Initial Project Report

MUSIC GENRE IDENTIFICATION USING MACHINE LEARNING

Done by

AKHIL ROCK BABU

Reg No: M23CA011

Under the guidance of

Prof. Manu M John

Project Synopsis

TOPIC : MUSIC GENRE IDENTIFICATION USING MACHINE LEARNING

Music genre classification is a fascinating and challenging task within the field of machine learning and audio signal processing. This project aims to classify music tracks into their respective genres using advanced machine learning techniques. With the ever-growing amount of digital music available, automated genre classification becomes increasingly relevant for various applications, such as music recommendation systems, music libraries, and streaming services.

In this project, we leverage machine learning algorithms to classify music genres based on features extracted from audio files. The dataset used for training and evaluation is the GTZAN dataset, available on Kaggle.The input to our system is raw audio file. Audio features are extracted from the audio signals and these features are essential as they capture the characteristics of the audio that are relevant for genre classification.We employ three different machine learning algorithms to classify the genres**: K-Nearest Neighbors (KNN), Support Vector Machine (SVM), and Naive Bayes classifier**. Each model is trained on the extracted features from the GTZAN dataset. The performance of these models is then evaluated to determine which algorithm provides the most accurate classification.Once trained, the models can classify new, unseen music tracks into their respective genres based on the extracted audio features.The output of the system is the predicted genre of the input music track.

The importance of this project lies in its potential applications in the music industry. Accurate genre classification can enhance user experience in music streaming services by providing personalized recommendations and organizing music libraries more efficiently. Moreover, it can aid in musicological research by automating the categorization of large collections of music. By exploring different machine learning algorithms, this project contributes to the ongoing development of more sophisticated and reliable music classification systems..

Dataset link: https://www.kaggle.com/datasets/andradaolteanu/gtzan-dataset-music-genre-classification

References:

1. Ndou, N., Ajoodha, R., & Jadhav, A. (2021). Music Genre Classification: A review of Deep-Learning and Traditional Machine-Learning Approaches. 2021 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS). https://doi.org/10.1109/iemtronics52119.2021.9422487

2. Ghildiyal, A., Singh, K., & Sharma, S. (2020). Music Genre Classification using Machine Learning. 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA). https://doi.org/10.1109/iceca49313.2020.9297444

3. Setiadi, D. R. I. M., Rahardwika, D. S., Rachmawanto, E. H., Sari, C. A., Irawan, C., Kusumaningrum, D. P., Nuri, N., & Trusthi, S. L. (2020). Comparison of SVM, KNN, and NB Classifier for Genre Music Classification based on Metadata. 2020 International Seminar on Application for Technology of Information and Communication (iSemantic). https://doi.org/10.1109/isemantic50169.2020.9234199

Submitted By:
Akhil Rock Babu
Reg No: M23CA011
2023 – 25 Batch MCA Department, MACE

Faculty Guide:
Prof. Manu M John
Associate Professor S3 MCA

LITERATURE REVIEW

Paper 1: Music Genre Classification: A Review of Deep-Learning and Traditional Machine-Learning Approaches

This paper provides a comprehensive review of both deep-learning and traditional machine-learning approaches for music genre classification. The study evaluates the strengths and weaknesses of each approach and discusses their respective accuracies and applications.

Title of the paper	Ndou, N., Ajoodha, R., & Jadhav, A. (2021). Music Genre Classification: A review of Deep-Learning and Traditional Machine-Learning Approaches. 2021 IEEE International IOT, Electronics and Mechatronics Conference(IEMTRONICS). https://doi.org/10.1109/iemtronics52119.2021.9422487
Area of work	This review is situated within the field of music information retrieval, specifically focusing on genre classification.
Dataset	The review considers various datasets used in existing literature, including the well-known GTZAN dataset, which is commonly used for benchmarking music genre classification models. The GTZAN dataset contains 60 columns and 10000 entries. https://www.kaggle.com/datasets/andradaolteanu/gtzan-dataset-music-genre-classification
Methodology/Strategy	The paper systematically reviews different methodologies employed in the classification of music genres. This includes discussing feature extraction techniques, such as MFCCs and chroma features, and comparing the effectiveness of various classification algorithms.
Algorithm	Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Convolutional Neural Network (CNN)
Result/Accuracy	**KNN : 92.69%** CNN : 72.40% SVM :80.80%
Advantages	Provides a thorough comparison of different approaches, highlighting the benefits of deep learning for feature extraction and classification accuracy.
Future Proposal	Suggests further research into hybrid models combining deep learning with traditional methods and exploring more diverse datasets for better generalization

Paper 2: Music Genre Classification using Machine Learning

This paper explores the application of machine learning techniques for the classification of music genres. The study investigates various algorithms and their performance in accurately categorizing music into different genres.

Title of the paper	Ghildiyal, A., Singh, K., & Sharma, S. (2020). Music Genre Classification using Machine Learning. 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA). https://doi.org/10.1109/iceca49313.2020.9297444
Area of work	This research falls under the domain of music information retrieval, specifically targeting the classification of music into genres using machine learning algorithms.
Dataset	The study utilizes the GTZAN dataset, which is a standard dataset for music genre classification containing 60 columns and 10000 entries. https://www.kaggle.com/datasets/andradaolteanu/gtzan-dataset-music-genre-classification
Methodology/Strategy	The methodology involves feature extraction from the audio tracks using techniques such as MFCCs, chroma features, and others. The extracted features are then used to train various machine learning models to classify the music genres.
Algorithm	Support Vector Machine (SVM), K-Nearest Neighbors (KNN) and Random Forest
Result/Accuracy	**SVM : 76.4%** RF : 69.6% KNN : 66.4%
Advantages	A key advantage of the study is its comparative analysis of different machine learning algorithms, providing insights into which models perform better for music genre classification.
Limitations	The paper mentions potential overfitting issues, especially with more complex models, and points out the limitations of the GTZAN dataset, which may not be sufficiently diverse or large enough to capture all the nuances of music genres.
Future Proposal	For future work, the authors suggest incorporating deep learning techniques, which have shown promise in other studies for music classification tasks. They also recommend using larger and more diverse datasets to improve the robustness and generalizability of the classification models.

Paper 3: Comparison of SVM, KNN, and NB Classifier for Genre Music Classification based on Metadata

This paper evaluates the performance of three different classifiers—Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Naive Bayes (NB)—for music genre classification based on metadata. The study aims to determine which classifier yields the highest accuracy when classifying music genres using metadata features.

Title of the paper	Setiadi, D. R. I. M., Rahardwika, D. S., Rachmawanto, E. H., Sari, C. A., Irawan, C., Kusumaningrum, D. P., Nuri, N., & Trusthi, S. L. (2020). Comparison of SVM, KNN, and NB Classifier for Genre Music Classification based on Metadata. 2020 International Seminar on Application for Technology of Information and Communication (iSemantic). https://doi.org/10.1109/isemantic50169.2020.9234199
Area of Work	The paper falls under the domain of music information retrieval, specifically focusing on genre classification based on metadata.
Dataset	The study uses metadata features extracted from Spotify music dataset from www.crowdai.org
Methodology/Strategy	The authors extracted metadata features and applied three classifiers—SVM, KNN, and NB—to classify music genres. They compared the performance of these classifiers to determine which one is most effective for this task.
Algorithm	Support Vector Machine (SVM) K-Nearest Neighbours (KNN) Naive Bayes classifier(NB)
Result/Accuracy	SVM : 80% KNN : 75.61% NB : 75.05%
Advantages	By comparing multiple classifiers, the study offers insights into the relative strengths and weaknesses of each approach, helping to identify the most suitable classifier for metadata-based music genre classification.
Limitations	Relying solely on metadata might not capture the full complexity of music genres, which can also be significantly influenced by audio features. This limitation suggests that the classifiers might not perform as well as they would with a more comprehensive feature set.
Future Proposal	The authors suggest that future research could benefit from combining metadata with audio features, which could lead to improved classification accuracy by leveraging the strengths of both types of data..

PROPOSED MODEL

Music Genre Classification Using k-Nearest Neighbors (KNN)

Introduction

Music genre classification is a crucial component in music information retrieval systems, significantly impacting music recommendation engines, playlist automation, and music library organization. While various machine learning algorithms have been explored for this task, k-Nearest Neighbors (kNN) has shown exceptional performance, particularly with short-duration audio features. This project aims to develop a high-accuracy music genre classification system using kNN, leveraging its strengths over other algorithms.

Objective

To develop and optimize a music genre classification system using the k-Nearest Neighbors (kNN) algorithm, demonstrating its superior accuracy and efficiency compared to other traditional and deep learning models.

Background and Motivation:

Recent research highlights the effectiveness of KNN in music genre classification, achieving higher accuracy than many traditional and deep learning algorithms. For example, in the study by Ndou, Ajoodha, & Jadhav (2021), KNN achieved an impressive accuracy of 92.69%, significantly outperforming Convolutional Neural Networks (CNNs). This high accuracy, combined with KNN's simplicity and ease of implementation, makes it an ideal choice for this project.

Why Choose KNN:

High Accuracy: KNN has demonstrated superior accuracy in recent studies, particularly for music genre classification tasks.

Simplicity and Efficiency: KNN is easy to implement and computationally efficient, making it suitable for real-time applications.

Effectiveness with Short-Duration Features: KNN performs exceptionally well with short-duration audio features, which are crucial for timely and accurate genre classification.

Methodology:

1. Data Collection and Preprocessing:

- Utilize the GTZAN dataset and extract relevant audio features such as Mel Frequency Cepstral Coefficients (MFCCs), chroma, and spectrograms using the Python Librosa package.

2. Model Development:

- Implement the k-Nearest Neighbors (KNN) algorithm, optimizing the number of neighbors (k) and distance metrics through cross-validation.

- Compare the performance of kNN with other algorithms such as SVM and Random Forest to validate its superiority.

3. Evaluation:

- Perform 10-fold cross-validation to ensure robustness and reliability of the model.

- Evaluate the model based on accuracy, precision, recall, F1-score, and computational efficiency.

Model Development:

1. Choosing the Algorithm: Based on literature review, the k-Nearest Neighbors (kNN) algorithm is chosen due to its high accuracy and efficiency.

2. Training the Model: The kNN model is trained using the feature matrix and corresponding labels. Cross-validation is used to optimize the number of neighbors (k) and the distance metric.

3. Validation: The model's performance is validated using 10-fold cross-validation to ensure robustness and avoid overfitting.

4. Hyperparameter Tuning: The model is fine-tuned by adjusting hyperparameters to achieve the best performance.

Prediction Process:

1. Input: An unseen audio file is input into the system.

2. Feature Extraction: Features are extracted from the input audio file using the same methods (Librosa) as during training.

3. Normalization: The extracted features are normalized to match the scale of the training data.

4. Model Prediction: The normalized feature vector is passed to the trained kNN model, which calculates the distance between this vector and all training samples, identifying the k-nearest neighbors.

5. Class Label Assignment: The model predicts the genre of the input audio file based on the majority class among the k-nearest neighbors.

Conclusion:

This project leverages the GTZAN dataset and the k-Nearest Neighbors (kNN) algorithm to develop a high-accuracy music genre classification system. By systematically extracting and normalizing relevant audio features, training a robust kNN model, and optimizing its parameters, the system is capable of accurately predicting the genre of new, unseen audio files. The combination of high accuracy, computational efficiency, and scalability makes this approach a powerful solution for music genre classification, with potential applications in music recommendation systems, automated playlist generation, and digital music libraries.

Dataset Overview:

The GTZAN dataset is a widely used benchmark for music genre classification tasks. It contains 1,000 audio tracks each 30 seconds long, divided into 10 genres: blues, classical, country, disco, hiphop, jazz, metal, pop, reggae, and rock. Each genre has 100 tracks, making the dataset balanced and suitable for classification tasks.

Source:

The GTZAN dataset is publicly available on Kaggle and was originally compiled by George Tzanetakis in 2002. It is a go-to dataset for researchers and practitioners working on music genre classification.

Features:

For music genre classification, various audio features can be extracted from the raw audio files. Using the Python Librosa package, we can extract the following key features:

1. Mel Frequency Cepstral Coefficients (MFCCs): Represent the short-term power spectrum of the audio and are commonly used in audio signal processing.

2. Chroma Features: Represent the energy distribution across 12 pitch classes, providing information on harmonic content.

3. Spectrogram: Visual representation of the spectrum of frequencies as they vary with time.

4. Zero-Crossing Rate: The rate at which the signal changes sign, indicating the noisiness of the signal.

5. Tempo: The overall estimated tempo of the audio track.

Class Labels:

The dataset is labeled with 10 distinct music genres:

1. Blues

2. Classical

3. Country

4. Disco

5. Hiphop

6. Jazz

7. Metal

8. Pop

9. Reggae

10. Rock