HINDUSTANI CLASSICAL MUSIC RECOMMENDER

RecSys Final Project

The team –

1. Ketaki Tamhanakar
2. Neha Tamhanakar
3. Vihan Vashishth
4. Vineet Priyedarshi

THE MOTIVATION

- There are many music recommendation systems currently in use, but not many which are fine-tuned to Hindustani Classical music. There are many intricacies of this form of music that can be considered while making recommendations.
- This recommender will be particularly useful to students who are learning classical music, and of course to enthusiasts of Hindustani Classical music.
- The system will recommend performances similar (in multiple different ways) to the user's taste.

A FEW DOMAIN-RELATED TERMS

- Note/Sur – A musical sound which has one major frequency perceived by the listener. (In classical music, these notes are denoted by Sa, Re, Ga, Ma, Pa, Dha, Ni, Sa)
- Raga - A set of notes/sur that are used to make compositions - governed by a few set of rules on the structure of their usage. Each raga brings a certain mood to the performance, the choice of notes matters.
- Taal - This is a rhythm, played on a percussion instrument, like Tabla, Pakhawaj, etc. There are some commonly used rhythms like Teentaal (16 beats) , Roopak (7 beats) , etc.
- Laya - Speed of the composition – Vilambit (Slow-paced), Madhya (Medium-paced), Drut (Fast-paced)
- Form – Khyal, Thumri, Tarana, etc

APPROACH CHALLENGES

- Data is not easily available on Hindustani Classical performances - especially meta-data on the performances.
- Lengths of performances are large, and each part of the performance adds a different flavor to the performance - therefore using transformers like music2vec become extremely time consuming, even for offline computation.
- Language models are hard to use here, as the language/terms used in classical music are quite specific to the art form.
- Lack of data makes it hard to train a neural network or transformer to recommend.

APPROACH

Due to the lack of data, our approach was to work with richer data, capturing different aspects that affect the similarity of two performances. We defined similarity metrics based on the following characteristics:

- Raag-based similarity
- Artist-based similarity
- Comparison based on miscellaneous details about the performance (Instruments used, rhythm, form of composition, etc.)
- Audio similarity

DATASET

- One available dataset, made for research purposes – Saraga: research datasets of Indian Art Music
  - This contained recordings and meta-data about around 100 Hindustani classical performance recordings.
  - The meta-data included : Artist name, Main Instrument, Raga, Taal, Laya, Accompanying Instruments, Form, etc.
- We additionally used performance recordings from the Darbar Festival, London performances, available on YouTube. Through web-scraping, we could extract some of the meta-data that we required. We manually added the rest of the meta-data.

PERFORMANCE DATASET

Performance dataset: Name of the main artist, and the instrument he/she plays Raga performed in the performance Taal (Rhythm) of the composition Laya of the compositions Form of the composition	We collected 316 performance recordings in total, as a combination of the two datasets mentioned. Quite a bit of the meta-data had to be manually added.

THE RAAG DATASET

- Each raag in the dataset was represented using the following attributes, which are based on Hindustani classical music theory:
  - Sur set - The set of surs/notes that the raga is defined on
  - Jaati – The number of notes in the ascent and descent of notes in a raga
  - Thaat (parent raag) – A 'parent' which the raga has been derived from
  - Vadi, Samvadi – 2 most important surs in a raag
  - Prahar –Time of the day when the raag should be sung

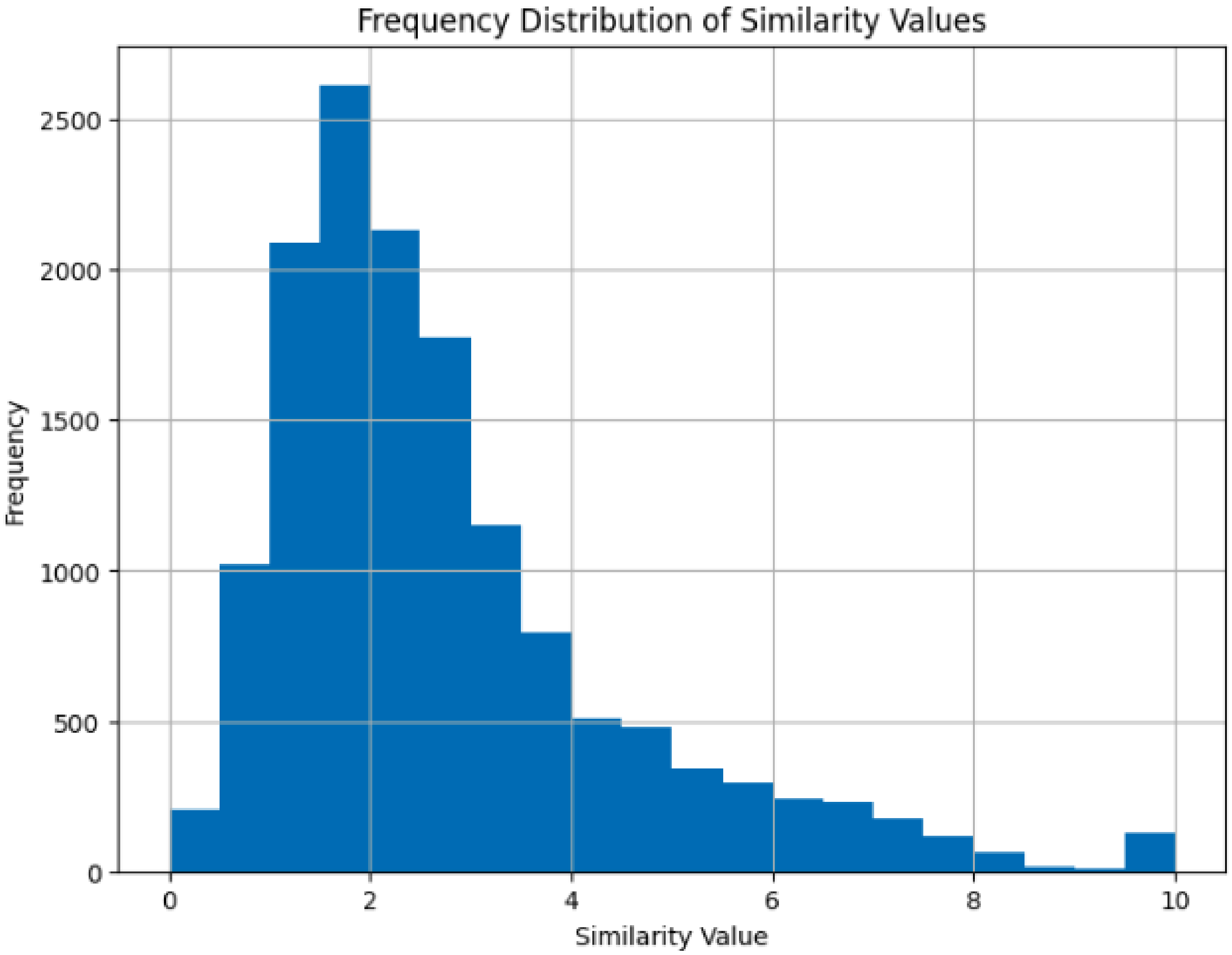
We collected information about 119 ragas, spread over all 10 thaats.

THE RAAG SIMILARITY METRIC

- Each of these attributes are one hot encoded (as there are limited number of possible values for each attribute).
- The metric considers the number of common attributes in both ragas.
- The similarity between sur sets and thaat were given more weightage as they play a more significant role in while comparing raags.
- In order to push the metric to be more discriminative, we applied the logistic function on the calculated similarity.

THE RAAG SIMILARITY METRIC

- Initially, the distribution of similarities between raags was not that discriminative; a lot of similarities appeared towards the lower end.



THE RAAG SIMILARITY METRIC

- We decided to apply the Logistic function to the values to push the high values higher and the low values lower to make the similarities more discriminative.

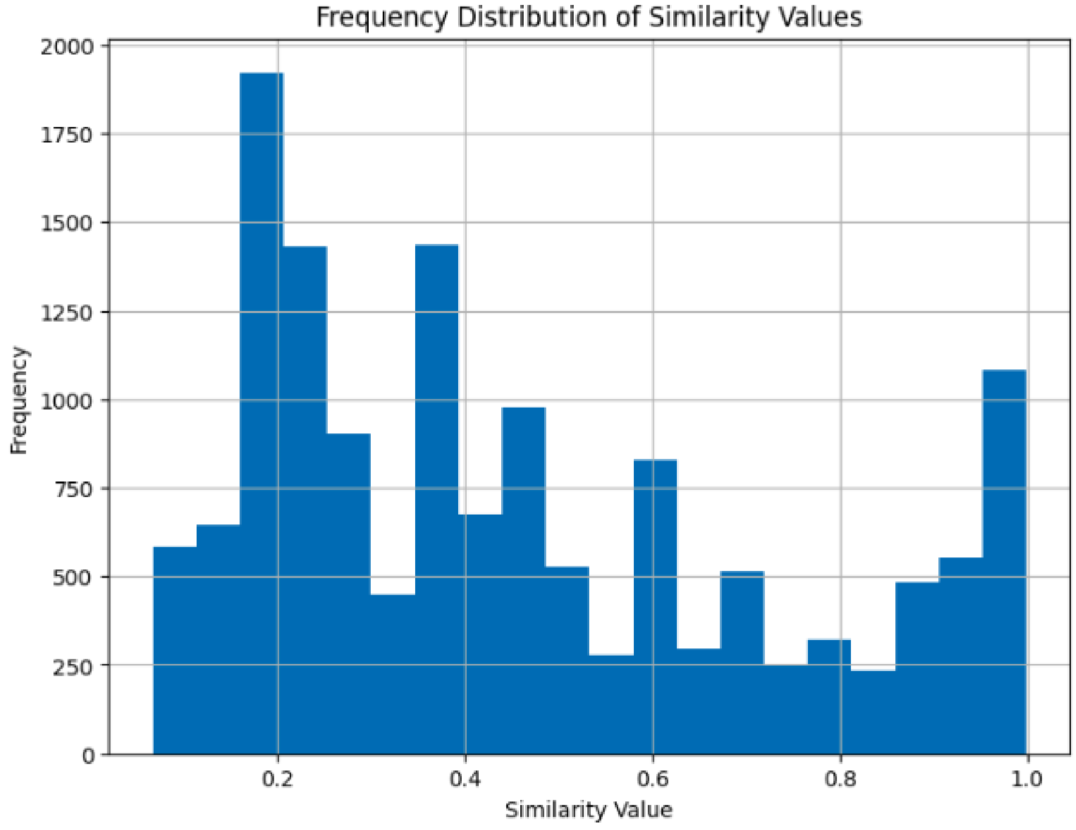
$$f(x) = \frac{1}{1 + e^{-x}}$$

(Here, our value of x was the computed similarity value between two ragas)

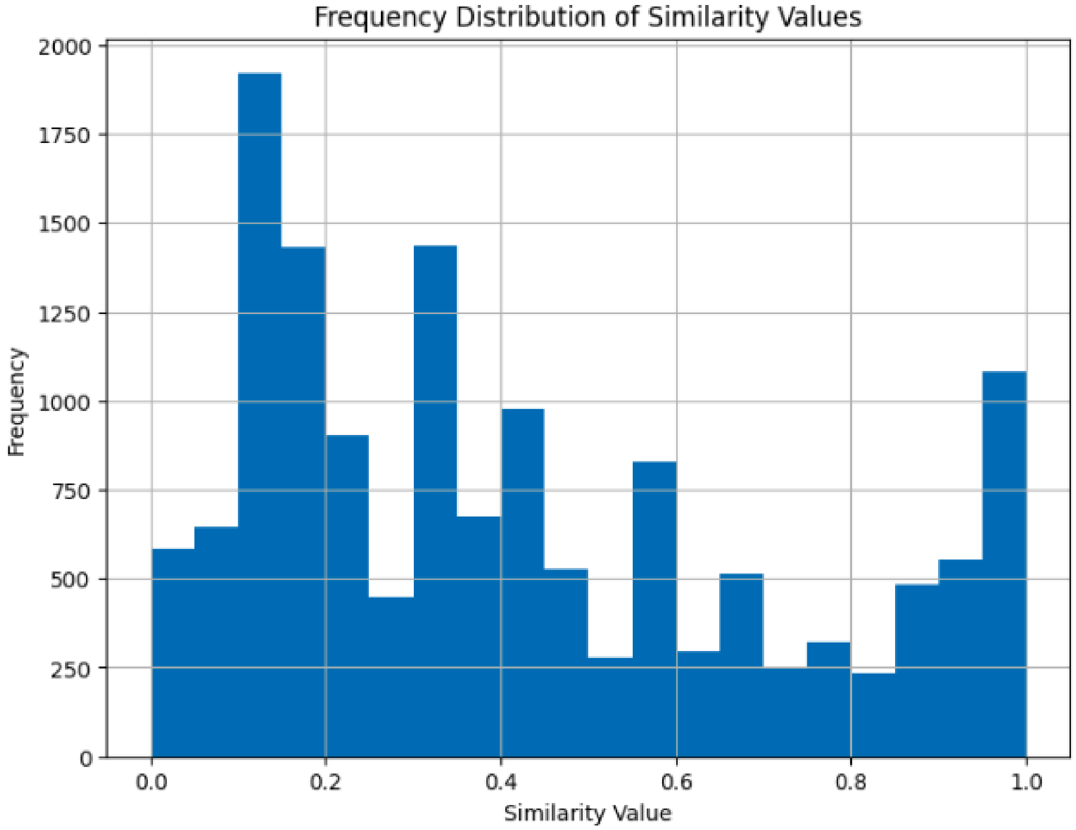
- We then normalized the values to lie between 0 and 1.

THE RAAG SIMILARITY METRIC

Output after applying the Logistic function



Output after normalizing



THE ARTIST DATASET

The Artist dataset consists of the following attributes

- Name: The name of the artist.
- Gharana: The school or tradition of Indian classical music to which the artist belongs.
- Guru: The teacher or mentor of the artist.
- Specialty: The specialization or expertise of the artist, such as vocal or instrumental.

- The dataset was scraped from the following website: https://www.swarganga.org/artistbase.php
- A lot of the artists who were present in the Performance dataset were not there in this dataset. Many artists' data had to be added manually.

APPROACHING THE ARTIST SIMILARITY METRIC

- The data we had collected on a given artist was the artist's speciality, gharana and guru (teacher).
- Gharana, a lineage of musicians which have cultivated a unique style, was one of the obvious measures in finding similarity between two artists.
- Gharana's used to be restricted to a particular region, but now as people travel quite more, different styles have spread to some extent to different regions in India.
- We decided to capture a more relevant metric – the relationship between two artists. There are three types of artists that will most likely have a very similar style –

1) Guru (Teacher)
2) Shishya (Student)
3) Guru - Bandhu (Students under the same teacher)

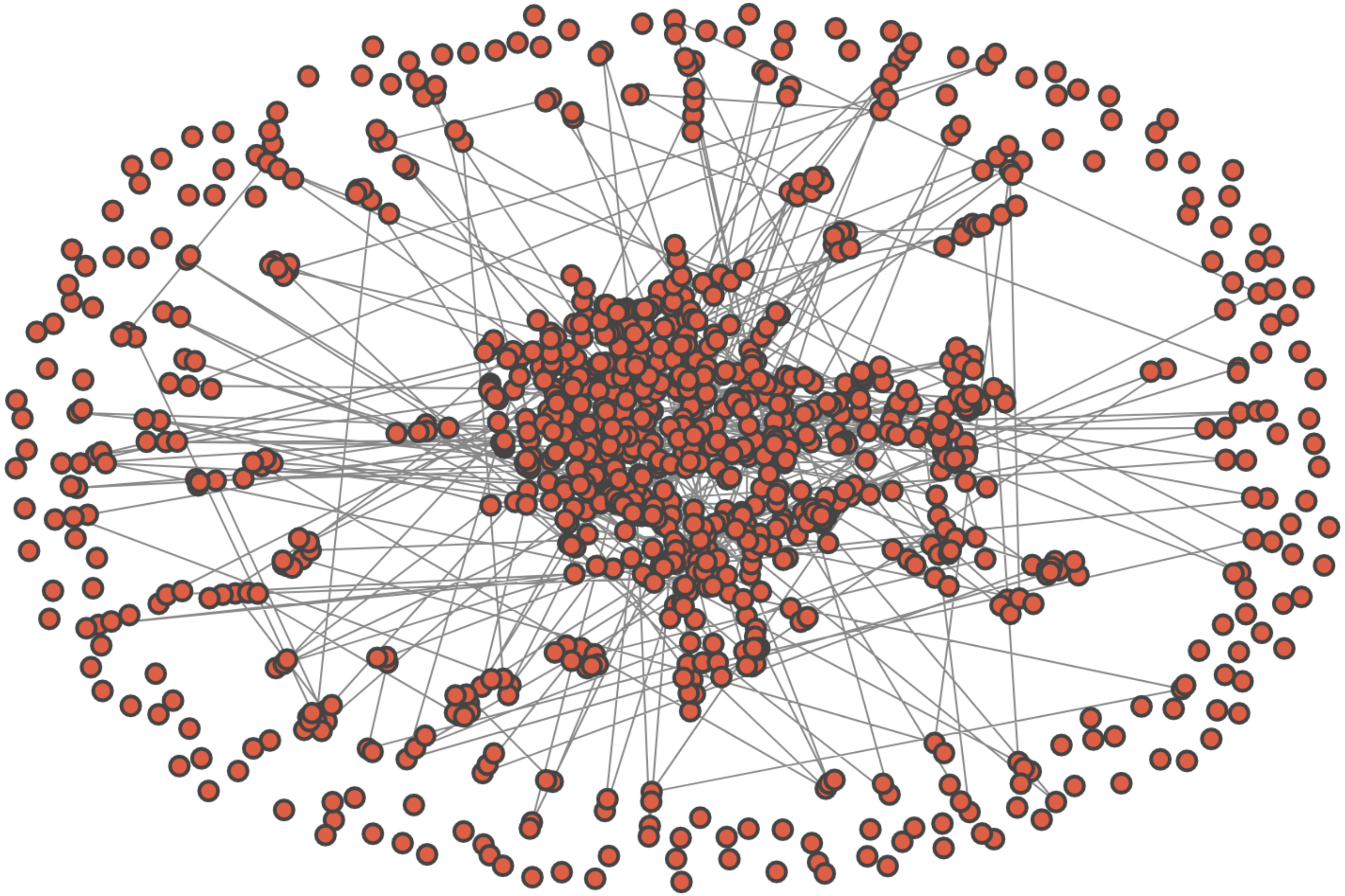
UNDERLYING ARTISTS RELATIONSHIPS

Guru - Shishya

Guru -Shishya

Guru Bandhu

ARTIST GURU RELATIONSHIP



Mean of component sizes: 4

Median of component sizes: 1

Biggest component size: 535

ARTIST SIMILARITY

- Guru-Student Relationship: The presence of shared teachers or students between artists contributes to their similarity score.
- Common Gurus: Artists sharing common gurus are deemed more similar, with the similarity score increasing based on the number of shared mentors.
- Gharana Affiliation: Similarity is enhanced when artists belong to the same gharana, representing a shared musical heritage or tradition.

ARTIST SIMILARITY

- Speciality Matching: If artists specialize in the same area, such as vocal or instrumental music, their similarity score is boosted.
- Connectedness: Additional weight is assigned if the two artists belong to the same connected component.

After quite a bit of trial and error, we set the following weights to the measures defined above:

weights = { 'guru_student': 0.1, 'gharana': 0.15, 'common_guru': 0.1, 'speciality': 0.125, 'connected': 0.075 }

MISCELLANEOUS DETAILS

- Other than the above defined similarities, here we tried to capture choices made in the performance.
- The following miscellaneous factors were taken into consideration:
  - Laya of the composition
  - Taal used in the composition
  - Instruments used in the performance
  - Form of composition

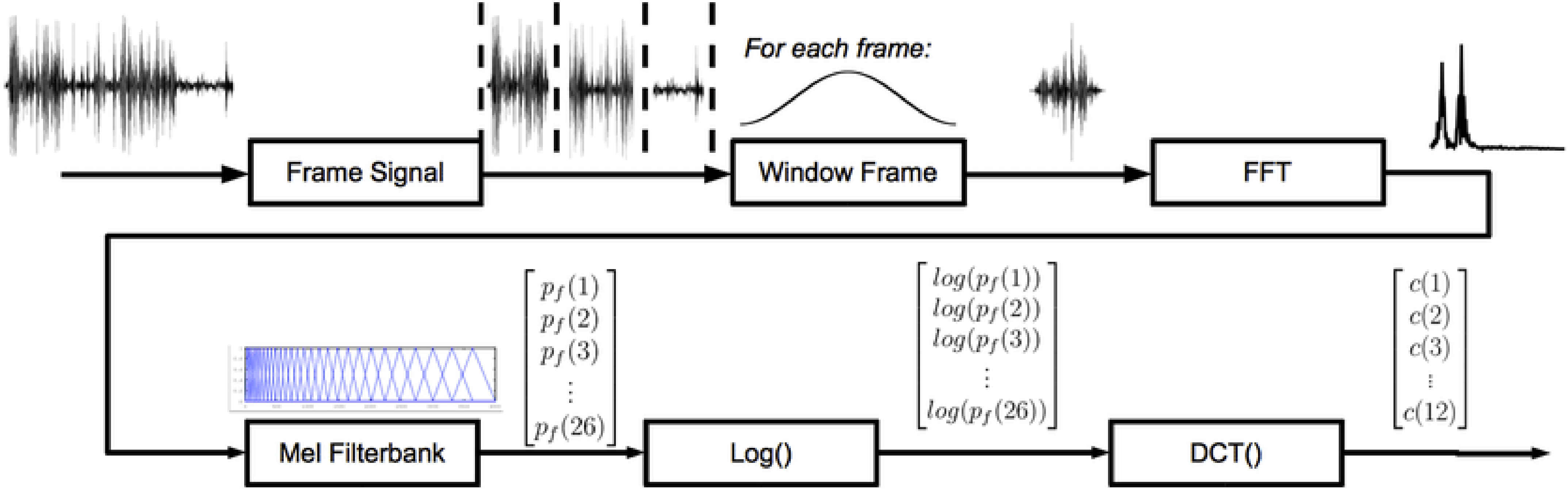
MISCELLANEOUS DETAILS SIMILARITY

- These attributes were also one-hot encoded as there are limited number of possible values for each attribute.
- This similarity, quite like the raga similarity is defined on the number of common attributes in both performances.

AUDIO SIMILARITY

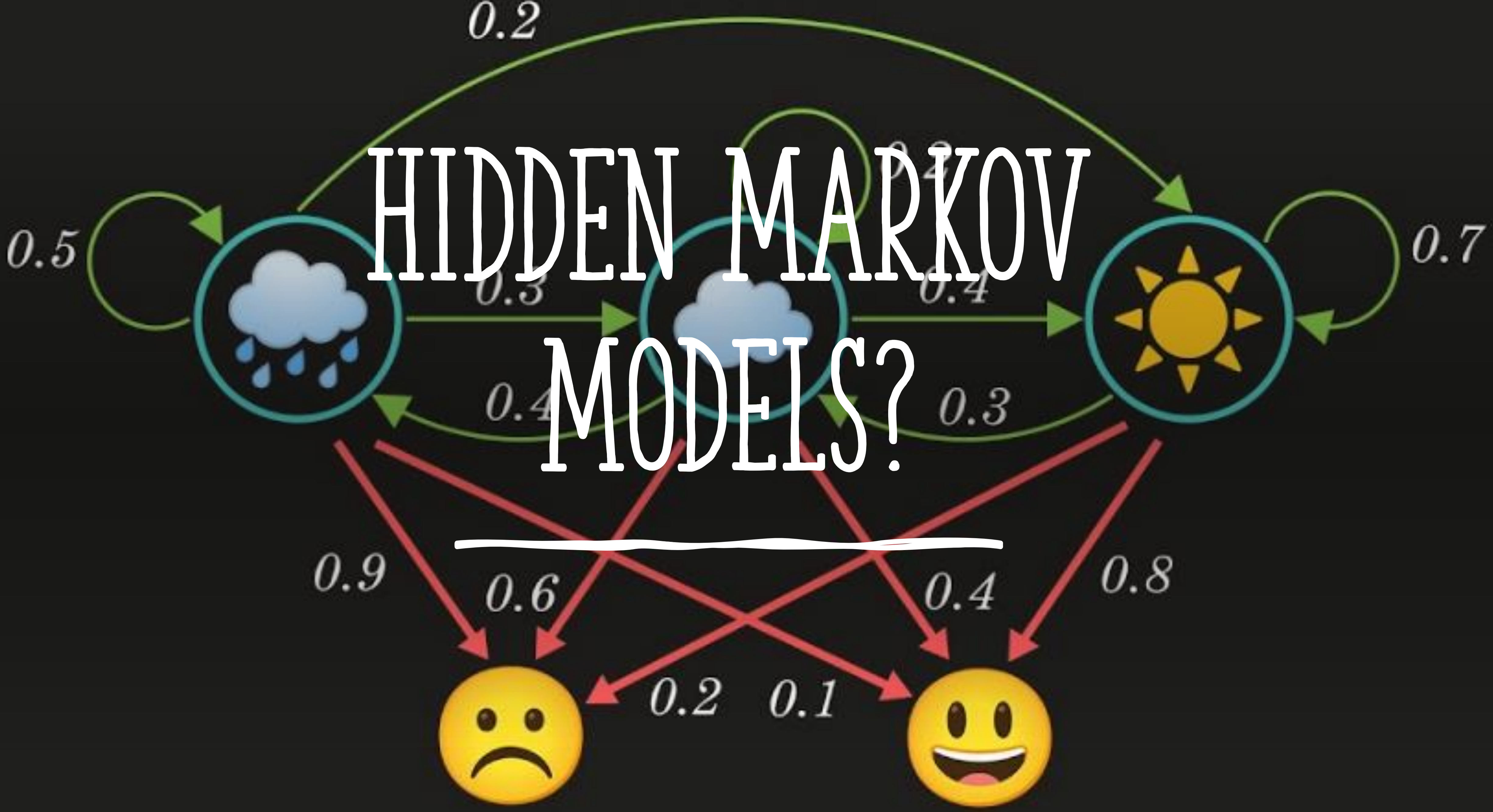
- Music preference analysis involves examining features such as pitch, timbre, rhythm, and harmony.
- Mel Frequency Cepstral Coefficients (MFCCs) encode these features as coefficient values.
- These coefficients, widely employed in speech recognition, represent properties extracted from sound waves.
- MFCC-based models provide insights into user music preferences, showcasing their versatility and effectiveness in capturing essential auditory characteristics.

MFFCS



For each frame:
Frame Signal
Window Frame
FFT
Mel Filterbank
$\begin{bmatrix} p_f(1) \\ p_f(2) \\ p_f(3) \\ \vdots \\ p_f(26) \end{bmatrix}$
Log()
$\begin{bmatrix} log(p_f(1)) \\ log(p_f(2)) \\ log(p_f(3)) \\ \vdots \\ log(p_f(26)) \end{bmatrix}$
DCT()
$\begin{bmatrix} c(1) \\ c(2) \\ c(3) \\ \vdots \\ c(12) \end{bmatrix}$

HIDDEN MARKOV MODELS?



AUDIO SIMILARITY

- MFCCs are utilized to build a music model, requiring a modeling technique to abstract piece features.
- Hidden Markov Model (HMM) is suitable for modeling as it can capture temporal patterns effectively.
- Once a model is generated, various operations like music similarity can be easily performed based on it.

Reference: *K. Kim, D. Lee, T. -B. Yoon and J. -H. Lee, "A music recommendation system based on personal preference analysis," 2008 First International Conference on the Applications of Digital Information and Web Technologies (ICADIWT), Ostrava, Czech Republic, 2008, pp. 102-106, doi: 10.1109/ICADIWT.2008.4664327.*

AUDIO SIMILARITY PIPELINE

Music Modeling with HMM: Music pieces are transformed into Hidden Markov Models (HMM) as part of the modeling process.

User Preference Analysis: Analyzing user preferences involves clustering similar music pieces that the user has listened to.

Recommendation Process: The system evaluates candidate music pieces for recommendation based on the analyzed user preferences.

Scoring System: Scores are assigned to candidate pieces based on the user's preferences, facilitating the creation of a recommendation list.

Modeling Music Similarity: MFCCs are extracted from sound wave data to build HMMs, enabling the comparison of music similarity by evaluating these models.

SIGNAL MODELING PROCESS

- Songs are divided into 2000 ms frames, and MFCCs are extracted from each segment. The 5 coefficients represent the sound waves as vectors in an 5-dimensional space.
- Hidden Markov Models are chosen for modeling each song due to their ability to handle sequential data. HMMs are derived from the series of feature vectors using the Baum-Welch algorithm.
- Each song is represented by its own HMM model, constructed from its respective series of feature vectors. This ensures that each song has a unique model capturing its specific characteristics.

MODEL SIMILARITY METRIC

- Similarity between music models implies similarity in the underlying audio content, which in turn suggests that the models can be used to predict user preferences or behavior. If a user enjoys songs that are similar to a given model, it is likely that they will also enjoy songs that are similar to the songs used to train that model.

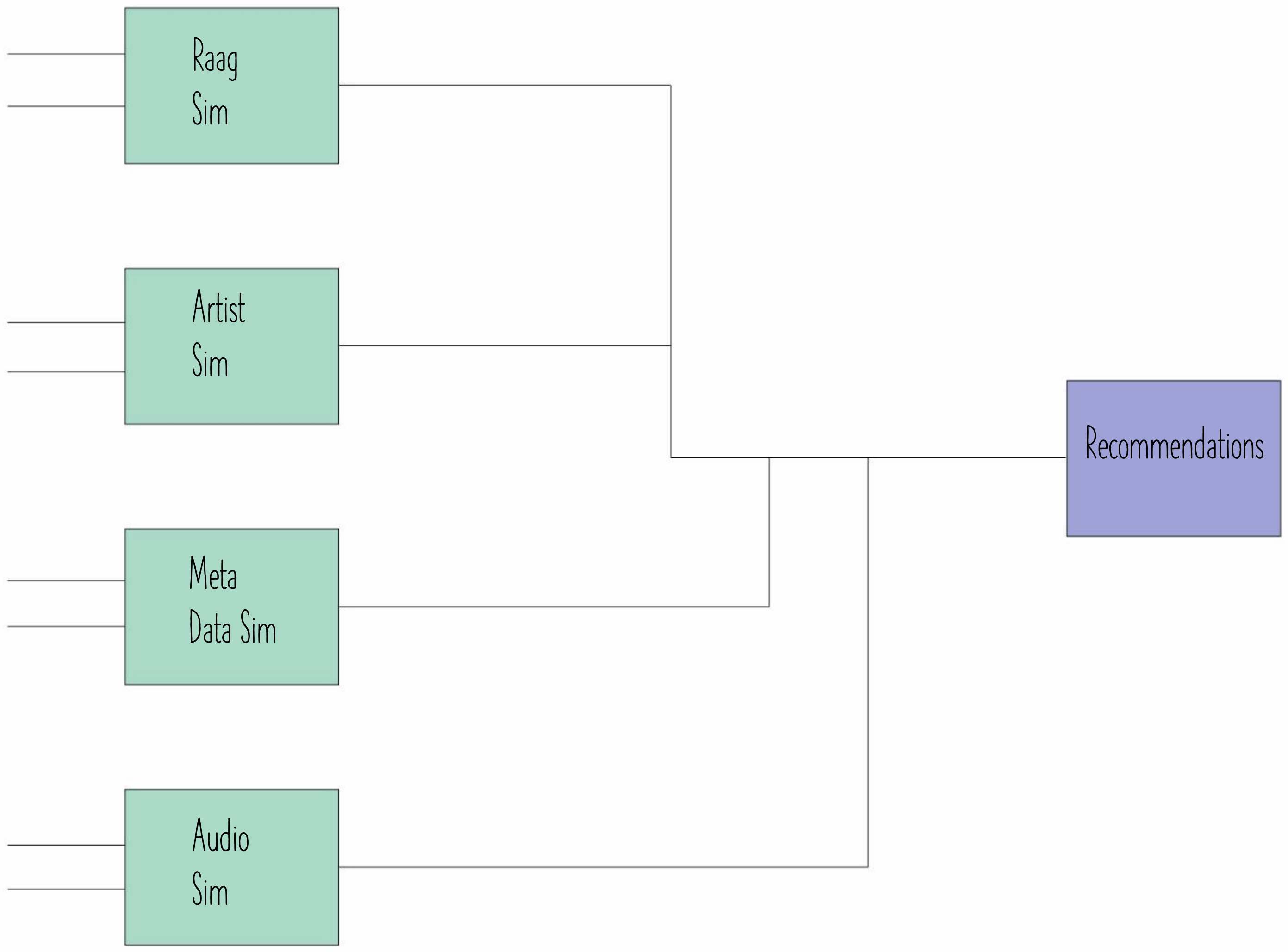
We define our similarity metric as

$$1 - \frac{-2}{\pi}\arctan\left(\frac{\pi}{2}x_{ij}k\right)$$

$$\longrightarrow x_{ij} = |\log(p(mfccs_i|M_j)) + \log(p(mfccs_j|M_i))|$$

FINAL SIMILARITY METRIC

- All 4 similarity metrics are combined into one similairty value, as a weighted average of the similarity measures.



OUR APPLICATION

- The application provides the following services:
  1. Initially, the user is asked for their raag and artist preference. Based on this, the top 10 recommendations are given
  2. The user then can select one of the following options:
     - Choose one of the recommended recordings
     - Search for a raga
     - Search for an artist

OUR APPLICATION

- Choosing one of the recommended recordings:
  - Once the user chooses a certain recording, 5 recordings are recommended. These are based on the chosen recording and the recordings that the user has listened to previously in that session.
  - A user representation is maintained by keeping a list of performances seen by the user. While recommending a recording the following similarity is considered:

    similarity = [sim(potential_recording , current_recording) + sim(potential_recording , user_representation)/5] / 2

- Searching for a raga/artist:
  - The recordings which contain the raga/artist that the user enters are listed. They are ordered based on the similarity with what the user has watched previously in that session.

CHALLENGES FACED

- Getting the data: There is very little data publicly available in this domain. Hence, we were not able to get much data and had to manually add a lot of it.
- Operating across datasets: We had to use multiple datasets and all datasets needed to be consistent with each other. Ensuring this consistency was a challenge as artist and raag names do not have standard spellings.
- Lack of computing power: Some operations we planned to do weren't feasible due to the excessive time it would take to run them
- Determining the weight to be given to each similarity factor

RECOMMENDATIONS

```
Enter an artist of your choice:  Kaushiki Chakraborty
Enter a raag of your choice:  Bhimpalasi
Recommendations based on user input:
 Id:62, Artist : Debasmita Bhattacharya , Main Instrument : Sarod , Raga : Bhimpalasi , Taal : nan , Laya : VilambitDrut
 Id:25, Artist : Ajoy Chakraborty , Main Instrument : Vocal , Raga : Piloo , Taal : Jatt , Laya : nan
 Id:9, Artist : Ajoy Chakraborty , Main Instrument : Vocal , Raga : Jog , Taal : EktaalJhaptaalTeentaal , Laya : MadhyaVi
lambit
 Id:18, Artist : Ajoy Chakraborty , Main Instrument : Vocal , Raga : Abhogi , Taal : JhaptaalTeentaalTeentaal , Laya : Ma
dhya
 Id:14, Artist : Ajoy Chakraborty , Main Instrument : Vocal , Raga : Shuddha Sarang , Taal : EktaalTeentaal , Laya : Vila
mbit
```

The above recommendations are quite good for the following reasons:

- The first recommendation is a performance of the same raga.
- Ajoy Chakraborthy is Kaushiki Chakraborthy's father and guru, and thus has a very similar style.
- The ragas recommended are also quite similar. Also most of the recommendations are vocal performances.

RECOMMENDATIONS

```
 Id:62, Artist : Debasmita Bhattacharya , Main Instrument : Sarod , Raga : Bhimpalasi , Taal : nan , Laya : VilambitDrut
 Id:25, Artist : Ajoy Chakraborty , Main Instrument : Vocal , Raga : Piloo , Taal : Jatt , Laya : nan
 Id:9, Artist : Ajoy Chakraborty , Main Instrument : Vocal , Raga : Jog , Taal : EktaalJhaptaalTeentaal , Laya : MadhyaVi
lambit
 Id:18, Artist : Ajoy Chakraborty , Main Instrument : Vocal , Raga : Abhogi , Taal : JhaptaalTeentaalTeentaal , Laya : Ma
dhya
 Id:14, Artist : Ajoy Chakraborty , Main Instrument : Vocal , Raga : Shuddha Sarang , Taal : EktaalTeentaal , Laya : Vila
mbit

Options:
1. Select one of the recommendations
2. Search for a raag
3. Search for an artist
4. Exit
Enter your choice:  2
Enter the raag you want to search for:  Todi
 Id:20, Artist : Ajoy Chakraborty , Main Instrument : Vocal , Raga : Todi , Taal : JhoomraTeentaal , Laya : Vilambit
 Id:183, Artist : Omkar Dadarkar , Main Instrument : Vocal , Raga : Todi , Taal : TeentaalTeentaal , Laya : Madhya
 Id:80, Artist : Indrani Mukherjee , Main Instrument : Vocal , Raga : Todi , Taal : nan , Laya : Vilambit
 Id:279, Artist : Shivputra Shidhharamayya Komkali 'Kumar Gandharva' , Main Instrument : Vocal , Raga : Todi , Taal : Ekt
aalTeentaal , Laya : nan

Options:
1. Select one of the recommendations
2. Search for a raag
3. Search for an artist
4. Exit
Enter your choice:
```