Music Recommendation System

Shail Patel
AU1940142
shail.p2@ahduni.edu.in
B.Tech CSE

Sahil Miskeen
AU1940267
sahil.m@ahduni.edu.in
B.Tech CSE

Varun Deliwala
AU1940034
varun.d@ahduni.edu.in
B.Tech CSE

Abstract—**Using past data of the listening preferences of the user in order to predict his future recommendations of songs. Use the different features related to songs (liveness, acoustic, dynamic etc) to predict what taste the user has in music is what drives all the music streaming companies. To get these predictions right is a tough task as the human mind keeps changing preferences continuously. We aim to use machine learning techniques in order to provide the best possible recommendations of song personalised according to the user.**

Key Words—**K-nearest neighbour, Machine Learning, Regression, Dependency, Recommendation**

I. INTRODUCTION

The number of hours an average person spends on listening to music, be it while working or studying or during one's leisure time, is substantially increasing day by day. With the increase of user base on applications like spotify, it often becomes difficult to discover new music that matches one's interests. So, the main aim of the project here is to identify the best solutions for music recommendation.

II. LITERATURE SURVEY

The approches taken till now for music recommendation systems are based on deep neural networks. These deep neural networks work on the basis of content based filtering taking into account different variables like the past listening history of the user, listening time of various genres etc. (1) Logistic regression has been applied to build a recommendation system. This takes into account the listening history of the user and predicts whether or not the user will like another song or not. Howevr this is a very strict recommendation system and the results are quite poor. (2) KNN is another ML algorithm which allows the recommendation system to make a comparison between K different songs and predict the next closest neighbor which corresponds to the song and gives us that song as a recommended song.

A. Logistic Regression

Logistic regression is commonly used in recommendation system where the model predicts a user will watch a movie or in our case, like a song suggested based on the movies he/she has watched previously. Sor for a given input pair (U, S) of user U and movie S, we want to predict whether the user will (0) or will not (1) like the song. As a logistic regression model can also predict the probability of the interaction in addition to a binary label, it is popularly used to the predict probilities to sort the object in terms of users and recommend some fixed number of top-ranked Objects, the objects ranging from movies, songs to food dishes too.

B. KNN Algorithm

k-Nearest neighbors is a machine learning algorithm which is commonly used in recommendation systems so as to find clusters of similar users, based on the common likings or experiences. The model makes predictions using the average distance (in our case the previous listens) of the top-k nearest neighbors. There can be a content based approach where the discrete characteristics of an item are utilized in order to recommend addition items with similar properties and there can be Collaborative filtering approach where the model becomes more accurate with increase in information about the user's past behavior. Majorly, models currently used in recommendation systems are based on a hybrid approach so as to improve accuracy.

III. IMPLEMENTATION

Understood the variables which we will be working with. (needs domain knowledge) and performing preprocessing on data. Visualization of the variables and relation between them. Removing the highly correlated variables to reduce redundancy of the dataset.

IV. RESULTS

Depicting the correlation between variables and the variable named "popularity" from the dataset and it can be seen in the following graph:
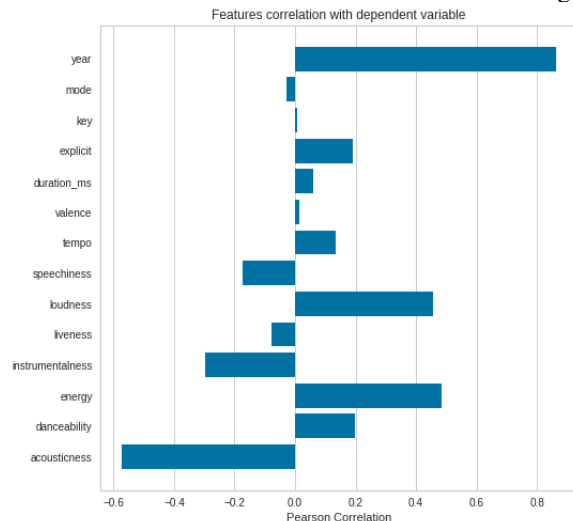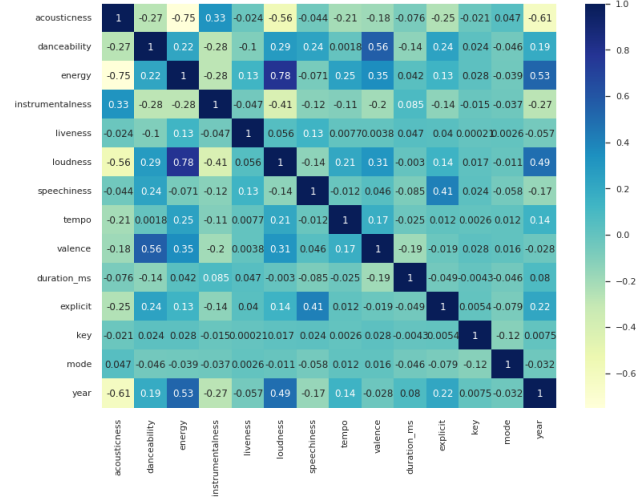
Fig-1: Correlation between variables and "popularity" variable

Understanding the correlation between variables and then remove the highly dependent variables:

Fig-2: Correlation Graph of variables

Capturing non linear relations between the variableshttps://www.overleaf.com/project/6236c4292cdb4f7188613f01
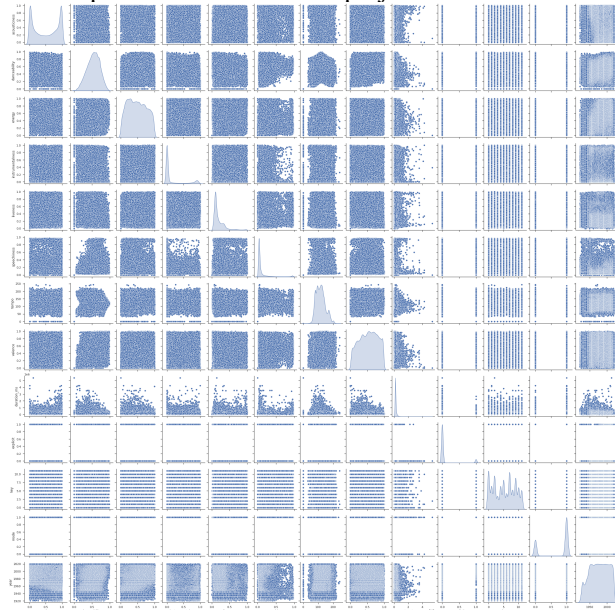

Fig-3: Correlation Graph of variables

V. CONCLUSION

As we can see from the above results that preprocessing of the raw data has helped us quite a lot in sorting the entries in a manner which can in turn help in better implementation of the model. After implementation of such a model, we will need to understand the efficiency of it and at the same time compare it with other existing models which might be content-based and provide better accuracy than the current one

REFERENCES

[1] Liao, K., 2018. Prototyping a Recommender System Step by Step Part 1: KNN Item-Based Collaborative Filtering. [online] Medium. Available at: ¡https://towardsdatascience.com/prototyping-a-recommender-system-step-by-step-part-1-knn-item-based-collaborative-filtering-637969614ea¿ [Accessed 20 March 2022].

[2] M. Schedl, "Deep Learning in Music Recommendation Systems," Frontiers, 01-Jan-1AD. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fams.2019.00044/full. [Accessed: 28-Mar-2022].

[3] "A music recommendation system based on logistic regression ..." [Online]. Available: https://www.researchgate.net/publication/336162555 A Music Recommendation System Based on logistic regression and eXtreme Gradient Boosting. [Accessed: 28-Mar-2022].