Signal And Systems Project

Research Paper

Machine Learning Based Music Classification

Mustafa Mansur YÖNÜGÜL, Ahmet Can ÖMERCİKOĞLU, Riad MEMMEDLİ

Düzce Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü.Düzce/Türkiye

Abstract: A music genre classifier is crucial for organizing and categorizing vast musical libraries, enabling efficient music discovery and personalized recommendations for users. It facilitates targeted content delivery, enhancing user experience by aligning listeners with their preferred genres and styles. In this project, it is aimed to classify music according to its genres by using machine learning algorithms. To achieve this goal, digital signal processing techniques were employed to extract features from music files. Machine learning algorithms were then applied to automatically detect music genres based on these extracted features. Subsequently, a recommendation system was developed, utilizing the obtained genre information. The GTZAN dataset was selected for use in this study to train and evaluate the model. Eleven different machine learning models were trained in the Matlab Classification Learner environment and the findings were compared. Among the tested models, the most successful result was obtained with the Quadratic SVM algorithm with an accuracy rate of 59,6%.

Keywords: Music, Classification, Machine Learning

1. Methods and Tools

In this study, machine learning methods were employed to determine the genre of the uploaded music in the system. The workflow diagram depicting the process followed in the study is presented in Figure 1.

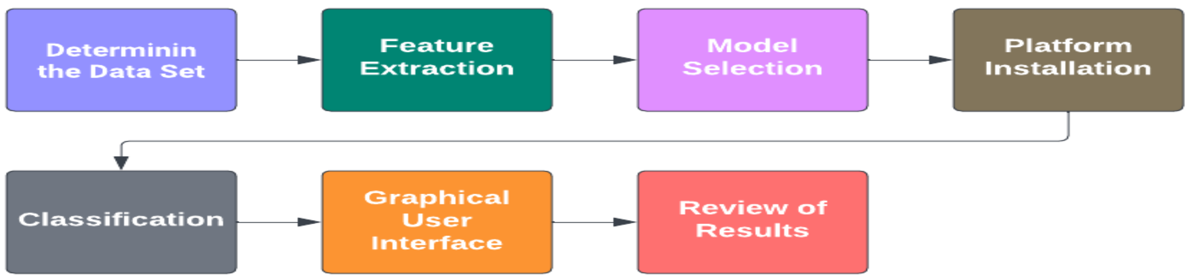


Figure 1. Work Flow Diagram

In the workflow, following the determination of the dataset, feature extraction was performed, and model selection was carried out. Subsequently, this model was trained on the platform, and the classification performance was evaluated, with the results examined through a user interface.

1.1 Data Set

In the study, the GTZAN dataset was chosen to compare the performances of the investigated methods. GTZAN is the most widely used dataset in music signal processing, proposed by G. Tzanetakis. This dataset encompasses a total of ten music genres, including blues, classical, country, disco, hip-hop, jazz, metal, pop, reggae, and rock (Table 1). However, in this project only 5 of the music genres are used - blues, classical, country, disco, hip-hop. The GTZAN dataset consists of two different .csv files. The first .csv file contains 100 audio files of each genre, each lasting 30 seconds. All tracks are .wav files with 22050Hz Mono 16-bit audio. The second .csv file follows the same structure, but the songs are segmented into 3-second audio files.

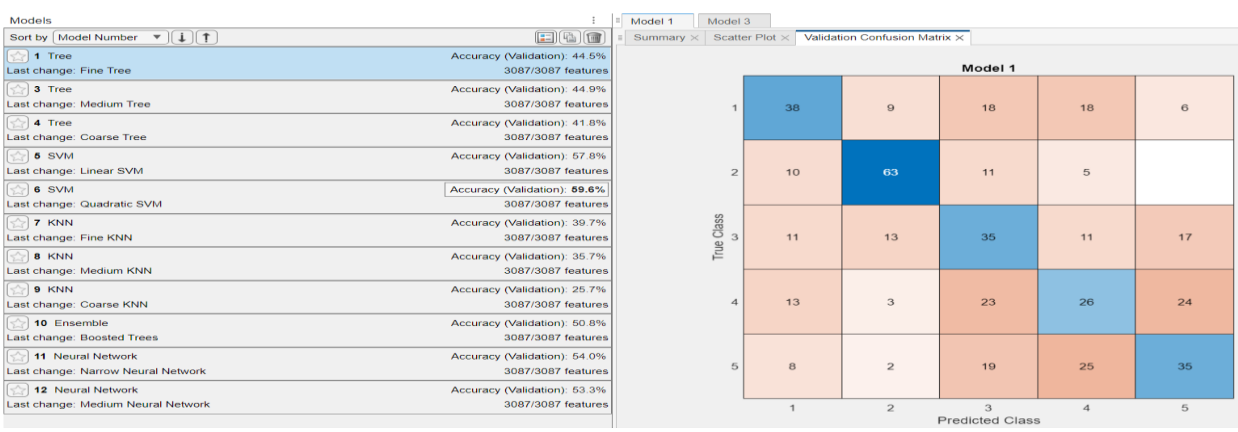
Table 1. GTZAN Data Set Content

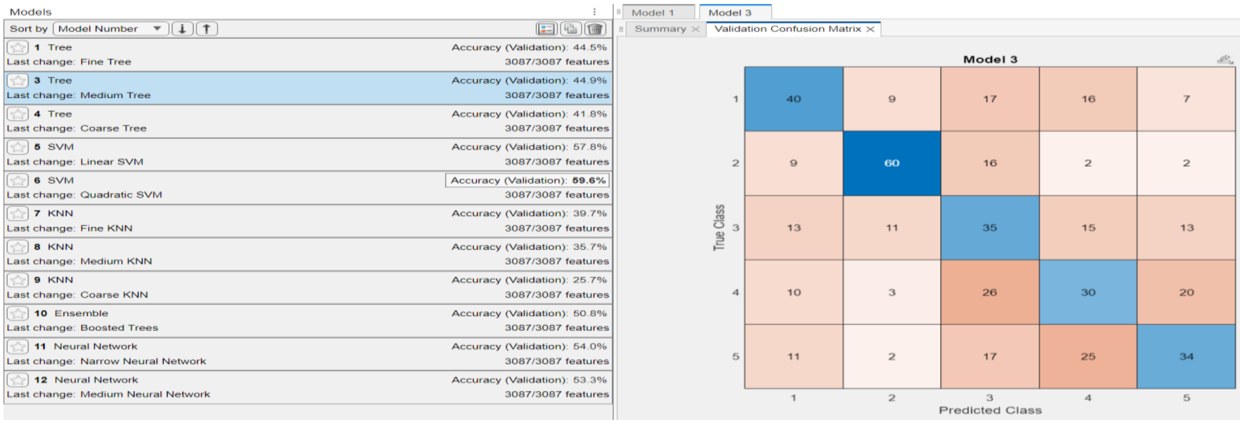
	Genres				
	Blues(1)	Classical(2)	Country(3)	Disco(4)	HipHop(5)
The number of file	100	98	97	99	97
Total	492				

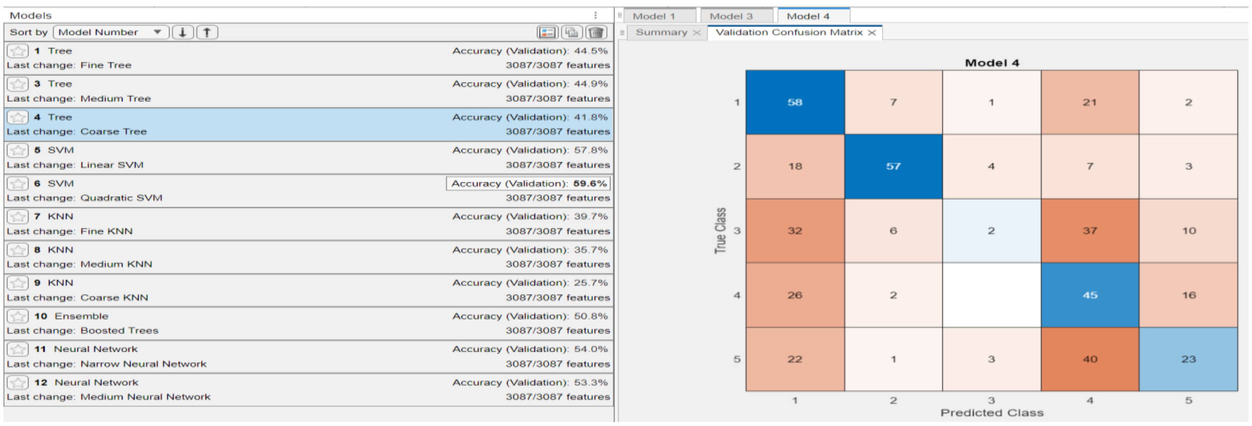
1.2 Feature Extraction

Feature extraction is a critical step in music classification, as it involves capturing relevant information from audio signals to facilitate the discrimination of different music genres. The selected features act as distinctive characteristics that contribute to the effectiveness of subsequent classification algorithms. In MATLAB, we employed various signal processing techniques for preprocessing audio data before feature extraction. This involved reading audio files, resampling, and normalizing the signals to ensure consistency across the dataset. For our music classification task, we focused on extracting key features known for their significance in differentiating genres. These included Mel-frequency cepstral coefficients (MFCCs) to capture spectral characteristics, rhythmic features like tempo and beat, and spectral contrast to highlight tonal variations. MATLAB's Signal Processing Toolbox was instrumental in implementing feature extraction. The extracted features were then used to train a classifier, and the results demonstrated improved accuracy compared to baseline models.

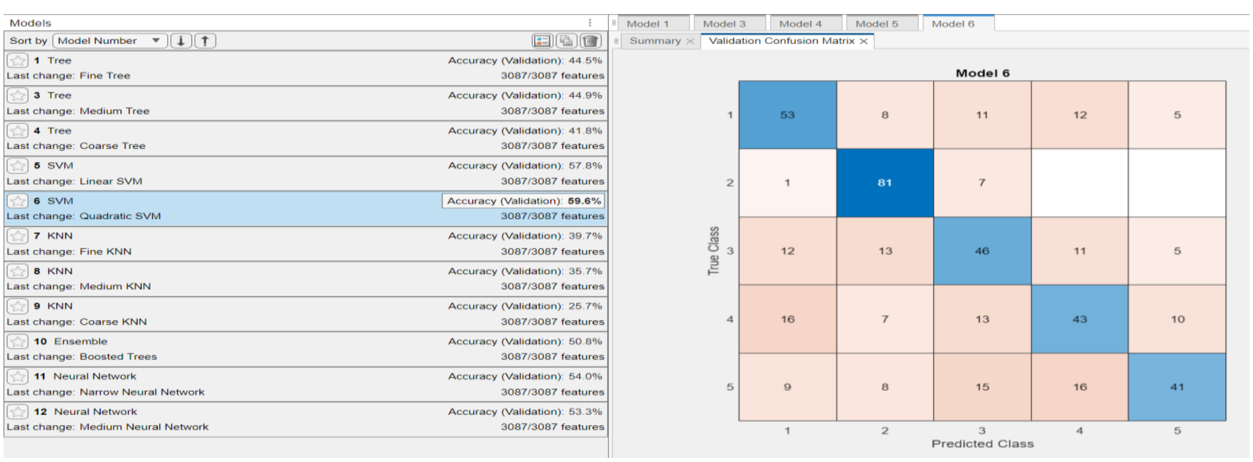
Confusion Matrix for Each Model

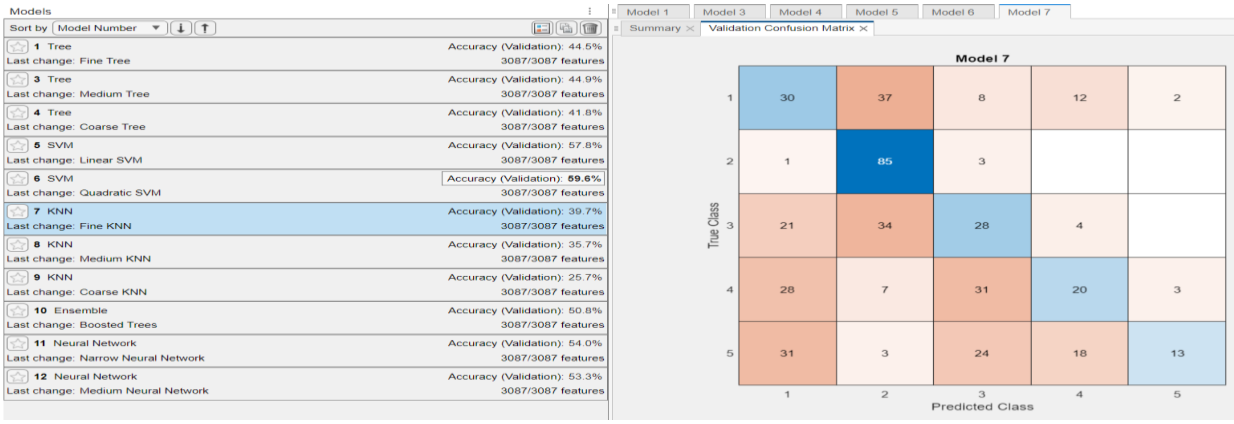


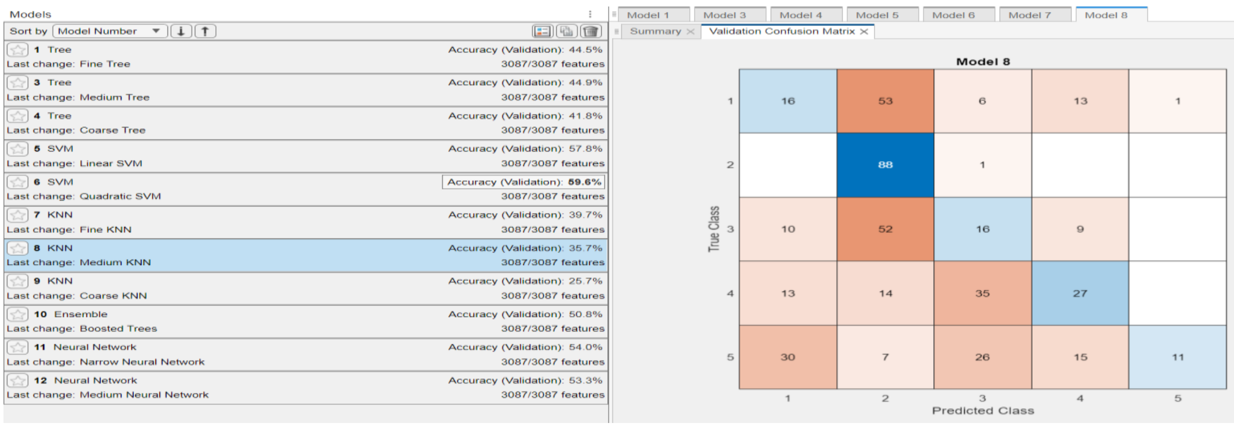

Models
Sort by Model Number
1 Tree Accuracy (Validation): 44.5%
Last change: Fine Tree 3087/3087 features
3 Tree Accuracy (Validation): 44.9%
Last change: Medium Tree 3087/3087 features
4 Tree Accuracy (Validation): 41.8%
Last change: Coarse Tree 3087/3087 features
5 SVM Accuracy (Validation): 57.8%
Last change: Linear SVM 3087/3087 features
6 SVM Accuracy (Validation): 59.6%
Last change: Quadratic SVM 3087/3087 features
7 KNN Accuracy (Validation): 39.7%
Last change: Fine KNN 3087/3087 features
8 KNN Accuracy (Validation): 35.7%
Last change: Medium KNN 3087/3087 features
9 KNN Accuracy (Validation): 25.7%
Last change: Coarse KNN 3087/3087 features
10 Ensemble Accuracy (Validation): 50.8%
Last change: Boosted Trees 3087/3087 features
11 Neural Network Accuracy (Validation): 54.0%
Last change: Narrow Neural Network 3087/3087 features
12 Neural Network Accuracy (Validation): 53.3%
Last change: Medium Neural Network 3087/3087 features
Model 1
Model 3
Summary
Validation Confusion Matrix
Model 3
True Class
Predicted Class
1: 40 9 17 16 7
2: 9 60 16 2 2
3: 13 11 35 15 13
4: 10 3 26 30 20
5: 11 2 17 25 34

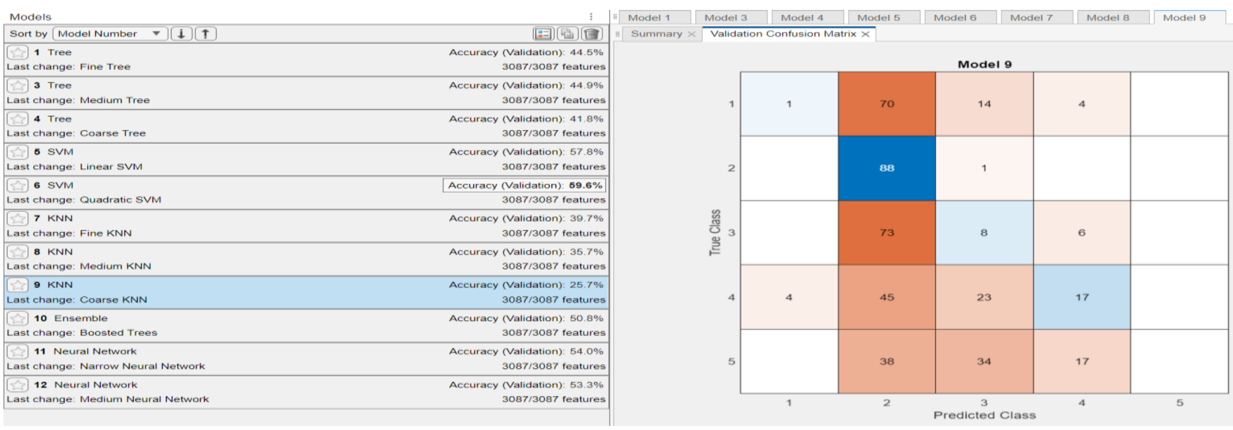

Models
Sort by Model Number
1 Tree Accuracy (Validation): 44.5%
Last change: Fine Tree 3087/3087 features
3 Tree Accuracy (Validation): 44.9%
Last change: Medium Tree 3087/3087 features
4 Tree Accuracy (Validation): 41.8%
Last change: Coarse Tree 3087/3087 features
5 SVM Accuracy (Validation): 57.8%
Last change: Linear SVM 3087/3087 features
6 SVM Accuracy (Validation): 59.6%
Last change: Quadratic SVM 3087/3087 features
7 KNN Accuracy (Validation): 39.7%
Last change: Fine KNN 3087/3087 features
8 KNN Accuracy (Validation): 35.7%
Last change: Medium KNN 3087/3087 features
9 KNN Accuracy (Validation): 25.7%
Last change: Coarse KNN 3087/3087 features
10 Ensemble Accuracy (Validation): 50.8%
Last change: Boosted Trees 3087/3087 features
11 Neural Network Accuracy (Validation): 54.0%
Last change: Narrow Neural Network 3087/3087 features
12 Neural Network Accuracy (Validation): 53.3%
Last change: Medium Neural Network 3087/3087 features
Model 1
Model 3
Model 4
Summary
Validation Confusion Matrix
Model 4
True Class
Predicted Class
1: 58 7 1 21 2
2: 18 57 4 7 3
3: 32 6 2 37 10
4: 26 2 45 16
5: 22 1 3 40 23

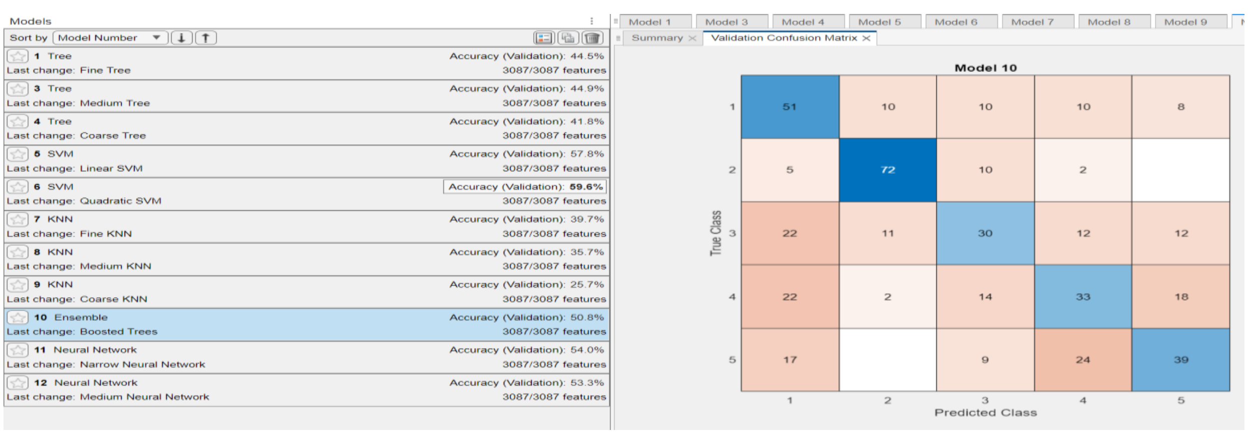

Models
Sort by Model Number
1 Tree Accuracy (Validation): 44.5%
Last change: Fine Tree 3087/3087 features
3 Tree Accuracy (Validation): 44.9%
Last change: Medium Tree 3087/3087 features
4 Tree Accuracy (Validation): 41.8%
Last change: Coarse Tree 3087/3087 features
5 SVM Accuracy (Validation): 57.8%
Last change: Linear SVM 3087/3087 features
6 SVM Accuracy (Validation): 59.6%
Last change: Quadratic SVM 3087/3087 features
7 KNN Accuracy (Validation): 39.7%
Last change: Fine KNN 3087/3087 features
8 KNN Accuracy (Validation): 35.7%
Last change: Medium KNN 3087/3087 features
9 KNN Accuracy (Validation): 25.7%
Last change: Coarse KNN 3087/3087 features
10 Ensemble Accuracy (Validation): 50.8%
Last change: Boosted Trees 3087/3087 features
11 Neural Network Accuracy (Validation): 54.0%
Last change: Narrow Neural Network 3087/3087 features
12 Neural Network Accuracy (Validation): 53.3%
Last change: Medium Neural Network 3087/3087 features
Model 1
Model 3
Model 4
Model 5
Summary
Validation Confusion Matrix
Model 5
True Class
Predicted Class
1: 57 12 8 9 3
2: 1 82 6
3: 23 13 33 12 6
4: 26 1 12 40 10
5: 13 1 15 16 44

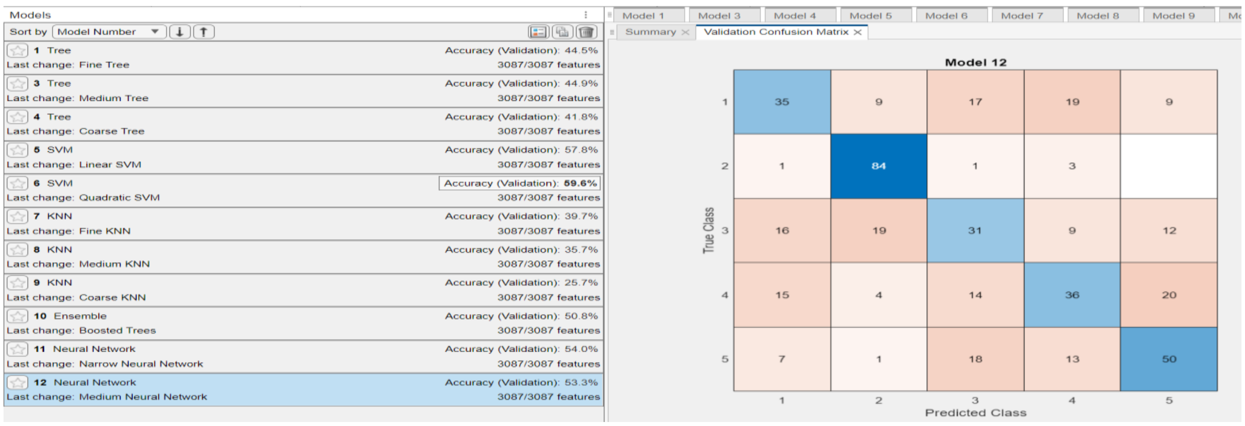

Models
Sort by Model Number
1 Tree Accuracy (Validation): 44.5%
Last change: Fine Tree 3087/3087 features
3 Tree Accuracy (Validation): 44.9%
Last change: Medium Tree 3087/3087 features
4 Tree Accuracy (Validation): 41.8%
Last change: Coarse Tree 3087/3087 features
5 SVM Accuracy (Validation): 57.8%
Last change: Linear SVM 3087/3087 features
6 SVM Accuracy (Validation): 59.6%
Last change: Quadratic SVM 3087/3087 features
7 KNN Accuracy (Validation): 39.7%
Last change: Fine KNN 3087/3087 features
8 KNN Accuracy (Validation): 35.7%
Last change: Medium KNN 3087/3087 features
9 KNN Accuracy (Validation): 25.7%
Last change: Coarse KNN 3087/3087 features
10 Ensemble Accuracy (Validation): 50.8%
Last change: Boosted Trees 3087/3087 features
11 Neural Network Accuracy (Validation): 54.0%
Last change: Narrow Neural Network 3087/3087 features
12 Neural Network Accuracy (Validation): 53.3%
Last change: Medium Neural Network 3087/3087 features
Model 1
Model 3
Model 4
Model 5
Model 6
Summary
Validation Confusion Matrix
Model 6
True Class
Predicted Class
1: 53 8 11 12 5
2: 1 81 7
3: 12 13 46 11 5
4: 16 7 13 43 10
5: 9 8 15 16 41


Models
Sort by Model Number
1 Tree
Last change: Fine Tree
Accuracy (Validation): 44.5%
3087/3087 features
3 Tree
Last change: Medium Tree
Accuracy (Validation): 44.9%
3087/3087 features
4 Tree
Last change: Coarse Tree
Accuracy (Validation): 41.8%
3087/3087 features
5 SVM
Last change: Linear SVM
Accuracy (Validation): 57.8%
3087/3087 features
6 SVM
Last change: Quadratic SVM
Accuracy (Validation): 59.6%
3087/3087 features
7 KNN
Last change: Fine KNN
Accuracy (Validation): 39.7%
3087/3087 features
8 KNN
Last change: Medium KNN
Accuracy (Validation): 35.7%
3087/3087 features
9 KNN
Last change: Coarse KNN
Accuracy (Validation): 25.7%
3087/3087 features
10 Ensemble
Last change: Boosted Trees
Accuracy (Validation): 50.8%
3087/3087 features
11 Neural Network
Last change: Narrow Neural Network
Accuracy (Validation): 54.0%
3087/3087 features
12 Neural Network
Last change: Medium Neural Network
Accuracy (Validation): 53.3%
3087/3087 features
Model 1
Model 3
Model 4
Model 5
Model 6
Model 7
Summary
Validation Confusion Matrix
Model 7
True Class
Predicted Class
1
2
3
4
5
30 37 8 12 2
1 85 3
21 34 28 4
28 7 31 20 3
31 3 24 18 13


Models
Sort by Model Number
1 Tree
Last change: Fine Tree
Accuracy (Validation): 44.5%
3087/3087 features
3 Tree
Last change: Medium Tree
Accuracy (Validation): 44.9%
3087/3087 features
4 Tree
Last change: Coarse Tree
Accuracy (Validation): 41.8%
3087/3087 features
5 SVM
Last change: Linear SVM
Accuracy (Validation): 57.8%
3087/3087 features
6 SVM
Last change: Quadratic SVM
Accuracy (Validation): 59.6%
3087/3087 features
7 KNN
Last change: Fine KNN
Accuracy (Validation): 39.7%
3087/3087 features
8 KNN
Last change: Medium KNN
Accuracy (Validation): 35.7%
3087/3087 features
9 KNN
Last change: Coarse KNN
Accuracy (Validation): 25.7%
3087/3087 features
10 Ensemble
Last change: Boosted Trees
Accuracy (Validation): 50.8%
3087/3087 features
11 Neural Network
Last change: Narrow Neural Network
Accuracy (Validation): 54.0%
3087/3087 features
12 Neural Network
Last change: Medium Neural Network
Accuracy (Validation): 53.3%
3087/3087 features
Model 1
Model 3
Model 4
Model 5
Model 6
Model 7
Model 8
Summary
Validation Confusion Matrix
Model 8
True Class
Predicted Class
1
2
3
4
5
16 53 6 13 1
88 1
10 52 16 9
13 14 35 27
30 7 26 15 11


Models
Sort by Model Number
1 Tree
Last change: Fine Tree
Accuracy (Validation): 44.5%
3087/3087 features
3 Tree
Last change: Medium Tree
Accuracy (Validation): 44.9%
3087/3087 features
4 Tree
Last change: Coarse Tree
Accuracy (Validation): 41.8%
3087/3087 features
5 SVM
Last change: Linear SVM
Accuracy (Validation): 57.8%
3087/3087 features
6 SVM
Last change: Quadratic SVM
Accuracy (Validation): 59.6%
3087/3087 features
7 KNN
Last change: Fine KNN
Accuracy (Validation): 39.7%
3087/3087 features
8 KNN
Last change: Medium KNN
Accuracy (Validation): 35.7%
3087/3087 features
9 KNN
Last change: Coarse KNN
Accuracy (Validation): 25.7%
3087/3087 features
10 Ensemble
Last change: Boosted Trees
Accuracy (Validation): 50.8%
3087/3087 features
11 Neural Network
Last change: Narrow Neural Network
Accuracy (Validation): 54.0%
3087/3087 features
12 Neural Network
Last change: Medium Neural Network
Accuracy (Validation): 53.3%
3087/3087 features
Model 1
Model 3
Model 4
Model 5
Model 6
Model 7
Model 8
Model 9
Summary
Validation Confusion Matrix
Model 9
True Class
Predicted Class
1
2
3
4
5
1 70 14 4
88 1
73 8 6
4 45 23 17
38 34 17


Models
Sort by Model Number
1 Tree
Last change: Fine Tree
Accuracy (Validation): 44.5%
3087/3087 features
3 Tree
Last change: Medium Tree
Accuracy (Validation): 44.9%
3087/3087 features
4 Tree
Last change: Coarse Tree
Accuracy (Validation): 41.8%
3087/3087 features
5 SVM
Last change: Linear SVM
Accuracy (Validation): 57.8%
3087/3087 features
6 SVM
Last change: Quadratic SVM
Accuracy (Validation): 59.6%
3087/3087 features
7 KNN
Last change: Fine KNN
Accuracy (Validation): 39.7%
3087/3087 features
8 KNN
Last change: Medium KNN
Accuracy (Validation): 35.7%
3087/3087 features
9 KNN
Last change: Coarse KNN
Accuracy (Validation): 25.7%
3087/3087 features
10 Ensemble
Last change: Boosted Trees
Accuracy (Validation): 50.8%
3087/3087 features
11 Neural Network
Last change: Narrow Neural Network
Accuracy (Validation): 54.0%
3087/3087 features
12 Neural Network
Last change: Medium Neural Network
Accuracy (Validation): 53.3%
3087/3087 features
Model 1
Model 3
Model 4
Model 5
Model 6
Model 7
Model 8
Model 9
Summary
Validation Confusion Matrix
Model 10
True Class
Predicted Class
1
2
3
4
5
51 10 10 10 8
5 72 10 2
22 11 30 12 12
22 2 14 33 18
17 9 24 39



1.3 Model Selection and Classification Algorithms

Model selection involves choosing an appropriate machine learning model or algorithm for the task at hand. In the context of music classification, you would want to select a model that can effectively learn patterns and relationships in the features extracted from audio signals. Once a model is selected, the next step is to train and evaluate it using the labeled dataset. MATLAB provides a rich set of functions and tools for implementing various classification algorithms. The music classification system achieved an accuracy of 59,6% with Quadratic SVM algorithm.

Fine Tree	Gini’s diversity index :100 splits	44.5
Medium Tree	Gini’s diversity index :20 splits	44.9
Coarse Tree	Gini’s diversity index :4 splits	41.8
Linear SVM	Kernel function : linear	57.8
Quadratic SVM	Kernel function : quadratic	59.6
Fine KNN	Euclidean distance : number of neighbours 1	39.7
Medium KNN	Euclidean distance : number of neighbours 10	35.7
Coarse KNN	Euclidean distance : number of neighbours 100	25.7
Boosted Trees	AdaBoost : 20 splits , 30 learners , learning rate 0.1	50.8
Narrow Neural Network	Activation : ReLU first layer size 10 , iteration limit 1000	54.0
Medium Neural Network	Activation : ReLU first layer size 25 , iteration limit 1000	53.3

1.4 Graphical User Interface

A graphical user interface has been designed in Matlab for the classification based on the selected machine learning method and recommendation of music in that genre. The Matlab App Designer has been used to design the interface in Matlab. Firstly, a form has been created, and labels, buttons have been added to the form (Figure 2).

By pressing the load button user can load music from data set. By machine learning the program shows the genre of the music and accuracy rate.

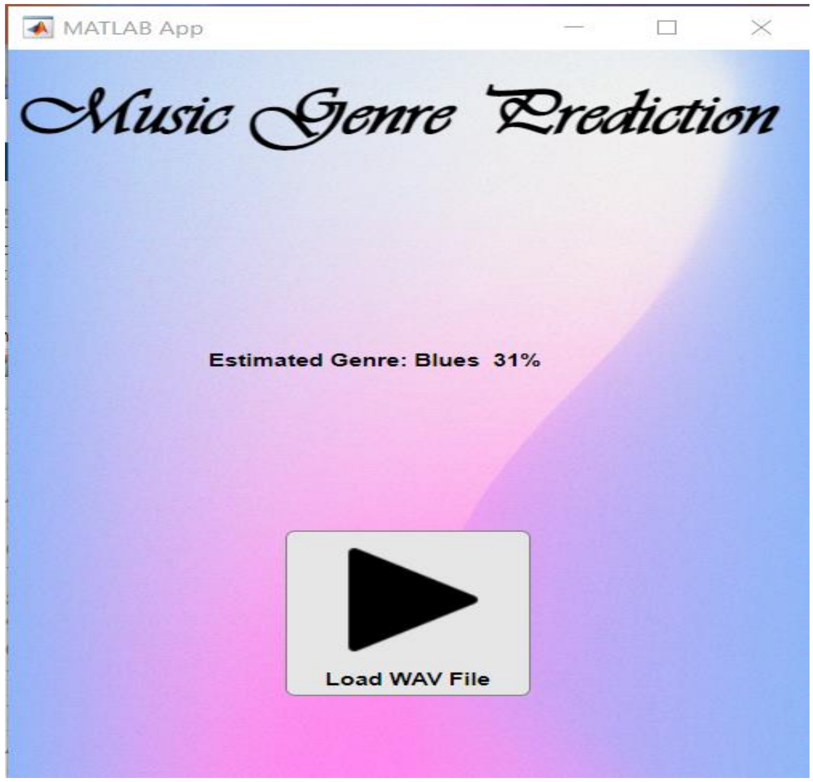


Figure 2. Graphical User Interface

As seen in the Figure 2, the user loads a music in the blues genre and the estimated genre shows that the music belongs to blues genre with the accuracy rate of 31%.

Authors' Contributions

The model training and software development processes were done by MM.Y , GUI and optimization parts were done by AC.Ö , report and visualization parts were done by R.M. The research paper was read and approved by all authors.