Creating Jazz with Deep Learning

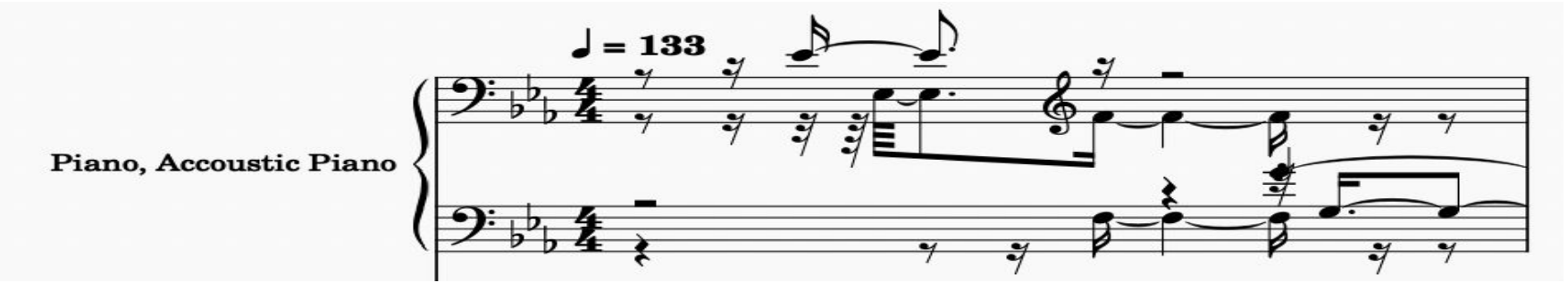
18 Dec 2021

Martin Yong

Let’s start with jazz!

Piano cover of “I’ve grown accustomed to her face” by Doug Mckenzie

But what is music, really? - A series of notes (pitch, octave, offset), chords, pauses



Approach

Dataset: 5 jazz piano covers in midi (.mid) format

Library: music21 from MIT

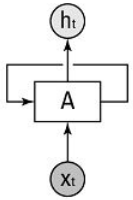
If NLP uses words, music uses notes

1 

Parse notes and chords from midi

2 X Y

Transform into:
- input (100 notes)
- target variables (the 101th note)

3 

Create neural network model and train

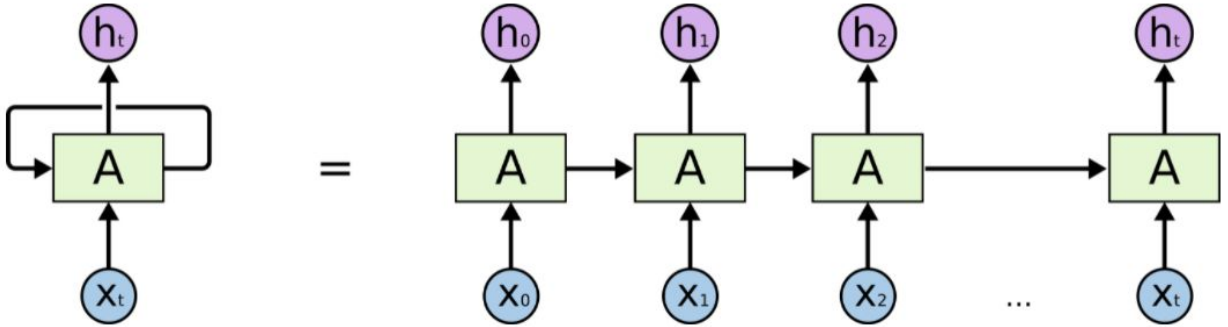
4 Fm

Generate 500 notes using the model

5 

Transform and store them into midi files

Modeling



Unrolled RNN model, credit: Christopher Olah's blog (see Appendix)

- Use RNN (Recurrent Neural Network) model, in particular Stacked LSTM (Long-Short Term Memory model)
- Use of batch normalization between layers to speed up learning
- Minibatch size of 32, running 400 epochs

Result

Epoch 1

Epoch 100

Epoch 382

Lessons Learned

- It takes a lottttt of epochs, a lot, before the result sounds… ok
- # of epochs are more important than the # of songs

Next Steps

- Ingest rest notes as inputs in order for the model to learn pauses in songs
- Create structure of actual songs, i.e. verse-chorus-verse-chorus-bridge
- Add an additional instrument, e.g. bass

Appendix

Appendix A - References

Shubham Gupta - https://www.hackerearth.com/blog/developers/jazz-music-using-deep-learning/

Sigurður Skúli - https://towardsdatascience.com/how-to-generate-music-using-a-lstm-neural-network-in-keras-68786834d4c5

Colah's blog - Understanding LSTM Networks: https://colah.github.io/posts/2015-08-Understanding-LSTMs/