

Machine Learning aplicado a la música: Aplicación de un modelo para predicción de género

Introducción

La música es, probablemente, la expresión artística más extensa, diversificada y consumida de todas. Y es también, por lógica, una industria gigantesca. Las formas y patrones de consumo han ido variando a lo largo de la historia: desde las percusiones tribales de los primeros homínidos hasta el día de hoy, donde escuchamos música con cualquier dispositivo electrónico mediante aplicaciones inteligentes como Spotify y Tidal, el ser humano y su música han recorrido un largo camino.

Motivación

El objetivo que este trabajo se propone es predecir el género musical de una canción, entre 11 clases diferentes, a partir de una serie intrínseca de características musicales que dicha canción posee. Lo que se busca es poder predecir cuales de esas variables y de qué forma me pueden determinar un género musical en particular.

Audiencia

El modelo se piensa como una herramienta de utilidad en diferentes ámbitos vinculados a la industria musical. Particularmente, se propone servir como parte dentro de un proceso de generación de *playlists* en aplicaciones musicales, de manera tal que cuando un artista o una distribuidora de música lance y publique una producción en alguna aplicación, luego la misma pueda ser catalogada e incluida en diferentes *playlists* de manera asertiva y coherente, dado que dicho sistema de múltiples *playlists* se constituye hoy en día como la principal forma en la que se reproduce (y promociona, claro) la música. Así mismo, puede ser de utilidad en la producción y lanzamiento de nuevas canciones por parte de discográficas y artistas independientes.

Metadata

El dataset elegido es el siguiente:

<https://www.kaggle.com/datasets/purumalgi/music-genre-classification>

Contiene 3 archivos: train.csv, test.csv y submission.csv. El archivo de entrenamiento y ajuste del modelo, es decir train.csv, posee 17.996 registros y 17 columnas. Los registros poseen información sobre el nombre de artista y de la canción y una serie variables inherentes a sus características y, finalmente, la variable “Class” que será nuestra variable *target*.

Los tipos de variables son los siguientes:

```
dtypes: float64(12), int64(3), object(2)
```

Preguntas de interés

- ¿Existe relación entre el género musical de una canción y sus diferentes características musicales?
- ¿Hay géneros musicales más populares que otros? ¿Cuáles serían?
- ¿Qué características comparten entre sí los diferentes géneros?

- ¿Cómo incide cada variable en su género? ¿Con qué intensidad? ¿En qué sentido?
- ¿Se pueden determinar patrones?