Stock Market Prediction using Support Vector Machine

KEVIN JOHN, DIVYA KUMAR, NELSON MORRIS, ANJALI SINGH

GUIDE : ANAGHA SHASTRI

Abstract

The aim of the project is to examine a number of different forecasting techniques to predict future stock returns based on past returns and numerical news indicators to construct a portfolio of multiple stocks in order to diversify the risk.

We do this by applying supervised learning methods for stock price forecasting by interpreting the seemingly chaotic market data.

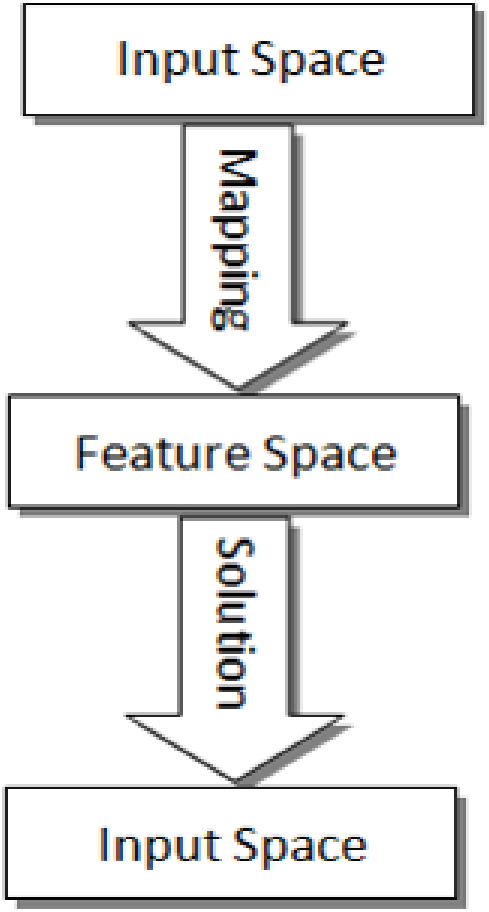
Based on the fact that the moving direction of these parameters have certain inertia within short-term period, we here explored the potential application of the moving trends of these parameters within 4 different time periods(5, 15, 30 and 45 trading days) respectively for forecasting the movement direction of stock price.

Problem Statement

- Forecasting the short-term trend of a stock market has been a big challenging task.
- Parameters of stock markets, including open/close prices, daily-high/low prices and trading volumes, were frequently used in previous studies to forecast the stock market.
- Our aim is to predict the future stock prices accurately using supervised learning algorithms, compare the results to predict the most accurate algorithm. This is to be achieved by building a software that analyses historical stock data.

Algorithm: SVM

- Well-known tool from the area of supervised Machine Learning
- Used both for classification and regression
- In High Dimensional Space, Vector Machines are very effective.
- SVM can cope up with noisy data.
- SVM is a deterministic algorithm.
- Using Kernel it can run very complex machines.
- Using wide margin in SVM can achieve high accuracy.
- Accuracy for neural networks is 81%.
- Accuracy for SVM is 97%.



Working

- In more mathematical terms, for the training set we have *N* features, for each of them we have *M* samples. We also have *M* responses.

$$\begin{pmatrix} F_1^1 & F_1^2 & F_1^3 & \dots & F_1^N \\ F_2^1 & F_2^2 & F_2^3 & \dots & F_2^N \\ F_3^1 & F_3^2 & F_3^3 & \dots & F_3^N \\ \dots & & & & \\ F_M^1 & F_M^2 & F_M^3 & \dots & F_M^N \end{pmatrix} \Rightarrow \begin{pmatrix} R_1 \\ R_2 \\ R_3 \\ \dots \\ R_M \end{pmatrix}$$

- Given a row of feature values, the left matrix, the SVM is trained to produce the response value.

- In our specific example, we have 500 samples and the corresponding responses.

$$\begin{pmatrix} F_1^1 & F_1^2 & F_1^3 & F_1^4 & F_1^5 \\ F_2^1 & F_2^2 & F_2^3 & F_2^4 & F_2^5 \\ F_3^1 & F_3^2 & F_3^3 & F_3^4 & F_3^5 \\ \cdots & & & & \\ F_{500}^1 & F_{500}^2 & F_{500}^3 & F_{500}^4 & F_{500}^5 \end{pmatrix} \Rightarrow \begin{pmatrix} R_1 \\ R_2 \\ R_3 \\ \cdots \\ R_{500} \end{pmatrix}$$

- We have five columns (features), each column corresponding to the returns with a different lag
(from 1 to 5).

- Once the SVM is trained on this set, we can start feeding it with sets of five features, corresponding to the returns for the five previous days, and the SVM will provide us with the response, which is the forecasted return.
- For example, after training the SVM on the previous 500 days, we will use the returns for days 500, 499, 498, 497 and 496 (these are ours as the input to obtain the forecasted return for day 501.

Results obtained :

Comparison of accuracy using different algorithms :

Classifier	Mean overall accuracy [%]			
ANN	61.8	80.2	87.4	87.1
RF	55.8	82.3	86.6	87.4
SVM-RBF	68.6	82.2	88.0	88.1
SVM-POLY	68.4	80.3	87.7	87.8
ML	45.0	75.8	79.1	78.9

Conclusion

In this project we attempted to highlight the main ideas underlying the SVM method. We applied supervised learning techniques in predicting the stock price trend. Our findings can be summarized into following:

- Supervised learning techniques like Logistic Regression(LR), Neural Network (NN), SVM have been used for the prediction and we found that SVM model can provide the highest predicting accuracy.
- Our feature selection analysis indicates that when use all of the 4 features we will get the highest accuracy. That's because the number of data points is much bigger than that of the features.
- It also gave us a brief idea regarding technologies that may help the users to predict the stock prices and invest at the right time

References

[1] Zhen Hu, Jie Zhu, and Ken Tse, (2013), “Stocks Market Prediction Using Support Vector Machine”

[2] Yuqing Dai, Yuning Zhang, “Machine Learning in Stock Price Trend Fore-casting”

[3] Xin Guo,How can machines learning help stock investment? Email: guoxin@stanford.edu

[4]https://www.academia.edu/11692137/”Analyzing_Different_Machine_Learning_Techniques_for_Stock_Market_Prediction”

[5] Xinjie Di , “Stock Trend Prediction with Technical Indicators using SVM”