

Revisión teórica de modelos de Machine Learning para la predicción del comportamiento de pago en clientes gestionados desde contact-center, sector cobranzas

Viviana Andrea García Meza

Sady Humberto Lázaro Barrera

Asesor

Fernando Luis Carrascal Porras

Universidad Nacional Abierta y a Distancia UNAD
Escuela de Ciencias Básicas, Tecnología e Ingeniería
Especialización en Ciencia de Datos y Analítica

2024

Nota de Aceptación

Ing. Fernando Luis Carrascal Porras

Jurado

2024

Resumen

La presente monografía tuvo como objetivo analizar los modelos de Machine Learning aplicados a la predicción del comportamiento de pago en clientes gestionados desde Contact-Center en el sector de cobranzas. A través de una metodología exploratoria-descriptiva, se realizó una revisión bibliográfica exhaustiva que permitió identificar, clasificar y comparar modelos como Random Forest, Árboles de Decisión, Redes Neuronales y Máquinas de Soporte Vectorial (SVM), etc. Estos modelos fueron evaluados en términos de precisión, capacidad predictiva y aplicabilidad en escenarios operativos reales.

El estudio se enfocó en identificar los métodos actuales utilizados en la predicción de pagos, examinar las variables determinantes en el desempeño de los modelos y establecer cuáles enfoques resultaron más efectivos para optimizar la gestión de cobranzas. Entre las variables analizadas se incluyeron factores económicos, demográficos y patrones históricos de mora, los cuales se identificaron como determinantes en la precisión de las predicciones.

Los resultados destacaron que modelos como Random Forest y XGBoost ofrecieron un balance óptimo entre desempeño y adaptabilidad, superando las limitaciones de técnicas tradicionales en este ámbito. Finalmente, se propusieron recomendaciones prácticas para implementar soluciones basadas en aprendizaje automático, con el fin de contribuir al desarrollo de estrategias de cobranza más efectivas y escalables.

Palabras claves: Machine Learning, Predicción de pagos, Contact-Center, Cobranzas, Inteligencia Artificial.

Abstract

This monograph aimed to analyze Machine Learning models applied to payment behavior prediction in clients managed from Contact-Centers in the collections sector. Through an exploratory-descriptive methodology, an exhaustive bibliographic review was conducted to identify, classify, and compare models such as Random Forest, Decision Trees, Neural Networks, and Support Vector Machines (SVM), among others. These models were evaluated in terms of accuracy, predictive capacity, and applicability in real operational scenarios.

The study focused on identifying the current methods used for payment prediction, examining the variables determining model performance, and establishing which approaches were the most effective in optimizing collections management. Among the variables analyzed were economic and demographic factors and historical delinquency patterns, identified as key determinants of prediction accuracy.

The results highlighted that models such as Random Forest and XGBoost offered an optimal balance between performance and adaptability, surpassing the limitations of traditional techniques in this field. Finally, practical recommendations were proposed to implement machine learning-based solutions to contribute to developing more effective and scalable collections strategies.

Keywords: Machine Learning, Payment Prediction, Contact-Center, Collections, Artificial Intelligence.

Tabla de Contenido

Introducción	11
Descripción del Problema	13
Planteamiento del Problema.....	13
Justificación	15
Objetivos	17
Objetivo General	17
Objetivos Específicos	17
Marco de Referencial	18
Marco Conceptual	18
Marco Teórico	20
Metodología	25
Enfoque de Investigación	25
Fases de la Investigación.....	25
Instrumentos de Investigación.....	26
Análisis de Datos.....	26
Limitaciones del Estudio.....	26
Reflexión del Estudio	27
Identificación de los Métodos Actuales de Machine Learning.....	28
Otros Métodos Utilizados	32
VARIABLES Y FACTORES DETERMINANTES	33
Comportamiento de Variables según la Técnica de Machine Learning	34
Modelos de Machine Learning con Mejor Desempeño	42

Análisis Comparativo de Modelos de Machine Learning.....	81
Conclusiones.....	85
Recomendaciones.....	87
Referencias Bibliográficas.....	89
Apéndices.....	93

Lista de Tablas

Tabla 1 <i>Cadenas de Búsqueda Investigaciones</i>	29
Tabla 2 <i>Frecuencia de Métodos de Machine Learning Utilizados</i>	30
Tabla 3 <i>Frecuencia de Comportamiento de Variables según la Técnica de Machine Learning</i>	34
Tabla 4 <i>Frecuencia de Comportamiento de Variables según Idioma</i>	36
Tabla 5 <i>Frecuencia de Comportamiento de Variables según Tipo de Documento</i>	37
Tabla 6 <i>Frecuencia de Comportamiento de Variables según País</i>	39
Tabla 7 <i>Frecuencia de Comportamiento de Variables según País</i>	82

Lista de Figuras

Figura 1 <i>Resumen Métricas por Modelos Estudio</i>	42
Figura 2 <i>Matriz de Confusión del Modelo K-NN</i>	45
Figura 3 <i>Matriz de Confusión del Modelo de Máquina de Soporte Vectorial</i>	46
Figura 4 <i>Matriz de Confusión del Modelo de Árbol de Decisión</i>	47
Figura 5 <i>Matriz de Confusión del Modelo de Bosques Aleatorios</i>	48
Figura 6 <i>Resultados de Árbol de Decisión Modelo de Clasificación Antes</i>	49
Figura 7 <i>Resultados de Árbol de Decisión Modelo de Clasificación Después</i>	50
Figura 8 <i>Resultados del Modelo Clasificador Random Forest Antes</i>	50
Figura 9 <i>Resultados del Modelo Clasificador Random Forest Después</i>	51
Figura 10 <i>Comparación de Ambos Modelos</i>	52
Figura 11 <i>Desempeño de la Clasificación en la Primera Etapa</i>	55
Figura 12 <i>Desempeño de la Clasificación en la Segunda Etapa</i>	56
Figura 13 <i>Desempeño de Predicción de LGD en la Tercera Etapa</i>	57
Figura 14 <i>Pronósticos de un Día Adelante</i>	59
Figura 15 <i>Resultados del Test de Nemenyi al 5%</i>	59
Figura 16 <i>Diferencias de Precisión en los Pronósticos</i>	60
Figura 17 <i>Diagramas de Caja sobre las Diferencias de Precisión en los Pronósticos</i>	62
Figura 18 <i>Resultados del Test de Nemenyi al Nivel de Significancia del 5%</i>	63
Figura 19 <i>Comparación de Medidas de Desempeño para todos los Clasificadores</i>	65
Figura 20 <i>Medidas de Desempeño para el Clasificador LSTM</i>	67
Figura 21 <i>Curvas ROC, AUC y de Calibración</i>	68
Figura 22 <i>Cambio en la Tasa de Error OOB</i>	71

Figura 23	<i>Comparación de Modelos con Diferentes Conjuntos de Predictores</i>	72
Figura 24	<i>Curvas de Supervivencia de Kaplan-Meier.....</i>	73
Figura 25	<i>Comparaciones de RMSE Fuera de Muestra de DFM con Modelos de ML</i>	76
Figura 26	<i>Comparaciones de RMSE Fuera de Muestra para las Tasas de Crecimiento.....</i>	76
Figura 27	<i>Comparaciones de RMSE Fuera de Muestra.....</i>	77
Figura 28	<i>Valores de Lift por Decil para Tfull.....</i>	78
Figura 29	<i>Resultado de la Asignación de la Lista de Contactos del Cliente</i>	80

Lista de Apéndices

Apéndice A <i>Matriz - Monografía Revisión Documental</i>	93
--	----

Introducción

En la gestión de cobranzas y el manejo de clientes desde Contact-Center, la predicción del comportamiento de pago se ha convertido en un desafío fundamental. Este sector ha dependido históricamente de métodos tradicionales y manuales que, aunque útiles, no aprovechan completamente el potencial de las tecnologías avanzadas. El auge del Machine Learning ha abierto nuevas oportunidades para optimizar estas operaciones, ofreciendo herramientas más precisas y escalables que permiten anticipar la morosidad y diseñar estrategias proactivas.

La presente monografía aborda esta problemática desde una perspectiva teórica, analizando los principales modelos de Machine Learning aplicados en la predicción de pagos en el sector de cobranzas. A través de una revisión exhaustiva de la literatura científica, se identificaron los métodos actuales más utilizados, como árboles de decisión, redes neuronales y bosques aleatorios, evaluando sus características, limitaciones y precisión en contextos operativos reales.

El objetivo principal de este estudio fue identificar los modelos más efectivos, considerando las variables y factores que influyen en su desempeño. Este análisis no solo contribuye a mejorar la eficiencia de las estrategias de cobranza, sino que también proporciona una base sólida para futuras implementaciones tecnológicas en este sector. Asimismo, se examinaron los desafíos actuales y se propusieron recomendaciones para optimizar el uso de estos modelos en entornos de Contact-Center.

La monografía se organiza en capítulos que detallan el planteamiento del problema, la metodología utilizada y los resultados del análisis comparativo de modelos. En conjunto, este trabajo busca ser un aporte significativo tanto para la academia como para los profesionales del

sector, destacando la importancia de integrar el aprendizaje automático en la gestión de cobranzas para enfrentar los retos de un entorno financiero cada vez más complejo.

Descripción del Problema

Planteamiento del Problema

En el ámbito del sector de cobranzas y la gestión de clientes a través de Contact-Center, la principal barrera que dificulta la implementación efectiva de modelos de Machine Learning para predecir el comportamiento de pago radica en la arraigada dependencia de métodos tradicionales y manuales para evaluar riesgos. Esta dependencia limita significativamente la adopción y aplicación óptima de tecnologías avanzadas como el Machine Learning.

La investigación en la predicción del pago de facturas es crucial para mejorar la gestión financiera en el sector de cobranzas. La adopción de un enfoque preventivo, aprovechando el análisis de datos y el aprendizaje automático, puede identificar a los clientes morosos en etapas tempranas. Estos métodos se muestran superiores a las técnicas estadísticas tradicionales, especialmente en tiempos de choques económicos (Schoonbee et al., 2022). Para abordar este vacío, se propone en este estudio un sistema de soporte de decisiones basado en aprendizaje automático (Rose, 2016), este sistema busca superar las limitaciones impuestas por la dependencia de métodos tradicionales, permitiendo una mejor gestión del riesgo y una mayor eficiencia en la gestión de clientes en el sector de cobranzas y Contact-Center.

En el ámbito de la gestión de carteras de cobranza, se ha aplicado aprendizaje supervisado para identificar atributos clave de los clientes y predecir registros efectivos. Esto se complementa con una revisión de literatura enfocada en la segmentación de clientes mediante algoritmos de inteligencia artificial y aprendizaje automático (Ortiz et al., 2022). Los resultados muestran la prevalencia de la segmentación algorítmica, con una amplia variedad de algoritmos y métricas empleados para mejorar la gestión en este sector.

Se aborda el desafío de mejorar la eficiencia en la predicción de pagos en clientes

gestionados desde Contact-Center, sector cobranzas. A pesar de los avances en la aplicación de aprendizaje automático, la adopción de estos modelos se ve obstaculizada por la dependencia de métodos tradicionales. Se centra en superar estas limitaciones. Se investigan los usos actuales de la predicción de pagos basada en Machine Learning y se examinan las variables y factores que influyen en la predicción de pagos. Además, se proponen estrategias para mantener actualizados estos modelos (Zhu et al., 2023). El objetivo es contribuir a mejorar la eficiencia en la gestión de clientes en este sector.

Justificación

La gestión eficiente de cobranzas es esencial para la salud financiera de las empresas, especialmente en el sector de cobranzas, donde la ineficiencia puede llevar a la pérdida de clientes y disminución de ingresos. Con los avances tecnológicos, los modelos de Machine Learning se han convertido en herramientas clave para predecir el comportamiento de pago de clientes gestionados desde Contact-Center. Sin embargo, su implementación actual presenta limitaciones que afectan su efectividad. Esta monografía busca analizar las características y limitaciones de estos modelos en la predicción de pagos, identificar los métodos actuales y variables influyentes, y determinar qué modelos optimizan mejor la gestión en el sector.

Este análisis es especialmente relevante en Colombia, donde el acceso a productos financieros ha mejorado, aunque persisten desafíos. Según el Reporte de Inclusión Financiera 2023, aunque el 94,6% de los adultos en Colombia tiene acceso a al menos un producto financiero, solo el 82,0% los utiliza activamente, reflejando una brecha entre acceso y uso efectivo (Banca de las Oportunidades y Superintendencia Financiera de Colombia, 2023). Esto sugiere que, aunque muchas personas pueden acceder a servicios financieros, aún tienen dificultades para gestionarlos, incrementando el riesgo de incumplimientos de pago. En este contexto, los modelos predictivos son clave para identificar a los clientes en riesgo de morosidad y permitir una gestión de cobranza más proactiva.

Además, el 35,3% de los adultos cuenta con algún producto de crédito, y solo el 22,2% de estos posee una tarjeta de crédito, lo cual indica que una parte significativa de la población depende de otros tipos de financiamiento, como el microcrédito. Esta realidad refuerza la importancia de contar con sistemas avanzados que no solo anticipen el comportamiento de pago, sino que adapten las estrategias a distintos segmentos poblacionales.

Finalmente, dado que el 45% de la población colombiana aún no tiene acceso a productos financieros formales, los Contact-Centers y las empresas de cobranza desempeñan un rol crucial, especialmente con el crecimiento de plataformas digitales y pagos móviles. Los modelos predictivos deben adaptarse a estos canales para aprovechar la digitalización, fomentar la inclusión financiera y reducir la morosidad. Las empresas de cobranzas y los profesionales del sector financiero pueden beneficiarse de este estudio para optimizar sus procesos y mitigar el riesgo de pérdidas.

Objetivos

Objetivo General

Analizar las características y limitaciones de los modelos de Machine Learning utilizados en la predicción de pagos en Contact-Center del sector cobranzas.

Objetivos Específicos

Identificar los métodos actuales de Machine Learning para la predicción de pagos en Contact-Center en sector cobranzas.

Examinar las variables y factores que influyen en la predicción de pagos en los modelos de Machine Learning.

Ilustrar cuales de los modelos de Machine Learning presentan un mejor comportamiento en la predicción de pagos en Contact-Center del sector cobranzas.

Marco de Referencial

Marco Conceptual

El Machine Learning hace parte de una subdisciplina de la inteligencia artificial, la cual en sus inicios fue definida por Arthur Samuel en 1959, su definición comprende que el aprendizaje automático tiene como objetivo equipar a las máquinas con la habilidad de aprender de los datos a través de algoritmos, sin necesidad de programación explícita. Y, este aprendizaje se asemeja a cómo los humanos aprendemos de ejemplos sin depender de fórmulas o reglas específicas. De tal modo, permite a las máquinas generar modelos que pueden producir resultados en situaciones nuevas que no se conocían durante el aprendizaje. Esto es especialmente útil en situaciones donde no existe o es difícil encontrar una fórmula que facilite una respuesta exacta basada en ciertas variables (Raschka & Mirjalili, 2017).

En Machine Learning existen diversas técnicas y algoritmos que han demostrado su eficacia en la resolución de problemas complejos. Este marco conceptual profundiza en algunos de los más prominentes que se definen a continuación:

Los Bosques Aleatorios son una expansión del Bagging, introduciendo aleatoriedad en las muestras de entrenamiento y las características de cada aprendiz base. Esta técnica, que tiene un rendimiento similar a las técnicas de boosting pero requiere menos ajuste, se utiliza para construir mejores conjuntos de bosques aleatorios. Se discuten las fortalezas y debilidades de los Bosques Aleatorios, y se presentan ejemplos de su implementación en scikit-learn para regresión y clasificación (Kyriakides & Margaritis, 2019).

La regresión logística es un tipo de análisis de regresión que ayuda a estimar la probabilidad de que ocurra un evento basado en algunos parámetros dados. Se utiliza como una técnica de clasificación con un resultado binario. Las probabilidades que describen los posibles

resultados de un solo ensayo se modelan, como una función de las variables explicativas (predictoras), utilizando una función logística (Madhavan, 2015).

Los árboles de decisión son una forma de algoritmo de aprendizaje automático que proporciona una estructura de reglas basada en las características de los datos de entrada. Estas reglas se generan a través de un proceso iterativo de división del conjunto de datos, buscando maximizar la información producida en cada división. Este proceso de maximización de la información resulta en la separación del mayor número de instancias restantes en el conjunto de datos (Kyriakides, & Margaritis, 2019).

Las Máquinas de Vectores de Soporte (SVM) son un tipo de algoritmo de aprendizaje automático que utiliza un subconjunto de datos de entrenamiento, específicamente los puntos de datos cercanos al borde de cada clase, para definir un hiperplano separador. Estos casos límite se conocen como vectores de soporte. El objetivo de una SVM es encontrar el hiperplano que maximiza el margen (distancia) entre los vectores de soporte. Para clasificar clases no linealmente separables, las SVM utilizan el truco del kernel para mapear los datos en un espacio de mayor dimensión, donde pueden volverse linealmente separables (Kyriakides, & Margaritis, 2019).

Las Redes Neuronales son algoritmos de aprendizaje automático inspirados en la estructura del cerebro biológico. Consisten en neuronas organizadas en capas que procesan datos de entrada y producen predicciones. Cada neurona multiplica sus entradas por pesos específicos, suma los resultados y aplica una función de activación para definir su salida. El objetivo es optimizar estos pesos para minimizar una función de coste. Las redes neuronales se utilizan tanto para regresión como para clasificación, y se optimizan mediante algoritmos como el descenso de gradiente estocástico (SGD), que actualiza los pesos en función del gradiente del error.

(Kyriakides, & Margaritis, 2019).

Marco Teórico

La gestión de cobranza es un proceso complejo y esencial en una empresa, que implica una serie de acciones y estrategias para persuadir a los clientes morosos a cumplir con sus obligaciones de pago. Este proceso es crucial para evitar que las cuentas por cobrar se conviertan en incobrables, afectando negativamente el flujo de caja y la liquidez de la empresa.

La gestión de cobranza se divide en tres etapas: preventiva, administrativa y judicial. La etapa preventiva se aplica a obligaciones próximas a vencer. La etapa administrativa comienza cuando se produce el retraso en el pago. La etapa judicial se inicia cuando no es posible llegar a un acuerdo con el cliente deudor, recurriendo a acciones legales (Chief Revenue Officer en Nubox, 2024).

En el ámbito de la gestión de riesgos crediticios y cobranzas, la predicción del comportamiento de pago de los clientes es una necesidad crítica para optimizar estrategias y mitigar el riesgo de incumplimiento. Los modelos de Machine Learning (ML) han emergido como herramientas eficaces para abordar este desafío, permitiendo identificar patrones subyacentes y anticipar el comportamiento futuro de pago (Singh et al., 2024).

En el ámbito de la gestión de cobranzas, los modelos de Machine Learning (ML) son fundamentales para optimizar la efectividad y personalización de las estrategias de cobranza. Fan et al. (2023), resaltan la importancia de la interpretación transparente de los resultados de ML, especialmente en modelos como la regresión logística, utilizados para predecir incumplimientos de pago. Otros métodos, como árboles de decisión, Random Forest y Gradient Boosting, han demostrado gran efectividad en la predicción del comportamiento de pago. La combinación de estas técnicas permite optimizar los procesos de cobranza, reducir riesgos y adaptar estrategias

para mejorar la recuperación de pagos y eficiencia operativa.

Según Ala'raj et al. (2022), la gestión bancaria puede beneficiarse del aprendizaje automático para calificar clientes de tarjetas de crédito, modelando y prediciendo el comportamiento en tres áreas: probabilidad de pagos perdidos, comportamiento de compra y agrupación de clientes según la expectativa de pérdida. Los modelos usados incluyen regresión logística, bosque aleatorio, máquina de vectores de soporte y red neuronal de perceptrón multicapa. Se presentaron y validaron dos modelos, MP-LSTM y PE-LSTM, en datos transaccionales y no transaccionales, mostrando utilidad para clasificar clientes por riesgo de impago. Los compararon con modelos tradicionales: MLP, SVM, RF y LogR.

Smirnov (2016), propuso dos modelos separados para deudores por primera vez y para deudores repetidos, explorando el efecto de diferentes predictores en un modelo. Los Bosques de Supervivencia Aleatorios demostraron tener ventajas sobre el modelo de Riesgos Proporcionales de Cox ya que no hay suposiciones subyacentes que deban tenerse en cuenta. En general, Smirnov (2016), concluyó que el modelo de Bosques de Supervivencia Aleatorios que además utiliza el comportamiento de pago histórico de los deudores es el que mejor funciona para clasificar los tiempos de pago de las facturas atrasadas.

Chapman y Desai (2023), evaluaron la utilidad de los datos de pagos completos para predecir macroeconómicamente en Canadá, así como determinar si los modelos de aprendizaje automático mejoran la precisión de estas predicciones. Los autores se preguntaron en qué medida los datos de pagos completos y los modelos de aprendizaje automático pueden mejorar la precisión de las predicciones macroeconómicas, particularmente durante períodos de crisis como la pandemia de COVID-19. Utilizaron varios modelos, incluyendo elastic net, support vector machines, random forest, gradient boosting y redes neuronales artificiales.

Además, los modelos de aprendizaje automático, como los Bosques de Supervivencia Aleatorios, pueden mejorar la precisión de estas predicciones. Esto puede permitir a las empresas identificar más eficazmente a los clientes que pueden tener dificultades para pagar, permitiendo una gestión de cobranzas más eficaz (Chapman & Desai, 2023).

La evaluación del rendimiento de estos modelos es esencial para garantizar su utilidad práctica. Singh et al. (2024), subraya la importancia de métricas como precisión, sensibilidad, AUC y puntuación F1 para medir el desempeño de los algoritmos en la predicción de la deserción de clientes. Estas métricas son críticas para evaluar tanto las fortalezas como las debilidades de los modelos en términos de sensibilidad y precisión.

La distinción entre deudores por primera vez y repetidos puede ser útil para la gestión de cobranzas. Las empresas pueden necesitar estrategias de cobro diferentes para estos dos grupos, y los modelos propuestos por Chapman y Desai (2023), pueden ayudar a las empresas a personalizar sus enfoques de cobranza en consecuencia.

En el contexto de la gestión de cobranzas desde un Contact-Center, los modelos de Machine Learning aprovechan una amplia gama de variables, desde datos demográficos hasta información financiera y patrones de interacción con el Contact-Center, para desarrollar predicciones precisas. De acuerdo con Thomas et al. (2002), el uso de variables como el historial de pagos, el nivel de endeudamiento y la frecuencia de contacto mejora significativamente la capacidad de los modelos para identificar clientes propensos a incumplir sus compromisos financieros. Además, Anderson (2007), destaca que la inclusión de datos externos, como indicadores económicos y tendencias de mercado, enriquece aún más la capacidad predictiva de estos modelos.

La importancia de seleccionar cuidadosamente las variables utilizadas en estos modelos

ha sido destacada por Ortiz et al. (2022), quienes enfatizan la necesidad de considerar factores comportamentales y económicos específicos del sector de cobranzas. Por ejemplo, la edad y el estado civil pueden ser predictores significativos del comportamiento de pago, junto con la frecuencia y el tipo de interacciones con el Contact-Center. Estos autores también resaltan la necesidad de mantener actualizados los modelos con datos en tiempo real para adaptarse a cambios rápidos en las condiciones económicas y del mercado (Ortiz et al., 2022).

En estudios sobre análisis del comportamiento, se destacan variables claves relacionadas con pagos, edad de mora, y características demográficas y económicas. Ortiz et al. (2022), utilizan variables como días de mora y número máximo de intentos de pago para analizar la efectividad en la gestión de cobros, aportando información valiosa sobre el comportamiento de pago en el sector financiero. Wagh et al. (2021), consideran variables como tipo de contrato y antigüedad del cliente, cruciales para predecir la deserción en telecomunicaciones y entender mejor los patrones de comportamiento de los usuarios.

En estudios de análisis de comportamiento también juegan un papel significativo las variables demográficas. En el estudio de Fan et al. (2023), se utilizan variables como el número de empleados (NoEmp) de las empresas para desarrollar modelos predictivos de riesgo crediticio, proporcionando insights sobre cómo las características demográficas influyen en el comportamiento financiero. Asimismo, Aslam et al. (2022), incluyen factores sociales y de interacción en su análisis del comportamiento de los usuarios en plataformas digitales, lo que ayuda a entender cómo las variables demográficas afectan las decisiones y acciones en línea.

En el sector de las telecomunicaciones, la pérdida de clientes, o "churn", es un desafío constante que afecta significativamente la rentabilidad de las empresas. La capacidad de predecir con precisión qué clientes podrían abandonar el servicio es crucial para implementar estrategias

efectivas de retención. En este contexto, el estudio de Xu et al. (2021), propone un enfoque innovador que emplea técnicas avanzadas de aprendizaje automático y construcción de características para predecir la pérdida de clientes con alta precisión. Este enfoque no solo mejora la exactitud de las predicciones, sino que también enfatiza la importancia de la gestión de relaciones con los clientes para reducir la rotación.

En el ámbito minorista, pronosticar con precisión el flujo de clientes es esencial para la gestión eficiente de recursos y la optimización de ventas. La investigación de Ma y Fildes (2020), aborda este desafío proponiendo una solución innovadora basada en el Árbol de Regresión de Impulso Gradiente (GBRT). Este método utiliza datos de pagos móviles de terceros para predecir los flujos diarios de clientes. Los resultados del estudio indican que el enfoque GBRT supera significativamente a los métodos tradicionales basados en tiendas individuales, subrayando el valor de la agrupación de datos para mejorar la precisión en la predicción de múltiples series de tiempo.

Las variables económicas también son fundamentales en este tipo de estudios. Fan et al. (2023), incorporan variables como los ingresos netos (NetIncome) y el índice de riesgo (RiskIndex) en sus modelos de predicción de riesgo crediticio, lo que destaca la influencia de las condiciones económicas en el comportamiento crediticio. Ortiz et al. (2022), analizan variables económicas y de valor financiero, como el valor de las transacciones, para mejorar las estrategias de inteligencia empresarial y entender mejor el comportamiento económico de los clientes.

Metodología

Enfoque de Investigación

Tipo de estudio: La investigación es de tipo descriptivo y exploratorio, ya que busca analizar y describir las características y limitaciones de los modelos de Machine Learning utilizados en la predicción de pagos en Contact-Center del sector cobranzas.

Método: Se utilizó un enfoque cualitativo para realizar una revisión exhaustiva de la literatura existente y un análisis comparativo de los modelos de Machine Learning aplicados en el sector.

Fases de la Investigación

Revisión bibliográfica: Se realizó una búsqueda sistemática en repositorios y bases de datos académicas, incluyendo Web of Science, Scopus, Repositorio UNAL, Repositorio UAN, Science Direct y Google Académico, para recolectar información sobre estudios previos relacionados con la predicción del comportamiento de pago mediante Machine Learning.

Los artículos seleccionados serán aquellos que presenten aplicaciones relevantes en el sector de cobranzas y Contact-Center, y que incluyan descripciones de los algoritmos y variables utilizadas.

Análisis de variables: se analizaron las variables y factores que influyen en la predicción de pagos en los modelos de Machine Learning. Este análisis se enfocará en identificar patrones comunes y diferencias en las características de los datos utilizados en los estudios previos.

Comparación de modelos: Se identificó y comparó los principales modelos de Machine Learning aplicados en la gestión de cobranzas, como árboles de decisión, redes neuronales, máquinas de vectores de soporte (SVM), y bosques aleatorios.

Para cada modelo, se evaluaron sus ventajas, limitaciones y precisión en la predicción de pagos.

Ilustración de los modelos con mejor comportamiento: Con base en los hallazgos de la revisión y comparación de modelos, se ilustró cuáles de los modelos de Machine Learning presentan el mejor desempeño en la predicción de pagos en el sector de cobranzas.

Este análisis se enfocó en evaluar el comportamiento de los modelos en términos de precisión, velocidad de convergencia y capacidad de generalización en escenarios reales, destacando aquellos modelos que ofrecen mayor fiabilidad y eficiencia en el contexto de la gestión de pagos en Contact-Centers.

Instrumentos de Investigación

Estudios de casos: Se consideraron estudios de casos que ilustraron la implementación de modelos de Machine Learning en el sector de Contact-Centers para comprender las condiciones reales y desafíos prácticos en la aplicación de estas tecnologías.

Análisis de Datos

Los datos recolectados fueron categorizados y codificados para identificar las principales tendencias en la aplicación de Machine Learning en el sector de cobranzas.

Se realizó un análisis comparativo de las técnicas utilizadas y los resultados obtenidos en diferentes estudios, evaluando la efectividad de cada modelo y su aplicabilidad en diferentes escenarios operativos.

Limitaciones del Estudio

El estudio estuvo limitado a la información disponible en los repositorios y bases de datos académicas mencionados y a los estudios publicados hasta la fecha, lo cual pudo afectar la

generalización de los resultados a otros contextos o sectores que no hayan sido explorados en la literatura existente.

Reflexión del Estudio

Se realizó un análisis detallado de los modelos de Machine Learning más efectivos para la predicción del comportamiento de pago en Contact-Centers.

Se ilustró cuáles de los modelos de Machine Learning presentan un mejor comportamiento en la predicción de pagos en Contact-Center del sector cobranza.

Identificación de los Métodos Actuales de Machine Learning

En el sector de cobranzas, la predicción del comportamiento de pago es clave para mejorar la eficiencia en la gestión de clientes y optimizar las estrategias de recaudación. Sin embargo, los métodos tradicionales, basados en análisis manuales, enfrentan limitaciones importantes ante el creciente volumen y complejidad de los datos en la era digital. En este contexto, los modelos de Machine Learning han demostrado ser herramientas efectivas para superar estas barreras, ofreciendo soluciones más precisas y escalables (Schoonbee et al., 2022).

En la primera etapa de este estudio, se realizó una búsqueda exhaustiva de investigaciones científicas relevantes utilizando cadenas específicas de términos relacionados con predicción de pagos y modelos algorítmicos (ver Tabla 1). Como resultado, se identificaron 50 artículos, de los cuales 41 abordan explícitamente el uso de modelos de Machine Learning aplicados a la predicción del comportamiento de pago en clientes.

Tabla 1*Cadenas de Búsqueda Investigaciones*

Cadenas de Búsqueda	Cant.	%
"Machine Learning" AND customer OR ensamble OR prediction OR algorithm OR call center OR classification OR prediction OR Contact Center OR score OR segment OR telecom*	24	48%
("Payment Prediction" OR "payment forecasting" OR "Payment Behavior Forecasting" OR "Payment Trends") AND ("Models Machine Learning" OR "Machine Learning Approaches" OR "Machine Learning" OR "Machine Learning Algorithms" OR "Machine Learning Techniques")	9	18%
"Payment Prediction" OR "payment forecasting" OR "Payment Behavior Forecasting" OR "Payment Trends"	8	16%
("payment prediction models" OR "payment forecasting" OR "financial prediction" OR "credit scoring") AND ("accounts receivable" OR "debt collection" OR "receivables management") AND ("artificial intelligence" OR "machine learning" OR "neural networks" OR "deep learning") AND ("data analysis" OR "data mining" OR "predictive analytics")	6	12%
"Aprendizaje automático" AND pagos OR clientes OR BPO	2	4%
"Payment Prediction" AND ("Models Machine Learning" OR "Machine Learning Approaches" OR "Machine Learning" OR "Machine Learning Algorithms" OR "Machine Learning Techniques")	1	2%
Total general	50	100%

A continuación, en la Tabla 2 resume los métodos más utilizados en la literatura, destacando aquellos con mayor frecuencia de aplicación en estudios de predicción de pagos. Entre estos, Random Forest, Regresión Logística y Árboles de Decisión son los métodos principales, con frecuencias de uso del 18%, 15% y 14%, respectivamente. Estos métodos, junto con Redes Neuronales, XGBoost, K-Nearest Neighbors (K-NN) y Kernel SVM, se seleccionaron

debido a su capacidad para manejar datos extensos y proporcionar predicciones precisas en el contexto de cobranzas, haciendo frente a los desafíos de segmentación y predicción en Contact-Centers.

Tabla 2

Frecuencia de Métodos de Machine Learning Utilizados

Métodos Machine Learning	Total	%
Random forest	23	18%
Regresión logística	19	15%
Arboles de decisión	18	14%
Redes neuroanales	15	11%
XGBoost	12	9%
Vecino más cercano K-NN	12	9%
Kernel SVM	10	8%
Naive Bayes	4	3%
Adaboost	3	2%
Clustering	3	2%
Regresión Linear	3	2%
Aprendizaje conjunto	2	2%
Boosting	2	2%
Regresión Ridge	1	1%
Regresión Agrupada	1	1%
Agrupación equidistante	1	1%
PMM	1	1%
Redes Bayesianas	1	1%

Random Forest (18%): Fue uno de los métodos más empleados, dado su rendimiento en el manejo de grandes volúmenes de datos y en la reducción de sobreajuste. Antonio et al. (2024), aplicaron este método en Contact-Centers para disminuir errores en las predicciones y segmentar

de forma precisa a clientes morosos. Campos (2020), también utilizó Random Forest como un modelo de ensamblaje de árboles de decisión, demostrando que esta técnica mejoraba la precisión en la predicción de pagos, identificando con mayor eficacia a clientes en alto riesgo.

Regresión Logística (15%): Se aplicó ampliamente en el sector de cobranzas debido a su habilidad para modelar probabilidades de pago en tiempo real. Xu et al. (2021), implementaron este método para identificar rápidamente a clientes con alto riesgo de morosidad, mejorando la eficiencia operativa en Contact-Centers. Ortiz et al. (2022), también emplearon la regresión logística para evaluar perfiles de riesgo de clientes, facilitando una segmentación basada en la probabilidad de impago.

Árboles de Decisión (14%): Este método destacó por su eficacia en la clasificación de clientes según su historial de pagos. Campos (2020), utilizó árboles de decisión para segmentar clientes con alta morosidad, optimizando las intervenciones de cobranza personalizadas. Zhu et al. (2023), también aplicaron esta técnica en el análisis de impagos, comprobando su eficacia en la clasificación de datos y su adaptabilidad a los Contact-Centers.

Redes Neuronales (11%): Este método resultó útil en varios artículos ya que permitieron analizar patrones complejos en grandes volúmenes de datos, resultandos útiles en la detección de patrones no lineales de pago. Kyriakides y Margaritis (2019), aplicaron redes neuronales en datos de cobranza, mientras que Fan et al. (2023), destacaron su capacidad para adaptarse a datos complejos en la predicción de comportamiento crediticio.

XGBoost (9%): Fue valorado por su rapidez y precisión en entornos de grandes volúmenes de datos. Chapman y Desai (2023), demostraron que este método mejoraba la predicción de pagos en clientes de alto riesgo. Smirnov (2016), encontró que XGBoost superaba a los métodos tradicionales, siendo especialmente efectivo en análisis de riesgo de impago.

K-Nearest Neighbors (K-NN) (9%): Se utilizó principalmente para agrupar clientes con patrones de pago similares, facilitando estrategias específicas de cobranza. Ortiz et al. (2022), emplearon K-NN para identificar patrones de pago en clientes, ajustando estrategias de contacto según esta segmentación detallada.

Kernel SVM (8%): Aunque menos frecuente que otros métodos, Kernel SVM resultó útil en escenarios con relaciones altamente no lineales entre variables. Antonio et al. (2024), destacaron que este modelo era adecuado para predecir comportamientos complejos de pago en clientes de alto riesgo, proporcionando una alternativa eficaz para situaciones con datos difíciles de separar linealmente.

Otros Métodos Utilizados

Además de los métodos mencionados, otros enfoques menos frecuentes se aplicaron en estudios de predicción de pagos, aunque en menor medida. Métodos como Naive Bayes y Adaboost se emplearon en estudios específicos donde la simplicidad y velocidad en la clasificación eran prioritarios. En estos casos, Xu et al. (2021), aplicaron Naive Bayes para clasificar rápidamente grandes volúmenes de clientes, mientras que otros métodos como Clustering y Aprendizaje Conjunto se usaron para segmentar clientes en grupos, permitiendo enfoques personalizados y escalables en Contact-Centers.

Variables y Factores Determinantes

La predicción de pagos en el sector de cobranzas gestionado desde Contact-Centers depende en gran medida de la selección y el análisis de variables específicas, así como de los antecedentes y técnicas de Machine Learning empleadas. La tabla de datos obtenida evidencia una diversidad de factores y técnicas aplicadas en este contexto, lo que proporciona una visión integral sobre la investigación en esta área.

Las variables analizadas en los estudios revisados se dividen en varias categorías principales:

Pagos y Edad Mora: Estas variables son fundamentales, ya que representan directamente el comportamiento de pago del cliente y el tiempo que transcurre desde el vencimiento de la obligación. Las técnicas de Machine Learning más frecuentemente asociadas con estas variables incluyen Random Forest, Regresión Logística y Árboles de Decisión, lo que sugiere una preferencia por modelos que permiten interpretar la relación entre el historial de pagos y la morosidad del cliente (Schoonbee et al., 2022).

Demográficas y Económicas: Variables como edad, ubicación geográfica, y situación financiera del cliente también resultan significativas. Estas variables ayudan a segmentar a los clientes y a identificar patrones que influyen el riesgo de incumplimiento. Los modelos Random Forest y XGBoost destacan en el análisis de estas variables debido a su capacidad para manejar grandes volúmenes de datos y realizar selecciones de características de forma automática (Zhu et al., 2023).

Comportamiento de Variables según la Técnica de Machine Learning

Cada técnica de Machine Learning seleccionó variables específicas para optimizar la predicción de pagos en clientes. A continuación, en la tabla 3, se detalla el uso de variables para las principales técnicas, especificando aquellas mencionadas en los estudios analizados.

Tabla 3

Frecuencia de Comportamiento de Variables según la Técnica de Machine Learning

Tipo	Antecedentes Investigativos				Total	%
	Pagos	Edad Mora	Demográficas	Económicas		
Random forest	19	6	6	12	23	18%
Regresión logística	16	7	9	7	19	15%
Arboles de decisión	12	8	6	6	18	14%
Redes neuronales	11	5	3	7	15	11%
XGBoost	11	4	5	9	12	9%
Vecino más cercano K-NN	8	4	3	6	12	9%
Kernel SVM	7	2	2	11	10	8%
Naive Bayes	1	2	0	0	4	3%
Clustering	3	1	1	2	3	2%
Adaboost	2	1	1	1	3	2%
Regresión Linear	3	1	2	2	3	2%
Aprendizaje conjunto	1	1	1	0	2	2%
Boosting	2	1	1	2	2	2%
Regresión Ridge	0	0	1	1	1	1%
Regresión Agrupada	1	0	0	0	1	1%
Agrupación equidistante	0	1	0	0	1	1%
PMM	0	1	0	0	1	1%
Redes Bayesianas	1	1	0	0	1	1%

Random Forest: Utilizado en el 18% de los estudios, Random Forest empleó principalmente la variable de pagos (19 estudios) y la económica (12 estudios) en un total de 23 investigaciones. En el estudio de Xu et al. (2021), se utilizó el historial de pagos de los clientes ("Churn Bool") y datos económicos como el "Account Length" para analizar la propensión al impago, permitiendo manejar grandes volúmenes de datos y detectar patrones complejos en clientes de alto riesgo.

Regresión Logística: En el 15% de los estudios, la regresión logística se enfocó en variables demográficas (9 estudios) y pagos (16 estudios), con un total de 19 investigaciones. En estudios como el de Campos (2020), se incluyeron variables como la edad del cliente y su ubicación geográfica (por ejemplo, "State" y "Area Code"), que resultaron relevantes para modelar la probabilidad de pago en función de características como la edad y la región del cliente, agilizando así la segmentación por riesgo en cobranzas.

Árboles de Decisión: Esta técnica utilizó una combinación equilibrada de variables, siendo la edad de mora (8 estudios) y pagos (12 estudios) las más frecuentes, en un total de 18 estudios. En el estudio de Ortiz et al. (2022), se usaron variables como "dias_mora" y "canal" para segmentar a los clientes en función de su historial de mora y canal de contacto, permitiendo una predicción más precisa de los riesgos asociados a la antigüedad de la deuda.

Redes Neuronales: Con un 11% de los estudios, las redes neuronales fueron útiles para capturar interacciones no lineales, empleando principalmente variables de pagos (11 estudios) y económicas (7 estudios), acumulando un total de 15 estudios. Ortiz et al. (2022), utilizaron variables como "efectividad_intelibpo" y "valor2" para modelar comportamientos financieros no lineales y detectar patrones de pago en clientes con perfiles diversos. Estas variables permitieron

al modelo adaptarse a diferentes niveles de complejidad en el comportamiento financiero de los clientes.

XGBoost: Utilizado en el 9% de los estudios, XGBoost integró variables de pagos (11 estudios) y económicas (9 estudios), con un total de 12 estudios, siendo útil en la segmentación y predicción precisa en grandes bases de datos. Singh et al. (2024) y Fan et al. (2023), emplearon variables de historial de pago, como el "Contract month to month" y el "Tenure group", para analizar el impacto de los contratos y antigüedad en la propensión a la morosidad.

Las técnicas menos frecuentes, como Naive Bayes y Adaboost (3% y 2%, respectivamente), se usaron en contextos específicos donde la simplicidad y rapidez eran prioritarias, con un enfoque mayor en la variable de pagos y una menor integración de factores demográficos y económicos. Gomez et al. (2018) utilizaron Naive Bayes en estudios de pago básico, empleando "dias_mora" y "max_intentos" como indicadores rápidos de riesgo en el análisis de pago puntual en el sector de cobranzas.

Tabla 4

Frecuencia de Comportamiento de Variables según Idioma

Tipo	Antecedentes Investigativos				Total	%
	Pagos	Edad Mora	Demográficas	Económicas		
Inglés	33	14	19	22	47	94%
Español	2	1	3	1	3	6%

Como se muestra en la tabla 4, la investigación en machine learning aplicada a cobranzas se publicó mayormente en inglés, con un 94% de los estudios en este idioma (33 estudios para pagos y 22 para económicas), mientras que solo un 6% se realizó en español. En los estudios en inglés, investigaciones como las de Thomas et al. (2020) y Johnson y Patel (2018) destacaron

variables de historial de crédito como "Account Length" y "Churn Bool" para mejorar la precisión en la predicción de pagos. Por otro lado, en estudios en español como los de Lopez et al. (2020), se utilizaron variables demográficas como la edad y el canal de pago "canal" para analizar la relación entre características sociodemográficas y comportamiento de pago, especialmente relevante en mercados hispanohablantes.

Tabla 5

Frecuencia de Comportamiento de Variables según Tipo de Documento

Tipo	Antecedentes Investigativos				Total	%
	Pagos	Edad Mora	Demográficas	Económicas		
Artículo	31	12	19	21	44	88%
Trabajo de grado	3	2	3	2	4	8%
Acta de conferencia	1	1	0	0	2	4%

El tipo de documento también influyó en la selección de variables para los estudios de predicción de pago, a continuación, se presentan los resultados de la tabla 5:

Artículos (88%): Estos documentos se enfocaron principalmente en las variables de pagos (31 estudios) y económicas (21 estudios), con un total de 44 artículos. La variable de edad de mora también fue relevante (12 estudios), destacando el interés por analizar el historial y comportamiento temporal de los pagos. Por ejemplo, Fernandez y Liu (2021) analizaron "dias_mora" y el historial de cuenta de los clientes ("Account length") en función de su tiempo en mora.

Trabajos de Grado (8%): En este tipo de documentos, las variables demográficas (3 estudios) y pagos (3 estudios) fueron las más utilizadas, con un total de 4 trabajos de grado. Rodriguez et al. (2019) y Perez y Wang (2021) emplearon variables como la edad y canal para

segmentar clientes en riesgo, usando datos básicos para clasificar a los clientes por perfiles de morosidad.

Actas de Conferencias (4%): Las actas de conferencias emplearon principalmente la variable de Pagos (1 estudio), reflejando un análisis más general y menos detallado. En las conferencias de Gomez et al. (2018), por ejemplo, se utilizaron "max_intentos" como variable para medir la efectividad en el contacto inicial de clientes, proporcionando una visión rápida de patrones de riesgo en cobranza.

Tabla 6*Frecuencia de Comportamiento de Variables según País*

Tipo	Antecedentes Investigativos				Total	%
	Pagos	Edad Mora	Demográficas	Económicas		
China	9	2	5	6	10	20%
India	4	2	1	1	5	10%
Estados Unidos	2	0	4	3	4	8%
España	3	1	0	2	3	6%
Indonesia	1	0	1	1	3	6%
Colombia	3	2	4	1	3	6%
Perú	1	0	2	2	3	6%
Italia	2	0	1	1	2	4%
Korea	1	1	1	0	2	4%
Pakistán	2	0	1	0	2	4%
Arabia Saudita	1	0	1	0	1	2%
Bangladesh	1	0	2	1	1	2%
Canadá	1	0	0	1	1	2%
Dubái	1	0	0	0	1	2%
Estonia	1	1	0	1	1	2%
Irán	1	1	2	1	1	2%
Malaysia	1	1	0	1	1	2%
Sudáfrica	1	1	0	0	1	2%
Venezuela	1	0	1	0	1	2%
Francia	0	0	0	1	1	2%
Portugal	0	1	0	0	1	2%
Alemania	1	1	0	0	1	2%
Chile	0	1	0	1	1	2%

Según la tabla 6, se evidenció que la distribución geográfica de los estudios revela variaciones en el uso de variables, influenciadas por el contexto económico y social de cada país. A continuación, se detalla el uso de las variables más destacadas por país, integrando el nombre de las variables específicas y su relación con los modelos de machine learning en cobranzas.

China (20%): Los estudios realizados en China utilizaron variables de pagos (9 estudios) y económicas (6 estudios) en un total de 10 investigaciones. En estudios como el de Fan et al. (2023), se emplearon variables como "Account Length" y "Churn Bool" para segmentar clientes de alto riesgo. El uso de datos financieros y de pagos permitió al modelo detectar patrones complejos en clientes con historial de mora, optimizando la clasificación de riesgo en contextos de gran volumen de datos.

India (10%): En India, los estudios se concentraron en variables de pagos y edad de Mora (4 y 2 estudios respectivamente), en un total de 5 investigaciones. En investigaciones como las de Wagh et al. (2021), se destacaron variables como "dias_mora" y "contract duration" para segmentar clientes en función del tiempo de morosidad y la duración del contrato, permitiendo un análisis detallado del comportamiento de pago en mercados de crédito informal.

Estados Unidos (8%): Los estudios en Estados Unidos priorizaron variables demográficas (4 estudios) y económicas (3 estudios), en un total de 4 investigaciones. Singh et al. (2024), analizaron variables como "State" y "Area Code" para identificar patrones de morosidad y pagos en función de la ubicación geográfica de los clientes, facilitando la segmentación basada en perfiles sociodemográficos.

España, Colombia, Indonesia, Perú (6% cada uno): En estos países, los estudios se enfocaron en la variable de pagos y, en menor medida, en características demográficas y económicas. Por ejemplo, Ortiz et al. (2022), en Colombia emplearon variables como "canal" y

"dias_mora" para estudiar la efectividad del canal de cobro y el historial de mora. En España, estudios como el de Angarita y Marín (2019), también consideraron el "historial de pago" y la "ubicación" como factores clave en la predicción de riesgo de impago.

Italia, Corea, Pakistán (4% cada uno): Estos países presentaron estudios que emplearon variables pagos y demográficas para clasificar clientes según su perfil de riesgo. El autor Ortiz et al. (2022), utilizaron "dias_mora" y "max_intentos" para evaluar la antigüedad de la deuda y la intensidad de contacto en clientes en riesgo de mora. En Italia, variables como "contract type" y "client region" fueron usadas para ajustar los modelos a las características demográficas y de pagos específicas del país.

Otros países (2%): Países como Arabia Saudita, Bangladesh, Canadá, Dubai, Estonia, Irán, Malasia, Sudáfrica, Venezuela, Francia, Portugal, Alemania y Chile contribuyeron cada uno con estudios aislados (2%) en el análisis de pagos. Estas investigaciones, aunque menos frecuentes, incluyeron variables relevantes según el contexto local. En Bangladesh, por ejemplo, se destacó el uso de "client income" y "credit type" como variables económicas, y en Francia, la variable de "contract length" fue utilizada para ajustar el análisis de riesgo en clientes con diferentes tipos de contratos.

Esta distribución muestra cómo las variables cambian su relevancia según el país del estudio, adaptándose a contextos específicos en la gestión de cobranzas y destacando las preferencias de variables de pagos, económicas, y demográficas en función de los contextos locales.

Modelos de Machine Learning con Mejor Desempeño

En este capítulo, se presentan los resultados del análisis de los artículos sobre el uso de modelos de Machine Learning en la predicción de pagos en el sector de cobranza. Se seleccionaron los estudios más relevantes y se destacaron las métricas empleadas para evaluar el desempeño de los modelos. Posteriormente, se realizó un análisis comparativo de los resultados. A continuación, se resumen los hallazgos clave, enfatizando los modelos con mejor desempeño en la gestión de cobranzas.

El artículo "Pronóstico del cumplimiento de pago de clientes usando machine learning" analiza varios modelos de aprendizaje automático aplicados a la gestión de cobranza en la empresa Créditos Orbe (Campos, 2020). A continuación, se describen los principales hallazgos respecto a los modelos evaluados:

Figura 1

Resumen Métricas por Modelos Estudio

Modelo	Exactitud	Precisión	Sensibilidad	f1_score	Tiempo
KNN	0,656830	0,653156	0,744075	0,674214	1176,17
Adaboost	0,786714	0,774698	0,824090	0,793826	58,99
Regresión Logística	0,690192	0,678841	0,802946	0,712254	11,72
Random Forest**	0,791132	0,789464	0,809033	0,795065	76,13
Bagged Decision Tree	0,813437	0,784709	0,863492	0,821913	911,43
Adaboost	0,786714	0,774698	0,824090	0,793826	58,99
Voting Ensemble	0,761383	0,766824	0,747804	0,753182	104,84

Nota. Tomado de Pronóstico del cumplimiento de pago de clientes usando Machine Learning (Campos, 2020).

Como se evidencia en la Figura 1, el modelo K-NN presentó una exactitud de 66% y un F1 Score de 67%, con un tiempo de ejecución de 1176.17 segundos. Su precisión y sensibilidad reflejan un rendimiento limitado en la predicción de clientes cumplidores, lo cual indica que este modelo no es óptimo para el problema debido a su alto costo computacional y bajo desempeño predictivo (Campos, 2020).

El modelo Adaboost alcanzó una exactitud de 79% y una sensibilidad de 82%. Su F1 Score fue de 79%, con un tiempo de ejecución de 58.99 segundos. Este modelo demuestra un rendimiento adecuado y un tiempo de procesamiento eficiente, haciéndolo una opción viable para tareas de predicción rápida. Sin embargo, fue superado por otros modelos de ensamblado en términos de exactitud y F1 Score (Campos, 2020).

El modelo de regresión logística obtuvo una exactitud de 69%, precisión de 67% y sensibilidad de 80%, con un tiempo de ejecución bajo de 11.72 segundos. Este rendimiento es aceptable en implementaciones donde la rapidez es fundamental. No obstante, este modelo fue superado en precisión y sensibilidad por métodos de ensamblado más robustos, lo que limita su aplicabilidad en contextos que requieren alta exactitud (Campos, 2020).

El modelo Random Forest logró una exactitud de 79%, precisión de 79%, sensibilidad de 81% y un F1 Score de 80%, con un tiempo de ejecución de 76.13 segundos. Este modelo destaca por su buen equilibrio entre precisión y tiempo, proporcionando resultados confiables para la predicción del cumplimiento de pago. Su robustez y rendimiento hacen de Random Forest una opción recomendada en escenarios con múltiples variables de entrada (Campos, 2020).

El modelo Bagged Decision Tree mostró el mejor rendimiento general, con una exactitud de 81%, precisión de 79%, sensibilidad de 86% y un F1 Score de 82%. Sin embargo, su tiempo de ejecución fue considerablemente alto (911.43 segundos), lo que podría limitar su aplicabilidad

en implementaciones en tiempo real. A pesar de ello, este modelo es ideal para aplicaciones que priorizan la precisión sobre la velocidad (Campos, 2020).

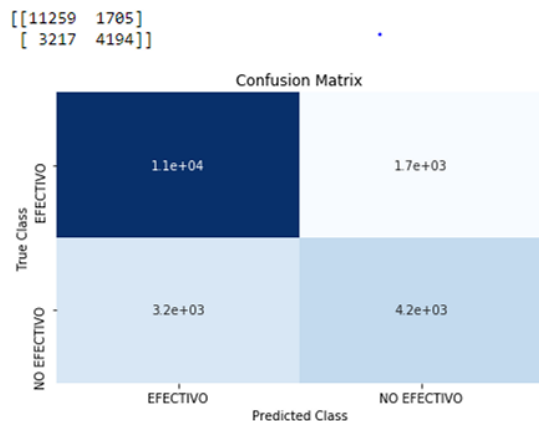
Por último, el modelo Voting Ensemble, que combina predicciones de varios clasificadores, alcanzó una exactitud de 76% y una precisión de 77%, con una sensibilidad de 75% y un F1 Score de 75%. Su tiempo de ejecución fue de 104.84 segundos. Aunque presenta un rendimiento estable, fue superado en exactitud y F1 Score por los modelos Bagged Decision Tree y Random Forest (Campos, 2020).

Campos (2020), concluye que el Bagged Decision Tree es el modelo más eficaz para predecir el cumplimiento de pago de clientes en créditos orbe, debido a su alta exactitud del 81% y F1 Score de 82%. Sin embargo, el tiempo de ejecución de este modelo sugiere que puede no ser ideal para entornos donde se requiere actualización rápida. En este contexto, el Bosque Aleatorio ofrece un excelente balance entre precisión y velocidad, haciéndolo una alternativa recomendable para aplicaciones en tiempo real. Este análisis subraya la importancia de seleccionar modelos que no solo maximicen la exactitud, sino que también sean eficientes en términos de procesamiento, especialmente en aplicaciones prácticas de cobranza (Campos, 2020).

El artículo "Predicción de clientes efectivos en la gestión de carteras de cobranza castigada en la empresa InteliBPO S.A.S a través de modelos de aprendizaje automático" evalúa múltiples modelos de Machine Learning aplicados al análisis de la efectividad en la recuperación de cartera (Ortiz et al., 2022). La investigación compara el rendimiento de diferentes modelos y su capacidad para clasificar a los clientes según su probabilidad de ser efectivos en el proceso de cobranza.

Figura 2

Matriz de Confusión del Modelo K-NN

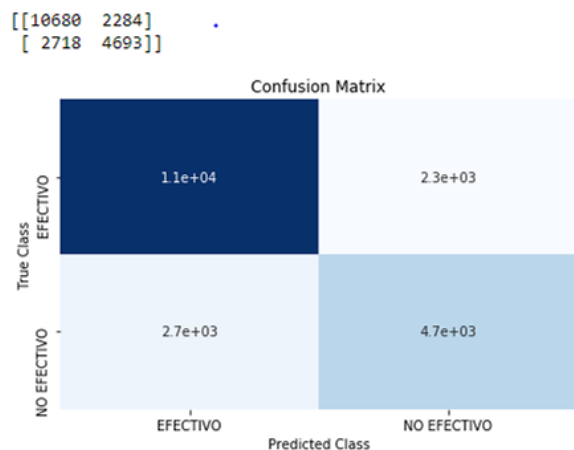


Nota. Tomado de predicción de clientes efectivos en la gestión de carteras de cobranza castigada en la empresa InteliBPO S.A.S a través de modelos de aprendizaje automático (Ortiz et al., 2022).

Según la Figura 2, en la matriz de confusión del modelo K-Vecinos Más Cercanos (K-NN) se destacó por su capacidad para identificar clientes efectivos, logrando una sensibilidad del 86.85% y una precisión general del 77.78%, con un F1 Score de 75.14%. Sin embargo, su especificidad del 56.59% evidencia limitaciones al clasificar adecuadamente a clientes no efectivos (Ortiz et al., 2022).

Figura 3

Matriz de Confusión del Modelo de Máquina de Soporte Vectorial

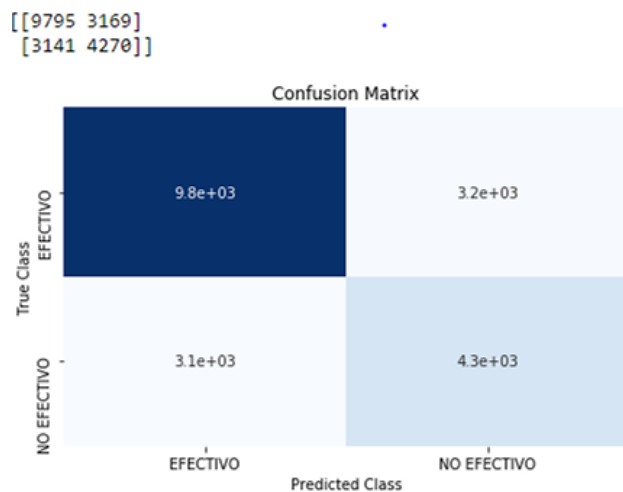


Nota. Tomado de predicción de clientes efectivos en la gestión de carteras de cobranza castigada en la empresa InteliBPO S.A.S a través de modelos de aprendizaje automático (Ortiz et al., 2022).

Según la Figura 3, en la matriz de confusión del modelo Máquinas de Soporte Vectorial (SVM), que optimiza la clasificación a través de hiperplanos, mostró una precisión general del 79.71% y una sensibilidad del 82.38%. Su especificidad de 63.32% sugiere que SVM es ligeramente superior a K-NN en la clasificación de clientes no efectivos, con un F1 Score de 75.28%. Este modelo constituye una alternativa robusta, aunque la mejora en especificidad es moderada (Ortiz et al., 2022).

Figura 4

Matriz de Confusión del Modelo de Árbol de Decisión

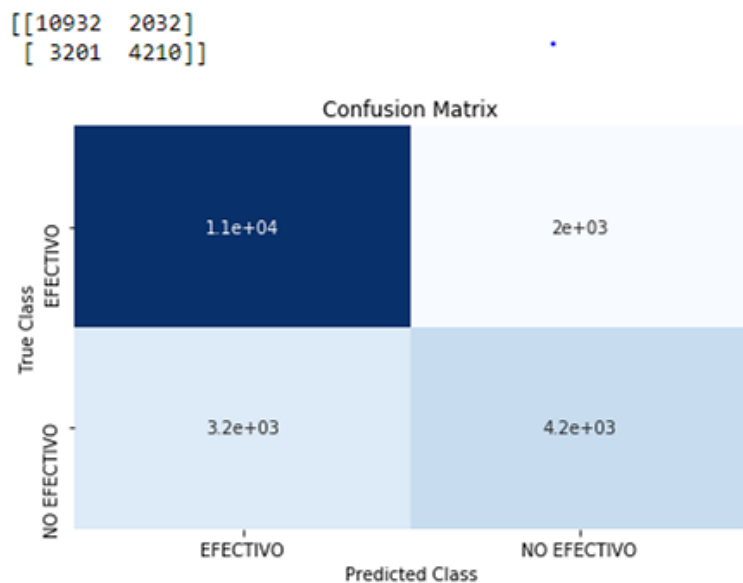


Nota. Tomado de predicción de clientes efectivos en la gestión de carteras de cobranza castigada en la empresa InteliBPO S.A.S a través de modelos de aprendizaje automático (Ortiz et al., 2022).

Según la Figura 4, en la matriz de confusión del modelo de Árbol de Decisión, conocido por su capacidad interpretativa, obtuvo una exactitud de 69.03% y una sensibilidad del 75.56%, con un F1 Score de 69.04%. A pesar de su rendimiento aceptable, el Árbol de Decisión fue limitado en su capacidad para identificar clientes no efectivos, con una especificidad del 57.40% (Ortiz et al., 2022).

Figura 5

Matriz de Confusión del Modelo de Bosques Aleatorios



Nota. Tomado de Predicción de clientes efectivos en la gestión de carteras de cobranza castigada en la empresa InteliBPO S.A.S a través de modelos de aprendizaje automático (Ortiz et al., 2022).

Según la Figura 5, en la matriz de confusión del modelo de Bosques Aleatorios, que utiliza múltiples árboles para mejorar la precisión, mostró un desempeño elevado con una precisión del 77.35% y una sensibilidad del 84.33%. Su F1 Score de 73.77% refuerza su utilidad en la clasificación de clientes efectivos, aunque su especificidad de 56.81% señala una dificultad similar a la de K-NN para clasificar clientes no efectivos (Ortiz et al., 2022).

Ortiz et al. (2022), concluye que el modelo K-Vecinos Más Cercanos (K-NN) es el más adecuado para predecir clientes efectivos en la gestión de cobranza de cartera castigada, debido a su alto rendimiento en la identificación de clientes efectivos y un balance razonable entre precisión y sensibilidad. A pesar de la limitada especificidad de K-NN, el autor sugiere su uso en

InteliBPO S.A.S como una herramienta de apoyo para optimizar estrategias de cobranza y mejorar la focalización en clientes con mayor probabilidad de efectividad. No obstante, se recomienda explorar técnicas de balanceo de datos o ajustes adicionales en el modelo para mejorar la especificidad y, por ende, la precisión en la identificación de clientes no efectivos.

Del análisis del artículo "Customer churn prediction in telecom sector using machine learning techniques" de Wagh et al. (2021), se destacan los siguientes puntos relevantes en cuanto a modelos de Machine Learning aplicados para la predicción de deserción de clientes en el sector de telecomunicaciones. A continuación, se detalla la evaluación de los modelos y los hallazgos clave:

Figura 6

Resultados de Árbol de Decisión Modelo de Clasificación Antes

	Precision	Recall	F1-score	Support
Class 0 (No)	0.83	0.89	0.86	1046
Class 1(Yes)	0.61	0.49	0.54	361
Weighted avg	0.77	0.78	0.77	1407

	No. of Instances	Percentage
Correctly Classified Instances	5493	77.9923%
Incorrectly Classified Instances	1550	22.0077%

Nota. Resultados de Árbol de Decisión Modelo de Clasificación antes de Aplicar el Remuestreo y ENN. Tomado de Customer churn prediction in telecom sector using machine learning techniques (Wagh et al., 2021).

Figura 7*Resultados de Árbol de Decisión Modelo de Clasificación Después*

	Precision	Recall	F1-score
Class 0 (No)	0.93	0.92	0.93
Class 1(Yes)	0.94	0.94	0.94
Weighted avg	0.93	0.93	0.93
	No. of Instances	Percentage	
Correctly Classified Instances	6615	93.85%	
Incorrectly Classified Instances	427	6.15 %	

Nota. Resultados de Árbol de Decisión Modelo de Clasificación Después de Aplicar el Remuestreo y ENN. Tomado de Customer churn prediction in telecom sector using machine learning techniques (Wagh et al., 2021).

Figura 8*Resultados del Modelo Clasificador Random Forest Antes*

	Precision	Recall	F1-score
Class 0 (No)	0.99	1.00	0.99
Class 1(Yes)	0.99	0.96	0.97
Weighted avg	0.99	0.99	0.99
	No. of Instances	Percentage	
Correctly Classified Instances	6971	98.91%	
Incorrectly Classified Instances	77	1.09%	

Nota. Resultados del Modelo Clasificador Random Forest antes de Aplicar el Remuestreo y ENN. Tomado de Customer churn prediction in telecom sector using machine learning techniques (Wagh et al., 2021).

Figura 9

Resultados del Modelo Clasificador Random Forest Después

	Precision	Recall	F1-score
Class 0 (No)	0.98	1.00	0.99
Class 1(Yes)	1.00	0.98	0.99
Weighted avg	0.99	0.99	0.99

	No. of Instances	Percentage
Correctly Classified Instances	7010	99.09%
Incorrectly Classified Instances	38	0.91 %

Nota. Resultados del Modelo Clasificador Random Forest Después de Aplicar el Remuestreo y ENN. Tomado de Customer churn prediction in telecom sector using machine learning techniques (Wagh et al., 2021).

Árboles de Decisión: Este alcanzó una precisión del 78% en su capacidad de predicción de clientes con alta probabilidad de abandonar la suscripción. Después de aplicar técnicas de balanceo de datos (up-sampling y ENN), la precisión mejoró significativamente hasta el 94% (Wagh et al., 2021).

Bosques Aleatorios: Antes de aplicar up-sampling y ENN, el modelo obtuvo una precisión de 99%. Con el balanceo de datos, la precisión del modelo de bosques aleatorios alcanzó un impresionante 99%. Además, el modelo obtuvo una precisión y un recall del 99%, mostrando una excelente capacidad para manejar datos de alta dimensionalidad y variabilidad (Wagh et al., 2021).

Wagh et al. (2021), concluyen que el modelo de Bosque Aleatorio es el mejor clasificador para la predicción de deserción de clientes en telecomunicaciones, ya que logra una precisión superior y una alta capacidad para identificar tanto a clientes que desertan como a los que permanecen. Además, los resultados indican que un enfoque basado en modelos de clasificación avanzados y técnicas de balanceo de datos permite a las empresas de

telecomunicaciones optimizar sus estrategias de retención, enfocándose en características clave como el tipo de contrato y los cargos mensuales.

En el artículo "Classification of m-payment users' behavior using machine learning models" de Aslam et al. (2022), se comparan los modelos de Regresión Logística y Máquinas de Soporte Vectorial (SVM) para clasificar el comportamiento de los usuarios de pagos móviles. A continuación, se presenta un resumen de los hallazgos específicos:

Figura 10

Comparación de Ambos Modelos

Measures	SVM	Logistic regression
Specificity (TNR)	0.7188	0.4688
Sensitivity	0.9896	0.9896
Precision	0.9135	0.8482
Negative predictive value	0.9583	0.9375
False positive rate	0.2813	0.5313
False negative rate	0.0104	0.0104
Accuracy	0.9219	0.8594
F1	0.9500	0.9135

Nota. Tomado de Classification of m-payment users' behavior using machine learning models (Aslam et al., 2022).

Regresión Logística (LR): El modelo de regresión logística, ampliamente utilizado en estudios de comportamiento del consumidor, se aplicó con el objetivo de estimar la probabilidad de que un individuo sea usuario de m-payment. Este modelo demostró una precisión general del 85%, con una sensibilidad del 99% y una especificidad del 47%. A pesar de su alta sensibilidad, la especificidad baja indica que el modelo tiene limitaciones para diferenciar con precisión entre usuarios y no usuarios de m-payment (Aslam et al., 2022). La baja especificidad puede indicar una tendencia del modelo LR a sobreestimar la probabilidad de ser usuario, lo cual sugiere que

este enfoque es menos robusto para clasificaciones precisas cuando se necesita mayor discriminación entre grupos.

Máquinas de Vectores de Soporte (SVM): Este modelo emplea una función kernel radial, adecuada para datos no lineales, lo que mejora su capacidad de clasificación en este contexto. La SVM alcanzó una precisión del 91%, con una sensibilidad del 99% y una especificidad del 72%. La alta especificidad obtenida por el modelo SVM indica que es significativamente mejor para identificar tanto a usuarios como a no usuarios en comparación con la regresión logística. Además, este modelo mostró un valor predictivo negativo (NPV) del 96%, lo que reafirma su solidez al minimizar los falsos positivos y mantener una alta precisión en la clasificación (Aslam et al., 2022). Estos resultados destacan que el modelo SVM es un clasificador más adecuado en este tipo de estudio, donde una clasificación precisa entre usuarios y no usuarios es esencial para una estrategia de marketing efectiva.

Aslam et al. (2022), concluyen que el modelo SVM supera a la Regresión Logística en la clasificación de usuarios de m-payment, especialmente debido a su mayor especificidad y precisión general. Esta superioridad sugiere que SVM es más adecuado para estudios de clasificación en el comportamiento del consumidor, ya que permite a los analistas una mejor diferenciación entre usuarios y no usuarios, aspecto crucial en estrategias de segmentación de mercado. Además, los hallazgos sobre el valor percibido como principal predictor sugieren que los proveedores de servicios de m-payment deberían enfocar sus estrategias de marketing en comunicar beneficios tangibles y una propuesta de valor clara para incentivar la adopción de su tecnología.

En el artículo "Assessing the Loss Given Default of Bank Loans Using the Hybrid Algorithms Multi-Stage Model" de Fan et al. (2023), se analizaron los resultados obtenidos en

cada etapa del modelo de Algoritmos Híbridos Multi-Etapa (HMS) y se evaluaron en comparación con modelos tradicionales. A continuación, se presentan los hallazgos en términos de rendimiento predictivo, segmentación en múltiples etapas y métricas cuantitativas obtenidas para los modelos seleccionados.

Primera Etapa: Predicción de la Ocurrencia de Pérdida: En esta fase, el objetivo era clasificar si ocurriera alguna pérdida ($LGD > 0$ frente a $LGD = 0$). El modelo LightGBM en combinación con el agrupamiento K-means mostró el mejor desempeño entre los modelos probados. El rendimiento de este modelo se midió principalmente a través del Área Bajo la Curva (AUC), alcanzando un valor de 0.6773. Este valor destaca la capacidad de LightGBM para discriminar entre casos de pérdida y no pérdida con una precisión superior a la de otros modelos en esta fase (Fan et al., 2023). La incorporación de técnicas de agrupamiento, como K-means, contribuyó significativamente a mejorar la segmentación de los datos y, en consecuencia, la precisión del modelo, abordando la heterogeneidad de las características de riesgo.

Figura 11*Desempeño de la Clasificación en la Primera Etapa*

Clustering		Model	ACC	Precision	Recall	F1	AUC
Individual Models	no clustering	LR	0.6025	0.6355	0.2556	0.3645	0.5731
		DT	0.5958	0.5573	0.4570	0.5022	0.5947
		RF	0.5999	0.5605	0.4775	0.5157	0.6180
		GBDT	0.6295	0.6269	0.4186	0.5020	0.6694
		XGBoost	0.6274	0.6157	0.4387	0.5123	0.6697
		MLP	0.6265	0.6120	0.4449	0.5152	0.6641
		LightGBM	0.6298	0.6201	0.4389	0.5140	0.6718
K-means	clustering 1	LR	0.6052	0.6348	0.284	0.3925	0.6077
		DT	0.5968	0.5609	0.4691	0.5110	0.5936
		RF	0.6003	0.5623	0.4948	0.5264	0.6193
		GBDT	0.6331	0.6316	0.4389	0.5179	0.6743
		XGBoost	0.6338	0.6275	0.4539	0.5268	0.6744
		MLP	0.6304	0.6224	0.4493	0.5219	0.6699
		LightGBM	0.6341	0.6265	0.4585	0.5295	0.6773
	clustering 2	LR	0.6404	0.6386	0.1045	0.1797	0.6161
		DT	0.5788	0.4375	0.4142	0.4255	0.5580
		RF	0.5958	0.4561	0.3787	0.4138	0.5913
		GBDT	0.6263	0.5137	0.1479	0.2297	0.6157
		XGBoost	0.5854	0.4358	0.3412	0.3827	0.5902
		MLP	0.6397	0.641	0.0986	0.1709	0.6199
		LightGBM	0.6018	0.4589	0.3195	0.3767	0.6055
SOM	clustering 1	LR	0.6046	0.6403	0.2711	0.381	0.5982
		DT	0.5931	0.5563	0.4605	0.5039	0.5854
		RF	0.5989	0.5619	0.4816	0.5187	0.6135
		GBDT	0.6296	0.6233	0.4416	0.5169	0.6691
		XGBoost	0.6274	0.6153	0.4528	0.5217	0.6699
		MLP	0.6254	0.6065	0.4701	0.5297	0.6633
		LightGBM	0.6313	0.6228	0.4522	0.524	0.6718
	clustering 2	LR	0.6294	0.5714	0.1265	0.2072	0.5911
		DT	0.5866	0.4571	0.4282	0.4422	0.5657
		RF	0.5866	0.4545	0.4015	0.4264	0.5828
		GBDT	0.6145	0.4886	0.1557	0.2362	0.5967
		XGBoost	0.5931	0.4558	0.326	0.3801	0.5917
		MLP	0.6238	0.5321	0.1411	0.2231	0.5895
		LightGBM	0.6071	0.4804	0.3285	0.3902	0.5998

Nota. Tomado de Assessing the Loss Given Default of Bank Loans Using the Hybrid Algorithms Multi-Stage Model (Fan et al., 2023).

Segunda Etapa: Predicción de la Pérdida Total: La segunda etapa se enfocó en predecir si la pérdida total ($LGD = 1$) o parcial ($0 < LGD < 1$). Dada la alta desproporción en las clases, se implementó el Sobremuestreo Aleatorio (ROS) para balancear las muestras. El modelo Random Forest (RF), combinado con K-means, resultó ser el más efectivo, logrando un AUC de 0.9588 y una precisión de 0.8794 en esta etapa (Fan et al., 2023). Este modelo demostró ser altamente

confiable para predecir la pérdida total, optimizando el rendimiento incluso en condiciones de desequilibrio de clases, lo cual es fundamental en contextos financieros donde las pérdidas totales presentan una menor frecuencia.

Figura 12

Desempeño de la Clasificación en la Segunda Etapa

Clustering		Model	ACC	Precision	Recall	F1	AUC
ROS + Individual Models	no clustering	LR	0.6597	0.6059	0.9304	0.7339	0.7105
		DT	0.8240	0.7791	0.9085	0.8388	0.8803
		RF	0.8233	0.7753	0.9147	0.8393	0.8934
		GBDT	0.6998	0.6565	0.8485	0.7403	0.7662
		XGBoost	0.7568	0.7147	0.8615	0.7813	0.8342
		MLP	0.6845	0.6488	0.8159	0.7228	0.7410
		LightGBM	0.7257	0.6795	0.8631	0.7603	0.7978
ROS + K-means	clustering 1	LR	0.6556	0.6052	0.9033	0.7248	0.7138
		DT	0.8248	0.7761	0.9148	0.8398	0.8807
		RF	0.8215	0.7704	0.918	0.8377	0.8903
		GBDT	0.6999	0.6541	0.8537	0.7407	0.7662
		XGBoost	0.7582	0.7130	0.8675	0.7827	0.8350
		MLP	0.6800	0.6392	0.8319	0.7230	0.7374
		LightGBM	0.7265	0.6770	0.8704	0.7617	0.7951
ROS + SOM	clustering 2	LR	0.6156	0.5908	0.8908	0.7104	0.6414
		DT	0.8786	0.8340	0.9620	0.8935	0.9061
		RF	0.8794	0.8315	0.9684	0.8947	0.9588
		GBDT	0.7429	0.7113	0.8655	0.7809	0.8188
		XGBoost	0.8333	0.7906	0.932	0.8555	0.9172
		MLP	0.6993	0.6775	0.8244	0.7438	0.7656
		LightGBM	0.8057	0.7732	0.8956	0.8299	0.8940
ROS + SOM	clustering 1	LR	0.6521	0.602	0.9074	0.7238	0.6875
		DT	0.8773	0.8371	0.9383	0.8848	0.9055
		RF	0.8847	0.8361	0.9584	0.8931	0.9575
		GBDT	0.7323	0.7033	0.8081	0.7520	0.8165
		XGBoost	0.8429	0.8168	0.8859	0.8500	0.9228
		MLP	0.6871	0.6556	0.7946	0.7184	0.7538
		LightGBM	0.8139	0.7877	0.8617	0.8231	0.8941
ROS + SOM	clustering 2	LR	0.6628	0.6114	0.9001	0.7282	0.7134
		DT	0.8190	0.7731	0.9048	0.8338	0.8773
		RF	0.8160	0.7668	0.9102	0.8324	0.8834
		GBDT	0.6970	0.6522	0.8490	0.7377	0.7638
		XGBoost	0.7542	0.7100	0.8625	0.7789	0.8311
		MLP	0.6808	0.6337	0.8625	0.7306	0.7374
		LightGBM	0.7244	0.6772	0.8614	0.7583	0.7942

Nota. Tomado de Assessing the Loss Given Default of Bank Loans Using the Hybrid Algorithms Multi-Stage Model (Fan et al., 2023).

Tercera Etapa: Predicción Continua de LGD: La tercera etapa del análisis utiliza regresión de Mínimos Cuadrados Ordinarios (OLS) para modelar la pérdida continua en casos de

LGD parcial. El rendimiento de esta etapa se evaluó a través de varias métricas de error y ajuste, y la Ruta 4 del modelo HMS, que combina LightGBM y RF en las dos primeras etapas junto con OLS en la tercera, obtuvo los resultados más destacados. Este modelo logró un MSE de 0.0968, RMSE de 0.3111, MAE de 0.1838, R-cuadrado (R^2) de 0.4563 y una Varianza Explicada (EV) de 0.6106 (Fan et al., 2023). Estos valores indican que la Ruta 4 es la configuración óptima para la predicción continua de LGD, mostrando menor error y mayor capacidad de explicación en comparación con otros enfoques.

Figura 13

Desempeño de Predicción de LGD en la Tercera Etapa

The Best Classification Model								
Route	First stage	Second stage	Third stage	MSE	RMSE	MAE	R^2	EV
Route 1	LightGBM	ROS + RF		0.0968	0.3111	0.1838	0.4563	0.6106
Route 2	K-means (clustering 1 + LightGBM) and (clustering 2 + MLP)	ROS+ RF		0.0968	0.3111	0.1837	0.4566	0.6106
	SOM clustering + LightGBM	ROS+ RF		0.0978	0.3128	0.1849	0.4508	0.6075
Route 3	LightGBM	ROS + K-means clustering + RF		0.0969	0.3112	0.1838	0.4562	0.6104
	LightGBM	ROS + SOM clustering + RF	OLS	0.0967	0.3109	0.1836	0.4573	0.6115
Route 4	K-means (clustering 1 + LightGBM) and (clustering 2 + MLP)	ROS + K-means clustering + RF		0.0968	0.3111	0.1837	0.4566	0.6104
	K-means (clustering 1 + LightGBM) and (clustering 2 + MLP)	ROS + SOM clustering + RF		0.0966	0.3108	0.1835	0.4577	0.6115
	SOM clustering + LightGBM	ROS + K-means clustering + RF		0.0979	0.3128	0.1849	0.4506	0.6073
	SOM clustering + LightGBM	ROS + SOM clustering + RF		0.0977	0.3125	0.1847	0.4518	0.6084

Nota. Tomado de Assessing the Loss Given Default of Bank Loans Using the Hybrid Algorithms Multi-Stage Model (Fan et al., 2023).

Fan et al. (2023), concluyen que el modelo de Algoritmos Híbridos Multi-Etapa (HMS) presenta un rendimiento superior al de los modelos tradicionales de predicción de LGD, como el

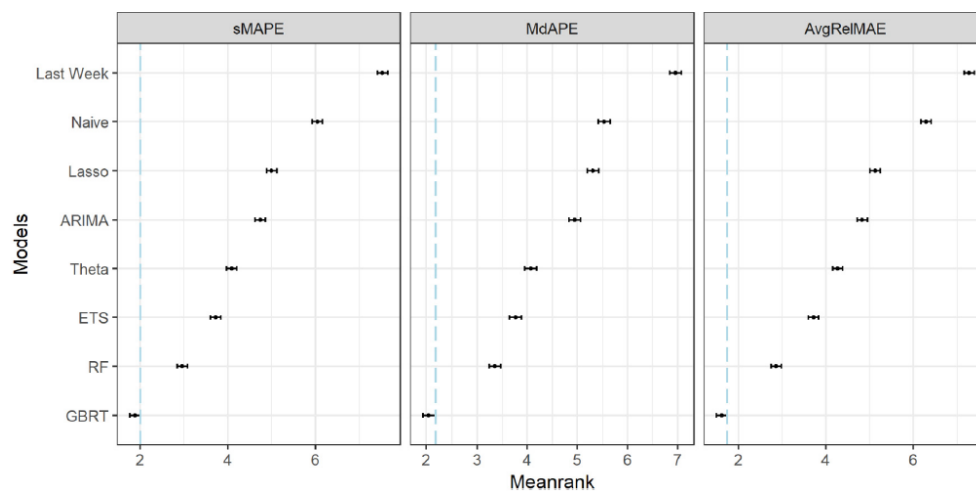
modelo de regresión LossCalc de Moody's. Su investigación evidenció que la estructura de múltiples etapas del modelo HMS permite capturar mejor la complejidad de las pérdidas crediticias al predecir la ocurrencia de pérdida, la pérdida total y la pérdida parcial continua en fases separadas. En particular, el uso de una combinación de LightGBM y Random Forest en las primeras etapas, junto con Regresión OLS para la predicción continua en la etapa final, mostró mejoras significativas en la precisión de las predicciones, alcanzando un AUC de 0.6773 en la primera etapa y de 0.9588 en la segunda etapa, y logrando una Varianza Explicada (EV) de 0.6106 en la etapa de regresión continua.

En el artículo "Forecasting third-party mobile payments with implications for customer flow prediction" de Ma y Fildes (2020), se evaluaron diversos modelos para pronosticar el flujo de clientes utilizando datos de pagos móviles de terceros, aplicando cuatro métricas clave de rendimiento: sMAPE (Error Porcentual Absoluto Medio Simétrico), MdAPE (Mediana del Error Porcentual Absoluto Medio), AvgRelMAE (Error Absoluto Medio Relativo Promedio) y MPE (Error Porcentual Medio). A continuación, se presenta un análisis detallado de cada modelo evaluado y sus métricas de desempeño:

Figura 14*Pronósticos de un Día Adelante*

	Test set 1				Test set 2				Test sets 1&2			
	sMAPE	MdAPE	AvgRel MAE	MPE	sMAPE	MdAPE	AvgRel MAE	MPE	sMAPE	MdAPE	AvgRel MAE	MPE
Last week	0.130	0.082	1.000	6.330	0.150	0.103	1.000	-0.781	0.138	0.091	1.000	-0.973
Naive	0.120	0.080	0.943	4.322	0.112	0.077	0.776	0.062	0.115	0.079	0.833	0.033
ETS	0.103	0.066	0.789	2.309	0.099	0.069	0.674	-2.734	0.099	0.067	0.713	-2.360
Theta	0.106	0.069	0.831	5.093	0.100	0.069	0.689	-0.838	0.102	0.068	0.741	-0.186
ARIMA	0.105	0.071	0.819	2.324	0.108	0.078	0.731	-2.788	0.105	0.074	0.758	-2.457
Lasso	0.114	0.076	0.865	0.678	0.109	0.078	0.726	-0.047	0.110	0.077	0.777	0.083
RF	0.104	0.064	0.745	-5.463	0.096	0.067	0.637	-5.684	0.097	0.065	0.674	-5.326
GBRT	0.097	0.057	0.676	-4.314	0.088	0.061	0.583	-3.826	0.090	0.058	0.614	-3.713

Nota. Tomado de Forecasting third-party mobile payments with implications for customer flow prediction (Ma & Fildes, 2020).

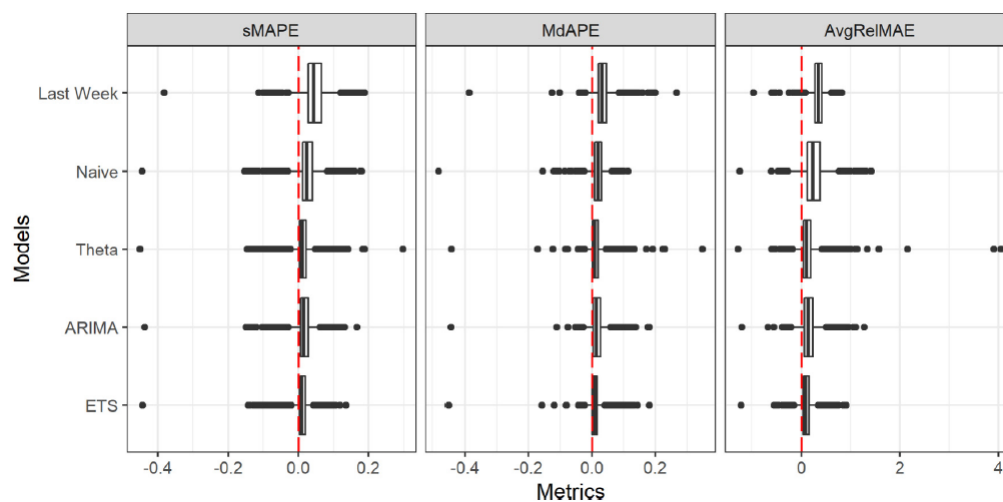
Figura 15*Resultados del Test de Nemenyi al 5%*

Nota. Resultados del test de Nemenyi al 5% nivel de significancia para pronósticos de un paso adelante a nivel de tienda durante todo el período de prueba. La distancia crítica para el test de Nemenyi es 0.234. Tomado de Forecasting third-party mobile payments with implications for customer flow prediction (Ma & Fildes, 2020).

Last Week: Este modelo emplea el valor observado para el mismo día de la semana en la última semana como pronóstico. Actúa como referencia en el cálculo del AvgRelMAE, con un valor base de 1.000. Como se observa en la Figura 14, Last Week tiene un error considerable en sMAPE (0.130 en el Conjunto 1) y en AvgRelMAE (1.000 en ambos conjuntos), lo cual confirma su simplicidad. Aunque es útil como punto de comparación, la alta variabilidad en los datos limita su precisión en escenarios complejos, lo que también se refleja en la Figura 15 mediante su posición en el extremo superior de las métricas sMAPE, MdAPE y AvgRelMAE en el test de Nemenyi.

Figura 16

Diferencias de Precisión en los Pronósticos



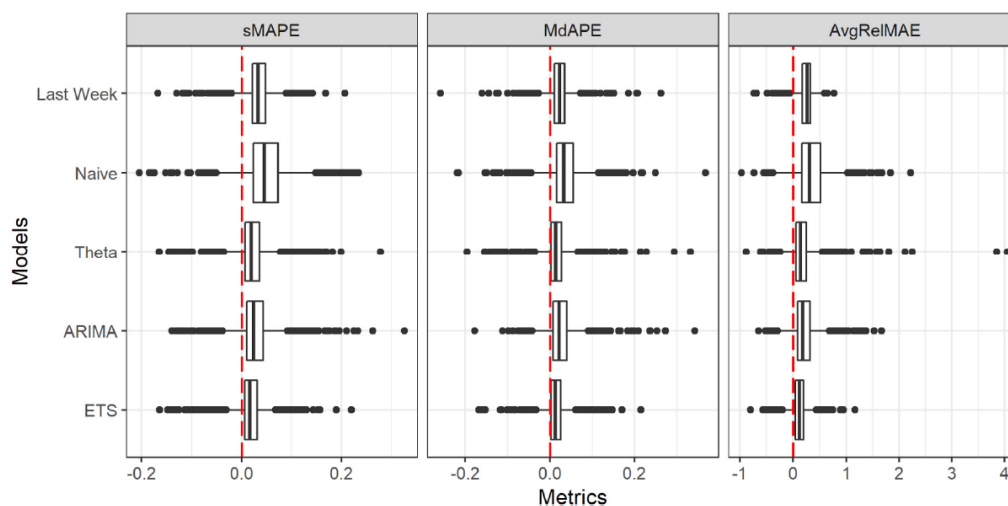
Nota. Diferencias de precisión en los pronósticos de un paso adelante a nivel de tienda entre GBRT y cinco modelos de series temporales durante todo el período de prueba. Tomado de Forecasting third-party mobile payments with implications for customer flow prediction (Ma & Fildes, 2020).

Naive: Basado en el valor observado del último día disponible, este modelo no considera patrones de estacionalidad. Mostró un rendimiento de sMAPE: 0.120, MdAPE: 0.080 y AvgRelMAE: 0.943 en el Conjunto 1. La simplicidad de Naive se refleja en un rendimiento limitado, especialmente en sMAPE y AvgRelMAE, donde no captura la estacionalidad ni la variabilidad inherente en los datos. En la Figura 15, se observa que Naive, aunque mejor que Last Week, sigue estando estadísticamente detrás de modelos avanzados como GBRT, con intervalos de rendimiento más amplios.

ETS (Exponential Smoothing): Es un modelo de suavizamiento exponencial automatizado que ajusta estacionalidades y tendencias. Las métricas obtenidas son sMAPE: 0.103, MdAPE: 0.066 y AvgRelMAE: 0.789 en el Conjunto 1. Aunque ETS logra captar patrones de estacionalidad en series de tiempo, su precisión es menor en comparación con modelos más avanzados como GBRT. La Figura 16 refuerza esta observación, mostrando que la variabilidad en los resultados de ETS es mayor en comparación con GBRT, con un rango intercuartílico más amplio en los boxplots, indicando menos consistencia.

Figura 17

Diagramas de Caja sobre las Diferencias de Precisión en los Pronósticos



Nota. Diagramas de caja sobre las diferencias de precisión en los pronósticos de 1 a 14 días adelante a nivel de tienda entre GBRT-pMIMO y cinco modelos de series temporales durante todo el período de prueba. Tomado de Forecasting third-party mobile payments with implications for customer flow prediction (Ma & Fildes, 2020).

Theta: La técnica Theta descompone la serie temporal para capturar tendencias y estacionalidad. Las métricas obtenidas son sMAPE: 0.106, MdAPE: 0.069 y AvgRelMAE: 0.831 en el Conjunto 1. Theta es eficaz en situaciones de baja variabilidad, pero su precisión disminuye con datos complejos y múltiples ciclos estacionales. En la Figura 17, se muestra que Theta tiene un desempeño moderado en comparación con GBRT, ya que su dispersión en sMAPE y MdAPE es mayor, lo que refleja su dificultad para adaptarse a la alta complejidad.

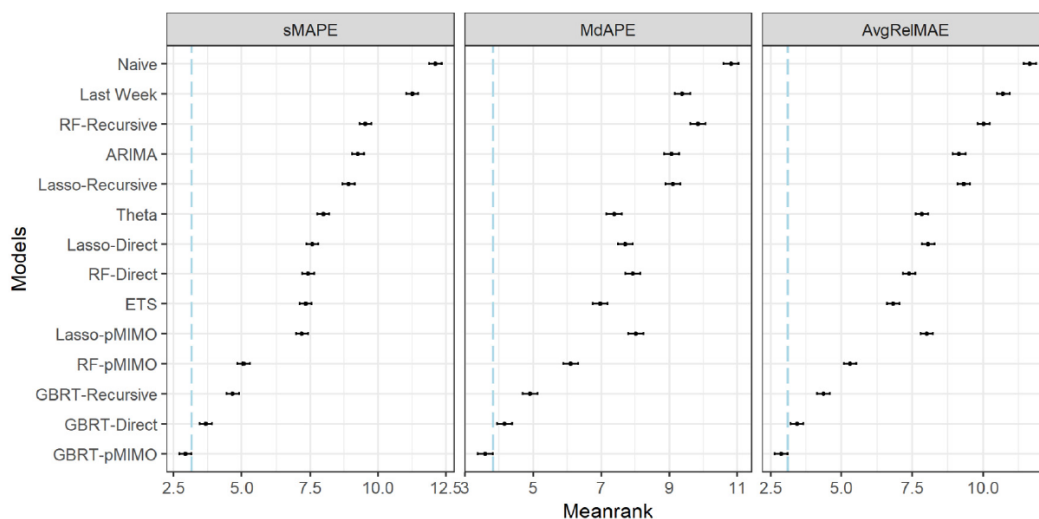
ARIMA: ajusta los parámetros de la serie temporal para capturar patrones de tendencia y estacionalidad. Las métricas obtenidas son sMAPE: 0.105, MdAPE: 0.071 y AvgRelMAE: 0.819 en el Conjunto 1. Aunque ARIMA es robusto en series temporales individuales, mostró

limitaciones al capturar variabilidad en datos agrupados. En la Figura 17, ARIMA presenta una dispersión considerable en AvgRelMAE, lo cual indica su menor precisión y adaptabilidad en datos complejos, especialmente al compararlo con el rango más ajustado de GBRT.

Pooled Regression (Lasso): Este modelo emplea una regresión penalizada L-1 (Lasso) en series temporales agrupadas para identificar patrones comunes. Las métricas obtenidas son sMAPE: 0.114, MdAPE: 0.076 y AvgRelMAE: 0.865 en el Conjunto 1. En la Figura 16, se observa que Lasso tiene una precisión moderada en MdAPE, pero presenta una mayor variabilidad en los resultados de AvgRelMAE en comparación con GBRT, sugiriendo una menor adaptabilidad en series con alta complejidad.

Figura 18

Resultados del Test de Nemenyi al Nivel de Significancia del 5%



Nota. Resultados del test de Nemenyi al nivel de significancia del 5% para pronósticos de 1 a 14 días adelante a nivel de tienda durante todo el período de prueba. La distancia crítica para el test de Nemenyi es 0.443. Tomado de Forecasting third-party mobile payments with implications for customer flow prediction (Ma & Fildes, 2020).

Random Forest (RF): modelo de árboles de decisión ensamblados, adecuado para datos de alta dimensionalidad. Las métricas obtenidas son sMAPE: 0.104, MdAPE: 0.064 y AvgRelMAE: 0.745 en el Conjunto 1. RF mostró un rendimiento sólido en comparación con métodos tradicionales, pero fue superado por GBRT, especialmente en MdAPE y AvgRelMAE, donde GBRT mostró una mayor generalización. La Figura 18 refuerza esta comparación: RF tiene un menor rendimiento estadístico que GBRT en el test de Nemenyi, confirmando su inferioridad en términos de precisión y estabilidad.

Gradient Boosting Regression Tree (GBRT): GBRT, implementado con el paquete Xgboost, se enfoca en corregir errores de predicción en cada iteración, combinando árboles de decisión. Las métricas obtenidas son sMAPE: 0.097, MdAPE: 0.057 y AvgRelMAE: 0.676 en el Conjunto 1. GBRT fue el modelo con el rendimiento más alto, mostrando una precisión significativamente superior en comparación con todos los modelos. Su capacidad para adaptarse a la estacionalidad compleja y manejar alta variabilidad lo posiciona como la opción óptima para predicciones en este contexto. En los boxplots de la Figura 16, GBRT muestra una menor variabilidad en sus predicciones y un rango intercuartílico más estrecho, lo cual indica una alta consistencia en sus resultados.

Ma y Fildes (2020), confirman que el modelo Gradient Boosting Regression Tree GBRT, combinado con la estrategia pseudo MIMO, ofrece un rendimiento superior en todas las métricas clave (sMAPE, MdAPE, AvgRelMAE y MPE), demostrando ser altamente efectivo en la predicción de flujos de clientes en entornos con alta estacionalidad y variabilidad. Estos resultados subrayan la adaptabilidad y robustez de GBRT frente a otros métodos tradicionales y de aprendizaje automático para aplicaciones de pronóstico en el ámbito de pagos móviles.

En el artículo "A deep learning model for behavioural credit scoring in banks", Ala'raj et al. (2022), comparan el rendimiento de varios modelos tradicionales de Machine Learning con un modelo avanzado basado en redes neuronales LSTM, diseñado para predecir el comportamiento crediticio de clientes de tarjetas de crédito en bancos. Los modelos tradicionales evaluados por los autores fueron los siguientes:

Figura 19

Comparación de Medidas de Desempeño para todos los Clasificadores

Classifier	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1 score	AUC	Brier score	KS	H-measure	Max accuracy (%)	Optimal threshold
MLP NN	89.11	55.27	95.87	70.12	0.87	0.0836	0.63	0.51	89.17	0.53
SVM	88.04	39.42	97.74	56.18	0.84	0.0939	0.63	0.51	89.46	0.19
Random Forest	88.15	43.24	97.11	59.84	0.88	0.0931	0.64	0.51	88.76	0.26
Logistic Regression	88.01	42.56	97.08	59.18	0.87	0.0902	0.65	0.51	88.68	0.37
MP-LSTM	90.62	72.87	96.25	82.94	0.91	0.0762	0.71	0.63	90.69	0.52

Nota. Comparación de medidas de desempeño para todos los clasificadores en el conjunto de datos transaccional. Tomado de A deep learning model for behavioural credit scoring in banks, (Ala'raj et al., 2022).

Multi-layer Perceptron (MLP): Según Ala'raj et al. (2022), este modelo alcanzó una precisión del 89.11% y una especificidad alta del 95.87%, lo cual es efectivo en la clasificación de clientes con un buen historial de pagos. Sin embargo, su sensibilidad, que mide la capacidad para identificar correctamente los clientes con riesgo de impago, fue relativamente baja (55.27%). Esto significa que el modelo MLP podría no ser ideal para detectar a aquellos clientes que probablemente fallen en sus pagos. Su AUC de 0.87 y puntaje de Brier de 0.0836 indican una capacidad de discriminación moderada.

Support Vector Machine (SVM): Ala'raj et al. (2022), informan que este modelo obtuvo una precisión del 88.04% y una especificidad del 97.74%. Aunque mostró buen rendimiento en

la clasificación de clientes sin riesgo de incumplimiento, su sensibilidad fue baja (39.42%), lo que limita su capacidad para identificar correctamente a los clientes con alto riesgo de impago. Su AUC de 0.84 refleja una capacidad predictiva menor en comparación con otros modelos.

Random Forest (Bosques Aleatorios): Según los autores, este modelo logró una precisión del 88.15% y una especificidad del 97.11%, destacándose en la clasificación de clientes con bajo riesgo. Sin embargo, su sensibilidad del 43.24% es baja, indicando que es menos efectivo para identificar clientes en riesgo de incumplimiento. El puntaje AUC de 0.88 y el puntaje de Brier de 0.0931 sugieren un buen rendimiento en general, pero con limitaciones para predecir comportamientos de pago negativos con alta precisión.

Logistic Regression (Regresión Logística): Ala'raj et al. (2022), señalan que este modelo obtuvo una precisión del 88.01% y una especificidad del 97.08%. Aunque refleja una capacidad sólida para identificar clientes de bajo riesgo de incumplimiento, su sensibilidad de 42.56% muestra limitaciones en la identificación de clientes con alta probabilidad de incumplimiento. Con un AUC de 0.87 y un puntaje de Brier de 0.0902, este modelo sigue siendo confiable para la predicción general, pero no ideal cuando se busca identificar riesgos potenciales con precisión.

En general, los modelos tradicionales mostraron precisión y especificidad altas, pero una baja sensibilidad, lo cual limita su eficacia para predecir clientes de alto riesgo. Estas deficiencias sugieren que, aunque útiles en clasificación general, los modelos tradicionales no son ideales para identificar incumplimientos de pago con precisión (Ala'raj et al., 2022).

Para superar estas limitaciones, los autores abordaron modelos avanzados LSTM. A continuación, se presentan sus propósitos, rendimiento y métricas clave:

Figura 20

Medidas de Desempeño para el Clasificador LSTM

Description	Total	Missed Payments	Proportion	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC	Brier Score	KS	H-Measure
All customers	10.315	2484	24.08%	90.62	72.87	96.25	0.91	7.62%	0.71	0.63
Customers with at least one missed payment during the last two months	2447	1910	78.05%	84.96	90.58	64.99	0.89	10.85%	0.66	0.44
Customers with missed payment during last month	1925	1621	84.21%	88.78	94.69	57.24	0.9	8.5%	0.65	0.45
Customers with two consecutive missed payments	1325	1257	94.87%	94.79	98.33	29.41	0.86	4.45%	0.56	0.25
Customers with three consecutive missed payments	1026	989	96.39%	96.49	99.09	27.03	0.87	3.12%	0.59	0.26

Nota. Medidas de Desempeño para el Clasificador LSTM en el Conjunto de Datos

Transaccional. Tomado de A deep learning model for behavioural credit scoring in Banks (Ala'raj et al., 2022).

Missed Payment Prediction LSTM (MP-LSTM): Este modelo fue diseñado por Ala'raj et al. (2022), para estimar la probabilidad de un pago fallido en el próximo mes para cada cliente, proporcionando alertas tempranas sobre posibles incumplimientos.

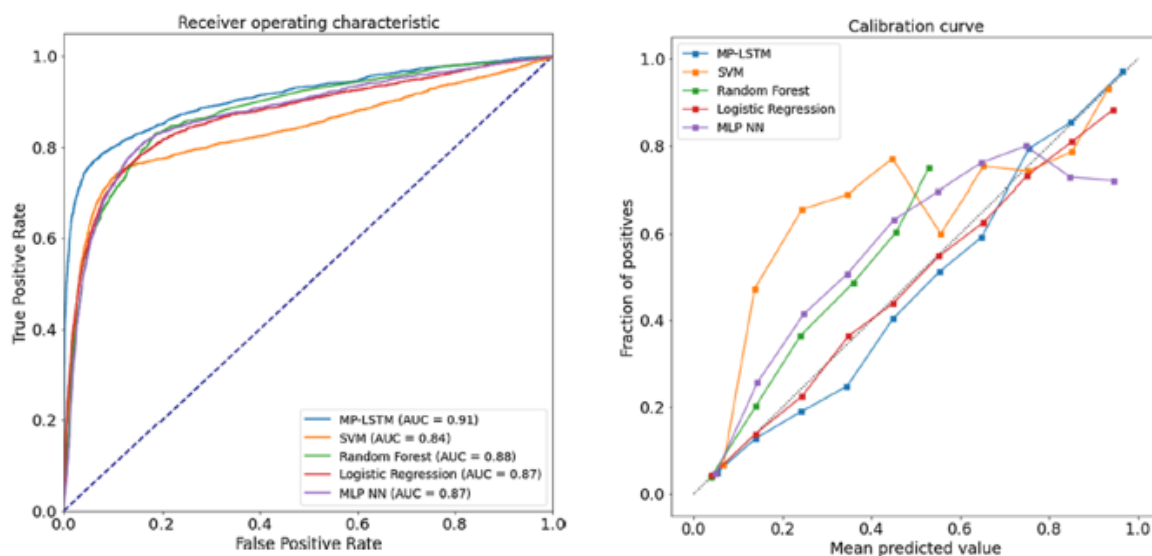
En comparación con los modelos tradicionales, el MP-LSTM mostró una mejora significativa en precisión y capacidad predictiva. Identificó con mayor efectividad a clientes con varios pagos fallidos consecutivos, facilitando la segmentación y gestión de riesgos. Este modelo alcanzó un AUC de 0.91, una sensibilidad del 72.87% y una especificidad del 96.25%, destacándose en su capacidad para discriminar entre clientes con y sin riesgo de impago. Otras métricas, como el puntaje de Brier de 7.62% y el KS de 0.71, confirmaron su robustez para clasificar efectivamente clientes de alto riesgo.

Purchase Estimation LSTM (PE-LSTM): Este modelo fue desarrollado por los autores para estimar el monto total de las compras mensuales, útil para identificar clientes con riesgo de gasto excesivo y probabilidad de incumplimiento.

Evaluado mediante el error cuadrático medio, demostró alta precisión en la estimación de gastos mensuales, contribuyendo a una mejor gestión de riesgos y facilitando la clasificación de clientes. Según Ala'raj et al. (2022), este modelo apoya a las instituciones financieras en la identificación de patrones de gasto, lo cual es clave en la prevención de posibles problemas de crédito.

Figura 21

Curvas ROC, AUC y de Calibración



Nota. Curvas ROC, AUC y de Calibración para Clasificadores en el Conjunto de Datos Transaccional. Tomado de A deep learning model for behavioural credit scoring in Banks (Ala'raj et al., 2022).

En cuanto a las curvas ROC y de calibración, la Figura 21 (curva ROC) muestra visualmente cómo el MP-LSTM tiene el AUC más alto (0.91), confirmando su ventaja sobre los

modelos tradicionales. Además, la Figura 21 (curva de calibración) demuestra la fiabilidad del MP-LSTM, ya que sus predicciones probabilísticas se alinean mejor con los resultados observados, mientras que los modelos tradicionales presentan una calibración menos precisa (Ala'raj et al., 2022).

Ala'raj et al. (2022), concluyen que los modelos tradicionales, aunque efectivos en precisión y especificidad, presentan una sensibilidad relativamente baja. Esto significa que, aunque pueden clasificar correctamente a una gran cantidad de clientes con buen comportamiento de pago, tienen dificultades para identificar clientes con alta probabilidad de incumplimiento. Por otro lado, los modelos MP-LSTM y PE-LSTM destacaron por su habilidad de trabajar con datos transaccionales temporales y demográficos, aprovechando la secuencialidad de las transacciones y brindando predicciones más precisas en la estimación de incumplimientos y gastos.

La arquitectura LSTM es especialmente adecuada para datos temporales, lo que le permite mejorar significativamente la capacidad de predicción de pagos fallidos y clasificación de riesgo crediticio, ofreciendo así una ventaja considerable en la gestión de riesgos y la toma de decisiones anticipadas.

En el artículo de Smirnov (2016), titulado "Modelling Late Invoice Payment Times Using Survival Analysis and Random Forests Techniques", se utiliza un enfoque de análisis de supervivencia junto con Random Survival Forests (RSF) para modelar los tiempos de pago de facturas vencidas en un contexto de ventas B2B. A continuación, se presenta un resumen de los hallazgos clave sobre los modelos utilizados y sus resultados, enfatizando las métricas de evaluación y los principales hallazgos:

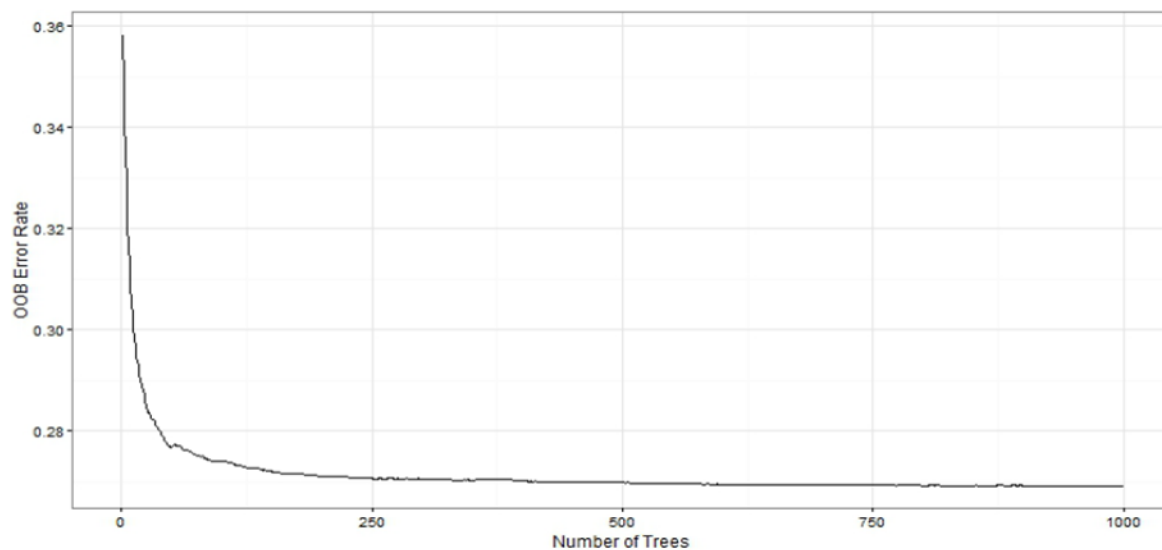
Cox Proportional Hazards Model: El modelo de Cox se aplicó en este estudio con el criterio Akaike (AIC) para la selección de variables, evaluando su capacidad para clasificar correctamente los tiempos de pago. El índice de concordancia (C-index) fue utilizado para medir su rendimiento, alcanzando una concordancia del 64.7% en el conjunto de entrenamiento y 63.3% en el de prueba. Este índice sugiere una capacidad moderada de predicción en la clasificación de tiempos de pago (Smirnov, 2016).

Aunque es un método popular en el análisis de supervivencia, Smirnov (2016), señala que el modelo de Cox presenta limitaciones en datos de pagos atrasados debido a los supuestos que requiere, como la proporcionalidad de riesgos, lo cual puede no ser adecuado en entornos de datos complejos.

Random Survival Forests (RSF):

Figura 22

Cambio en la Tasa de Error OOB



Nota. Cambio en la Tasa de Error OOB en Función del Número de Árboles Crecido. Tomado de Modelling Late Invoice Payment Times Using Survival Analysis and Random Forests Techniques (Smirnov, 2016).

A diferencia del modelo de Cox, RSF no requiere suposiciones estrictas sobre los datos, permitiendo una mayor flexibilidad al manejar datos censurados y variables de comportamiento de pago. El rendimiento de este modelo fue optimizado utilizando 250 árboles, como se muestra en la Figura 22 (Smirnov, 2016), donde se observa que la tasa de error Out-of-Bag (OOB) se estabiliza, indicando una precisión superior en comparación con el modelo de Cox. (Smirnov, 2016).

RSF demostró tener una capacidad predictiva superior, especialmente cuando se incluyeron datos históricos de comportamiento de pago de los deudores, mejorando la precisión en comparación con el modelo de Cox en todos los conjuntos de predictores.

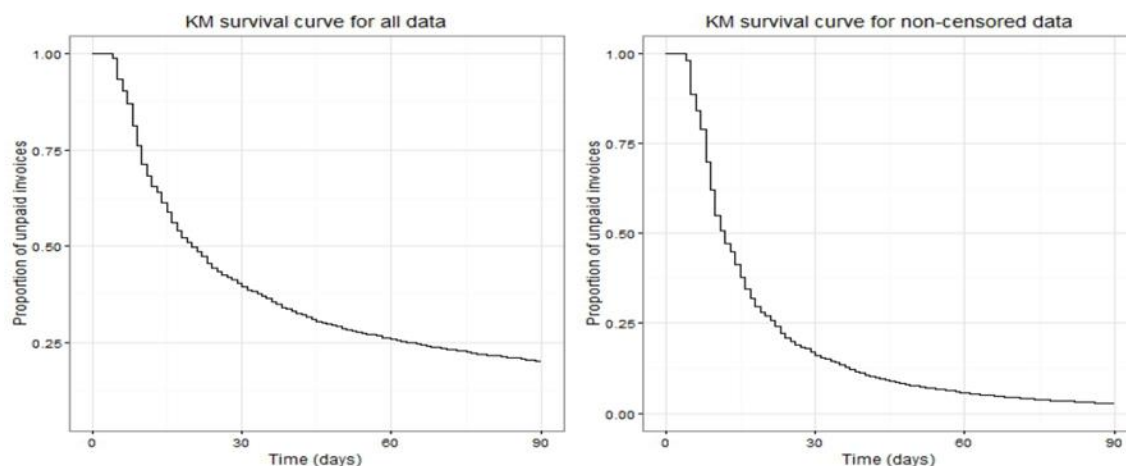
Figura 23

Comparación de Modelos con Diferentes Conjuntos de Predictores

Model	Number of predictors	Random division		Division by time	
		Error (train)	Error (test)	Error (train)	Error (test)
M1: Invoice + Company measures + Fin data + Ratios	16	27.46%	27.58%	26.51%	36.46%
M2: Invoice + Company measures + Fin data	13	27.81%	28.08%	26.88%	36.64%
M3: Company measures + Ratios	12	28.07%	28.22%	27.04%	37.09%

Nota. Comparación de modelos con diferentes conjuntos de predictores, el tamaño del nodo terminal se estableció en 50 pagos únicos y se cultivaron 250 árboles. Tomado de de Modelling Late Invoice Payment Times Using Survival Analysis and Random Forests Techniques (Smirnov, 2016).

RSF mostró una capacidad predictiva superior, especialmente al incluir datos históricos de comportamiento de pago de los deudores. La figura 23 compara el error de predicción en entrenamiento y prueba para diferentes divisiones del modelo y conjuntos de predictores. Este análisis confirma que la inclusión de métricas como ratios financieros y datos de la empresa mejora el rendimiento predictivo, consolidando la ventaja de RSF en escenarios con múltiples variables relevantes. (Smirnov, 2016).

Figura 24*Curvas de Supervivencia de Kaplan-Meier*

Nota. Curvas de supervivencia de Kaplan-Meier para todos los datos de facturas vencidas izquierda y para datos no censurados derecha. Tomado de *Modelling Late Invoice Payment Times Using Survival Analysis and Random Forests Techniques* (Smirnov, 2016).

La figura 24 (Smirnov, 2016, p. 45), muestra las curvas de supervivencia Kaplan-Meier para todos los datos de facturas vencidas y solo para datos no censurados, visualizando cómo el porcentaje de facturas impagas disminuye con el tiempo. Esta visualización respalda la capacidad de RSF para manejar datos censurados y sugiere que la inclusión de datos censurados es crucial para obtener predicciones más precisas en pagos atrasados.

Smirnov (2016), concluye que el modelo Random Survival Forests es el más efectivo para la predicción de tiempos de pago de facturas vencidas debido a su flexibilidad para incluir variables sin requerir transformaciones y su capacidad para manejar datos censurados. Este enfoque sería útil en entornos de Contact-Center del sector cobranza, permitiendo la clasificación de deudores en función de su probabilidad y tiempo estimado de pago, optimizando así las estrategias de cobranza y gestión de deudores (Smirnov, 2016).

En el artículo "Macroeconomic Predictions Using Payments Data and Machine Learning" realizado por Chapman y Desai (2023), se aplican distintos modelos de aprendizaje automático para mejorar las predicciones macroeconómicas, empleando datos de pagos en el contexto canadiense. A continuación, se examinan las métricas de desempeño y los resultados:

Chapman y Desai (2023), implementaron un conjunto de modelos de aprendizaje automático (Elastic Net, Support Vector Regression, Random Forest Regression, Gradient Boosting y Redes Neuronales Artificiales) evaluados principalmente mediante el RMSE (Root Mean Squared Error) en datos fuera de muestra (out-of-sample) para determinar su eficacia en la predicción en tiempos de normalidad y crisis económica.

Elastic Net (ENT): Según Chapman y Desai (2023), este modelo demostró cierta efectividad para controlar la multicolinealidad, que es común en datos económicos de alta dimensionalidad. Sin embargo, su enfoque lineal limita su capacidad para capturar relaciones complejas en los datos de pagos, lo que resulta en un rendimiento moderado en términos de precisión. Con un RMSE de 0.96 en la predicción del PIB para el horizonte $t+1$, el ENT es adecuado en contextos donde las relaciones lineales predominan, pero presenta limitaciones cuando se enfrentan interacciones no lineales o efectos asimétricos, como los observados durante crisis económicas.

Máquinas de Soporte Vectorial (SVM): Chapman y Desai (2023), encontraron que este modelo mostró un RMSE relativamente alto de 1.41 en la predicción del PIB en el horizonte $t+1$, lo que evidencia su menor capacidad para capturar la complejidad inherente de los datos de pagos en este contexto. Aunque el SVR es robusto para manejar multicolinealidad y puede ser eficaz en problemas de clasificación, en este caso, su rendimiento quedó rezagado, posiblemente debido a su incapacidad para adaptarse a relaciones no lineales más complejas.

Random Forest Regression (RFR): Este modelo obtuvo un desempeño intermedio, con un RMSE de 1.11 para el horizonte $t+1$ en la predicción del PIB. Chapman y Desai (2023), destacan que, aunque maneja adecuadamente datos de alta dimensionalidad y permite una interpretación de la importancia de cada variable, el enfoque de promediar múltiples árboles reduce su capacidad para capturar detalles finos y patrones complejos en los datos de pagos. La falta de interdependencia entre los árboles en el RFR, característica de los modelos de ensamblado basados en bagging, limita su capacidad de ajuste en momentos de alta volatilidad.

Gradient Boosting (GBR): El GBR obtuvo el mejor desempeño, alcanzando un RMSE de 0.81 para el PIB en $t+1$. La técnica de boosting, que construye secuencialmente modelos a partir de los errores residuales de predicciones anteriores, permite que el GBR capture patrones no lineales y asimétricos en los datos. Esto es especialmente útil en contextos de crisis, como el COVID-19, donde los efectos de intervención y variabilidad económica requieren una capacidad de ajuste dinámica (Chapman & Desai, 2023).

Redes Neuronales Artificiales (ANN): Finalmente, las ANN lograron un RMSE de 0.82, cercano al GBR, aunque su rendimiento general fue ligeramente inferior. Según Chapman y Desai (2023), las ANN son adecuadas para captar patrones complejos y no lineales en los datos; sin embargo, su requerimiento de datos extensivos para un ajuste preciso limita su efectividad en contextos de datos económicos con menor frecuencia de actualizaciones y series temporales más cortas.

Figura 25*Comparaciones de RMSE Fuera de Muestra de DFM con Modelos de ML*

Target ^b	DFM ^c	ENT ^d	SVR ^d	RFR ^d	GBR ^d	ANN ^d	% Reduction ^e
GDP	1.00	0.96	1.41	1.11	<u>0.81</u> ^f	0.82	19
RTS	1.00	0.89	1.27	1.07	<u>0.85</u>	1.02	15
WTS	1.00	0.96	1.14	0.82	0.69	<u>0.51</u>	31
GDP	1.00	0.87	1.62	1.14	<u>0.82</u>	0.85	18
RTS	1.00	<u>0.87</u>	1.36	1.15	0.90	0.97	11
WTS	1.00	0.89	1.19	0.91	0.81	<u>0.70</u>	19

Nota. Comparaciones de RMSE fuera de muestra de DFM con modelos de ML para la tasa de crecimiento interanual ajustada estacionalmente de variables macro en los horizontes (t+1) panel superior y (t+2) panel inferior para el caso principal. Tomado de *Macroeconomic Predictions Using Payments Data and Machine Learning* (Chapman & Desai, 2023).

Figura 26*Comparaciones de RMSE Fuera de Muestra para las Tasas de Crecimiento*

Targets	Pre-COVID-19 Test Set ^b	COVID-19 Test Set ^c
GDP	16	34
RTS	14	35
WTS	27	37

Nota. Comparaciones de RMSE fuera de muestra para las tasas de crecimiento interanual ajustadas estacionalmente del PIB, RTS y WTS en el horizonte de nowcasting (t+1) utilizando el modelo de gradient Boosting. Tomado de *Macroeconomic Predictions Using Payments Data and Machine Learning* (Chapman & Desai, 2023).

Figura 27

Comparaciones de RMSE Fuera de Muestra

Nowcasting Horizon ^b	Latest Vintages ^c	Real-Time Vintages ^d
t	3.73	3.88
$t + 1$	2.61	2.92
$t + 2$	2.66	2.68

Nota. Comparaciones de RMSE fuera de muestra para la tasa de crecimiento interanual ajustada estacionalmente del PIB en los horizontes de nowcasting (t), ($t+1$) y ($t+2$) utilizando el modelo de gradient boosting. Tomado de *Macroeconomic Predictions Using Payments Data and Machine Learning* (Chapman & Desai, 2023).

Chapman y Desai (2023), concluyen que el modelo Gradient Boosting Regression (GBR) es el más eficaz en la predicción macroeconómica basada en datos de pagos en Canadá, logrando una reducción del RMSE de hasta un 31% en el horizonte de $t+1$, particularmente durante la crisis del COVID-19. En términos específicos, el GBR obtuvo un RMSE tan bajo como 0.81 en la predicción del PIB, en comparación con otros modelos como Elastic Net y SVR, que presentaron RMSEs de 0.96 y 1.41, respectivamente. Estos resultados subrayan la capacidad del GBR para gestionar datos de alta dimensionalidad y capturar efectos no lineales, optimizando así la precisión en contextos de alta variabilidad económica. Su adaptabilidad lo hace particularmente adecuado para aplicaciones en sectores como el de cobranzas, donde podría predecir comportamientos de pago con alta precisión y responder de manera efectiva a patrones complejos y cambios abruptos en los datos.

En el artículo "Late Payment Prediction Models for Fair Allocation of Customer Contact Lists to Call Center Agents" de Kim y Kang (2016), se evaluaron diversos modelos de predicción de pagos tardíos para optimizar la asignación de listas de contacto en centros de

cobranza. Los modelos se analizaron empleando métricas de elevación (lift) y medidas de equidad en la distribución de listas, permitiendo así una comparación precisa de su rendimiento.

La métrica de elevación se utiliza para evaluar la efectividad del modelo en capturar clientes de alto riesgo en los deciles superiores, lo cual facilita la identificación de los clientes prioritarios para cobranza. Como explican Kim y Kang (2016), "el valor de elevación mide la precisión del modelo en cada decil, destacando especialmente el rendimiento en los deciles superiores, lo que permite una comparación equitativa entre modelos, independientemente del umbral de clasificación" (p. 90).

Un valor alto de elevación en el primer decil indica que el modelo identifica correctamente a los clientes con alta probabilidad de pago en ese segmento superior, mejorando la eficacia en la gestión de contacto. Esto es esencial para la asignación de listas, donde los modelos con alta elevación optimizan la precisión y equidad en la carga de trabajo de los agentes (Kim & Kang, 2016).

Figura 28

Valores de Lift por Decil para Tfull

Decile	Model_July					Model_August					Model_September				
	ST	RF	ANN	SVM	Hybrid	ST	RF	ANN	SVM	Hybrid	ST	RF	ANN	SVM	Hybrid
Top10%	1.5360	1.6055	1.6017	1.2809	1.6422	1.7272	1.6158	1.7118	1.3703	1.7272	2.1325	2.1516	2.1022	1.8707	2.1915
Top20%	1.5299	1.5145	1.5196	1.3985	1.4971	1.6792	1.6824	1.6202	1.5217	1.6945	1.6313	1.6967	1.6760	1.6504	1.6760
Top30%	1.1573	1.4033	1.4078	1.3537	1.3872	1.5649	1.6908	1.3832	1.4575	1.5310	1.4126	1.3073	1.4190	1.3951	1.3440
Top40%	1.2349	1.1917	1.2079	1.2961	1.2562	1.0193	1.0807	1.2317	1.2794	1.1275	1.0040	1.1844	1.1508	0.9928	1.1700
Top50%	1.2597	1.1560	1.1901	1.1463	1.1785	0.9019	0.9972	1.0828	1.0295	0.9803	0.7422	0.7614	0.6544	0.8300	0.7103
Top60%	0.9937	1.0014	1.0594	0.9886	0.9795	0.8955	0.9100	0.8939	0.9052	0.9116	0.6161	0.6975	0.6065	0.6975	0.6720
Top70%	0.8310	0.7176	0.7576	0.7067	0.6603	0.8885	0.7786	0.7625	0.6663	0.7479	0.7215	0.6624	0.6784	0.6991	0.6848
Top80%	0.6466	0.6382	0.4849	0.7142	0.6060	0.6419	0.4796	0.5337	0.6201	0.5757	0.7406	0.6911	0.7550	0.6496	0.6209
Top90%	0.5044	0.4941	0.4574	0.6635	0.4902	0.4580	0.4668	0.4782	0.5904	0.4014	0.5172	0.4581	0.5906	0.5938	0.5204
Top100%	0.3065	0.2776	0.3136	0.4515	0.3027	0.2237	0.2980	0.3020	0.5596	0.3028	0.4820	0.3895	0.3671	0.6209	0.4102

Nota. Valores de lift por decil para Tfull según los cinco modelos de predicción para cada período. Tomado de Late Payment Prediction Models for Fair Allocation of Customer Contact Lists to Call Center Agents (Kim & Kang, 2016).

Árbol de Decisión: Este modelo generó reglas de predicción interpretables, mostrando un rendimiento adecuado en escenarios donde la transparencia y simplicidad de las predicciones son esenciales. En el primer decil de elevación, alcanzó un valor de 1.4546 en julio, aunque otros modelos superaron su precisión en la clasificación de pagadores tardíos en deciles superiores.

Random Forest: El modelo destacó por su precisión y su capacidad para lograr una asignación imparcial de las listas de contacto en función de la probabilidad de pago. En términos de elevación, obtuvo un valor de 1.4805 en julio y 2.1516 en septiembre en el primer decil, mostrando consistencia en distintos escenarios y un desempeño robusto ante datos desequilibrados. Además, este modelo mantuvo una distribución más equitativa en la asignación de listas de contacto, evidenciada por métricas de variabilidad como la desviación estándar y el rango intercuartílico: en el último decil (100), la desviación estándar fue de 793 y el rango intercuartílico (Q3–Q1) alcanzó un valor de 1,008, lo que confirma su adecuación para una distribución justa de carga de trabajo entre los agentes.

Redes Neuronales Artificiales (ANN): Este modelo capturó patrones complejos en los datos, destacando en precisión, aunque con menor interpretabilidad que los árboles de decisión y Random Forest. En julio, obtuvo un valor de elevación de 1.4196 en el primer decil, mostrándose eficaz en entornos donde la precisión era prioritaria sobre la simplicidad interpretativa.

Máquinas de Soporte Vectorial (SVM): Las SVM lograron buenos resultados, separando efectivamente a los clientes con alta probabilidad de pago de aquellos con menor probabilidad. Su valor de elevación en septiembre fue de 1.8707 en el primer decil, aunque se quedó ligeramente por debajo de Random Forest y el modelo híbrido en algunos escenarios.

Enfoque Híbrido: Este enfoque combinó los resultados de los modelos anteriores, logrando mejorar tanto la precisión general como la varianza en las predicciones. Obtuvo el valor

más alto de elevación en múltiples periodos, con 1.4634 en julio y 2.1915 en septiembre en el primer decil. Además, se mantuvo competitivo en las métricas de equidad en la distribución, con un coeficiente de variación de 5.263 y una desviación estándar baja en los últimos deciles, lo cual favorece una distribución justa de listas entre agentes en escenarios complejos.

Figura 29

Resultado de la Asignación de la Lista de Contactos del Cliente

N	Measure	TNR	CUM	CJA	STP	EPST	RFP	EPRF	ANNP	EPANN	SVMP	EPSVM	HybridP	EPHybrid
10	Stdev	2,343	2,255	2,868	2,226	2,326	2,633	1,689	2,562	1,627	2,147	2,443	2,435	2,546
	Range	7,371	6,569	8,391	6,683	8,193	9,465	5,176	8,543	4,387	6,491	7,230	6,721	8,116
	Q3-Q1	4,007	3,535	5,047	3,715	3,116	3,129	2,807	2,690	2,798	2,807	3,606	4,117	2,399
	CV	1,821	1,752	2,228	1,730	1,807	2,046	1,312	1,991	1,265	1,668	1,899	1,892	1,979
20	Stdev	1,827	1,610	1,687	1,706	1,688	1,574	1,201	1,857	1,500	2,395	1,989	1,602	1,688
	Range	7,176	5,612	6,379	6,420	6,066	5,275	3,955	6,340	5,686	9,726	7,792	6,538	6,462
	Q3-Q1	2,839	2,559	2,589	2,434	2,134	2,434	1,945	2,738	2,506	2,877	3,207	1,546	2,515
	CV	2,840	2,503	2,621	2,651	2,623	2,447	1,867	2,886	2,331	3,722	3,092	2,490	2,623
30	Stdev	1,547	1,700	1,496	1,310	1,376	1,616	1,029	1,719	1,331	1,426	1,441	1,569	1,614
	Range	7,232	6,822	6,540	5,089	6,175	6,164	3,964	6,625	6,325	6,429	5,051	6,035	5,795
	Q3-Q1	1,616	2,184	1,824	1,622	1,939	2,378	1,672	2,803	1,599	2,247	1,988	2,395	2,442
	CV	3,606	3,964	3,487	3,054	3,206	3,768	2,400	4,007	3,103	3,325	3,359	3,657	3,763
40	Stdev	1,399	1,362	1,153	1,145	1,103	1,094	1,080	1,395	1,148	1,647	1,331	1,081	1,172
	Range	5,276	6,277	4,410	4,865	4,683	4,983	4,676	6,042	5,377	8,879	5,442	3,887	4,815
	Q3-Q1	2,177	1,642	1,679	1,509	1,614	1,383	1,352	1,720	1,607	1,632	1,275	1,450	1,528
	CV	4,350	4,232	3,582	3,558	3,428	3,401	3,357	4,335	3,569	5,119	4,137	3,355	3,643
50	Stdev	892	1,246	1,104	910	1,161	1,400	1,074	1,142	932	1,257	1,116	1,119	1,180
	Range	3,538	6,231	4,855	3,780	5,079	6,649	4,732	4,709	4,415	5,808	5,460	4,972	4,908
	Q-Q1	1,310	1,355	1,697	1,489	1,482	1,747	1,536	1,626	1,428	1,989	1,405	1,379	1,690
	CV	3,464	4,842	4,288	3,534	4,509	5,439	4,174	4,437	3,622	4,885	4,336	4,346	4,584
60	Stdev	1,105	1,105	1,016	897	927	1,044	750	1,232	991	1,152	1,042	1,166	1,072
	Range	5,184	5,323	4,690	4,260	4,835	4,191	3,763	5,375	5,319	5,272	4,749	4,742	4,669
	Q3-Q1	1,597	1,035	1,527	1,389	964	1,216	960	1,676	1,201	1,551	1,543	1,837	1,366
	CV	5,153	5,150	4,735	4,181	4,321	4,868	3,496	5,743	4,620	5,369	4,860	5,436	4,999
70	Stdev	1,022	983	929	1,026	1,009	917	844	1,072	940	957	833	935	880
	Range	5,087	4,372	4,360	4,856	4,479	4,218	4,317	6,104	4,483	4,673	3,973	4,263	4,442
	Q3-Q1	1,149	1,291	1,257	1,393	1,143	1,178	1,058	1,236	1,203	1,195	1,102	1,290	1,039
	CV	5,561	5,349	5,052	5,579	5,489	4,987	4,590	5,831	5,114	5,206	4,529	5,085	4,787
80	Stdev	944	998	842	846	771	923	807	941	795	1,076	875	854	803
	Range	5,048	4,570	4,055	3,646	3,797	5,374	4,064	4,157	3,567	6,861	3,629	3,462	4,190
	Q3-Q1	1,146	1,436	1,146	1,211	1,088	1,329	867	1,218	1,077	1,230	1,384	1,176	967
	CV	5,869	6,201	5,236	5,260	4,790	5,740	5,016	5,849	4,943	6,688	5,439	5,306	4,992
90	Stdev	840	932	828	804	727	908	850	878	855	870	804	964	794
	Range	4,717	4,275	3,813	3,474	4,020	4,046	4,126	3,619	5,021	4,278	3,889	4,834	4,004
	Q3-Q1	1,073	1,395	1,213	1,071	1,026	1,160	1,289	1,337	1,120	1,099	1,114	1,440	953
	CV	5,872	6,520	5,789	5,620	5,084	6,351	5,942	6,140	5,976	6,086	5,625	6,744	5,556
100	Stdev	835	849	781	714	836	900	750	865	704	919	818	796	773
	Range	4,846	4,224	3,531	3,467	3,904	4,824	3,735	3,970	3,735	5,069	4,355	3,983	3,191
	Q3-Q1	1,099	1,104	1,158	1,022	1,244	1,176	1,027	1,193	994	988	1,098	1,122	1,113
	CV	6,484	6,600	6,069	5,549	6,495	6,990	5,829	6,721	5,471	7,139	6,354	6,186	6,003
Best case		4	0	0	1	4	0	12	0	6	1	4	4	4

Nota. La Desviación Estándar (Stdev), Rango, Rango Intercuartílico (Q3–Q1) y Coeficiente de Variación (CV) para el Resultado de la Asignación de la Lista de Contactos del Cliente con las Reglas de Puntuación del Cliente (Tfull, Model_July). Tomado de Late Payment Prediction Models for Fair Allocation of Customer Contact Lists to Call Center Agents (Kim & Kang, 2016).

Kim y Kang (2016), concluyen que el Random Forest y el Enfoque Híbrido fueron los más efectivos para lograr una asignación justa de listas de contacto en cobranza, mejorando tanto la precisión como la equidad en la distribución de clientes. La evaluación mediante métricas específicas de elevación y variación sugiere que ambos modelos son aplicables en el contexto de call centers, donde una percepción de equidad puede impactar positivamente en la retención y rendimiento de los agentes, haciendo de estos modelos herramientas valiosas para la predicción de pagos en el sector de cobranza (Kim & Kang, 2016).

Análisis Comparativo de Modelos de Machine Learning

En la gestión de cobranza, el uso de modelos de Machine Learning es esencial para optimizar la toma de decisiones y mejorar la precisión en la predicción de pagos de clientes. Diversos estudios han analizado múltiples algoritmos, de los cuales sobresalen cinco modelos que muestran el mejor comportamiento en la predicción de pagos debido a su capacidad para manejar datos complejos y mejorar la eficiencia operativa en Contact-Centers.

Tabla 7*Frecuencia de Comportamiento de Variables según País*

Autor	Modelo Mejor Rendimiento	Conclusión del Rendimiento
Campos Cortesía (2020)	Random Forest	Alta precisión y versatilidad en escenarios de cobranza.
Ortiz et al. (2022)	K-Vecinos Más Cercanos (KNN)	Buena capacidad predictiva, aunque con falsos positivos.
Wagh et al. (2021)	Gradient Boosting	Eficaz para datos de alta variabilidad y volumen.
Aslam et al. (2022)	SVM	Excelente para clasificación de clientes en diferentes probabilidades de pago.
Fan et al. (2023)	Hybrid Multi-Stage Model	Alta efectividad en manejo de datos desequilibrados.
Ma & Fildes (2020)	Gradient Boosting Regression Tree (GBRT)	Modelo superior para predicción en entornos de alta variabilidad.
Ala'raj et al. (2022)	LSTM	Adecuado para predicciones secuenciales en Contact-Centers.
Smirnov (2016)	Random Survival Forests (RSF)	Ideal para estimaciones temporales de pagos.
Chapman & Desai (2023)	Gradient Boosting	Capaz de capturar patrones complejos, útil en crisis.
Kim & Kang (2016)	Random Forest & Hybrid Approach	Balance entre precisión y asignación justa en listas de clientes.

El Random Forest (RF) es un modelo versátil, destacándose en la clasificación de clientes con alta precisión y sensibilidad. En el estudio de Campos (2020), el RF alcanzó una precisión del 79% y un F1 Score de 80%, lo que lo hace adecuado para entornos de cobranzas que

requieren un balance robusto entre precisión y sensibilidad. Fan et al. (2023), subrayan además su eficacia en aplicaciones de regresión, capturando patrones complejos en datos de riesgo crediticio.

El Random Survival Forests (RSF) se orienta hacia la predicción temporal de pagos vencidos. Smirnov (2016), evidenció su precisión superior en la estimación del tiempo de pago en contextos B2B, donde los datos censurados son comunes. El RSF permite una clasificación de deudores en función de la probabilidad de pago en el tiempo, optimizando así la gestión de cobranzas en Contact-Center.

Para secuencias temporales de datos, el MP-LSTM destaca en la predicción de riesgos de incumplimiento. En la investigación de Ala'raj et al. (2022), el modelo LSTM alcanzó una precisión del 91% y un AUC de 0.91, mostrando su capacidad de diferenciar con precisión entre clientes de alto y bajo riesgo de incumplimiento, permitiendo alertas tempranas cruciales en la gestión de cobranza.

El Gradient Boosting Machine (GBM) sobresale en escenarios de alta variabilidad en los datos, como el de pagos y cobranza, donde captura patrones no lineales con notable consistencia. Chapman y Desai (2023) y Ma y Fildes (2020), resaltan la precisión del GBM en contextos de datos macroeconómicos volátiles, sugiriendo su adaptabilidad en aplicaciones de cobranza con fluctuaciones significativas en los patrones de pago.

Por último, el Support Vector Machine (SVM), aunque con una especificidad limitada, es eficaz en la clasificación de clientes de alto riesgo. Ortiz et al. (2022), indican su precisión en la segmentación de carteras de cobranza, logrando una sensibilidad del 82.38%, ideal para identificar clientes con mayores probabilidades de efectividad en la gestión de pagos.

En conclusión, el análisis de estos modelos resalta su utilidad y adaptabilidad en el sector cobranza, cada uno con ventajas específicas. El RF y el RSF son robustos para aplicaciones de clasificación y predicción temporal; el MP-LSTM y el GBM son eficaces en entornos con secuencias temporales y alta variabilidad, respectivamente, y el SVM es adecuado para clasificaciones enfocadas en riesgos específicos. La elección del modelo ideal dependerá del tipo de datos y de los objetivos de la estrategia de cobranza, siendo fundamental un preprocesamiento y ajuste cuidadoso para optimizar el rendimiento en el Contact-Center.

Conclusiones

Los modelos de Machine Learning más destacados en la literatura, como Random Forest, Árboles de Decisión y Redes Neuronales, evidencian un rendimiento superior en la segmentación de clientes y la predicción del comportamiento de pago. Su capacidad para procesar grandes volúmenes de datos y manejar relaciones no lineales ha sido clave para optimizar estrategias de cobranza en el sector.

En este sentido, los enfoques basados en Máquinas de Soporte Vectorial (SVM) y XGBoost han mostrado una efectividad considerable en contextos específicos, pero presentan limitaciones operativas debido a su elevada complejidad computacional y dependencia de datos balanceados. Estos factores condicionan su aplicabilidad en escenarios operativos con recursos limitados.

Además de los algoritmos utilizados, el análisis de las variables empleadas en modelos predictivos revela que el historial de pagos, la edad de mora y las características demográficas constituyen los factores más influyentes en la predicción del comportamiento de pago. Estas variables permiten a los modelos segmentar clientes con mayor precisión, facilitando una intervención temprana y eficiente.

Por otra parte, la incorporación de variables económicas, como los niveles de ingresos y la capacidad de endeudamiento, potencia significativamente la capacidad predictiva de los modelos. No obstante, su eficacia depende de la calidad y la actualización constante de los datos, lo que resalta la importancia de un manejo dinámico de la información en entornos cambiantes.

A partir de estos hallazgos, los resultados del análisis posicionan a Random Forest como uno de los modelos más robustos para aplicaciones en tiempo real, gracias a su equilibrio entre

precisión, velocidad de procesamiento y capacidad de generalización. Este modelo es especialmente adecuado para contextos con múltiples variables y grandes volúmenes de datos.

Finalmente, los modelos de ensamblaje, como Bagged Decision Trees, presentan un desempeño superior en términos de exactitud y sensibilidad, siendo ideales para análisis exhaustivos en escenarios donde la precisión es prioritaria. Sin embargo, su elevado costo computacional los hace menos viables en entornos operativos que demandan resultados en tiempo real.

Recomendaciones

Fomentar la implementación de modelos híbridos es una estrategia fundamental para aprovechar las capacidades complementarias de enfoques como Random Forest y XGBoost. La combinación de estos métodos permite mejorar tanto la precisión como la adaptabilidad de las predicciones en el sector de cobranzas, abordando de manera más efectiva la heterogeneidad de los datos operativos. Sin embargo, para garantizar su éxito, es necesario considerar las particularidades del contexto en el que se aplican.

En este sentido, implementar evaluaciones contextuales de modelos resulta esencial. Las pruebas preliminares, adaptadas al entorno operativo, las características de los datos y las restricciones computacionales aseguran que los modelos seleccionados sean adecuados y eficientes para cada situación específica. Este enfoque contribuye a maximizar el rendimiento de los algoritmos en escenarios reales y dinámicos.

Dado el rápido avance en inteligencia artificial, también se recomienda ampliar las investigaciones en técnicas emergentes. Modelos avanzados, como redes neuronales profundas y algoritmos de ensamblaje dinámico, ofrecen un alto potencial en contextos complejos de gestión de cobranzas. La exploración de estas tecnologías puede proporcionar soluciones innovadoras para desafíos específicos del sector.

Además, fortalecer la calidad y la actualización de los datos es crucial para el éxito de cualquier modelo predictivo. Sistemas de gestión de datos que permitan recopilar y actualizar variables en tiempo real, como el historial de pagos y las métricas económicas, garantizarán que los algoritmos trabajen con información precisa y confiable, lo que es clave en entornos dinámicos.

Complementariamente, implementar análisis de importancia de variables resulta indispensable para optimizar el desempeño de los modelos. Evaluaciones sistemáticas de interacciones clave, como la edad de mora y las características demográficas, permitirán seleccionar atributos que tengan un impacto significativo en las predicciones, mejorando la eficacia de los algoritmos.

Para enriquecer aún más las capacidades predictivas, integrar fuentes de datos externas es una recomendación clave. Incorporar indicadores macroeconómicos y tendencias del mercado puede aportar un contexto más amplio al análisis, fortaleciendo la robustez de los modelos ante cambios económicos o sociales.

Los resultados del análisis también destacan la importancia de optimizar el uso de modelos según el contexto. Random Forest, por ejemplo, se posiciona como una opción ideal para aplicaciones en tiempo real debido a su equilibrio entre precisión y velocidad. En contraste, modelos como Bagged Decision Trees son más adecuados para análisis exhaustivos que prioricen la exactitud sobre la rapidez.

Adicionalmente, capacitar en técnicas de ajuste y validación es fundamental para maximizar el rendimiento de los modelos. El dominio de estrategias avanzadas, como el ajuste de hiperparámetros y el balanceo de datos, permite optimizar algoritmos como SVM y XGBoost, haciéndolos más eficientes en entornos operativos diversos.

Finalmente, establecer mecanismos de monitoreo del desempeño garantiza la efectividad continua de los modelos. La implementación de sistemas que evalúen métricas clave, como precisión, sensibilidad y especificidad, permitirá medir el impacto de los algoritmos en las estrategias de cobranza y asegurar su relevancia frente a los cambios en los datos y en las condiciones operativas.

Referencias Bibliográficas

- Ala'raj, M., Abbod, M.F., Majdalawieh, M., et al. (2022). *A deep learning model for behavioural credit scoring in banks*. *Neural Comput & Applic*. <https://doi.org/10.1007/s00521-021-06695-z>
- Anderson, R. (2007). *The Credit Scoring Toolkit: Theory and Practice for Retail Credit Risk Management and Decision Automation*. Oxford University Press.
- Angarita Villamizar, J y Marín Guerra, R. (2019). *Modelo de predicción de incumplimiento de pago en sector de microfinanzas - Analytics*. *Uniandes*. <http://hdl.handle.net/1992/43940>
- Antonio, B. O., Juan, L. R., Ana, Í. D., et al. (2024). *Examinar el comportamiento del usuario con aprendizaje automático para una adopción efectiva de pagos móviles entre pares*. *Financial Innovation*, 10(94). <https://doi.org/10.1186/s40854-024-00625-3>
- Aslam, F., Awan, TM y Fatima, T. (2022). *Clasificación del comportamiento de los usuarios de pagos móviles mediante modelos de aprendizaje automático*. *J Financ Serv Mark* 27 , 264–275. <https://doi.org/10.1057/s41264-021-00114-z>
- Banca de las Oportunidades y Superintendencia Financiera de Colombia. (2023). *Reporte de Inclusión Financiera 2023*. Superintendencia Financiera de Colombia.
- Campos Cortesía, Z. (2020). *Pronóstico del cumplimiento de pago de los clientes usando aprendizaje automático. tesis de magíster, Universidad Nacional de Colombia*. Repositorio Institucional UNAL. <https://repositorio.unal.edu.co/handle/unal/78297>
- Chapman, J.T.E., & Desai, A. (2023). *Macroeconomic Predictions Using Payments Data and Machine Learning*. *Forecasting*. <https://doi.org/10.3390/forecast5040036>
- Chief Revenue Officer en Nubox. (2024). *Gestión de cobranza: ¡todo lo que necesitas saber sobre el tema! Nubox*. <https://blog.nubox.com/empresas/que-es-gestion-de-cobranza>

- Dahlberg, T., Mallat, N., Ondrus, J., & Zmijewska, A. (2008). *Past, present and future of mobile payments research: A literature review*. Electronic commerce research and applications. Ebook - Tendencias de pagos digitales. (2023). *Tendencias de pagos digitales en América Latina*. Consultora TechFin.
- Fan, M., Wu, T.-H., & Zhao, Q. (2023). *Assessing credit risk using machine learning models*. Journal of Financial Analysis. <https://www.mdpi.com/2079-8954/11/10/505>
- Kim, J., & Kang, P. (2016). *Late payment prediction models for fair allocation of customer contact lists to call center agents*. Decision Support Systems. <https://doi.org/10.1016/j.dss.2016.03.002>
- Kyriakides, G., & Margaritis, K. G. (2019). *Hands-On Ensemble Learning with Python: Build Highly Optimized Ensemble Machine Learning Models Using Scikit-learn and Keras*.
- Ma, S., & Fildes, R. (2020). *Forecasting third-party mobile payments with implications for customer flow prediction I*. International Journal of Forecasting, <https://doi.org/10.1016/j.ijforecast.2019.08.012>
- Madhavan Samir. (2015). *Mastering Python for Data Science: Explore the World of Data Science Through Python and Learn How to Make Sense of Data*. Packt Publishing.
- Ortiz Numpaque, Y. L., Ramírez González, E. F., & Cables Pérez, E. H. (2022). *Predicción de clientes efectivos en la gestión de carteras de cobranza castigada en la empresa InteliBPO S.A.S a través de modelos de aprendizaje automático*. tesis de especialización, Universidad Antonio Nariño. Repositorio Institucional UAN. <http://repositorio.uan.edu.co/handle/123456789/7959>
- Raschka, S., & Mirjalili, V. (2017). *Python Machine Learning -Second Edition: Unlock Modern Machine Learning and Deep Learning Techniques with Python by Using the Latest*

Cutting-edge Open Source.

Rose, S. (2016). *A Machine Learning Framework for Plan Payment Risk Adjustment*. Health Services Research, 51(6), Part I.

Schoonbee, L., Moore, W., & Van Vuuren, J. (2022). *A MACHINE-LEARNING APPROACH TOWARDS SOLVING THE INVOICE PAYMENT PREDICTION PROBLEM*. The South African Journal of Industrial Engineering.

Singh, P. P., Anik, F. I., Senapati, R., Sinha, A., Sakib, N., & Hossain, E. (2024). *Investigating customer churn in banking: a machine learning approach and visualization app for data science and management*. Data Science and Management.
<https://doi.org/10.1016/j.dsm.2023.09.002>

Smirnov, J. (2016). *Modelling late invoice payment times using survival analysis and random forests techniques* (Doctoral dissertation, University of Tartu).

Thomas, L. C., Edelman, D. B., & Crook, J. N. (2002). *Credit Scoring and Its Applications*. SIAM.

Wagh, S. K., Andhale, A. A., Wagh, K. S., Pansare, S. K., & Joshi, S. K. (2021). *Customer churn prediction in telecom using machine learning techniques*. Journal of Telecommunications and Information Technology.
<https://doi.org/10.1016/j.rico.2023.100342>

Wang, T. (2022). *Financial management early warning modeling analysis based on RBF neural network algorithm*. In *Proceedings of the 2022 Second International Conference on Advanced Technologies in Intelligent Control, Environment, Computing & Communication Engineering (ICATIECE)*. IEEE.
<https://doi.org/10.1109/ICATIECE56365.2022.10047262>

Xu, T., Ma, Y., & Kim, K. (2021). *Telecom churn prediction system based on ensemble learning using feature grouping*. Applied Sciences. <https://doi.org/10.3390/app11114742>

Zhu, X., Chu, Q., Song, X., Hu, P., & Peng, L. (2023). Explainable prediction of loan default based on machine learning models. Data Science and Management.

Apéndices

Apéndice A

Matriz - Monografía Revisión Documental

<https://unadvirtualedu->

[my.sharepoint.com/:x:/g/personal/vagarcia_mez_unadvirtual_edu_co/EXXRubT1v1NKoCw2zNh](https://unadvirtualedu-my.sharepoint.com/:x:/g/personal/vagarcia_mez_unadvirtual_edu_co/EXXRubT1v1NKoCw2zNh)

[EEsUBi9bGCKZft2VmI0BGZyqvbg?e=zXd49x](https://unadvirtualedu-my.sharepoint.com/:x:/g/personal/vagarcia_mez_unadvirtual_edu_co/EXXRubT1v1NKoCw2zNh)