IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

# Music Recommendation System Based On Musical Notations

Sahil Saklani, Kshitiz Goel, Priyanshu Bhatt,
Nikhil Aswal
Department of Computer Science Engineering
*Graphic Era Deemed to be University
Dehradun, Uttarakhand, India-248001

**Abstract**—In the realm of music streaming services, algorithms play a pivotal role in helping users discover new music based on their preferences. This paper introduces a unique project—an innovative music recommendation system solely based on the sound content of songs. Unlike traditional recommendation systems, our approach explores the mathematical representation of timbre, rhythm, and style in the sonic landscape.

It can forecast the potential music a listener would be drawn to using machine learning algorithms. Furthermore, our algorithm may recommend some music based on the user's search terms. We also intend to provide a function that allows users to create their own personal libraries and authenticate themselves on our website. Based on what they have listened to, the user can also be directed to a particular song using this methodology.

Our song recommendation model aims to bridge the gap between user and his choice of taste in music, making expert guidance readily available to all. It empowers users with valuable information, about his music taste with indication about what it would like from his other music taste enhancing the overall listening experience. However, we emphasize the importance of using our platform responsibly. The goal is to provide best music taste for the user.

**Keywords:** ML Algorithm, User Interface, Song Navigation.

## I. INTRODUCTION

In the current era of rapid multimedia technology development and the ubiquity of big data, the music industry is faced with the challenge of efficiently navigating massive datasets. This paper explores the landscape of personalized music recommendation systems, acknowledging the significant growth in music data volume and user demand. The evolution of music recommendation algorithms is traced from early user preference-based models to contemporary approaches emphasizing mutual recommendations among users. The focus is on manipulating big data to predict user choices, enhancing the capability of song retrieval from the immense sea of available music resources.

The love of music is a universal sentiment. Streaming platforms like Spotify, Gaana uses sophisticated algorithms to recommend songs based on user preferences. This project is based on sonic similarities that pushes the boundaries of music recommendation by exploring the intrinsic characteristics of sound, offering users a unique way to discover music.Traditional recommendation systems consider various factors, including user preferences, collaborative filtering and metadata collaborative filtering. This project fills that gap for users who are primarily interested in the sonic qualities of music, by creating a recommendation system that focuses mainly on the sound content.

Sound, a common phenomenon, refers to the vibrations of acoustical waves that we detect through our ears. Frequency, which has a significant impact on how we perceive pitch, is the key characteristic of sound. We delve into the complexities of sound, examining the various ways in which different instruments and components contribute to the overall auditory experience.

To analyze sound data effectively, we utilize the Fourier transform to decode the intricate frequencies present in audio signals. The resulting spectrograms provide a visual representation, illustrating the amplitude of frequencies over time. Our focus on the mel scale refines this representation, aligning it with human auditory perception.

**Fourier Series**

Consider a periodic signal $g(t)$ be periodic with period T, then the Fourier series of the function $g(t)$ is defined as in equation 1,

$$g(t) = \sum_{n=-\infty}^{\infty} C_n e^{jn\omega_0 t} \quad ....(1)$$

Where, $Cn$ ( u se d i n t he e q 1 ) is the Fourier series coefficient and is given by,

$$C_n = \frac{1}{T}\int_{\frac{-T}{2}}^{\frac{T}{2}} g(t)e^{-jn\omega_0 t}dt \quad ....(2)$$

Derivation of Fourier Transform from Fourier Series

Let $x(t)$ be a non-periodic signal and let the relation between $x(t)$ and $g(t)$(used in eq2) is given by,

$$X(t) = \lim_{T\to\infty} g(t) \quad .....(3)$$

Where, T is the time of the periodic signal $g(t)$.

By rearranging eq. (2), we get,

$$TC_n = \int_{\frac{-T}{2}}^{\frac{T}{2}} g(t)e^{-jn\omega_0 t}dt$$

The term $Cn$ represents the magnitude of the component of frequency nω0.

Let nω0 = ω at $T \to \infty$. Then, we have,

$$\omega_0 = \frac{2\pi}{t}|_{T\to\infty} \to 0$$

Thus, the discrete Fourier spectrum becomes continuous and hence the summation becomes integral and [$g(t) \to x(t)$]. Therefore, at $T \to \infty$,

$$TC_n = \lim_{T\to\infty}\int_{\frac{-T}{2}}^{\frac{T}{2}} g(t)e^{-j\omega t}dt$$

$$\Rightarrow TC_n = \int_{-\infty}^{\infty}[\lim_{T\to\infty} g(t)]e^{-j\omega t}dt \quad .....(4)$$

From equations (3) & (4), we have,

$$\Rightarrow TC_n = \int_{-\infty}^{\infty} x(t)e^{-j\omega t}dt = X(\omega) \quad ....(5)$$

Therefore, the Fourier transform of the non-periodic signal is.

$$X(\omega) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t}dt \quad ....(6)$$

The function X(ω) represents the frequency spectrum of function $x(t)$ and is called the spectral density function.

## II. MOTIVATION

The evolving nature of music consumption and the unrealized possibility of utilizing rich musical structures for individualized suggestions are the driving forces behind the research study on song recommendation based on musical notations. This research is motivated by several primary goals:

- **Untapped Potential of Notation in Music**:
  A composition's essence is captured in musical notation, which offers a thorough and organized depiction of all of its constituent parts. Notations have received little attention in the field of recommendation systems, despite their importance in the composition of music.
- **Diverse User Preferences Beyond Audio Features**: Although most traditional recommendation algorithms concentrate on aural characteristics like key, tempo, and genre, there is

a large community of music lovers whose choices are based on the subtleties of musical structure.

- **Personalized Music Experience**:
  Now, music consumption has evolved into a more personalized and selective process. Using advanced technology, the recommendation system analyzes frequency and pitch to identify user preferences accurately. As a result, each individual receives music suggestions tailored to their unique taste, resulting in a highly personalized music listening journey.
- **Enriching User Experience and Engagement:**

  Understanding and valuing the intricate details of musical notations exist enhancing the overall user experience by offering suggestions that better match an individual's distinct musical preferences.

  - Grammatical Mistakes:
  - exist enhancing

## III. RELATED WORKS AND RESEARCH GAP

### *A. LITERATURE SURVEY*

Many studies have made contribution to the interest of music recommendation systems and the incorporation of musical notations. Fundamental concept like Collaborative filtering has emerged in music recommendation systems, utilizing user behavior and preferences to create personalized music suggestions. This technique has greatly increased user satisfaction through customized music recommendations [1]. Furthermore, the idea of prodviding top-n recommendations has been widely studied due to its importance in maintaining user engagement on music platforms. Effective methods for generating top-n recommendations, including evaluating precision and recall, are essential for creating personalized music playlists [2].

Concept of content-based recommendation systems for music, recent advancements delve into the specifics of sound content and its mathematical representations. A notable contribution is the work of Aucouturier and Pachet, who suggested a music recommendation system based on timbre, modeling the perceptual similarity of songs using Gaussian Mixture Models for timbral descriptors.

Expanding content-based recommendation systems, which concentrate on the inherent qualities of items, have gained popularity. In the music context, Tzanetakis and Cook [3] developed a content-based music recommendation system based on audio characteristics. They utilized signal processing techniques to extract features like timbre, rhythm, and pitch, laying the groundwork for further investigations into the acoustic landscape[4].

In a similar vein, Lerch [5] explored the utilization of audio features for measuring music similarity, incorporating elements like spectral contrast and key strength. This research illustrate that how overall similarity between songs is influence by the distinct sounds characteristics.

## IV. METHODOLOGY

Music recommendation system suggests personalized songs in response to users' ever-more-individualized demands and the fast growing volume of information available. One of the most popular uses of personalized recommendation at the moment is music suggestion based on user preferences. The significance of recommendation systems is demonstrated by the fact that personalized recommendation is essential on all significant music websites, both domestically and internationally. Recommendation algorithms are classified into two primary categories based on their development from the original basis: content recommendation algorithms and context recommendation algorithms.

To kickstart the project, we leveraged the Spotify Public API to scrape comprehensive song information. This included essential details such as track name, artist, and album, which served as the foundation for our recommendation system.
The focus of our methodology was to capture the sonic essence of each song. To achieve this, we extracted 30second mp3 previews for every track. These previews provided concise yet representative segments of the songs, ensuring computational efficiency without compromising on the richness of the audio content.

To enable the application of deep learning techniques, we converted the waveform data from the mp3 previews into Mel spectrograms as described in figure 1. The Fourier transform played a crucial role in extracting the frequencies present in each audio segment. The resulting spectrograms were then transformed into the mel scale to better represent the perceived distance between frequencies.

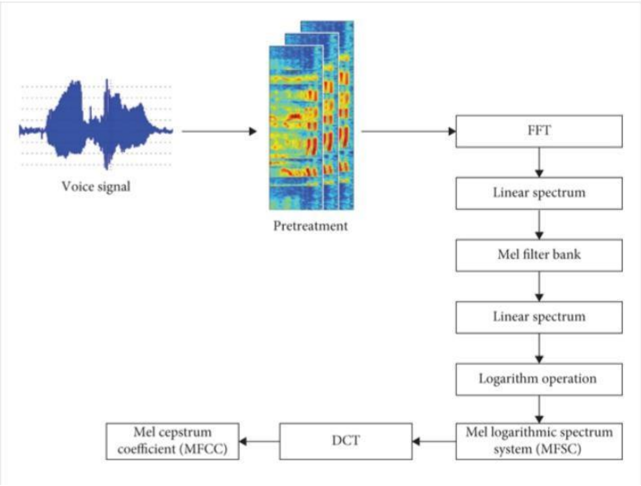


Figure 1 MFCC feature parameter extraction process[6]

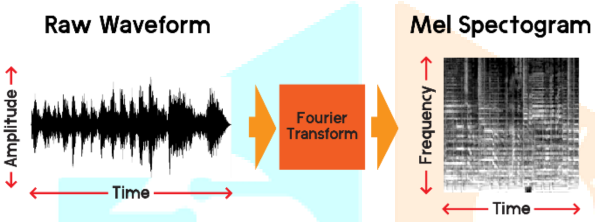


Figure 2 Raw Waveform to Mel Spectogram[7]

The cornerstone of our recommendation system is the autoencoder neural network. The architecture was carefully designed to balance the complexity required for feature extraction and the necessity for computational efficiency.

One noticeable aspect of our autoencoder model is the unique split encoder structure. Instead of using traditional two-dimensional convolutional layers, we decided to utilize one-dimensional convolutional layers. This change enabled us to incorporate two distinct encoders—one focusing on the time axis and the other on the frequency axis. Through training on X and its transpose, we captured a more comprehensive representation of the audio content.

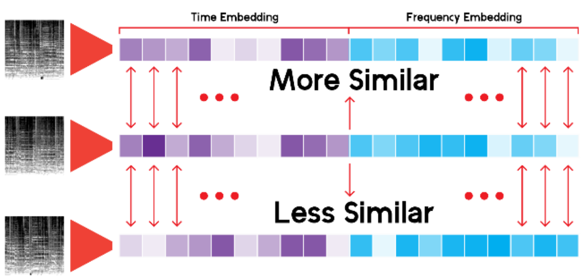


Figure 2 Grouping with similarities

The training process of the autoencoder focuses on minimizing the mean squared error between the original Mel reconstructed output. The model went through 20 epochs, with the training set consisting of 125,440 images and the validation set having 2,560 images. The final mean squared error impressively decreased to 0.0037 for both the training and validation sets.For the creation of the latent space, we introduced a customized 'LatentSpace' class. This class streamlined the inference procedure using the trained

encoder, resulting in a 256-dimensional vector for each Mel spectrogram. This vector encapsulated the fundamental characteristics defining the sonic content of each piece.

To venture into the abstract domain of the latent space, we employed dimensionality reduction methods. UMAP, a robust tool similar to T-SNE, was chosen to convert the high-dimensional vectors into a two-dimensional space.

The recommendation process hinged on cosine similarity in the latent space. For a given query, we utilized the Spotify API to retrieve the mp3 preview of the first result. After processing the preview through the encoder, we compared its vector representation with the vectors of the entire track database using cosine similarity. The top ten highest scores were then sorted and returned as recommendations.

## V. RESULT AND DISCUSSION

Our project of sound content-based recommendation systems opens up a fascinating avenue for understanding the nuances of musical expression. By bypassing conventional metadata such as genre labels or artist information, Sonufy delves directly into the essence of music as perceived through auditory sensations. This approach not only challenges traditional paradigms of music categorization but also underscores the richness and diversity of sonic landscapes across different genres, eras, and cultures.

One of the notable strengths of Sonufy lies in its ability to capture the subtleties of musical timbre and rhythm through mel spectrograms and autoencoder models. By compressing complex audio signals into compact feature vectors, Sonufy's recommendation system transcends language and cultural barriers, offering a universal framework for exploring musical similarities and connections. This universality is particularly evident in the project's emphasis on the continuous spectrum of musical language, where genre. Distinctions blur, and individual tracks find their place based solely on sonic resemblance.

However, Sonufy also grapple with certain challenges inherent in. sound-based recommendation systems. The reliance on sound content alone may overlook contextual factors such as lyrical themes, cultural significance, or listener preferences, which play crucial roles in shaping musical experiences. Additionally, the project acknowledges the limitations imposed by the availability and diversity of data, with certain genres or artists underrepresented in the dataset. These limitations underscore the importance of ongoing data collection and curation efforts to ensure the robustness and inclusivity of the recommendation system.

Furthermore! Sonufy's emphasis on the mathematical representation of music opens up intriguing possibilities for interdisciplinary exploration. By bridging the gap between music theory. Signal processing, and machine learning, Sonufy exemplifies the convergence of art and science in understanding and appreciating music. This interdisciplinary. Approach not only enriches our understanding of musical phenomena but also inspires new avenues for research and innovation in fields ranging from artificial intelligence to musicology.

We have been endlessly searching through this recommendation system, and I am satisfied that the model can pick out very interesting connections between different but also similar musical sounds. Here are some of my conclusions:

The recommendations are more connected than can be heard.
What I mean by this is that the model is making recommendations based on the sound content in each song, but it is not listening to the song. It creates a mel spectrogram and makes a mathematical comparison.

Sometimes the system will make a recommendation for a song based on its age. If a song was recorded a long time ago, those particular frequencies of the recording material or equipment will be picked up by the model, and display the results.

Also, the model is very good at picking up voice or particular instruments. Because of this, if a song has a lot of talking or talking-singing, it might only recommend spoken word tracks. Also, if there is a lot of distortion in a song, it might recommend rain sounds or bird songs.

Some genres or artists are underrepresented in the data. Some track previews are unavailable in the Spotify API, as pointed out in my initial EDA. Therefore, their contribution to the model is also missing and won't be a recommendation when they might be a perfect fit for one. For example, there are no songs by James Brown, the Beatles, or Prince. Needs more data.

The system is using over 278.000 previews to make recommendations, and that's still not enough. Looking at the UMAP projection for all tracks, there is a lot of continuity in the data, but there are some holes, Ideally, the system could use a lot more data to draw on. It is a feature of a recommendation system and not an ensemble model.
Also, music has no barriers. Most times, when querying a song in the recommendation system, results will come from all different eras and all different places. Since none of the metadata of a song is an input for the autoencoder, results are based on their sonic similarity, and nothing more.

In conclusion, Sonufy represents a pioneering effort to reimagine music recommendation systems through the lens of sound content analysis. By harnessing the power of advanced algorithms and mathematical frameworks, Sonufy offers a glimpse into the future of personalized music discovery experiences. While the project grapples with certain challenges and limitations, its overarching vision of music as a universal language transcends boundaries and opens up new horizons for exploration and appreciation. As the project continues to evolve and integrate additional features and data sources, it promises to shape the future of how we discover, explore, and connect with music in the digital age.

## REFERENCES


1. Herlocker, J. L., Konstan, J. A., Borchers, A., & Riedl, J. (2019). "An Algorithmic Framework for Performing Collaborative Filtering." Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval.
2. Cremonesi, P., Koren, Y., & Turrin, R. (2016). "Performance of recommender algorithms on top-n recommendation tasks." Proceedings of the Fourth ACM Conference on Recommender Systems.
3. Tzanetakis, G., & Cook, P. (2020). Musical genre classification of audio signals. IEEE Transactions on Speech and Audio Processing, 10(5), 293-302.
4. Aucouturier, J. J., & Pachet, F. (2013). Representing musical genre: A state of the art. Journal of New Music Research, 32(1), 83-93.
5. Lerch, A. (2012). An introduction to audio content analysis: Applications in signal processing and music informatics. John Wiley & Sons.

6 https://www.hindawi.com/journals/misy/2022/3387598/fig2/

7. https://www.geeksforgeeks.org/music-recommendation-systemusing-machine-learning/