Data Incubator Presentation

Fernando Araiza Gonzalez
GitHub: Faraigonz

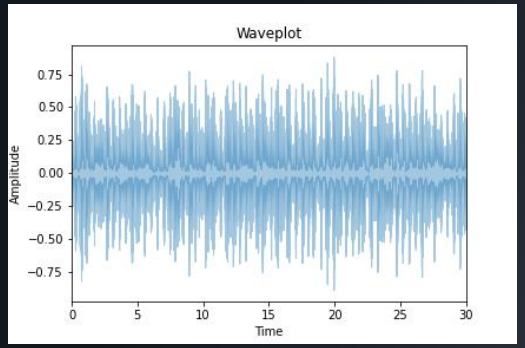
October 29, 2020

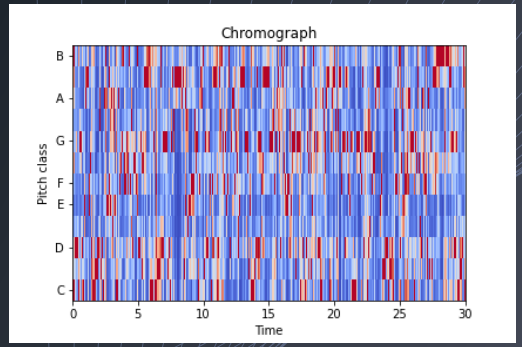
Music Classification using Neural Networks

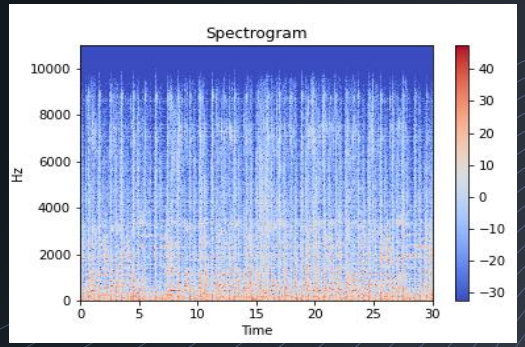
- Motivation
  - Build a neural network trained on music classified as ‘good’ and ‘bad’, to predict whether I would enjoy unclassified music – a neural network to predict musical taste.
- Application
  - User recommendation system
  - Copyright violation detection
  - Speech recognition
  - Sound event detection

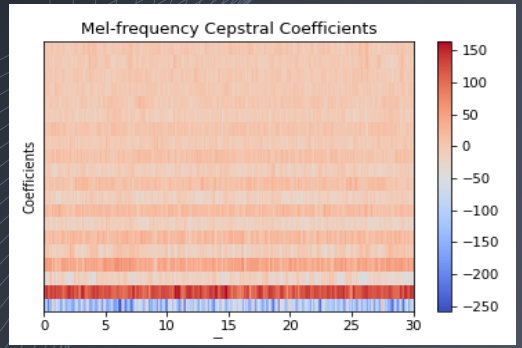
The Old Way

- Create a time averaged, 1D feature vector to pass to NN
  - Amplitude waveplots
  - Chromograph
  - Spectrogram
  - MFCCs
- Misses entire temporal component of music – chord, melody progressions



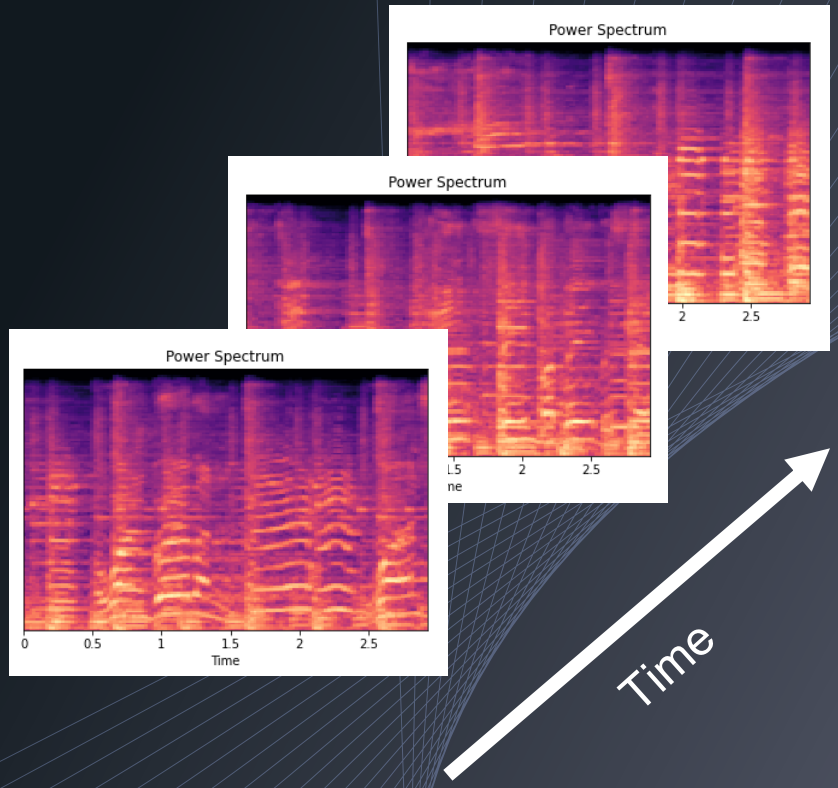






A New Way

- Use full time-dependant power spectrogram as the feature for a Convolutional Recurrent Neural Network (CRNN)
- CRNN
  - Convolutional layers learn global features
  - Recurrent layers learn temporal features

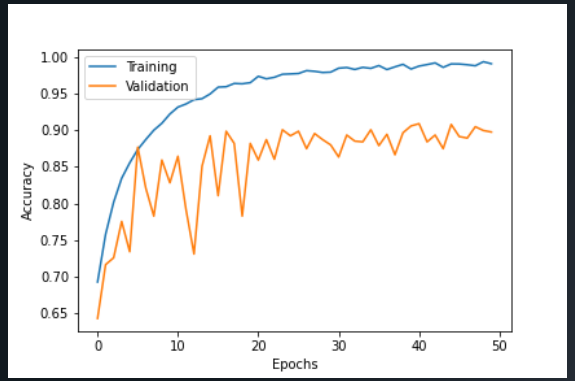


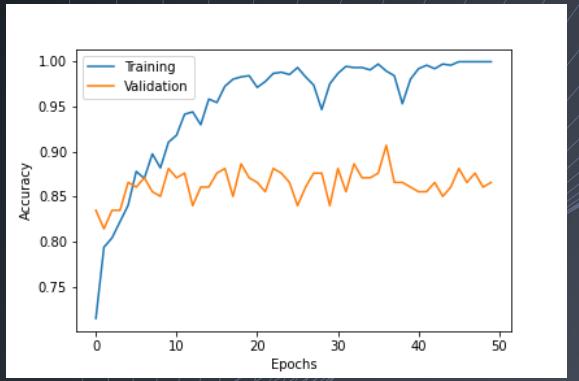
Resource: **“Music Artist Classification with Convolutional Recurrent Neural Networks”**
Zain Nasrullah, Yue Zhao.

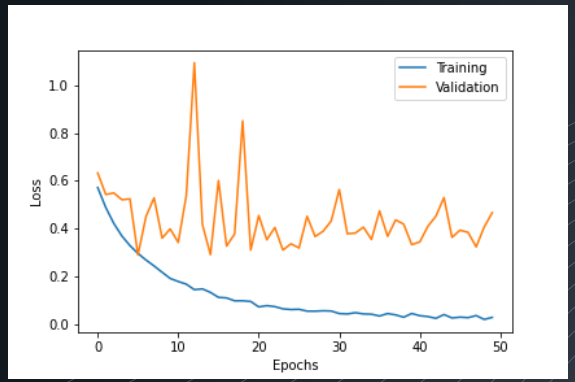
Data and Results

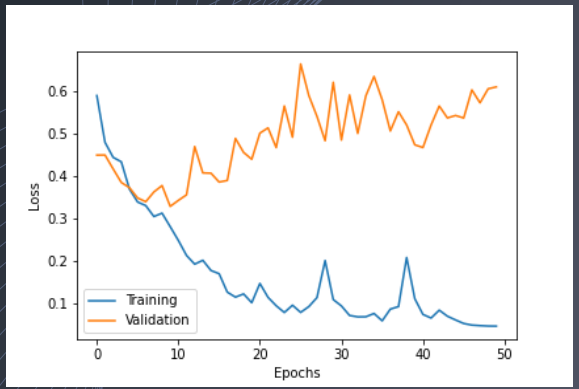
- GTZAN Dataset
  - 1.3 GB of data
  - 10,000 songs
  - 10 different genres, 100 30 second song files

Author	Method	Accuracy
Nasrullah, Zhao	CRNN	93.7%
Mandel	MFCC	83.9%
Me	CRNN	89.8%
Me	MFCC	86.6%









CRNN Approach

MFCC Approach

Conclusion and Next Steps

- The CRNN approach significantly improves loss over the MFCC approach, but only modestly improves the accuracy
- Fitting the CRNN model is quite costly – 50 epochs took $\sim 7\ hours$
- Fitting the MFCC model is very cheap – 50 epochs to $\sim 10\ seconds$
- A combination of both methods may produce good results with less computation
  - Pre-train the CRNN model using 1D feature vector to reduce processing time

Links and References

- Papers
  - [1] “Music Artist Classification with Convolutional Recurrent Neural Networks”
  - [2] “The GTZAN dataset: Its contents, its faults, their effects on evaluation, and its future use”
  - [3] “Song-Level Features and Support Vector Machines for Music Classification”
- Data
  - [4] GTZAN
- GitHub
  - [5] CRNN Model
  - [6] MFCC Model

Thank You