**Music Genre Classification**

**Submitted for**

**Artificial Intelligence and Machine Learning CSET301**

Submitted by:

**E23CSEU0302 Harshit Dhar**

**E23CSEU0304 Aryan Tyagi**

**E23CSEU0312 Jayesh Anand**

Submitted to

**DR. YAJNASENI DASH**

**Jan-Apr 2025**

**SCHOOL OF COMPUTER SCIENCE AND ENGINEERING**

A close-up of a logo

Description automatically generated

**INDEX**

|  |  |  |
| --- | --- | --- |
| **Sr.No** | **Content** | **Pg No** |
| 1 | [Abstract](#_ABSTRACT) | 3 |
| 2 | [Introduction](#_INTRODUCTION) | 3 |
| 3 | [Related work](#_Related_Work) | 4 |
| 4 | [Methodology](#_METHODOLOGY) | 4-5 |
| 5 | [Hardware/software required](#_HARDWARE/SOFTWARE_REQUIRED) | 5 |
| 6 | [Experimental results](#_EXPERIMENTAL_RESULTS) | 6 |
| 7 | [Conclusion](#_CONCLUSION) | 6 |
| 8 | [Future scope](#_FUTURE_SCOPE) | 7 |
| 9 | [Github link](#_Github_Link) | 7 |

# **ABSTRACT**

Music genre classification is an important and prominent task in music information retrieval that helps in music recommendation, organization, and personalization systems. This project aims to classify audio tracks into genres using machine learning and deep learning techniques. The GTZAN dataset is a dataset widely used in music genre classification research. Mel-spectrogram features are extracted from audio files and used to train a Convolutional Neural Network (CNN) model for classification. The model achieved promising accuracy, demonstrating the effectiveness of deep learning in understanding complex audio patterns.

# **INTRODUCTION**

With the exponential growth of digital music libraries, automatic music genre classification has become essential for efficient music indexing and recommendation. Traditionally, genre classification was performed manually or by simple rule-based systems, which often lacked accuracy and scalability. With advances in artificial intelligence and deep learning, automatic classification systems can now analyze music patterns more effectively. This project implements an AI-based model that classifies songs into one of ten predefined genres using the GTZAN dataset.

# **RELATED WORK**

Several approaches have been explored in the past for music genre classification:

* **Traditional Machine Learning Models**: Support Vector Machines (SVM), k-Nearest Neighbors (k-NN), and Random Forests using handcrafted features like MFCCs, chroma, and tempo.
* **Deep Learning Models**: CNNs and RNNs trained on spectrograms or raw audio signals have shown superior performance due to their ability to automatically learn relevant features.
* **Hybrid Models**: Some studies have used combinations of CNN and LSTM to capture both spatial and temporal features in music.

# **METHODOLOGY**

**Dataset: GTZAN**

* Contains 1000 audio tracks (30 seconds each) in WAV format.
* 10 genres: blues, classical, country, disco, hiphop, jazz, metal, pop, reggae, and rock each having 100 samples.

**Data Preprocessing**

* Audio files converted to **Mel-spectrograms** using Librosa.
* Normalization and resizing of input to a fixed shape (e.g., 128x128).

**Model Architecture**

* **Conv2D** layers with ReLU activation
* **MaxPooling2D** layers for downsampling
* **Dropout** layers for regularization
* Fully connected **Dense** layers leading to a softmax output layer

**Training**

* Optimizer: Adam
* Loss function: Categorical Crossentropy
* Epochs: 30
* Batch size: 32
* Validation split: 20%

# **HARDWARE/SOFTWARE REQUIRED**

**Hardware:**

* Processor: Intel i5 or higher (recommended: GPU-enabled system for faster training)
* RAM: 8 GB or more
* Optional: NVIDIA GPU with CUDA support

**Software:**

* Python 3.x
* Libraries: TensorFlow/Keras, Librosa, NumPy, Matplotlib, Scikit-learn
* Jupyter Notebook or Google Colab for development and visualization

# **EXPERIMENTAL RESULTS**

| **Metric** | **Value** |
| --- | --- |
| Training Accuracy | 98.5% |
| Validation Accuracy | 92.4% |
| Test Accuracy | ~91% (avg) |
| Loss (Final Epoch) | ~0.2 |

* + - * Confusion Matrix showed that classical, jazz, and metal were most accurately classified.
* Some overlap was seen between rock and pop genres due to their acoustic similarity.

# **CONCLUSION**

The project successfully demonstrated that deep learning models, particularly CNNs, are effective for music genre classification when trained on well-processed audio features such as Mel-spectrograms. The model achieved over 90% accuracy, showing its robustness across multiple genres. However, there are still challenges in distinguishing acoustically similar genres, which can be addressed with more diverse datasets and hybrid models.

# **FUTURE SCOPE**

* **Use of Larger and More Diverse Datasets**: Incorporating datasets like FMA or MSD for better generalization.
* **Real-Time Classification**: Integrate the model into a mobile or web application for live audio classification.
* **Transfer Learning**: Utilize pre-trained audio models such as YAMNet or OpenL3 for better performance.
* **Hybrid Architectures**: Combine CNNs with LSTMs to learn temporal dynamics in music more effectively.
* **Multilingual/Multi-Cultural Genre Detection**: Extend to genres from non-Western music traditions.

# **GITHUB LINK**

[harshitdhar9/Music\_Genre\_Classification: Music genre classification using Gtzan dataset from Kaggle.](https://github.com/harshitdhar9/Music_Genre_Classification)