

DETECCIÓN DE FRAUDE EN TARJETAS DE CRÉDITO

PROYECTO HÍBRIDO KNIME Y PYTHON

* Problema

Las transacciones fraudulentas representan un costo operativo y de reputación enorme. El principal desafío técnico es detectarlas en un mar de datos legítimos, ya que en nuestro dataset solo el 0.17% de las transacciones eran fraudulentas. Un modelo que no aborde este desbalance extremo es ineficaz.

* Solución

01

PROTOTIPADO RÁPIDO EN KNIME:

Se usó para el Análisis Exploratorio de Datos (EDA), preprocessamiento, y la optimización de un modelo base con XGBoost, estableciendo un benchmark sólido de forma rápida y visual.

02

REFINAMIENTO EN PYTHON:

Se migró el flujo a un entorno profesional (VSCode, Jupyter) para aplicar Ingeniería de Características avanzada, creando nuevas "pistas" a partir de las variables de tiempo (transformación cíclica) y monto (indicador de valor cero).



Visualización destacada

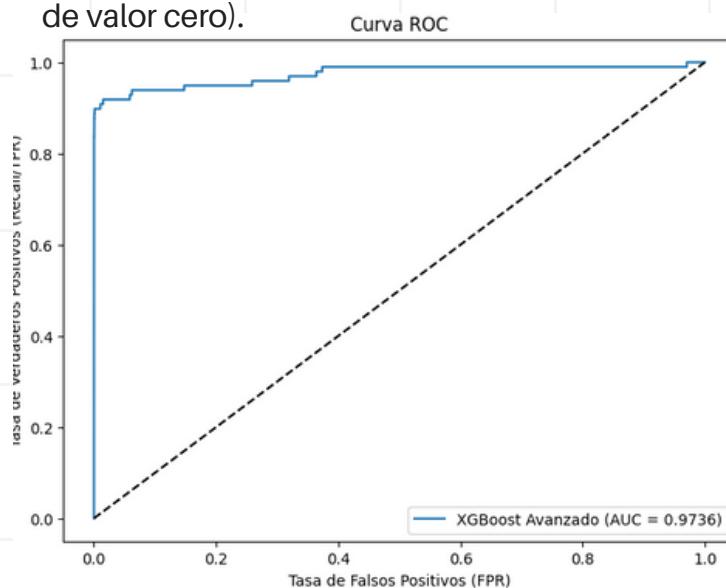
COMPARATIVA DE MODELOS:

La curva ROC compara el rendimiento del modelo base con el modelo avanzado. La curva azul (modelo avanzado) está notablemente por encima, demostrando una capacidad de discriminación superior gracias a la ingeniería de características.



Herramientas utilizadas

- KNIME
- Python
- Pandas
- Scikit-learn & Imbalanced-learn
- XGBoost
- Git & GitHub



github.com/raigosa720/deteccion-fraude-tarjeta-credito

