

Prueba Técnica para Científico de Datos

Anti-Money-Laundering

Instrucciones generales:

Tu objetivo en esta prueba es cubrir las fases de un proyecto de ciencia de datos, sus tareas respectivas y las relaciones entre estas tareas, en este caso orientado a ayudar en la detección de actividades sospechosas.

Asegúrate de documentar y justificar detalladamente cada decisión a lo largo de la prueba. Incluye el proceso completo que desarrolles. Por ejemplo EDA, selección de variables, establecimiento de métricas y cualquier cosa que creas pertinente.

Entregable mínimo: jupyter notebook con el código, explicaciones y conclusiones.

Parte 1: Modelo Supervisado

Datos: Se te proporciona un conjunto de datos centrado en los usuarios. Cada fila representa un usuario con características derivadas de sus transacciones, y una etiqueta que indica si el usuario es sospechoso. El nombre de las columnas es auto-indicativo, por lo tanto explica a qué se refiere cada columna.

Objetivo: Crea y entrena un modelo de regresión que proporcione una probabilidad de que un usuario sea sospechoso.

Parte 2: Modelo No Supervisado

Datos: Se te proporciona el conjunto de datos de transacciones. Cada fila representa una transacción con sus características inherentes. El nombre de las columnas es auto-indicativo.

Objetivo: Crea y entrena un modelo de clustering para estos datos. Interpreta los grupos identificados y propone cómo podrían ser utilizados para detectar transacciones o usuarios sospechosos.