

Data Scientist Challenge Open Banking



Data Scientist Challenge | Open Banking

Objetivo: Evaluar habilidades en el manejo de datos heterogéneos, análisis exploratorio, modelado y generación de insights en el contexto de **Open Banking**.

Caso Práctico

Identificación de Oportunidades de Cross-Selling en Open Banking

Contexto

Una fintech que opera bajo regulación de Open Banking quiere mejorar su estrategia de **cross-selling** (venta cruzada) de productos financieros (ej: tarjetas de crédito, seguros, inversiones) para sus clientes.

Tienes acceso a datos de:

1. **Transacciones bancarias** (de APIs de Open Banking).
2. **Productos contratados** por los clientes (de la base de datos interna).
3. **Datos demográficos** (provenientes de un CRM).

Problema:

¿Cómo podemos identificar qué clientes tienen mayor probabilidad de adquirir un nuevo producto financiero (ej: un seguro) basado en su comportamiento transaccional y perfil?

Datos Provistos

1. Datos de Transacciones (API Open Banking)
2. Productos Contratados (Base Interna)
3. Datos Demográficos (CRM)

https://drive.google.com/file/d/1jMPH425I_ACudNDSvxXcZM3mcFFtvDVR/view?usp=sharing

Tareas a Realizar

1. Análisis Exploratorio (EDA)

- Identificar patrones de gasto por categoría.
- Relacionar transacciones con productos ya contratados (ej: clientes con tarjeta de crédito gastan más en viajes).
- Analizar distribución de ingresos vs. productos contratados.

2. Feature Engineering

- Crea variables relevantes. Por ejemplo: `avg_monthly_spend` (gasto promedio mensual) o `favorite_category` (categoría más frecuente en transacciones).

3. Modelado Predictivo

- Construir un modelo (ej: **Random Forest / XGBoost**) para predecir la probabilidad de que un cliente contrate un **seguro**.
- Usar técnicas de clasificación y evalúa con métricas relevantes (AUC-ROC, precisión, recall).

4. Recomendaciones de Negocio

- ¿Qué características son más importantes para predecir cross-selling?
- ¿Qué segmentos de clientes deberían ser priorizados en campañas?

Entrega Esperada

- ✓ **Código** (Python/R + SQL si aplica) en Jupyter Notebook o script.
- ✓ **Visualizaciones** claves (ej: heatmap de correlación, importancia de features).
- ✓ **Breve informe** (2-3 páginas) con hallazgos y recomendaciones.

Bonus Points

- Uso de **APIs de Open Banking simuladas** (ej: Mockoon).
- Análisis de **series de tiempo** para predecir cambios en patrones de gasto.
- Diseño de un **dashboard** (Streamlit/Power BI) para el equipo comercial