Stock Market Prediction Using Machine Learning

Research Proposal · November 2018

DOI: 10.13140/RG.2.2.12300.77448

1 author:

Kranthi Sai Reddy Vanukuru
Sreenidhi Institute of Science & Technology
6 PUBLICATIONS **25** CITATIONS

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 05 Issue: 10 | Oct 2018 www.irjet.net p-ISSN: 2395-0072

Stock Market Prediction Using Machine Learning

V Kranthi Sai Reddy[1]

[1]*Student, ECM, Sreenidhi Institute of Science and Technology, Hyderabad, India*

Abstract - *In the finance world stock trading is one of the most important activities. Stock market prediction is an act of trying to determine the future value of a stock other financial instrument traded on a financial exchange. This paper explains the prediction of a stock using Machine Learning. The technical and fundamental or the time series analysis is used by the most of the stockbrokers while making the stock predictions. The programming language is used to predict the stock market using machine learning is Python. In this paper we propose a Machine Learning (ML) approach that will be trained from the available stocks data and gain intelligence and then uses the acquired knowledge for an accurate prediction. In this context this study uses a machine learning technique called Support Vector Machine (SVM) to predict stock prices for the large and small capitalizations and in the three different markets, employing prices with both daily and up-to-the-minute frequencies.*

Key Words: Stock Market, Machine Learning, Predictions, Support Vector Machine

1. INTRODUCTION

Basically, quantitative traders with a lot of money from stock markets buy stocks derivatives and equities at a cheap price and later on selling them at high price. The trend in a stock market prediction is not a new thing and yet this issue is kept being discussed by various organizations. There are two types to analyze stocks which investors perform before investing in a stock, first is the fundamental analysis, in this analysis investors look at the intrinsic value of stocks, and performance of the industry, economy, political climate etc. to decide that whether to invest or not. On the other hand, the technical analysis it is an evolution of stocks by the means of studying the statistics generated by market activity, such as past prices and volumes.

In the recent years, increasing prominence of machine learning in various industries have enlightened many traders to apply machine learning techniques to the field, and some of them have produced quite promising results.

This paper will develop a financial data predictor program in which there will be a dataset storing all historical stock prices and data will be treated as training sets for the program. The main purpose of the prediction is to reduce uncertainty associated to investment decision making.

Stock Market follows the random walk, which implies that the best prediction you can have about tomorrow's value is today's value. Indisputably, the forecasting stock indices is very difficult because of the market volatility that needs accurate forecast model. The stock market indices are highly fluctuating and it effects the investor's belief. Stock prices are considered to be a very dynamic and susceptible to quick changes because of underlying nature of the financial domain and in part because of the mix of a known parameters (Previous day's closing price, P/E ratio etc.) and the unknown factors (like Election Results, Rumors etc.). There has been numerous attempts to predict stock price with Machine Learning. The focus of each research projects varies a lot in three ways. (1) The targeting price change can be near-term (less than a minute), short-term (tomorrow to a few days later), and a long-term (months later), (2) The set of stocks can be in limited to less than 10 particular stock, to stocks in particular industry, to generally all stocks. (3) The predictors used can range from a global news and economy trend, to particular characteristics of the company, to purely time series data of the stock price.

The probable stock market prediction target can be the future stock price or the volatility of the prices or market trend. In the prediction there are two types like dummy and a real time prediction which is used in stock market prediction system. In Dummy prediction they have define some set of rules and predict the future price of shares by calculating the average price. In the real time prediction compulsory used internet and saw current price of shares of the company.

Computational advances have led to introduction of machine learning techniques for the predictive systems in financial markets. In this paper we are using a Machine Learning technique i.e., Support Vector Machine (SVM) in order to predict the stock market and we are using Python language for programming.

2. Methodology

In this project the prediction of stock market is done by the Support Vector Machine (SVM) and Radial Basis Function (RBF).

2.1 Support Vector Machine

A Support Vector Machine (SVM) is a discriminative classifier that formally defined by the separating hyperplane. In other words, the given labeled training data (supervised learning), the algorithm outputs the optimal hyperplane which categorizes new examples. In the two-dimensional space this hyperplane is a line dividing a plane into two parts where in each class lay in either side.

Support Vector Machine (SVM) is considered to be as one of the most suitable algorithms available for the time series prediction. The supervised algorithm can be used in

both, regression and classification. The SVM involves in plotting of data as point in the space of n dimensions.

These dimensions are the attributes that are plotted on particular co-ordinates. SVM algorithm draws a boundary over the data set called as the hyper-plane, which separates the data into two classes as shown in the Fig 1.

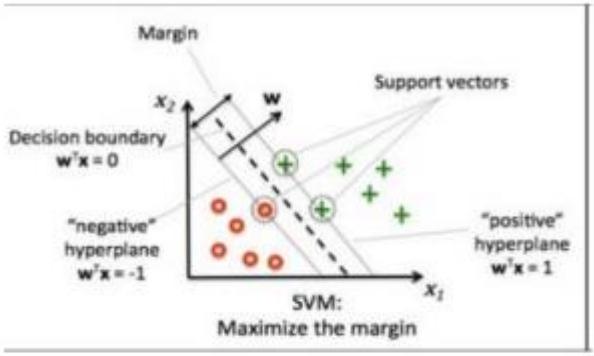


Fig -1: The Support Vector Machine Decision Making Boundary

The hyper-plane is a decision boundary which is later extended or maximized on either side between the data points. Considering the same figure, if μ is some unknown data point and w is vector which is perpendicular to the hyper-plane, then the SVM decision rule will be

$\bar{w}\bar{\mu} + b \geq 0$..... (1)

The width w of the hyper-plane must be maximized the spread

w = [2/ || w ||] ..... (2)

w = (max [2/|| w ||]) ..... (3)

Applying lagrange's multiplier as

L = 0.5|| w ||^2 →-$\sum\alpha_i[y_i(\omega_i x_i + b) - 1]$..... (4)

L = $\sum\alpha_i - 0.5\Sigma_i\Sigma_j\alpha_i\alpha_j y_i y_j x_i x_j$..... (5)

The updated decision rule will be

$(\sum\alpha_i y_i x_i)\mu+b \geq 0$

2.2 Radial Basis Function (RBF)

In the machine learning, the radial basis function kernel, or RBF kernel, is a popular kernel function used in the various kernelized learning algorithms. In particular, it is most commonly used in support vector machine classification.

A radial basis function is the real-valued function whose value depends only on the distance from the origin, so that; or alternatively on the distance from some other point, called a center, so that. Any function which satisfies the property is a radial function.

RBF = Local Response Function

The RBF Kernel is nothing more than a low-band pass filter, which is well known in Signal Processing as a tool to smooth images. RBF Kernel acts as the prior that selects out smooth solutions.

The Radial basis function kernel, is also called as the RBF kernel, or Gaussian kernel, is a kernel that is in the form of a radial basis function (more specifically, a Gaussian function). The RBF kernel is defined as

$K_{RBF}(x, x') = \exp[-y\,||x - x'||^2]$

Where y is the parameter that sets "spread" of the kernel

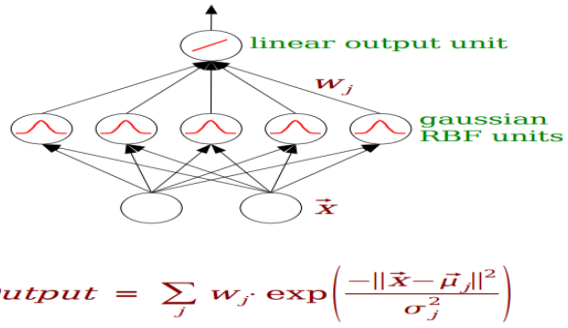


Fig -2: RBF Network

The RBF units provide a new basis set for synthesizing the output function. The radial basis functions are not orthogonal and are overcomplete.

2.3 The Learning Environment

The Weka and the YALE Data Mining Environments were used for carrying out the experiments. The general setup used is as follows:

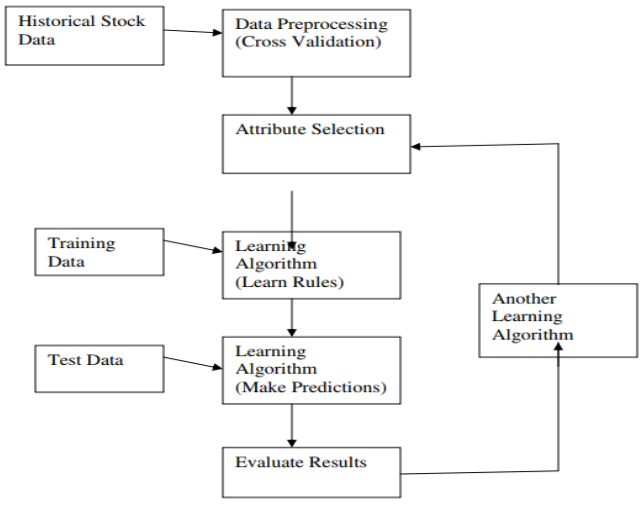


Fig -3: Learning Environment

3. Model Creation and Evaluation Methods

In this paper we focus on predicting the Stock Market using Machine Learning model i.e., Support Vector Machine (SVM) by RBF kernel.

3.1 Feature Selection

In this project we use four features to predict stock price direction – price volatility, price momentum, sector volatility, and sector momentum. More details are provided in Table 1, styled in the form used by Kim [4].

Table 1: Features used in SVM

Feature Name	Description	Formula
σs	Stock price volatility. This is an average over the past n days of percent change in the given stock's price per day.	$\frac{\sum_{i=t-n+1}^{t} \frac{C_i - C_{i-1}}{C_{i-1}}}{n}$
Stock Momentum	This is an average of the given stock's momentum over the past n days. Each day is labeled 1 if closing price that day is higher than the day before, and −1 if the price is lower than the day before.	$\frac{\sum_{i=t-n+1}^{t} y}{n}$
σi	Index volatility. This is an average over the past n days of percent change in the index's price per day.	$\frac{\sum_{i=t-n+1}^{t} \frac{I_i - I_{i-1}}{I_{i-1}}}{n}$
Index Momentum	This is an average of the index's momentum over the past n days. Each day it is labeled 1 if closing price that day is higher than the day before, and −1 if the price is lower than the day before	$\frac{\sum_{i=t-n+1}^{t} d}{n}$

3.2 Steps for Stock Market Prediction

Step 1: This step is important for the download data from the net. We are predicting the financial market value of any stock. So that the share value up to the closing date are download from the site.

Step 2: In the next step the data value of any stock that can be converted into the CSV file (Comma Separate Value) so that it will easily load into the algorithm.

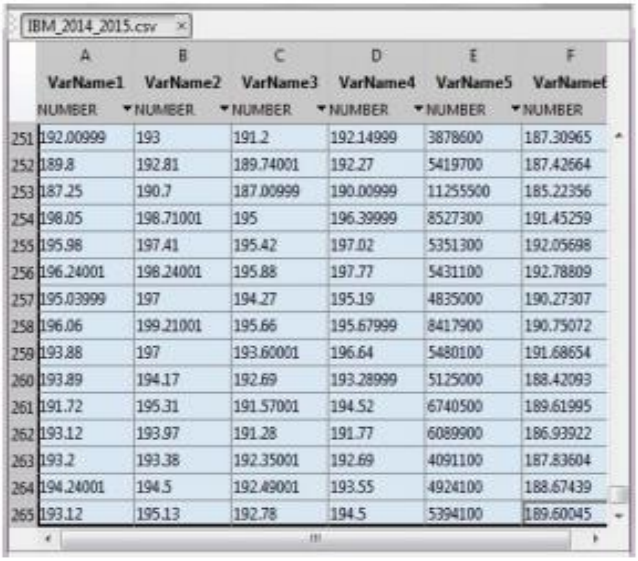
Step 3: In the next step in which GUI is open and when we click on the SVM button it will show the window from which we select the stock dataset value file.

Step 4: After selecting the stock dataset file from the folder it will show graph Stock before mapping and stock after mapping.

Step 5: The next step algorithm calculated the log2c and log2g value for minimizing error. So, it will predict the graph for the dataset value efficiently.

Step 6: In final step algorithm display the predicted value graph of select stock which shows the original value and predicted value of the stock.

4. RESULTS



IBM_2014_2015.csv

	A	B	C	D	E	F
	VarName1	VarName2	VarName3	VarName4	VarName5	VarName6
	NUMBER	NUMBER	NUMBER	NUMBER	NUMBER	NUMBER
251	192.00999	193	191.2	192.14999	3878600	187.30965
252	189.8	192.81	189.74001	192.27	5419700	187.42664
253	187.25	190.7	187.00999	190.00999	11255500	185.22356
254	198.05	198.71001	195	196.39999	8527300	191.45259
255	195.98	197.41	195.42	197.02	5351300	192.05698
256	196.24001	198.24001	195.88	197.77	5431100	192.78809
257	195.03999	197	194.27	195.19	4835000	190.27307
258	196.06	199.21001	195.66	195.67999	8417900	190.75072
259	193.88	197	193.60001	196.64	5480100	191.68654
260	193.89	194.17	192.69	193.28999	5125000	188.42093
261	191.72	195.31	191.57001	194.52	6740500	189.61995
262	193.12	193.97	191.28	191.77	6089900	186.93922
263	193.2	193.38	192.35001	192.69	4091100	187.83604
264	194.24001	194.5	192.49001	193.55	4924100	188.67439
265	193.12	195.13	192.78	194.5	5394100	189.60045

Fig -4: Stock dataset for IBM Inc. in CSV file

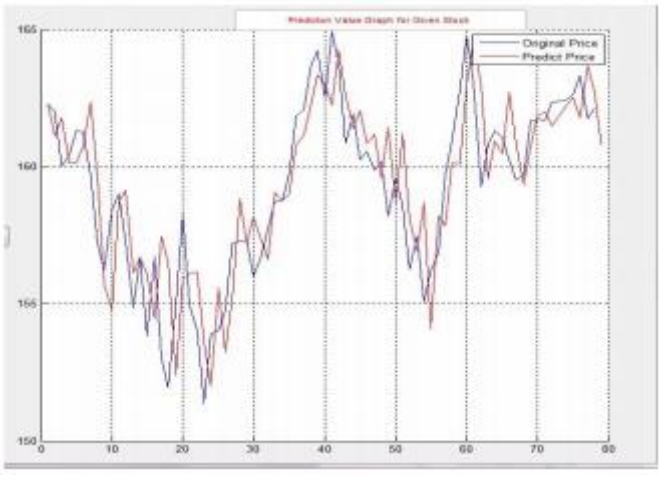


Fig -5: Predicted Output by SVM for IBM Inc.

5. CONCLUSION

In the project, we proposed the use of the data collected from different global financial markets with machine learning algorithms in order to predict the stock index movements. SVM algorithm works on the large dataset value which is collected from different global financial markets. Also, SVM does not give a problem of over fitting. Various machine learning based models are proposed for predicting the daily trend of Market stocks. Numerical results suggest the high efficiency. The practical trading models built upon our well-trained predictor. The model generates higher profit compared to the selected benchmarks.

REFERENCES

[1] Zhen Hu, Jibe Zhu, and Ken Tse "Stocks Market Prediction Using Support Vector Machine", 6th International Conference on Information Management, Innovation Management and Industrial Engineering, 2013.M.

[2] Wei Huang, Yoshiteru Nakamori, Shou-Yang Wang, "Forecasting stock market movement direction with support vector machine", Computers & Operations Research, Volume 32, Issue 10, October 2005, Pages 2513–2522.

[3] N. Ancona, Classification Properties of Support Vector Machines for Regression, Technical Report, RIIESI/CNR-Nr. 02/99.

[4] K. jae Kim, "Financial time series forecasting using support vector machines," Neurocomputing, vol. 55, 2003.

[5] Debashish Das and Mohammad shorif uddin data mining and neural network techniques in stock market prediction: a methodological review, international journal of artificial intelligence & applications, vol.4, no.1, January 2013