PROPUESTA INICIAL DEL PROYECTO
DETECCIÓN DE ANOMALÍAS EN TRANSACCIONES DE TARJETAS DE CRÉDITO

GRUPO 25
Andrés Felipe Gualdrón Gutiérrez
Jersson Hernán Morales Hernández
Carina Lizebeth Ordoñez Araque

Aprendizaje No Supervisado – MIAD 2024-14
Septiembre 2024

1 RESUMEN

Este informe presenta la propuesta inicial del proyecto para detección de anomalías en tarjetas de crédito utilizando técnicas avanzadas de aprendizaje de máquina no supervisado, específicamente, mediante el uso de los algoritmos **Isolation Forest** y **One-Class SVM**. El objetivo es abordar el problema del fraude transaccional, que representa una amenaza significativa para las instituciones financieras debido a su creciente sofisticación e impacto en el sistema financiero. El proyecto busca aplicar modelos de detección de anomalías para identificar transacciones fraudulentas haciendo uso de un conjunto de datos extremadamente desbalanceado de transacciones con tarjetas de crédito realizadas en Europa en septiembre de 2013 para el entrenamiento y prueba de estos modelos.

La metodología incluye la exploración y el preprocesamiento de datos, la implementación de los modelos seleccionados, y la comparación de su desempeño mediante métricas como **PRECISIÓN, RECALL Y F1-SCORE,** junto con el análisis de la curva **ROC** y el valor **AUC**. La evaluación también pretende contemplar el impacto financiero de las predicciones incorrectas. Adicionalmente, para la interpretación de los resultados se hará uso de técnicas de reducción de dimensiones como **t-SNE** para facilitar la visualización de las anomalías.

En la fase inicial de exploración del conjunto de datos que se va usar para el entrenamiento de los modelos, se determinó que el número de transacciones registradas es de 284,807 de las cuales 492 fueron clasificadas como fraudulentas mostrando un desbalance considerable en las clases de la variable de respuesta. Las 30 variables dependientes son numéricas, y a excepción de las variables TIME y AMOUNT, las restantes 28 variables corresponden a los componentes principales de los datos originales, considerados confidenciales. El conjunto no contiene datos faltantes ni categóricos.

La propuesta incluye estudios relevantes que respaldan la elección de los algoritmos, incluyendo enfoques híbridos y técnicas de optimización de hiperparámetros. Finalmente, concluye con una propuesta metodológica para implementar y evaluar los modelos de detección de anomalías, destacando la importancia del procesamiento para el desarrollo de los modelos y del posprocesamiento para facilitar el reporte de resultados.

2 INTRODUCCIÓN

En el ámbito del aprendizaje automático, el curso de Aprendizaje No Supervisado ha brindado una sólida base teórica y práctica sobre técnicas avanzadas para la detección de anomalías. Este trabajo final tiene como objetivo aplicar estos conocimientos en un proyecto real, centrado en la detección de fraudes en tarjetas de crédito utilizando modelos de aprendizaje no supervisado.

El fraude en tarjetas de crédito representa una amenaza significativa para las instituciones financieras y para sus clientes debido a su creciente sofisticación y al impacto financiero asociado. Por lo tanto, la elección adecuada de los algoritmos de detección es crucial para identificar transacciones fraudulentas con precisión y eficacia.

En este proyecto, se utilizarán modelos avanzados de detección de anomalías como **ISOLATION FOREST** y **ONE-CLASS SVM**. Estos modelos son seleccionados debido a su capacidad para manejar datos desbalanceados y su eficacia en la identificación de patrones anómalos. Se presentarán estudios relevantes que justifican esta elección, incluyendo un análisis comparativo de algoritmos de aprendizaje no supervisado y la aplicación de técnicas de optimización de hiperparámetros.

La metodología propuesta abarca desde la exploración y el preprocesamiento de los datos hasta la implementación y evaluación de los modelos seleccionados. Se utilizarán técnicas de

aprendizaje no supervisado para detectar anomalías en las transacciones de tarjetas de crédito. Además, se evaluarán los modelos utilizando métricas de desempeño y se analizará el impacto financiero de las predicciones incorrectas.

Este enfoque metodológico permitirá no solo implementar técnicas avanzadas de detección de anomalías, sino también comparar su eficacia en el contexto de datos reales. La visualización de las anomalías se realizará mediante técnicas de reducción de dimensiones, facilitando la interpretación y análisis de los resultados obtenidos.

3 REVISIÓN PRELIMINAR DE ANTECEDENTES EN LA LITERATURA

Para fundamentar la selección de ISOLATION FOREST y ONE-CLASS SVM como modelos de predicción en nuestro proyecto, se consideraron varios estudios relevantes que demuestran la eficacia y ventajas de estos algoritmos en la detección de anomalías en tarjetas de crédito. A continuación, se presentan las principales investigaciones que respaldan esta elección:

"Hybrid Machine Learning Approach for Data Anomaly Detection in Credit Card Transactions"
Este artículo aborda la detección de anomalías en transacciones de tarjetas de crédito mediante un enfoque híbrido que combina Isolation Forest (IF) y One-Class Support Vector Machine (OCSVM), ambos algoritmos no supervisados, con un clasificador Random Forest. Los resultados muestran que la combinación de IF y OCSVM mejora la precisión en la detección de anomalías en comparación con el uso individual de cada algoritmo. La investigación valida la eficacia de estos métodos no supervisados para manejar grandes volúmenes de datos y detectar patrones anómalos.

"Credit Card Fraud Detection Using Machine Learning Algorithms"
Este artículo explora técnicas avanzadas para detectar fraude en tarjetas de crédito mediante aprendizaje automático. Destaca el uso de agrupamiento de patrones históricos y una ventana deslizante para analizar datos de transacciones en streaming. Incluye mecanismos para manejar el "concept drift" y emplea algoritmos como Local Outlier Factor e Isolation Forest sobre datos transformados con PCA. Este enfoque integral es relevante para el proyecto al combinar técnicas de agrupamiento y detección de anomalías, mejorando la identificación de fraudes en datos desbalanceados.

"Enhanced Particle Swarm Optimization-Based Hyperparameter Optimized Stacked Autoencoder for Credit Card Fraud Detection"
Este artículo presenta un enfoque avanzado que combina la optimización de hiperparámetros con autoencoders apilados para la detección de fraude. Aunque se enfoca en el aprendizaje profundo, la metodología aplicada puede ser relevante para mejorar el rendimiento de modelos no supervisados como Isolation Forest y One-Class SVM. La mejora en la precisión del modelo y la comparación con otros enfoques destacan la importancia de ajustar los parámetros para obtener mejores resultados en la detección de anomalías.

"Survey of Fraud Detection Techniques for Credit Card Transactions"
Este artículo ofrece una revisión exhaustiva de las técnicas actuales para la detección de fraude en tarjetas de crédito, incluyendo metodologías tradicionales y modernas. Proporciona un contexto general para comparar diversas técnicas de detección de fraude y destaca los desafíos asociados, como el manejo de datos desbalanceados y la adaptación a nuevas técnicas de fraude. La revisión ayuda a contextualizar la elección de Isolation Forest y One-Class SVM, al proporcionar una base para evaluar su eficacia en comparación con otras metodologías.

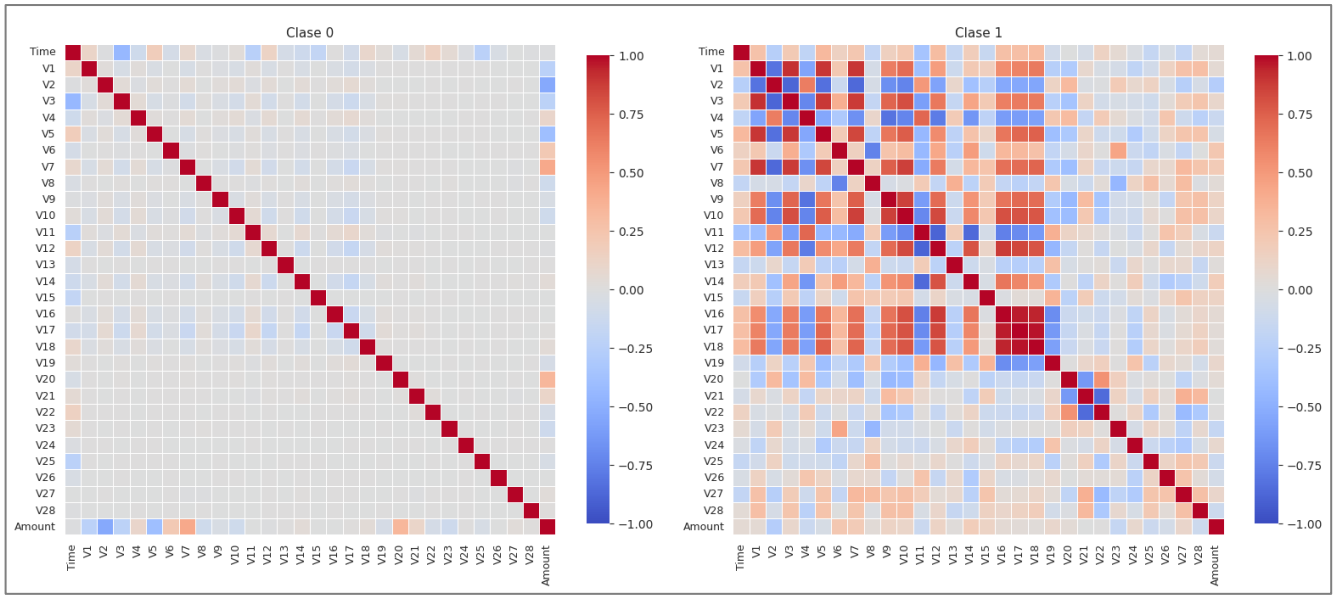
4 DESCRIPCION DETALLADA DE LOS DATOS

Para entrenar y probar los modelos de detección de fraudes, utilizamos un conjunto de datos de transacciones con tarjetas de crédito realizadas en septiembre de 2013 por titulares en Europa. Este conjunto de datos incluye el registro de las transacciones durante dos días, lo que corresponde a 284,807 transacciones de las cuales 492 fueron determinadas como casos de fraude lo que hace que el conjunto de datos sea extremadamente desbalanceado.

El dataset está compuesto únicamente por variables numéricas, que fueron transformadas mediante Análisis de Componentes Principales (PCA) para proteger la confidencialidad de la información original. Las variables **V1, V2, ..., V27, V28** son los componentes principales generados por el algoritmo PCA, mientras que las variables **Time** y **Amount** son las únicas variables no transformadas en la base de datos para un total de 30 variables independientes. La descripción de las variables conocidas se muestra a continuación.

- **Time**: Indica los segundos transcurridos desde la primera transacción registrada en el dataset.
- **Amount**: Es el monto de la transacción registrada. Su unidad es desconocida pero no es necesaria para el análisis.
- **Class**: Es la variable de respuesta que indica si una transacción es clasificada como fraudulenta (1) o no (0).

La visualización de la matriz de correlación para las observaciones clasificadas como no fraudulentas muestra cierto nivel de correlación entre algunos de los componentes principales y las variables **Time** y **Amount**; y no muestra ninguna o una mínima correlación entre los componentes principales. Sin embargo, la matriz de correlación para las transacciones clasificadas como fraudulentas muestra un comportamiento totalmente opuesto, una fuerte correlación entre los componentes principales.



Por otro lado, intervalo entre transacciones es bastante pequeño, inclusive muchos de los registros son simultáneos. El intervalo entre transacciones máximo es de 34 segundos y un promedio de 0.6067 segundos. Respecto al monto de las transacciones, estas están en un rango desde 0 hasta 25,000, con una media de 88 y una mediana de 22, lo que concluye una distribución asimétrica.

En su mayoría, el rango de los componentes principales tiene mínimos en las decenas de valores negativos y máximos en las decenas de valores positivos con medias en cero (0), y medianas que dependen de la presencia de valores atípicos. (Refiérase al Anexo 2)

Al determinar la distribución de cada una de las variables discriminando la clase, es notorio que, para los componentes principales, las transacciones clasificadas como no fraudulentas, Clase 0, siguen, prácticamente, una distribución normal con una dispersión, en su mayoría, baja; mientras que las clasificadas como fraudulentas, muestran una asimetría con una dispersión alta, condición que puede ser una ventaja al momento de la detección de anomalías.

Las asimetrías con alta dispersión podrían deberse a una gama más amplia de comportamientos financieros en las transacciones fraudulentas (Clase 1), mientras que la Clase 0 presenta una distribución más controlada. La similitud en las distribuciones de algunas variables como V13 y V15 entre las dos clases sugiere que no todas las variables serán igualmente útiles para la predicción de fraudes.

Respecto a la variable **Time**, esta no muestra diferencias significativas entre las clases en cuanto a su distribución; ambas clases tienen rangos similares y carecen de valores atípicos prominentes. Sin embargo, al momento de graficar la densidad de probabilidad, se destacan picos a ciertas horas de día, lo que indicaría que las transacciones fraudulentas pueden ocurrir en momentos específicos del día o en ciertos intervalos de tiempo. (Refiérase al Anexo 3)

Respecto al análisis estadístico de los datos de cada variable, de la variable TIME se destaca que las transacciones tienen ocurren principalmente durante el día y entrada la noche, con menos ocurrencia en la madrugada. (Refiérase al Anexo 4)

5 PROPUESTA METODOLÓGICA

La metodología propuesta para acometer el proyecto se resume a continuación:

5.1 Exploración y Preprocesamiento de los Datos:

Mediante un análisis descriptivo de los datos, obtener previamente características que permitan su comprensión y que puedan afectar su modelamiento y los resultados. Por ejemplo, la presencia de datos faltantes, el desbalance de las clases, el rango de los valores, entre otros.

Respecto a preprocesamiento, estandarizar las variables requeridas, determinar el número de componentes de acuerdo con un porcentaje de varianza explicada, eliminar duplicaciones, balancear las clases, entre otros, que puedan aplicar para acondicionar los datos como prerrequisito para la implementación de los modelos.

5.2 Implementación de los Modelos:

Para la detección de anomalías se escogieron preliminarmente, dos algoritmos de aprendizaje no supervisado: ISOLATION FOREST y ONE CLASS SVM.

El primero, se basa en el principio de que las anomalías son pocas y diferenciables, haciéndolas fáciles de identificar y aislar de las observaciones normales. Tiene ventajas al ser computacionalmente eficiente en bases de datos con una gran cantidad de dimensiones al no basarse en métricas de distancias o densidades. Una desventaja, es la selección adecuada del *Parámetro de Contaminación* para un modelo de predicción efectivo.

El segundo, se basa en el aprendizaje de una frontera de decisión hiperesférica que determina como anomalía las observaciones que se ubican fuera de ella. Al igual que el algoritmo ISOLATION FOREST, es eficiente en bases de datos con una gran cantidad de dimensiones, sin embargo, es exigente computacionalmente en bases de datos con numerosas observaciones, además de que su efectividad está supeditada a selección adecuada del kernel y sus parámetros.

5.3 Comparación del Desempeño de los Modelos:

Para la comparación de los modelos, o la determinación de un modelo conjunto, se determinan previamente las métricas adecuadas como *PRECISION, RECALL, o F1-SCORE* para medir el desempeño de estos. El análisis de las métricas junto con la curva ROC y el valor AUC serán usadas para medir el desempeño de los modelos.

5.4 Evaluación de los Modelos:

La evaluación de los modelos se determinará calculando el impacto financiero de predicciones incorrectas como el costo de falsos positivos (transacciones legítimas consideradas como fraude).

5.5 Visualización de las Anomalías:

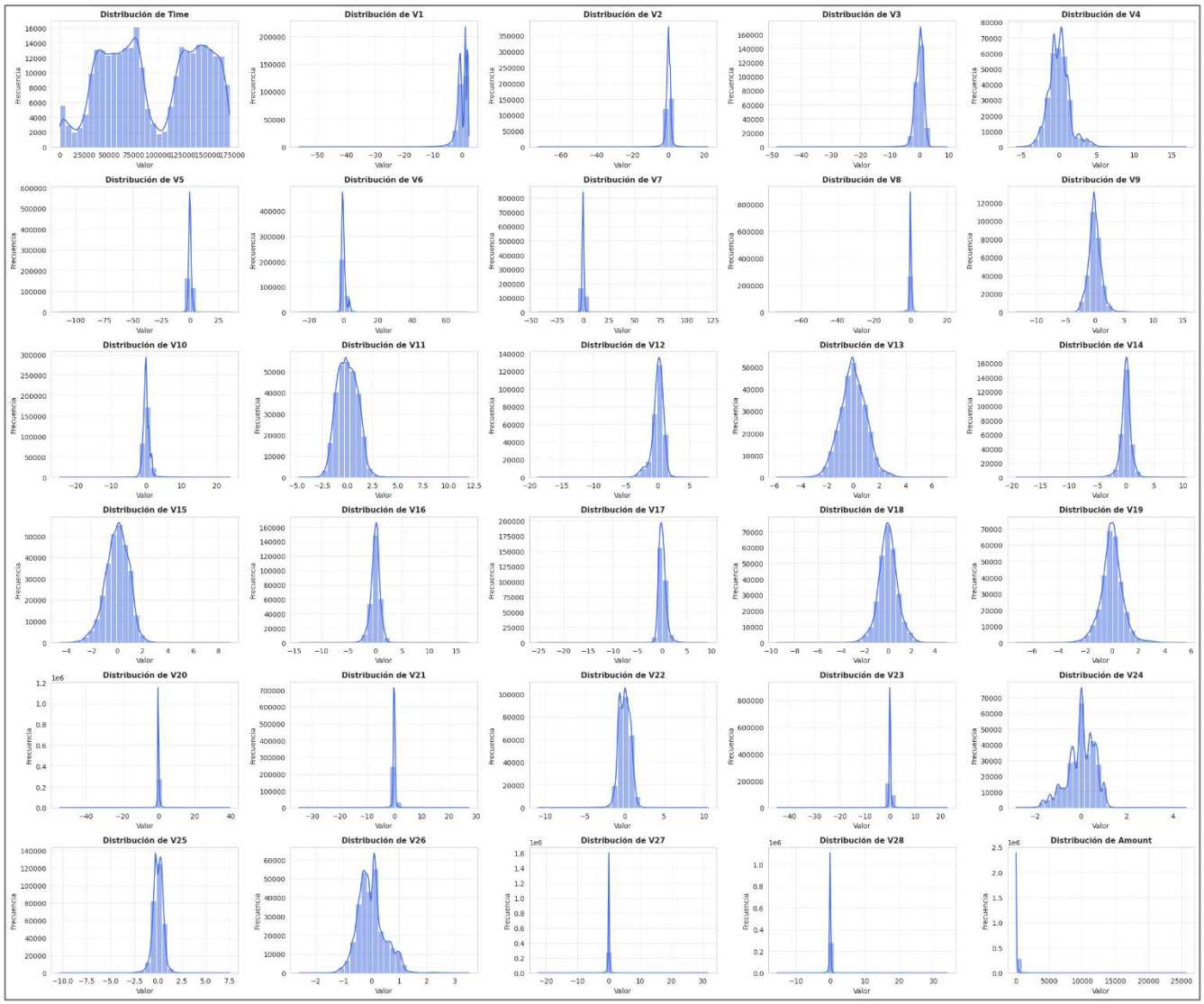
Los resultados se acondicionaron para su visualización apropiada mediante la reducción de las dimensiones, por ejemplo, t-SNE, PCA, etc. en gráficas bidimensionales o tridimensionales.

6 ANEXO 1 - BIBLIOGRAFÍA

La siguiente bibliografía fue consultada para la preparación de la propuesta inicial del proyecto:

- **Alshameri, F., & Xia, R. (September 2024).** An evaluation of variational autoencoder in credit card anomaly detection. *Big Data Mining and Analytics, 7*(99), 1-12.
- **ObaidAli, Ola Imran, & Al-Sultan, Yakoob. (July 2024).** Survey of fraud detection techniques for credit card transactions. *Journal of University of Babylon for Pure and Applied Sciences.*
- **Smairi, Nadia & Abadlia, Houda. (March 2024).** Enhanced particle swarm optimization-based hyperparameter optimized stacked autoencoder for credit card fraud detection. *International Journal of Data Science and Analytics.*
- **Kowsalya, K., Vasumathi, M., & Selvakani, S. (March 2024).** Credit card fraud detection using machine learning algorithms. *EPRA International Journal of Multidisciplinary Research (IJMR).*
- **Bhakta, S. S., Ghosh, S., & Sadhukhan, B. (December 2023)**. Credit card fraud detection using machine learning: A comparative study of ensemble learning algorithms. *In Proceedings of the 9th International Conference on Smart Computing and Communications (ICSCC), Kochi, Kerala, India*.
- **Lavanya, K. (April 24, 2024)**. Isolation Forest for Unsupervised Anomaly Detection. Techdevathe. https://medium.com/techdevathe/isolation-forest-for-unsupervised-anomaly-detection-4d4594e13451
- **SPX. (March 25, 2024)**. Anomaly Detection with Isolation Forest. https://medium.com/@SPX701/anomaly-detection-with-isolation-forests-367e28e6a74e
- **Castagno, P. (February 1, 2024)**. Isolation Forest Concept and Pseudocode. https://patriziacastagnod.medium.com/isolation-forest-if-70e4b6860fde
- **Peters, M. (October 11, 2023)**. Understanding Support Vector Machine (SVM) and One-Class SVM. https://www.geeksforgeeks.org/understanding-one-class-support-vector-machines/
- **Pierobon, G. (August 8, 2023).** One-Class SVM (Support Vector Machine) for Anomaly Detection. https://medium.com/@gabrielpierobon/one-class-svm-support-vector-machine-for-anomaly-detection-a2e00c742ad7
- **Kuo, C., & Dataman, Dr. (Ocotber 9, 2022).** Handbook of Anomaly Detection (6) – One-Class SVM. https://medium.com/dataman-in-ai/handbook-of-anomaly-detection-with-python-outlier-detection-6-ocsvm-f746dae9f450
- **Moser, R. (March 29, 2019).** Fraud Detection with Cost-Sensitive Machine Learning. https://towardsdatascience.com/fraud-detection-with-cost-sensitive-machine-learning-24b8760d35d9
- **Kuo, C., & Dataman, Dr. (August 10, 2018).** Feature Engineering for Credit Card Fraud Detection. Dataman in AI. https://medium.com/dataman-in-ai/how-to-create-good-features-in-fraud-detection-de6562f249ef
- **AltexSoft Co. (January 12, 2021).** Credit Card Fraud Detection: How Machine Learning Can Protect Your Business from Scams. https://www.altexsoft.com/blog/credit-card-fraud-detection/

7 ANEXO 2 – DISTRIBUCION DE LOS DATOS DE LAS VARIABLES.



8 ANEXO 3 – DENSIDAD DE PROBABILIDAD DE LOS DATOS DE LAS VARIABLES.

9 ANEXO 4 – DISTRIBUCION DE LAS TRANSACCIONES EN EL PERIODO.