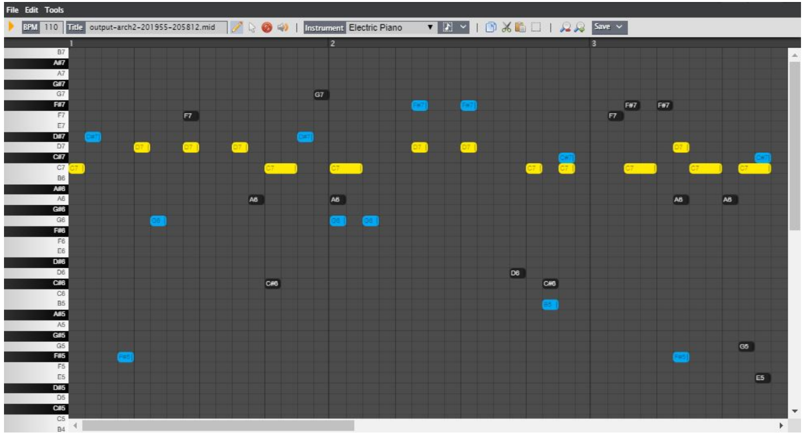
KRAMZ MUSIC GENERATOR

Kevin Ray, Ronald Moore, Austin Irvine, Matthew Taylor, Zach McGrath

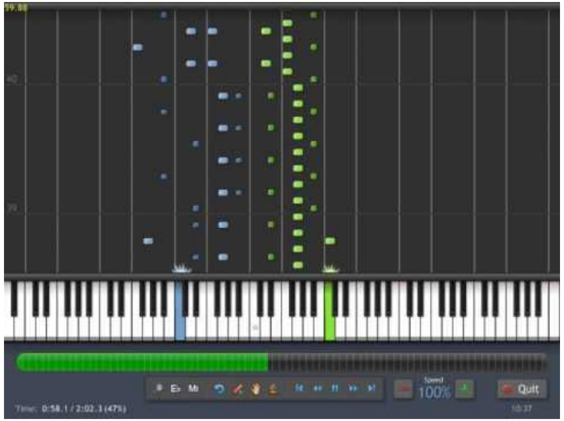
Objective

Use machine learning to generate music with multiple instruments

Other implementations only use one instrument/ignore rests



BPM 110 Title output-arch2-201955-205812.mid Instrument Electric Piano



Approach

Use Jazz MIDI (.mid) files to load music

Separate file into the three common jazz instruments: Piano, Bass, Saxophone

Use Tensorflow and Keras to implement LSTM Networks in three different architectures

Produce MIDI files as output

Data

A Jazz MIDI dataset with over 800 files was found on Kaggle

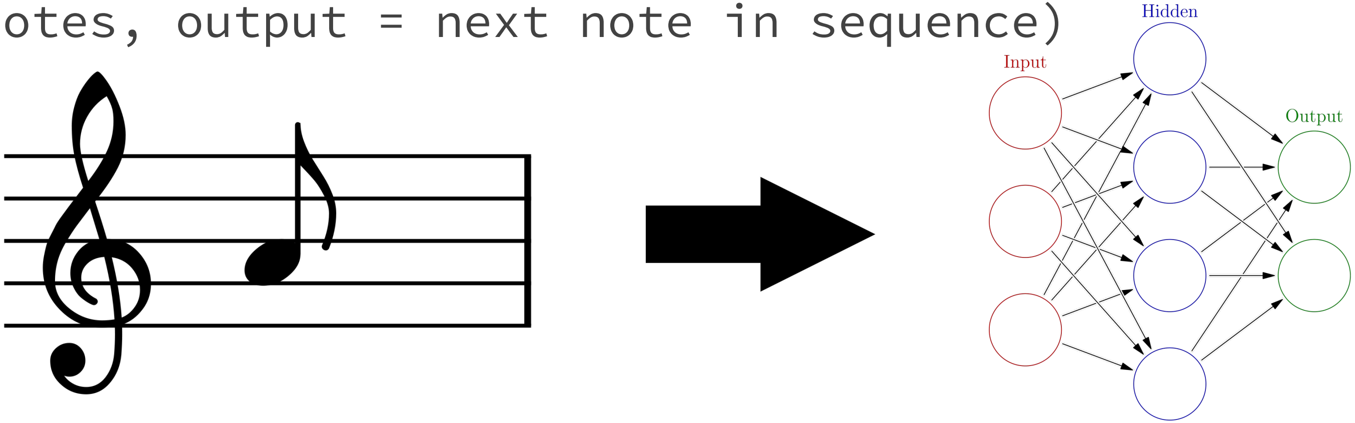
Most songs had only one instrument, only about 30 had our three desired instruments

The dataset featured a wide variety of songs, not exclusively jazz

MIDI and music21

music21 is a music theory library created for Python, it provided an easy way to convert between MIDI files and neural network inputs and outputs

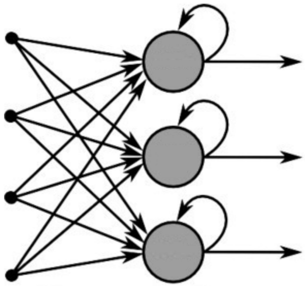
Able to extract 46,000 training samples (input = a sequence of 100 notes, output = next note in sequence)



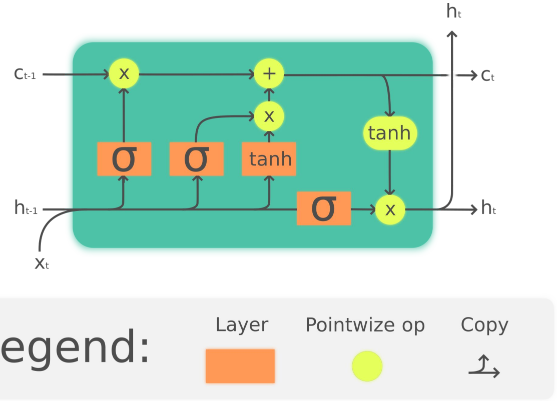
RNN and LSTM

Recurrent Neural Networks are good for sequential data

Long Short-Term Memory Networks improve on this by being able to keep context of events that happened long ago while also being able to ‘forget’ events that should not impact output



Recurrent Neural Network

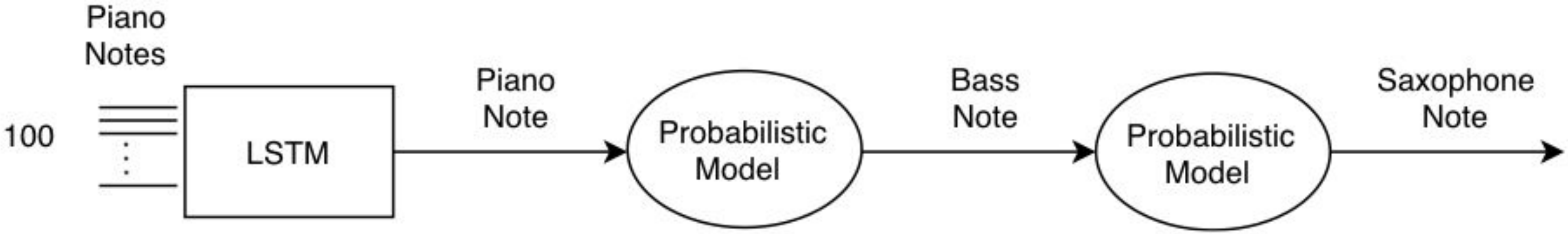


Architectures

1. One LSTM with a base instrument and probabilistic models for the other two
2. Three LSTMs with input of the next dependent on the output of the previous
3. Three LSTMs with the input of one being the output of the previous

All of the architectures used the outputs as probability distributions to sample from.

Architecture 1

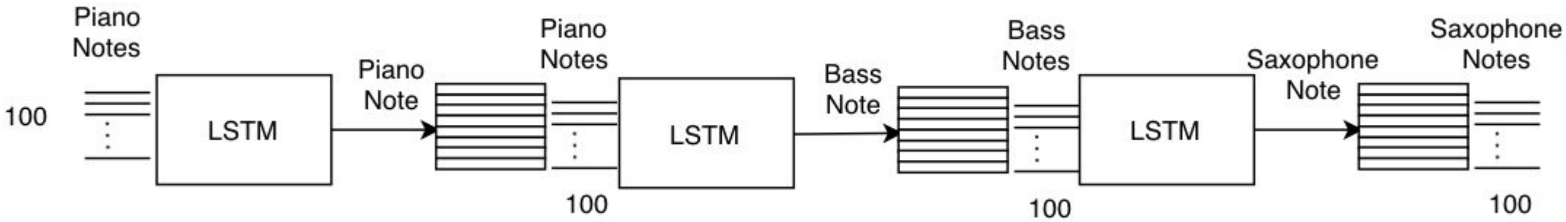


LSTM takes in 100 sequence of piano notes and produces next note

A probabilistic relationship between piano notes and bass notes generates a bass note

A probabilistic relationship between bass notes and saxophone notes generates a saxophone note

Architecture 2

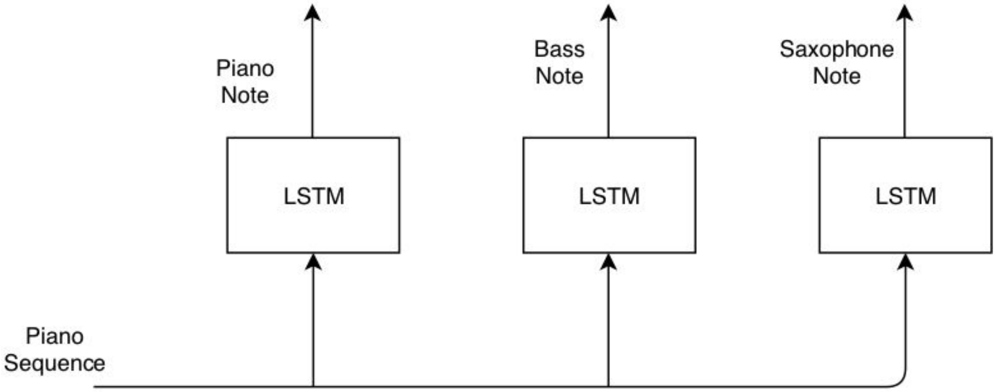


1st LSTM takes in sequence of 100 piano notes and produces a piano note

2nd LSTM takes in sequence of 100 piano notes from 1st LSTM and produces a bass note

3rd LSTM takes in sequence of 100 bass notes from 2nd LSTM and produces a saxophone note

Architecture 3



All LSTMs take in sequence of piano notes to produce the next note of their respective instrument.

Resulting Songs

Architecture 1:

Architecture 2:

Architecture 3:

Analysis of Results

The song that sounds the best is up to the listener.

- Architecture 1 (Not Good 💩)
- Architecture 2 (Cinematic ⭐)
- Architecture 3 (Not Bad 🆗)

Design Issues

Lots of them

Data cleaning/feature engineering

Rest to note ratio is very high

Understanding Music Theory

Design Issues

First attempt

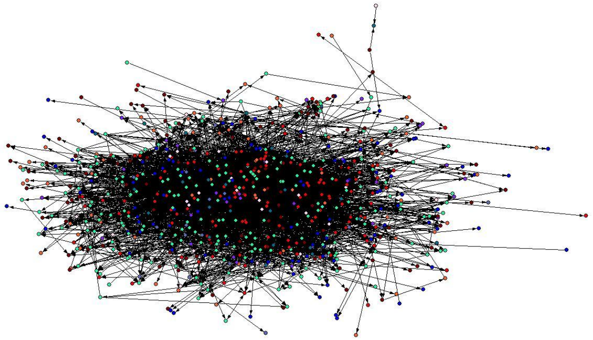
- Data cleaning with durations

Second attempt

- Number of classes

Third attempt

- Chords produced non-harmonious rhythms



Alternative Designs/Future Work

Harmony

Note volume

Chords

Combination of notes

Tempo