Learning of SVM Centered Indian Stock Market Forecast Procedures

[1]Nitin Rameshrao Talhar, [2]Anwesha Tomar, [3]Tapasya Ghorpade, [4]Namrata Sakore, [5]Mayuri Zile

[1]Assistant Professor, [2,3,4,5]Students
AISSMS COE, COMPUTER DEPARTMENT

***Abstract*: Investing currency has never been a risk-free procedure. Many prototypes have been planned for the forecast of stock market revenues. In this survey paper, we extant an investigation of the numerous works done in the field of sustenance vector apparatuses for the forecast of stock market earnings. Exactitudes of various approaches are investigated and the best accomplishment model is preferred. We then present our planned model, explicate its organizations and possibility. The numerous variables, dataset and their impression on the correctness of the calculation are elucidated for each model. Thus serving depositors to select their desired model for forecast.**

***Keywords*: Stock Market, Support Vector Machines, forecast, Calculation, Indian Stock**

I. INTRODUCTION

The stock market is an intricate and volatile scheme [1]. Stock market is affected by many factors such as political events, economic circumstances and investors soppiness [10]. This field of financial forecasting is characterized by data intensity, noise, non-stationarity, unstructured nature, hidden relationships and a high degree of uncertainty [1],[2].Therefore, forecasting in the field of finance is considered to be very difficult. Typically financial fore-casting using support vector machines involves two main methods: feature selection and constructing the prediction model [1].Selecting the right features is just as important as selecting the right model for the prediction of the stock market. Various researches have been conducted into finding the right model for the accurate prediction of the stock market. Multiple models have been developed using Support Vector Machines, Relevant Vector Machines, Deep Learning and Artificial Neural Networks [1],[3],[5],[6]. Although various techniques have been used for the prediction of stock market Support Vector Machines has gained popularity and has outperformed Artificial Neural Network [10].

II. RELATED WORKS

A. Zhen Hu, Jie Zhu, and Ken Tse, Stocks Market Prediction Using Support Vector Machine:

It uses SVM for prediction of the stock market, six macroeconomic and four company-specific factors are used for prediction. Fifteen random companies were selected, for all years from 1995-200. Ten values were used in testing data while 78 values were used in training data. Each stock is labelled as good or bad to specify if it is a good or bad investment. It also makes use of six macroeconomic variables and four company specific variables for the purpose of prediction .The accuracy was found to be 96.15 %.

B. Yuling LIN, Haixiang GUO and Jinglu HU, An SVM-based Approach for Stock Market Trend Prediction:

The approach used consists of two parts: feature selection and predictive model. A quasi-linear SVM is used for the prediction of stock market trends. This model fixes three major problems. The first problem is of overfitting, most SVM models suffer from this problem when flexible non-linear kenels are used in high-dimensional and high-noise data sets. The second problem is in feature selection, even though feature selection helps reduce noise they have their limitations. If we select a feature subset that is highly correlated with the output but are unrelated to each other this can lead to problems. The third problem is most models ignore evaluation features, and even the ones that do consider it are effected by a subjective factor and multicollinearity.

C. Vivek Rajput and Sarika Bobde, Stock Market Prediction Using Hybrid Approach:

The proposed method is completely different from all the past methodologies and makes use of two different techniques in combination for the prediction of the stock market. The model uses sentiment analysis for extracting words from the available text and then finds words that are related to the particular field and tests the influence on the clusters that are formed using the second technique that is the clustering technique. Two models are developed namely the sentiment analysis model and clustering model. The output that is produced by the sentiment analysis model is then given as an input to the second model that is the clustering model. The combination of these model gives greater accuracy.

D. Honghai Yu and Haifei Liu, Improved Stock Market Pre-diction by Combining Support Vector Machine and Empirical Mode Decomposition:

The model uses a two-stage neural network architecture by combining support vector machines and empirical mode decomposition. The basic principle of EMD is to decompose a time series into a sum of intrinsic mode functions(IMFs).SVM is used to predict

different IMFs. The data is taken from Shanghai Composite Index of Wind database period from November 27, 2008 to February 22,2010. The data-set included one and a half years worth of daily data. Error analysis is done to find the accuracy of the model, 0.10%-1.91% is the percentage error for the model.

E. Fangqiong Luo, Jiansheng Wu and Kesong Yan, A Novel Nonlinear Combination Model Based on Support Vector Machine for Stock Market Prediction:

The author suggests that the change of the stock mar-ket is more disorder, it is hardly exactly speaking that it is merely linear or nonlinear system. Combination model based on Support Vector Macine regression combining the linear regression of traditional statistical model with the nonlinear regression of Neural Network model. Using 400 training samples and 40 testing samples from Shanghai Stock Exchange, the two models SVM-CM and LR-CM(linear regression combination) are tested. The result of study was that SVM-CM gives accuracy of 75% while LR-CM gives 52%.

III. LIMITATIONS OF EXISTING MODEL

Every model that uses support vector machines suffers from the problem of over-fitting, this problem has been han-dled by using quasi-linear support vector machines. But this model suffers from the disadvantage of selecting parameters that may not always be of optimal value or produce the best results. The hybrid model that uses sentiment analysis and clustering has the limitation of not using enough variables for the evaluation or prediction of the stock market. When a support vector machine is used without combining it with any other model it gives only 70% accuracy on testing data. Thus we propose a model which is a combination of support vector machines and decision trees, this model will not only assist in the prediction of the Indian stock market but will also aid in the analysis of risk management by using decision trees.

IV. PROPOSED METHOD

The proposed model is a combination of support vector machines and decision trees. The support vector machines are used for the prediction of the future value of the Indian stock market, while decision trees are used for the analysis of this prediction. The decision tree will take the output of support vector machines as input and give the output as either risky or not risky to symbolize whether or not the investment in that stock is risky or risk free.

A. SUPPORT VECTOR MACHINES

Since the stock market is complex, uncertain and diffi-cult to predict it is hardly a liner or non-linear system[4]. SVM algorithm developed by Vapnik is based on statistical learning theory[2]. It can be used for both classification and regression tasks[2]. The use of the linear model to implement the nonlinear class boundaries in which the vector is fed into the high dimensional feature space[3]. SVM is an algorithm that is known as the algorithm that finds a special kind of linear model, it has maximum margin hyper plane[3]. Therefore, SVM can be used for accurate prediction of the stock market.

ADVANTAGES:: Support vector machine training involves optimization of a convex function with linear con-straint, it also has a unique global minimum which in turn overcome strucking to local minima observed in a neural network, which reduces the computational cost[3].

1) DISADVANTAGES:: SVM is more likely to avoid the problem of falling into local minimum[3]. Support vector machine accuracy mostly ranges from 59.35% and 71.43% only[3].

B. DECISION TREES

Decision trees may be defined as a directed tree with a node which has no incoming edges, each internal node tests an attribute whereas each branch corresponds to attribute value[7]. In our model, we will train the decision tree to find out if the investment in that particular stock is risky or not. The value predicted by SVM will be used as an input value for the decision tree and then the output of the decision tree will be used to understand if the investment is risky or not.

C. DATA SET

The data set is obtained from the Bombay Stock Ex-change(BSE), five different sectors are decided and five companies are selected per sector. Seventy percent of this data is used for training purposes and the other thirty percent data is used for testing purposes. Then this output produced from the SVM is used to feed it into the decision tree.

D. FACTORS

In order to obtain accurate results, the right factors are to be used, in our model, we will attempt to use four company-specific factors and six macroeconomic factors. While selecting factors it is important to pick ones that will directly contribute to the prediction of the future value of the stock, but we also need to ensure that these factors do not cause overfitting.

1) Company Specific factors:

Net revenue Net income
Price per earning ratio of stock Diluted earnings per share

2) Macroeconomic factors:

Consumer earnings Customer investment Unemployment rate Inflation rate

Federal funds

Dow Jones industrial average

V. CONCLUSIONS

Support vector machines can be used for finding out or predicting the future value of stocks but the prediction is never accurate. In all the papers that have been analyzed support vector machines is used in combination with some other algorithm namely, sentiment analysis, neural networks and clustering. But there are still factors in the above methods that cause inaccuracies, in an attempt to eliminate such inaccuracies we recommend the use of of a hybrid model that uses support vector machines in order to perform fitting and analysis of data for the purpose of predicting the future values of the stock market, and the use of decision trees for predicting if investing in such a stock is risky or risk free. This model will have a lower error rate ratio since we will be combining two different classification models.

REFERENCES

[1] Yuling LIN, Haixiang GUO and Jinglu HU, ”An SVM-based Ap-proach for Stock Market Trend Predict”,2013.

[2] Zhen Hu, Jie Zhu, and Ken Tse, ”Stocks Market Prediction Using Support Vector Machine”, 2013.

[3] M. Sreemalli, P.Chaitanya, K. Srinivas,”Comparative Analysis of Machine Learning Techniques on Stock Market Prediction”, 2016.

[4] Fangqiong Luo, Jiansheng Wu and Kesong Yan,”A Novel Nonlinear Combination Model Based on Support Vector Machine for Stock Market Prediction”, 2010.

[5] Masaya Abe and Hideki Nakayama,”Deep Learning for Forecasting Stock Returns in the Cross-Section”, 2018.

[6] Karthik HS, Nishanth VA and Manikandan J,”Stock Market Prediction using Optimum Threshold based Relevance Vector Machines”, 2016.

[7] S.S.Panigrahi and Dr.J.K.Mantri, ”A Text Based Decision Tree Model for Stock Market Forecasting”, 2015.

[8] Keerti. S.Mahajan, Dr. R .V. Kulkarni,”Stock Market Prediction and Investment Portfolio Selection Using Computational Approach”, 2015.

[9] Vivek Rajput and Sarika Bobde,”Stock Market Prediction Using Hy-brid Approach”, 2016.

[10] Honghai Yu and Haifei Liu,”Improved Stock Market Prediction by Combining Support Vector Machine and Empirical Mode Decompo-sition”, 2012