iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** IV **Month of publication:** April 2023

DOI: https://doi.org/10.22214/ijraset.2023.50890

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

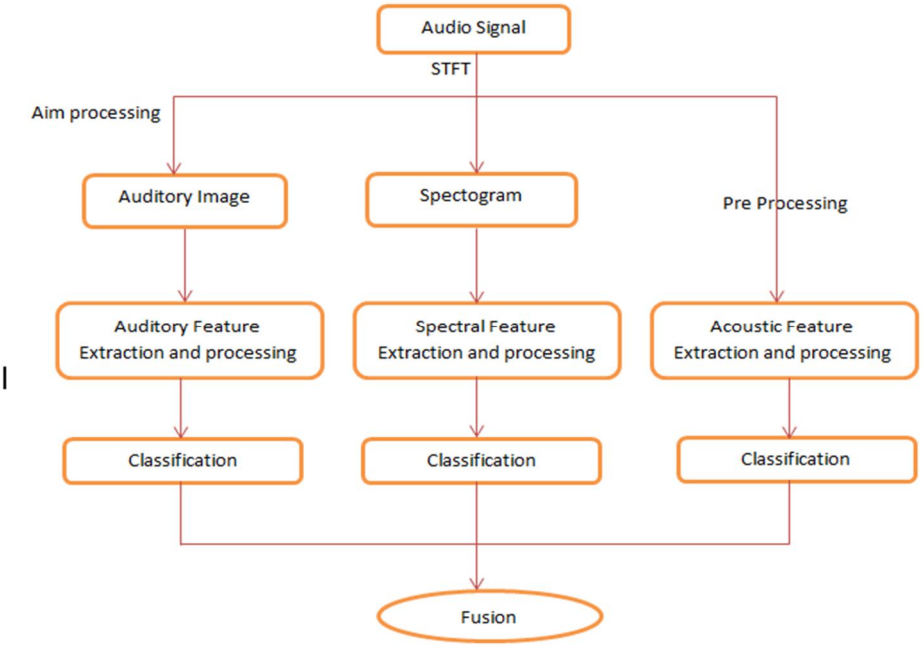
International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue IV Apr 2023- Available at www.ijraset.com

Song Classification using Machine Learning

Ritika Dhyani[1], Priyansh Vatsal[2], Priyanshu Goel[3], Prafull Chauhan[4], Prince Chauhan[5], Pratham Chauhan[6]
B. Tech Candidates, Department of Computer Sciences, IMS Engineering College, Ghaziabad, Uttar Pradesh

Abstract: ***The classification of music by genre is crucial in the modern world since the number of music tracks, both online and offline, is growing quickly. We must appropriately index them in order to have greater access to them. To retrieve music from a vast collection, automatic music genre classification is crucial. The majority of the current methods for categorising music genres rely on machine learning. We give a music dataset with ten distinct genres in this article. The system is trained and classified using a Deep Learning technique. Convolution neural networks are employed in this instance for training and classification. For audio analysis, feature extraction is the most important step. For sound samples, the Mel Frequency Cepstral Coefficient (MFCC) is employed as a feature vector. The suggested technique uses feature vector extraction to categorise music into different genres. Our findings indicate that our system's accuracy level is approximately 76%, which will significantly increase and facilitate the automatic classification of musical genres.***

Keywords: Classification of music genres, deep learning, convolutional neural networks, and neural networks.



I. INTRODUCTION

With the abundance of music at consumers' fingertips throughout the globe, there is a growing need for automatic classification of music for indexing of music and easier retrieval, which is frequently done manually by specialists in the field. In a nutshell, the issue statement for our project may be stated as follows: Given a number of audio recordings, the job is to classify each audio file into a specific category, such as audio that belongs to happy, sad, etc. Audio processing is one of the more difficult data science projects compared to image processing and other classification techniques.One such use is the classification of music genres, which seeks to place audio files in the appropriate sound groups to which they belong. Because classifying music manually requires listening to each song for the entirety, the application is crucial and needs automation to reduce manual error and time. Therefore, we will employ machine learning and deep learning techniques to automate the procedure.

In a nutshell, the issue statement for our project may be stated as follows: Given a number of audio files, the job is to classify each audio file into a specific genre, such as disco, hip-hop, etc.

©IJRASET: All Rights are Reserved

A classification algorithm uses a dataset of labelled examples as inputs to create a model that can automatically categorise unlabeled examples when presented with new, unlabeled data. A binary classification problem is one where there are just two labels (such as "calm" or "rock"). The challenge of multi-class classification arises when there are three or more labels in the set. We are looking at a multi-class problem because the set contains a variety of genres.

II. LITERATURE REVIEW

1) When listening to brief musical samples, humans are very adept at identifying the song's author, title, and even genre. Numerous NN approaches have been used to try and replicate these skills, with various degrees of success [7]. The mobile app Shazam is a well-known example of an application that uses music data to automatically identify an artist and a song's title. Shazam is well renowned for its ability to identify a song's title and artist from just a few seconds of audio. According to Shazam, a song's trademark consists of its spectrogram's prominent amplitude peaks. In place of (latitude, longitude, height), it's like compiling the positions of the highest mountain peaks in a region. For these noticeable peaks, we have (time, frequency, amplitude) [8]. Using two fully connected layers and a final classification output layer containing genre labels, Shazam's Tim O'Brien created a NN. This seems to be a pretty "vanilla" multiclass classifier model. He scored in the low 90% level for test accuracy. He was able to somewhat enhance the model by combining his NN with Sharath Pingula's (another Shazam employee) track-level collaborative filtering features. This article provides a summary of the machine learning research and application work done with regard to musical genre classification. For the purposes of the research, songs were divided into brief time segments. These time segments were then represented by the accompanying spectrogram images.These spectrograms were each assigned a music genre label before being used as inputs into a CNN. Six convolutional layers, a fully connected layer, a softmax function, and a one-hot array of genre classifications were the components of the NN. The softmax function was used to determine the likelihood that each genre would be recognised. On the basis of the test data, the results were 85% accurate.

2) We contrast the effectiveness of two kinds of models in this study. The first method uses deep learning to train a CNN model from beginning to end to predict an audio signal's genre label simply based on its spectrogram. The second method makes use of specially created time- and frequency-domain features. These features are used to train four conventional machine learning classifiers, and we evaluate how well they perform. The characteristics that are most helpful in this classification process are determined. For audio streaming services like Spotify and iTunes, being able to automatically categorise and assign tags to the music that is currently in a user's collection based on genre would be advantageous. In this study, the use of machine learning (ML) algorithms to recognise and categorise the genre of an audio recording is explored. Convolutional neural networks [2] are used in the first model that is discussed in this research. It is trained end-to-end on the MEL spectrogram of the audio input. In the second section of the investigation, we extract features from the audio signal's time domain and frequency domain. These features are then supplied to well-known machine learning models, such as Support Vector Machines, Gradient Boosting, Random Forests, and Logistic Regression, which are trained to categorise the given audio file. On the Audio Set dataset, the models are assessed [1]. We contrast the suggested models and research the relative significance of certain variables. It can be seen that with only the top 10 features, the model performance is surprisingly good, and that the model with the top 30 features only slightly performs worse than the full model, which has 97 features. We study how much performance in terms of AUC and accuracy, can be obtained by just using the top N while training the model.

3) In this study, music's acoustic characteristics were extracted using digital signal processing techniques, and music genre classification was subsequently carried out using neural networks. The process of grouping related types of information into a single identity (depending on the rhythm instrument used or the harmonic content) and naming that identity is known as genre classification. The genre, which is distinguished by some distinctive elements of the music, is one way to classify and arrange songs. Music genre classification has been a hotly debated topic ever since the early days of the Internet. Since they result from a complex interplay between the general audience, marketing, historical, and cultural variables, musical genres lack specific definitions and boundariesSome academics have proposed the definition of a new genre classification system specifically for the purposes of music information retrieval as a result of this observation. [4][12] 2003 [13] genre of music is reportedly the best source of general knowledge for deciphering the music's substance, according to Aucouturier and Pachet. For audio streaming services like Spotify and iTunes, being able to automatically categorise and assign tags to the music that is currently in a user's collection based on genre would be advantageous. Recent deep learning techniques make use of spectrograms, which are visual representations of the audio signal. Convolutional neural networks (CNNs) are fed data from these visual representations.[14] The authors of Lidy and Rauber (2005)[13] talk on the use of psychoacoustic properties for classifying

musical genres, particularly the significance of STFT measured using the Bark Scale (Zwicker and Fastl, 1999). Among the features used by (Tzanetakis and Cook, 2002)[4] were spectral contrast, spectral roll-off, and mel-frequency cepstral coefficients (MFCCs). In Nanni et al. (2016), SVM and AdaBoost classifiers are trained using a combination of audio and visual information.

III. METHODOLOGY

Convolutional neural networks (CNNs) are fed data from these visual representations.[14] The authors of Lidy and Rauber (2005)[13] talk on the use of psychoacoustic properties for classifying musical genres, particularly the significance of STFT measured using the Bark Scale (Zwicker and Fastl, 1999). Among the features used by (Tzanetakis and Cook, 2002)[4] were spectral contrast, spectral roll-off, and mel-frequency cepstral coefficients (MFCCs). In Nannietal. (2016), SVM and AdaBoost classifiers are trained using a combination of audio and visual information.

A. Common ML Algorithms

A few of the algorithms are described below.

1) *Artificial Neural Network (ANN):* ANNs are effective parallel-processing mathematical modelling systems that may simulate biological neural networks by using interconnected neuron units. The most well-liked learning algorithms in ML are ANNs, which are well-known for their adaptability, efficiency, and ability to represent complex flood processes with high fault tolerance and precise approximation. As a result, ANNs are regarded as trustworthy data-driven tools for developing black-box models of intricate and nonlinear interactions between rainfall and flooding as well as forecasting river flow and discharge. Numerous flood prediction applications, such as streamflow forecasting, river flow, rainfall-runoff, precipitation-runoff modelling, water quality, evaporation, river stage prediction, low-flow estimation, flood mapping and susceptibility, and river time series, have already been successfully implemented using artificial neural networks (ANNs). Iterative parameter adjustment is one of the main drawbacks of ANN use.
2) *Support Vector Machine (SVM):* Flood modelling makes extensive use of SVM, a supervised learning machine that operates on the principles of structural risk minimization and statistical learning theory. The SVM's training process creates models that assign new non-probabilistic binary linear classifiers that, by using inverse problem-solving, minimise the empirical classification error and maximise the geometric margin. Based on training from historical data, SVM is used to predict a quantity going forward in time. SVMs are now recognised as reliable and effective ML flood prediction systems. As ML alternatives to ANNs, SVM and SVR have gained appeal among hydrologists for flood prediction. As a result, they are used to predict floods in a variety of situations with promising results, superior generalisation ability, and higher performance when compared to ANNs, such as in cases of extreme rainfall, precipitation, rainfall-runoff, reservoir inflow, streamflow, flood quantiles, flood time series, and soil moisture.
3) *K- Nearest Neighbour (KNN):* Problems involving classification and regression can both be solved using this approach. It appears that the solution of categorization issues is more frequently applied within the Data Science business. It is a straightforward algorithm that sorts new instances by getting the consent of at least k of its neighbours and then saves all of the existing cases.This calculation is made using a distance function.By drawing parallels between.KNN and actual life, it is simple to comprehend. For instance, it makes sense to speak with a person's friends and co-workers if you want to learn more about them.

Before using the K Nearest Neighbours Algorithm, keep the following points in mind:

- KNN is computationally expensive
- Variables should be normalised to prevent greater range variables from skewing the algorithm
- Data still needs to be pre-processed.

4) *Convolutional Neural Network (CNN):* Convolutional neural network is a Deep Learning method built specifically for working with photos and videos. It uses photographs as inputs, extracts and learns the image's attributes, then categorises the images using the learned features. This programme takes its cues from how the Visual Cortex functions in the human brain. Processing of visual data from the outside world is carried out by the visual cortex, a region of the human brain. It has many levels, and each layer functions independently, extracting different information from images or other visuals. Once all the information from the various layers has been merged, the picture or visual is then evaluated or classed.

A neural network type called a convolutional neural network, or CNN or ConvNet, is particularly adept at processing input with a grid-like architecture, like an image. A binary representation of visual data is a digital image. It is made up of a grid-like arrangement of pixels, each of which has a pixel value to indicate how bright and what colour it should beIn CNN, rather than all the neurons in the fully linked layer, a layer's neurons will only be connected to a tiny portion of the layer.

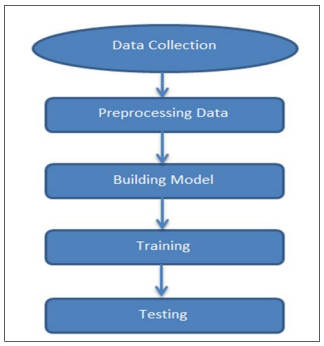
IV. RANDOM FOREST CLASSIFICATION

The supervised classification approach known as the random forest can be applied to both classification and regression issues. As the name implies, this algorithm builds a forest out of several trees.

A. Input Data Set

Three Types of Music Metadata.

1) *Descriptive Metadata:* With objective text tags like song title, duration in milliseconds, danceability, acousticness, energy, instrumentalness, and other information, descriptive metadata describes the contents of the recording. Every time someone has to search, arrange, sort, or display the music, descriptive information is used.
2) *Ownership/Performing Rights Metadata:* The cash will be split among a number of parties, including performing artists, lyricists, producers, and songwriters, whether we're talking digital streams, airplay, or movie synch. Therefore, ownership metadata is required, describing the legal arrangements supporting the release for the purpose of calculating (and allocating) royalties.
3) *Recommendation Metadata:* Metadata for recommendations differs. It primarily consists of subjective tags intended to reflect the recording's content and characterise its sound. To connect tracks in a meaningful way and fuel recommendation engines, recommendation information is used, such as mood labels, generative genre tags, and song similarity scores. There are several songs in the dataset. There are labels on the songs that are collected. One of the output classes—Happy, Sad, Energetic, Calm—includes labels. Additionally, each of these songs is examined, its parameters are retrieved, and a numerical value on a scale of 1 to 10 is assigned.



The picture classification model will be created, trained, and tested using the Python programming language. The model could be categorised roughly into:

a) Importing libraries and getting data ready
b) Model definition
c) Report on classification
d) Confusion Matrix
e) Last classified photos

V. CONCLUSION

Our application successfully categorises playlists according to mood with the aid of machine learning, giving users a categorised playlist. When a playlist is being listened to, the listener feels more at ease and filled with emotions, which boosts their mood and improves their mental condition. Marilyn Manson once said that "Music is the strongest form of magic" because music has the power to heal people and transform their emotions, which is equivalent to any form of magic.

Different music from your mood can make you feel stressed and unhappy, which can lead to low energy or inappropriate actions. However, this application's playlist perfectly matches the user's mood. The right music energises and inspires people to combat or handle their current predicament.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

VI. FUTURE SCOPE

With more research in this area, we will be able to use different machine learning algorithms, compare accuracies, and make even more accurate predictions while also learning how other models function and their benefits.

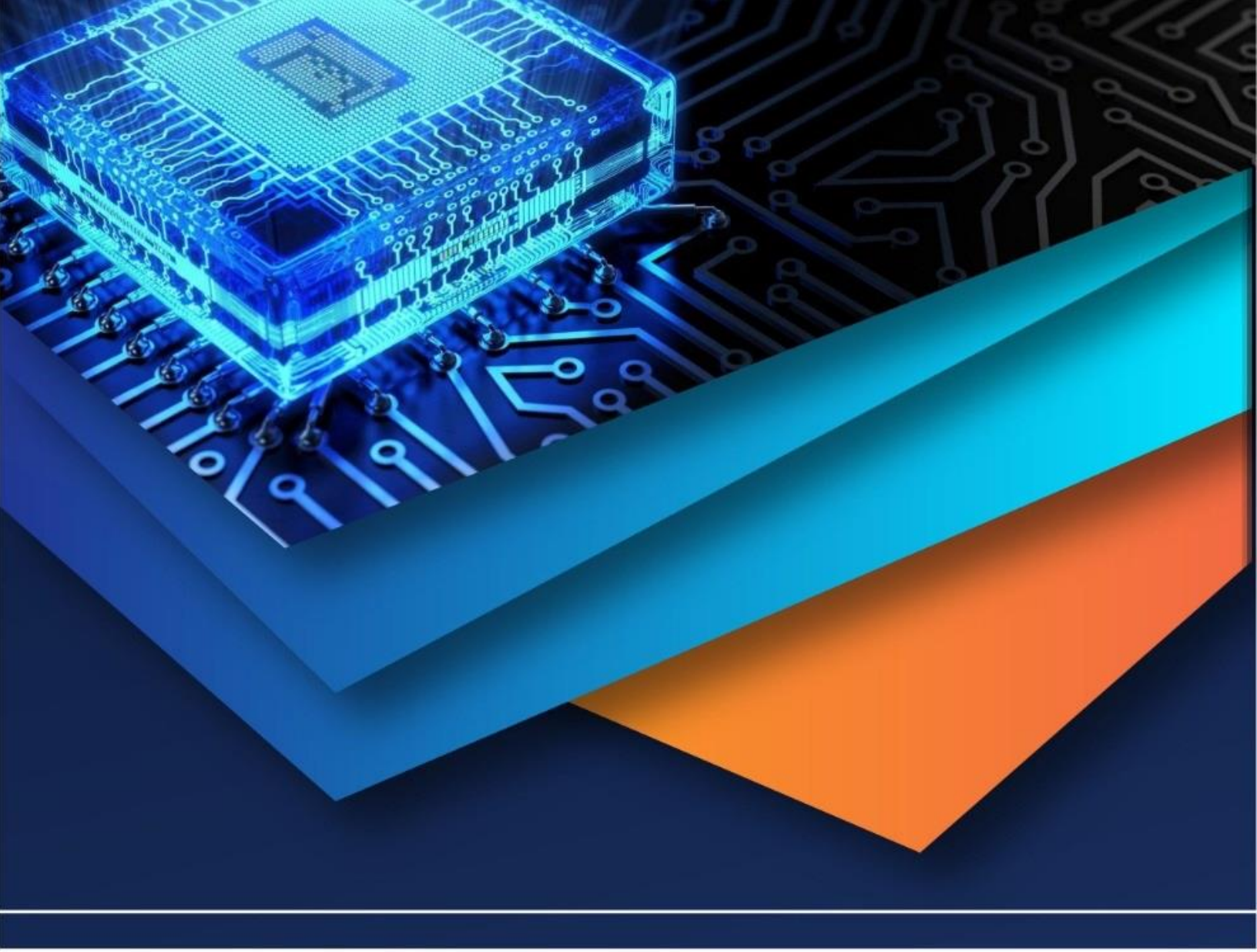
The classification of music into genres is a fundamental component of a powerful recommendation system. The major objective is to develop a machine learning model that categorises music samples into various genres in a more methodical manner.

Automating music classification can make it easier to locate important information like trends, popular genres, and performers.

REFERENCES

[1] Neural Network Music Genre Classification des genres de par reseau-neuronal (Nikki Pelchat).
[2] Music Genre Classification using Machine Learning Techniques (by Hareesh Bahuleyan)
[3] Music Genre Classification Using Deep Learning (by Navneet Parab, Shikta Das, Gunj Goda, Ameya Naik)
[4] George Tzanetakis and Perry Cook. 2002. Musical genre classification of audio signals. IEEE Transactions on speech and audio processing 10(5):293– 302.
[5] Y. M. Costa, L. S. Oliveira, and C. N. Silla, "An evaluation of convolutional neural networks for music classification using spectrograms," Appl. Soft Comput., vol. 52, pp. 28–38, Mar. 2017. Accessed: Dec. 16, 2018. [Online]
[6] "On Combining Diverse Models for Lyrics-Based Music Genre Classification ,Caio Luiggy Riyoichi SawadaUeno;Diego Furtado Silva, 2019 8th Brazilian Conference on Intelligent Systems (BRACIS).
[7] J. Despois. Finding the Genre of a Song With Deep Learning— A.I. Odyssey Part. 1. Accessed: Dec. 27, 2018. [Online]. Available: https://hackernoon.com/finding-the-genre-of-a-song-with-deep-learningda8f59a61194.
[8] F. Pachet and D. Cazaly, "A classification of musical genre," in Proc. RIAO Content-Based Multimedia Information Access Conf., Paris, France, Mar. 2000.
[9] S. Gollapudi, Practial Machine Learning. Birmingham, U.K.: Packt, 2016.
[10] T. O'Brien. (2017). Learning to Understand Music From Shazam. Accessed: Dec. 19, 2018. [Online]. Available: https://blog.shazam. com/learning-to-understand-music-from-shazam-56a60788b62
[11] T. Feng. Deep learning for music genre classification. 2014.
[12] R. Panda and R. P. Paiva, "Mirex 2012: Mood classification tasks submission," Machine Learning, vol. 53, no. 1-2, pp. 23–69, 2003

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)