International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 12 | December - 2022 Impact Factor: 7.185 ISSN: 2582-3930

Music Genre Classification Systems – A Computational Approach

Srishti Gupta[1*]
Computer Science and Engineering
Panipat Institute of Engineering and Technology,Kurukshetra University
Haryana India

Srishti Miglani[2*]
Computer Science and Engineering
Panipat Institute of Engineering and Technology, Kurukshetra University
Haryana, India

Dr. S.C.Gupta[3]
HOD, Department of Computer Science and Engineering, Panipat Institute of Engineering and Technology, Haryana India

Abstract

A music genre is a conventional category that identifies some pieces of music as belonging to a shared tradition or set of conventions. It is to be distinguished from musical form and musical style. Music can be divided into different genres in many different ways. The popular music genres are Pop, Hip-Hop, Rock, Jazz, Blues, Country, and Metal. Categorizing music files according to their genre is a challenging task in the area of music information retrieval (MIR). Automatic music genre classification is important to obtain music from a large collection. It finds applications in the real world in various fields like automatic tagging of an unknown piece of music (useful for apps like Saavan, Wynk, etc.). Companies nowadays use music classification, either to be able to place recommendations to their customers or simply as a product. Determining music genres is the first step in the process of music recommendation. Most of the current music genre classification techniques use machine learning techniques.

I. Introduction

Data mining is a technique for extracting and detecting patterns in huge data sets that combines machine learning, statistics, and database systems. Data mining is an interdisciplinary subject of computer science and statistics with the purpose of extracting information from a data collection using intelligent methods and transforming it into an intelligible structure for subsequent use. With numerous real-world applications, genre classification is an important task. As the amount of music released on a daily basis continues to rise, particularly on internet platforms like Soundcloud and Spotify (a 2018 estimate suggests that tens of thousands of songs were released every month on Spotify),the need for accurate meta-data for database management and search/storage purposes rises in tandem. The ability to rapidly identify songs in any given playlist or collection by genre is a critical feature for any music streaming/purchasing business, and the statistical analysis potential provided by accurate and full labeling of music and audio is virtually endless. A specific dataset can be used to classify genres, and the model can be trained tomatch the dataset

In today's world, an individual's music collection generally contains hundreds of songs, while a professional collection normally contains tens of thousands of music files. Music databases are incessantly gaining a reputation in relation to specialized archives and private sound collections. With improvements in internet services and increases in network bandwidth, there is also an increase in the number of people accessing the music database. Dealing with extremely large music databases is exhausting and time-consuming. Since manually classifying each track of a large music database according to its genre is a tedious task, Machine Learning Techniques to perform Automatic Music Genre Classification are used.

The following is the rest of the document. Section 2 discusses the literature that was analyzed for the project. The proposed approach for music genre classification and recommendation is discussed in Section 3. In Section 4, the emphasis is on implementation. Section 5, which outlines the reach and probable directions, brings the investigation to a close.

II. LITERATURE REVIEW

Music Genre Classification is an area that has attracted the interest of many researchers. This section will provide details about some of the research work already done in this field. Vishnupriya S and K Meenakshi [1] have proposed a Neural Network Model to perform the classification. Tzanetakis and Cook [2] pioneered their work on music genre classification using a machine-learning algorithm. They created the GTZAN dataset which is to date considered a standard for genre classification. Changsheng Xu et al have shown how to use support vector machines (SVM) for this task. Matthew Creme, Charles Burlin, and Raphael Lenain from Stanford University have used 4 different methods to perform the classification. They have used Support Vector Machines, Neural Networks, Decision Trees, and K-Nearest Neighbours methods to perform classification. Tao shows the use of restricted Boltzmann machines and arrives at better results than a generic multilayer neural network by
generating more data out of the initial dataset, GTZAN. We show four types of qualities that are used to categorize music genres. In order to execute accurate and informed categorization, the right set of features must be chosen.

1) A characteristic depending on the magnitude
2) Features dependent on time
3) Features based on pitch, and
4) Features based on the chord progression.

The GTZAN dataset was used in each of the three research done. This GTZAN dataset is an ensemble of 1000 thirty-second-long snippets. The 1000 pieces of music are divided into ten genres, each containing 100 pieces of music. In this work, feature extraction was used for two purposes: Dimensionality reduction: raw data dimensions are usually too huge to handle efficiently, e.g., an entire raw audio file may be too enormous. According to related research, a feature set is utilized to provide data with fewer values, and a single feature value for an entire audio signal can be produced. Meaningful representation: while the raw audio file has all of the information we could ever extract and use, it's critical that we express the musical aspects in a way that machines or people can understand. The classification results of the Convolutional Neural Network (CNN) when trained on spectrograms, three-second features, and thirty-second features are presented in this subsection. The three-second feature set provides more training data, which may explain why it achieves higher accuracy. The CNN implementation had the lowest accuracy, at 53.50 percent, because of the thirty-second duration features. This study was divided into three phases: 'phase A,' 'phase B,' and 'phase C.' Each step had an importance that corresponded to the research's contribution to the present body of knowledge. We offer music genre categorization using machine-learning and deep-learning methodologies, as well as a comparison of machine-learning and deep-learning models' accuracy in completing the classification assignment.

III.PROPOSED SYSTEM

The GTZAN dataset served as the basis for training and testing the models. The GTZAN genre collection dataset was compiled in the years 2000-2001. It consists of 1000 audio files, each of which lasts 30 seconds. Each class has 100 audio tracks and is divided into ten categories (10 music genres). All of the tracks are in.wav format. It features audio songs from the following eleven genres:

- Blues
- Classical
- Country
- Disco
- HipHop
- Jazz
- Metal
- Pop
- Reggae
- Rock

We'll utilise Convolutional Neural Networks because they've demonstrated to be the most effective for this task in numerous studies. Convolutional Neural Networks (CNN) are a type of artificial neural network that uses convolutional neural networks. CNNs are excellent at detecting design elements in input images, such as lines, gradients, circles, and even eyes and faces. CNN is made up of multiple convolutional layers stacked on top of each other, each capable of identifying more complex structures

1. **Music Genre Classification** A music genre classifier is a piece of software that determines what type of music is being played in an audio file. These devices are utilized for things like automatically labelling music for services like Spotify and Billboard, as well as choosing appropriate background music for events.
2. **Music Recommendation System**: The Recommender System is a software application and algorithm that makes recommendations for the products that a user is most interested in. Recommendations are used in a variety of real-world situations, such as deciding what products to buy, listening to music, or reading the latest news. Three algorithms are used by the recommendation engine:

- **Popularity filtering** It's a simple model that ranks the songs in the training set in order of popularity and proposes the most popular ones. This strategy does not take into account the preferences of the users.

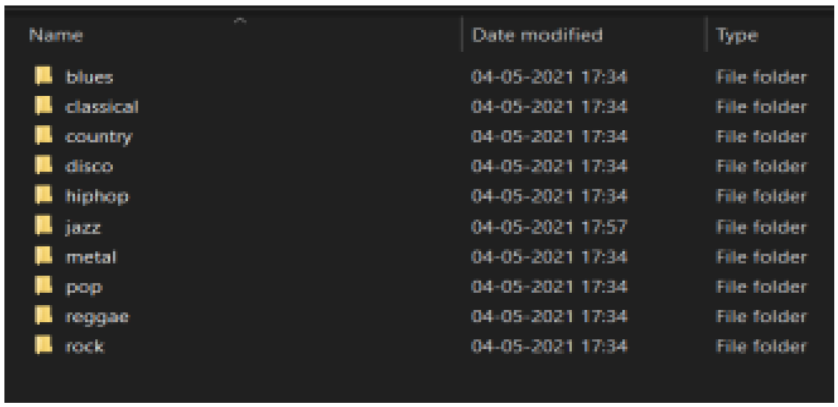
- **Content-Based filtering** Filtering methods based on content are based on the item's description and a profile of the user's preferences. These strategies work best when there is known information about an item (name, location, description, etc.) but not about the user. Suggestion is approached as a user-specific classification problem, with content-based recommenders developing a classifier based on product attributes for the user's likes and dislikes.

- **Collaborative filtering** It is founded on the idea that people who have agreed in the past will agree again in the future, and that they will enjoy comparable products. Only information about rating profiles for various persons or things is used to generate suggestions by the algorithm. It creates recommendations using this neighborhood by seeking peer users/items with a rating history similar to the current user or item.

IV. IMPLEMENTATION

1. Convolutional Neural Networks: A CNN, or convolutional neural network, is a deep learning neural network designed to analyze structured arrays of data-like representations. CNN's are excellent at detecting design elements in input images, such as lines, gradients, circles, and even eyes and faces. It is because of this feature that convolutional neural networks are so effective in computer vision. CNN does not require any pre-processing and can run directly on an underdone image. A feed-forward neural network with up to 20 layers is known as a convolutional neural network. The goal of this field is to train machinesto see the world in the same way that humans do, to perceive it in the same way, and to use that knowledge for a variety of tasks such as image and video recognition, image inspection, and classification, media recreation, recommendation systems, natural language processing, and so on.

2. Dataset: For training purposes, any Machine Learning model requires some data. Because future predictions are fully dependent on the training data, the data is tremendously crucial. The ensuing predictions would be abnormal if the dataset was not adequately filtered. The GTZAN dataset (as shown in fig 3) is the most often used public dataset for music genre recognition evaluation in machine learning research. The files were collected in 2000-2001 from a variety of sources, including personal CDs, radio, and microphone recordings, in orderto represent a variety of recording settings.



VI. CONCLUSION AND FUTURE SCOPE

In this project, we created a classifier that can predict the genre of audio recordings. The GTZAN music genre classification dataset is used in this study. This project explains how to use audio files to extract relevant information.. We used the Convolutional Neural Network Model in this research. The system will sort the music into one of ten genres: Blues, Classical, Country, Metal, Hip Hop, Jazz, Disco, and Pop. Reggae and rock are two genres of music. The recommendation algorithm suggests music to the user that is similar to what they are listening to. The project can now be implemented in a graphical user interface (GUI) format. It can then be turned into a fully functional website or application.

References

[1] 'The Echo Nest - Powering Intelligent Digital Music Applications.' The Intelligent Music Application Platform. Web. 07 May 2012. http://the.echonest.com/

[2] 'Last.fm.' Last.fm. Web. 07 May 2012. http://www.last.fm/

[3] Category-based intrinsic motivation,Proceedings of the Ninth International Conference on Epigenetic Robotics,Rachel Lee, Ryan Walker, Lisa Meeden, and James Marshall (2009)

[4] A growing neural gas learns topologies Advances in Neural Information Processing Systems 7 , Bernd Fritzke (1995)

[5] 'Musical Genre Classification of Audio Signals',George Tzanetakis and Perry Cook,IEEE Transactions on Speech and Audio Processing, 10(5), July 2002

[6] Mr. Ndiatenda Ndou, Mr. Ritesh Ajoodha and Ms. Ashwini Jadhav, "Music Genre Classification: A Review of Deep-Learning and Traditional Machine-Learning Approaches", 2021 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS), 2021.

[7] Jongseol Lee; Saim Shin; Dalwon Jang; Sei-Jin Jang; Kyoungro Yoon, "Music Recommendation System Based on Usage History and Automatic Genre Classification", 2015 IEEE International Conference on Consumer Electronics (ICCE), 2015.