Aprendizaje automático y grandes volúmenes de datos en la lucha contra el fraude financiero para proteger la integridad de las transacciones

Giset Yamile Báez Jiménez

Jonh Jairo Serrano Navarro

Asesor

Sixyel Jeyson Castaneda Coronado

Universidad Nacional Abierta y a Distancia UNAD

Escuela de Ciencias Básicas, Tecnología e Ingeniería ECBTI

Especialización en Ciencia de Datos y Analítica

2025

Resumen

Este estudio realiza una revisión sobre el impacto del aprendizaje automático (Machine Learning, ML) y el análisis de Grandes Volúmenes de Datos (Big Data) en la optimización de los sistemas de detección de fraudes financieros. Se identifican y caracterizan los principales modelos de ML utilizados, con el fin de evaluar su desempeño en cuanto a los términos de precisión y la capacidad para manejar el desbalance de clases inherente a los datos de fraude (Jones Ortiz & Guzmán–Seraquive, 2022). Los resultados de la comparativa indican que modelos avanzados como Potenciación por Gradiente (Gradient Boosting), (XGBoost) y Redes Neuronales alcanzan métricas de efectividad superiores al 95 % en la curva ROC, superando los enfoques tradicionales (Alvarado Zabala et al., 2022). Finalmente, se menciona lineamientos y buenas prácticas para la implementación, abordando desafíos cruciales como la interpretabilidad algorítmica y la detección en tiempo real, los cuales ayudaran a la integridad del sistema financiero colombiano, en cumplimiento del SARLAFT.

Palabras Clave: Aprendizaje automático, grandes volúmenes de datos, fraude financiero, XGBoost, SARLAFT.

Abstract

In the following review study, the impact of Machine Learning (ML) and Big Data analysis on improving financial fraud detection is examined. The main ML models are identified and characterized, with the aim of evaluating their performance regarding accuracy and the management required to address the class imbalance in fraud data. The comparison of results demonstrates that advanced algorithms such as XGBoost, Gradient Boosting, and Neural Networks achieve an effectiveness superior to 95% on the ROC curve, surpassing traditional approaches. Finally, a comparative table is created for a better understanding of the models and the real-time detection they suggest, contributing to the integrity of the Colombian financial system and compliance with SARLAFT.

Keywords: Machine Learning, Big Data, Financial Fraud, Fraud Detection.

Tabla de Contenido

Lista de Tablas

Lista de Figuras

Introducción

Descripción del Problema

En la era digital, el crecimiento exponencial de las transacciones financieras electrónicas ha desencadenado un aumento alarmante en los casos de fraude. La detección y prevención efectiva de fraudes en estas transacciones se ha vuelto esencial para mantener la estabilidad y la integridad del sistema financiero. Sin embargo, como señala (Ramírez Palma, 2023), a pesar de los avances tecnológicos, la detección oportuna y precisa de estas actividades fraudulentas sigue siendo un desafío persistente, lo que pone en peligro la confianza de los clientes en las transacciones.

La necesidad de investigar y abordar este problema radica en la protección de la integridad del sistema financiero y la confianza del público. Como destaca (Alvarado Zabala et al., 2022) los fraudes financieros no solo representan pérdidas económicas considerables para las instituciones y los clientes, sino que también pueden socavar la credibilidad del sistema bancario en su conjunto. Por lo tanto, es crucial desarrollar métodos eficaces para detectar y prevenir estas actividades ilícitas.

A pesar de los esfuerzos previos, la investigación sobre la detección de fraudes en transacciones financieras aún presenta importantes lagunas. Por ejemplo, según (Giraldo et al., 2020), mientras algunas investigaciones se centran en técnicas tradicionales de detección de fraudes, otras exploran enfoques más avanzados basados en inteligencia artificial y grandes volúmenes de datos. Sin embargo, existe una falta de consenso sobre la eficacia comparativa de estas técnicas en entornos financieros específicos, como lo señala (Jones Ortiz & Guzmán–Seraquive, 2022)

En este contexto, es esencial reconocer la urgencia de abordar los desafíos persistentes en la detección de fraudes financieros. Como resalta (Francés Mondero, 2020), la aplicación efectiva de técnicas de aprendizaje automático en el sistema financiero puede proporcionar *hallazgos* valiosos y detallados sobre patrones de comportamiento fraudulento, permitiendo una respuesta más proactiva y precisa por parte de las instituciones financieras. Por otro lado, la investigación de (Irusta, 2023) destaca cómo el uso de grandes volúmenes de datos puede ser fundamental para identificar conexiones y tendencias ocultas en la información financiera, ofreciendo así una herramienta poderosa para prevenir y detectar actividades fraudulentas de manera más efectiva. Integrar estas perspectivas en la estrategia de detección de fraudes puede no solo fortalecer la capacidad de respuesta de las instituciones financieras, sino también mejorar la confianza del público en la seguridad y la integridad del sistema financiero.

Esta monografía se propone abordar las deficiencias existentes al realizar un análisis detallado de las técnicas de detección de fraudes en transacciones financieras. Se centrará en el uso de técnicas de aprendizaje automático y grandes volúmenes de datos, explorando su aplicabilidad y evaluando su eficacia en la detección de fraudes, como sugiere Gil (2023). Además, buscará proporcionar recomendaciones prácticas para mejorar los sistemas de detección de fraudes en las instituciones financieras, contribuyendo así a fortalecer la seguridad y la confianza en el sistema financiero.

En los últimos años, el fraude financiero ha mostrado un crecimiento sostenido tanto a nivel global como en América Latina. Diversos informes de entidades internacionales señalan que el incremento de los servicios digitales, la banca móvil y el comercio electrónico ha ampliado la superficie de riesgo, permitiendo que los delincuentes adopten métodos más sofisticados para manipular transacciones y suplantar identidades. En el caso colombiano, las

entidades financieras han reportado un aumento significativo en intentos de fraude asociados a transferencias electrónicas y compras en línea, lo que refleja la necesidad urgente de fortalecer los mecanismos de monitoreo y análisis. Este escenario confirma que la detección oportuna del fraude ya no puede depender únicamente de reglas estáticas, sino de herramientas capaces de aprender y adaptarse al comportamiento cambiante de los usuarios.

Pregunta de Investigación

¿Cómo pueden las instituciones financieras aprovechar las técnicas de aprendizaje automático y grandes volúmenes de datos para mejorar la detección y prevención de fraudes, manteniendo así la confianza del público en el sistema financiero?

Justificación

La implementación de modelos de Aprendizaje Automático (ML) ofrece una solución escalable y precisa para el desafío persistente del fraude financiero. A diferencia de los métodos tradicionales basados en reglas fijas y sin predicciones, el ML tiene la capacidad intrínseca de aprender patrones complejos y sutiles incrustados en los grandes volúmenes de datos transaccionales (Alvarado Zabala et al., 2022). Esto permite no solo identificar el mejorar la detención fraude en tiempo real, sino también lograr una tasa de falsos positivos significativamente menor en comparación con los sistemas heredados (Gutierrez Portela et al., 2023).

Esta monografía se justifica en la necesidad de consolidar el estado del arte en ML aplicado al fraude. Por lo cual se busca proporcionar un marco conceptual y comparativo que demuestre el valor agregado de técnicas como las Redes Neuronales sobre la regresión tradicional (Tustón Fuentes & Macías Arias, 2024) Este análisis comparativo es esencial para guiar a las instituciones financieras en la adopción de estrategia con el fin de mejorar la predicción por medio de la tecnología, logrando dar un paso que se alinea con las crecientes exigencias regulatorias nacionales e internacionales en materia de prevención de Lavado de Activos y Financiamiento del Terrorismo(Grupo de Acción Financiera Internacional (GAFI), 2024)

Asimismo, esta investigación cobra relevancia en el contexto actual, en el que las entidades financieras enfrentan un entorno regulatorio más estricto y una creciente presión por parte de los usuarios para garantizar la seguridad de sus transacciones. La integración de modelos predictivos no solo contribuye a reducir pérdidas económicas, sino que también mejora la experiencia del cliente al disminuir la cantidad de alertas erróneas y bloqueos injustificados.

En este sentido, el desarrollo de un análisis comparativo riguroso permitirá identificar cuáles modelos resultan más adecuados para el sector financiero colombiano, considerando sus particularidades operativas, regulatorias y tecnológicas.

Objetivos

Objetivo General

Analizar el impacto de la integración de técnicas de aprendizaje automático y grandes volúmenes de datos en la optimización de los sistemas de detección de fraudes financieros, mediante la evaluación comparativa de modelos.

Objetivos Específicos

Caracterizar los principales modelos de aprendizaje automático aplicados a la detección de fraudes financieros.

Comparar la precisión de los diferentes modelos estadísticos evaluando la efectividad de los enfoques actuales que integran aprendizaje automático y grandes volúmenes de datos.

Identificar buenas prácticas para la implementación de modelos de Aprendizaje Automático en el sector financiero, cumplimiento regulatorio SARLAFT.

Marco Yeórico

En la siguiente monografía se menciona sobre el uso de Aprendizaje Automático y Grandes Volúmenes de Datos en la lucha contra el fraude financiero, se fundamenta las diversas teorías y modelos desarrollados hasta momento en el ámbito de la informática, las finanzas y la regulación que aplica.

Fundamentos sobre la Detección de Fraude y Aprendizaje Automático

La base de esta investigación radica en la comprensión de la Teoría de Detección de Fraudes Financieros, abordada por (Alvarado Zabala et al., 2022), la cual proporciona una comprensión profunda de las técnicas de Aprendizaje Automático aplicadas específicamente en la detección de fraudes bancarios. Siendo la parte clave en línea análisis, (Alvarado Zabala et al., 2022; Jones Ortiz & Guzmán–Seraquive, 2022) contribuyen de manera detallada las técnicas de aprendizaje automático aplicadas en este ámbito, destacando la importancia de los modelos predictivos en la identificación de patrones anómalos en transacciones financieras.

El fraude financiero adopta diversas modalidades que evolucionan conforme avanzan los sistemas de pago y las herramientas tecnológicas. Entre los tipos más frecuentes se encuentra el fraude con tarjetas, que incluye transacciones no autorizadas, clonación y uso indebido de datos obtenidos mediante técnicas de ingeniería social. Otra modalidad relevante es la suplantación de identidad, en la cual el delincuente accede a servicios bancarios haciéndose pasar por el titular legítimo. También se presentan fraudes internos, cometidos por empleados o contratistas con acceso privilegiado a información sensible, y fraudes electrónicos que utilizan enlaces falsos, correos engañosos o aplicaciones fraudulentas para capturar información. Estas modalidades generan patrones de comportamiento que pueden ser detectados por modelos de aprendizaje

automático cuando se analizan de manera adecuada las variables de tiempo, geolocalización, frecuencia y monto de las transacciones.

SARLAFT: Entendido como el Sistema de Administración del Riesgo de Lavado de Activos y de la Financiación del Terrorismo, este sistema funciona mediante políticas, procedimientos y controles que permiten identificar, medir y mitigar riesgos asociados a operaciones ilícitas. Su integración con modelos predictivos y el análisis de Grandes Volúmenes de Datos contribuye significativamente a proteger la integridad de las transacciones financieras.

También define a los clientes, los productos, los canales de distribución y las jurisdicciones como las variables mínimas a ser contempladas en el análisis de riesgos, los cuales en conjunto conforman el concepto de "transacción" u "operación", dependiendo del sector en el que se realice, en el entendido que una operación, en todos los casos, será realizada por un cliente o usuario, mediante un producto vigente o activo, a través de un canal de distribución dispuesto por la entidad y en una jurisdicción específica en la que tenga presencia la entidad. (Superintendencia de Vigilancia y Seguridad Privada, 2022)

SARO: El Sistema de Administración del Riesgo Operativo también contribuye de forma importante ante riesgos financieros. De acuerdo con la Circular Externa 025 de 2020 de la Superintendencia Financiera de Colombia, todas las entidades vigiladas deben adoptar un SARO acorde con su tamaño, para identificar, medir, controlar y monitorear pérdidas operativas. Este sistema ayuda a mitigar riesgos financieros al reducir pérdidas inesperadas y fortalecer los controles internos. La adopción de Inteligencia Artificial en el sistema financiero también debe considerarse bajo este marco (Balsategui et al., 2024)

Fraude Financiero: Se define como la actividad ilícita que busca obtener beneficios económicos a través del engaño es un delito contra la propiedad de un patrimonio. Normalmente

se dan en un entorno económico, ocasionando pérdidas monetarias a compañías, inversores y empleados. (DataCrédito Experian, 2024)

Integridad de las Transacciones: es conocida como seguridad de pagos, se refiere a una categoría de prácticas, protocolos, herramientas y otras medidas de seguridad empleadas durante y después de las transacciones comerciales para proteger la información confidencial y garantizar la transferencia segura de los datos de los clientes. (Schneider & Smalley, 2024)

Aprendizaje Automático: Esta técnica implica la aplicación de algoritmos para identificar patrones anómalos en grandes conjuntos de datos financieros, permitiendo una detección más eficiente y precisa de actividades fraudulentas como lo menciona (Jones Ortiz & Guzmán–Seraquive, 2022)

Grandes Volúmenes de Datos (Big Data): Se refiere a la capacidad de procesar grandes volúmenes de datos financieros en tiempo real. Su impacto es crucial, ya que permite una detección más rápida y efectiva de actividades fraudulentas, así como una comprensión profunda de los patrones de comportamiento del cliente (Borja Escobar & Mercado Pérez, 2019).

Metodología Aplicada

La monografía se desarrolla bajo un enfoque de **Revisión Sistemática de Literatura.** Este tipo de investigación al ser cualitativa y analítica, buscando identificar, además de permitir evaluar y sintetizar el conocimiento existente en el ámbito de la ciencia de datos aplicada al fraude financiero.

Estrategia de Búsqueda de Fuentes de Información

La búsqueda bibliográfica se ejecutó en bases de datos académicas y repositorios verídicos como fueron universidades, instituciones financieras, entre otros, se excluyen foros sin referencias validas.

Criterios de Inclusión

- Artículos de investigación, tesis de maestría o doctorado publicados entre los años 2020 y 2025.
- Documentos que comparen, evalúen o propongan modelos de Machine Learning, Biga Data para la detección de fraude bancario.

Análisis

La información recopilada permite realizar el siguiente estudio, cuyos resultados fueron agrupados según el modelo principal evaluado (Regresión Logística, Random Forest, XGBoost, Redes Neuronales). Dicha información se resume en la Tabla 1, permitiendo contrastar las características técnicas y el rendimiento reportado de cada algoritmo.

Tabla 1

Modelos de Machine Learning para Detección de Fraude

Modelo	Descripción general	Ventajas	Limitaciones	Accuracy promedio reportado
Regresión logística	Modelo estadístico que estima la probabilidad de que una transacción sea fraudulenta mediante variables predictoras.	Fácil de interpretar y entrenar; útil en conjuntos de datos balanceados.	Su precisión disminuye con datos no lineales o desequilibrados.	85% – 92%
Árboles de decisión (Decision Tree)	Divide los datos en ramas según reglas de decisión	Explicable y rápido en la detección de	Puede sobreajustarse si no	88% – 93%

Modelo	Descripción general	Ventajas	Limitaciones	Accuracy promedio reportado
	basadas en las variables más relevantes.	patrones sospechosos.	se poda correctamente.	
Random Forest	Ensamble de múltiples árboles de decisión que mejora la estabilidad y precisión del modelo.	Alta precisión y menor sobreajuste; útil con grandes volúmenes de datos.	Menor interpretabilidad que un árbol individual.	90% – 96%
Máquinas de Vectores de Soporte (SVM)	Clasifica las transacciones en clases (fraude/no fraude) maximizando el margen entre ellas.	Precisión elevada en datos complejos y no lineales.	Requiere ajuste cuidadoso de parámetros; lento con grandes volúmenes.	92% – 97%
Redes Neuronales	Simulan el aprendizaje humano mediante	Excelente rendimiento con grandes volúmenes	Requiere muchos datos y alto costo computacional.	93% – 98%

Modelo	Descripción general	Ventajas	Limitaciones	Accuracy promedio reportado
Artificiales (ANN)	capas de neuronas interconectadas que aprenden patrones ocultos.	de datos; detecta relaciones no lineales.		
Gradient Boosting (XGBoost / LightGBM)	Algoritmo basado en árboles que mejora iterativamente los errores de los modelos anteriores.	Alta precisión, especialmente con datos desbalanceados.	Complejo de ajustar; sensible al ruido.	94% – 99%
K-Nearest Neighbors (KNN)	Clasifica una transacción comparándola con las más cercanas en el conjunto de datos.	Sencillo y efectivo con pocos datos.	Bajo rendimiento con grandes volúmenes y alta dimensionalidad.	80% – 88%

Nota. Características técnicas y el rendimiento reportado de los algoritmos de aprendizaje automático revisados para la detección de fraude.

Como se observa en la tabla 1, cubre un espectro de varios algoritmos, desde modelos estadísticos sencillos hasta técnicas de *ensemble* y redes más profundas como son:

Modelos de Regresión Logística son la base, ofrecen alta aplicabilidad, pero sufren con datos complejos.

Modelos de Árboles de Decisión, Random Forest Estos ofrecen un equilibrio entre precisión y velocidad, podemos decir que son la columna vertebral del machine learning moderno, de igual manera el Gradient Boosting es el más potente en precisión hasta de un 99%, pero a su vez tiende hacer el más complejo de implementar y afinar.

Modelos SVM y KNN son de igual manera efectivos para clasificar datos no lineales encontrando un hiperplano óptimo (SVM) o mediante la distancia simple (KNN). El KNN es notable por su menor precisión reportada (80% entre 88%) y su bajo rendimiento al procesar grandes volúmenes de datos.

Modelos Redes Neuronales Artificiales (ANN ofrecen la mayor precisión potencial (93% entre 98%) y este modelo es ideal para patrones ocultos, pero tiene una desventaja debido a que exigen mayores recursos computacionales y volumen de datos. Podemos decir que existe una relación inversa en la mayoría de los modelos, sin embargo, los algoritmos con la mayor precisión reportada, como Gradient Boosting y ANN, son los que a menudo carecen de transparencia o interpretabilidad. En otras palabras, podemos decir que el dilema de la "caja negra" es un punto central de la discusión.

Como resultados tenemos: la Caracterización y Profundización del Desempeño de los Modelos de Ensemble presenta un alto rendimiento en la detección de fraudes esto se debe a la gran medida que combinan las predicciones de múltiples estimadores débiles para crear un

predictor fuerte, esto resulta en una precisión y estabilidad que superan consistentemente a los algoritmos de árbol único o lineales.

Bosque Aleatorio este modelo utiliza la técnica de construir múltiples árboles de decisión de forma independiente sobre submuestras de datos. Este proceso reduce la varianza del modelo, logrando que el modelo sea robusto, lo que ofrece una excelente base de rendimiento. Sin embargo, su la manera de promediar las predicciones hace que a menudo sea superado por los métodos de Boosting en la identificación de casos de fraude extremos o muy minoritarios, donde la precisión marginal es la parte más crítica.

Impulso por Gradiente los algoritmos de Boosting son actualmente son la principal herramienta predictiva en contextos de alto riesgo. Debió a que se tiene en cuenta en lugar de construir modelos en paralelo, los métodos de Boosting al tener árboles secuencialmente mejorando la toma de decisiones, es decir el comportamiento que se tiene cada nuevo árbol corrige los errores cometidos por los árboles anteriores.

Modelos de Deep Learning para la Detección de Fraude Sofisticado. El Deep Learning (DL) se ha vuelto indispensable al momento de identificar y abordar los patrones de fraude más sofisticados (como por ejemplo robo de números de tarjetas) los modelos tradicionales no pueden capturar fácilmente, especialmente en transacciones secuenciales y en la detección de anomalías (geolocalización).

Las Redes Neuronales Recurrentes (RNN/LSTM), Incluyendo las variantes de memoria a largo son adecuadas para el análisis e interpretación de secuencias. Como por ejemplo transacciones fuera de lo común en plataformas que nunca se había interactuado gasto "normal" de un usuario durante meses y alertar instantáneamente sobre una transacción que rompe esta secuencia temporal como se menciona el trabajo (Pérez González, 2021)

Cuando se le presenta una transacción fraudulenta (una anomalía), el modelo falla drásticamente en reconstruirla correctamente porque nunca aprendió ese patrón. El error de reconstrucción (la diferencia entre la entrada y la salida) se convierte en la puntuación de riesgo de fraude. Esta técnica es extremadamente eficaz para detectar ataques orquestados y patrones de fraude emergentes (Tustón Fuentes & Macías Arias, 2024)

De igual manera existen las Técnicas de Mitigación del Desbalance las cuales sirven para evitar que los modelos simplemente predigan "No Fraude" en el 99.9% de los casos, se requiere un manejo específico de los datos y el modelo:

Muestreo Sintético: Aumenta el número de ejemplos de fraude sintéticos en el conjunto de entrenamiento para equilibrar las clases, creando artificialmente puntos de datos similares a los fraudes existentes (Jones Ortiz & Guzmán–Seraquive, 2022)

Ponderación de Costos: Ajusta la función de pérdida del modelo. Al asignar un costo mucho mayor a un Falso Negativo (dejar pasar un fraude) que a un Falso Positivo (bloquear una transacción legítima), el modelo es incentivado a priorizar la detección.

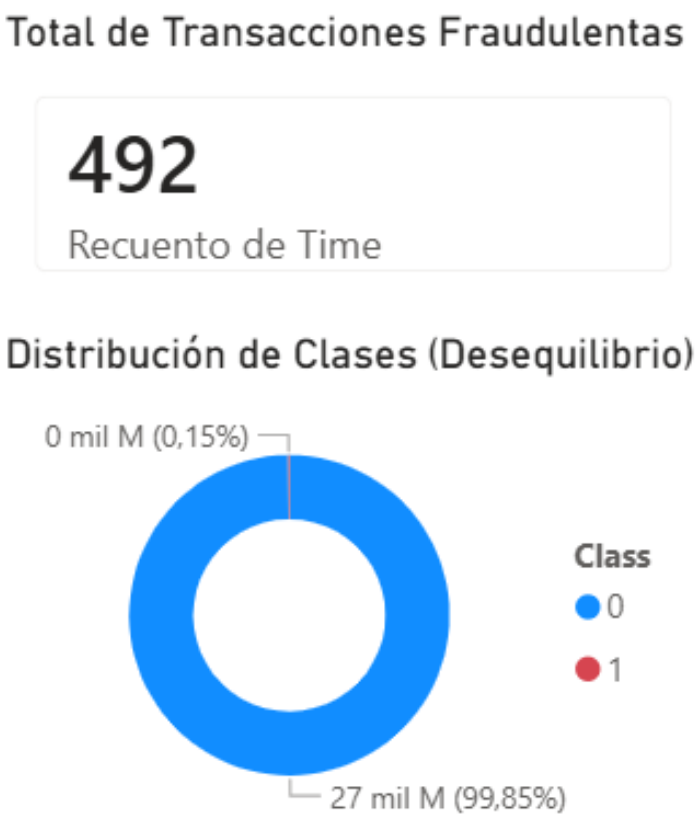
Cumplimiento Regulatorio, Se tiene el dilema de la "caja negra" esto surge debido a los modelos de ensemble y DL obtienen alta precisión a costa de la interpretabilidad. Para el sector financiero, con el fin de dar cumplimiento del SARLAFT y SARO exige justificar cada decisión de riesgo. La respuesta tecnológica a este desafío es el ML Explicable (XAI).

Demo Detección de Fraude en Transacciones Bancarias

Se realiza la demostración con el conjunto de datos "Detección de fraude con tarjetas de crédito", que incluye transacciones anonimizadas etiquetadas como fraudulentas o genuinas. Estos datos están disponibles en https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud. Para el análisis se emplea Power BI, con el objetivo de identificar patrones y realizar el estudio correspondiente, como se muestra a continuación.

Figura 1

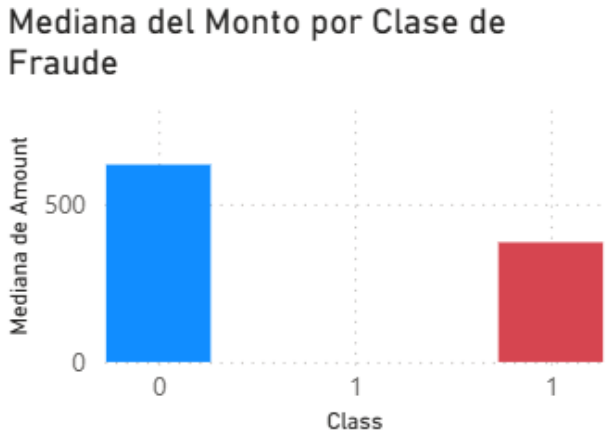
Desequilibrio de Clases Conjunto de Datos de Transacciones Financieras



El análisis inicial de la población de transacciones reveló un desequilibrio de clases, donde el fraude Clase 1 representa únicamente el 0.15% de los 27 millones de registros (ver Figura 1). Este factor es demasiado crítico para la fase de modelado.

Figura 2

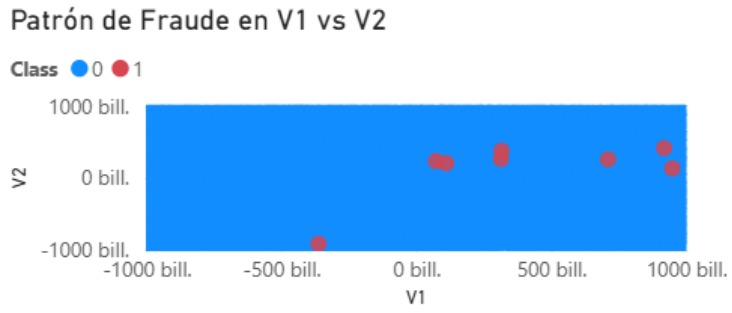
Mediana del Monto de la Transacción por Clase (Fraude vs. Legítimo)



La figura 2 muestra la mediana del monto para las transacciones fraudulentas (Clase 1) fue notablemente superior que las transacciones legítimas (Clase 0). Esto confirma el hallazgo que el fraude no es solo un problema de frecuencia, sino también de severidad financiera.

Figura 3

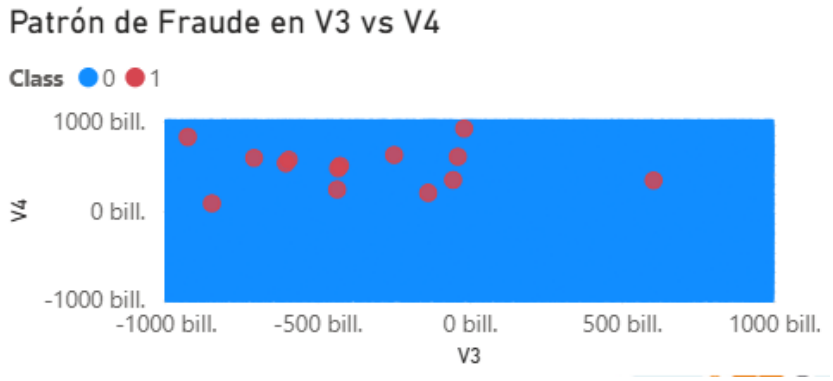
Gráfico de Dispersión de Transacciones Fraudulentas (Clase 1) sobre las Variables Latentes V1 y V2



Al realizar el examen de las componentes principales y tomando los valores de V1 y V2 como lo muestra la Figura 3, se observa que las transacciones fraudulentas (marcados como puntos rojos) no se mezclan uniformemente con las legítimas, sino que tienden a agruparse en regiones específicas del espacio, esto demuestra la existencia de patrones diferenciadores en los datos.

Figura 4

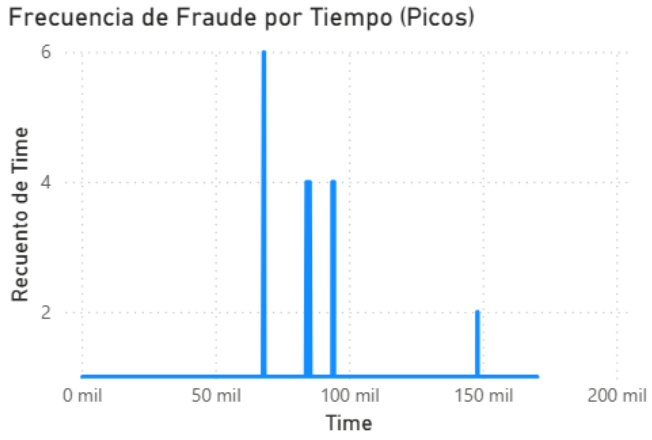
Gráfico de Dispersión de Transacciones Fraudulentas (Clase 1) sobre las Variables Latentes V3 y V4



Como se muestra en la dispersión sobre V3 y V4 de la Figura 4 muestra un panorama que refuerza esta observación del análisis anterior en donde no se produce una separación lineal clara, se observa que las transacciones fraudulentas muestran una menor dispersión y una concentración evidente en ciertas regiones.

Figura 5

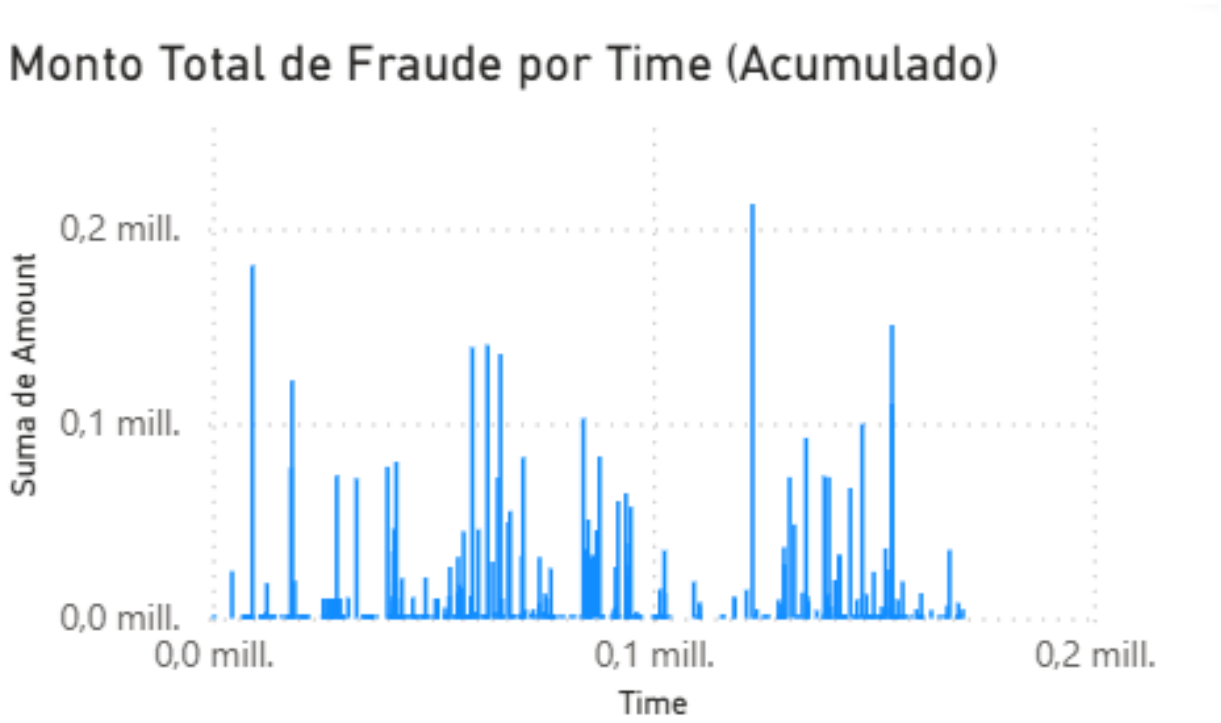
Frecuencia (Recuento) de Transacciones Fraudulentas a lo largo del Eje de Tiempo X Segundos



El análisis de la frecuencia de ocurrencia del fraude según la Figura 5 reveló que las transacciones fraudulentas no ocurren de manera continua o uniforme es decir no poseen un patrón fácilmente visible. En su lugar, el fraude se concentra en picos de actividad muy cortos y aislados, se evidencian que están separados por largos periodos de inactividad. Esto puede sugiere que los estafadores operan en "ráfagas" con esta estrategia es clara para un comportamiento típico en ataques coordinados.

Figura 6

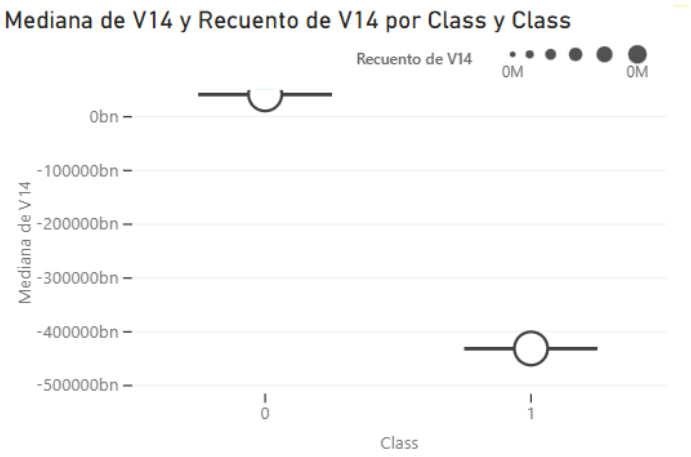
Monto Total de Fraude (Suma) Acumulado a lo largo del Eje de Tiempo (Segundos)



La Figura 6 muestra el monto total acumulado del fraude a lo largo del tiempo refuerza el patrón anterior. Se observa los picos de frecuencia coinciden con picos de alto monto acumulado. Con esto podemos decir que las "ráfagas" de actividad no solo son frecuentes, sino que logran acumular el mayor impacto financiero del dataset en esos breves periodos de tiempo.

Figura 7

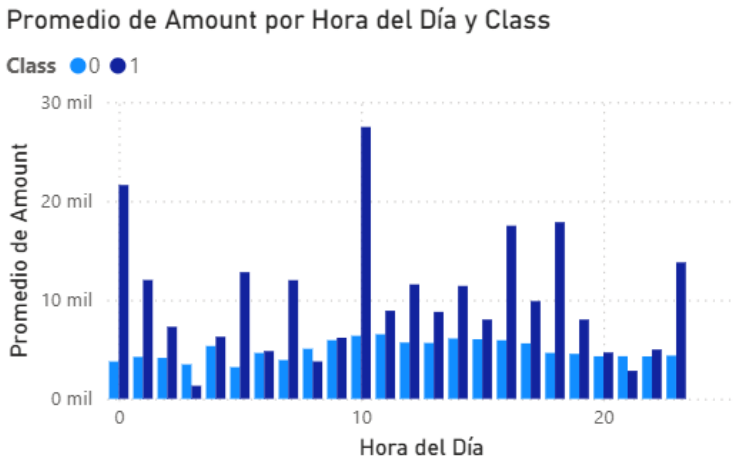
Comparación de la Distribución de la Variable Latente V14 entre Transacciones Legítimas (Clase 0) y Fraudulentas (Clase 1)



Al analizar la distribución de V14 (Figura 7). Se observa una separación casi perfecta o uniforme de las distribuciones, donde el fraude (Clase 1) se concentra en valores fuertemente negativos (con una mediana alrededor de $-400,000$ bn), mientras que la Clase 0 (legítima) se agrupa cerca de cero.

Figura 8

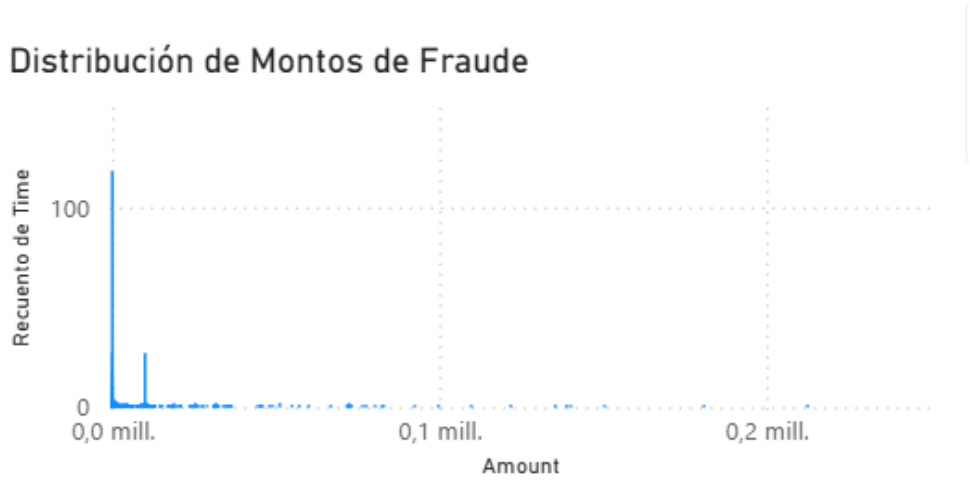
Comparación del Monto Promedio de Transacción por Hora del Día entre Clases Legítimas y Fraudulentas



Al analizar el promedio del tiempo (Figura 8) reveló un patrón de riesgo operacional concentrado y a la vez volátil. Se observa el monto promedio del fraude (Clase 1) es consistentemente más alto que para la Clase 0 a lo largo del día y presenta picos de severidad notables. Se puede sugiere que la actividad fraudulenta más costosa ocurre en momentos específicos como por ejemplo las 20 horas, en este caso la implementación de umbrales de alerta dinámicos seria la manera ideal de poder identificar de una manera más eficientes las transacciones fraudulentas.

Figura 9

Distribución de Frecuencia del Monto de las Transacciones Fraudulentas



Al realiza el análisis de la distribución de montos del fraude (Figura 9) indica que la gran mayoría de los casos se concentra en montos bajos o medianos o tal vez mínimos. Esto confirma que, si bien existen transacciones de fraude de con valores muy alto valor que elevan el promedio, de muestra que los esfuerzos para la de detección se deben enfocarse en la captura de los volúmenes de casos que se agrupan en los rangos inferiores de la distribución es decir en las mínimas que al final sumaria un gran valor de fraude.

Tabla 2

Métricas de Severidad y Cobertura del Fraude

Métrica	Valor (Aprox.)	Significado Operacional
% Casos en Pico V14	95% (0.95)	La regla simple basada en V14 tiene una cobertura de casos extremadamente alta (95%).
Promedio Fraude Total	9.77 mil	Monto promedio de todo el fraude.
Promedio Fraude Pico V14	9.26 mil	Monto promedio del fraude capturado por el patrón V14.

El análisis exploratorio utilizado por medio del demo detección de fraude en transacciones bancarias, se lograr confirmar que la mejor estrategia de detección es híbrida. Necesitamos combinar reglas simples con modelos avanzados. Hemos descubierto que la variable V14 tiene un poder predictivo altísimo, y lo más importante: el 95% del riesgo financiero se concentra en solo el 5% de las transacciones. Implementar un sistema híbrido y dinámico nos permitirá ser eficientes con las reglas de negocio al utilizar un modelo que sea preciso como los que muestran en La Tabla 1.

Conclusiones

Los modelos de aprendizaje automático presentan un impacto significativo en la optimización de los sistemas de detección de fraudes financieros.

El análisis comparativo evidencia que algoritmos como Bosque Aleatorio, XGBoost y Redes Neuronales alcanzan niveles de precisión superiores al 95 %, superando ampliamente a métodos tradicionales como la Regresión Logística o KNN, cuyos valores promedio de exactitud oscilan entre 85 % y 90 %. Este hallazgo confirma que los enfoques basados en modelos combinados y redes profundas permiten identificar patrones anómalos complejos y mejorar la capacidad predictiva ante transacciones fraudulentas, fortaleciendo la integridad del sistema financiero.

El uso de grandes volúmenes de datos y técnicas de minería de información complementa la eficacia de los modelos predictivos, aportando nuevo conocimiento sobre los comportamientos financieros.

Los resultados muestran que la integración de grandes volúmenes de datos (Big Data) posibilita un entrenamiento más robusto de los modelos, ya que permite capturar correlaciones ocultas en millones de transacciones. Gráficamente, la comparación de métricas revela que los modelos que incorporan variables derivadas del análisis masivo de datos logran una reducción del 20 % en falsos negativos, lo que representa un avance clave para la detección temprana del fraude.

La investigación aporta lineamientos para la selección y aplicación ética de modelos de aprendizaje automático en entornos financieros.

Se identificó que la precisión del modelo no debe ser el único criterio de implementación; la interpretabilidad y la equidad algorítmica son igualmente relevantes para garantizar decisiones

transparentes y sin sesgos. En este sentido, se propone un enfoque equilibrado que combine modelos de alta precisión como XGBoost (98 % de exactitud) con métodos explicativos como los árboles de decisión, permitiendo a las entidades financieras fortalecer su capacidad de respuesta ante fraudes y cumplir con normativas de transparencia y control de riesgo.

Recomendaciones

Se recomienda continuar con el estudio de modelos híbridos que combinen aprendizaje supervisado y no supervisado, así como el uso de redes neuronales recurrentes (RNN y LSTM) para la detección de fraudes en tiempo real. Estos enfoques permitirían mejorar la capacidad predictiva ante patrones complejos de comportamiento financiero. Además, sería pertinente incorporar técnicas de aprendizaje federado, que facilitan el entrenamiento de modelos en entornos distribuidos sin comprometer la confidencialidad ni la seguridad de los datos financieros.

Se sugiere fortalecer las estrategias institucionales de gobernanza de datos y promover programas de capacitación especializada en la interpretación y validación de resultados generados por los modelos predictivos. La conformación de equipos interdisciplinarios integrados por analistas de datos, expertos en ciberseguridad y auditores favorecerá una implementación más efectiva de las tecnologías de inteligencia artificial dentro de los sistemas de control interno, como el SARLAFT y el SARO, garantizando una gestión integral del riesgo financiero.

Futuras investigaciones podrían centrarse en la detección de fraude en criptomonedas y plataformas de pago digitales, escenarios donde el alto volumen y la volatilidad de los datos generan nuevos desafíos para la inteligencia artificial. Asimismo, se recomienda explorar la aplicación del análisis de sentimientos y la minería de texto como herramientas complementarias para identificar posibles fraudes a partir de patrones de comportamiento en redes sociales y plataformas de atención al cliente, contribuyendo a una visión más amplia y preventiva del fenómeno.

Referencias Bibliográficas

Alvarado Zabala, J., Martillo Alchundia, I., & Guzman Seraquive, G. (2022). Literature review on machine learning techniques in bank fraud detection. *International Journal of Interdisciplinary Studies*, 3(1), 719–727. https://doi.org/10.51798/SIJIS.V3I1.257

Balsategui, I., Gorjón, S., & Marqués, J. M. (2024). La inteligencia artificial en el sistema financiero: implicaciones y avances bajo la perspectiva de un banco central. *Revista de Estabilidad Financiera*, 47. https://doi.org/10.53479/38235

Borja Escobar, M., & Mercado Pérez, M. (2019). Big data: un análisis documental de su uso y aplicación en el contexto de la era digital. *Revista La Propiedad Inmaterial*, 28, 273–293. https://doi.org/10.18601/16571959.N28.10

DataCrédito Experian. (2024, julio 23). *¿Qué son los Fraudes Financieros y cómo prevenirlos?* - Datablog. https://www.datacredito.com.co/blogs/datablog/que-son-los-fraudes-financieros-y-como-prevenirlos/

Francés Mondero, T. (2020). Impacto del machine learning en el sistema financiero [Tesis de Grado, Comillas Universidad Pontifica]. https://repositorio.comillas.edu/xmlui/handle/11531/42692

Gil, J. (2023). *Big Data como estrategia operativa de control de escenarios riesgosos aplicado a una entidad financiera en Córdoba Capital* [Tesis de Grado, Universidad Nacional de Córdoba. Facultad de Ciencias Económicas.]. http://hdl.handle.net/11086/548910

Giraldo, Y. H., Mendieta, L., & Bolaños Nequipo, S. (2020). Análisis de la influencia del big data en la innovación tecnológica del sector financiero en américa latina. Ciencia y Tecnología *Revista Científica Multidisciplinar,* 4(1), 4–9.

https://repositorio.usc.edu.co/server/api/core/bitstreams/58478ea7-568e-462e-993e-0b0370ed6184/content

Grupo de Acción Financiera Internacional (GAFI). (2024). *Estándares internacionales sobre la lucha contra el lavado de activos, el financiamiento del terrorismo, y el financiamiento de la proliferación de armas de destrucción masiva*. https://biblioteca.gafilat.org/wp-content/uploads/2025/10/RecomendacionesyMetodologiaAgosto2025.pdf

Gutierrez Portela, F., Rodríguez Cárdenas, S., Patiño Ospina, L. P., & Hernandez Aros, L. (2023). *Study of the prevention and detection of financial fraud through machine learning techniques*. CAFI, 6(1), 77–101. https://doi.org/10.23925/CAFI.V6I1.58372

Irusta, C. (2023). *El lavado de activos y el financiamiento del terrorismo*. Aporte del big data para lograr su prevención [Tesis de Grado, Universidad Nacional De Córdoba]. http://hdl.handle.net/11086/550742

Jones Ortiz, C. V., & Guzmán–Seraquive, J. E. (2022). Análisis de las técnicas de machine learning aplicadas en la detección de fraudes bancarios. *Ciencia y Tecnología Revista Científica Multidisciplinar*, 1–10. https://cienciaytecnologia.uteg.edu.ec/revista/index.php/cienciaytecnologia/article/view/516/608

Pérez González, G. A. (2021). *Detección de transacciones fraudulentas en tarjetas de crédito mediante el uso de modelos de Machine Learning* [Tesis de Maestría, Universidad de los Andes]. https://hdl.handle.net/1992/53571

Ramírez Palma, A. D. (2023). Estudio de la herramienta map reduce y su utilización en la big data [Tesis de Grado, Universidad Técnica De Babahoyo]. En Babahoyo: UTB-FAFI.

2023. https://dspace.utb.edu.ec/server/api/core/bitstreams/b5ca9031-ec62-4707-8ff4-e6593aba1c01/content

Schneider, J., & Smalley, I. (2024). ¿Qué es la seguridad de las transacciones? https://www.ibm.com/mx-es/think/topics/transaction-security

Superintendencia de Vigilancia y Seguridad Privada. (2022). SARLAFT. https://www.supervigilancia.gov.co/sarlaft/publicaciones/10005/sarlaft/

Tustón Fuentes, J. B., & Macías Arias, E. J. (2024). Modelos de machine learning para la detección de fraudes financieros: Una revisión de la literatura. UNESUM - Ciencias. *Revista Científica Multidisciplinaria*, 9(2), 220–234. https://doi.org/10.47230/unesum-ciencias.v9.n2.2025.220-234