October 28th, 2020

**Project On Bollywood Music Generation Using Deep Learning**

As an independent study under prof. Lim, I undertook a project on an application of Deep Learning that’s always keenly interested me ever since I dived into the rabbit-hole of [OpenAI Jukebox’s AI-produced music](https://youtu.be/uSWTSML9hwY) — music generation.

Having already done an 11-week Introduction to Machine Learning course before, I decided to complete the 16-week deeplearning.ai specialization in order to gain all the required skills to undertake this project.

After completing the specialization, I dived into the literature of music generation to understand the specific Deep Learning methods and optimizations applied to make an AI produce music. Quite simply, music generation models first train on existing music to understand the features of music as data sequences. Recurrent Neural Networks (RNNs) are the common standard for sequence data models as they are specifically made to fit sequence data.

For music, this sequence is simply the ‘sheet music’ format of a given music piece/song, with the information of the notes encoded as keys on a time axis, with additional information like durations and offsets (how long a key is played for) as well encoded into the sequence.



Figure 1: Simplified Example Of How Music Can Be Encoded As Sequence Data. Source: <https://towardsdatascience.com/practical-tips-for-training-a-music-model-755c62560ec2>

Specifically, the RNN model learns to predict the next note given the exact last ‘n’ notes and information about all the notes before it, and in doing so, it effectively learns the patterns and sequences that occur across a given music piece. As it learns multiple songs, it sees and learns commonalities and understands the patterns of music in general.

To actually produce music, after training the model with existing songs/music pieces, we then take any random sample of ‘n’ notes from a given song and use them as the first input for the model. The model then does the same thing as before: given the previous ‘n’ notes, it predicts the next note, except this time, this produced note is then again re-fed into the model to predict the note after that, and on and on to make it generate music.

Many of the methods of manipulating audio data and the initial version of the model are based on [this](https://hedonistrh.github.io/2018-04-27-Music-Generation-with-LSTM/) project, which is a music generator for single-instrument classical music. Much of the project involved fine-tuning and readjusting many of the model’s parameters and variables to fit my data of multi-instrument Bollywood music and modifications like considering each instrument in a song as its own unique music sequence.

After training the final model using 71 Bollywood songs, here is a sample song it produced:

<embedded: <https://youtu.be/ckcQ1K9efvk>>

It should be noted that the output of the model is also in the ‘sheet music’ format, which can then be imported into a Digital Audio Production software to add instruments in order to produce the final samples. But this also means that we can make it sound like a completely different genre (like Synthwave) simply by using different instruments!

<embedded: <https://youtu.be/pqya_VlJ7Vs> >

<embedded: <https://youtu.be/95FGbHAVr8E> >