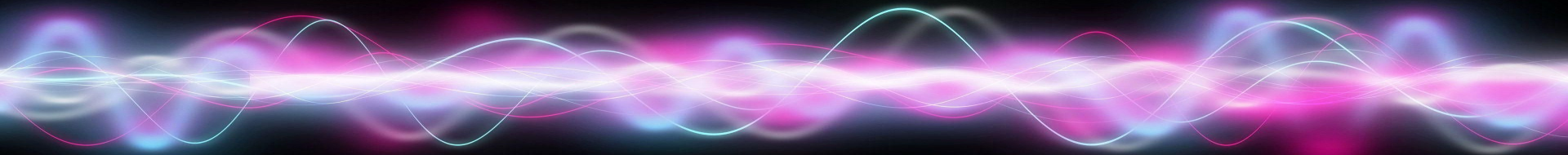
Genre Classification with a Neural Network

By Blake Franey



Overview

- Intro
- Overview of Audio Content Analysis
  - Key Terms
  - Fundamentals
  - Classification
- My Model
- Results and Conclusions


SUPERB
WALLPAPERS
Audio Content Analysis

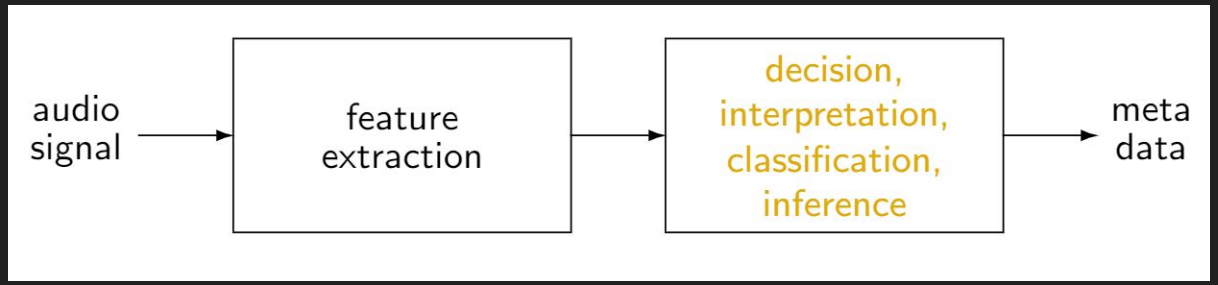
Key Terms

- Music Information Retrieval (MIR)
- Machine Listening
  - Recognition and understanding of music
- Computational Auditory Scene Analysis (CASA)
  - Human perception and cognition

Key Terms

- Timbral
  - Sound Quality
- Tonal
  - Pitch
- Temporal
  - Rhythm and Tempo
- Spectral
  - A color spectrum that is ordered by a characteristic like frequency

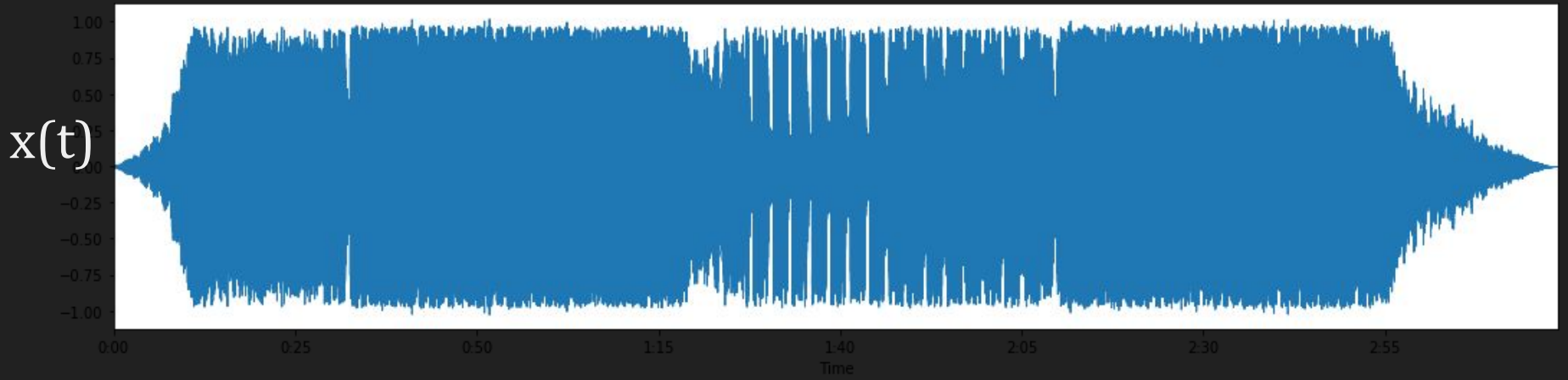
Fundamentals



From 1

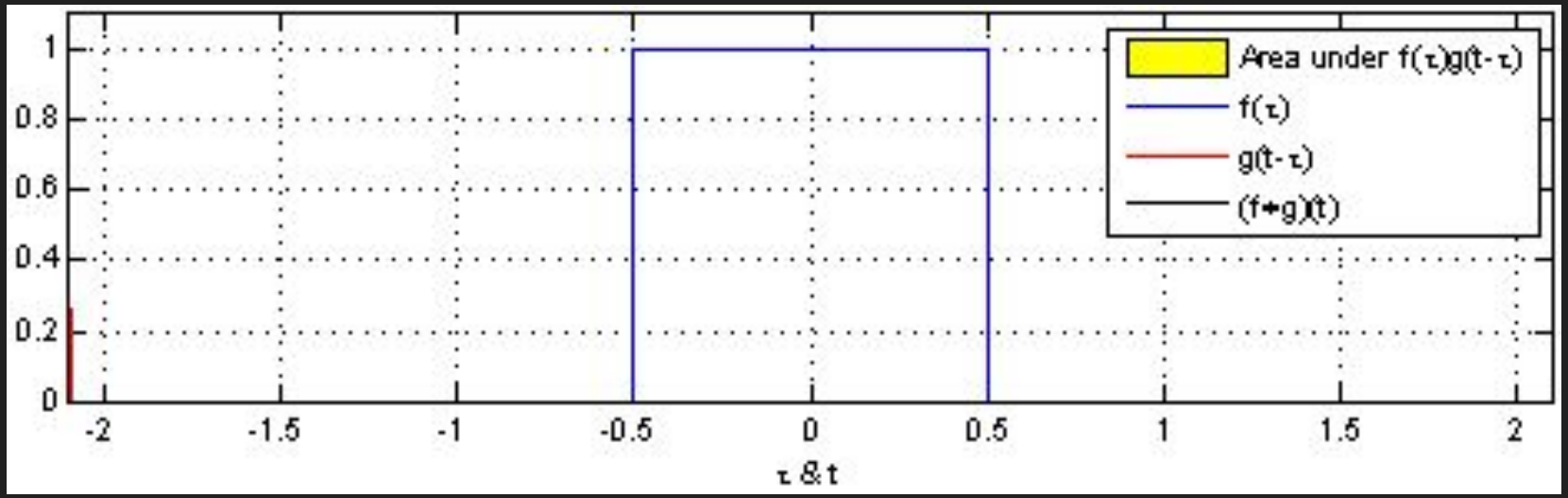
- The schematic of audio analysis.
- Features are completely dependent to the end goal.
- Can be “custom” features that are made for your specific purpose.

Fundamentals: Audio Time Series



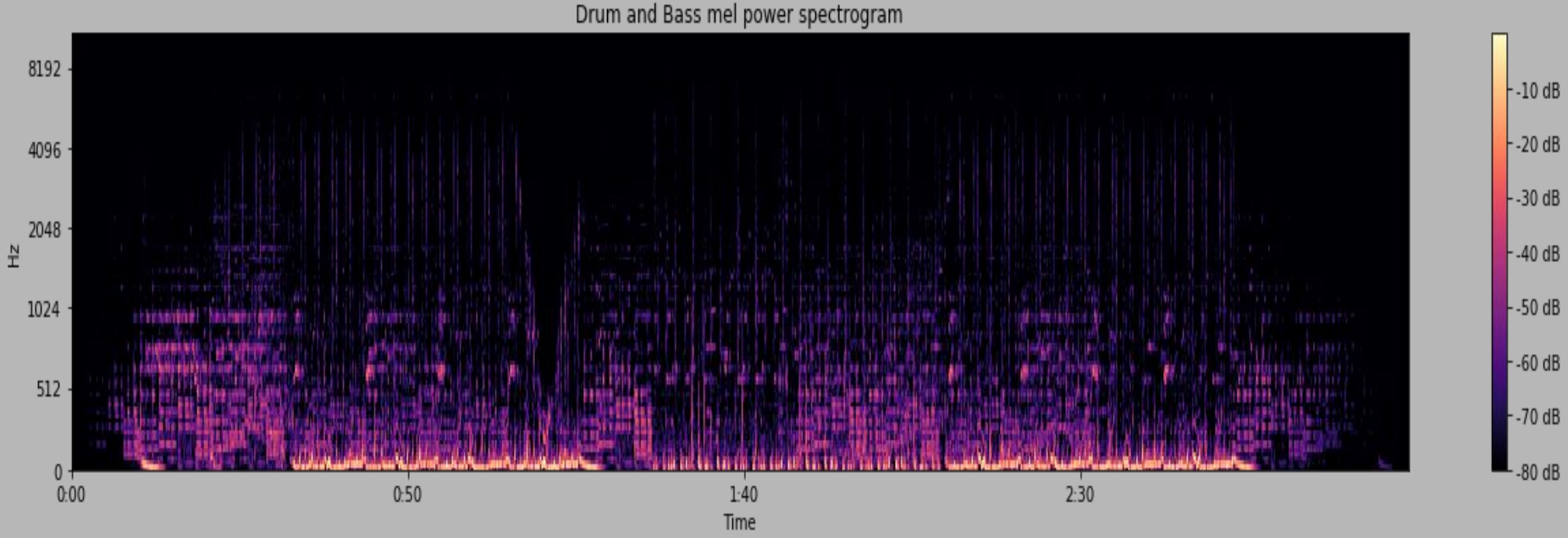
Time [s]

Fundamentals: Convolution

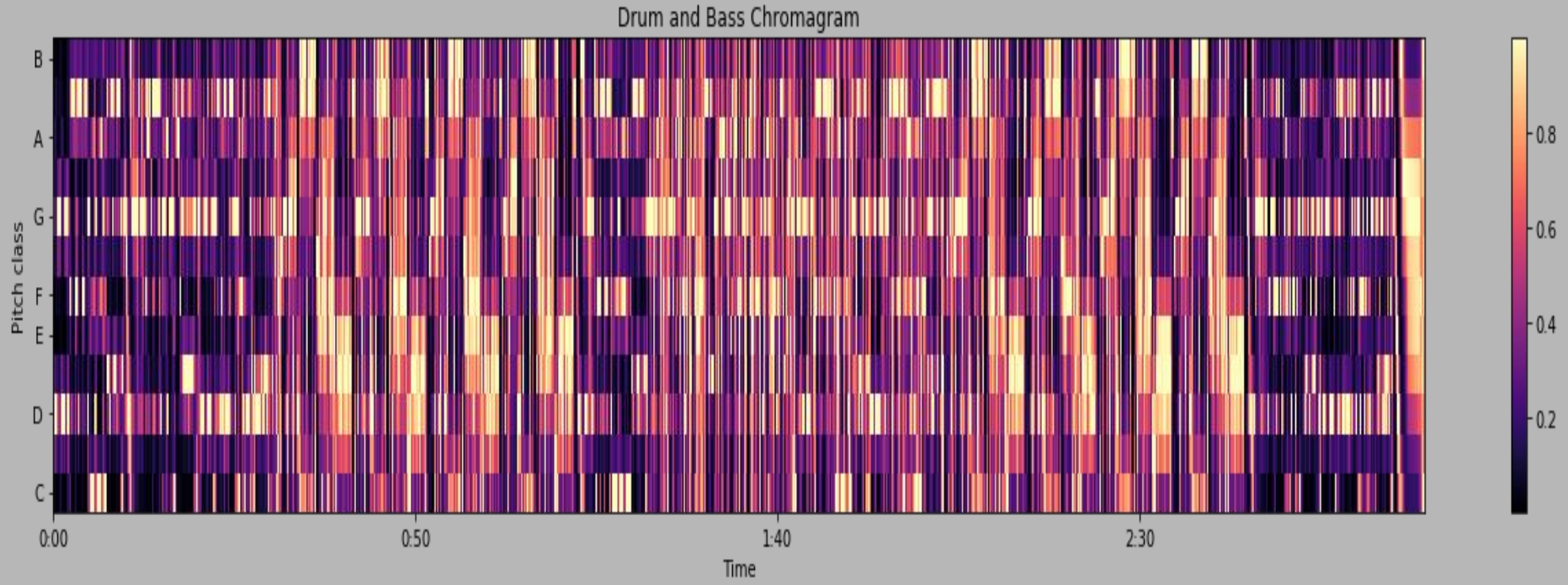


From 2

Fundamentals: Mel-frequency Cepstral Coefficients

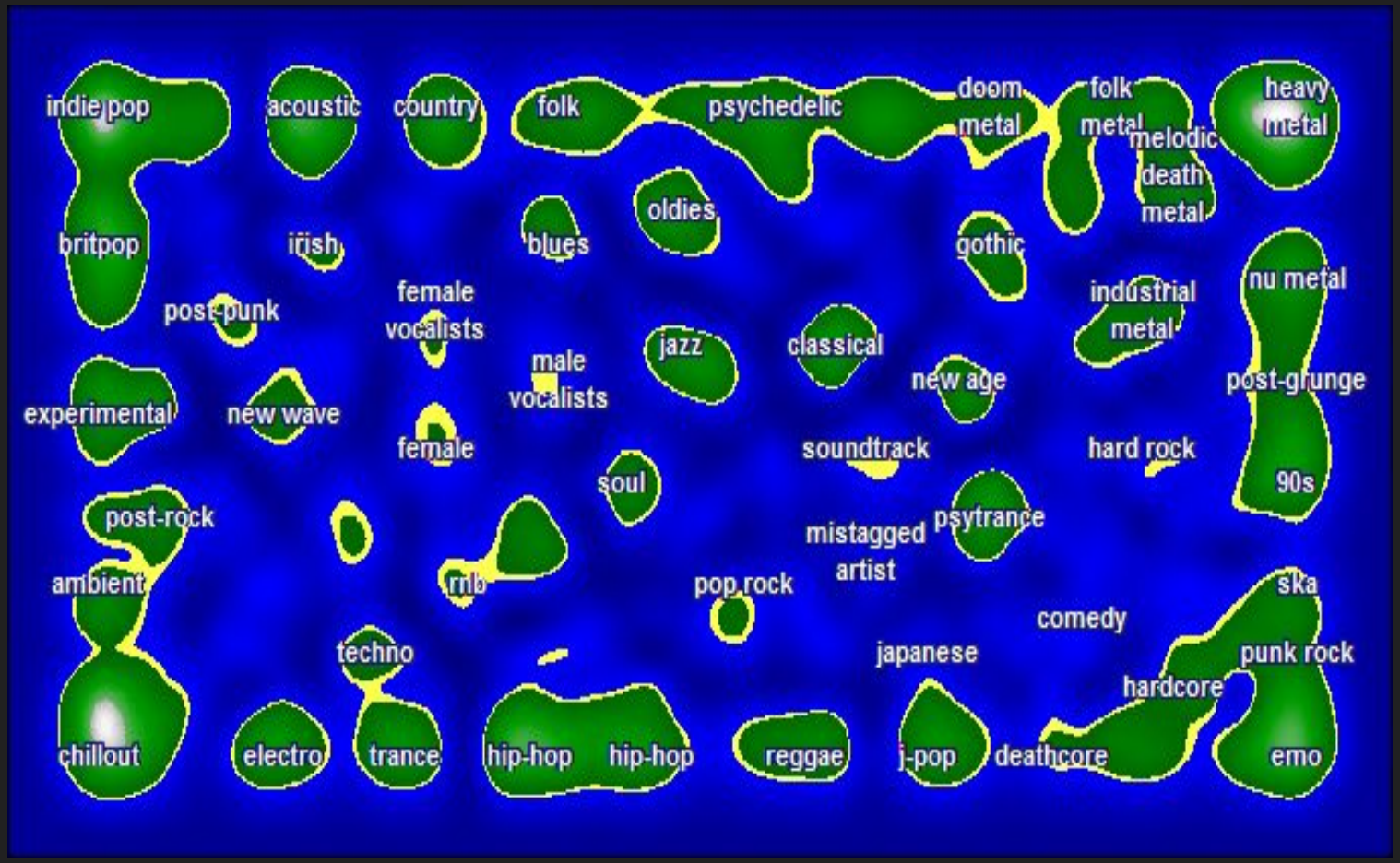


Fundamentals: Pitch Chroma



Classification!

- Machine Learning!
  - k-Nearest Neighbors
  - Gaussian Mixture Models
  - Neural Networks
  - Self-organizing Map
  - k-Means Clustering
- Musical Similarity
- Genre Classification
- Mood
- Audio Fingerprinting

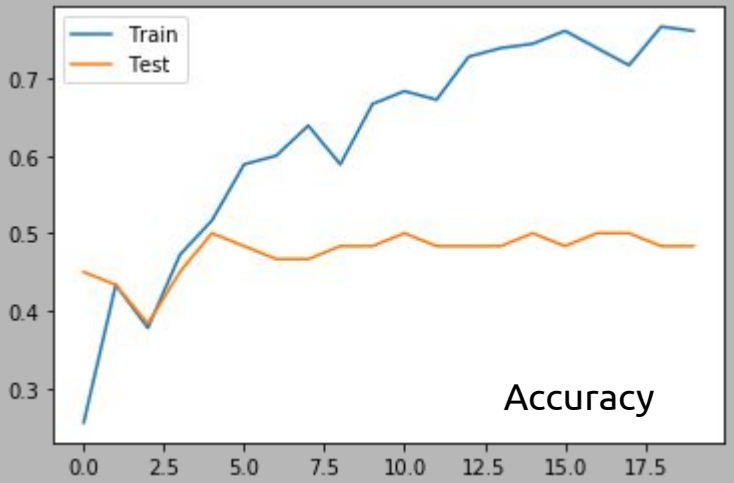


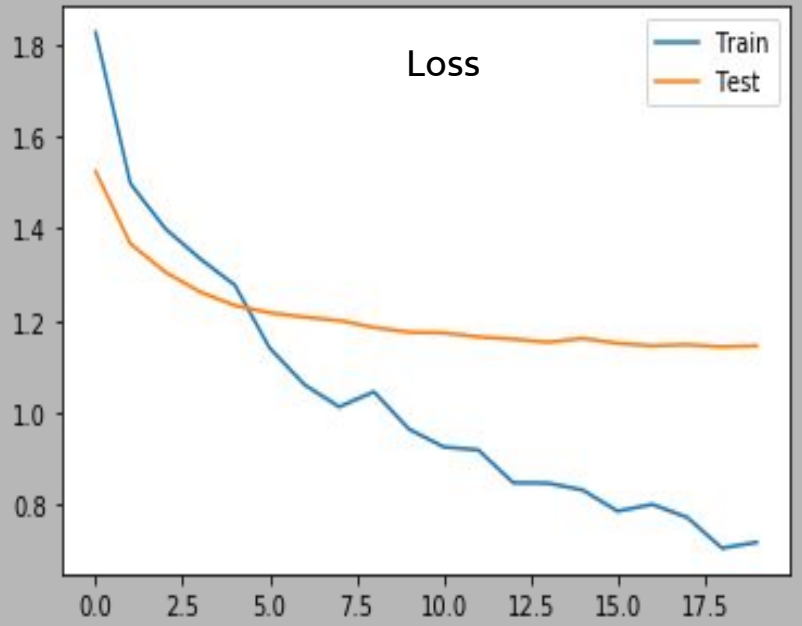
From Pampalk 2001


My Model
mirror_mod.use_y = True
mirror_mod.use_z = False
elif operation == "MIRROR_Z":
mirror_mod.use_x = False
mirror_mod.use_y = False
mirror_mod.use_z = True
selection at the end -add back the deselected mirror modifier object
mirror_ob.select= 1
modifier_ob.select=1
bpy.context.scene.objects.active = modifier_ob
print("Selected" + str(modifier_ob)) # modifier ob is the active ob
From 3

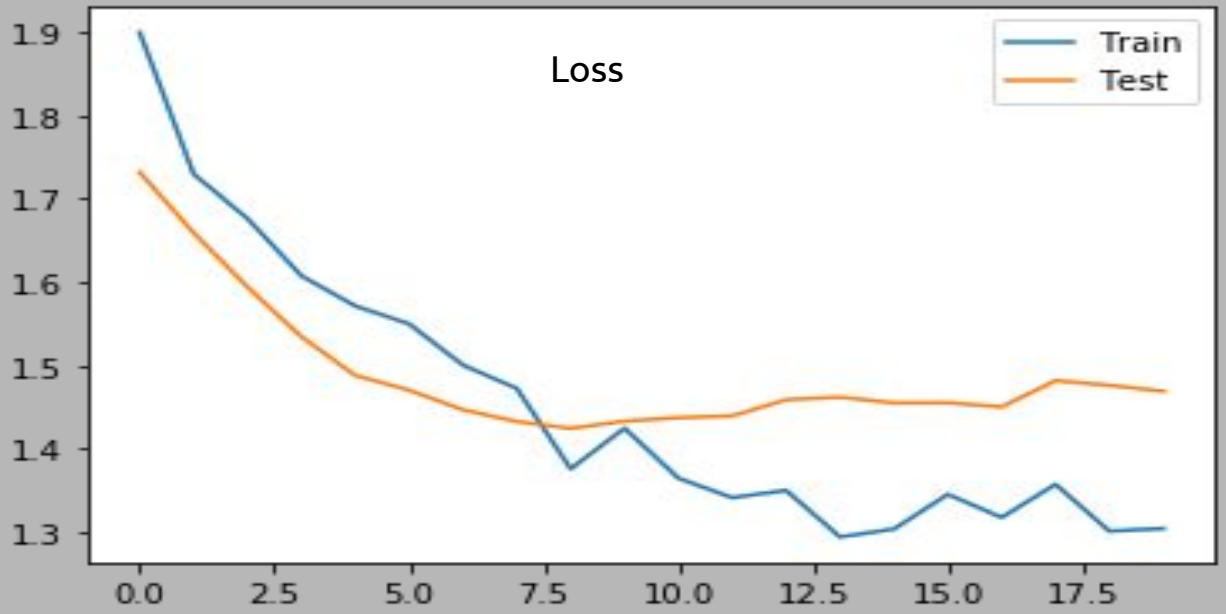
The Data

- 1376 final count of songs (~13 GB)
- Trained on 240 songs using only the MFCCs and chromas separately
- Baseline Accuracy ~17%
- Improving accuracy is the goal

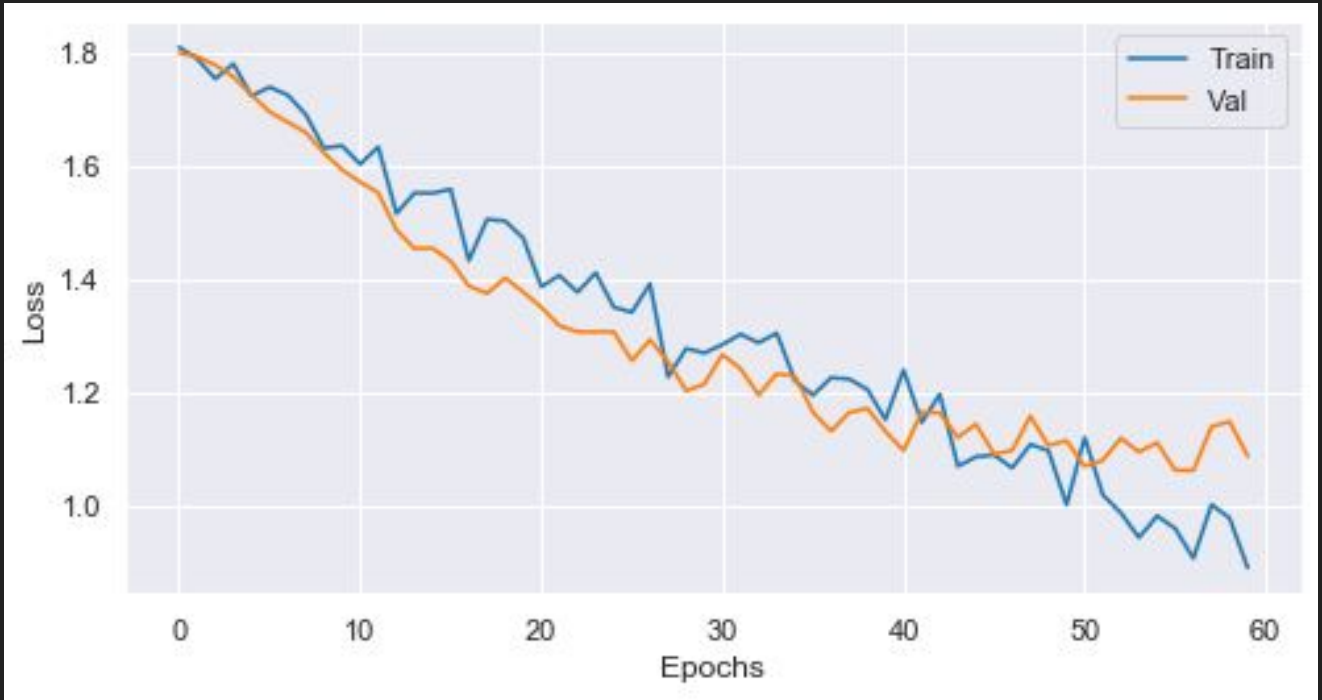




Pitch Chroma Model

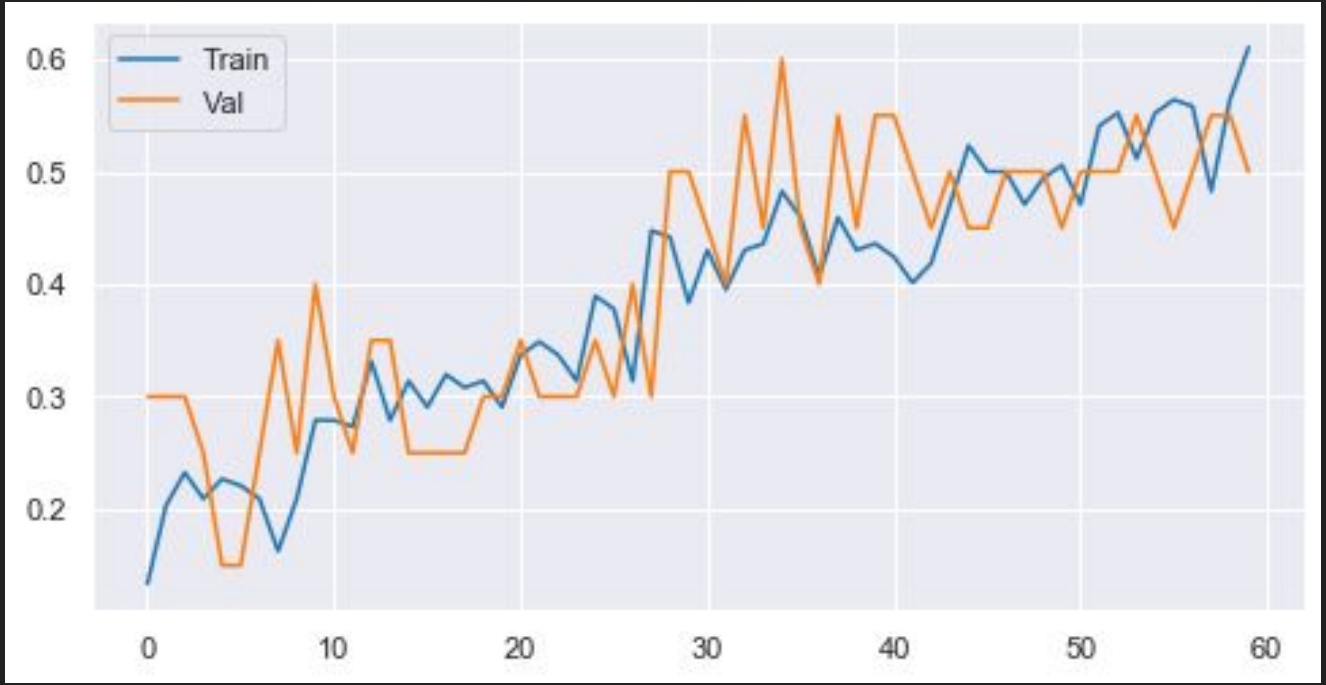


First Model

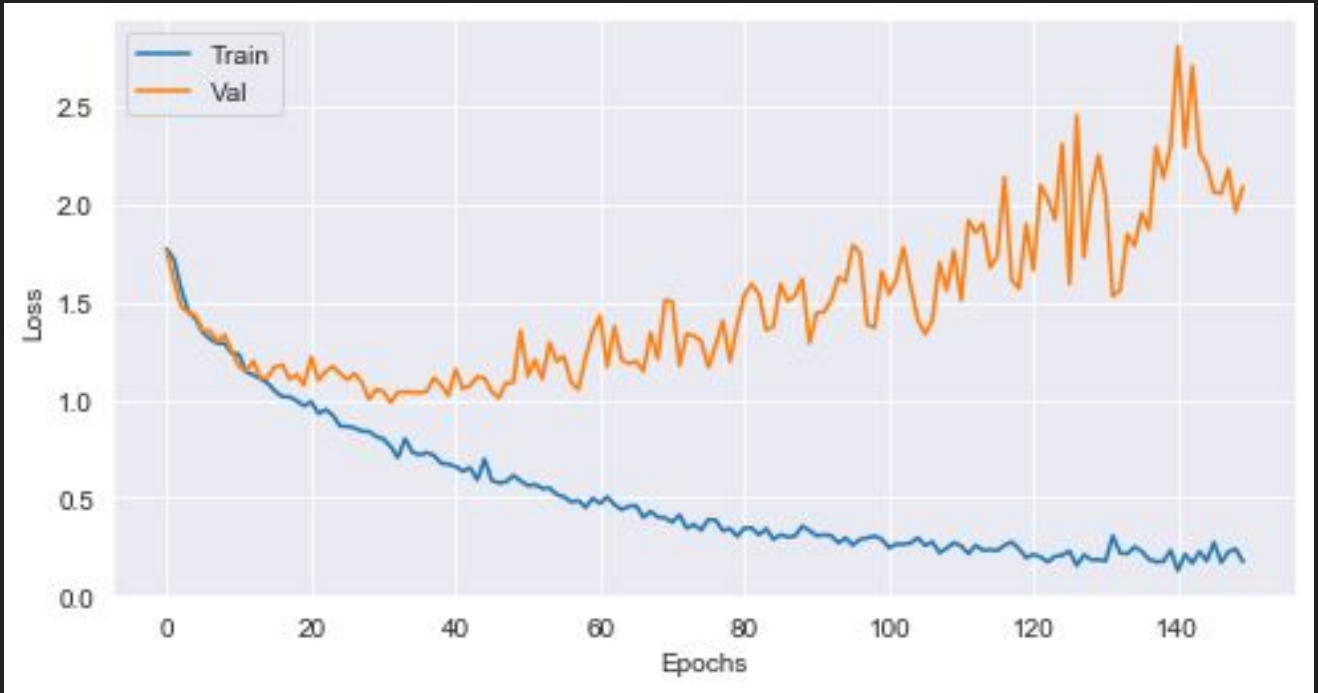


First Model

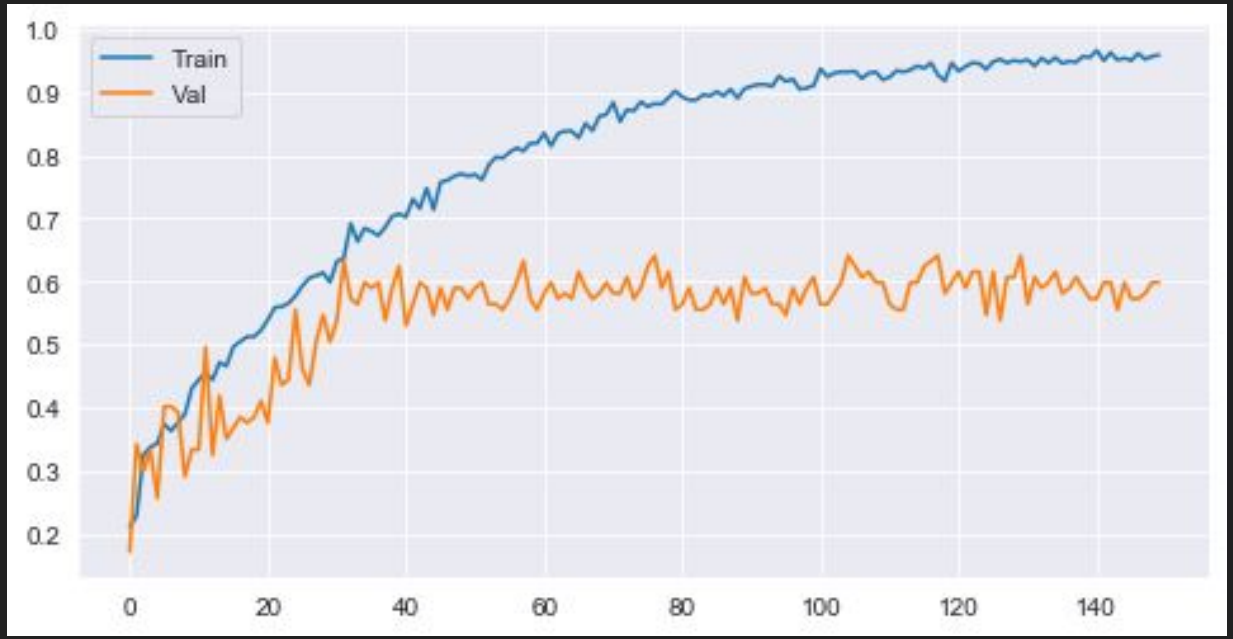
Accuracy



Final Model



Final Model



Results and Conclusions

Audio Content Analysis is hard. So are Artificial Neural Networks.

- My initial goal was to build a classifier that could be further broken down into subgenres and eventually a recommender.
- There are a lot of aspects to consider and multiple approaches, both 1- and 2-D features, are necessary to properly classify genres.
- There was a lot of subject and arbitrary choices that introduced bias.

Thank you!

References/Citations:

1. Audio Content Analysis: audiocontentanalysis.org
2. Wikipedia: https://en.wikipedia.org/wiki/Convolution
3. Security Industry: https://www.securityindustry.org/2019/02/25/machine-learning-artificial-neural-networks-deep-learning-in-the-security-industry/
4. Sound Wave Wallpaper: https://wallpapercave.com/sound-wave-wallpaper