UNIVERSIDAD DEL VALLE DE GUATEMALA

Facultad de Ingeniería

Trabajo Práctico PlusTI: Detección de fraudes en servicios digitales de alto riesgo

Ingeniería en Ciencias de la Computación y Tecnologías de la Información
Universidad del Valle de Guatemala - UVG
Erick Stiv Junior Guerra Muñoz 21781

Guatemala, junio 2025

Tabla de contenido

Resumen

Este estudio tuvo como objetivo desarrollar e implementar modelos de aprendizaje automático para la detección de transacciones fraudulentas en plataformas digitales de alto riesgo, como criptomonedas, juegos en línea y servicios de apuestas. Utilizando un conjunto de datos transaccionales, se evaluaron tres estrategias de modelado: un modelo base (LightGBM estándar), un modelo ajustado para priorizar la detección de fraudes (Custom1) y otro especializado en fraudes de alto riesgo (Custom2). Los resultados indicaron que el modelo base logró el mejor equilibrio general (AUC-ROC: 0.9967, F1-Score: 0.8080), mientras que el modelo Custom2 redujo ligeramente los falsos positivos sin sacrificar significativamente la capacidad de detección. No obstante, ningún modelo superó el 90% de tasa de detección, lo que evidencia oportunidades de mejora en el análisis aplicado para optimizar el alcance del objetivo planteado inicialmente.

Abstract

This study aimed to develop and implement machine learning models for detecting fraudulent transactions on high-risk digital platforms such as cryptocurrencies, online gaming, and betting services. Using a transactional dataset, three modeling strategies were evaluated: a baseline model (standard LightGBM), a model adjusted to prioritize fraud detection (Custom1), and a model specialized in high-risk fraud (Custom2). Results showed that the baseline model achieved the best overall balance (AUC-ROC: 0.9967, F1-Score: 0.8080), while Custom2 slightly reduced false positives without significantly sacrificing detection performance. However, none of the models surpassed a 90% detection rate, highlighting opportunities for improvement in the applied analysis to better achieve the originally stated objective.

Metodología

En la búsqueda de alcanzar el objetivo planteado, se aplicó un procedimiento estructurado de procesamiento, análisis y predicción de datos a través de distintos modelos. Dicha metodología incluye los siguientes apartados:

- **Análisis exploratorio de datos (EDA):** Se realizó una limpieza exhaustiva de los datos, normalización, categorización y transformación temporal. Además, se aplicaron técnicas avanzadas de ingeniería de variables que permitieron detectar patrones de comportamiento inusual en las transacciones.
- **Balanceo del conjunto de datos:** Dado el desbalance natural entre transacciones legítimas y fraudulentas, se aplicaron técnicas como sobremuestreo y ajuste de umbrales para mejorar la sensibilidad del modelo sin afectar desproporcionadamente la precisión.
- **Entrenamiento de modelos:** Se entrenaron 3 modelos de clasificación supervisada. Las estrategias comparadas fueron:
    - Modelo LightGBM base
    - Modelo priorizando detección sobre precisión
    - Modelo optimizado para fraudes de alto riesgo
- **Evaluación y selección de umbrales:** Se evaluó cada modelo usando métricas como F1-Score, AUC-ROC, matriz de confusión y PR AUC. También se exploraron distintos umbrales de decisión para maximizar la detección manteniendo un ratio de falsos positivos aceptable.

Descripción de la implementación práctica

La solución desarrollada, una vez identificado el mejor modelo de los planteados, está pensada para integrarse a plataformas de servicios digitales:

1. **Modelo para Producción:** Si bien no se logró alcanzar la meta de 90% en la tasa de detección, se recomienda desplegar el modelo *Custom2*, el cual se basa en la detección de fraudes de alto riesgo, ya que obtuvo mejores resultados en la reducción de falsos positivos, a la vez que mantiene una alta detección de fraudes reales identificados.
2. **Monitoreo continuo:** Una vez desplegado el modelo para su funcionamiento en un entorno real, las predicciones realizadas se monitorean para incorporar nuevos fraudes detectados al set de entrenamiento, fortaleciendo el modelo de manera iterativa. A su vez, en caso de que se filtre algún caso de fraude no detectado por el modelo, se recomienda revisar estas transacciones para ajustar características y mejorar el modelo.

Conclusiones

- Se identificó que el modelo base presentó el mejor equilibrio entre precisión y sensibilidad (F1-Score de 0.8080 y AUC-ROC de 0.9967), sin embargo, su tasa de detección no supera el 80%.
- Se obtuvo que, al ajustar el modelo para priorizar la detección, se logró aumentar levemente la cobertura de fraudes, aunque con un incremento en los falsos positivos (de 40 a 47).
- Se concluye que el modelo optimizado para fraudes de alto riesgo es el seleccionado para su despliegue en producción, ya que logró el mejor compromiso al mantener una alta sensibilidad (F1: 0.8080, PR AUC: 0.8471), con una reducción en falsos positivos (39) y una matriz de confusión más balanceada.

Análisis de resultados

Modelo	F1-Score	AUC-ROC	PR AUC	Falsos Positivos	Ratio FP	Recall (detección)
Base	0.8080	0.9967	0.8485	40	1.1980	78.3%
Custom1	0.7726	0.9832	0.7921	47	1.2448	74.42%
Custom2	0.8120	0.9832	0.8471	39	1.1921	78.7%

- El modelo base sigue siendo el más sólido para contextos donde la precisión general es prioritaria.
- El modelo optimizado para fraudes de alto riesgo ofrece el mejor rendimiento si el objetivo es prevenir fraudes severos en plataformas donde el costo de un falso positivo es asumible (como suspensiones temporales automáticas, validaciones manuales, etc.).