UNIVERSIDAD DEL VALLE DE GUATEMALA

Security Data Science
Sección 10
Ing. Jorge Yass

Detección de Fraude en Comercios Nuevos

Sebastián Solorzano

GUATEMALA, 02 de junio de 2025

Resumen

El presente trabajo aborda el problema de la detección de transacciones fraudulentas en comercios recién afiliados. Dado que este tipo de comercios no cuenta con suficiente historial, los modelos tradicionales tienden a tener un bajo rendimiento en este grupo. A través de ingeniería de características, definición de métricas personalizadas y optimización del modelo mediante LightGBM y Optuna, se logró mejorar significativamente el desempeño del sistema en la identificación de fraudes en nuevos comercios. Se evaluaron distintas estrategias, incluyendo métricas penalizadas y enfoques balanceados, y se diseñó una métrica enfocada únicamente en este segmento.

Metodología

El enfoque seguido incluyó las siguientes etapas principales:

1. Análisis Exploratorio de Datos (EDA): Se exploraron las características de las transacciones, distribución de clases y comportamiento de los comercios.
2. Ingeniería de Variables: Se crearon variables específicas para mejorar la capacidad predictiva del modelo, enfocándose en la antigüedad del comercio, tasas históricas de fraude y comportamiento en los primeros días.
3. Modelo Base: Se utilizó LightGBM para entrenar un modelo con métricas estándar como AUC y F1-score, utilizando el mes de diciembre de 2020 como conjunto de prueba.
4. Métricas Personalizadas: Se definieron métricas como false_positive_penalty_ratio y balanced_metric para reflejar la importancia de reducir falsos positivos y balancear recall y precisión.
5. Optimización con Optuna: Se buscaron hiperparámetros óptimos para maximizar las métricas personalizadas. Además, se ajustó el umbral de decisión para mejorar el rendimiento práctico.
6. Evaluación Especializada: Se analizó el rendimiento específicamente en comercios nuevos, definidos como aquellos cuya primera transacción ocurrió en enero de 2019.

Descripción de la Implementación

La implementación técnica se desarrolló en Python utilizando las siguientes tecnologías:

- Pandas, NumPy: para el procesamiento de datos.
- Scikit-learn: para métricas, validación y preprocesamiento.
- LightGBM: como motor principal de modelado.
- Optuna: para búsqueda bayesiana de hiperparámetros.

Variables Derivadas

- merchant_age_days, is_new_merchant, merchant_fraud_rate, entre otras, capturan el contexto histórico de los comercios.
- Variables como std_amt_first_10_days o avg_distancia_first_10_days modelan el comportamiento inicial.

Entrenamiento del Modelo

Se usó LGBMClassifier con validación estratificada. El modelo se entrenó con 1000 iteraciones y early_stopping. La métrica principal fue AUC.

Análisis de Resultados y Comparación de Estrategias

Resultados Generales

El modelo base entrenado con LightGBM sin modificaciones específicas mostró un rendimiento adecuado con un AUC-ROC de 0.8593, un recall para la clase fraude del 65.12% y un F1 Score de 0.7015. Aunque estos valores indican que el modelo es capaz de detectar correctamente una buena proporción de fraudes, el recall todavía es insuficiente para aplicaciones financieras, donde no detectar fraudes puede conllevar pérdidas económicas importantes.

Mejora con Métricas Personalizadas y Optimización

Al introducir métricas personalizadas, diseñadas para penalizar más fuertemente los falsos positivos o balancear mejor precisión y recall, y optimizar hiperparámetros con Optuna, el desempeño del modelo mejoró.

- Modelo Balanceado: alcanzó un recall de 74.03% y una precisión de 86.04%, con un F1 Score de 0.7958. Este aumento en recall representa una mejora crítica en la detección de fraudes, con un nivel aceptable de falsos positivos (31), adecuado para aplicaciones en las que la identificación del fraude tiene prioridad.

- Modelo con Penalización de Falsos Positivos: mostró un desempeño ligeramente inferior en recall (72.87%) y F1 Score (0.7883), pero con una reducción marginal en falsos positivos. Esto puede ser útil en contextos donde el costo de una alerta falsa es alto, aunque se sacrifica un poco de sensibilidad.

Comparación Cuantitativa

MODELO	RECALL	FPR	FP	TP	FN
BALANCED	0.9690	0.0056	1569	907	29
FP PENALTY	0.9690	0.0052	1464	907	29

Ambos modelos optimizados alcanzan un recall muy alto (96.9%) al usar un umbral ajustado (0.01), lo que implica detectar casi todos los fraudes reales. La reducción del falso positivo rate es leve, pero relevante en sistemas con gran volumen de transacciones.

Foco en Comercios Nuevos

El análisis específico en el subconjunto de comercios nuevos, un grupo especialmente difícil por la falta de historial, mostró que el modelo final alcanzó:

- Recall: 70.71%
- Precisión: 97.22%
- F1 Score: 81.87%

Estos resultados indican un muy buen equilibrio, con alta capacidad para detectar fraudes reales (recall), evitando alarmas innecesarias (precision), y un F1 Score que refleja buen desempeño general.

Conclusiones

El desarrollo de este modelo demuestra que es posible mejorar significativamente la detección de fraudes en comercios recién afiliados mediante una combinación de ingeniería de variables, métricas personalizadas y optimización de hiperparámetros. Se logró alcanzar un recall superior al 70% en comercios nuevos, lo cual es notable considerando la escasez de historial disponible.

El modelo balanceado se destaca por ofrecer un excelente equilibrio entre recall y precisión, ideal en escenarios donde es más crítico detectar fraudes que evitar falsas alarmas. Por otro lado, el modelo con penalización de falsos positivos puede ser más útil en contextos donde los costos operativos por alertas innecesarias son altos.

Finalmente, el ajuste del umbral de clasificación se identificó como un factor clave en la mejora del rendimiento, resaltando la importancia de calibrar los modelos para su aplicación específica y no depender únicamente de métricas estándar como la AUC-ROC.