Decision Support System for Investment in Stock Market using OAA-SVM

Himanshu Shrimalve, Shreya Sulakhe, Madhavi Sontakke, Mayuri Thakare and Sopan Talekar[*]

NDMVPS's KBT College of Engineering, Department of Computer Engineering Nashik - 422013, Maharashtra, India; talekar.sopan@kbtcoe.org

Abstract

The investment in stock market could be high-risk. Stock price depends on several factors, such as currency exchange rate, economics situation, and flow of funds. In stock market, depending on a stock selection and a suitable time on trading successful investors can earn maximum profits. Because of the financial crisis and scoring profits, it is mandatory to have a secure prediction of the values of the stocks. Generally, investors use two Statistical techniques for making a decision,which are the fundamental analysis and the technical analysis and many no of machine learning models have been investigated for stock prediction such as Genetic Algorithm (GA), Support Vector Machine (SVM) and Neural Network(NN). In proposed system, we have used the binary tree SVM algorithm which is one of the mainstream algorithms for multi-class classification in the fields of pattern recognition and machine learning. In order to reduce the training and testing time of system, One-Against-All SVM (OAA-SVM) algorithms will be proposed for multi-class classification.

Keywords: Hyper-Plane, Optimization, EMA, MACD, OAA, RSI, SVM

1. Introduction

The Indian stock market is considered to be one of the earliest in Asia, which is in operation since 1875. However, it remained largely outside the global integration process until 1991. Stock market prediction is the act of trying to determine the future value of a company stock or other financial instruments traded on an exchange. The successful prediction of a stock's future price could yield significant profit. The efficient market hypothesis suggests that stock prices reflect all currently available information and any price changes that are not based on newly revealed information thus are inherently unpredictable. We are using Technical analysis technique to enhance investors for trading, it is a statistical technique using an opening price, a closing price and a volume in each day for stock trading. This technique can provide a trading signal (buying or selling) to investors. Furthermore, technical analysis is able to generate technical indicators in order to understand trends of stock price and the momentum of buying and selling a stock, which is represented by a stock chart. Artificial Intelligence (AI) which is a technique to increase intelligence on the information system has been developed continuously. Machine Learning (ML), which is an AI tool, can be employed to learn and recognize related data patterns on the classification problems. It can be also applied for solving problems in a stock market. Currently, there are several types of ML model used such as Support Vector Machines (SVM). The model has used the preprocessed data set of closing value of NSE and BSE India First, we will take dataset from NSE & BSE India Official website which contains historical stock informa-

---

tion of various industries. Then by using real time values of stocks and News the algorithm i.e. SVM will generate the correct predicted values and graph or chart which will be prediction of stock of specific company and helping Investors for making decision[1]. We are using ML technique because it will automatically predict the results that company shares are good to buy/sell or not. So that customer will get basic idea related to company's stock value. Customer will see prediction and will decide to invest his money into specific company.

2.   Related Work

In day to day life many peoples make invest in stocks, but some people have not get sufficient knowledge about stock due to which people have lost their money in stocks. There are many myths related about investing in stock. To give people sufficient knowledge and analysis of stock many researchers had contributed their knowledge and came up with the solution.

Radu Iacomin, proposed a system "Stock Market Prediction" which   uses various machine learning algorithm such as Support Vector Machine with feature selection algorithm but this technique will be suitable for only single class classification[2].

Sabaithip Boonpeng, Piyasak Jeatrakul proposed a system in which OAA-Neural Network is used for classifying the stock data into 3 different classes such as, buying data, selling data and holding data. However there are several limitations that, training time required for neural network are more, it required large memory to store the data and neural network not suitable for large dataset[3].

Binoy B. Nair, *et al.* proposed a system using hybrid Decision tree and Neuro Fuzzy which has required large searching time and memory.[4]

QIU Mingyue, *et al.* proposed a system using hybrid Genetic Algorithm and Artificial Neural Network which has high convergence rate.[5]

To overcame the drawbacks of all above system we have proposed OAA-SVM technique invented by Xiaowei Yang, *et al.*[6] It will be useful for classifying various dataset. For development of our system that classify data in 3 classes such as, buying data, selling data and holding data. As we have combine Support Vector Machine with One-Against-All multiclass classification algorithm, it gives an optimal solution with less training and testing time and also requires less space to store the data. It is summarized in the Table 1.

Table 1.   Related work

Title	Year	Technology Used	Limitations
Stock Market Prediction	IEEE (2016)	SupportVector Machine	Single Class classification technique
Decision Support System for Investing in Stock Market by using OAA-Neural Network	2016	OAA-Neural Network	More training time required.Large memory to store data.Not suitable for large datasets.
Multi-class pattern classification using neural networks	2006	Artificial Neural Network	Ambiguityand uncovered space regions.
A Stock market trend prediction system using a hybrid decision tree-neuro-fuzzy system	IEEE (2016)	HybridDecision tree andNeuro-Fuzzy	Large searching time and large Memory required.
Application of the Artificial Neural Network in predicting the direction of stock market index	IEEE (2016)	Hybrid Genetic Algorithm and Artificial Neural Network	High convergence rate

3. Proposed Framework

Architecture of proposed system is shown in figure 1. It is prepared by using two years data, retrieved from NSE and BSE India. Value of technical indicators such as RSI, MACD, SOT is determined by preprocessing data. The data is divided into training and testing datasets. After the training database, data is extracted and then it is normalized to create input dataset. Using the cross-validation technique, the datasets (training and testing) goes to the SVM model. This process is performed by the training phase and the remaining part is for testing phase to increase the performance. Together the training and testing phase, SVM plays an important role to improve the accuracy followed by the prediction. After this process, the final output is displayed in the graphical form or spread-sheet form.
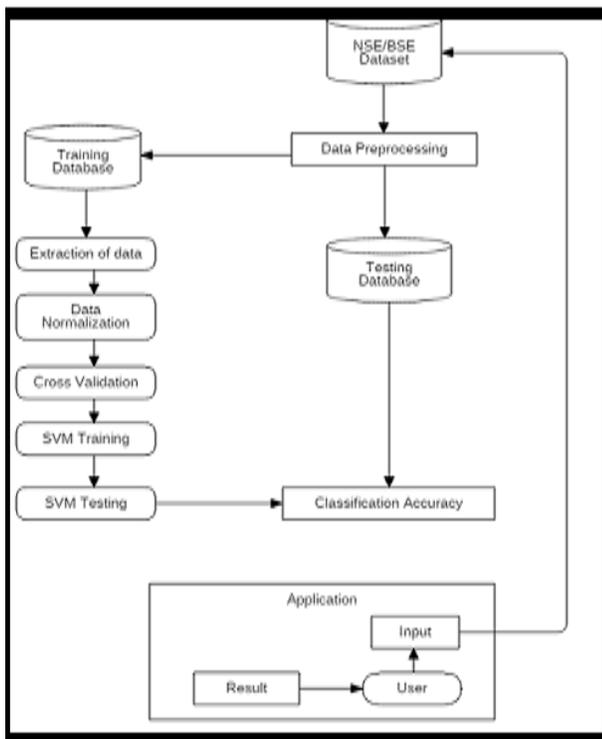
Figure 1. Architecture of system.

The proposed system will be divided into several independent modules as follows:

- **Prediction System Module**
This module fetches the input data/stock from NSE/BSE India according to user requirements. The data is nor-

malized using technical indicators. After that data will be divided into training and testing data for better accuracy, using cross-validation technique. Training dataset is supplied to algorithm (OAA-SVM) for classifying and predicting the stock values (i. e Buying, Holding and Selling class).

- **Graphical or Spreadsheet Representation Module**
After Prediction of classes of stock, data is displayed on spreadsheet or in graphical format for better user interface and understanding.

- **News Monitoring Module**
Daily news updates related to stocks are fetched by new monitoring module. This module is useful for user to get an idea for future condition of stock.

- **Analysis Module**
In analysis module classified data is supplied to OAA-SVM testing. Investment in the company stock is predicted by OAA-SVM testing.

- **Data Storing Module**
Historical data is stored in data storing module and news is accessed by user.

4. Algorithmic Solution

Methodology for development of our proposed system is as follows:

Step 1: Fetch data of stock from NSE of last 2 years.
Step 2: Calculate the value of Technical indicators such as, RSI, MACD, SOT etc.
Step 3: Train model using these indicators and training dataset (Using OAA-SVM Algorithm).
Step 4: Test the model using testing dataset (Using OAA- SVM Algorithm).
Step 5: Evaluate system using various evaluation techniques.

4.1 Technical Indicators

4.1.1 Relative Strength Index (RSI)

RSI indicates trading signal based on the relationship of stock price in the current day and stock price in the past period. RSI value can be calculated by the following for mula:

$$RSI = 100 - [100 / 1 + RS[3] \qquad (1)$$

$$RS = \textit{Average Of } 14 \textit{ days closes up / Average Of } 14 \textit{ days closes up} \qquad (2)$$

4.1.2 Stochastic Oscillatory Technique (SOT)

Stochastic Oscillator technique compares the current close price of a stock with its price range over a period in the past. It uses two lines for generating trading signal, which are %K line and %D line. These lines are calculated by the following formula:

$$K\% \textit{ line} = 100 - (Receentcloses - LowestLow(n)) / (HighestHigh(n) - LowestLow) \qquad (3)$$

$$D\% \textit{ line} = 3 \textit{ period moving average of } \% K \textit{ line } [3] \qquad (4)$$

Stochastic technique indicates trading signals when % K line and %D line is across each other above 80 or below 20. The selling signal is generated when %K line and %D line across each other above 80 while the buying signal is indicated when %K line and %D line across each other below 20.

4.1.3 Moving Average Convergence and Divergence (MACD)

MACD technique is proposed by Gerald Appel. This technique can indicate trading signals (buying and selling) and also the trends of a stock price (upward trend, downward trend and sideway trend). MACD technique indicates trading signal based on the difference of two moving average lines. These are 12 days and 26 days. MACD value is calculated by the following formula.

$$MACD \textit{ Line} = EMA(12) - EMA(26) [3] \qquad (5)$$

MACD technique indicates trading. When MACD crosses over zero line, the buying signal is indicated and the trend of stock price is starting an upward trend. While MACD crosses under zero line, the selling signal is generated and the downward trend is beginning

4.2 Classification Algorithm

4.2.1 Support Vector Machine (SVM)

Support Vector Machine (SVM) is Supervised Learning technique of Machine Learning (ML). In SVM, support vectors are shown in figure 2. It analyzes the data and also recognize the patterns, which are used for regression and classification analysis. SVM training algorithm builds a model that assigns new examples to one category or the other, making it a non-probabilistic binary linear classifier. An SVM model is the representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall.

SVM builds hyper-plane or set of hyperplanes, which can be used for regression and classification as shown in figure 3. Good separation is done by hyperplane that also has largest distance, if margin is higher generalization of error of classifier is lower.
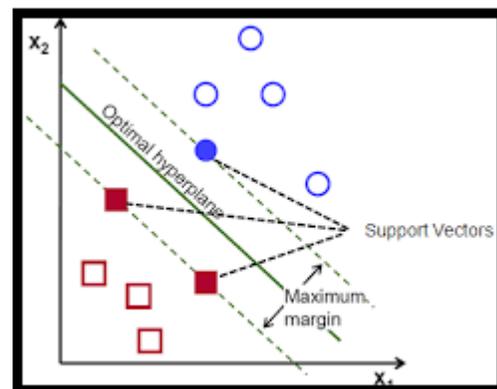

Figure 2. Support Vector Machine.

Mathematical Equations

- For 2D dataset let,

D is the dataset containing $\{x_i, y_i\}$ no of tuples where,

x_i = Set of attributes/tuples

y_i = Set of classes

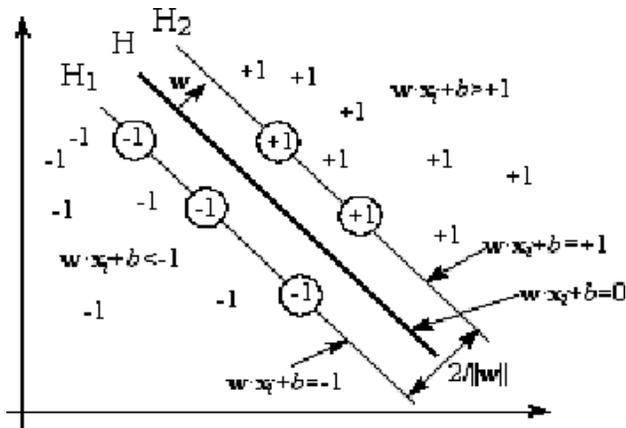As, dataset is 2D, y_i = +1/–1

- Equation of hyperplane:

$$w.x + b = 0 \qquad (6)$$

Where,

w = Weight vector (Distance of attribute from hyperplane)

x = No of Attribute

b = Scalar/ Bias

Figure 3. **Maximum Margin of Hyperplane**

- Equation of attributes lies on hyperplane:

$$W0 + W1.X1 + W2. X2 = 0 \qquad (7)$$

- Equation of attributes lies above the hyperplane:

$$H1: W0 + W1.X1 + W2. X2 \geq 0 \qquad (8)$$

- Equation of attributes lies below the hyperplane:

$$H2: W0 + W1.X1 + W2. X2 \leq -1 \qquad (9)$$

- Combining equation of attribute above hyperplane (H1) and below the hyperplane (H2), we have obtained following equation:

$$Yi(W0 + W1.X1 + W2. X2) \, \forall I \qquad (10)$$

- Maximum Marginal Hyperplane (MMH) given by,

$$\frac{2}{\|w\|} \qquad (11)$$

where $\|w\|$ *Eucledian norm of W*

ii. One Against All – Support Vector Machine (OAA-SVM)

OAA-SVM algorithm is useful to divide the dataset into 'n' no of classes. The conceptual map of OAA & SVM as shown in figure 4. This algorithm overcomes the drawback of SVM, which only suitable to divide data into 2 classes. OAO-SVM needs to construct C(C-1)/2 SVM models where each one is trained on data points from two classes. For the training data points from the ith class and the jth class, we need to solve the following binary classification problem[6].
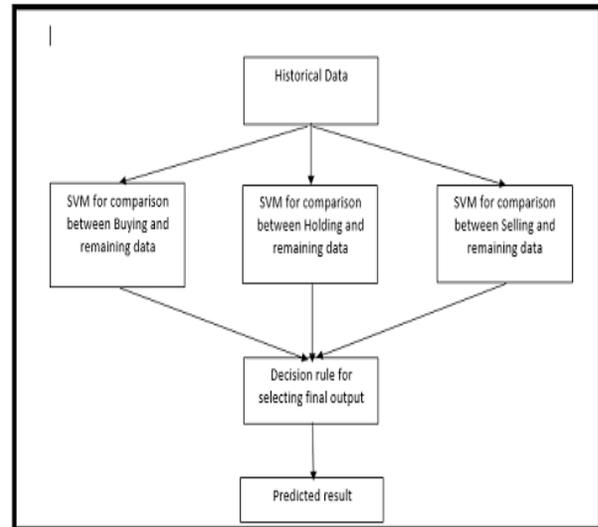
Figure 4. Conceptual map of OAA-SVM.

$$Min \, j(wij, bij, \zeta) = \frac{1}{2}(wij)T.wij + \gamma \sum_{i=1}^{l} \zeta tij \qquad (12)$$

S.T

$$(wij)T \, \Phi(Xt) + bij \geq 1 - \sum_{t}^{ij} Yt = i \qquad (13)$$

$$(wij)T \, \Phi(Xt) + bij \geq 1 + \sum_{t}^{ij} Yt = i \qquad (14)$$

Where,

$$\sum_{t}^{ij} Yt \geq 0, t = 1........l$$

5. Results

The predicted result using OAA-SVM model for the daily analysis of stocks is shown in Table 2. As shown in figure 5. the analysis is represend using graph.

Table 2.   Analysis of Daily Stocks

symbol	Date	Opening price	Closing price	Average price	Actual result	Predicted result
Ashoka	11-sep-2017	188.95	196.35	196.89	Up	Down
Ashoka	12-sep-2017	197.7	199.95	200.42	Up	Down
Ashoka	13-sep-2017	200	198	199.88	Down	Down
Ashoka	14-sep-2017	198.3	198.9	197.57	Up	Up
Ashoka	15-sep-2017	198	197.85	200.64	Down	Down
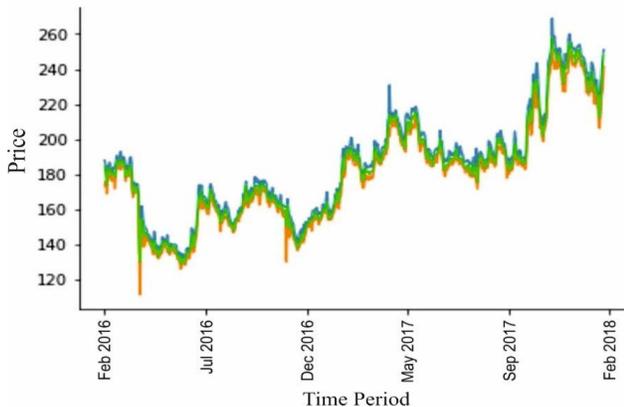
Figure 5.   Analysis of stocks by graph.

Graph shows the daily analysis of stocks where Blue, Green, Orange line indicates opening price, Average price and Closing price respectively. And X-axis and Y-axis shows Dates on which stock is generated and Stock price obtained respectively.

6. Conclusion

The proposed system using OAA-SVM predicts the results which helps to the investors for making the decisions that company shares are good to buy/sell or not.System model developed in present work is useful forsuccessful prediction of a stock's future price that couldyield significant profit. The market hypothesis suggeststhat stock prices reflect all currently available informa-tion and any price changes that are not based on newlyrevealed information thus are inherently unpredictable.The present System model has been also used for predic-tion of direction of the movement of the closing value ofthe index. Purpose of this system is to help investor formaking best suitable decision according to stock valueand help them to suffer from money losses.

7. References

1. Available from: www.nseindia.com
2. RaduIacomin. Stock Market Prediction. Romani: Faculty of Automatic Control and Computers University POLITEHNICA of Bucharest, Romania. 2016. p. 200-205.
3. Sabaithip Boonpeng and Piyasak Jeatrakul. Decision Support System for Investing in Stock Market by using OAA-Neural Network School of Information Technology Mae Fah Luang University Muang, Chiang Rai, IEEE digital xplore , 11 April 2016.
4. Binoy B Nair, Mohana Dharini N, Mohandas VP. A Stock market trend prediction system using a hybrid decision tree-neuro-fuzzy system. Coimbatore - 641105, Tamilnadu, India: Department of Electronics and Communication Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham. IEEE xplore. Oct 2010.p.16-17
5. Qiu Mingyue, Li Cheng, Song Yu. Application of the Artificial Neural Network in predicting the direction of stock market index. Fukuoka, Japan: Department of System Management, Fukuoka Institute of Technology.
6. Xiaowei Yang, Qiaozhen Yu, Lifang He, Teng Jiao Guo. The one-against-all partition based binary tree support vector machine algorithms for multi-class classification. Guangzhou 510641, PR China: School of Computer Science and Engineering, South China University of Technology. August 2013: (113):1-7.
7. Guobin Ou, Yi Lu Murphey. Multi-class pattern classification using neural networks. Dearborn, USA: Department of Electrical and Computer Engineering, The University of Michigan-Dearborn. 2006