Deep Learning Individual Report – Group 1
Name- Sairam Venkatachalam

<u>Topic: Classical Music Generation</u>

INTRODUCTION

The fusion of art and technology has ushered in a new era of musical creation, where Neural Network models are employed to craft classical compositions. This project explores the predictability of musical patterns in classical music, aiming to create authentic compositions using advanced techniques and algorithms. Leveraging a dataset of classical music compositions in MIDI format, our goal was to push the boundaries of music generation and explore the potential of Neural Network models in creating compositions that resonate with the rich legacy of classical music.

PROJECT FLOW

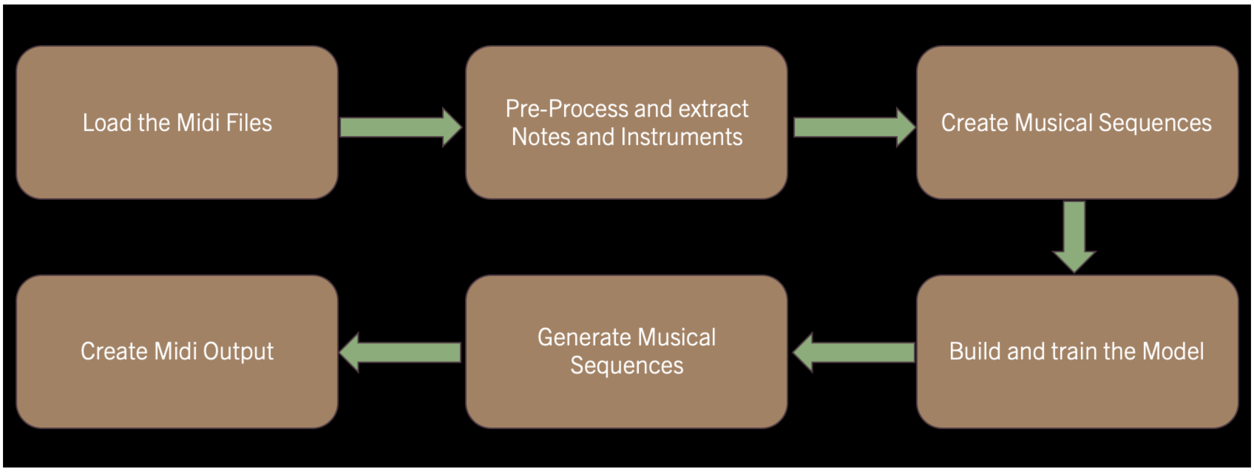

Figure.1. Project FLow

We start by loading MIDI files of a selected composer's classical music compositions. Then, we preprocess these files to extract important musical elements like notes and instruments. Next, we organize these elements into sequential representations, setting the stage for model training. Our neural network model, powered by LSTM, learns temporal dependencies within these sequences during the training phase. Once trained, the model autonomously generates new musical sequences, drawing from its learned patterns. Finally, these generated sequences are converted into MIDI format, resulting in unique classical compositions.

Data Preparation

One of my contributions to the project focused on data preparation, which played a pivotal role in understanding and harnessing the intricacies of classical music compositions. Using Python libraries such as PrettyMIDI, I loaded MIDI files corresponding to a specific composer and extracted important musical elements such as instruments and notes. This involved meticulously extracting attributes like pitch, velocity, start time, and end time for each musical note, laying the foundation for constructing meaningful musical sequences.

I had explored different ways to extract the data, such as looking at the piano roll directly , as well as looking at waveforms to extract more information, however, I was set on using a next note generation approach, as it mimics methods used in Large Language models for next tok

Setting Up the Models

In addition to data preparation, I played a key role in setting up the models for the project. This involved defining a sequence length, which governs the temporal span of input sequences and significantly influences the depth of contextual information considered during sequence generation. I also experimented with hyperparameters for sequence length and embedding dimensions to optimize the model's performance. Additionally, I explored an initial regression approach for predicting pitch values, which proved to be ineffective to generate note predictions, however, it was a crucial step in the learning process as I explored the use of embeddings and converted a regression problem to a classification problem.

Studying the Use of Embedding

One of the critical aspects of my contribution was studying the use of embedding to capture semantic information in the musical sequences. Embedding transformed categorical MIDI pitch values into continuous vector representations, enabling the model to capture nuanced relationships between musical elements more effectively. This transformation streamlined the learning process, leading to more coherent and expressive musical sequences.

Creating the Duration Model

Furthermore, I took the initiative to create the duration model, which extended our model to predict both pitch and duration of musical notes. This involved introducing a one-hot encoding scheme for duration prediction and building a separate LSTM model to handle duration prediction. Despite the challenges faced, such as setting appropriate sequence length and seed values, the duration model enhanced the model's capabilities, allowing us to experiment with diverse generation settings.

Conclusion

In conclusion, my contributions to the project focused on data preparation, setting up the models, and exploring innovative approaches to enhance the model's performance. By leveraging advanced techniques and algorithms, we were able to push the boundaries of music generation and create compositions that resonate with the rich legacy of classical music. Moving forward, there are several promising avenues for future research, including experimenting with diverse model architectures and incorporating attention mechanisms to capture long-range dependencies more effectively.

Code References

I leveraged minimal code from online sources, mainly for extracting notes from midi files and converting midi to wav in streamlit.

Code percentage = (32-8)/(32 + 288) = 7.5%