Journal of Business and Management Studies
ISSN: 2709-0876
DOI: 10.32996/jbms
Journal Homepage: www.al-kindipublisher.com/index.php/jbms

JBMS
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

Deep Learning Models for Stock Market Forecasting: A Comprehensive Comparative Analysis

Md Salim Chowdhury[1], Norun Nabi[2], Md Nasir Uddin Rana[3], Mujiba Shaima[4], Hammed Esa[5], Anik Mitra[6], Md Abu Sufian Mozumder[7], Irin Akter Liza[8], Md Murshid Reja Sweet[9] ✉ and Refat Naznin[10]

[1]*College of Graduate and Professional Studies, Trine University, USA*
[2]*Department of Information Technology, Washington University of Science and Technology, Alexandria, Virginia, USA*
[34]*Department of Computer Science, Monroe College, New Rochelle, New York, USA*
[5]*Department of Business Administration, International American University Los Angeles, California, USA*
[6]*College of Engineering and Science, Louisiana Tech University, Ruston, Louisiana*
[7]*College of Business, Westcliff University, Irvine, California, USA*
[8]*College of Graduate and Professional Studies (CGPS), Trine University, USA*
[9]*Department of Management Science and Quantitative Methods, Gannon University, USA*
[10]*Department of Business Administration National University, Gazipur, Bangladesh*

Corresponding Author: Md Murshid Reja Sweet, **E-mail**: sweet006@gannon.edu

| ABSTRACT

This study presents a comprehensive comparative analysis of deep learning models for stock market forecasting using data from two prominent stock exchanges, the National Stock Exchange (NSE) and the New York Stock Exchange (NYSE). Four deep neural network architectures—Multilayer Perceptron (MLP), Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), and Convolutional Neural Network (CNN)—were trained and tested on NSE data, focusing on Tata Motors in the automobile sector. The analysis included data from sectors such as Automobile, Banking, and IT for NSE and Financial and Petroleum sectors for NYSE. Results revealed that the deep neural network architectures consistently outperformed the traditional linear model, ARIMA, across both exchanges. The Mean Absolute Percentage Error (MAPE) values obtained for forecasting NSE values using ARIMA were notably higher compared to those derived from the neural networks, indicating the superior predictive capabilities of deep learning models. Notably, the CNN architecture demonstrated exceptional performance in capturing nonlinear trends, particularly in recognizing seasonal patterns within the data. Visualizations of predicted stock prices further supported the findings, showcasing the ability of deep learning models to adapt to dynamic market conditions and discern intricate patterns within financial time series data. Challenges encountered by different neural network architectures, such as difficulties in recognizing certain patterns within specific timeframes, were also analyzed, providing insights into the strengths and limitations of each model.

| KEYWORDS

Deep Learning Models; Stock Market Forecasting; Multilayer Perceptron; Recurrent Neural Networks; New York Stock Exchange

| ARTICLE INFORMATION

ACCEPTED: 15 March 2024 **PUBLISHED:** 02 April 2024 **DOI:** 10.32996/jbms.2024.6.2.9

1. Introduction

The stock market is a complex and dynamic ecosystem where investors navigate through fluctuating trends and unpredictable shifts. Accurately forecasting stock prices amidst this volatility remains a formidable challenge, especially with the inherent non-

linearity of market data. Traditional approaches, such as statistical models and technical indicators, often fall short of capturing the intricate patterns and sudden changes characteristic of stock market movements.
In recent years, however, the advent of deep learning models, particularly deep neural networks, has revolutionized stock market prediction. These models leverage advanced data processing techniques to uncover hidden patterns from raw data, offering a promising solution to the challenges posed by traditional methods. Deep reinforcement learning (DRL), in particular, has emerged as a novel approach for predicting stock market movements by learning optimal decision-making strategies based on historical data.

This paper explores the application of deep learning techniques, with a focus on deep neural networks and DRL, in forecasting stock market trends. Drawing from a comprehensive review of the literature, we delve into the methodologies employed in training and testing neural network architectures using real-world stock market data. Specifically, we investigate the performance of various deep learning models, including Multilayer Perceptron (MLP), Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), and Convolutional Neural Network (CNN), in predicting stock prices across different sectors and markets.

The methodology section outlines the process of training neural networks using historical data from the National Stock Exchange (NSE) of India, with a diverse set of stocks representing the Automobile, Banking, and IT sectors. We discuss the importance of data normalization and parameter tuning in optimizing the performance of deep learning models. Additionally, we conduct sensitivity analyses to assess the impact of network designs and hyperparameters on prediction accuracy.

Furthermore, we present the results and discussions of our experiments, highlighting the comparative performance of deep learning models against traditional linear models like ARIMA. Through a detailed analysis of Mean Absolute Percentage Error (MAPE) and visualizations of predicted stock prices, we demonstrate the superior predictive capabilities of deep neural networks in capturing non-linear trends and adapting to dynamic market conditions.

In conclusion, we underscore the potential of deep learning models as a transformative tool in stock market forecasting, offering insights into future market trends and aiding investors in making informed decisions. We discuss the challenges and limitations associated with implementing deep learning techniques and propose avenues for future research to enhance the accuracy and robustness of predictive models. Overall, this interdisciplinary study bridges the gap between finance and machine learning, paving the way for advancements in financial forecasting and risk management.

2. Literature Review

Khan et al. (2023) explore the application of deep reinforcement learning (DRL) as a novel method for predicting stock market movements. Traditional approaches to stock price prediction often rely on statistical models or technical indicators, which may struggle to capture the non-linear patterns and sudden shifts in stock prices. In contrast, deep reinforcement learning has gained prominence for its ability to learn intricate patterns from raw data and make sequential decisions. The proposed framework for stock price prediction incorporates a deep neural network as a function approximator. This neural network is trained using the Q-learning algorithm, allowing it to learn optimal actions for buying, selling, or holding stocks based on historical stock price data. The neural network outputs Q-values, representing expected rewards for different actions at each time step. The decision-making process involves selecting the best course of action in each market state based on these Q-values.

A sensitivity analysis was conducted to explore the impact of various network designs and hyperparameters on the effectiveness of the DRL-based strategy. The findings highlighted the significant influence of hyperparameters, such as learning rate and exploration rate, on performance. Fine-tuning these hyperparameters proved to be crucial in further improving prediction accuracy.

3. Methodology

The study's approach involves initially training neural networks using data specifically from Tata Motors on the NSE, then evaluating their performance on various companies from both the NSE and NYSE. The dataset covers stocks from the Automobile, Banking, and IT sectors, focusing on highly traded ones. Results indicate that the chosen neural network architectures surpass the traditional linear model, ARIMA, particularly in capturing the nonlinear patterns present in financial time series data. During training, window sizes are carefully considered, with a window size of 200 deemed optimal for predicting 10 days ahead. Comparative analysis with ARIMA underscores the neural networks' superior ability to detect intricate nonlinear trends, highlighting their advantage over traditional linear models.

3.1. Artificial Neural Network and Feed Forward Network

Artificial Neural Networks (ANNs), as described in the reference, function similarly to biological neurons, aiming to identify patterns in data and generalize from them. ANNs are recognized for their nonlinear statistical capabilities, effectively modeling complex relationships between inputs and outputs. One major advantage of ANNs is their ability to learn underlying data patterns, a feature often absent in traditional methods. Typically, ANNs comprise three layers: input, hidden, and output layers. Nonlinear activation functions are applied in all nodes within the hidden and output layers. Each node in the input layer connects with every neuron in the subsequent hidden layer, followed by the output layer. The feed-forward network, illustrated in the reference, is equivalent to the Multilayer Perceptron (MLP) and serves as a straightforward example of a neural network. The connection between input neurons and subsequent hidden layer neurons is established through a weighted matrix denoted as A1B2. The network consists of three layers: input, hidden, and output layers. Artificial neurons, as mentioned in the reference, exist in both the hidden and output layers. These neurons receive inputs from the preceding layer, with the important distinction that neurons within the same layer do not connect with each other; rather, they establish connections with neurons in the subsequent layer.

3.2 Dataset

The dataset utilized in this study comprises highly traded stocks from three distinct sectors: Automobile, Banking, and IT, specifically Maruti, Axis Bank, and Hcltech from the NSE. Each dataset contains information such as stock symbol, series, date, and various price metrics like previous closing, opening, high, low, last closing, average prices, total traded quantity, turnover, and number of trades. The analysis focuses on extracting the daily closing prices of each stock, as investors typically base their decisions on buying or selling stocks on these prices. The chosen training dataset is TATA MOTORS from the Automobile sector, covering the period from January 1, 1996, to June 30, 2015, including closing prices over 4,861 days ranging from 58.79 to 1365.15. Normalization is applied to standardize the data range between 0 and 1. Normalizing the data is crucial for handling stock data from different markets, aiming to standardize the data range. This normalization process is conducted using a specific equation.

$Q_{norm} = (Q - Q_{min}) / (Q_{max} - Q_{min}$--------(1)

In this context, Qnorm signifies the normalized value, while Qmin and Qmax represent the minimum and maximum values within the training dataset. The normalization equation is utilized to obtain normalized data, which is then used as input for the network. A fixed window size of 200 days is employed to predict stock values for the subsequent 10 days. After generating output from the network, a de-normalization process is applied to obtain the original predicted values. The network undergoes training for 1000 epochs. Determination of the optimal window size involves calculating errors for various window sizes ranging from 50 to 250. Among these, a window size of 200 is identified as optimal, displaying the minimum error compared to other window sizes.

Table 1 illustrates the Mean Absolute Percentage Error (MAPE) for different window sizes and prediction days. The rows represent window sizes (50, 100, 150, 200, and 250), while the columns represent prediction days (10, 20, 30, and 40). Analysis reveals that the lowest MAPE occurs with a window size of 200 for a 10-day prediction period. Therefore, a window size of 200 is selected for predicting 10 days.

Table 1 provides details on the window size and the range of days for which predictions are made.

PREDICTION DAYS	50	100	150	200	250
10	7.5 6.	5.34	4.62	4.17	4.18
20	8.53	6.05	5.88	5.61	5.16
30	9.49	7.84	6.97	5.11	5.86
40	10.65	9.34	6.84	5.86	7.06

During the testing phase, data from the primary sectors—Automobile, Banking, and IT—were selected, corresponding to stocks including Maruti, Axis Bank, and HCL Technologies, respectively. Similar to the training phase, we extracted day-wise closing prices for each stock, normalized the data, and applied a de-normalization process to the predicted outputs, following the procedures used in the training dataset. The test datasets covered the period from October 5, 2007, to June 30, 2017. Determination of the input window size for stock data during testing involved error calculations across various window sizes, resulting in a window size of 200. To evaluate the accuracy of the predicted output, mean absolute percentage error (MAPE) was utilized and calculated using the specified formula.

3.3 Verifying Dataset.

To evaluate whether the models can identify common patterns across different stock exchanges, predictions were conducted using NYSE stock data obtained from Yahoo Finance. Specifically, attention was given to the top two active stocks on the New York Stock Exchange: Bank of America (BAC) and Chesapeake Energy (CHK). The dataset covers the period from January 3, 2011, to December 30, 2016, with values denoted in US dollars. Only the day-wise closing prices were extracted from this dataset. For the testing phase, day-wise closing prices for each company within the timeframe of January 3, 2011, to December 30, 2016, were selected. Subsequently, the extracted data was normalized using equation (1) to standardize it before presenting it as input to the network.

4. Result and Discussion

The research involved analyzing data from two prominent stock markets: the NSE (National Stock Exchange) and NYSE (New York Stock Exchange). Four deep neural network architectures—MLP, RNN, LSTM, and CNN—were utilized for this analysis. These networks underwent training using NSE data, specifically focusing on Tata Motors in the automobile sector. Following training, the networks were tested using data from both NSE and NYSE. For NSE, data from the automobile, financial, and IT sectors were chosen, while for NYSE, data from the financial and petroleum sectors were selected. The study included the ARIMA model as a linear benchmark to compare linear and non-linear models. A prediction period of 400 days was considered for both ARIMA and neural network models to assess their performance over a specific timeframe. The results obtained from these analyses are presented in Tables 2 and 3.

Table 2 displays the Mean Absolute Percentage Error (MAPE) incurred in forecasting NSE values for MARUTI, HCL, and AXIS BANK using the ARIMA model over 400 days.

COMPANY	MAPE
MARUTI	22.56
HCL	21.70
MARUTI	19.89

Table 2 shows the MAPE obtained for predicting closing price for 400 days using ARIMA model.

COMPANY	RNN	LSTM	CNN	MLP
MARUTI	5.63	6.33	4.30	4.71
HCL	5.45	5.50	4.33	3.74
MARUTI	11.45	4.90	5.32	5.45

Table 3 illustrates the Mean Absolute Percentage Error (MAPE) derived from the neural network for a 400-day prediction period. Upon comparing the results in Table 2 and Table 3, it is evident that the neural network architecture outperforms ARIMA. This difference in performance may be attributed to the fact that ARIMA struggles to discern the nonlinearities inherent in the data, while neural network architectures excel in identifying and capturing these nonlinear trends within the dataset.

Table 4 illustrates the Mean Absolute Percentage Error (MAPE) generated in predicting NSE values for MARUTI, HCL, and AXIS BANK using a Deep Learning (DL) network.

COMPANY	RNN	LSTM	CNN	MLP
MARUTI	8.63	8.36	7.30	6.71
HCL	7.45	7.88	5.33	7.74
MARUTI	9.45	8.16	7.32	8.45

Table 4 presents the Mean Absolute Percentage Error (MAPE) values acquired during the testing of MARUTI, HCL, and AXIS BANK for the period spanning from October 5, 2007, to June 30, 2017.

5. Conclusion

In conclusion, forecasting stock market trends remains a daunting task due to its dynamic and unpredictable nature. Traditional linear models like AR, ARMA, and ARIMA have been commonly used but have limitations when it comes to the complex and fluctuating time series data of the market. The emergence of deep learning models, especially deep neural networks, represents a significant step forward in tackling these challenges. These models offer a more adaptable and comprehensive approach to capturing the intricate patterns inherent in financial data.

Deep neural networks provide promise in overcoming the shortcomings of linear models, thanks to their flexibility and ability to handle non-linear data patterns. However, implementing these models comes with its own set of challenges, including the need for substantial computational resources and the risk of overfitting. Success also depends on the quality and quantity of available data. As financial markets evolve, continuous refinement and adaptation of deep learning techniques will be essential for improving their forecasting accuracy.

This interdisciplinary research, merging finance and machine learning, opens up new avenues for exploration. Advancements in deep learning architectures, combined with a deeper understanding of financial market dynamics, will help refine forecasting models and mitigate risks associated with stock market predictions.

In summary, the fusion of deep learning models with financial forecasting represents a promising frontier in navigating the complexities of the stock market. As these models evolve, their integration into real-world financial decision-making processes has the potential to improve accuracy and efficiency in predicting market trends, benefiting investors and financial institutions in navigating the ever-changing landscape of finance.

Funding: This research received no external funding.
Conflicts of Interest: The authors declare no conflict of interest.
Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations or those of the publisher, the editors, and the reviewers.

References

[1] Sako K, Mpinda BN, Rodrigues PC (2022). Neural Networks for Financial Time Series Forecasting. Entropy (Basel). 2022 May 7;24(5):657. doi: 10.3390/e24050657. PMID: 35626542; PMCID: PMC9141105.

[2] Shen, J., Shafiq, M.O. (2020). Short-term stock market price trend prediction using a comprehensive deep learning system. *J Big Data* 7, 66. https://doi.org/10.1186/s40537-020-00333-6

[3] Singh, T., Kalra, R., Mishra, S. (2023). An efficient real-time stock prediction exploiting incremental learning and deep learning. *Evolving Systems* 14, 919–937 (2023). https://doi.org/10.1007/s12530-022-09481-x

[4] Murkute, A., & Sarode, T. (2015). Forecasting market price of stock using artificial neural network. *International Journal of Computer Applications*, *124*(12), 11-15.

[5] Chong, T. T. L., & Ng, W. K. (2008). Technical analysis and the London stock exchange: testing the MACD and RSI rules using the FT30. *Applied Economics Letters*, *15*(14), 1111-1114.

[6] Huang, W., Nakamori, Y., & Wang, S. Y. (2005). Forecasting stock market movement direction with support vector machine. *Computers & operations research*, *32*(10), 2513-2522.

[7] Mizuno, H., Kosaka, M., & Yajima, H. (1998). Application of Neural network to technical analysis of stock market prediction, Retrieved Sep. 8, 2004.

[8] Herrera, M., Torgo, L., Izquierdo, J., & Pérez-García, R. (2010). Predictive models for forecasting hourly urban water demand. *Journal of Hydrology*, *387*(1-2), 141-150.

[9] Vijh, M., Chandola, D., Tikkiwal, V. A., & Kumar, A. (2020). Stock closing price prediction using machine learning techniques. *Procedia computer science*, *167*, 599-606.

[10] Bhavani, D. D., & Sarada, C. (2021). Analysis of Stock Market Price Prediction of Indian Finance Companies using Artificial Neural Network Approaches. *CVR Journal of Science and Technology*, *20*(1), 43-49.