Analysing Music and Acoustics through the Lens of Machine Learning

ABSTRACT

Recommendation systems leverage several types of information relating to a recommendable item. The recommendation methods are often based on the analysis of how a set of users associate or rate a given set of items, but they can also focus on the analysis of how the content of the items is Related. In very recent years, the development of a music recommendation system has been a more heated problem due to a higher level of digital song consumption and the advancement of machine learning techniques. Some traditional approaches such as collaborator filtering have been widely used in recommendation systems, and have helped the music recommendation system to give music listeners quick access to the music. However, collaborative filtering or model-based algorithms have limitations in giving a better result with the ignorance of the combination factor of acoustic cues and genre. Therefore, in our project we propose an improved algorithm for music recommendation based on studying a song's acoustic features as valence, tempo, energy, beats per minute, etc. and using metrics like t-SNE, HAC, and Cosine similarity to create a better recommendation system, and allowing us to gain insights and be able to do a Time-Series analysis of the songs, and present them as a dashboard for the user. Moreover, we also aim to cluster various songs and genres in an unsupervised manner.

Keywords: music recommendation; music information retrieval; time-series analysis;

INTRODUCTION

Fostered by the advancement of digital technologies, both catalogs of music distributors and personal music collections have grown to sizes that call for automated methods to manage them. Classification algorithms, for instance, help grouping music according to a given taxonomy. Here we are using a song's features collected in a dataset to create an algorithm that will help to recommend song artists in a much better and advanced way as compared to the present metrics used in industry. Not only this but we also are using this data to analyse and present a time series forecast to predict and analyse the trends in song to find out the next major genre in the coming time.

With the explosion of the internet in the past decades, the internet has become the major source of retrieving multimedia information such as video, books, and music etc. People have considered that music is an important aspect of their lives and they listen to music, an activity they engage in frequently. Previous research has also indicated that participants listened to music more often than any of the other activities.

However, the problem now is to organise and manage the millions of music titles produced by society. Music Recommendation techniques have been developed to solve problems such as genre classification, artist identification, and instrument recognition Additionally, music recommender is to help users filter and discover songs according to their tastes.

A good music recommender system should be able to automatically detect preferences and generate playlists accordingly. Meanwhile, the development of recommender systems provides a great opportunity for industry to aggregate the users who are interested in music. More importantly, it raises challenges for us to better understand and model users' preferences in Music Currently, based on users' listening behaviour and historical ratings, collaborative filtering algorithm has been found to perform well. Combined with the use of a content-based model, the user can get a list of similar songs by low level acoustic features such as rhythm, pitch or high-level features like genre, instrument etc.

Music is subjective and universal. It not only can convey emotion, but also can it modulate a listener's mood. The tastes in music are varied from person to person, therefore, the previous approaches cannot always meet the users' needs. An emotion-based model and a context-based model have been proposed.The latter one collects other contextual information such as comments, music review, or social tags to generate the playlist.

Though hybrid music recommender systems would outperform the conventional models, the development is still at a very early stage. Due to recent studies in psychology, signal processing, machine learning and musicology, there is much room for future extension.

Moreover, music based computational and statistical modelling is severely lacking in current research. Music and acoustic cues are an important data resource which can be studied to analyse how music changes over time and how we can improve content based recommendation systems.

➔ **Limitations**

  Though it is fast and accurate, the drawbacks are obvious. First of all, the user has to know about the editorial information for a particular music item. Secondly, it is also time consuming to maintain the increasing metadata.

  Moreover, the recommendation results are relatively poor, since it can only recommend music based on editorial metadata and none of the users' information has been considered.

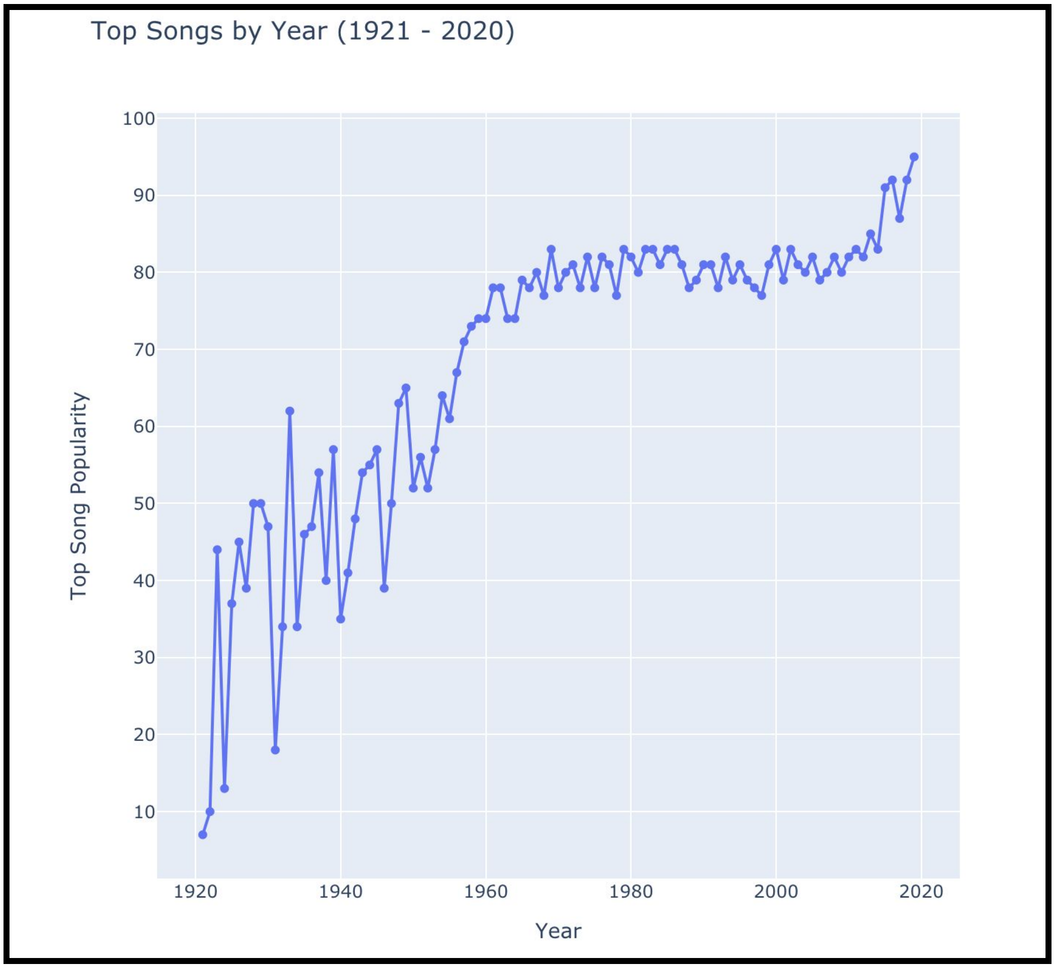
OBJECTIVE

This necessitates a thorough analysis of music data and study how it changes over time. Previous works have lacked a complete overview and deep dive of musical data and only focusing on a tunnel vision based analysis. This makes it imperative to build a complete framework that incorporates all the aspects of musical data analysis and statistical tools and overcomes previous research works.

This framework comprises 3 parts.In this project, we aim to build a unified multi-tier framework to study, analyse and investigate music and the acoustic cues associated with it. The framework consists of:-

- **<u>Recommendation System</u>**
  - Using the acoustic features of songs described below, we aim to build recommendations for similar songs and similar artists. This enables us to compute rich semantic similarities based on the inherent musical cues in the data.
- **<u>Time-Series Forecasting</u>**
  - We aim to analyse the time series aspect of the data to study how music and genres change over time. This also enables us to forecast how music may change over a period of time.
- **<u>Clustering</u>**
  - We also cluster the genres and artists to build unsupervised clustering algorithms based on the acoustic features.

Using the features described, we conduct a thorough exploratory analysis to study how music can be explored through the eyes of statistical computation.



DATASET

We use a dataset of 1,60,000 songs released between 1920-2020.

The Dataset is available from Spotify via it's Developer API, the features of the dataset have been described in here.

- **duration_ms** The duration of the track in milliseconds.
- **key** The estimated overall key of the track. Integers map to pitches using standard Pitch Class notation . E.g. 0 = C, 1 = C♯/D♭, 2 = D, and so on. If no key was detected, the value is -1.
- **mode** Mode indicates the modality (major or minor) of a track, the type of scale from which its melodic content is derived. Major is represented by 1 and minor is 0.

- **acousticness** A confidence measure from 0.0 to 1.0 of whether the track is acoustic. 1.0 represents high confidence the track is acoustic.

- **danceability** Danceability describes how suitable a track is for dancing based on a combination of musical elements including tempo, rhythm stability, beat strength, and overall regularity. A value of 0.0 is least danceable and 1.0 is most danceable.

- **energy** Energy is a measure from 0.0 to 1.0 and represents a perceptual measure of intensity and activity. Typically, energetic tracks feel fast, loud, and noisy. For example, death metal has high energy, while a Bach prelude scores low on the scale. Perceptual features contributing to this attribute include dynamic range, perceived loudness, timbre, onset rate, and general entropy.

- **instrumentalness** Predicts whether a track contains no vocals. “Ooh” and “aah” sounds are treated as instrumental in this context. Rap or spoken word tracks are clearly “vocal”. The closer the instrumentalness value is to 1.0, the greater likelihood the track contains no vocal content. Values above 0.5 are intended to represent instrumental tracks, but confidence is higher as the value approaches 1.0. The distribution of values for this feature look like this:

- **liveness** Detects the presence of an audience in the recording. Higher liveness values represent an increased probability that the track was performed live. A value above 0.8 provides strong likelihood that the track is live.

- **loudness** The overall loudness of a track in decibels (dB). Loudness values are averaged across the entire track and are useful for comparing relative loudness of tracks. Loudness is the quality of a sound that is the primary psychological correlate of physical strength (amplitude). Values typical range between -60 and 0 db.

- **speechiness** Speechiness detects the presence of spoken words in a track. The more exclusively speech-like the recording (e.g. talk show, audio book, poetry), the closer to 1.0 the attribute value. Values above 0.66 describe tracks that are probably made entirely of spoken words. Values between 0.33 and 0.66 describe tracks that may contain both music and speech, either in sections or layered, including such cases as rap music. Values below 0.33 most likely represent music and other non-speech-like tracks.

- **valence** A measure from 0.0 to 1.0 describing the musical positiveness conveyed by a track. Tracks with high valence sound more positive (e.g. happy, cheerful, euphoric), while tracks with low valence sound more negative (e.g. sad, depressed, angry).

- **tempo** The overall estimated tempo of a track in beats per minute (BPM). In musical terminology, tempo is the speed or pace of a given piece and derives directly from the average beat duration.

Methods/Materials/Tools

- **PCA / t-SNE**
  - We use these statistical dimensionality reduction algorithms to embed the acoustic features in a latent space.
- **Prophet**
  - This is a linear model based statistical time forecasting algorithm which we plan to use to model the dynamics of music data and the associated time stamps.
  - Additive Regression model
- **HAC**
  - We use hierarchical agglomerative clustering to cluster artists/songs based on their latent space embeddings/feature vectors.

Similarity Based Metrics that are being used in this project:

Based on the extracted features, the distance between songs are measured. Three typical similarity measurements are listed below.

- **K-means clustering with Earth-Mover's Distance:** It computes a general distance between Gaussian Mixture Models (GMM) by combining individual distance between gaussian components.
- **Average Feature Vectors with Euclidean Distance**: It calculates loworder statistics such as mean and variance over segments.
- **Cosine Similarity:** Cosine similarity is a measure of similarity between two non-zero vectors of an inner product space. It is defined to equal the cosine of the angle between them, which is also the same as the inner product of the same vectors normalized to both have length 1.

$$similarity(A,B) = \frac{A \cdot B}{\|A\| \times \|B\|} = \frac{\sum_{i=1}^{n} A_i \times B_i}{\sqrt{\sum_{i=1}^{n} A_i^2} \times \sqrt{\sum_{i=1}^{n} B_i^2}}$$

Technologies used -

- ➔ Python, Numpy, Pandas, scikit-learn
- ➔ Plotly for data visualisation, flask framework for building web applications.

TIME FRAME(GANTT CHART)

Activity	August		September			October		November	Start Date	Finish Date
Project Discussion									21/08/2020	28/08/2020
Collect Facts									21/08/2020	28/08/2020
Analysis and Development									21/08/2020	28/08/2020
Proposal and Approal									21/08/2020	28/08/2020
Literature Review									28/08/2020	11/09/2020
Algorithm Designing									04/09/2020	09/10/2020
Estimated First Progress report									18/09/2020	25/09/2020
Publication Work									09/10/2020	13/11/2020
Report Writing									09/10/2020	13/11/2020
Estimated Final Presentation									06/11/2020	13/11/2020

*All dates are tentative and subject to change.

REFERENCES

[1]:https://www.kaggle.com/yamaerenay/spotify-dataset-19212020-160k-tracks

Spotify Dataset 1921-2020

[2]:https://ieeexplore.ieee.org/abstract/document/6395003

Genre Classification for Musical Documents Based on Extracted Melodic Patterns and Clustering

[3]:https://www.researchgate.net/publication/279495550_Time_Series_Analysis_as_a_Method_to_Examine_Acoustical_Influences_on_Real-time_Perception_of_Music

Time Series Analysis as a Method to Examine Acoustical Influences on Real-time Perception of Music

[4]:http://webprojects.eecs.qmul.ac.uk/marcusp/papers/Pearce2011.pdf

Time-series analysis of Music: Perceptual and Information Dynamics

[5]:https://unsworks.unsw.edu.au/fapi/datastream/unsworks:500/SOURCE01?view=true

Measurement and Time Series Analysis of Emotion in Music

[6]:https://rstudio-pubs-static.s3.amazonaws.com/249909_e9f45fd3d3f44cfea9b37a935b286b81.html

Recommender System for Music Recommendation

[7]:https://towardsdatascience.com/music-artist-recommender-system-using-stochastic-gradient-descent-machine-learning-from-scratch-5f2f1aae972c

Music artist Recommender System using Stochastic Gradient Descent | Machine Learning from Scratch

[8]:https://www.frontiersin.org/articles/10.3389/fams.2019.00044/full

Deep Learning in Music Recommendation Systems

[9]:https://arxiv.org/abs/1812.10613

Generative Adversarial User Model for Reinforcement Learning Based Recommendation System

[10]:https://www.researchgate.net/publication/284219925_The_Use_of_Machine_Learning_Algorithms_in_Recommender_Systems_A_Systematic_Review

The use of machine learning algorithms in recommender systems: A systematic review

[11]:https://www.researchgate.net/publication/277714802_A_Survey_of_Music_Recommendation_Systems_and_Future_Perspectives

A Survey of Music Recommendation Systems and Future Perspectives

[12]:https://ieeexplore.ieee.org/abstract/document/5230332

Music Recommendation Based on Acoustic Features and User Access Patterns

[13]:https://facebook.github.io/prophet/

Prophet

[14]:https://peerj.com/preprints/3190

Forecasting at Scale

[15]:https://www.researchgate.net/publication/277714802_A_Survey_of_Music_Recommendation_Systems_and_Future_Perspectives

A Survey of Music Recommendation Systems and Future Perspectives