Elinor Velasquez
Capstone Three
Project Proposal

I wish to predict stock market behavior by modeling the VWAP: Volume Weighted Average Price, the target variable to predict. VWAP is a trading benchmark used by traders that gives the average price the stock has traded at throughout the day, based on both volume and price.

My clients are Traders who trade stocks and want to be able to predict the behavior and outcome of stock market behavior so that their trading activity will be successful. My client may buy or sell energy stocks according to my analysis of the prediction of how the energy stocks will behave in the future.

I am using a subset of the Indian stock market data, downloaded from the Kaggle website. I am constraining the problem by using only the data from the publicly traded Indian companies which are energy-focussed.

I have already acquired the data by downloading the data from the Kaggle website, https://www.kaggle.com/rohanrao/nifty50-stock-market-data, and selecting only the energy companies for my data.

I will solve the problem by using time series forecasting with the Long Short-Term Memory (LSTM) Network in Python, via Keras. I will also consider traditional time series models such as ARIMA. I will also need to tune the LSTM in order to optimize the model via the hyper-parameters of the LSTM. I will also try simpler time series models such as a random walk and ARIMA to provide a model of the data. I will also use cross validation with a training set and a test set prepared from the data, which was already normalized/rescaled. Thus, for this project, I will employ deep learning and time series for predicting the VWAP.

My deliverables will be code (uploaded to GitHub), a project report, and a slide deck (presentation slides).

My criteria for success will be a good model of the data, given by successful prediction of the VWAP, and evaluation of metrics taken from the metrics package of SciKit-Learn, such as mean squared error, R-square, as well as data visualization for a global estimate of the best model.