Music Generation

A Transformer-Based
Chord Generation Model

RYAN GHIMIRE & ANIKETH LUCHMAPURKAR

Problem

- Melodies without chords/accompaniments sound empty
- Many struggle to compose accompaniments
- Current models generate full songs
    - These models restrict, not foster, creativity

Related Work

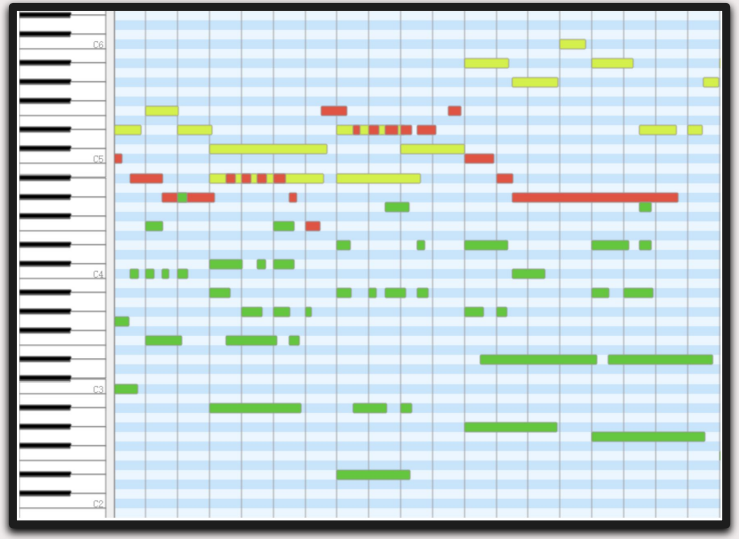
- Music generation began with statistical modeling to generate melodies (Conklin, 2003)
- Transformer and Recurrent Neural Network (RNN) models
    - DeepBach,
        - RNN model focused solely on generating classical music pieces from scratch (Hadjeres et al., 2017).
- These models have further evolved to generating:
    - Varied music styles (Mao et al., 2018)
    - Songs based on melodic feature input (Copet et al., 2024).
- Google Magenta
    - Specialty in chord generation
- Spotify CTRL working on AI tools for music creation assistance

Solution

- Model that generates chords for melodies
- Predicts chord from melody fragment and past chords, ensuring...
  - Chord matches notes
  - Consistent chord progressions

Dataset & Preprocessing

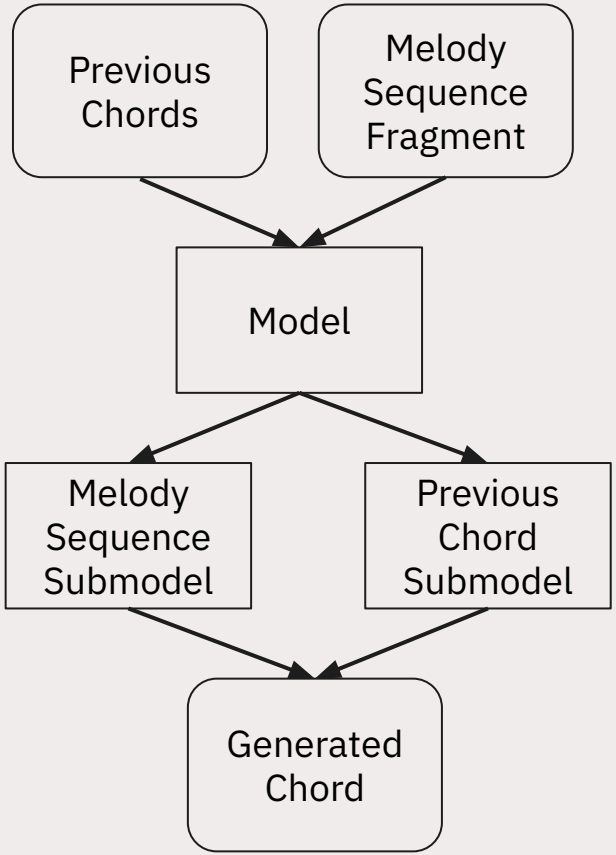
- POP909 dataset: 909 pop melodies and accompaniments
- Broke melody into fragments by associated chord
- Constructed CSV
  - Each instance: <Key, Melody Sequence, Past N Chords, Chord>
- Numericized “Past N Chords” (feature) and “Chord” (labels) with label encoder



The Model

Split into sub-models:

- Melody Sequence Sub-Model
- Previous Chord Sub-Model

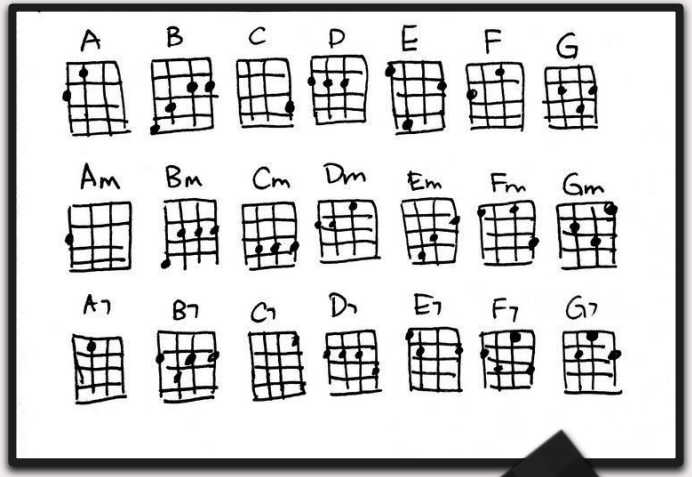


Melody Sub-Model

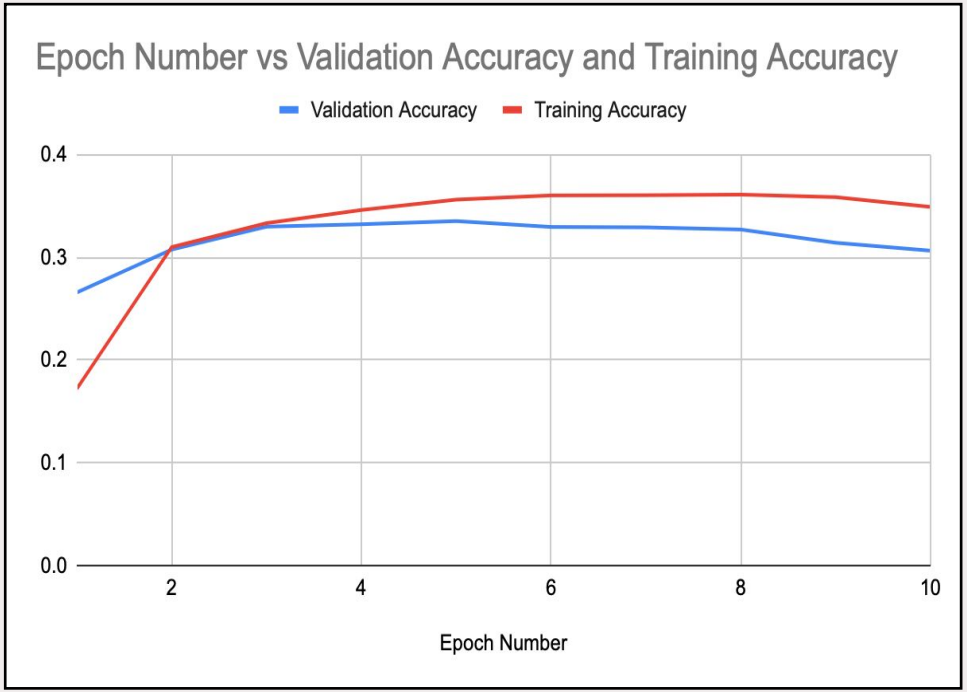
- Input layer
- Embedding layer
  - Lowers dimensionality to make it able to draw relationships between parts of input
- Transformer Layer
  - Self-attention
    - Discovers what is important in melody
    - Develops correlations and relationships
  - Multi-head attention
    - Parallel runs of self-attention
- Global Pooling Layer
  - Average into singular vector

Previous Chord Sub-Model

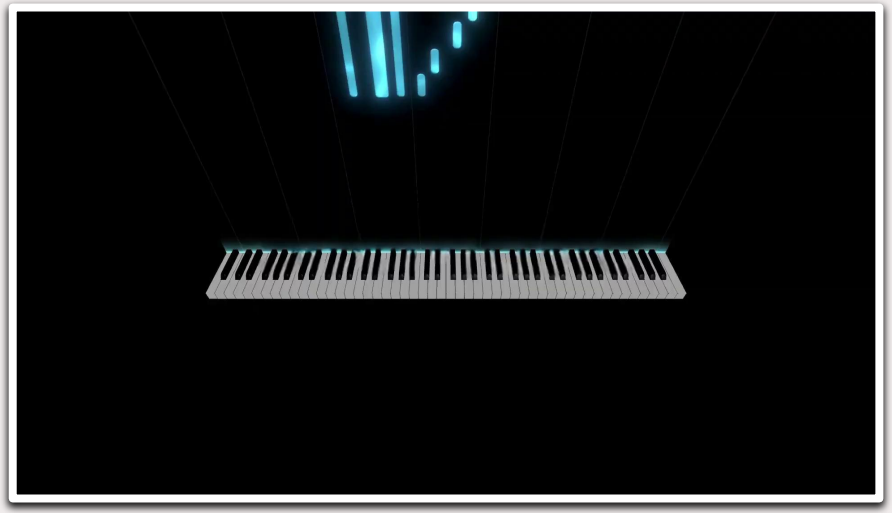
- Input Layer
- Embedding Layer
- Flatten Layer
  - Multidimensional Vector → 1D vector (flattening)
- Dense Hidden Layer
  - Finds correlations/patterns in flattened data

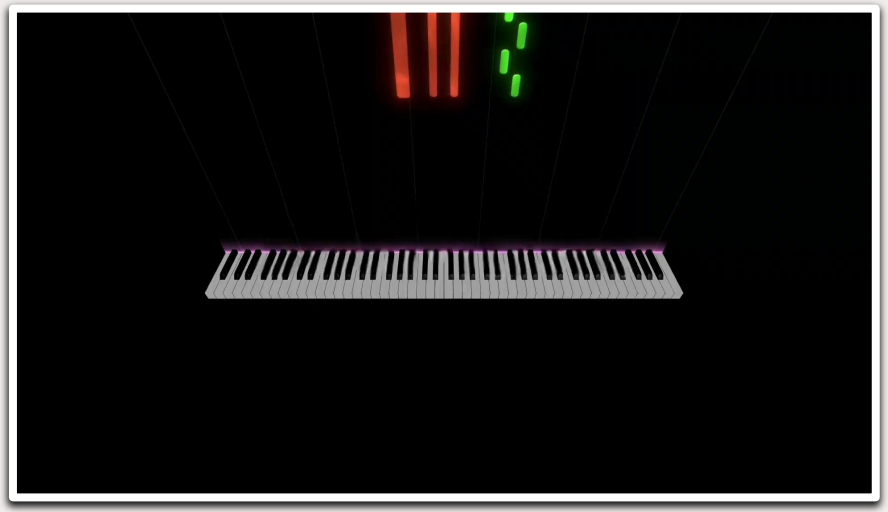


Results


Epoch Number vs Validation Accuracy and Training Accuracy
Validation Accuracy
Training Accuracy
0.4
0.3
0.2
0.1
0.0
2
4
6
8
10
Epoch Number

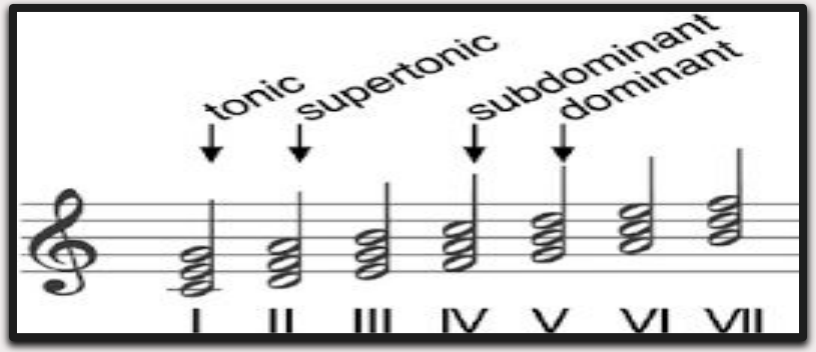
Results





Conclusion

- Model predicts harmonically fitting chords
- Struggles with repetitive chord progressions
- Needs better focus on musical structure
- Future Improvements
  - Incorporate harmonic function data
  - Expand to tempo and dynamic variation



Thank You!

Any Questions?