

UNIVERSIDAD MAYOR DE SAN ANDRÉS
FACULTAD DE CIENCIAS PURAS Y NATURALES
CARRERA DE INFORMÁTICA



Objetivo de Investigación

Top 200 Spotify Songs Dataset

Responsable: Badani Dávalos Kassandra Andrea

Asignatura: Inteligencia Artificial (INF-354)

Fecha: 9 de diciembre

La Paz – Bolivia

2024

Objetivo de Investigación a partir del Dataset Elegido

El objetivo principal de esta investigación es desarrollar un **modelo predictivo para estimar la popularidad de las canciones en Spotify**, utilizando el conjunto de datos **Top 200 Spotify Songs Dataset**. A través del entrenamiento de un modelo basado en técnicas de **inteligencia artificial**, este aprenderá a predecir la **probabilidad de que una canción se convierta en popular** en función de una combinación de características **musicales, demográficas y contextuales**.

Para medir la popularidad de las canciones, se utilizará el atributo **Points (Total)**, que refleja el desempeño de cada canción en las listas diarias de Spotify, y es considerado un indicador de su popularidad. Este atributo captura la cantidad de puntos acumulados por las canciones, lo que está directamente relacionado con la frecuencia y el impacto de su reproducción, compartición y escucha en la plataforma. Así, el modelo estará diseñado para clasificar las canciones en **populares** y **no populares**, basándose en su **puntuación total**, identificando aquellas que tienen una mayor probabilidad de generar un alto nivel de interacción y, por lo tanto, éxito en la plataforma.

El conjunto de datos seleccionado contiene una amplia gama de atributos que se consideran cruciales para la clasificación de canciones en función de su éxito. Estos atributos incluyen características musicales, como la **danceability, energy, loudness, y valence**, así como detalles demográficos y de contexto, tales como el **número de nacionalidades** representadas en las canciones y la **región geográfica** de los artistas.

A través de este análisis, se busca identificar patrones y factores clave que influyen en el éxito de una canción, proporcionando una comprensión más profunda de qué atributos contribuyen significativamente a la popularidad en Spotify. Con esta información, el modelo permitirá **anticipar el potencial de popularidad de nuevas canciones** a partir de sus características, lo que tiene aplicaciones prácticas en la industria musical, como la mejora de las recomendaciones en plataformas de streaming o el pronóstico del éxito de lanzamientos musicales.

El dataset seleccionado para este análisis es accesible en Kaggle, a través del siguiente enlace:

[Top 200 Spotify Songs Dataset](https://www.kaggle.com/datasets/brunoalarcon123/top-200-spotify-songs-dataset):

<https://www.kaggle.com/datasets/brunoalarcon123/top-200-spotify-songs-dataset>