

MUSIC RECOMMENDATION SYSTEM



Disha Jain	(PES1UG20CS132)
Isha Adiraju	(PES1UG20CS169)
Juhi Bhattacharya	(PES1UG21CS182)

Abstract

What is the problem that you are solving ?

To analyze music currently being played and suggest similar music.
The program suggests music based on audio features and lyrics of song being played.

Describe the need and significance of the problem.

A music recommendation system is essentially a solution that allows music streaming platforms to offer their users relevant music recommendations in real-time. It provides personalization and thus boots user engagement.



Scope

What is inside the scope ?

- Recommendation system seek to predict or filter preferences according to user's choices. Useful model for various domains like e-commerce movies ,vidoes , news ,songs etc.
- We have employed ML algorithms and Data analytics to implement the same for music recommendation.

What is Outside the scope ?

- The model will not be able to sense or empathise with the user's real emotions [using sentiment analysis] and sometimes this can be a limitation since the model can only recommend music based on the content filtering and popularity .

Dataset

Dataset Source:

Million Songs Dataset contains of two files: triplet_file and metadata_file. The triplet_file contains user_id, song_id and listen time. The metadata_file contains song_id, title, release, year and artist_name. Million Songs Dataset is a mixture of song from various website with the rating that users gave after listening to the song.

Preprocessing & Data Analytics techniques used :

- Two features are combined into a new feature called song
- Title and artist column can be eliminated for cleaner results
- Shortening the data set for quicker processing
- List of the most listened song in ascending order
- Percentage determines how much the song was listened by users in the data

song_id					
	A	B	C	D	E
1	song_id	title	release	artist_name	year
2	SOQMMH	Silent Night	Monster B	Faster Puss	2003
3	SOVFVAK1	Tanssi vaa	Karkuteilla	Karkkiauto	1995
4	SOGTUKN	No One Can	Butter	Hudson M	2006
5	SOBNYVR1	Si Vos Que	De Culo	Yerba Brav	2003
6	SOHSBXH1	Tangle Of	Rene Abla	Der Mystic	0
7	SOZVAPQ1	Symphony	Berwald: S	David Mor	0
8	SOQVRHI1	We Have C	Strictly Th	Sasha / Tu	0
9	SOEYRFT1	2 Da Beat	Da Bomb	Kris Kross	1993
10	SOPMIYT1	Goodbye	Danny Boy	Joseph Lo	0
11	SOJCFMH1	Mama_m	March to	The Sun H	0
12	SOYGNWF	L'antarctic	Des cobra	3 Gars Su'	2007
13	SOLJTLX12	El hijo del	32 Grande	Jorge Negr	1997
14	SOQQESG	Cold Beer	Internatio	Danny Dia	0
15	SOMPVQB	Pilots	The Loyal	Tiger Lou	2005
16	SOGPCJI12	N Gana	Afropea 3	Waldemar	0
17	SOSDCFG1	6	Lena 20 A	Lena Philip	1998
18	SOBARPM	(Looking F	Cover Girl	Shawn Co	1994
19	SOKOVRQ	Ethos of C	Descend I	Dying Fetu	2009
20	SOIMMJJ1	Rock-N-Ru	I'm Only A	Emery	2007

Machine Learning related Details

Machine learning algorithms planning to use:

- Popularity filtering
- Content-Based filtering
- Collaborative filtering

ML Library used and other library used:

- pandas
- numpy
- Recommenders

Design Approach / Methodology / Planning of work

Understanding the problem statement:

This would include intense study of the problem statement. Recognizing the type of coding employed and mathematical knowledge required for the same.

Task Management:

Dividing the coding as well as research part equally amongst the team members to make as much progress as possible in less time.

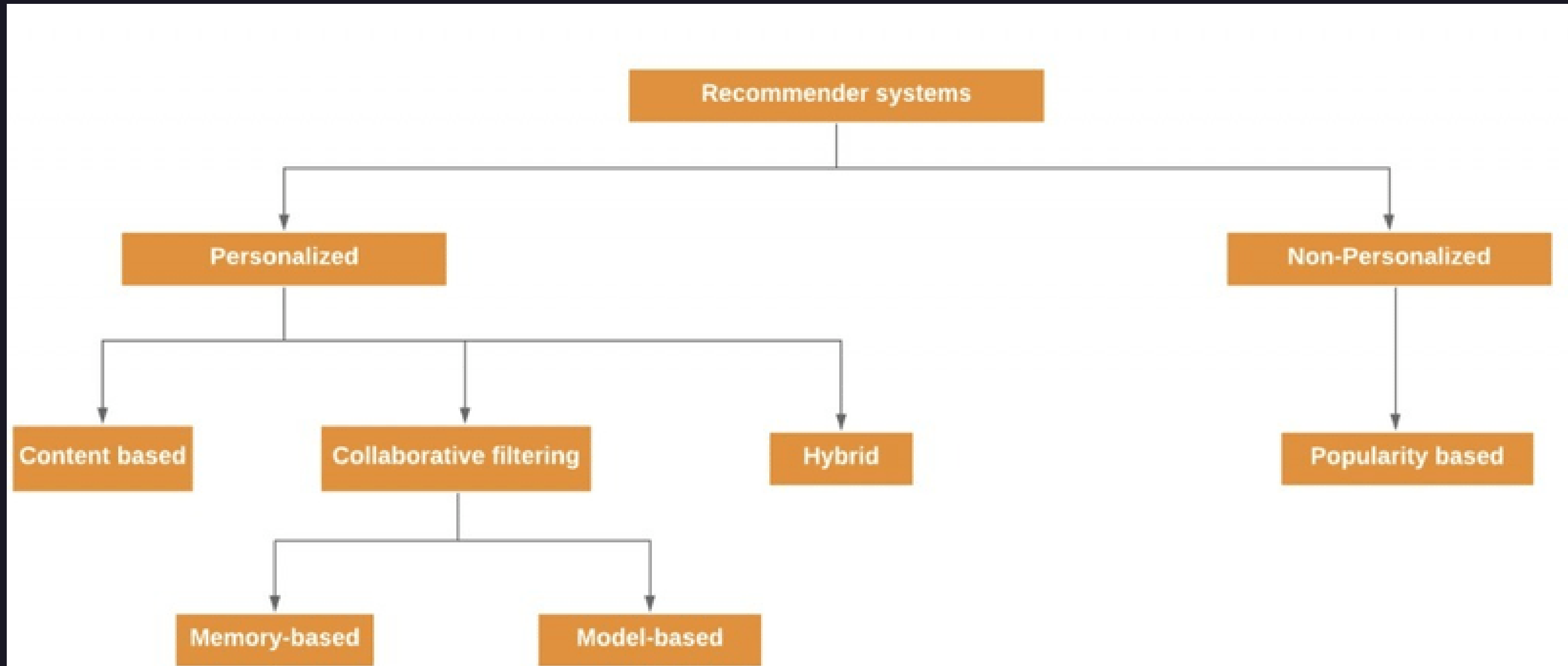
Implementing Algorithms:

Implementing algorithms mentioned and coding additional modules required to implement the project.

Fixing bugs and errors.

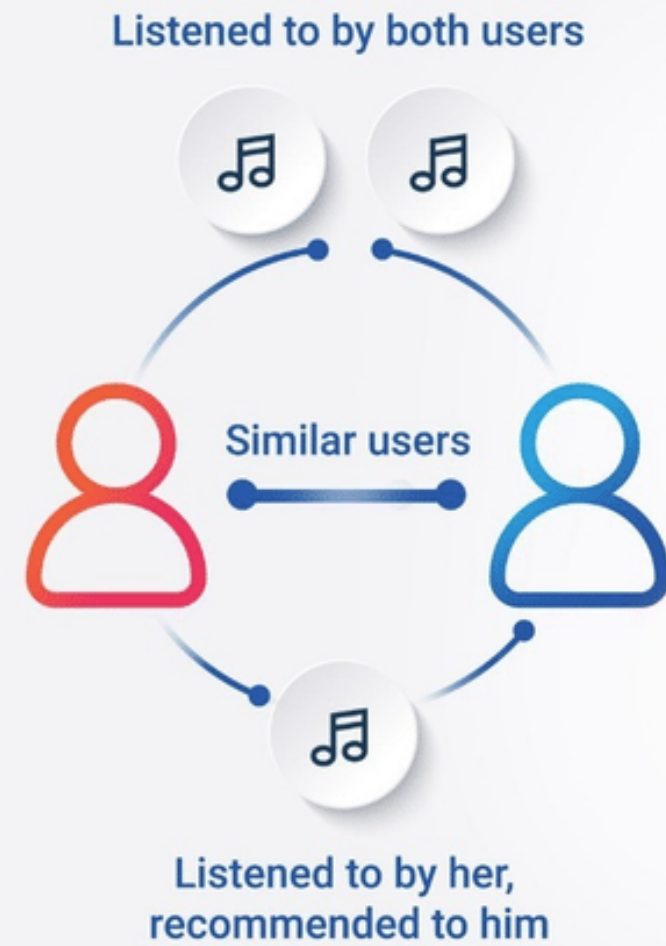


Design Details

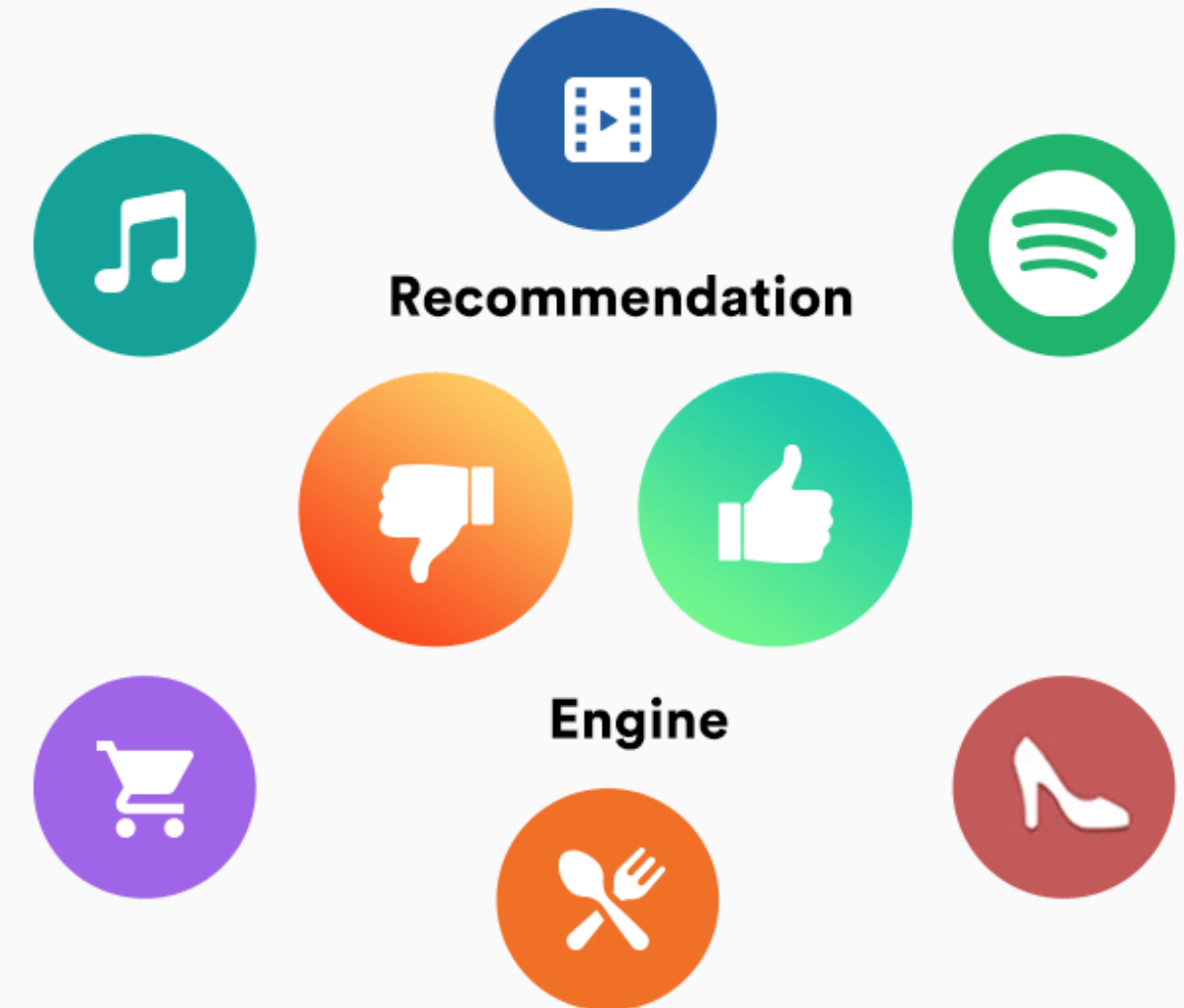
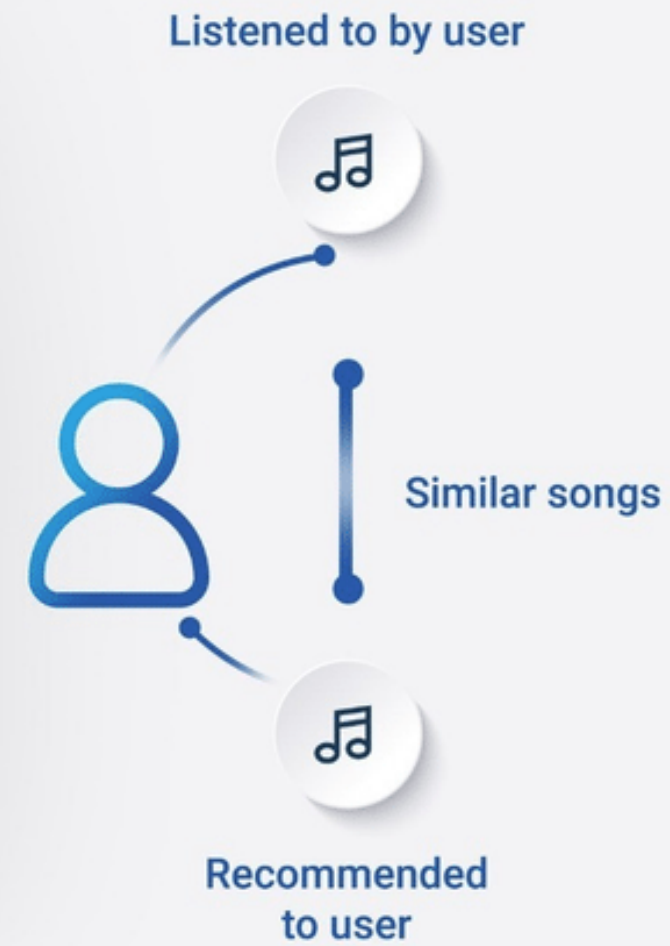


Workflow of the project/Model Architecture

COLLABORATIVE FILTERING



CONTENT-BASED FILTERING





Literature Survey

Paper 1

Year of Publication: 2019

Authors: Namitha S J

Dataset: Million Song dataset, Musixmatch dataset, and Lastfm dataset

Abstract:

Approach used in this paper focuses on various features of songs. The user interface allows user to play songs and at the same time recommends songs based on the current song being played. The recommendations are stored using firebase firestore and firebase storage.

Methodology:

KNN and ANN used to predict the similarity score between pairs of songs.

Merits:

- The system allows users to create playlists, add songs to it and stream it whenever they are logged in.
- Recommendations are also made based on the same artist.

Demerits:

- It does not recommend quality content to the users.
 - Does not consider new user cold start problem.
- Not very efficient in terms of performance

Paper 2

Year of Publication: CME, MS&E, Stanford University (2017)

Authors: Emilien Dupont, Isabelle Rao, William Zhang

Dataset: Million Song Dataset.

Abstract:

The goal of this project was to recommend songs to users based solely on their listening histories, with no information about the music.

Methodology:

Applied various Collaborative Filtering methods to achieve this:

- user-user neighborhood models
- item-item neighborhood models
- latent factor models.

Merits:

-Simple filtering methods like neighbourhood models and latent factor models used.

Demerits:

- Limitation on number of users.
- It creates a large unwanted bias towards highly played songs.

Paper 3

Year of Publication: Indian Institute of Technology, Kanpur (2014)

Authors: Shefali Garg & Fangyan SUN

Dataset: Million Song Data Challenge hosted by Kaggle

Abstract:

They used Million Song Dataset provided by Kaggle to find correlations between users and songs and to learn from the previous listening history of users to provide recommendations for songs which users would prefer to listen most.

Methodology:

Algorithms used - [Popularity based model, Collaborative based model, SVD model, KNN Model]

Merits:

- SVD Based latent factor model gives better results than popularity based model.

Demerits:

- Memory based instead of content based.
- KNN model did not work well and performs worse than even the popularity model.

Paper 4

Abstract: Collaborative filtering and content filtering recommendation algorithm to combine the output of the network with the log files to recommend music to the user in a personalized music recommendation system.

Methodology:

- SVM model
- KNN model
- Decision Tree

Merits:

- The music Recommender is able to check plagiarism in the dataset taken by generating the similarity score for each recommended song.
- Advanced Classification methods used.

Demerits:

- The complex nature of the machine learning systems like the Music Recommendation System can't have a standardized structure because different music recommender systems work in different way.
- Organising and managing the millions of music tiles that society produces.

Paper 5

Year of Publication: Dept. of Information Technology, PCE, Navi, 2021

Authors: Varsha Verma, Ninad Marathe, Parth Sanghavi, Dr. Prashant Nitnaware

Abstract:

Sample data set of songs is used to find correlations between users and songs so that a new song will be recommended to them based on their previous history. A front end with flask that will show us the recommended songs when a specific song is processed.

Methodology:

- Libraries like NumPy, Pandas.
- Cosine similarity along with CountVectorizer is used.

Merits:

- It overcomes the lack of personalization involved with non-personalized recommender systems.
- It is domain independent.

Demerits:

- The recommendations are not very specific. It still lacks personalization.
- The computational time is low.

REFERENCES:

PAPER 1:

<https://www.ijert.org/music-recommendation-system>

PAPER 2:

https://cs229.stanford.edu/proj2015/138_report.pdf

PAPER 3: https://cse.iitk.ac.in/users/cs365/2014/_submissions/shefalg/project/report.pdf

PAPER 4:

<https://www.ijert.org/music-recommendation-system-with-advanced-classification>

PAPER 5:

https://www.researchgate.net/publication/357600972_Music_Recommendation_System_Using_Machine_Learning