

Universidad del Valle de Guatemala
SECURITY DATA SCIENCE



Detección de Fraude en Transacciones Móviles: Análisis Comparativo de Estrategias de Optimización

Andres Montoya - 21552

Guatemala, 26 de Junio del 2025

Resumen

Este proyecto presenta un sistema de detección de fraude especializado en transacciones móviles utilizando LightGBM con funciones de evaluación personalizadas. Se identificaron 420,304 transacciones móviles basándose en criterios específicos: categorías de compra típicamente móviles (shopping_net, food_dining, entertainment, travel), horarios no comerciales, y distancias significativas entre cliente y comercio (>50km). El modelo final alcanzó un AUC de 0.9814 en validación cruzada, con una precisión del 90% en la detección de fraudes, aunque presentó desafíos en el balance de falsos positivos.

Metodología

Identificación de Transacciones Móviles

Se desarrolló una función específica para identificar transacciones móviles considerando:

- **Categorías móviles:** shopping_net (30.6%), entertainment (28.9%), food_dining (28.1%), travel (12.4%)
- **Patrones temporales:** Transacciones fuera del horario comercial (antes de 9 AM o después de 5 PM)
- **Distancia geográfica:** Separación >50km entre cliente y comercio como indicador de movilidad

Estrategias de Balanceamento

Se evaluaron dos enfoques principales:

1. **SMOTE (Synthetic Minority Over-sampling Technique):** Inicialmente implementado con ratio 0.5, descartado por generar una tasa excesiva de falsos positivos
2. **RandomUnderSampler:** Estrategia final adoptada con ratio 1.0, resultando en 2,754 muestras por clase

Funciones de Evaluación Personalizadas

Se diseñaron tres métricas de evaluación específicas para optimización:

1. Custom False Positive Ratio

python

```
def custom_false_positive_ratio(y_true, y_pred):  
    return (tp + fp) / tp if tp > 0 else 0
```

2. Custom Fraud Detection Metric

python

```
def custom_fraud_detection_metric(y_true, y_pred):  
    fp_penalty = fp / (tn + fp) if (tn + fp) > 0 else 1  
    tp_reward = tp / (tp + fn) if (tp + fn) > 0 else 0  
    return tp_reward * (1 - fp_penalty)
```

3. AUC Estándar: Como métrica de referencia

Descripción de la Implementación Práctica

Ingeniería de Características

Se crearon 15 características adicionales incluyendo:

- Edad del cliente calculada desde fecha de nacimiento
- Tasa de gasto por hora y por categoría
- Distancia promedio por cliente
- Frecuencia de transacciones por categoría
- Velocidad de transacciones y patrones de gasto

Arquitectura del Modelo

LightGBM con parámetros optimizados:

- `num_leaves`: 7 (reducido para controlar sobreajuste)
- `max_depth`: 3
- `learning_rate`: 0.005 (conservador)
- Regularización fuerte: `lambda_l1=2.0`, `lambda_l2=2.0`
- `feature_fraction=0.4`, `bagging_fraction=0.4` (submuestreo agresivo)

Variables Más Importantes

1. `num_amt` (36.5%): Monto de la transacción
2. `num_amt_month_shopping_net_spend` (20.7%): Gasto mensual en compras online
3. `num_hour` (8.5%): Hora de la transacción
4. `cat_category_shopping_net` (4.5%): Categoría shopping_net
5. `num_amt_month` (4.2%): Gasto mensual total

Análisis de Resultados de la Evaluación

Rendimiento del Modelo Final

- **Accuracy:** 97% (consistentemente alto en todas las configuraciones)
- **AUC Cross-Validation:** 0.9884 (± 0.0033)
- **Recall para fraudes:** 90%
- **Precisión para fraudes:** 6%

Matriz de Confusión (Conjunto de Prueba)

- Verdaderos Negativos: 30,194
- Falsos Positivos: 966
- Falsos Negativos: 7
- Verdaderos Positivos: 64
- Tasa de detección: $64 / (64 + 7) = 90.14\%$

Comparativo de Estrategias de Balanceamiento

Estrategia	Ventajas	Desventajas	FP Rate
SMOTE	Preserva información original	Tasa excesiva de FP	>20%
RandomUnderSampler	Control efectivo de FP	Pérdida de información	~3.1%

Análisis de Funciones de Evaluación

1. **False Positive Ratio:** Efectiva para controlar alarmas falsas, pero puede comprometer la detección
2. **Fraud Detection Metric:** Mejor balance entre penalización de FP y recompensa de TP
3. **AUC:** Estable pero no específica para el objetivo de negocio debido a que este modelo prioriza el balance general entre sensibilidad y especificidad sin considerar los costos asimétricos del fraude. En contextos de detección de fraude, un falso negativo (fraude no detectado) tiene un costo significativamente mayor que un falso positivo (transacción legítima bloqueada).

Conclusiones

Hallazgos Principales

1. **Accuracy Persistentemente Alto:** El accuracy del 97% se mantiene constante debido a la naturaleza extremadamente desbalanceada del dataset (0.67% fraudes), donde el modelo puede lograr alta precisión simplemente clasificando la mayoría como legítimos.
2. **Efectividad del Criterio de Distancia:** La inclusión de distancia >50km como criterio para transacciones móviles demostró ser efectiva para capturar patrones de fraude geográficamente anómalos.
3. **Superioridad del RandomUnderSampler:** Aunque SMOTE es teóricamente superior, en este caso específico el undersampling controló mejor los falsos positivos manteniendo 90% de detección.

Estrategia Óptima Identificada

La **Custom Fraud Detection Metric** combinada con **RandomUnderSampler** logró el mejor balance:

- Mantiene 90% de detección de fraudes (objetivo cumplido)
- Reduce falsos positivos al 3.1%
- Balanceo efectivo entre precisión operacional y detección

Limitaciones y Trabajo Futuro

- **Sesgo de submuestreo:** Pérdida de información por undersampling
- **Generalización:** Validar en otros períodos temporales
- **Características temporales:** Explorar patrones estacionales más complejos

Recomendaciones Operacionales

1. Implementar umbrales dinámicos basados en la Custom Fraud Detection Metric
2. Monitoreo continuo de la ratio de falsos positivos
3. Actualización periódica del modelo con nuevos patrones de fraude móvil
4. Considerar ensemble methods para mejorar robustez

El modelo desarrollado cumple con el objetivo de 90% de detección manteniendo una tasa de falsos positivos operacionalmente manejable, representando una solución viable para detección de fraude en transacciones móviles.