Eternal Echoes: Reviving Chopin through LSTM Networks

1st Farhan Ahmad
Computer Scienece
FAST NUCES
Islamabad, Pakistan

2nd Sohail Shahbaz
Computer Scienece
FAST NUCES
Islamabad, Pakistan

3rd Mubeeb Ur Rehman
Computer Scienece
FAST NUCES
Islamabad, Pakistan

***Abstract*—This document explores the innovative application of Long Short-Term Memory (LSTM) networks in the field of music and lyric generation. Through a comprehensive review of current research, we identify various approaches that utilize LSTM models to generate both musical compositions and corresponding lyrics, highlighting the diversity and capability of this technology in handling complex sequence data. We particularly focus on the application of LSTM networks to understand and generate music that reflects specific emotions, styles, or thematic content. The literature reveals a growing trend towards using LSTM for creating coherent and emotionally resonant music, often involving detailed preprocessing of datasets and rigorous model tuning to achieve desired results. Furthermore, the document discusses our methodology for leveraging LSTM networks to generate lyrics and music, providing insights into the technical challenges, dataset preparations, and the efficacy of different LSTM configurations. The synthesis of this knowledge underscores the potential of LSTM networks to revolutionize music generation, making it a promising area for further research and application.**

I. Literature Review

A. *LSTM Based Music Generation System*

Research Question or Problem: The paper addresses the problem of automating music generation using recurrent neural networks, specifically focusing on the use of LSTM networks to generate musical notes. [1]

Methods: The authors propose and implement a single-layered LSTM network model to learn sequences of polyphonic musical notes. They discuss methods for preprocessing MIDI file data for use in the model, and provide a detailed description of the model's architecture and the training process.

Datasets: The dataset used consists of MIDI files, which represent musical notes and their sequences. These files serve as the primary input for training the LSTM model.

Accuracy Metrics: The paper details the evaluation of the model based on the distribution of weights and biases, loss metrics, and accuracy measurements for each training batch. However, specific numeric accuracy metrics such as precision or recall are not mentioned.

Contribution to Understanding the Problem and Potential Solutions: This paper contributes to the field by demonstrating the potential of LSTM networks in learning and generating complex sequences of musical notes, thus offering a viable solution for automated music generation. The detailed discussion of LSTM architecture and its application to music provides insights that could be beneficial for further research in both music technology and neural network design.

Identify applicable funding agency here. If none, delete this.

B. *Music Generation with LSTMs*

Research Question or Problem: The paper explores the use of LSTM networks in music generation, focusing on overcoming the complexity of training neural networks to generate rhythmic music that is appealing across different musical tastes. [2]

Methods: The method discussed involves using TensorFlow and Keras to implement an LSTM model for generating music. The article also details the use of the pretty_midi library for handling MIDI files, essential for training the LSTM model.

Datasets: The dataset mentioned is the MAESTRO dataset, which consists of multiple MIDI files with numerous piano notes, suitable for training the LSTM model.

Accuracy Metrics: The article does not specify particular accuracy metrics but discusses the importance of model training, loss metrics, and the visual inspection of generated music.

Contribution to Understanding the Problem and Potential Solutions: This article contributes by detailing the steps and code necessary to implement an LSTM music generation model, providing a practical example of how deep learning can be applied to music generation. The use of MIDI files and detailed code snippets offers a valuable resource for similar projects.

C. *Generating Music using an LSTM Network*

Research Question or Problem: The paper investigates the capability of LSTM networks to generate polyphonic music, focusing on the ability of these models to retain and utilize musical structure and coherence over time. [3]

Methods: The researchers propose a Bi-axial LSTM network architecture, which is trained using a method that integrates elements of convolutional kernels. This approach aims to enhance the network's ability to model and predict complex musical sequences.

Datasets: While the specific dataset details are not mentioned in the abstract, it typically involves polyphonic music data suitable for training LSTM models.

Accuracy Metrics: The performance of the proposed model is assessed through both quantitative and qualitative analyses, although specific metrics like precision or recall are not detailed in the abstract.

D. Music Generation Using LSTM and Its Comparison with Traditional Method

Research Question or Problem: The paper explores the effectiveness of LSTM neural networks in generating music compared to traditional RNN methods, addressing issues like vanishing and exploding gradients. [4]

Methods: The authors utilize an LSTM model for music generation, detailing their approach in preprocessing MIDI files, model architecture, and the specific configurations used for training. They emphasize the model's training loss in relation to the number of epochs to optimize performance.

Datasets: The study uses MIDI files specifically containing piano music, focusing on both notes and chords for training the LSTM network.

Accuracy Metrics: The paper evaluates the LSTM model using training loss metrics and visual inspections of the learning curves and generated sheet music.

Contribution to Understanding the Problem and Potential Solutions: This paper highlights the advantages of LSTM over traditional RNN by demonstrating improvements in handling gradient issues and enhancing music generation quality. It provides a comparative analysis that supports LSTM's superiority for complex sequence learning tasks like music generation.

E. Music Generation Using an LSTM

Research Question or Problem: This study explores the application of Long Short-Term Memory (LSTM) networks for generating music, addressing the challenge of creating coherent and complex musical sequences. [7]

Methods: The authors developed an LSTM-based Recurrent Neural Network (RNN) to generate music sequences, detailing the network architecture, training process, and specific challenges like managing overfitting and optimizing sequence prediction.

Datasets: The dataset includes MIDI files, particularly focusing on sequences of pitches and durations to construct songs. Details about the exact size and diversity of the dataset are generalized in the abstract.

Accuracy Metrics: The paper discusses qualitative measures, such as the coherence and aesthetic quality of generated music, rather than traditional quantitative metrics.

Contribution to Understanding the Problem and Potential Solutions: The paper contributes insights into the use of LSTM for music generation, particularly its capability to retain and utilize musical structures over time effectively. It also highlights the potential future improvements to enhance the model's performance.

F. LSTM based Lyrics and Music Generation Model

Research Question or Problem: The article discusses the development of a model for generating lyrics and music using Long Short-Term Memory (LSTM) networks, aimed at producing songs that reflect specific emotional themes. [8]

Methods: The author describes the use of LSTM networks for both lyrics and music generation. The model takes sequences of words as input to generate lyrics and then uses these lyrics along with melodic accompaniments to generate music, all processed through LSTM networks.

Datasets: The dataset consists of lyrics collected from various artists across different music genres, categorized by emotions. The lyrics and their corresponding musical notes were collected from websites like All Music and songmeanings.com, and processed into training data.

Accuracy Metrics: While specific metrics like precision or recall are not detailed, the author discusses the qualitative assessment of the lyrics generated by the LSTM model and describes the need for tuning the model to improve output quality.

Contribution to Understanding the Problem and Potential Solutions: The article contributes to the field by showcasing how LSTM networks can be effectively used for generating complex content like music and lyrics, reflecting a nuanced understanding of language and music theory. It also discusses challenges like model tuning and the need for extensive preprocessing of data.

II. Data Set

This section outlines the dataset that will be utilized for the LSTM-based music generation project. The focus is on classical music MIDI files, specifically designed to train machine learning models capable of generating music.

A. Classical Music MIDI

[5] **Description:** This dataset comprises MIDI files from well-known classical composers. It is specifically curated to foster research in the field of algorithmic music generation using deep learning techniques. [6]

Contents:

- **Number of Files:** Approximately 400 MIDI files.
- **Composers:** Includes compositions from renowned composers such as Beethoven, Chopin, and Bach, among others.
- **Instrumentation:** Although the dataset contains various instruments, this project will focus solely on piano compositions (Chopin).
- **Data Type:** Sequential MIDI files suitable for processing and learning by LSTM networks.

Justification for Selection: The selection of this dataset is motivated by its focus on classical music, which provides structured and complex sequences ideal for training LSTM models. The diversity in composition styles and the high quality of MIDI encodings are crucial for exploring the depth of

music generation capabilities of LSTM networks. Additionally, focusing on Chopin's compositions allows for a detailed exploration of a single composer's style, enhancing the model's ability to generate stylistically coherent and harmonically rich outputs.

REFERENCES

[1] Mangal, S., Modak, R., & Joshi, P. (2019). "LSTM Based Music Generation System," *IARJSET*, vol. 6, no. 5, pp. 47-54. Available: https://doi.org/10.17148/IARJSET.2019.6508

[2] Bharath K. (2022). "Music Generation with LSTMs," *Paperspace Blog*. Available: https://blog.paperspace.com/music-generation-with-lstms/

[3] Kotecha, N., & Young, P. (2018). "Generating Music using an LSTM Network," *arXiv preprint arXiv:1804.07300*. Available: https://arxiv.org/abs/1804.07300

[4] Arya, P. K., Kukreti, P., & Jha, N. (2022). "Music Generation Using LSTM and Its Comparison with Traditional Method," *Advanced Production and Industrial Engineering*, doi:10.3233/ATDE220793.

[5] Rakshit, S. (Year). Classical Music MIDI. *Kaggle*. Available: https://www.kaggle.com/datasets/soumikrakshit/classical-music-midi

[6] MIDI World. (Year). Classical Archives. *MIDI World*. Available: https://www.midiworld.com/classic.htm

[7] Conner, M., Gral, L., Adams, K., Hunger, D., Strelow, R., & Neuwirth, A. (Year). "Music Generation Using an LSTM," *Milwaukee School of Engineering*. Available: https://example.com/full-link-to-paper

[8] Kumar, A. S. (2020). "LSTM based Lyrics and Music Generation Model," *Medium*. Available: https://medium.com/@aishwaryasathyakumar/lstm-based-lyrics-and-music-generation-model-article-link