Classification of fMRI Brain Images Based on Music Genres

Andy Jiang **Chloe Cho** **Aryan Jindal** **Leonardo Medina**

July 25, 2024

Abstract

Functional magnetic resonance imaging (fMRI) brain images were analyzed to investigate the neural correlates of music processing in response to five distinct genres: Ambient, Country, Heavy Metal, Rock 'n Roll, and Classical Symphonic. Building upon previous work with a subset of 20 participants, we employed multiple machine learning algorithms (K-Nearest Neighbors, Support Vector Machine, and regression models) to improve classification accuracy. Our results demonstrate a significant improvement in accuracy, highlighting the potential of music-based analysis for therapeutic applications. The study contributes to the growing body of research on the cognitive effects of music, offering a novel methodology for analyzing fMRI data.

***K*eywords** Music, fMRI, Machine Learning, Brain, Cognition

1 Introduction

The interplay between music and neural activity is a topic long explored by researchers in which recognizing the effects of music may help in collaborative therapeutic settings with the aim of improving emotional, physical, and mental health. Understanding how different musical genres influence neural activity can provide important insights into human cognition and emotional processing. Functional Magnetic Resonance Imaging (fMRI) allows researchers to look into these factors by capturing real-time brain activity during music listening tasks.

We base our framework around the fMRI data collected from participants listening to music across five distinct genres: Ambient, Country, Heavy Metal, Rock 'n Roll, and Classical Symphonic. Each fMRI image is characterized by a set of features extracted from raw brain scans, accompanied by corresponding genre labels encoded numerically (0 to 4). This dataset includes training data (train_data.csv), and labels (train_labels.csv), which have been split into training and testing sets (X_train, X_test, y_train, y_test) for model training and evaluation.

The goal of this study is to evaluate and compare the performance of various classifying models such as K-Nearest Neighbors (KNN), Random Forest Model, Support Vector Machines (SVM), and regression models to accurately classify music genres based on fMRI data. The accuracy and other metrics of these models will allow further insight into the different neural activities associated with different music genres.

The output of this project includes predicted genre labels for the test dataset, detailed evaluation metrics (such as accuracy, confusion matrix, Mean Squared Error (MSE), and R-squared), and visualizations to illustrate model performance and comparative analyses.

1.1 Approach

Our approach involves a multi-step methodology to classify music genres based on fMRI data. Initially, we preprocess the dataset to ensure that it is suitable for model training and evaluation. This involves standardizing the features and, if needed, applying dimensionality reduction methods such as Principal Component Analysis (PCA). However, early testing revealed that the nature of the data limited the usefulness of PCA. We then explored and compared several classification models, as well as regression models, to determine which performs best in accurately predicting music genres from the fMRI data. To maximize performance, each model goes through a comprehensive cross-validation and tuning process. Ultimately, we assess the models based on multiple criteria, including confusion matrix and accuracy, providing a thorough evaluation of their performance in this task.

2 Methodology

2.1 Dataset

The input data for this project comprises fMRI brain images taken while subjects were listening to music from five different genres. The data is provided in two files: 'train_data.csv' and "train_labels.csv". Each fMRI image is represented by a set of features extracted from the raw brain scan data, and the corresponding labels indicate the genre of music the subject was listening to.

train_data.csv: This file contains the training data, with each row representing an fMRI image's features. There are no headers in this file, and the data is structured in a way that each column represents a different feature extracted from the fMRI images. This file has 200 samples, each with 22036 features.

train_labels.csv: This file contains the labels corresponding to the training data. Each row in this file represents the music genre label for the corresponding fMRI image in train_data.csv. The genre is encoded as integers from 0 to 4, where 0 = Ambient, 1 = Country, 2 = Heavy Metal, 3 = Rock, and 4 = Classical Symphonic.

The study also provided testing_data.csv, which contained data from 50 samples. However, without the labels, it is difficult to utilize the testing data.

2.2 Preprocessing

We split the given training data into training and testing data using a 80-20 split. Since there are vastly more features than samples, we used Principal Component Analysis (PCA) to reduce the dimensionality of the dataset. However, this led to a decrease in accuracy by 0.2 on average, suggesting that there are too many important features. We utilized k-fold cross-validation to compare models based on accuracy metrics. We used a higher number of folds because we have a relatively small data set.

2.3 KNN

We constructed several KNN models to find the k with the highest accuracy. Across multiple tests, $k = 8$ consistently gave the highest accuracy, ranging from 0.45 to 0.55. The accuracy was inconsistent for higher values of k because of the small sample size, since the models with a high k value tended to underfit.

The model performed very well at correctly classifying 0 (Ambient), which suggests that ambient music has a very distinct effect on blood oxygenation levels. KNN may not be the best model for this specific task because it is limited by the small dataset. Models that create generalizations, such as SVM, are better suited for small datasets.

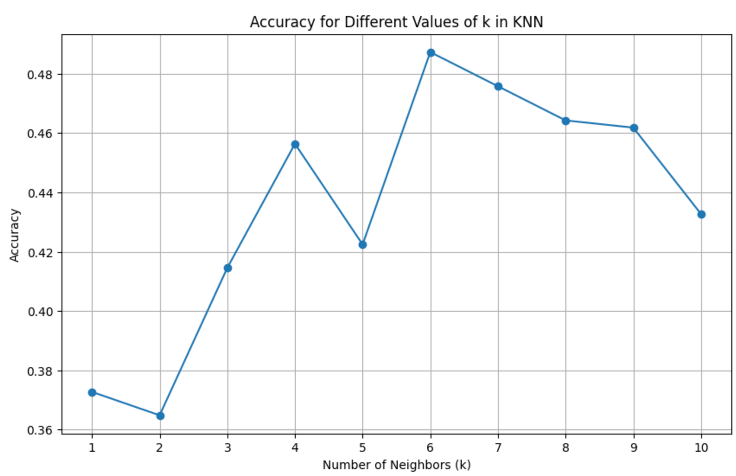

Figure 1: Accuracy for Different Values of k

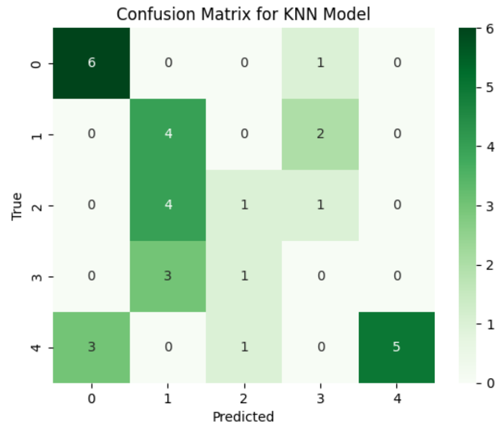

Figure 2: KNN Confusion Matrix

2.4 Random Forest

Our results show that the combination of 100 decision trees produced the highest level of accuracy among all tested configurations. Increasing the number of trees generally led to improved accuracy, but this was achieved at the cost of increased computational requirements.

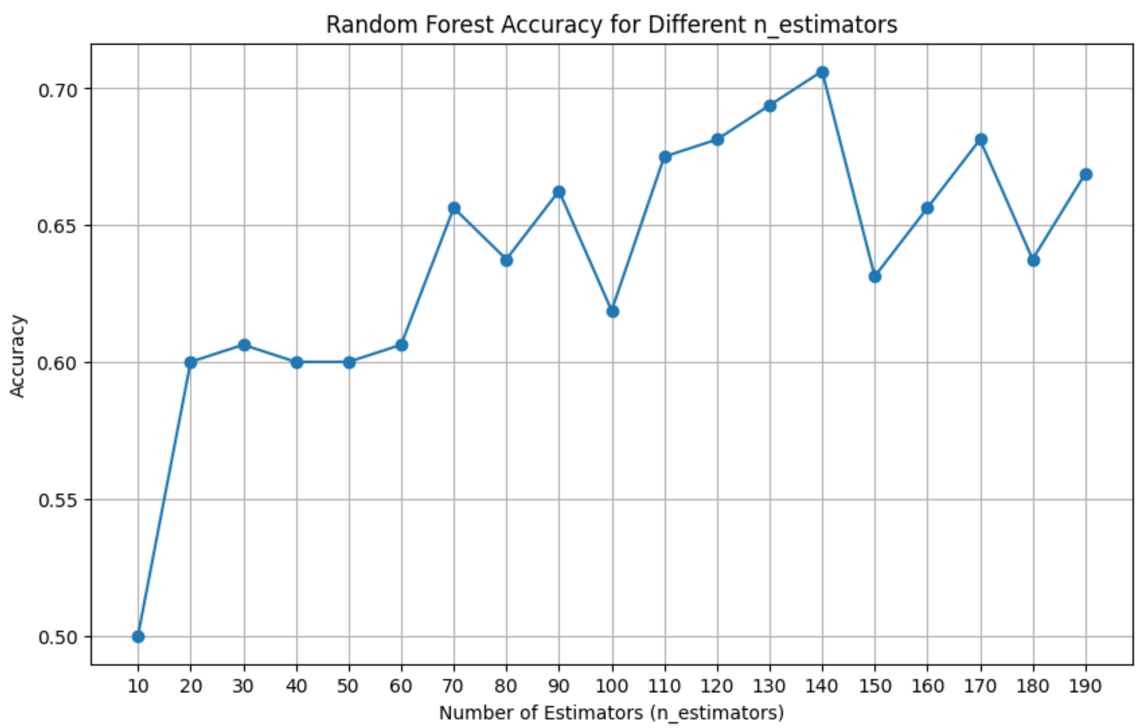

Figure 3: Accuracy for Number of Trees

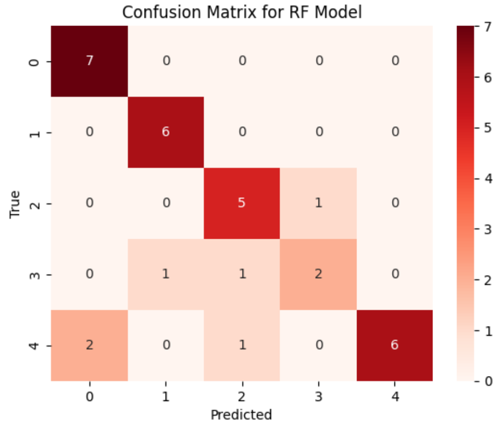

Figure 4: Random Forest Confusion Matrix

2.5 SVM

The SVM model was evaluated on a dataset with a linear kernel yielding an accuracy of 81.25%, which was the highest out of all of the kernels. The polynomial kernel gave less than 40%, and the RBF and sigmoid kernel yielded results between 50% and 60%. The dataset preview highlights significant variability in the feature values across instances. The model's decision boundary was influenced by the high-dimensional feature space, with the coefficients indicating the weights assigned to each feature. The label distribution shows a balanced representation of classes, with the first four instances labeled as 0 and the fifth as 1. The accuracy of 81% suggests that the model performs well in distinguishing between the classes, although further validation on a larger and more diverse dataset might be more suitable to ensure the validity of the model.

2.6 Evaluation

In order to evaluate the performance of our classification models, we employed several metrics, each providing unique insights into the model's effectiveness. The primary metric used was accuracy, which measures the proportion of correctly predicted labels out of the total number of predictions. Additionally, for our regression model, we used Mean Squared Error (MSE) and R-squared (R^2) as evaluation metrics.

3 Results

Model	Accuracy
KNN (k = 8)	0.45 - 0.55
Random Forest (100 trees)	0.65 - 0.70
SVM (Linear)	0.81

a. **Random Forest**: The Random Forest algorithm is a powerful ensemble learning method that aggregates multiple decision trees to enhance predictive accuracy and mitigate overfitting, making it particularly robust to noisy data, a common issue in fMRI datasets. Its ability to handle extensive feature sets makes it an excellent fit for the high-dimensional nature of fMRI data, allowing it to manage the complexity and variability inherent in brain imaging data, which is relevant to our project.

b. **K-Nearest Neighbors (KNN)**:
As a simple, non-parametric algorithm, KNN makes predictions based on the proximity of data points in the feature space, making it adept at capturing local patterns in the data. This ability can be crucial for identifying subtle differences in neural responses to different music genres. Although KNN can be computationally expensive with large datasets, its simplicity and effectiveness in capturing local patterns justify its inclusion in our study.

c. **Support Vector Machine (SVM)**: SVM excels in high-dimensional spaces and is effective when the number of features exceeds the number of samples, a characteristic that is particularly advantageous for our fMRI dataset, which contains numerous features extracted from brain images. SVM is also particularly effective for classification tasks with clear margin separations between classes, aligning well with our goal of distinguishing between different music genres based on neural activity.

4 Conclusion and Future Work

In this comprehensive study, advancements in machine learning-empowered music were explored, explicitly focusing on classification models. By implementing KNN, Random Forest, and SVM models, the authors compared which model served to be most effective in classifying fMRI brain images based on five different music genres. Using these models, the authors sought to understand how different musical genres influence neural activity and how they can provide important insights into human cognition. The SVM model emerged as the most accurate, with an accuracy of 81%, suggesting its superior efficacy in this context. With this machine learning model, we aim to contribute to the applications of music genres in therapeutic settings, particularly by contributing to research for the most effective ways of analyzing fMRI data.

Although this project was successful, we believe that we could expand upon numerous aspects to further research in this field. For example, we believe that if we employ feature extraction techniques and advanced architectures like CNNs and RNNs to analyze fMRI data for music genre classification, our model would be more accurate. Furthermore, a limitation that the dataset experienced was that there were only twenty subjects in the original project. If there had been more subjects, the accuracy rates could have been higher. Conducting this study over time would also be beneficial for researchers to observe changes in brain responses to music and to assess the impact of repeated exposure to different genres. With a larger subject group, we could incorporate participants of various backgrounds. By including a varied demographic of participants, we can explore how age, cultural background, and musical preferences influence neural responses.

5 Acknowledgements

We would like to thank Miss. Haripriya, Mr. Shokhruz and Mehta+ for assisting us with this research paper.

References

Casey, M. A., *Music of the 7ts: Predicting and decoding multivoxel fmri responses with acoustic, Schematic, and categorical music features*. Frontiers, June 2017.