International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 11 | Nov - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Music Genre Classification

Bhavyashree R

ABSTRACT

Automatic music genre classification has become increasingly important for digital music management, recommendation systems, streaming platforms, and automated tagging. This work presents a deep learning–driven framework that classifies music genres using Mel-Frequency Cepstral Coefficients (MFCCs) along with complementary spectral–temporal features such as Root Mean Square (RMS) energy, Spectral Centroid, Tempo (BPM), and Zero Crossing Rate (ZCR). Convolutional Neural Networks (CNN), LSTMs, and hybrid CNN–LSTM architectures are employed to learn high-level audio representations. The proposed system demonstrates improved accuracy and scalability compared to conventional machine-learning approaches and reduces the dependency on manual feature engineering.

INTRODUCTION ~ TO DOMAIN

In recent years, Deep Learning (DL) has emerged as a transformative tool for analyzing complex data modalities, including audio signals. Unlike traditional machine-learning algorithms, deep neural networks automatically learn hierarchical patterns that are difficult to extract manually. Within the domain of music information retrieval, Convolutional Neural Networks (CNNs) have proven particularly effective, as they can process time-frequency representations such as spectrograms and MFCCs with high precision. Their ability to capture spatial and temporal patterns makes them ideal for genre classification tasks.

This research leverages CNN-based architectures, enriched spectral features, and audio preprocessing techniques to construct an end-to-end system capable of robust and efficient music genre prediction.

PROBLEM STATEMENT

Manually categorizing music based on genre presents numerous challenges:

- It is slow and labor-intensive.
- Judgments are often subjective, leading to inconsistent results.
- Accuracy is highly dependent on the expertise of the evaluator.

Traditional machine-learning models struggle with:

- Large variations in rhythm, tempo, and instrumentation,
- Limited representation of temporal dependencies,
- Poor scalability for big audio datasets,
- Limited generalization to emerging or hybrid genres.

These limitations highlight the need for an automated, scalable, and more accurate deep learning–based solution.

LITERATURE SURVEY PAPER 1

TITLE : A Study on Music Genre Classification using Machine Learning **Author :** Partha Ghosh, Soham Mahapatra, Subhadeep Jana, Ritesh Jha YEAR: 2023

DESCRIPTION : This paper presents a comparative study on music genre classification using various machine learning models. The authors explore multiple supervised learning techniques including SVM, KNN, Random Forest, and Decision Tree to classify audio samples into distinct genres. The work emphasizes the importance of feature extraction using techniques like MFCCs, chroma, and zero-crossing rate. Results indicate that appropriate preprocessing and model selection can significantly impact genre classification accuracy.

PAPER 2

TITLE: Optimizing the Configuration of Deep Learning Models for Music Genre Classification

Author : Teng Li YEAR: 2024
DESCRIPTION : The paper investigates optimization strategies for deep learning architectures in the domain of music genre classification. The author analyzes convolutional and recurrent neural networks, leveraging grid search and evolutionary algorithms to optimize hyperparameters such as layer size, learning rate, and dropout. The study uses melspectrogram representations of audio and achieves improved accuracy through model fine-tuning, demonstrating the potential of deep neural networks in audio classification tasks.

PAPER 3

TITLE : Machine Learning for Music Genre Classification Using Visual Mel Spectrum

Author : Y-H. Cheng, C.-N. Kuo YEAR:2022
DESCRIPTION : This research introduces a visual-based approach for classifying music genres using Mel spectrogram images. The authors employ convolutional neural networks (CNNs) to process visual representations Of audio data. The study compares several CNN architectures and discusses their performance across various genre datasets. The findings suggest that treating audio as image data can enhance classification efficiency and model robustness.

PAPER 4

TITLE : Music Genre Classification Using Deep Learning with KNN

Author : Dr. S. Ponlatha, Mathisalini B YEAR:2021
DESCRIPTION:The paper proposes a hybrid deep learning approach for music genre classification by integrating convolutional neural networks with the K-Nearest Neighbors algorithm. Feature extraction is carried out using CNN layers while KNN handles final classification. The system demonstrates promising accuracy and scalability, making it suitable for applications in music recommendation and digital libraries.

PAPER 5

TITLE : Automated Music Genre Classification through Deep Learning Techniques

Author : Mamidi Kiran Kumar YEAR:2023
DESCRIPTION : This paper explores the application of deep learning methods for automated music genre classification. The study utilizes Convolutional Neural Networks (CNNs) to extract features from audio spectrograms and classify them into various genres. Emphasis is placed on minimizing manual feature engineering and achieving high accuracy through model training on diverse datasets. The proposed technique demonstrates strong potential for use in intelligent music recommendation systems.

EXISTING SYSTEM

Current genre classification systems primarily rely on traditional machine-learning algorithms such as SVMs, Decision Trees, and KNN. These models often depend solely on limited features like MFCCs, ignoring key spectral characteristics such as brightness, loudness, or rhythmic intensity. Furthermore:

- Rule-based systems are rigid and fail to adapt to evolving genres.
- Shallow models lack the ability to capture complex temporal structures present in music.
- Performance drops significantly for diverse or large audio datasets.

PROPOSED SYSTEM

The proposed system integrates modern deep learning approaches to overcome limitations of traditional methods. Its distinctive features include:

- CNN, LSTM, and hybrid CNN-LSTM architectures for high-level feature learning.
- Rich feature extraction using MFCCs, RMS, Spectral Centroid, BPM, and ZCR.
- Normalization and preprocessing of audio signals for improved reliability.
- Data augmentation to enhance model robustness.
- A web-based (Flask) interface supporting real-time prediction of uploaded audio files.

The model is trained on multiple musical categories including Rock, Pop, Jazz, Hip-Hop, Blues, Metal, Classical, and Country.

OBJECTIVES

1. Design a deep learning model capable of accurately classifying music genres.
2. Enhance predictive accuracy by incorporating spectral-temporal audio features.
3. Assess the effect of adding RMS, Centroid, BPM, and ZCR features.
4. Compare CNN, LSTM, and hybrid architectures to determine the most effective model.
5. Automate the entire feature extraction pipeline.
6. Visualize and analyze feature distributions across various genres.
7. Build a scalable system capable of handling multiple music genres.

METHODOLOGY

1. **Dataset Acquisition:** Datasets such as GTZAN and FMA are used. All audio files are converted to mono with a 22,050 Hz sampling rate for consistency.

Feature Extraction

The system extracts:

- MFCCs
- RMS loudness
- Spectral Centroid
- Tempo (BPM)
- Zero Crossing Rate (ZCR)

These features provide a balanced combination of temporal, spectral, and rhythmic characteristics.

2. Model Design:

- CNN learns spatial patterns from MFCCs.
- LSTM captures sequential behavior across time frames.
- CNN-LSTM hybrid leverages strengths of both.

The models are optimized with the Adam/RMSprop optimizers and trained using cross-entropy loss.

3. Model Evaluation

Metrics used:

- Accuracy
- Precision
- Recall
- F1-score
- Confusion Matrix

4. Deployment: A lightweight Flask interface enables real-time predictions. Users can upload audio files and receive instant genre classification results.

SYSTEM REQUIREMENTS

Hardware Requirements

RAM : 2 GB

Hard disk : 100 GB

Process : 32/64 Pentium

Software Requirements

IDE : Flask

Language : Python.

Tool : Jupyter Notebook

Software : Anaconda

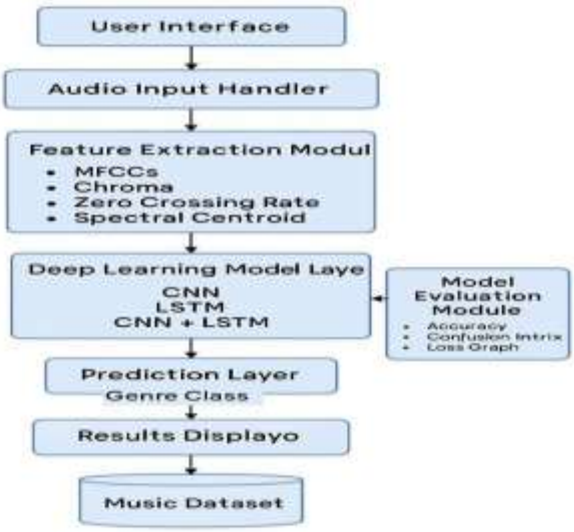
Front End : Html,CSS

Libraries : Tensorflow, keras, numpy,panda

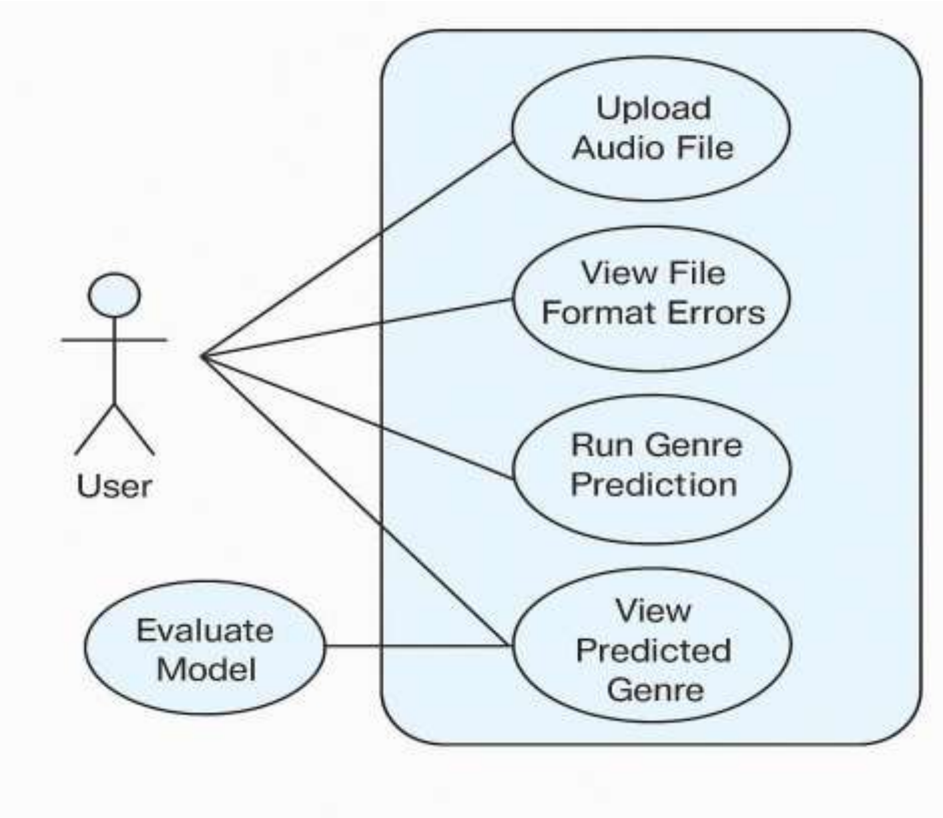
ALGORITHMS USED

Deep learning forms the core of this system. CNNs extract spatial and spectral patterns from MFCCs, while LSTMs focus on temporal correlations. Their combination produces a robust architecture capable of learning detailed musical characteristics without manual intervention.

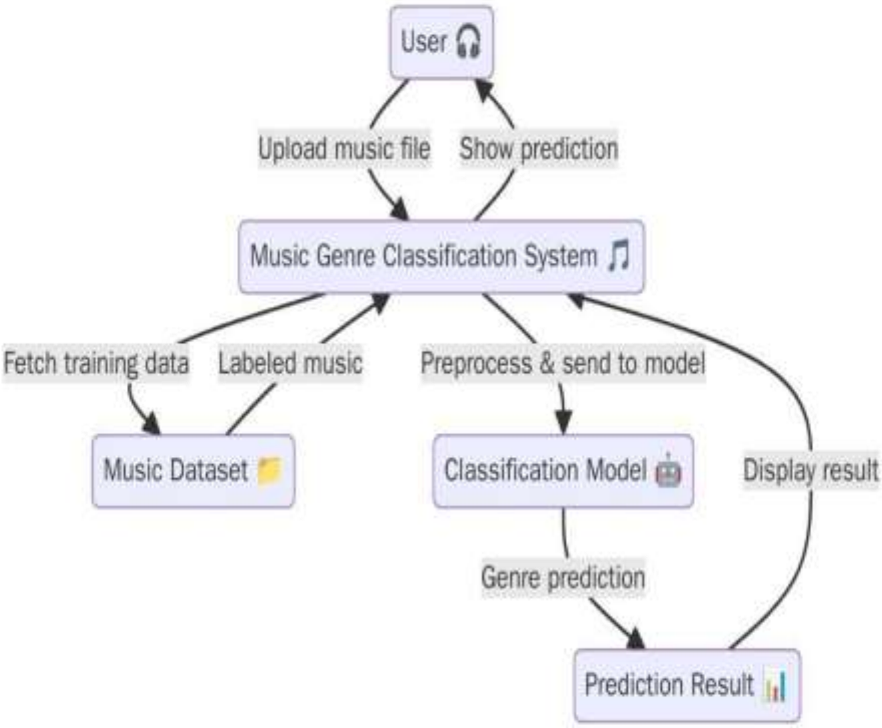
ARCHITECTURE DIAGRAM



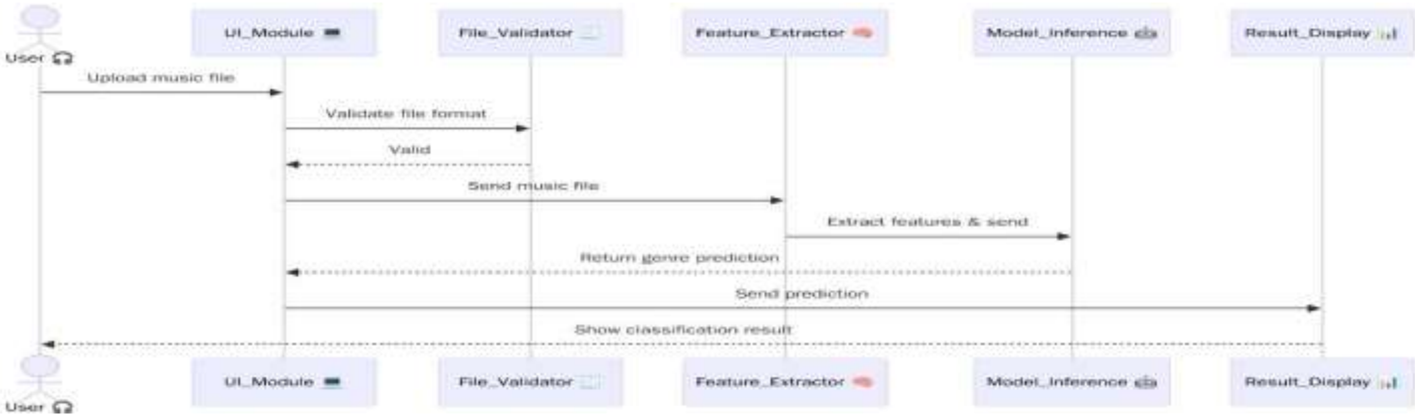
USE CASE DIAGRAM



DATA FLOW DIAGRAM



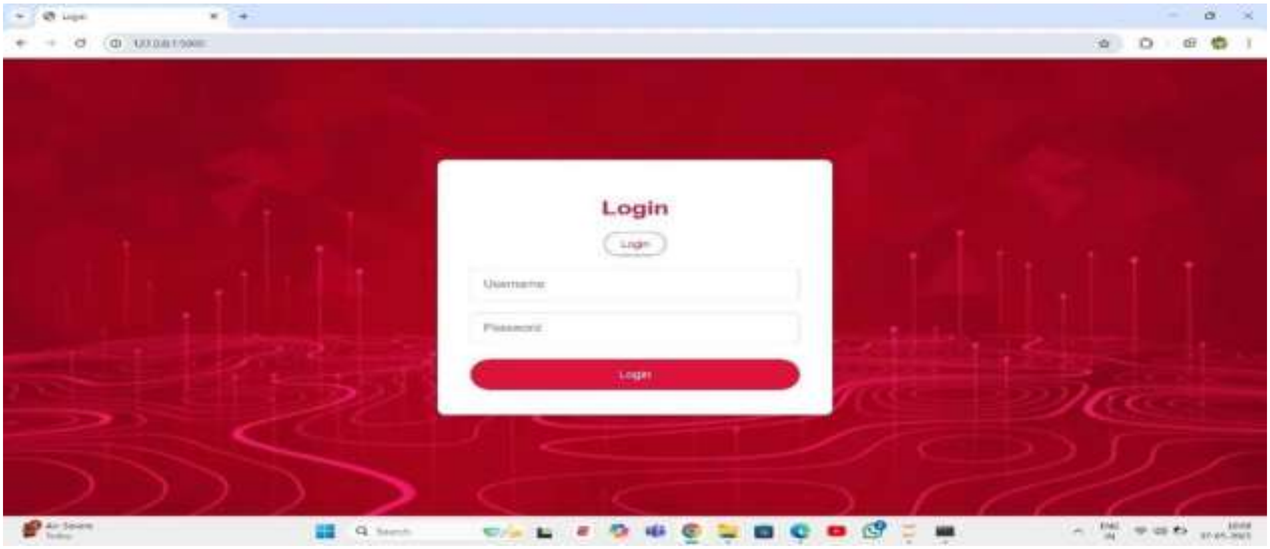
SEQUENCE DIAGRAM

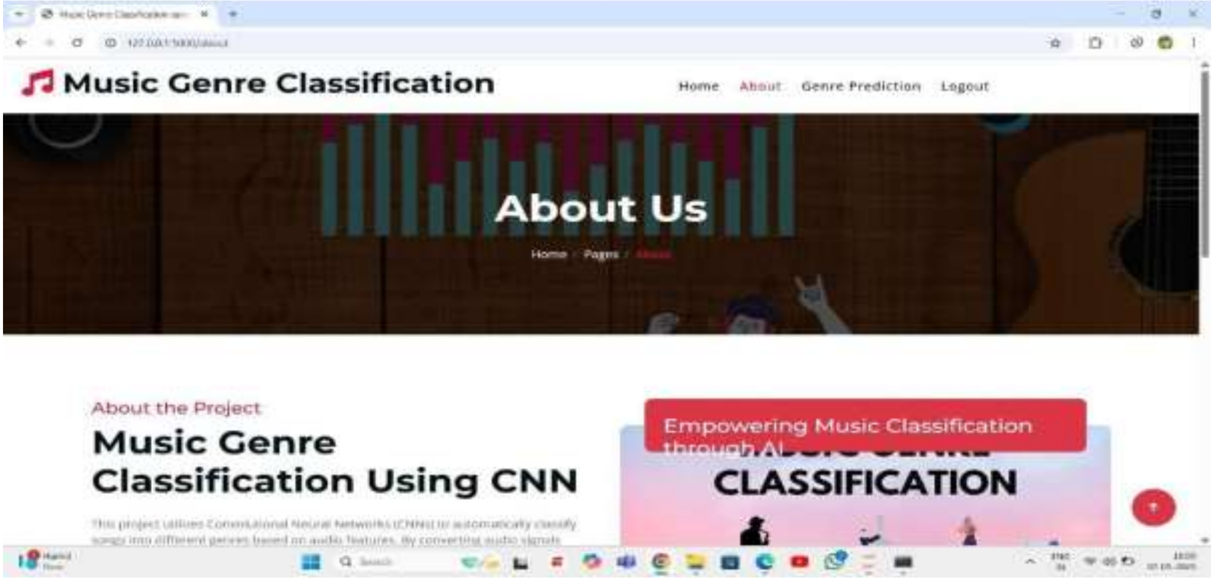


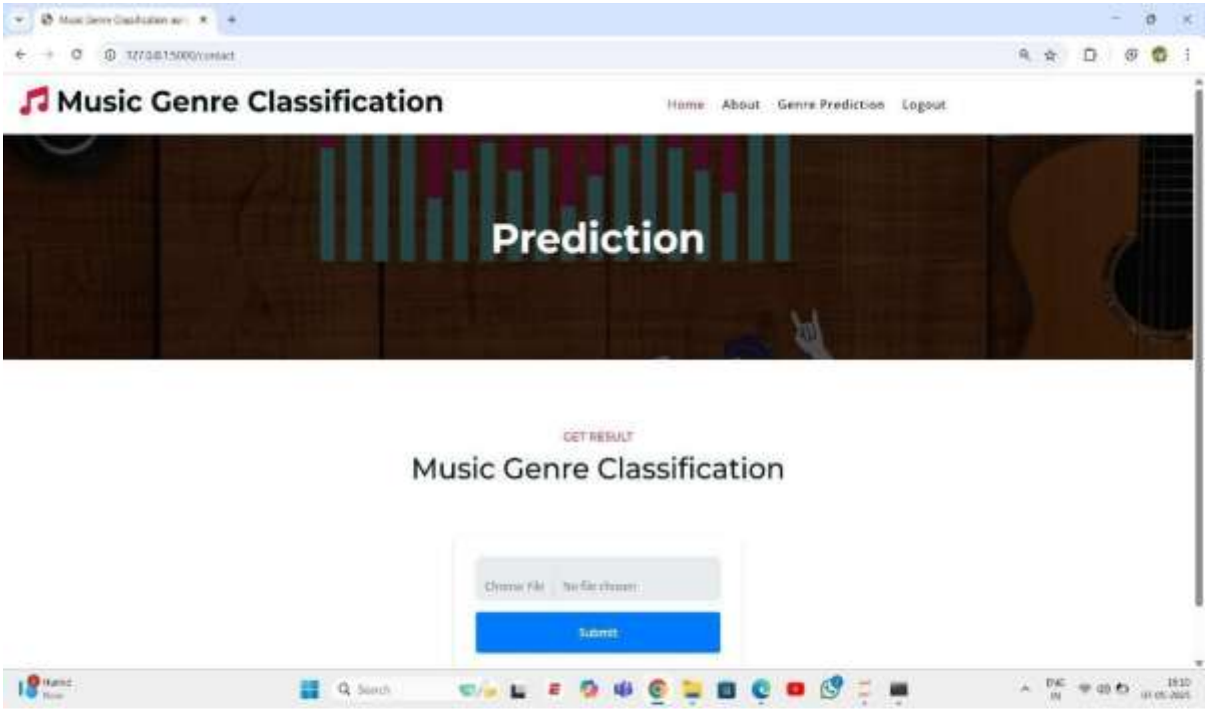
CONCLUSION

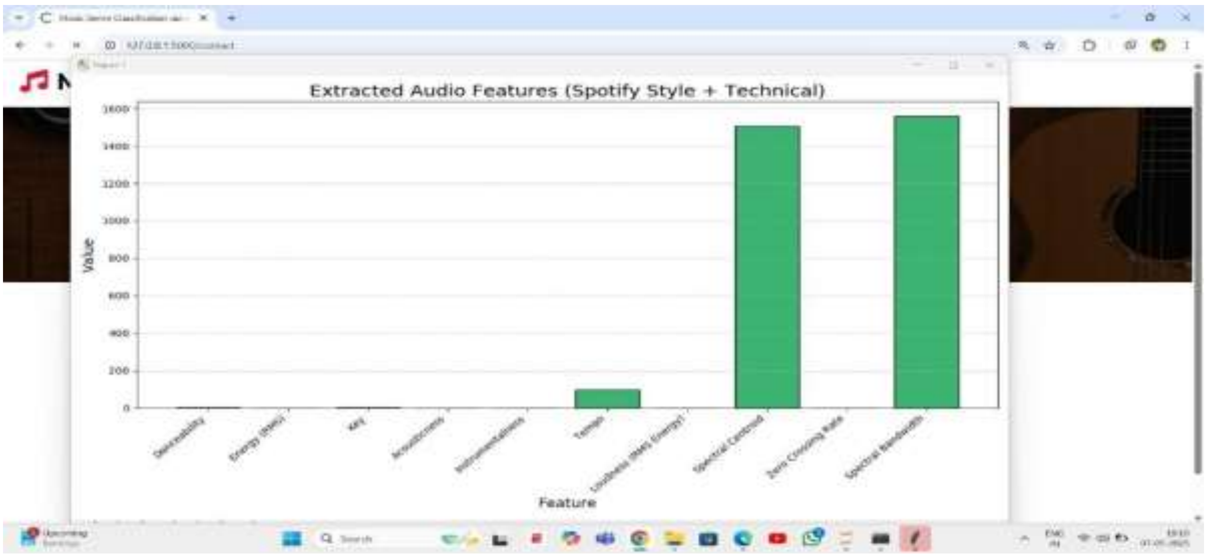
This research successfully demonstrates the potential of deep learning for music genre classification. By integrating CNN-based architectures with MFCCs and complementary spectral-temporal features, the system achieves reliable and high-accuracy predictions. The web-enabled interface further extends the applicability of the model for real-time usage. Overall, the approach is scalable, adaptable, and capable of supporting modern music retrieval and recommendation systems.

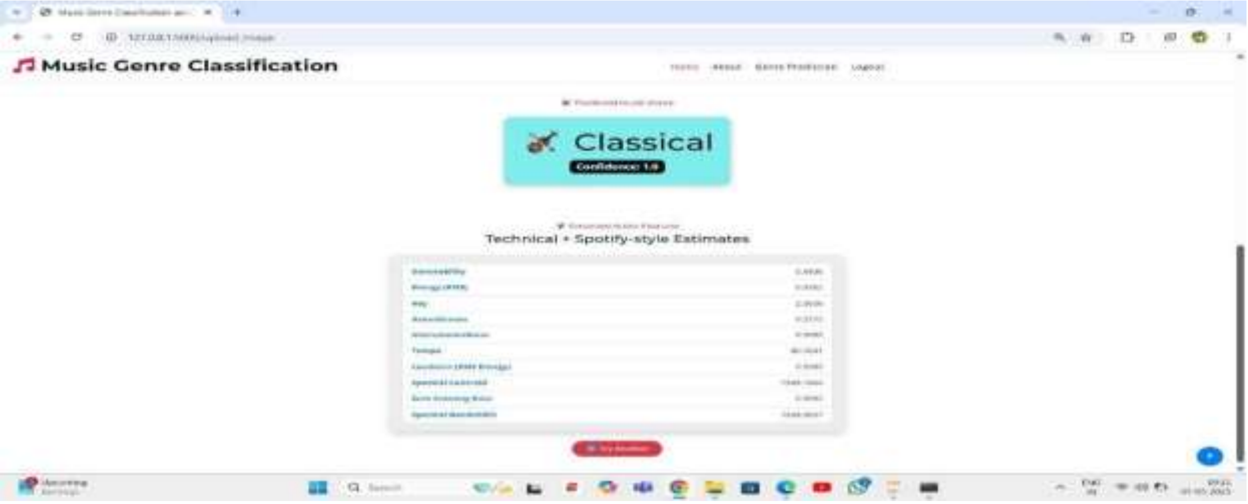
SNAPSHOTS




Music Genre Classification
Home
About
Genre Prediction
Logout
About Us
About the Project
Music Genre Classification Using CNN
Empowering Music Classification through AI
CLASSIFICATION


Music Genre Classification
Home
About
Genre Prediction
Logout
Prediction
GET RESULT
Music Genre Classification
Submit


Extracted Audio Features (Spotify Style + Technical)
Value
Feature
1600
1400
1200
1000
800
600
400
200
0
Danceability
Energy (RMS)
Key
Acousticness
Instrumentalness
Tempo
Loudness (RMS Energy)
Spectral Centroid
Zero Crossing Rate
Spectral Bandwidth



REFERENCES

[1] Ghosh, Partha & Mahapatra, Soham & Jana, Subhadeep & Jha, Ritesh. (2023). A Study on Music Genre Classification using Machine Learning. International Journal of Engineering Business and Social Science. 1. 308-320. 10.58451/ijebss.v1i04.55.

[2] Teng Li Optimizing the configuration of deep learning models for music genre classification https://doi.org/10.1016/j.heliyon.2024.e24892 Received 6 May 2023; Received in revised form 2

[3] Cheng, Y.-H.; Kuo, C.-N. Machine Learning for Music Genre Classification Using Visual Mel Spectrum. *Mathematics* **2022**, *10*, 4427. https://doi.org/10.3390/math10234427

[4] Dr. S. Ponlatha, Mathisalini B, Music Genre Classification Using Deep Learning with KNN International Journal of Advanced Research in Science, Communication and Technology (IJARSCT) Volume 12, Issue 2, December 2021

[5] Mamidi Kiran Kumar1 Automated Music Genre Classification through Deep Learning Techniques E3S Web of Conferences , 010 (2023)ICMPC 2023 430 33 https://doi.org/10.1051/e3sconf/202343001033

[6] Chaudhury, M.; Karami, A.; Ghazanfar, M.A. Large-Scale Music Genre Analysis and Classification Using Machine Learning with Apache Spark. *Electronics* **2022**, *11*, 2567. https://doi.org/10.3390/electronics11162567

[7] Jeffery Ho Kin Pou Music Genre Classification using Machine Learning link: https://www.riverpublishers.com/pdf/ebook/chapter/RP_9788770229005C53.p df

[8] LOW-RESOURCE MUSIC GENRE CLASSIFICATION WITH CROSS- MODAL NEURAL MODEL REPROGRAMMING arXiv:2211.01317v3 [cs.SD] 3 May 2023