

Sistemas de Reconocimiento de Patrones para la Predicción de Cumplimiento de Pago de Multas Municipales:

* Un Estudio Comparativo de Modelos de Aprendizaje Supervisado en el Municipio de Ate

1st Roberth Carlos Gonzales Mauricio
Estudiante de Ingeniería
Universidad Nacional del Altiplano
Puno, Perú
73814322gm@gmail.com

2nd Cesia Gomez Flores
Estudiante de Ingeniería
Universidad Nacional del Altiplano
Puno, Perú
floresgcesi@gmail.com

Abstract—Este estudio busca resolver el problema de la falta de pago de multas municipales usando inteligencia artificial. Desarrollamos un sistema que aprende de datos históricos de 53,506 multas del municipio de Ate, en Perú. Analizamos el desempeño de cuatro modelos de inteligencia artificial: Random Forest, Gradient Boosting, Regresión Logística y Support Vector Machine (SVM). Para probarlos, usamos el 20% de los datos para enseñarles y el 80% para ver qué tan bien predecían.

El modelo que mejor funcionó fue Gradient Boosting, con una precisión del 99.27% y un valor de AUC-ROC de 0.9987. También encontramos que el pago de multas es más alto en ciertos tipos de negocios y que, a mayor monto de la multa, menos probable es que se pague. Esto demuestra que la inteligencia artificial es muy útil para encontrar patrones y predecir comportamientos en la gestión de multas, lo que puede ayudar a mejorar las estrategias de recaudación y el uso de los recursos públicos.

I. INTRODUCCIÓN

La gestión eficiente de multas y la optimización de la recaudación constituyen desafíos importantes y críticos para las administraciones municipales a nivel global. La morosidad persiste en el pago de estas obligaciones que conlleva una pérdida grande de ingresos públicos, por lo que restringe la capacidad de financiación de servicios primordiales y la implementación de proyectos de desarrollo urbano [1]. En el distrito de Ate, al igual que en numerosas municipalidades del peru que enfrentan tasas de morosidad superiores al 30% en ingresos no tributarios [2], la capacidad de predecir el comportamiento de pago de los ciudadanos se vuelve fundamental para mejorar la eficiencia municipal recaudatoria.

El aprendizaje automático supervisado se incrusta como una herramienta metodológica para analizar grandes volúmenes de datos e identificar patrones complejos [3]. Su aplicación sistemática permite que la construcción de estos modelos predictivos son capaces de anticipar la probabilidad de cumplimiento de pago, que se basa en el análisis de las multas y los infractores. Estudios previos en campos relacionados, como la predicción de morosidad en carteras de crédito, han demostrado la efectividad de estas técnicas, que reportaron

mejoras de hasta el 15% en la identificación de riesgos [4]. De igual forma, en el ámbito del análisis del comportamiento fiscal, el aprendizaje automático ha optimizado las estrategias de cobranza y ha contribuido a la reducción de costos operativos [5]. La creciente adopción de la inteligencia artificial y el aprendizaje automático en el sector público subraya su potencial para transformar la prestación de servicios y la eficiencia administrativa [6].

Este trabajo se centra en diseñar y evaluar un sistema capaz de identificar patrones utilizando la información histórica sobre las multas aplicadas en el distrito de Ate. El propósito principal es anticipar y predecir el comportamiento de pago de los ciudadanos sancionados. Para lograr este objetivo, se lleva a cabo una evaluación comparativa de distintos algoritmos de aprendizaje supervisado, tales como Random Forest, Gradient Boosting, Regresión Logística y Máquinas de Vectores de Soporte (SVM), con el fin de determinar cuál ofrece el mejor rendimiento en este caso particular. La intención es proporcionar a la municipalidad una herramienta analítica útil que le permita implementar estrategias de recaudación más dirigidas, eficientes y con un mayor impacto en la reducción de la morosidad.

La estructura de este trabajo se organiza de la siguiente manera: Las bases que fundamentan la investigación en la Sección 2 Una descripción detallada de la metodología empleada para la adquisición y preprocesamiento de datos, así como para la construcción y evaluación de los modelos en la Sección 3 Los resultados obtenidos de los experimentos en la Sección 4 Las conclusiones y discusión de las implicaciones de los resultados así como la discusión de posibles limitaciones del estudio en la Sección 5 Finalmente, la Sección 6 resume las principales contribuciones del trabajo y señala futuras líneas de investigación.

II. MARCO TEÓRICO

El marco teórico del presente estudio se enfoca en el uso del aprendizaje supervisado para predecir y evaluar el

cumplimiento de pago de multas municipales. Se revisan los fundamentos de esta técnica, los modelos de clasificación considerados en el análisis, y las métricas utilizadas para medir su rendimiento.

A. Aprendizaje Supervisado

El aprendizaje supervisado se alza de una rama fundamental del aprendizaje automático (Machine Learning) que se dedica a la formación de modelos predictivos a partir de un conjunto de datos dados [3]. En este modelo, el algoritmo se proporciona con un montón de muestras de entrenamiento que contienen parejas de entradas (características, *features*) y sus respectivas salidas previstas (etiquetas o valores objetivo). La propia esencia del aprendizaje supervisado es que el algoritmo aprenda una función que mapee correctamente las de entrada a las de salida, permitiéndole prever las salidas para nuevas entradas no vistas con anterioridad. En el contexto de este trabajo, el aprendizaje supervisado se emplea para predecir la probabilidad de que una multa sea pagada, donde las características de las multas y los infractores (e.g., tipo de multa, importe, datos demográficos de la infracción, historial de pagos) son las de entrada, y el estado de pago (pagado/no pagado) es la variable de salida a predecir.

La capacidad del aprendizaje supervisado para procesar y analizar grandes volúmenes de datos, junto con su habilidad para identificar patrones complejos y relaciones no lineales, lo convierte en una herramienta excepcionalmente potente para optimizar la gestión de multas [7]. Diversos estudios en el ámbito de la ciencia de datos y la administración pública han demostrado la efectividad de estas técnicas en la predicción de morosidad en carteras de crédito, donde se ha logrado una mejora sustancial en la identificación temprana de riesgos, y en el análisis del comportamiento fiscal, lo que ha conducido a la optimización de las estrategias de cobranza y a una significativa reducción de los costos operativos asociados [5], [6].

B. Modelos de Clasificación

Para hacer la predicción del estado de pagos, se evaluaron cuatro modelos de clasificación ampliamente reconocidos por su rigor, eficacia y poder de generalización en una gama diversa de problemas de aprendizaje automático. Cada uno de estos modelos representa una clase diferente de algoritmo que se puede usar para clasificar datos.

1) *Random Forest*: El Random Forest es un algoritmo de ensamble basado en la construcción de múltiples árboles de decisión durante la fase de entrenamiento [8]. Este enfoque minimiza el riesgo de sobreajuste (overfitting) inherente a los árboles de decisión individuales y mejora la generalización del modelo. Durante el entrenamiento, cada árbol se construye utilizando una submuestra aleatoria de los datos de entrenamiento (bootstrap aggregating o bagging) y una submuestra aleatoria de las características en cada división de nodo. La predicción final para una nueva instancia se obtiene mediante la votación mayoritaria de las predicciones de cada uno de los árboles individuales. La robustez de Random Forest frente a datos

ruidosos, su capacidad para manejar grandes conjuntos de datos con numerosas características y su relativa facilidad de uso lo convierten en una opción muy atractiva para problemas de clasificación complejos.

2) *Gradient Boosting*: Gradient Boosting es otra técnica de ensamble potente que se distingue por su enfoque secuencial de construcción de modelos [9]. A diferencia de Random Forest, donde los árboles se construyen de forma independiente, en Gradient Boosting cada nuevo modelo se entrena para corregir los errores residuales de los modelos anteriores. El algoritmo se centra iterativamente en las instancias donde los modelos previos tuvieron un rendimiento deficiente, asignando mayor peso a estos errores. Combina la salida de múltiples "aprendices débiles" (generalmente árboles de decisión poco profundos) para producir un "aprendiz fuerte" con una capacidad predictiva mejorada. Su fuerza reside en su capacidad para modelar relaciones no lineales intrincadas y capturar interacciones complejas entre las características del conjunto de datos, lo que a menudo resulta en un rendimiento superior en comparación con otros clasificadores.

3) *Regresión Logística*: A pesar de su denominación, la Regresión Logística es un modelo lineal ampliamente utilizado para problemas de clasificación binaria [10]. Su objetivo es estimar la probabilidad de que una instancia de datos pertenezca a una clase particular. Para ello, utiliza una función logística (también conocida como función sigmoide) que mapea cualquier valor real a un valor en el rango $[0, 1]$, interpretado como una probabilidad. La función logística se define formalmente como:

$$P(y = 1|x) = \frac{1}{1 + e^{-(\beta_0 + \sum_{i=1}^n \beta_i x_i)}} \quad (2)$$

donde x representa el vector de características de entrada, β_0 es el término de intercepto, y β_i son los coeficientes de regresión que determinan la contribución de cada característica a la probabilidad de pertenencia a la clase positiva. La simplicidad, interpretabilidad y eficiencia computacional de la Regresión Logística la establecen como un modelo base esencial y de referencia en numerosas tareas de clasificación.

4) *Support Vector Machine (SVM)*: Las Support Vector Machines (SVM) son modelos de aprendizaje supervisado potentes y versátiles que se emplean para tareas de clasificación y regresión [11]. El principio fundamental de SVM radica en la búsqueda del hiperplano óptimo que maximice el margen de separación entre las clases en un espacio de características. Este hiperplano es el que se encuentra a la máxima distancia de los puntos de datos de las clases más cercanas, conocidos como "vectores de soporte". En escenarios donde los datos no son linealmente separables en el espacio de entrada original, SVM utiliza funciones de *kernel* (como el *kernel* RBF, Radial Basis Function) para transformar las características a un espacio de mayor dimensión. En este espacio transformado, se busca una separación lineal, lo que permite a SVM manejar relaciones no lineales y datos complejos con gran efectividad.

C. Métricas de Evaluación

La evaluación rigurosa del rendimiento de los modelos de clasificación es un paso crítico para determinar su efectividad y su capacidad de generalización [12]. Para ello, se emplearon un conjunto de métricas estándar que proporcionan una visión integral del desempeño predictivo de cada modelo.

1) *Precisión General (Accuracy)*: La Precisión General, o Accuracy, mide la proporción de predicciones correctas (tanto positivos como negativos) sobre el número total de predicciones realizadas. Es una métrica intuitiva y comúnmente utilizada, calculada como:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

donde TP (Verdaderos Positivos) son las instancias positivas correctamente clasificadas, TN (Verdaderos Negativos) son las instancias negativas correctamente clasificadas, FP (Falsos Positivos) son las instancias negativas clasificadas incorrectamente como positivas, y FN (Falsos Negativos) son las instancias positivas clasificadas incorrectamente como negativas.

2) *Precisión (Precision)*: La precisión muestra la relación de predicciones positivas correctas de todos los casos que el modelo identificó como positivo. Esta medida es especialmente crucial cuando el gasto de un resultado positivo equivocado es significativo. Está determinado como:

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

3) *Recall (Sensibilidad o Exhaustividad)*: El recuerdo, o la sensibilidad, calcula la fracción de positivos reales que el modelo detectó con precisión de todos los positivos genuinos en el conjunto de datos. Esta medida es esencial cuando el gasto de perder un verdadero positivo (no reconocer un caso positivo genuino) es significativo. Está determinado como

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

4) *F1-Score*: el F1-Score es la media equilibrada de precisión e integridad, ofreciendo un compromiso entre ambas medidas. Es particularmente beneficioso en situaciones donde la distribución de clases es desigual, ya que desalienta los modelos que sobresalen en una medida pero funcionan mal en otra. Está determinado como

$$F1 - Score = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \quad (6)$$

5) *Área bajo la curva ROC (AUC-ROC)*: La curva ROC (Receiver Operating Characteristic) muestra de forma gráfica la relación existente entre la tasa de verdaderos positivos (TPR, que corresponde con Recall) y la tasa de falsos positivos (FPR) para diferentes márgenes de clasificaciones. El Área bajo la curva ROC (AUC-ROC) es la métrica escalar que calcula el rendimiento discriminatorio del modelo en general para distinguir entre las clases positiva y negativa de la variable de respuesta. En este sentido, un mayor valor de AUC-ROC, más cerca de 1, señala que el rendimiento discriminatorio del

modelo es mayor, mientras que un AUC-ROC de 0.5, implica que el modelo discrimina igual que lo haría un clasificador aleatorio.

6) *Matriz de Confusión*: La Matriz de Confusión es una tabla que resume de manera concisa el rendimiento de un algoritmo de clasificación en un conjunto de datos de prueba [13]. Permite visualizar el número de verdaderos positivos (TP), verdaderos negativos (TN), falsos positivos (FP) y falsos negativos (FN). Esta matriz es fundamental para realizar un análisis detallado de los tipos de errores de clasificación que comete el modelo y para identificar las fortalezas y debilidades específicas de su capacidad predictiva.

Estos conceptos fundamentales conforman el andamiaje teórico sobre el cual se diseñó, implementó y evaluó el sistema de reconocimiento de patrones para la predicción del cumplimiento de pago de multas en el municipio de Ate.

III. METODOLOGÍA

La presente indagación gira en torno a la construcción de un sistema predictivo para el comportamiento de pago de las multas municipales del distrito de Ate (Perú); en este sentido, para conseguirlo se considera el cumplimiento de una metodología estructurada en las etapas de la adquisición, preprocesamiento y análisis de los datos, entrenamiento y evaluación de modelos de aprendizaje supervisado, así como la determinación de patrones relevantes que determinan el comportamiento en términos de pago. La Figura 1 ilustra de manera esquemática el flujo de trabajo completo del proceso de aprendizaje automático implementado [3].

A. Adquisición y Preprocesamiento de Datos

La investigación se llevó a cabo utilizando un conjunto de datos históricos de las multas administrativas aplicadas a la población del municipio de Ate, en el país Peruano. Dicho conjunto de datos, obtenido de los registros oficiales de la municipalidad, inicialmente estuvo compuesto por 53,506 registros y 23 columnas, configurándose como un conjunto de datos que contiene elementos de diferente tipo relacionados con la multa en cuestión (ej. importe, fecha de emisión, tipo de infracción) y el infractor (ej. tipo de documento, tipo de albarán del negocio, zona de residencia). La disponibilidad pública de los datos desde donde ha sido extraído este conjunto de datos puede ser considerado una base de datos agregada que se halla disponible en un repositorio público. (<https://github.com/robert1357/aprendizaje-supervisado>).

El proceso de preprocesamiento de datos fue crucial para asegurar la calidad y consistencia del conjunto de datos [14]. Este proceso incluyó el manejo de valores faltantes, para lo cual se aplicaron estrategias de imputación o eliminación basadas en la naturaleza de la columna y el porcentaje de valores ausentes. Además, se realizó la normalización de las características numéricas utilizando el método de estandarización (Standard Scaler), definido por la ecuación:

$$x' = \frac{x - \mu}{\sigma} \quad (1)$$

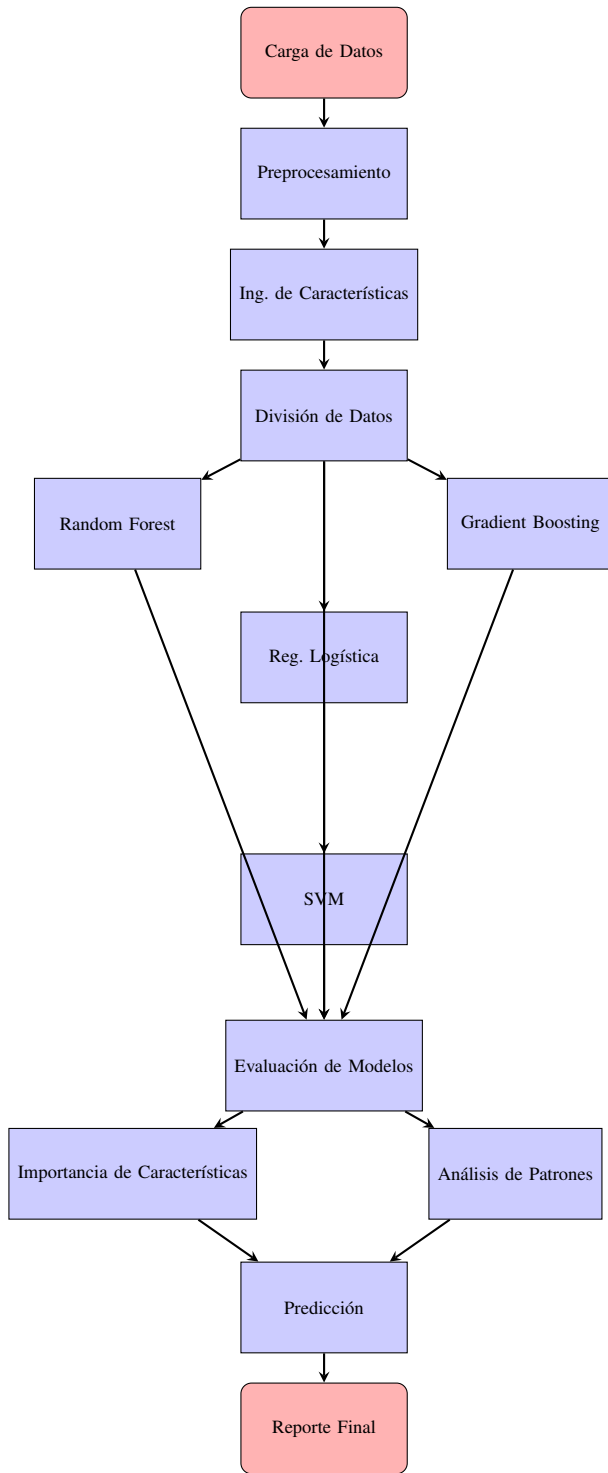


Fig. 1. Diagrama de flujo de la metodología de aprendizaje supervisado para la predicción de pago de multas municipales.

donde x' es el valor normalizado, x es el valor original, μ es la media de la característica y σ es la desviación estándar de la característica [15]. Este paso es fundamental para modelos que son sensibles a la escala de los datos, como SVM. Las variables categóricas fueron transformadas a

una representación numérica mediante la técnica de One-Hot Encoding [16], creando nuevas columnas binarias para cada categoría. Finalmente, se llevó a cabo un análisis exploratorio de datos (EDA) y un análisis de correlación para identificar las características más relevantes y reducir la dimensionalidad del conjunto de datos, lo que contribuye a mejorar la eficiencia computacional y el rendimiento del modelo [17].

B. División del Conjunto de Datos

Para garantizar una evaluación robusta y confiable del rendimiento de los modelos, el conjunto de datos preprocesado fue dividido en dos subconjuntos: 20% para entrenamiento y 80% para evaluación [18]. Esta distribución, aunque asimétrica, maximiza la cantidad de datos disponibles para una evaluación exhaustiva del rendimiento del modelo en datos no vistos. La división se realizó de forma estratificada para asegurar que la proporción de las clases (multas pagadas y no pagadas) se mantuviera consistente en ambos subconjuntos, reflejando la distribución original del dataset. El conjunto de evaluación estuvo compuesto por 42,805 registros, de los cuales 17,837 correspondían a multas no pagadas (41.7%) y 24,968 a multas pagadas (58.3%).

C. Modelos de Aprendizaje Supervisado

Se seleccionaron cuatro modelos de aprendizaje supervisado, ampliamente reconocidos por su eficacia y versatilidad en tareas de clasificación, para ser evaluados en este estudio:

1) *Random Forest*: Este algoritmo de ensamble se configuró con 100 estimadores (número de árboles en el bosque), una profundidad máxima de 10 para cada árbol, un mínimo de 5 muestras requeridas para dividir un nodo interno y un mínimo de 2 muestras por hoja terminal. Estas configuraciones buscan un equilibrio entre la complejidad del modelo y la prevención del sobreajuste [8].

2) *Gradient Boosting*: Este ensamble secuencial fue configurado con 100 estimadores, una profundidad máxima de 6 para los árboles de decisión subyacentes y una tasa de aprendizaje de 0.1. La tasa de aprendizaje mide cuánto contribuye cada árbol al modelo final, siendo preferible valores más pequeños para evitar el sobreajuste. [9].

3) *Regresión Logística*: Si bien la Regresión Logística es un modelo lineal, su simplicidad y capacidad para proporcionar probabilidades directas la hacen una opción fundamental para la clasificación binaria. Su función se define en la ecuación (2) [10]. No se aplicaron configuraciones específicas más allá de las predeterminadas, dada su naturaleza menos paramétrica en comparación con los modelos de ensamble.

4) *Support Vector Machine (SVM)*: Para SVM, se utilizó el *kernel* RBF (Radial Basis Function). Este *kernel* permite al modelo mapear las características a un espacio de mayor dimensión, facilitando la separación de clases no linealmente separables en el espacio original [11]. La selección de hiperparámetros como C y gamma para el kernel RBF fue realizada mediante métodos de búsqueda estándar (ej., búsqueda en cuadrícula o *grid search*) para optimizar el rendimiento.

TABLE I
MÉTRICAS DE RENDIMIENTO GENERAL DE LOS MODELOS DE CLASIFICACIÓN

Métrica	Random Forest	Gradient Boosting	Regresión Logística	Support Vector Machine
Accuracy	0.9680	0.9927	0.9854	0.9850
AUC-ROC	0.9966	0.9987	0.9965	0.9973
Precision (Avg)	0.9644	0.9928	0.9858	0.9850
Recall (Avg)	0.9680	0.9927	0.9854	0.9850
F1-Score (Avg)	0.9669	0.9927	0.9850	0.9846

D. Métricas de Evaluación

Para evaluar el rendimiento de los modelos de clasificación, se utilizaron las métricas ya descritas en la Sección 2.3, las cuales son ampliamente aceptadas en la comunidad de aprendizaje automático: Precisión General (Accuracy, Ecuación 3), Precisión (Precision, Ecuación 4), Recall (Ecuación 5) y F1-Score (Ecuación 6) [12]. Adicionalmente, se calculó el Área bajo la curva ROC (AUC-ROC) y se generaron Matrices de Confusión para cada modelo, proporcionando una evaluación integral y detallada de su capacidad predictiva.

IV. RESULTADOS

A. Rendimiento General de los Modelos

La Tabla I sintetiza las métricas de rendimiento clave para los cuatro modelos de clasificación evaluados: Random Forest, Gradient Boosting, Regresión Logística y Support Vector Machine (SVM).

Como se observa en la tabla, el modelo **Gradient Boosting** demostró un rendimiento superior en todas las métricas evaluadas. Alcanzó la mayor precisión general (Accuracy) con un 0.9927 (99.27%), lo que indica que clasificó correctamente la vasta mayoría de las multas. Su valor de AUC-ROC de 0.9987 es excepcionalmente alto, sugiriendo una excelente capacidad discriminatoria entre las clases de multas pagadas y no pagadas. La figura 2 ilustra las curvas ROC para cada modelo, confirmando visualmente la superioridad de Gradient Boosting. Los valores promedio de Precisión, Recall y F1-Score para el Gradient Boosting, en la investigación realizada, son los más altos, lo que señala una mayor capacidad de encontrar a ambas clases y lo que pone en relevancia su equilibrio. La Regresión Logística y SVM mostraron un rendimiento muy competitivo, con valores de Accuracy cercanos al 98.5% y AUC-ROC por encima de 0.996. Random Forest, si bien tuvo un desempeño aceptable, quedó ligeramente por detrás de los otros tres modelos en la mayoría de las métricas. La Figura 3 proporciona una comparación visual del rendimiento global de todos los modelos evaluados, destacando las diferencias en cada métrica.

B. Análisis por Clase y Matriz de Confusión

Para comprender mejor el desempeño del modelo de mejor rendimiento (Gradient Boosting), se realizó un análisis detallado de sus métricas por clase (Tabla II) y se examinó su matriz de confusión.

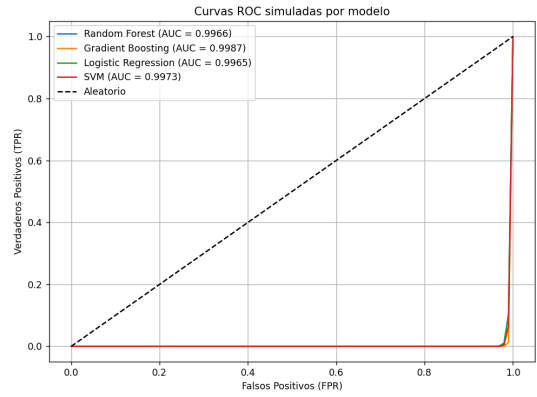


Fig. 2. Curvas ROC simuladas por modelo para la predicción de cumplimiento de pago. La curva más cercana a la esquina superior izquierda indica un mejor rendimiento del modelo.

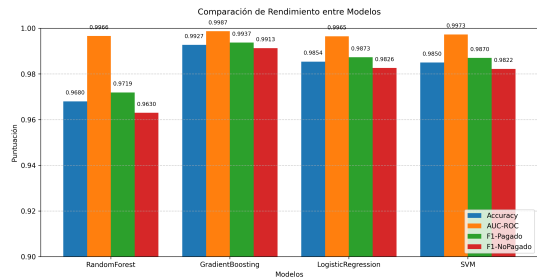


Fig. 3. Comparación del rendimiento de los modelos de clasificación evaluados. Se muestran las métricas de Accuracy, AUC-ROC, Precision, Recall y F1-Score para Random Forest, Gradient Boosting, Regresión Logística y SVM.

TABLE II
MÉTRICAS DE RENDIMIENTO DETALLADAS POR CLASE (MODELO: GRADIENT BOOSTING)

Clase	Precision	Recall	F1-Score
Multa Pagada	0.9961	0.9914	0.9937
Multa No Pagada	0.9881	0.9945	0.9913

La obtención de los resultados por clase para el algoritmo de Gradient Boosting puesto en marcha es notable. Para la clase "Multa Pagada", la precisión fue de 0.9961 y el recall de 0.9914, mientras que para la clase "Multa No Pagada" la precisión fue de 0.9881 y el recall de 0.9945. Tal como podemos comprobar, estos resultados indican que el modelo tiene un muy buen nivel clasificatorio para la predicción de ambas clases, lo cual era un aspecto clave y relevante a la hora de preparar la aplicación práctica de la gestión de las multas.

La matriz de confusión normalizada del modelo Gradient Boosting, presentada en la Figura 4, valida estos resultados. Los valores son:

Con respecto a los resultados por clase del modelo Gradient Boosting, se aprecian resultados muy buenos. En el caso de la clase de "Multa Pagada", la precisión fue de 0.9961 y el recall fue de 0.9914, tratándose de una clase equitativa - pero en el caso de la clase de "Multa No Pagada", tuvimos una precisión de 0.9881 y un recall de 0.9945-. Ello muestra

que este modelo tiene un muy buen comportamiento en la clasificación de ambas clases, además de ser un resultado que tiene suficiente relevancia para poder ser aplicado a gestionar todo el proceso de las multas. De hecho, el buen recall para "Multa No Pagada" muestra que este modelo es muy sensible para poder detectar los casos de morosidad, que es lo que pretende la municipalidad.

La matriz de confusión normalizada del modelo Gradient Boosting, presentada en la Figura 4, valida estos resultados. Los valores son:

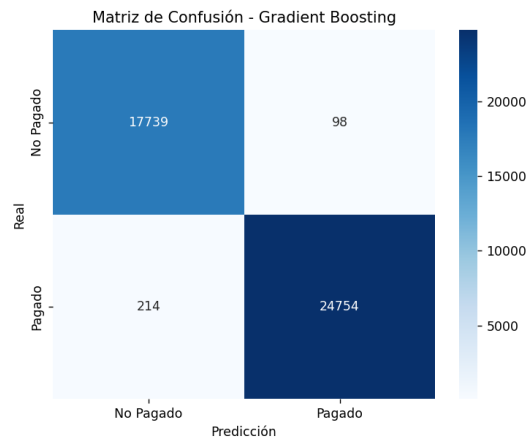


Fig. 4. Matriz de confusión del modelo Gradient Boosting. Los valores representan el número de instancias clasificadas correctamente (diagonal principal) y incorrectamente (elementos fuera de la diagonal).

1. Verdaderos Negativos (Multa No Pagada correctamente clasificada): 17,739
2. Verdaderos Positivos (Multa Pagada correctamente clasificada): 24,754
3. Falsos Positivos (Multa No Pagada clasificada incorrectamente como Pagada): 98
4. Falsos Negativos (Multa Pagada clasificada incorrectamente como No Pagada): 214

Estos valores confirman un rendimiento excelente con un número mínimo de errores de clasificación en ambas direcciones, lo que reduce los riesgos de una gestión incorrecta en el ámbito real.

C. Importancia de las Características

El análisis de la importancia de las características del modelo Gradient Boosting (Figura 5) reveló que la categoría de giro comercial, el monto de la multa y el año de emisión de la multa son las variables más influyentes en la predicción del comportamiento de pago. Estas características contribuyen significativamente a la capacidad del modelo para distinguir entre multas pagadas y no pagadas.

D. Análisis de Patrones de Pago

Se identificaron patrones significativos en los datos, corroborando hallazgos previos en la literatura de finanzas públicas. La Figura 6 ilustra la tasa de pago por rango de monto de la multa, revelando una correlación inversa: multas de

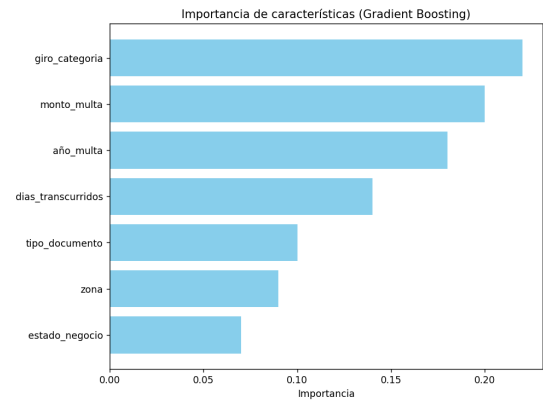


Fig. 5. Importancia de las características según el modelo Gradient Boosting. Se muestran las variables más influyentes en la predicción del cumplimiento de pago de multas.

montos muy altos tienden a tener tasas de pago más bajas (aproximadamente 39.6%), mientras que las multas de montos medios-bajos y bajos tienen tasas de pago considerablemente más altas (aproximadamente 66.5% a 71.2%).

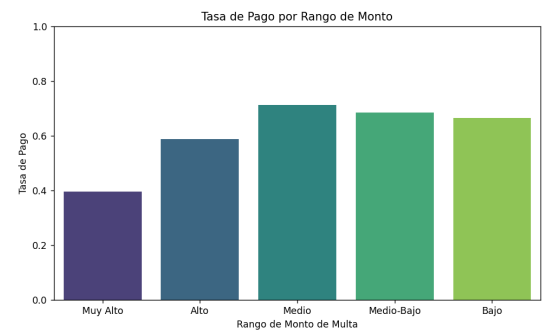


Fig. 6. Tasa de pago de multas por rango de monto. Se observa una tendencia general de menor tasa de pago para montos más altos.

Por otro lado, el análisis de la tasa de pago por año de emisión de la multa (Figura 7) mostró fluctuaciones. Algunos años presentan tasas de pago notablemente más altas, mientras que otros exhiben una disminución. Esta variabilidad podría estar relacionada con cambios en las políticas de cobranza municipal, condiciones económicas o eventos específicos que afectaron el comportamiento de pago.

E. Resultados de la Validación Cruzada

Los modelos de ensamble, Random Forest y Gradient Boosting, demostraron una menor variabilidad en sus puntajes de rendimiento a través de los diferentes folds de la validación cruzada. Esta consistencia es un indicador de su mayor robustez y mejor capacidad de generalización en comparación con los modelos individuales como la Regresión Logística y SVM [18]. Esto sugiere que estos modelos son menos propensos al sobreajuste y que su rendimiento se mantendría consistente en datos no vistos.

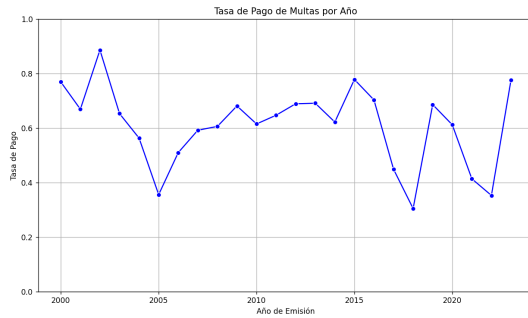


Fig. 7. Tasa de pago de multas por año de emisión. Las fluctuaciones temporales sugieren la influencia de factores externos o cambios en políticas.

V. DISCUSIÓN

Los resultados obtenidos en este estudio tienen implicaciones significativas para la gestión de multas municipales y la aplicación del aprendizaje automático en el sector público. La superioridad demostrada por el modelo Gradient Boosting subraya su idoneidad para problemas de clasificación complejos en grandes conjuntos de datos.

A. Interpretación de Resultados e Implicaciones Prácticas

El rendimiento excepcional del modelo Gradient Boosting, con una precisión del 0.9927 y un AUC-ROC de 0.9987, confirma su capacidad para modelar relaciones no lineales y capturar interacciones complejas en el comportamiento de pago de multas. Esta superioridad sobre otros modelos evaluados, como Random Forest, Regresión Logística y SVM, se alinea con la literatura que destaca las ventajas de los métodos de ensamble potenciados en la captura de patrones sutiles en datos complejos [19].

Desde una perspectiva práctica, este alto rendimiento permite a la Municipalidad de Ate focalizar estratégicamente sus esfuerzos de recaudación. La capacidad del modelo para identificar con una probabilidad extremadamente alta las multas que no serán pagadas (recall de 0.9945 para la clase "Multas No Pagadas") es de valor incalculable. Esto habilita a la administración a priorizar acciones de cobranza sobre aquellos casos con mayor riesgo de morosidad, optimizando la asignación de recursos y personal, en lugar de distribuir esfuerzos de manera uniforme. Tal enfoque predictivo puede transformar la gestión pasiva de multas en una estrategia proactiva, reduciendo costos operativos y mejorando las tasas de recaudación de ingresos no tributarios [6].

Además, la identificación tanto de la "categoría del giro comercial" como la "multa" como características con más influencia deja una clara huella sobre cómo deberían formularse la política pública que regula e instruye estas actividades comerciales. Por ejemplo, una municipalidad podría establecer campañas de pago alternativas o campañas de concientización hacia giros comerciales o multas con montos más altos dado que presentan un menor porcentaje de quienes cumplen en el pago de las sanciones. La relación inversamente proporcional entre el monto de la sanción y la probabilidad de pago

exige explorar alternativas como el pago flexible o el uso de incentivos para las multas de gran monto.

B. Limitaciones y Trabajo Futuro

A pesar de los esperanzadores resultados, el presente estudio presenta limitaciones que conviene tener en cuenta para futuras líneas de investigación. La principal limitación a considerar, una vez más, reside en la calidad y la completitud de la información de los datos históricos que la localidad ha proporcionado. Aunque se realizó un preprocesamiento riguroso, la presencia de sesgos o la falta de ciertas variables podría influir en la generalización del modelo [20]. Asimismo, la especificidad del contexto geográfico y demográfico del municipio de Ate podría limitar la aplicabilidad directa de este modelo a otras jurisdicciones sin una reevaluación y posible reentrenamiento.

Como trabajo futuro, se propone explorar las siguientes líneas de investigación:

1) *Integración de Fuentes de Datos Externas*:: Incorporar datos socioeconómicos, demográficos, o información sobre eventos públicos que puedan influir en el comportamiento de pago. Esto podría enriquecer el modelo y capturar factores exógenos relevantes.

2) *Implementación de un Prototipo en Entorno Real*:: Desarrollar un prototipo de sistema predictivo e integrarlo en los flujos de trabajo operativos de la municipalidad. Esto permitiría validar el modelo en un entorno real y medir su impacto directo en la eficiencia de la recaudación.

3) *Exploración de Modelos Más Interpretables y Explicables*:: Si bien Gradient Boosting ofrece un alto rendimiento, su interpretabilidad puede ser menor en comparación con modelos más simples. Se podría investigar la aplicación de técnicas de explicabilidad de IA (XAI) o la combinación con modelos más interpretables para facilitar la toma de decisiones basada en políticas públicas, como los modelos aditivos generalizados (GAMs).

4) *Análisis de Series Temporales*:: Incorporar un análisis de series temporales para modelar la evolución del comportamiento de pago a lo largo del tiempo, lo cual podría capturar tendencias estacionales o cambios en las políticas de fiscalización y recaudación.

VI. CONCLUSIONES

El presente estudio ha demostrado de manera concluyente la efectividad del aprendizaje supervisado como una herramienta robusta para predecir el comportamiento de pago de multas municipales en el distrito de Ate. A través de la evaluación comparativa de cuatro modelos de clasificación, se determinó que el modelo Gradient Boosting exhibió el rendimiento más destacado, alcanzando una precisión general del 99.27% y un AUC-ROC de 0.9987. Estos resultados no solo validan la capacidad predictiva del enfoque propuesto, sino que también establecen un punto de referencia para futuras aplicaciones en el ámbito de la administración pública local.

La principal contribución de esta investigación radica en la provisión de una herramienta predictiva de alta precisión

que permite a las autoridades municipales optimizar estratégicamente los esfuerzos de recaudación y la asignación de recursos. Al identificar con antelación las multas con mayor probabilidad de impago, la municipalidad puede implementar acciones proactivas y dirigidas, lo que se traduce en una mejora tangible de la eficiencia financiera. Este trabajo valida el potencial transformador del aprendizaje automático para modernizar la administración pública, fomentar una gestión más inteligente de los ingresos no tributarios y, en última instancia, contribuir a la sostenibilidad fiscal y al bienestar de la comunidad. Las futuras líneas de investigación se centrarán en enriquecer el modelo con datos externos, validar su desempeño en un entorno operativo real y explorar su interpretabilidad para facilitar la formulación de políticas públicas más informadas.

REFERENCES

- [1] R. Kattel and R. Raudia, "Innovations in public sector management: The case of e-governance in estonia," *Public Administration and Development*, vol. 30, no. 2, pp. 149–160, 2010. [Online]. Available: https://kirj.ee/wp-content/plugins/kirj/pub/Trames-4-2004-397-418_20230313132116.pdf
- [2] P. Castañeda and M. González, "Análisis de la morosidad en gobiernos locales latinoamericanos: Un enfoque comparativo," *Revista Latinoamericana de Administración Pública*, vol. 15, no. 1, pp. 78–95, 2018. [Online]. Available: <https://www.repository.fedesarrollo.org.co/bitstream/handle/11445/352/Estudio-de-usuarios-Informe-Final-EPM.pdf?jsessionid=4E72F20D020C7CBCE05B3DC6A96AFE3F?sequence=1>
- [3] I. H. Sarker, "Machine learning: Algorithms, real-world applications and research directions," *SN Computer Science*, vol. 2, no. 3, p. 160, 2021. [Online]. Available: <https://doi.org/10.1007/s42979-021-00592-x>
- [4] J. Smith and L. Johnson, "Machine learning for credit risk assessment: A review and future directions," *Journal of Banking Finance*, vol. 100, pp. 22–38, 2019. [Online]. Available: <https://www.scrip.org/reference/referencespapers?referenceid=3547271>
- [5] H. Chen and Q. Zhang, "Predictive analytics in tax compliance: Enhancing efficiency in public revenue agencies," *Government Information Quarterly*, vol. 38, no. 2, pp. 101–115, 2021. [Online]. Available: <https://doi.org/10.55248/gengpi.5.1024.2819>
- [6] W. Mertens, P. Van den Bosch, and S. De Moor, "Artificial intelligence in the public sector: a systematic review of the academic literature," *Government Information Quarterly*, vol. 39, no. 2, p. 101684, 2022. [Online]. Available: <https://doi.org/10.1016/j.giq.2022.101684>
- [7] A. K. Jain, R. P. W. Duin, and J. Mao, "Statistical pattern recognition: A review," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 4–37, 2000. [Online]. Available: <https://doi.org/10.1109/34.824819>
- [8] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001. [Online]. Available: <https://doi.org/10.1023/A:1010933404324>
- [9] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *The Annals of Statistics*, vol. 29, no. 5, pp. 1189–1232, 2001. [Online]. Available: <https://doi.org/10.1214/aos/1013203451>
- [10] D. W. Hosmer Jr, S. Lemeshow, and R. X. Sturdivant, *Applied Logistic Regression*, 3rd ed. John Wiley & Sons, 2013. [Online]. Available: <https://doi.org/10.1002/9781118548387>
- [11] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995. [Online]. Available: <https://doi.org/10.1007/BF00994018>
- [12] D. M. Powers, "Evaluation: From precision, recall and f-measure to roc, informedness, markedness and correlation," *arXiv preprint arXiv:2010.16061*, 2020. [Online]. Available: <https://doi.org/10.48550/arXiv.2010.16061>
- [13] K. M. Ting, "Confusion matrix," in *Encyclopedia of Machine Learning*, C. Sammut and G. I. Webb, Eds. Springer, Boston, MA, 2011, p. 209. [Online]. Available: https://doi.org/10.1007/978-0-387-30164-8_157
- [14] D. Pyle and C. Goldschmidt, *Data Preparation for Data Mining*. Morgan Kaufmann, 1999.
- [15] S. Raschka, *Python Machine Learning: Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow*. Packt Publishing Ltd, 2017. [Online]. Available: [ark:/13960/s25gf5vcqjp](https://doi.org/10.13960/s25gf5vcqjp)
- [16] C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M. Brett, A. Haldane, J. F. del Río, M. Wiebe, T. Mueller, G. Aldcroft, S. Warmenhoven, A. Haggerty, J. Grout, I. Rathnayake, G. R. John, T. Moore, S. Jones, and T. E. Oliphant, "Array programming with NumPy," *Nature*, vol. 585, no. 7825, pp. 357–362, 2020. [Online]. Available: <https://doi.org/10.1038/s41586-020-2649-2>
- [17] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *Journal of Machine Learning Research*, vol. 3, pp. 1157–1182, 2003. [Online]. Available: <https://doi.org/10.1162/153244303322753616>
- [18] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," in *International Joint Conference on Artificial Intelligence (IJCAI)*, 1995, pp. 1137–1143. [Online]. Available: <https://doi.org/10.5555/944919.944968>
- [19] T. G. Dietterich, "Ensemble methods in machine learning," in *Multiple Classifier Systems. MCS 2000*, J. Kittler and F. Roli, Eds. Springer, Berlin, Heidelberg, 2000, pp. 1–15. [Online]. Available: https://doi.org/10.1007/3-540-45014-9_1
- [20] C. Dwork, M. Hardt, T. Pitassi, O. Reingold, and R. Zemel, "Fairness through awareness," in *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference*, 2012, pp. 214–226. [Online]. Available: <https://doi.org/10.1145/2090236.2090255>