



Enhancing Accuracy of Stock Market Prediction Using Linear Regression Algorithm by Comparing with Additive Regression

Myneni. Keerthi Pranav¹, S. Ashok Kumar²

¹Research Scholar, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Tamil Nadu, India.

²*Project Guide, Corresponding Author, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Tamil Nadu, India.

ABSTRACT

Aim: The Aim of the study is to develop an innovative method for stock market prediction using Linear Regression algorithm in comparison with Additive Regression algorithm. **Materials and Methods:** In stock market prediction a Data Mining technique Additive Regression algorithm and Linear Regression has been proposed and developed which helps in easy analysis and prediction of stock price. The Sample size was determined by using the G power calculator and found to be 20 per group. Totally 40 groups are used and Pretest Power is 25%. **Results:** Based on the analysis of the historic data using the innovative methods we can say that Linear Regression Algorithm has maximum accuracy of 85% whereas Additive Regression has 79%. The results achieved with $p=0.539$ ($p>0.05$) shows that two groups are statistically insignificant. **Conclusion:** The study proves that the proposed algorithm Linear Regression Algorithm method has higher accuracy than Additive Regression Algorithm using all the innovative approaches.

Keywords: Linear Regression, Additive Regression, Prediction, Stock Market, Accuracy, Investors

INTRODUCTION

For many researchers and business analysts stock market price prediction is always a challenging task. The stock market prices estimation and analysis using innovative algorithms is a challenging and interesting area of research (Bosco and Khan 2018). Time variant, and nonlinear are main characteristics of the data associated with the stock market. Stock market prediction plays a vital role in stock business (Wielechowski and Czech 2021). The lack of information and knowledge of the stock market by investors and investing can suffer a greater loss (Rao, Srinivas, and Mohan 2020). The most critical issue in the

present day financial market is to find a basic way to deal with and visualize the predictions in the stock market to be made by individuals in order to attain the greatest benefit by investments. (Cornell University, INSEAD, and WIPO 2020). The analysis of stock market prediction is an interesting subject as it influences the lives of the investors as they need to decide upon the choice of stocks to invest on a daily basis. The stock market prediction involves a lot of risk analysis as the lives of the involved in it. However, if the right decision is made using the innovative methods and algorithms then the investors get a lot of profit (Sharaff and Choudhary

2018). The advent of Deep learning techniques, numerous approaches have been proposed using innovative methods for stock market prediction(Prasad et al. 2021). Stock market predictions using linear regression are implemented for prediction to benefit stock investors; nearly 98 in IEEE explorer and 176 in Google scholar. The articles that are mostly cited based on the most viewed are one of them depending on the study of stock market prediction. There are many sources and organizations which have their main role to predict the future stock prices for a period of time(Biswas et al. 2021).These stock prices are predicted using a regression and classification model. This research paper states that stock market prediction benefits both the seller and investor as it has been the trendier method for money investment. Stock market prediction needs to bring out the output of accurate values as it deals with the financial mark(Biswas et al. 2021; Nova et al. 2021).(Parakh et al. 2020; Pham et al. 2021; Perumal, Antony, and Muthuramalingam 2021; Sathiyamoorthi et al. 2021; Devarajan et al. 2021; Dhanraj and Rajeshkumar 2021; Uganya, Radhika, and Vijayaraj 2021; Tesfaye Jule et al. 2021; Nandhini, Ezhilarasan, and Rajeshkumar 2020; Kamath et al. 2020) The limitations of stock investors is to get the accurate values and accurate prediction.In this research it is observed that Linear Regression has brought the outcome of more accuracy than Additive Regression Algorithm(Abhishek et al. 2012). The main aim of study is to predict the prices of stock using Linear regression model.

MATERIALS AND METHODS

The research work is carried out in DBMS Laboratory, Department of Computer Science Engineering, Saveetha School of Engineering. In this study 2 sample groups were taken. The group 1 was Linear Regression model and group 2 was Additive Regression Algorithm. Sample size is calculated using G power software considering the pretest power to be 80% and CI of 80%(Hendra, Lesmana, and Sabrina 2019) . The work has been carried out with 3500 records which was taken from a kaggle data set.The accuracy in predicting loyalties was initiated by two different groups.Totally 10 iterations were analyzed and performed on each group to accomplish maximum accuracy. Dataset contains 3000 instances and 7 attributes named Date, High, Low, Close, Open, Volume, OpenInt. Here data is from kaggle website (Kaggle: Your Machine Learning and Data Science Community).

Linear regression is one of the techniques in the regression model that are used to predict future prices of the stock market .This regression model mostly needs the dataset as the abstract for price prediction.This linear regression is classified into two types based on the explanatory variables,such as simple linear regression and multiple linear regression. Linear regression model uses two different types such as Independent variables and Dependent variables named as x and y respectively(Thrane 2019). The equation is represented as $Y=\beta_0+\beta_1X$. Additive Regression Algorithm can be defined as one of the intelligent models in data mining techniques which can be used for predicting the stock market prediction.This Additive Regression Algorithm includes a set of threshold functions. These functions are trained on the historical data and after

connecting each other with adaptive weights, they are used to make future predictions (Thrane 2019; Hoffmann 2021). The tool used to execute the program is Google colab and the databases are directly imported when the commands are instructed in the command prompt. Algorithm is implemented in the python code and the accuracy is obtained based on the dataset.

Linear Regression(LR) Algorithm

In Deep learning algorithms Linear Regression is based on the model of the relationship between two variables by fitting a linear equation to observe data, when one variable is considered to be an exploratory variable.

Pseudo Code

Input: Training Data

Output: Accuracy

Begin

For all the existing data

Read and Load the data set.

Extract Data features of a sound file imported.

Add the technical indicators.

Initialize the LR classifier.

Train the classifier

Predict the test set

Return the Accuracy

End for

Additive Regression (AR) Algorithm

In Deep learning algorithms, Additive Regression is one of the simplest machine learning algorithms based on Supervised Learning techniques that can be used for both classification and regression problems. Additive regression Algorithm is a Nonparametric Regression method

Pseudo Code

Input: Training Data

Output: Accuracy

Begin

For all the existing data

Read and Load the data set.

Extract Data features of a sound file imported.

Calculate the mean and standard deviation of the predictor variables in each class.

Initialize the Additive Regression classifier.

Train the classifier

Predict the test set

Return the Accuracy

End for

Statistical Analysis

The analysis of stock price prediction was done using IBM SPSS software. Independent sample t test is carried out for analysis. Independent variables are dataset and dependent variable is accuracy (Maguluri et al. 2020).

RESULTS

The Additive Regression algorithm shows that databases are the abstract for determining the accurate predictions. Table 1. represents the group statistics analysis which include the total no. of. Values, mean, standard deviation, standard error mean. Linear regression method and Additive Regression algorithm have an accuracy of 85% and 79% respectively. Standard error mean for the Additive Regression Algorithm method is less than the Linear regression method.

In Table 2, the results achieved with $p=0.539$ ($p>0.05$) shows that two groups are statistically insignificant.

Figure 1 represents the graph that explains the comparison of the accuracy value with the algorithm Linear regression method and

the Additive Regression algorithm, where the accuracy of Linear regression method is 85% and the accuracy value of the Additive Regression algorithm is 79%.

DISCUSSION

The data evolution was performed using IBM SPSS version 21. To analyze the data, Independent sample T-test and group statistics can be carried out. In this research study it is proved that the linear regression method has got the more accurate values than the Additive regression algorithm (Shahvaroughi Farahani and Razavi Hajiagha 2021).

Linear regression is a method which has lower time complexity when compared to other regression and classification methods. This method takes less time for compiling to get the accurate values (Smithers 2022). Linear regression performs exceptionally well for linearly separable data. This method is easier to implement, interpret and efficient to train (Inthachot, Boonjing, and Intakosum 2016). Whereas in Additive Regression Algorithm choosing a good kernel function is not an easy task as it consumes most time. The Additive Regression Algorithm model works slow when it works on large datasets (Rangvid 2021). The accuracy of output depends on the quality of the data. The Additive Regression Algorithm is difficult to understand its variable weights and individual impact (Cheng 2015).

Although the proposed methodology obtained satisfactory results in stock market prediction, There are limitations of the investors are the accurate stock price prediction and less efficiency. On analyzing innovative methods we can say that stock

market prediction is easier using linear regression (Qiu and Song 2016). The main aim of the study is to provide accurate stock price prediction using linear regression method. There are various applications where linear regression can be used such as predicting weather, forecasting results, academic reports, etc (Qiu and Song 2016; Bosco and Khan 2018).

CONCLUSION

In the proposed model, the accuracy percentage of predicting the stock market prices using linear regression model is 85%, whereas in the comparison model the Additive Regression algorithm has got the accuracy of 79% only.

DECLARATION

Conflict of Interests

No conflict of interest in this manuscript.

Authors Contribution

Author MKP was involved in data collection, data analysis, and manuscript writing. Author SAK was involved in the action process. Data verification and validation and Criteria review of manuscript.

Acknowledgement

The authors would like to express their gratitude towards Saveetha School of Engineering, Saveetha Institute of Medical And Technical Sciences (Formerly known as Saveetha University) for providing the necessary infrastructure to carry out this work successfully.

Funding

We thank the following organizations for providing financial support that enabled us to complete the study.

1. Oracle Tech Solutions Pvt.Ltd
2. Saveetha University

3. Saveetha Institute of Medical and Technical Sciences
4. Saveetha School of Engineering

REFERENCES

1. Abhishek, K., A. Khairwa, T. Pratap, and S. Prakash. 2012. "A Stock Market Prediction Model Using Artificial Neural Network." In *2012 Third International Conference on Computing, Communication and Networking Technologies (ICCCNT'12)*. IEEE. <https://doi.org/10.1109/icccnt.2012.6396089>.
2. Biswas, Milon, Atanu Shome, Md Ashraful Islam, Arafat Jahan Nova, and Shamim Ahmed. 2021. "Predicting Stock Market Price: A Logical Strategy Using Deep Learning." In *2021 IEEE 11th IEEE Symposium on Computer Applications & Industrial Electronics (ISCAIE)*. IEEE. <https://doi.org/10.1109/iscaie51753.2021.9431817>.
3. Bosco, Joish, and Fateh Khan. 2018. *Stock Market Prediction and Efficiency Analysis Using Recurrent Neural Network*. GRIN Verlag.
4. Cheng, Siwei. 2015. "Analysis, Evaluation, and Prediction of China's Stock Market." *The Chinese Stock Market Volume II*. https://doi.org/10.1057/9781137464699_1.
5. Cornell University, INSEAD, and WIPO. 2020. *Global Innovation Index 2020: Who Will Finance Innovation?* WIPO.
6. Devarajan, Yuvarajan, Beemkumar Nagappan, Gautam Choubey, Suresh Vellaian, and Kulmani Mehar. 2021. "Renewable Pathway and Twin Fueling Approach on Ignition Analysis of a Dual-Fuelled Compression Ignition Engine." *Energy & Fuels: An American Chemical Society Journal* 35 (12): 9930–36.
7. Dhanraj, Ganapathy, and Shanmugam Rajeshkumar. 2021. "Anticariogenic Effect of Selenium Nanoparticles Synthesized Using Brassica Oleracea." *Journal of Nanomaterials* 2021 (July). <https://doi.org/10.1155/2021/8115585>.
8. Hendra, Edwin, Theresia Lesmana, and Sasya Sabrina. 2019. "Market Reaction on Reverse Stock Split Announcement: Empirical Evidence in Indonesian Stock Market." *Proceedings of the 2nd Economics and Business International Conference*. <https://doi.org/10.5220/0009200601540161>.
9. Hoffmann, John P. 2021. "Indicator Variables in Linear Regression Models." *Linear Regression Models*. <https://doi.org/10.1201/9781003162230-7>.
10. Inthachot, Montri, Veera Boonjing, and Sarun Intakosum. 2016. "Artificial Neural Network and Genetic Algorithm Hybrid Intelligence for Predicting Thai Stock Price Index Trend." *Computational Intelligence and Neuroscience* 2016 (November): 3045254.
11. Kamath, S. Manjunath, K. Sridhar, D. Jaison, V. Gopinath, B. K. Mohamed Ibrahim, Nilkantha Gupta, A. Sundaram, P. Sivaperumal, S. Padmapriya, and S. Shantanu Patil. 2020. "Fabrication of Tri-Layered Electrospun Polycaprolactone Mats with Improved Sustained Drug Release Profile." *Scientific Reports* 10 (1): 18179.
12. Maguluri, Lakshmana, Annamalai University, Ragupathy Rengaswamy, and Annamalai University. 2020. "An

- Efficient Stock Market Trend Prediction Using the Real-Time Stock Technical Data and Stock Social Media Data.” *International Journal of Intelligent Engineering and Systems*. <https://doi.org/10.22266/ijies2020.0831.28>.
13. Nandhini, Joseph T., Devaraj Ezhilarasan, and Shanmugam Rajeshkumar. 2020. “An Ecofriendly Synthesized Gold Nanoparticles Induces Cytotoxicity via Apoptosis in HepG2 Cells.” *Environmental Toxicology*, August. <https://doi.org/10.1002/tox.23007>.
14. Nova, Arafat Jahan, Zahada Qurashi Mim, Sanjida Rowshan, Md Riad Ul Islam, Md Nurullah, and Milon Biswas. 2021. “Stock Market Prediction on High-Frequency Data Using ANN.” *Asian Journal of Research in Computer Science*, July, 1–12.
15. Parakh, Mayank K., ShriramUlaganambi, Nisha Ashifa, Reshma Premkumar, and Amit L. Jain. 2020. “Oral Potentially Malignant Disorders: Clinical Diagnosis and Current Screening Aids: A Narrative Review.” *European Journal of Cancer Prevention: The Official Journal of the European Cancer Prevention Organisation* 29 (1): 65–72.
16. Perumal, Karthikeyan, Joseph Antony, and Subagunasekar Muthuramalingam. 2021. “Heavy Metal Pollutants and Their Spatial Distribution in Surface Sediments from Thondi Coast, Palk Bay, South India.” *Environmental Sciences Europe* 33 (1). <https://doi.org/10.1186/s12302-021-00501-2>.
17. Pham, Quoc Hoa, SupatChupradit, Gunawan Widjaja, Muataz S. Alhassan, Rustem Magizov, Yasser Fakri Mustafa, Aravindhan Surendar, AmirzhanKassenov, Zeinab Arzehgar, and WanichSuksatan. 2021. “The Effects of Ni or Nb Additions on the Relaxation Behavior of Zr55Cu35Al10 Metallic Glass.” *Materials Today Communications* 29 (December): 102909.
18. Prasad, Venkata Vara, Srinivas Gumparthy, Lokeswari Y. Venkataramana, S. Srinethe, R. M. Sruthi Sree, and K. Nishanthi. 2021. “Prediction of Stock Prices Using Statistical and Machine Learning Models: A Comparative Analysis.” *Computer Journal*, July. <https://doi.org/10.1093/comjnl/bxab008>.
19. Qiu, Mingyue, and Yu Song. 2016. “Predicting the Direction of Stock Market Index Movement Using an Optimized Artificial Neural Network Model.” *PloS One* 11 (5): e0155133.
20. Rangvid, Jesper. 2021. “The Stock Market over the Business Cycle.” *From Main Street to Wall Street*. <https://doi.org/10.1093/oso/9780198866404.003.0009>.
21. Rao, Polamuri Subba, K. Srinivas, and A. Krishna Mohan. 2020. “A Survey on Stock Market Prediction Using Machine Learning Techniques.” In *Lecture Notes in Electrical Engineering*, 923–31. Lecture Notes in Electrical Engineering. Singapore: Springer Singapore.
22. Sathiyamoorthi, Ramalingam, Gomathinayakam Sankaranarayanan, Dinesh Babu Munuswamy, and Yuvarajan Devarajan. 2021. “Experimental Study of Spray Analysis for Palmarosa Biodiesel-diesel Blends

- in a Constant Volume Chamber.” *Environmental Progress & Sustainable Energy* 40 (6). <https://doi.org/10.1002/ep.13696>.
23. Shahvaroughi Farahani, Milad, and Seyed Hossein Razavi Hajiagha. 2021. “Forecasting Stock Price Using Integrated Artificial Neural Network and Metaheuristic Algorithms Compared to Time Series Models.” *Soft Computing*, April, 1–31.
24. Sharaff, Aakanksha, and Meenakshi Choudhary. 2018. “Comparative Analysis of Various Stock Prediction Techniques.” In *2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI)*. IEEE. <https://doi.org/10.1109/icoei.2018.8553825>.
25. Smithers, Andrew. 2022. “Valuing the US Stock Market.” *The Economics of the Stock Market*. <https://doi.org/10.1093/oso/9780192847096.003.0015>.
26. Tesfaye Jule, Leta, Krishnaraj Ramaswamy, Nagaraj Nagaprasad, Vigneshwaran Shanmugam, and Venkataraman Vignesh. 2021. “Design and Analysis of Serial Drilled Hole in Composite Material.” *Materials Today: Proceedings* 45 (January): 5759–63.
27. Thrane, Christer. 2019. “Linear Regression with Several Independent Variables Multiple Regression.” *Applied Regression Analysis*. <https://doi.org/10.4324/9780429443756-4>.
28. Uganya, G., Radhika, and N. Vijayaraj. 2021. “A Survey on Internet of Things: Applications, Recent Issues, Attacks, and Security Mechanisms.” *Journal of Circuits Systems and Computers* 30 (05): 2130006.
29. Wielechowski, Michał, and Katarzyna Czech. 2021. “Companies’ Stock Market Performance in the Time of COVID-19: Alternative Energy vs. Main Stock Market Sectors.” *Energies*. <https://doi.org/10.3390/en15010106>.

TABLES AND FIGURES

Table 1. Group Statistics results (Mean of Linear Regression is 85 which is greater when compared to AR Algorithm is 73 and Standard error mean for Linear Regression is 1.024 and KNN algorithm is 0.933)

	Algorithms	N	Mean	Std. Deviation	Std. Error Mean
Accuracy	Linear Regression	20	85	3.239	1.024
	AR Algorithm	20	79	2.951	0.933

Table 2. The significance value $p=0.539$ ($p<0.05$) shows that two groups are statistically insignificant.

	Levene's test for equality of variables		T-test for Equality of Mean					
	F	Sig	t	df			Std. Error	95% confidence

Enhancing Accuracy of Stock Market Prediction Using Linear Regression Algorithm by Comparing with Additive Regression

						Sig(2-tailed)	Mean difference	difference	interval of the difference	
									Lower	Upper
Accuracy	Equal variance assumed	0.391	0.539	4.474	18	0.000	6.200	1.386	3.289	9.111
	Equal variances not assumed			4.474	17.847	0.000	6.200	1.386	5.287	9.113

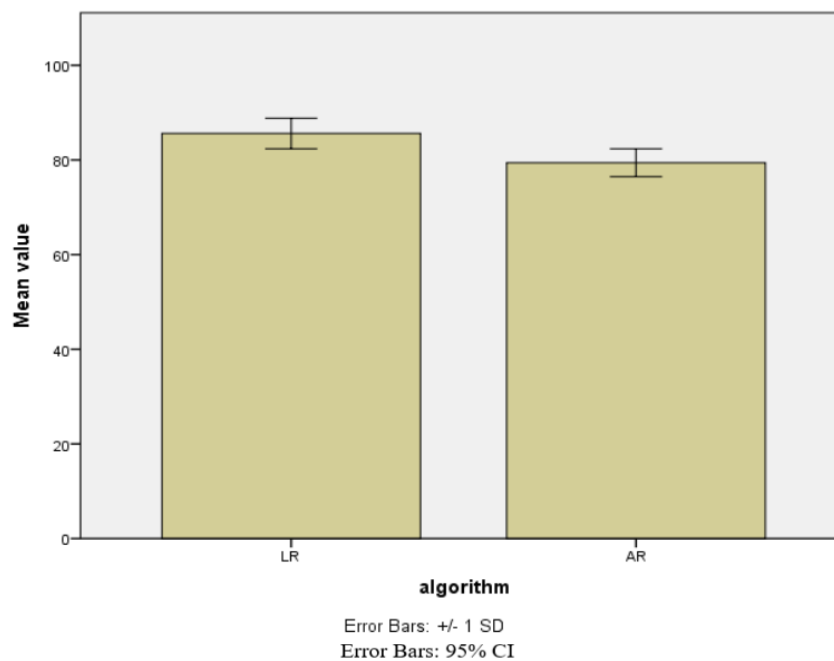


Fig. 1. Bar Chart representing the comparison of Mean Accuracy of Linear Regression and Additive regression Algorithm. Mean accuracy of Linear Regression is 85% appears to be better than Additive Regression Algorithm which is 79%. The X-axis represents Logistic Regression and Additive Regression Algorithm and the Y-axis represents the mean accuracy ± 1 SD.