

# **BITACORA DE EJECUCIÓN**

**TITULO:** Creación de un modelo predictivo de Machine Learning para predecir categóricamente la demanda de taxis en la ciudad de Nueva York.

**OBJETIVO:** Crear una tabla a pedido para el modelo de Machine Learning y disponibilizar el modelo al cliente utilizando Render.

## **ARCHIVOS:**

- modelo\_demanda.ipynb: Notebook para concatenar los datos de uber y yellow taxis
- modelo\_demanda.parquet: Contiene los datos de uber y yellow taxis.
- main.py: Es el archivo en el que está creada la API para deployar la en render.
- machine\_learning.ipynb: Es el archivo que utilizamos para evaluar el modelo más eficiente.
- tabla\_2024.parquet: Es un archivo parquet que va a ser usado para realizar las predicciones en vertex AI.
- taxi+\_zone\_lookup (1).csv: Es un archivo que contiene las id de las zonas junto con los nombres de las zonas y el borough.

## **PROCEDIMIENTO PARA RENDER**

1. Realizar un buen EDA y ETL para limpiar datos y ver tendencias
2. Unir los dataframes a usar (En este caso el de uber y yellow taxis)
3. Averiguar qué tipo de modelo voy a usar (ejemplo: categorización o regresión)
4. Preparar los datos y las columnas a usar
5. Dividir los datos en entrenamiento y testeo (en este caso utilizamos 80% entrenamiento y 20% testeo)
6. Comenzar a evaluar el rendimiento y optimización de distintos modelos
7. Probar diferentes hiper parámetros para un mejor rendimiento
8. Preparar el archivo main.py para deployar el modelo.
9. Entrar a la web donde esta hosteado el modelo y realizar las predicciones

## **EN VERTEX AI**

1. Realizar un buen EDA y ETL para limpiar datos y ver tendencias
2. Unir los dataframes a usar (En este caso el de uber y yellow taxis)
3. Preparar los datos y las columnas a usar
4. Subir el archivo con los datos a Cloud Storage
5. Pasar los archivos de Cloud Storage a Big Query
6. Cargar el conjunto de datos en Vertex AI
7. Entrenar el modelo
8. Realizar las predicciones