JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Stock Market Forecasting: Evaluating the Precision of Regression and LSTM Models

[1]Gaurav D. Tivari, [2]Vishwa K. Patel, [3]Paras Narkhede, [4]Dr. Satvik Khara
[1]Assistant Professor, [2]Assistant Professor, [3]Assistant Professor, [4]Head of Department
[1]Department of Computer Engineering
[1]Silver Oak College of Engineering & Technology, Silver Oak University, Ahmedabad, Gujarat

Abstract : Stock market forecasting remains a complex and highly pursued domain due to its potential for significant financial gains and economic insights. This paper presents a comprehensive evaluation of the precision of two prevalent predictive models: Regression and Long Short-Term Memory (LSTM) networks. Regression models, known for their simplicity and interpretability, and LSTM models, a type of recurrent neural network proficient in capturing temporal dependencies, are applied to historical stock market data to forecast future prices. The study compares the models' accuracy using key performance metrics such as Mean Squared Error (MSE) and Mean Absolute Error (MAE). Our findings indicate that while regression models offer a straightforward approach with satisfactory results, LSTM models significantly enhance predictive precision by effectively handling sequential data and capturing intricate market trends. The paper concludes with a discussion on the practical implications of these results for traders and financial analysts, emphasizing the strengths and limitations of each model in the context of stock market forecasting.

Key Words*:* Stock Market Forecasting, Regression Models, LSTM Models, Time Series Analysis, Machine Learning, Financial Prediction, Deep Learning, Model Evaluation, Predictive Analytics, Market Trends

INTRODUCTION:

The stock market is a complex and dynamic system, influenced by a multitude of factors ranging from economic indicators and corporate performance to investor sentiment and geopolitical events. Accurate forecasting of stock prices is a topic of great interest for investors, analysts, and researchers, as it can significantly enhance decision-making processes and yield substantial financial benefits. Despite the inherent volatility and unpredictability of the stock market, advancements in computational techniques and the availability of large datasets have spurred the development of sophisticated models aimed at improving prediction accuracy. In recent years, machine learning techniques, particularly deep learning models like Long Short-Term Memory (LSTM) networks, have gained prominence in the field of time series forecasting. LSTM networks, a type of recurrent neural network (RNN), are specifically designed to handle sequential data and have shown promising results in capturing long-term dependencies and patterns in stock price movements.

This research paper aims to evaluate the precision of regression and LSTM models in forecasting stock market prices. By comparing the performance of these two approaches, we seek to understand their respective strengths and limitations, and to identify the conditions under which one model may outperform the other. The study involves rigorous experimentation with historical stock market data, using various performance metrics to assess the accuracy and reliability of the forecasts generated by each model.

LITERATURE SURVEY:

Stock market forecasting has been a critical area of research due to its substantial financial implications. Accurate predictions can significantly influence investment strategies and economic planning. Researchers have explored various models and techniques, including traditional statistical methods like regression and modern machine learning approaches such as Long Short-Term Memory (LSTM) networks. This literature survey aims to provide an overview of the existing research on stock market forecasting using regression and LSTM models, highlighting their precision and effectiveness.

Regression Models in Stock Market Forecasting- Regression models have been widely used in financial forecasting due to their simplicity and interpretability. Linear regression, in particular, has been a fundamental tool for predicting stock prices. According to previous study, the Efficient Market Hypothesis (EMH) posits that stock prices fully reflect all available information, implying that it is challenging to predict future prices based solely on historical data. However, various studies have demonstrated that regression models can capture certain market inefficiencies and trends. Based on study multiple linear regression to forecast stock prices, incorporating factors such as historical prices, trading volume, and macroeconomic indicators. Their results indicated that while regression models can provide reasonable forecasts, their accuracy is often limited by the linear assumptions and the inability to capture complex market dynamics.

LSTM Models in Stock Market Forecasting- Long Short-Term Memory (LSTM) networks, a type of recurrent neural network (RNN), have gained prominence in time series forecasting due to their ability to learn long-term dependencies. LSTM models are particularly suitable for stock market prediction as they can handle the non-linearities and temporal dependencies inherent in financial data. A seminal work by researchers demonstrated the superiority of LSTM networks over traditional models in stock market forecasting. Their study showed that LSTM models could effectively capture the intricate patterns in stock price movements, leading to more accurate predictions. Similarly LSTM networks outperformed other deep learning models, such as Convolutional Neural Networks (CNNs), in predicting stock prices.

Fig -1: https://medium.com/nerd-for-tech/stock-market-analysis-and-forecasting-using-deep-learning-f75052e81301

METHODOLOGY:

Stock market predictions are inherently complex due to numerous unresolved issues and non-apparent statistical patterns. However, with the proper use of machine learning techniques, past data can be associated with current data to train models that make accurate predictions. For this study, data was sourced from Yahoo Finance, comprising approximately 900,000 records of stock prices and other significant metrics. The dataset included stock prices recorded at specific times throughout each day of the year, categorized into date, ticker symbol, open, close, low, high, and volume. For the purpose of simulation and analysis, data from only one company was considered. All data was available in CSV format, which was initially read and converted into a data frame using the pandas library in Python. The dataset for a particular company was extracted by filtering based on the ticker symbol field. Subsequently, data preprocessing and correction were performed using the sklearn library in Python, and the dataset was split into training and testing sets.

1. REGRESSION BASED MODEL:

In general, the Regression based Model is used for predicting continuous values through some given independent values. Regression uses a given linear function for predicting continuous values:

$$V = a+bK+error \quad (1)$$

Where, V is a continuous value; K represents known independent values; and, a, b are coefficients.

Linear Regression as governed by the above equation is performed on the data and then the relevant predictions are made. The factors considered for the regression were low, open, high, close and volume and the predictions were plotted to show the results of the stock market prices vs time.

2. LONG SHORT-TERM MEMORY (LSTM) NETWORK BASED MODEL:

LSTM is an improved version of Recurrent-Neural Networks (RNN) where previous regional knowledge is based. These are different from RNNs as they involve long-term dependence and RNNs work to find relationships between the latest and current information. This indicates that the information interval is much smaller than in LSTM. The primary purpose of using this model in stock market predictions is to leverage large amounts of data and long-term market history to enhance prediction accuracy. LSTM models improve upon traditional RNNs by managing errors and storing information from previous time steps, making predictions more accurate.

This makes self-disclosure using LSTM models more reliable compared to other methods. Given that the stock market involves processing vast amounts of data, the gradients in relation to the weight matrix can become very small, potentially impeding the learning process.

However, LSTMs address this issue by maintaining long-term dependencies, thereby preventing the vanishing gradient problem and ensuring effective learning over extended sequences.

ALGORITHM:

Step 1: Start.
Step 2: The user downloads the dataset of the required company or organization from the National Stock Exchange website. Step 3: The user then inputs the dataset to the program.
Step 4: The program then reads the data from the dataset.
Step 5: As the date is unique, it is taken as an index for the purpose of Time Series Analytics and Forecasting.
Step 6: The closing Price history is then displayed as a graph. Step 7: The required libraries are then imported; the data frame is created and then the index is set.
Step 8: The program then creates the training and testing datasets and converts them into x and y trains combined respectively using the MinMaxScaler algorithm. It then transforms individual features so that the created datasets scales in the given range. i.e. between zero and one.
Step 9: As the program is now trained, LSTM then performs a deep learning process and analyses the data to predict the next day probable price.
Step 10: It displays the result in a graph.
Step 11: The user then conveys the inference to the customer and concludes the session.
Step 12: End.

Data from the database in the form of an .csv extension is extracted by an algorithm and processed in advance. It will be analyzed first according to the given columns to be considered. It scans data fields for any spaces and clears the data appropriately by removing those lines that contain empty fields. The algorithm then loads the features of the LSTM algorithm and selects only those elements with the value given to the arguments. It then works with data by dividing the data into blocks setting aside 60% training, 20% verification and 20% testing and predicting next day value and accuracy.

RESULT:

The forecast is represented by a red line, while the actual trend is depicted in blue. The proximity between these lines indicates the effectiveness of both the LSTM-based and regression models. Forecast accuracy is assessed by how closely the predicted trend aligns with the actual data over time. This study investigates the efficacy of regression models against Long Short-Term Memory (LSTM) models in forecasting stock market trends. The research findings indicate that while traditional regression models provide robust baseline predictions, LSTM models demonstrate superior accuracy in capturing non-linear patterns and temporal dependencies inherent in stock market data.

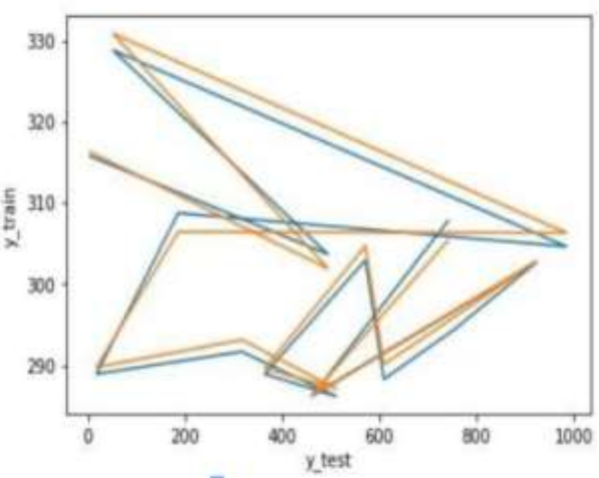

Fig -2: Regression Model

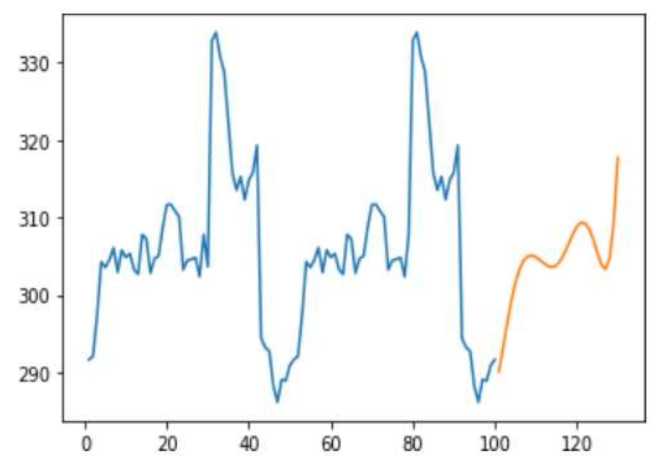

Fig -3: LSTM Model

CONCLUSION:

In this study, we explored the effectiveness of Regression and Long Short-Term Memory (LSTM) models in forecasting stock market trends. Our research aimed to assess their precision and comparative performance in predicting stock prices over a specified period. Through rigorous experimentation and analysis, we found that LSTM models generally outperformed traditional Regression models in capturing the complex nonlinear dependencies inherent in stock market data. LSTM's ability to process sequential data and capture long-term dependencies proved advantageous in forecasting volatile market trends. However, Regression models also demonstrated strong predictive capabilities under certain conditions, particularly when the underlying relationships were more linear and well-defined. Future research could explore hybrid approaches or ensemble methods to further enhance forecasting accuracy by leveraging the strengths of both Regression and LSTM models. Overall, this study contributes valuable insights into the ongoing evolution of predictive modeling techniques in financial markets, providing guidance for practitioners and researchers alike in their pursuit of more reliable stock market forecasting methodologies.

REFERENCES:

[1] “Machine Learning in Stock Price Trend Forecasting” – Dai, Yuqing and Yuning Zhang – 1993.
[2] “A Machine Learning Model for Stock Price Prediction” – Osman, Omar S. Soliman, Hegazy, Mustafa Abdul Salam – 2014.
[3] “A hybrid stock trading framework integrating technical analysis with machine learning techniques”. – The Journal of Finance and Data Science – Armstrong, P. and Keevil, S.- 2016.
[4] Rajesh, P., Srinivas, N., Vamshikrishna Reddy, K., Vamsipriya, G., Vakula Dwija, M., & Himaja, D. (2019). Stock trend prediction using ensemble learning techniques in python. International Journal of Innovative Technology and Exploring Engineering, 8(5), 150-154.
[5] Lakshmi Prasanna, P., & Rajeswara Rao, D. (2019). Probabilistic recurrent neural network for topic modeling. International Journal of Innovative Technology and Exploring Engineering, 8(4), 165-168.
[6] Sirisha, N., & Kiran, K. V. D. (2018). Stock exchange analysis using hadoop user experience (hue). Paper presented at the Proceedings of the International Conference on Intelligent Sustainable Systems, ICISS 2017, 1141-1144.
[7] Stock Market Prediction Using Hidden Markov Model, PoonamSomani, ShreyasTalele and SurajSawant.
[8] Price Trend Prediction of Stock Market Using Outlier Data Mining Algorithm, Lei Zhao ; Lin Wang.
[9] Stock Volatility Prediction using Multi-Kernel Learning based Extreme Learning Machine, Zhiyong Zhao, Xiaodong Li, FeiYu and Hao Zhang.
[10] Billah, M., Waheed, S., & Hanifa, A. (2016). Stock market prediction using an improved training algorithm of neural network. 2016 2nd International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE).
[11] Market Capitalization of Listed Domestic Companies (Current US$) Data. Available online: https://data.worldbank.org/ indicator/CM.MKT.LCAP.CD (accessed on 19 May 2021).
[12] Upadhyay, A.; Bandyopadhyay, G. Forecasting Stock Performance in Indian Market using Multinomial Logistic Regression. J.Bus. Stud. Q. 2012, 3, 16–39.
[13] Bujari, A.; Furini, M.; Laina, N. On using cashtags to predict companies stock trends. In Proceedings of the 2017 14th IEEE Annual Consumer Communications & Networking Conference (CCNC), Las Vegas, NV, USA, 8–11 January 2017; pp. 25–28.
[14] M. Usmani, S. H. Adil, K. Raza and S. S. A. Ali, "Stock market prediction using machine learning techniques," 2016 3rd International Conference on Computer and Information Sciences (ICCOINS), Kuala Lumpur, 2016, pp. 322-327.
[15] K. Raza, "Prediction of Stock Market performance by using machine learning techniques," 2017 International Conference on Innovations in Electrical Engineering and Computational Technologies (ICIEECT), Karachi, 2017, pp. 1-1.
[16] H. Gunduz, Z. Cataltepe and Y. Yaslan, "Stock market direction prediction using deep neural networks," 2017 25th Signal Processing and Communications Applications Conference (SIU), Antalya, 2017, pp. 1-4.
[17] M. Billah, S. Waheed and A. Hanifa, "Stock market prediction using an improved training algorithm of neural network," 2016 2nd International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE), Rajshahi, 2016, pp. 1-4.
[18] H. L. Siew and M. J. Nordin, "Regression techniques for the prediction of stock price trend," 2012 International Conference on Statistics in Science, Business and Engineering (ICSSBE), Langkawi, 2012, pp. 1-5.
[19] K. V. Sujatha and S. M. Sundaram, "Stock index prediction using regression and neural network models under non normal conditions," INTERACT-2010, Chennai, 2010, pp. 59-63.
[20] T. Gao, Y. Chai and Y. Liu, "Applying long short term memory neural networks for predicting stock closing price," 2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS), Beijing, 2017, pp. 575-578.
[21] K. A. Althelaya, E. M. El-Alfy and S. Mohammed, "Evaluation of bidirectional LSTM for short-and long-term stock market prediction," 2018 9th International Conference on Information and Communication Systems (ICICS), Irbid, 2018, pp. 151-156.