

DESAFÍO DS

Una empresa de telecomunicaciones ha estudiado una pérdida de clientes, por eso se acerca a nuestra empresa para construir un modelo para calcular la probabilidad de abandono (churn) de los clientes y generar un ranking de clientes basado en esta probabilidad. Sin embargo, necesitan que este sea rápido, por lo que se nos ha pedido que no construyamos modelos basados en redes neuronales.

Se entrega el siguiente archivo `data_users.csv` en donde se encuentra el dataset con el que construir la solución.

El objetivo del ejercicio es:

1. Construir un modelo para calcular la probabilidad de abandono (churn) de los clientes y en base a estas, generar un ranking de los clientes
2. Además de las probabilidades, se requiere de un análisis de las métricas de evaluación del modelo utilizadas
3. Describir paso a paso qué se está haciendo y por qué.

Luego del desarrollo, necesitamos que respondas lo siguiente:

1. Le ha entregado los resultados del modelo a la empresa de telecomunicaciones. Ahora, debe explicar cómo esta información puede ser utilizada para una mejor toma de decisiones estratégicas.
2. ¿Cómo se podría complementar la solución para el cliente?
3. ¿Cómo se podría simplificar la tarea o hacerla más eficiente sin sacrificar la calidad del modelo?
4. ¿Cuáles son las limitaciones y riesgos de utilizar un modelo de predicción de churn en un entorno empresarial? ¿Cuáles son las fuentes potenciales de sesgo en los resultados?

Entregables:

1. Entregar las probabilidades y la clasificación de los usuarios del dataset: data_usuarios.csv en formato CSV.
2. Un Jupyter Notebook autocontenido que incluya:
 - Paso a paso de todo el proceso.
 - Preprocesamiento y modelado.
 - Evaluación del modelo con métricas, indicadores o KPIs
 - Respuestas a las preguntas especificadas en el desafío y análisis complementario