Class: CS 760 - Machine Learning Fall 2018
Prof: Dr. Yingyu Liang
Date: Monday - October 22nd, 2018
RE: Final Project Proposal

Team Name:

A Irate ML Coke Ninja

Team Members:

Name	Wisc ID
Ankit Vij	vij2
Amanpreet Singh Saini	saini5
Eric Brandt	ebrandt
Joel Haynie	jhaynie

Topic Summary:

Our project will focus on the current state of training machine learning systems using audio clips as the training data. Currently, there seems to be differing opinions within the ML community concerning the suitability of convolutional neural networks for audio-based training data. For example, Hershey et al (Jan 2017) states, “Convolutional Neural Networks (CNNs) have proven very effective in image classification and show promise for audio,” while Rothmann (Mar 2018) offers a dissenting opinion: “Essentially, these techniques [CNN] apply machine vision in order to do machine hearing. I believe this poses an essential problem which could be hindering the progress of AI-assisted technologies in audio processing.”

To investigate this further, our project will be composed of three separate but related components:

1. We will survey existing literature on how audio clips are used in training machine learning systems. We will summarize the important milestones and methods in our final report.
2. We will discuss and provide insights into the differences of opinion between Hershey’s and Rothmann’s opinions quoted above.
3. We will implement a small scale ML system that learns to classify music clips into one of a fixed set of nominal labels containing the mood of the performer. We will prepare a training set and evaluate the performance of our implementation using stratified n-fold cross validation. Based on the results of this classification, we may explore novel uses of the classification such as suggesting relevant advertisements to listeners of songs classified with particular moods.

Background:

Starting point for Background reading:

1. CNN Architectures for Large-Scale Audio Classification; Shawn Hershey, et al.
2. What's wrong with CNNs and spectrograms for audio processing?; Daniel Rothmann
3. Inside the spectrogram: Convolutional Neural Networks in audio processing; Monika Dörfler, et al.
4. How do deep convolutional neural networks learn from raw audio waveforms?; Yuan Gong, et al.
5. Getting Started with Audio Data Analysis using Deep Learning; Faizan Shaikh
6. Hearing AI: Getting Started with Deep Learning for Audio on Azure; Xiaoyong Zhu, et al.

Project Timeline:

Project Proposal Due Date: Monday - October 22nd, 2018

- Begin literature review - October 23rd, 2018
- Data collection and cleaning of data - November 9th, 2018
- Collect and label Ads dataset (can be done in parallel) - November 14th, 2018
- Find audio features for the data - November 12th, 2018
- Do feature selection - November 16rd, 2018
- Train the model - November 21th, 2018
- Test and validate mood - November 25th, 2018
- Complete Literature Review - December 1st, 2018
- Finalize results and report - December 11th , 2018

Project Due Date: Wednesday - December 12th, 2018

Tool Chains:

- Report authoring:
  - Latex environment
  - using the style files of NIPS'18
- Development & Execution Environment:
  - PyCharm v2018.1.2 (community edition)
  - Visual Studio 2015/2017
- Libraries / Packages:
  - python v3.6
  - scipy v1.1.0
  - ipython v6.3.1
  - matplotlib v2.2.2
  - pandas v0.22.0