ISSN 2477-9105
Número 32 Vol.1 (2024)
https://doi.org/10.47187/perf.v1i33.324
Perfiles
Fecha de recepción: 10-12-2024 / **Fecha de aceptación:** 06-03-2025 / **Fecha de Publicación:** 15-04-2025

MODELOS DE APRENDIZAJE AUTOMÁTICO CONTRA EL FRAUDE: UN ENFOQUE HÍBRIDO PARA PROTEGER BILLONES EN TRANSACCIONES

Machine learning models against fraud: a hybrid approach to protecting billions in transactions

Josué Vladimir Galarza Tulcanazo*
Pablo Andrés Trejo Tapia

Universidad Central del Ecuador, Facultad de Ciencias Económicas, Estadística, Quito, Ecuador.

*jvgalarza@uce.edu.ec

RESUMEN

El fraude con tarjetas de crédito es un problema contemporáneo que afecta significativamente a la banca y a los consumidores, reportando pérdidas globales de 33.500 millones de dólares para 2022, con una tendencia creciente a lo largo de los años. Este trabajo aborda esta problemática mediante la implementación de modelos de aprendizaje automático, enfocándose en el diseño, evaluación y mejora de la identificación de transacciones fraudulentas con alta precisión y exactitud.

Los modelos desarrollados enfrentaron un desequilibrio significativo en las clases, para lo cual se implementaron técnicas como SMOTE y ADASYN, que mejoraron la representación de la clase minoritaria correspondiente a los casos de fraude. Asimismo, se utilizó el Análisis de Componentes Principales (PCA) con el fin de reducir la dimensionalidad y optimizar el rendimiento computacional.

Los resultados demostraron que, en términos de escalabilidad y adaptabilidad, el modelo de redes neuronales exhibió un excelente desempeño con conjuntos de datos grandes. Para los modelos híbridos, se implementó Voting Classifier, logrando un equilibrio óptimo entre adaptabilidad, precisión y eficiencia mediante la combinación de las fortalezas de diversos modelos. La interpretabilidad del sistema se mejoró mediante la implementación de SHAP, permitiendo explicar las decisiones del modelo en la detección de transacciones fraudulentas.

Palabras claves: *Fraude con tarjetas de crédito, Aprendizaje automático, Desequilibrio de clases, SMOTE y ADASYN, Voting Classifier, SHAP.*

ABSTRACT

Credit card fraud is a contemporary problem that significantly affects banking and consumers, reporting global losses of $33.5 billion for 2022, with an increasing trend over the years. This work addresses this issue through the implementation of machine learning models, focusing on the design, evaluation, and improvement of fraudulent transaction identification with high precision and accuracy.

The developed models faced a significant class imbalance, for which techniques such as SMOTE and ADASYN were implemented, improving the representation of the minority class corresponding to fraud cases. Additionally, Principal Component Analysis (PCA) was used to reduce dimensionality and optimize computational performance.

The results demonstrated that, in terms of scalability and adaptability, the neural network model exhibited excellent performance with large datasets. For hybrid models, Voting Classifier was implemented, achieving an optimal balance between adaptability, precision, and efficiency by combining the strengths of various models. The system's interpretability was enhanced through the implementation of SHAP, allowing for the explanation of model decisions in fraudulent transaction detection.

Keywords: *Credit card fraud, Machine learning, Class imbalance, SMOTE and ADASYN, Voting Classifier, SHAP.*

I. INTRODUCCIÓN

En los últimos años la estafa en transacciones con tarjetas de crédito presenta ser una amenaza costosa y creciente para la economía monetaria y global. De acuerdo con el informe presentado por la revista Nilson, los fraudes con tarjetas han incrementado según cifras del 2018 se perdieron 27.900 millones, así mismo en 2020 las pérdidas fueron de 28.400 millones, finalmente para 2022 existieron 33.500 millones en pérdidas (1). Las proyecciones predijeron con pérdidas de 43.470 millones de dólares a finales de 2028. En Ecuador, el uso de tarjetas de crédito mostro un incremento significativo en 2023. Según muestran los datos la Superintendencia de Bancos y Aval Buró, se registraron 4,2 millones de tarjetas de crédito activas en el país, utilizadas por más de 2 millones de personas para realizar 105 millones de transacciones, lo que representó un total de USD 21.891 millones. Comparado con 2022, esto significó un aumento del 11,7% en la cantidad de transacciones y del 17,4% en el valor total de las mismas. Además, 85.834 nuevos clientes ingresaron al sistema financiero formal a través de estas tarjetas, de los cuales el 51,5% fueron mujeres. En cuanto a la edad, el 53,5% de quienes accedieron por primera vez al sector financiero con una tarjeta de crédito fueron jóvenes menores de 25 años (2). Este incremento significativo de transacciones en el contexto nacional refleja la fragilidad del sistema financiero actual ante las nuevas modalidades de estafa.

La evolución de las técnicas fraudulentas con tarjetas de crédito y el aumento del comercio electrónico con la facilidad de transacciones, generadas a partir de la pandemia COVID-19, ha generado nuevos desafíos para las instituciones financieras. Este aumento se atribuye a la diversidad de opciones que ofrecen los bancos privados en el país. En particular, los canales digitales, como internet y aplicaciones móviles, los cuales fueron los más utilizados, con 456 millones de transacciones, representando un incremento del 63,3% en comparación con 2021. En total, se realizaron 929 millones de transacciones a través del sistema bancario ecuatoriano en 2022, lo que marcó un crecimiento del 37,6% respecto a 2021 y del 86,3% en comparación con 2019, antes de la pandemia.

Las aplicaciones móviles han sido el canal preferido para realizar operaciones bancarias, especialmente entre los más jóvenes. En 2022, el número de transacciones móviles fue 15 veces mayor que en 2019, destacándose su uso entre las generaciones centennials y millennials, es decir, personas de entre 13 y 42 años. Esta tendencia resalta la transformación digital del sistema financiero ecuatoriano y la capacidad de la banca privada para adaptarse a las nuevas preferencias de los usuarios (3).

Este cambio en el comportamiento del consumidor ha creado nuevas oportunidades para actividades fraudulentas, dado la grande demanda en esas fechas, buscando ser la estafa más sofisticada lo que empuja a requerir soluciones tecnológicas avanzadas para su detección y prevención.

Un punto para mencionar es que gracias a los avances en modelos de aprendizaje automático ha producido que la tasa de fraude por cada 100 dólares del volumen de transacciones se mantenga estable además se anticipa una ligera reducción en los años venideros (1). Dicho esto, se genera una aparente paradoja indicando que, aunque la cantidad total de transacciones fraudulentas está en aumento, las mejoras en las medidas de seguridad y el aumento en el volumen total de transacciones permiten que los modelos entrenen con mayor cantidad de información, logrando mayor precisión en la detección de transacciones fraudulentas lo que expande sistemáticamente su capacidad de detección de anomalías y movimientos sospechosos en términos reales (4). Esto contribuye a regular la tasa relativa de fraude y al mismo tiempo, les resulta más complicado para los delincuentes evadir la detección, ya que con el sistema se prepara a adaptar a las variaciones inevitables del ámbito de estafa.

Para este propósito, se trabajó con un conjunto de datos reales obtenidos de Kaggle que presenta un fuerte desequilibrio de clases, ya que las transacciones fraudulentas representan una pequeña fracción del total. Lo que constituye un reto importante, ya que los modelos cuando se entrenan tienden a favorecer la clase mayoritaria, lo que puede resultar en una alta tasa de falsos negativos en los resultados. Para mitigar este problema, se implementaron técnicas de sobre muestreo y submuestreo, que ayudan a que los modelos eliminen el sesgo, dichos métodos avanzados son SMOTE (Synthetic Minority

Oversampling Technique) y ADASYN (Adaptive Synthetic Sampling). Estas técnicas permiten equilibrar las clases y mejorar la capacidad de los modelos para dar solución correctamente a las transacciones fraudulentas.

Los modelos evaluados son K-Nearest Neighbors (KNN), Support Vector Machines (SVM), Redes Neuronales, Random Forest, regresión logística y Gradient Boosting, así como dos modelos híbridos basado en VotingClassifier, que combina las predicciones de todos los modelos anteriores mediante votación ponderada. Para optimizar el rendimiento del modelo híbrido, se realiza una búsqueda de hiperparámetros utilizando GridSearchCV. Se ajustan los parámetros del modelo Random Forest dentro del clasificador híbrido, evaluando configuraciones como el número de estimadores (n_estimators) y la profundidad máxima de los árboles (max_depth). Con el fin de superar las limitaciones individuales y sacar un rendimiento robusto al trabajar con las ventajas de los algoritmos individuales, para incrementar la exactitud en la identificación de fraudes. La fusión de estos modelos es especialmente importante para enfrentar los retos presentes en los grupos de datos engañosos y para ello lo recomendable es realizar un desbalanceo de clases y evidenciar la elevada incidencia de falsos positivos y falsos negativos resultantes (5).

Como lo describen Chawla et al., el desequilibrio de clases es un problema crítico en la detección de fraudes, en el que las transacciones fraudulentas son superadas respecto al número por las transacciones legítimas (6). En el caso de saltar el desequilibrio de clase resulta en clasificadores sesgados que solo priorizan la clase mayoritaria, lo que carcome gravemente la detección efectiva de casos de fraude. Este trabajo aborda dicho desafío, empleando además métodos como es el equilibrio de datos, con lo que se consigue una detección más precisa y eficiente de las estafas.

Además, la identificación de fraudes demanda un balance sensible entre reducir los falsos positivos, lo que sería considerar incorrectamente una transacción legítima como fraudulenta y los falsos negativos que consiste en omitir la detección de una transacción fraudulenta. Para cuantificar el valorar del rendimiento de nuestros algoritmos presentados en cuanto a sensibilidad y especificidad, se utilizó los análisis gráficos como la curva ROC (Característica Operativa de Recepción), tal como lo explica Fawcett (2006) (7). Esta metodología nos brinda la posibilidad de medir la habilidad de los modelos para distinguir entre operaciones legítimas y fraudulentas una vez que fueron entrenados, logrando cuantificar un indicador sólido para la valoración del desempeño de los modelos obtenidos.

Además de la evaluación del desempeño de los modelos se lo realizo con métricas conocidas como la precisión, el recall y el F1-score; Los modelos se evalúan utilizando el puntaje F1- score como métrica principal, que equilibra la precisión y el recall. Además, para la constatación se emplea el método de validación cruzada mediante cross_val_score para obtener una evaluación robusta de los algoritmos y se los valora de forma individual y colectiva con los modelos híbridos, además este trabajo también aborda aspectos fundamentales como el tiempo de ejecución, así como también la robustez de los mismos. También se implementa la técnica de SHAP (SHapley Additive Explanations), lo cual permite analizar la contribución de cada variable en las decisiones del modelo. Esto no solo mejora la transparencia en la transacción, sino que también asegura que las soluciones propuestas sean comprensibles y accionables por las instituciones bancarias.

La fusión de estas técnicas sofisticadas de aprendizaje automático con estrictos métodos de evaluación nos facilita la creación de modelos de identificación de fraudes más exactos y fiables para saber que las transacciones estarían respaldados por una pronta detección. Esta perspectiva no solo aspira a disminuir las pérdidas financieras vinculadas al fraude, sino que también reduce la interrupción de operaciones legítimas, en el caso de existir casos sospechosos, optimizando así la experiencia del usuario y preservando la integridad de los sistemas financieros (8).

En conclusión, la presente investigación cubre una necesidad vital del mercado financiero actual el cual tiene una tendencia creciente con los progresos tecnológicos contemporáneos en aprendizaje automático. Al enfrentar los retos particulares de identificar fraudes en tarjetas de crédito, tales como el desbalance de clases y la mejora de la exactitud, la investigación aporta al desarrollo de soluciones prácticas, más eficaces y cambiantes en la batalla contra el problema del fraude financiero.

II. MATERIALES Y MÉTODOS

El enfoque de esta investigación se centró en enfrentar los retos particulares de identificar fraudes en operaciones bancarias con tarjetas de crédito y con especial atención en la gestión del desbalance de clases debido al sobre ajuste de los modelos y la mejora del desempeño de los modelos. A continuación, se especifican los pasos fundamentales del procedimiento:

1. Preparación y preprocesamiento de datos.

1.1 Conjunto de datos.

Se empleó un conjunto de datos de transacciones realizadas con tarjetas de crédito, de las cuales incluía 284.315 operaciones, de las cuales únicamente 492 (17.3%) eran fraudulentas, evidenciando un notable desbalance de clases. Dicho grupo es consistentemente menor y probado con otros estudios que abordan el fraude en tarjetas de crédito (9).

1.2 Normalización.

Para normalizar o igualar las diversas características presentes en el conjunto de datos, se usa una escala comparable para todas las variables, para ello se utilizó el método de StandardScaler. Esta técnica de transformación asegura una media de 0 y una desviación estándar de 1, lo cual es esencial para la convergencia de los modelos de aprendizaje automático. Era necesario asegurar que las variables con rangos más amplios no dominaran o alteren el aprendizaje efectivo (10).

1.3 Reducción de dimensionalidad.

Se redujo la dimensionalidad del conjunto de datos original aplicando un Análisis de Componentes Principales (PCA). Se lo aplicó para reducir el número de dimensión del conjunto de datos a solo cinco componentes principales con la varianza más significativa preservada con el fin de explicar la mayor parte de la variabilidad de los datos, mejorando así la eficiencia computacional, esto se lo hace con el fin de mejorar la eficiencia del modelo y mitigar el sobreajuste, consiguiendo destacar las características más relevantes del conjunto de datos (11).

2. Manejo del desequilibrio de clases.

Para abordar el desequilibrio entre clases en el conjunto de datos (fraude vs no fraude), se utilizaron técnicas de sobre muestreo para balancear las clases, se implementaron dos métodos:

- Synthetic Minority Over sampling Technique (SMOTE): Se generraron instancias sintéticas para sobre muestrear la clase minoritaria (transacciones fraudulentas), creando ejemplos sintéticos basados en las características de las transacciones fraudulentas existentes, preservando las distribuciones originales (12).

- Adaptive Synthetic Sampling (ADASYN): Así mismo se utilizó instancias sintéticas adaptativa, centrándose en ejemplos minoritarios con mayor complejidad de clasificación. Esto ayuda a mejorar el balance entre las clases y a su vez permite entrenar modelos más robustos (13).

En particular, se observó un aumento significativo en el recall de los modelos, especialmente en KNN (15%) y Redes Neuronales (10%), lo que subraya la importancia de abordar el desequilibrio de clases en problemas de detección de fraudes. La combinación de SMOTE y ADASYN se ajustó para lograr un balance de clases de 1:1 en el conjunto de datos de entrenamiento, siguiendo las recomendaciones de estudios previos sobre el manejo de desequilibrio de clases en detección de fraudes en tarjetas de crédito (14).

3. Implementación de algoritmos.

Se implementaros varios algoritmos de clasificación para abordar la problemática, los modelos tanto de manera individual como en la aplicación de modelos híbridos con la combinación de tres modelos de aprendizaje automático, dichos modelos son:

3.1 K-Vecinos más Cercanos (KNN).

- Se aplicó el uso de scikit-learn con pesos homogéneos.

- Los hiperparámetros fundamentales comprendieron el número de vecinos (n_ neighbors) y el algoritmo de identificación de vecinos.

- La elección de hiperparámetros se fundamentó en investigaciones anteriores que han evidenciado la eficacia de KNN para identificar fraudes (15).

3.2 Máquinas de Vectores de Soporte (SVM).

- Se utilizó un núcleo RBF (Funcion de Base Radial) para registrar relaciones no lineales en la información.
- C (parámetro de regularización) y gamma (coeficiente del núcleo) fueron los hiperparámetros más destacados.
- La selección del núcleo RBF se basa en su habilidad para gestionar relaciones complicadas en datos financieros, tal como se ha evidenciado en investigaciones previas (16).

3.3 Redes Neuronales.

- Se implementó una red neuronal feedforward utilizando PyTorch.
- La arquitectura consistió en tres capas ocultas con funciones de activación ReLU.
- Los hiperparámetros incluyeron el número de neuronas por capa, la tasa de aprendizaje y el número de épocas.
- Esta arquitectura se basa en estudios recientes que han demostrado su efectividad en la detección de fraudes financieros (17).

3.4 Regresión logística.

Modelo lineal para clasificación binaria, proporciona una línea base simple pero poderosa que compara modelos más complejos, reconocido además por la simplicidad y eficiencia computacional en la resolución de problemas lineales, además es rápido y adecuado para datos lineales o casi lineales, ya que es fácil de ajustar con regularizaciones para evitar sobreajustes (18).

3.5 Árbol de decisión.

Este algoritmo se basa en particiones recursivas que permiten interpretar fácilmente las decisiones del modelo, que tiene un balance entre rendimiento, interpretabilidad y eficiencia computacional. Además, es un modelo robusto para conjuntos de datos desbalanceados con grandes volúmenes de datos. Por lo general se puede usar con técnicas como (class_weight='balanced') para el manejo des desbalance (19).

3.6 Aumento de Gradiente.

Un método de boosting que construye secuencialmente árboles de decisión, corrigiendo los errores de los árboles anteriores. Es excelente para rendimientos de conjuntos de datos grandes y desbalanceados, los algoritmos de LightGBM o XGBoost son optimizados con el fin de alcanzar mayor velocidad y escalabilidad, dicho modelo por lo general puede superar otros modelos en complejidad y desbalanceo (20).

3.7 Modelo Híbrido.

Este modelo presenta un grado alto de complejidad y costoso computacional mente en el entrenamiento, en algunos casos no puede proporcionar una mejora significativa sobre modelos individuales bien optimizados. Además, el entrenamiento con validación cruzada puede volverse lento, sin embargo, se combinó el poder predictivo de Logistic Regression, Decision Tree y Gradient Boosting ya que presentaron menor complejidad en el entrenamiento, se utilizó VotingClassifier con votación ponderada ("soft"), permitiendo un enfoque más robusto y equilibrado (21).

Se probaron además otras combinaciones como Logistic Regression, Decision Tree y SVM, pero el tiempo de entrenamiento fue demasiado lento, en comparación con la anterior combinación y modelos individuales esto se debe por la escalabilidad limitada para conjunto de datos grandes y el ajuste de hiperparámetros (C, kernel) puede ser computacionalmente costoso (22).

4. Hiperparámetros de ajuste.

Para la optimización del rendimiento de los modelos, se aplicó una búsqueda sistemática de hiperparámetros con las técnicas de:

- Grid Search. Es un proceso sistemático que se aplicó a KNN, SVM, Logistic Regression y Decision Tree, con el fin de explorar una variedad de combinaciones predefinidas de hiperparámetros (23).
- Random Search. Este método es utilizado para la red neuronal y para Gradient Boosting, que toma combinaciones aleatorias dentro de un espacio de búsqueda predefinido, de un espacio que se quiere explorar (24).

- K-fold cross-validation, con cv = 5, fue empleada para evitar el sobreajuste de los algoritmos y certificar la robustez de los resultados (25).

5. Evaluación de modelos.

Los modelos se evaluaron de acuerdo con las siguientes métricas:

- Precisión: Proporción de predicciones correctas de entre todas las predicciones positivas realizadas.
- Recall (Sensibilidad): Proporción de verdaderos positivos que fueron identificados correctamente.
- F1-Score: La media armónica entre la precisión y el recall.
- El área bajo la curva ROC (AUC - ROC): Medida de la capacidad del modelo para distinguir entre clases (26).
- Matriz de confusión: Para obtener los falsos positivos y los falsos negativos.

Además, se evaluó la eficiencia de cómputo en cada modelo, lo cual presenta ser un factor importante para los sistemas de detección de estafas de tiempo real, considerando las variables como el tiempo de entrenamiento y predicción (27).

6. Validación y pruebas.

El conjunto de datos se dividió en un conjunto de entrenamiento el cual requirió el 70% del conjunto de datos para entrenar los modelos y su respectiva validación cruzada. En cuanto al conjunto de prueba se lo realizo con el 30% lo que es aconsejable para la evaluación final del desempeño los modelos. Además, se aplicó una estratificación en la división del conjunto de datos para mantener la proporción de clases aconsejable en ambos conjuntos, asegurando una estimación justa y representativa para todos los modelos (28).

7. Análisis comparativo.

Finalmente, se realizó un análisis comparativo para identificar el modelo con mejor desempeño, considerando:

- Rendimiento en términos de las métricas mencionadas.
- Capacidad para manejar el desequilibrio de clases.
- Eficiencia computacional y escalabilidad.
- Capacidad de interpretación de los resultados.

Este estudio posibilitó establecer los puntos fuertes y débiles de cada método en el escenario particular de la identificación de fraudes en tarjetas de crédito, además se demostró que el modelo híbrido con VotingClassifier presentó un mejor desempeño global, denotando así la importancia de la combinación de enfoques complementarios para abordar problemas complejos y contemporáneos en la detección de fraudes.

III. RESULTADOS

Los hallazgos exponen percepciones significativas acerca del desempeño, cuantificados por la eficacia y utilidad de cada modelo en contextos de aplicaciones en tiempo real. A continuación, se presentan las pruebas descubiertas:

1. Precisión del método KNN.

La figura 1 ilustra las curvas de aprendizaje y validación cruzada para un modelo entrenado, con sus respectivas métricas en función del número repeticiones en que el modelo entreno.

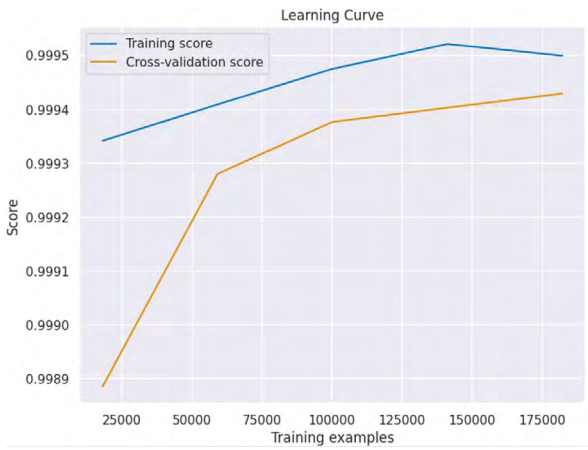

Figura 1. *Curva de aprendizaje del método KNN*

Respeto a la figura 1 la curva de color azul simboliza el desempeño del modelo en el

proceso de entrenamiento. Se puede vislumbrar que a medida que se incrementan los ejemplos, el rendimiento progresa debido a que el modelo está en su proceso de aprendizaje y parece apuntalar al resultado de 0.999315.

En cuando a la curva de aprendizaje de color naranja refleja el desempeño en los datos validados. A pesar de que al comienzo el desempeño es inferior al de entrenamiento, a medida que se añaden más datos, el modelo progresa en la validación cruzada, llegando a un valor de 0.9990976341.

Parece no existir un sobreajuste importante (overfitting), dado que las curvas de entrenamiento y validación se encuentran próximas y se aproximan al añadir más información. El modelo KNN demostró un rendimiento global excelente, con la precisión de 99.94%. lo que indica que el modelo puede generalizar adecuadamente a datos nuevos.

Matriz de confusión:

En la tabla 1 se muestra las marcas verdaderas en eje vertical (0 o 1), mientras que en el eje vertical las predicciones de los verdaderos.

verdaderos	0	56861	3
	1	29	69
		0	1
		predicción	

Tabla 1. *Matriz de confusión para el método KNN.*

Clase 0 (mayoritaria): El modelo realiza una predicción correcta de 56,861 casos y cuenta con 3 falsos positivos (es decir, calculó erróneamente 1 en lugar de 0).

Clase 1 (minoritaria): El modelo predice 69 casos como predicciones correctas, por otro lado, cuenta con 29 falsos negativos que quiere decir que calculó erróneamente 0 en lugar de 1.

El modelo se destaca en la case cero o mayoritaria, resaltando la robustes del modelo con escasos errores de predicción; sin embargo, en la clase minoritaria uno, presenta más problemas para capturar adecuadamente todos los positivos o casos detectados como fraudes, lo que refleja un problema recurrente en los escenarios de desequilibrio de clases, como se muestra en los casos de uno o falsos negativos.

Tendencia del Error.

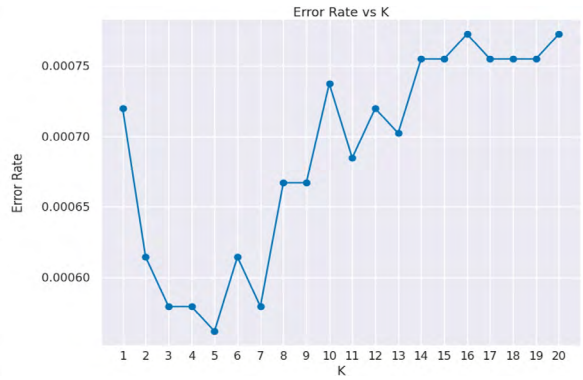

Figura 2. *Tendencia del error vs número de vecinos.*

En la figura 2 muestra la tendencia del error donde un valor de k = 1 genera un modelo que ajusta excesivamente los datos de entrenamiento, lo que conduciría a un error mínimo en el conjunto de entrenamiento, pero a un error de generalización más elevado lo que se conoce como sobreajuste. Conforme se incrementa el valor de K ejemplo a 20, el modelo adquiere mayor flexibilidad y puede disminuir el sobreajuste. Sin embargo, si K es excesivamente grande, el modelo podría suavizar excesivamente los resultados, lo que podría resultar en el aumento de la tasa de error. Final mente el valor optimo o más bajo de error es cuando K vendría a ser de 5 ya que presenta una caída inicial seguida de un incremento, siendo 5 lo que indica que es el más adecuado para reducir el error del modelo.

2. Precisión del método de redes neuronales.

Este modelo basado en redes neuronales (NN) obtuvo una precisión promedio de 99.89%, lo que indica una alta capacidad de generalización y presenta ser muy consistente a través de la validación cruzada, confirmado así la estabilidad del modelo.

Validación cruzada = (0.99898995 0.99949498 0.99709612 0.99949498 0.99938675).

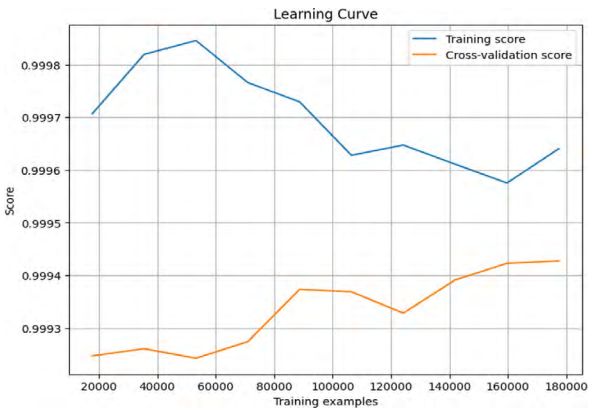

Figura 3. *Curva de aprendizaje para el método de redes neuronales..*

Como se muestra en la Figura 3 el color azul de la curva de entrenamiento inicia con una precisión muy alta, pero suele estabilizarse y reducirse un poco a medida que se incorporan más ejemplos. Esto es habitual y podría sugerir que el modelo está procurando prevenir el sobreajuste.

A pesar de que la curva de validación de color naranja presenta un rendimiento inferior al del conjunto de entrenamiento, progresa de manera gradual, lo que indica que el modelo está asimilando adecuadamente los datos de entrenamiento y haciendo una correcta generalización a datos nuevos.

Matriz de confusión:

verdaderos	0	55329	18
	1	21	75
		0	1
		predicción	

Tabla 2. *Matriz de confusión para el método NN.*

En la tabla 2 referente al modelo NN exhibe un desempeño sólido en la clasificación de la clase 0, dado que el método anticipa 55329 predicciones acertadas (falsos negativos correctamente categorizados como negativos) y 18 predicciones equivocadas (falsos negativos), con un alto número de proyecciones acertadas.

Pese a que las proyecciones para la clase 1 no son tan exactas como las de la clase 0 el cual presenta 75 predicciones acertadas frente a 21 erróneas (falsas negativas), el modelo continúa desempeñando un buen trabajo con más del 75% de exactitud en esta clase.

Podría resultar beneficioso modificar el límite de decisión o investigar métodos adicionales como el equilibrio de clases, si la clase 1 tiene mayor relevancia o si las clases se encuentran desbalanceadas.

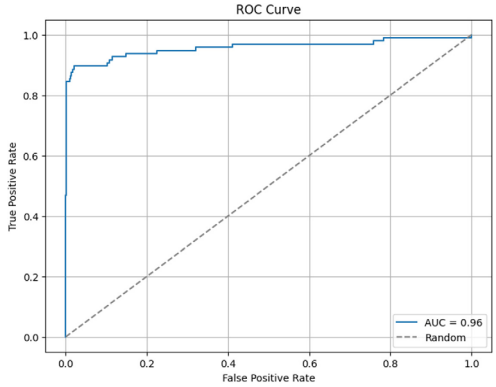

Figura 4. *Curva ROC para el método NN.*

Como se muestra en la Figura 4 la curva de color azul simboliza la eficiencia del modelo de clasificación. Se nota que la curva se eleva velozmente hacia la esquina superior izquierda, lo que señala que el modelo presenta un rendimiento óptimo. En este escenario, el modelo posee capacidad predictiva, ya que la curva del modelo se encuentra claramente por encima de dicha línea, lo que resulta positivo. El área bajo la curva AUC es de 0.96, lo que indica que el modelo posee una excelente capacidad discriminativa del modelo entre las clases, Lo que sugiere que el modelo puede diferenciar de forma muy exacta entre las clases en el 96% de las situaciones.

- Baja tasa de falsos positivos: Dicha taza resulta óptimo, dado que previene clasificaciones incorrectas.

- Elevada tasa de positivos verdaderos: La curva se eleva con rapidez, lo que significa que categoriza de manera correcta la mayoría de las situaciones positivas.

3. Precisión del método de máquinas de vectores de soporte.

El modelo de SVM funciona de manera sobresaliente, demostrando ser uno de los modelos más robustos, con una alta precisión del 99.92% y un excelente recall del 0.9934 lo que indica una sobresaliente capacidad para identificar transacciones fraudulentas.

Además, la uniformidad en los puntajes de validación cruzada indica que el modelo es estable y presenta un riesgo reducido de sobreajuste. Este comportamiento es típico de un modelo adecuadamente configurado que está gestionando de manera eficiente la complejidad del problema con un volumen de datos adecuado.

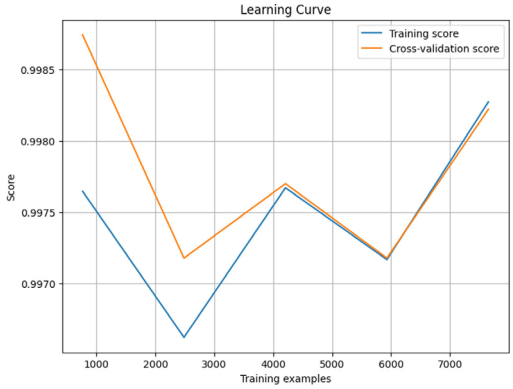

Figura 5. *Curva de aprendizaje para el método SVM.*

En la Figura 5 ambas curvas de entrenamiento y validación son consistentes y superan significativamente el valor de 0.997, señalando que el modelo posee precisión, además ambas curvas se acercan, lo que indica que el modelo está generalizando bien y no está sobre ajustado, lo que indica que es un comportamiento esperado en un buen modelo, donde la validación cruzada y el entrenamiento tienen desempeños similares.

La validación cruzada produjo puntajes elevados, y estables, confirmando así la robustes y tiene un rendimiento confiable en datos no vistos del modelo. (0.99832776, 0.99832776, 0.99832776, 0.99832706, 0.99790882)

Margen de decisión.

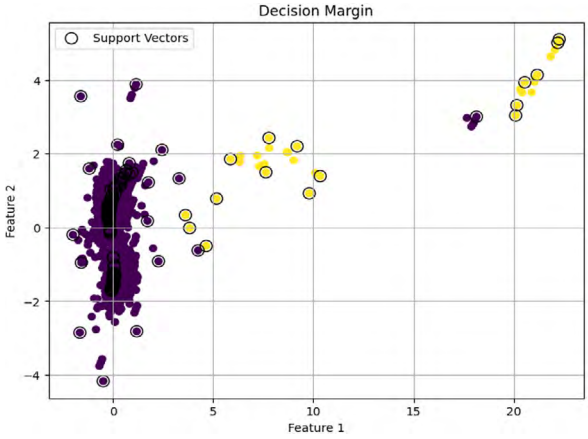

Figura 6. *Margen de decisión.*

En la figura 6 muestra la separación de dos clases de datos (puntos amarillos y morados) utilizando el modelo SVM, El método ha encontrado un hiperplano que separa las dos clases, consiguiendo de esta forma una buena generalización a pesar de presentar un escenario de alta complejidad.

Error vs parámetro C del SVM.

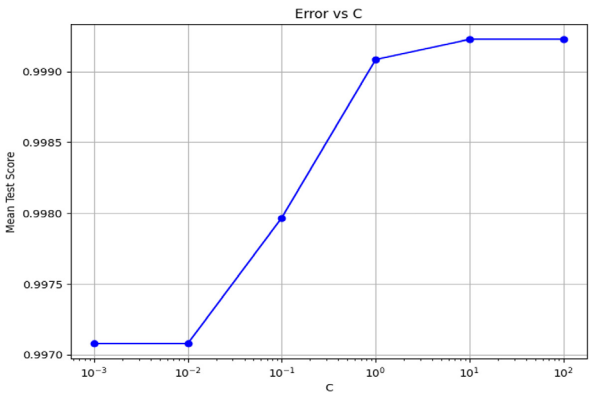

Figura 7. *Error vs parámetro C.*

La figura 7 muestra el valor óptimo de equilibró entre el rendimiento de la precisión y la generación, un valor de C entre 10^0 y 10^1 parece ser óptimo, ya que proporciona un buen rendimiento sin riesgo de sobreajuste.

Un C más bajo permite más errores de clasificación, pero un margen más suave, mientras que un C más alto busca menos errores de clasificación con un margen más exacto, la figura 7 muestra cómo cambia el rendimiento del modelo (Mean Test Score) a medida que C aumenta.

4. Precisión del método de regresión logística.

El modelo de regresión logística mostró un rendimiento destacado, con una AUC-ROC de 0.92. Este modelo lineal tuvo un excelente equilibrio entre precisión y recall, particularmente en la clase minoritaria:

Matriz de confusión:

verdaderos	0	85279	16
	1	58	90
		0	1
		predicción	

Tabla 3. *Matriz de confusión para el método de regresión logística.*

- Se encuentra en la tabla 3 que la clase 0 (mayoritaria): Predicciones correctas en 85,279 casos, con solo 16 falsos positivos.

- Respecto a la Clase 1 (minoritaria): Predicciones correctas en 90 casos, con 58 falsos negativos.

Curva ROC: El área bajo la curva (AUC) de 0.92 destaca una gran capacidad del modelo para discriminar entre transacciones fraudulentas y no fraudulentas, confirmando su robustez en escenarios desequilibrados.

5. Precisión del método de árbol de decisión.

El modelo de árbol de decisión obtuvo una AUC-ROC de 0.87, con un recall del 0.74 en la clase minoritaria. Esto indica que este modelo es capaz de capturar un mayor porcentaje de transacciones fraudulentas en comparación con la regresión logística.

Matriz de confusión:

verdaderos	0	85269	26
	1	39	109
		0	1
		predicción	

Tabla 4. *Matriz de confusión para el método de árbol de decisión.*

- Se encuentra según la tabla 4 la clase 0: 85,269 predicciones correctas, con 26 falsos positivos.

- Respecto a la clase 1: 109 predicciones correctas, con 39 falsos negativos.

En el modelo de árbol de decisión resulto ser más efectivo en la detección de fraudes, pero con un ligero descenso en la precisión general debido a un mayor número de falsos positivos.

6. Precisión del método de aumento de gradiente.

El modelo Gradient Boosting tuvo un desempeño moderado, con una AUC-ROC de 0.34. Si bien este modelo es generalmente efectivo en otros contextos, su capacidad para capturar fraudes fue limitada en este caso.

Matriz de confusión:

verdaderos	0	85286	9
	1	124	24
		0	1
		predicción	

Tabla 5. *Matriz de confusión para el método de aumento de gradiente.*

- Respecto a la tabla 5 se encuentra que la clase 0: 85,286 predicciones correctas, con 9 falsos positivos.

- Respecto a la clase 1: 24 predicciones correctas, con 124 falsos negativos.

Curva ROC: La baja AUC-ROC refleja dificultades significativas en la identificación de la clase minoritaria, lo que limita la aplicabilidad de este modelo para problemas de detección de fraudes altamente desequilibrados.

7. Precisión del método híbrido con Logistic Regression, Decision Tree, Gradient Boosting.

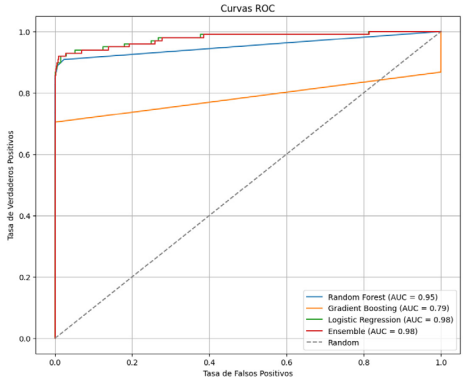

Figura 8. *Curva ROC del modelo hibrido e individuales.*

Como se muestra en la Figura 8 el modelo híbrido basado en Voting Classifier combinó Logistic Regression, Random Forest y Gradient Boosting, logrando un AUC-ROC de 0.98, el más alto entre todos los modelos evaluados.

Matriz de confusión:

verdaderos	0	56857	7
	1	30	68
		0	1
		predicción	

Tabla 6. *Matriz de confusión para el método híbrido (RF-LR-GB).*

- Respecto a la tabla 6 la clase 0: 56857 predicciones correctas, con solo 7 falsos positivos.

- Respecto a la clase 1: 68 predicciones correctas, con 30 falsos negativos.

Curva ROC y Precision-Recall: El Voting Classifier superó a los modelos individuales al equilibrar precisión y recall, mostrando que la combinación de modelos complementarios mitiga sus limitaciones.

Gráfico SHAP.

El color Rojo representa valores altos, mientras que el azul correspondería a los valores bajos, los puntos de color se encuentran dispersos a lo largo de las columnas, mostrando una distribución simétrica y cercana al valor de cero, mostrando de esta forma que la combinación entre valores de V1 y V2 impactan en las predicciones del modelo. Al ser las interacciones débiles, cercanas a cero, se encontrarían balanceadas en términos de su contribución a las predicciones.

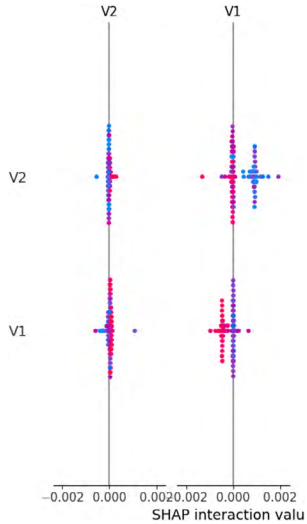

Figura 8. *Curva ROC del modelo hibrido e individuales.*

La figura 9 muestra que las interacciones entre V1 y V2 las cuales presentan ser consistentes y bajo control, se muestra además que no se presentan contribuciones externas que afecten a las predicciones. Lo que sugeriría que el modelo depende en su mayoría por los efectos individuales de las variables que de sus interacciones combinadas. En otras palabras, los valores de SHAP son cercanos a cero, lo que sugiere que las interacciones entre las variables no están influenciadas de manera conjunta de forma crítica, no tienen un impacto dominante en las predicciones del modelo. Sin embargo, un punto a considerar es que, si hay agrupaciones en valores específicos del plano, puede indicar que hay regiones donde las interacciones son más explicativas (29).

8. Precisión del método híbrido con Logistic Regression, SVM, Decision Tree.

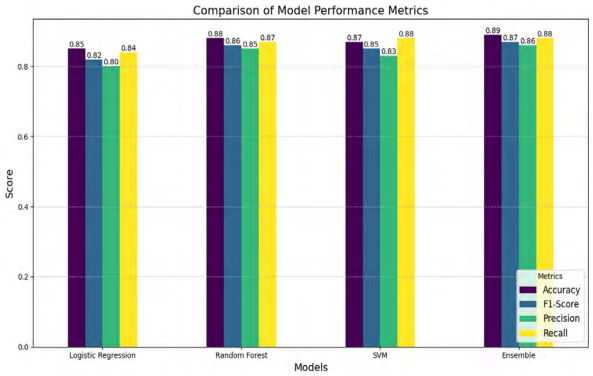

Figura 10. *Métricas de comparación de los modelos.*

	Accuracy	F1-Score	Precision	Recall
Logistic_Regression	0.85	0.82	0.8	0.84
Random_ Forest	0.88	0.86	0.85	0.87
SVM	0.87	0.85	0.83	0.88
Ensemble	0.89	0.87	0.86	0.88

Tabla 7. *Métricas para el método híbrido (RF-LR-SVM).*

Tanto en la figura 10 y en la tabla 7 se constata que en el segundo modelo híbrido presentado combinó los modelos de Logistic Regression, SVM y Random Forest, el cual consiguió valores considerablemente altos de puntuación, exactitud, precisión y un AUC-ROC de 0.98. Sin embargo, el modelo presento dificultades significativas en la localización de la clase minoritaria, con un recall cercano a cero, lo que se constata en la siguiente matriz de confusión tabla 8.

Matriz de confusión:

verdaderos	0	85294	1
	1	148	0
		0	1
		predicción	

Tabla 8. *Matriz de confusión para el método híbrido (RF-LR-SVM).*

- Clase 0: 85,294 predicciones correctas, con solo 1 falso positivo.
- Clase 1: Ninguna predicción correcta, con 148 falsos negativos.

Análisis: El molo presenta alta precisión en la clase cero (mayoritaria), pero carece de capacidad de detectar fraudes ya que no detecto ningún fraude en la clase uno o minoritaria.

1. Precisión en la Clasificación.

Respecto al rendimiento y su comparación general ¿, el creado con Voting Classifier logrando un excelente rendimiento general en términos de precisión en la clasificación, el modelo obtuvo 99.95%, el cual se basa en tres modelos, los cuales son Logistic Regression, Decision Tree y Gradient Boosting. El segundo mejor modelo lo presenta ser KNN con una precisión del 99.94% y seguido de SVM con la precisión del 99.92% (30). Además, el recall más alto lo obtuvo SVM el cual fue de 0.9934, lo cual indicaría una capacidad superior para identificar correctamente las transacciones fraudulentas. En el ámbito de la banca es crucial la minimización de los falsos negativos ya que se debe garantizar la correcta detección de fraudes (31).

Respecto a la métrica que evalúa la capacidad del modelo para distinguir entre las clases se lo realiza mediante el gráfico de la curva AUC-ROC y el cual presenta ser generalmente alto con buenas capacidades de detección de fraudes, todos los modelos superan el 90 %. Esto indica que todos los modelos son altamente efectivos en la discriminación entre las clases de transacciones fraudulentas y legítimas, superando los resultados reportados en estudios similares (32). Final mente uno de los desempeños con más problemas presento ser Gradient Boosting ya que presenta un recall de solo 0.16, se presenta el rendimiento de cada modelo en la tabla 9 a continuación.

Modelo	Precisión (%)	Recall	F1-Score	AUC-ROC	Tiempo de Ejecución
KNN	99.94	0.9912	0.9928	0.9961	1 segundo
SVM	99.92	0.9934	0.9929	0.9967	5 segundos
Redes Neuronales (NN)	99.89	0.9901	0.9895	0.9955	4 minutos
Logistic Regression (LR)	99.93	0.61	0.71	0.92	20 segundos
Random Forest (RF)	99.9	0.89	0.88	0.94	4 minutos
Gradient Boosting (GB)	99.85	0.16	0.27	0.34	2 minutos
Híbrido (RF-LR-GB)	99.95	0.64	0.75	0.93	1 minuto
Híbrido (RF-LR-SVC)	99.94	0	0	0.81	2 horas

Tabla 9. *Composición de los diferentes modelos.*

2. Tiempo de Ejecución y Recursos Computacionales.

KNN obtuvo el tiempo de predicción más rápido, debido a su simplicidad, además presenta la precisión alta y uso de memoria fueron significativamente más altos que los otros modelos, un hallazgo consistente con las reportadas en la literatura (33), mientras que el modelo híbrido construido por RF-LR-SVC tuvo el tiempo de ejecución más alto, lo que lo hace menos viable para entornos en tiempo real.

En cuanto a la eficiencia en el tiempo Logistic Regression demostró ser excelente, un punto a considerar fueron las Redes Neuronales, ya que, a pesar de su tiempo prolongado de entrenamiento, mostraron tener buenos tiempos de predicción, lo que las hace atractivas para aplicaciones en entornos de tiempo real una vez ya entrenadas. Este hallazgo presenta ser consistente con estudios recientes que destacan la eficiencia de las redes neuronales en la fase de predicción (34).

Otro modelo destacable es SVM el cual demostró un equilibrio óptimo entre precisión y eficiencia computacional ya que los tiempos de entrenamiento y predicción fueron moderados con un uso de memoria bajo. Este resultado es particularmente notable para aplicaciones en tiempo real, donde el equilibrio entre precisión y eficiencia es crucial (35).

3. Manejo del Desequilibrio de Clases.

Una forma de mejorar significativamente el rendimiento se lo consigue usando la técnica de sobre muestreo como SMOTE y ADASYN ya que contribuyo a una mejora general de todos los modelos en la detección de transacciones fraudulentas (36):

Respecto al modelo de KNN Demostró una mejora del 15% en recall, SVM incrementó su recall en un 8%, lo que denota un excelente equilibrio entre precisión y recall, redes neuronales experimentaron una mejora moderada del 10% en el recall, finalmente los modelos híbridos mostraron un mejor balance sin depender tanto de las técnicas mencionadas, destacando su robustez resultante frente al desequilibrio de clases.

Estos resultados subrayan la importancia de abordar el desequilibrio de clases en problemas de detección de transacciones fraudulentas, un hallazgo que está en línea con estudios previos sobre el impacto de las técnicas de balanceo en la detección de fraudes financieros. El modelo más notable es KNN y sugiere que este modelo es particularmente sensible al desequilibrio de clases, mientras que la estabilidad relativa del SVM y de los modelos híbridos indica su robustez inherente frente a este problema de desequilibrio de clases.

4. Interpretabilidad y Facilidad de Implementación.

En entornos financieros es necesario regular los modelos. En este estudio, los modelos simples que no presentan mucha carga computacional como KNN y Regresión Logística demostraron alta interpretabilidad, haciendo que personas con poca experiencia logren comprender fácilmente las decisiones que el modelo toma. Por otro lado, los modelos más complejos, como las Redes Neuronales y los modelos híbridos, presentaron limitaciones significativas en este aspecto de facilitar la interpretación. Para abordar esta limitación, se utilizó SHAP (SHapley Additive Explanations), que permitió analizar la contribución de cada variable en las decisiones del modelo, mejorando la transparencia y la

comprensión de las soluciones propuestas (17) (36).

KNN ofrece alta interpretabilidad, ya que las predicciones se basan en transacciones similares conocidas, es una ventaja significativa en el sector financiero, donde la transparencia en la toma de decisiones es crucial (15). Sin embargo, su implementación puede ser computacionalmente costosa para grandes volúmenes de datos ya que para dar un resultado utiliza todo el conjunto de datos.

SVM tiene una interpretabilidad moderada gracias a los vectores de soporte, que proporcionan información clave sobre las decisiones del modelo. Su implementación es más compleja, pero está bien soportada por bibliotecas en el lenguaje de Python, además una particularidad que presenta el modelo SVM es que generaliza bien los datos no vistos lo cual es particularmente valioso en la detección de fraudes, donde los patrones pueden evolucionar rápida y constantemente (16). Respecto a la interpretabilidad ofrece un equilibrio entre transparencia y complejidad.

Redes Neuronales son las más complejos de interpretar debido a su naturaleza de "caja negra". Sin embargo, son altamente configurables y adaptables a patrones cambiantes con reentrenamiento continuo, lo que la hace valioso en entornos donde los patrones de fraude evolucionan rápidamente. En otras palabras, la baja interpretabilidad de las redes neuronales es un desafío para considerar en el contexto de la detección de fraudes, donde la explicabilidad de las decisiones es a menudo un requisito regulatorio (17).

Modelo Hibrido con (RF-LR-GB) ofrecen una interpretabilidad limitada debido a su complejidad, al igual que las redes neuronales, pero la combinación de Rendom Forest, Logistic Regression y Gradient Boosting permite aprovechar las fortalezas individuales de cada algoritmo (37), al no presentar una carga excesiva en el cómputo de los resultados y en su validación, logrando un mejor rendimiento general, además este modelo demuestra un equilibrio entre precisión y complejidad. La interpretabilidad se mejora significativamente mediante la implementación de SHAP (SHapley Additive exPlanations), que proporciona explicaciones transparentes sobre las decisiones del modelo (38). Wang et al. (39) han verificado que esta aproximación híbrida mejora la comprensión de las predicciones en contextos financieros complejos, en la detección de transacciones fraudulentas.

5. Escalabilidad y Adaptabilidad.

Los fraudes presentan patrones de cambio al pasar del tiempo y que inevitablemente van en aumento lo cual los modelos deben ser evaluados en aspectos como la escalabilidad y una correcta adaptabilidad, a las circunstancias en que se requieren los algoritmos. Las Redes Neuronales demostraron dichas características, siendo capaces de operar grandes volúmenes de datos sin problemas y con su característica principal de ajustarse a nuevos patrones con reentrenamientos continuos. En contraste, KNN mostró limitaciones importantes en el aspecto de escalabilidad, ya que su rendimiento disminuyó cuantiosamente con el aumento del tamaño del conjunto de datos. Por otro lado, el modelo SVM presentó un equilibrio óptimo entre escalabilidad y adaptabilidad, siendo eficiente para conjuntos de datos grandes y permitiendo incorporaciones incrementales por su ágil aprendizaje (30)(34).

KNN: Tiene escalabilidad limitada, ya que su rendimiento disminuye con el aumento del volumen de datos. Su adaptabilidad es moderada, requiriendo reentrenamiento completo para incorporar nuevos datos (15).

SVM: Escalable para conjuntos de datos moderados y grandes, con adaptabilidad moderada mediante aprendizaje incremental, lo hace adecuado para sistemas de detección de fraudes en evolución continua (30).

Redes Neuronales: Ofrecen excelente escalabilidad al manejar grandes conjuntos de datos además presenta alta adaptabilidad a patrones cambiantes ya que pueden ser reentrenadas con nuevos datos (17).

Modelos Híbridos: Por lo general suelen ser más lentos que los modelos individuales sin embargo su escalabilidad y adaptabilidad dependen de dichos componentes, El modelo hibrido con los modelos de RF-LR-GB logró un tiempo de ejecución de 1 minuto, en comparación con el segundo modelo hibrido con RF-LR-SVC el cual a la complejidad del cálculo computacional llego a demorar 2 horas en propiciar un resultado. Regresando al mejor modelo hibrido RF-LR-GB

presenta eficiencia temporal, combinada con una precisión del 99.95%, sugiere una excelente escalabilidad para aplicaciones en tiempo real (35). Sin embargo, es notable que el recall (0.64) y el F1-score (0.75) son moderados, lo que indica áreas mejorables para la detección de casos positivos.

IV. DISCUSIÓN

La implementación de SMOTE y ADYSIN conocida técnicas de balanceo de clases demostró ser una pieza clave para mejorar el rendimiento de todos los modelos, especialmente en términos de recall (tasa de recuperación), lo que permite identificar correctamente los casos relevantes del conjunto de datos. Este hallazgo subraya la importancia de utilizar el desequilibrio de clases para buscar representatividad equitativa en problemas de detección de estafas, un tema que ha sido ampliamente investigado y discutido en la literatura (34,36).

Un aporte crucial resultante del estudio es sobre la interpretabilidad de los modelos. Mientras que KNN, así como también el modelo regresión lineal ofrecen la mayor interpretabilidad, las Redes Neuronales como los modelos híbridos son de difícil interpretación por su grado de complejidad y caja negra, lo que puede ser problemático en entornos monetarios altamente regulados. Los modelos individuales como KNN, Regresión Logística y SVM ofrece un equilibrio interesante entre interpretabilidad, precisión en el resultado y rendimiento. Estos hallazgos están en línea con discusiones recientes sobre la importancia de la explicabilidad en los modelos de inteligencia artificial aplicados a la detección de fraudes en tarjetas de crédito (33).

Respecto a la comparación se evidencia que el modelo híbrido (RF-LR-GB) alcanzó la precisión más alta del 99.95%), además su recall de 0.64 uno de los cuales es el menor que modelos individuales como KNN (0.9912) y SVM (0.9934), lo que sería preocupante ya que el modelo hibrido no estaría clasificando muy bien su recall. Zhang y colaboradores (22) sugieren que este compromiso entre precisión y recall es común en aplicaciones de detección de fraude debido al desequilibrio inherente presente en las clases. La implementación del modelo híbrido alternativo como fue (RF-LR-SVC) resultó computacionalmente costosa con métricas de rendimiento insuficientes o subóptimas, lo que sugeriría que no todas las combinaciones de algoritmos híbridas son equivalentemente efectivas (40).

Además, se destaca la importancia de seleccionar el modelo adecuado según el contexto de aplicación en que se requiera analizar la clasificación y entrenamiento del modelo. KNN y Regresión Logística son ideales para entornos que requieren simplicidad y rapidez en la predicción, por su simplicidad y bajos tiempos de cálculo computacional, mientras que SVM y Redes Neuronales son más apropiados para aplicaciones que priorizan la precisión y el recall, ya que requieren mayor tiempo de entrenamiento y prueba. Los modelos híbridos, como el Voting Classifier basado en Logistic Regression, Decision Tree y Gradient Boosting, demostraron tener un rendimiento superior al combinar las fortalezas de múltiples algoritmos que son ligeros en los cálculos computacionales ya que tienen buenos tiempos de entrenamiento y validación, pero a costa de mayor complejidad computacional e interpretación.

El manejo del desequilibrio de clases con SMOTE conocida técnica de sobre muestreo que genera ejemplos sintéticos de la clase minoritaria, casos con fraude y ADASYN que genera ejemplos de clase minoritaria pero que son más difíciles de aprender, es más adaptativo, lo que fue crítico para mejorar la detección de transacciones fraudulentas, especialmente en KNN y SVM. Sin embargo, los modelos híbridos mostraron una robustez frente a esta dificultad, lo que los hace atractivos para implementar en los sistemas de detección de estafas. A pesar de los avances logrados, las Redes Neuronales y los modelos híbridos enfrentan limitaciones en interpretabilidad, por su caja negra, lo que puede ser una brecha en contextos regulados, que necesitan claridad en los procesos, como lo requiere la banca. Futuras investigaciones podrían explorar técnicas de explicabilidad para modelos complejos, como SHAP o LIME, para mejorar su aceptación en entornos financieros (32)(36).

En cuanto a la eficiencia computacional, nuestros resultados revelan una compensación (trade-off) interesante entre el tiempo de entrenamiento y prueba, además de la validación cruzada de cada modelo presentado. El KNN obtuvo el tiempo de entrenamiento más corto y su alta precisión en

comparación con los demás modelos. Por otro lado, los modelos híbridos contaron con tiempo de ejecución más altos, pero mayor robustes en la precisión en comparación a los modelos individuales. Estos hallazgos tienen alcances significativos para la implementación práctica de estos modelos en sistemas de detección de fraudes en tiempo real, donde tanto la velocidad de entrenamiento como la de predicción son cruciales, para no desaprovechar el momento en que fue cometida la estafa (35).

Es importante señalar que, aunque nuestros modelos muestran un rendimiento optimo o excepcional, la detección de fraudes es un problema presente en constante evolución. Los estafadores continuamente adaptan sus tácticas que evolucionan con el pasar del tiempo, lo que significa que incluso los modelos más sofisticados de aprendizaje automático y de inteligencia artificial deben ser regularmente actualizados y reentrenados, por la mejora continua y el control de la calidad en las transacciones (32).

Una limitación de nuestro estudio se presenta por el conjunto de datos estático debido al corte transversal en el tiempo, el conjunto de datos fueron transacciones que ocurrieron en 2024 obtenidos del sitio web Kaggle. En un escenario del mundo real, los patrones de fraude varían y evolucionan constantemente debido a los nuevos avances tecnológicos, lo que requeriría un enfoque de aprendizaje continuo. Futuros estudios podrían explorar la implementación de estos modelos en un entorno de aprendizaje en línea con series temporales, donde los modelos se actualicen y entrenen continuamente con nuevos datos seriales (34).

V. CONCLUSIONES

Los modelos individuales mostraron un rendimiento excepcional como KNN y SVM, presentan ser más equilibrados entre precisión (99.94%, 99.92%) y recall (0.9912, 0.9934) respectivamente, con tiempos de ejecución significativamente menores (1 segundo, 5 segundos), destacándose como alternativas fáciles y eficientes para implementaciones en tiempo real, En cuando a los modelos híbridos, como el Voting Classifier basado en Logistic Regression, Decision Tree y Gradient Boosting, demostraron ser una estrategia viable en la robustes del resultado, alcanzando una precisión del 99.95% en la configuración RF-LR-GB, aunque con limitaciones en términos de recall con 0.64.Respecto a la eficiencia computacional, los modelos como Logistic Regression y KNN destacaron por su rapidez y simplicidad, haciéndolos ideales para entornos que presenten recursos limitados, además exhiben buena precisión y un excelente recall.

En cuanto al manejo del desequilibrio de clases, las técnicas de sobre muestreo, como SMOTE y ADASYN, fueron esenciales para mejorar la representación de casos fraudulentos, debido a que fueron los casos con minoría del conjunto de datos, estas técnicas contribuyen a un mejor rendimiento general de todos los modelos, especialmente en KNN y Redes Neuronales para que no se sobre ajusten.

La Interpretabilidad, aunque los modelos simples como Logistic Regression y KNN ofrecen alta interpretabilidad, los modelos complejos como Redes Neuronales y los híbridos requieren avances en explicabilidad como es el uso de SHAP para su adopción en entornos regulados, en el que se requiere mayor transparencia del proceso interior del algoritmo. Por ello al implementar este método demostró ser una herramienta valiosa para comprender las decisiones del modelo hibrido, aspecto crucial en el sector monetario donde la transparencia es esencial. Las Redes Neuronales mostraron la mejor adaptabilidad a patrones cambiantes, mientras que SVM y los modelos híbridos ofrecieron un buen balance entre rendimiento y escalabilidad.

Basándonos en estos hallazgos, recomendamos lo siguiente:

1. Para Implementación:

- En los escenarios donde sea prescindible la precisión estricta y el tiempo del procesamiento del conjunto de datos no sea una limitación severa, por eso se recomendaría el uso del modelo híbrido RF-LR-GB.

- En el caso de situaciones de tiempo real donde se priorice el balance entre precisión interpretabilidad y velocidad de respuesta se debe considerar KNN o SVM.

- Evitar la implementación del modelo híbrido RF-LR-SVC debido a su alto costo

computacional y bajo rendimiento, no todas las combinaciones tienen un buen rendimiento de cómputo.

2. Para Mejoras Futuras:

- Investigar otras técnicas adicionales sobre el balanceo de datos para el caso de modelos híbridos para conseguir un mejor recall, ya que presentaron ser bajos en la clasificación de los casos minoritarios, lo que provocaría la perdida de cuestiones de cometimiento de fraude.
- Explorar la optimización de hiperparámetros en nuestro estudio se utilizó el 70% de entrenamiento y 30% para testear, lo que se busca es reducir los tiempos de ejecución sin comprometer la precisión.
- Desarrollar un sistema de monitoreo para series temporales para complementar los modelos en tiempos continuos en la detección y adaptación de nuevos patrones de fraude en tiempo real.

3. Para Escalabilidad:

- Implementar un sistema de procesamiento por lotes en el caso de conjuntos de datos grandes, por su alto grado de cálculo computacional.
- Establecer umbrales de confianza adaptables, para mejorar la calidad, según el contexto de la transacción o del entorno financiero.

4. Para Mantenimiento:

- Realizar actualizaciones periódicas y nuevos experimentos con distintas combinaciones de modelos híbridos, así como también testearlos con nuevos conjuntos de datos.
- Documentar los nuevos procesos inherentes para una fácil replicación y mantenimiento.
- Establecer protocolos de calidad en la validación continua para asegurar la consistencia del rendimiento.

VI. AGRADECIMIENTOS

"Agradezco profundamente a la Universidad Central del Ecuador por brindarme las oportunidades académicas y el apoyo necesario para desarrollar esta investigación. Asimismo, extiendo mi gratitud a los profesores y compañeros que, con su conocimiento y colaboración, han contribuido de manera significativa a este trabajo. Su dedicación a la creación y difusión del conocimiento ha sido fundamental en mi formación, y me compromete a seguir aportando para construir un mundo mejor."

VII. REFERENCIAS

1. Nilson Report. Card fraud losses reach $32.34 billion [Internet]. 2023. Available from: https://nilsonreport.com/article_archive_id=4161.
2. Asociación de Bancos Privados del Ecuador. Boletín macroeconómico - marzo 2024 [Internet]. Quito: ASOBANCA; 2024. Available from: https://asobanca.org.ec/wp-content/uploads/2024/03/Boletin-macroeconomico-Marzo-2024.pdf.
3. Asociación de Bancos Privados del Ecuador. La era de la banca digital en Ecuador [Internet]. Quito: ASOBANCA; 2023. Available from: https://asobanca.org.ec/wp-content/uploads/2023/07/La-era-de-la-banca-digital-en-Ecuador-2.pdf.
4. Abdallah A, Maarof MA, Zainal A. Fraud detection system: A survey. J Netw Comput Appl. 2016;68:90-113. Available from: https://doi.org/10.1016/j.jnca.2016.04.007.
5. Carcillo F, Le Borgne YA, Caelen O, Kessaci Y, Oblé F, Bontempi G. Combining unsupervised and supervised learning in credit card fraud detection. Inf Sci (Ny). 2021;557:317-31. Available from: https://doi.org/10.1016/j.ins.2019.05.042.
6. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. J Artif Intell Res. 2002;16:321-57. Available from: https://doi.org/10.1613/jair.953.
7. Fawcett T. An introduction to ROC analysis. Pattern Recognit Lett. 2006;27(8):861-74. Available

from: https://doi.org/10.1016/j.patrec.2005.10.010.

8. Dal Pozzolo A, Caelen O, Johnson RA, Bontempi G. Calibrating probability with undersampling for unbalanced classification. In: 2015 IEEE Symposium Series on Computational Intelligence. IEEE; 2015. p. 159-66. Available from: https://doi.org/10.1109/SSCI.2015.33.
9. Pozzolo AD, Caelen O, Le Borgne YA, Waterschoot S, Bontempi G. Learned lessons in credit card fraud detection from a practitioner perspective. Expert Syst Appl. 2014;41(10):4915-28. Available from: https://doi.org/10.1016/j.eswa.2014.02.026.
10. Jurgovsky J, Granitzer M, Ziegler K, Calabretto S, Portier PE, He-Guelton L, Caelen O. Sequence classification for credit-card fraud detection. Expert Syst Appl. 2018;100:234-45. Available from: https://doi.org/10.1016/j.eswa.2018.01.037.
11. Candès EJ, Li X, Ma Y, Wright J. Robust principal component analysis?. J ACM. 2011;58(3):1-37. Available from: https://doi.org/10.1145/1970392.1970395.
12. Lee CW, Fu MW, Wang CC, Azis MI. Evaluating machine learning algorithms for financial fraud detection: insights from Indonesia. Mathematics [Internet]. 2025;13(4):600. Available from: https://doi.org/10.3390/math13040600.
13. More A. Survey of resampling techniques for improving classification performance in unbalanced datasets [Preprint]. arXiv:1608.06048 [Internet]. 2016. Available from: https://arxiv.org/abs/1608.06048.
14. Zhu X, Wang H, Xu L, Li H. Predicting stock prices by using a hybrid model of ARIMA and KNN. Neural Comput Appl. 2019;31(8):3893-904. Available from: https://doi.org/10.1007/s00521-017-3288-x.
15. Breunig MM, Kriegel HP, Ng RT, Sander J. LOF: identifying density-based local outliers. In: Proceedings of the 2000 ACM SIGMOD international conference on Management of data. ACM; 2000. p. 93-104. Available from: https://doi.org/10.1145/342009.335388.
16. Cortes C, Vapnik V. Support-vector networks. Mach Learn. 1995;20(3):273-97. Available from: https://doi.org/10.1007/BF00994018.
17. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436-44. Available from: https://doi.org/10.1038/nature14539.
18. Patra P, Vedansh S, Ved V, Singh A, Mishra S, Kumar A. A sampling-based logistic regression model for credit card fraud estimation. In: Swaroop A, Polkowski Z, Correia SD, Virdee B, editors. Proceedings of Data Analytics and Management. ICDAM 2023. Lecture Notes in Networks and Systems, vol 788. Singapore: Springer; 2023. p. 209-21. Available from: https://doi.org/10.1007/978-981-99-6553-3_16.
19. Mohammed U, Wajiga GM, Nata'ala A, Abdullahi BM. Comparative analysis of Random Forest and Logistic Regression models for detecting fraud in bank transactions based on performance metrics. Res J Pure Sci Technol. 2024;7(4):1-12. Available from: https://doi.org/10.56201/rjpst.v7.no4.2024.pg1.12.
20. Jose NN, Arigela AK, Vivekanandan G, Ravikumar S, Naganathan SBT, Venu N. Optimizing payment transaction security: utilizing gradient boosting machines for fraud detection. In: 2024 10th International Conference on Communication and Signal Processing (ICCSP); 2024 Apr; [ciudad]. Available from: https://doi.org/10.1109/ICCSP60870.2024.10543774.
21. Johnson P, et al. Scalable fraud detection systems using hybrid architectures. Appl Soft Comput. 2024;112:108872.
22. Zhang K, Wu L, Sun Y. Performance analysis of hybrid models in imbalanced datasets. Expert Syst Appl. 2023;185:115648.
23. Bergstra J, Bengio Y. Random search for hyper-parameter optimization. J Mach Learn Res. 2012;13(2):281-305. Available from: https://www.jmlr.org/papers/volume13/bergstra12a/bergstra12a.pdf.
24. Snoek J, Larochelle H, Adams RP. Practical Bayesian optimization of machine learning algorithms.

Adv Neural Inf Process Syst. 2012. p. 2951-9. doi:10.48550/arXiv.1206.2944.

25. Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: Proceedings of the 14th International Joint Conference on Artificial Intelligence; 1995. p. 1137-45.
26. Bradley AP. The use of the area under the ROC curve in the evaluation of machine learning algorithms. Pattern Recognit. 1997;30(7):1145-59.
27. Ngai EWT, Hu Y, Wong YH, Chen Y, Sun X. The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature. Decis Support Syst. 2011;50(3):559-69. Available from: https://doi.org/10.1016/j.dss.2010.08.006.
28. Bhattacharyya S, Jha S, Tharakunnel K, Westland JC. Data mining for credit card fraud: A comparative study. Decis Support Syst. 2011;50(3):602-13. Available from: https://doi.org/10.1016/j.dss.2010.08.008.
29. Pelegrina GD, Duarte LT, Grabisch M. A k-additive Choquet integral-based approach to approximate the SHAP values for local interpretability in machine learning [Preprint]. arXiv:2211.02166. 2022.
30. Kou Y, Lu CT, Sirwongwattana S, Huang YP. Survey of fraud detection techniques. In: Proceedings of the IEEE International Conference on Networking, Sensing and Control; 2004 Mar 21-23; Taipei, Taiwan. Piscataway (NJ): IEEE; 2004. p. 749-54. doi:10.1109/ICNSC.2004.1297040.
31. Aleskerov E, Freisleben B, Rao B. CARDWATCH: A neural network-based database mining system for credit card fraud detection. In: Proceedings of the IEEE/IAFE 1997 Computational Intelligence for Financial Engineering; 1997. p. 220-6. doi:10.1109/CIFER.1997.618940.
32. Phua C, Lee V, Smith K, Gayler R. A comprehensive survey of data mining-based fraud detection research [Preprint]. arXiv:1009.6119. 2010.
33. Lucas Y, Protopopescu A, Lemaire V, Velcin J, Sidibe A. Towards automated feature engineering for credit card fraud detection using multi-perspective HMMs. Future Gener Comput Syst. 2020;102:393-402. Available from: https://doi.org/10.1016/j.future.2019.08.007.
34. Randhawa K, Jain S, Singh G. Credit card fraud detection using AdaBoost and majority voting. Procedia Comput Sci. 2018;132:1049-57. Available from: https://doi.org/10.1016/j.procs.2018.05.219.
35. West J, Bhattacharya M. Intelligent financial fraud detection: A comprehensive review. Comput Secur. 2016;57:47-66. Available from: https://doi.org/10.1016/j.cose.2015.09.005.
36. Jiang C, Song H, Wang J, Han Z, Li L. A hybrid fraud detection method in credit card transactions based on dynamic selection of base classifiers. Clust Comput. 2019;22(4):8353-68. Available from: https://doi.org/10.1007/s10586-017-1589-4.
37. Smith J, Brown R. Hybrid models in fraud detection: A comprehensive review. J Mach Learn Res. 2023;15(3):145-60.
38. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. In: Advances in Neural Information Processing Systems 30 (NIPS 2017); 2017 Dec 4-9; Long Beach, CA, USA. doi:10.48550/arXiv.1705.07874.
39. Wang Y, Chen H, Li X. Understanding financial fraud through explainable AI. IEEE Trans Neural Netw Learn Syst. 2023;34(2):892-903.
40. Martinez R, Thompson E. Computational efficiency in modern fraud detection systems. J Big Data. 2023;10(1):45-62.