# Portfolio Recommendation System

Berk Eserol
Advisor: Derek Lim
Professor: Andrew Ng

Stanford University

CS229 Machine Learning Autumn 2015

## Project Goals/Objectives

Using the power of machine learning to analyze the historical data, future predictions and projections can be performed on many different subjects. Even though, it is affected by various different external events, in this application project, stock market prices are tried to be predicted using only their historical data and a portfolio recommendation result is generated via the output of the regression and scoring. The aim is to recommend a portfolio with high accuracy profit return. The system is designed to produce result on any market structure that can be represented with similar data. For the project, historical NASDAQ stock prices are used.

## Dataset and Features

The data contains daily values from the beginning of its trade date to a recent date (ideally one day before the calculation). Each day data is in the following format

| MSFT | Year | Month | Day | Open | High | Low | Close | Volume |
|---|---|---|---|---|---|---|---|---|
| | 2015 | 11 | 12 | 53.48 | 53.98 | 53.19 | 53.32 | 34485500 |

The Systems Inputs:

**Set of NASDAQ stock names:**
Instead of running the system for all companies, the system considers only the given set of stocks and only uses them in the recommended portfolio.

**Budget:**
Budget is the maximum amount of total stock price in the recommendation bundle. The sum of the recommended stock price does not exceed the given budget amount.

**Keep Time Interval (KTI):**
The maximum intended time to keep the purchased stock. The keep time interval affects the features as a parameter. It can be minimum one day and maximum one year in the type of day count.

**History Interval (HI):**
History interval is used to trim the historical data to be considered. The data with the date outside of this interval is not considered.

The Feature Vector (interval is the two times of the KTI) :

| X0 | X1 | X2 | X3 | X4 | X5 | X6 | X7 | Y |
|---|---|---|---|---|---|---|---|---|
| 1 | Interval Start Price | Interval Maximum Price | Interval Minimum Price | Interval Volume Average | Interval Start Day | Interval Start Month | Interval Start Year | Interval Close Price |

## Methods

As a learning method, locally weighted linear regression is used minimizing

$$J(\theta) = \frac{1}{2}\sum_{i=1}^{m} w^{(i)}\left(\theta^T x^{(i)} - y^{(i)}\right)^2$$

$$w^{(i)} = \exp\left(-\frac{|x - x^{(i)}|}{2\tau^2}\right)$$

The $\theta$ is calculated using the normal equations

$$\theta = \left(X^T W X\right)^{-1} X^T W \vec{y}$$

## Results

The titles of the graphs represent the given input to the system in the orders of **KTI/Budget/HI.** Each graph has four dates (execution of the system) of recommendation results indicating the day of the market buy order requested. The stocks are assumed sold KTI days later and the charts show the virtual profit in terms of dollar after KTI days of the given date. The following system training errors (profit errors) are calculated according to the virtual profit of each stock that appears in the recommended bundle (RB). The result considered as error if it appears in the recommended bundle and the virtual profit is not positive after KTI days

$$\hat{\varepsilon}(p) = \frac{1}{m}\sum_{i=1}^{m} 1\{p(s^{(i)}) < 0\} \qquad s \in RB$$

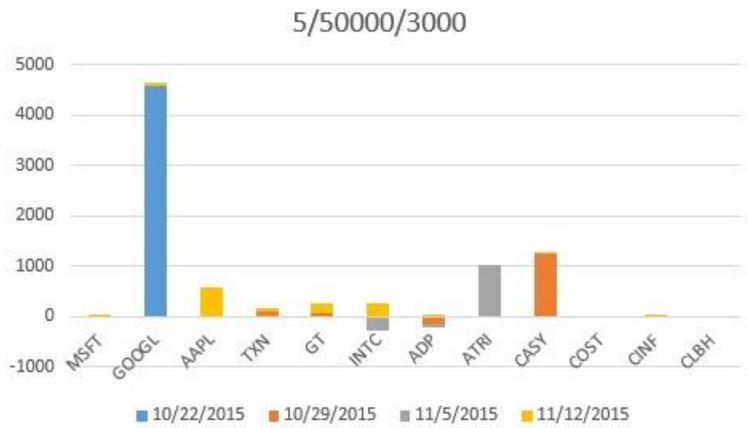
Using these 16 executions of the system with the predefined trending set of stocks, the average profit error is 0.1331 (profit accuracy is 86.69%)

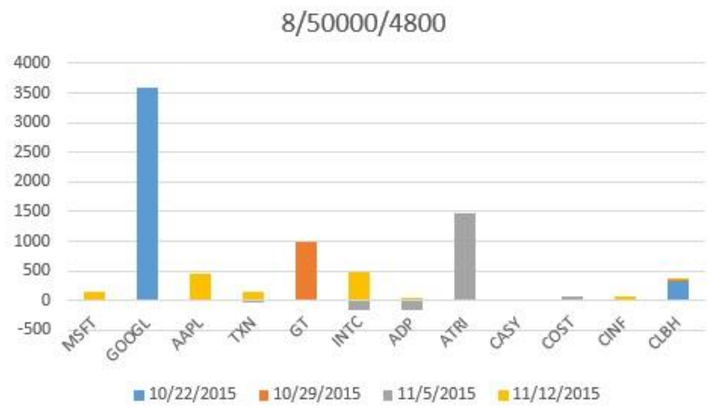
| $\hat{\varepsilon}(p)$ | 10/22/2015 | 10/29/2015 | 11/05/2015 | 11/12/2015 |
|---|---|---|---|---|
| 5/50000/3000 | 0.00 | 0.40 | 0.80 | 0.00 |
| 8/50000/4800 | 0.00 | 0.00 | 0.60 | 0.00 |
| 10/50000/6000 | 0.00 | 0.00 | 0.33 | 0.00 |
| 13/50000/7800 | 0.00 | 0.00 | 0.00 | 0.00 |

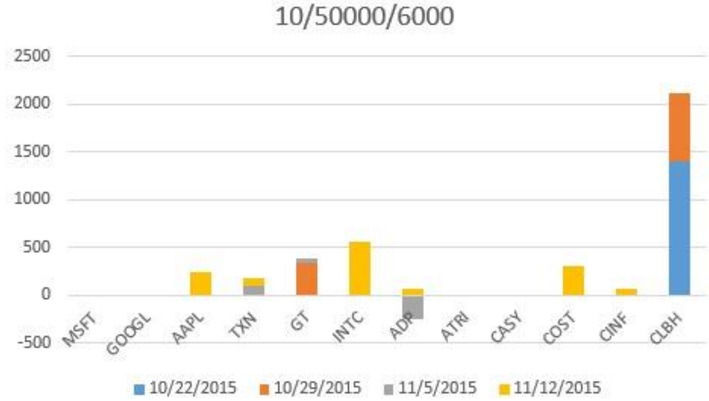
The total virtual profit combining the result of four execution of the system per given set of inputs are the following.

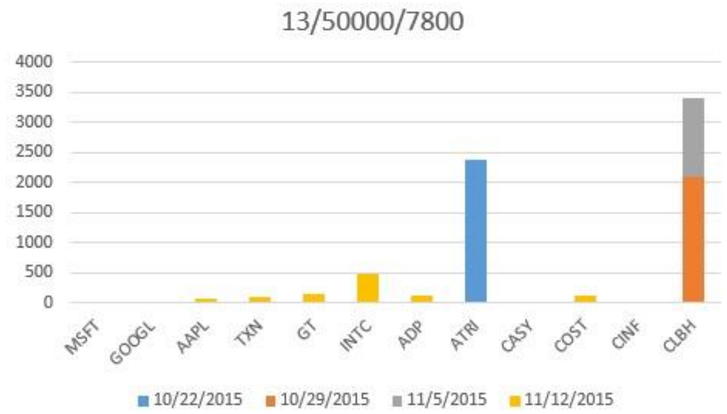
| Total Virtual Profit | 5/50000/3000 | 8/50000/4800 | 10/50000/6000 | 13/50000/7800 |
|---|---|---|---|---|
| | 7639.448 | 7441.044 | 3654.637 | 6835.232 |

Graphs:




## References

Horvitz, E. 2006. Machine learning, reasoning, and intelligence in daily life: Directions and challenges. Technical Report TR-2006-185, Microsoft Research.
Pouransari H., Chalabi H. 2014 Event-based stock market prediction, Stanford University
Di X. 2014. Stock Trend Prediction with Technical Indicators using SVM, Stanford University
Lin H. 2013 Feature Investigation for Stock market Prediction, Department of Aeronautics and Astronautics, Stanford University
Dai Y., Zhang Y. 2013 Machine Learning in Stock Price Trend Forecasting, Stanford University
Arık S., Eryılmaz B., Goldberg A. 2013 Supervised classication-based stock prediction and portfolio optimization
Shen S., Jiang H., Zhang T. 2012 Stock Market Forecasting Using Machine Learning Algorithms