

Analysing the Trend of Stock Market and Evaluate the performance of Market Prediction using Machine Learning Approach

Mathanprasad L and Gunasekaran M

Department of Computer Science, Government Arts College, Periyar University, Salem-636007,
E-mail : madanprasad@hotmail.com drmgunacs@gmail.com

Abstract- The main focus of the research work is to explore the current Stock market data values based on a real-time data, in which stock market data value fluctuates depends on time. The stock market prediction and analysing the future stock values is still considered as the challenging task in research field. The motivation of the current research work is the stock market data values varies time to time according to the subject risk. so it is the need to develop a computational automated methodology for predict the stock market data values. The detailed information regarding the existence of stock market variations is identified by collecting the information from previous historical data for making the choice of predictions strategy by considering the data with the stock analyst's experts of the stock market analysis system. The performance improvement of the newly developed by machine learning classification approach method is analysed through the comparative analysis report, which ensures the accurate prediction of the proposed method. This machine learning classification algorithm have been used, this predicts the stock market price and stock market movement changes from this performance can be evaluated and also proposed work suggested and recommended to the user, so user can easily find out the which stock will be in market longer period. The prediction accuracy of the stock exchange has analysed and improved to 94.17% using machine learning algorithms. As the results, this prediction will helpful to investor, to judge the current value and future prediction values of company's stock market rate.

Keyword: Stock market, Machine learning Algorithm, Company Stock Prediction, Accuracy

I. INTRODUCTION

In the modern world, the stock price is very tough to predict the values. So this research work focuses on future market trend predictions of company's stock values

The proposed research work applied the machine learning technique to predict the stock values of various company's based on previous historical data. The Steps and techniques involved in the stock market prediction are illustrated in the Table 1.

Table 1. Steps and Techniques of Stock Market Prediction

| Step involved in this work | Techniques Applied |
|---|--|
| Step 1 – Data Acquisition Process | Web-Worker Database - Quandl |
| Step 2 – Feature Extraction and Data set Creation | Historical stock values - Django templates |
| Step 3 – Machine Learning Approach | Categories the data as High open, High close, average movement. |
| Step 4 – Algorithm Comparison | To determine the best accuracy, the performance metrics are evaluated. |
| Step 5 - Prediction | Future Company stock predicted, useful for Investor. |

1.1 The Proposed Work Key features

Initially Data acquisition process done by Web-worker-Quandl, this will provide the different sources of stock data from several source. The data values that is feature extraction The previous stock market data are extracted using Web- worker techniques and converted into data required for machine learning classification techniques. Next process is the machine learning classification techniques for categorizing the High open, High close, average movement changes of stock values using historical data. Machine learning classification techniques named as, Random Forest, Support Vector Machine and Neural Network have been used to compare the proposed algorithms in terms of accuracy, sensitivity and specificity.

1.2 The Objective of research work

The main aim of the research work is Analysing the Trend of Stock Market and Evaluate the performance of Market Prediction using Machine Learning Approach

- To predict the stock market price considered as major challenging task.
- To study the stock market price variations changes due to several reasons.
- To predict the future values of company's stock market rate.
- To analysis the several factors like time-series values, equity, time-delay values during earlier stage itself using previous historical stock data.
- To evaluate the machine learning classification algorithms that applied to determine the prediction of stock market data from the Web datasets.
- The three different classifier have been used and compared the proposed algorithm in terms of performance metrics.
- This is considered as an important aspect because prediction will helpful to investor, to judge the current value and future prediction values of company's stock market rate.

II LITERATURE REVIEW

Literature survey which is related to the work adopted in this investigation is presented in this chapter.

Ananthi, et al. (2021) made research with Artificial Intelligent Techniques for predict the stock market data values using micro blogging sentiment analysis. By this technique the author predicts the future stock values overturn [1]. They have designed the hybrid algorithm of Artificial Neural Network with Genetic Algorithms optimization. This optimization algorithm gives the better results of the number of time delays in stock market data using Genetic Algorithm [2]. Also used machine learning algorithm named as support vector machine and random forest for predicting the Stock market. Earlier, the pre-processing was achieved by analysing the stock market data, extracted the features and finally by applying the machine learning approach the stock market data was predicted. By this approach the Random Forest yields the better accuracy [3]. Author [4] briefly provide the comprehensive survey about the relationship

with society and also recommended that the stock market data values which will be useful for stock analysts. Also, provides the stock movement values and society sentiment in the huge market media landscape of China. Initially data acquisition done by social media platform such as web forums, blogs & chat box. Finally predict the stock market by using chat box to analyst the stock market current directions which will be useful for trading purpose [5]. Finally, performed new methodology for stock market data analysing using deep convolutional neural network. Here NSE(India) stock market data and NYSE (New York) stock market data applied and yields the better accuracy for prediction using the deep convolutional neural network [6]. The accurately predicted stock market data value using newspaper data and past historical stock market data values. Here the machine learning approach applied for classify the text data from newspaper. As the results the Naïve based machine learning approach provides the better accuracy. This approach based on extracting the data text features from the newspaper. The machine learning approach such as multi-layer perceptron algorithm to analyse the current stock market data in the Karachi city. This algorithm provides the accuracy of 77% which will be based on the feature data.

The pipelined the stock market data and data acquisition and pre-processing achieved by Big data Hadoop concept, finally data are classified by machine learning approach. The main drawback of this approach the results yields the very less accuracy. Overall, the risk belonging to the stock market prediction and categorizing the severity risk level, those challenges are taken into consideration and find out the best solution, which has a direct effect on the user or investor to prevent the several losses. In this section, detailed domain knowledge and recent report on various stock prediction and classify the severity level by several machine learning techniques were also discussed. Based on the literature report, the performance measures are still needs to be improved based on the evidence of these report. Figure 1. Represents the Current Stock values of Infosys Ltd.



Figure 1. Current Stock values of Infosys Ltd.

III PROPOSED PREDICTION METHODOLOGY

The Figure 2. Illustrates the architecture of proposed methodology

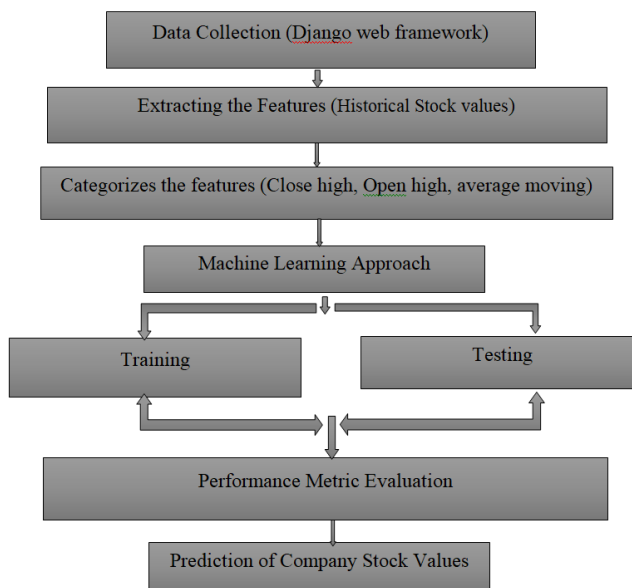


Figure 2. Illustrates the architecture of proposed methodology

IV IMPLEMENTATION

4.1 Neural Network Algorithm (NN)

Figure 3 illustrates the Neural Network with Input,

Hidden and Output layer. This is very important for the elegant design of ANN. This has a multiple layered neural mechanism for functioning. There are three specific layers

- Input layer - To get inputs 1, 2, 3.... N
- Hidden layer
- Output layer - To deliver the output

This NN is having the link to the biological NN for performing the functions in separate or in collection by units. There is no demarcation apparently of subtasks for the diversion units that is apportioned. NN is the representation of the model which has statistical employment, cognitive thinking of psychology and artificial intelligence. The framework of ANN has a mandate that it has a layer, which is single shrouded with one to ten neurons in hidden way. For statistical reporting, the equation of ANN gives the representation of the layers of input and then makes the signification of the parameter of output layer.

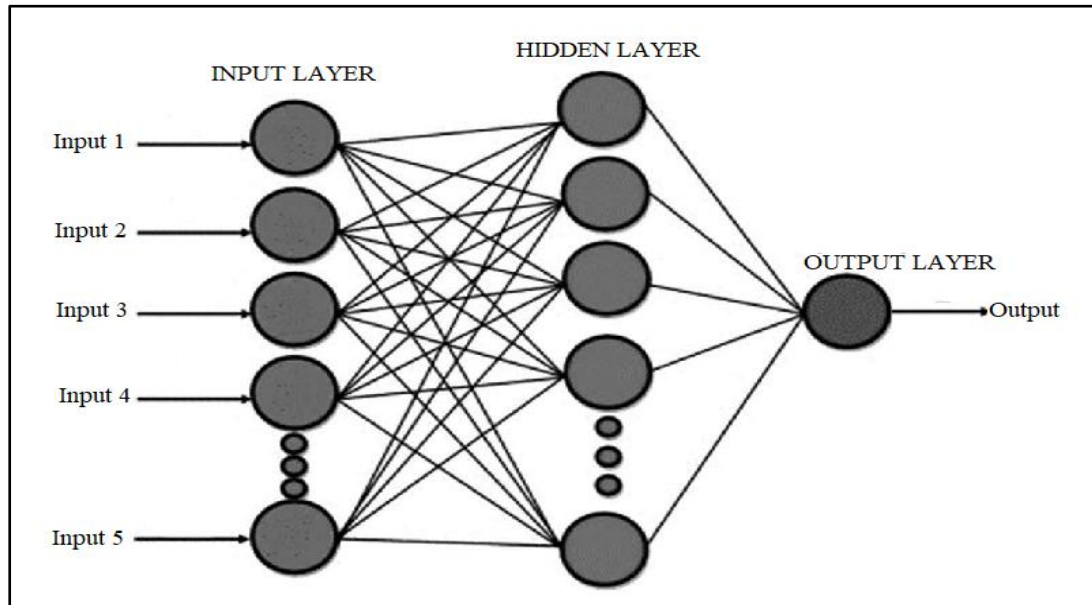


Figure 3. Neural Network

4.2 Support Vector Machine (SVM)

Figure 4 represents the Multiclass Support Vector Machine. The objective of SVM is to classify the data into groups based on their closeness to one another. Support Vector Machine is used for doing the classification of metrics of stock market. The binary techniques (Support Vector Machine - SVM) for the linear classification filters the hyper-plane which is the hyper-space of qualities. This performs the class separation in a prepared set. The ideal position that is a guideline for the kernel function's ideal position is the prime assignment.

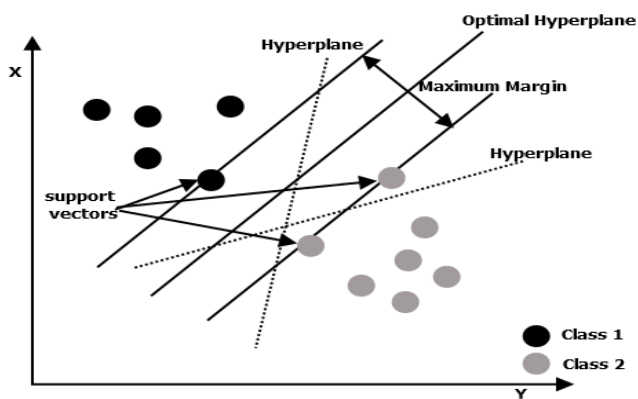


Figure 4 Multiclass Support Vector Machine

4.3 Random Forest

Random forests are a machine learning classification techniques for regression, and other methods that implement by generating training time of random forest and producing the output results of the individual trees.

Here, Random Forest, Support Vector Machine and Neural Network machine learning approach have been applied to find out the classification metrics of stock market. The system has been trained the stock market data values based on the past learning historical values and then testing the stock data values and finally automatically predict the results. The performance metric is evaluated with includes the parameters of Accuracy, Specificity and Sensitivity.

Both sensitivity and specificity are the most essential parameters for performance metrics calculation of any classifier.

| | | | |
|-------------------------------------|---|---|--|
| | Has the condition | Does not have the condition | |
| Positive Result from screening test | T+ True positive | F+ False positive | Row entries for determining positive predictive value |
| Negative | F- False negative | T- True negative | Row entries for determining negative predictive value |
| | Column entries for determining sensitivity | Column entries for determining specificity | |

$$\text{Sensitivity(or) Recall} = \frac{TP}{TP + FN}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Specificity} = \frac{TN}{TN + FP}$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

$$F1\ Score = 2 * \frac{Precision * Recall}{Precision + Recall}$$

V. EXPERIMENTAL ANALYSIS

The proposed implemented by open source available dataset using MATLAB software. Figure 5 represents the Classifier Performance. Table 2 represents the Accuracies in % of different classifiers. Table 3 shows the Performance Evaluation metrics includes the Specificity rate, Sensitivity rate, Precision value and F1 score of three different classifiers. Table 4. Represents the Prediction of NSE Stock Market values using Machine Learning Approach. The best accuracy which yields 94.17% accuracy from Levenberg-Marquardt method of Neural Network

Classification. Figure 6 represents the Levenberg-Marquardt method Neural Network Classification Results. Figure 7 represents the Levenberg-Marquardt method Neural Network Performance Analysis. Figure 8 shows the Levenberg-Marquardt method Neural Network Training State Analysis. Figure 9 represents the Levenberg-Marquardt method Neural Network Regression Analysis. Figure 10 represents the Levenberg-Marquardt method Neural Network Regression State.

Table 2 Accuracies in % of different classifiers

| SVM | RF | NN |
|--------|--------|--------|
| 86.67% | 73.33% | 94.17% |

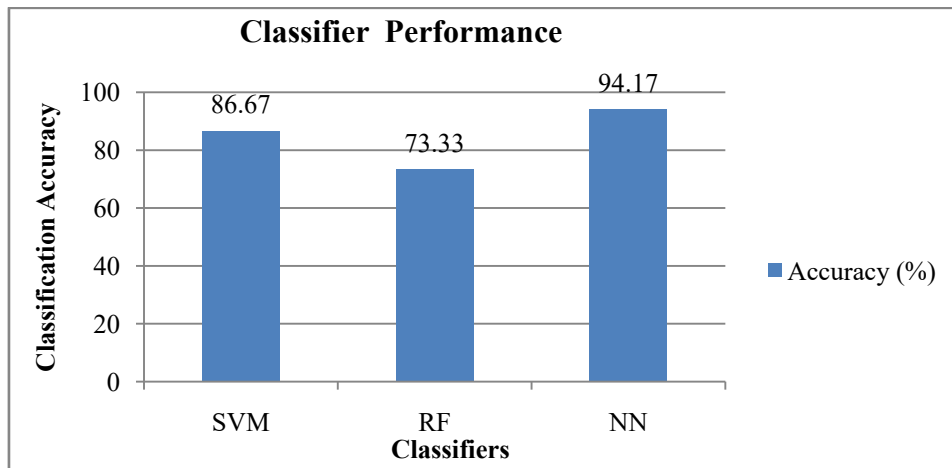


Figure 5 Classifier Performance

Table 3 Performance Evaluation metrics includes the Specificity rate, Sensitivity rate, Precision value and F1 score.

| Classifiers | Sensitivity | F1 Score | Specificity | Precision |
|------------------------|-------------|----------|-------------|-----------|
| Support Vector Machine | 0.8363 | 0.8449 | 0.9667 | 0.8579 |
| Random Forest | 0.7402 | 0.7152 | 0.9221 | 0.7678 |
| Neural Network | 0.9225 | 0.9443 | 0.9779 | 0.9550 |

Table 4. Prediction of NSE Stock Market values using Machine Learning Approach

| Company | SVM | RF | Neural Network |
|----------------|----------|----------|----------------|
| TCS | 2850.769 | 2930.067 | 1962.287 |
| Infosys | 1245.445 | 1281.533 | 1877.448 |
| Tech Mahendra | 549.018 | 521.178 | 568.766 |
| MindTree | 712.435 | 733.664 | 765.667 |
| Hexaware | 312.355 | 355.672 | 389.459 |
| Oracle Finance | 3688.668 | 4321.457 | 3500.67 |
| Wipro | 389.6488 | 377.6477 | 489.4673 |
| HCL | 965.6746 | 842.588 | 879.666 |

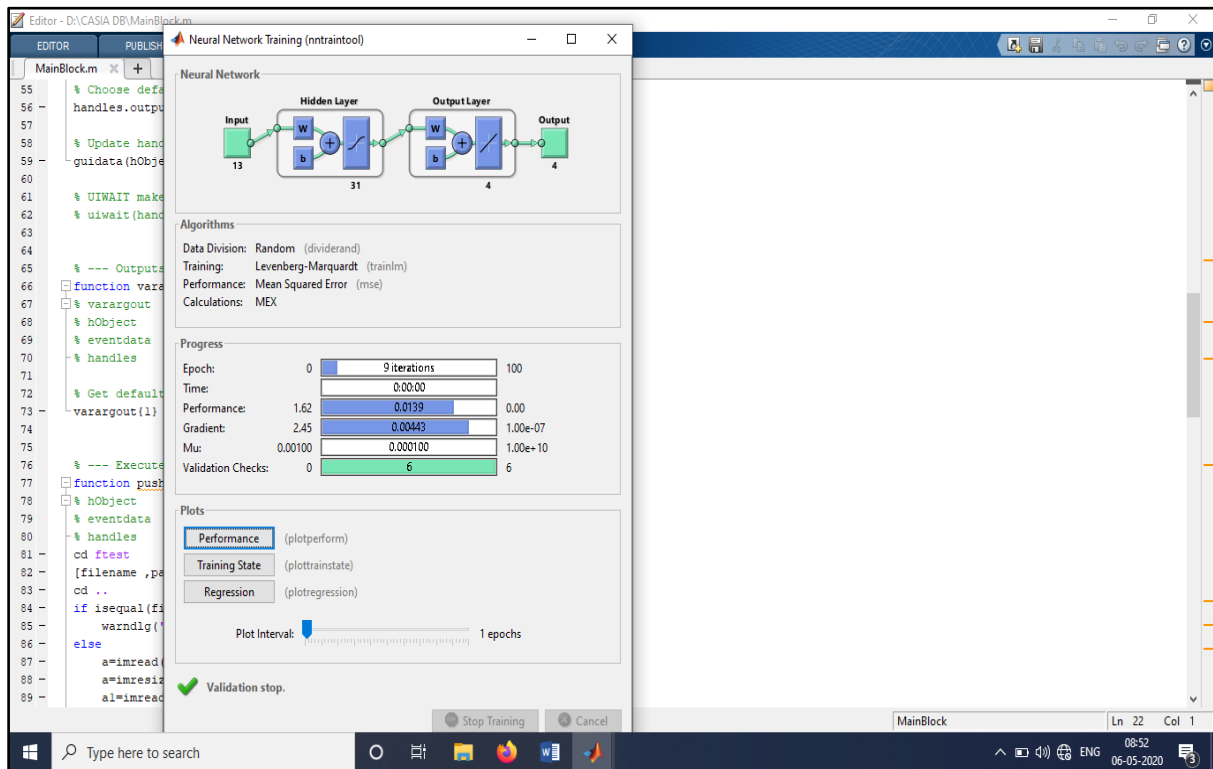


Figure 6 Levenberg-Marquardt methodNN Classification Results

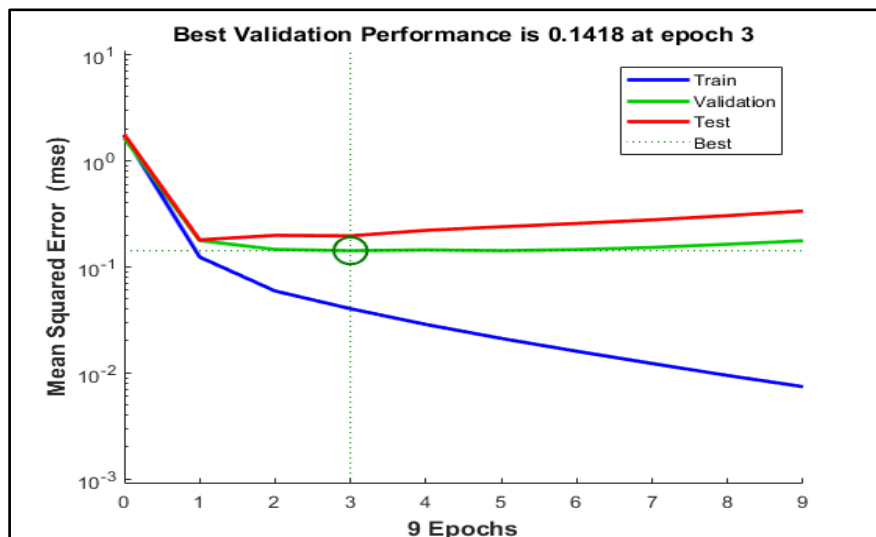


Figure 7 Levenberg-Marquardt methodNN Performance Analysis

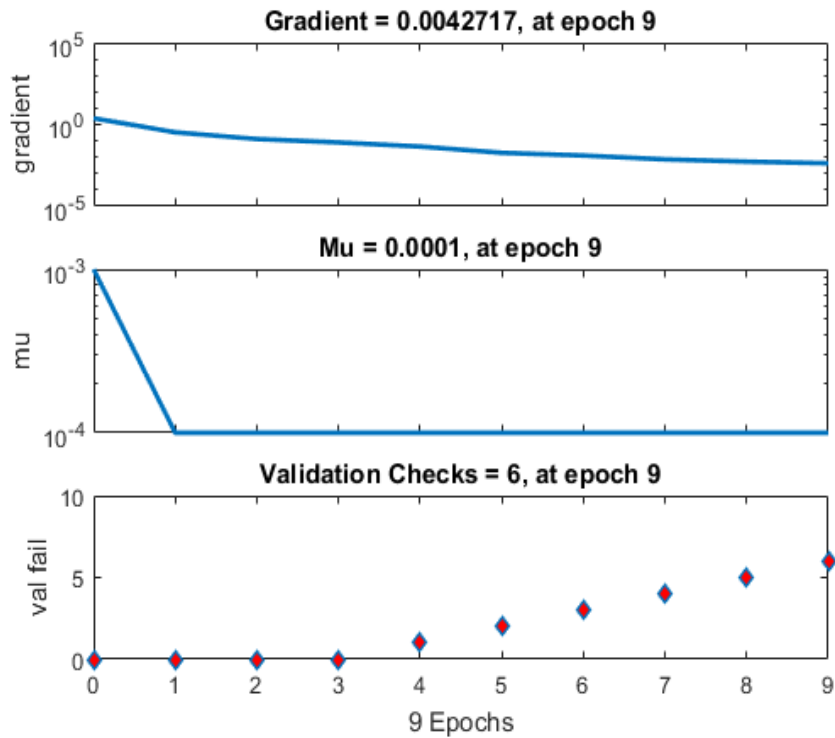


Figure 8 Levenberg-Marquardt methodNN Training State Analysis

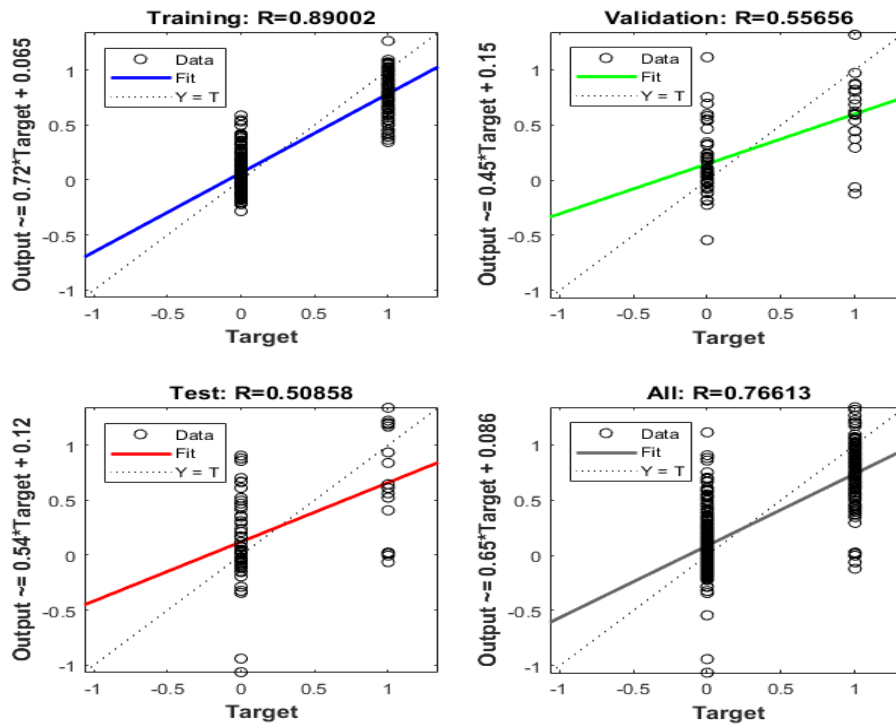


Figure 9 Levenberg-Marquardt methodNN Regression Analysis

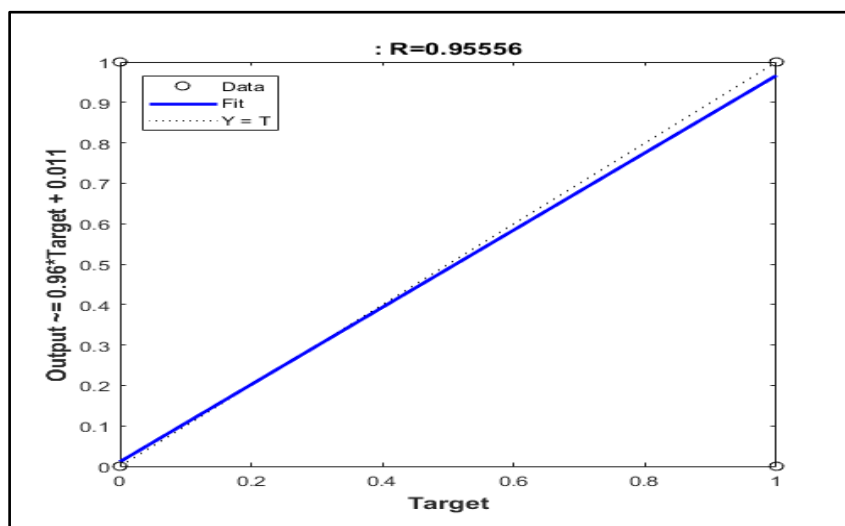


Figure 10 Levenberg-Marquardt method NN Regression State

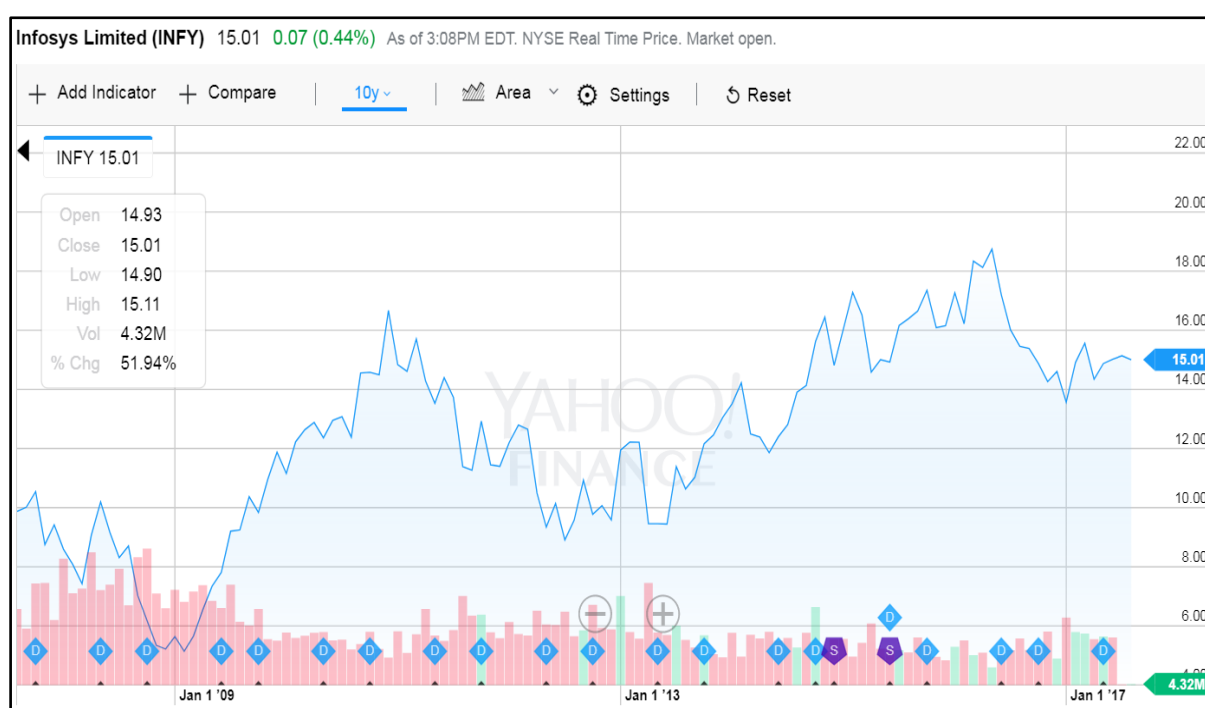


Figure 11 Stock value movement change in RSI

Figure 11 represents the Stock price variation with RSI (Relative Strength Index). Where, the movement in stock changes is calculated by stock closing values.

VI CONCLUSION

This computational approach solves the long-outstanding stock market problems and find out the solutions in stock market prediction research. As the result, the current research report is useful for Investor to predict the stock values movement

variations. By Levenberg-Marquardt method Neural Network yields the best accuracy of 94.17% among the various machine learning algorithm. As the results, this prediction will help to investor, to judge the current value and future prediction values of company's stock market rate.

REFERENCES

- [1] Ananthi, M., Vijayakumar, K. Stock market analysis using candlestick regression and market trend prediction (CKRM). . Journal of Ambient Intelligent and Humanized Computing, 12, 4819–4826 (2021). <https://doi.org/10.1007/s12652-020-01892-5>
- [2] Zhao J, Sun N, Cheng W (2019) Logistics forum based

- prediction on stock index using intelligent data analysis and processing of online web posts. ISSN: 1868-5145. <https://doi.org/10.1007/s12652-019-01520-x>
- [3] Z. Wang, S. Ho and Z. Lin, "Stock Market Prediction Analysis by Incorporating Social and News Opinion and Sentiment," 2018 IEEE International Conference on Data Mining Workshops (ICDMW), 2018, pp. 1375-1380, doi: 10.1109/ICDMW.2018.00195.
- [4] Peng Z et al (2019) Stock analysis and prediction using big data analytics. In: 2019 international conference on intelligent transportation, big data & smart City (ICITBS), pp 309–312.
- [5] Vijayakumar K, Arun C (2019) Continuous security assessment of cloud based applications using distributed hashing algorithm in SDLC. Cluster Computing. <https://doi.org/10.1007/s10586-017-1176-x>.
- [6] Vijayakumar K, Arun C (2017) Automated risk identification using NLP in cloud based development environments. Journal of Ambient Intelligent and Humanized Computing, <https://doi.org/10.1007/s12652-017-0503-7>.