Richard Cloutier

Machine Intelligence

**Abstract**

Predicting the stock market has been the end goal of every financial analyst, trader, and stockbroker since the opening of the stock market in 1817. Only recently have engineers and computer scientists joined them in their pursuit of predicting the unpredictable. In this project, various classification models and neural networks were used to predict future stock prices, with support vector machines (SVM) being heavily favored. Additionally, news articles were analyzed to determine correlations between key words and their impact on a company’s stock price. The application of SVM classification with historical stock opening and closing prices on forecasted stock prices resulted in a test accuracy of ***XX***% over 12 stocks. The addition of the news article scanner resulted in a test accuracy of ***YY***%.

**Expected Deliverables**

1. Algorithm that uses historical price data to generate a model of prices for various stocks.
   1. Comparison between various classification/regression models
2. Scan historical and real time news articles containing names of companies and market sectors to compare to their effect on stock prices.
3. Increase the number of stocks being tested from 12 to 30.
4. Use all stocks listed in Dow Jones IA.
5. Real time trading algorithm with a non-negative ROI.
   1. End goal is ROI percentage >= savings account or index fund ROI
6. Comparison between classification/regression models and neural network implemented in TensorFlow. (Stretch goal, most likely will not be implementing TensorFlow/neural networks)

**References**

1. W. Huang, Y. Nakamori, and S.-Y. Wang, “Forecasting stock market movement direction with support vector machine,” ScienceDirect, 25-May-2004. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0305054804000681>. [Accessed: 09-Nov-2018].

Huang et al. analyzes various classification models and neural networks to determine the best algorithm for predicting stock prices. They determine that the combination of the support vector machine (SVM) algorithm and other classification methods produce the best prediction algorithm.

1. D. Mirvish, “The Hathaway Effect: How Anne Gives Warren Buffett a Rise,” The Huffington Post, 07-Dec-2017. [Online]. Available: <https://www.huffingtonpost.com/dan-mirvish/the-hathaway-effect-how-a_b_830041.html>. [Accessed: 09-Nov-2018].

Mirvish explores how a poorly implemented stock price forecaster with news article scanning support mistakes mentions of actress Anne Hathaway with those of holding company Berkshire Hathaway.

1. Shen and H. Jiang, “Stock Market Forecasting Using Machine Learning Algorithms,” Semantics Scholar. [Online]. Available: <https://pdfs.semanticscholar.org/b68e/8d2f4d2c709bb5919b82effcb6a7bbd3db37.pdf>. [Accessed: 09-Nov-2018].

Shen et al. compares SVM, linear, and GLM classification models augmented with temporal correlation between global markets for price forecasting on the NASDAQ, Dow Jones Industrial Average, and the S&P 500.

1. C.-Y. Yeh, C.-W. Huang, and S.-J. Lee, “A multiple-kernel support vector regression approach for stock market price forecasting,” ScienceDirect, 15-Aug-2010. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0957417410007876>. [Accessed: 09-Nov-2018].

Yeh et al. investigates the application of multiple-kernel learning to remove the need to manually adjust the hyperparameters of commonly used SVM models. This allows for a model with greater accuracy with a higher number of hyperparameters adjusted than what would otherwise have been used.

1. L. Yu, H. Chen, S. Wang and K. K. Lai, "Evolving Least Squares Support Vector Machines for Stock Market Trend Mining," in IEEE Transactions on Evolutionary Computation, vol. 13, no. 1, pp. 87-102, Feb. 2009. URL: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4632148&isnumber=4769002>

Yu et al. investigates the usage of an evolving least squares multi-kernel SVM algorithm to predict stock market prices. Multiple genetic algorithms are used to select input features and tune the parameters of the SVM model.