

EVALUACIÓN FINAL: ANÁLISIS DE PREFERENCIAS MUSICALES A NIVEL GLOBAL

Trabajas en un centro de investigación de tendencias culturales y te han encargado un análisis sobre las preferencias musicales en distintos países. Tu tarea es aplicar técnicas de **aprendizaje de máquina no supervisado** para extraer patrones de escucha y agrupar países según sus similitudes en consumo musical.

OBJETIVOS DE LA EVALUACIÓN

- Aplicar **clusterización** para encontrar grupos de países con patrones de escucha similares.
- Utilizar **reducción de dimensionalidad** para visualizar la estructura de los datos.
- Evaluar diferentes algoritmos y justificar sus ventajas y desventajas.
- Interpretar los resultados en un contexto cultural y social.

INSTRUCCIONES

1. Carga y exploración de datos (1 punto)

- Carga el dataset proporcionado, que contiene información sobre la popularidad de distintos géneros musicales en países como Chile, EE.UU., México, Corea, Japón, Alemania, Rusia e Italia.
- Analiza las características del dataset, identificando distribuciones y tendencias iniciales.

2. Aplicación de algoritmos de clusterización (5 puntos)

K-Means:

- Aplica el algoritmo K-Means con un valor inicial de $K=3$.
- Determina el valor óptimo de K utilizando el **método del codo** y el **coeficiente de silueta**.

Clustering jerárquico:

- Genera un **dendrograma** y determina el número óptimo de clusters.
- Aplica clustering jerárquico y compara con los resultados de K-Means.

DBSCAN:

- Aplica DBSCAN con diferentes valores de **eps** y **MinPts**.
- Justifica la elección de los parámetros y analiza si DBSCAN identifica agrupaciones significativas.

3. Aplicación de reducción de dimensionalidad (3 puntos)

PCA:

- Aplica PCA y determina cuántos componentes principales explican al menos el **90% de la varianza**.
- Visualiza los países en un gráfico bidimensional con las primeras dos componentes principales.

t-SNE:

- Aplica t-SNE para visualizar la relación entre los países en un espacio de 2D.
- Experimenta con diferentes valores de **perplexity** y analiza cómo afectan la representación.

4. Análisis de resultados y conclusiones (1 punto)

Comparación de métodos:

- Explica las diferencias entre K-Means, clustering jerárquico y DBSCAN. ¿Cuál funcionó mejor en este caso y por qué?
- Compara los resultados obtenidos con PCA y t-SNE. ¿Cuál técnica permitió visualizar mejor la relación entre los países?

Interpretación:

- ¿Los clusters obtenidos reflejan similitudes culturales o geográficas en la música?
- Relaciona los resultados con tendencias globales en consumo musical.



INSTRUCCIONES ADICIONALES:

- **Puntos totales:** 10 puntos.
- Descarga el material complementario que se encuentra disponible en la plataforma.
- Comprime el archivo en formato .zip o .rar.
- Incluye un documento de texto con las reflexiones analíticas.
- Sube el archivo a la plataforma.