Journal of Advanced Zoology

ISSN: 0253-7214
Volume 44 Issue S-6 Year 2023 Page 943:948

Music Generation Using LSTM Neural Network

Asad Ahmad[1], Amal Mirwaiz[2], Avanish C S[3], Aravapalli Pavan Kalyan[4], Prof. Jeno Lovesum S P[5]

[1,2,3,4,5]Department of C & IT, Reva University
Email: [1]asad.12dec@gmail.com, [2]Amalmrwz@gmail.com, [3]gowda0876@gmail.com, [4]pavankalyan31708@gmail.com, [5]jenolovesum.sp@reva.edu.in

**Corresponding author's E-mail: asad.12dec@gmail.com*

Article History

Received: 06 June 2023
Revised: 05 Sept 2023
Accepted: 30 Nov 2023

Abstract

With the advancement of machine learning, Neural Networks are utilized in various fields like music, writing and pictures. Music generation is a troublesome undertaking and has been effectively investigated since decades. In this paper we are proposing a system through which we can generate music automatically using Recurrent Neural Networks. The existing system like the markov model or graph-based minimization methods lack thematic structures and they are usually repetitive sequences of the same nodes. They typically overlook the data in the negative time heading, which is non-minor in the music expectation task, so we propose a bidirectional LSTM model to create the note succession. The system which we are proposing would produce unique and coherent compositions. We used a Long Short-Term Memory (LSTM) network. They are a kind of Recurrent Neural Network that can proficiently learn through inclination plunge. Utilizing a gating instrument, LSTMs can perceive and encode long haul designs. LSTMs are very helpful to take care of issues where the network needs to recall data for an extensive stretch of time similar to the case in music and text age.

Keywords: *Neural Networks, Lstm, markov model, node.*

1. Introduction

Humans are the most intelligent species on earth. Inspired by the human brain, the scientists and engineers decided to give humans like intelligence to computers. This Artificial intelligence in Computers which allows it to take decisions and learn from previous experiences just like humans is called Artificial Intelligence. Through this artificial intelligence the computer is able to perform calculations and take decisions on its own. Nowadays, Artificial Intelligence is used in various fields. It is used in self-driving cars, in medical fields for performing various surgeries and is especially helpful in performing operations where minor incisions must be made as the chances of a computer making an error is almost zero. Not only that nowadays AI is being used in handwriting recognition, image processing, chatbots, google maps etc.

Though AI is used in various fields it does not mean it is flawless. There are certain downsides of AI number one being the high cost of implementation. Given the complications that go into building an AI, setting up an AI entails huge cost of implementation. An AI also lacks creativity, and it does not improve with experiences thus cannot replace humans.

The human brain works by sending signals through neurons. These signals are called synapses. Neuron is the basic working of the human brain. Similarly in Machine Learning we have a network of artificial neurons and these are called artificial neural networks. It works on the same principle except that the inputs and outputs in a computer are binary. The objects that do calculations are called perceptron. There are various types of Neural Networks like Feed forward Neural Network, Multilayer perceptron, Modular Neural Network, Recurrent Neural Networks example LSTM which is Long Short-Term Memory etc. These Neural Networks are used in various fields like sales forecasting, stock price prediction, data validation, music generation etc.

Just like Artificial Intelligence even the Neural Networks have certain downsides the major one being the requirement of large amounts of data and high computing power. Neural networks require much more data than traditional algorithms and as they require a lot of data, so the computers consume a lot of time to process them unless the computing power of computer is massive. The specialty of organizing sounds on schedule to deliver an arrangement is called music. Earlier people used to think that music is an analog signal and it has to be generated manually but nowadays with the advancement of technology and machine learning music can be generated automatically. Though there are techniques to generate music without using neural networks, those techniques lack thematic structures and are usually repetitive sequences of the same node.

A Recurrent Neural Network is a class of fake neural networks wherein associations between hubs structure a coordinated chart along a transient succession which is a replication of the human brain. RNN are used by Apple's Siri and Google voice search. In this paper we are proposing to use RNN for automatic music generation. The music thus generated using Recurrent Neural Networks will be unique and not just a sequence of notes repeating over a period of time.

Related Works

[1] was published in 2019 by three authors Sanidhya Mangal, Rahul Modak, Poorva Josh. Here they used one LSTM to predict both the time and node sequence. Here they have used a midi file format to where the output would be stored.[2] was published in 2019 by Tianyu Jiang; Qinyin Xiao; Xueyuan. They tried testing old style piano sets and concluded that music can be generated using Neural Networks. [3] was published by Jean Piere Brot in 2020. He designed an algorithm to generate computer based music using AI, Deep learning and Neural networks.[4] was published by J. P. Briot, G. Hadjeres, F.D. Pachet in 2019.[4] talked about the different types of Neural networks and their architectures. Then it was shown how these networks can be used to produce interesting compositions.[5] was published by Dr. P. S. Rani, S .V. Praneeth , V. R. K Reddy , M. J. Satish in 2020. Here they have used LSTM and GRU network architectures to generate music. They have taken 84 characters and passed it through the system to generate 84 random characters which can be used to generate musical nodes. [6] was published in 2019 by Majid Farzaneh, Rahil Mahdian Toroghi. Here they used an evolutionary algorithm and a BiLSTM to generate music. [7] was published by Gautam Mittal, Jesse Engel, Curtis Hawthrone, Ian Simon in 2021.Here they have used a diffusion model to generate music.[8] was published by Korneel van dan Broek in 2021. He used convolutional GAN to generate one minute long raw audio which can be converted to midi files.[9] was published by Nabil Hewahi, Salman Alsaigal, Sulaiman Aljahani in 2019. They used LSTM Neural Network to take midi files as input and convert it into songs after the augmentation process and training.[10] was published by Aishwarya Bhave, Mayank Sharma, Rekh Ram Jhengein 2019. They used Deep Learning techniques to produce interesting compositions.

Proposed Work

Objectives:

1. To have a convenient condition where musical components could be generated entirely by a neural network rather than manually by users.

2. To produce compositions that sound unique and musically coherent.

Block diagram:

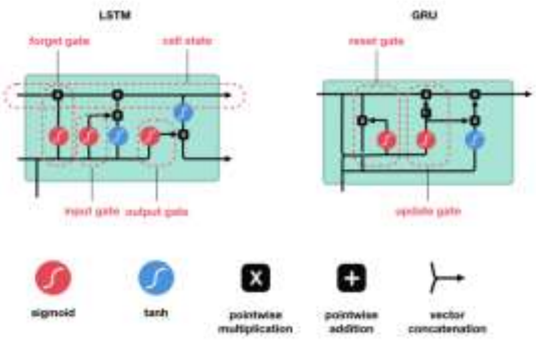

fig3.1

A neural network takes in information, goes through numerous layers of covered up neurons (little capacities with extraordinary coefficients that should be learned), and predicts an output based on the training it received. Neural networks are prepared by training them with datasets. As we can see in fig3.1 that there are three parts in an LSTM cell. These are called gates. The first one is the forget gate, the second one is called the input gate and the last one is the output gate.

Flowchart:

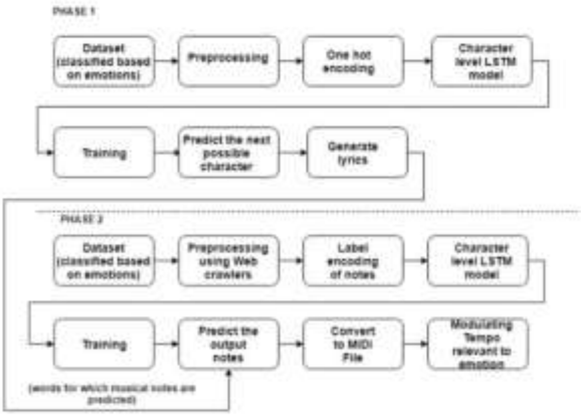

fig 3.2

The flowchart in fig 3.2 shows the working of a LSTM network. It happens in two stages. In the first phase it takes a dataset as an input which undergoes preprocessing. After the encoding is done the dataset undergoes training after which the LSTM predicts the output. In the first phase this training is done on characters and in the second phase these characters are converted into notes which are converted into midi files which can be played on any midi player.

With the expansion in computational assets and late progressions in network structures, novel music creation may now be pragmatic for enormous scope corpuses. Beginning with, we will make a mapping function which will convert the data in the form of string to integers. We are doing this because Neural Networks are efficient in numerical data when compared with strings. At that point, we need to make input arrangements for the framework and their individual yields. The yield for every data arrangement will be the chief note or concordance that comes after the progression of notes in the data gathering in our summary of notes.

The System uses LSTM and GRU network architectures to generate music entirely by a RNN. LSTM is a high level RNN, a consecutive network that permits data to continue. It is capable of dealing with the vanishing gradient issue faced by RNN. RNN is utilized for persistent memory. Likewise RNNs work, they recollect the past data and use it for preparing the current information. The shortcoming of RNN is, they can not recollect Long term conditions because of vanishing gradients. LSTMs are designed to keep away from long-term reliance issues. There are three parts in the LSTM cell known as gates. Input, output and forget gates are the three types of gates. Actually like a normal RNN, a LSTM likewise has a hidden state where H(t-1) addresses the hidden condition of the previous timestamp and Ht is the hidden condition of the current timestamp. Notwithstanding that LSTM likewise has a cell state addressed by C(t-1) and C(t) for previous and current timestamp individually.

As the application is built entirely on neural networks it takes time to learn from experiences. In powerful computers the data sets can be trained faster. To create music with RNN we have to analyse the next chord. For this reason the array we are using for analyzing should contain each and every note that is encountered while training the datasets.Next we need to stress over where we need to put the notes. As by far most that have checked out music have seen, notes generally have fluctuating stretches between them. You can have various notes with scarcely a respite in the middle and a while later followed by a rest period where no note is played for a brief timeframe. The last layer ought to consistently contain a similar measure of hubs as the number various outputs our framework has. This guarantees that the output of the network will plan straightforwardly to our classes.

To ensure that we can stop the preparation anytime without losing the entirety of our diligent effort, we will utilize model designated spots. Model designated spots furnish us with an approach to save the loads of the organization hubs to a document after each generation. This permits us to quit running the neural networks whenever we are content with the loss value without stressing over losing the weights. Else we would need to stand by until the framework has got done with going through each of the 200 epochs before we could find the opportunity to save the weights to a document.

3. Results and Discussion

It takes 200 epochs to train our datasets but we have designed it in such a way that we can stop the training after any number of epochs and the training will be saved till that point and a midi file will be generated upto the point it received the training.

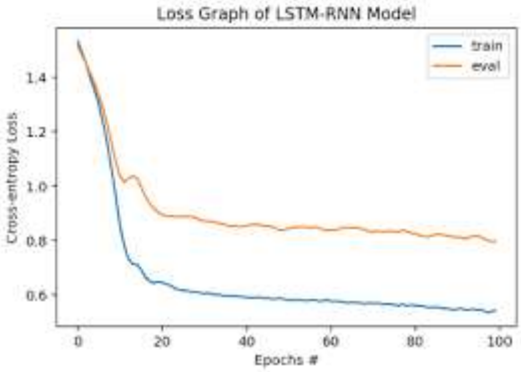

Figure 4.1

The figure 4.1 shows Loss graph of an LSTM RNN model. The blue line shows the trained model and the orange line shows evaluated model.

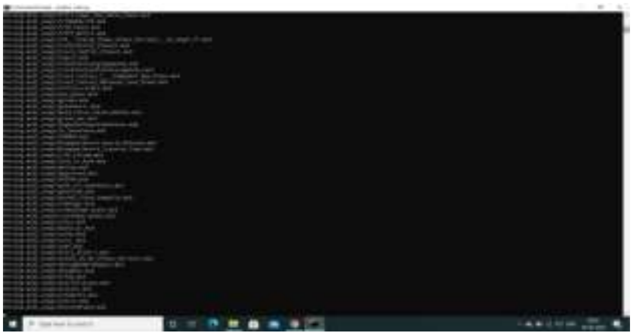

Figure 4.2

Figure 4.3

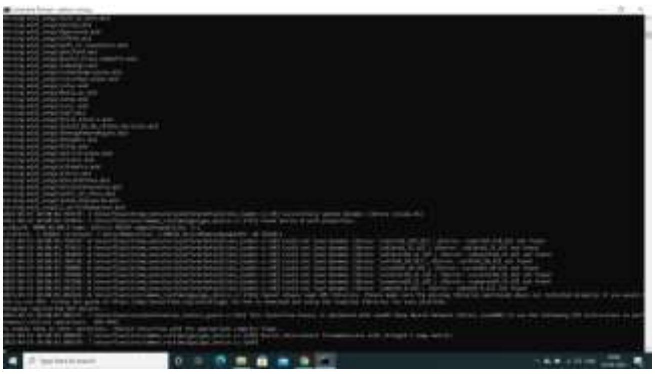

Figure 4.4

The above figures 4.2, 4.3 & 4.4 shows the training of different datasets for the generation of music which will be generated in a midi file which can then be played on any midi player.

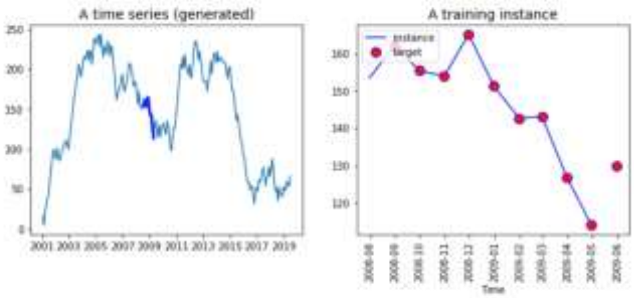

Figure 4.5

The figure 4.5 shows a time series graph of a recurrent Neural Network in the training instance.

Training the datasets: We trained our data with approximately 20 songs and the results we got were pretty impressive. It took a lot of time to generate the midi file which could then be played on any midi player or online midi player. In powerful computers it will take a lot less time. We have designed the system in such a way that after every epoch a checkpoint is made and we can stop the training after certain epochs. These checkpoints ensure that the time is not wasted while training. It saves the training till the last epoch it has received.

4. Conclusion

With the increase in computational resources and advancement of technologies music generation using neural networks has become a practical thing.The LSTM and GRU architectures can be used as they are recurring neural networks and can recall data for an extensive stretch of time. While the results may not be awesome, they are truly imperative notwithstanding and shows us that neural networks can make music and may really be used to help make seriously bewildering melodic pieces. We have achieved amazing results yet we can develop in certain areas.

First and foremost, the execution we have right currently doesn't maintain changing length of notes and different offsets notes. To achieve that we could add more classes for each remarkable length and add rest classes that address the rest time span between notes. To achieve satisfying outcomes with more classes added we would moreover have to grow the significance of the LSTM organization, which would require a basically more exceptional PC.

Second, add beginnings and endings to pieces. As the organization is as of now there is no capability between pieces, as such the organization doesn't have even the remotest clue where one piece closes and another beginning. This would allow the organization to deliver a piece start to finish rather than completing the made piece unexpectedly as it does now.

Third, add a methodology to manage dark notes. As it is by and by the organization would enter a bomb state in case it encounters a note that it doesn't have even the remotest clue. An expected technique to handle that issue is find the note or amiability that is for the most part like the dark note.

Finally, adding more instruments to the dataset. As it is right now, the organization just supports pieces that simply have a single instrument. It is intriguing to check whether it might be reached out to help a whole group.

Acknowledgment

Our appreciation goes to Dr. Jeno Lovesum S P Professor of Computer Sciences, the university of Reva for her strong ongoing support for the research, so that, amid a maze of possibilities and shortcomings, we can achieve a planned objective.

References:

1. LSTM Based Music Generation System Sanidhya Mangal, Rahul Modak, Poorva Josh.(2019). https://arxiv.org/abs/1908.01080
2. Music Generation Using Bidirectional Recurrent Network. T. Jiang, Q. Xiao, X. Yin(2019) https://ieeexplore.ieee.org/abstract/document/8839399
3. From artificial neural networks to deep learning for music generation: history, concepts and trends Jean Piere Biot(2020) https://link.springer.com/article/10.1007/s00521-020-05399-0
4. Deep Learning Techniques for Music Generation -- A Survey Jean-Pierre Briot, Gaëtan Hadjeres, François-David Pachet(2019) https://arxiv.org/abs/1709.01620
5. Music generation using deep learning by Dr.P.SHOBHA RANI, S.V.PRANEETH, VATTIGUNTA REVANTH KUMAR REDDY, M.J.SATHISH(2020)https://www.irjet.net/archives/V7/i3/IRJET-V7I3989.pdf

6. Music Generation Using an Interactive Evolutionary Algorithm Majid Farzaneh, Rahil Mahdian Toroghi(2019) https://link.springer.com/chapter/10.1007/978-3-030-37548-5_16
7. Symbolic Music Generation with Diffusion Models Gautam Mittal, Jesse Engel, Curtis Hawthrone, Ian Simon (2021) https://paperswithcode.com/paper/symbolic-music-generation-with-diffusion
8. MP3net: coherent, minute-long music generation from raw audio with a simple convolutional GAN Korneel van dan Broek(2021) https://paperswithcode.com/paper/mp3net-coherent-minute-long-music-generation
9. Generation of music pieces using machine learning: long short-term memory neural networks approach by Nabil Hewahi, Salman Alsaigal, Sulaiman Aljahani(2019) https://www.tandfonline.com/doi/full/10.1080/25765299.2019.1649972
10. Music Generation Using Deep Learning Aishwarya Bhave, Mayank Sharma, Rekh Ram Jhengel (2019) https://link.springer.com/chapter/10.1007/978-981-13-3393-4_21
11. Midi Net: A convolutional Generative Adversarial Network for Symbolic Domain Music Generation. Li Chia Lang, Su Zu Chau, Yi Husang Yang (2019)
12. Wu Y, Li W (2018) Automatic audio chord recognition with MIDI-trained deep feature and BLSTM-CRF sequence decoding model. IEEE TASLP 27(2):355–366. https://doi.org/10.1109/taslp.2018.2879399
13. Arumugam M, Kaliappan M (2018) Feature selection based on MBFOA for audio signal classification under consideration of Gaussian white noise. IET Signal Proc 12(6):777–785.
14. Wang D, Deng S, Xu G (2018) Sequence-based context-aware music recommendation. Inf Retr J 21(2–3):230–252.
15. GRUV: Algorithmic Music Generation using Recurrent Neural Networks https://cs224d.stanford.edu/reports/NayebiAran.pdf
16. An Improved RNN-LSTM based Novel Approach for Sheet Music Generation. https://www.sciencedirect.com/science/article/pii/S1877050920310152
17. Music Generation by Deep Learning - Challenges and Directions (2020) Jean Piere Briot and Francois Pachet.
18. Continual lifelong learning with neural networks(2019) Parisi GI, Kemker R, Kanan C
19. A style specific Music Composition Neural Network.(2020) Cong Jin, Yun tie, Yong Bai.
20. Understanding a convolutional Neural Network(2018) Saad Albawi, Tareq Abed Mohammed, Saad Al Zawi.