Predicting Stock Price Movements Using Supervised Learning Algorithms

Zhao Yang, Nikolaos Karianakis, Ameet Talwalkar
Department of Computer Science, University of California, Los Angeles, CA 90025 USA

Abstract and Introduction

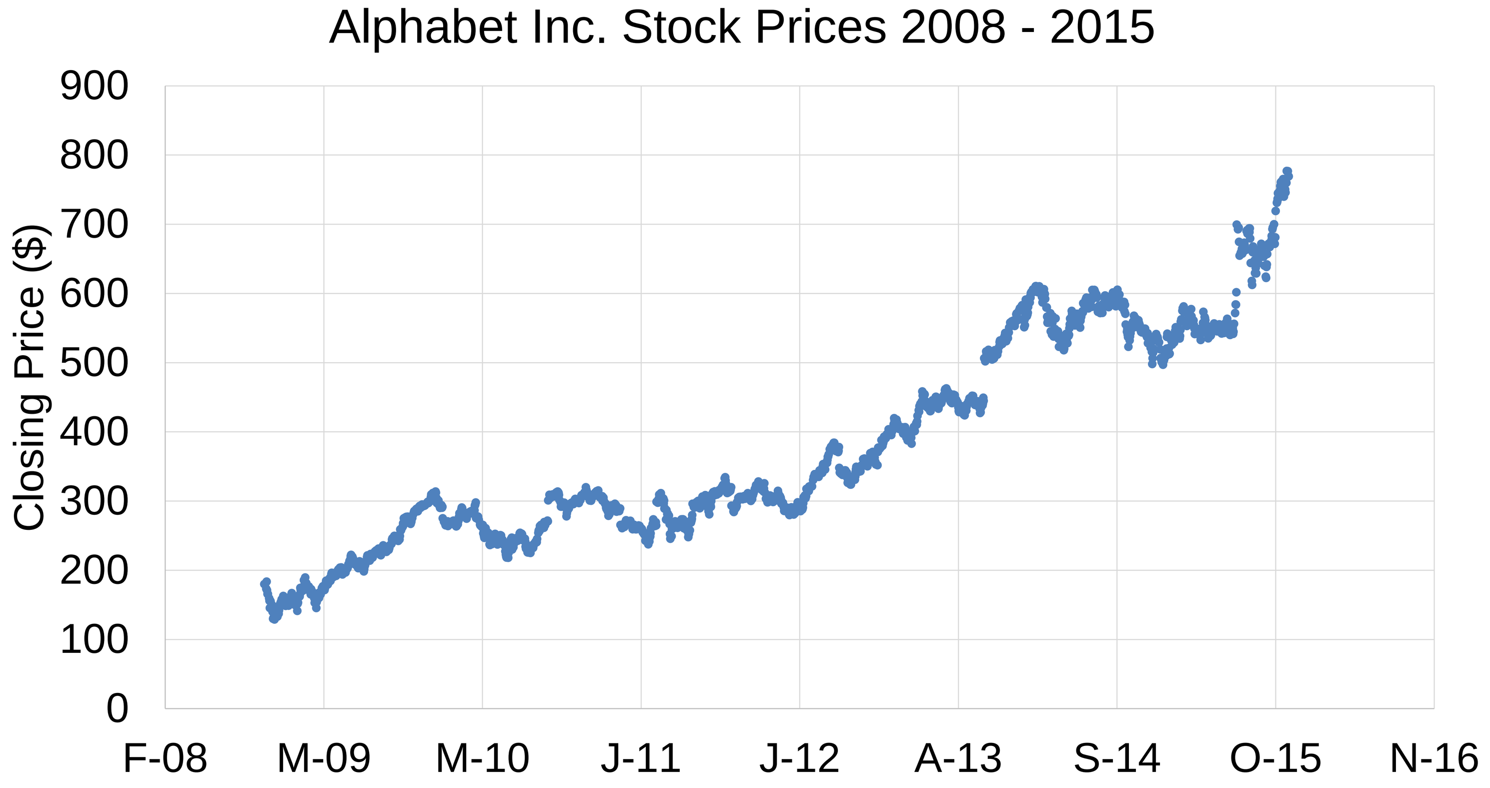
Stock price forecasting is an essential part of developing trading strategies. Existing prediction methodologies are fundamental analysis and technical analysis, both requiring human expert knowledge. Fundamental analysis aims to extract and analyze fundamental data that reflects the health and potential of a business, while technical analysis makes predictions on future stock prices using past prices and other trading variables.

Both methods have their limitations. Data required by fundamental analysis is often sparse, incomplete, biased, and difficult to obtain. Technical analysis heavily relies on personal preferences in the interpretation of data, with different beliefs yielding different predictions. In addition, the efficient market theory questions the validity of technical analysis by stating that current stock price already reflects all available information regarding the value of a company.

In this study, I pose a challenge to the efficient market theory and present an automated mechanism for stock price movement forecasting. Using supervised statistical learning algorithms, I achieved ~90% accuracy for prediction period longer than 25 days as compared to the baseline accuracy of ~60%. This result suggests that past prices indeed contain information relevant to future stock prices, and automatic predictions on price trends can be accurate. In the future, I plan to generalize my experiments to more stocks and analyze the implications further.

Data and Features

Alphabet Inc. Stock Prices 2008 - 2015

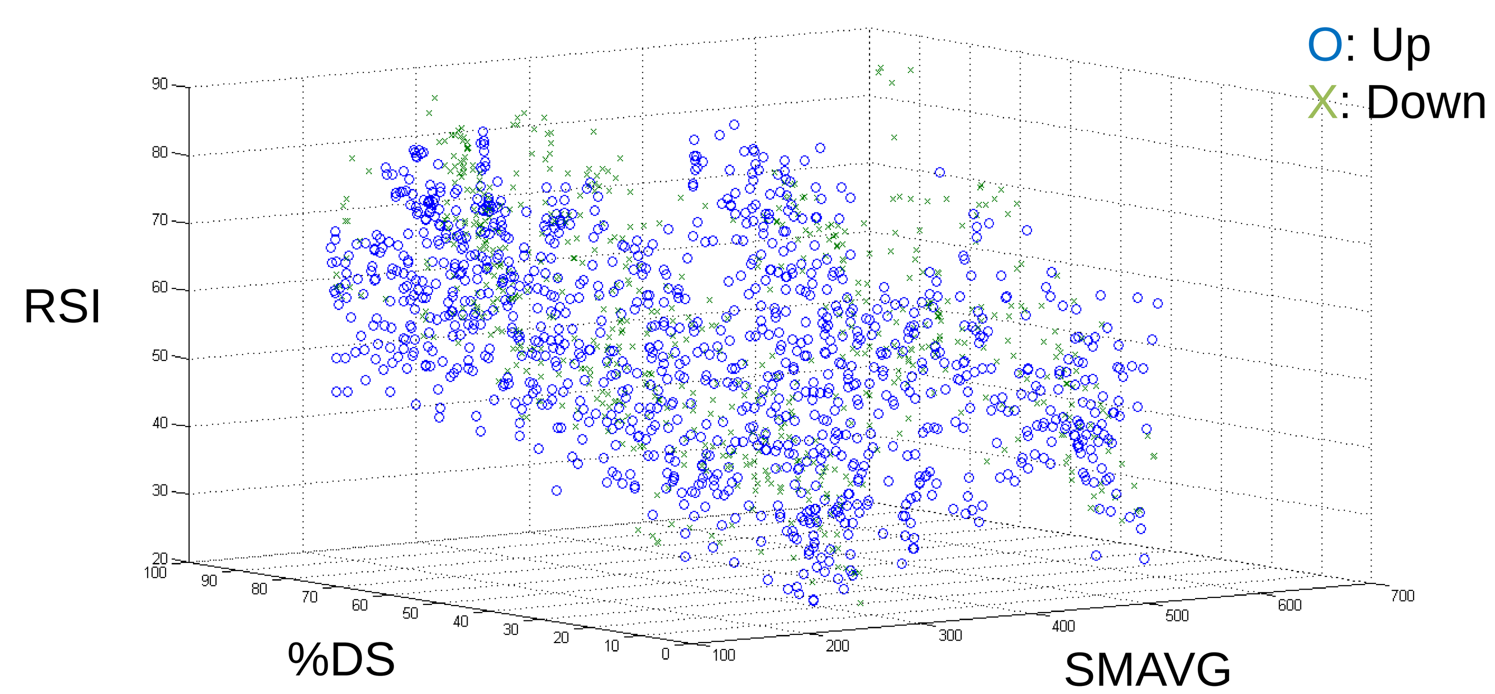


- Company: Alphabet Inc.
- Stock Type: Class A shares
- Time Frame: Nov.2008 - Nov.2015
- Daily closing prices: generation of price movement labels.
- Technical indicators: relative strength index, simple moving average, moving average oscillator, etc.
- Financial data: PE ratio, enterprise value, overridable alpha, etc.

Data and Features (cont')

Date	%B	SMAVG(5)	SMAVG(50)	Osc	PTPS	RSI	%DS	Close	Price Movement (1-day)
11/20/2015	1	758.67	700.52	2.512	734.91	68.64	82.837	777	
11/19/2015	1	751.284	698.086	-2.2333	764.617	63.37	83.121	759.94	1
11/18/2015	1	750.602	695.909	-1.9094	765.855	63.41	85.427	760.01	-1

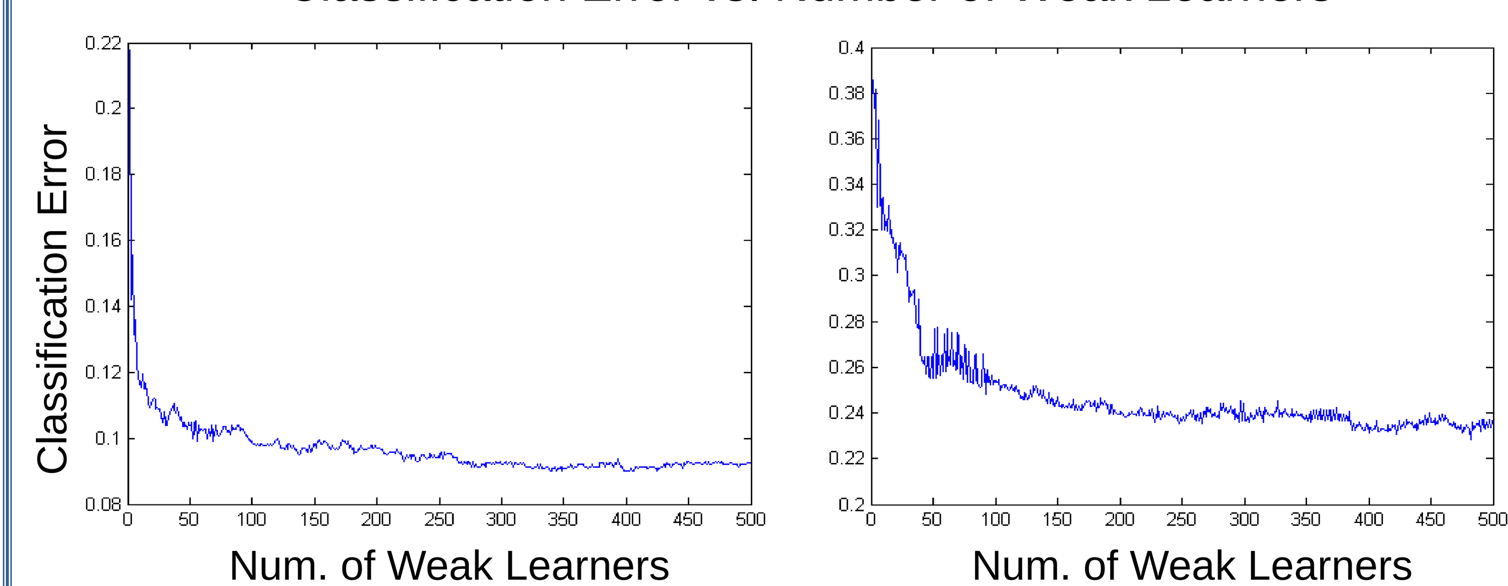
- Label definition: 1 for upward movement; -1 for downward movement
- Label calculation: current label = sign(future price – current price)



- No visible boundary in three-dimensional space
- Cross validation and forward feature selection
- The best (greedy) feature combination: seven technical indicators

Algorithm Tuning

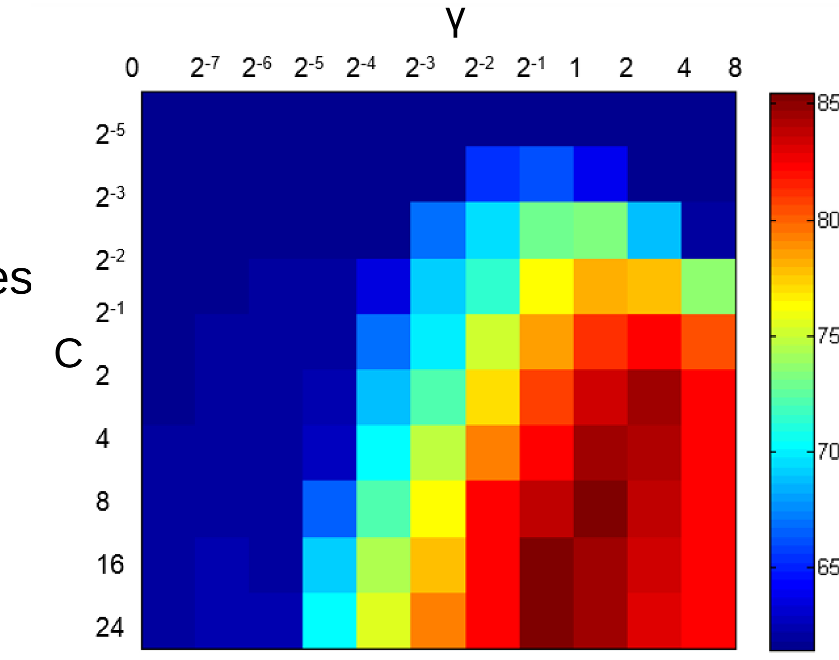
Classification Error vs. Number of Weak Learners



- Top Left: Bootstrap Aggregation
- Top Right: AdaboostM1
- Weak learner: Decision Tree
- Number of trees: 500
- Time window: 30 days

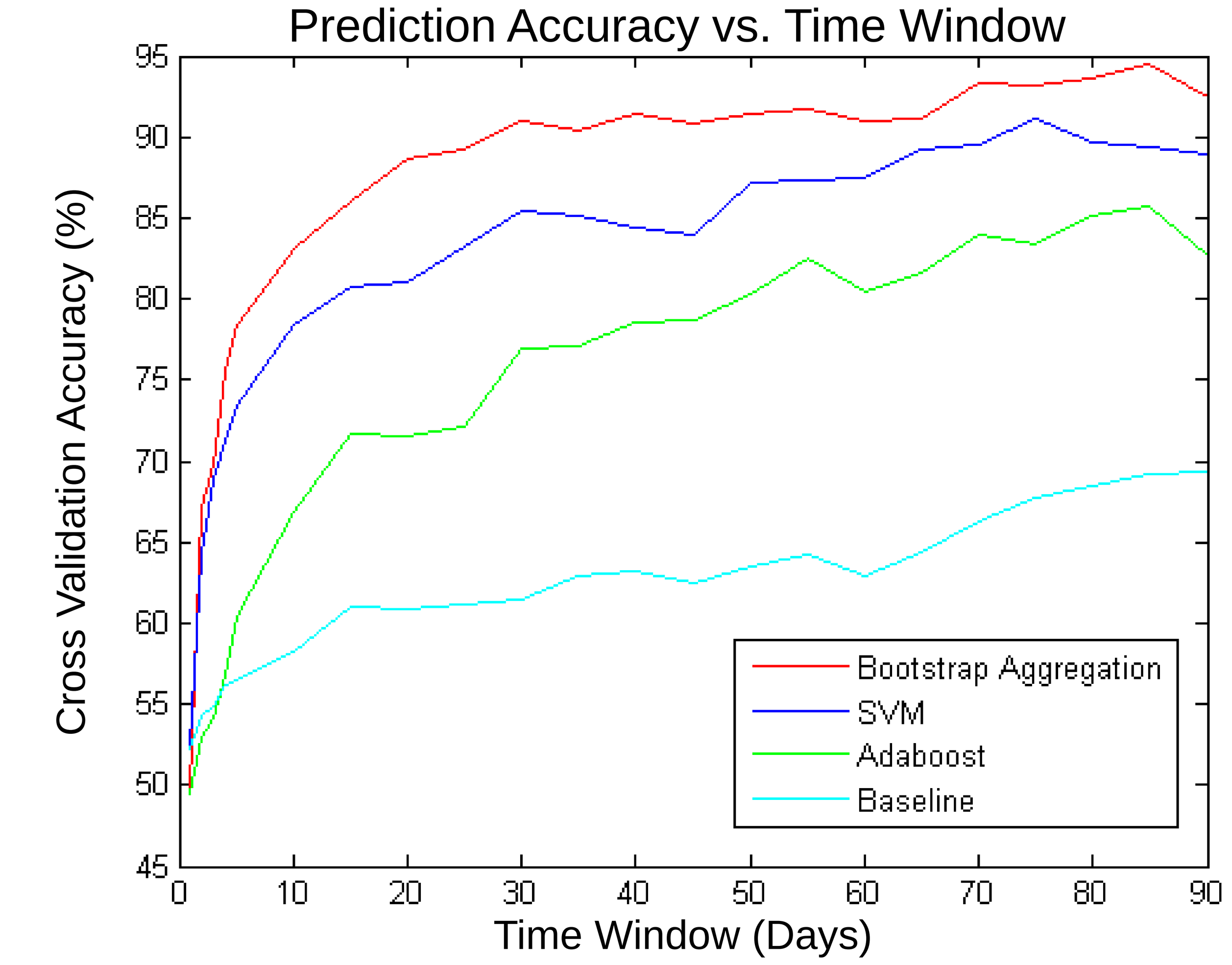
- Right: Support Vector Machines
- Kernel: Radial Basis Function
- Optimization:
- 5-fold Cross validation
- Grid Search
- Time window: 30 days

Validation Accuracy w.r.t. C and γ



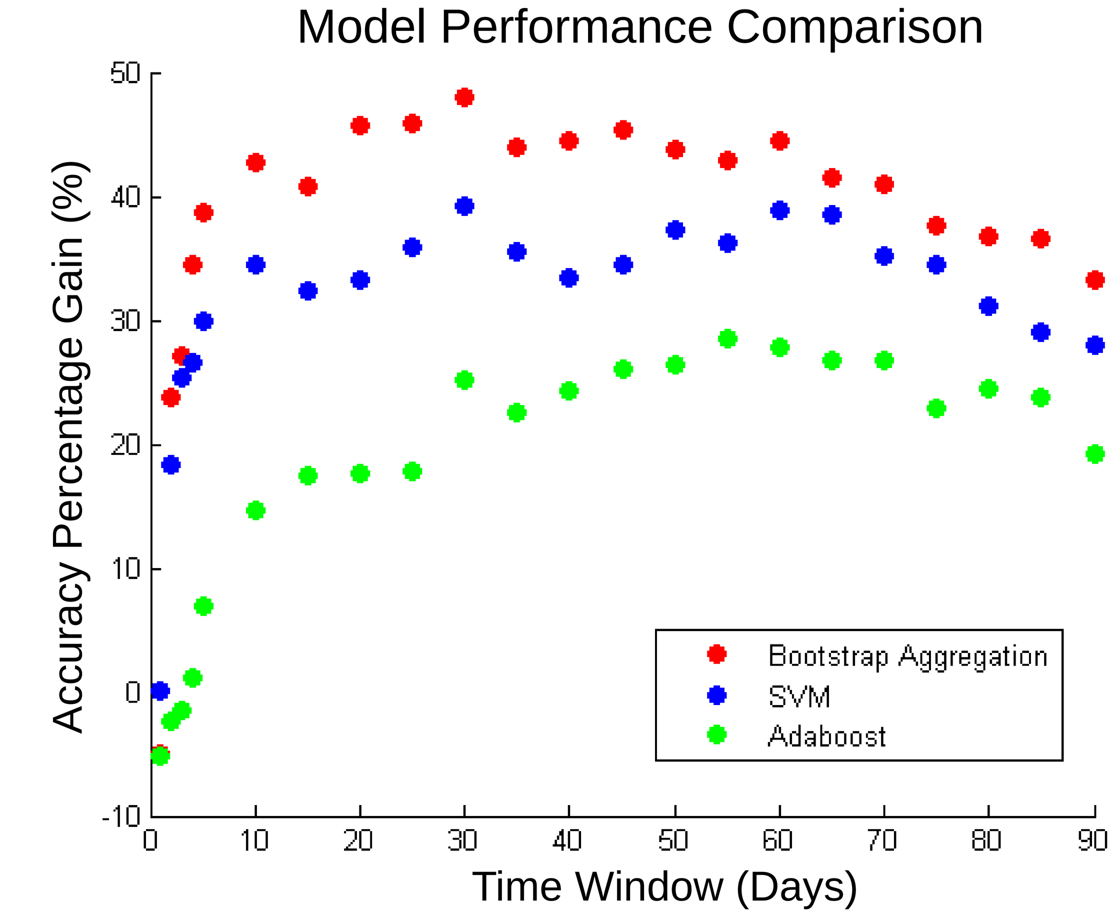
Results and Conclusion

Prediction Accuracy vs. Time Window



- The highest validation accuracy of %94.45 is obtained at 85-day interval using bootstrap aggregation (Bagging)
- The baseline accuracy is the higher of the proportions of the two labels, representing the accuracy when guessing using one label
- Bagging demonstrates strong predictive power for predictions over 3 days or longer
- Predictions over periods shorter than 3 days are equivalent to random guesses

Model Performance Comparison



- Accuracy percentage gain is the percentage increase in prediction accuracy over the baseline achieved by the algorithm
- Bagging delivers the best performance at the 30-day time window with 48% increase in accuracy
- By applying supervised learning to features derived from past stock prices, I achieved great long-term accuracy on predicting stock price movements.
- In the future, I would like to use limit order book to predict short-term stock price changes.