IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 7.39 ⁜ Vol. 11, Issue 6, June 2022

DOI: 10.17148/IJARCCE.2022.11626

A Hybrid Approach for Modern Music Recommendation System using Neural Networks and Feature Level Fusion

Dr Maria Manuel Vianny[1], Polireddy Nishith Reddy[2], Ramu Velaguri[3], Busi Yashwanth Reddy[4], Diran Srinath Reddy Dodda[5]

[1]Assistant Professor (CS-DS), JAIN University, Bengaluru Karnataka

[2-5]B.Tech (CS-DS), JAIN University, Bengaluru Karnataka

Abstract: The custom music recommender supports users' favorite songs, which are stored in a huge music database. To predict only the user's favorite songs, the management of the user's preference information and the genre rating is required. In our study, a very short feature vector obtained from a low-dimensional projection and already developed audio features is used for the music genre classification problem. We apply a metric distance learning algorithm to reduce the dimensionality of the feature vector with little performance degradation. We propose the system through the automatic management of user preferences and gender classification in the personalized music system. This Recommender System uses a feature level fusion to combine multiple perspectives and gives an outcome that suits all types of users. The performance of this system is compared with existing legacy system.

INTRODUCTION:

Ever wonder how Spotify recommends songs and playlists based on your listening history? Wondering how Spotify manages to find songs that sound similar to the ones you've already listened to? Interestingly, Spotify has a web API that developers can use to get audio characteristics and metadata about songs, such as the song's popularity, tempo, volume, key, and year of release. We can use this data to create music recommendation systems that recommend songs to users based on the audio capabilities and metadata of the songs they have been listening to. One of the most popular features of Spotify is Discover Playlist, a playlist created each week based on the user's listening habits. As a Spotify user, I have found that these playlists are sometimes very accurate, but sometimes I am not completely satisfied with my custom selections. I wanted to take a closer look at how Spotify and other music companies make these recommendations. With online music streaming becoming the dominant medium for listening to your favorite songs, music streaming services can now collect vast amounts of data about their customers' listening habits. These streaming services like Spotify, Apple Music or Pandora use this data to make recommendations to their listeners. These music recommendation systems are part of a broader class of recommendation systems that filter information to predict a user's preferences for a particular item. Think of Netflix movie recommendations or Pandora radio. This great Wikipedia article on the subject divides recommender systems into two classes that also apply to music-specific recommender systems. These two classes or approaches to recommender systems are collaborative filtering and Content based filtering.

LITERATURE SURVEY:

According to the paper, customized tune advice gadget (PMRS) primarily based totally at the convolutional neural networks (CNN) method. The CNN method classifies tune primarily based totally at the audio sign beats of the tune into one-of-a-kind genres. In PMRS, we suggest a collaborative filtering (CF) advice set of rules to integrate the output of the CNN with the log documents to endorse tune to the person. The log record carries the records of all customers who use the PMRS. The PMRS extracts the person’s records from the log record and recommends tune beneath Neath every style. We use the million-track dataset (MSD) to assess the PMRS. To display the running of the PMRS, we evolved a cellular application (an Android version). We used the self- belief rating metrics for one-of-a-kind tune style to test the overall performance of the PMRS. The customized tune recommender helps the user favorite songs saved in a massive tune database. In order to are expecting most effective person-favorite songs, handling person choices records and style class are necessary. In our study, a completely brief characteristic vector, acquired from low dimensional projection and already evolved audio features, is used for tune style class problem. We implemented a distance metric getting to know set of rules with the intention to lessen the dimensionality of characteristic vector with a bit overall performance degradation.

ISO 3297:2007 Certified ⁜ Impact Factor 7.39 ⁜ Vol. 11, Issue 6, June 2022

DOI: 10.17148/IJARCCE.2022.11626

We suggest the gadget approximately the automated control of the person choices and style class in the customized tune gadget.

METHODOLOGY:

Our goal is to construct a complete music recommendation system, which overcomes the limitations of the previous systems. The simplest service of music recommendation is accomplished by using the keyword-based filtering approach to notify users when appropriate music objects arrive. The mechanism for this notification service is described as follows. For an incoming music object, the corresponding description is manually attached to the music object, such as the music genre, title, and composer. The users are required to specify their preferences in music terms. The users' preferences will be compared with the descriptions of the music objects. Once matched, the system will send a notification of the matched music objects to the users. However, the task of manual description takes tremendous efforts when available music objects grow explosively. Therefore, this kind of music recommendation is impractical in the real world. Ringo in Sharad Anand and Maes (1995) is a pioneer collaborative music recommendation system. In Ringo, each user is requested to make ratings for some music objects. These ratings constitute the personal profile. Several algorithms are proposed to measure the similarity between two users' profiles. For collaborative recommendation, only the ratings of the users whose profiles are similar to the target user are considered. Whether a music object will be recommended is then based on the weighted average of the ratings considered. On the contrary, a content-based personalized music filtering system is introduced in Kuo and Shan (2002). The system learns the user's preferences by mining the melody patterns from the music objects in the user's access history. Using these melody patterns, a melody preference classifier is then constructed for each user. An incoming music object will be recommended to the user if it is classified into the preferred class. In this system, only the pitch information is considered for feature extraction. Ignoring other information, e.g., duration and loudness, provided in the music objects limits the system to deal with other kinds of user preferences. For example, the user may prefer the music objects with slower tempo. However, the system in Kuo and Shan (2002) cannot actually reflect this preference for the users during recommendation. Moreover, due to the inherence of the content-based filtering approach, this system cannot provide any surprising recommendation results as in the collaborative approach. In this paper, we propose an alternative way of music recommendation, which overcomes the limitations of the previous works. Instead of textual descriptions, we fully consider the perceptual properties of music objects, such as pitch, duration, and loudness, which can be directly extracted from the music objects. Based on the various features derived from these perceptual properties, music objects are then grouped automatically. For users, their interests and behaviors are derived from the access histories and recorded in user profiles for further user grouping. Based on the favorite degrees of the users to the music groups, and the user groups they belong to, three recommendation methods are proposed to satisfy different needs of the users. Both the perceptual properties of the music objects and the derivation of user interests and behaviors are considered in the proposed system. As a result, our recommendation system can be more satisfactory to users than the previous systems. The rest of this paper is organized as follows.

RECOMMENDER SYSTEM:

The Music Recommendation System (MRS) is a website which provides the service of music recommendation based on music grouping and user grouping. The music objects in the database of MRS, as well as the incoming music objects, are candidates for music recommendation. As shown in figure 1, the system consists of seven function blocks, namely, the track selector, the feature extractor, the classifier, the profile manager, the recommendation module, the interface, and the database. When a new music object is inserted in the database of the MRS, it goes through the track selector and feature extractor. Using the features extracted from a music object, a dynamic classifier is designed for music grouping. Each incoming music object is properly assigned to certain music group by the classifier. The profile manager is implemented for the purpose of updating the access histories of the users. When the user accesses a music object from the list of music objects or the recommendation results, the profile manager records the information of the music object in the user's access history. Using the information recorded by the profile manager, three recommendation mechanisms, i.e., the content-based (CB) method, the collaborative (COL) method, and the statistics-based (STA) method, are designed for different users' needs.

3. CONVOLUTIONAL RECURRENT NEURAL NETWORKS

In this section, we will explain our research steps in detail, with the focus to describe Convolutional Recurrent Neural Networks (CRNNs) architecture.

3.1. Collecting Music Data

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 7.39 ⁜ Vol. 11, Issue 6, June 2022

DOI: 10.17148/IJARCCE.2022.11626

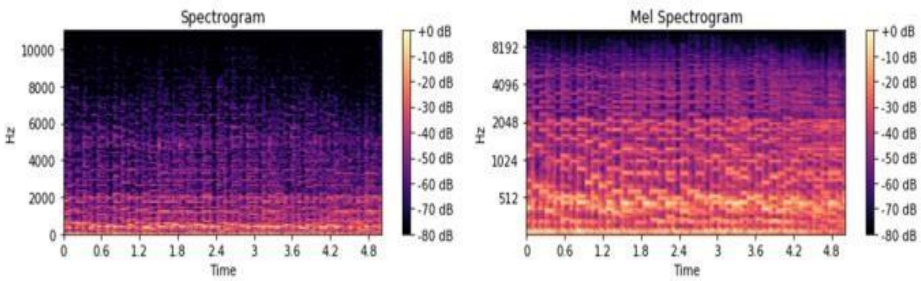
Data is collected by downloading available datasets. The dataset comes from the Free Music Archive (FMA) 14 music collection that is legal to download. For the research, it is used fma_medium dataset which data volume is 22 GB and total collections are 25,000 music in mp3 format. This dataset is imbalance with 16 genres of music and duration of each music is 30 seconds. For our research, we only used a part of the FMA dataset by choosing the good quality records and maintaining the balance of data for each genre.

3.2. Data Cleaning and Selection

First, data that has been collected is cleaned up by discarding music with low audio quality and music which has wrong genre label. The cleaning process is done easily by sorting the music based on the creator, because music with the same creator has the same audio quality in the dataset. We also removed the music which formed of two or more genres, such as combination of pop and folk. Finally, we chose only seven genres for maintaining the data balance, that is: classical, electronic, folk, hip-hop, instrumental, jazz and rock. The following table is the amount of data for each music genres after the process of data cleaning and data selecting.

3.3. Audio Preprocessing

We need a suitable audio representation as input for neural networks architecture. Therefore, data in the form of audio signal is converted into spectrogram image. First, to do this conversion, the audio signal is processed with a certain sampling rate. Next, we applied Short Time Fourier Transform (STFT) at each sampling rate with a certain window length and hop length, thus the frequency of the audio signal can be calculated at a specific time. To smooth the frequency and avoid the spectral leakage, a window function is used. Finally, the results from STFT are converted into Mel scale. Melspectrograms is considered the most effective representation compared to other audio representations.



3.4. Modeling Neural Networks:

To classify the music genres, we used two architectures, that is: Convolutional Neural Networks (CNNs) and Convolutional Recurrent Neural Network (CRNNs). CRNNs is used because the model can extract the important features for the prediction results. Not only looking at frequency features on the spectrogram, CRNNs also can look at time sequence patterns. Finally, feature vectors that produce before the classification layer can be used as a basis for recommendations.

The following is an explanation of the training and testing stages on the CNNs and CRNNs models. First, we divided Mel-spectrogram data into training, validation, and testing dataset, which the proportion is 8:1:1 respectively. Because the model is binary classification, labelling is done by using one hot encoding. It can be done by giving a value of 1 to a chosen music genre and 0 to other genres.

The CRNNs architecture used two layers of RNN with Gated Recurrent Units (GRU) to summarize 2D temporal patterns from the results of four CNN layers. CNNs was first performed on this model, which was used for extracting local features. The results of sub-sampling are feature maps with a large N × 1 × 15 (number of feature maps × frequency × time) that will be used for two layers of RNNs. Figure 4 shows illustration of CRNN structure with dimensional change for each layer.

CONCLUSION:

First, music recommender system should consider the music genre information to increase the quality of music recommendations. Second, CRNNs that considers both the frequency features and time sequence patterns has overall better performance. It indicates the effectiveness of its hybrid structure to extract the music features. Based on our analyses, we can suggest for future research to add other music features in order to improve the accuracy of the recommender system, such as using tempo gram for capturing local tempo at a certain time.

ISSN (O) 2278-1021, ISSN (P) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⋇ Impact Factor 7.39 ⋇ Vol. 11, Issue 6, June 2022
DOI: 10.17148/IJARCCE.2022.11626

REFERENCES:

[1] https://sci-hub.ru/https://link.springer.com/article/10.1007/s00500-020-05364-y
[2] https://sci-hub.ru/https://www.sciencedirect.com/science/article/pii/S0140366419311399
[3] https://sci-hub.ru/https://dl.acm.org/doi/abs/10.1145/3383583.3398623
[4] https://sci-hub.ru/https://dl.acm.org/doi/abs/10.1145/3419635.3419669
[5] https://www.sciencedirect.com/science/article/pii/S1877050919310646
[6] https://www.hindawi.com/journals/complexity/2021/5579086/
[7] https://dl.acm.org/doi/abs/10.1145/3365843
[8] https://dl.acm.org/doi/abs/10.1145/3407190
[9] https://www.mdpi.com/2076-3417/10/15/5324
[10] https://www.mdpi.com/2078-2489/11/9/439
[11] https://link.springer.com/chapter/10.1007/978-3-030-72113-8_3
[12] https://elisabethlex.info/docs/2021fntir-psychology.pdf

