

L'ADN dels èxits musicals: Anàlisi i recomanació basada en patrons

EVOLUCIÓ TFG

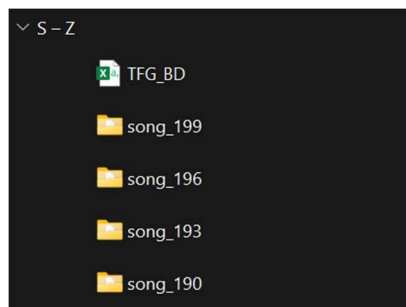
Seguint la cronologia de l'informe inicial, us comento com va evolucionant el projecte:

- SETMANA (17/03)

BASE DE DADES

Aquesta setmana vaig proposar acabar la base de dades i processar les dades per entrenar els models de clustering. Finalment, la base de dades consta de les 200 cançons extretes del top 200 Global de la llista de cap d'any de Billboard. He triat aquesta llista, perquè reuneix les cançons més populars del 2024 de forma global.

Tinc un directori amb diferents subcarpetes que representen les cançons, dins aquestes subcarpetes es troba la cançó original i el processament de cada cançó.



A més, també tinc un fitxer amb les metadades de les cançons, que em poden ajudar en el sistema de recomanació. Aquest fitxer conté característiques de les cançons com: l'autor, el nom de la cançó, l'identificador, els gèneres, i la duració.

PROCESSAMENT DE LES DADES

El processament de les cançons, el continuo investigant, perquè hi ha diferents mètodes per triar. He començat segmentant les cançons en fragments de 12 segons, vaig intentar fer-ho detectant els canvis de ritme de la cançó, però hi havia cançons que generaven molts intervals, i no em donava els resultats que volia.

Un cop fragmentades les cançons, vaig investigar sobre l'extracció de característiques d'àudio per poder calcular la similitud entre cançons. De moment, he triat algunes mètriques que he vist que s'utilitzen, per fer un primer model d'agrupació.

1- AUDIO FEATURES

- **Zero-crossing rate (ZCR):**

És una eina d'anàlisi en el domini temporal que ens ajuda a calcular la freqüència d'àudio. Mesura la quantitat de vegades que un senyal d'àudio creua l'eix zero, mesura quan el senyal canvia de positiu a negatiu i al contrari.

- **Spectral centroid:**

Aquesta mètrica ens indica on se situa el "centre de la massa" d'un so i es calcula com la mitjana ponderada de les freqüències presents en el so. En una cançó on tenim freqüències altes al final, el centroide estarà cap al final, en canvi, en una cançó on tenim freqüències iguals al llarg de tot l'espectre, el centroide es situarà en el centre.

- **Spectral bandwidth:**

Mesura l'amplada de banda espectral d'ordre p, que és una mètrica de dispersió de l'espectre respecte al centroide espectral. En altres paraules, serveix per determinar el rang de freqüències d'un senyal.

- **Spectral Rolloff:**

Ens diu la freqüència per sota de la qual es troba un percentatge especificat de l'energia espectral total.

- **Mel-Frequency Cepstral Coefficients**

Aquesta mètrica és molt important, i s'utilitza molt en projectes on es tracta amb senyals d'àudio. Tracta sobre un conjunt de característiques (entre 10-20) que descriuen de forma concisa la forma general de l'espectre.

- **Chroma STFT**

Mostra com varia l'energia de les 12 notes musicals al llarg del temps, sense importar l'octava en la qual es troben. És útil per detectar acords, i identificar la tonalitat de la cançó.

- **Root Mean Square energy (RMS):**

Mesura l'energia de la cançó.

2- MIDI FEATURES

Després de trobar aquestes característiques d'àudio. Vaig parlar amb la Lucía Carmona, que s'encarrega de generar les cançons d'èxit, i em va comentar que ella necessitaria arxius MIDI per poder treballar. Són uns arxius on no s'emmagatzema àudio, però Sí instruccions musicals de com tocar la cançó. Després d'investigar vaig veure que podia extreure característiques com:

- La distribució de notes
- El tempo i els beats
- Els instruments i la intensitat (velocitat)

Per tant, he convertit cada fragment d'àudio a format MIDI amb una llibreria de Python. I he programat un codi per generar un vector de característiques on combino les característiques de l'àudio i les de l'arxiu MIDI.

CLUSTERING

Un cop tenia les dades processades, necessitava escollir un mètode de clustering que s'adaptés al meu projecte. Vaig trobar-ne un que complia amb les meves necessitats. DBSCAN és un mètode d'aprenentatge no supervisat que ens ofereix agrupar dades sense identificar el nombre de grups prèviament. S'especifiquen dos paràmetres: l'èpsilon que determina la distància màxima entre dos punts perquè es considerin veïns, i el nombre de punts necessaris per formar una regió densa.

1 – PUNT DE PARTIDA

L'algoritme comença en un punt arbitrari no visitat del conjunt de dades i recupera tots els punts dins de la distància èpsilon del punt inicial. Quan el nombre de punts és inferior a l'especificat, el punt s'etiqueta com a soroll, si és més gran, aleshores s'etiqueta com a nucli.

2 – EXPANSIÓ DEL CLUSTER

Tots els veïns del punt central s'afegeixen al clúster, i per cada veí, si està etiquetat com a punt central, els seus veïns s'afegeixen al grup de forma recursiva. Quan no és un punt central, es marca com un punt de frontera i es para l'expansió.

3 – REPETICIÓ DEL PROCÉS

L'algoritme es desplaça al següent punt no visitat del conjunt de dades, i es repeteixen els passos 1 i 2 fins a acabar de visitar tots els punts.

RESULTATS | PLANTEJAMENT ACTUAL

He fet un parell de proves amb el DBSCAN, detecta molt soroll, però hi ha algunes cançons que si detecta com a similars, i ho fa bé, per exemple ha detectat cançons “Rap” com a similars.

He pensat a investigar més a com realitzar l'extracció de característiques i com integrar-ho en el mètode DBSCAN.

En el cas que no vegi bons resultats, iniciaré un segon mètode on intentaria segmentar les cançons per canvis de ritme (ONSET DETECTION). I a més, podria provar de separar l'àudio en stems (ÀUDIO, PIANO, BASE, PERCUSSIÓ..), per detectar similituds més concretes.

Realitzant aquesta investigació i desenvolupant el projecte, m'han sorgit alguns dubtes:

Per una banda, no entenc molt bé com validar el clustering perquè jo puc mirar els clústers i escoltar les cançons per veure si s'assemblen o no, però no sé com demostrar-ho amb resultats.

Per altra banda, mirant cap a l'experiment que vull fer amb EEG, no se molt bé com funciona i que hauria de fer escoltar a la gent que porti per fer l'experiment. També dir, que la setmana del (7/04) no crec que pugui començar amb l'EEG per exàmens, i sobretot pels resultats del clustering que encara no tinc, ho allargaria una setmana.