Project Proposal

**Requirements**

* Find a problem worth solving, analyzing, or visualizing
* Use ML in the context of technologies learned.
* You must use: Scikit-Learn and/or another machine learning library.
* You must use at least two of the below:
  + Python Pandas HTML/CSS/Bootstrap Python Matplotlib JavaScript Plotly
  + JavaScript D3.js
  + JavaScript Leaflet, SQLite Database, MongoDB Database
  + Google Cloud SQL, Amazon AWS/Heroku, Tableau
* Prepare a 15-minute data deep-dive or infrastructure walkthrough that shows machine learning in the context of what we’ve already learned.

# Name of the project - KnowMe: Learn about yourself with Machine Learning

# Summary

* Graphology is a method of identifying, evaluating and understanding human personality traits through the strokes and patterns revealed by handwriting. Handwriting reveals the true personality including emotional outlay, fears, honesty, defenses and many others. Professional handwriting examiners called graphologists often identify the writer with a piece of handwriting. Accuracy of handwriting analysis depends on how skilled the Analyst is. Although human intervention in handwriting analysis has been effective, it is costly and prone to error. Hence the proposed methodology focuses on developing a system that can predict the personality traits with the aid of machine learning without human intervention. To make this happen, we consider seven handwriting features: (i) size of letters, (ii) slant of the writing, (iii) baseline, (iv) pen pressure, (v) spacing between letters, (vi) spacing between words and (vii) top margin in a document to predict eight personality traits of a writer. After extracting all these features from the image containing the handwriting, eight support vector machines are trained which output each personality trait of the writer.

# Team Members:

* Deepmala Agarwal
* Sunil Agarwal
* Michael Hissey

# Description/Outline

In this project, we will use Deep Learning to read the handwriting of particular subjects. Then, we will apply Machine Learning to surface their personality traits. In the process of doing so, we analyze particular features:

* Baseline
* Individual Lines
* Letter Size
* Line Spacing
* Word Spacing
* Top Margin
* Pen Pressure
* Slant of Letters

# Datasets to be used

* MNIST db
* Personal handwriting samples

# Libraries Used

* Keras / Open CV Library (CV2 Library)
* SVM (Support Vector Machine)
* MatPlotLib
* Numpy

# Application Design

* Pre-Processing of Image (Image\_Cleanser.py)
* Load Data into SQLite DB
* Create a model.h5 to define the model
  + Read the handwriting (DL\_hand.ipynb)
  + Analyze the handwriting (ML\_person.ipynb)
    - There will be 8 files, one for each feature (listed above)
    - We will read through documentation to find correct models
* Create flask application (app.py) that will test the model
* We will deploy with Heroku & Charts.js
  + HTML & app.js
* Loss and Accuracy will help us determine how strong our model is
* Compare accuracy of two different libraries