



INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

MUSIC RECOMMENDER SYSTEM

Sarfraz Ahmad*1, Sagori Halder*2, Tanfeez Ahmad*3,

Vishnu Singh*4, Prof. Mr. Sreenu Banoth*5

* 1 Student, Department of Computer Science & Engineering, IIMT College of Engineering, Greater Noida, Uttar Pradesh India.

*2 Student, Department of Computer Science & Engineering, IIMT College of Engineering Greater Noida, Uttar Pradesh India

*3 Student, Department of Computer Science & Engineering, IIMT College of Engineering, Greater Noida, Uttar Pradesh, India

*4 Student, Department of Computer Science & Engineering, IIMT College of Engineering, Greater Noida, Uttar Pradesh, India

*5 Professor, Department of Computer Science & Engineering, IIMT College of Engineering, Greater Noida, Uttar Pradesh, India

1.ABSTRACT

For building the music recommender system it requires very much information or data. So, it is retrieval task. In this project, we will be using a sample data set of songs to find the correlations between users and songs so that a new song will be recommended to them based on their previous history. We applied three approaches to build this music recommender system i.e., Popularity Based Filtering, Collaborative Filtering and Content Based Filtering. We will implement this project using libraries like NumPy, Pandas. We will also be using Cosine similarity along with Count Vectorizer. Along with this, a front end with flask that will show us the recommended songs when a specific song is processed.

Keywords: NumPy, Pandas, Python, Machine Learning.

2.INTRODUCTION

Nowadays lots of music industries like amazon music, wink music, gaana.com are using recommender systems and the old-fashioned way of selling music has changed to a totally different cloud based. Now all the music resources are present in their cloud and users can listen to the songs directly from the cloud. But the issue is there are lot of songs present in the cloud system. so, we need to classify all the songs based on different genres, artists locations, age groups, languages and the main goal is to classify these set of songs in accordance to the taste of the user. Because user expects valuable return after the investment of time as well as money thereby, we can attract a lot of customers by providing various valuable services of their interests for this project we are using various machine learning algorithms as well as data mining techniques. We have implemented various algorithms and compared the results with one another to find the effective algorithm that suits our model. The most common approaches people have used implementing various recommender systems are collaborative filtering and content-based models. These algorithms aim to find similarity between various user's various songs and artists. Other than these algorithms we have also used random forest algorithm and decision tree algorithm. Both of these algorithms aim to predict a decision based on different attributes. For improving the efficiency and accuracy of the model we have also used cross validation

technique in which we have recursively changed the training and testing data set to get optimum results. We have faced a lot of problems which making this project like the data we have chosen was too big (1.2 GB) so we had to create a subset of that data set for optimum usage. They were lot of outliers present in that data set which required complex data pre-processing.



Figure 1.1(Recommendation Systems)

With the promotion of the Internet and the advancement of E-trade, the Ecommerce destinations offer a great many items available for purchasing. Picking among such a large number of choices is challenging for buyers. So, clients typically lose all sense of direction in the huge space of ware data and can't discover the products they truly need. Recommender frameworks have risen in light of this issue. A recommender framework for an E-trade site prescribes items that are probably going to meet client's requirements. Suggestion frameworks have rapidly changed the way in which the life less sites can now interact and speak with their clients and users. As opposed to provide a constant involvement in which clients look for and conceivably purchase items, recommender frameworks increment cooperation to give a more extravagant or deal. Recommender architecture is used by the EBusiness objective to propose and suggest items and services which are similar to their clients. There are many constraints and parameters on which an internet service provider recommends a user certain choices and options depending upon some restrained set of parameters. These parameters can include language, age, nationality, history, likes, ratings, purchase and many more. The items can be prescribed and suggested dependent on the best generally speaking and interacting vender on a particular site, in the presence of the social and economic constraints of the client, or dependent on an examination of the past purchasing conduct of the client as an expectation for future purchasing conduct. These parameters enable the service providers to analyze and determine a set of choices and preferences. Moreover, these strategies have been used extensively and are a piece of personalization on a site, since this enables the site to adjust to every client in accordance to the client or user. Moreover, these famous and successful online service providers use personalization as a key component while recommendations because generalization cannot be more accurate under defined parameters. Recommender frameworks mechanize personal services and platforms on the internet, which facilitates and empowers singular personalization for every client.

3.BRIEF LITERATURE SURVEY

The proposition issue in the music region has additional challenges as person & music perception depends upon various parameters and constraints. In research it was found that songs acumen is impacted by the setting of the customer. They found that music preference mainly differs on the basis of age differences, locations and languages. These parameters further can be classified into sub age groups, countries, states, regional languages and many more. It was reported that artists of similar sounds do not necessarily have the similar music and taste of listeners may differ. Music can be near or undeniable to the extent in every way that really matters any property that can be used to depict music, for instance, sort, melody, beat, arrive starting and instrumentation, which makes it possible to answer the subject of similarity between two skilled workers from different perspectives.

In research it was found that most of the music listeners are in the age between 16 to 45 years of age and that was further divided into sub - groups:

- i. Broad taste:** People whose melodic learning are exceptionally broad. They contributed 7 percent of total division.
- ii. Enthusiasts:** There are lot of people in this world who believes that music is life and they are crazy for music. Indeed, music is the most relaxing thing in this world. They include 21 percent of this division.
- iii. Casual music listeners:** People who casually listen to music in their free time include 32 percent of this division.
- iv. Indifferent:** They have different mindset about music and including 40 percent of this age group.

As per research every person requires unique set of suggestions. Academics is exceptionally urgent and are along these lines the most troublesome audience members to give suggestions to. They require unsafe and shrewd proposals rather than famous ones. Lovers then again value a harmony between fascinating, obscure, and commonplace proposals. Casuals and indifferent, who speak to 72% of the populace, don't require confused proposals and famous standard music that they can without much of a stretch relate to would accommodate their melodic needs. In this way, it is critical for a recommender framework to have the capacity to recognize the kind of client and act as needs be. The objective was to enhance suggestion precision by including more sound information from numerous melodies. For this purpose, songs from similar collection and similar artists were analyzed to find the correlation and was named as "collection effect". As of late, be that as it may, inquire about on recommenders utilizing communitarian separating has picked up a greater prominence in the music space. The main music recommender framework utilizing community oriented. It utilized a compelled person connection for computing similarity effect which corresponds to total like content. Then again, slithered client related to an enormous account of robotized tunes that drove web looks of blueprints to somebody's gifted specialists. They utilized system sifting techniques on the information to diagram proposals. Then they utilized substance based and synergistic sifting approaches unreservedly to support music subject to music and client social gatherings. The music packs contained melodies the client was beginning late enchanted by and client bunches set clients with comparative interests. They accomplished higher precision with the substance-based framework, yet the synergistic secluding system gave all the moreover dumbfounding suggestions. Sanchez-Moreno et al. (2016) proposed a total separating method that utilized listening coefficients as an approach to manage area the decrease sheep issue of synergistic sifting. In order to distinguish between clients, the listening clients conduct with respect to specialists they tune into which is utilized to describe the clients dependent on the exceptionalness of their inclinations. The proposed strategy fundamentally surpassed more customary community oriented sifting technique.

Personalized Recommender Systems:

Personalization issues adapting to the individual desires, interests, and preferences of every user. They are tools for suggesting things to users.

Non-Personalized Recommender Systems:

As is evident from the name itself, a non-personalized recommender system is a generic recommender system which provides recommendations based on the opinions and the feedbacks of the other users.

Further these are divided into Popularity Based Filtering, Collaborative Filtering and Content Based Filtering:

Personalized Popularity Based Filtering: As the name suggests popularity-based recommender system works with the trend. It basically uses the items which are in trend right now.

Collaborative Filtering: The basic idea of collaborative filters that similar users tend to like similar items. It is based on the assumption that if some users have had similar interests in the past, they will also have similar tests in the future.

Content Based Filtering: The basic idea of content-based filtering is to map relevant content based on certain features or characteristic. Content here refers to the content or attributes of the products you like.

4. PROBLEM FORMULATION

NEED:

As the web moved from a proprietor model to an open publicly supporting model and enabled individuals to contribute unreservedly, it saw an exponential ascent in the measure of substance accessible, which was something to be thankful for. Be that as it may, this prompted two noteworthy issues:

i. Aggregation: The measure of data turned out to be large to the point that it inspired extreme to oversee it while as yet having the capacity to run a web benefit that was reachable to all parts of the world. This issue was tackled by building overall substance conveyance and dissemination systems, helped by the ascent of NoSQL Database frameworks and diminishing stockpiling costs.

ii. Searching: The second significant issue was the means by which to guarantee that the data is inside the scope of the client and that the client does not become mixed up in the immense information dumps accessible. This turned out to be a significantly more concerning issue than accumulation since the information troves are tremendous and every client carries alongside him/her a remarkable point of view and consequently a one-of-a-kind pursuit design. We are as yet attempting to take care of this issue today and are a long way from accomplishing an ideal answer for it. This is the place recommender frameworks become possibly the most important factor.

More or less, a recommender framework is a framework that predicts client reaction to an assortment of choices. Anticipating what the client may get next is the basic point of a recommender framework. There is a broad class of web applications that include foreseeing the client's reaction to choices.

Such an office is known as a recommender framework.

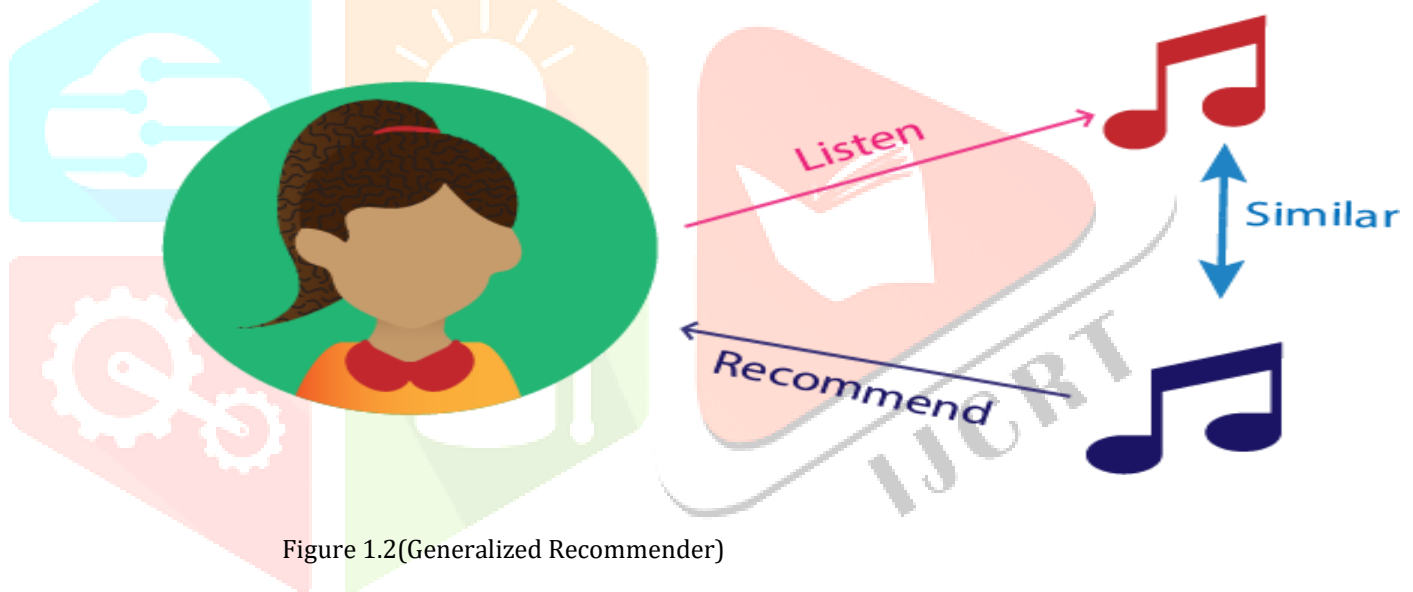


Figure 1.2(Generalized Recommender)

Recommender systems need information for functioning, data about a particular user. This particular data can be fetched directly or indirectly. Directly collecting data means that user of a particular service gives feedback and review of the item. Indirectly means that system will analyze the user's interaction with the particular service consisting of history and present services.

PROBLEM STATEMENT:

The main goal is suggesting best set of options to the user. For a specific user we had their song history frequency list liked songs. From all this information we had to predict what songs user might like then the question comes: how can we use all this information to achieve our goal. As it not a straight forward task to find the relevance between various songs it might be possible that one song which looks similar to other may be completely different and users may dislike that song or may be that song is not of users' taste. There are lots of users around the world and lots of songs so making a relevance between songs and users is a tedious task.

5.OBJECTIVES

The objective is to provide recommendation systems that fit listeners profile in terms of music universe, content popularity, familiarity, new releases, appropriation cycle (discovery, repetition, pleasure, saturation), diversity of genres, surprise, continuation of past exploration (including outside the music platform).

The main object in terms of outcome was to create a framework for users which can help them suggesting the right songs for them. This project aims to find the correlation and similarity between different music lovers their tastes and various songs so that if a user's taste is similar to the other one, we can recommend the songs of one to another on the basis of similar taste. Or if a song is similar to the other, we can suggest that song to the user that listen the first one. One of the objects of this project is to reduce the time that user generally wastes on looking for the right song.

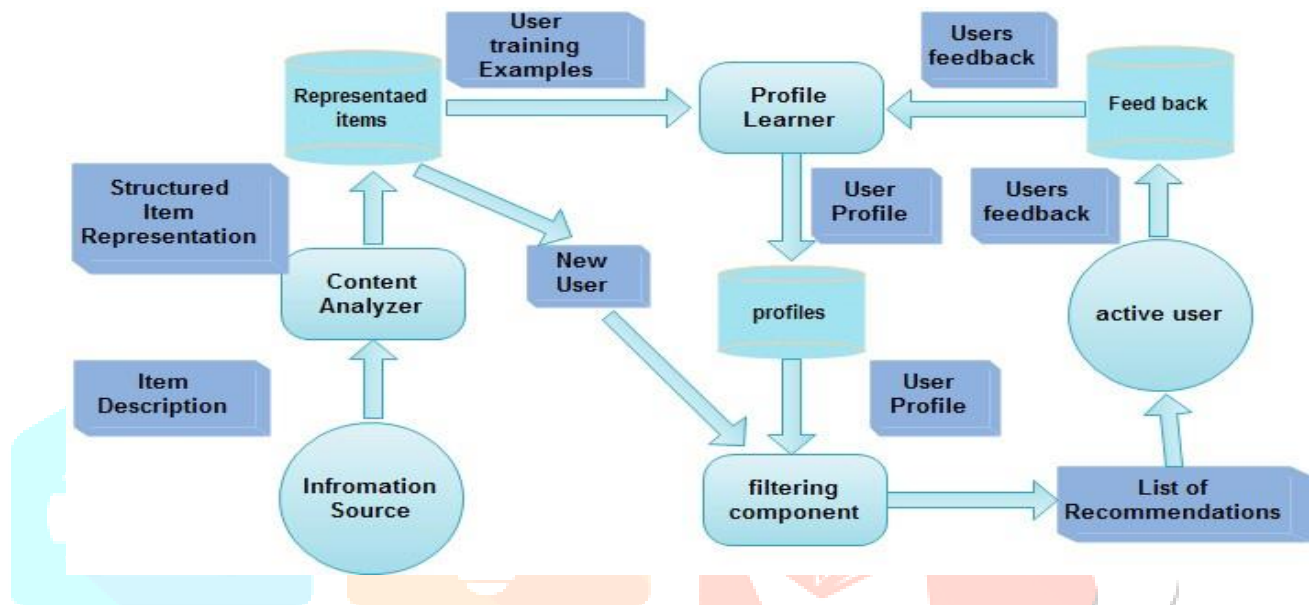


Figure 1.3(Content Based Architecture)

6.METHADODOLOGY

FUNCTIONAL REQUIREMENT:

The functional requirement specification of the project are mainly categorized as user requirements, security requirements, and device requirement each of which are explained in detail below:

- i. **User Requirement:** User ought to have account on framework and client must have somewhere around one song listened to investigate the identity for the music suggestion.

NON-FUNCTIONAL REQUIREMENT:

- i. **Performance:** The framework will have a speedy, exact and dependable outcomes.
- ii. **Capacity and Scalability:** The framework will have the capacity to store identity registered by the framework into the database.
- iii. **Availability:** The framework will be accessible to client whenever at whatever point there is an Internet association.
- iv. **Recovery:** if there should arise an occurrence of breaking down or inaccessibility of server, the framework ought to have the capacity to recuperate and keep any information misfortune or excess.

- V. Flexibility and Portability:** System will be available whenever from any areas Sentiment analysis process of Twitter data.

7.LIMITATIONS

- i. Significant Investments Required:** Recommendation engines are a big investment, not only financially, but in terms of time, too: it takes a long time and deep expertise to build an effective recommendation engine in-house. Besides the requisite data scientists and other assorted specialist staff, you will need to factor in costs for the discovery and analysis phase.
- ii. Lack Of Data Analytics Capability:** Like all AI-based technologies, recommendation engines rely on data – if you do not have high-quality data, or cannot crunch and analyze it properly, you will not be able to make the most of the recommendation engine.
- iii. The Cold Start Problem:** Relying on user data has its downsides, one of which is the issue of ‘cold start’. This is when a new user enters the system or new items are added to the catalogue, and therefore, it will be difficult for the algorithm to predict the taste or preferences of the new user, or the rating of the new items, leading to less accurate recommendations.
- iv. Privacy Concern:** The more the algorithm knows about the customer, the more accurate its recommendations will be. However, many customers are hesitant to hand over personal information, especially given several high-profile cases of customer data leaks in recent years. However, without this customer data, the recommendation engine cannot function effectively. Therefore, building trust between the business and customers is key.

8.CONCLUSION

We had a great learning experience doing this project. We have learnt about data mining and data cleaning. This is the very first task of a machine learning model to remove all the problem creating objects from the dataset. Data cleaning and data exploration were very useful to making dataset algorithm ready. We have learnt to create machine learning model, train the model and then doing testing on it. Most common algorithms used before by others were content and collaborative but the hybrid of the two gave extraordinary results. Best suited algorithm for this project was random forest algorithm.

9.REFERENCES

- [1] McFee, B., BertinMahieux, T. Ellis, D. P., Lanckriet, G. R. (2012, April). The million-song dataset challenge.
- [2] In Proceedings of the 21st international conference companion on World Wide Web (pp. 909916).ACM.
- [3] Aioli, F. (2012). A preliminary study on a recommender system for the million songs dataset challenge. PREFERENCE LEARNING: PROBLEMS ANDAPPLICATIONS IN AI
- [4] <https://maroo.cs.umass.edu/getpdf.php?id=1289>
- [4]Spotify. (2013, March) Spotify Music for everyone. [Online]. <https://www.spotify.com/us/> [10] (2013, March) Wikipedia. [Online]. http://en.wikipedia.org/wiki/LastFM#User_accounts
- [5](2013,March)GooglePatents.[Online]. <http://www.google.com/patents/US7003515?dq=7.003.515>