



EJERCICIO ML SUPERVISADO:CHURN TELCO



Enunciado ejercicio

El objetivo de este ejercicio es estimar la probabilidad de un cliente a abandonar el servicio de la compañía telefónica, en base a un modelo predictivo de abandono ("churn prediction").

El objetivo de la compañía es basar toda su estrategia de retención de clientes en aquellos clientes que más probablemente abandonen el próximo mes (mes+1). Para desarrollar el modelo, hemos extraído de la base de datos una muestra que incluye clientes que han abandonado y clientes que no han abandonado en un mes especifico (mes 0). Disponemos de sus atributos como cliente en el mes anterior (mes-1): tipo de cliente, características sociodemográficas, uso del servicio, productos contratados, etc.

Se pide completar en Colab el flujo completo Supervised ML: Preprocessing, Train, Validation, Test, Interpretability y Predict.

- Teneis a vuestra disposición en GitHub un template del flujo ML con un checklist de los pasos a completar. https://github.com/isolway/TEMPLATE-Sup-ML. Para utilizarlo debeis hacer git clone en la carpeta del entregable y trabajar sobre este template que incluye 3 notebooks separados (Preprocessing, Model, Predict). Trabajareis el ejercicio en equipos del TFM.
- Teneis en la carpeta el **dataset** para train-test y otro para predict, así como un diccionario de variables del dataset.





EJERCICIO ML SUPERVISADO:CHURN TELCO



Instrucciones de entrega

El ejercicio se completará en Colab y se debe guardar en Google Drive en una carpeta con el nombre del equipo: "DSC{nº edicion} – Entregable SupML - {1ºApellido de cada miembro}"

En la misma carpeta se debe incluir:

- .txt con los nombres completos del equipo
- Los 3 notebooks del template SupML ejecutados (que se vea el print de los resultados de cada paso)
- El modelo resultado guardado como "model.pkl"
- Comentarios en Markdowns (celdas de texto) en el preprocessing y en la parte de intepretabilidad para comentar

Los notebooks deben guardarse con los outputs de las celdas a la vista (prints y plots). También deben poder ejecutarse al completo sin errores dándole al botón "Ejecutar todas las celdas".





EJERCICIO ML SUPERVISADO:CHURN TELCO



Instrucciones de entrega

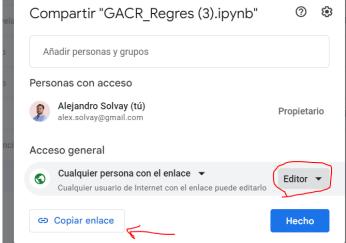
El deadline de la entrega es el 21 de Diciembre 2022.

Para entregar el ejercicio se debe habilitar el acceso compartido, habilitar el rol de Editor y compartirme el enlace por Slack.

No se editará el ejercicio, es simplemente para poder ejecutarlo. Aun así se recomienda guardar una copia del

entregable.







EJERCICIO ML SUPERVISADO: CHURN TELCO



Cómo se corregirá el ejercicio



SE EVALUARÁ...

- **El flujo:** debe ejecutarse correctamente y ser capaz de entrenar el modelo, testearlo y realizar predicciones en nuevos datos.
- Preprocessing: debe ser coherente y <u>argumentarse</u> en markdowns. Tener especial cuidado con las variables de alta cardinalidad.
- Test: la estrategia de partición train-test debe ser valida y las métricas adecuadas.
- **Train:** deben probarse al menos 3 algoritmos ML.
- Explicabilidad/Interpretabilidad: se debe comentar los resultados de las top 5 features.

SE VALORARÁ POSITIVAMENTE



- Código limpio y explicado.
- Uso de funciones vistas en clase sklearn, pandas, etc.
- Análisis exploratorio de los datos
- Feature Engineering: Extracción de la máxima información posible a partir de los literales.
- Hiperparametrización (OJO que puede tardar mucho, limitar a 20 iteraciones como maximo).