Music Generation using LSTM

By

Group No: 04

Roll No.	Name
1604033	Gogoi Haradhruchitrajit Dhrubajit
1604113	Tawde Rahul Rajendra
1604129	Rayudu Novina Balaramakrishna

Problem Statement:

Need to generate melodious music with the techniques of deep learning using LSTM, which could be similar to the music created by music composer.

Challenges:

- Understanding the specification for MIDI format.
- Differentiate between notes and chords.
- Processing MIDI files to get feature network for the vector.
- Re-writing "predicted" notes and chords objects back to MIDI file.

Dataset:

0	0	Header	1	8	480		
1	0	Start_track					
1	0	Title_t	"Espana Op. 165"				
1	0	Title_t	"Tango"				
1	0	Copyright	"Copyright © 2001 by Bernd Krueger"				
1	0	Text_t	"Isaac Albeniz"				
1	0	Text_t	"Andantino"				
1	0	Text_t	"Fertiggestellt 28.01.2001\012"				
1	0	Text_t	"Normierung: 23.12.2002\012"				
1	0	Text_t	"Update am 29.8.2010\012"				
1	0	Text_t	"Dauer: 2:08 Minuten\012"				
1	0	SMPTE_of	96	0	3	0	0
1	0	Time_sigr	2	2	24	8	
1	0	Key_signa	2	"major"			
1	0	Tempo	958466				
1	360	Tempo	932836				

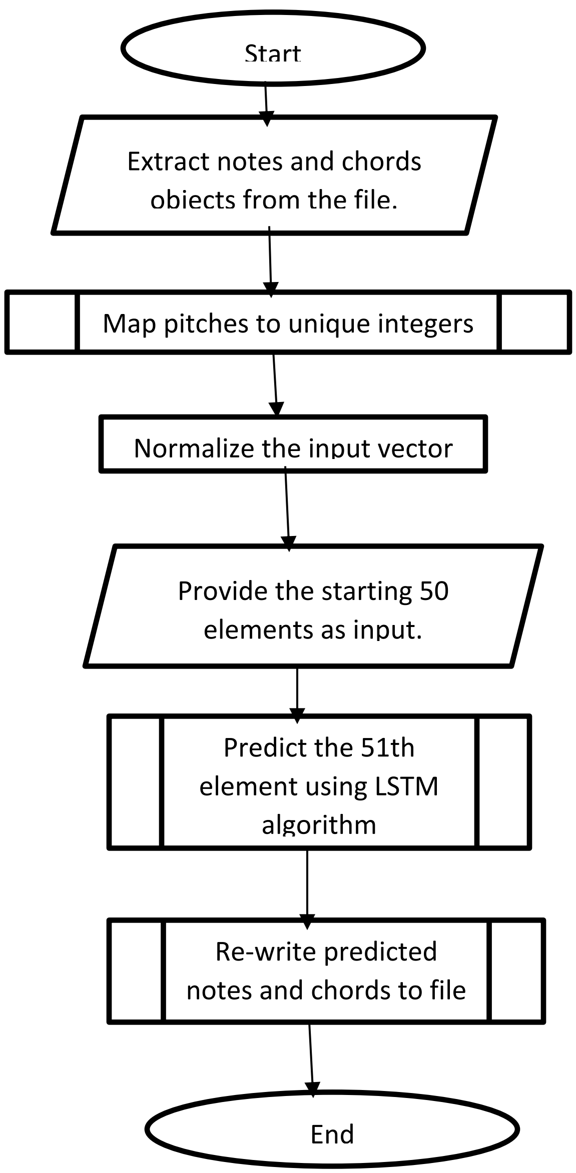
Tools:

- Keras
- Sequential model
- Music21 toolkit: for processing MIDI file.

Work Done:

- Created right data representation for the network.
- Processed MIDI file and extracted notes and chords objects.
- Pitches are in string form, so mapped them to unique integers.
- Reshaped input vector in batch size, normalized it.
- Provided starting 50 elements of MIDI file as input and predicted the 51th element.
- Re-write the predicted notes and chords to MIDI file.

Flow Chart:



Results:

- We are able to show that multi-layered LSTM, character level language model applied to dad representation is capable of generating music that is atleast comparable to sophisticated time series probability density techniques prevalent in literature.
- We showed this model is able to learn meaningful musical structure.
- After 200 epochs we got an accuracy of 54.26%.
- Epochs list of start and end are as shown below:

  Epoch 1/200

  505/505 [==============================] - 5s 10ms/step - loss: 4.2653 - accuracy: 0.0297

  Epoch 2/200

  505/505 [==============================] - 4s 8ms/step - loss: 4.0914 - accuracy: 0.0614

  Epoch 3/200

  505/505 [==============================] - 4s 8ms/step - loss: 3.9306 - accuracy: 0.0634

  Epoch 4/200

  505/505 [==============================] - 4s 8ms/step - loss: 3.8949 - accuracy: 0.0574

  Epoch 5/200

  505/505 [==============================] - 4s 8ms/step - loss: 3.8952 - accuracy: 0.0574

  Epoch 196/200

  505/505 [==============================] - 4s 8ms/step - loss: 1.4468 - accuracy: 0.4475

  Epoch 197/200

  505/505 [==============================] - 4s 8ms/step - loss: 1.4483 - accuracy: 0.4931

  Epoch 198/200

  505/505 [==============================] - 4s 8ms/step - loss: 1.3782 - accuracy: 0.4673

Epoch 199/200

505/505 [==============================] - 4s 8ms/step - loss: 1.4005 - accuracy: 0.4911

Epoch 200/200

505/505 [==============================] - 4s 8ms/step - loss: 1.3188 - accuracy: 0.5426

References:

[1] Nicolas Boulanger-Lewandowski, Yoshua Bengio, and Pascal Vincent. Modeling temporal dependencies in high-dimensional sequences: Application to polyphonic music generation and transcription. Proceedings of the 29th International Conference on Machine Learning, (29), 2012.

[2] Chun-Chi J. Chen and Risto Miikkulainen. Creating melodies with evolving recurrent neural networks. Proceedings of the 2001 International Joint Conference on Neural Networks, 2001.

[3] Douglas Eck and Jurgen Schmidhuber. A first look at music composition using lstm recurrent neural networks. Technical Report No. IDSIA-07-02, 2002.

[4] Daniel Johnson. Composing music with recurrent neural networks.