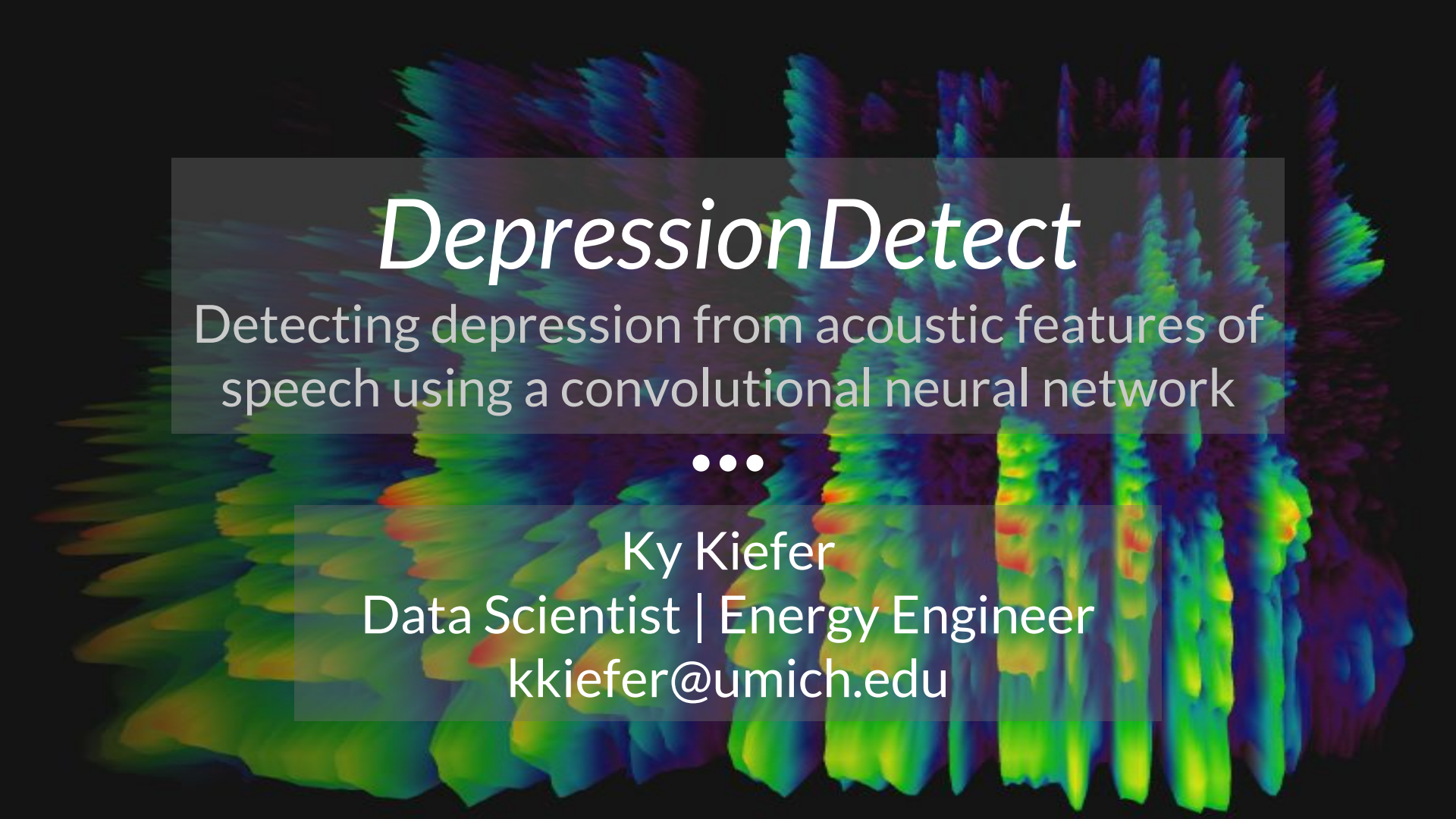


DepressionDetect

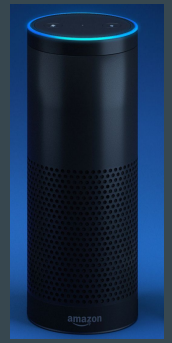
Detecting depression from acoustic features of speech using a convolutional neural network

•••

Ky Kiefer
Data Scientist | Energy Engineer
kkiefer@umich.edu

Motivation

- 50% do not seek treatment for major depression.



Solution

- Wearable or home devices providing and early and objective detection.

DepressionDetect

Discontent | Awareness | Apathy

Can you detect depression from speech?

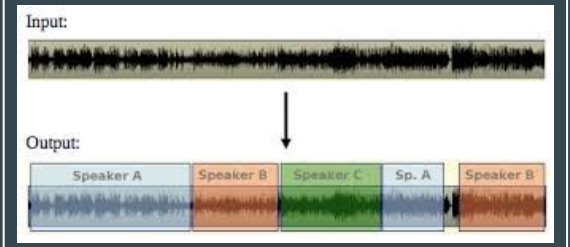
Not semantics, acoustics!

Data

- 50 hours of virtual interviews from USC.
- 189 participants.
- Depression classification (0/1) from psychiatric survey.

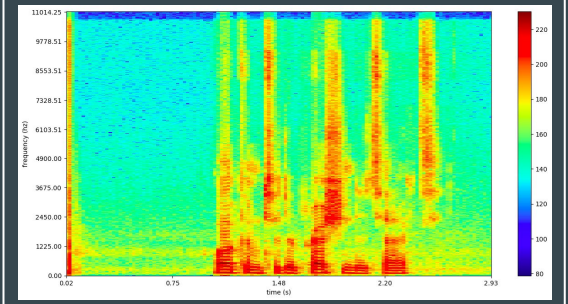
Workflow

1 | Speaker Diarization



Support Vector Machine

2 | Feature Extraction

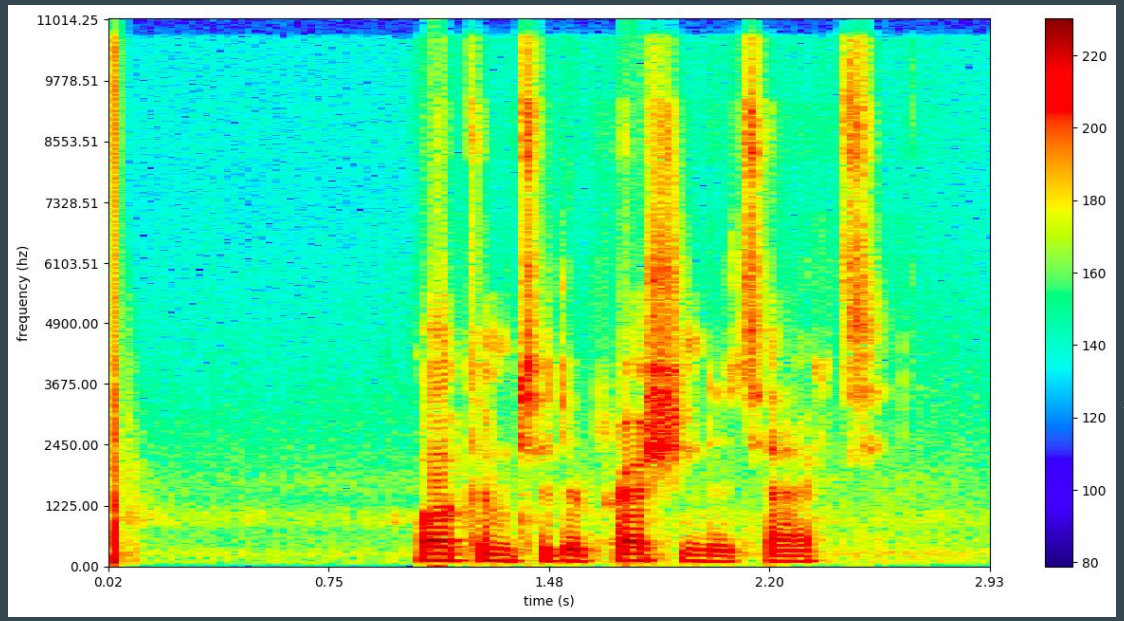


Spectrogram

3 | Classification

Convolutional Neural Network

Seeing sound. Spectrograms.



Frequency (Hz)

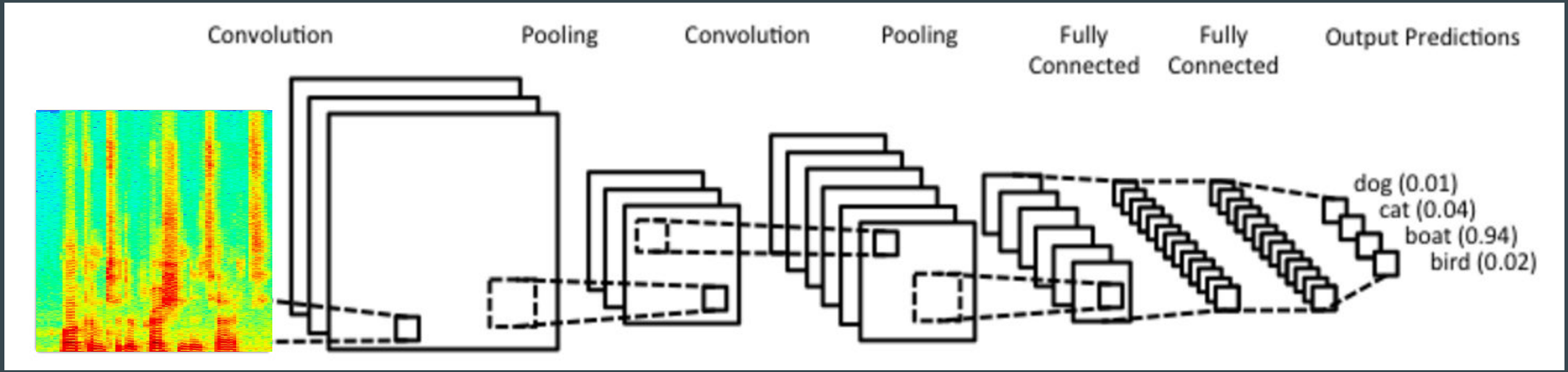
Time (s)

Decibels (dB)

Short-time Fourier transform

Live visual: musiclab.chromeexperiments.com/Spectrogram

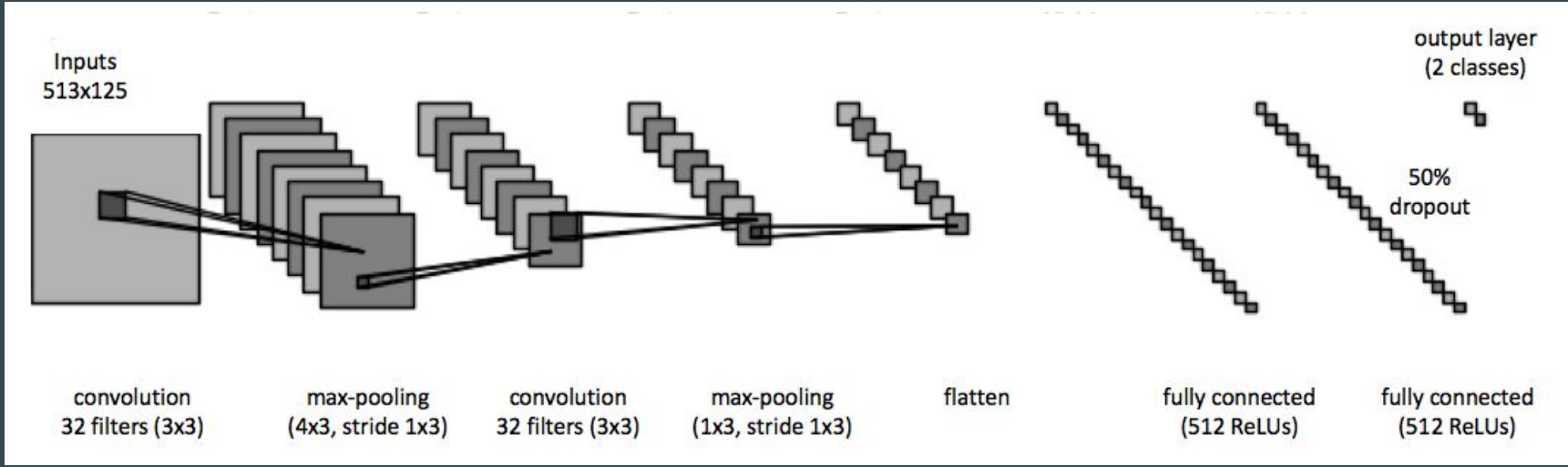
Convolutional Neural Networks



Resources

- Video: https://www.youtube.com/watch?v=FmpDIaiMIeA
- Stanford's CS231n: http://cs231n.github.io/convolutional-networks/

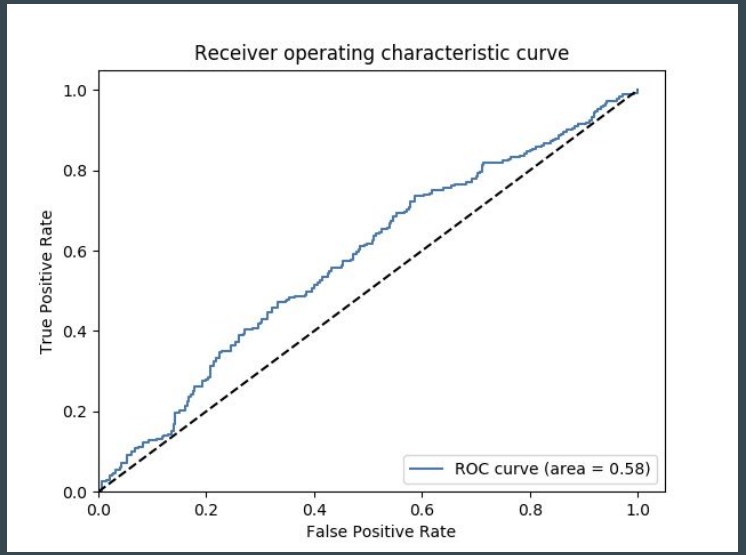
DepressionDetect Architecture



- 513 frequency bins (0 to 8kHz)
- 125 time bins (totaling 4s of audio)

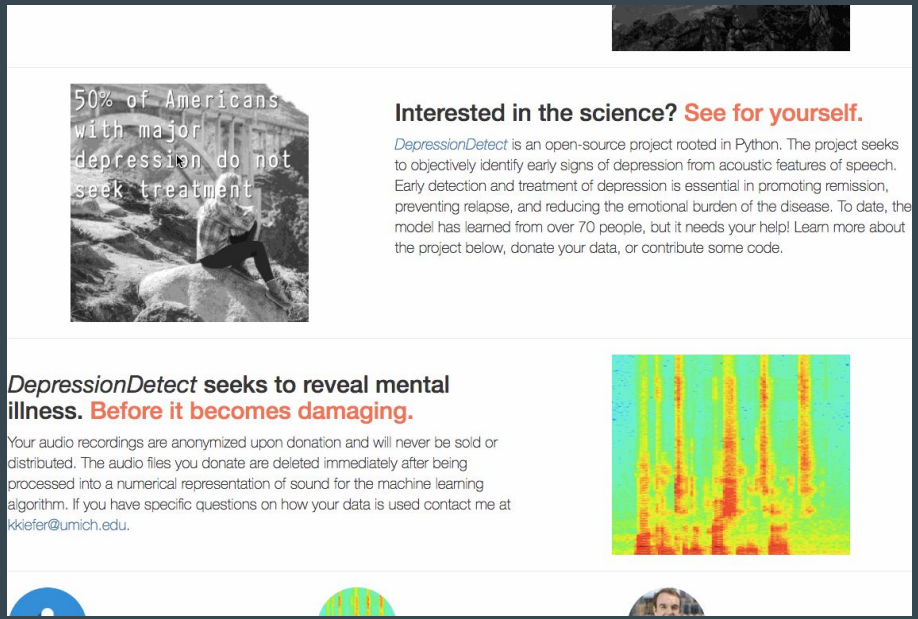
Results

- 55% accuracy predicting on 4-seconds of audio.
- 64% accuracy when taking majority vote across 40 4-second clips (2m 40s).



Become a data donor!

- Speech data is inherently noisy.
- Robust models rely on 1,000s hours of data.



DataStopsDepression.com

Questions?

Visit DataStopsDepression.com to become a data donor!

Ky Kiefer
kkiefer@umich.edu
linkedin.com/in/kykiefer
github.com/kykiefer

Tech Stack

pyAudio Analysis
python
Flask
SciPy
NumPy
matplotlib
IPython Interactive Computing
pandas
scikit learn
theano
Keras
amazon web services