International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-8 Issue-4, April 2019

Stock Market Prediction using Machine Learning Concepts

Vidhyavani. A, Deepak Adithya. K. N., Sateesh N., Vignesh Kannan

Abstract: ***Stock market plays a vital role in deciding the Economy of the company as well as the country, since some part of GDP of the country depends upon MNC's present in the country. This paperwork is done in order to predict the future of the stock market using machine learning for accurate and profitable outcome. Stock market analysis is one of the toughest analysis since it not only depends upon the previous values of stock but also depends upon numerous other factors like sentiments of the company, projects undertaken by the company and many more. To predict stock market we used three algorithms. First is Least square support vector machine (LS-SVM) which will analyse the previous data of the stock market, second is Autoregressive moving average model (ARMA) for getting multiple predicted outcomes in terms of polynomial function, and the third is Particle swap optimization (PSO) for optimising the value which is obtained from Autoregressive moving average model. In accordance with the proposed system it is expected to get an accuracy of 25% - 30% in predicting the stock market. In future where earning money will be difficult but it still defines world's view we can use this project to earn money through stock exchange and MNC's. Investors can use this prediction to invest money on the Companies.***

Key terms – Stock market analysis, Least square support vector machine (LS-SVM), Autoregressive moving average model (ARMA), Particle swap optimization (PSO)

I. INTRODUCTION

Stock exchange play a major roles in the economic development of the country as well as the economic development of the company. The change and estimation of the stock market and its datais of clear interest to all shareholders in the market.According to the Efficient Market Hypothesis (EMH), the stock market depends upon various aspects like natural calamity, sentiments of the company, projects undertaken by the company, accidents in the company and thus theprediction naturally uses information from varies sources, which can be roughly categorized into quantitative data.It is tough to deal with qualitative data as they don't have any distinct structure and it is not conservative thus finding out important and useful data from that is difficult[1].

Manuscript published on 30 April 2019.
***** Correspondence Author (s)
Vidhyavani*, Assistant professor, Computer science department, SRM Institute of Science and Technology, Ramapuram, Chennai
Deepak Adithya.K.N., Batchelor of Technology in Computer Science and Engineering, SRM Institute of Science and Technology, Ramapuram, Chennai
Sateesh N, Batchelor of Technology in Computer Science and Engineering, SRM Institute of Science and Technology, Ramapuram, Chennai
Vignesh Kannan, Batchelor of Technology in Computer Science and Engineering, SRM Institute of Science and Technology, Ramapuram, Chennai

Many researchers from different fields pay more and more attention on it. With the development of computer science technology and software field, some methodologies based on data mining and machine learning have been introduced to predict the stock price movement automatically.

The major issue with the stock exchange prediction is that the data which we receive is highly changing and the change in data can be so drastic due to some minor factor. The prediction can never be more 60 % of accuracy due to the various factors like the other stocks in the market, the ups and down in the company expensive and major issue is, loss faced due to natural disaster and system malfunction in the industries. The solution to these problem is we predict more the one answer and solve it down to one answer of more accuracy. While predicting them we must take care of the special causes like natural disaster to maximise the predicted values accuracy. In this paperwork we predicted the trend of stock exchanges by using three different algorithms. We use last 10 years values and run them through three algorithm for getting the predicted value.

II. LITERATURE SURVEY

Stock market prediction is one of the toughest prediction. Stock market is very unstable in nature. This depends upon many factors such as projects undertaken by the company, sentiments of the company, rules and regulations of the company, company investments, social pressure, news and reports[6,8].Due to these reasons prediction of stock market becomes very difficult. There are many proposed systems which tried to predict the stock market with high accuracy, but most of the proposed systems failed to do so, at times prediction is accurate and at time prediction goes wrong totally. Hence it is very difficult to predict the stock market with accuracy all the time. Most of the time machine learning is used for predicting stock market, machine learning consist of many algorithms which when used at a right place can predict the stock up to an accuracy of approximately 50%, remaining 50% depends on the sentiments of the society[8]. Commonly, Regression is used to predict the stock market because regression minimises the data efficiently[1]. But the only disadvantage of regression is that it takes more time to optimise huge amount of data. Back propagation neural network is also used in some projects.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

Back propagation neural network is very slow because it requires small learning rates for stable learning[7].Artificial Neural Networks is also used in some places this is also inefficient because Artificial neural networks produces a probing solution that does not give a clue as to why and how. This reduces trust in the network.Relevance Vector Machines are also used in some systems. Multi-kernel learning based extreme learning machine is used in most of the places but the biggest disadvantage of this is that Choosing a good kernel function is not easy and it takes Long training time for large datasets[4].Principal component analysis is also used for this to be successful the data must be independent but in case of stock market most of the time data is dependent on its previous data.

III. PROPOSED METHODOLOGY

The proposed system contain three algorithms. All the three algorithms are widely used for prediction of the result and optimising the obtained result. Three algorithms are as follows:

A. Support-Vector Machine (SVM).
B. Autoregressive–moving-average model (ARMA).
C. Particle swarm optimization (PSO).

Stock exchange data is collected from various sources like National Stock Exchange (NSE), NASDAQ, London Stock Exchange (LSE), New York Stock Exchange (NYSE) and many more.

A. Support Vector Machine (SVM)

Machine learning has lot of algorithms but the right algorithms should be chosen at right place. The most important operation of Machine learning is Regression. This is the operation where all the data which is fed is optimised initially and this is one of the crucial part of the operation.

Regression is capable of limiting or optimising the data efficiently[3], but incapable of dealing with highly complex data. The only disadvantage of support vector machine is that large amount of data needs more time for optimisation this means that as the data feed increases the amount of time taken to optimise the data increase. We cannot easily overcome this problem.Support Vector Machine (SVM) uses machine learning algorithms to predict and classify the solutions[3]. This also uses machine learning to maximise the predictive accuracy. This also automatically avoids over fitting of data. Support Vector machines is a systems which use hypothesis space of a linear functions in a high dimensional feature space, this is trained with algorithms to optimise the linear solutions and predict the solution with high accuracy[2]. SVM performs better in term of not over generalization when the neural networks might end up over generalizing easily.

B. Autoregressive–moving-average model (ARMA).

Autoregressive–moving-average model is also one of the most used algorithm in machine learning. Based on the value fed this algorithm generates two polynomial one by using Auto regression and another by using moving average when we combine both of these obtained polynomial we get a complex polynomial which has to be optimised to get the desired result.An autoregressive integrated moving average model is a algorithm which uses regressive analysis of data whose strength of dependent variable is relative to the changing factors around the variable for more accurate future prediction of the values[7]. These value will be in both real and complex we use them we bring the accurate prediction[2].The reason behind choosing this algorithm is that it creates a polynomial and polynomial results are easy to solve and the time complexity and the space complexity of the polynomial is less compared to others. This algorithm states that the value which is obtained is dependent on the current and the previous value[2].

The process **{xt;t∈ Z}** is an autoregressive moving average process of order **(p, q)**, denoted with

$$xt \sim ARMA(p, q),$$

if,

$$x_t - \varphi 1 x_{t-1} - .. - \varphi p x_{t-p} = ut + \theta 1 u_{t-1} + .. + \theta_q u_{t-q} \forall t \in Z,$$

where, $u_t \sim WN(0, \sigma^2_u)$, and $\varphi_1, ..., \varphi_p, \theta_1, ..., \theta_q$ are **p + q** constants and the polynomials

$$\varphi(z) = 1 - \varphi_1 z - ... - \varphi p z(p)$$

and

$$\theta(z) = 1 + \theta 1 z... + \theta q z q$$

have no common factors.

For **q = 0** the process reduces to an autoregressive process of order p, denoted with **xt~AR(p),**

$$x_t - \varphi_1 x_{t-1} - ... - \varphi_p x_{t-p} = u_t \forall t \in Z,$$

For **p = 0** to a moving average process of order q, denoted with, $x_t \sim MA(q)$

$$x_t = u_t + \theta_1 u_{t-1} + ... + \theta_q u_{t-q} \forall t \in Z$$

It is important to underline that if we consider the set of autocorrelation functions there is not a one-to-one correspondence between the parameters of a causal ARMA (p,q) process and the autocorrelation function**.**

We note that an AR(p) process is always invertible, even if it is non-stationary, while an MA(q) process is always stationary, even if it is non-invertible. The invertibility can be used in order to ensure the identifiability of MA processes[2].

C. Particle swarm optimization (PSO).

Particle Swarm Optimization is a very simple algorithm and easy to understand algorithm.Particle swarm optimisation also finds the roots of the polynomial.In this algorithm we do over a number of iterations where a group of variables are adjusted in such a way where the values are closer to the member whose value is closest to the required outcome/prediction at any given moment[2].

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

The algorithm uses same logic and it's easy to implement in the project. The algorithm keeps track of three global variables:

1. Target value or condition
2. Global best (gBest) value indicating which particle's data is currently closest to the Target
3. Stopping value indicating when the algorithm should stop if the Target isn't found

Each particle consists of:

1. Data representing a possible solution.
2. The amount of data that can be changed is indicated by a Velocity value.
3. A personal best (pBest) value indicating the closest the particle's Data has ever come to the Target.

IV. WORKING PRINCIPLE

Based on the data fed, from various resources like National Stock Exchange (NSE), NASDAQ, London Stock Exchange (LSE), New York Stock Exchange(NYSE), least square support vector machine will analyse the data and data patterns. It will filter the data as much as possible[3] because the next algorithm which is used in this has a fairly high time complexity.Auto regressive moving average model states that the future value depends upon current and previous values. Considering all the previous and current values, Auto regressive will give one polynomial and moving average model will give another polynomial combing these two polynomial will give a complex polynomial which will be optimised further. Particle swarm optimisation will find the roots of the polynomial given by auto regressive moving average and find the roots of the polynomial and give us optimised roots of the polynomial which is the predicted value on stock market. As per the testing the accuracy of the predicted value is 25% to 30%. These are the predicted graphs using the three algorithms

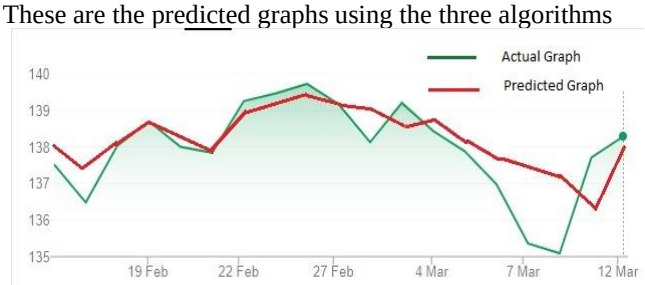

Fig 1. IBM Stock

Fig 2. Boeing Stock

Table 1. Stock Market prediction of Tata Motors

Date	Actual value		Predicted value	
	High	Low	High	Low
24-03-2019	24.8	22.8	25	23.5
25-03-2019	25.9	23.75	25.8	24
26-03-2019	27	26	26.5	26
27-03-2019	28.8	28.8	27.56	28.5

V. ARCHITECTURE

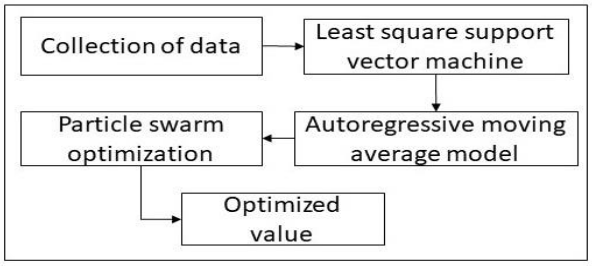

Fig 3.Architecture

The following work is done using three algorithm for accurate and single outcome. The algorithm used are LS-SVM, ARMA, PSO. First we collect the data of the last 10 years of the stock market from the internet and categorised them into year and products separately.The first algorithm we use is Least square support vector machine (LS-SVM). In this algorithm is sent the last 10 years data for predicted values. From here we use Autoregressive moving average model for getting the polynomial. Then we use the last algorithm for getting the optimized final value which accuracy will lie between 25-30%.

VI. CONCLUSION

The stock market prediction is expected to show and accuracy of 25% to 30%. Stock market prediction is one of the toughest prediction. This paper uses three machine learning algorithms viz, Least square support vector machine (LS-SVM) which reduces and learns the data obtained from sources like NSE, NASDAQ, Autoregressive moving average model (ARMA) which converts the data into polynomial function as the time complexity of polynomial solving is less we used this algorithm, Particle swap optimization (PSO) which will find the roots of the polynomial and optimise the roots. The optimised roots are the predicted values. Further this system can be updated by including two algorithms to increase the accuracy of the prediction. Stock market prediction is very difficult because it depends upon various aspects. We can increase the accuracy of the same by adding sentimental analysis. Since sentimental analysis plays an important role in predicting the stock market. Adding sentimental analysis can be a future enhancement. If sentimental analysis is added then this will show an accuracy of 60% to 70% which is huge in number.

REFERENCE

1. Lin Chen, ZhilinQiao, Minggang Wang, Chao Wang, Ruijin Du, H. Eugene Stanley "Which artificial intelligence algorithm better predicts the Chinese stock market?". JOURNAL OF LATEX CLASS FILES, VOL. 14, NO. 8, AUGUST 2015
2. Umadevi.K.S, AbhijitsinghGaonka, Ritwik Kulkarni, R. Jagadeesh Kannan "Analysis of Stock Market using Streaming data Framework".2018 IEEE
3. Honghai Yu, Haifei Liu "Improved Stock Market Prediction by Combining Support Vector Machine and Empirical Mode Decomposition". 2012 Fifth International Symposium on Computational Intelligence and Design.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

4. Zhiyong Zhao, Xiaodong Li, Fei Yu, and Hao Zhang "Stock Volatility Prediction using Multi-Kernel Learning based Extreme Learning Machine". 2014 International Joint Conference on Neural Networks (IJCNN) July 6-11, 2014, Beijing, China.
5. Mehak Usmani1, Syed Hasan Adil, Kamran Raza and Syed SaadAzhar Ali, "Stock Market Prediction Using Machine Learning Techniques". 2016 3rd International Conference On Computer And Information Sciences (ICCOINS).
6. Tong Sun, Jia Wang, Pengfei Zhang, Yu Cao, Benyuan Liu, Degang Wang, "Predicting Stock Price Returns using Microblog Sentiment for Chinese Stock Market". 2017 3rd International Conference on Big Data Computing and Communications.
7. Ashish Sharma, Dinesh Bhuriya, Upendra Singh. "Survey of Stock Market Prediction Using Machine Learning Approach". International Conference on Electronics, Communication and Aerospace Technology ICECA 2017.
8. TejasMankar, TusharHotchandani, Manish Madhwani, AkshayChidrawar, Lifna C.S, "Stock Market Prediction based on Social Sentiments using Machine Learning". IEEE 2018.

AUTHORS PROFILE

Vidhyavani, Assistant professor, Computer science department, SRM Institute of Science and Technology,Ramapuram, Chennai - 89.

Deepak Adithya.K.N., Batchelor of Technology in Computer Science and Engineering, SRM Institute of Science and Technology, Ramapuram, Chennai - 89.

Sateesh N, Batchelor of Technology in Computer Science and Engineering, SRM Institute of Science and Technology, Ramapuram, Chennai - 89.

Vignesh Kannan, Batchelor of Technology in Computer Science and Engineering, SRM Institute of Science and Technology, Ramapuram,
Chennai - 89.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)