

Project Plan

CS 410: Explorations in Data Science

Objective

Inquire and explore steps in the data scientist pipeline including: data wrangling, data cleaning, and predictive modeling with AI. We will use the [Speech Recognition](#) data set, found on Kaggle.com, in order to build an algorithm that recognizes simple speech commands.

Approach

Our approach is broken up into three phases: research, modeling, and evaluation. In the research phase we will gather and clean the data. We must also determine which type of machine learning models are best practice for this data set (regression, clustering, ...). Search for Kaggle Kernels relating to our topic to avoid mistakes others have made. Then, implement a handful of machine learning models to obtain an accuracy score. Finally, examine our results to determine the precision/accuracy of our models.

Team Structure

Team D is divided into two sub groups: data visualization and machine learning. We represent the machine learning subgroup and the research is divided as follows.

- Parker Moore: Machine Learning & Neural Networks
- Aaron Squier: Deep Learning & Convolutional Networks
- Merlin Carson: Data Cleaning & Pre-processing

Project Milestones

1. Research and understand given topics. Present Research Presentation
2. Gather and clean data
3. Apply machine learning model to dataset and record results
4. Apply deep learning model to dataset and record results
5. Investigate results
6. Present final project