

Detección de transacciones fraudulentas en tarjetas de crédito

Bastián Aceitón - Vicente Frías

Universidad Técnica Federico Santa María

Septiembre, 2024

- 1 Descripción del Usuario
- 2 Requerimientos del Negocio
- 3 Requerimientos Funcionales
- 4 Estado de la investigación

- 1 Descripción del Usuario
- 2 Requerimientos del Negocio
- 3 Requerimientos Funcionales
- 4 Estado de la investigación

Usuario

Compañías de tarjetas de crédito

Necesitan detectar transacciones fraudulentas para proteger a sus clientes y minimizar pérdidas.

Contexto

La base de datos tiene las siguientes características:














- Contiene 284.807 transacciones de tarjetas de crédito europeas de septiembre de 2013.
- Incluye información sobre el tiempo desde la primera transacción, el monto en euros, y si la transacción fue fraudulenta o no.

Objetivos

El objetivo principal es minimizar la no detección de fraudes y maximizar la precisión en la detección de estos para evitar bloquear tarjetas por transacciones legítimas.

- 1 Descripción del Usuario
- 2 Requerimientos del Negocio**
- 3 Requerimientos Funcionales
- 4 Estado de la investigación

Customer Journey

Fases	Entendimiento del problema	Acceso y preparación de los datos	Procesamiento de los datos	Selección de modelos y entrenamiento	Evaluación de modelos	Implementación y monitoreo	Mejora continua
Objetivo	Identificar transacciones fraudulentas para prevenir pérdidas financieras	Acceder y preparar el dataset para el análisis	Limpiar y procesar los datos para el posterior entrenamiento de los modelos	Seleccionar los mejores y más eficientes modelos posibles	Evaluar el rendimiento de los modelos mediante métodos cuantitativos, como las métricas	Desplegar el modelo en producción y monitorear su rendimiento	Monitorear y mejorar el modelo continuamente
Actividad	Analizar el dataset para detectar fraudes	Cargar y explorar el dataset, manejar el desbalance de clases	Estandarizar características, manejar desbalance de clases, dividir el dataset	Elegir modelos robustos y entrenarlos con los datos procesados	Usar métricas tales como AUC-ROC	Integrar el modelo en el pipeline de transacciones	Recoger feedback, re-entrenar el modelo de ser necesario, implementar técnicas avanzadas
Sensaciones							
conclusiones							
	Es necesario automatizar la detección de fraudes	Es necesario implementar técnicas para lidiar con el desbalance de clases	Es necesario una preparación adecuada de los datos para mejorar la calidad de los modelos	Es necesario hallar el equilibrio entre buenos modelos predictivos y modelos costo eficientes	Es necesario el uso de distintas métricas para poder elegir el modelo que funcione en el mayor abanico de escenarios	Sistemas de alerta para notificar transacciones potencialmente fraudulentas	Uso de aprendizaje no supervisado para detectar nuevos patrones de fraude

- 1 Descripción del Usuario
- 2 Requerimientos del Negocio
- 3 Requerimientos Funcionales**
- 4 Estado de la investigación

- **Automatización de la detección de fraudes:** Se debe poder identificar fraudes en tiempo real y sin ninguna intervención manual.
- **Preparación y acceso a datos:** Cargar y procesar los datos para poder gestionar los desbalances de clases.
- **Procesamiento de datos:** Se limpian y estandarizan los datos de modo que se asegure un buen entrenamiento para el modelo.

- **Selección de modelos y métricas:** Se entrenan distintos modelos y, según las métricas escogidas, se comparan.
- **Integración en producción:** Para asegurar el funcionamiento en tiempo real, implementar el modelo en el pipeline de transacciones.
- **Monitoreo y actualización:** Monitorear el modelo e ir actualizando según la problemática que se presente.
- **Sistemas de alerta:** Notificar automáticamente al cliente y al equipo de seguridad sobre transacciones sospechosas.

- 1 Descripción del Usuario
- 2 Requerimientos del Negocio
- 3 Requerimientos Funcionales
- 4 Estado de la investigación**

En la detección de fraudes con tarjetas de crédito existen numerosas metodologías y herramientas. Un recurso clave es el *Fraud Detection Handbook*¹, el cual proporciona fundamentos teóricos y herramientas prácticas.

¹Fraud Detection Handbook.

Hallazgos Clave

- **Simulador de datos:** Crea datos basados en las características de las transacciones reales con el fin de testear modelos.
- **Recomendación de métricas:** Se recomienda el uso de la métrica *Área bajo la curva de Precision-Recall* (AUPRC) por sobre la usual *Accuracy*.
- **Técnicas recomendadas:** Se recomiendan modelos como Random Forest, Gradient Boosting y redes neuronales. Además de técnicas como el oversampling o el undersampling para equilibrar las clases y mejorar la detección de fraudes.

Thank You