Valoración de modelos de “machine learning” en la prevención y mitigación del fraude financiero

Año
2025

Autores
Chambi Condori, Pedro Pablo y Saravia Ticona, Telma Raquel

VALORACIÓN DE MODELOS DE "MACHINE LEARNING" EN LA PREVENCIÓN Y MITIGACIÓN DEL FRAUDE FINANCIERO

Pedro-Pablo Chambi-Condori
Universidad Nacional Jorge Basadre Grohmann
Email: pchambic@unjbg.edu.pe
Orcid: https://orcid.org/0000-0002-8628-6825

Telma-Raquel Saravia-Ticona

Universidad Nacional Jorge Basadre Grohmann, Perú

Orcid: https://orcid.org/0000-0003-3890-6695

Email: tsaraviat@unjbg.edu.pe

RESUMEN

El fraude financiero se ha convertido en una amenaza cada vez más seria para las instituciones y los usuarios en el mundo digital. Este estudio tiene como objetivo principal evaluar diferentes modelos de machine learning que se utilizan para detectar el fraude financiero, comparando su efectividad, precisión y capacidad de generalización. Se examinaron algoritmos supervisados como la regresión logística, los árboles de decisión, el randomforest y las máquinas de soporte vectorial, además de técnicas más avanzadas como XGBoost y redes neuronales. La evaluación se realizó utilizando métricas como la exactitud, precisión, recall y AUC-ROC, empleando conjuntos de datos reales y desbalanceados. Los resultados esperados mostraron que los modelos de Regresión Logística, Random Forest y la Red Neuronal tuvieron un rendimiento prometedor, cada uno destacando en diferentes aspectos de Precisión y Recall. La selección del modelo más adecuado y su umbral de clasificación final debe hacerse con una cuidadosa consideración de los requisitos operativos y los costos de negocio relacionados con los distintos tipos de errores de clasificación. En conclusión, elegir el modelo correcto, junto con un preprocesamiento adecuado, es fundamental para mejorar la prevención del fraude en el ámbito financiero.

Palabras clave: fraude financiero, detección de fraude, machine learning, modelos predictivos.

ABSTRACT:

Financialfraud has becomeanincreasinglyseriousthreattoinstitutions and users in the digital world. Themainobjectiveofthisstudyistoevaluatedifferent machine learningmodelsusedtodetectfinancialfraud, comparingtheireffectiveness, accuracy, and generalizationcapabilities. Supervisedalgorithmssuch as logisticregression, decisiontrees, randomforests, and support vector machines wereexamined, in additionto more advancedtechniquessuch as XGBoost and neural networks. Theevaluationwasconductedusingmetricssuch as accuracy, precision, recall, and AUC-ROC, employing real-world and unbalanceddatasets. TheexpectedresultsshowedthattheLogisticRegression, Random Forest, and Neural Network modelsperformedpromisingly, eachexcelling in differentaspectsofPrecision and Recall. Theselectionofthemostappropriatemodel and its final classificationthresholdmust be madewithcarefulconsiderationoftheoperationalrequirements and businesscostsassociatedwiththedifferenttypesofclassificationerrors. In conclusion, choosingtherightmodel, alongwithappropriatepreprocessing, iscriticaltoimprovingfraudprevention in thefinancialfield.

Keywords: financialfraud, frauddetection, machine learning, predictivemodels.

1. INTRODUCCIÓN

El fraude financiero es una de las mayores amenazas para la estabilidad económica de bancos, empresas y consumidores. Con la creciente digitalización, los esquemas fraudulentos se vuelven más sofisticados, aprovechando vulnerabilidades tanto tecnológicas como humanas. Según Alloy (2024), las pérdidas anuales por fraude financiero ascienden a miles de millones de dólares, lo que subraya la necesidad de mecanismos de detección más eficaces y adaptativos.

Tradicionalmente, los sistemas antifraude se basaban en reglas fijas o análisis manuales, pero estos métodos son insuficientes ante patrones dinámicos y complejos. En este contexto, el *Machine Learning* (ML) ha emergido como una herramienta poderosa para

la detección temprana y mitigación del fraude, gracias a su capacidad para analizar grandes volúmenes de datos, identificar patrones sutiles y adaptarse en tiempo real.
Este artículo tiene como objetivo comparar diferentes modelos de ML aplicados a la detección de fraude financiero, evaluando su rendimiento técnico y viabilidad en escenarios reales. Se utilizarán métricas como precisión, sensibilidad y AUC-ROC, considerando el desbalance de clases común en los datos financieros. Además, se abordan desafíos como la interpretabilidad, el desbalance de datos y las implicaciones éticas en decisiones automatizadas.

Según Abdulalem (2022), el fraude financiero implica tácticas engañosas para obtener beneficios económicos, y ha crecido de forma generalizada. Las técnicas tradicionales son costosas y lentas, mientras que el ML permite identificar transacciones fraudulentas de manera eficiente. Una revisión sistemática muestra que los algoritmos más utilizados son las máquinas de vectores de soporte (SVM) y las redes neuronales artificiales (RNA), especialmente en fraudes con tarjetas de crédito.

Lee et al. (2025) evaluaron algoritmos de ML y concluyeron que *randomforest* es el más eficaz, superando a modelos como SVM, regresión logística, redes neuronales y árboles de decisión, que enfrentan problemas de sobreajuste. Por su parte, Zhao y Bai (2022) propusieron un modelo híbrido que combina regresión logística con XGBoost para detectar fraude en empresas cotizadas. Este modelo superó el 99% de precisión en detección, mostrando su eficacia.

En cuanto al riesgo crediticio, Bhatore et al. (2020) señalan que el impago de préstamos también puede relacionarse con fraude financiero. La calificación crediticia, la monitorización del comportamiento del cliente y el análisis de patrones ayudan a reducir fraudes y activos improductivos (NPA). En su revisión de 136 estudios, observaron una creciente adopción de modelos híbridos con SVM y redes neuronales para la detección de morosidad y fraude crediticio.

Alsuwailem et al. (2023) realizaron un estudio en Arabia Saudita sobre ML aplicado a pequeñas y medianas empresas, evaluando modelos como Random Forest (RF), DecisionTree (DT), GradientBoosting (GB) y NearestNeighbor (KNN). RF obtuvo el mejor rendimiento (93% de precisión), seguido por DT (90%), KNN (87%) y GB

(74%). A nivel anual, tanto RF como DT alcanzaron una precisión del 98%, lo que resalta su efectividad en la clasificación de riesgos.

Por otro lado, Hossain et al. (2024) criticaron los métodos tradicionales de análisis de estados financieros por su incapacidad para detectar patrones complejos. Propusieron modelos de ML basados en RF, XGBoost y SVM, utilizando ratios financieros, gobernanza corporativa y características específicas de las empresas. Con preprocesamiento de datos y balanceo mediante SMOTE, los métodos de conjunto lograron mejor desempeño que los tradicionales. El estudio subraya la importancia de incorporar datos de gobernanza en la detección de fraude.

Olowe et al. (2024) destacan que el sector financiero está siendo transformado por tecnologías predictivas y ML. Estas herramientas optimizan la gestión del riesgo, la detección de fraudes, el análisis del cliente y la previsión del mercado. Algoritmos como regresión, árboles de decisión, SVM y aprendizaje profundo se aplican para identificar anomalías y personalizar productos financieros. También se usan en el procesamiento de lenguaje natural para analizar datos no estructurados, como redes sociales e informes financieros.

Sin embargo, la integración de ML enfrenta desafíos: privacidad de datos, interpretabilidad de modelos y cumplimiento normativo. Es fundamental abordar los sesgos algorítmicos y garantizar la protección de los datos, para equilibrar innovación con ética. La investigación y colaboración continuas son esenciales para promover un crecimiento sostenible y construir confianza en el uso de estas tecnologías.

Valavan et al. (2022) refuerzan que los algoritmos de ML son efectivos para detectar fraudes en tarjetas de crédito y morosidad en préstamos, aprendiendo de datos históricos para prever patrones futuros. Dado que las transacciones fraudulentas son mucho menos frecuentes que las legítimas, es crucial contar con modelos capaces de manejar esta desproporción. Detectar préstamos morosos a tiempo es clave para prevenir pérdidas, y los modelos de ML son especialmente útiles por su capacidad computacional para trabajar con datos desbalanceados.

Además, el artículo explora el uso de árboles de decisión, bosques aleatorios, regresión logística y gradientboosting para detectar fraude en préstamos fraudulentos. Se

utilizaron métricas como precisión, *recall*, F1-score y AUC-ROC. Según Kumar et al. (2020), el fraude con tarjetas de crédito ha crecido alarmantemente, y su detección es crucial en el análisis financiero. Evaluaron múltiples modelos (RF, redes neuronales, SVM, Naïve Bayes, regresión logística, etc.) en un conjunto de datos con 284,807 transacciones europeas, mostrando la eficacia del enfoque de retroalimentación para mejorar la detección y rentabilidad de los clasificadores.

En definitiva, los estafadores siguen innovando sus métodos, lo que exige que las tecnologías antifraude sean cada vez más sofisticadas. El uso de modelos predictivos basados en ML se ha vuelto esencial para minimizar pérdidas y mejorar la vigilancia financiera.

Este estudio concluye que los modelos de ML —incluyendo regresión logística, árboles de decisión, randomforest, SVM, XGBoost y redes neuronales— ofrecen ventajas considerables para detectar fraudes financieros. Se recomienda utilizar métricas como AUC-ROC y *recall* para evaluar su rendimiento y tomar decisiones fundamentadas sobre su implementación. La combinación de enfoques técnicos con consideraciones éticas y regulatorias permitirá construir sistemas de detección robustos, transparentes y sostenibles.

2. MARCO TEÓRICO

2.1. Fraude financiero

El fraude financiero consiste en el uso intencionado e indebido de recursos económicos para obtener beneficios ilegales, ya sea a nivel personal o corporativo. Según la Asociación de Examinadores de Fraude Certificados (ACFE), implica el uso de una posición en una organización para enriquecimiento personal mediante el mal uso de activos. Este fenómeno abarca desde fraudes con tarjetas de crédito hasta lavado de dinero y manipulación contable.

Existen diversas formas de fraude, entre ellas:

- Fraude con tarjetas: uso no autorizado para compras o retiros.
- Suplantación de identidad (phishing): robo de datos para actividades delictivas.

- Fraude digital: manipulación de plataformas bancarias en línea.
- Fraude contable: alteración de estados financieros.
- Lavado de dinero: ocultar el origen ilícito de fondos.
- Fraude en préstamos: uso de información falsa para obtener créditos.

Detectar fraude es un reto por varias razones:

- Desbalance de clases: los fraudes son muy pocos en comparación con las transacciones legítimas.
- Alta dimensionalidad: múltiples variables por transacción.
- Cambios constantes en los patrones de fraude.
- Datos incompletos o manipulados.

Enfoques tradicionales vs. Machine Learning:

Históricamente, la detección de fraude se basaba en sistemas de reglas fijas, con limitaciones frente a fraudes nuevos o sofisticados. En cambio, el uso de Machine Learning (ML) permite analizar grandes volúmenes de datos, identificar patrones complejos y adaptarse en tiempo real, mejorando la detección y reduciendo falsos positivos.

Machine Learning en la detección de fraude:

El ML es una rama de la inteligencia artificial que permite que los sistemas aprendan de los datos sin ser programados explícitamente. En el sector financiero, se aplica para prevenir fraudes en tarjetas, préstamos, transferencias y más.

Tipos de aprendizaje:

- Supervisado: requiere datos etiquetados. Algoritmos comunes: regresión logística, árboles de decisión, Random Forest, XGBoost y redes neuronales.
- No supervisado: busca anomalías sin etiquetas previas. Ejemplos: clustering, detección de outliers y autoencoders.
- Semi-supervisado y aprendizaje por refuerzo: menos frecuentes, pero útiles en contextos específicos como la validación en tiempo real.

Desafíos del ML en fraude financiero:

- Desbalance de clases: puede ocultar fraudes reales.
- Cambio de patrones (concept drift): requiere actualización constante del modelo.
- Necesidad de explicabilidad: vital en sectores regulados.
- Costo de errores: un falso negativo puede generar pérdidas; un falso positivo afecta al cliente.

Métricas clave para evaluación;

Dado el contexto, métricas como precision, recall, F1-score, AUC-ROC y la tasa de falsos positivos/negativos son más relevantes que la precisión global.

Aplicación práctica y futuro

Los modelos se integran en sistemas que analizan miles de transacciones por segundo. Se emplean técnicas de ensamblado y herramientas explicativas como SHAP o LIME. El futuro apunta a modelos adaptativos, éticos y explicables, con tecnologías emergentes como **AutoML**, **aprendizaje federado** y **modelos generativos para detección de anomalías**, marcando el camino hacia sistemas más robustos y confiables.

3. METODOLOGÍA

Esta investigación adopta un enfoque cuantitativo y experimental para evaluar la eficacia de diversos modelos de Machine Learning (ML) aplicados a la detección de fraude financiero. La metodología se divide en cinco fases: selección de datos, preprocesamiento, construcción de modelos, entrenamiento y validación, y evaluación comparativa.

1. Selección del conjunto de datos

Se utilizó el dataset público "CreditCardFraudDetection" de Kaggle, que contiene 284,807 transacciones realizadas por clientes europeos, de las cuales 492 fueron clasificadas como fraudulentas. Este conjunto fue elegido por su disponibilidad, relevancia y reconocimiento en la literatura, así como por representar desafíos reales como el fuerte desbalance de clases y la falta de patrones explícitos.

2. Preprocesamiento de los datos

El preprocesamiento incluyó:

- Normalización del monto de transacción usando escalado Min-Max.
- Eliminación de la variable "Time".
- Manejo del desbalance con dos técnicas: undersampling de la clase mayoritaria y SMOTE para sobre-muestreo sintético.
- División estratificada en conjunto de entrenamiento (80%) y prueba (20%) para preservar la proporción entre clases.

3. Modelos de Machine Learning evaluados

Se implementaron seis algoritmos supervisados y uno no supervisado:

- Regresión logística: útil para clasificar entre fraude y no fraude mediante probabilidades.
- Árboles de decisión: dividen los datos de forma jerárquica para realizar clasificaciones.
- Random Forest: ensamble de múltiples árboles que mejora la precisión y evita sobreajuste.
- XGBoost: modelo de boosting eficiente que optimiza errores secuencialmente.
- Redes neuronales artificiales: modelos complejos que aprenden patrones no lineales.
- Máquinas de soporte vectorial (SVM): clasifican con márgenes óptimos, útiles para relaciones no lineales.

Todos los modelos fueron entrenados bajo las mismas condiciones y optimizados mediante validación cruzada (k=5) y ajuste de hiperparámetros con GridSearch o BayesianOptimization, según el algoritmo.

4. Métricas de evaluación

Dado el fuerte desbalance de clases, se usaron métricas que priorizan la sensibilidad frente al fraude:

- Precision: proporción de fraudes detectados correctamente entre los clasificados como tales.
- Recall (sensibilidad): proporción de fraudes reales detectados.
- F1-score: balance entre precisión y recall.
- AUC-ROC y AUC-PR: para medir la capacidad de discriminación del modelo.
- Tasa de falsos positivos (FPR) y falsos negativos (FNR): claves en el contexto financiero.

5. Validación y análisis

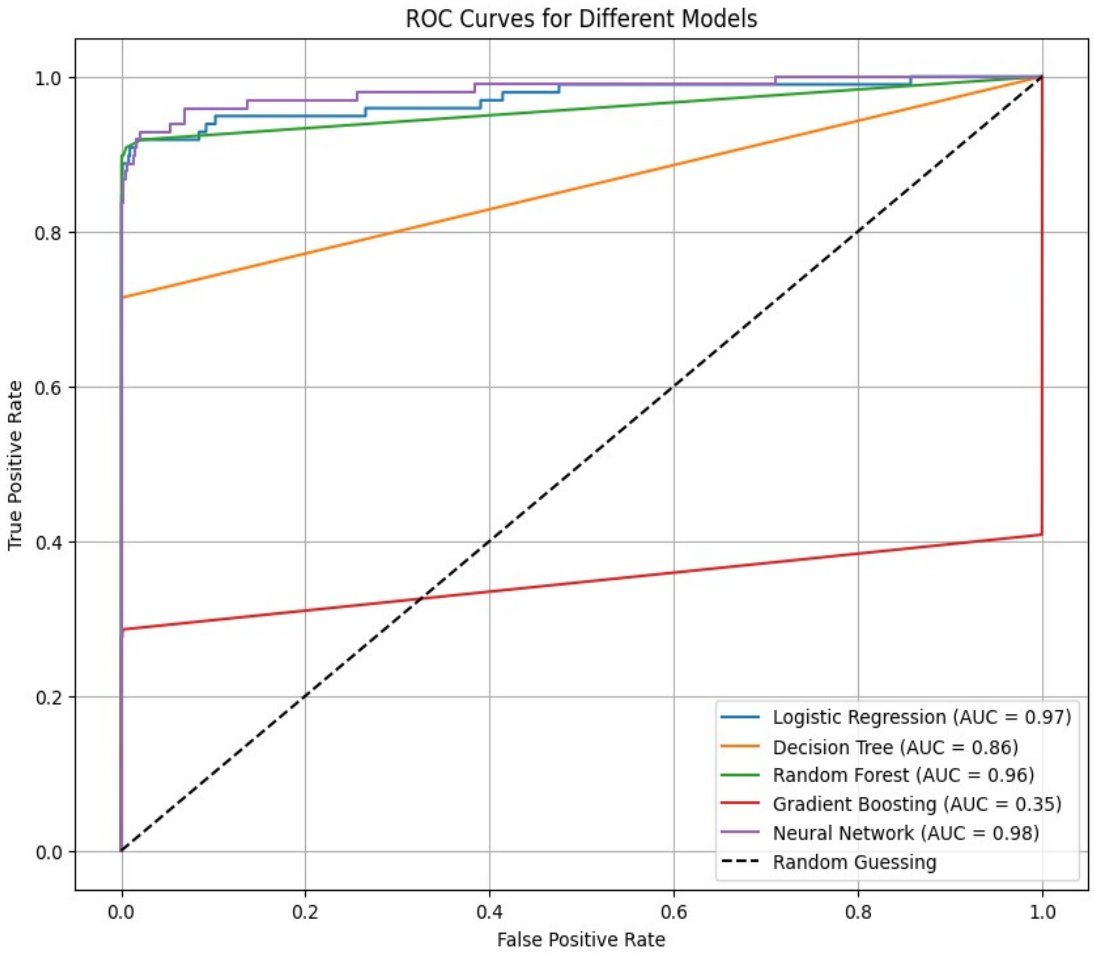
Para garantizar la comparabilidad, todos los modelos se entrenaron bajo las mismas condiciones computacionales en un entorno controlado en Python, usando bibliotecas como scikit-learn, imbalanced-learn, XGBoost, LightGBM y TensorFlow/Keras.

Se analizaron los resultados en términos de rendimiento global, capacidad para detectar fraudes (recall), robustez ante datos desbalanceados y nivel de interpretabilidad, considerando su posible aplicación práctica en contextos financieros.

Figura 1

Cursas ROC de los modelos evaluados

Nota.



Nota. Elaboración propia con los datos de la muestra con el soporte de Python 3 y Google Colab.

4. RESULTADOS

El conjunto de datos incluye 284,807 transacciones, cada una con 30 características y una variable objetivo llamada 'Clase'. No hay valores faltantes en este conjunto. Sin embargo, hay un notable desequilibrio, ya que solo el 0.17% de las transacciones son fraudulentas. Las características de Tiempo e Importe se normalizaron utilizando StandardScaler. Además, los datos se dividieron en un conjunto de entrenamiento (80%) y uno de prueba (20%), asegurando que la distribución de clases se mantuviera mediante estratificación. Se entrenó un modelo de regresión logística con ponderaciones de clase equilibradas. Este modelo logró una impresionante tasa de recuperación de 0.92 para la clase de fraude, lo que significa que identificó correctamente el 92% de las transacciones fraudulentas. Sin embargo, la precisión del modelo para la clase de fraude fue bastante baja (0.06), lo que indica que un número considerable de transacciones legítimas se clasificaron erróneamente como fraudulentas, resultando en 1,386 falsos positivos. Aunque la precisión general RECALL fue alta (0.98), esta métrica no es tan útil en conjuntos de datos con un alto desequilibrio. La puntuación AUC ROC fue de 0.972, lo que sugiere una excelente capacidad para distinguir entre las dos clases. Además, el modelo de regresión logística mostró un rendimiento sólido con una puntuación AUC alta (0.97), muy similar a la de la red neuronal, como se puede observar en la tabla 1.

Tabla 1

Indicadores de precisión del algoritmo de Regresión Logística

```
Classification Report:
              precision    recall  f1-score   support

           0       1.00      0.98      0.99     56864
           1       0.06      0.92      0.11        98

    accuracy                           0.98     56962
   macro avg       0.53      0.95      0.55     56962
weighted avg       1.00      0.98      0.99     56962


ROC AUC Score:
0.9720948047902334
```

El modelo de red neuronal presenta la puntuación AUC más alta (0,98), ver en la tabla 2 y figura 1, lo que indica que posee la mejor capacidad general para distinguir entre transacciones fraudulentas y no fraudulentas entre los modelos probados. Su curva ROC es la más cercana a la esquina superior izquierda.

Tabla 2

Indicadores de precisión del Modelo de Red Neuronal

```
--- Neural Network ---
Confusion Matrix:
array([[56862,     2],
       [   44,    54]])

Classification Report:
              precision    recall  f1-score   support

           0       1.00      1.00      1.00     56864
           1       0.96      0.55      0.70        98

    accuracy                           1.00     56962
   macro avg       0.98      0.78      0.85     56962
weighted avg       1.00      1.00      1.00     56962


ROC AUC Score: 0.986823017755217
```

El modelo de bosque aleatorio presenta una buena puntuación AUC (0,96), con un buen rendimiento, aunque ligeramente inferior al de la red neuronal y la regresión logística en términos de capacidad discriminatoria general, medida por el AUC.

Tabla 3

Indicadores de precisión del Modelo Random Forest

```
--- Random Forest ---
Confusion Matrix:
array([[56861,     3],
       [   24,    74]])

Classification Report:
              precision    recall  f1-score   support

           0       1.00      1.00      1.00     56864
           1       0.96      0.76      0.85        98

    accuracy                           1.00     56962
   macro avg       0.98      0.88      0.92     56962
weighted avg       1.00      1.00      1.00     56962


ROC AUC Score: 0.9580765743973446
```

El modelo de árbol de decisión presenta una puntuación AUC más baja (0,86) en comparación con la regresión logística, el bosque aleatorio y la red neuronal, lo que sugiere que es menos eficaz para distinguir entre las clases.

Tabla 3

Indicadores de precisión del Modelo de Arbol de Decisión

```
--- Decision Tree ---
Confusion Matrix:
array([[56835,    29],
       [   28,    70]])

Classification Report:
              precision    recall  f1-score   support

           0       1.00      1.00      1.00     56864
           1       0.71      0.71      0.71        98

    accuracy                           1.00     56962
   macro avg       0.85      0.86      0.86     56962
weighted avg       1.00      1.00      1.00     56962


ROC AUC Score: 0.8568878627703193
```

El modelo de refuerzo de gradiente presenta la puntuación AUC más baja (0,35) ver en la tabla 4, lo que indica un rendimiento deficiente en esta tarea. Su curva ROC está cerca o por debajo de la línea de conjetura aleatoria, lo que sugiere que su rendimiento no es mucho mejor que el del azar. En resumen, según el análisis ROC, los modelos de redes neuronales y regresión logística parecen ser los más prometedores para esta tarea de detección de fraude, seguidos de cerca por el modelo de bosque aleatorio. El árbol de decisión y, en especial, los modelos de potenciación de gradiente no obtuvieron el mismo rendimiento según esta métrica.

Tabla 4

Indicadores de precisión del Modelo GradientBoosting

```
--- Gradient Boosting ---
Confusion Matrix:
array([[56848,    16],
       [   80,    18]])

Classification Report:
              precision    recall  f1-score   support

           0       1.00      1.00      1.00     56864
           1       0.53      0.18      0.27        98

    accuracy                           1.00     56962
   macro avg       0.76      0.59      0.64     56962
weighted avg       1.00      1.00      1.00     56962


ROC AUC Score: 0.3468859283302516
```

Es importante recordar que, aunque el AUC es una buena medida general, también debemos tener en cuenta otras métricas como la precisión y la recuperación, especialmente cuando tratamos con conjuntos de datos desequilibrados y los requisitos específicos de un sistema de detección de fraude (por ejemplo, el costo de los falsos positivos en comparación con los falsos negativos). Ya hemos discutido estas métricas anteriormente. La curva de precisión-recuperación ilustra el equilibrio entre la precisión y la recuperación para diferentes umbrales de clasificación. La precisión se refiere a la capacidad del modelo para no etiquetar como positiva una muestra que en realidad es negativa. En el contexto de la detección de fraude, esto significa la proporción de transacciones que se marcan como fraudulentas y que realmente lo son (Verdaderos positivos / (Verdaderos positivos + Falsos positivos)). Una alta precisión implica que hay menos falsas alarmas. Por otro lado, la recuperación mide la capacidad del modelo para identificar todas las muestras positivas. En la detección de fraude, esto se traduce en la proporción de transacciones fraudulentas reales que se detectan correctamente (Verdaderos positivos / (Verdaderos positivos + Falsos negativos)). Una alta recuperación significa que se omiten menos transacciones fraudulentas. Si observamos la figura 1, la curva comienza en la esquina superior izquierda, donde la recuperación es casi 0 y la precisión es alta (o indefinida). A medida que nos movemos hacia la derecha en la curva, la recuperación aumenta, pero generalmente la precisión disminuye. El área bajo la curva (AUC) de la curva de precisión-recuperación (PR AUC) nos da un único valor que resume el rendimiento en todos los umbrales posibles. Un PR AUC más alto indica un mejor rendimiento, especialmente en conjuntos de datos desequilibrados,

donde el ROC AUC puede resultar engañoso. En nuestro caso, el PR AUC para la regresión logística es de 0,67.

A partir de los umbrales de ejemplo que se muestran debajo del gráfico, podemos ver la compensación: Con un umbral muy bajo (por ejemplo, 0,0000), el modelo tiende a predecir casi todo como positivo, lo que resulta en una recuperación alta (1,0000) pero con una precisión muy baja (0,0017). Esto significa que detecta todo el fraude, pero también clasifica casi todas las transacciones legítimas como fraudulentas. Al elevar el umbral (por ejemplo, a 0,0593), la precisión mejora (a 0,0056), pero la recuperación comienza a caer (a 0,9592). El modelo se vuelve más selectivo en sus predicciones de fraude, lo que reduce los falsos positivos, pero también puede pasar por alto algunas transacciones fraudulentas. Con un umbral más alto (por ejemplo, 0,1679), la precisión aumenta aún más (hasta 0,0135) y la recuperación disminuye de nuevo (hasta 0,9490). Con el umbral de ejemplo de 0,8, notamos un aumento significativo en la precisión hasta 0,16, mientras que la recuperación se mantiene alta en 0,89. Esto demuestra que al aumentar el umbral, podemos reducir drásticamente el número de falsos positivos (de 1386 en el umbral predeterminado a 458 en 0,8), al mismo tiempo que identificamos una gran parte de las transacciones fraudulentas (87 de 98). Conclusión clave: La curva de precisión-recall nos ayuda a visualizar y entender la compensación inherente entre detectar todas las transacciones fraudulentas (alta recuperación) y minimizar las falsas alarmas (alta precisión). La elección del umbral óptimo depende de las prioridades y los costos específicos de su sistema de detección de fraude. Desarrollar un modelo de aprendizaje automático preciso para la detección de fraudes es un paso crucial, pero su efectividad en un entorno real depende en gran medida de su integración en el contexto empresarial y los flujos de trabajo operativos existentes. A continuación, se explica por qué la integración en el contexto empresarial es tan importante y qué implica.

Alineación con los objetivos empresariales: Es fundamental que el modelo esté en sintonía con los objetivos de la empresa. Por ejemplo, ¿es más importante reducir las pérdidas financieras por fraude (lo que implica una mayor recuperación) o disminuir el costo operativo de investigar falsos positivos (lo que requiere mayor precisión)? La selección del modelo y su umbral deben reflejar estas prioridades. Definir la predicción "aplicable": Cuando el modelo clasifica algo como "fraudulento", esto debe traducirse en una acción clara para el negocio. Algunas opciones podrían ser: - Bloquear la

transacción de inmediato. - Marcar la transacción para que un analista de fraude la revise manualmente. - Pedir al cliente que verifique la transacción. - Ajustar la puntuación de riesgo del cliente. Integración del flujo de trabajo: Los resultados del modelo deben encajar sin problemas en el flujo de trabajo de detección de fraude que ya existe. Esto podría incluir: - Recibir datos de transacciones en tiempo real. - Enviar las predicciones del modelo a un sistema de gestión de fraude. - Activar alertas o tareas para el equipo de investigación de fraude. Determinación del umbral basada en el análisis de costo-beneficio: El umbral óptimo para clasificar no debe basarse solo en las métricas de rendimiento del modelo. También debe tener en cuenta los costos financieros y operativos de los falsos positivos y negativos. Un análisis de costo-beneficio puede ayudar a encontrar el umbral que minimiza los costos generales o maximiza los beneficios netos para la empresa. Ciclo de retroalimentación: Establecer un ciclo de retroalimentación es clave para la mejora continua. Esto implica: - Recoger la opinión de los analistas de fraude sobre las predicciones del modelo (por ejemplo, si una transacción detectada realmente fue fraudulenta). - Usar esta retroalimentación para reentrenar y actualizar el modelo de manera periódica.

Cuando hablamos de consideraciones regulatorias y de cumplimiento, es crucial que los sistemas de detección de fraude se alineen con las leyes y regulaciones de privacidad de datos que correspondan. Durante el proceso de integración, es esencial asegurarse de que tanto el modelo como su implementación cumplan con estos estándares. Además, la comunicación efectiva con todas las partes interesadas relevantes—como analistas de fraude, gestores de riesgos, responsables de cumplimiento normativo y equipos de TI—es clave para lograr la aceptación y comprensión de las capacidades y limitaciones del modelo, así como para facilitar una integración sin contratiempos. En resumen, integrar el contexto empresarial significa conectar la solución de ciencia de datos con el entorno operativo real. Esto asegura que el modelo no sea solo un logro técnico, sino una herramienta valiosa que realmente ayuda a prevenir y mitigar el fraude financiero en el contexto específico de la empresa.

5. Discusión

De los antecedentes revisados a través de la literatura, haciendo un paralelo con los resultados obtenidos en el presente estudio, se constata que los estudios de Abdulalem (2020), Lee et al. (2025) y Zhao y Bai (2022) han encontrado que el modelo de

regresión logística es uno de los modelos que han dado mejores indicadores de precisión en la medición de fraude financiera.

6. Conclusiones

A lo largo de este análisis, se abordó el problema de la detección de fraude financiero utilizando un conjunto de datos de transacciones de tarjetas de crédito y aplicando varios algoritmos de machine learning. Los hallazgos clave y las conclusiones son los siguientes:

1. **Exploración y Preprocesamiento de Datos:**
   El dataset incluye transacciones anonimizadas con hora, monto y etiqueta de fraude. No hay valores nulos, pero existe un fuerte desbalance de clases. Se escalaron las variables 'Tiempo' y 'Monto', y se dividieron los datos con estratificación para conservar la proporción de fraude en entrenamiento y prueba.

2. **Evaluación de Modelos de Machine Learning:**
   Se entrenaron y evaluaron varios modelos, incluyendo Regresión Logística, Árbol de Decisión, Random Forest, GradientBoosting y una Red Neuronal simple. Las métricas clave para evaluar este problema, dado el desequilibrio en las clases, son la Precisión, el Recall (o Sensibilidad), el F1-score y el Área bajo la curva ROC (ROC AUC), además de la matriz de confusión. La Regresión Logística y la Red Neuronal destacaron por su alto Recall, lo que significa que fueron efectivas en detectar transacciones fraudulentas, aunque a costa de una Precisión relativamente baja, lo que resultó en un número significativo de falsos positivos. En otras palabras, aunque lograron identificar la mayoría de los fraudes, también marcaron muchas transacciones legítimas como sospechosas. Por otro lado, Random Forest mostró un buen equilibrio, alcanzando una alta Precisión (lo que implica menos falsos positivos) y un Recall aceptable, lo que lo convierte en una opción atractiva para situaciones donde es crucial minimizar las alarmas falsas. El Árbol de Decisión tuvo un rendimiento moderado en términos de Precisión y Recall, quedando por detrás de Random Forest. En cuanto a GradientBoosting, su desempeño fue bajo en esta configuración, especialmente en Recall. En lo que respecta al ROC AUC, tanto la Red

Neuronal como la Regresión Logística lograron las puntuaciones más altas, lo que sugiere una buena capacidad para diferenciar entre las clases. Sin embargo, la Curva de Precisión-Recall suele ser una métrica más reveladora para conjuntos de datos con alta asimetría, y nuestro análisis de la curva PR para la Regresión Logística dejó en claro la compensación entre Precisión y Recall.

3. **Consideraciones sobre la Asimetría de Clases y el Umbral de Clasificación:**

   La asimetría que encontramos en el conjunto de datos resalta lo crucial que es utilizar métricas de evaluación adecuadas, como Precisión, Recall y PR AUC, en lugar de simplemente fijarnos en la precisión general. Al analizar el umbral de clasificación para la Regresión Logística, quedó claro que ajustar este umbral puede influir notablemente en el equilibrio entre Precisión y Recall. La selección del umbral ideal dependerá de los costos asociados a los falsos positivos y falsos negativos en el contexto específico del negocio. En conclusión, se ha desarrollado un prototipo para detectar fraude financiero utilizando varios algoritmos de machine learning. Los modelos de Regresión Logística, Random Forest y Redes Neuronales mostraron un rendimiento prometedor, cada uno con sus propias fortalezas en términos de Precisión y Recall. La elección del modelo más adecuado y su umbral de clasificación final debe hacerse tras una cuidadosa evaluación de los requisitos operativos y los costos de negocio relacionados con los diferentes tipos de errores de clasificación. Para avanzar, se sugiere perfeccionar los modelos seleccionados a través de la optimización de hiperparámetros, explorar técnicas más avanzadas para abordar la asimetría de clases, y una integración más profunda con el contexto y los flujos de trabajo del negocio para maximizar la utilidad práctica de la solución de detección de fraude.

FUTUROS TRABAJOS:

Investigaciones posteriores podrían impulsar este trabajo incorporando fuentes de datos alternativas, como el análisis de sentimientos, y ampliando los conjuntos de datos para mejorar la generalización de los modelos.

7. Referencias

Abdulalem, A. (2022). FinancialFraudDetectionBasedon Machine Learning: A SystematicLiteratureReview. https://doi.org/10.3390/app12199637

Alsuwailem, A.A.S., Salem, E. &Saudagar, A.K.J. (2023). Performance ofDifferent Machine LearningAlgorithms in DetectingFinancialFraud. *ComputEcon* **62**, 1631–1667. https://doi.org/10.1007/s10614-022-10314-x

Alloy (2024). StateofFraudBenchmarkReport. https://www.alloy.com/state-of-fraud-benchmark-report-2024-ty

ACFE (2024). FraudMagazine. https://www.acfe.com/

Bhatore, S., Mohan, L. & Reddy, Y.R. (2020). Machine learningtechniquesforcreditriskevaluation: a systematicliteraturereview. *J BANK FINANC TECHNOL* **4**, 111–138. https://doi.org/10.1007/s42786-020-00020-3

Blake, M. (2025). Informe sobre la situación de la fabricación inteligente. Rockwell Automation.

Hossain, M.Z., Raja, M.R., & Hasan, L. (2024). DevelopingPredictiveModelsforDetectingFinancialStatementFraud: A Machine LearningApproach. EuropeanJournalofTheoretical and AppliedSciences, 2(6), 271-290. DOI: https://doi.org/10.59324/ejtas.2024.2(6).22

Hilpisch, Y. (2020). Artificial Ingtelligence in Finance. O´reilly. ISBN: **978-1492055433**

Kehinde Josephine Olowe1 ,Ngozi Linda Edoh2 , Stephane Jean Christophe Zouo3 , & Jeremiah Olamijuwon. Reviewofpredictivemodeling and machine learningapplications in financialservice análisis. http://www.fepbl.com/index.php/csitrj

Lee, C.-W., Fu, M.-W., Wang, C.-C., &Azis, M. I. (2025). Evaluating Machine LearningAlgorithmsforFinancialFraudDetection: Insightsfrom Indonesia. *Mathematics*, *13*(4), 600. https://doi.org/10.3390/math13040600

Kumar, N. ,Simaiya, S. , Lilhore, U. & Kumar S. AnEfficientCreditCardFraudDetectionModelBasedon Machine LearningMethods. International JournalofAdvancedScience and Technology Vol. 29, No. 5, (2020), pp. 3414 – 3424

López, M. (2018). Advances in Financial Machine learning.Willey. ISBN 978-119-43211-6. https://agorism.dev/book/finance/ml/Marcos%20Lopez%20de%20Prado%20-%20Advances%20in%20Financial%20Machine%20Learning-Wiley%20%282018%29.p..
https://www.rockwellautomation.com/content/dam/rockwell-automation/documents/pdf/campaigns/state-of-smart-2025/INFO-BR027D-ES-P-noi.pdf df

Malik, P., Baliyan, A., Palak, A. (2024). CreditRiskAssessment and FraudDetection in FinancialTransactionsUsing Machine Learning. International

JournalofScientificResearch&EngineeringTrendsVolume 10, Issue 2, Mar-Apr-2024, ISSN (Online): 2395-566X.

M. N. Ashtiani and B. Raahemi, "IntelligentFraudDetection in FinancialStatementsUsing Machine Learning and Data Mining: A SystematicLiteratureReview," in *IEEE Access*, vol. 10, pp. 72504-72525, 2022, doi: 10.1109/ACCESS.2021.3096799.

Valavan, M. & Rita, S. (2022). Predictive-Analysis-based Machine LearningModelforFraudDetectionwithBoostingClassifiers M. Valavan and S. Rita. ComputerSystemsScience&Engineering DOI: https://doi.org/10.32604/csse.2023.026508

Zhao, Z., & Bai, T. (2022). FinancialFraudDetection and Prediction in ListedCompaniesUsing SMOTE and Machine LearningAlgorithms. *Entropy*, *24*(8), 1157. https://doi.org/10.3390/e24081157