilidad del 86%, frente a un 5% de ser considerado moroso. Las variables que más contribuyeron a esta decisión fueron VALOR_DEUDA = 2.63 (+0.19), VALOR_FACTURACION = 2.63 (+0.14) y VALOR_RECAUDACION = 2.58 (+0.11), todas asociadas a montos bajos que redujeron el riesgo proyectado. En contraste, factores como V_VENCIDO_ENERGIA = 0 (+0.08), DESC_TERCERA_EDAD = 0 (+0.03) y MESES_DEUDA = 1 (+0.02) tuvieron un efecto leve hacia la morosidad, pero sin suficiente peso para alterar la clasificación. Esta interpretación confirma la consistencia del modelo al otorgar mayor peso a variables económicas directas en su decisión final, como se muestra en la Figura 15.

Figura 15

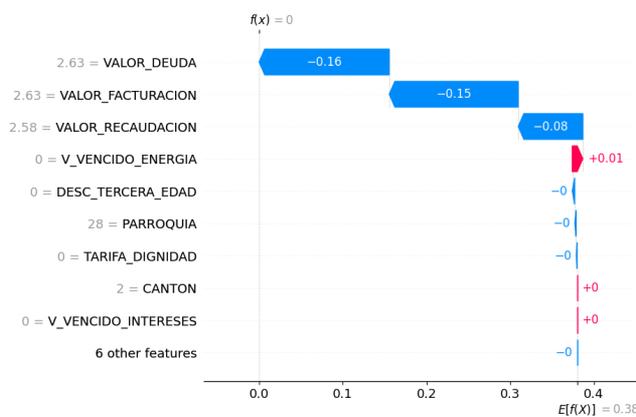
LIME aplicado al modelo VotingClassifier



SHAP aplicado al modelo Gradient Boosting indicó que la predicción final fue 0, a partir de un valor esperado del modelo de 0.38, lo que refleja una fuerte inclinación hacia la clasificación de No Moroso. Las variables que más contribuyeron a esta disminución fueron VALOR_DEUDA = 2.63 (-0.16), VALOR_FACTURACION = 2.63 (-0.15) y VALOR_RECAUDACION = 2.58 (-0.08), evidenciando que valores bajos en estos indicadores financieros son determinantes para reducir el riesgo. En contraste, V_VENCIDO_ENERGIA = 0 aportó ligeramente a la morosidad (+0.01), pero sin impacto relevante. Otras variables como DESC_TERCERA_EDAD, PARROQUIA, CANTON, TARIFA_DIGNIDAD y V_VENCIDO_INTERESES no influyeron significativamente. Esta interpretación muestra que el modelo priorizó información cuantitativa sobre deuda y facturación como factores clave para una decisión confiable, como se representa en la Figura 16.

Figura 16

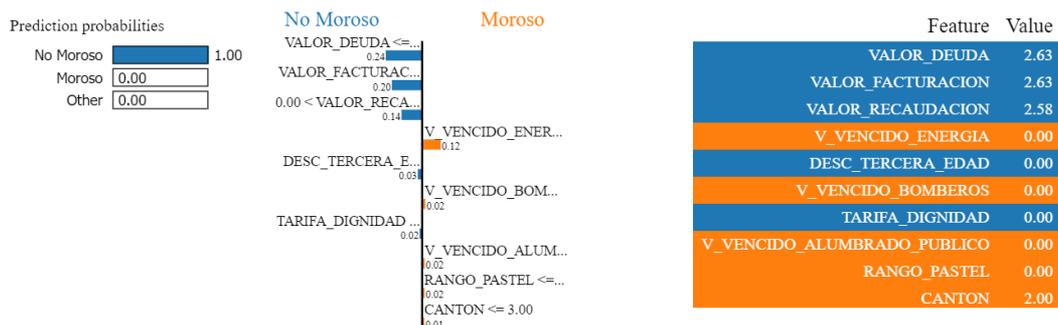
SHAP aplicado al modelo Gradient Boosting



Mientras que LIME mostró que el cliente fue clasificado como No Moroso con una probabilidad del 100%. Las variables que más contribuyeron a esta decisión fueron VALOR_DEUDA = 2.63 (+0.24), VALOR_FACTURACION = 2.63 (+0.20) y VALOR_RECAUDACION = 2.58 (+0.14), indicando que estos montos bajos o adecuados tuvieron un peso determinante en la reducción del riesgo proyectado. En contraste, V_VENCIDO_ENERGIA = 0.00 (+0.12) fue la única variable con una contribución moderada hacia la morosidad, mientras que otras como DESC_TERCERA_EDAD = 0, V_VENCIDO_BOMBEROS = 0, CANTON = 2 y RANGO_PASTEL = 0 mostraron impactos muy bajos (+0.01 a +0.02). Esta explicación confirma que el modelo priorizó los indicadores económicos directos para tomar la decisión, descartando variables estructurales por su escasa influencia en esta observación, como se muestra en la Figura 17.

Figura 17

LIME aplicado al modelo Gradient Boosting



Modelo matemático

Finalmente, se propone un modelo matemático formulado para abordar el problema de clasificación multiclase en la predicción del riesgo de morosidad. Este modelo busca proporcionar una base teórica rigurosa que respalde la validación técnica del enfoque implementado y, al mismo tiempo, facilite su replicabilidad en otros entornos operativos con características comparables (Akinjole et al., 2024; Gao & Balyan, 2022; Qin et al., 2021).

- Clase 0 (Riesgo Bajo): Clientes con baja probabilidad de incumplimiento.
- Clase 1 (Riesgo Medio): Clientes con una probabilidad intermedia de incumplimiento.
- Clase 2 (Riesgo Alto): Clientes con alta probabilidad de incumplimiento.

Se define $X_i \in R^n$ como el vector de características asociadas al cliente i , donde n es el número de variables predictoras relevantes (monto de deuda, meses de mora, consumo de energía, tipología tarifaria, entre otras).

Se busca encontrar una función de clasificación multiclase:

$$f: R^n \rightarrow \{0,1,2\}$$

Donde $y_i = f(x_i)$ representa la clase de riesgo asignada al cliente.

Estructura del sistema ensemble

El sistema ensemble implementado agrupa m clasificadores individuales f_j , cada uno entrenado de manera independiente bajo metodologías basadas en árboles y boosting. Entre estos se encuentran: Random Forest, Gradient Boosting, AdaBoost.

Cada clasificador f_j produce una probabilidad vectorial sobre las tres clases posibles:

$$p_j(x_i) = [p_j^{(0)}(x_i), p_j^{(1)}(x_i), p_j^{(2)}(x_i)]$$

Donde:

- $p_j^{(k)}(x_i) \in [0,1]$ es la probabilidad de que el cliente i pertenezca a la clase, $k \in \{0,1,2\}$ según el clasificador j ,
- $\sum_{k=0}^2 p_j^k(x_i) = 1$ para cada j e i .

Votación ponderada (Soft Voting Multiclase)

Ensemble integra las predicciones individuales mediante una votación ponderada, asignando un peso w_j a cada clasificador base, donde $w_j > 0$ y $\sum_{j=1}^m w_j = 1$.

Para cada clase k la probabilidad agregada es:

$$p^k(x_i) = \sum_{j=1}^m w_j \cdot p_j^k(x_i), \quad k = 0,1,2$$

La asignación de la clase al cliente i se realiza seleccionando la clase con la probabilidad conjunta máxima:

$$f(x_i) = \operatorname{arg\,max}_{k \in \{0,1,2\}} p^k(x_i)$$

La optimización del ensemble multiclase se realiza minimizando la pérdida logarítmica categórica (categorical cross-entropy):

$$L = -\frac{1}{N} \sum_{i=1}^N \sum_{k=0}^2 \mathbb{I}\{y_i = k\} \cdot \log P^k(x_i)$$

Donde:

- $\mathbb{I}\{y_i = k\}$ es una función indicadora que vale 1 si el cliente i pertenece a la clase k y 0 en caso contrario.
- $P^k(x_i)$ es la probabilidad estimada de la clase k para el cliente i ,
- N es el total de observaciones en el conjunto de entrenamiento.

Gradient Boosting Multiclase

Gradient Boosting construye una función de decisión compuesta como suma de árboles de decisión:

$$F^k(x_i) = \sum_{t=1}^T r_t^k h_t^k(x_i) \quad \text{Para } k \in \{0,1,2\}$$

Donde:

- $h_t^k(\cdot)$ es el árbol base número t especializado de la clase k ,
- r_t^k es el peso del árbol h_t^k
- T es el número total de iteraciones (boosting rounds).

La función de predicción utiliza la normalización softmax para convertir los scores $F^k(x_i)$ en probabilidades:

$$P^k(x_i) = \frac{e^{F^k(x_i)}}{\sum_{j=0}^2 e^{F^j(x_i)}}, \quad \text{Para } k = 0,1,2$$

La clase predicha es la que tiene la mayor probabilidad estimada:

$$f(x_i) = \operatorname{arg\,max}_{k \in \{0,1,2\}} p^k(x_i)$$

Función de Pérdida

El entrenamiento de Gradient Boosting optimiza la función de pérdida categórica (log-loss) sobre N ejemplos del conjunto de entrenamiento:

$$L = -\frac{1}{N} \sum_{i=1}^N \sum_{k=0}^2 \mathbb{I}\{y_i = k\} \cdot \log P^k(x_i)$$

- $\mathbb{I} \{y_i = k\}$ es una función indicadora que vale 1 si el cliente i pertenece a la clase k y 0 en caso contrario.
- $P^k(x_i)$ es la probabilidad estimada de la clase k para el cliente i .

Discusión

El presente estudio permitió identificar el modelo de aprendizaje automático con mayor capacidad predictiva para anticipar el riesgo de morosidad en clientes de CNEL EP Unidad de Negocio Bolívar, a partir de un enfoque metodológico riguroso sustentado en Design Science Research (DSR). Esta perspectiva metodológica facilitó la integración de evidencia empírica con la construcción de un artefacto tecnológico alineado a las necesidades operativas institucionales, identificadas en el Ciclo de Relevancia. La aplicación de DSR resultó adecuada para el contexto de la ingeniería aplicada, al permitir la solución de problemas prácticos mediante la construcción y validación iterativa de artefactos, como señalan Goecks et al., 2021; Gregor, 2021. Asimismo, la relevancia y madurez de este enfoque en entornos organizacionales complejos ha sido ratificada por estudios recientes que reconocen su contribución tanto al desarrollo de conocimiento científico como a la transformación digital de procesos en sectores intensivos en datos (Akoka et al., 2023).

Se evidencia un desempeño sobresaliente del modelo Gradient Boosting, el cual alcanzó métricas cercanas a la perfección (Accuracy: 0.9982; F1 Macro: 0.9957; AUC ROC: 1.000). El modelo VotingClassifier presentó un rendimiento igualmente elevado (Accuracy: 0.9983; F1 Macro: 0.9959; AUC ROC: 1.000). Al comparar estos resultados con hallazgos previos en la literatura, se constata una mejora significativa. (Xu, 2024) reportó para Gradient Boosting un Accuracy del 92% y un AUC de 0.85, mientras que (Lai, 2020) informó un Accuracy del 95.2% y un AUC de 0.94 para XGBoost. Asimismo, (Akinjole et al., 2024) alcanzaron un Accuracy del 93.7% y un AUC de 96.5% utilizando un enfoque ensemble basado en VotingClassifier. Estos antecedentes respaldan la superioridad de los modelos desarrollados en este estudio, particularmente en contextos caracterizados por desbalance de clases y elevada exigencia en capacidad discriminativa.

La aplicación de técnicas de preprocesamiento como SMOTE y PCA resultó fundamental para la mejora del desempeño predictivo. En particular, SMOTE contribuyó significativamente al equilibrio de clases, incrementando la capacidad del modelo para identificar correctamente observaciones pertenecientes a clases minoritarias, como ha sido evidenciado también por (Akinjole et al., 2024), quienes reportan mejoras sustanciales al combinar métodos de sobremuestreo con algoritmos ensemble. Por su parte, el uso de PCA permitió reducir la dimensionalidad del conjunto de datos manteniendo el 90.21% de la varianza explicada, lo cual optimizó la eficiencia computacional sin afectar la precisión, coincidiendo con lo planteado por (Liu et al., 2024), quienes subrayan los beneficios que la transformación basada en árboles aporta al desempeño de los modelos logísticos. Asimismo, la calidad de la segmentación previa al modelado supervisado fue validada mediante un índice de Silhouette de 0.681, reforzando la solidez de las agrupaciones generadas y la pertinencia del enfoque adoptado.

Una limitación relevante del estudio es la falta de validación externa en contextos distintos al conjunto de entrenamiento. Aunque se aplicaron validación cruzada y pruebas de estrés, se requiere contrastar los resultados en otros entornos operativos para confirmar la generalización del modelo, como recomiendan (Alonso Robisco & Carbó Martínez, 2022) en aplicaciones financieras reguladas.

Desde una perspectiva aplicada, los modelos desarrollados presentan un alto potencial para ser integrados como herramientas de apoyo a la toma de decisiones en la gestión de cartera, permitiendo la priorización de intervenciones en función del nivel de riesgo estimado. Para investigaciones futuras, se sugiere evaluar su desempeño en entornos productivos reales, a fin de validar su eficacia operativa. Asimismo, se recomienda la incorporación de técnicas de interpretabilidad, como SHAP o LIME, que faciliten la comprensión de las predicciones individuales, lo cual es particularmente relevante en contextos regulados. Finalmente, se plantea como línea de trabajo futura la exploración de modelos híbridos basados en arquitecturas neuronales o enfoques secuenciales que capturen la dinámica temporal del comportamiento del cliente, siguiendo propuestas recientes como las de (Zhang & Wang, 2023).

En síntesis, la investigación demuestra la viabilidad de adaptar metodologías avanzadas de ciencia de datos al sector eléctrico público, logrando soluciones predictivas precisas, replicables y contextualizadas operativamente. Estos resultados respaldan las metas de sostenibilidad tanto operativa como financiera de CNEL EP, y se relacionan con estudios previos que destacan la efectividad de los modelos ensemble en dominios de riesgo crediticio (Akinjole et al., 2024; Dastile et al., 2020), así como con la aplicabilidad del enfoque de Design Science Research para generar artefactos útiles en contextos reales (Delpont et al., 2024).

Conclusiones

El presente estudio tuvo como objetivo identificar el modelo de aprendizaje automático más preciso para predecir el riesgo de morosidad en los clientes de CNEL EP, Unidad de Negocio Bolívar enmarcándose en la necesidad institucional de mejorar la sostenibilidad financiera a través de la optimización de la gestión de cartera.

Se aplicaron diversas técnicas de preprocesamiento y modelos de clasificación multiclase, destacando Gradient Boosting y VotingClassifier como los más eficaces. Ambos modelos alcanzaron métricas de desempeño sobresalientes (Accuracy: 0.9982–0.9983; F1 Macro: 0.9957–0.9959; AUC ROC: 1.000) incluso bajo condiciones simuladas de ruido, desbalance y pérdida de datos.

Los resultados refuerzan la validez de aplicar técnicas de aprendizaje automático en contextos distintos al financiero, como el sector eléctrico público. Además, contribuyen al cuerpo de conocimiento al evidenciar que, mediante PCA, SMOTE y modelos ensemble, es posible desarrollar soluciones predictivas robustas y replicables que superan el rendimiento reportado en estudios previos.

Una limitación clave del estudio es la ausencia de validación externa en otros entornos geográficos o temporales. Aunque se realizaron pruebas de estrés y validación cruzada, se recomienda aplicar el modelo en otros contextos reales para evaluar su generalización.

Los modelos propuestos pueden ser integrados como herramientas de apoyo en la gestión comercial de CNEL EP, permitiendo priorizar intervenciones según el riesgo de morosidad. Para investigaciones futuras, se sugiere validar el modelo en tiempo real, incorporar variables temporales y explorar arquitecturas híbridas con redes neuronales que capturen la dinámica del comportamiento del cliente.

Este estudio demuestra la factibilidad de adaptar técnicas avanzadas de ciencia de datos al sector eléctrico público, generando artefactos predictivos precisos y útiles. Su contribución se manifiesta no solo en el fortalecimiento de la gestión operativa de CNEL EP, sino también

en la ampliación del alcance de la inteligencia artificial explicable a contextos no financieros, promoviendo una toma de decisiones más informada y transparente.

Reconocimientos

Deseo expresar mi más sincero agradecimiento a la Escuela Superior Politécnica de Chimborazo, así como a nuestros docentes, el distinguido Dr. Wilson Gustavo Chango Sailema por su valiosa colaboración y apoyo en la realización de este estudio.

Los autores declaran la contribución y participación equitativa de roles de autoría para esta publicación.

Referencias

- Akinjole, A., Shobayo, O., Popoola, J., Okoyeigbo, O., & Ogunleye, B. (2024). Ensemble-Based Machine Learning Algorithm for Loan Default Risk Prediction. *Mathematics*, 12(21), 3423. <https://doi.org/10.3390/MATH12213423>
- Akoka, J., Comyn-Wattiau, I., & Storey, V. C. (2023). Design Science Research: Progression, Schools of Thought and Research Themes. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 13873 LNCS, 235–249. https://doi.org/10.1007/978-3-031-32808-4_15
- Alonso Robisco, A., & Carbó Martínez, J. M. (2022). Measuring the model risk-adjusted performance of machine learning algorithms in credit default prediction. *Financial Innovation*, 8(1), 1–35. <https://doi.org/10.1186/S40854-022-00366-1>
- Alvi, J., Arif, I., & Nizam, K. (2024). Advancing financial resilience: A systematic review of default prediction models and future directions in credit risk management. *Heliyon*, 10(21), e39770. <https://doi.org/10.1016/J.HELIYON.2024.E39770>
- Aurona, G., & Richard, B. (2023). Design Science Research for a New Society: Society 5.0. *18th International Conference on Design Science Research in Information Systems and Technology, DESRIST 2023, Pretoria, South Africa, May 31 – June 2, 2023, Proceedings*, 13873. <https://doi.org/10.1007/978-3-031-32808-4>
- Dastile, X., Celik, T., & Potsane, M. (2020). Statistical and machine learning models in credit scoring: A systematic literature survey. *Applied Soft Computing*, 91, 106263. <https://doi.org/10.1016/J.ASOC.2020.106263>
- Delport, P. M. J., Solms, R. Von, & Gerber, M. (2024). Methodological Guidelines for Design Science Research. *Procedia Computer Science*, 237, 195–203. <https://doi.org/10.1016/J.PROCS.2024.05.096>
- Gao, B., & Balyan, V. (2022). Construction of a financial default risk prediction model based on the LightGBM algorithm. *Journal of Intelligent Systems*, 31(1), 767–779. <https://doi.org/10.1515/JISYS-2022-0036>
- Goecks, L. S., De Souza, M., Librelato, T. P., & Trento, L. R. (2021). Design Science Research in practice: review of applications in Industrial Engineering. *Gestão & Produção*, 28(4), e5811. <https://doi.org/10.1590/1806-9649-2021V28E5811>
- Gregor, S. (2021). Reflections on the Practice of Design Science in Information Systems. *Engineering the Transformation of the Enterprise: A Design Science Research Perspective*, 101–113. https://doi.org/10.1007/978-3-030-84655-8_7
- Ha, T., Xiao, D., Katsikis, V. N., Khan, H., Li, S., Zhu, M., Shia, B.-C., Su, M., & Liu, J. (2024). Consumer Default Risk Portrait: An Intelligent Management Framework of Online Consumer Credit Default Risk. *Mathematics* 2024, Vol. 12, Page 1582, 12(10), 1582. <https://doi.org/10.3390/MATH12101582>

- Kim, H., Cho, H., & Ryu, D. (2020). Corporate Default Predictions Using Machine Learning: Literature Review. *Sustainability* 2020, Vol. 12, Page 6325, 12(16), 6325. <https://doi.org/10.3390/SU12166325>
- Lai, L. (2020). Loan Default Prediction with Machine Learning Techniques. *Proceedings - 2020 International Conference on Computer Communication and Network Security, CCNS 2020*, 5–9. <https://doi.org/10.1109/CCNS50731.2020.00009>
- Liu, J., Liu, J., Wu, C., & Wang, S. (2024). Enhancing credit risk prediction based on ensemble tree-based feature transformation and logistic regression. *Journal of Forecasting*, 43(2), 429–455. <https://doi.org/https://doi.org/10.1002/for.3040>
- Montevechi, A. A., Miranda, R. de C., Medeiros, A. L., & Montevechi, J. A. B. (2024). Advancing credit risk modelling with Machine Learning: A comprehensive review of the state-of-the-art. *Engineering Applications of Artificial Intelligence*, 137, 109082. <https://doi.org/10.1016/J.ENGAPPAI.2024.109082>
- Mujo, A., Nikolla, S., Hoxha, E., & Pelivani, E. (2025). Explainable AI in Credit Scoring: Improving Transparency in Loan Decisions. *Journal of Information Systems Engineering and Management*, 10(27s), 506–515. <https://doi.org/10.52783/JISEM.V10I27S.4437>
- Qin, C., Zhang, Y., Bao, F., Zhang, C., Liu, P., & Liu, P. (2021). XGBoost optimized by adaptive particle swarm optimization for credit scoring. *Mathematical Problems in Engineering*, 2021. <https://doi.org/10.1155/2021/6655510>
- Ren, H. (2025). Machine Learning-Based Prediction of Customer Churn Risk in E-commerce. *Advances in Economics, Management and Political Sciences*, 153(1), 47–52. <https://doi.org/10.54254/2754-1169/2024.19473>
- Rios, S. B., & Arbeláez, D. H. (2023). Construcción de un modelo para predecir la morosidad de cartera. *Cuaderno Activa*, 15(1). <https://doi.org/10.53995/20278101.1229>
- scikit-learn.org. (n.d.-a). 3.1. Cross-validation: evaluating estimator performance — *scikit-learn 1.6.1 documentation*. Retrieved May 4, 2025, from https://scikit-learn.org/stable/modules/cross_validation.html
- scikit-learn.org. (n.d.-b). 3.2. Tuning the hyper-parameters of an estimator — *scikit-learn 1.6.1 documentation*. Retrieved May 4, 2025, from https://scikit-learn.org/stable/modules/grid_search.html
- Shahid, A., Hussain, M., & Iqbal, A. (2023). Machine Learning Based Improved Customer Churn Prediction Model for Telecommunications Industry. *18th IEEE International Conference on Emerging Technologies, ICET 2023*, 147–153. <https://doi.org/10.1109/ICET59753.2023.10375037>
- Sheikh, M. A., Goel, A. K., & Kumar, T. (2020). An Approach for Prediction of Loan Approval using Machine Learning Algorithm. *Proceedings of the International Conference on Electronics and Sustainable Communication Systems, ICESC 2020*, 490–494. <https://doi.org/10.1109/ICESC48915.2020.9155614>
- Wang, Z., Zhang, H., Wang, J., Jiang, C., He, H., & Ding, Y. (2025). Forecasting time to risk based on multi-party data: An explainable privacy-preserving decentralized survival analysis method. *Information Processing and Management*, 62(1). <https://doi.org/10.1016/j.ipm.2024.103881>
- Xu, T. (2024). Comparative Analysis of Machine Learning Algorithms for Consumer Credit Risk Assessment. *Transactions on Computer Science and Intelligent Systems Research*, 4, 60–67. <https://doi.org/10.62051/R1M3PG16>
- Yemmanuru, P. K., Yeboah, J., & Nti, I. K. (2024). Customer Credit Risk: Application and Evaluation of Machine Learning and Deep Learning Models. *2024 IEEE 3rd International Conference on Computing and Machine Intelligence, ICMI 2024 - Proceedings*. <https://doi.org/10.1109/ICMI60790.2024.10585896>
- Zhang, L., & Wang, L. (2023). An Ensemble Learning-Enhanced Smart Prediction Model for Financial Credit Risks. <https://doi.org/10.1142/S0218126624501299>, 33(7). <https://doi.org/10.1142/S0218126624501299>