

# From Music to Wherever 2025

**Realizado por:**

Marina Xiuping Garrido Castaño

Proyecto Final del Bootcamp de Data Science  
y Machine Learning



# Índice

## O3 **Introducción**

Contextualización

## O4 **Desarrollo del proyecto**

Dataset y Análisis exploratorio.

Preprocesado

Enfoque de Machine Learning

## O7 **Evaluación y Mood Mapping**

## O9 **Recomendador de canciones y Aplicación**

## 10 **Estado actual del proyecto y Conclusión**

Conclusión general

Futuro

Bibliografía

# INTRODUCCIÓN

## Contextualización

¿Cuántas cosas en nuestro día a día están condicionadas por nuestro estado de ánimo? ¿Influye este en decisiones cotidianas?

La música es un gran medio para la expresión emocional. Sin embargo, muchos sistemas de recomendación dependen de metadatos, etiquetas de género o comportamientos del usuario lo que introduce sesgos y limita la exploración sonora.

En este proyecto se propone un enfoque alternativo basado en el análisis directo del sonido mediante *audio-features* y técnicas de *Machine Learning no supervisado*, de manera que es el modelo quien descubre patrones acústicos sin la imposición de etiquetas previas.

El **objetivo principal** es construir un sistema capaz de representar canciones en un espacio acústico latente, agruparlas según su similitud y utilizarlas para la recomendación musical entre sus muchas aplicaciones. A parte, se plantea una capa estética que traduce el *mood* detectado a una propuesta visual, concebida explícitamente como prueba de concepto.

# DESARROLLO DEL PROYECTO

## **Dataset y análisis exploratorio**

Para poder realizar este proyecto, partimos de un dataset compuesto por *audio-features* continuas extraídas de Spotify. La idea inicia era obtenerlas mediante la API de Spotify sin embargo, se carecía de las credenciales necesarias. A pesar de esto debo señalar que existe la posibilidad de añadirlo en el futuro. De aquí extraemos una serie de variables como *energy*, *valence*, *tempo*, *danceability*, *acousticness*, *loudness* y *speechiness*, las cuales describen las propiedades acústicas del sonido.

No se emplean letras, géneros musicales ni métricas de popularidad. Esta decisión reduce sesgos semánticos y permite que el modelo se centre exclusivamente en el sonido, independientemente del idioma o el contexto cultural.

Al no contar con la API de Spotify se optó por trabajar con datasets públicos obtenidos de esta misma API y unificarlos en un mismo dataset con las columnas que nos interesaban.

Durante el análisis exploratorio se observaron distribuciones heterogéneas y la presencia de valores extremos, especialmente en variables relacionadas con intensidad y velocidad, lo cual influyó directamente en las decisiones de preprocesado.

---

## Preprocesado

Una vez obtenido un dataset con más de 165000 filas pasamos a su limpieza. Se eliminaron valores nulos y duplicados, y se seleccionaron las columnas relevantes. Para el escalado se empleó **RobustScalers**, ya que es menos sensible a outliers al basarse en la mediana y el rango intercuartílico, lo que nos resuelve el problema detectado en el análisis exploratorio en relación a valores extremos.

# Enfoque de Machine Learning

El problema se aborda como un problema **no supervisado**, ya que no existen etiquetas objetivas de mood musical. El pipeline consta de tres etapas principales:

- O1 **Escalado de audio-features.**
- O2 **Reducción de dimensionalidad mediante UMAP.**

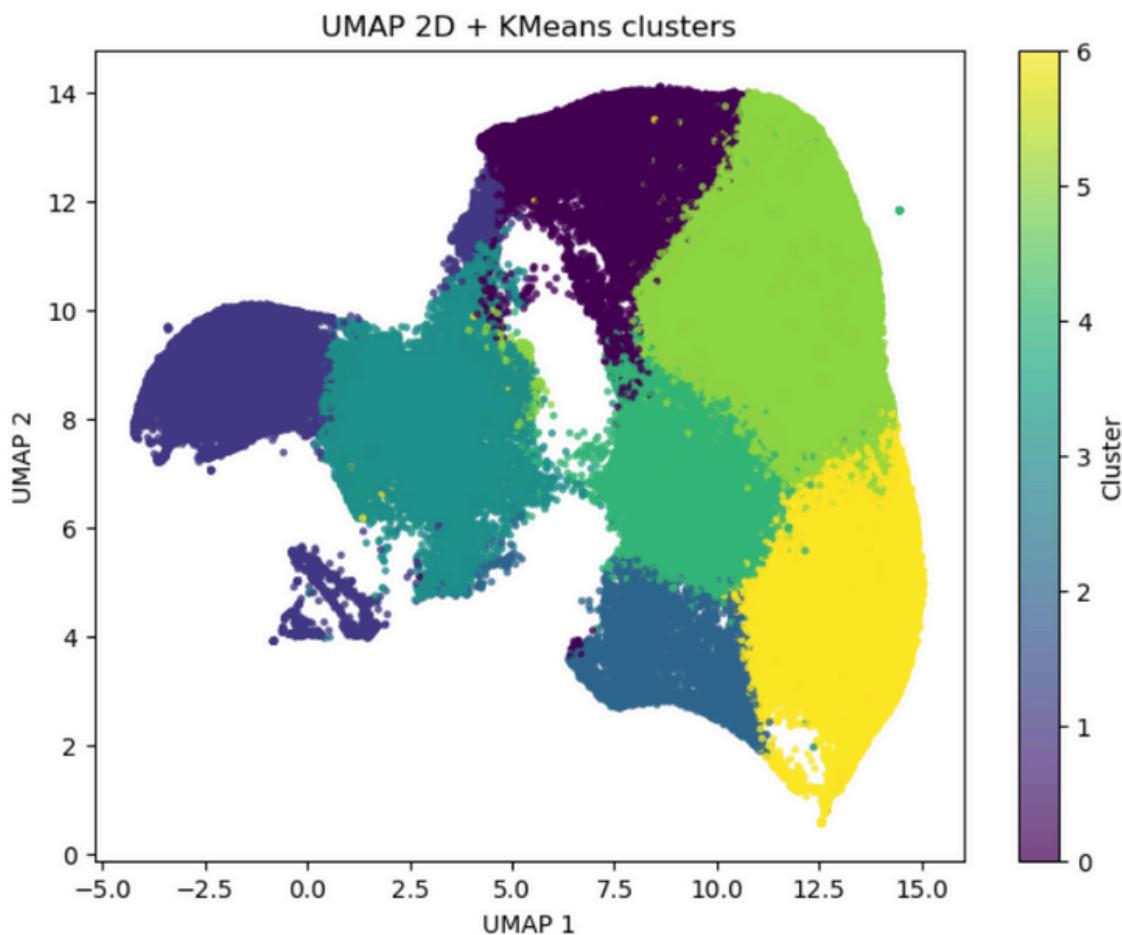
UMAP se usa para obtener una representación latente no lineal del sonido. El embedding se entrena en 9 dimensiones al contar con 9 audio-features, y se proyecta a 2D únicamente para visualizarlo.
- O3 **Clustering con KMeans**

Sobre este espacio se entrena KMeans con  $k=7$ , valor elegido tras evaluar distintas configuraciones y priorizar la coherencia interpretativa de los clusters frente a la maximización de métricas.

# EVALUACIÓN Y MOOD MAPPING

La evaluación del clustering se realizó mediante **Silhouette Score** tanto en train como en test con el objetivo de analizar la estabilidad estructural del modelo.

Dado el carácter creativo y emocional del problema, no se esperan métricas elevadas por lo que la evaluación cuantitativa se complementa con análisis cualitativo.



Una vez obtenidos los clusters, se llevó a cabo un proceso de ***mood mapping*** mediante el análisis de centroides y ejemplos representativos de canciones.

Este proceso permitió asignar descriptores emocionales. Sin embargo debemos destacar que se trata de un proceso interpretativo que se realiza a posteriori y no forma parte del entrenamiento del modelo.

Cluster	Mood	Notes
0		Concierto / Euforia Energía muy alta, loudness alto, speechiness alta y acousticness baja
1		Intenso / Dramático Energía alta, loudness alto, valence baja, speechiness alta
2		Instrumental / Concentración Acousticness alta, instrumentalness alta, speechiness baja, energía media-baja
3		Chill Groove / Urbano Energía media, acousticness media, danceability media, instrumentalness media-alta
4		Happy / Buen Rollo Danceability alta, valence muy alta, loudness alto, speechiness baja
5		Fiesta / Subidón Tempo alto, loudness alto, energía media-alta, acousticness baja
6		Triste / Melancólico Valence muy baja, energía baja, acousticness alta, tempo bajo

# RECOMENDADOR DE CANCIONES Y APLICACIÓN

El sistema permite recomendar canciones similares utilizando la distancia euclídea entre embeddings UMAP, lo que posibilita la exploración musical basada exclusivamente en similitud acústica.

El resultado se implementa en una aplicación interactiva desarrollada con **Streamlit** que permite buscar canciones, predecir el mood correspondiente, modificar los valores de audio.features y obtener recomendaciones musicales.

Como extensión exploratoria, se incluye una capa estética que traduce el mood a paletas de color y propuestas de outfit. Esta capa se plantea explícitamente como una prueba de concepto y no como un sistema validado con usuarios.

# ESTADO ACTUAL DEL PROYECTO Y CONCLUSIONES

## Conclusión general

El núcleo técnico del proyecto centrado en la clasificación de las canciones por audio-features se encuentra validado. La capa estética se concibe como una línea de evolución futura que demuestra la extensibilidad del sistema.

Este proyecto demuestra cómo es posible estructurar el sonido musical mediante técnicas no supervisadas y extraer significados emocional sin depender de etiquetas previas, abriendo la puerta a sistemas de recomendación más exploratorios y menos condicionados por metadatos tradicionales.

## Futuro

Como líneas futuras se plantea la validación con usuarios reales, la personalización basada en feedback, la integración con plataformas de streaming y la exploración de modelos multimodales y contextuales.

## Bibliografía

- McInnes, L. et al. *UMAP: Uniform Manifold Approximation and Projection.*
- Scikit-learn Documentation