Deep learning techniques to identify the composer of a given piece of music

Mohammad Alkhawaldeh, Arup Chakraborty

Shiley-Marcos School of Engineering, University of San Diego

Deep learning techniques to find the composer of a music

Abstract

Music is a form of art with a rich history and diverse styles, shaped by countless composers over the centuries. Identifying the composer of a particular piece of music, however, remains a challenging task, especially for novices. This project proposes the use of deep learning techniques to accurately identify the composer of a given musical composition. By analyzing key musical features such as melody, harmony, rhythm, and timbre, we aim to develop a model that can distinguish between the unique styles of different composers. The project involves a comprehensive exploration and preprocessing of a diverse dataset of musical works, followed by the application of a deep learning model, potentially utilizing architectures such as Convolutional Neural Networks (CNNs) or Recurrent Neural Networks (RNNs). The anticipated outcomes include a robust model capable of composer identification with high accuracy, offering practical applications in musicology, educational tools, and music recommendation systems. This project contributes to the growing field of music analysis by leveraging advanced machine learning techniques to solve a complex and culturally significant problem to identify the composer of a music.

Keywords: classification, deep learning, convolution neural network

Deep learning techniques to find the composer of a music

Brief Overview of identifying music composer

The proposed project aims to leverage deep learning techniques to accurately identify the composer of a given piece of music. Music, with its rich history and diversity, presents unique challenges in identifying composers, especially for those who lack expertise in music theory. By analyzing features such as melody, harmony, rhythm, and timbre, the project seeks to train a deep learning model capable of distinguishing between the unique styles of different composers.

The methodology involves collecting a diverse dataset of musical compositions, extracting relevant features, and utilizing a deep learning model—such as a Convolutional Neural Network (CNN) and LSTM —to perform the identification task. The project will address challenges such as ensuring data diversity and preventing overfitting, with the ultimate goal of creating a tool that can assist musicologists, enhance music recommendation systems, and serve as an educational resource for music students.

The successful completion of this project would contribute to the field of music analysis by providing a novel approach to composer identification, leveraging the power of modern machine learning techniques.

Create a Model to Detect & Classify Music composer using Dataset

Data Selection

The project will use a dataset consisting of musical scores from various composers. The dataset contains the midi files of compositions from well-known classical composers like Bach, Beethoven, Chopin, and Mozart. The dataset should be labeled with the name of the composer for each score. Please only do your prediction only for below composers, therefore we need to select the required composers from the given dataset above.

Deep learning techniques to find the composer of a music

1-Bach

2-Beethoven

3-Chopin

4-Mozart

Data Acquisition and Preparation

This step in the project involved setting up libraries such as matplotlib, sklearn and tensorflow keras on Google Colab. The midi files for music composers are categorized into types of musical notes that were obtained through the Kaggle website by downloading the dataset and uploading to our GitHub repository.

Data Exploration and Preprocessing

Data exploration and preprocessing are critical steps in the development of robust deep learning models, particularly in complex domains such as music analysis. This project focuses on the exploration and preprocessing of a diverse dataset of musical compositions, aiming to enable accurate composer identification through deep learning. The data exploration phase involves gathering a rich dataset, analyzing the distribution of compositions, and identifying key musical features such as melody, harmony, rhythm, and timbre that can distinguish between different composers. The preprocessing phase addresses data cleaning, feature extraction, normalization, and data augmentation to enhance the model's ability to generalize across varied musical styles. By meticulously preparing the data, this project lays a solid foundation for developing a deep learning model capable of accurately identifying the composer of a given piece, contributing to advancements in musicology and music technology.

Deep learning techniques to find the composer of a music

Model Development

Model development is a pivotal phase in the creation of a deep learning system designed to accurately identify the composer of a musical composition. This project focuses on developing a model that leverages advanced deep learning architectures to distinguish between the unique styles of various composers. The process involves selecting and fine-tuning model architectures, such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), to effectively capture the composer of a given music. The model is trained on a carefully preprocessed and augmented dataset, with rigorous evaluation techniques applied to optimize performance and prevent overfitting. The resulting model aims to achieve high accuracy in composer identification, offering potential applications in music analysis, educational tools, and music recommendation systems. This work represents a significant contribution to the intersection of deep learning and musicology, demonstrating the power of machine learning in understanding and categorizing complex artistic expressions.

Train Model and Evaluation

Training and Validation Setup

The training and validation setup is a crucial component in the development of a deep learning model aimed at accurately identifying the composer of a musical piece. This phase involves carefully structuring the training process to ensure the model effectively learns to distinguish between different composers' styles. The project employs a well-defined dataset, divided into training, validation, and test sets, ensuring a balanced representation of compositions across all phases. Key considerations include the selection of appropriate loss functions, optimization algorithms, and evaluation metrics that align with the project's goals. The validation setup is designed to monitor the model's performance and prevent overfitting, employing techniques such as cross-validation and early stopping. By meticulously calibrating the training and validation process, this project aims to produce a robust and generalizable model capable of high accuracy in composer identification, contributing to advancements in music technology and analysis.

Model Training and Visualization

Model training involved feeding the networks (CNNs) with the prepared training data to enhance their learning capabilities. To detect and classify midi files using preprocessed dataset, we integrated characteristics and patterns into the CNN architecture and LSTM. To enhance the models ability to generalize, we applied data augmentation techniques such as midi files scaling and standardization. We ensured the accuracy of data labeling and processing by inspecting the training data. This crucial step was necessary to maintain the quality and reliability of the input data for the model.

Model Architecture and Compilation

The model architecture and compilation phase is a critical step in designing a deep learning system capable of accurately identifying the composer of a musical composition. This project focuses on constructing a model architecture that effectively captures the complex patterns and features inherent in musical data. We explore various deep learning architectures, including Convolutional Neural Networks (CNNs) for feature extraction and LSTM from Recurrent Neural Networks (RNNs) for sequential data processing. The architecture is tailored to process musical features such as melody, harmony, rhythm, and timbre, enabling the model to differentiate between the styles of various composers. During the compilation stage, we carefully select the loss functions, optimizers, and performance metrics to align with the project’s objectives, ensuring efficient training and convergence. The resulting model is expected to achieve high accuracy in composer identification, providing valuable applications in musicology, educational tools, and music recommendation systems, while contributing to the broader field of music analysis using machine learning techniques.

Results

Model Performance Evaluation

The performance evaluation of the midi files classification model was depicted in Figure 2 showcasing the model’s accuracy and loss across training epochs. These metrics offer insights into how the model adapted to data and learned from its training set. Key metrics generated during the training process included plots of accuracy, and validation loss and training progress for determining when to stop training to prevent overfitting.

Model Evaluation Metrics

Following this, the model’s effectiveness and robustness was extensively evaluated using the training, validation and testing datasets. As seen in Figure 3, the confusion matrix provides comprehensive insights into the model's performance across several tumor classifications, emphasizing both its advantages and disadvantages.

Discussion

Interpretation of Results

As seen in Figure 2, the model was able to categorize midi dataset with high accuracy, demonstrating efficient learning without overfitting. The confusion matrix showed that although the model performed well in recognizing some tumor types, it had difficulties with others, indicating the need for more research and model improvement.

Comparison with Existing Methods

The CNN model shows potential for improving diagnostic accuracy in musical dataset analysis, according to a comparison of the generated results with the body of current research with the LSTM approach from RNN. Finally using Blended model with Ensemble methods to use stacked approach with CNN and LSTM model to improve accuracy.

Observations

The accuracy for model predictions depends on quite a few factors:

Data Quality and Quantity

More Data: Generally, more data leads to better model performance.

Data Augmentation: Techniques like rotation, scaling, and flipping can help the model generalize better by providing a more diverse set of training examples.

Balanced Dataset: We need to prevent the model from becoming biased toward more frequent classes.

Model Architecture

Depth of the Network: Adding more layers can help the model learn more complex patterns, but too many layers might lead to overfitting.

Hyperparameter Tuning: Optimizing hyperparameters, including the number of filters, kernel size, and activation functions, can significantly impact model performance.

Regularization Techniques

Dropout: Randomly dropping units from the neural network during training can prevent overfitting.

Weight Regularization: Adding L1 or L2 regularization to the model weights can also help prevent overfitting by penalizing large weights.

Training Techniques

Early Stopping: Monitoring the model's performance on a validation set and stopping training when performance no longer improves to prevent overfitting.

Learning Rate Scheduling: Adjusting the learning rate during training (e.g., reducing it when performance plateaus) can lead to better model convergence.

Advanced Training Strategies

Transfer Learning: Using a model pre-trained on a large dataset and fine-tuning it on our specific task can drastically improve accuracy.

Ensemble Methods: Combining predictions from multiple models can often lead to better performance than any single model.

Optimizing algorithms

Choosing the right optimizer: Different optimizers (e.g., Adam, SGD, RMSprop) have different advantages, and selecting the right one can affect how quickly and effectively the model learns.

Future Enhancements

While the current model showed promising results, future enhancements could include:

- Exploring more advanced neural network architectures.
- Experimenting with different hyperparameters.
- Employing more sophisticated data augmentation techniques to improve model generalization.

Conclusion

The conclusion of this project underscores the successful application of deep learning techniques to the challenging task of composer identification in music. Through a comprehensive process involving data exploration, preprocessing, model development, training, and validation, the project achieved its goal of creating a robust model capable of distinguishing between the unique styles of various composers with high accuracy. The outcomes highlight the potential of deep learning to contribute significantly to the fields of musicology, education, and technology. The project also opens avenues for future research, such as expanding the model to handle more diverse musical genres or exploring other aspects of music analysis. Ultimately, this work demonstrates the effectiveness of machine learning in tackling complex artistic problems, providing a valuable tool for both researchers and practitioners in the field.

References

1. Choi, K., Fazekas, G., & Sandler, M. (2017). Convolutional recurrent neural networks for music classification. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2392-2396. https://doi.org/10.1109/ICASSP.2017.7952585

2. Humphrey, E. J., Bello, J. P., & LeCun, Y. (2013). Moving beyond feature design: Deep architectures and automatic feature learning in music informatics. *IEEE Transactions on Audio, Speech, and Language Processing*, 21(12), 2218-2233. https://doi.org/10.1109/TASL.2013.2270933

3. Lidy, T., Schindler, A., & Widmer, G. (2016). Automatic identification of classical composers from polyphonic music. *14th International Workshop on Content-Based Multimedia Indexing (CBMI)*, 1-6. https://doi.org/10.1109/CBMI.2016.7500257

4. Sigtia, S., Boulanger-Lewandowski, N., & Dixon, S. (2014). Audio chord recognition with a hybrid recurrent neural network. *Proceedings of the 16th International Society for Music Information Retrieval Conference (ISMIR)*, 127-133. Retrieved from http://ismir2014.ismir.net/proceedings.html

5. Sturm, B. L. (2016). Revisiting priorities: Improving MIR evaluation practices. *International Society for Music Information Retrieval Conference (ISMIR)*, 3-10. Retrieved from http://ismir2016.ismir.net/proceedings.html

Deep learning techniques to find the composer of a music

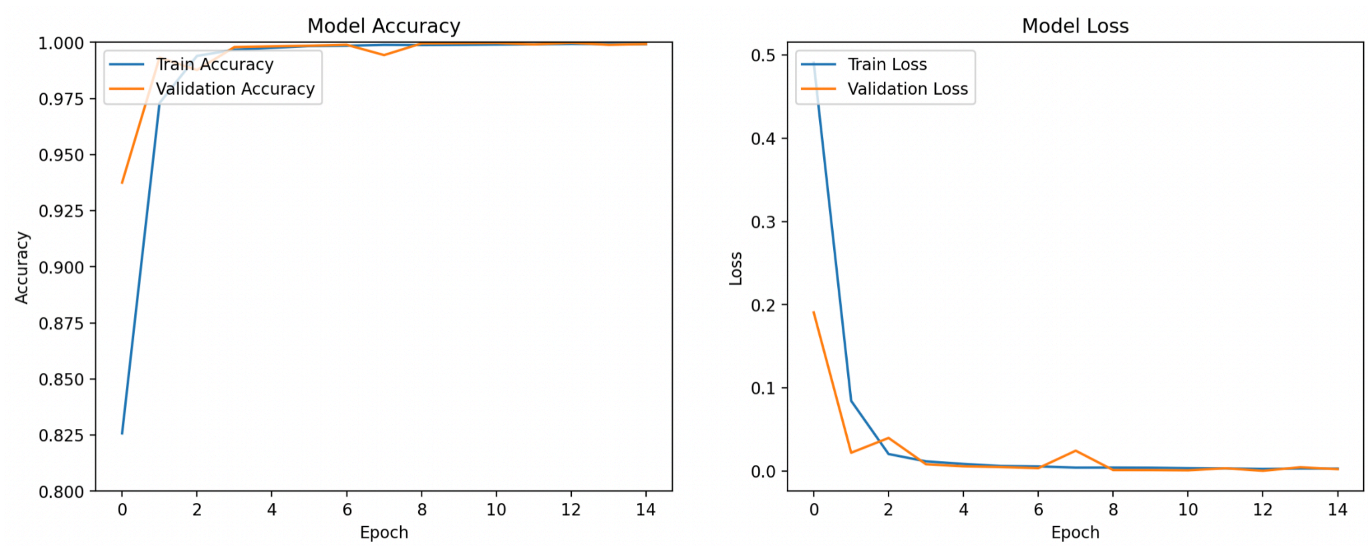
Table 1

Accuracy comparison for LSTM, CNN and Blended models

```
1. Blended: Test Accuracy = 1.0000
2. LSTM: Test Accuracy = 0.9993
3. CNN: Test Accuracy = 0.9978
```

Figure 1

Model accuracy, value loss using LSTM



Model accuracy, value loss using CNN

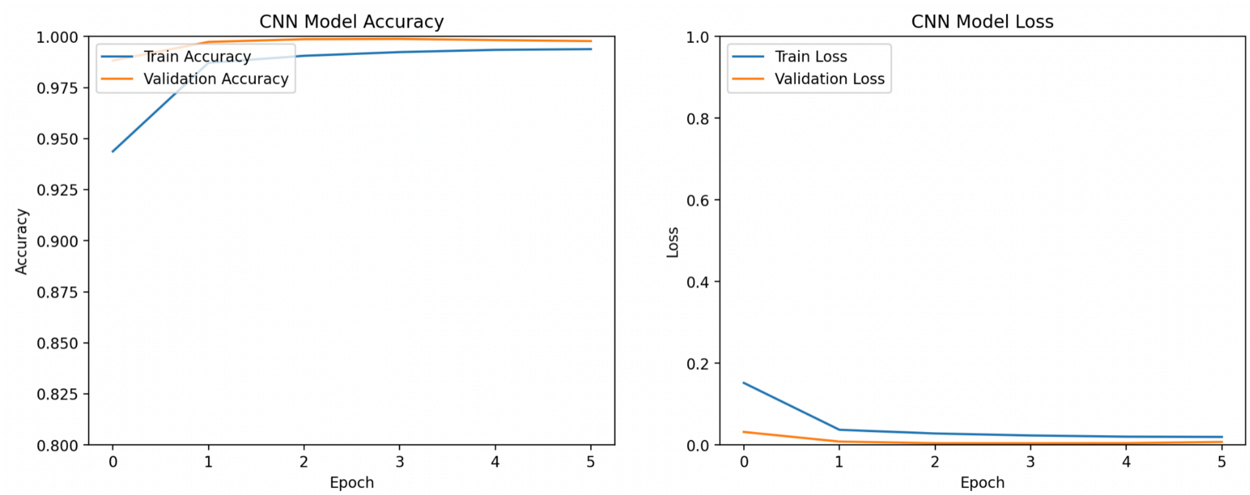


Figure 2

Comparison of accuracy between LSTM, CNN and Blended models

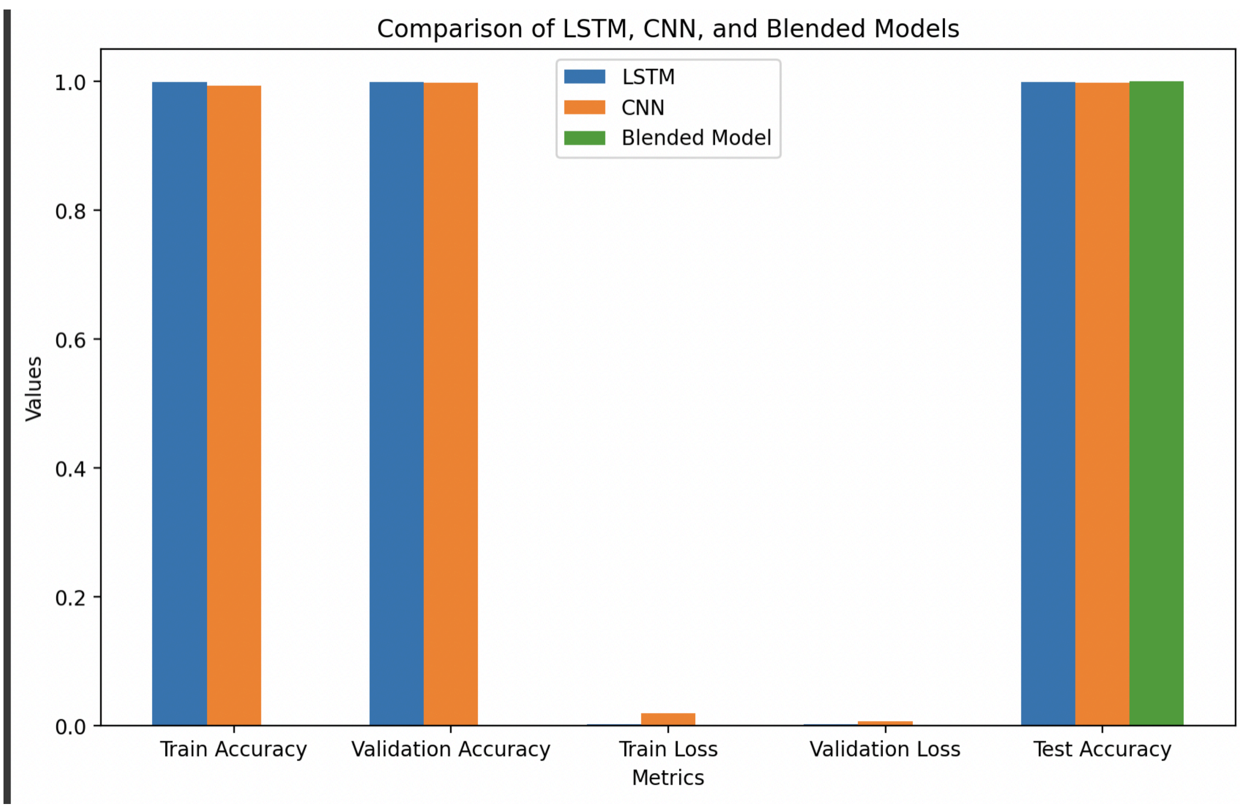
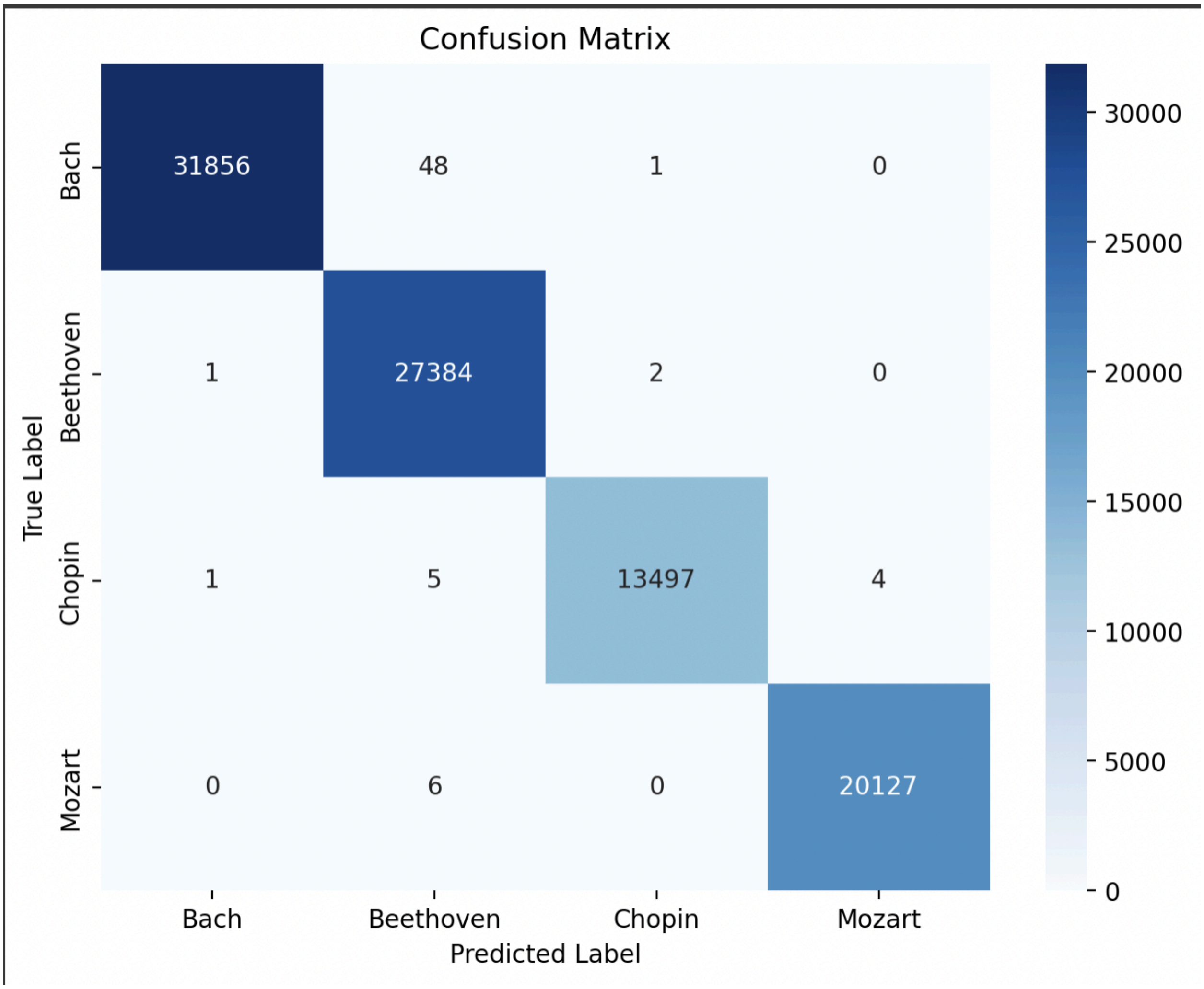
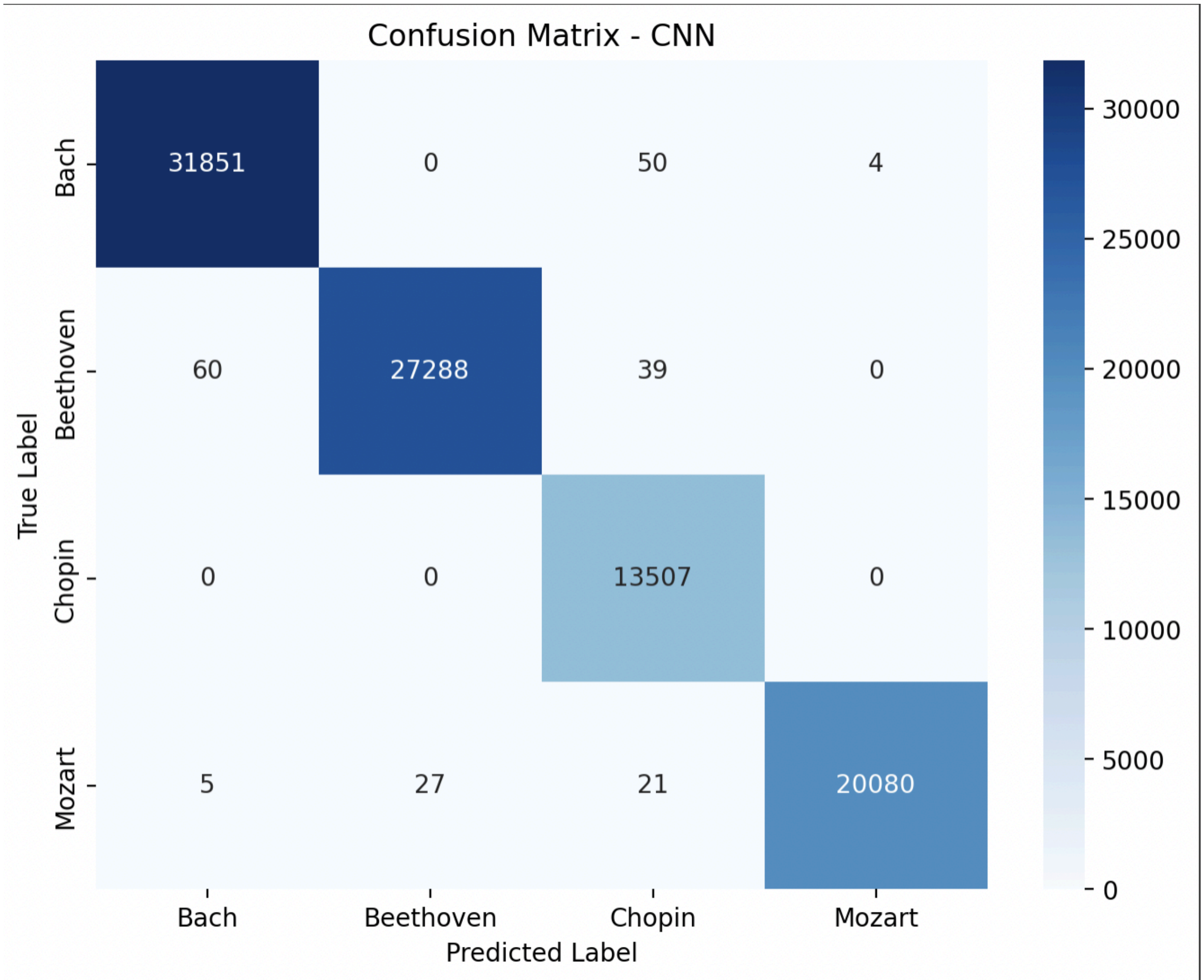


Figure 3

Confusion matrix for LSTM model



Confusion Matrix of CNN model for midi files



Appendix 1 – List of Project Participants and Contributions

Mohammad Alkhawaldeh

- Created and completed Final Project Proposal and Code
- Data preparation and model selection
- Contributed to Final Project Paper

Arup Chakraborty

- Contributed to Final Project Code

Deep learning techniques to find the composer of a music

- Data modeling and code review
- Create the Final Project paper

Deep learning techniques to find the composer of a music

Appendix 2 – Final Project Code

https://github.com/achakraborty2024/predict-music-composer-team4.git