Music Synthesis via Chord Progression Generation

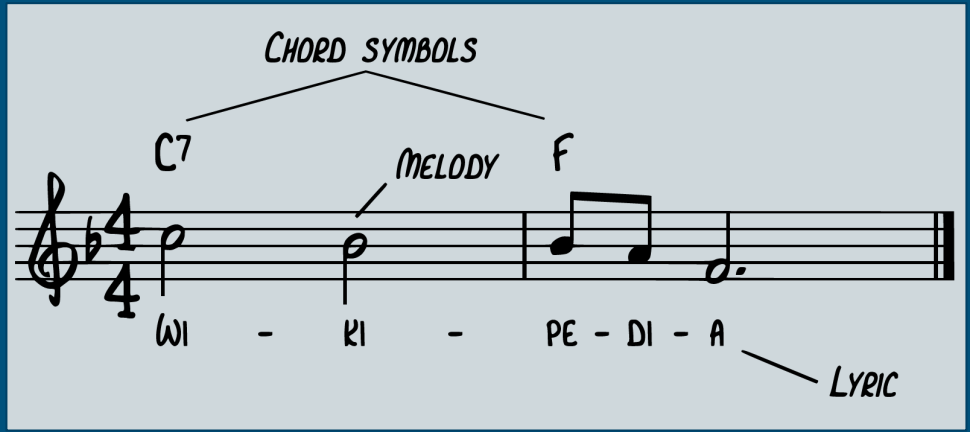
Benjamin Geyer
bgeyer3@gmu.edu

Problem Overview

- Goal:
  - Generate pleasing sounding music
- Problem
  - Tough to evaluate new music without human scoring
- Solution
  - Utilize music theory to produce theoretically good music
- Proposed Method
  - Train a model to generate chord progressions for a given song melody

Dataset

- Lead Sheet Database
  - 11,000 Songs Total
  - Pianoroll: timesteps x pitches
    - 1300 Songs



Input:

[{"note":"A", "duration":2},

{"note":"B", "duration":1},

{"note":"C", "duration":0.25},

{"note":"D", "duration":0.5},

{"note":"A", "duration":1}]

Output:

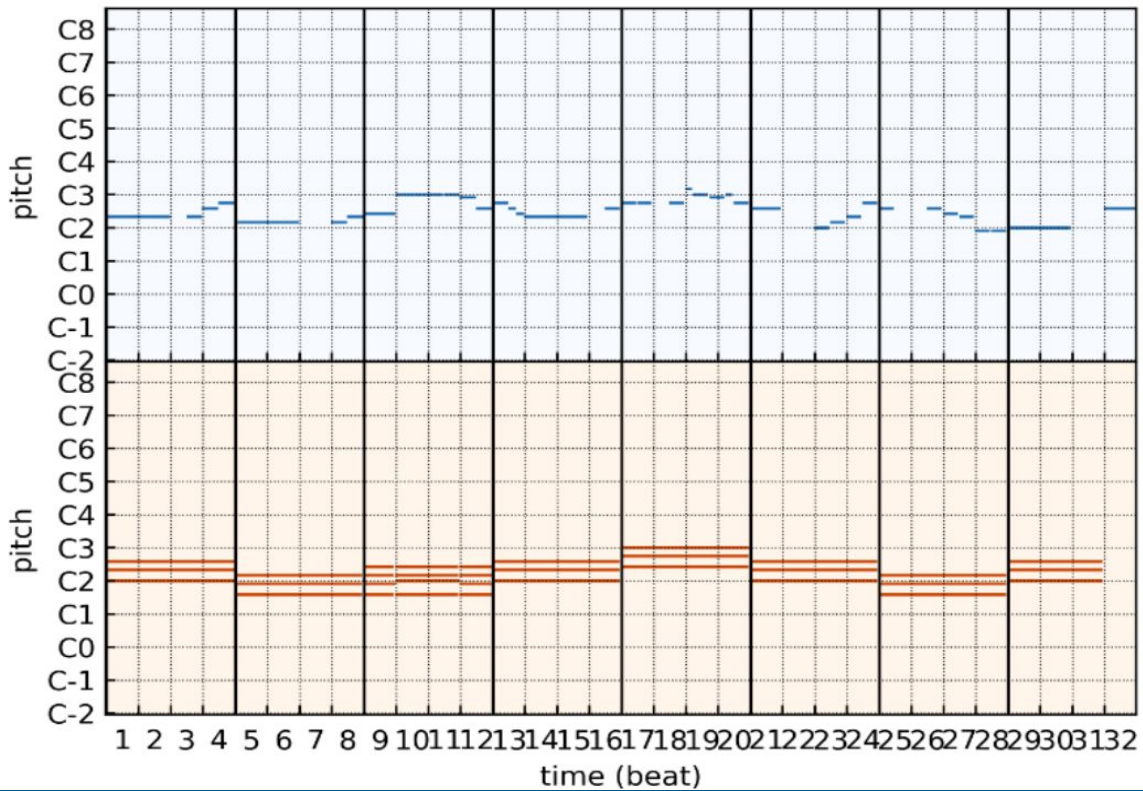
[{"chord":"C.E.G", "duration":2},

{"chord":"A.D.E", "duration":1},

{"chord":"C.F.G", "duration":0.25},

{"chord":"A.D.G", "duration":0.5}]

Example


pitch
C8
C7
C6
C5
C4
C3
C2
C1
C0
C-1
C-2
pitch
C8
C7
C6
C5
C4
C3
C2
C1
C0
C-1
C-2
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32
time (beat)

Methodology

- Feed melody to the network
- Get a chord progression as output

- Evaluate quality:
  - Compare generated chords to the known chords of the song
  - Also investigated predicting melody along with chords
    - Next note per timestep

Model Architecture

- Embedding from notes/chords to vectors
- LSTM / BiLSTM
- Hidden Layers repeated
  - Dropout, Dense, ReLU
- Melody, Chord Prediction Layers
  - Take hidden layer output
  - Dropout, Dense, ReLU
  - Predict most likely next note/chord via softmax activation

Evaluation

- Loss: Negative Log Likelihood
  - Chord Predictions
  - Melody Predictions

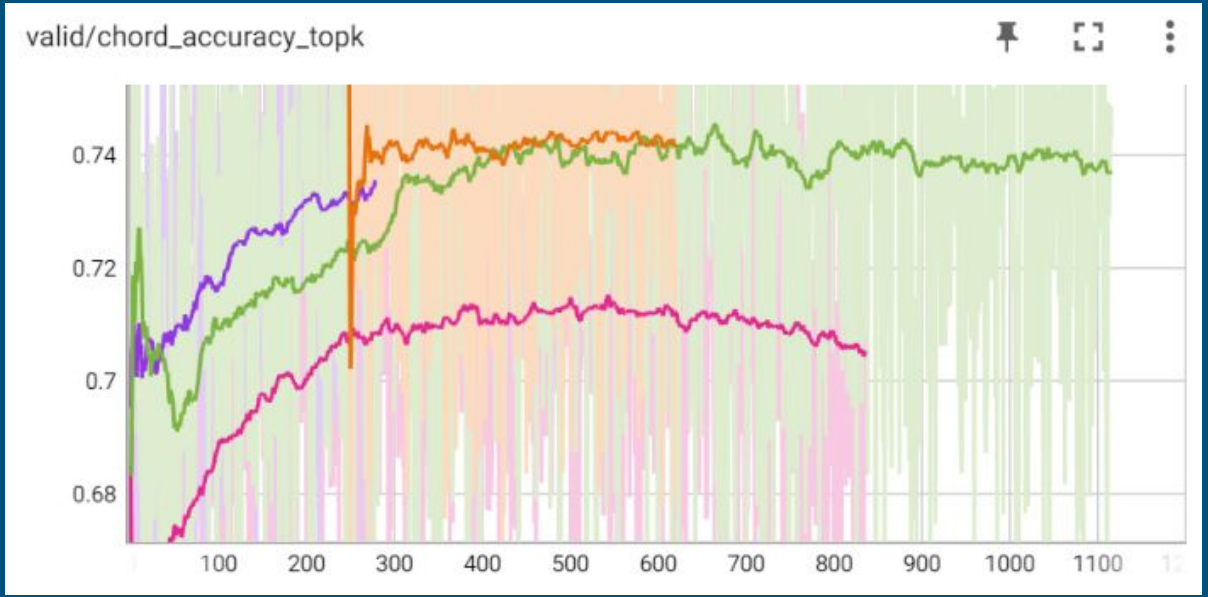
- Metrics:
  - Accuracy
  - Top-K Accuracy (K=3)

Experiments

- Combined loss for chord and melody vs just chords
- BiLSTM vs LSTM

- Tensorboard for tracking and visualization
- Up to 100 epochs, patience of 10 epochs

Results



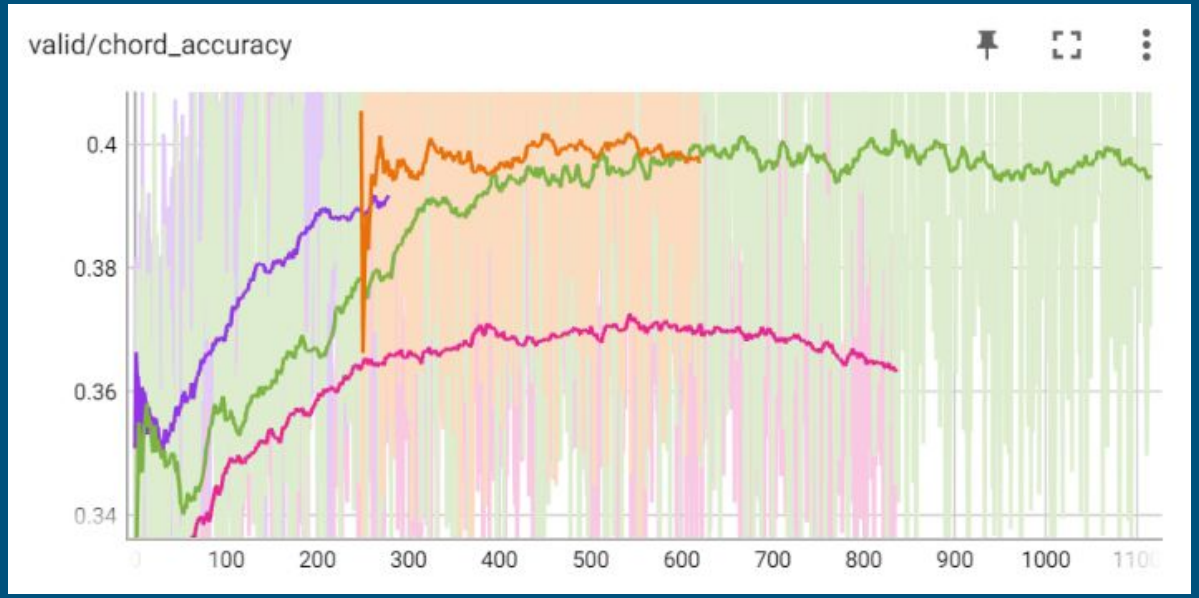
LSTM Chord & Melody

BiLSTM Chord & Melody Part 1

BiLSTM Chord & Melody Part 2

BiLSTM Only Chord

Results



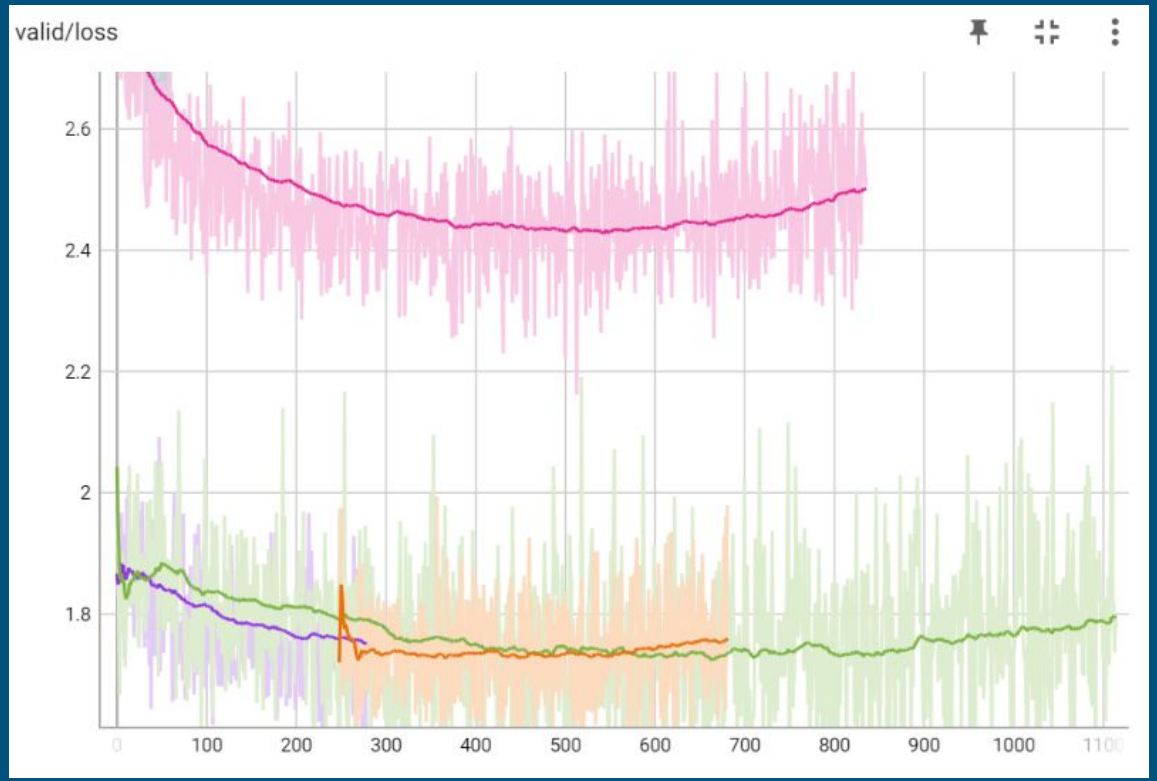
LSTM Chord & Melody

BiLSTM Chord & Melody Part 1

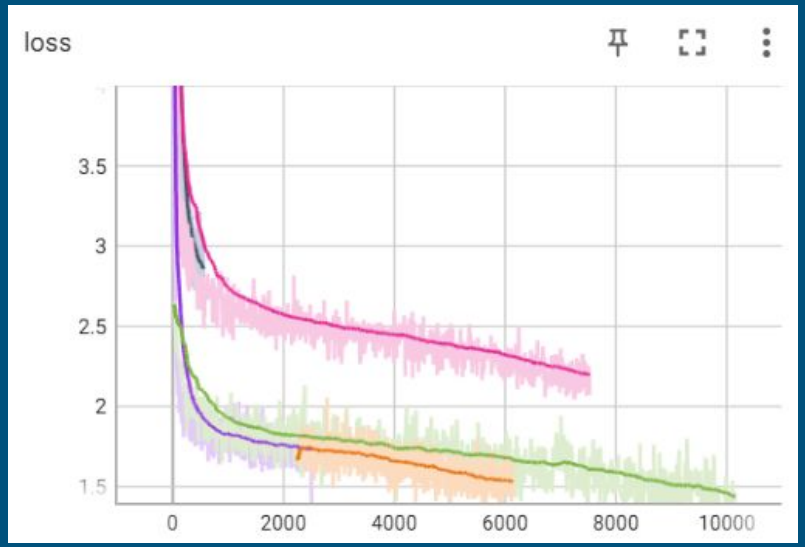
BiLSTM Chord & Melody Part 2

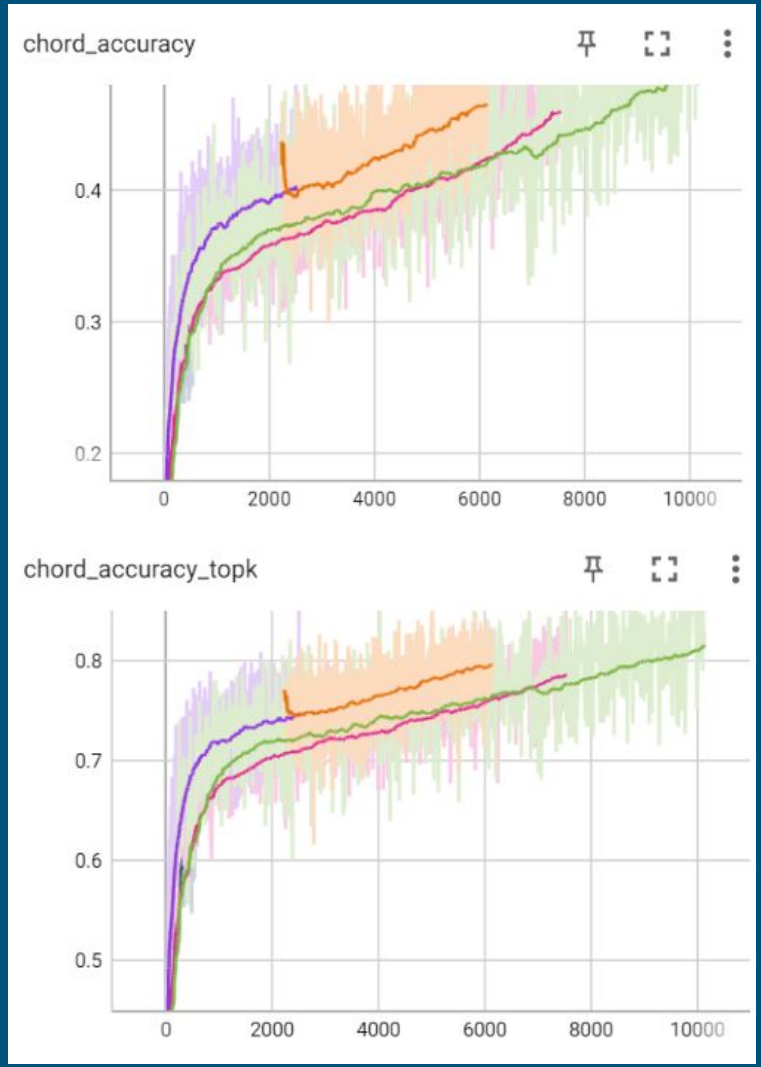
BiLSTM Only Chord

Results


valid/loss
2.6
2.4
2.2
2
1.8
0
100
200
300
400
500
600
700
800
900
1000
LSTM Chord & Melody
BiLSTM Chord & Melody Part 1
BiLSTM Chord & Melody Part 2
BiLSTM Only Chord

Overfitting





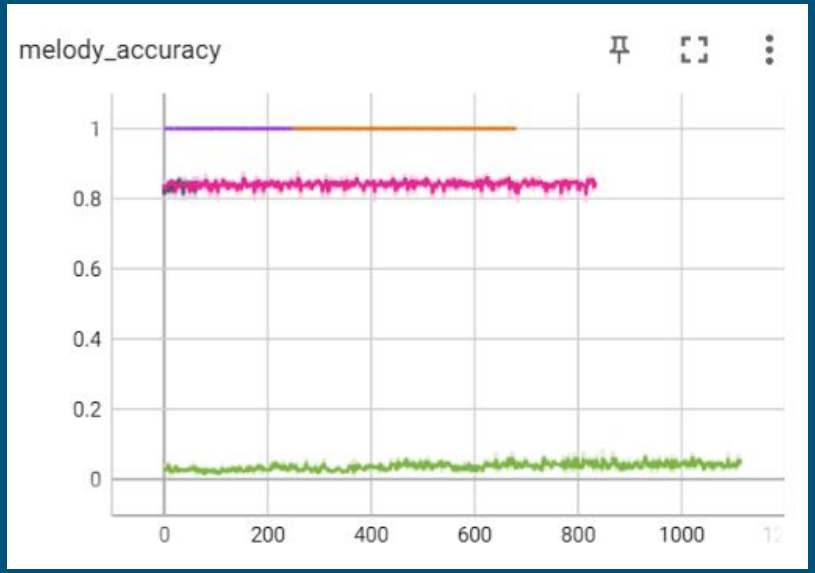
Melody Metrics

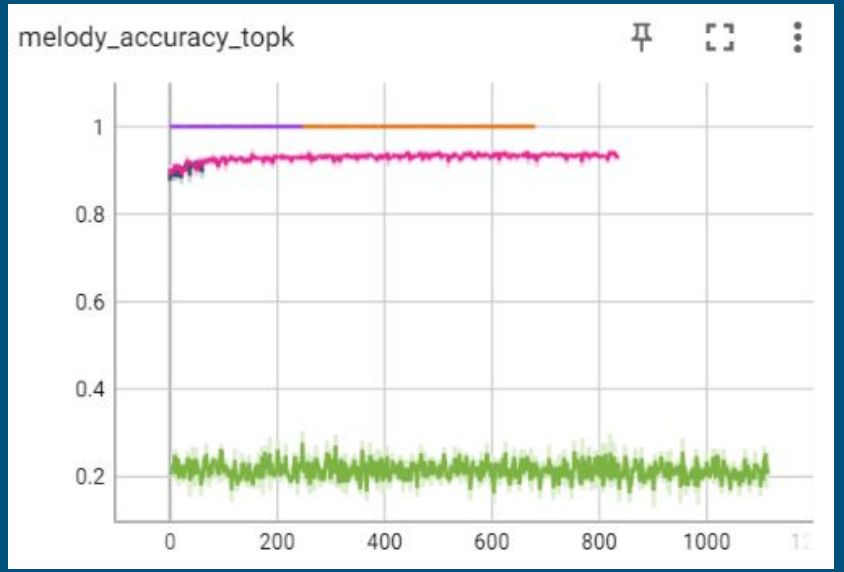
LSTM Chord & Melody

BiLSTM Chord & Melody Part 1

BiLSTM Chord & Melody Part 2

BiLSTM Only Chord





Findings and Discussion

- Combination of Chord and Melody loss was best
- BiLSTM was better than LSTM
  - Only applicable in some situations
- Lots of Overfitting

Conclusion

- Yes, a machine can learn to generate pleasing music

- Caveats for my project:
  - Limited to two pieces, melody and a chord progression
  - Not polyphonic / multiple instruments
  - Live composition is tricky with BiLSTM

Future Work

- Overfitting prevention such as Dropout
- BiLSTM with masking for use in live scenarios
- Adaptive Loss Function
  - Melody + Chord loss to just Chord as time goes on
- More advanced model architecture
  - Attention, Transformers, etc

Thank You