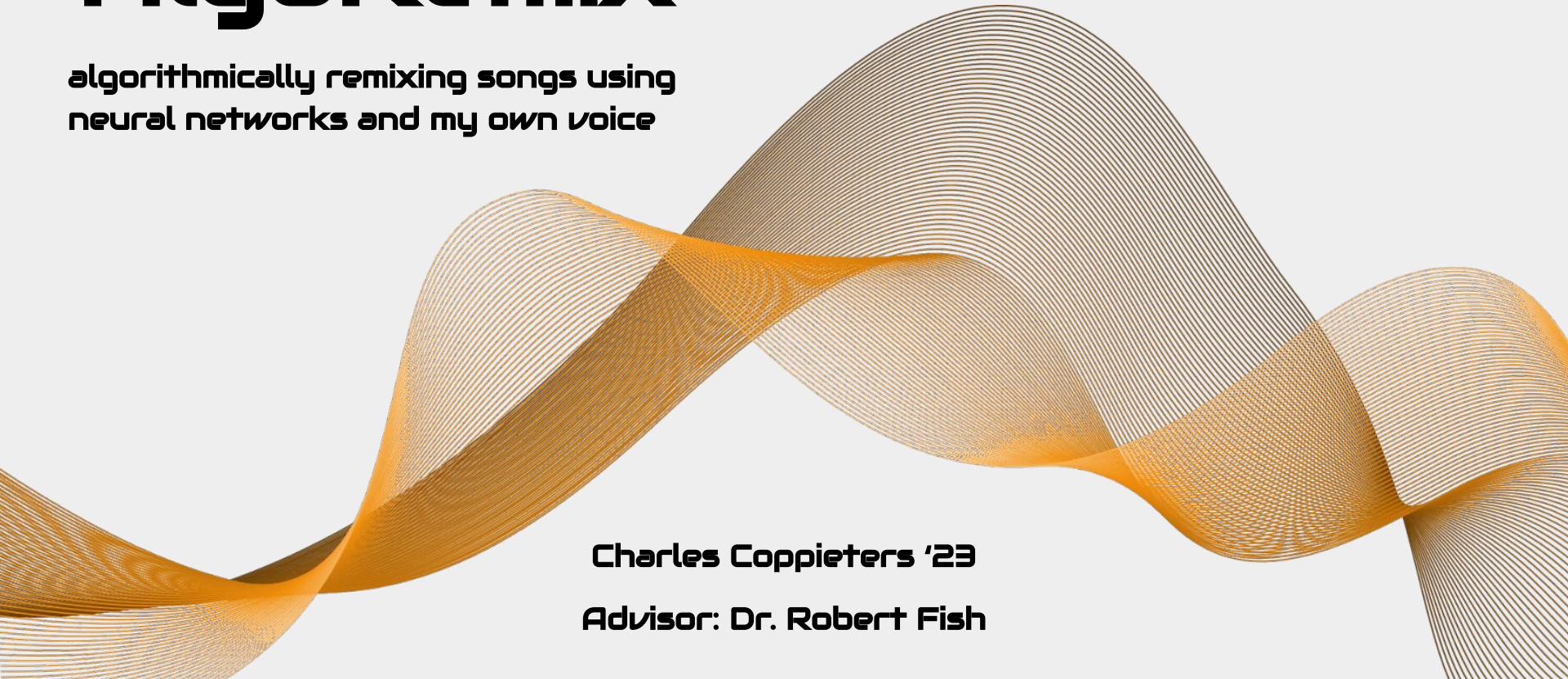
AlgoRemix

algorithmically remixing songs using neural networks and my own voice



Charles Coppieters ‘23

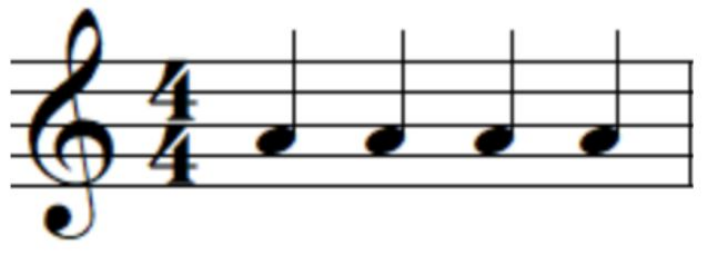
Advisor: Dr. Robert Fish

Motivation

Where is the AI generated music?

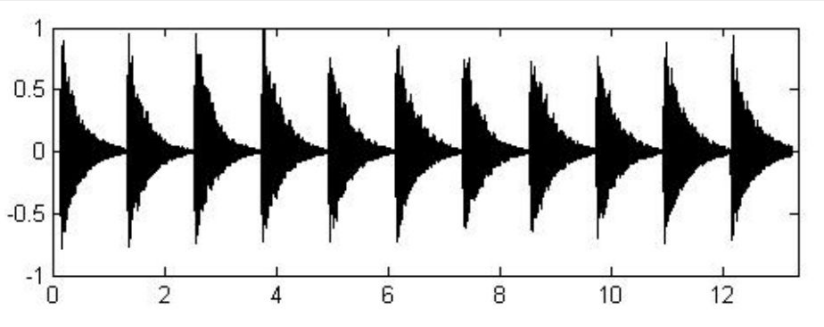
Background

- Beats and measures
- Notes
  - Pitch, duration, offset
- Key and harmonic context

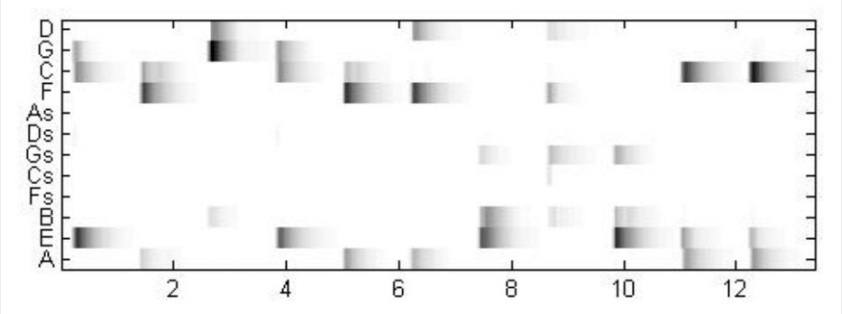


Background

- Audio files



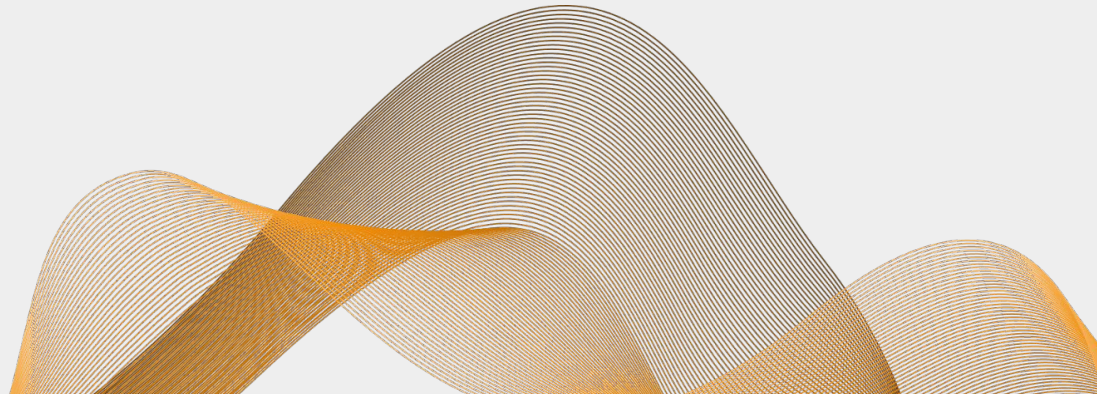
- MIDI



Related Work

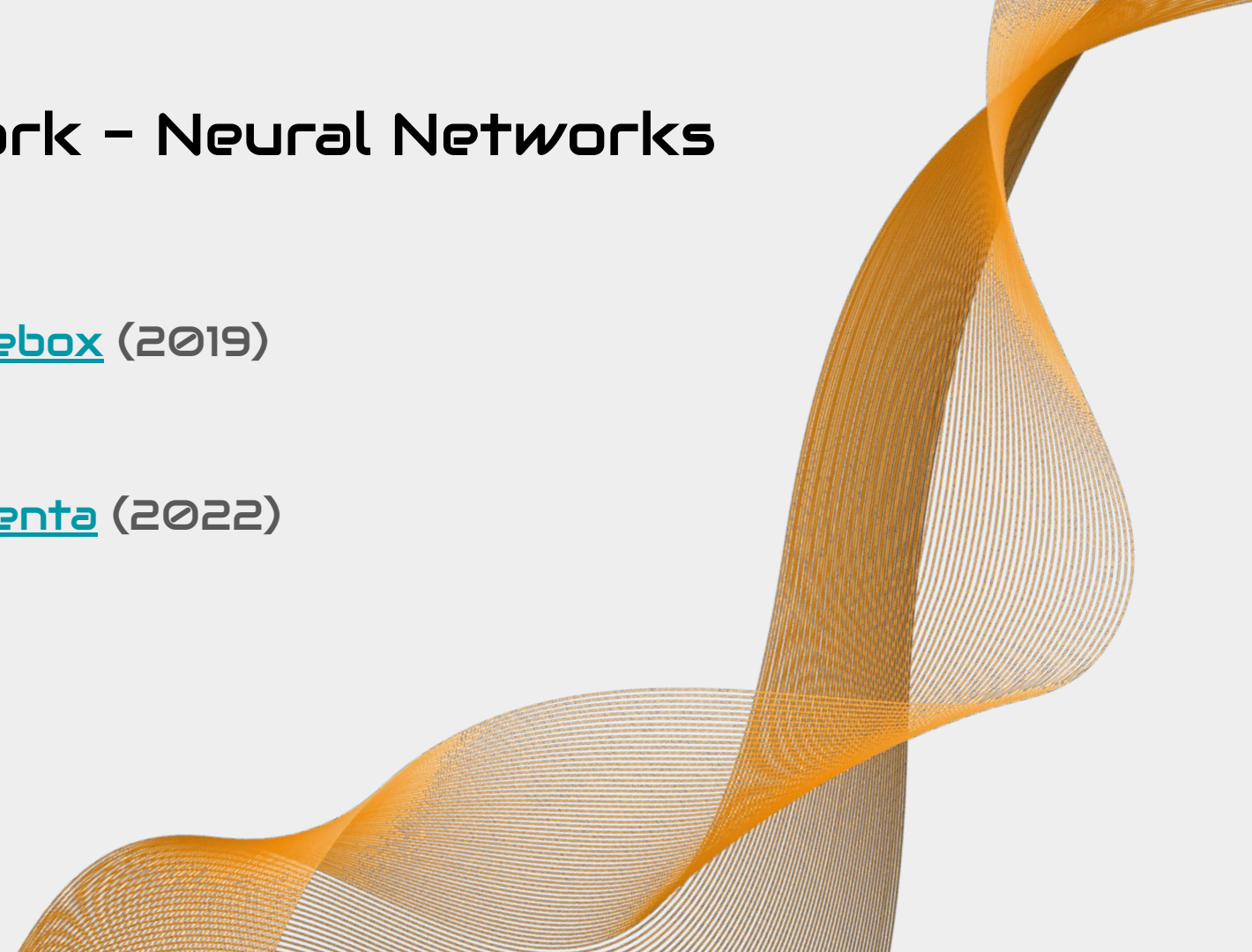
- Alan Turing (1951)

- Lejaren Hiller and Leonardo Isaacson's "Illiac Suite" (1957)



Related Work - Neural Networks

- OpenAI Jukebox (2019)
- Google Magenta (2022)

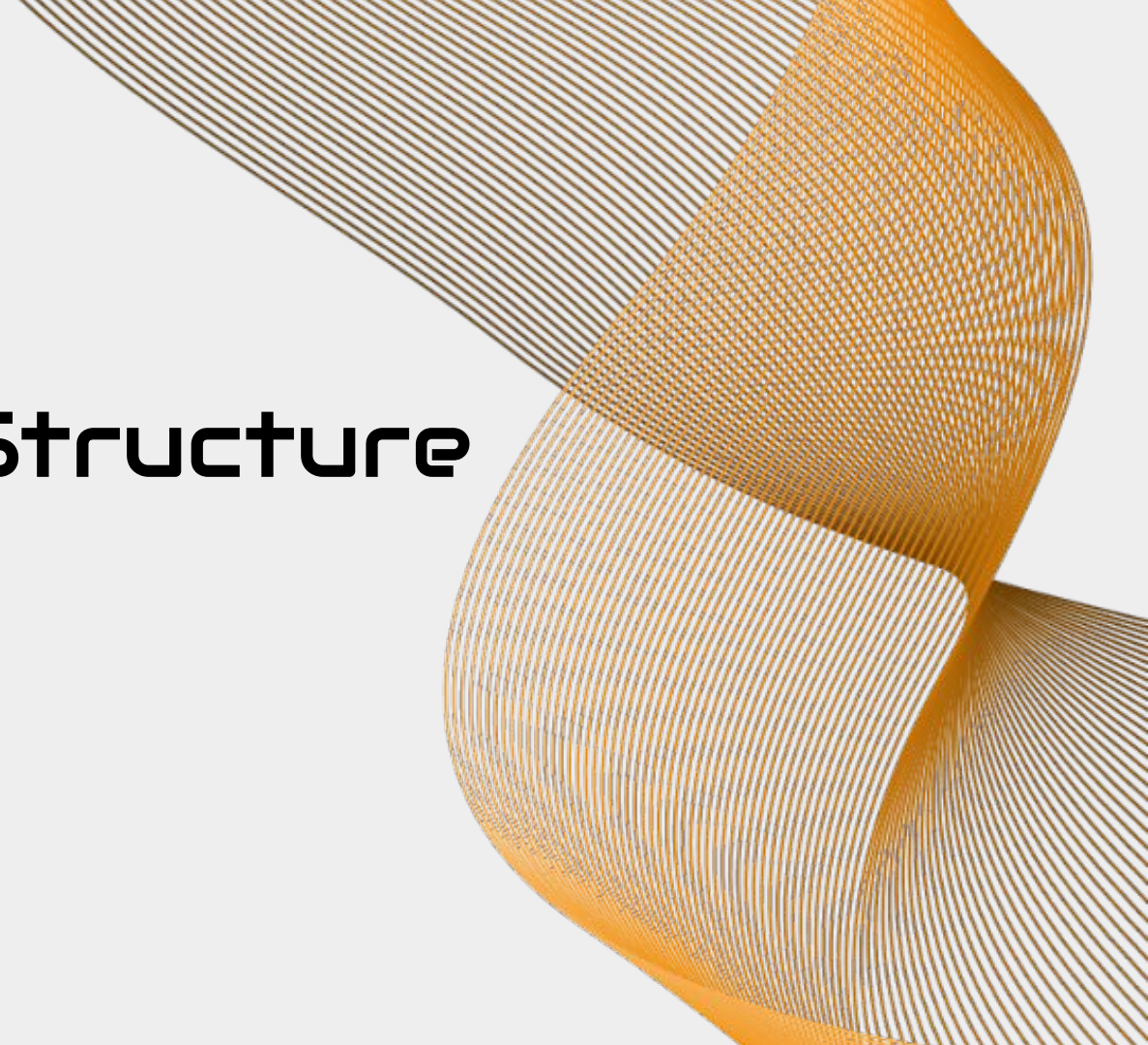


Related Work - Diffusion Models

- Tencent Diffsound (2022)
- Google Noise2Music (2023)

The Existing Problem:

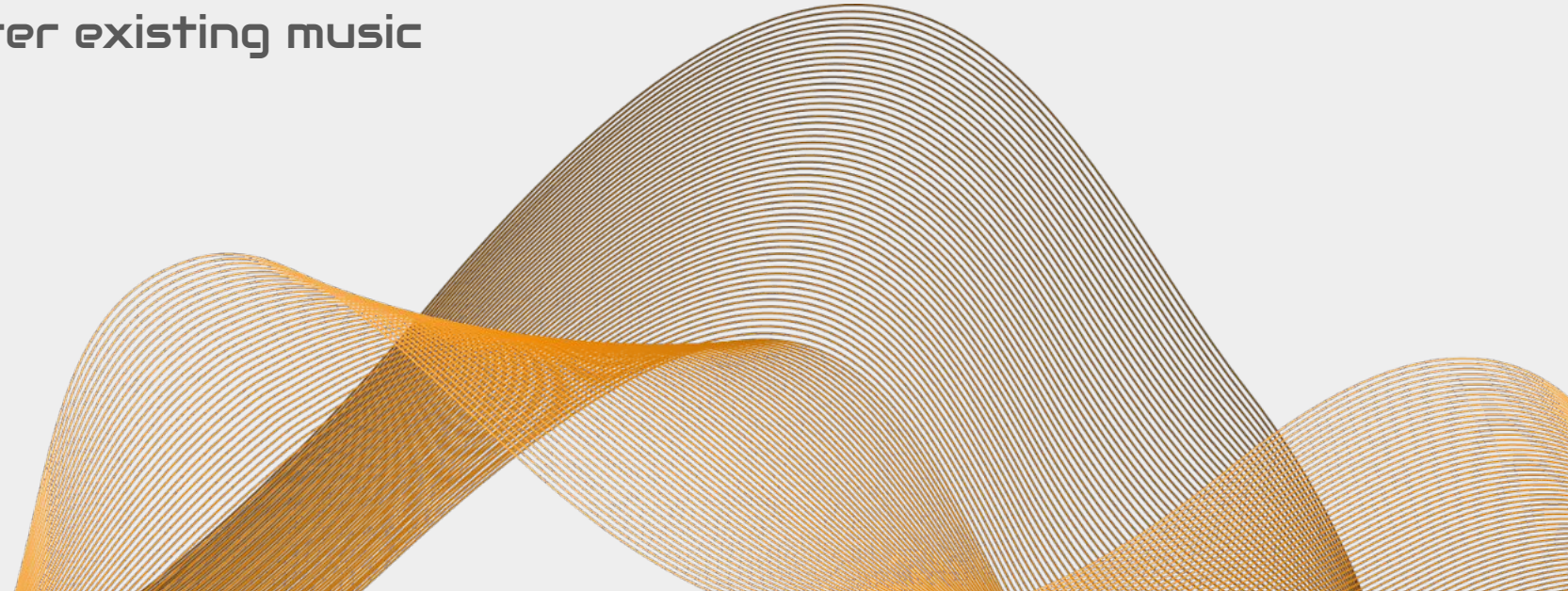
Long Term Structure



Goal

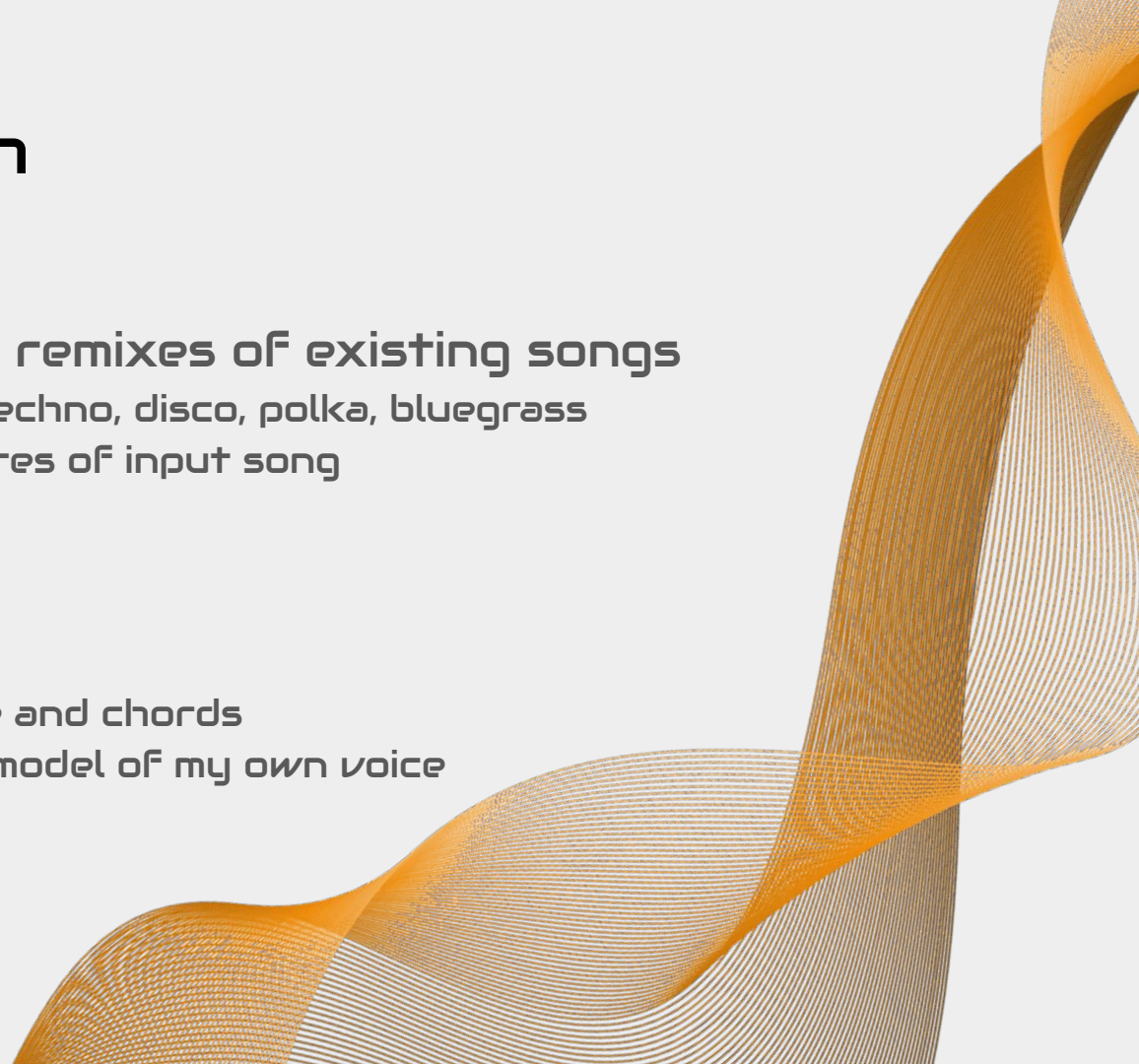
- Work around long term structure issue

- Alter existing music



Goal and Approach

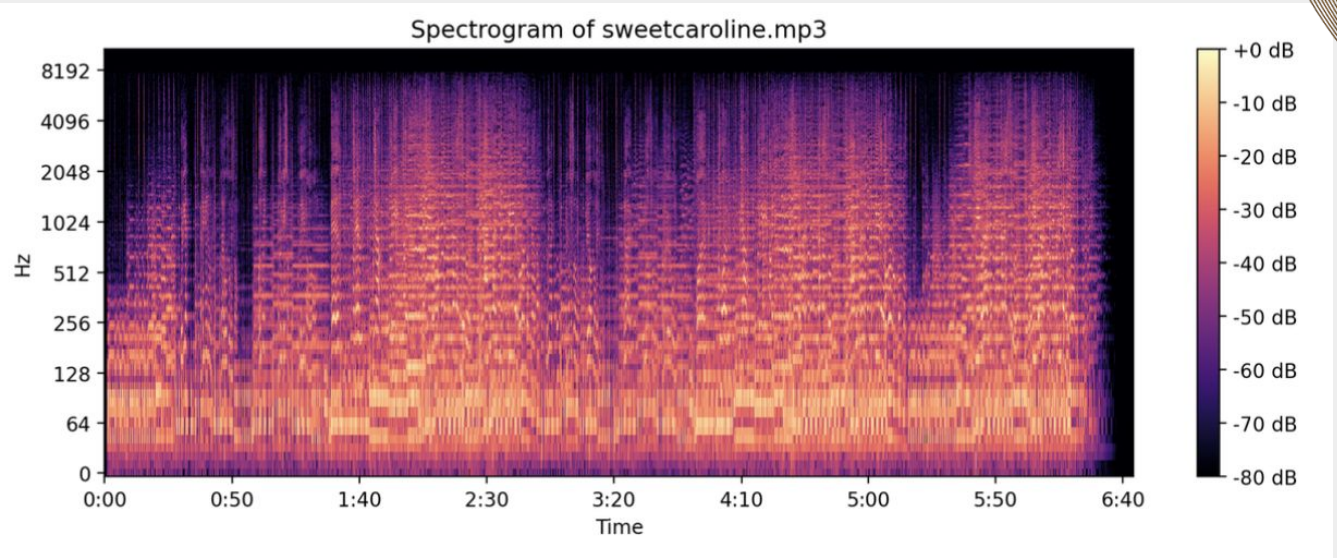
- The Wub Machine
- ML model to generate remixes of existing songs
  - 5 genres: bossa nova, techno, disco, polka, bluegrass
  - Analyzes musical features of input song
    - BPM
    - Beat Grid
    - Key
    - Chord Progression
  - Generates *new* bassline and chords
  - Generates lyrics using model of my own voice



Implementation

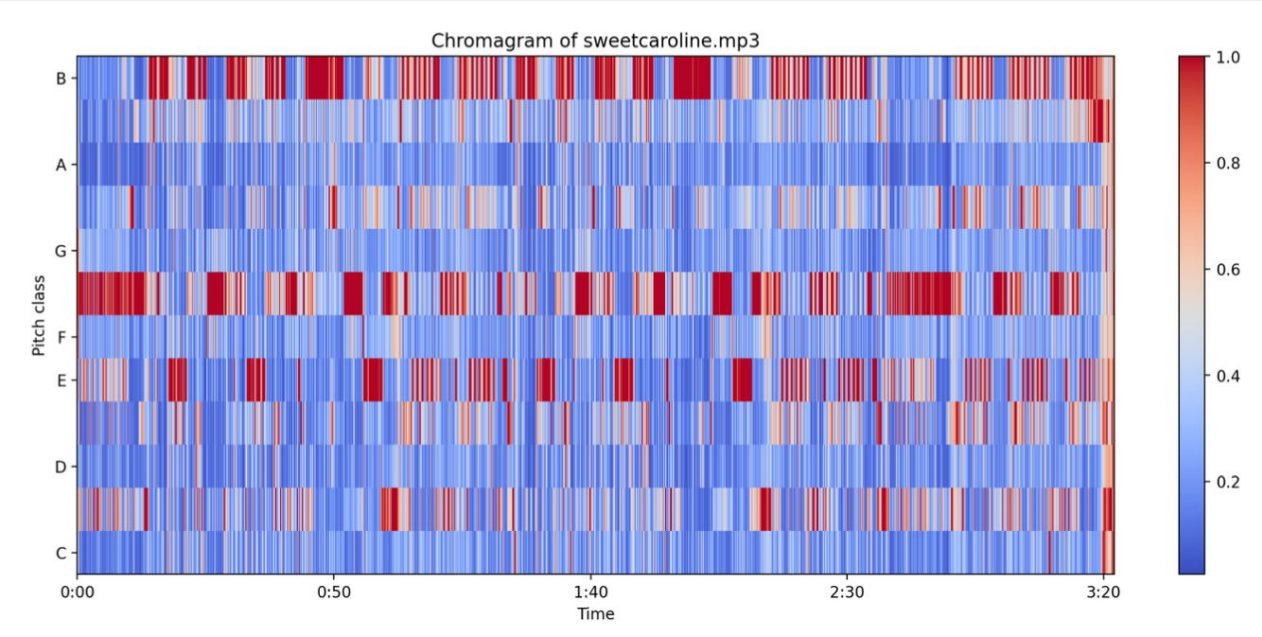
Song Analysis - Preprocessing

- Spectrogram



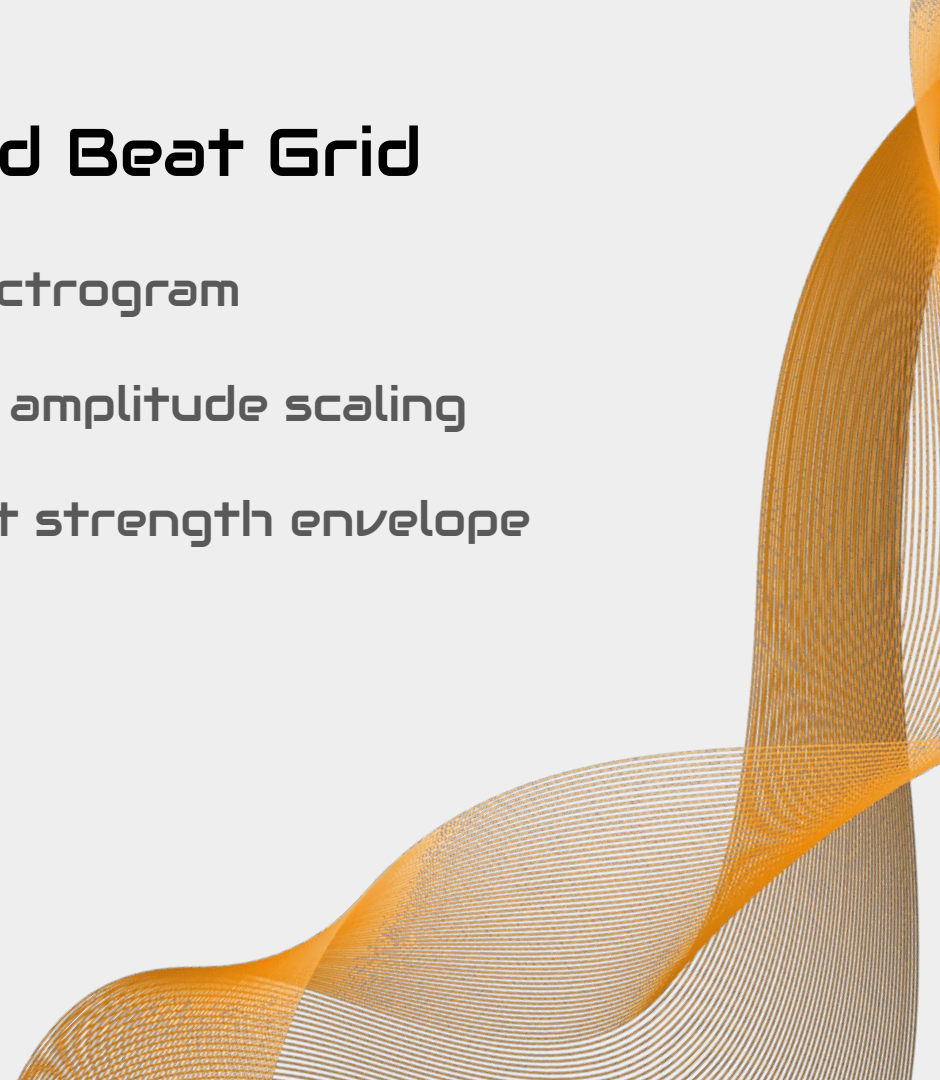
Song Analysis - Preprocessing

- Chromagram



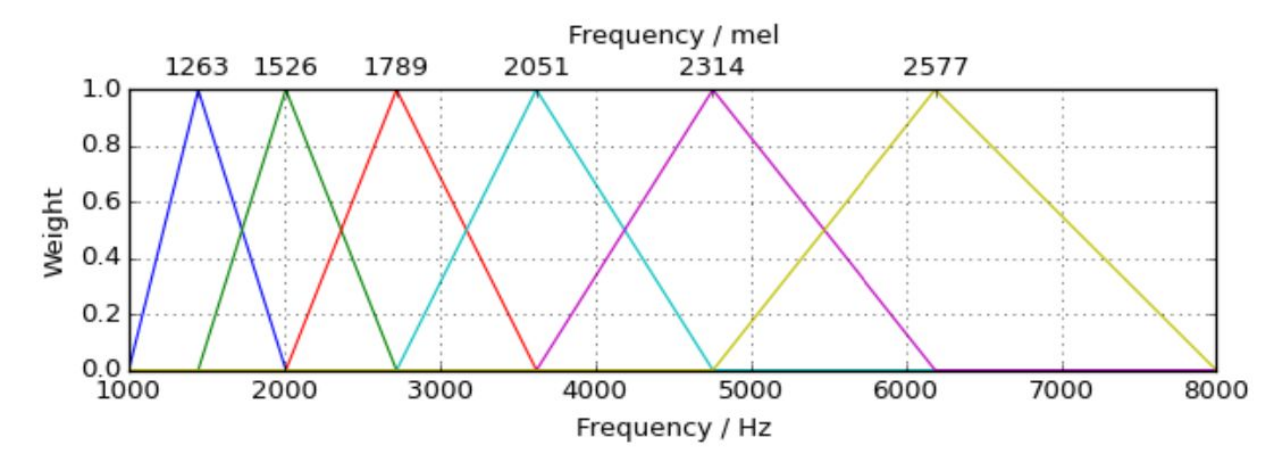
Song Analysis - BPM and Beat Grid

1. Spectrogram -> Mel-scale spectrogram
2. Mel-scale spectrogram -> Log amplitude scaling
3. Log amplitude scaling -> Onset strength envelope
4. BPM estimation



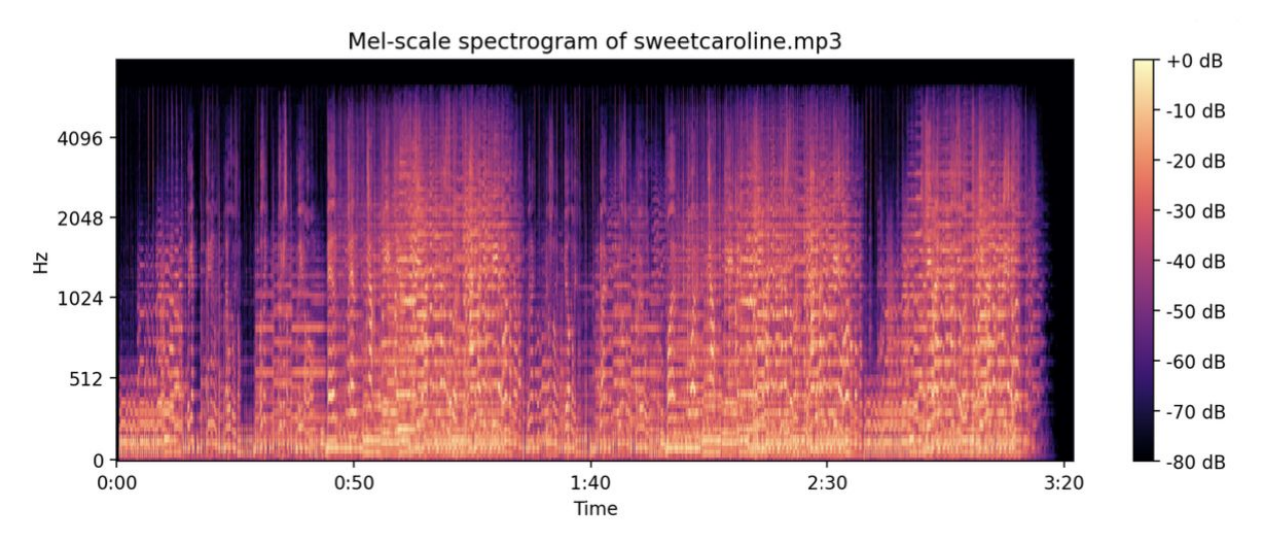
Song Analysis - BPM and Beat Grid

1. Spectrogram -> Mel-scale spectrogram



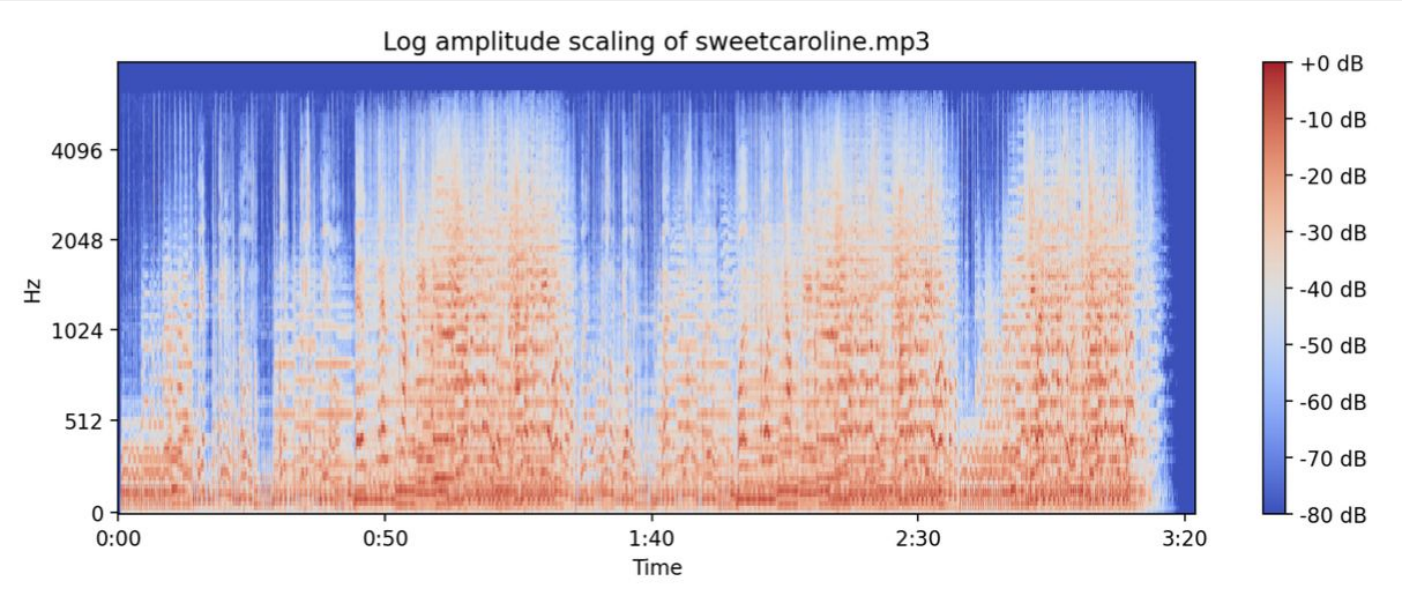
Song Analysis - BPM and Beat Grid

1. Spectrogram -> Mel-scale spectrogram



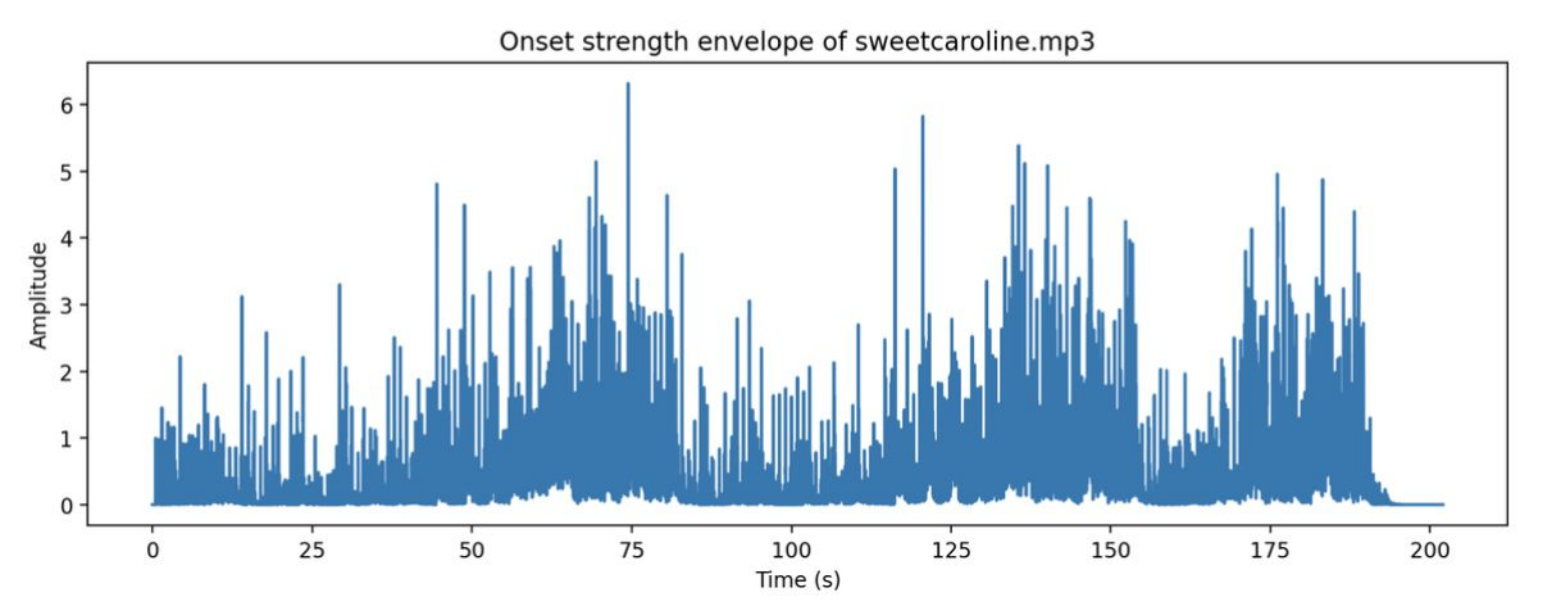
Song Analysis - BPM and Beat Grid

2. Mel-scale spectrogram -> Log amplitude scaling



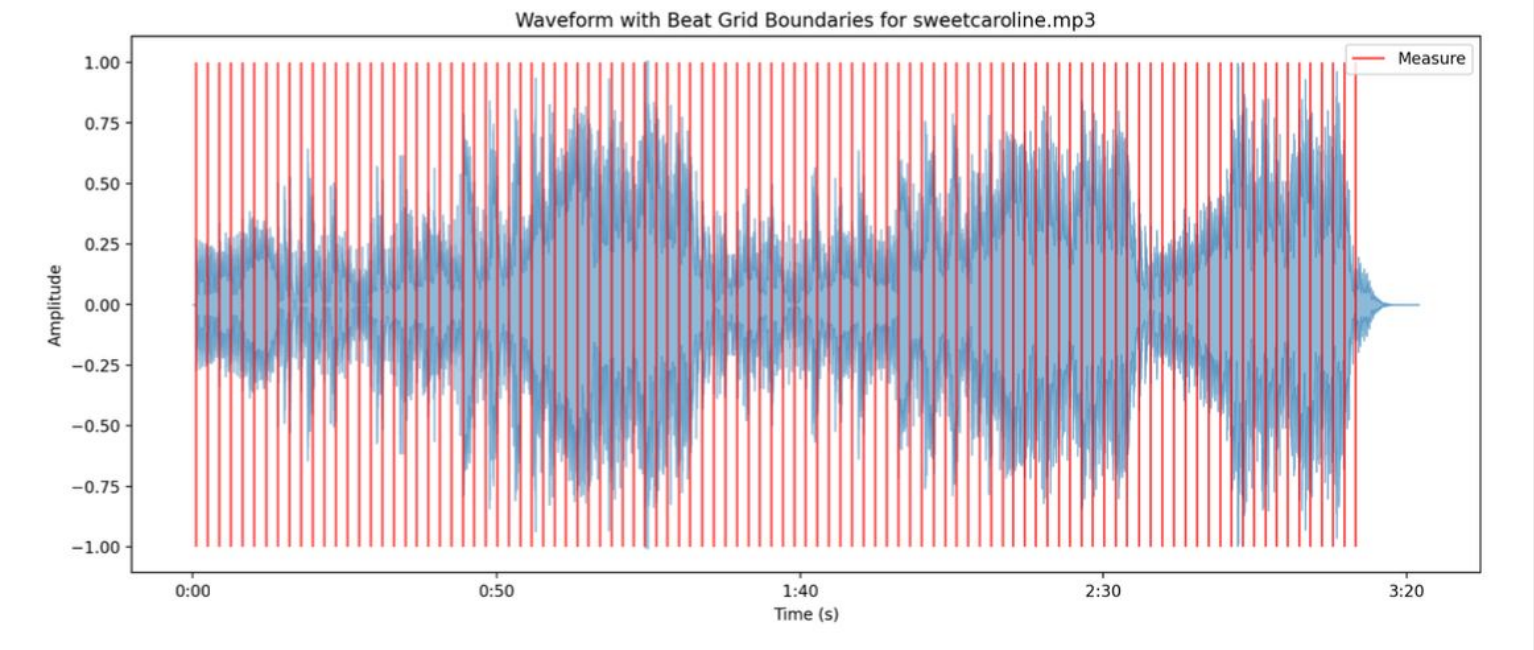
Song Analysis - BPM and Beat Grid

3. Log amplitude scaling -> Onset strength envelope



Song Analysis - BPM and Beat Grid

4. Onset strength envelope -> BPM Estimation



Song Analysis - Key

1. Mean chroma values from chromagram
2. Krumhansl-Schmuckler Algorithm

Major Key Profiles

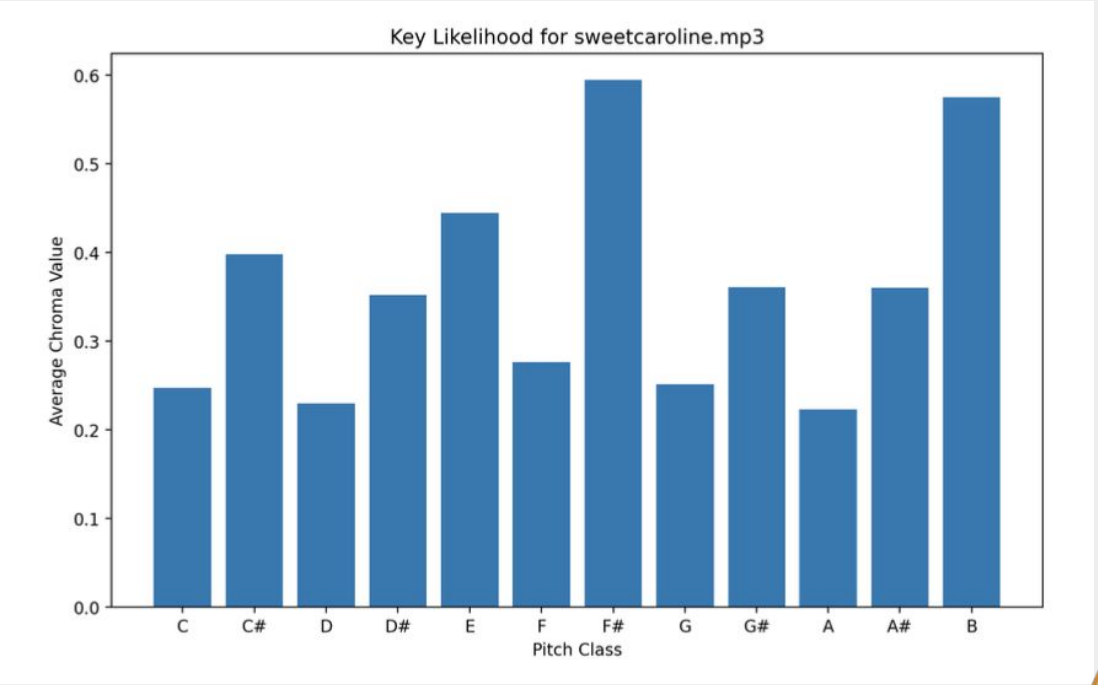
Key	C	C#	D	D#	E	F	F#	G	G#	A	A#	B
Freq	6.35	2.23	3.48	2.33	4.38	4.09	2.52	5.19	2.39	3.66	2.29	2.88

Minor Key Profiles

Key	C	C#	D	D#	E	F	F#	G	G#	A	A#	B
Freq	6.33	2.68	3.52	5.38	2.60	3.53	2.54	4.75	3.98	2.69	3.34	3.17

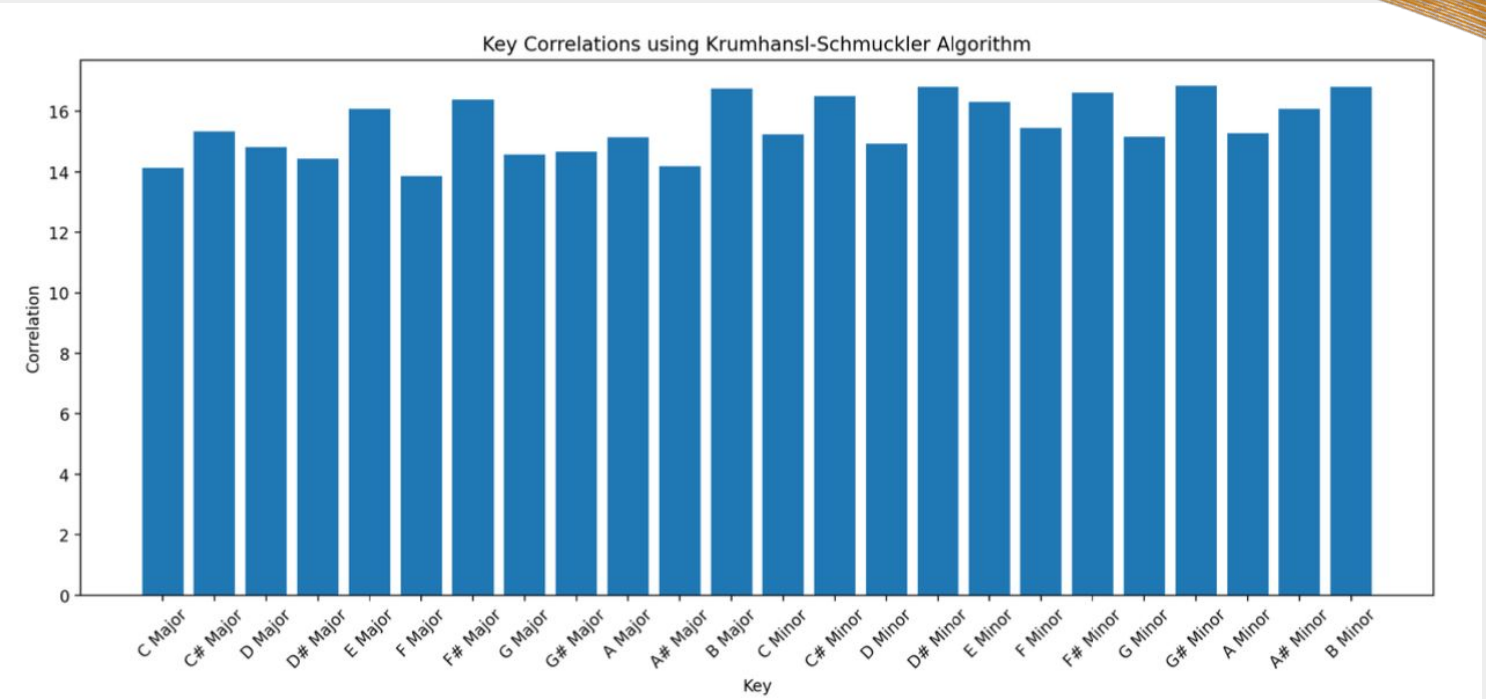
Song Analysis - Key

1. Mean chroma values from chromagram



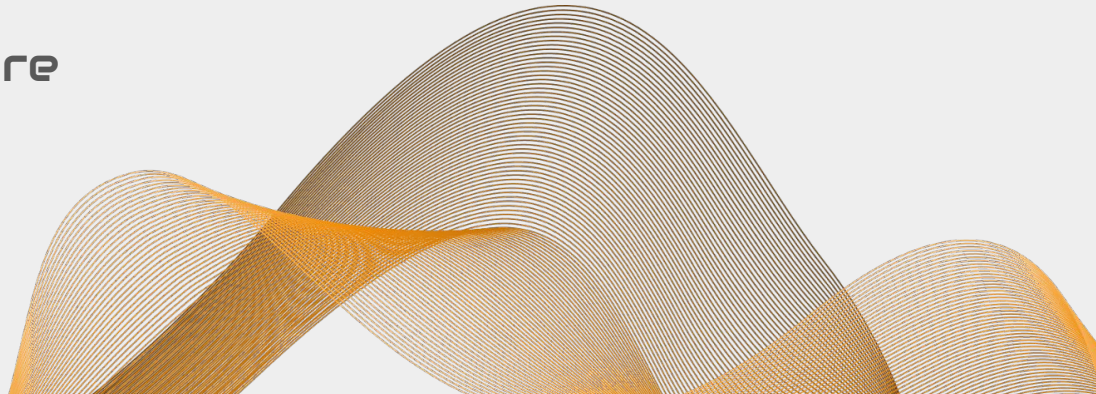
Song Analysis - Key

2. Krumhansl-Schmuckler Algorithm

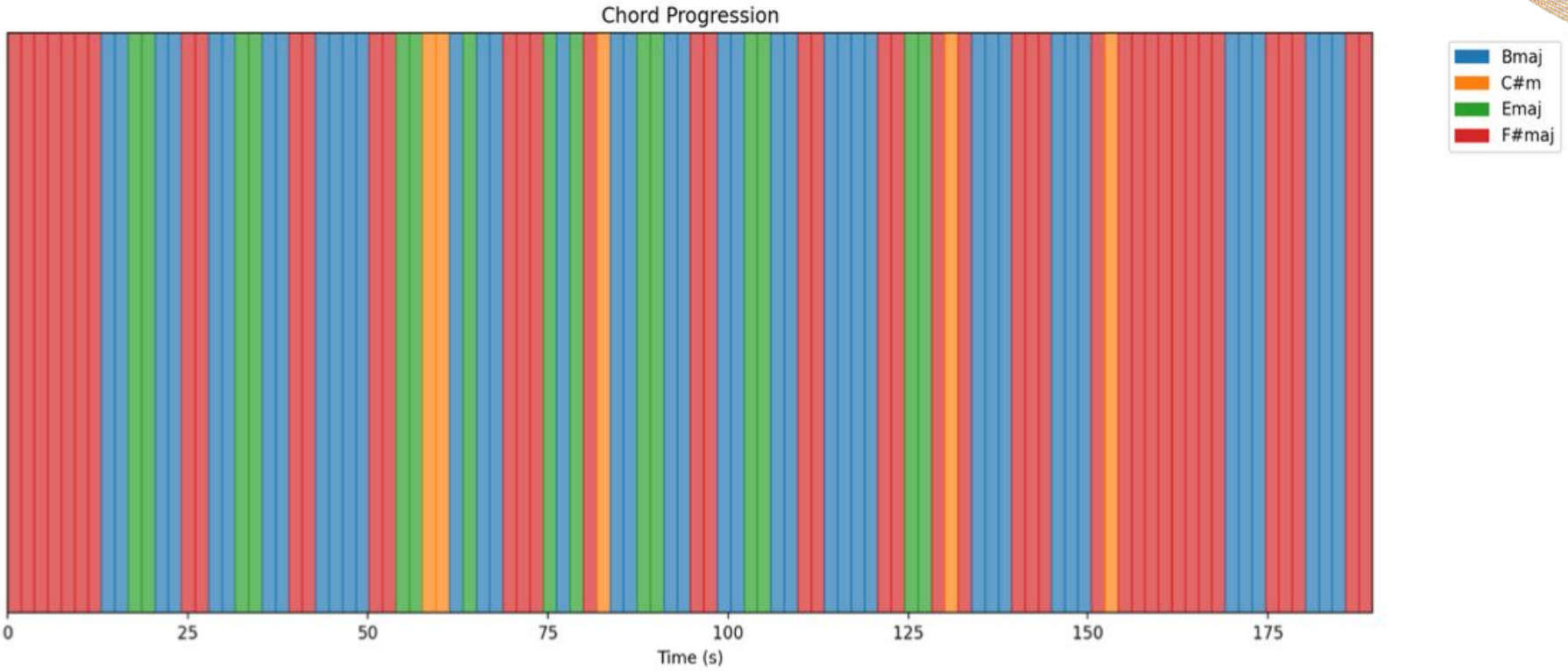


Song Analysis - Chord Progression

1. Sync avg chroma values with beat grid
2. Estimate root note by calculating highest chroma value for each measure
3. Estimate chord quality based on key
4. Repeat for every measure



Song Analysis - Chord Progression



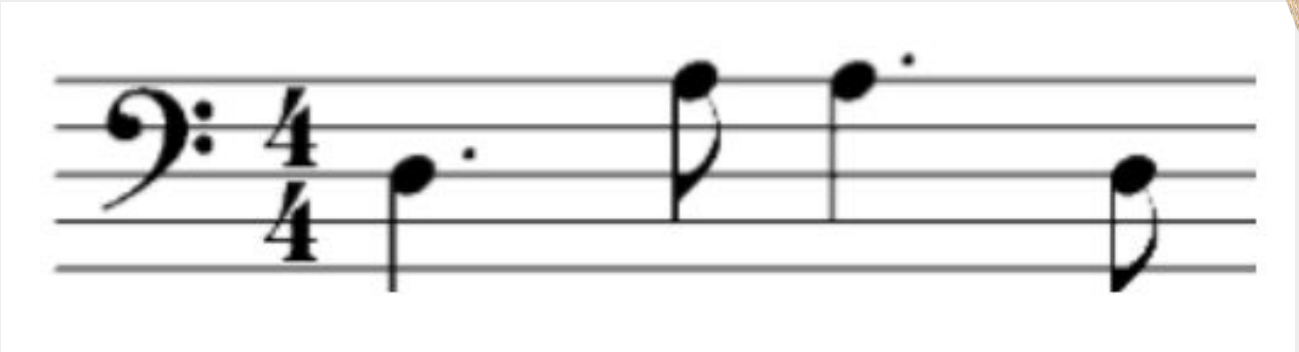
Remix Generation

Data

- MIDI files recorded by me
  - Labeled by relevant chord (“cmaj.midi”)
- Numerical encoding
  - Label: 0-35 (for each chord)
  - Notes: tuple of normalized start time, end time, pitch, velocity

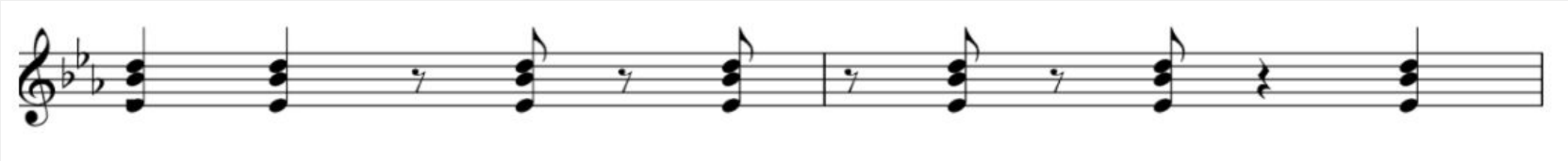
Bassline Data

- Notes and rhythms

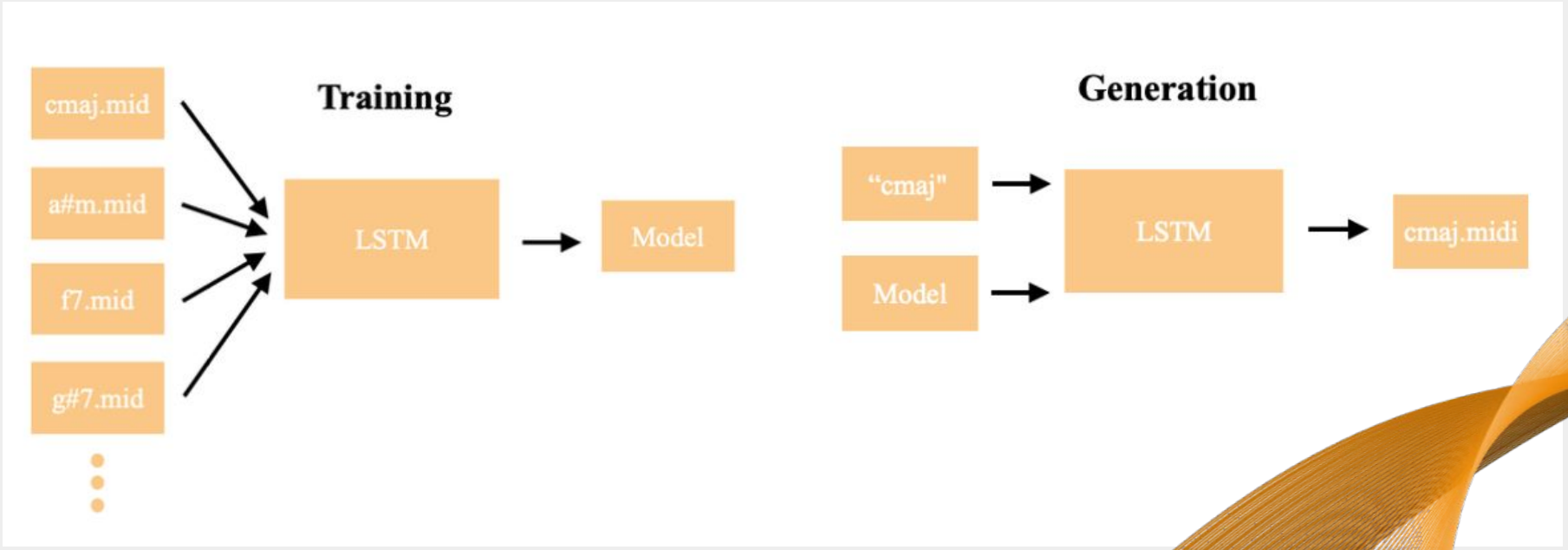


Chord Data

- Rhythms
- Extrapolation of full chords after generation



Model – LSTM


cmaj.mid
a#m.mid
f7.mid
g#7.mid
Training
LSTM
Model
Generation
“cmaj”
Model
LSTM
cmaj.midi

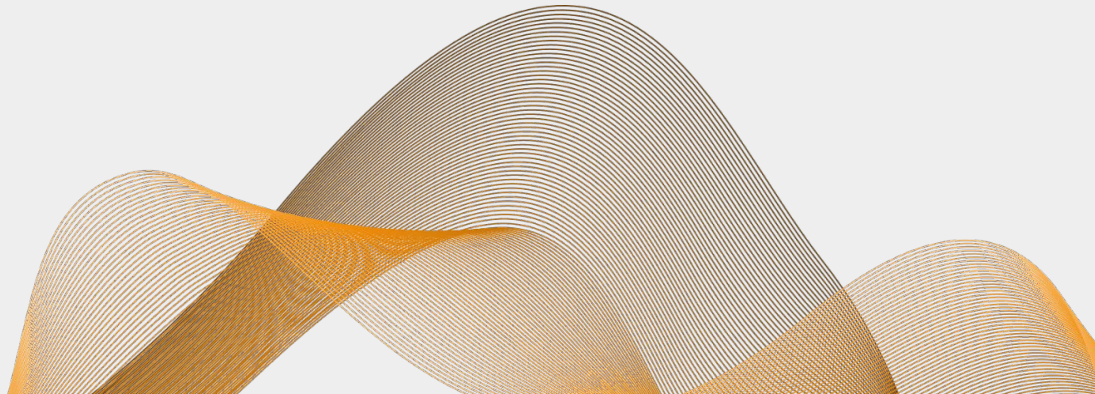
Model - LSTM

```
model = Sequential([
    Embedding(input_dim=num_classes, output_dim=128, input_length=1),
    LSTM(128, return_sequences=True),
    LSTM(128, return_sequences=False),
    Dense(max_len * 4, activation='relu')
])
```

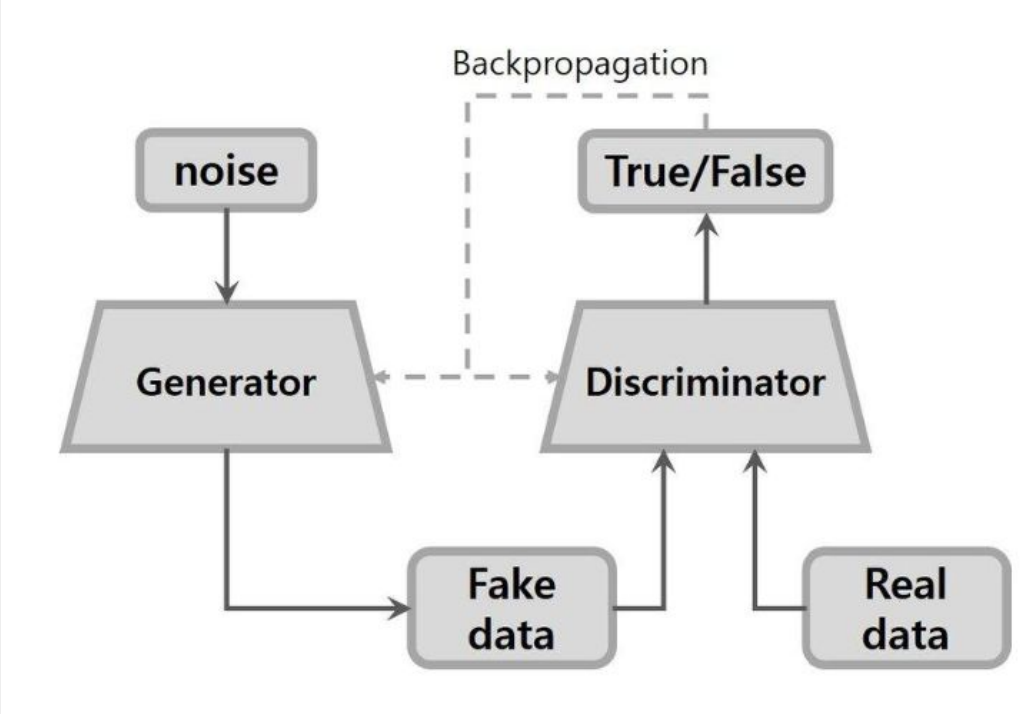
Vocal Generation

Data

- Audio files recorded by me
  - Harvard Sentences
  - 10 second chunks
- Preprocess files into spectrograms



Model – SoftVC VITS Voice Conversion

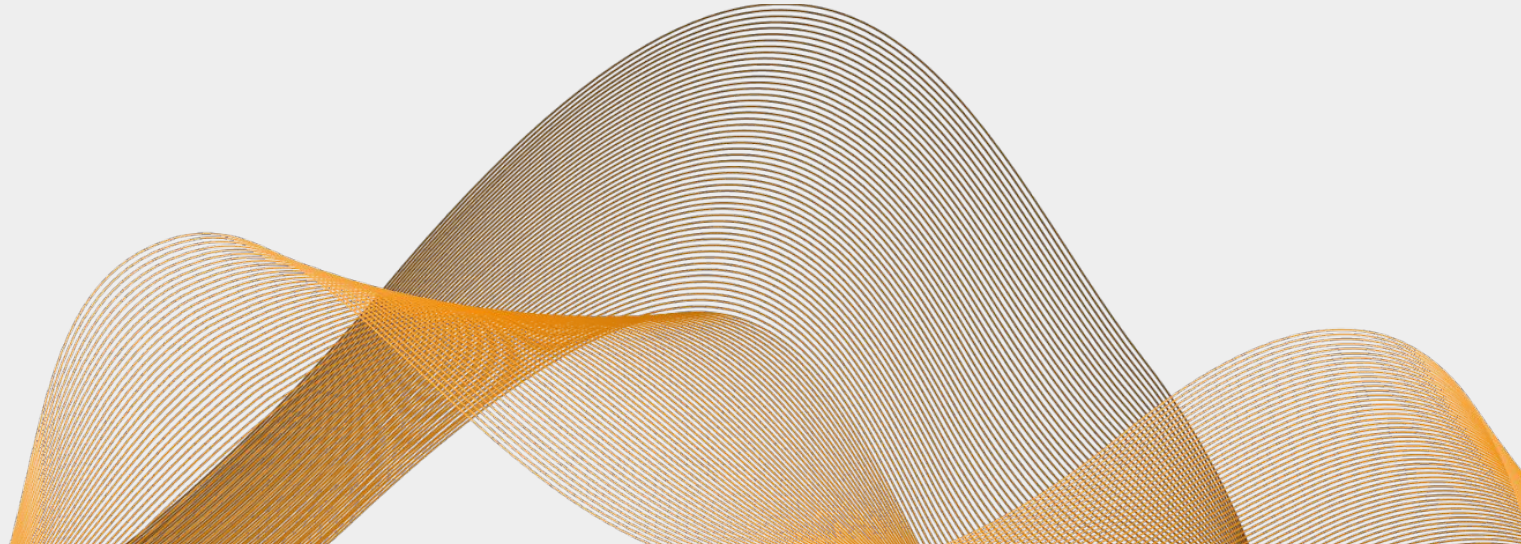


Results

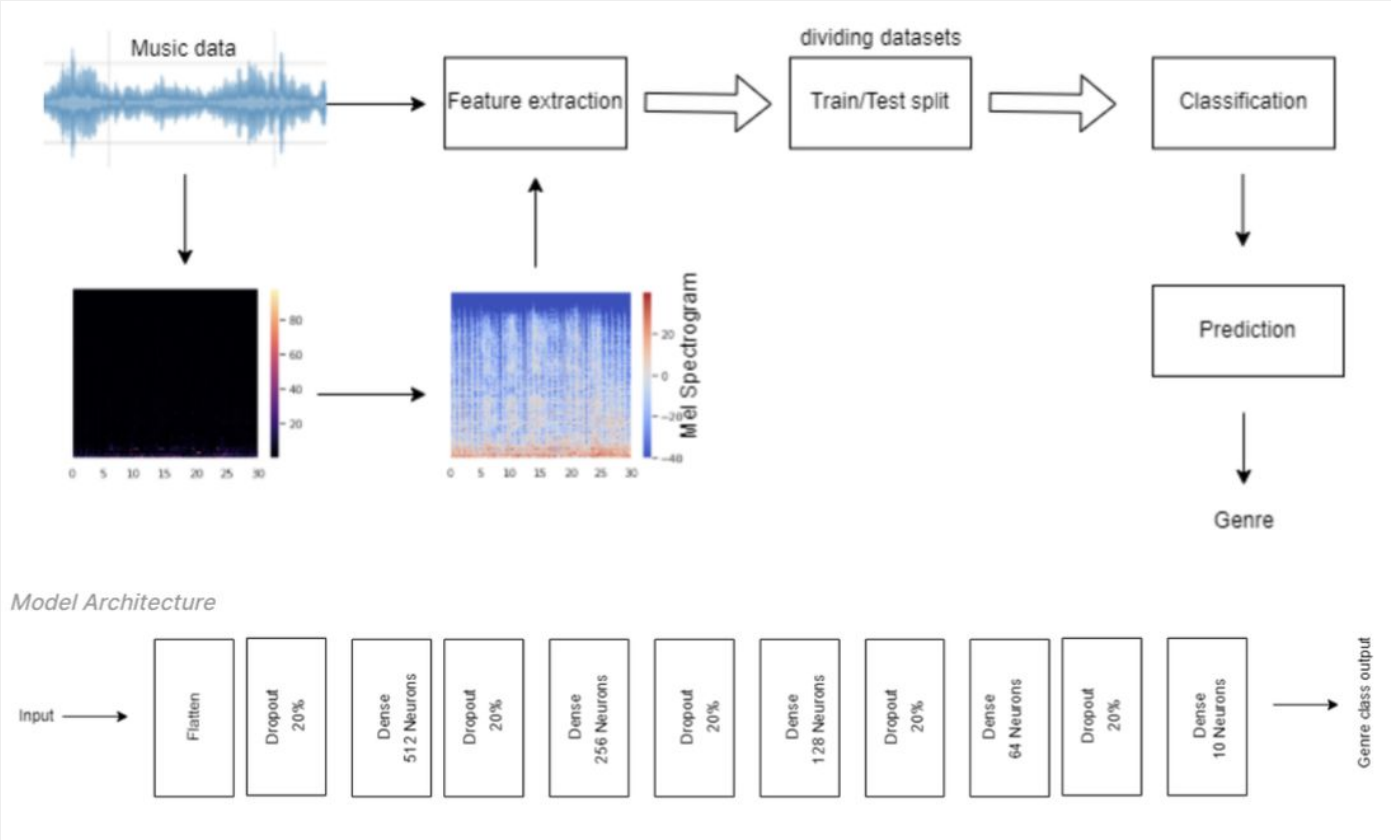
https://github.com/cwallant/AlgoRemix/tree/main/Output_Examples

Evaluation

1. Genre Classification Evaluation
2. Subjective Listening Evaluation



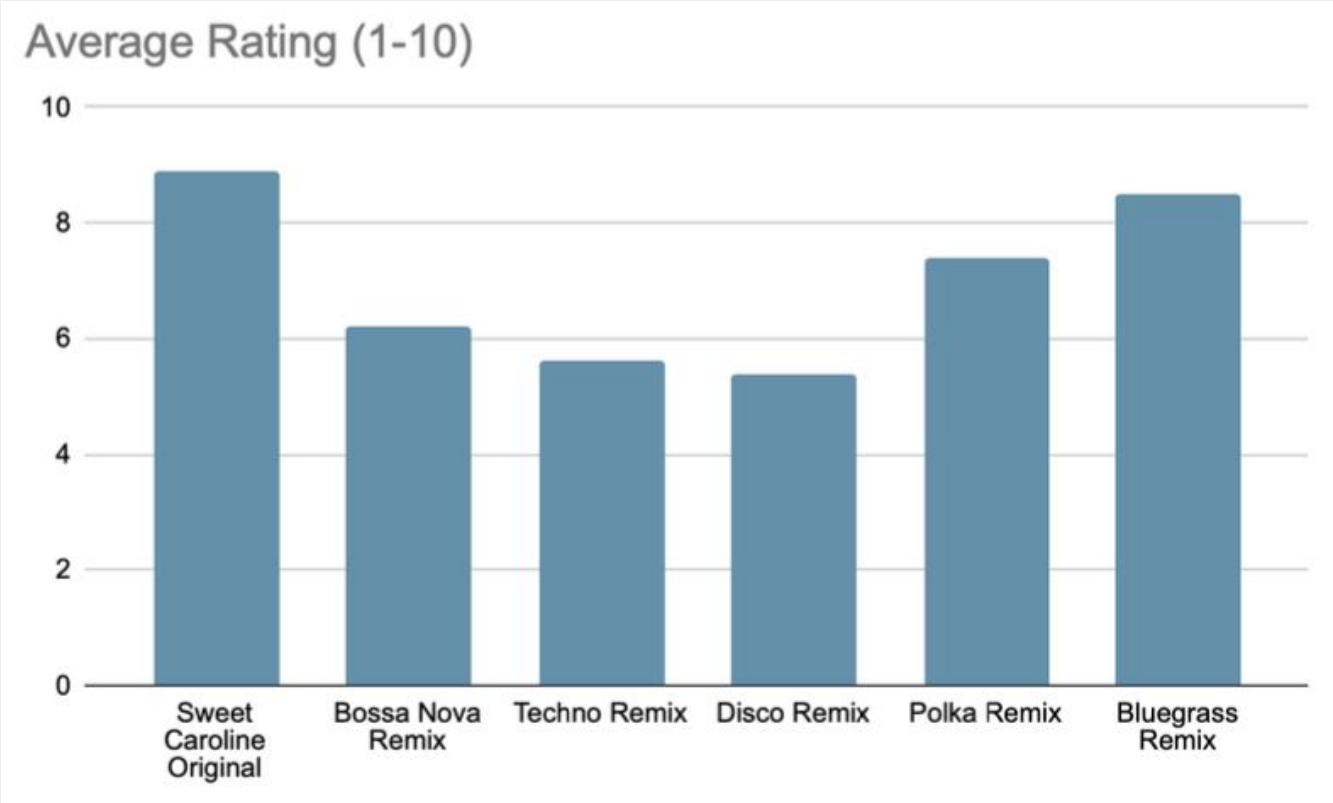
Evaluation - Genre Classification



Evaluation - Genre Classification

Song Type	Original Song	Bossa Nova Remix	Techno Remix	Disco Remix	Polka Remix	Bluegrass Remix
Estimated Genre	Pop	Bossa Nova	Techno	Disco	Polka	Polka

Evaluation - Subjective Listening



Conclusion

- Trade off between long term structure and repetitiveness
- Vocal limitations
- Altering music as the future of music