



## “STOCK MARKET FORECASTING USING HYBRID DEEP LEARNING APPROACH”

Priya Pradeep Barhalikar<sup>1\*</sup>, Prof. Chandrakant Navdeti<sup>2</sup>, Dr. B.S. Shetty<sup>3</sup>

### Abstract—

The Stock Market and its tendencies are very erratic in the finance sector. Recent studies have demonstrated the enormous influence that information stories and social analyses may have on investors' perceptions of the financial marketplace. Consequently, The goal of this study is to investigate how stock market news mood movements are related use of data from various news outlets, business organization publications, and commercial enterprise websites. By utilizing prior information about the model's structure, In comparison to the auto-regressive integrated moving average model as well as vector autoregression technique, this will be study offers an implementation of The Bayesian Structural Time Series (BST) Model that has greater transparency and permits better handling of uncertainty. The assumption of linearity is one of the model's major flaws. A nonlinear model, The Long Short Term Memory model had been potential to represent a variety of nonlinear properties present in the data set. The suggested approach uses a hybrid model that incorporates the long-short-term memory and Bayesian Structural Time Series models, as well as a regression component that gathers data from different news outlets to find market predictors. The suggested approach recognizes unexpected behavior or abnormal patterns in stock price movement, making it preferable to previous methods.

**Index Terms**—long-short-term memory (LSTM),Support Vector Machine (SVM), Hybrid Approach, prediction, stock market.

---

<sup>1\*</sup>Department Of Information Technology Shri Guru Gobind Singhji Institute of Engg. & Technology Vishnupuri, Nanded, India email: pbarhalikar@gmail.com

<sup>2\*</sup>Department Of Information Technology Shri Guru Gobind Singhji Institute of Engg. & Technology Vishnupuri, Nanded, India email: cpnavdeti@sggs.ac.in

<sup>3\*</sup>Department Of Information Technology Shri Guru Gobind Singhji Institute of Engg. & Technology Vishnupuri, Nanded, India email: head.it@sggs.ac.in

**\*Corresponding Author:** Priya Pradeep Barhalikar

\*Department Of Information Technology Shri Guru Gobind Singhji Institute of Engg. & Technology Vishnupuri, Nanded, India email: pbarhalikar@gmail.com

**DOI:** 10.48047/ecb/2023.12.si10.00355

## I. INTRODUCTION

Use acts the try to anticipate an upcoming value of the firm stocks or another financial instrument listed along annexe change is known as predicting The Stock market and analysis. The Stock Market is a key aspect of the country’s economy and plays an important role in the expansion of the country’s industry and commerce, which eventually influences the country's economy. This Stock Market Prediction system is an online application that attempts to predict the value of stock markets using Linear Regression, Support Vector Machine, and LSTM. This research aims to resolve the economic quandary that has arisen among those who wish to expand in the Stock Market. With the rise of Web access, Social media platforms, as well as online social interactions, predicting everyday user behavior has become a viable career. As a result, our issue content is to create a public service that incorporates old data and user prediction to create a more powerful model that benefits everyone. The auto-regressive Integrated Moving Average (ARIMA) Model forecasted the stock market in early 2009. ARIMA is an acronym that stands for The autoregressive integrated moving average. It describes a model of time series in which the time series forecasts using the series’ departed values. The Model is designed to deal with statistical difficulties. Though the model is capable of predicting periodicity and the latest developments, outliers are challenging to foresee for ARIMA since they fall outside of the broad movement represented by the model. The ARIMA model can’t perform sentiment analysis. The technology will be used to create as well as construct the system for detecting the Sentiment Analysis in news articles and social media for stock investment.

## II. LITERATURE REVIEW

Bhaswati Ganguli et al.[1] proposed the system with the help of this paper To discover market predictors, the authors suggest a hybrid model that combines the Long - short term memory model with the BST model, as well as the regression component that incorporates data from several news sources. The suggested approach recognizes unexpected behavior or abnormal patterns in stock price movement, making it preferable to previous methods. Our new hybrid model accumulates errors at a reduced rate and outperforms some of the different current hybrid models, including the AR - MLP, ARIMA - LSTM, and VAR - LSTM models.

Ching - Te Wang et al.[2] proposed the system with the help of this technology can forecast stock prices for users. After establishing the appropriate rule, an internet robot can begin collecting content

immediately as Possible. To accelerate the process of the computations among data analysis and data crawled, The System creates a centralized distribution of the web robot system that separates the web robots, the data servers, and the analysis servers. As a result, the technology automates the procedure and effectively increases the reaction rate. The major feature of this system is that it automatically collects data by employing web crawlers and establishes the regular expressions or the X path to analyzing web pages.

According to The experimental data, the System performed better. As a result, the Approach can trawl the important data, efficiently analyze large amounts of information, and predict stock market prices. Sahaj Singh Maini et al.[3] proposed the system with the help of This study, The Predictive analysis was performed along the Dow Jones industrial average index, which can be comprised of the thirty firms and is mostly owned by the S P Global. This is important for traders since such an analysis might positively affect their decision to purchase or sell an item. We have discussed two statistical Machine Learning(ML)Models, The Support Vector Machine(SVM) and the Random Forest Model(RFM).

Which are used to anticipate stock market movements using historical data. Based on the data, we can conclude that both models performed admirably in predicting the direction of the stock index. For text analysis, The accuracy score for The Random Forest Model with 1 grams model was 84.33 percent. Chin Kim et al.[4] proposed the system with the help of the In This study gives an overview of the use of Artificial Neural Network (ANN) in stock market forecasting. Although Neural Network (NN) has produced respectable results, many researchers are still attempting to modify The Stock Market Prediction Accuracy by employing a hybrid method fetching into consideration additional external elements to make much high fidelity predictions. It is common knowledge that a variety of external factors play a role in the non-linear, volatile, and dynamic stock market environment. People who desire to conduct research on stock market prediction through NN can use this paper as a starting point

RaduI com in et al.[5] proposed the system with the help of the In This study produced the latest algorithms for predicting stock market movements. As demonstrated in this research, on that point are a variety of ways for forecasting signals using machine learning algorithms including mathematical methods. The principal component-Support Vector Machine (PCA-SVM) was used to

remove incorrect predictions as well as to decide which features are relevant. When compared to simple SVM approaches and moving to PCA-SVM and Support Vector Machine(GA-SVM), The solution for primary challenges and sub-issues becomes more than exhibits and efficient promising outcomes for real-world forecasting utilizing New data sets.

Omkar Chikane et al.[6] proposed the system with the help of the In This study An Investigation of Various Data Mining Techniques for Crop Yield Prediction. In India, the majority of people’s income comes from agriculture and industries that are related to it. The nation’s primary industry for earning money is agriculture. It is also among the nations affected by severe natural disasters like drought and agricultural devastation.

This prediction system is a great choice for data mining applications since predictions of crops deal with a wide variety of databases. Data mining allows us to draw knowledge from vast amounts of data. This report describes a study of all of the different data mining approaches used to estimate crop productivity.

The performance of a particular crop yield prediction system is strongly dependent on how correctly the characteristics are extracted and how well classifiers are used. This document summarizes the findings of numerous algorithms employed by various authors for agricultural production prediction, as well as their accuracy and recommendations.

S. Ahmad et al. [7] proposed the system with the help of the In This study Agriculture now is not the same as it was in the past. The substantial Climatic changes caused by numerous factors such as global warming make understanding climatic conditions challenging. As a result, farmers are unable to determine which crop to plant to increase productivity. The Farmer will be able to take the proper Crop at the exact site by knowing the climate and soil conditions utilizing these data mining systems, which will increase yields. Farmers can easily determine which crop to grow amid unpredictable weather circumstances. Using data mining technologies, this research will aid in the resolution of various agricultural issues.

In recent years, forecasting stock price movements have gained significant research attention [9]. Ariyo et al. [10] utilized the ARIMA model [17] to predict stock prices using data from NYSE and NSE. Deju Zhang and Xiaomin Zhang successfully predicted the stock market index in 2009 with

Markov properties and analyzed stock prices within the Markov model, offering a reference model for investors. Scott and Varian [20] and Brodersen et al. [15] employed the BST series model in 2015 to capture trends and seasonality in time series data, enabling future value forecasting. Recent research revealed that social media platforms like blogs and Twitter can notably influence economic [17] and commercial indicators [14]. Sentiment derived from tweets, classified into mood dimensions, exhibited a significant correlation with the DJIA index. To address inaccuracies during special events, we propose a novel BST-LSTM model with sentiment scores for anomaly detection and improved stock market forecasts.

Sentiment analysis plays a crucial role in classifying opinions based on language and determining positive, negative, or neutral views on a topic or news. While traditional share price prediction relies on technical and fundamental analysis, recent research highlights the strong correlation between stock prices and corporate news [24]. Chen and Lazer [26] explored the relationship between Twitter feeds and stock market activity, anticipating market changes and sentiment impact. Zhang’s research linked stock prices and Twitter sentiment, comparing machine learning methods based on correlation outcomes [27]. DL algorithms, particularly LSTM, effectively recognize patterns and correlations for stock market predictions [28]. Sentiment analysis of text data, including events and social opinions, aids trend determination [29]. While lexicon-based approaches show effectiveness, overcoming new word challenges with statistical and machine-learning techniques is crucial [31]. Structured methods provide comprehensive insights into events’ impact on stock market trends [32].

### III. PROPOSED METHODOLOGY

The System is intended to forecast stock prices. The System boundary can be written down as follows:

- 1) The goal of the research area of forecasting time series is to address several problems, especially in the financial industry.
- 2) Negative Regression By creating an input-output mapping function, issues are resolved.
- 3) When compared to conventional Support Vector Regression (SVR), the Least - Square Support Vector Regression (LSSVR) algorithm significantly reduces the amount of computation required while increasing efficiency.
- 4) Firefly Algorithm (FA), world world-inspired meta-heuristic technique, has lately demonstrated exceptional performance in handling a variety of optimization issues.

### A. Prediction system for Machine Learning(ML)

Machine Learning represents the branch of computer science that enables computer systems to form a sense of information in the same manner that human beings do. In a nutshell, machine learning is similar to an artificially intelligent system that uses an algorithm as well as a method to extract patterns from raw data. The primary goal of machine learning is to enable computer systems to learn from experiences without having to be

implicitly programmed or requiring human intervention. One of the most challenging tasks is predicting how the stock markets will perform. There are numerous aspects involved in prediction, such as physical vs. physiological factors, rational vs. irrational behavior, and so on. All of these factors contribute to making share prices unpredictable and difficult to anticipate with great accuracy.

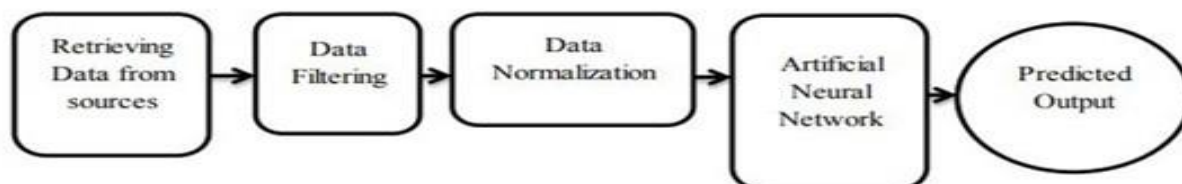


Fig. 1. Prediction System

### B. The System Architecture for Systems

- Data Sources
- This project attempts to predict the stock value concerning the stock’s previous value and trends. It requires historical data on the stock market as the project also emphasizes data mining techniques. So, it is necessary to have a trusted source having relevant and necessary data required for the prediction
- Dataset Creation
- First of all, a dataset is created for training the Artificial neural network. In this study, daily stock

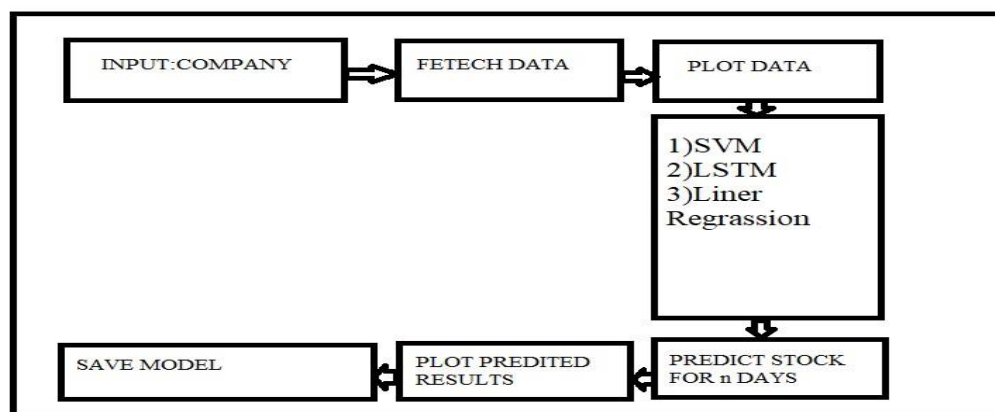


Fig. 2. System Architecture

- Data Normalisation
- Before entry, the data is normalized. The training data input vectors are normalized so that all features have zero mean as well as unit variance. By applying the min-max function, the target values were normalized so that all values are now between 0 and 1. The numbers 0 and 1 represent the minimum as well as maximum values, respectively.

### IV. STOCK PREDICTION SYSTEM

Prediction using LSTM and Bayesian methods in a hybrid manner For forecasting, The system employs a BST time series model with the feature of selections. A statistical method known as the Bayesian Structural Time Series (BSTS) model is utilized for a variety of Statistical tasks, including selecting features, now casting, Time Series Forecasting, causative impact inference, as well as more. The model is intended to be used with the time series data. The concept has also shown

promise in the realm of analytic marketing. The System also uses LSTM to better the Accuracy of data from social group media. Because there might be a delay of an undetermined period between the critical occurrence in the time series, Long Short Memory (LSTM) networks are well suited to categorize, process, and make predictions based on the time series data. Since there can be an interval of undiscovered length between the significant events in the time series, Long-term memory networks are somewhat well suited to classifying, processing, and production forecasting based on time series data. long short-term memories were highly developed to address the difficulty of vanishing change that can arise when training traditional recurrent neural networks. We are interested in inferring the following for this implementation. If we incorporate all of the components

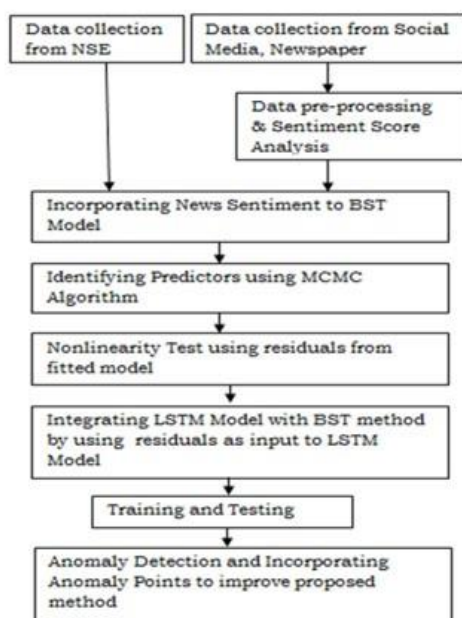


Fig. 3. Stock Prediction System

that influence stock performance and feed them to a machine-learning system with appropriate data preparation and filtering, We will eventually be able to develop a model soon after training the network with data. This improves stock forecasting, has a high degree of accuracy in predicting stock momentum, and boosts financial institutions’ profits. self-contained.

$$RMSE = \sqrt{\frac{1}{N} * \sum_{t=1}^N (At - Ft)^2}$$

Mean Absolute Percentage Error (MAPE), which on the other hand, analyses the error regarding the true significance; instead of taking into consideration an actual difference, it determines how far the predicted values are from

reality. If we can manage too small or too big values, this is a good way to control the error ranges. In these values within the range of 10e6, for example, Root Mean Square Error (RMSE) may blow out of proportion, and MAPE will maintain an error in a fixed range of values.

$$MAPE = \frac{1}{N} * \sum_{t=1}^N \left| \frac{At - Ft}{At} \right|$$

## V. RESULTS & DISCUSSIONS

### A. The Axis Bank Historical Price Visualization

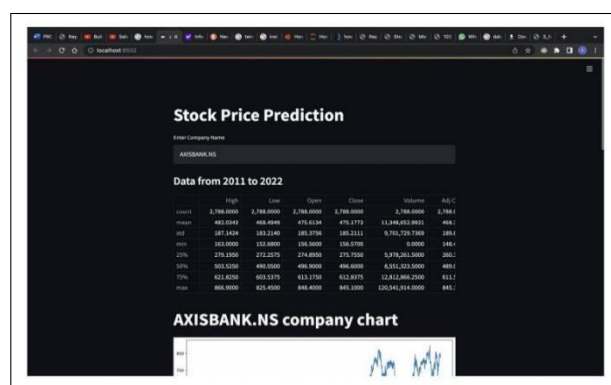


Fig. 4. The Axis Bank Historical Price Visualization

### B. The Axis Bank Dataset

Date	Open	High	Low	Close	Volume	Dividends	Stock Splits
1999-11-18	27.70822016	30.44859644	24.35887553	26.79476357	62546380	0	0
1999-11-19	26.14772471	26.18578675	24.24468944	24.5872364	15234146	0	0
1999-11-22	25.15815086	26.79476357	24.39693595	26.79476357	6577870	0	0
1999-11-23	25.88129697	26.56639082	24.35886765	24.35886765	5975611	0	0
1999-11-24	24.43499033	25.53875193	24.35886789	25.00590134	4843231	0	0
1999-11-26	24.80172751	25.27233499	24.81560667	25.08203125	1729466	0	0
1999-11-29	24.96784832	25.84324447	24.70142213	25.65294075	4074751	0	0
1999-12-01	25.57681711	26.14772825	24.92978512	25.69099998	4310034	0	0
1999-12-01	25.69099408	26.45220877	25.50069041	26.14772224	2957329	0	0
2023-07-05	118.9499969	120.2799988	118.6100006	119.5	1246200	0	0
2023-07-06	116.3600006	118.8899994	115.6600037	118.4300003	1626600	0	0
2023-07-07	117.6699982	119.0800018	116.8899994	117	1493300	0	0
2023-07-10	116.7300034	119.5199966	116.5599976	118.3600006	1889200	0	0
2023-07-11	118.3000031	118.9800034	117.6399994	118.4800034	1096700	0	0
2023-07-12	120	120.4899979	119.1500015	120	1595400	0	0
2023-07-13	120.7600021	120.8700027	118.9899979	119.3499985	1229800	0	0
2023-07-14	119.5400009	120.3899994	118.25	119.75	1225400	0	0
2023-07-17	119.9499969	119.9800034	117.6600037	118.25	1123000	0	0
2023-07-18	118.3799973	119.7900009	118	119.5599976	1362700	0	0
2023-07-19	120.0100021	123.3399963	119.8600006	123.0800018	2425400	0	0
2023-07-20	123.5100021	124.9000015	123.1299973	124.5500031	2278800	0	0
2023-07-21	124.5500031	127.7099991	123.1399994	127.5250015	1964392	0	0

Fig. 5. The Axis Bank Dataset

### C. Axis Bank 100 Days Average The Price



Fig. 6. Axis Bank 100 Days Average The Price

### D. Axis Bank Predicted Price

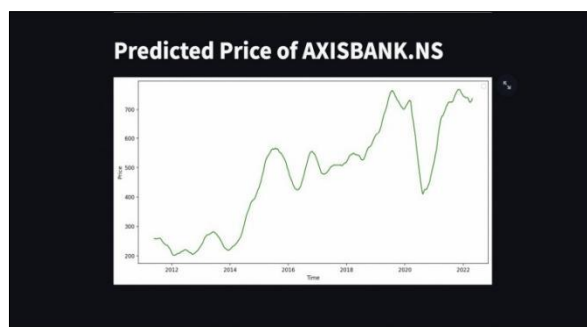


Fig. 7. Axis Bank Predicted Price

E. Infosys Data visualization from 2011 to 2022

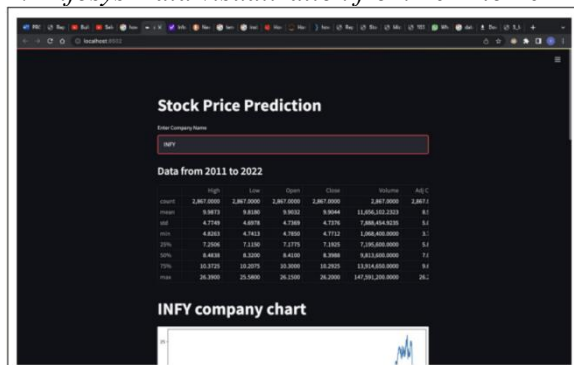


Fig. 8. Infosys Data visualization from 2011 to 2022

F. Infosys Price Chart

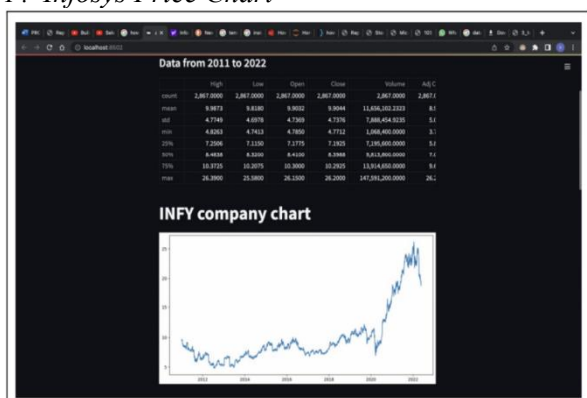


Fig. 9. Infosys Price Chart

G. Infosys Average Price



Fig. 10. Infosys Average Price

H. Infosys.



Fig. 11. Infosys.

I. Infosys The Predicted Price.

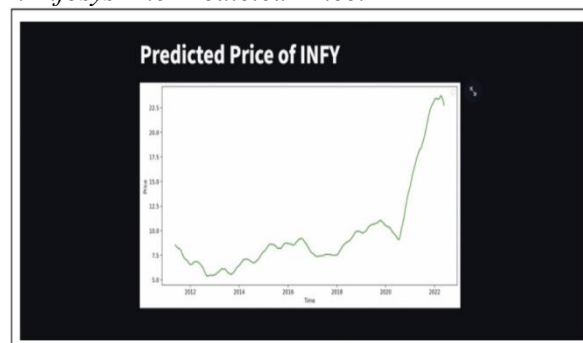


Fig. 12. Infosys The Predicted Price.

VI. CONCLUSION

Accurately predicting the stock’s future price could take a sizable profit. The cost-efficient market hypothesis holds that the Stock Prices accurately represent all currently available information and that any price fluctuations that are not dependent on the most recently disclosed information are, as a result, inherently undetermined. These three market experts hold radically different opinions about the prospects of the current bull market, which has been intensified by the stock market’s recent decline. This research uses various neural networks and machine learning to show how to forecast stock market trends. The outcome will demonstrate how historical data can be used to reasonably anticipate stock movement. Additionally, we shall conclude that LSTM outperforms Backpropagation and SVM after analyzing the results of the T-test. For this execution, we would like to infer the following. If we incorporate all factors that influence the execution of stocks and input them into a neural network(NN) with suitable filtering data preparation, we will eventually be able to develop a model soon after training the network. This improves stock forecasting, has a high degree of accuracy in predicting stock momentum, and boosts financial institutions’ profits. self-contained.

## REFERENCES

1. Ray, Paramita, Bhaswati Ganguli, and Amlan Chakrabarti. "A Hybrid Approach of Bayesian Structural Time Series With LSTM to Identify the Influence of News Sentiment on Short-Term Forecasting of Stock Price." *IEEE Transactions on Computational Social Systems* (2021).
2. Ching-Te Wang, "The Prediction System for Data Analysis of Stock Market by Using Genetic Algorithm" pp. 1721-1725, 2015 IEEE.
3. Sahaj Singh Maini, "Stock Market Prediction using Data Mining Techniques" *Proceedings of the International Conference on Intelligent Sustainable Systems (ICISS 2017)* pp. 654-661, 2017 IEEE.
4. Chang SimVui, Gan Kim Soon, Chin Kim On, "A Review of Stock Market Prediction with Artificial Neural Network (ANN)" 2013 IEEE International Conference on Control System, Computing and Engineering, 29 Nov.-1Dec. 2013, Penang, Malaysia, pp.477-482.
5. Radulacomin, "Stock Market Prediction" 2015 19th International Conference on System Theory, Control and Computing (ICSTCC), October 14-16, Cheile Gradistei, Romania.
6. Yogesh Gadge, Sandhya, "A Study Of Various Data Mining Techniques For Crop Yield Prediction", IEEE, 2017.
7. Omkar Chikane, Nilesh Dumbre, Gitesh More, "System For Agriculture Recommendation Using Data Mining", IERJ, Dec 2015
8. S. Ahmad, A. Lavin, S. Purdy, and Z. Agha, "Unsupervised real-time anomaly detection for streaming data," *Neurocomputing*, vol. 262, pp. 134–147, Nov. 2017.
9. W. Antweiler and M. Z. Frank, "Is all that talk just noise? The information content of Internet stock message boards," *J. Finance*, vol. 59, no. 3, pp. 1259–1294, Jun. 2004.
10. A. A. Ariyo, A. O. Adewumi, and C. K. Ayo, "Stock price prediction using the ARIMA model," in *Proc. 16th Int. Conf. Comput. Model. Simul. (UKSim-AMSS)*, Mar. 2014, pp. 106–112.
11. E. A. Augustine, "Spoons: Netflix outage detection using micro text classification," Dept. Comput. Sci., California Polytech. State Univ., San Luis Obispo, CA, USA, Tech. Rep., 2013.
12. D. Banerjee, "Forecasting of Indian stock market using timeseriesARIMA model," in *Proc. 2nd Int. Conf. Bus. Inf. Manage. (ICBIM)*, Jan. 2014, pp. 131–135.
13. D. Barry and J. A. Hartigan, "A Bayesian analysis for change point problems," *J. Amer. Stat. Assoc.*, vol. 88, no. 421, pp. 309–319, Mar.1993.
14. J. Bollen, H. Mao, and X. Zeng, "Twitter mood predicts the stock market," *J. Comput. Sci.*, vol. 2, no. 1, pp. 1–8, Mar. 2011.
15. K. H. Brodersen, F. Gallusser, J. Koehler, N. Remy, and S. L. Scott, "Inferring causal impact using Bayesian structural time-series models," *Ann. Appl. Statist.*, vol. 9, no. 1, pp. 247–274, Mar. 2015.
16. R. Chalapathy and S. Chawla, "Deep learning for anomaly detection: A survey," 2019, arXiv:1901.03407.
17. M. J. Culnan, P. J. McHugh, and J. I. Zubillaga, "How large US companies can use Twitter and other social media to gain business value," *MIS Quart. Executive*, vol. 9, no. 4, pp. 1–18, 2010.
18. R. Killick and I. A. Eckley, "Changepoint: AnRPackage for changepoint analysis," *J. Stat. Softw.*, vol. 58, no. 3, 2014.
19. J. Roman and A. Jameel, "Backpropagation and recurrent neural networks in financial analysis of multiple stock market returns," in *Proc. 29th Hawaii Int. Conf. Syst. Sci. (HICSS)*, 1996, pp. 454–460.
20. S. L. Scott and H. R. Varian, "Bayesian variable selection for nowcasting economic time series," *Nat. Bur. Econ. Res.*, Cambridge, MA, USA, Working Paper 19567, 2013. [Online]. Available:
21. SentiWordNet: URL. Accessed: Jul. 15, 2019. [Online]. Available: <https://github.com/aesuli/sentiwordnet>.
22. S. Taylor and B. Letham, "Prophet: Forecasting at scale," Facebook, Menlo Park, CA, USA, Tech. Rep., 2017.
23. Y. Yetis, H. Kaplan, and M. Jamshidi, "Stock market prediction by using artificial neural network," in *Proc. World Automat. Congr. (WAC)*, 2014, pp. 718–722.
24. Mohan, Saloni, Sahitya Mullapudi, Sudheer Sammeta, Parag Vijayvergia, and David C. Anastasiu. 2019. "Stock Price Prediction Using News Sentiment Analysis." *IEEE Xplore*. April 1, 2019. <https://doi.org/10.1109/BigDataService.2019.00035>.
25. Pavithya, M. B. D., G. S. D. Perera, S. L. Munasinghe, and S. N. Karunarathna. 2021. "Quantitative Analysis and Sentiment Analysis for Stock Price Forecast: The Case of Colombo Stock Exchange." 2021 10th International Conference on Information and Automation for Sustainability (ICIAfS), August. <https://doi.org/10.1109/iciafs52090.2021.9606034>.

26. Chen, Ray and Marius Lazer. “Analysis of Twitter Feeds for the Prediction of Stock Market Movement.” (2011).
27. Zhang, Linhao. 2013. “Sentiment Analysis on Twitter with Stock Price and Significant Keyword Correlation.” Repositories. lib. texas.edu. 2013. <http://hdl.handle.net/2152/20057>
28. Ar’evalo, Andr’es, Jaime Niño, German Hern’andez, and Javier Sadoval. 2016. “High-Frequency Trading Strategy Based on Deep Neural Networks.” In *Intelligent Computing Methodologies*, 424–36. [https://doi.org/10.1007/978-3-319-42297-8\\_40](https://doi.org/10.1007/978-3-319-42297-8_40). <https://doi.org/10.1109/iiia-aaai.2018.00128>., India, 2020, pp. 1-6, doi: 10.1109/ICCSEA49143.2020.9132928
29. Y Bakshi, Rushlene Kaur, Navneet Kaur, Ravneet Kaur, and Gurpreet Kaur. “Opinion mining and sentiment analysis.” 2016 3rd International Conference on Computing for Sustainable Global Development (INDACom) (2016): 452-455.
30. Musto, Cataldo, Giovanni Semeraro and Marco Polignano. “A Comparison of Lexicon-based Approaches for Sentiment Analysis of Microblog Posts.” *DART@AI\*IA* (2014).
31. Oliveira, Nuno, Paulo Cortez, and Nelson Areal. 2014. “Automatic Creation of Stock Market Lexicons for Sentiment Analysis Using StockTwits Data.” *Proceedings of the 18th International Database Engineering Applications Symposium on - IDEAS ’14*.
32. Ding X, Zhang Y, Liu T, Duan J. Using structured events to predict stock price movement: An empirical investigation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*.2014.