Distinguishing sounds

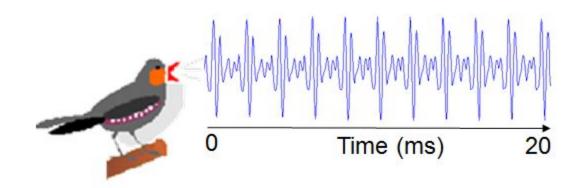## with deep learning

Whoami

Data Scientist @ Kelvin Inc.

Background in software engineering/architecture
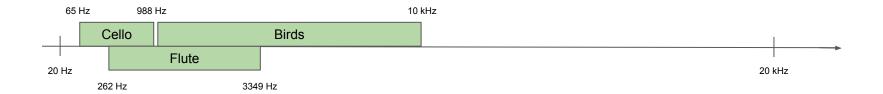
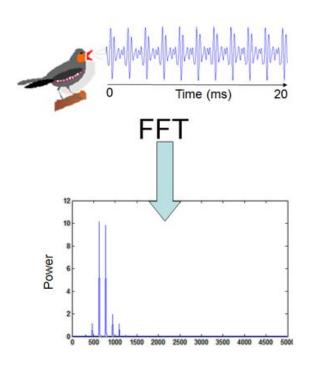Builder of real-time dataflow systems
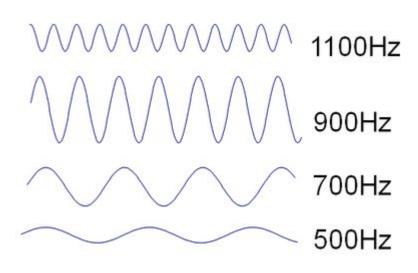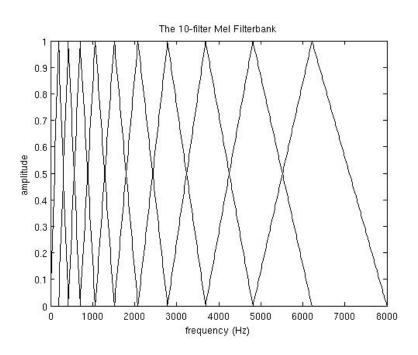
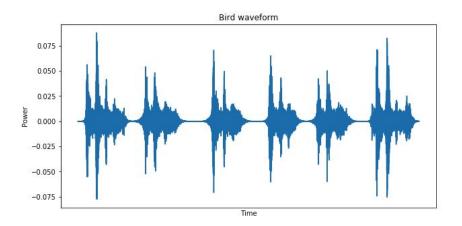Runner

What is sound

What is sound

Fourier transform

Mel scale

The 10-filter Mel Filterbank

Bird singing

Bird waveform
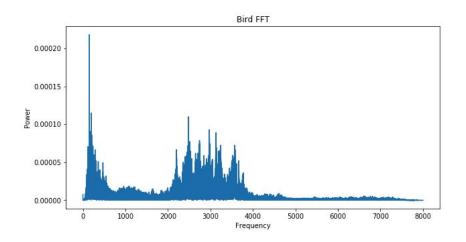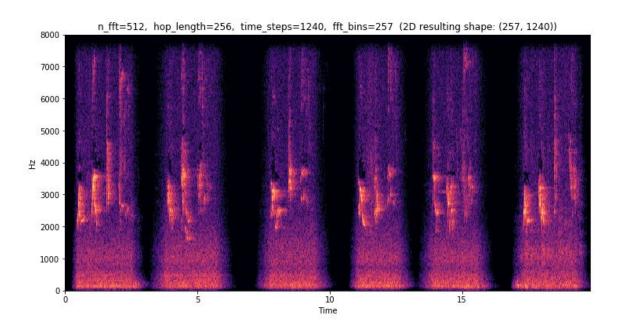
Bird singing

Bird singing

n_fft=512, hop_length=256, time_steps=1240, fft_bins=257 (2D resulting shape: (257, 1240))

The model

Input   Conv + Maxpool   Conv + Maxpool   Conv + Maxpool   Conv + Maxpool   FC   FC   Output

Thank you
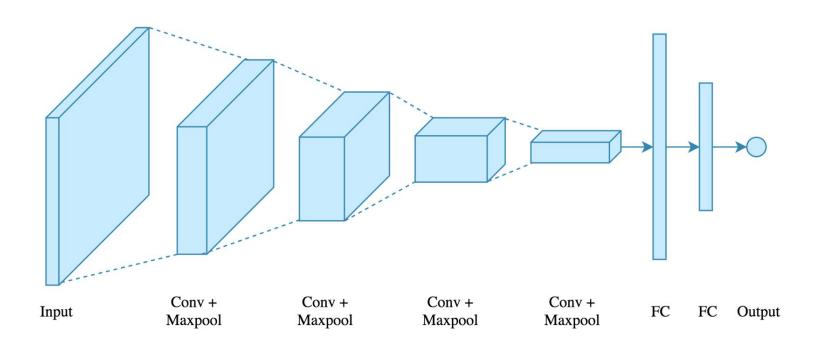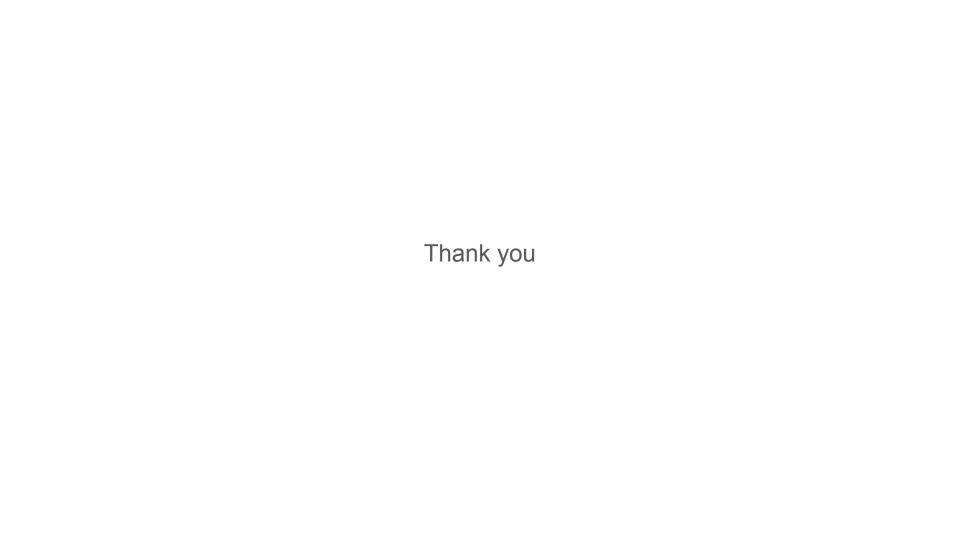
Resources

- https://www.kaggle.com/fizzbuzz/beginner-s-guide-to-audio-data
- http://practicalcryptography.com/miscellaneous/machine-learning/guide-mel-frequency-cepstral-coefficients-mfccs/
-