

# Music Genre Classification

Swarnava Mandal<sup>1</sup>

Saumitra Prasad Kulkarni<sup>1</sup>

Ayush Singh<sup>1</sup>

<sup>1</sup>IIT Jodhpur

{B23EE1072, B23EE1065, B23EE1008}@iitj.ac.in

<https://github.com/Saumi18/Music-Genre-Classification>

## Abstract

This project focuses on classifying music genres using a hybrid machine learning approach. We explore various models like Random Forest, SVM, ANN, and K-Nearest Neighbors (KNN), choosing the best model at end. We also include dimensionality reduction using PCA. Our implementation is from scratch and evaluated on the GTZAN dataset using metrics like accuracy and confusion matrix.

**Keywords:** Music Genre Classification, Random Forest, Neural Network, PCA + GMM, KNN, SVM.

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Citing Paper . . . . .	1
<b>2</b>	<b>Approaches Tried</b>	<b>2</b>
<b>3</b>	<b>Experiments and Results</b>	<b>2</b>
3.1	Dataset . . . . .	2
3.2	Experimental Setup . . . . .	2
3.3	Results Summary . . . . .	3
3.4	Cloud Exploration . . . . .	5
<b>4</b>	<b>Summary</b>	<b>5</b>
<b>5</b>	<b>Future Work</b>	<b>5</b>
<b>A</b>	<b>Contribution of each member</b>	<b>5</b>

## 1 Introduction

Music genre classification plays a vital role in music recommendation systems and audio content organization. The goal of this project is to classify music samples into predefined genres using machine learning techniques. We developed and evaluated multiple models from scratch, including KNN, Random Forest, SVM (One-vs-All), and a Neural Network. Finally, we implemented the best model with highest accuracy.

### 1.1 Citing Paper

For reference and benchmarking, we referred to prior comparative studies such as @articlegarg2022, author = Garg, Bhavya and Tiwari, Deepti and Jadon, Shivansh, title = Music Genre Classification using Machine Learning: A Comparative Study, journal = ResearchGate, year = 2022, note = Available at [https://www.researchgate.net/publication/362619781\\_Music\\_Genre\\_Classification\\_using\\_Machine\\_Learning\\_A\\_Comparative\\_Study](https://www.researchgate.net/publication/362619781_Music_Genre_Classification_using_Machine_Learning_A_Comparative_Study)

## 2 Approaches Tried

We implemented and tested the following approaches:

1. **KNN from scratch:** We implemented KNN using Euclidean distance to classify a sample based on the majority label of its closest neighbors. It is a non-parametric, instance-based learning algorithm that makes no assumptions about data distribution. The model was evaluated using different values of  $k$  to optimize classification accuracy. KNN achieved good performance on the GTZAN dataset due to its simplicity and effectiveness on clean, labeled data.
2. **Random Forest from scratch:** We implemented a Random Forest from scratch using parallel processing and manual tree building. Each decision tree was grown using bootstrap samples and random feature selection for node splits. Gini impurity was used to determine the best splits, and trees stopped growing based on depth or purity. The model predicted outputs via majority voting across all trees, with support for both labels and probabilities. It performed robustly on the GTZAN dataset and provided interpretability while demonstrating ensemble learning's power.
3. **Random Forest using SKLearn:** Implemented using the sklearn library with hyperparameter tuning using GridSearchCV and evaluated using Stratified K-Fold cross-validation for optimal performance. The final model was selected based on the best combination of parameters.
4. **SVM (One-vs-All):** We implemented a binary Support Vector Machine (SVM) from scratch using gradient descent and hinge loss. Then extended it to handle multiclass classification using the One-vs-All strategy, training one classifier per class. Each classifier distinguishes one genre from all others and makes predictions based on maximum margin separation. The model showed good interpretability and control over training dynamics via learning rate and regularization. We evaluated it using a confusion matrix, giving insights into how well each genre was distinguished.
5. **Neural Network:** We designed a deep neural network with dropout, batch normalization, and multiple hidden layers for robust learning. It was trained using the Adam optimizer with learning rate scheduling and evaluated on a multiclass music genre dataset. The network learns complex feature interactions, improving generalization and accuracy compared to simpler models. The model achieved strong results and was visualized using a confusion matrix for class-level performance analysis. Finally, we saved the trained model for reuse, making it practical for deployment or further tuning.
6. **(PCA + GMM):** Applied Gaussian Mixture Model (GMM) with Expectation Maximization (EM) for clustering, using PCA for dimensionality reduction and SelectKBest for feature selection to predict music genre classification.

Each model was evaluated on its accuracy and confusion matrix.

## 3 Experiments and Results

### 3.1 Dataset

We used the **GTZAN dataset**, which contains 1000 audio tracks equally distributed among 10 music genres. We extracted features from 3-second audio segments using pre-processed `features_3_sec.csv` file.

### 3.2 Experimental Setup

- Environment: Python 3.10
- Libraries: NumPy, Pandas, Scikit-learn (for metrics only)
- Custom Implementations: KNN, Random Forest, SVM, Neural Network, GMM
- Evaluation Metric: Accuracy, Confusion Matrix

### 3.3 Results Summary

- KNN Accuracy: 91%

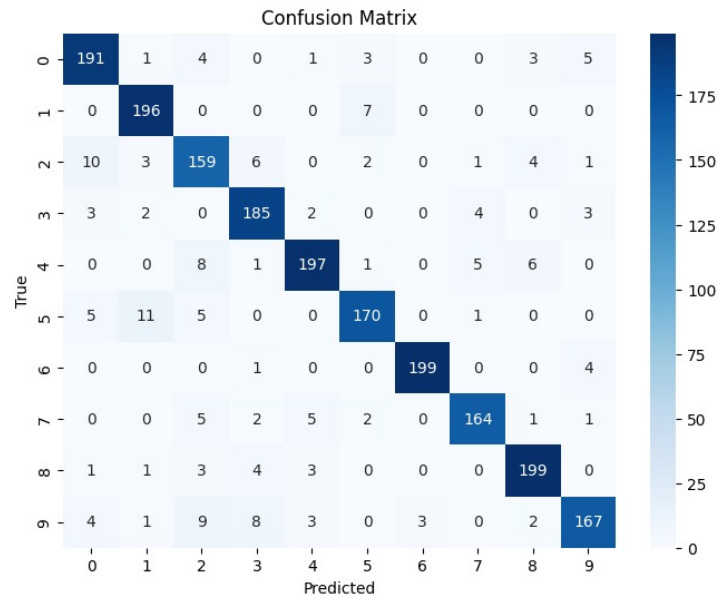


Figure 1: Confusion Matrix for KNN

- Random Forest Accuracy (From Scratch): 86%

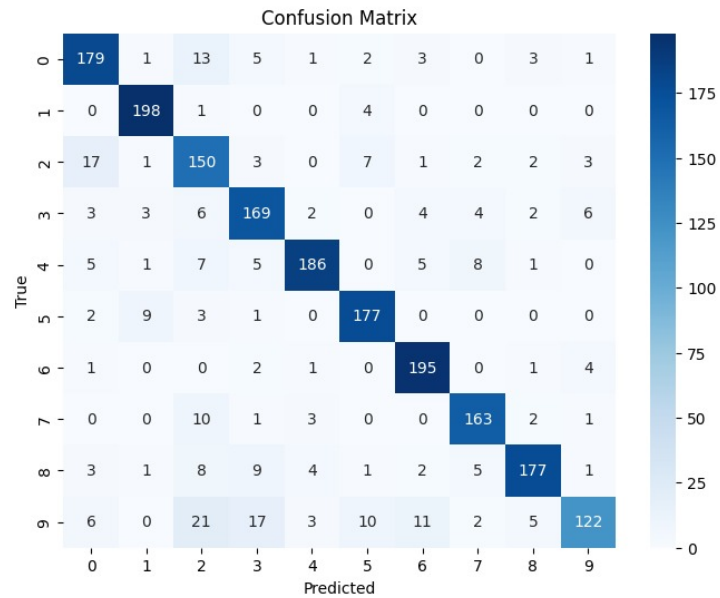


Figure 2: Confusion Matrix for Random Forest (From Scratch)

- Random Forest Accuracy (Using SKLearn): 89%
- GMM Accuracy: 40%

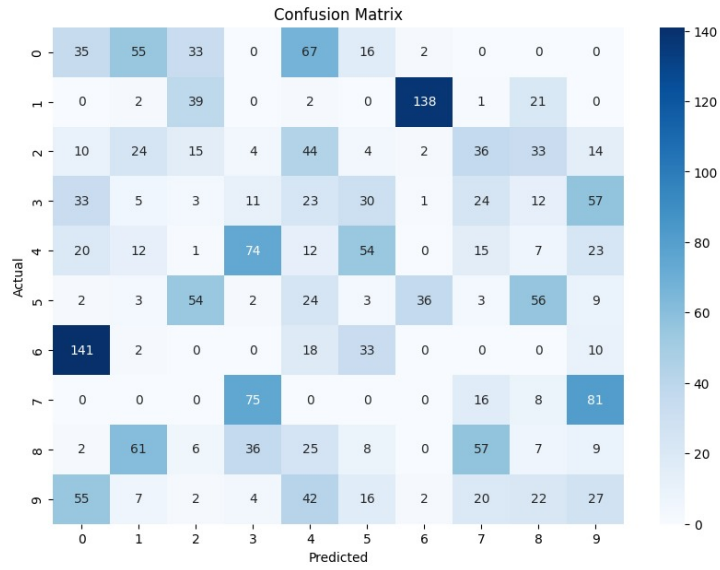


Figure 3: Confusion Matrix for GMM

- SVM Accuracy: 50%

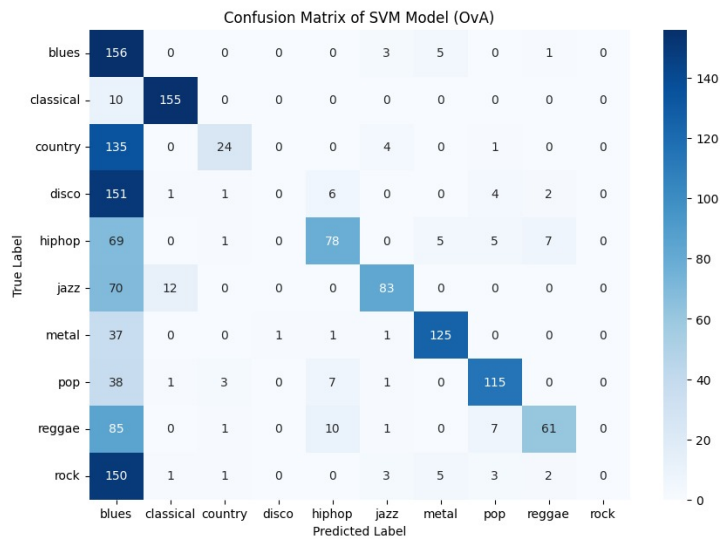


Figure 4: Confusion Matrix for SVM

- Neural Network Accuracy: 93.5%

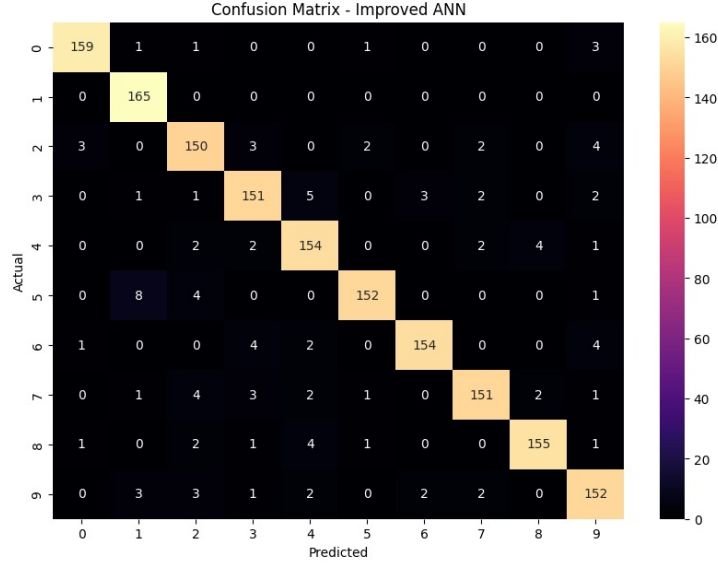


Figure 5: Confusion Matrix for Neural Network

### 3.4 Cloud Exploration

During the course of this project, we also explored cloud-based platforms such as **Google Colab** and **Kaggle Kernels** for faster training and experimentation. These platforms provided access to GPU resources and enabled more efficient development and testing of deep learning models, especially the Artificial Neural Network implemented using PyTorch.

## 4 Summary

In this project, we explored and implemented several machine learning models for the task of music genre classification using the GTZAN dataset. Our approach began with traditional algorithms such as K-Nearest Neighbors (KNN) and Random Forest, where we not only used scikit-learn implementations but also developed a custom Random Forest model from scratch to understand the underlying mechanics. We then moved on to implement a Support Vector Machine (SVM) from scratch using a One-vs-All strategy and gradient descent optimization. To further improve classification performance, we designed and trained a feedforward Artificial Neural Network (ANN) using PyTorch, incorporating enhancements like batch normalization, dropout, and learning rate scheduling for better generalization and stability during training. Each model was evaluated using accuracy scores and confusion matrices to analyze performance. The neural network model ultimately yielded the highest accuracy of 93.5 percent, outperforming all other models. Based on these results, we selected the ANN as the final model for music genre classification. This project demonstrates a systematic approach to model experimentation and comparison, where we not only evaluated multiple models but also fine-tuned and improved them step by step, both from theoretical and practical standpoints.

## 5 Future Work

- Fine-tune model hyperparameters for better accuracy.
- Experiment with deep learning models (CNN, RNN).
- Improve feature extraction using raw audio and spectrograms.

## A Contribution of each member

1. Swarnava Mandal, Ritik Singh: Implemented Custom Random Forest, and Custom KNN and performed evaluations and plotted confusion matrices, created Project Page, created Report.

2. Kukatlapalli William Samuel, Saumitra Prasad Kulkarni: Implemented Neural Networks, PCA + GMM, created Minutes of Meeting, report, Spotlight Video.
3. Prince Kumar, Ayush Singh: Implemented SVM, handled dataset processing, Web Demo, Spotlight Video.