

## ACTIVIDAD SESIÓN APRENDIZAJE DE MÁQUINA NO SUPERVISADO

La empresa **Drop Music**, una plataforma de streaming musical, ha recolectado información de usuarios en distintos países sobre cómo interactúan con las canciones: cantidad de reproducciones, duración promedio escuchada, saltos, número de artistas reproducidos, entre otros.

Te han contratado para realizar un análisis exploratorio y aplicar **técnicas de aprendizaje no supervisado** con el fin de:

- **Segmentar a los usuarios en grupos de comportamiento similar.**
- **Reducir la dimensionalidad del dataset** para visualizar mejor los patrones.
- **Detectar posibles comportamientos anómalos.**

### INSTRUCCIONES:

#### 1. Carga y exploración del dataset (1 punto)

- Descarga y carga el archivo usuarios\_musica.csv.
- Muestra las primeras 5 filas y el resumen estadístico de los datos.

#### 2. Limpieza y preprocesamiento (1 punto)

- Revisa valores faltantes o atípicos y aplica imputación o eliminación según sea conveniente.
- Normaliza los datos numéricos con StandardScaler.

#### 3. Aplicación de K-Means para clusterización (3 puntos)

- Aplica K-Means con un valor  $k=3$ .
- Agrega la columna "cluster" al DataFrame con el resultado.
- Visualiza los clusters usando una gráfica 2D (usa PCA para reducir dimensiones si es necesario).

#### 4. Reducción de dimensionalidad con PCA (2 puntos)

- Aplica **PCA** para reducir las dimensiones a 2 componentes principales.
- Interpreta brevemente qué representa cada componente principal.

- Grafica los puntos proyectados en 2D, coloreados por su cluster.

#### 5. Detección de anomalías (2 puntos)

- Calcula la **distancia euclidiana** entre cada punto y el centro de su cluster.
- Marca como *anomalías* a los 5 usuarios más alejados de sus respectivos centroides.
- Agrega una columna "anomalía" con valores True/False.

#### 6. Reflexión final (1 punto)

- ¿Qué hallazgos consideras más útiles para el equipo de marketing?
- ¿Cómo podrían usar esta segmentación en campañas personalizadas?

#### INSTRUCCIONES ADICIONALES:

- Puntos totales = 10 puntos.
- Comprimir el archivo en formato .zip o .rar.
- Subir el archivo a la plataforma.
- Descargar material complementario usuarios\_musica.csv