Stock Predictor Using Machine Learning

Prakash Dhimal, Kevin Sanford
Faculty Advisor: Dr. Eman El-Sheikh
Department of Computer Science, University of West Florida

Introduction

Predicting the stock price trend by analysing historical market data has been a long-time attractive topic to both investors and researchers from various fields.

Stock Predictor is a Stock prediction program written in Python that provides predictions for a given stock's current trading day closing price.

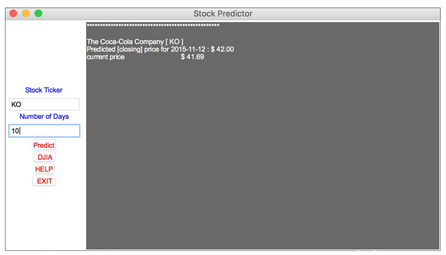


Fig: Stock Predictor Graphical User Interface

- Data from Yahoo Finance
- AI Implementation:
  - Machine Learning
    - Support Vector Machine

- To predict a specific stock, user enters a company's stock ticker (e.g. GOOG) and number of days they wish to train the Support Vector Machine.

Methods

- **Data Collection**
  - We used yahoo-finance, a python module to get data from Yahoo Finance.
  - We maintained a database of 27382 company names and their ticker symbols.

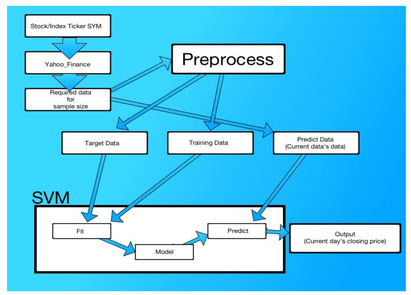


Fig: Flow chart for Stock Predictor Program

- **Support Vector Machine (SVM)**
  - Fit training set to target data
  - Predict using current trading day's features

Data Preprocessing:

- An important step in implementation of any machine learning algorithm is data preprocessing.
- Transforming raw market data to get it ready for the Support Vector Machine.
- Data Standardization
  - All data in range -1 to 1
  - Standard deviation = 1
  - Mean = 0
- Data Separation:
  - Training data (features)
    - Opening price, daily low, daily high, price change
  - Target data
    - Closing price
- Sample size
  - Number of trading days in the data set

Results

Combination of various inputs like sample size, SVM kernel, gamma, and C, support vector machine provided results up-to 60% on predicting the right direction of stock price (up or down)

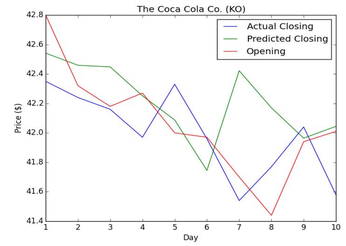
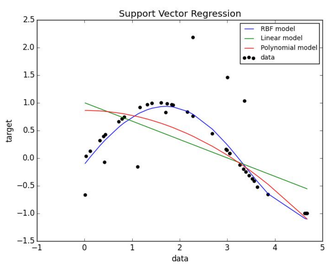


Fig: Predicted closing price vs actual closing price



- Radial basis function (RBF) kernel with gamma= 0.001 and C=1e3 produced the maximum results.
- Average result was 50% on a 10 day prediction run.

Technical Analysis

- Forecasting direction of prices using historical market data
  - Opening Price, Closing Price, Daily Low, Daily High, Volume, Spread, etc
- Different from fundamental analysis which measures company's intrinsic value
- Machine learning provides powerful tools to extract trends from market data

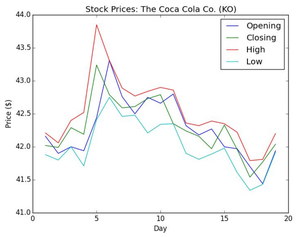


Fig: Stock price movements for The Coca Cola Co.


Fig: Technical Analysis[2]

Support Vector Machine

- Support Vector Machines (SVMs) are a set of supervised learning methods used for classification, regression, and outliers detection.[1]
- When comparing two or more companies, we standardized the data to better compare the fluctuation of stock prices. (e.g. opening prices)

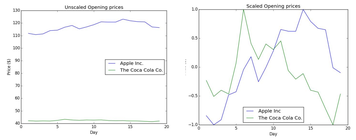


Fig: Unscaled data vs Scaled data

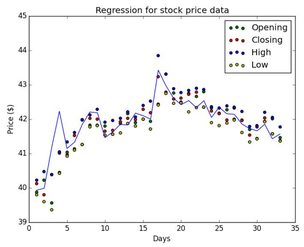


Fig: Regression for stock price data using RBF kernel

- Scikit Learn is an open source machine learning library written in python
- Provides Support vector regression to solve regression problems
- Different kernel implementations:
  - Linear and nonlinear
- Fit model
  - Fits SVM model based on given data
- Predict
  - Performs regression on prediction data

Conclusion

We made a stock predictor using Support Vector Machine in python. Unfortunately, with any prediction method it can't predict the unforeseeable future(i.e. market crashes) and only uses past information as its basis -- which may also include tragic events that affected the company's value. As it stands, our stock predictor can't exclude these outliers from the data. So for now the stock predictor tends to overestimate a stock's price movement. On average it predicts about 50% correct which is the same as flipping a coin.

Future Works

- Predict a stock's price a week and a month in the future.
- Adjust SVM parameters C and gamma based on the sample size

Acknowledgements

- Dr. Kevin Krieger, Professor of Finance, University of West Florida
- Yahoo Finance
- Scikit Learn

References

[1] "1.4. Support Vector Machines." *1.4. Support Vector Machines — Scikit-learn 0.17 Documentation*. Web. 12 Nov. 2015.
[2] FOREX TRADING. (n.d.). Retrieved November 14, 2015, from http://www.forextradeoracle.com/technical-analysis.php