Workshop ML

# Consigna:

**Iniciaremos realizando un breve EDA del dataset:**

1. Importar las librerías necesarias para la realización del desafío (Ej: pandas, numpy, matplotlib, etc) y el dataset propuesto para la ejercitación.

2. Imprimir las 5 primeras filas.

3. Imprimir las 5 últimas filas.

4. Imprimir las dimensiones del dataset.

5. Obtener la totalidad de registros por columnas.

6. Analizar el tipo de dato de cada columna.

7. Obtener el tipo de estructura del dataset. Ej: si es o no un DataFrame.

8. Listar los nombres de las columnas.

**Machine Learning - Modelo de Clasificación:**

1. Separar los features en una variable X y la variable target en y.
2. Separar los datos en Train (70 %) y Test (30 %) con un random\_state=42.
3. Crear un objeto de Árbol de Decisión para un modelo de clasificación, con los siguientes parámetros predefinidos:
   * 1. Nivel de Profundidad = 2
     2. Semilla para la reproducibilidad = 42
4. Entrenar el modelo creado.
5. Realizar la predicción del modelo en Train y Test.
6. Calcular la métrica de Accuracy en Train y Test para la comparación posterior de los modelos.
7. Graficar las variables de importancia del Árbol de Decisión creado (feature\_importances\_).

## Recomendaciones:

Antes de realizar cualquier modelo de clasificación, es muy importante conocer en profundidad a nuestro datasest, es por ello que lo primero que debemos hacer es un EDA para entender a nuestros datos. Luego, es necesario crear el modelo de Árbol de Decisión para la clasificación de la variable target (Survived) haciendo uso de diversas funciones de Machine Learning integradas en la librería “sklearn”. También se solicita agregar una breve interpretación de las conclusiones obtenidas luego de realizar el workshop propuesto.

## Formato:

Notebook desarrollado en Python con extensión **.ipynb**