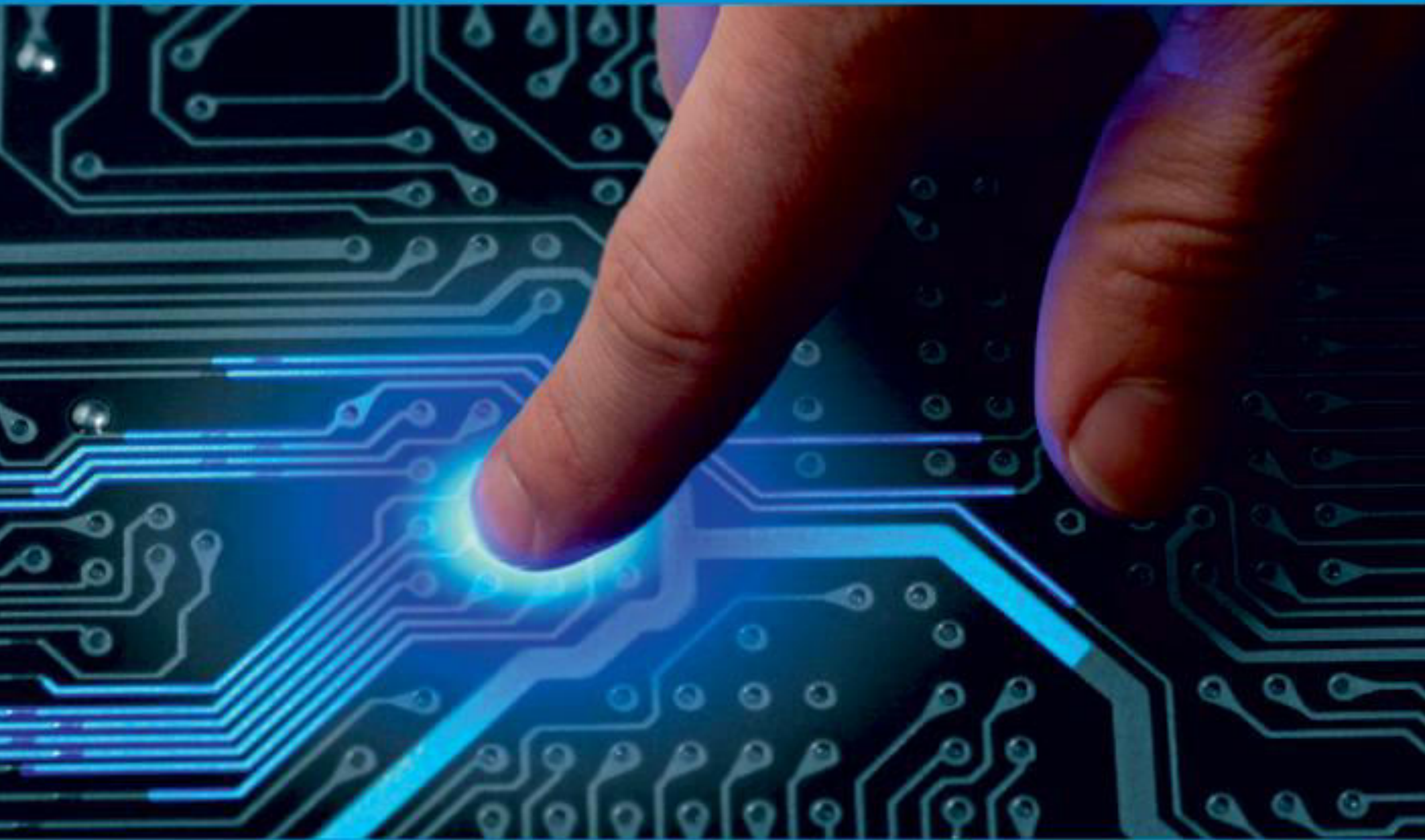
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

## IN COMPUTER & COMMUNICATION ENGINEERING

Volume 10, Issue 7, July 2022

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.165

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165

|| Volume 10, Issue 7, July 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1007080 |

# A Musical Composition Assistant System using LSTM

**Poonam G. Fegade[1], Trupti Jagtap[2]**

Assistant Professor, Department of MCA, K. K. Wagh Institute of Engineering Education & Research, Nashik, India[1]

Student, Department of MCA, K. K. Wagh Institute of Engineering Education & Research, Nashik, India[2]

**ABSTRACT**: Universally, music is one of the elements that create harmony in this world. Traditionally playing music or creating it has always been seen as a manual task. Music can be created with the help of instruments, voices and sounds. In the Modern Era of technology and AI, we can automate this process. For this, we must understand the basic structure of how music forms and view it scientifically. For the proposed system, we are using RNN-LSTM algorithm to generate music from given sample inputs. A Musical Composition Assistant System (MCAS) involves transformation of music scores into time series representation, encoding the music. A model is designed to execute this algorithm where data is represented with the help of musical instrument digital interface (MIDI) file format for easier access and better understanding. Pre-processing of data before feeding it into the model, revealing methods to read, process and prepare MIDI files for input are used. This system creates good music pieces in MIDI format with given input. The proposed system uses Flask API to interact with the frontend

**KEYWORDS**: AI, RNN-LSTM, MIDI.

## I. INTRODUCTION

Music is an art which is thoroughly spread globally. Music Industry has many creative artists. Apart from this there are professionals, learners, students who are also part of this music community. Sheet music is a handwritten or printed form of musical notation that uses musical symbols to indicate the pitches, rhythms, or chords of a song or instrumental musical piece. Sheet music can be used as a record of, a guide to, or a means to perform, a song or piece of music. Sheet music enables instrumental performers who are able to read music notation (a pianist, orchestral instrument players, a jazz band, etc.) or singers to perform a song or piece. Our proposed system is able to creative new and innovative music ideas and presents them in the form of Sheet music [1]. This is very beneficial for the music community. The main objective of the system is to generate music. Input music file is provided to the system in the MIDI format. MIDI (Musical Instrument Digital Interface) is a protocol designed for recording and playing back music on digital synthesizers that is supported by many makes of personal computer sound cards. Three important aspects of the system are: Input music sequence file, RNN LSTM model and generated output. Following are the basic tasks of the system.

1. Accept an Input from the user in MIDI format.
2. If the file has any error or is corrupt then warn the user about this.
3. If format of file is correct then input it to the model for generating a new sequence.
4. Create the output file.
5. User should be able to download the output file

**Basic Terminologies related to Music:**
**Melody**: A melody is a collection of musical tones that are grouped together as a single entity

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165

|| Volume 10, Issue 7, July 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1007080 |

**MIDI Notation:** MIDI (Musical Instrument Digital Interface) is a technical standard that de- scribes a communications protocol, digital interface, and electrical connectors that connect a wide variety of electronic musical instruments, computers, and related audio devices for playing, editing, and recording music. It maps note names to numbers.

**Music Representation used for our model**

- We are representing music in the time series format.
- We are using MIDI note numbers whenever any key occurs.
- For held notes we use ” ” symbol.
- We use ”r” for the rest.

[ “60”, “_”, “_”, “_”,
“62”, “_”, “_”, “_”,
“64”, “_”, “64”, “_”,
“65”, “_”, “62”, “_”,
…]

Figure 1: Time Series Representation

## II. RELATED WORKS

Multiple works has been done on the generation of music using neural networks and other statistical models. Some of the research shows better output accuracy when evaluated. To conduct and understand this research in detail, previous work is critically reviewed in this section. In the past few years, there has been an increased interest in machine-learning-created music. Ranging from Open Jukebox which focuses on making complete soundtracks using AI[2], to the Magenta research project which touches of many different aspects of music production. In addition to these rather large scale projects on music, the more successful approaches in music creation have been results of neural networks. There have been smaller projects looking particularly at the MIDI-build-up of music. Daniel Johnson managed to make a simple piece of music using a specially designed recurrent neural network (RNN) [3], and RNNs have repeatedly been shown to give good and useful results in music generation. Drawing inspiration from these works, we will also apply RNN algorithms to model sequences of music but in the form of RNN LSTM. Because RNN has its own limitations which can be overcome by RNN-LSTM [4]

As the revolution in the field of Artificial Intelligence is at a faster pace, there are various ways to work with these technologies. In this modern age every aspect of life is getting blended with AI using machine learning [5]. So does the music. Music is an important part of many people’s lives. This system is an experiment in which we have tried to combine Music with AI. Following are the needs for the system:
1. This system is targeted toward Music Enthusiast who is willing to learn music in any form, who are musicians and people who want some creative ideas related to melodies.
2. There are not many existing systems which provide automated melody creation. This system provides a simple UI which can be understood by any person.
3. It is important to experiment with different algorithms in order to find new techniques in the creative music field.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165

|| Volume 10, Issue 7, July 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1007080 |

## III. METHODS

Melody Palette is AI system which creates music from given input by the user. We are using RNN-LSTM algorithm for generating the model. This system is a web platform and built using Python, Flask API, HTML, CSS, and Bootstrap. Basic objective is to generate a melody from given 4-5 notes. It takes and gives output in the form of MIDI. This system is useful for music community especially the music learners, professionals and music lovers. System Receive a MIDI music file from user as a sample input. Neural Network algorithm is applied to create new melody. System shall generate a MIDI file which user can download.

### A. MACHINE LEARNING AND DEEP LEARNING ALGORITHMS

#### 1) Neural Network

Neural Networks are set of algorithms which closely resemble the human brain and are designed to recognize patterns. They interpret sensory data through a machine perception, labeling or clustering raw input. They can recognize numerical patterns, contained in vectors, into which all real-world data (images, sound, text or time series), must be translated.

Artificial neural networks are composed of a large number of highly interconnected processing elements (neuron) working together to solve a problem.

An ANN usually involves a large number of processors operating in parallel and arranged in tiers. The first tier receives the raw input information analogous to optic nerves in human visual processing. Each successive tier receives the output from the tier preceding it, rather than from the raw input in the same way neurons further from the optic nerve receive signals from those closer to it. The last tier produces the output of the system

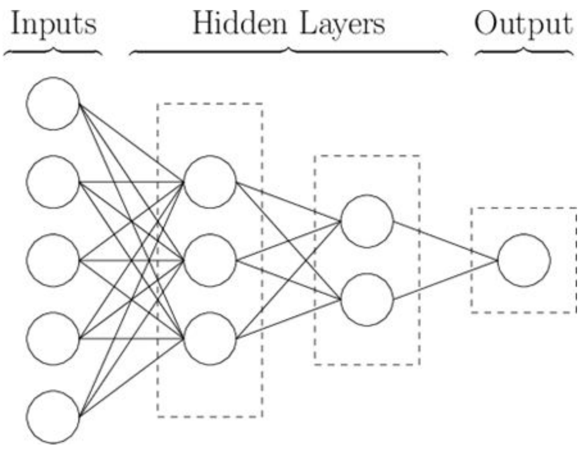



Figure 2 :SimpleNeuralNetwork

#### 2) RNN

Arecurrentneuralnetwork(RNN)isaclassofartificialneuralnetworkswhereconnectionsbetweennodesformadirectedorundirectedgraphalongatemporalsequence.Thisallowsittoexhibittemporaldynamicbehavior.Derived from feed forward neural networks, RNNs can use theirinternalstate(memory)toprocessvariablelengthsequencesofinputs.In RNN, all the inputs are related to each other. First, it takes the X(0) from the sequence of input and then it outputs h(0) which together with X(1) is the input for the next step. So, the h(0) and X(1) is the input for the next step. Similarly, h(1) from the next is the input with X(2) for the next step and so on. This way, it keeps remembering the context while training.

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165

|| Volume 10, Issue 7, July 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1007080 |

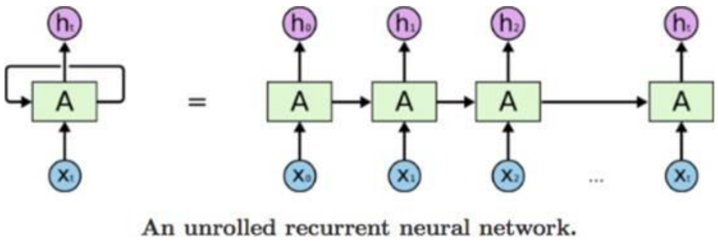



Figure 3:Recurrent Neural Network

The formula for the current state is

$$h_t=f(h_{t-1},x_t)$$

Applying Activation Function:

$$h_t=tanh(W_{hh}h_{t-1}+W_{xh}X_t)$$

W is weight, h is the single hidden vector, Whh is the weight at previous hidden state, Whx is the weight at current input state, tanh is the Activation function, that implements a Non-linearity that squashes the activations to the range[-1.1]

Output :

$$yt = Whyht$$

**3) LSTM**

Long Short-Term Memory (LSTM) networks are a type of recurrent neural network capable of learning order dependence in sequence prediction problems. This is a behavior required in complex problem domains like machine translation, speech recognition, and more. LSTMs are a complex area of deep learning. The main difference between RNN and LSTM is in terms of which one maintain information in the memory for the long period of time. Here LSTM has advantage over RNN as LSTM can handle the information in memory for the long period of time as compare to RNN.

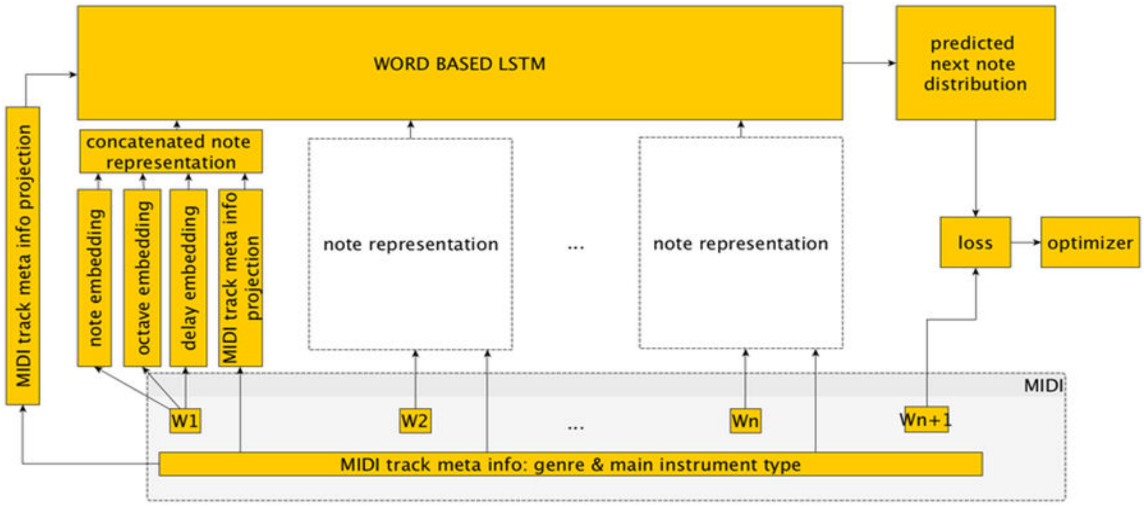



Figure 4:Working of RNN LSTM

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165

|| Volume 10, Issue 7, July 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1007080 |

**B. Data**

Dataset we are using is called deutschl from KernScores. This dataset is the collection of folksongs from Germany. The kern representation can be used to represent basic or core information for common Western music. The kern scheme can be used to encode pitch and duration, also other common score related information.


Figure 5:KernRepresentationvsMusicscore

**C.Tools**

1)**MuseScore**

MuseScore is a score writer for Windows, macOS, and Linux supporting a wide variety of file formats and input methods. We can create, play back and print beautiful sheet music with free and easy to use music notation software MuseScore. We have used it to read and play the output audio sequence.

2) **Music21**

Music21 is a Python-based toolkit for computer-aided musicology. People use music21 to answer questions from musicology using computers, to study large datasets of music, to generate musical examples, to teach fundamentals of music theory, to edit musical notation, study music and to compose music(both algorithmically and directly)

## IV. Implementation

Experiment and model training were conducted on Google Colab, running over Google Cloud Platform with deep learning .Code is implemented in Keras (using Tensorflow as backend). A fully trained model was used to generate a suite of music.

1) Preprocessing

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165

|| Volume 10, Issue 7, July 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1007080 |

Data preprocessing is a process of preparing the raw data and making it suitable for a machine learning model. It is the first and crucial step while creating a machine learning model. We have used music21 library to transform the .krn files into numerical representation encoded format. Then this transformed dataset is used to train the model[6].

2) Training Model

i) The LSTM network is trained on polyphonic musical notes to acquire the knowledge of probability of occurrence of a musical note at current time. The output of the network at a time step t, conditioned on the previous notes' state till time step t-50 are fed into the input unit to recall past details and structure of notes. The LSTM layer depends on a selected input. Only selected notes are used to train LSTM model, which are useful for effectively tuning the model that results in efficient information gain.
ii) With these inputs, LSTM layer learns the mapping and correlation between notes and their projection. Next to the LSTM layer, Dropout layer is used to create generalizations in the model. Once the model has learned the probability distributions of notes and sequences, all the LSTM cells were combined with each other. This gap is subordinated with the help of dense layer.
iii) Dense layer ensures that the model is fully connected. At the endpoint, the Activation layer is added to the model, which helps in deciding, which neurons (LSTM cells) should be activated and whether the information gained by the neuron is relevant, making activation function highly important in a deep neural network
iv) After training the LSTM network, the model is ready to generate a new sequence of musical notes. To ensure better prediction and diverse output of sequences, a large and varied dataset was elicited with different variations in the structural composition of musical notes. The aim was to expose the model with diverse dataset which would lead to a better tuning of the model. The MIDI file format was used to extract dataset. MIDI files played an important role in extracting information about note sequence, note velocity and the time component.


Figure 6: Melody Generator Training

3) Procedure

Two python scripts were created as Input conversion and Melody Generator. Input conversion script encodes input melody. The given input is first received by flask app. Flask app calls the initialize function from Melody Generator class. This class first checks if the values of input file are within the range of values used in model. Then the given melody is transposed to C Major Scale if it is major else A minor mode. Transposed melody is passed to generate melody function which actually generates melody using trained neural network. The output given by neural network is again parsed back to MIDI format and sent back to the user using flask app.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165

|| Volume 10, Issue 7, July 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1007080 |

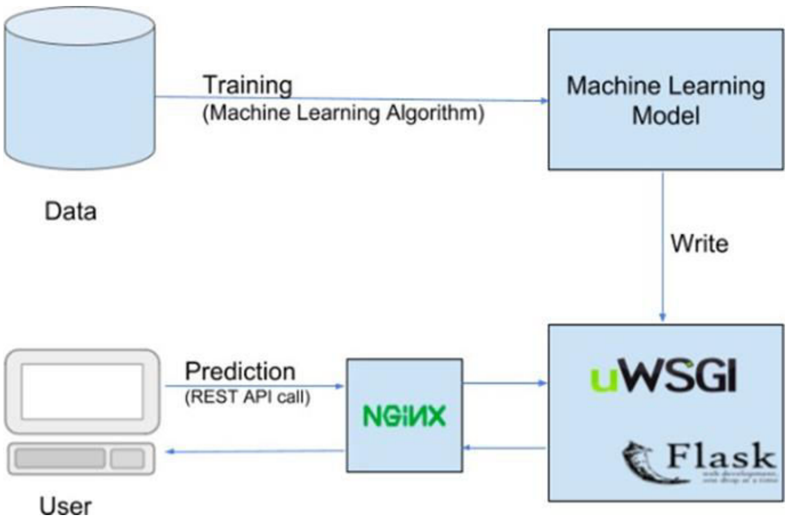



Figure 7:InteractionwithFlask

## IV. RESULT

MCAS achieves the goal of designing a model which can be used to generate music and melodies automatically without any human intervention. We trained the model using 1700 songs. The model is capable to recall the previous details of the dataset and generate a polyphonic music using a single layered LSTM model, proficient enough to learn melodic note sequence from MIDI files of German Folk music. The model design is described with a perception of functionality and adaptability. Induction and method of training dataset for music generation is achieved through this work. Moreover, analysis of the model is also impersonated for better insights and understanding. Current system is generating a simple melody from given input. Users are able to upload input and download output. The accuracy of the output generated is mostly dependent on humans as humans are expert at music than a machine [7].

This system works on all kind of screen resolutions including desktop as well as phone screen. The generated output in the MIDI format can be viewed in MuseScore Software. Screenshots of actual result is given here.

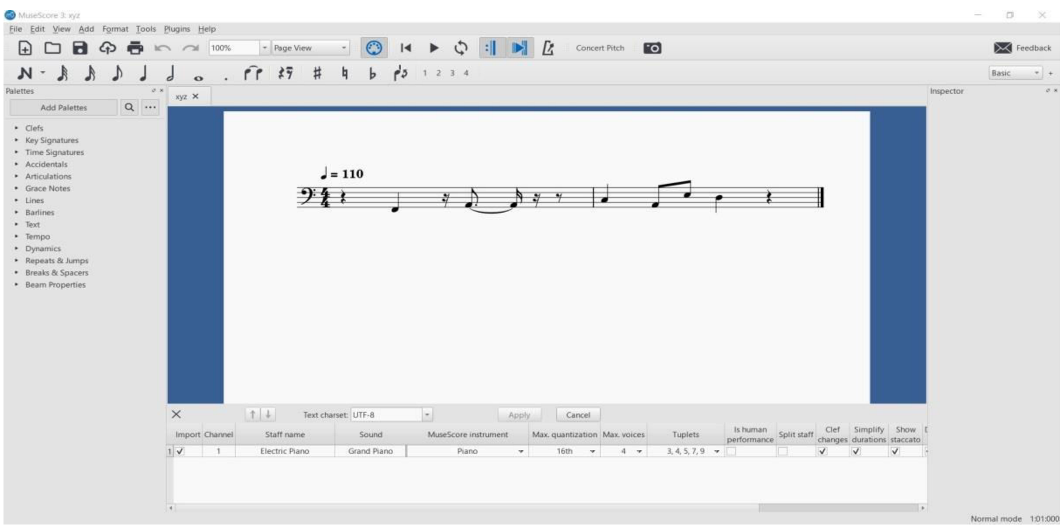


Figure 8: Input MIDI File

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165

|| Volume 10, Issue 7, July 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1007080 |

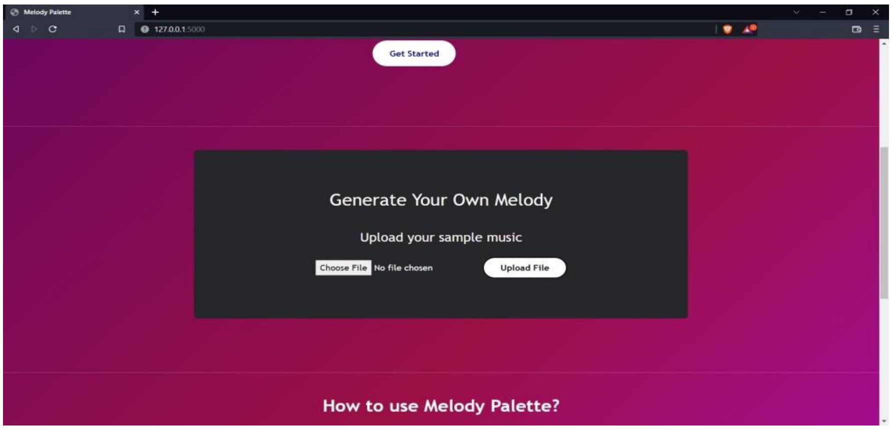



Figure 9: Melody Palette Home

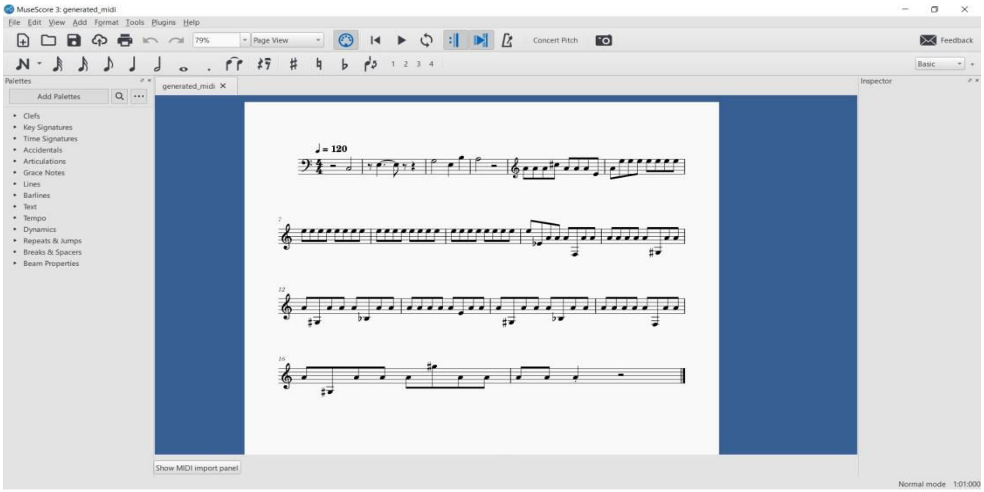



Figure 10: Generated Melody in MuseScore

## V. CONCLUSION

Proposed system for generating music with the help of AI is a valuable addition for the field of Music plus modern technology. It proves that there are no limits to music and musical research. The proposed system used RNN-LSTM model to create melodies. As the music is all about taste of the person, accuracy of the model should rather be tested and acknowledged by humans. Basic functionality of the system includes uploading and downloading music files, processing. This system will be an advantage for the people who play or learn any musical instrument and it is also beneficial for professionals to get some of the inspiring musical ideas.

## Future Scope

We have generated a fine quality music but there is huge room to improve this. Current system is working with 4/4 bar pattern music. A bar (or measure) is a single unit of time containing a specific number of beats played at a particular tempo. In the future we can generate 3/4 and other patterns too.

We have generated model using 1700 musical tunes. If we increase the number of tunes then accuracy and diversity will increase. Also Melody Palette is base on Melody. We need harmoney which can be achieved by using chords and multiple instrument. We can add these features in the future.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165

|| Volume 10, Issue 7, July 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1007080 |

Also current system only takes one file as an input. In future we can increase the number of Inputs so that output will be versatile.

## REFERENCES

[1] Lin, Chern-Sheng & Lin, Chi & Yang, Shih-Wei & Liou, Shuo-Rong & Lay, Yun-Long. (2012). A Musical Composition Assistant for the People with Severe Physical and Visual Disabilities. Assistive technology : the official journal of RESNA. 24. 123-31. 10.1080/10400435.2012.659327

[2] Prafulla Dhariwal, Heewoo Jun, Christine Payne, Jong Wook Kim, Alec Radford, Ilya Sutskever,, Jukebox: A Generative Model for Music, OpenAI, San Francisco

[3] Johnston, D. (2015, Aug). Composing music with recurrent neural networks. Daniel Johnson. Retrieved from http://www.hexahedria.com/2015/08/03/composing -music-with-recurrent-neural-networks

[4] Maria Klara Je¸drzejewska, Adrian Zjawinski, Bartłomiej Stasiak, Generating musical expression of MIDI music with LSTM neural network, 978-1-5386-5024-0/18/$31.00 ©2018 IEEE

[5] David Kang Stanford dwkang Jung Youn Kim Stanford jyk423 Simen Ringdahl Stanford ringdahl, Project milestone: Generating music with Machine Learning, stanford.edu

[6] Qibin Lou, Music Generation Using Neural Networks, qibinlou@stanford.edu

[7] Yumi Isowa1, Nami Iino2,3, Shigeki Okino2, Yasuki Iizuka1, A Musical Composition Assistant System using a Stochastic Musical Model, 2017 6th IIAI International Congress on Advanced Applied Informatics