Music Stream Splitting
Vikul Gupta, Nidhi Manoj, and Sandhini Agarwal
CS 230, Stanford University 2018

MedleyDB & Problem Definition
- MedleyDB is a collection of 108 songs with melody annotations. 50 songs have isolated drums streams which are used to train our neural network model.
- Splitting a mixed audio (wav) music file of instruments and vocal in order to isolate the vocal stream.

Challenges
- Splitting up music into a variable number of instruments/sources is an extremely difficult problem and not many research papers have had success.
- Simply feeding in time steps into an LSTM causes vanishing gradient issues.
- Need to compute a measure of error: we compare the spectrogram of the predicted drums and true drums audio files.

Approach: Fully Connected Layers & LSTM
- Found the spectrogram for each song file for 2560 timesteps and got input dimensions of 399x288x513.
- Encoded by feeding these into two fully connected layers in 43 chunks.
- Passed layers output into a bi-directional LSTM of 60 units.
- Decoded the output of LSTM by passing it through two fully connected layers.

Approach: Baseline (Classifier)
- Built a logistic regression model that would classify songs as piano vs. not piano.
- Achieved 80% accuracy.

Approach: ICA
- Built an ICA model that would separate the song into two channels.
- The first channel output corresponded to the instrument channel and the second one to the vocals channel.
- Compared difference between the channels of the input song and the output song.

Evaluation Metric
- Summed the difference between drum spectrum (b) and our output (a).

Results and Analysis
- Our accuracy went from 0.0081 in the first epoch, to 0.9727 in the tenth epoch.
- Our error rate reduced from ~2390 to ~1632 (after scaling both by 10^15).
- We think that one reason the model may be giving us a lower accuracy may be due high similarity between the drum channel and the mixed song.
- We aim to test this model with vocals to see if we get better results.

References
- Michael Galanin: Logistic regression using python.