

Stock Prediction Using Machine Learning Algorithms

¹Ashutosh Kumar, ²Hemlata

¹Research Scholar, ²Assistant Professor
Department of Computer Science
BITS, Sonapat.

Abstract- Every day more than 5000 trade companies enlisted in Bombay stock Exchange (BSE) offer an average of 24,00,00,000+ stocks, making an approximate of 2000Cr+ Indian rupees in investments. Thus analyzing such a huge market will prove beneficial to all stakeholders of the system. An application which focuses on the patterns generated in this stock trade over the period of time, and extracting the knowledge from those patterns to predict future behavior of the BSE stock market is essential. An application representing the information in visual form for user interpretation to buy and to sell a specific company's stock is a key requirement. Such an application based on machine learning algorithms is the right choice in current scenario. This paper surveys the machine learning algorithms suitable for such an application; as well it discusses what are the current tools and techniques appropriate for its implementation.

Key Words: Support Vector Machine (SVM), Support Vector Regression (SVR) and stock market.

I. INTRODUCTION

Machine learning can be defined as the data which is obtained by knowledge extraction. Machines don't have to be programmed explicitly instead they are trained to make decisions that are driven by data. Instead of writing a code for every specific problem, data is provided to the generic algorithms and logic is developed on the basis of that data. When a machine improves its performance based on its past experiences it can be said that machine has truly learnt.

The technique for most accurate prediction is by learning from past instances, and to make a program to do this is best possible with machine learning techniques. Any machine learning technique (supervised or unsupervised) is efficient enough to generate rules for programs, in consideration with present ones to take a better decision. In this scenario, the decision is whether the stock will increase or decrease (Stock analysis). A stock market is an open market for companies or for individuals to raise money. Stock market helps companies to purchase or sell their shares. The cost of shares relies on the interest and supplies of shares. This procedure of purchasing and selling of shares is called trading/exchanging; just the Listed Companies are permitted to do exchanging. Recently huge amounts of investment are exchanged via stock market over the world. National economies are firmly connected and intensely affected the execution of their Stock Markets. Therefore they are identified with macroeconomic parameters, as well as they impact ordinary life in a more straightforward manner. Hence they constitute a mechanism which has vital and immediate social effects.

The nature of stock market in common is unpredictable which may depend on the long and short term future state. This is unpleasant and also unavoidable for the speculator when Stock Market is chosen as an investment tool. The main aim is to reduce this unpredictability and the Stock Market Anticipation (or Forecasting) is used in this process. Stock market forecast is the procedure of attempting to focus the future stock estimation of an organization. The effective forecast of a stock's future cost could return huge benefit. Stock value movements are administered by the speculations random walk hypothesis and efficient-market hypothesis. The predictors of stock market concentrate on creating methodologies which effectively foresee stock costs utilizing overall characterized trading methodologies. A fruitful anticipation model is the particular case that works with best exactness having least input requirements and minimum complex model. Speculators and government associations depend on anticipating techniques to safeguard against threats and to monitor market changes. For analysts, these serve as a source of perspective for investigations of financial issues like estimating money related subsidiaries and portfolio determination.

At present, the stock brokers rely on upon Intelligent Trading Systems which support them in anticipating costs focused around different circumstances and conditions, which also help them in immediate investment decision. Stock market price are thought to be exceptionally dynamic and capable of fast changes since the basic nature of the financial space and partially due to the mix of known parameters (P/E Ratio, Previous Day's Closing Price etc.) and some different components. A trained trader would foresee the stock value and purchase a stock before the cost of stock climbs, or sell it before its esteem decreases. It is difficult to supplant the expertise that an experienced trader has picked up from his experience but an exact forecast algorithm can straightforwardly come into high benefits, individual experts, which shows an immediate relationship between the precision of the prediction algorithm and the benefit produced using utilizing the algorithm.

The key benefit of prediction is that the provision of different stakeholders with profitable data that can be utilized to take decision regarding the future. It is practically impossible to precisely anticipate the future prediction of future events comes with a margin of error. The margin of error augments particularly when anticipating in deep into the future, or when foreseeing. Variables and their normal impact may change (with social, monetary and political change) and new variables may develop. These mistakes emerge as a consequence of the level of error of the base information utilized and the strategy used to estimate into future. This makes the selection of anticipation method significant when foreseeing into future. As a rule prediction uses quantitative information instead of subjective information which rely on upon the judgment of specialists. Anyhow prediction that uses quantitative information is more exact than that of subjective information. Future price movement can be predicted using the following methodologies: the fundamental or intrinsic value analysis and the “chartist” or technical analysis and traditional time series prediction.

II. MACHINE LEARNING ALGORITHMS

2.1 Unsupervised learning

When the dataset is not well defined or very hard for interpretation, it is called unsupervised learning. The labels for the data are not defined. There no right way to divide data set except performing iterations. Thus, in supervised learning the input is used to generate a structure by looking at the relation of the input itself.

For example, Classification of animals. [4]

According to this research, unsupervised learning is not advisable for prediction.

2.2 Supervised learning

Supervised learning can be said as function approximation, training examples lead to function generation. If the learning is done with right training set, a well behaved function can be expected. Supervised learning grows consistently with the data. It is a type of induction learning, and it causes biased supervised learning sometimes.

E.g.: The function generated with supervised learning will be X^2 , if X is the input value and the output is self-multiplied.

Since, there is well defined data available from BSE itself and which is in well-defined numeric form it would be beneficial to use supervised learning algorithms. Supervised learning algorithms are of two variants: [3]

1. Regression.
2. Classification

III. PROPOSED ALGORITHM

In the proposed algorithm, data statistics are analyzed using Machine learning algorithm. In this machine learning neural network is used to apply the system. A number of data sets for a year is predicted & evaluated the performance parameter. There are following step of proposed algorithm.

1. Load Dataset

The data and info-vending services of National Stock Exchange (NSE) are provided through NSE Data & Analytics Limited (formerly known as Dot Ex International Limited) which is a separate professional set-up dedicated solely for this purpose. It provides NSE's market quotes and data for Capital Market Segment (CM), Futures and Options Segment (F&O), Currency Derivative Market Segment (CDS), Wholesale Debt Market Segment (WDM), Corporate Data, Corporate Bond Market data and Securities Lending & Borrowing Market (SLBM). NSE's real time data is provided in three levels (level 1, level 2, level 3 and tick by tick). Level 1 provides best bid and ask price, Level 2 provides market depth data upto 5 best bid and ask prices and Level 3 provides market depth data upto 20 best bid and ask prices.

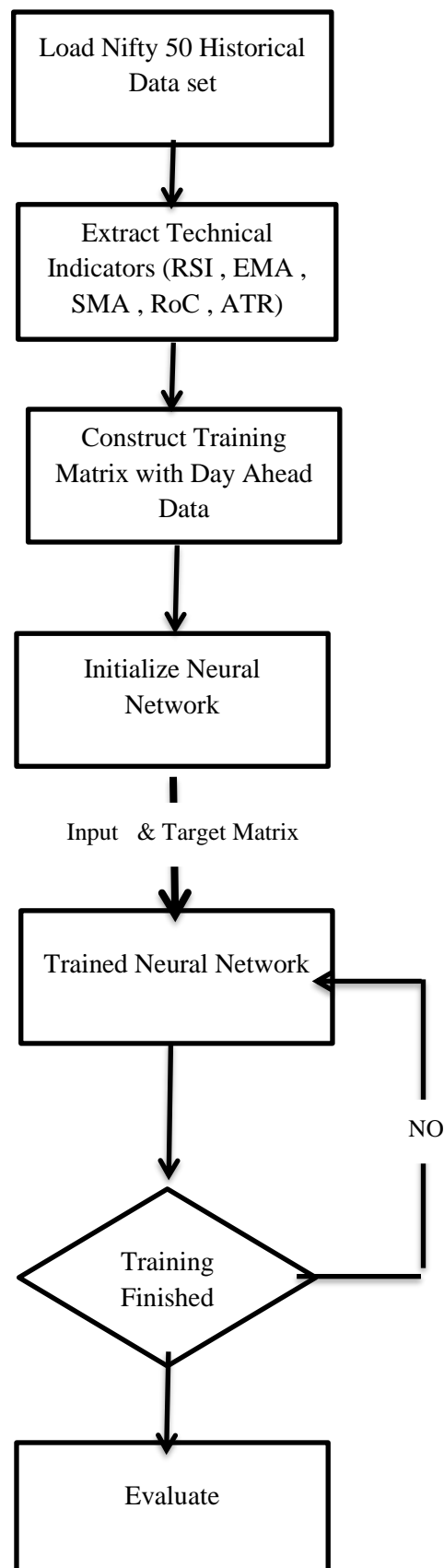


Fig 1 Proposed Methodology

The real time data feed is provided in TCP-IP format. It is provided on-line through a dedicated 2-10 mbps channelized E1 private leased line circuits. This line shall be owned by the customer and the line should be from National Stock Exchange, Mumbai to the premises of the customer. Alternatively, the customer can take the data from one of our authorized data vendors.

2. Extract Technical Indicator

Technical indicators are based on mathematical equations that produce a value that is then plotted on your chart. For example, a moving average calculates the average price of a currency pair in the past and plots a point on your chart. As your currency chart moves forward, the moving average plots new points based on the updated price information it has. Ultimately, the moving average gives you a smooth indication of which direction the currency pair is moving. In proposed work, RSI, EMA, SMA, Roc, ATR parameter are extracted using proposed algorithm.

3. Construct Training Matrix

The purpose of this section is to present a system, referred to as the Training Design Matrix, for streamlining the design of instructional environments in a way that integrates the instructional design elements. This system expands the use of job-analytic information to better link needs assessment to the design of instructional environments. Specifically, the TDM shows how determining an instructional objective's standing along four training issues—extensiveness of training, location of training, nature of transfer performance. These plans involve design elements such as trainee-to-trainer ratio, feedback delivery, hands-on practice, and modeling, structure of practice, amount of training, meaningfulness, pre- and post-training interventions, and general training methods.

4. Initialize Neural Network

When training and building a neural network, a number of subtle but important decisions needs to be taken. Zeroing down on the loss function to be used, the number of layers, kernel size, and the stride for each convolution layer, best-suited optimization algorithm for the network, etc. Compared to all these things, the choice of initialization of weights may seem trivial pre-training detail. But weight initialization contributes as a significant factor on the final quality of a network as well as its convergence rate. This paper discusses different approaches to weight initialization and compares their results on few datasets to find out the best technique that can be employed to achieve higher accuracy in relatively lower duration.

5. Trained Neural Network & Finished

Neural networks can be used without knowing precisely how training works, just as one can operate a flashlight without knowing how the electronics inside it work. Most modern machine learning libraries have greatly automated the training process. Owing to those things and this topic being more mathematically rigorous, you may be tempted to set it aside and rush to applications of neural networks. But the intrepid reader knows this to be a mistake, because understanding the process gives valuable insights into how neural nets can be applied and reconfigured. Moreover, the ability to train large neural networks eluded us for many years and has only recently become feasible, making it one of the great success stories in the history of AI, as well as one of the most active and interesting research areas.

6. Evaluate Performance Parameter

The different parameter like price and RSI are calculated from the proposed algorithm. RSI is a highly useful and popular oscillator. The RSI compares the magnitude of stock's recent profit to the magnitude of its recent losses and turns that information into a number that ranges from 0 to 100.

IV. RESULT ANALYSIS

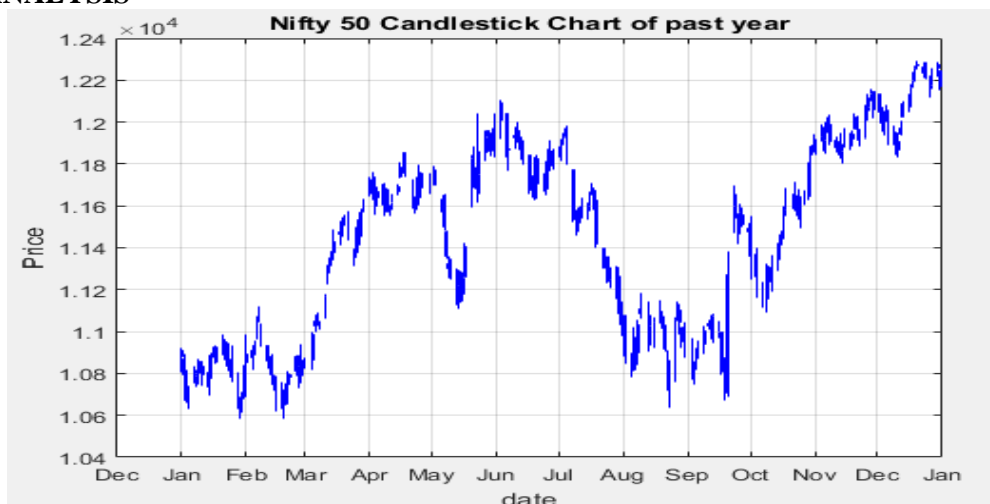


Fig 2 Nifty 50 candlestick for complete Year

The stock market graph is shown in the figure 2. In this graph, Nifty 50 candlestick for complete year is shown. The price of is varies in accordance to year. The stock market is depend upon economy of the country.

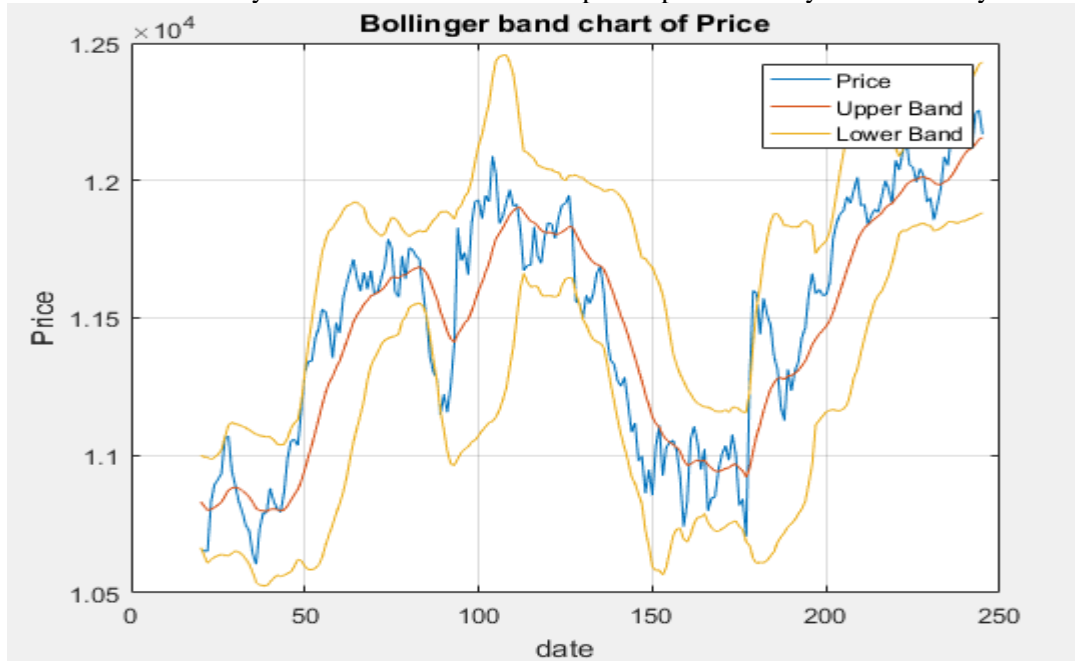


Fig 3 Bollinger Band chart of Price

The Bollinger band chart of price represent in fig 3. The graph showed the comparative analysis between lower band, upper band and price analysis.

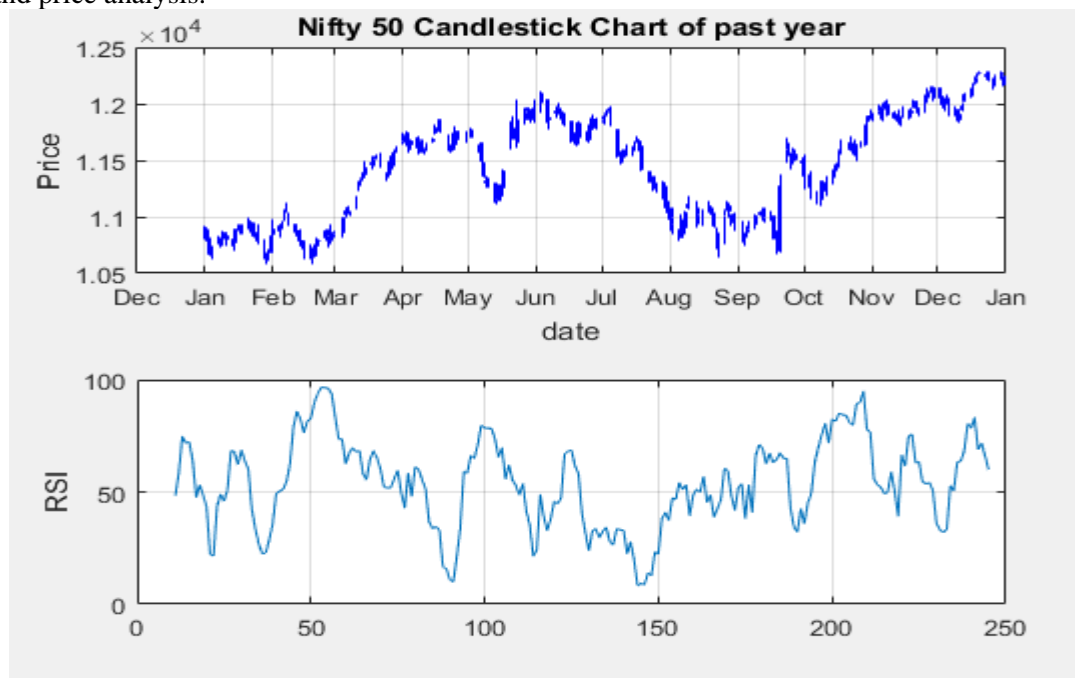


Fig 4 Nifty 50 candlestick Price & RSI chart for complete Year

Nifty 50 candlestick Price & RSI chart for complete year shown in fig 4. RSI is compare with price as shown in above figure.

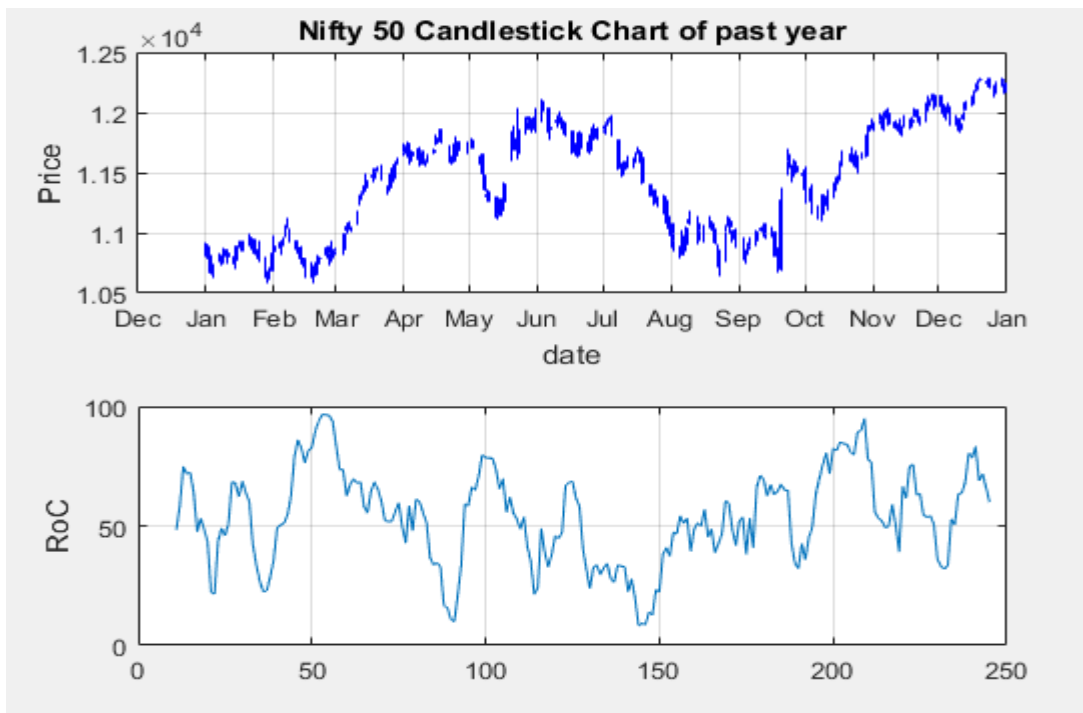


Fig 5 Nifty 50 candlestick Price & ROC chart for complete Year

Nifty 50 candlestick Price & ROC chart for complete year shown in fig 5. ROC is compare with price as shown in above figure.

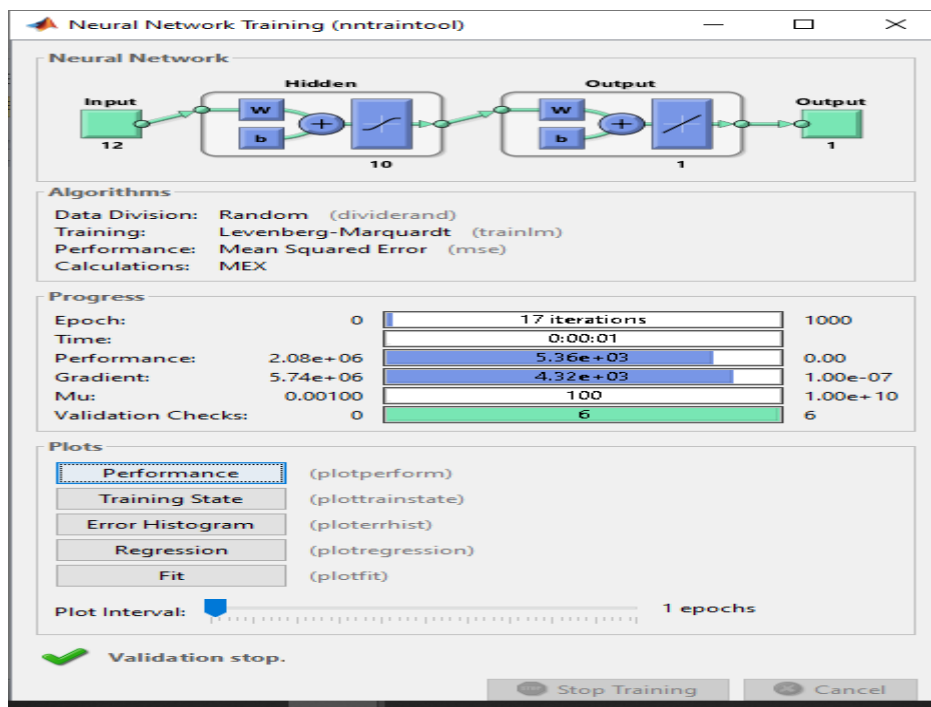


Fig 6 Neural Network Training

Neural Network trainer is shown in fig 6. The system have three layers, input layer , output layer and hidden layer. The neural network have various types. One of the neural network type Levenberg – Marquardt trainer is used in proposed system.

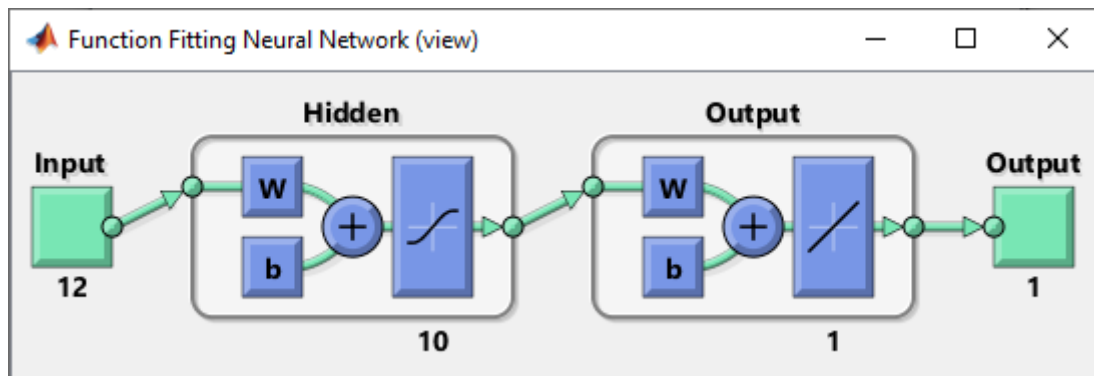


Fig 7 Function Fitting Neural Network

Neural Network fitting function is shown in the fig 7. These are used to obtain desired output of the function.

V. CONCLUSION

This paper summarizes important techniques in machine learning which are relevant to stock prediction. The outcome of this research concludes that the machine learning algorithms can be used to predict the increase or decrease in the stock market performance. It verifies the dependency of NSE on the factors taken in the study. Our findings confirm that the dependency of NSE is highest on the product rate, since the correlation factor is highest. Also, the correlation factor is lowest for product rate, showing least dependency of Nifty -50 on it. Of all the machine learning algorithms used, Neural Network shows the highest accuracy.

REFERENCES

1. Abhishek Gupta, Samidha & Sharma,D 2014, 'Clustering-Classification Based Prediction of Stock Market Future Prediction', International Journal of Computer Science and Information Technologies, vol. 5 , no.3, pp.2806-2809
2. Adebisi Ayodele, Ayo Charles, Adebisi Marion & Otokiti Sunday 2012, 'Stock Price Prediction using Neural Network with Hybridized Market Indicators', Journal of Emerging Trends in Computing and Information Sciences, vol. 3, no. 1, pp.1-9
3. Amit Ganatr & Kosta,YP 2010, 'Spiking Back Propagation Multilayer Neural Network Design for Predicting Unpredictable Stock Market Prices with Time Series Analysis', International Journal of Computer Theory and Engineering, vol.2, no.6, pp.963-971
4. Ao,SI 2006,'A framework for neural network to make business forecasting with hybrid VAR and GA components', Engineering Letters, vol.13,no.1, pp. 24-29
5. Bjoern Krollner, Bruce Vanstone & Gavin Finnie 2010,'Financial Time Series Forecasting with Machine Learning Techniques: A Survey', Symposium on Artificial Neural Networks - Computational Intelligence and Machine Learning
6. Carol Hargreaves & Yi Hao 2013,'Prediction of Stock Performance Using Analytical Techniques', Journal of Emerging Technologies in Web Intelligence, vol. 5, no. 2, pp.136-142
7. Cerny , V 1985,'Thermo dynamical Approach to the Traveling Salesman Problem: An Efficient Simulation Algorithm', Journal of Optimization Theory and Application, vol.45, no.1, pp.41-51
8. Chen & Mu-Yen 2011,'Predicting corporate financial distress based on integration of decision tree classification and logistic regression', Expert Systems with Applications , vol.38, no.9, pp.11261-11272
9. Chen,TL , Cheng,CH & Teoh,HJ 2008, 'High-order fuzzy time-series on multi-period adaptation model for forecasting stock markets', Physical A, vol.387, pp. 876-888
10. Chen,Y , Mabu,S , Shimada,K & Hirisawa,K 2009,'A Genetic Network Programming with Learning Approach for Enhanced Stock Trading Model', Expert Systems with Applications, vol.36, pp. 12537-12546
11. Cheng,CH , Chen,TL & Chiang,CH 2006,'Trend-Weighted Fuzzy Time-Series Model for TAIEX Forecasting', Neural Information Processing, Lecture Notes in Computer Science, vol. 4234, pp.469-477
12. Cheng,CH , Chen,TL & Wei,LY 2010, 'A hybrid model based on rough sets theory and genetic algorithms for stock price forecasting', Information Sciences , vol.180, pp. 1610-1629
13. Ching-Hsue Cheng, Hia Jong Teoh & Tai-liang Chen 2007,'Forecasting Stock Price Index Using Fuzzy Time-Series Based on Rough Set', Fuzzy Systems and Knowledge Discovery, vol.3, pp.336-340
14. Chi - chen wang 2011,' A comparison study between fuzzy time series model and ARIMA model for Forecasting Taiwan export', Expert system with Application , vol. 38, pp. 9296 - 9304

15. Ching Long Su, ChuenJyh Chen & Shih Ming Yang 2010, 'A self-organized neuro-fuzzy system for stock market dynamics modeling and forecasting', WSEAS Transactions on Information Science and Applications, vol.7.8, no.9, pp.1137-1149
16. Chiu,D & Chian,S 2010,'Exploring stock market dynamism in multi-nations with genetic algorithm, support vector regression, and optimal technical analysis', Proceedings of the Sixth International Conference on Networked Computing and Advanced Information Management, pp. 694-699
17. Choudhry,R & Garg,K 2008, 'A hybrid neuro-genetic approach for stock forecasting', World Academy of Science, Engineering and Technology, vol.15, pp. 315-318
18. Dase,RK & Pawar,DD 2010, 'Application of Artificial Neural Network for stock market predictions: A review of literature', International Journal of Machine Intelligence, vol. 2, no. 2, pp-14-17
19. Dassanayake,MMK & Chandima Tilakarathme 2010,'Predicting Trading Signals of Sri Lankan Stock Market Using Genetic Algorithms and Neural Networks', Technological Developments in Networking, Education and Automation, pp. 269-273
20. De Faria,EL , Marcelo, Albuquerque,P , Gonzalez,JL , Cavalcante,JTP , Marcio & Albuquerque,P 2009,'Predicting the Brazilian stock market through neural networks and adaptive exponential smoothing methods', Expert Systems with Applications, pp.1-4
21. Debashish Das & Mohammad Shorif Uddin 2013,'Data Mining And Neural Network Techniques In Stock Market Prediction: A Methodological Review', International Journal of Artificial Intelligence & Applications, vol.4, no.1, pp.117-127
22. Devroye L 1981,'On the equality of Cover and Hart in nearest neighbour discrimination', IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.3, pp. 75-78
23. Devroye & Wagner 1982, 'Nearest neighbour methods in discrimination, In Classification, Pattern Recognition and Reduction of Dimensionality', Handbook of Statistics, vol.2, pp. 193–197. North-Holland, Amsterdam
24. Devroye, Gyorfı, Krzyzak & Lugosi 1994,'On the strong universal consistency of nearest neighbour regression function estimates', The Annals of Statistics, vol.22, pp. 1371–1385
25. Dian Palupi Rini, Siti Mariyam Shamsuddin, Siti Sophiyati Yuhaniz 2011, 'Particle Swarm Optimization: Technique, System and Challenges', International Journal of Computer Applications, vol.14, no.1, pp.19-27
26. Ding, Song & Zen 2008,'Forecasting financial condition of chinese listed companies based on support vector machine', Expert Systems and Application, vol.34, no.4, pp.3081–3089
27. Enke,D & Thawornwong,S 2005,'The use of data mining and neural networks for forecasting stock market returns', Expert Systems with Applications, vol.29, pp. 927 - 940