Lecture Notes in Networks and Systems 788

Abhishek Swaroop
Zdzislaw Polkowski
Sérgio Duarte Correia
Bal Virdee *Editors*

Proceedings of Data Analytics and Management

ICDAM 2023, Volume 4

Springer

Lecture Notes in Networks and Systems

Volume 788

The series “Lecture Notes in Networks and Systems” publishes the latest developments in Networks and Systems—quickly, informally and with high quality. Original research reported in proceedings and post-proceedings represents the core of LNNS.

Volumes published in LNNS embrace all aspects and subfields of, as well as new challenges in, Networks and Systems.

The series contains proceedings and edited volumes in systems and networks, spanning the areas of Cyber-Physical Systems, Autonomous Systems, Sensor Networks, Control Systems, Energy Systems, Automotive Systems, Biological Systems, Vehicular Networking and Connected Vehicles, Aerospace Systems, Automation, Manufacturing, Smart Grids, Nonlinear Systems, Power Systems, Robotics, Social Systems, Economic Systems and other. Of particular value to both the contributors and the readership are the short publication timeframe and the world-wide distribution and exposure which enable both a wide and rapid dissemination of research output.

The series covers the theory, applications, and perspectives on the state of the art and future developments relevant to systems and networks, decision making, control, complex processes and related areas, as embedded in the fields of interdisciplinary and applied sciences, engineering, computer science, physics, economics, social, and life sciences, as well as the paradigms and methodologies behind them.

Indexed by SCOPUS, INSPEC, WTI Frankfurt eG, zbMATH, SCImago.

All books published in the series are submitted for consideration in Web of Science.

For proposals from Asia please contact Aninda Bose (aninda.bose@springer.com).

Time Series Forecasting of NSE Stocks Using Machine Learning Models (ARIMA, Facebook Prophet, and Stacked LSTM)

Prabudhd Krishna Kandpal, Shourya, Yash Yadav, and Neelam Sharma

Abstract It is widely recognised and acknowledged among market observers and analysts that the stock market, by its very nature, exhibits a tremendous degree of volatility, resulting in frequent and substantial fluctuations. Consequently, the ability to accurately anticipate and forecast market trends assumes paramount importance when it comes to making well-informed decisions regarding the buying and selling of stocks. To achieve such predictive capabilities, the focus of this particular research endeavour is specifically centred around leveraging advanced machine learning models, including but not limited to AutoRegressive Integrated Moving Average (ARIMA), Prophet, as well as deep learning models such as Long Short-Term Memory (LSTM). Root Mean Squared Error (RMSE) is utilised to assess the performance and efficacy of these models. Therefore, the results emanating from this meticulously conducted study contribute invaluable insights and shed light on the comparative effectiveness of different models within the realm of time series forecasting. Importantly, the prevailing body of evidence strongly supports the notion that deep learning-based algorithms, such as LSTM, hold a distinct advantage over traditional statistical methods like the ARIMA model, thereby reinforcing their superiority in this domain.

Keywords Deep learning · Long Short-Term Memory (LSTM) · AutoRegressive Integrated Moving Average (ARIMA) · Prophet · Time series forecasting · Machine learning

P. K. Kandpal (✉) · Shourya · Y. Yadav · N. Sharma
Department of Artificial Intelligence and Machine Learning, Maharaja Agrasen Institute of Technology, Delhi, India
e-mail: prabudhd2003@gmail.com

N. Sharma
e-mail: neelamsharma@mait.ac.in

A. Swaroop et al. (eds.), *Proceedings of Data Analytics and Management*,
Lecture Notes in Networks and Systems 788,
https://doi.org/10.1007/978-981-99-6553-3_24

1 Introduction

There are numerous techniques available for addressing time series forecasting problems. While several procedures assist in drawing conclusions, they do not guarantee accurate results. To make an informed decision, it is crucial to precisely forecast which option to choose. Therefore, it is necessary to carefully evaluate the pros and cons of each method before applying it.

The ARIMA model is commonly used for stock predictions due to its simplicity, ability to identify time-dependent patterns, and provision of statistical insights. However, it has limitations in capturing nonlinear patterns, requires stationary data, does not consider exogenous variables, and is less accurate for long-term forecasts. Combining ARIMA models with advanced methodologies can improve accuracy and overcome these limitations in real-world usage.

Prophet offers a user-friendly and efficient means of conducting time series analysis for stock forecasting. It provides a quick and simple solution without requiring substantial modification. However, its rudimentary assumptions and limited control might restrict its applicability in complex scenarios that demand better modelling methodologies or detailed integration of external elements.

LSTM models for stock predictions offer benefits such as long-term forecasting, handling sequential data, incorporating exogenous factors, and capturing complex patterns. However, they require a large amount of data, are computationally demanding, are highly susceptible to overfitting, and lack interpretability. Despite these shortcomings, LSTM models are often used in combination with various methods to enhance accuracy for stock forecasting purposes.

This paper follows a clear and organised structure. It begins by introducing the dataset used in the research and providing the necessary background information. The different models utilised are then listed and explained, offering a comprehensive overview of the methodologies employed. The proposed methodology is presented in detail, describing the specific approach taken in the research. Subsequently, the paper presents the results obtained from implementing these models, emphasising the outcomes of the analysis. A thorough examination and analysis of these results are conducted to gain a deeper understanding of the findings. Finally, the paper concludes by summarising the findings of the proposed model, discussing their implications, and acknowledging the limitations and potential for future research in the field.

The motivations of this paper can be summarised as follows:

1. This study aims to compare the performance of three popular forecasting models, namely ARIMA, Facebook Prophet, and LSTM, in predicting stock prices. By analysing the results across multiple companies with varying levels of stability and volatility, the study aims to provide insights into the strengths and weaknesses of each model.
2. Accurate stock predictions are crucial for financial decision-making. The study aims to emphasise the importance of accurate forecasting for long-term predictions and highlights how advanced models like LSTM can enhance accuracy.

3. The study also aims to provide insights into the robustness and adaptability of these models by examining their performance in both stable and volatile market conditions.

Overall, this study contributes to the field of stock price prediction by comparing the performance of different forecasting models and highlighting the potential of deep learning-based algorithms. The findings have practical implications for researchers and practitioners, paving the way for further advancements in the field of time series forecasting.

2 Literature Review

Siami-Namini et al. have suggested that forecasting time series data presents a formidable challenge, primarily due to the ever-evolving and unpredictable nature of economic trends and the presence of incomplete information. Notably, the increasing volatility witnessed in the market over recent years has raised significant concerns when it comes to accurately predicting economic and financial time series. Consequently, it becomes imperative to evaluate the precision and reliability of forecasts when employing diverse forecasting methodologies, with a particular focus on regression analysis. This is crucial since regression analysis, despite its utility, possesses inherent limitations in its practical applications [1]. Pang et al. have discovered that neural networks have found extensive applications across a range of fields, including pattern recognition, financial securities, and signal processing. Particularly in the realm of stock market forecasting, neural networks have garnered considerable acclaim for their effectiveness in regression and classification tasks. Nevertheless, it is important to note that conventional neural network algorithms may encounter challenges when attempting to accurately predict stock market behaviour. One such obstacle arises from the issue of random weight initialisation, which can lead to a susceptibility to local optima and, consequently, yield incorrect predictions [2]. This study aims to accurately predict the closing price of various NSE stocks using machine learning methods like ARIMA, Prophet, and deep learning models, namely the LSTM model. Stock market forecasting has traditionally relied on linear models such as AutoRegressive (AR), AutoRegressive Moving Average (ARMA), and AutoRegressive Integrated Moving Average (ARIMA). However, a notable limitation of these models is their specificity to a particular time series dataset. In other words, a model that performs well for forecasting the stock market behaviour of one company may not yield satisfactory results when applied to another company. This can be attributed to the inherent ambiguity and unpredictable nature of the stock market, which inherently carries a higher level of risk compared to other sectors. Consequently, this inherent complexity and risk associated with stock market prediction significantly contribute to the difficulty of accurately forecasting stock market trends [3]. There are several reasons why deep learning models have come to be significantly successful in comparison to traditional machine learning and statistical models, and their usage has been on the rise for several decades. Models such as the LSTM model possess the ability to take into account the temporal dependencies present in time series data. Secondly, these models are very successful in extracting features from raw data, eliminating the need for manual feature extraction. Moreover, deep learning models are capable of accommodating both univariate and multivariate time series data, with even irregular and unevenly spaced data points. With the latest advancements in parallel computing and GPUs, deep learning models can be trained and optimised on large-scale data. Saiktishna et al. [4] focus on the utilisation of the FB Prophet model for historical analysis of stock markets and time series forecasting. It explores the techniques, conclusions, and limits of previous research in this field. The evaluation emphasises FB Prophet's ability to capture market patterns and seasonality, as well as future research possibilities. It contains useful information for scholars and practitioners who want to use FB Prophet for stock market study and forecasting.

Numerous publications in the literature have made an effort to investigate the hybrid modelling of financial time series movement using various models. He et al. [5] utilised a hybrid model using the ARMA and CNN-LSTM model to accurately predict the financial market by applying it to three different time series with different levels of volatility. They presented that optimisations are still possible to machine learning and deep learning models, given the rapid development in the aforementioned fields. Fang et al. [6] proposed a novel approach using the dual-LSTM approach, which consisted of two LSTM layers with batch normalisation which addressed the problem of sharp point changes by capturing significant profit points using an adaptive cross-entropy loss function, enhancing the model's prediction capabilities. Gajamannage et al. [7] have emphasised the importance of real-time forecasting and presented its importance in risk analysis and management. They have put forth a sequentially trained dual-LSTM model which has addressed the issue of semi-convergence in a recurrent LSTM setup and has validated their results based on various diverse financial markets. Patil [8] highlights the use of machine learning approaches for stock market forecasting, including ARIMA, support vector machines, random forest, and recurrent neural networks (RNNs). The paper explores the strengths and limitations of these models, emphasising the importance of feature engineering and selection in improving prediction accuracy. It also includes case studies and empirical research to show how these models may be used in stock price forecasting.

It is evident from the above examples that LSTM is a robust model that is excellent for the purpose of time series analysis and forecasting. It has displayed its capabilities in various other fields, including energy consumption forecasting, wind speed forecasting, carbon emissions forecasting, and aircraft delays forecasting.

3 Dataset Description

The historical stock data used in this research study have been sourced from Yahoo Finance. The dataset encompasses the stock prices of four prominent companies: Reliance Industries, Tata Steel LLC, ICICI Bank, and Adani Enterprise.

The selection of Reliance Industries, Tata Steel LLC, ICICI Bank, and Adani Enterprise for this research study was based on specific criteria. Reliance Industries, Tata Steel LLC, and ICICI Bank were chosen due to their stable stock performance and a general upward trend observed over time. These stocks have demonstrated consistent growth and are considered relatively stable investments. In contrast, Adani Enterprise was included in the dataset because it is known for its high volatility, with stock prices being heavily influenced by market reports and external factors. By including a mix of stocks with different characteristics, such as stability and volatility, we aim to accurately assess the capabilities of our predictive models in handling various stock market scenarios and understanding their effectiveness in different market conditions.

Within the dataset, two crucial components play a pivotal role in our modelling approach: the 'Close' and 'Date' variables. The 'Close' variable signifies the last recorded price at which a particular stock was traded during the regular hours of a trading session. It serves as the target variable in our predictive models, as we aim to forecast future stock prices based on historical trends and patterns. On the other hand, the 'Date' variable acts as the predictor variable, providing temporal information to aid in the prediction process. For the purpose of this research, the dataset spans from the inception of each respective company until May 1, 2023, thereby covering a substantial period of historical data (Fig. 1).

	Open	High	Low	Close	Volume
Date					
2023-04-24	2375.0	2380.899902	2348.000000	2358.000000	5970048
2023-04-25	2366.0	2380.600098	2350.500000	2376.050049	4262471
2023-04-26	2379.0	2386.100098	2354.050049	2362.100098	3977129
2023-04-27	2375.0	2384.000000	2364.000000	2377.050049	4230627
2023-04-28	2382.0	2423.899902	2381.750000	2420.500000	7183342

Fig. 1 Sample dataset of reliance industries obtained from Yahoo finance

Table 1 Number of time series observations

Stock	Observations		Total
	Train 90%	Test 10%	
Reliance Industries	6182	687	6869
Tata Steel LLC	6184	688	6872
ICICI Bank	4656	518	5174
Adani Enterprise	4658	518	5176

4 Data Preparation

The financial time series datasets were divided into two parts: a training dataset and a test dataset. The training dataset consisted of 90% of each dataset and was used to train the models. The remaining 10% of each dataset was allocated to the test dataset to evaluate the accuracy of the models. The number of time series observations for each dataset is provided in Table 1.

5 Assessment Metric

We have utilised the Root Mean Square Error (RMSE) to evaluate the precision of our model's predictions. RMSE measures the differences or residuals between the predicted and actual values. By employing RMSE, we have been able to compare prediction errors within the same dataset across different models rather than between different datasets. The formula used to calculate RMSE is as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(x_i - \widehat{x_i}\right)^2}, \tag{1}$$

where N is the total number of stocks, x_i is the actual value of the stock, whereas $\widehat{x_i}$ is the value predicted by our model.

6 Models

6.1 ARIMA

ARIMA is a widely used time series forecasting model that combines autoregressive (AR), differencing (I), and moving average (MA) components to capture linear relationships, stationarity, and dependencies within the data [9]. We have made use of rolling ARIMA to perform forecasting. It means that we have refitted the model at

Table 2 ARIMA model parameters

Stock	ARIMA model (p, d, q)	RMSE
Reliance Industries	(5, 2, 0)	41.62
Tata Steel LLC	(2, 1, 2)	18.62
ICICI Bank	(4, 1, 4)	11.28
Adani Enterprise	(4, 2, 4)	83.27

RMSE values are for the entire test set in Table 2

each iteration as new data becomes available. This allowed our model to continuously adapt to the most recent data, enhancing its accuracy and robustness of the forecast. In ARIMA modeling, the notation ARIMA (p, d, q) is commonly used, where [1]:

- 'p' represents the number of lag observations used in training the model (i.e. lag order).
- 'd' denotes the number of times differencing is applied (i.e. degree of differencing).
- 'q' indicates the size of the moving average window (i.e. order of moving average).

To determine the appropriate values for these parameters, we utilised the Autocorrelation Function (ACF) graph and Partial Autocorrelation (PACF) graphs [10]. The ACF graph provided insights into the correlation between the current observation and lagged observations at various time lags. Meanwhile, the PACF graphs helped us assess the correlation between the current observation and the residuals from previous observations, taking into account the effects of intermediate observations. By carefully examining these graphs, we were able to estimate the optimal values for the AR, MA, and differencing components of the ARIMA model for each dataset. They are listed in Table 2.

6.2 Facebook Prophet

Prophet is an open-source library developed by Facebook and is generally used for univariate time series forecasting. Prophet [11] is a decomposable framework that reduces an elaborate problem, such as time series data prediction, into simpler ones and does so by taking three factors into account: seasonality, holidays, and trend.

$$y(t) = g(t) + s(t) + h(t) + \varepsilon. \tag{2}$$

Here, $g(t)$ represents the trend, $s(t)$ represents seasonality, $h(t)$ represents holidays, and t represents the error rate. The trend parameter monitors two additional parameters: saturation growth and change points. Seasonality is another factor that Prophet considers, which uses the Fourier series to create a precise end model.

$$s(t) = \sum_{n=1}^{N}\left(a_n \cos\left(\frac{2\pi nt}{p}\right) + b_n \sin\left(\frac{2\pi nt}{p}\right)\right). \quad (3)$$

$s(t)$ denotes seasonality, and P denotes the time period, which might be monthly, weekly, daily, quarterly, or even annual. N is the frequency of change, and the parameters a_n and b_n are dependent on it. The Prophet [12] is adept at detecting lost data, changing trends, and treating outsiders often. Compared to previous time series forecasting techniques, the prophet makes it evident how it generates a faster forecast that is more precise.

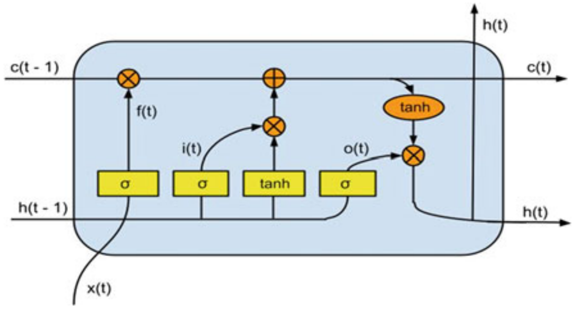
6.3 LSTM

Long Short-Term Memory (LSTM) is a variation of the Recurrent Neural Network (RNN) which is often used in time series forecasting due to its ability to take the temporal dependencies of the time series data into account. Before diving into LSTM, we will first go over the following to develop a better understanding:

(a) Layered-formatted neurons make up the core of Feedforward Neural Networks (FFNNs) [13]. Each neuron updates its values using an optimisation algorithm, such as Gradient Descent, Adam optimiser, and computes values based on randomly initialised weights. FFNNs are loop-free and completely linked. Every FFNN has three layers of neurons: an input layer that receives input from users, a hidden layer that allows the network to learn complex patterns and relationships in data, and an output layer that produces the output based on the input from the last layer. In an FFNN, each layer of neurons feeds information to the layer above it.
(b) Recurrent Neural Networks (RNNs) [1] are special neural networks where the outputs are partially dependent on a series of outputs obtained in the previous stages. The hidden layers in an RNN network work as memory units that hold this information and use it during computation. The only drawback of RNNs is that they are only capable of learning a small number of previous stages, which makes them incapable of remembering long sequences of data. The LSTM model solves this issue by introducing a 'memory' line.
(c) Long Short-Term Memory (LSTM) [14] is an improvement of the RNN model. It is equipped with input, output, and forgetting gates to accommodate the effects of longer time intervals and delays while also solving the problem of vanishing gradient and exploding gradient. The structure of an LSTM cell is given in Fig. 1:

In Fig. 2, $h(t)$ and $h(t-1)$ represent the outputs of the current and previous cell, $x(t)$ represents the input of the current cell, and $c(t)$ and $c(t-1)$ represent the current and previous states of the neuron at t. $i(t)$ represents the input threshold which determines the information gain with the sigmoid function, and $o(t)$ represents the output

Fig. 2 LSTM cell



threshold which determines the output neuron state using the sigmoid function and the tanh activation function. $f(t)$ represents the forgetting threshold which controls the information that is discarded with the help of the sigmoid function.

7 Proposed Methodology

For the purpose of this study, we perform univariate time series forecasting, following these steps:

Data Collection: The historical time series data relevant to your problem were gathered. Factors like data quality, missing values, outliers, and potential seasonality or trends in the data were considered.

Data Preprocessing: The data were prepared for modelling by performing various preprocessing steps. Missing values were handled by deciding on a strategy to fill or impute them, such as forward/backward filling, interpolation, or using statistical methods. Scaling and normalisation techniques were applied to normalise the data to a common scale, such as Min–Max scaling, to improve model convergence and performance.

Train–Test Split: The data were split into training and testing sets. Typically, a larger portion was allocated for training, while a smaller portion was kept for evaluating the model's performance. The split maintains the temporal order of the data to simulate real-world forecasting scenarios.

Model Selection: The appropriate machine learning or deep learning models were chosen for the time series forecasting task. These models were chosen for time series forecasting:

Autoregressive model (ARIMA) captures the dependency of the current observation on previous observations.

Long Short-Term Memory (LSTM) model is a type of RNN model specifically designed to capture long-term dependencies in time series data.

The Prophet model combines the flexibility of generalised additive models with the simplicity of traditional forecasting methods. It incorporates seasonality, trend

Table 3 RMSEs of ARIMA, Prophet, and LSTM models or the last 100 days of forecast

Stock	RMSE			Lowest RMSE
	ARIMA	Prophet	LSTM	
Reliance Industries	32.38	286.16	14.81	LSTM
Tata Steel LLC	11.29	34.63	4.86	LSTM
ICICI Bank	9.73	33.91	9.36	LSTM
Adani Enterprise	155.54	1035.87	81.2	LSTM

changes, and holiday effects, making it effective for predicting time series data with intuitive and interpretable results.

Model Training: The selected model was trained using the training dataset. During training, the model learns to capture patterns, trends, and seasonality in the data. Hyperparameters (e.g. learning rate, batch size, number of layers) were adjusted through experimentation.

Model Evaluation: The trained model's performance on the testing dataset was evaluated using appropriate evaluation metrics for time series forecasting, such as Root Mean Squared Error (RMSE). The model's ability to generalise and make accurate predictions on unseen data was assessed.

Model Refinement: If the initial model's performance was not satisfactory, the model was refined by adjusting hyperparameters, trying different architectures, or employing regularisation techniques to improve the model's accuracy and generalisation capabilities. This step was iterated until the desired performance was achieved.

8 Observations and Results

The experimental findings have been summarised in Table 3, which provides a comprehensive overview of the performance of the three models across the selected set of four stocks (Figs. 3, 4, 5, and 6).

9 Result Analysis

Considering the higher level of comparability between the predictions of the ARIMA and Stacked LSTM models compared to Prophet, we have done a detailed performance evaluation of these two models. To gain more profound insights into the accuracy of these models for stock price prediction, we have conducted our analysis individually for each company. This comprehensive approach has allowed us to thoroughly assess the performance of ARIMA and Stacked LSTM models across all four companies. Additionally, we have focused on the final 30 days of the forecasted

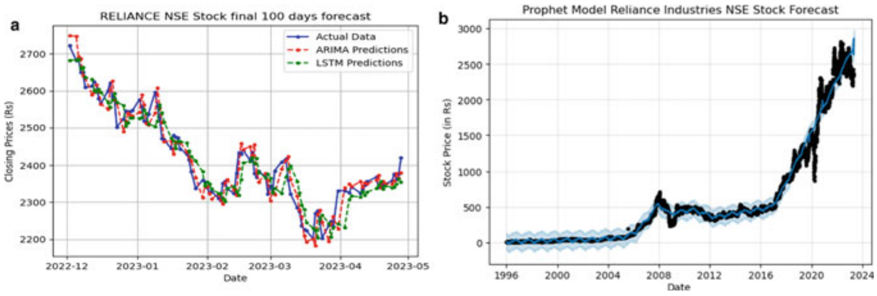


Fig. 3 **a** ARIMA and LSTM models' last 100 days' forecast on RELIANCE NSE stock. **b** Facebook Prophet model's forecast on RELIANCE NSE stock

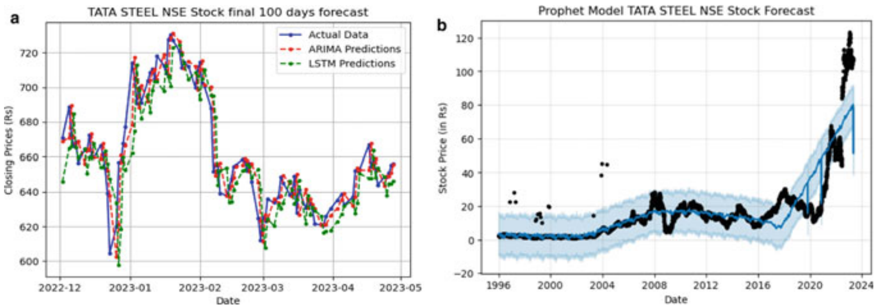


Fig. 4 **a** ARIMA and LSTM models' last 100 days' forecast on TATA STEEL NSE stock. **b** Facebook Prophet model's forecast on TATA STEEL NSE stock

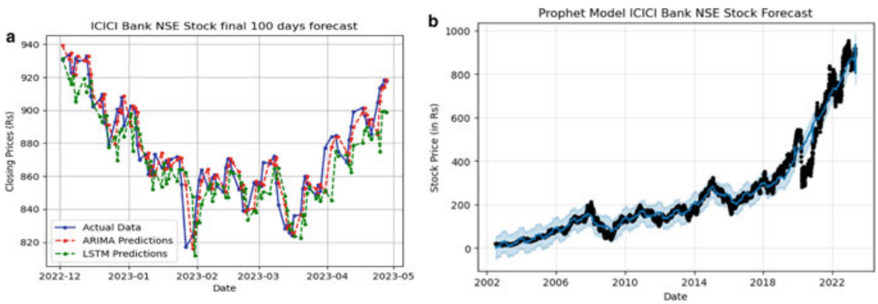


Fig. 5 **a** ARIMA and LSTM models' last 100 days' forecast on ICICI BANK NSE stock. **b** Facebook Prophet model's forecast on ICICI bank NSE stock

period for each company, enabling us to make precise evaluations of these models' effectiveness.

A. Reliance Industries

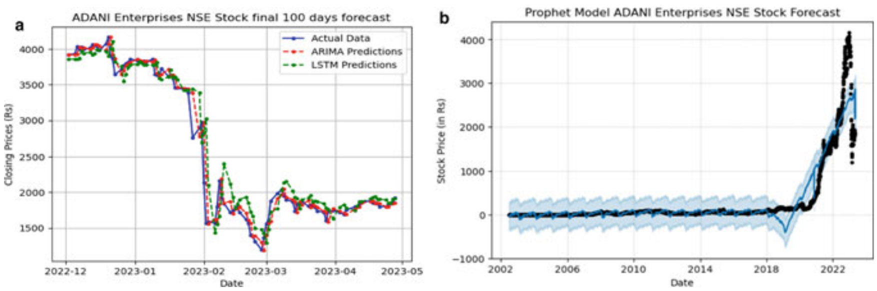


Fig. 6 **a** ARIMA and LSTM models' last 100 days' forecast on ADANI NSE stock. **b** Facebook Prophet model's forecast on ADANI NSE stock

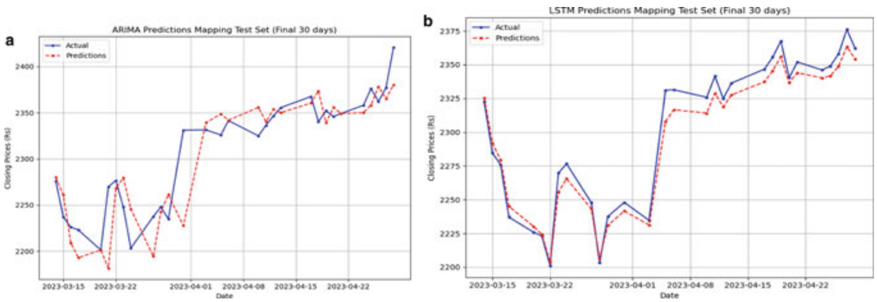


Fig. 7 **a** ARIMA model's last 30 days' forecast on RELIANCE NSE stock. **b** LSTM model's last 30 days' forecast on RELIANCE NSE stock

ARIMA: The ARIMA model for Reliance Enterprises achieved an RMSE of 32.26, indicating an average deviation of approximately Rs 32.26 between the predicted and actual stock prices (Fig. 7).

Stacked LSTM: The LSTM model for Reliance Enterprises achieved a lower RMSE of 9.26, implying an average deviation of approximately Rs 9.26.

Comparing the two models, both LSTM and ARIMA effectively captured the trends in Reliance's stock prices, as evident from the graphs. However, the LSTM model appears to be more accurate in mapping these trends. The significantly lower RMSE of 9.26 for the LSTM model suggests that it better captured the underlying patterns in Reliance's stock prices compared to ARIMA.

B. Tata Steel LLC:

ARIMA: The ARIMA model for Tata Steel achieved an RMSE of 8.1, indicating an average deviation of approximately Rs 8.1 between the predicted and actual stock prices (Fig. 8).

Stacked LSTM: The LSTM model for Tata Steel achieved a lower RMSE of 4.26, implying an average deviation of approximately Rs 4.26.

Comparing the two models, both LSTM and ARIMA effectively captured the trends in Tata Steel's stock prices, as evident from the graphs. However, the LSTM

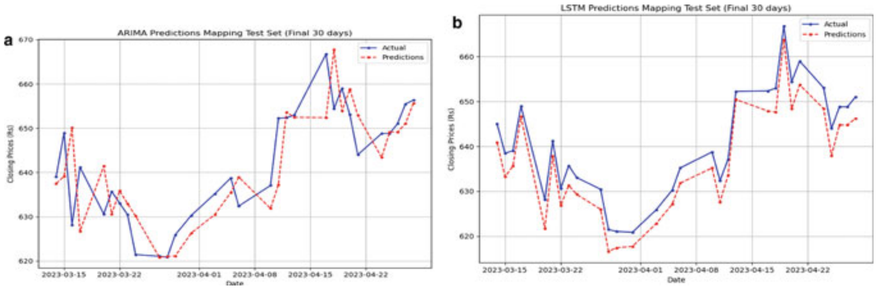


Fig. 8 **a** ARIMA model's last 30 days' forecast on TATA STEEL NSE stock. **b** LSTM model's last 30 days' forecast on TATA STEEL NSE stock

model demonstrated greater accuracy in mapping these trends. The significantly lower RMSE of 4.26 for the LSTM model suggests that it better captured the underlying patterns in Tata Steel's stock prices compared to ARIMA.

C. ICICI Bank:

ARIMA: The ARIMA model for ICICI Bank achieved an RMSE of 8.89, indicating an average deviation of approximately Rs 8.89 between the predicted and actual stock prices (Fig. 9).

Stacked LSTM: The LSTM model for ICICI Bank achieved a lower RMSE of 7.93, implying an average deviation of approximately Rs 7.93.

Comparing the two models, both LSTM and ARIMA effectively captured the trends in ICICI Bank's stock prices, as evident from the graphs. However, the LSTM model demonstrated greater accuracy in mapping these trends. The significantly lower RMSE of 7.93 for the LSTM model suggests that it better captured the underlying patterns in ICICI Bank's stock prices compared to ARIMA.

D. Adani Enterprises:

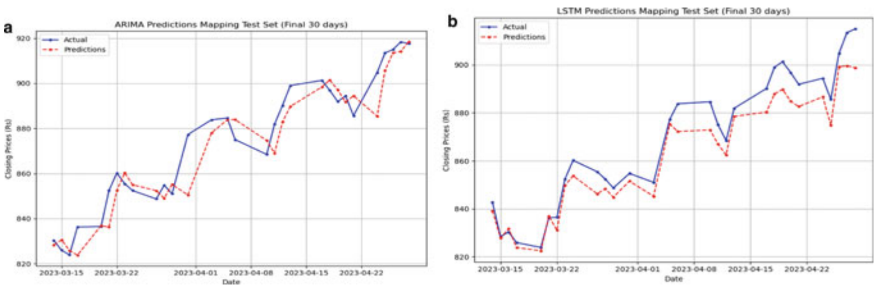


Fig. 9 **a** ARIMA model's last 30 days' forecast on ICICI bank NSE stock. **b** LSTM model's last 30 days' forecast on ICICI bank NSE stock

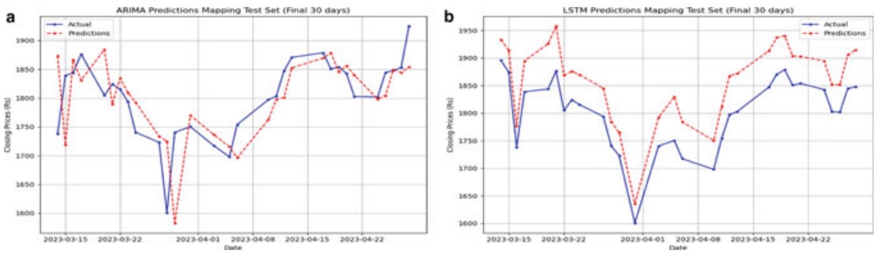


Fig. 10 **a** ARIMA model's last 30 days' forecast on ADANI NSE stock. **b** LSTM model's last 30 days' forecast on ADANI NSE stock

ARIMA: The ARIMA model for Adani Enterprises achieved an RMSE of 58.77, indicating an average deviation of approximately Rs 58.77 between the predicted and actual stock prices (Fig. 10).

Stacked LSTM: The LSTM model for Adani Enterprises achieved a slightly lower RMSE of 58.0, implying an average deviation of approximately Rs 58.0.

Both the LSTM and ARIMA models achieved some success in capturing the trends in Adani Enterprises' stock prices, as seen in the graphs, but had high RMSE values, indicating notable deviations from the actual prices. Adani's stock is highly volatile, influenced by market reports and external factors, making accurate predictions challenging. Although the Stacked LSTM model performed the best among the two, accurately forecasting Adani Enterprises' stock prices still has room for improvement due to the inherent uncertainty and complexity of market dynamics.

10 Limitations

While the findings of this research provide valuable insights into the performance of LSTM, ARIMA, and Facebook Prophet for stock price prediction, there are a few limitations to consider:

1. Lack of External Factors: The models used in this study solely relied on historical stock price data as input. External factors, such as company-specific news, industry trends, or global economic events, were not incorporated into the analysis. These external factors can significantly influence stock prices and might enhance the accuracy of the predictions if considered.
2. Limited Generalizability: The study focused on a specific set of companies, namely Reliance Industries, Tata Steel LLC, ICICI Bank, and Adani Enterprise. The findings may not apply to other stocks or industries. The performance of the models could vary when applied to different datasets with diverse characteristics.
3. Limited Scope of Model Selection: While the research compared LSTM, ARIMA, and Facebook Prophet models, it is important to note that there are several other advanced forecasting algorithms available, such as gradient

boosting machines and Long Short-Term Memory networks with attention mechanisms. Exploring a wider range of forecasting techniques could provide additional insights and potentially reveal alternative models that could yield different results.

Addressing these limitations and conducting further research would enhance the robustness and applicability of the findings, leading to a more comprehensive understanding of the strengths and weaknesses of different forecasting models in stock price prediction.

11 Conclusion

The findings of this study highlight the superior performance of LSTM, a deep learning-based algorithm, in comparison to ARIMA and Facebook Prophet for stock price prediction across Reliance Industries, Tata Steel LLC, and ICICI Bank. These companies, known for their stability, exhibited low RMSE values when analysed using both ARIMA and Stacked LSTM models. However, when applied to the highly volatile stock of Adani Enterprise, the models yielded higher RMSE values. Despite this, the models were successful in capturing the general trend of the stock. It is worth noting that while ARIMA demonstrated good overall performance, LSTM consistently outperformed it in terms of accuracy. The limitations of Facebook Prophet in handling time series with little or no seasonality, such as stock prices, were also evident in this study.

This research highlights the advantages of deep learning-based algorithms in analysing economic and financial data, providing valuable insights for finance and economics researchers and practitioners. It calls for further exploration of these techniques in different datasets containing varying features, expanding our understanding of the improvements that can be achieved through deep learning in various domains.

In summary, this study contributes to the comparative performance analysis of ARIMA, Prophet, and LSTM models in stock price prediction. It supports the notion that deep learning-based algorithms, particularly LSTM, show promise in enhancing prediction accuracy. It also recognises the reliability of ARIMA for stock price prediction and acknowledges the limitations of Prophet for time series lacking strong seasonality.

12 Social Impact

The study conducted in this paper, although not achieving highly precise stock price prediction, has demonstrated the effectiveness of two models: ARIMA and Stacked LSTM, in accurately forecasting market trends. This is evident from the 30-day forecast presented in Sect. 10. While predicting the exact stock price is considered an extremely challenging task, the ability to forecast market trends can provide valuable assistance to society in the following ways:

1. Risk Management: Accurate trend forecasting helps investors and institutions manage stock market risks, optimising investment strategies to minimise losses and maximise returns.
2. Market Timing: Understanding market trends enables effective investment timing and optimising buying and selling decisions to capitalise on opportunities and enhance investment performance.
3. Strategic Planning: Accurate trend forecasting informs businesses' strategic planning, aligning product development, marketing, and expansion strategies with market dynamics for competitive advantage and informed resource allocation.
4. Economic Analysis: Trend forecasting contributes to understanding the overall economy, providing insights into industry health, market sentiments, and potential economic shifts, and aiding policymakers in decision-making.
5. Algorithmic Trading Strategies: The findings of this study can benefit algorithmic trading developers, enhancing trading algorithm performance for more profitable automated strategies.

In conclusion, while precise stock price prediction may be challenging, the ability to forecast market trends, as demonstrated by the ARIMA and Stacked LSTM models in this study, offers significant benefits to society. From risk management and strategic planning to economic analysis, accurate trend forecasting supports informed decision-making in various domains, contributing to better financial outcomes and market understanding.

13 Future Scope

This research opens up several avenues for future investigation and expansion. Here, we outline some potential directions and areas of exploration that can contribute to the advancement of this field:

1. Incorporating External Factors: To enhance the predictive accuracy of the models, future research can consider integrating external factors such as company-specific news, industry trends, macroeconomic indicators, and market sentiment into the analysis. This can provide a more comprehensive understanding of the factors influencing stock prices and improve the models' ability to capture complex market dynamics.
2. Comparing with Other Advanced Forecasting Techniques: While this research focused on LSTM, ARIMA, and Facebook Prophet, there are numerous other advanced forecasting techniques available. Future studies could expand the model selection and compare the performance of additional algorithms, such as gradient boosting machines, support vector machines, or ensemble methods. This comparative analysis can shed light on various forecasting approaches relative to strengths and weaknesses of various forecasting approaches.

3. Real-Time Prediction and Adaptive Models: Another interesting avenue for future research is to evaluate the performance of the models in real-time prediction scenarios. This would involve updating the models with the latest available data and assessing their ability to adapt to changing market conditions. Developing adaptive models that can adjust their predictions dynamically based on new information can be valuable for investors and financial institutions.
4. Integration of Hybrid Models: Hybrid models that combine the strengths of different forecasting techniques can be explored. For example, integrating the strengths of LSTM and ARIMA in a hybrid model may provide improved forecasting accuracy. Investigating the effectiveness of such hybrid models can contribute to the development of more robust and accurate prediction systems.

By addressing these future directions, researchers can further advance the field of stock price prediction and deepen our understanding of the capabilities and limitations of different forecasting models.

References

1. Siami-Namini S, Tavakoli N, Siami Namin A (2018) A comparison of ARIMA and LSTM in forecasting time series. In: 2018 17th IEEE international conference on machine learning and applications (ICMLA), pp 1394–1401
2. Pang X, Zhou Y, Wang P, Lin W, Chang V (2020) An innovative neural network approach for stock market prediction. J Supercomput 76(3):2098–2118. https://doi.org/10.1007/s11227-017-2228-y
3. Hiransha, Gopalakrishnan, Menon VK, Soman (2018) NSE stock market prediction using deep-learning models. Proc Comput Sci 132:1351–1362. https://doi.org/10.1016/j.procs.2018.05.050
4. Saiktishna C, Sumanth NS, Rao MM, Thangakumar J (2022) Historical analysis and time series forecasting of stock market using FB prophet. In: 2022 6th International conference on intelligent computing and control systems (ICICCS), pp 1846–1851
5. He K, Yang Q, Ji L, Pan J, Zou Y (2023) Financial time series forecasting with the deep learning ensemble model. Mathematics 11(4):1054. https://doi.org/10.3390/math11041054
6. Fang Z, Ma X, Pan H, Yang G, Arce GR (2023) Movement forecasting of financial time series based on adaptive LSTM-BN network. Expert Syst Appl 213:119207
7. Gajamannage K, Park Y, Jayathilake DI (2023) Real-time forecasting of time series in financial markets using sequentially trained dual-LSTMs. Expert Syst Appl 223:119879. https://doi.org/10.1016/j.eswa.2023.119879
8. Patil R (2021). Time series analysis and stock price forecasting using machine learning techniques 19. https://doi.org/10.1994/Rajat/AI
9. Jamil H (2022) Inflation forecasting using hybrid ARIMA-LSTM model. Laurentian University of Sudbury
10. Zhang R, Song H, Chen Q, Wang Y, Wang S, Li Y (2022) Comparison of ARIMA and LSTM for prediction of hemorrhagic fever at different time scales in China. PLoS ONE 17(1):e0262009. https://doi.org/10.1371/journal.pone.0262009
11. Lilly SS, Gupta N, Anirudh RRM, Divya D (2021) Time series model for stock market prediction utilising prophet. Turk J Comput Math Educ (TURCOMAT) 12(6):4529–4534. https://turcomat.org/index.php/turkbilmat/article/view/8439
12. Kaninde S, Mahajan M, Janghale A, Joshi B (2022) Stock price prediction using Facebook prophet. ITM Web Conf 44:03060. https://doi.org/10.1051/itmconf/20224403060

13. Staudemeyer RC, Morris ER (2019) Understanding LSTM—a tutorial into long short-term memory recurrent neural networks. arXiv [cs.NE]. http://arxiv.org/abs/1909.09586
14. Zhang J, Ye L, Lai Y (2023) Stock price prediction using CNN-BiLSTM-attention model. Mathematics 11(9):1985. https://doi.org/10.3390/math11091985