IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-7, Issue-4, April 2019 E-ISSN: 2347-2693

Estimating Efficiency of Support Vector Machine Based Model in Prediction of The Direction of Future Stock Price During Different Trends of Market

Shailja Kashyap

Department of Computer Science & Engineering, Assam Don Bosco University , Assam, India

Corresponding Author: shailjaswati@gmil.com, Tel.:9461059381

DOI: https://doi.org/10.26438/ijcse/v7i4.259262 | Available online at: www.ijcseonline.org

Accepted: 15/Apr/2019, Published: 30/Apr/2019

***Abstract*—** The main objective of this research is to find the efficiency of prediction model based on support vector machine for predicting the stocks in NIFTY 50 during different trends of the market in last 10 years. The prediction model takes different features and predict the next day price as up or down as compared to previous day price. The features used in the model are difference of current day and previous day low-high, open, close, and moving average of open, close, high and low. Label is the difference between open and close prices of the current day. It is observed that the implemented support vector machine algorithm performs very well in predicting the stock price when there is drop in price , irrespective of market trend , but performance reduces significantly in predicting the stock price when there is increase in price ,when market follows upward trend.

***Keywords*—**support vector machine, stock , market trend, prediction , efficiency .

I. Introduction

Number of people involved in trading /investing in stocks has increased significantly, in last few decades ,which includes both professionals and non-professionals . The financial market is a complex, evolutionary, and non-linear dynamical system [1]. It has become crucial to be accurately predict the stock prices to avoid risks.

Many researchers have been studying and have introduced stock market forecasting methods based on machine learning techniques, where ANN and SVM models have been the most widely discussed , compared and used. It has been stated that the SVM gives the best performance for financial time series data.[2,3]

This motivated us to build a SVM based prediction model and check its predictability in different trends of market .In this paper I have introduced a SVM based model and check the accuracy of the model in predicting the increasing stock prices (higher than previous day) when market was following downward trend and the decreasing stock prices (lower than previous day) when market was following upward trend.

I have organized this paper in following sections , section 1 is the introduction to the proposed work, section 2 explains the theory of the algorithm used and experiment design . Section 3 contains the results and discussions of the research .Section 4 talks about the conclusion and future work .

II. Related Work

Bruno Miranda Henrique, Vinicius Amorim Sobreiro , Herbert Kimura ,[4], suggest that SVR has good predictive power especially when using strategy of uploading the model periodically .Results of their experiments also show an increased prediction precision during lower volatility periods . Malkiel BG, Fama EF. ,in[5] claims that it is impossible to generate profits in long terms using prediction models for stock markets .The theory have been challenged by many researchers since then , Nayak RK, Mishra D, Rath AK ,in [6] used hybrid model of SVM -KNN to have better prediction capability . Tay FE, Cao L. in [7] examines the feasibility time series forecasting by comparing it with a multi-layer back propagation .As per the results of experiment SVM was found to be more advantageous to forecast financial time series .

III Methodology

A. SVM theory

Support vector machine is a supervised classification and regression technique .

In classification target is to find the optimal hyper plane such that it efficiently separates the samples from different classes in space . If we take vector w , as the vector normal / perpendicular to the hyper plane , with unknown |w| . we can define the hyperplane as

$$w^T . u^T = c$$

, for some constant c

or

$$w^T . u^T + b = 0 \text{ , for } b= -c \qquad (1)$$

u^T is the vector of any input sample . The projection of u^T on w^T , decides the class to which the input sample belongs .

In this experiment we try to map the input samples , based on features discussed below, to two classes that is $y_i=\{0,1\}$, where 0 denotes fall of stock price , and 1 denotes rise in stock price , as compared to previous day price. Basically {0,1} are the labels .

Considering (1) , we can say that

for all the samples belonging to class 1

$$w^{T.} .x^T{}_1 + b \geq 1 \qquad (2)$$

for all the samples belonging to class 0

$$w^{T.} .x^T{}_0 + b \geq -1 \qquad (3)$$

Samples closest to the hyper plane are known as support vectors

For obtaining optimal hyper plane , we try to maximize the margin between the two classes or say we try to minimize the norm of vector w , denoted by $|w|$.

Let $x^T{}_0$ and $x^T{}_1$ the support vectors or gutter points in the space .

then the width or the margin between the two classes can be given by margin ,

$$M = (x^T{}_0 - x^T{}_1) \cdot w^T \div |w| \qquad (4)$$

which is equal to

$$M = 2 \div |w| \qquad (5)$$

using Lagrange multipliers ,

we get

$$\min L(|w|) = \frac{1}{2}|w|^2 \qquad (6)$$

subject to, $y_i(w^{T.} .x^T{}_i + b) \geq 1$ for all i,

where y_i represents the all of the labels of the training samples.

For non-linear classification , we replace $x^T{}_i$ with a transformation function , known as kernel function so the equation becomes

$$y_i(w^{T.} .\varphi_i + b) \geq 1 \qquad (7)$$

there are three different kernals that we can use and are defined as

RBF

$$K(a,b) = \exp(-(a-b)^2 / 2\sigma^2) \qquad (8)$$

Polynomial

$$K(a,b) = (1 + \Sigma a_i b_i)^d \qquad (9)$$

Sigmoid

$$K(x,y) = \tanh(\alpha x^T y + h) \qquad (10)$$

In our experiment we have used RBF Kernel.

B. Experiment design

1) Data Set

The model is applied on data of NIFTY 50 collected from NSE . We have taken the stocks data from Jan 2008 to July 2018 for our experiment.

2) Features and label

Using SVM based prediction model we predict the next day price based on last 90 days. The features used in the model are difference of current day and previous day low-high, open, close, and moving average of open, close, high and low. Label is the difference between open and close prices of the current day If the difference is less than 0 , it shows that the stock price has closed at a lower price than opening price in morning , and we denote it by 0 label and if the difference is greater than 0 , we denote it by 1 label.

Table 1 Feature Description

Feature	Description	Formula
Open price difference	It is the difference between previous day and current day opening price	$(Open_i - Open_{i-1})$,where i denotes the current day
Close price difference	It is the difference between previous day and current day closing price	$(Close_i - Close_{i-1})$,where i denotes the current day
Low price difference	It is the difference between previous day and current day lowest price	$(Low_i - Low_{i-1})$,where i denotes the current day
High Price Difference	It is the difference between previous day and current day opening price	$(Open_i - Open_{i-1})$,where i denotes the current day
Open price Moving average	MA_{t+1} is defined as the moving average with the n time length is calculated at time t+1M	$MA_{t+1} = MA_t - open_{t-n+1}/n + open_t/n$ $open_{t-n+1}$ and $open_t$ are the observed values of stock opening price at time t.
Close Price moving average	MA_{t+1} is defined as the moving	$MA_{t+1} = MA_t - close_{t-n+1}/n +$

	average with the n time length is calculated at time t+1M	$close_t$/n $close_{t-n+1}$ and $close_t$ are the observed values of stock closing price at time t.
Low Price moving average	MA_{t+1} is defined as the moving average with the n time length is calculated at time t+1M	$MA_{t+1} = MA_t - low_{t-n+1}/n + low_t/n$ low_{t-n+1} and low_t are the observed values of lowest stock price at time t.
High price moving average	MA_{t+1} is defined as the moving average with the n time length is calculated at time t+1M	$MA_{t+1} = MA_t - high_{t-n+1}/n + high_t/n$ $high_{t-n+1}$ and $high_t$ are the observed values of highest stock price at time t.

3) Training and Testing data

I have used the last 90 days(from prediction day) data to train SVM prediction model, to predict the next day price. Model is tested for each day in data set except first 90 days data.

III. Results and Discussion

Table 2. shows the percentage accuracy of stock price prediction for 10 selected sticks of NIFTY 50. The tables shows accuracy of SVM based algorithm in two trends of market ,one in which the market was showing an upward trend and other in which the market was showing an downward trend.

In this experiment it was observed that the model shows high accuracy in predicting the prices when stock price drops irrespective of the market trend . Whereas accuracy of model reduces to an average of 0.70 ,when the stock prices increases , in case of upward trend of market . The accuracy further decreases to an average of 0.67 when the stock prices increase ,in case of downward trend of market .

Table 2. Model Accuracy

Stock	**Upward Trend**		**Downward Trend**	
	UP (increase in stock price) ACCURACY	**DOWN (drop in stock price)ACCURACY**	**UP (increase in stock price) ACCURACY**	**DOWN (drop in stock price) ACCURACY**
GAIL	0.70	0.93	0.66	0.93
BPCL	0.73	0.92	0.69	0.96
GRASIM	0.83	0.87	0.69	0.87
SBIN	0.72	0.93	0.62	0.95
TATASTEEL	0.85	0.95	0.84	0.92
YESBANK	0.80	0.92	0.018	0.99
RELIANCE	0.79	0.91	0.71	0.93
ITC	0.78	0.76	0.61	0.94
TECHM	0.83	0.91	0.58	0.96

IV. Conclusion and Future Scope

Based on our experiment results it is found that the svm based model shows higher accuracy in predicting stock prices during downward trend of market for both drop and increase of stock price as compared to when there is upward trend. The results suggest that the implemented model is downward biased and it is not efficient to use it for trading in real life but the prediction accuracy is high hence if we include more features from advance technical analysis indicators or combine different machine learning models then it can be improved to use in real world trading.

References

[1] Abu-Mostafa YS, Atiya AF. "Introduction to financial forecasting". Applied Intelligence 1996;6:pp205–13.

[2] Puchong Praekhaow,"Determination of Trading Points using the Moving Average Methods". International Conference for a Sustainable Greater Mekong Subregion,August 2010

[3] 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, pp 2513–2522.

[4] Bruno Miranda Henrique, Vinicius Amorim Sobreiro*, Herbert Kimura "Stock price prediction using support vector regression on daily and up to the minute prices" . The Journal of Finance and Data Science 4 (2018) pp 183e201

[5] Malkiel BG, Fama EF," Efficient capital markets: a review of theory and empirical work". J Finance. 1970;25(2):pp383e417

[6] Nayak RK, Mishra D, Rath AK. A Naïve SVM-KNN based stock market trend reversal analysis for Indian benchmark indices. Appl Soft Comput. 2015;35(1):pp670e680.

[7] Tay FE, Cao L. Application of support vector machines in financial time series forecasting. Omega. 2001;29(4):pp309e317

[8] Qu H, Zhang Y. A new kernel of support vector regression for forecasting high-frequency stock returns. Math Probl Eng. 2016;2016(1):pp1e9

[9] Ping-Feng Pai∗, Chih-Sheng Lin," A hybrid ARIMA and support vector machines model in stock price forecasting".2004 Omega, 33(6), pp 497–505.

[10] Zhang, Z., Shi, C., Zhang, S., & Shi, Z. (2006). "Stock Time Series Forecasting Using Support Vector Machines Employing Analyst", 2006 ,Springer ,pp 452–457.

Authors Profile

Ms. Shailja Kashyap pursed Bachelor of Technology from Rajasthan Technical University, in 2012 and Master of Technology from Maharshi Dayananad University in year 2017. She is currently working as Assistant Professor in Department of Computer Science and Engineering, Assam Don Bosco University , Assam since 2017. She is a member of Institution of Engineers(India). She has 5 years of teaching experience and 2 years of Research Experience. Her main areas of interest are Data analysis , Machine Learning and IOT.

