Song Recommendation System

Namit Shah | Martand Javia | Suhanee Patel | Devarsh Patel

***Abstract*—Music is one of the most popular and important cultural aspects of the society. Over the time, numerous audio streaming platforms have been developed to satisfy our pursuit for music. Evidently, people have different musical taste and hence it is crucial to have a good recommendation engine as per one's preferences. The idea is simple - to build a recommendation system which can recommend songs based on user's preferences by means of unsupervised learning mechanisms.**

***Index Terms*—spotify, unsupervised, clustering, recommendation**

I. Introduction

Spotify is one of the leading audio streaming platforms, hosting almost all song collections present in the world. We selected an independent dataset derived from Spotify Public API comprising of more than 1.2 Million songs [1], to design a machine learning model for the song recommendation system. The prediction is done based on different audio features applying unsupervised learning algorithms to classify the data into different clusters. The system will predict songs belonging to a particular set of clusters based on different features of the song. The dataset consists of various attributes and features of the songs, which helps in grouping them based on certain similar premises.

II. Literature Survey

Most of the platforms for streaming songs have their own recommendation systems. These recommendation systems analyze the user's preference by merely averaging the properties of music within the user's list. P. Cano presented "content-based system" where he extracted descriptions related to instrumentation, rhythm and harmony from music signal using similarity metrics [5]. G. Tzanetakis presented a way to extract feature from music, and divide musical genre automatically [4]. Another approach for recommendation includes collaborative filtering. Here, recommendations are made considering user preferences into account along with the attributes of the songs. Other models also include content based filtering which includes keyword searching, user actions and feedback. Though there are numerous methods for developing predictive models for songs, item based filtering is the most standard approach for creating an elementary recommendation system.

III. Implementation

A. Preprocessing Dataset

In order to derive valid inferences, it is important to make sure that the dataset does not have biased error terms. As a result, the selected dataset was initially pre-processed and cleaned to extract and remove all the indifferent values. Moreover, incomplete data values were removed from the dataset by employing different analysis functions.

B. Observing trends in the audio features

Generally, every dataset observes some sort of trend or pattern to derive abstract intuitions for the analysis. After processing the dataset, we tried to visualise the trends and relationships between different features in the dataset. In this case, it helped finding patterns between different types of songs, and visualising the evolution of songs and music in last couple of decades. This process can also be used to study the change in people's song preference over the course of time.

C. Applying Unsupervised Learning Algorithm

To build a recommendation system, the dataset should be classified into several groups comprehending the relationship between the features. As the selected dataset does not contain any feature which can be considered as a target variable, we adopted unsupervised learning mechanisms. K-Means algorithm was implemented to cluster the dataset into different groups. K-means basically divides the data into K different clusters taking into consideration various features and the nature of the dataset. Deciding the optimal value of K is very important considering the complexity of algorithm and size of the dataset. For this, elbow method was implemented which compares different values of K with their run-time performance to provide an optimal K-value.

D. Visualising clusters

Since the dataset has many features, it is very difficult to visualize the dataset after the clusters are formed. There are clusters in the feature space (properties of each songs) which are only visible via dimensionality reduction. Thus, for reducing this high dimensional feature space, we use PCA (Principal Component Analysis), as PCA in conjunction with k-means is a powerful method for visualizing high dimensional data. Reducing all the features of the songs down to the principal components and then visualizing the clusters in those principal components using k-means hints the clustering of the dataset.

IV. Results

For the implementation, the data with null values were eliminated and the features of each song is processed, so that the range of the data is in the numerical values. The results obtained by implementing the algorithms discussed previously are presented below along with its description of their representation.

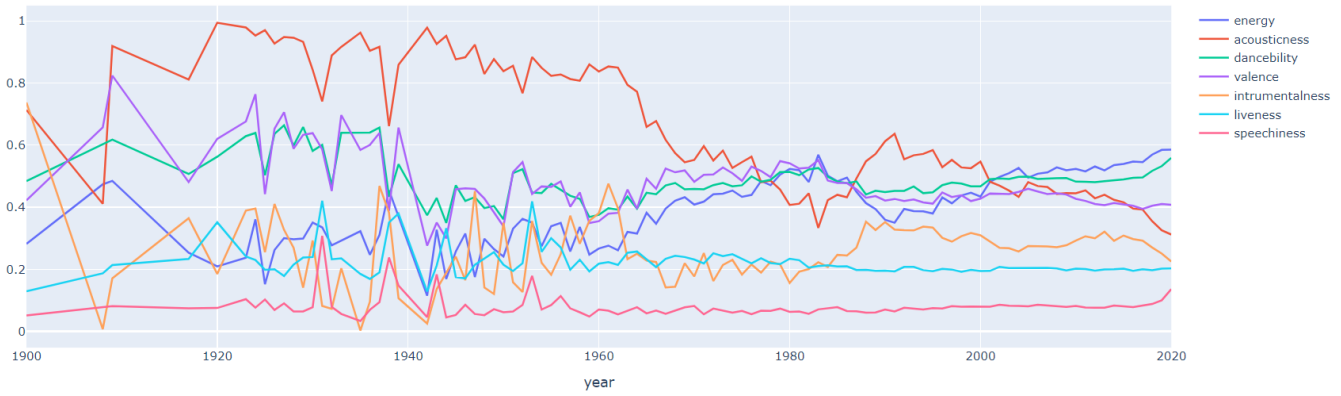

Fig. 1. Trend of audio features over decades

Fig. 1 shows the normalized behavior of various features in the dataset. It shows the mean value of all the features of the songs listed in the dataset for a particular year. Analyzing the plots, we can infer that there has been a noticeable change in the composition of music/songs over the past couple of years. For example, newer songs tend to have lower acousticness and instrumental effects as compared to older songs.

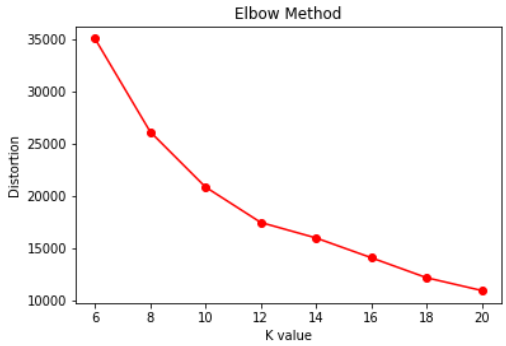

Fig. 2. Elbow Method: $distortion\ \ vs\ \ k$

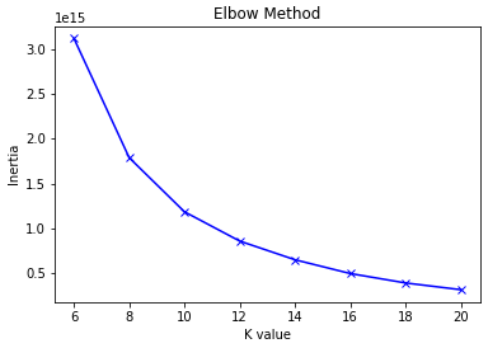

Fig. 3. Elbow Method: $inertia\ \ vs\ \ k$

Fig. 2 shows the results of elbow method for the chosen dataset. Fig. 2 is the plot of distortion vs. k (number of clusters), where distortion here means the average of the squared distances (considering euclidean distance) from the data points to the mean of the respective clusters & Fig. 3 is the plot of inertia vs. k (number of clusters), where inertia is the sum of squared distances of samples to the mean of data points of the closest intra-cluster.

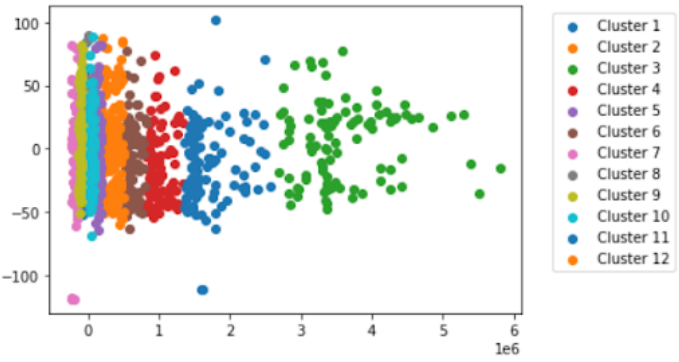

Fig. 4. PCA for 2 components

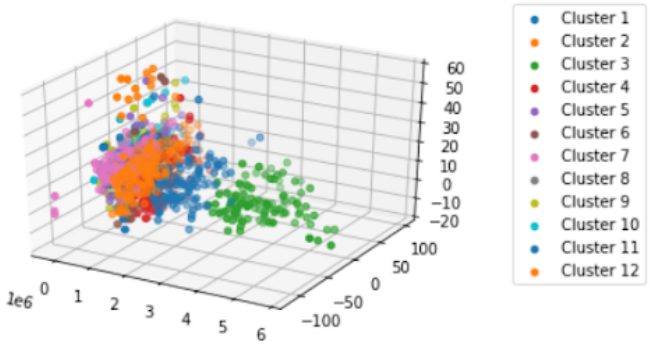

Fig. 5. PCA for 3 components

Fig. 4 shows the data points obtained after performing PCA for 2 principal components scattered in 2D plane & Fig. 5 shows the data points obtained after performing PCA for 2 principal components scattered in 2D plane through which we can visualize the clusters formed by k-means.

V. Conclusion

Using classical unsupervised learning algorithms, the dataset can be classified uniformly to differentiate various types of songs. Visualising the dataset showed that the features are weakly correlated and hence, large number of clusters are required to classify the data. Implementing elbow method to determine the optimal number of clusters returned k=12 as the elbow value. As the k-value is quite large, Principal Component Analysis (PCA) has to be performed on the dataset for proper data visualization. Two dimensional PCA and three dimensional PCA does not help differentiate between clusters accurately.

References

[1] Kaggle.com. 2021. Spotify 1.2M+ Songs. [online] Available at: https://www.kaggle.com/rodolfofigueroa/spotify-12m-songs.

[2] Need, T., 2021. K Means Clustering — K Means Clustering Algorithm in Python. [online] Analytics Vidhya. Available at: https://www.analyticsvidhya.com/blog/2019/08/comprehensive-guide-k-means-clustering

[3] Chen, H. and Chen, A., 2001. A music recommendation system based on music data grouping and user interests. Proceedings of the tenth international conference on Information and knowledge management - CIKM'01.

[4] Tzanetakis, G. and Cook, P., 2002. Musical genre classification of audio signals. IEEE Transactions on Speech and Audio Processing, 10(5), pp.293-302.

[5] Cano, P., Koppenberger, M. and Wack, N., 2005. Content-based music audio recommendation. Proceedings of the 13th annual ACM international conference on Multimedia - MULTIMEDIA '05.