Maestría en

CIENCIA DE DATOS Y MÁQUINAS DE APRENDIZAJE CON MENCIÓN EN INTELIGENCIA ARTIFICIAL

Trabajo previo a la obtención de título de Magister en Ciencia de Datos y Máquinas de Aprendizaje con mención en Inteligencia Artificial

AUTORES:

Brigith Angelina Ullauri Guaranga

Rodrigo Iván Ullauri Guaranga

Darwin Rolando Ipiales Bunci

Joffre Enrique Velasco Robalino

TUTOR/ES:

Alejandro Cortés López

Iván Reyes

TEMA

Detección de fraudes en cuentas bancarias a través del análisis de transacciones financieras mediante modelos de IA y machine learning para mejorar la seguridad en instituciones bancarias del Ecuador.

Quito - Ecuador

REINVENTEMOS EL FUTURO

Certificación de Autoría

Nosotros, Brigith Angelina Ullauri Guaranga, Rodrigo Iván Ullauri Guaranga, Darwin Rolando Ipiales Bunci y Joffre Enrique Velasco Robalino**,** declaramos bajo juramento que el trabajo aquí descrito es de nuestra autoría; que no ha sido presentado anteriormente para ningún grado o calificación profesional y que se ha consultado la bibliografía detallada.

Cedemos nuestros derechos de propiedad intelectual a la Universidad Internacional del Ecuador (UIDE), para que sea publicado y divulgado en internet, según lo establecido en la Ley de Propiedad Intelectual, su reglamento y demás disposiciones legales.

Firma del graduado

Brigith Angelina Ullauri Guaranga

Firma del graduado

Rodrigo Iván Ullauri Guaranga

Firma del graduado

Darwin Rolando Ipiales Bunci

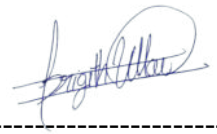
Firma del graduado

Joffre Enrique Velasco Robalino

Autorización de Derechos de Propiedad Intelectual

Nosotros, Brigith Angelina Ullauri Guaranga, Rodrigo Iván Ullauri Guaranga, Darwin Rolando Ipiales Bunci y Joffre Enrique Velasco Robalino, en calidad de autores del trabajo de investigación titulado "*Detección de fraudes en cuentas bancarias a través del análisis de transacciones financieras mediante modelos de IA y machine learning para mejorar la seguridad en instituciones bancarias del Ecuador"*, autorizamos a la Universidad Internacional del Ecuador (UIDE) para hacer uso de todos los contenidos que nos pertenecen o de parte de los que contiene esta obra, con fines estrictamente académicos o de investigación. Los derechos que como autores nos corresponden, lo establecido en los artículos 5, 6, 8, 19 y demás pertinentes de la Ley de Propiedad Intelectual y su Reglamento en Ecuador.

D. M. Quito, (julio 2025)



Firma del graduado

Brigith Angelina Ullauri Guaranga

Firma del graduado

Rodrigo Iván Ullauri Guaranga

Firma del graduado

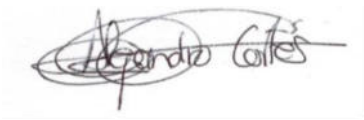
Darwin Rolando Ipiales Bunci

Firma del graduado

Joffre Enrique Velasco Robalino

Aprobación de Dirección y Coordinación del Programa

Nosotros, Alejandro Cortés López Director EIG e Iván Reyes Coordinador UIDE, declaramos que: Brigith Angelina Ullauri Guaranga, Rodrigo Iván Ullauri Guaranga, Darwin Rolando Ipiales Bunci y Joffre Enrique Velasco Robalino son los autores exclusivos de la presente investigación y que ésta es original, auténtica y personal de ellos.



Alejandro Cortés López

Director de la

Maestría en Ciencia de Datos y Máquinas de Aprendizaje con mención en Inteligencia Artificial

---

Iván Reyes Chacón

Coordinador/a de la

Maestría en Ciencia de Datos y Máquinas de Aprendizaje con mención en Inteligencia Artificial

Dedicatoria

Brigith Angelina Ullauri Guaranga.

Dedico este trabajo con profundo amor y gratitud a mi familia, quienes con su esfuerzo y apoyo incondicional impulsan cada uno de mis logros y por estar siempre presente en cada etapa de este camino. Este proyecto representa no solo un avance académico, sino también el compromiso con el desarrollo de soluciones que protejan a nuestra sociedad de amenazas tecnológicas como el fraude financiero.

Rodrigo Iván Ullauri Guaranga.

A Dios, por darme la sabiduría y fortaleza necesarias para culminar este proceso. A mi familia y amigos queridos, por su constante respaldo y palabras de aliento. Dedico este esfuerzo a los profesionales que luchan por una banca más segura y confiable para todos los ecuatorianos.

Darwin Rolando Ipiales Bunci.

Dedico este trabajo, con profundo amor y gratitud, a mi madre, una mujer ejemplar, valiente y perseverante, cuyo esfuerzo, sabiduría y amor incondicional han sido mi mayor fuente de inspiración. Gracias por enseñarme con tu ejemplo el valor del sacrificio y la importancia de nunca rendirse, este logro es tan suyo como mío. También lo dedico a mis amigos de maestría, por su compromiso y apoyo en esta etapa académica, además de que, este trabajo sirva como un paso más hacia una banca más segura y respaldada por soluciones basadas en ciencia de datos e inteligencia artificial.

Joffre Enrique Velasco Robalino.

A mi esposa Lili y a mis hijas Antonela y Mile, porque siempre cuento con su respaldo y apoyo en los momentos difíciles y mi alegría en los triunfos. Gracias por su paciencia,

comprensión y amor. También lo dedico a los profesionales del sector bancario que día a día enfrentan desafíos complejos en la lucha contra el fraude.

Agradecimiento

Queremos comenzar expresando nuestra más profunda gratitud a Dios, fuente de sabiduría, fortaleza y esperanza, por guiarnos en cada etapa de este proyecto.

A nuestras familias, por su amor incondicional, su apoyo, por las palabras de ánimo y por ser nuestro pilar emocional durante todo este proceso. Este logro también es suyo.

A nuestros docentes y tutores, quienes nos acompañaron con su profesionalismo y compromiso, gracias por compartir su conocimiento, por exigirnos dar lo mejor de nosotros y por guiarnos en el desarrollo de un trabajo con sentido y relevancia para la sociedad ecuatoriana.

Finalmente, agradecemos a todas las personas e instituciones que, de una u otra manera, contribuyeron a que este proyecto se hiciera realidad. Confiamos en que el conocimiento aquí generado será útil para mejorar la seguridad financiera en nuestro país y contribuirá al desarrollo tecnológico con propósito y responsabilidad.

Resumen

El fraude financiero no solo impacta a las entidades bancarias en términos de pérdidas económicas, sino que también afecta la confianza de los clientes y la estabilidad del sistema financiero. En Ecuador, los reportes de la Asociación de Bancos Privados del Ecuador (ASOBANCA, 2023) indican que el fraude electrónico representa el 75% de los incidentes de seguridad financiera, lo que subraya la importancia de estrategias innovadoras para la detección y prevención de estos eventos. Los métodos tradicionales de detección de fraude, basados en reglas heurísticas y revisiones manuales, presentan limitaciones en términos de escalabilidad y capacidad de adaptación a nuevas tácticas fraudulentas. La implementación de técnicas de ciencia de datos, como análisis de series temporales, modelos de clasificación basados en redes neuronales convolucionales y arquitecturas de aprendizaje profundo, permiten una detección más efectiva y oportuna, reduciendo la exposición al riesgo. Dado que el acceso a datos transaccionales reales está restringido por normativas de privacidad y confidencialidad, este proyecto empleará datasets sintéticos generados con algoritmos de modelado generativo, así como fuentes abiertas de datos financieros. Esto permitirá validar el desempeño de los modelos sin comprometer la integridad y confidencialidad de la información financiera real, garantizando su aplicabilidad en entornos productivos. La contribución de este proyecto radica en el desarrollo de un modelo adaptable al contexto bancario ecuatoriano, con un enfoque replicable a otras instituciones financieras. Se espera que los resultados permitan diseñar un sistema de detección temprana de fraude que optimice la seguridad bancaria y reduzca los riesgos asociados a transacciones fraudulentas.

Palabras Claves: Fraude bancario, Modelos predictivos, Modelos supervisados, Detección de anomalías, Transacciones financieras.

Abstract

Financial fraud not only impacts banking institutions in terms of economic losses, but also affects customer trust and the stability of the financial system. In Ecuador, reports from the Association of Private Banks of Ecuador (ASOBANCA, 2023) indicate that electronic fraud accounts for 75% of financial security incidents, highlighting the importance of innovative strategies for detecting and preventing such events. Traditional fraud detection methods, based on heuristic rules and manual reviews, present limitations in terms of scalability and adaptability to new fraudulent tactics. The implementation of data science techniques, such as, time series analysis, classification models based on convolutional neural networks and deep learning architectures, allows for more effective and real-time detection, reducing risk exposure. Given that access to real transactional data is restricted by privacy and confidentiality regulations, this project will use synthetic datasets generated with generative modeling algorithms, as well as open financial data sources available on platforms like Kaggle and IEEE-CIS Fraud Detection. This will enable us to validate the performance of the models without compromising the integrity of real financial information, ensuring their applicability in production environments. The contribution of this project lies in the development of a model adaptable to the context of the Ecuadorian banking sector, with an approach that can be replicated by other financial institutions at the region. The results are expected to support the design of an early fraud detection system that optimizes banking security and reduces the risks associated with fraudulent transactions.

Keywords:

Bank fraud, Supervised models, Predictive models, Anomaly detection, Financial transactions.

Tabla de Contenidos

Lista de Tablas

Lista de Figuras

Capítulo 1

1. Introducción

1.1. Definición del Proyecto

Las transacciones con tarjetas de crédito han aumentado debido a los rápidos avances tecnológicos y la comodidad de los servicios electrónicos (Lebichot, Paldino, Siblini, He-Guelton, Oblé y Bontempi, 2021), (Zhang, Han, Xu y Wang, 2021); y, en consecuencia, se ha producido un aumento de los problemas de seguridad, como el fraude con tarjetas de crédito, que se ha convertido en una preocupación importante tanto para las instituciones financieras como para los clientes (Bakhtiari, Nasiri y Vahidi, 2023), (Yang, Luo, Vijayalakshmi y Shalinie, 2022). El fraude financiero representa una amenaza global significativa para la estabilidad y confiabilidad del sector bancario, especialmente en entornos donde las transacciones electrónicas han tenido un crecimiento exponencial en los últimos años. Según el informe de Nielsen, las pérdidas por fraude con tarjetas de crédito en 2019, 2020 y 2021 ascendieron a aproximadamente 28.650, 28.500 y 32.340 millones de dólares, respectivamente (Wang, Liu, Kou, Xiao, Wang y Tang, 2023), (El-Naby, Hemdan y El-Sayed, 2023), (Alarfaj, Malik, Khan, Almusallam, Ramzan y Ahmed, 2022). Además, las pérdidas por fraude con tarjetas de crédito a nivel mundial se han triplicado en la última década, pasando de 9.840 millones de dólares en 2011 a 32.340 millones de dólares en 2021 (Islam, Uddin, Aryal y Stea, 2023). En Ecuador, el incremento en el uso de plataformas digitales ha generado un escenario favorable para la proliferación de patrones fraudulentos, incluyendo la suplantación de identidad, fraudes en transferencias electrónicas y transacciones no autorizadas (ASOBANCA, 2023). Según la Superintendencia de Bancos del Ecuador (2022), los fraudes financieros electrónicos representaron un incremento del 40% en los últimos tres años, evidenciando la necesidad de implementar estrategias robustas de análisis,

detección y prevención (Superintendencia de Bancos del Ecuador, 2022).

En este contexto, la ciencia de datos y el aprendizaje automático han demostrado ser herramientas altamente efectivas para la identificación temprana de actividades fraudulentas mediante la detección de patrones anómalos o inusuales en transacciones financieras.

1.2.Justificación e Importancia del Trabajo de Investigación

El fraude financiero no solo impacta a las entidades bancarias en términos de pérdidas económicas, sino que también afecta la confianza de los clientes y la estabilidad del sistema financiero. En Ecuador, los reportes de la Asociación de Bancos Privados del Ecuador (ASOBANCA, 2023) indican que el fraude electrónico representa el 75% de los incidentes de seguridad financiera, lo que subraya la importancia de estrategias innovadoras para la detección y prevención de estos eventos (ASOBANCA, 2023).

Dado que el acceso a datos transaccionales reales está restringido por normativas de privacidad y confidencialidad, este proyecto considera datasets sintéticos generados con algoritmos de modelado generativo o datos anonimizados mediante PCA, así como fuentes abiertas de datos financieros disponibles en plataformas como Kaggle y IEEE-CIS Fraud Detection. Esto nos permitirá validar el desempeño de los modelos sin comprometer la integridad de la información financiera real y de esta manera garantizamos su aplicabilidad en entornos productivos.

Los métodos tradicionales de detección de fraude, basados en reglas heurísticas y revisiones manuales, presentan limitaciones en términos de escalabilidad y capacidad de adaptación a nuevas tácticas fraudulentas. La implementación de técnicas de ciencia de datos, como análisis de series temporales, modelos de clasificación, redes neuronales y arquitecturas de aprendizaje profundo, permiten una detección más efectiva y en tiempo real, reduciendo la

exposición al riesgo.

Además, el uso de técnicas de detección de anomalías mediante modelos de clustering como DBSCAN o K-Means, puede complementar los enfoques supervisados, permitiendo la identificación de patrones de comportamiento inusuales en grandes volúmenes de datos financieros (Zhang, Han, Xu y Wang,2021).

El enfoque de este proyecto radica en la utilización de un conjunto de datos de acceso abierto que contine alrededor de 284 mil datos sintéticos, además el desarrollo de un modelo de detección de fraude que mejore la seguridad bancaria en las transacciones con tarjetas de crédito permitiendo reducir los riesgos asociados con escenarios fraudulentos.

1.3.Alcance

Este trabajo contempla la implementación y evaluación de modelos de clasificación aplicados a datos transaccionales con el objetivo de detectar fraudes en operaciones con tarjetas de crédito. Debido a las restricciones legales sobre el uso de datos reales, se utilizarán fuentes de datos abiertas y sintéticas que simulen escenarios reales. El estudio abarca la evaluación tanto de algoritmos de aprendizaje supervisado como no supervisado, con el fin de identificar la mejor aproximación para detectar patrones anómalos en las transacciones. Además, se desarrollará un modelo de detección que contribuya a mejorar la seguridad bancaria mediante la reducción de fraudes y la minimización de falsos negativos.

1.4.Objetivos

1.4.1. Objetivo General.

El objetivo de este proyecto es desarrollar un modelo de clasificación basado en aprendizaje automático para la detección de fraude en transacciones bancarias con tarjetas de crédito, priorizando la reducción de transacciones fraudulentas no detectadas y asegurando la

adaptabilidad, escalabilidad y replicabilidad del modelo; empleando metodologías de ciencia de datos.

1.4.2. Objetivos Específicos.

- Identificar conjuntos de datos sintéticos y fuentes de acceso abierto con información relevante al fraude con tarjetas de crédito, y así para garantizar la calidad y relevancia de los datos.
- Evaluar distintos algoritmos y determinar la combinación óptima que permita clasificar las transacciones fraudulentas y no fraudulentas en transacciones electrónicas con tarjetas de crédito.
- Evaluar el desempeño de los modelos de clasificación desarrollados mediante métricas y detección de falsos negativos, con el fin de determinar su efectividad en la detección de fraudes en transacciones con tarjetas de crédito.
- Proponer un modelo de detección de fraude orientado a minimizar los falsos negativos, optimizando la sensibilidad del sistema, y que esté preparado para su integración y despliegue en una solución tecnológica.

Capítulo 2

2. Revisión de literatura

2.1.Estado del arte

El fraude en transacciones con tarjetas de crédito representa un desafío creciente para las instituciones financieras, dada la expansión del comercio electrónico y los pagos digitales. Ante esta situación, se ha incrementado la investigación en métodos automatizados de detección que permitan identificar comportamientos anómalos con alta precisión y mínima tasa de falsos positivos (Bhattacharyya, Jha, Tharakunnel y Westland, 2011). Los métodos de aprendizaje

automático (ML) se han utilizado ampliamente para la detección de fraudes con tarjetas de crédito (CCFD), alcanzando rendimientos muy importantes (Dang, Tran, Tuan y Tiep, 2021), (Alfaiz y Fati, 2022), (Mienye y Jere, 2024). Los algoritmos de ML se pueden clasificar en aprendizaje supervisado, no supervisado, semisupervisado o de refuerzo (Dong, Xia y Peng, 2021). El método de ML más utilizado para identificar fraudes con tarjetas de crédito es el aprendizaje supervisado (SL) (Btoush, Zhou, Gururajan, Chan, Genrich y Sankaran, 2023). El aprendizaje supervisado implica entrenar un algoritmo de ML utilizando un conjunto de datos donde cada punto de datos tiene una etiqueta. La etiqueta indica la clase específica a la que pertenece el punto de datos, como fraude o no fraude. Las técnicas de SL tienden a aprender la relación entre las características de entrada (o variables independientes) y las etiquetas de salida (variables dependientes).

Inicialmente, los enfoques tradicionales para la detección de fraude se basaban en reglas definidas manualmente y en algoritmos estadísticos, como regresión logística, máquinas de vectores de soporte (SVM) y árboles de decisión. Sin embargo, estas técnicas muestran limitaciones en entornos de datos desbalanceados y ante patrones de fraude dinámicos y cada vez más sofisticados (Sahin y Duman, 2011).

Varios estudios han demostrado la capacidad de las redes neuronales para identificar transacciones fraudulentas en datos complejos de tarjetas de crédito (Bin Sulaiman, Schetinin y Sant, 2022), (Osegi y Jumbo, 2021). Una red neuronal es un tipo de aprendizaje automático con un proceso que imita el cerebro humano y puede ser supervisado o no supervisado (Wang, Fan y Wang, 2021). Las redes neuronales con múltiples capas, también llamadas aprendizaje profundo (DL), pueden extraer progresivamente características de alto nivel y analizar patrones complejos con predicciones mejoradas. Se han utilizado enfoques de DL para identificar transacciones

fraudulentas en datos de tarjetas de crédito. Por ejemplo, Mienye y Sun (Mienye y Sun, 2023), desarrollaron un enfoque para la detección de fraudes con tarjetas de crédito utilizando un conjunto apilado de redes de memoria a largo plazo (LSTM) y unidades recurrentes cerradas (GRU), con un perceptrón multicapa (MLP) como aprendiz base. De manera similar, senogho, Mienye, Swart, Aruleba y Obaido (Esenogho, Mienye, Swart, Aruleba y Obaido, 2022), propusieron un enfoque basado en aprendizaje automático (DL) para la detección de fraudes con tarjetas de crédito utilizando la red neuronal LSTM como aprendiz base en la implementación de refuerzo adaptativo (AdaBoost), logrando un excelente rendimiento en la clasificación. Además, diversos estudios han utilizado redes neuronales convolucionales (Gambo, Zainal y Kassim, 2022), (Berhane, Melese, Walelign y Mohammed, 2023).

Mientras tanto, las redes neuronales recurrentes (RNN) y sus variantes, como LSTM y GRU, son las redes basadas en DL más utilizadas para modelar y analizar transacciones con tarjetas de crédito debido a su capacidad para aprender datos secuenciales y detectar relaciones temporales (Sehrawat y Singh, 2023), (Raval, Bhattacharya, Jadav, Tanwar, Sharma, Bokoro, Elmorsy, Tolba y Raboaca, 2023), (Benchaji, Douzi, El Ouahidi y Jaafari, 2021). La red LSTM es útil para aprender dependencias a largo plazo en una secuencia. Es potente porque puede recordar información de intervalos de tiempo anteriores y olvidar o actualizar selectivamente dicha información a medida que se procesan nuevas entradas. Al igual que LSTM, GRU puede actualizar u omitir selectivamente datos de intervalos de tiempo anteriores gracias a su mecanismo de control, lo que lo hace adecuado para el modelado de series de tiempo. En estos últimos años, el uso de algoritmos de aprendizaje profundo ha cobrado relevancia por su capacidad de extraer representaciones complejas y adaptarse al gran volumen de datos financieros. De acuerdo con Nguyen (Nguyen, Phan y Tran, 2024), las redes neuronales

profundas (DNN), redes recurrentes LSTM, redes GRU y redes convolucionales (CNN) han demostrado alto desempeño en la clasificación de transacciones fraudulentas, especialmente en conjuntos de datos etiquetados (Nguyen, Phan y Tran, 2024). Sin embargo, a pesar de la robustez de las técnicas de aprendizaje profundo, su uso para la detección de fraudes con tarjetas de crédito presenta ciertas ventajas y limitaciones.

En paralelo, se han explorado modelos no supervisados que no requieren datos etiquetados, lo cual es particularmente útil en contextos donde es difícil acceder a registros verificados de fraude. Entre estos destacan los autoencoders, que permiten detectar anomalías mediante el error de reconstrucción, e Isolation Forest, que identifica outliers mediante un enfoque de aislamiento estadístico (Nguyen, Pham y Tran, 2024). El uso combinado de modelos supervisados y no supervisados ha sido propuesto como un enfoque robusto para mejorar la precisión y generalización de los sistemas de detección (Roy, Sun, Mahoney y Harms, 2018).

Debido a restricciones legales y éticas en el uso de datos reales, muchos estudios emplean conjuntos sintéticos o bases de datos abiertas, como los proporcionados por la comunidad de investigación (por ejemplo, el dataset de fraude de Kaggle o el Synthetic Financial Dataset For Fraud Detection). Estas fuentes permiten simular escenarios realistas y evaluar modelos sin comprometer información sensible (Carcillo, Le Borgne, Caelen y Bontempi, 2018).

Finalmente, las métricas de evaluación juegan un papel crucial para validar el rendimiento de los modelos. En contextos con clases desbalanceadas, métricas como precision, recall y F1-score son preferidas sobre la simple exactitud, ya que reflejan mejor la capacidad del modelo para detectar casos de fraude sin generar una cantidad excesiva de falsos positivos (Nguyen, Pham y Tran, 2024).

2.2.Marco teórico

2.2.1. Fraude en transacciones con tarjetas de crédito.

El fraude con tarjetas de crédito ocurre cuando un usuario no autorizado obtiene acceso a los datos de la tarjeta de crédito de otra persona y realiza transacciones (Gold, 2014). Si bien las transacciones se realizan frecuentemente en línea, también pueden realizarse con la tarjeta en caso de extravío o robo. Los estafadores utilizan diferentes métodos para obtener la información del titular de la tarjeta, incluyendo el phishing, donde un estafador se hace pasar por un funcionario financiero para obligar al usuario a revelar información personal, y los skimmers utilizan una interfaz con un cajero automático o dispositivo de punto de venta que puede leer la tarjeta directamente (Ambashtha y Kumar, 2023), (Guers, Chowdhury y Rifat, 2022). Este tipo de delito representa una amenaza significativa para la banca y los consumidores, debido al incremento del comercio digital y la sofisticación de las técnicas de suplantación de identidad. El objetivo de los sistemas de detección de fraude es identificar transacciones inusuales que se desvíen del comportamiento legítimo, minimizando al mismo tiempo las falsas alarmas que puedan afectar negativamente la experiencia del usuario. Detectar el fraude con tarjetas de crédito es esencial para garantizar la seguridad de las finanzas y la información financiera de los consumidores.

Los sistemas automatizados son más populares debido a su capacidad para procesar grandes volúmenes de datos de forma rápida y eficiente, a través de modelos estadísticos y de aprendizaje automático avanzados (Van Belle, Baesens y De Weerdt, 2023). Además, los algoritmos de aprendizaje automático, incluidas las redes neuronales, se emplean ampliamente para detectar fraudes con tarjetas de crédito. Por ejemplo, Mienye y Sun (Mienye y Sun, 2023) propusieron un método CCFD que utiliza una selección híbrida de características basada en

algoritmos genéticos y ganancia de información, y el algoritmo de aprendizaje fue la máquina de aprendizaje extremo (ELM). La función de aptitud del algoritmo genético empleada en el estudio fue la media geométrica, que se utilizó para abordar el problema de desequilibrio de clases, lo que mejoró el rendimiento de la clasificación. De igual manera, Karthik (Karthik, Mishra y Reddy, 2022), propusieron un enfoque de conjunto híbrido para la detección de fraudes con tarjetas de crédito con el fin de resolver el problema de desequilibrio de clases. El estudio combinó métodos de boosting y bagging, es decir, boosting adaptativo (AdaBoost) y bosque aleatorio, respectivamente, logrando un rendimiento superior en comparación con los clasificadores individuales. Otros ejemplos de algoritmos de aprendizaje automático (ML) para la detección de fraudes con tarjetas de crédito incluyen bosque aleatorio (Aburbeian y Ashqar, 2023), XGBoost (Kafhali y Tayebi, 2022), red neuronal convolucional (CNN) (Illanko, Soleymanzadeh y Fernando, 2022), (Karthika y Senthilselvi, 2023), RNN (Fanai y Abbasimehr, 2023), LSTM (Raval, Bhattacharya, Jadav, Tanwar, Sharma, Bokoro, Elmorsy, Tolba y Raboaca, 2023), (Xie, Liu, Yan, Jiang, Zhou y Li, 2024), (Wang, Kim y Joe, 2023), GRU (Karthika y Senthilselvi, 2023) y unidad recurrente bidireccional con compuerta (BiGRU) (Prabhakaran y Nedunchelian, 2023).

Diversos trabajos de investigación han examinado la detección del fraude en numerosas revisiones y encuestas publicadas en artículos revisados por pares. Por ejemplo, Modi y Dayma (Modi y Dayma, 2017), presentaron revisiones sobre la aplicación del ML para detectar el fraude con tarjetas de crédito. Lucas y Jurgovsky (Lucas y Jurgovsky, 2020) analizaron las dificultades para detectar el fraude con tarjetas de crédito. Se centraron en los métodos propuestos para abordar la desviación conceptual y los problemas de desequilibrio, que constituyen dos desafíos importantes al analizar los datos de transacciones con tarjetas de crédito. La desviación

conceptual se produce cuando las propiedades estadísticas de los datos utilizados para entrenar un modelo de ML cambian con el tiempo. Como resultado, el modelo puede funcionar de forma diferente a la prevista o generar predicciones menos precisas. La revisión proporcionó un análisis detallado de la desviación conceptual, la clasificación de desequilibrios y los enfoques para abordarlos. Al-Hashedi y Magalingam (Al-Hashedi y Magalingam, 2021) realizaron una revisión exhaustiva de la detección de fraudes, incluyendo fraudes de seguros, tarjetas de crédito y otros fraudes financieros. La revisión describió los métodos de aprendizaje automático (ML) empleados para los diferentes problemas de detección de fraude. Además, se analizaron conjuntos de datos y métricas de evaluación del rendimiento.

El artículo también enumera las ventajas y desventajas de cada método de ML. Sin embargo, la revisión se limita a las siguientes técnicas de ML: SVM, regresión logística, red neuronal artificial, k-vecino más cercano (KNN), AG, red bayesiana, árbol de decisión, lógica difusa y modelo oculto de rkov. Popat y Chaudhary (Popat y Chaudhary, 2018) examinaron varios estudios de CCFD basados en aprendizaje automático (ML), centrándose en las dificultades que encuentran los modelos de ML para detectar fraudes. Los métodos estudiados incluyen regresión logística, SVM, árbol de decisión, redes neuronales artificiales (RNA) y redes bayesianas. Ryman-Tubb (Ryman-Tubb, Krause y Garn, 2018) realizaron una revisión y analizaron las técnicas actuales para detectar fraudes con tarjetas mediante el volumen de transacciones. Los métodos revisados incluyen SVM, KNN, CNN, redes neuronales multicapa (MLP), árbol de decisión y bosque aleatorio. Pandey (Pandey, Sachan y Ganpatrao, 2021), revisaron el CCFD, centrándose en los diferentes tipos y estadísticas de fraude con tarjetas de crédito en India. Alamri y Ykhlef (Alamri y Ykhlef, 2022) presentaron un estudio sobre estudios de detección de fraude con tarjetas de crédito que emplearon técnicas de muestreo tras identificar

el problema de desequilibrio de clases como el principal desafío que enfrentan los investigadores al construir modelos CCFD.

El estudio consideró técnicas de sobremuestreo, como la técnica de sobremuestreo sintético minoritario (SMOTE) y la técnica Borderline-SMOTE; métodos de submuestreo, como la técnica de submuestreo aleatorio (RUS) y los enlaces Tomek; y métodos de muestreo híbrido, como los enlaces SMOTE-ENN y SMOTE-Tomek. El estudio identificó los métodos de muestreo híbrido como más eficientes para abordar el problema de desequilibrio de clases en CCFD, si bien señaló que las técnicas de sobremuestreo pueden provocar sobreajuste y el submuestreo puede descartar muestras esenciales.

2.2.2 Aprendizaje supervisado.

El aprendizaje supervisado es una técnica que requiere un conjunto de datos etiquetados, en el que se conoce el resultado esperado para establecer un modelo de clasificación (por ejemplo, si una transacción es fraudulenta o no). Los modelos aprenden a predecir la clase de nuevas observaciones en función de las características de entrada. Algoritmos como redes neuronales profundas, árboles de decisión y máquinas de soporte vectorial se emplean comúnmente para este fin. En el contexto de detección de fraude, estos modelos aprenden a diferenciar patrones legítimos de los fraudulentos con base en transacciones históricas etiquetadas.

2.2.3 Aprendizaje no supervisado.

A diferencia del enfoque supervisado, el aprendizaje no supervisado no requiere etiquetas. Estos modelos buscan identificar patrones inusuales o grupos ocultos en los datos. En la detección de fraude, se utilizan técnicas como autoencoders, clustering (agrupamiento) y algoritmos de aislamiento como Isolation Forest para detectar anomalías. Estas técnicas son

especialmente útiles cuando los casos de fraude son escasos o no están claramente identificados, permitiendo encontrar desviaciones respecto al comportamiento normal sin conocimiento previo de lo que constituye un fraude.

2.2.4 Datos sintéticos y abiertos en entornos bancarios.

Al existir normativa con relación a garantizar la privacidad y la protección de datos personales, los conjuntos de datos de tarjetas de crédito no son fácilmente accesibles. Por ello, el uso de conjuntos sintéticos o de acceso abierto se ha convertido en una alternativa válida para entrenar y validar modelos de detección. Estos conjuntos permiten replicar la distribución y estructura de datos reales, sin comprometer la información confidencial de los usuarios. Su uso facilita la investigación y la comparación objetiva de modelos de clasificación.

El conjunto de datos seleccionado para el desarrollo de un modelo de detección de fraude es el conjunto de datos de tarjetas de crédito europeas. El conjunto de datos de tarjetas de crédito europeas (Credit Card Fraud Detection, 2017), ha sido ampliamente utilizado por investigadores para la creación de modelos CCFD robustos. Este conjunto de datos está compuesto por 284.807 transacciones realizadas en países europeos, clasificadas en dos categorías: legitimas (clase 0) y fraudulentas (clase 1). Cada transacción se describe mediante 30 características: 28 de ellas son variables anónimas derivadas de la aplicación de un proceso de reducción de dimensionalidad mediante Análisis de Componentes Principales (PCA), una corresponde al importe de la transacción (Amount) sin escalar ni anonimizar, y la última es una variable temporal (Time) que indica el número de segundos transcurridos desde la primera transacción registrada hasta la última, abarcando un período de 48 horas. El conjunto de datos utilizado se ha consolidado como un estándar de referencia para el estudio y la evaluación de algoritmos de detección de fraude, principalmente debido a su marcada desproporción entre clases: de las 284.807 transacciones

registradas, únicamente el 0,17% corresponde a la clase 1 (transacciones fraudulentas), mientras que el 99,83% restante pertenece a la clase 0 (transacciones legítimas). Esta distribución altamente desbalanceada representa un desafío considerable para los modelos de aprendizaje automático, ya que puede afectar negativamente la capacidad de los algoritmos para detectar fraudes. Por ello, es fundamental aplicar estrategias y técnicas adecuadas durante el desarrollo y la validación del modelo de clasificación para abordar correctamente el problema del desbalance de clases.

2.2.5 Modelos de clasificación.

La determinación de los modelos aplicables para la clasificación de fraudes constituye una etapa fundamental, particularmente debido al marcado desbalance de clases presente en los datos. En este contexto, las métricas tradicionales como la exactitud pueden resultar engañosas, ya que un modelo que clasifica todas las transacciones como legítimas alcanzaría un alto valor de exactitud sin detectar fraudes reales. Por este motivo, métricas como la precisión o sensibilidad (recall) y la medida F1 (F1-score) se consideran más apropiadas.

La precisión refleja la proporción de transacciones clasificadas como fraude que efectivamente lo son, mientras que la sensibilidad (recall) indica la proporción de fraudes reales que el modelo logra identificar correctamente. Por su parte, la F1-score proporciona una media armónica entre precisión y recall, permitiendo evaluar el equilibrio entre ambas métricas, lo cual es especialmente relevante cuando el objetivo es minimizar tanto los falsos positivos como los falsos negativos en escenarios altamente desbalanceados.

2.2.6 Esquema conceptual del proceso de detección de fraude.

La metodología CRISP-DM (Cross-Industry Standard Process for Data Mining) es ampliamente utilizada para estructurar proyectos de ciencia de datos y resulta adecuada para

implementar modelos de detección de fraude con aprendizaje automático. El proceso abarca desde la comprensión del negocio hasta el despliegue de soluciones, asegurando un desarrollo ordenado y reproducible.

La detección automatizada de fraude con tarjetas de crédito se beneficia significativamente de una aproximación estructurada, como la que proporciona la metodología CRISP-DM. Este marco guía el desarrollo de soluciones de ciencia de datos, asegurando que cada fase del proceso esté alineada con los objetivos del negocio y las mejores prácticas analíticas.

El proceso inicia con una profunda comprensión del negocio, donde se identifican tanto los objetivos estratégicos como las necesidades específicas de la organización financiera. En esta etapa, se define el impacto esperado de un sistema automatizado de detección de fraudes, así como los criterios que determinarán su éxito. A continuación, se realiza una exploración exhaustiva de los datos disponibles sobre transacciones, analizando su calidad, la presencia de desbalances entre clases y la relevancia de las variables con respecto al objetivo de clasificación. Esta fase es esencial para comprender el contexto y las limitaciones inherentes a los datos.

Posteriormente, el flujo de trabajo avanza hacia la preparación de los datos, una fase crítica que implica la limpieza, transformación, reducción de dimensionalidad y selección de características pertinentes. Estas tareas garantizan que los modelos reciban únicamente información relevante y de calidad, lo que maximiza la efectividad de los algoritmos de aprendizaje automático aplicados en la etapa siguiente. El modelado consiste en seleccionar y ajustar los modelos más adecuados incluyendo la optimización de hiperparámetros y la aplicación de técnicas de balanceo para abordar el fuerte desbalance de clases y reducir los falsos negativos, que representan un riesgo significativo en el contexto financiero.

La evaluación rigurosa del desempeño de los modelos es fundamental antes de su despliegue. Para ello, se emplean métricas especializadas como precisión, recall, F1-score y AUC, las cuales permiten medir la capacidad del sistema para identificar fraudes reales sin incrementar excesivamente los falsos positivos.

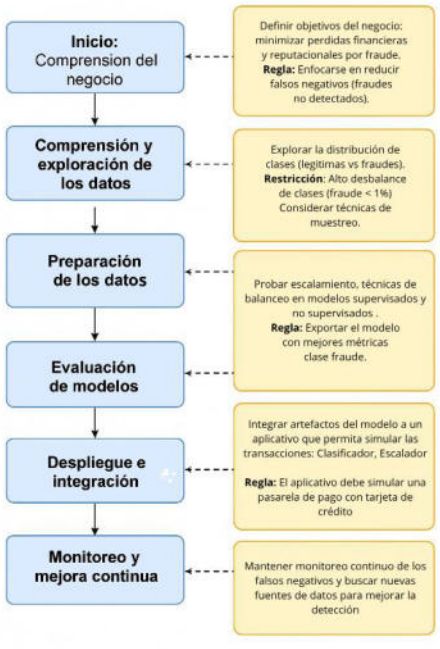
Finalmente, el modelo entrenado y validado se integra en una solución tecnológica que opera en tiempo real. Este sistema procesa nuevas transacciones bancarias y, al detectar patrones anómalos, genera alertas automáticas que permiten la revisión inmediata o el bloqueo preventivo de las operaciones sospechosas.

De este modo, la aplicación asegura que el sistema de detección de fraude sea no solo técnicamente sólido, sino también alineado con los objetivos del negocio y preparado para responder eficazmente a los desafíos del entorno financiero digital.

A continuación, en la **Figura 1** se expone el proceso propuesto.

Figura 1

Flujo del proceso para detección de fraudes



Nota. El diagrama ilustra el proceso adaptado de CRISP-DM para la detección de fraude con tarjetas de crédito mediante aprendizaje automático, destacando reglas de negocio, restricciones y consideraciones clave para el manejo de datos desbalanceados y la priorización de la reducción de falsos negativos. El proceso abarca desde la comprensión del negocio hasta el monitoreo y la mejora continua, integrando la validación de modelos supervisados y no supervisados y el despliegue en un entorno simulado de pasarela de pagos.

2.2.7 Comparación de modelos aplicados a la detección de fraude

A continuación, en la **Tabla 1** se presenta una comparación conceptual entre enfoques supervisados, no supervisados e híbridos para la detección de fraudes. Los modelos supervisados como Regresión Logística, Random Forest y XGBoost; los modelos no supervisados como MLP, LSTM y GRU; y los modelos híbridos que combinan algoritmos como Random Forest, XGBoost y redes neuronales (MLP, LSTM).

Tabla 1

Comparación de modelos, ventajas y limitaciones

Tipo de entrenamiento	Modelo	Ventajas principales	Limitaciones
Supervisado	Regresión Logística, Random Forest, XGBoost	Alta precisión si hay etiquetas; captura patrones complejos.	Requiere grandes volúmenes de datos etiquetados.
No Supervisado	MLP, LSTM, GRU	Detecta anomalías sin etiquetas; útil ante clases raras.	Puede generar falsos positivos si no se ajustan bien los umbrales.
Híbrido	Random Forest + XGBoost + MLP + LSTM	Combina generalización y secuencialidad	Mayor complejidad computacional

Nota. Los modelos supervisados ofrecen alta precisión cuando se dispone de etiquetas, mientras que los no supervisados permiten detectar patrones anómalos debido a que no se cuenta con datos etiquetados. Los modelos híbridos combinan lo mejor de ambos modelos anteriores, aunque para eso requiere de una mayor complejidad computacional.

Capítulo 3

3. Desarrollo

3.1. Desarrollo del trabajo

El presente trabajo se enfoca en el diseño, implementación y evaluación de modelos de aprendizaje automático para la detección de fraudes en transacciones bancarias, específicamente en pagos con tarjetas de crédito. La metodología adoptada combina un enfoque científico-experimental con fundamentos técnicos, empleando modelos supervisados, no supervisados e híbridos, en función del tipo de datos disponibles y del contexto operativo. La estructura metodológica para obtener el modelo adecuado para la detección de fraudes con tarjetas de crédito se compone de las siguientes etapas:

Etapa 1: Comprensión del Negocio: Detección de Fraude en Transacciones con Tarjetas de Crédito

La primera etapa se centra en la comprensión del negocio bancario relacionado con las transacciones realizadas mediante tarjetas de crédito. Esta fase resulta esencial para contextualizar el problema de detección de fraude y establecer objetivos claros, realistas y alineados con las prioridades de las entidades financieras. En particular, el interés radica en reducir las pérdidas económicas y reputacionales derivadas de transacciones fraudulentas, así como en cumplir con las regulaciones vigentes y proteger la confianza de los clientes.

Durante esta etapa, se analizan los procesos operativos involucrados en la gestión y autorización de transacciones, definiendo con precisión los criterios que caracterizan una operación fraudulenta dentro del flujo de trabajo bancario. Asimismo, se determinan las reglas de negocio relevantes, como la necesidad de minimizar los falsos negativos (fraudes no detectados), que constituyen uno de los principales retos para la organización. Esta comprensión inicial permite orientar el despliegue de la solución y fundamentar la posterior selección de técnicas de ciencia de datos y aprendizaje automático que serán empleadas para abordar el problema de manera eficaz y replicable.

Etapa 2: Obtención y exploración de conjuntos de datos relevantes a fraudes bancarios

En esta segunda etapa, se centra la atención en la identificación, obtención y exploración de conjuntos de datos pertinentes para abordar la problemática de la detección de fraude en transacciones con tarjetas de crédito. El acceso a datos de calidad es fundamental para el desarrollo y la validación de modelos de aprendizaje automático capaces de distinguir entre transacciones legítimas y fraudulentas.

Se priorizan fuentes de datos abiertas y reconocidas en la literatura especializada, destacando especialmente el conjunto de datos “Credit Card Fraud Detection” disponible en Kaggle, ampliamente empleado en estudios académicos y competiciones internacionales por su relevancia, grado de anonimización y representatividad de escenarios reales. Asimismo, se consideran fuentes adicionales de interés, como el repositorio IEEE-CIS Fraud Detection y bases de datos institucionales, tales como las proporcionadas por la Superintendencia de Bancos del Ecuador, que podrían enriquecer futuras investigaciones.

Para el presente trabajo, la selección recae sobre el dataset de Kaggle debido a sus características técnicas y al cumplimiento de los requisitos para la construcción, entrenamiento y evaluación de modelos supervisados y no supervisados, garantizando así la reproducibilidad y comparabilidad de los resultados obtenidos.

Etapa 3: Preparación de los datos

Durante la fase de preparación de los datos, se implementaron actividades orientadas a garantizar la calidad, integridad y adecuación del conjunto de datos para su posterior análisis mediante técnicas de aprendizaje automático. En primer lugar, se evaluó la completitud de los datos, verificando la ausencia de valores nulos o perdidos y asegurando la coherencia entre registros y variables. Posteriormente, se realizó un análisis detallado de los rangos de cada variable para detectar posibles inconsistencias, valores atípicos (outliers) o anomalías que pudieran distorsionar los resultados de los modelos.

Adicionalmente, se llevó a cabo una revisión que incluyó la visualización de la distribución de las variables, tanto individuales como combinadas, con el fin de identificar patrones, sesgos y el grado de desbalance de clases inherente al problema de detección de fraude. Esta revisión permitió ajustar estrategias específicas para el manejo del desbalance.

En cuanto al preprocesamiento, se aplicaron técnicas de limpieza y normalización de los datos, asegurando la homogeneidad de las escalas y la correcta interpretación de los algoritmos.

Estas actividades de preparación resultaron fundamentales para garantizar que el conjunto de datos estuviera adecuadamente depurado y estructurado, habilitando así un entrenamiento robusto y una evaluación confiable de los modelos de detección de fraude desarrollados en fases posteriores.

Etapa 4: Evaluación de modelos:

En esta fase, se procedió a comparar de manera sistemática distintos enfoques de clasificación supervisada y no supervisada para la detección de fraude en transacciones con tarjetas de crédito. Entre los algoritmos evaluados se incluyeron modelos clásicos como Regresión Logística, SVM y Random Forest, así como métodos avanzados como XGBoost y diferentes variantes de redes neuronales multicapa (MLP). Adicionalmente, se exploraron técnicas específicas de optimización orientadas a la reducción de falsos negativos, dada la criticidad de este indicador en el contexto bancario, mediante estrategias como el ajuste de umbrales de decisión, el uso de ensambles conservadores y la aplicación de técnicas sensibles al costo.

El proceso de evaluación se apoyó en un pipeline automatizado de entrenamiento y validación, que permitió comparar métricas clave como la precisión (precision), la exhaustividad (recall), la puntuación F1 (F1-score) y el área bajo la curva ROC (AUC-ROC), priorizando el recall de la clase minoritaria para mitigar el riesgo de fraudes no detectados.

Adicionalmente, se consideró el balance entre falsos positivos y falsos negativos, evaluando el impacto operativo de cada modelo.

Para la selección del modelo final, se adoptó un enfoque integral que combinó el rendimiento cuantitativo, la interpretabilidad y la eficiencia computacional, seleccionando aquel modelo que ofrecía el mejor compromiso entre alta capacidad de detección de fraude y viabilidad de despliegue en un entorno productivo. Esta metodología permitió identificar la arquitectura óptima para el caso de estudio, asegurando la robustez y escalabilidad de la solución propuesta.

Fase 5: Despliegue e integración.

En la fase de despliegue e integración, se aborda la transición desde el entorno de experimentación y desarrollo hacia una solución funcional y reutilizable. Una vez seleccionado el modelo óptimo, se procedió a la exportación de los artefactos clave, incluyendo los modelos entrenados, los escaladores, el orden de las características y las estadísticas de referencia necesarias para la generación de datos sintéticos y la predicción. Estos artefactos fueron serializados, garantizando su portabilidad y facilitando su carga eficiente en ambientes productivos.

Para la demostración y validación práctica de la solución, se cuenta con una aplicación web, la que permita simular en tiempo real la detección de fraude en transacciones con tarjetas de crédito. La aplicación integra todos los componentes del pipeline, desde el preprocesamiento de los datos de entrada y la aplicación del modelo de predicción, hasta la visualización intuitiva de los resultados y métricas clave. Esta integración facilita el acceso de usuarios finales y equipos técnicos a la funcionalidad de detección de fraude, demostrando la escalabilidad y flexibilidad de la arquitectura propuesta para su potencial despliegue en sistemas reales de monitoreo transaccional.

Etapa 6: Monitoreo y mejora continua.

Una vez que la aplicación de detección de fraude se encuentre operando en producción durante un periodo prudencial, se iniciará la fase de monitoreo y mejora continua.

En ese momento, se realizará una evaluación sistemática del rendimiento del sistema, utilizando los artefactos previamente exportados (modelos entrenados, escaladores y configuraciones de variables) y la aplicación desplegada. El objetivo es identificar posibles desviaciones en los indicadores clave de desempeño, como la tasa de falsos negativos, el recall y

la precisión, lo que permite anticipar la aparición de nuevos patrones de fraude o cambios en los datos transaccionales. Con base en estos hallazgos, se planificará la actualización y el reentrenamiento de los modelos, integrando datos recientes y evaluando nuevas técnicas o arquitecturas si fuera necesario. De este modo, se garantiza que la solución tecnológica se mantenga alineada con los requerimientos operativos y continúe siendo efectiva ante la evolución de los riesgos en el entorno bancario real.

Capítulo 4

4. Análisis de Resultados

4.1. Pruebas de Concepto (PoC)

El capitulo documenta el desarrollo y validación de las diferentes pruebas de concepto, siguiendo la estructura propuesta en el Capítulo 3 para la construcción de soluciones basadas en ciencia de datos e identificación de transacciones fraudulentas con tarjetas de crédito.

4.2. Establecimiento del entorno de trabajo para los experimentos y PoC

El proceso inicia con la preparación del entorno de trabajo, que comprende la configuración de los recursos computacionales necesarios, la instalación de las librerías y dependencias requeridas, así como la organización de los directorios y artefactos empleados a lo largo del ciclo de vida del proyecto. Esta etapa es fundamental para garantizar la reproducibilidad de los experimentos y la trazabilidad de los resultados obtenidos.

El entorno de trabajo se estableció sobre Jupyter Notebook, empleando el lenguaje Python y diversas bibliotecas especializadas en análisis de datos y visualización. La preparación comenzó con la importación de las librerías más utilizadas en ciencia de datos, como: pandas, numpy, matplotlib y seaborn, que permiten la manipulación eficiente de datos, la realización de cálculos numéricos y la creación de gráficos descriptivos.

Posteriormente, se procedió a obtener el conjunto de datos directamente desde Kaggle con el objeto de no perder integridad empleado la librería kagglehub, luego utilizando la función read_csv de pandas sobre la ruta local asignada a la variable csv_file_path permitió acceder a la información de las transacciones del conjunto de datos creditcard.csv. Tras la carga, se utilizó el método head() para visualizar las primeras filas del conjunto de datos, lo cual sirvió para confirmar la correcta lectura del archivo y obtener una vista preliminar de la estructura y las variables disponibles.

De este modo, el entorno estuvo preparado para la ejecución de los análisis exploratorios y las siguientes etapas del proceso de análisis de datos, incluyendo la identificación de valores atípicos, la revisión de la completitud de los datos y la preparación para el entrenamiento de modelos de clasificación enfocados en minimizar los falsos negativos.

4.3.Análisis exploratorio de datos

A continuación, se realiza un análisis exploratorio de datos (EDA) para comprender la estructura y distribución de las variables presentes en el conjunto de datos, identificar patrones relevantes, anomalías y la magnitud del desbalance de clases, aspecto crítico en problemas de detección de fraude.

En primera instancia del EDA, se centró en la verificación del conjunto de datos, asegurando que la información fuera adecuada para su posterior procesamiento y modelado. Se empleó la biblioteca pandas para la manipulación eficiente de los datos tabulares. El archivo principal de transacciones se encontraba en formato CSV, por lo cual se aplicó una transformación en un DataFrame ejecutando el comando df = pd.read_csv('creditcard.csv'). Esta instrucción permitió disponer de todas las transacciones para su análisis posterior.

Para confirmar la integridad del conjunto de datos, se empleó el método df.shape, que reveló la existencia de 284,807 registros y 31 variables por transacción. Seguidamente, mediante df.info(), se verificó que todas las columnas presentan el tipo de dato esperado (mayoritariamente valores numéricos en formato float64 y la variable objetivo como int64). Además, la comprobación de valores nulos con df.isnull().sum() arrojó como resultado la ausencia total de datos faltantes en el dataset, lo cual es crucial para evitar sesgos o errores durante el entrenamiento de los modelos de aprendizaje automático.

Estos resultados pueden consultarse en la **Tabla 2**, que sintetiza la estructura y tipología de las variables contenidas en el conjunto de datos.

Tabla 2

Estructura y tipos de variables del conjunto de datos de transacciones de tarjetas de crédito

Variable	Tipo de dato	Valores nulos
Time	float64	0
V1	float64	0
...	...	...
V28	float64	0
Amount	float64	0
Class	int64	0

Nota. Elaboración propia a partir de la función df.info() y el análisis de valores nulos.

Esta primera exploración garantiza que el conjunto de datos es completo, consistente y apto para los análisis posteriores, sin necesidad de aplicar técnicas de imputación ni eliminación de registros, y cumpliendo con los requisitos básicos de calidad de datos recomendados en la literatura para proyectos de ciencia de datos.

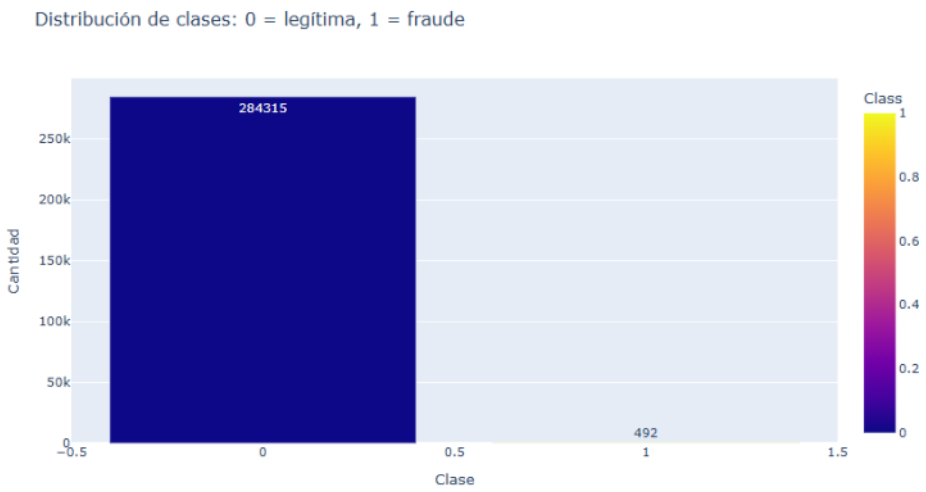
Una vez validada la integridad del conjunto de datos, se realizó el análisis exploratorio de la variable objetivo (Class), que indica si una transacción es legítima (0) o fraudulenta (1). El

propósito de esta etapa es comprender el grado de desbalance de clases, un aspecto crítico en la detección de fraude con machine learning.

El conteo de observaciones para cada clase se obtuvo utilizando la instrucción df['Class'].value_counts(). Los resultados confirmaron una alta desproporción, con 284,315 transacciones legítimas frente a solo 492 fraudulentas, lo que corresponde a un 0.17% de fraudes sobre el total. Esta distribución se visualizó mediante un gráfico de barras generado con plotly, como se indica en la Figura 2.

Figura 2

Distribución de la variable objetivo en el conjunto de datos



Nota. Desbalance de las clases en el Dataset.

El gráfico evidenció el severo desbalance entre clases, aspecto que representa un desafío importante para los algoritmos de clasificación, ya que los modelos tienden a favorecer la clase mayoritaria (transacciones legítimas) y pueden pasar por alto los casos de fraude.

Este diagnóstico inicial subraya la necesidad de aplicar técnicas específicas de re-muestreo (sampling) y estrategias de evaluación que prioricen la detección de la clase

minoritaria, evitando una falsa sensación de buen desempeño asociada a la alta exactitud global, pero con bajo recall en la clase de fraude.

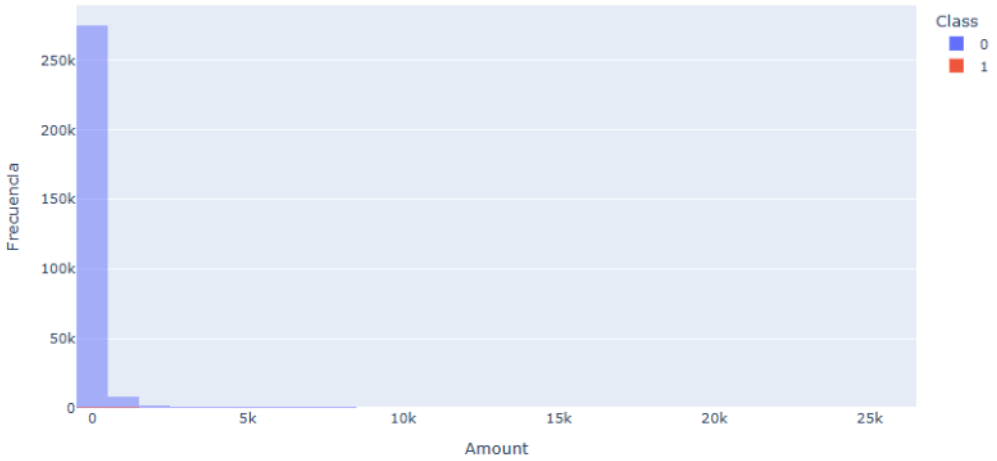
La severidad del desbalance de clases justifica el uso de métricas alternativas a la exactitud y anticipa la aplicación de metodologías especializadas en las siguientes fases del proyecto.

El análisis realizado permitió comprender el comportamiento de las transacciones y anticipar posibles retos para el modelado. Con el objeto de plantear estrategias de modelado de datos se procedió con el estudio de las características Amount (monto de la transacción) y Time (tiempo transcurrido desde la primera transacción), esto en razon que son las únicas que no se encontraban tratadas previamente con algún tipo de reducción de dimensionalidad.

Para explorar la distribución del monto de las transacciones, se utilizó un histograma segmentado por clase, el cual muestra una fuerte asimetría positiva en la Figura 3, donde la mayoría de las transacciones se concentran en montos bajos, con pocos valores atípicos en el rango de valores elevados. Esta distribución sugiere la conveniencia de normalizar o escalar la variable antes de su uso en modelos de machine learning, para evitar que los algoritmos sean influenciados por los valores extremos.

Figura 3

Distribución de Amount según la Clase



N*ota.* Desbalance de la variable Amount.

En el mismo sentido de análisis, la variable Time refleja el número de segundos transcurridos desde la primera transacción registrada en el conjunto de datos. Su distribución, visualizada de igual forma con un histograma segmentado por clase, en la Figura 4 indica que las transacciones están agrupadas en patrones periódicos, probablemente asociados a ciclos de actividad financiera a lo largo de los dos días registrados. Sin embargo, dado que Time es una variable relativa y no necesariamente aporta discriminación directa para la detección de fraude, su utilidad debe evaluarse cuidadosamente en el proceso de selección de variables.

Figura 4

Distribución de Tiempo según la Clase

Nota. El gráfico representa la distribución temporal de las transacciones, diferenciadas por clase. Se evidencia que los fraudes no siguen un patrón regular en el tiempo y pueden ocurrir de forma dispersa, lo que sugiere que la variable Tiempo tiene baja capacidad predictiva directa, pero puede aportar en combinación con otras características.

Con la orientación hacia fortalecer el análisis, se revisa la presencia de valores atípicos puede afectar negativamente el entrenamiento de los modelos. Por tal motivo, se calcularon *"outliers"* empleando dos criterios: Z-score (mayor a 3 desviaciones estándar) e IQR (fuera del rango intercuartílico 1.5x). Los resultados, resumidos en la **Tabla 3**, muestran la cantidad de outliers por variable:

Tabla 3

Comparación de la cantidad de valores atípicos detectados por Z-score y por IQR

Variable	Outliers (Z-score > 3)	Outliers (IQR)
Time	0	0
V1	3701	7062
V2	4318	13526
V3	1987	3363
V4	3094	11148
V5	2945	12295
V6	4652	22965

Variable	Outliers (Z-score > 3)	Outliers (IQR)
V7	3401	8948
V8	4221	24134
V9	2293	8283
V10	3488	9496
V11	684	780
V12	3393	15348
V13	1192	3368
V14	3380	14149
V15	1254	2894
V16	2077	8184
V17	2515	7420
V18	1685	7533
V19	3399	10205
V20	4645	27770
V21	4064	14497
V22	1222	1317
V23	3364	18541
V24	657	4774
V25	2809	5367
V26	1047	5596
V27	4771	39163
V28	3264	30342
Amount	4076	31904
Class	492	492

Nota. La tabla presenta la cantidad de valores atípicos detectados por dos métodos distintos: Z-score (valores con puntuación Z mayor a 3) e IQR (valores fuera del rango intercuartílico). El total de outliers identificados en el conjunto de datos fue de 84,090 mediante Z-score y 370,864 mediante IQR. Esta comparación permite evidenciar la sensibilidad diferencial de cada técnica en la detección de atípicos.

Los resultados obtenidos de la clasificación de clases y outliers permiten comparar y aproximar el comportamiento de las transacciones, y si las fraudulentas presentan algún patrón

diferenciador respecto a las legítimas, información útil para el desarrollo o planteamiento de estrategias para el modelo respecto a las reglas de negocio.

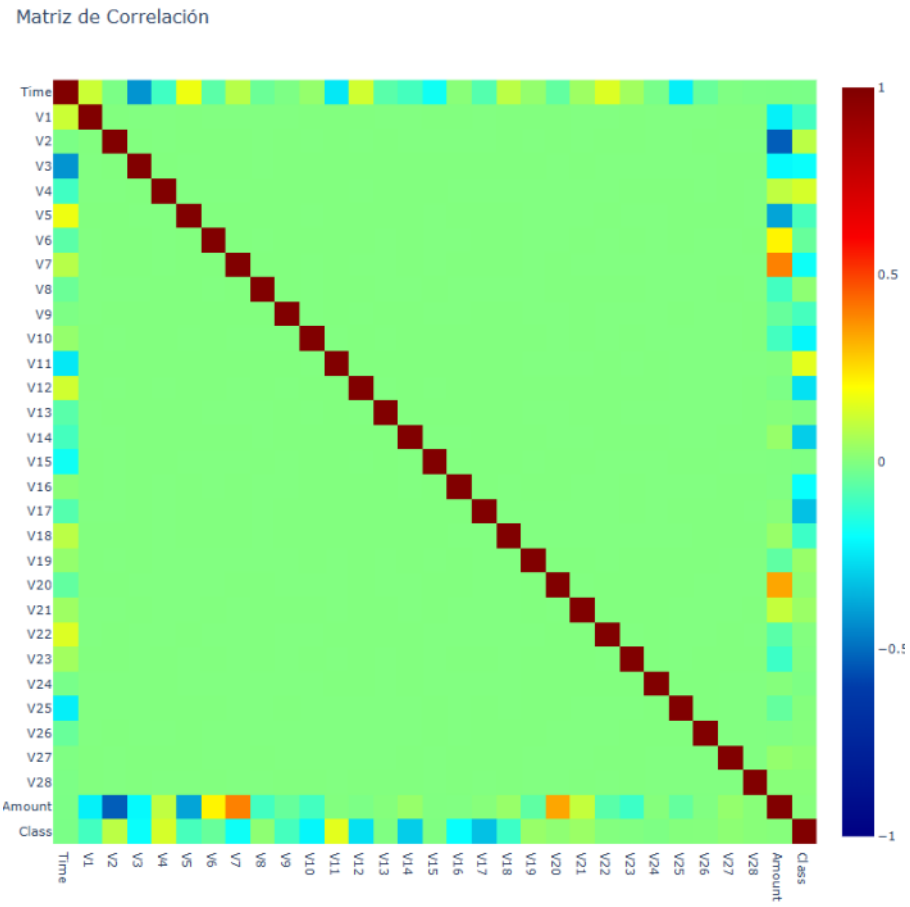
4.4.Análisis de características

Previo a la construcción de modelos de detección de fraude con aprendizaje automático es fundamental realizar una identificación de las variables que presentan mayor relevancia predictiva. En conjuntos de datos altamente dimensionales, como el utilizado en este estudio (con 28 variables transformadas y 3 originales), la correcta selección y comprensión de la importancia de cada característica puede marcar la diferencia en el desempeño del modelo.

Se procedió con el calculó la matriz de correlación de Pearson entre todas las variables, incluyendo la variable objetivo Class. Esto permite detectar redundancias, relaciones lineales y posibles multicolinealidades. El cálculo se realizó con: columns_corr = df.corr(), mismo que permitió generar una matriz de visualización, mediante un mapa de calor representado en la Figura 5.

Figura 5

Matriz de Correlación



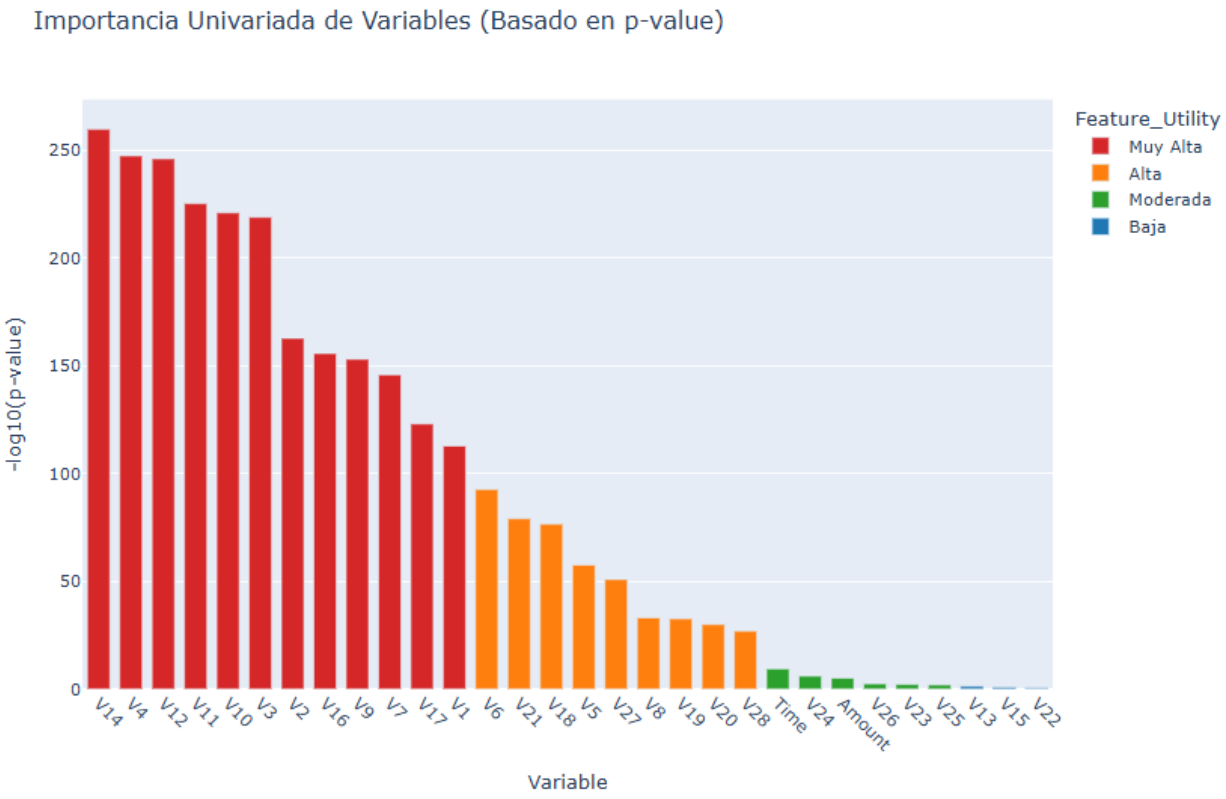
Nota. La matriz ayuda a identificar grupos de variables altamente correlacionadas, lo cual puede ser útil para reducir la dimensionalidad o evitar el uso de variables redundantes en el desarrollo del modelo.

Esta visualización grafica reveló que variables como V14, V4, V12, V11, V10 y V3 destacan posiblemente como las más relevantes en términos de su relación estadística con los eventos de fraude, lo que sugiere que deberían ser consideradas prioritarias en los modelos supervisados.

Adicionalmente, en la **Figura 6** se evaluó la importancia de cada variable mediante pruebas estadísticas (como el p-valor de la diferencia de medias) y métricas de utilidad (Feature Utility), clasificándolas en categorías: Muy Alta, Alta, Moderada o Baja.

Figura 6

Gráfico de barras de la importancia univariada de las variables



Nota. El gráfico muestra la importancia individual de las variables predictoras, calculada mediante análisis univariado. Se observa que un grupo reducido de variables concentra la mayor relevancia estadística, lo que respalda su priorización en los modelos de detección de fraude.

4.5.Establecimiento de condiciones para el entrenamiento supervisado

A partir del análisis exploratorio realizado sobre el conjunto de datos, se identificó un marcado desbalance entre las clases, así como la presencia significativa de valores atípicos en múltiples variables. Sin embargo, al profundizar en la distribución de los registros de fraude y legítimos, se observó que los eventos de fraude no se concentran ni dependen exclusivamente de los rangos extremos o atípicos detectados por criterios de Z-score o IQR. Por el contrario,

muchas transacciones fraudulentas se distribuyen a lo largo del rango completo de valores y su identificación no puede sustentarse únicamente en la exclusión de outliers.

Del mismo analisis, se desprende el análisis univariado el cual revela que ciertas variables (particularmente algunas componentes de V1 a V28) exhiben una mayor capacidad discriminativa para detectar fraudes, actualmente no se cuenta con información descriptiva precisa sobre el significado u origen de cada una de ellas, ya que son el resultado de una transformación PCA realizada por los autores del conjunto de datos original. Esta falta de interpretación limita la posibilidad de realizar una selección informada de características, pues eliminar alguna de estas variables podría inadvertidamente excluir información relevante y deteriorar el desempeño predictivo del modelo, especialmente en un contexto donde los eventos de fraude son extremadamente escasos.

Por este motivo, y siguiendo las mejores prácticas para el tratamiento de variables anonimizadas o transformadas, se opta por conservar todas las variables V1 a V28 incluido sus outliers para el modelado, asegurando así que el algoritmo de aprendizaje automático supervisado tenga acceso a la totalidad de la información representada en el dataset. La única excepción es la variable Time, misma que mostró una correlación limitada con la clase objetivo y, dado su carácter relativo (representando segundos transcurridos desde la primera transacción), su contribución informativa es marginal frente al resto de las variables transformadas. Por consiguiente, se decide excluir Time de los procesos de entrenamiento y predicción.

Finalmente, se establece como regla técnica fundamental que solo la variable Amount debe ser escalada antes de entrenar los modelos, en virtud de que su rango y dispersión son significativamente mayores en comparación con el resto de las variables, y su normalización

evita posibles sesgos durante el aprendizaje. Las variables resultantes de PCA se asumen ya normalizadas.

Estas decisiones garantizan que no se eliminen instancias relevantes para la detección de fraude, lo que resulta fundamental dado el carácter altamente desbalanceado del problema y la baja representatividad de la clase minoritaria. Así, se prioriza la máxima retención de información útil para la posterior etapa de modelado, evitar sesgos que podrían comprometer la capacidad del modelo para detectar fraudes en escenarios reales, maximizar el potencial de los algoritmos supervisados y minimizar el riesgo de perder patrones clave para la detección de fraudes poco frecuentes.

4.6.Identificación de modelos de clasificación candidatos.

Tras el análisis exploratorio, la revisión del desbalance y la caracterización del dataset, el siguiente paso consiste en la construcción, entrenamiento y evaluación de diferentes modelos de clasificación supervisada. El objetivo principal de este proceso es identificar el enfoque que permita maximizar la detección de fraudes (minimizando falsos negativos) manteniendo un adecuado balance entre sensibilidad y especificidad, aspecto crítico en entornos reales de prevención del fraude bancario (El-Naby, Hemdan, y El-Sayed, 2023) (Alfaiz y Fati, 2022) (Bin Sulaiman, Schetinin y Sant, 2022) (Carcillo, 2019).

La selección de los algoritmos y técnicas a emplear se fundamenta tanto en la literatura reciente como en las particularidades del dataset, optando por métodos probados en el contexto de la detección de fraude con tarjetas de crédito (Alarfaj, Malik, Khan, Almusallam, Ramzan y Ahmed, 2022) (Btoush, Zhou, Gururajan, Chan, Genrich, y Sankaran, 2023) (Lucas, y Jurgovsky, 2020) (Popat y Chaudhary, 2018) (Ryman-Tubb, Krause y Gran, 2018).

En este sentido, se han considerado los siguientes modelos basados en algoritmos de machine learning tradicionales, enfoques de ensamble y arquitecturas de redes neuronales.

A continuación, se detallan y justifican la selección de modelos tradicionales y de redes neuronales:

- **Modelo tradicional - Dummy Classifier:** Utilizado como línea base, representa un modelo sin capacidad predictiva real, que permite contrastar el desempeño de los modelos desarrollados con un clasificador aleatorio o de mayoría (Bhattacharyya, Jha, Tharakunnel y Westland, 2011).
- **Modelo tradicional - Regresión Logística:** Modelo lineal clásico ampliamente utilizado en detección de fraude debido a su interpretabilidad y buen comportamiento como primera aproximación en el contexto de alto desbalance (Bhattacharyya, Jha, Tharakunnel y Westland, 2011) (Sahin y Duman, 2011) (Bahnsen, 2016).
- **Modelo tradicional - Random Forest:** Algoritmo de ensamble basado en árboles de decisión que ha demostrado alta robustez ante outliers y ruido, así como un rendimiento superior en datasets desbalanceados (Bakhtiari, Nasiri y Vahidi, 2023) (Alarfaj, Malik, Khan, Almusallam, Ramzan y Ahmed, 2022) (Alfaiz y Fati, 2022) (Aburbeian y Ashqar, 2023).
- **Modelo tradicional - Support Vector Machine (SVM):** Algoritmo efectivo en espacios de alta dimensionalidad y con buenos resultados en escenarios de fraude, especialmente tras una adecuada selección y escalado de características (Mienye y Jere, 2024) (Lucas y Jurgovsky, 2020).

- **Modelo tradicional - XGBoost:** Método de boosting basado en árboles de decisión que se caracteriza por su alta precisión, escalabilidad y eficiencia computacional, ampliamente adoptado en competiciones y estudios de detección de fraude (Zhang, Han, Xu y Wang, 2021) (Kafhali y Tayebi, 2022) (Chen y Guestrin, 2016).
- **Modelo tradicional - LightGBM:** Alternativa de boosting que optimiza el uso de recursos y tiempo de entrenamiento, destacando por su capacidad de manejar grandes volúmenes de datos y por su rendimiento sobresaliente en tareas de clasificación financiera (Ke, Meng, Finley, Wang, Chen, Ma y Liu, 2017).
- **Modelo tradicional - K-Nearest Neighbors (KNN):** Utilizado como benchmark no lineal, puede captar relaciones complejas, aunque suele ser sensible al desbalance de clases y a la escala de los datos, por lo que se incorpora a modo comparativo (Alfaiz y Fati, 2022) (Pandey, Sachan y Ganpatrao, 2021).
- **Red neuronal - Perceptrón Multicapa (MLP):** Esquema de red neuronal densa tradicional, probado en el contexto de fraude para capturar relaciones no lineales entre las variables (Alfaiz y Fati, 2022) (Esenogho, Mienye, Swart, Aruleba y Obaido, 2022) (Gambo, Zainal y Kassim, 2022).
- **Redes Neuronales - Recurrentes (RNN, LSTM, GRU):** Estos modelos, y sus variantes bidireccionales (BiLSTM, BiGRU), se han empleado para modelar la secuencialidad y dependencias temporales en transacciones,

incluso cuando la variable tiempo está anonimizada o eliminada (Alarfaj, Malik, Khan, Almusallam, Ramzan y Ahmed, 2022) (Benchaji, Douzi, El Ouahidi y Jaafari, 2021) (Fanai y Abbasimehr, 2023) (Xie, Liu, Yan, Jiang, Zhou y Li, 2024) (Raval, Bhattacharya, Jadav, Tanwar, Sharma, Bokoro, Elmorsy, Tolba y Raboaca, 2023) (Wang, Kim y Joe, 2023) (Sehrawat y Singh, 2023).

4.7.Construcción y evaluación del modelo de clasificación candidatos

Una vez seleccionados los modelos candidatos, se procede con la construcción y evaluación de los modelos supervisados, para lo cual se descartó la variable *Time* por su baja correlación y escaso aporte predictivo, mientras que la variable Amount fue escalada mediante normalización estándar en todos los experimentos, siguiendo buenas prácticas reportadas en la literatura (Zhang, Han, Xu y Wang, 2021) (El-Naby, Hemdan y El-Sayed, 2023) (Alfaiz y Fati, 2022) (Credit Card Fraud Detection, 2017). La **Tabla 4** presenta las principales métricas y la **Tabla 5** muestra las matrices de confusión de los modelos tradicionales de clasificación y los modelos basados en redes neuronales.

Tabla 4

Desempeño de modelos tradicionales y de redes neuronales en detección de fraude

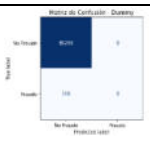
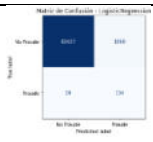
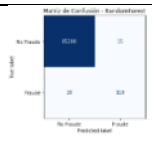
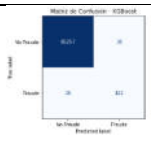
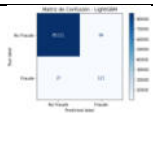
Modelo	Recall fraude	F1-score	ROC AUC	Tiempo entrenamiento (s)
Dummy	0.0000	0.0000	0.5000	0.01
Logistic Regression	0.9054	0.1247	0.9687	0.02
Random Forest	0.8041	0.8440	0.9763	1.32
SVM	0.8514	0.1510	0.9511	360.07

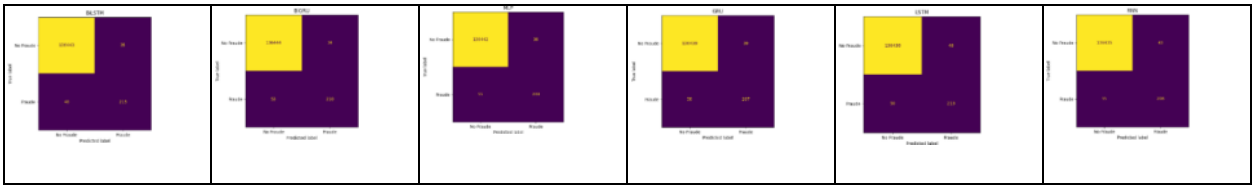
Modelo	Recall fraude	F1-score	ROC AUC	Tiempo entrenamiento (s)
XGBoost	0.8243	0.7922	0.9781	13948.20
LightGBM	0.8176	0.6856	0.9611	1.76
KNN	0.8514	0.6316	0.9252	2.31
BiLSTM	0.8176	0.8384	0.9819	0.07
BiGRU	0.7987	0.8289	0.9776	94.33
MLP	0.7912	0.8200	0.9701	91.63
GRU	0.7872	0.8127	0.9785	44.86
LSTM	0.8101	0.8124	0.9825	72.34
RNN	0.7907	0.8096	0.9787	64.19

Nota. La tabla muestra los resultados de desempeño para los modelos candidatos para la clasificación de fraudes. Se presentan las métricas y el tiempo de entrenamiento (en segundos, redondeado a dos decimales). El modelo Dummy se incluye como referencia base.

Tabla 5

Comparativa de matrices de confusión de modelos candidatos

Modelos tradicionales					
Dummy	**Logistic Regression**	**Random Forest**	**SVM**	**XGBoost**	**LightGBM**
					
Modelos de redes neuronales					
BiLSTM	**BiGRU**	**MLP**	**GRU**	**LSTM**	**RNN**



Nota. La figura presenta las matrices de confusión generadas por modelos tradicionales (fila superior) y modelos de redes neuronales (fila inferior) para la detección de fraude con tarjetas de crédito. Las clases están representadas como *No Fraude* y *Fraude*, evaluadas sobre un conjunto de prueba no balanceado.

4.8.Determinación y selección del modelo de clasificación

La evaluación de los resultados de desempeño **Tabla 4** y matrices de confusión **Tabla 5** de los modelos candidatos entrenados revela diferencias notables en el desempeño, particularmente en la sensibilidad hacia la clase de fraude (recall), el equilibrio general entre precisión y exhaustividad (F1-score) y la eficiencia computacional.

Los modelos de referencia, como Dummy, muestran una incapacidad total para identificar fraudes reales (recall fraude = 0.00), lo que se evidencia en matrices de confusión con nula detección positiva. Por el contrario, modelos tradicionales como la Regresión Logística destaca por alcanzar el recall más alto para la clase fraude (0.9054), lo que indica su elevada sensibilidad para identificar transacciones fraudulentas. Sin embargo, su F1-score es considerablemente bajo (0.1247), reflejando un desbalance entre precisión y exhaustividad, lo cual se observa claramente en la matriz de confusión, el modelo identifica la mayoría de los fraudes, pero a costa de un alto número de falsos positivos, afectando la precisión global. Este fenómeno es típico en entornos desbalanceados, donde el ajuste del umbral de decisión se vuelve crítico para controlar el coste operativo de las alertas erróneas.

Por otro lado, el Random Forest logra un balance más robusto entre recall (0.8041) y F1-score (0.8440), demostrando una mejor capacidad para distinguir entre clases y una mayor

estabilidad frente al desbalance de los datos, como lo evidencian los valores de la matriz de confusión y el ROC AUC de 0.9763. La precisión en la predicción de fraudes y la moderación en los falsos positivos sugieren que este modelo captura patrones relevantes de las transacciones fraudulentas sin comprometer excesivamente la precisión global. Además, mantiene un tiempo de entrenamiento razonable, lo que refuerza su viabilidad en contextos productivos donde la actualización y el despliegue rápido son necesarios.

Modelos como XGBoost y LightGBM presentan un rendimiento similar en ROC AUC, aunque el primero implica un coste computacional significativamente mayor (más de 13,000 s de entrenamiento), lo que limita su viabilidad para despliegues rápidos o en entornos con restricciones de recursos, tal como se lo menciona en (Kafhali y Tayebi, 2022) (Ke, Meng, Finley, Wang, Chen, Ma y Liu, 2017).

Por otro lado, el modelo KNN si bien alcanza un recall equivalente a SVM (0.851), su F1-score (0.632) y velocidad de entrenamiento lo sitúan como una alternativa viable para escenarios donde la simplicidad y el bajo costo computacional sean prioritarios, aunque su rendimiento en el balance general es inferior al de Random Forest y la Regresión logistica.

Respecto a los modelos basados en redes neuronales, arquitecturas como BiLSTM, BiGRU, GRU, LSTM y RNN demostraron desempeños competitivos, con valores de F1-score que oscilan entre 0.81 y 0.84 y ROC AUC superiores a 0.97, lo que confirma la capacidad de estas redes para modelar secuencias temporales y relaciones no lineales. Sin embargo, el tiempo de entrenamiento y la complejidad inherente de estos modelos suelen ser mayores, y la ventaja práctica e interpretabilidad frente a Random Forest resulta menos evidente para el caso de estudio, considerando la naturaleza tabular y anonimizada de las variables V1-V28.

En síntesis, los resultados de desempeño muestran que Random Forest y la Regresion Logistica son candidatos que podrían generar una solución robusta y eficiente para la detección de fraude con tarjetas de crédito, logrando un equilibrio óptimo entre sensibilidad, precisión y costo computacional. Su desempeño es consistente con la literatura, que respalda el uso de este tipo de modelos como opciones preferentes para problemas de clasificación con alta desbalanceo de clases y datos estructurados (Bakhtiari, Nasiri y Vahidi, 2023) (Alarfaj, Malik, Khan, Almusallam, Ramzan y Ahmed, 2022) (Kafhali y Tayebi, 2022) (Ke, Meng, Finley, Wang, Chen, Ma y Liu, 2017).

4.9. Construcción del modelo de clasificación de fraudes

Una vez establecido los modelos candidatos para la construcción de la propuesta de solucion adecuada para el caso de estudio, se procede a ejecutar un conjunto de experimentos que incluyen la aplicación de diversas estrategias de preprocesamiento, balanceo de clases, ajuste de hiperparámetros y comparación de métricas clave.

Como parte de la fase experimental final y previo al despliegue del modelo definitivo, se llevó a cabo una serie de pruebas orientadas a identificar la configuración óptima del sistema de detección de fraude. Para tal fin, se seleccionó y particionó una muestra representativa de 50,000 transacciones a partir del dataset original, garantizando la proporción natural entre clases (492 fraudes y 49,508 transacciones legítimas). Esta reducción del tamaño de la muestra responde a la necesidad de lograr una mayor eficiencia computacional en los experimentos, facilitar la repetibilidad de las pruebas y acelerar los ciclos de ajuste y validación, tal como recomiendan diversos trabajos para escenarios de experimentación y prototipado rápido (Zhang, Han, Xu y Wang, 2021) (Bakhtiari, Nasiri y Vahidi, 2023) (Alarfaj, Malik, Khan, Almusallam, Ramzan y Ahmed, 2022) (Carcillo, 2019). Además, una muestra de este tamaño permite mantener la

representatividad estadística necesaria para la evaluación robusta de los modelos, sin incurrir en sesgos ni comprometer la calidad de las conclusiones obtenidas como se menciona en (Credit Card Fraud Detection, 2017) (Dal Pozzolo, 2015).

El conjunto de datos reducido se dividió en subconjuntos de entrenamiento (40,000 registros) y prueba (10,000 registros), replicando la proporción de clases original y asegurando así la validez y la capacidad de generalización de los experimentos realizados. Esta estrategia permitió evaluar de manera sistemática múltiples técnicas de preprocesamiento, balanceo y modelado bajo condiciones controladas y comparables, sentando las bases para la selección de la mejor alternativa para su posterior integración en el sistema final.

Durante esta etapa, se evaluaron múltiples combinaciones sobre los modelos candidatos Regresión Logística y Random Forest en donde se incluyó técnicas de balanceo de clases, como SMOTE y submuestreo, con el objetivo de maximizar la capacidad de detección de fraudes y, en especial, minimizar la tasa de falsos negativos, tal como sugieren los lineamientos de la literatura (Lebichot, Paldino, Siblini, He-Guelton, Oblé y Bontempi, 2021) (Bakhtiari, Nasiri y Vahidi, 2023) (Bhattacharyya, Jha, Tharakunnel y Westland, 2011) (Carcillo, 2019).

Los resultados de los experimentos rápidos muestran que la Regresión Logística entrenada sobre los datos reducidos logra el mejor desempeño global, con un AUC ROC de 0.9837 y un Average Precision (AP) de 0.8929, superando tanto a las variantes balanceadas como a los modelos de Random Forest Tabla 6. El informe de clasificación sobre el conjunto de prueba revela una precision para la clase fraude de 0.96, un recall de 0.84 y un F1-score de 0.90, indicando un excelente equilibrio entre la detección efectiva de fraudes y el control de los falsos positivos. Esto se evidencia además en la matriz de confusión, donde de 98 fraudes reales, 82

fueron correctamente identificados, y solo 3 transacciones legítimas fueron clasificadas erróneamente como fraude.

Tabla 6

Resultados comparativos de modelos y técnicas de balanceo para la detección de fraude

Modelo	Técnica de Balanceo	AUC ROC	AP (Precisión Promedio)
Logistic Regression	Original	0.9837	0.8929
Logistic Regression	SMOTE_Fast	0.9787	0.8825
Logistic Regression	UnderSample	0.9807	0.8838
Random Forest Fast	Original	0.9705	0.8784
Random Forest Fast	SMOTE_Fast	0.9788	0.8831
Random Forest Fast	UnderSample	0.9768	0.8756

Nota. La tabla muestra los valores de AUC ROC (área bajo la curva ROC) y Precisión Promedio (AP) para cada modelo bajo diferentes técnicas de balanceo. Todos los modelos se entrenaron y evaluaron sobre una muestra estratificada de 50,000 transacciones, conservando la proporción original entre clases.

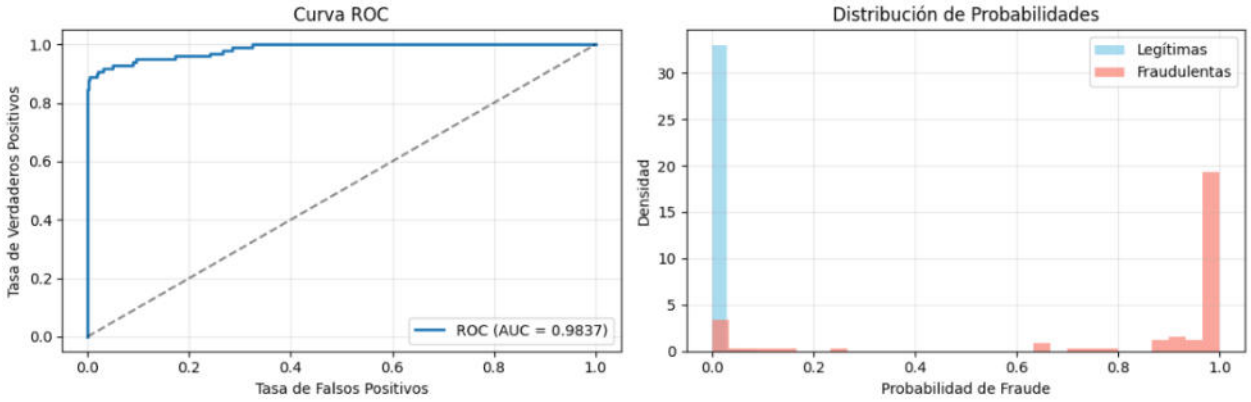
El análisis visual presentado en la **Figura 7**, mediante la curva ROC con un AUC de 0.9837, confirma la sobresaliente capacidad discriminativa del modelo Logistic Regression. Esta métrica sitúa la curva significativamente por encima de la diagonal de referencia, lo que indica que el modelo es capaz de mantener una baja tasa de falsos positivos incluso al incrementar la tasa de verdaderos positivos, aspecto fundamental en contextos de detección de fraude, donde los errores de falsos negativos tienen un alto costo operativo y reputacional. La proximidad de la curva ROC al vértice superior izquierdo del gráfico evidencia que el modelo logra distinguir

eficazmente entre transacciones legítimas y fraudulentas, alineándose con lo reportado en la literatura reciente para tareas de clasificación en dominios desbalanceados (Alarfaj, Malik, Khan, Almusallam, Ramzan y Ahmed, 2022) (Dang, Tran, Tuan y Tiep, 2021) (Btoush, Zhou, Gururajan, Chan, Genrich y Sankaran, 2023).

Adicionalmente, la distribución de probabilidades respalda este desempeño, mostrando que el modelo asigna probabilidades próximas a uno para la mayoría de las transacciones fraudulentas y valores cercanos a cero para las legítimas. Esta separación clara entre ambas clases refuerza la fiabilidad del modelo, ya que minimiza las ambigüedades y facilita la toma de decisiones automáticas en entornos de producción. Este comportamiento es especialmente relevante considerando el elevado desbalance de clases característico del problema, donde la correcta identificación de eventos anómalos representa un reto considerable . En conjunto con la matriz de confusión obtenida Figura 8 permiten estimar que la solución propuesta ofrece una robusta capacidad de discriminación, justificando que la selección del modelo de regresión logística es la opcion idónea para su posterior optimización e integración en el sistema de detección de fraudes.

Figura 7

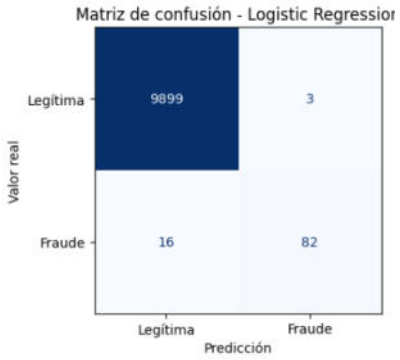
Curva ROC y distribución de probabilidades para el modelo Logistic Regression



Nota. La figura ilustra, a la izquierda, la curva ROC con un área bajo la curva (AUC) de 0.9837, que refleja la alta capacidad discriminativa del modelo para distinguir entre transacciones legítimas y fraudulentas. A la derecha, la gráfica de densidad muestra la clara separación entre las probabilidades asignadas a cada clase, con valores cercanos a 1 para fraudes y a 0 para transacciones legítimas, lo que evidencia la precisión del modelo en la asignación de riesgos.

Figura 8

Matriz de confusión del modelo Logistic Regression sobre el conjunto de prueba



Nota. La matriz de confusión muestra la clasificación de 10,000 transacciones, donde el modelo identifica correctamente la mayoría de las transacciones legítimas y fraudes, manteniendo una baja tasa de falsos negativos.

Luego de identificar a la regresión logística como la opción más robusta para la detección de fraude, gracias a su alta capacidad discriminativa y estabilidad en experimentos previos, se procedió a un proceso sistemático de optimización, dirigido específicamente a reducir la cantidad de fraudes no detectados (falsos negativos). Este enfoque responde a la alta sensibilidad requerida en escenarios reales, donde la omisión de un fraude podria implica un costo no solo económico significativamente mayor que la clasificación errónea de una transacción legítima.

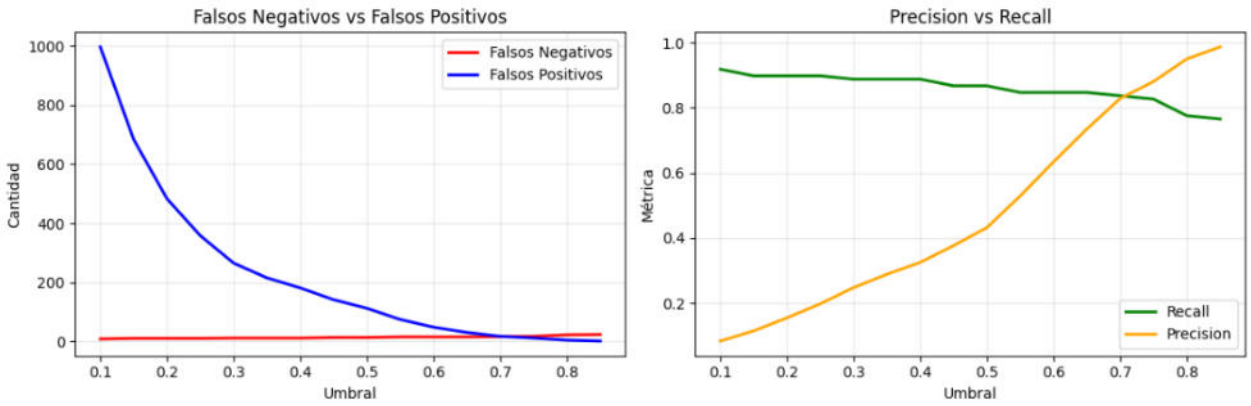
En primera instancia, se evaluaron diversas estrategias avanzadas de resampling, incluyendo técnicas como BorderlineSMOTE, ADASYN, SMOTETomek y SMOTEENN. Estas

acciones buscaron mejorar la capacidad del modelo para aprender patrones de la clase minoritaria (fraudes), balanceando el conjunto de entrenamiento de forma sintética. El uso de BorderlineSMOTE logró un balance favorable entre recall (0.8367) y precisión (0.9111), mientras que ADASYN alcanzó el mayor recall (0.8673) a costa de una importante reducción en la precisión (0.4315), evidenciando la clásica disyuntiva entre sensibilidad y exactitud en contextos de alta desbalance de clases.

Posteriormente, se buscó una optimización del umbral de decisión ajustando el valor por defecto (0.5) a 0.6, lo que permitió alcanzar un recall de 0.8469 con una precisión mejorada (0.6336). Los gráficos de la **Figura 9** ilustran este fenómeno: a medida que se incrementa el umbral de corte, los falsos positivos disminuyen drásticamente, pero los falsos negativos tienden a aumentar solo levemente dentro de un rango de umbrales críticos. Esto evidencia la importancia de la calibración del umbral, no solo en función del ROC AUC, sino considerando la relación de costo-beneficio propia del caso de uso.

Figura 9

Optimizacion de umbral



Nota. A la izquierda: Evolución de falsos negativos (línea roja) y falsos positivos (línea azul) según el umbral de decisión; a la derecha: curvas de precisión y recall en función del umbral. El

análisis visual evidencia el trade-off inherente a la optimización en tareas de detección de fraude, permitiendo seleccionar el umbral más adecuado para los objetivos del proyecto.

Adicionalmente, se exploró también la implementación de estrategias ensemble conservadoras, combinando varios modelos de Random Forest optimizados para alta sensibilidad. El ensemble, empleando un umbral de votación bajo, logró la mejor tasa de recall (0.8878), reduciendo los falsos negativos a 11 casos, aunque incrementando el número de falsos positivos a 209. Si bien la precisión resultante (0.2939) es menor, este resultado puede ser aceptable en contextos donde el costo de perder un fraude supera con creces el de investigar una alerta errónea.

Finalmente, se ensayaron técnicas de aprendizaje sensible a costos, en las que la penalización por errores en la clase minoritaria fue intensificada (ratios de hasta 1:25). Aunque se logró un recall igualmente alto (0.8878), la precisión cayó considerablemente, lo que puede saturar de falsas alarmas los sistemas de monitoreo humano.

En síntesis Tabla 7, el proceso de optimización confirma que la reducción de falsos negativos exige sacrificar parte de la precisión y que la mejor estrategia identificada es el uso de un ensemble conservador, ajustando el umbral de decisión y priorizando la sensibilidad sobre la especificidad. Estos hallazgos están alineados con la literatura reciente, donde se recomienda la aplicación de modelos robustos combinados con técnicas de resampling y ajuste de umbrales para el manejo efectivo de problemas de desbalance de clases en fraude financiero.

Tabla 7

Resultados comparativos de estrategias de optimización orientadas a la reducción de falsos negativos

Estrategia	TP	TN	FP	FN	Recall	Balanced Accuracy	F1-score
BorderlineSMOTE	82	9894	8	16	0.8367	0.9179	0.8723
ADASYN	85	9790	112	13	0.8673	0.9285	0.5795
SMOTETomek	83	9888	14	15	0.8469	0.9228	0.8513
SMOTEENN	83	9887	15	15	0.8469	0.9227	0.8469
SMOTE_Aggressive	83	9891	11	15	0.8469	0.9229	0.8646
OptimalThreshold (ADASYN)	83	9854	48	15	0.8469	0.9211	0.7220
Conservative Ensemble	87	9693	209	11	0.8878	0.9334	0.4413
CostSensitive (1:15)	85	9664	238	13	0.8673	0.9216	0.4041
CostSensitive (1:20)	85	9584	318	13	0.8673	0.9176	0.3406
CostSensitive (1:25)	87	9504	398	11	0.8878	0.9239	0.2974

Nota. La tabla resume los resultados de distintas estrategias para optimizar la reducción de fraudes no detectados (falsos negativos). El mejor desempeño en términos de recall se alcanzó con el ensemble conservador, sacrificando precisión a cambio de máxima sensibilidad, lo que es recomendable cuando el costo de perder un fraude es elevado. Fuente: elaboración propia a partir de resultados experimentales.

4.10. Implementación del modelo de clasificación

Finalizado el proceso de experimentación, optimización y validación de modelos, se procedió a la implementación práctica de la solución de detección de fraude, orientando los

esfuerzos hacia la preparación de los artefactos necesarios para su despliegue en entornos productivos y para la futura generación de datos sintéticos.

Para asegurar la portabilidad y reproducibilidad del sistema, se exportaron los modelos entrenados, junto con los escaladores y el orden de columnas exacto utilizado en cada pipeline. En concreto, los artefactos generados y almacenados incluyen:

- **Modelo base:** best_base_model.pkl
- **Escalador del modelo base:** amount_scaler.pkl
- **Orden de columnas (features) del modelo base:** base_feature_order.pkl
- **Modelo optimizado (mejor estrategia):** best_strategy_model.pkl
- **Escalador del modelo optimizado:** strategy_scaler.pkl
- **Orden de columnas del modelo optimizado:** strategy_feature_order.pkl

Estos archivos, exportados en formato **Pickle** mediante la biblioteca joblib, garantizan la posibilidad de reutilizar los modelos en cualquier plataforma Python compatible, facilitando así la integración en aplicaciones web, APIs REST o pipelines automatizados. El orden de las columnas es fundamental para evitar errores en la etapa de inferencia y asegurar la consistencia de los datos de entrada, el uso en producción consiste en:

1. **Carga de artefactos:** Utilizar joblib.load('ruta/al/archivo.pkl') para restaurar cada objeto (modelo, escalador, orden de columnas).
2. **Preparación de datos:** Antes de la predicción, el DataFrame con nuevas transacciones debe reordenarse: df = df[feature_order].
3. **Escalado selectivo:** Escalar únicamente la columna 'Amount' con el escalador correspondiente, replicando el preprocesamiento aplicado durante el entrenamiento.
4. **Aplicación de umbral óptimo:** En el caso de utilizar el modelo optimizado, es necesario

emplear el threshold determinado durante la optimización para clasificar las transacciones.

Adicionalmente, se desarrolló y almacenó el artefacto synth_stats.pkl, que contiene las estadísticas clave (medias y matrices de covarianza) para cada clase (legítima y fraude), así como el orden de las variables empleadas. Este recurso es fundamental para la generación de transacciones sintéticas compatibles, ya sea para pruebas de estrés, simulaciones, o para balancear datasets en nuevas fases de entrenamiento. Su uso consiste en:

- Cargar el artefacto con joblib.load().
- Generar muestras mediante la función numpy.random.multivariate_normal(), respetando el orden de columnas especificado en feature_order.

La correcta gestión de estos artefactos es indispensable para la robustez y reproducibilidad de la solución en producción, permitiendo una actualización ágil ante la llegada de nuevos datos o ante la necesidad de reentrenar el sistema frente a patrones de fraude emergentes.

Este enfoque potencia la capacidad del sistema para adaptarse a escenarios cambiantes, evaluar robustez ante nuevas amenazas y facilitar la investigación en ambientes controlados.

Con el objetivo de facilitar el acceso, la usabilidad y la interpretación de la solución de detección de fraudes, se desarrolló una aplicación web interactiva utilizando la plataforma Streamlit. Esta herramienta permitio la implementación de interfaces gráficas, asegurando que tanto usuarios técnicos como no técnicos puedan simular, analizar y comprender el funcionamiento del sistema de detección de fraudes en tiempo real.

4.11. Despliegue del modelo de clasificación

La aplicación se diseñó para cargar dinámicamente los artefactos exportados en la fase anterior (modelos, escaladores y configuraciones de features), garantizando la consistencia entre los procesos de entrenamiento y de inferencia. La arquitectura central incluye:

- **Carga de artefactos:** Al inicio, la aplicación busca y carga los modelos entrenados (best_base_model.pkl y best_strategy_model.pkl), los escaladores correspondientes para la variable Amount y los archivos con el orden de columnas esperado. Esto asegura que cualquier transacción analizada reciba el preprocesamiento idéntico al aplicado durante el entrenamiento.
- **Simulación y análisis de transacciones:** La aplicación permite la simulación de transacciones tanto legítimas como fraudulentas. Los usuarios pueden ingresar datos de una tarjeta, monto y parámetros contextuales (categoría de comercio, ubicación). Además, se genera automáticamente un conjunto de variables sintéticas (V1-V28 y Amount) usando estadísticas reales del dataset (synth_stats.pkl) o, en su defecto, perfiles demo calibrados. Esta flexibilidad asegura que la solución pueda ser evaluada en escenarios controlados o reales.
- **Procesamiento y predicción:** Una vez recibida la transacción, se reordenan las columnas, se escala la variable Amount y se realiza la predicción de fraude utilizando tanto el modelo base como el modelo optimizado (con threshold ajustable). La aplicación reporta la probabilidad de fraude, la predicción de cada modelo y un score de riesgo compuesto, brindando información clave para la toma de decisiones.
- **Interpretabilidad del modelo:**

  La aplicación incorpora técnicas de interpretabilidad avanzada, mostrando al usuario la

importancia de las variables más relevantes a través de gráficos interactivos generados con SHAP y LIME. Además, integra una función de explicación automatizada mediante lenguaje natural, soportada por modelos de lenguaje de Hugging Face, que permite ofrecer explicaciones claras y comparativas de las decisiones de ambos modelos, adaptadas a usuarios no técnicos.

- **Interfaz visual y experiencia de usuario:**

  Se ofrece una vista previa personalizada de la tarjeta, métricas instantáneas de evaluación, visualizaciones comparativas de probabilidades, y un resumen detallado de las características de la transacción evaluada. El panel lateral permite ajustar parámetros, consultar el estado de los artefactos cargados y acceder a estadísticas de uso de la sesión.

- **Gestión y trazabilidad:**

  Todas las transacciones procesadas se contabilizan, registrando la cantidad de fraudes detectados y permitiendo así el monitoreo del desempeño de la aplicación en diferentes contextos de prueba.

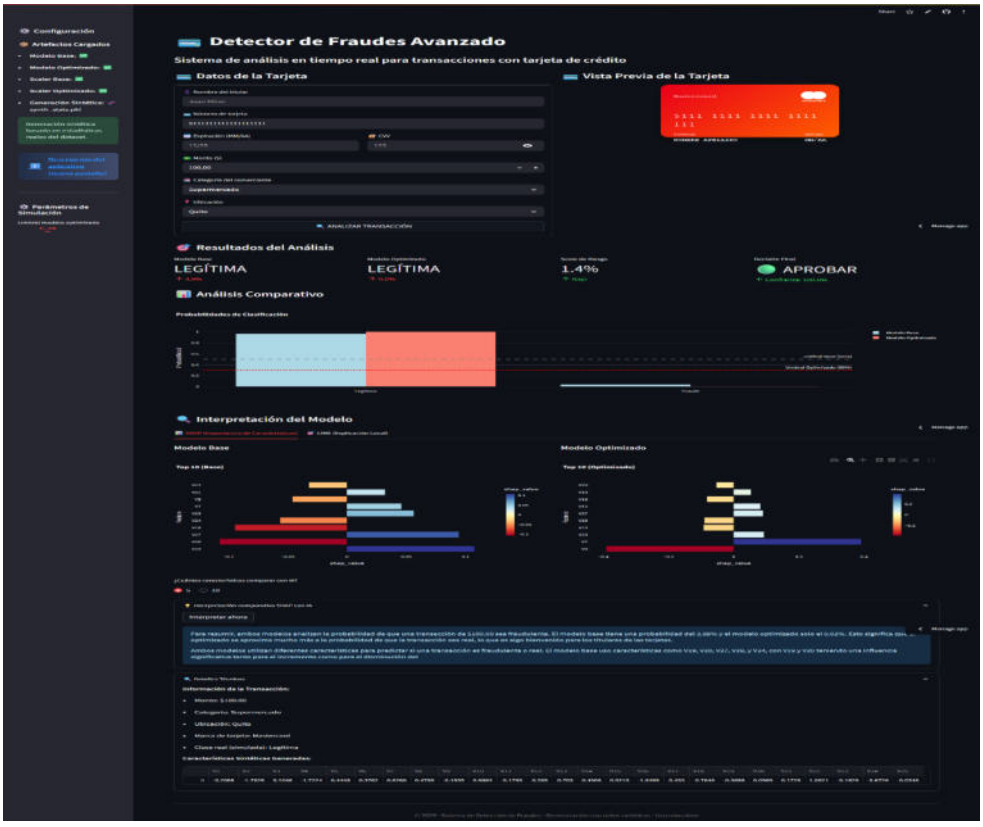
El despliegue llevado a cabo como demo ofrece varias ventajas:

- Permite el análisis de nuevas transacciones de forma inmediata, facilitando el diagnóstico, la auditoría y la capacitación de usuarios finales.
- Favorece la reproducibilidad y portabilidad de la solución, pues basta con disponer de los artefactos exportados para levantar la aplicación en cualquier entorno Python.
- La integración de explicaciones interpretables y visualizaciones interactivas mejora la transparencia del sistema, incrementando la confianza en las predicciones del modelo y facilitando la toma de decisiones basada en evidencia.

A continuación, la **Figura 10** muestra la interfaz principal de la aplicación web desarrollada para la detección avanzada de fraudes en tarjetas de crédito. A través de esta plataforma, los usuarios pueden ingresar los datos de una transacción y visualizar, en tiempo real, el análisis y la decisión automática del modelo implementado (aprobación o bloqueo), junto con métricas de riesgo, probabilidades de clasificación y gráficos de interpretabilidad (SHAP y LIME) para ambos modelos (base y optimizado). El sistema también despliega detalles técnicos relevantes y facilita la comprensión mediante explicaciones generadas con inteligencia artificial.

Figura 10

Interfaz de la aplicación de detección de fraudes en producción



Nota. La figura presenta la versión en producción de la aplicación Streamlit para la evaluación en línea de transacciones con tarjetas de crédito, mostrando la integración de los modelos exportados y los artefactos de interpretabilidad desarrollados en este estudio. La visualización

incluye todas las funcionalidades implementadas y corresponde a una simulación con datos sintéticos, usada exclusivamente con fines educativos y demostrativos.

Capítulo 5

5. Conclusiones y Recomendaciones

Conclusiones

1. El desarrollo y despliegue de un sistema de detección de fraude en transacciones con tarjetas de crédito, apoyado en técnicas de machine learning, ha evidenciado la complejidad inherente a los problemas de desbalance de clases y la importancia de una correcta selección y procesamiento de variables. En este estudio, se determinó que la variable "Time" no aporta valor predictivo significativo al problema de detección de fraude, presentando baja correlación con el evento de interés y nula relevancia para la discriminación entre transacciones legítimas y fraudulentas.
2. Una de las principales prioridades fue reducir la tasa de falsos negativos (fraudes no detectados), pues estos representan un riesgo operativo y financiero directo para las instituciones emisoras. Sin embargo, la disminución de los falsos negativos suele incrementarse a expensas de un aumento en los falsos positivos, lo que puede traducirse en costos adicionales asociados al bloqueo o revisión manual de transacciones legítimas. Este trade-off exige un análisis económico y de riesgo ajustado al contexto de cada organización, pues la tolerancia a falsos positivos depende de la estrategia de negocio y la capacidad operativa para gestionar alertas.
3. Los experimentos realizados demuestran que la estratificación en la división de los conjuntos de entrenamiento y prueba es indispensable para preservar la proporción real de fraudes en los distintos subconjuntos, asegurando la validez estadística de la evaluación y

la reproducibilidad de los resultados. Ignorar esta recomendación puede generar sobreestimaciones de desempeño e inducir a error en la toma de decisiones, especialmente en datasets altamente desbalanceados como el analizado en este trabajo

4. Contrario a lo esperado, se observó que modelos de alta complejidad (por ejemplo, redes neuronales profundas o ensembles sofisticados) no necesariamente superan a algoritmos más sencillos —como la regresión logística o Random Forest— cuando se aplican sobre conjuntos de datos severamente desbalanceados o de naturaleza tabular, tal como ha sido reportado en múltiples trabajos recientes (Alarfaj, Khan, Almusallam, Ramzan y Ahmed, 2022) (Alfaiz y Fati, 2022) (Bin Sulaiman, Schetinin y Sant, 2022) (Lucas y Jurgovsky, 2020). En particular, la regresión logística, debidamente calibrada y combinada con técnicas de balanceo y ajuste de umbral, ofreció resultados competitivos, alta interpretabilidad y mayor robustez frente a posibles sesgos o sobreajuste.
5. El trabajo desarrollado en este estudio proporciona un marco metodológico robusto y replicable para la detección automatizada de fraudes en transacciones con tarjetas de crédito, lo cual reviste especial importancia para las entidades financieras del Ecuador en el contexto actual de digitalización y aumento de amenazas cibernéticas. La integración de técnicas avanzadas de preprocesamiento, modelado supervisado y optimización orientada a la reducción de falsos negativos permite no solo mejorar la tasa de detección de fraudes, sino también minimizar el impacto operativo y reputacional asociado a este tipo de situaciones de fraude.
6. El desarrollo de los artefactos generados en este trabajo, tanto modelos como escaladores y esquemas de orden de variables, representa una ventaja estratégica clave para las instituciones financieras. Gracias al enfoque modular y estandarizado en la exportación,

portabilidad e integración de los modelos, la solución desarrollada no se limita exclusivamente a su aplicación en pasarelas de pago o aplicaciones web como las presentadas, sino que puede ser fácilmente adaptada e integrada en infraestructuras tecnológicas más amplias de seguridad bancaria, tales como sistemas de monitoreo transaccional en tiempo real, plataformas antifraude centralizadas o módulos de autenticación adaptativa. Esta flexibilidad facilita la interoperabilidad, reduce los costos de adopción y maximiza el impacto potencial de la herramienta en la prevención y detección temprana de fraudes en diferentes puntos críticos de las entidades financieras.

Recomendaciones

1. Se recomienda excluir variables como “Time” cuando no demuestran aporte informativo significativo, priorizando la inclusión de atributos con correlación o importancia demostrada en el proceso de detección de fraude. Sin embargo, ante la falta de información detallada sobre el significado de ciertas variables anónimas (V1-V28), es recomendable conservarlas en el modelado para evitar la pérdida de patrones útiles.
2. La selección del umbral de decisión debe considerar el balance entre recall y precisión, integrando análisis de costo-beneficio que cuantifique el impacto económico de los falsos negativos versus los falsos positivos. En entornos bancarios y financieros, una reducción drástica de los falsos negativos puede justificar un mayor número de falsos positivos, siempre que exista capacidad operativa para su gestión y el costo de un fraude no detectado sea superior al de una alerta falsa.
3. Mantener la proporción real de clases en todas las fases de entrenamiento y validación es fundamental para garantizar la representatividad y robustez del modelo en escenarios

reales. Se recomienda aplicar validación cruzada estratificada y técnicas de resampling controlado.

4. No siempre los modelos más sofisticados logran mejores desempeños en problemas desbalanceados. Algoritmos como la regresión logística y Random Forest, combinados con balanceo de clases, escalado selectivo de variables y ajuste de umbral, pueden ser más efectivos, interpretables y menos costosos computacionalmente. Se sugiere priorizar la sencillez y la transparencia del modelo en la fase de producción.
5. La estrategia de optimización debe alinearse con la política de riesgos de la organización. Se recomienda una revisión periódica de los parámetros de configuración y métricas objetivo, incorporando análisis de impacto financiero, satisfacción de usuarios y capacidad de respuesta operativa.
6. Se recomienda que futuros trabajos exploren técnicas avanzadas de ingeniería de características, incluyendo la creación de variables derivadas a partir de las existentes (feature construction) y la utilización de métodos automáticos de selección e interpretación de variables, como SHAP o LIME, para identificar atributos con mayor aporte discriminativo. Profundizar en la comprensión y documentación del significado de cada variable podría permitir el diseño de modelos aún más eficientes y transparentes, así como facilitar la transferencia de resultados a otros dominios o entidades financieras.
7. A las entidades financieras del Ecuador, en coordinación con organismos reguladores, asociaciones gremiales y la academia, impulsen la creación de repositorios de datos abiertos y anonimizados sobre transacciones y fraudes en el contexto local. Esta iniciativa, alineada con las mejores prácticas internacionales en ciencia de datos y transparencia, facilitaría el acceso de investigadores y desarrolladores a información representativa de la

realidad ecuatoriana. Así, sería posible diseñar y validar modelos de detección de fraude más precisos, adaptados al comportamiento transaccional y a las amenazas específicas del entorno nacional, potenciando la innovación, la colaboración interinstitucional y la efectividad de las estrategias de prevención de fraude en el sistema financiero.

Referencias

Aburbeian, A. M., and Ashqar, H. I. (2023). *Credit card fraud detection using enhanced random forest classifier for imbalanced data*. In Proc. Int. Conf. Adv. Comput. Res. Cham, Switzerland: Springer, pp. 605–616.

Alamri, M., and Ykhlef, M. (2022). *Survey of credit card anomaly and fraud detection using sampling techniques*. Electronics, vol. 11, no. 23, p. 4003.

Alarfaj, F. K., Malik, I., Khan, H. U., Almusallam, N., Ramzan, M., and Ahmed, M. (2022). *Credit card fraud detection using state-of-the-art machine learning and deep learning algorithms*. IEEE Access, vol. 10, pp. 39700–39715.

Alfaiz, N. S. and Fati, S. M. (2022). *Enhanced credit card fraud detection model using machine learning*. Electronics, vol. 11, no. 4, p. 662.

Al-Hashedi, K. G., and Magalingam, P. (2021). *Financial fraud detection applying data mining techniques: A comprehensive review from 2009 to 2019*. Comput. Sci. Rev., vol. 40, Art. no. 100402.

Ambashtha, K. L., and Kumar P. (2023). *Online fraud, in Financial Crimes: A Guide to Financial Exploitation in a Digital Age*. Berlin, Germany: Springer, pp. 97–108.

ASOBANCA. (2023). *Informe sobre seguridad bancaria y fraude financiero en Ecuador*. Asociación de Bancos Privados del Ecuador.

Bahnsen, A. (2016). *Feature engineering strategies for credit card fraud detection*. Expert Systems with Applications, 51, 134–142.

Bakhtiari, S., Nasiri, Z., and Vahidi, J. (2023). *Credit card fraud detection using ensemble data mining methods*. Multimedia Tools Appl., vol. 82, no. 19, pp. 29057–29075.

Benchaji, I., Douzi, S., El Ouahidi, B., and Jaafari, J. (2021). *Enhanced credit card fraud detection based on attention mechanism and LSTM deep model*. J. Big Data, vol. 8, no. 1, pp. 1–21.

Berhane, T., Melese, T., Walelign, A., and Mohammed, A. (2023). *A hybrid convolutional neural network and support vector machine-based credit card fraud detection model*. Math. Problems Eng., vol. 2023, pp. 1–10.

Bhattacharyya, S., Jha, S., Tharakunnel, K., and Westland, J. C. (2011). *Data mining for credit card fraud: A comparative study*. Decision Support Systems, 50(3), 602–613.

Bin Sulaiman, R., Schetinin, V., and Sant P. (2022). *Review of machine learning approach on credit card fraud detection*. Hum.-Centric Intell. Syst., vol. 2, no. 1, pp. 55–68, 2022.

Btoush, E. A. L. M., Zhou, X., Gururajan, R., Chan, K. C., Genrich, R., and Sankaran, P. (2023). *A systematic review of literature on credit card cyber fraud detection using machine and deep learning*. PeerJ Comput. Sci., vol. 9, p. e1278.

Buda, M., Maki, A., and Mazurowski, M. A. (2018). *A systematic study of the class imbalance problem in convolutional neural networks*. Neural Networks, 106, 249–259.

Carcillo, F. (2019). *Combining unsupervised and supervised learning in credit card fraud detection*. Information Sciences, 557, 317–331.

Carcillo, F., Le Borgne, Y. A., Caelen, O., and Bontempi, G. (2018). Scarff: *A scalable framework for streaming credit card fraud detection with spark*. Information Fusion, 41, 182–194.

Chen, T., and Guestrin, C. (2016). XGBoost: *A scalable tree boosting system*. In Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining (pp. 785-794).

Credit Card Fraud Detection. (17 de octubre de 2017). https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud

Dal Pozzolo, A. (2015). *Learned lessons in credit card fraud detection from a practitioner perspective*. Expert Systems with Applications, 41(10), 4915–4928.

Dang, T. K., Tran, T. C., Tuan, L. M., y Tiep, M. V. (2021). *Machine learning based on resampling approaches and deep reinforcement learning for credit card fraud detection systems*. Appl. Sci., vol. 11, no. 21, p. 10004.

Dong, S., Xia, Y., and Peng, T. (2021). *Network abnormal traffic detection model based on semi-supervised deep reinforcement learning*. IEEE Trans. Netw. Service Manag., vol. 18, no. 4, pp. 4197–4212.

El-Naby, A. A., Hemdan, E. E.-D., and El-Sayed, A. (2023). *An efficient fraud detection framework with credit card imbalanced data in financial services*. Multimedia Tools Appl., vol. 82, no. 3, pp. 4139–4160.

Esenogho, E., Mienye, I. D., Swart, T. G., Aruleba, K., and Obaido, G. (2022). *A neural network ensemble with feature engineering for improved credit card fraud detection*. IEEE Access, vol. 10, pp. 16400–16407.

Fanai, H., and Abbasimehr, H. (2023). *A novel combined approach based on deep autoencoder and deep classifiers for credit card fraud detection*. Exp. Syst. Appl., vol. 217, Art. no. 119562.

Gambo, M. L., Zainal, A., and Kassim, M. N. (2022). *A convolutional neural network model for credit card fraud detection.* In Proc. Int. Conf. Data Sci. Appl. (ICoDSA), pp. 198–202.

Gold, S. (2014). *The evolution of payment card fraud*. Comput. Fraud Secur., vol. 2014, no. 3, pp. 12–17.

Guers, K., Chowdhury, M. M., and Rifat, N. (2022). Card skimming: *A cybercrime by hackers*. In Proc. IEEE Int. Conf. Electro Inf. Technol. (eIT), pp. 575–579.

Illanko, K., Soleymanzadeh, R., and Fernando, X. (2022). *A big data deep learning approach for credit card fraud detection in Computer Networks, Big Data and IoT*. Cham, Switzerland: Springer, pp. 633–641.

Islam, M. A., Uddin, M. A., Aryal, S., and Stea, G. (2023). *An ensemble learning approach for anomaly detection in credit card data with imbalanced and overlapped classes*. J. Inf. Secur. Appl., vol. 78, Art. no. 103618.

Jaiswal, S., and Jadav, N. J. (2021). *Credit Card Fraud Detection using Machine Learning Algorithms*. SSRN Electronic Journal.

Kafhali S. E., and Tayebi, M. (2022). *XGBoost based solutions for detecting fraudulent credit card transactions*. In Proc. Int. Conf. Adv. Creative Netw. Intell. Syst. (ICACNIS), pp. 1–6.

Karthik, V. S. S., Mishra, A., and Reddy, U. S. (2022). *Credit card fraud detection by modelling behaviour pattern using hybrid ensemble model*. Arabian J. Sci. Eng., vol. 47, no. 2, pp. 1987–1997.

Karthika J., and Senthilselvi, A. (2023). *An integration of deep learning model with navo minority over-sampling technique to detect the frauds in credit cards*. Multimedia Tools Appl., vol. 82, no. 14, pp. 21757–21774.

Karthika, J., and Senthilselvi, A. (2023). *Smart credit card fraud detection system based on dilated convolutional neural network with sampling technique*. Multimedia Tools Appl., vol. 82, no. 20, pp. 31691–31708.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., and Liu, T. Y. (2017). LightGBM: *A highly efficient gradient boosting decision tree*. Advances in neural information processing systems, 30.

Lebichot, B., Paldino, G. M., Siblini, W., He-Guelton, L., Oblé, F., and Bontempi, G. (2021). *Incremental learning strategies for credit cards fraud detection*. Int. J. Data Sci. Anal., vol. 12, no. 2, pp. 165–174.

Lucas Y., and Jurgovsky, J. (2020). *Credit card fraud detection using machine learning: A survey*. arXiv:2010.06479.

Mienye, I. D. and Jere, N. (2024). *A survey of decision trees: Concepts, algorithms, and applications*. IEEE Access, vol. 12, pp. 86716–86727.

Mienye, I. D., and Sun, Y. (2023). *A deep learning ensemble with data resampling for credit card fraud detection*. IEEE Access, vol. 11, pp. 30628–30638.

Mienye, I. D., and Sun, Y. (2023). *A machine learning method with hybrid feature selection for improved credit card fraud detection*. Appl. Sci., vol. 13, no. 12, p. 7254.

Modi, K., and Dayma, R. (2017). *Review on fraud detection methods in credit card transactions*. In Proc. Int. Conf. Intell. Comput. Control (IC), pp. 1–5.

Nguyen, M., Pham, T., y Tran, H. (2024). *Deep Learning for Credit Card Fraud Detection: A Review of Algorithms, Challenges, and Solutions*. https://www.researchgate.net/publication/382187222

Osegi, E. N., and Jumbo, E. F. (2021). *Comparative analysis of credit card fraud detection in simulated annealing trained artificial neural network and hierarchical temporal memory*. Mach. Learn. Appl., vol. 6. Art. no. 100080.

Pandey, K., Sachan, P., and Ganpatrao, N. G. (2021). *A review of credit card fraud detection techniques*. In Proc. 5th Int. Conf. Comput. Methodologies Commun. (ICCMC), pp. 1645–1653.

Popat, R. R., and Chaudhary, J. (2018). *Survey on credit card fraud detection using machine learning.* In Proc. 2nd Int. Conf. Trends Electron. Informat. (ICOEI), pp. 1120–1125.

Pozzolo, A. D., Caelen, O., Le Borgne, Y. A., Waterschoot, S., and Bontempi, G. (2015). *Calibrating probability with undersampling for unbalanced classification*. In 2015 IEEE Symposium Series on Computational Intelligence.

Prabhakaran, N., and Nedunchelian, R. (2023). *Oppositional cat swarm optimization-based feature selection approach for credit card fraud detection*. Comput. Intell. Neurosci., vol. 2023, pp. 1–13.

Raval, J., Bhattacharya, P., Jadav, N. K., Tanwar, S., Sharma, G., Bokoro, P. N., Elmorsy, M., Tolba, A., and Raboaca, M. S. (2023). RaKShA: *A trusted explainable*

LSTM model to classify fraud patterns on credit card transactions. Mathematics, vol. 11, no. 8, p. 1901.

Roy, A., Sun, J., Mahoney, W., and Harms, M. (2018). *Deep learning detecting fraud in credit card transactions*. Proceedings of the 17th IEEE International Conference on Machine Learning and Applications.

Ryman-Tubb, N. F., Krause, P., and Garn, W. (2018). *How artificial intelligence and machine learning research impacts payment card fraud detection: A survey and industry benchmark*. Eng. Appl. Artif. Intell., vol. 76, pp. 130–157.

Sahin, Y., and Duman, E. (2011). *Detecting credit card fraud by ANN and logistic regression*. Proceedings of the International Symposium on Innovations in Intelligent Systems and Applications.

Sehrawat, D., and Singh, Y. (2023). *Auto-encoder and LSTM-based credit card fraud detection*. Social Netw. Comput. Sci., vol. 4, no. 5, p. 557.

Superintendencia de Bancos del Ecuador. (2022). *Reporte de incidentes financieros y tendencias de fraude en el sector bancario ecuatoriano*. Superintendencia de Bancos del Ecuador.

Van Belle, R., Baesens, B., and De Weerdt, J. (2023). CATCHM: *A novel network-based credit card fraud detection method using node representation learning*. Decis. Support Syst., vol. 164, Art. no. 113866.

Wang, J., Liu, W., Kou, Y., Xiao, D., Wang, X., and Tang, X. (2023). *Approx-SMOTE federated learning credit card fraud detection system.* In Proc. IEEE 47th Annu. Comput., Softw., Appl. Conf. (COMPSAC), pp. 1370–1375.

Wang, P., Fan, E., and Wang, P. (2021). *Comparative analysis of image classification algorithms based on traditional machine learning and deep learning*. Pattern Recognit. Lett., vol. 141, pp. 61–67.

Wang, Z., Kim, S., and Joe, I. (2023). *An improved LSTM-based failure classification model for financial companies using natural language processing*. Appl. Sci., vol. 13, no. 13, p. 7884.

Xie, Y., Liu, G., Yan, C., Jiang, C., Zhou, M., and Li, M. (2024). *Learning transactional behavioral representations for credit card fraud detection*. IEEE Trans. Neural Netw. Learn. Syst., vol. 35, no. 4, pp. 5735–5748.

Yang, M.-H., Luo, J.-N., Vijayalakshmi, M., and Shalinie, S. M. (2022). *Contactless credit cards payment fraud protection by ambient authentication*. Sensors, vol. 22, no. 5, p. 1989.

Zhang, X., Han, Y., Xu, W., and Wang, Q. (2021). HOBA: *A novel feature engineering methodology for credit card fraud detection with a deep learning architecture*. Inf. Sci., vol. 557, pp. 302–316.