Comparative Analysis of Stock Price Prediction Based on ARIMA and LSTM Models

Qi Wu*

International Business College, South China Normal University, Foshan, Guangdong, 528225, China

*20223638040@m.scnu.edu.cn

Abstract. This study delves into the application of the Autoregressive Integrated Moving Average Model (ARIMA), Long Short-Term Memory Network (LSTM), and their hybrid models of stock price prediction. It offers a comparative analysis of these models' performances, highlighting the significance of accurate forecasting in the volatile world of stock markets. The paper begins by underscoring the importance of stock market forecasting and the limitations of traditional time series analysis techniques, particularly the ARIMA model's struggle with long-term dependencies and nonlinear features. It then explores the LSTM model's prowess in capturing long-term patterns within time series data and the potential of hybrid models that marry the strengths of both LSTM and ARIMA to enhance forecasting precision. Through case analysis, the research evaluates these models' performance across different market environments, culminating in comprehensive recommendations for model selection. The study concludes that hybrid models effectively integrate the strengths of ARIMA and LSTM, significantly boosting the accuracy and robustness of time series forecasting. It advocates for future research to explore a range of datasets to affirm the model's universality and robustness and to probe the impact of varying data characteristics on predictive power, thereby offering more targeted guidance for practical applications.

Keywords: Time series analysis; stock price prediction; ARIMA model; LSTM model; hybrid model

1 Introduction

The stock market, a complex nonlinear dynamic system, is a tapestry of price fluctuations influenced by a myriad of factors such as macroeconomic indicators, market sentiment, and policy shifts. For investors navigating this financial landscape, the ability to predict stock prices with precision is not just beneficial—it's essential. It's the compass that guides investment strategies, risk management, and asset allocation. Traditional time series analysis methods, like the ARIMA model, have long been the stalwarts of linear time series data processing. However, as deep learning technologies have unfurled, the LSTM model has stepped into the limelight for its unique ability to

© The Author(s) 2025
B. Siuta-Tokarska et al. (eds.), *Proceedings of the 2025 3rd International Academic Conference on Management Innovation and Economic Development (MIED 2025)*, Advances in Economics, Business and Management Research 348,
https://doi.org/10.2991/978-94-6463-835-6_3

unravel long-term dependencies in time series data. Moreover, the hybrid model, a fusion of ARIMA's linear analysis and LSTM's nonlinear processing capabilities, has shown promise in the arena of stock price prediction.

The ARIMA model, a titan in time series analysis, is lauded for its adept handling of linear time series data [1]. Yet, the nonlinear nature of stock market data often leaves the ARIMA model grappling to capture the market's intricate dynamics, thereby diluting its accuracy and utility in stock price forecasting. As the financial world has grown more complex, so too has the demand for sophisticated predictive tools. Machine learning models have been enlisted for their robust nonlinear fitting capabilities [2]. Among these, LSTM, a deep learning model, has garnered attention for its proficiency in managing time series data [3]. It's the detective that uncovers long-term dependencies, crucial for detecting market cycles and movements. For instance, Mehtab et al. proposed an LSTM-based regression model that, through grid search optimization of hyperparameters, enhanced prediction accuracy for the NIFTY 50 index [4]. Similarly, Sunny et al. dissected the performance of LSTM and Bidirectional LSTM models in stock price forecasting, revealing that fine-tuning hidden layers and optimizing hyperparameters can yield highly accurate predictions [5]. However, LSTM models don't always outperform traditional statistical models, especially in short-term forecasting, where data volume and feature selection can be limiting factors.

Recognizing the limitations of standalone models, researchers have begun to explore hybrid approaches that intertwine traditional statistical models with machine learning models. These hybrids aim to harness ARIMA for linear trends and seasonal components while leveraging LSTM to grasp nonlinear patterns and long-term dependencies [6]. For example, Zhang Yi introduced an ARIMA-LSTM hybrid model that merges ARIMA's long-term trend forecasting with LSTM's nonlinear fitting ability, further enhanced by attention mechanisms and optimization algorithms to significantly uplift prediction accuracy [7]. Likewise, Guan Xueying's study found the ARIMA-RNN hybrid model superior to standalone recurrent neural network models in predicting the CSI 300 index [8]. A comparative study by Su Siqi also spotlighted the ARIMA-LSTM model's strong performance in forecasting the CSI 300 index [9]. Although research on ARIMA-LSTM hybrid models is nascent, their potential to amplify predictive accuracy has been established, beckoning further exploration and practical application.

Through case research, this study seeks to scrutinize the predictive capabilities of the LSTM model, the ARIMA time series model, and the ARIMA-LSTM hybrid model in stock price forecasting. The insights gleaned aim to offer fresh perspectives on stock market forecasting, serving as a beacon for scholars and professionals treading similar paths.

2 Methodology

2.1 ARIMA Model

The Theoretical Foundation of the ARIMA Model. The ARIMA model, also the Box-Jenkins model, is a revered time series forecasting method that Box and Jenkins

introduced in the early 1970s. It's a harmonious blend of three fundamental components: Autoregression (AR), which leans on historical data points to forecast the present value; Differencing (I), which tackles non-stationarity through differencing techniques; and Moving Average (MA), which refines predictions by weaving in previous forecasting errors. Characterized by three parameters—p (the order of autoregressive terms), d (the degree of differencing), and q (the order of moving average terms)—the ARIMA(p, d, q) model can be mathematically expressed as:

$$\left(1 - \sum_{i=1}^{p} \phi_i L^i\right)(1 - L)^d y_t = \mu + (1 + \sum_{i=1}^{q} \theta_i L^i)\epsilon_t \quad (1)$$

Where y_t represents the time series, ϕ_i denotes the autoregressive coefficients, θ_i represents the moving average coefficients, L is the lag operator, d is the degree of differencing, μ is the constant term, and ϵ_t refers to the white noise error term.

2.1.1 Modeling Steps of the ARIMA Model.

Stationary Test. The ARIMA model demands that the input time series be stationary, necessitating stationarity testing before forecasting and model construction. Common methods, such as the Augmented Dickey-Fuller (ADF) test, are employed to test stationarity. The presence of a unit root signals non-stationarity, serving as the null hypothesis in the ADF test. The alternative hypothesis posits that the series is stationary. If the p-value from the ADF test dips below the chosen significance level, typically 0.05, the null hypothesis is rejected, affirming the series as stationary.

ARIMA Model Selection and Order Identification. Once the time series' stationarity is confirmed, the next quest is to pinpoint the ARIMA model's order—specifically, the values of p, d, and q. This journey can be aided by the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots. The ACF plot assists in discerning the order q for the MA component, while the PACF plot is instrumental in identifying the order p of the AR component. Moreover, model selection can be steered by information criteria like the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC), which aid in determining the most suitable order by balancing the goodness of fit with model complexity and model complexity.

Residual Diagnostic Test. To affirm the model's adequacy, a residual diagnostic test is imperative once the ARIMA model is constructed. Residuals, the discrepancies between the model's predicted values and the actual observed outcomes, should ideally resemble white noise for a well-fitted model, indicating the absence of autocorrelation among them. The Ljung-Box test can be deployed to scrutinize the residuals for autocorrelation. Should the residuals fail to mimic white noise, it may be necessary to reevaluate the model's order or parameters and make the requisite adjustments.

2.2 LSTM Model

The Theoretical Foundation of the LSTM Model. LSTM, a breed of Recurrent Neural Network (RNN) conceived by Sepp Hochreiter and Jürgen Schmidhuber in 1997, is designed to navigate the vanishing or exploding gradient issues that plague typical

RNNs when processing lengthy sequences. Its unique ability to learn long-term dependencies is achieved through the ingenious introduction of three gating mechanisms—the input gate, forget gate, and output gate. These gates empower the LSTM to dynamically decide, at each step of the sequence, which information to retain, discard, or update, making it a potent tool for time series analysis.

Architecture and Parameter Settings of the LSTM Model. The architecture of an LSTM model typically consists of an input layer, one or more LSTM layers, and an output layer. Key parameters include: the size of the hidden layer, which refers to the number of neurons in the LSTM layer and affects the model's learning capacity; the number of layers, where adding more layers can enhance the model's learning ability but may also introduce the risk of overfitting; the learning rate, which influences the model's convergence speed; batch size, which affects the stability of model training and memory consumption; and the number of epochs, which refers to the number of iterations during the training process.

A pivotal aspect of model training is the loss function, which quantifies the divergence between predicted and observed values, steering the optimization process. By minimizing the loss function, the model fine-tunes its parameters to pare down prediction errors and enhance accuracy. The Mean Squared Error (MSE) is a popular choice for the loss function. When it comes to optimizers, the Adam optimizer is a favorite due to its adaptive learning rate and momentum attributes. It adeptly adjusts the learning rate, accelerates the model's convergence, and bolsters the model's generalization prowess. The update rule for the Adam optimizer is as follows:

$$\theta_{t+1} = \theta_t - \eta \nabla_\theta \mathscr{L}(\theta_t) \tag{2}$$

Where θ represents the model parameters, η is the learning rate, $\mathscr{L}$ is the loss function, and ∇_θ represents the gradient of the parameters.

2.2.1 Modeling Steps of the LSTM Model.

Data Preprocessing. Before the LSTM model is erected, the time series data must undergo preprocessing. This entails data cleansing (banishing missing values or outliers), standardization or normalization (to bestow upon the data zero mean and unit variance), and crafting lagged features for the time series. These steps are pivotal because the LSTM model is sensitive to the scale of input data and necessitates the time series data to be presented in a specific window format.

Constructing the Sequential Dataset. The time series data is metamorphosed into a format amenable to input into the LSTM model. This typically involves forging a dataset where each input sequence corresponds to an output value.

Defining the LSTM Model Architecture. The LSTM model's structure is delineated by determining the number of LSTM layers, the number of hidden units within each layer, and configuring the output layer. This often requires leveraging deep learning frameworks like TensorFlow or PyTorch to construct a sequential model comprising LSTM layers and fully connected layers.

Model Training. The LSTM model is trained using the meticulously prepared dataset. This involves selecting an apt loss function (e.g., MSE) and optimizer (e.g.,

Adam), as well as setting the batch size and the number of iterations. During the training odyssey, the model learns the patterns within the time series data and tweaks its parameters to minimize the prediction error.

Residual Diagnostic Test. Post-training, a residual diagnostic test is conducted to ensure the model's predictive prowess. This typically involves verifying if the residual sequence approximates white noise, i.e., there is no autocorrelation between the residuals. If the residuals do not dance to the tune of white noise, the model structure may need to be tweaked, or the model may require retraining.

Model Evaluation and Optimization. Model evaluation is commonly conducted through metrics like MSE and MAE (Mean Absolute Error) to assess its effectiveness. Depending on the evaluation's findings, the model's predicted accuracy could need a boost by modifying its hyperparameters or refining the data preprocessing procedures.

2.3 ARIMA-LSTM Hybrid Model

ARIMA Model Parameter Estimation and Forecasting. In the tapestry of constructing the ARIMA+LSTM hybrid model, the training dataset is first fitted to distill the parameters of the ARIMA model. Once the model order is etched, the ARIMA model is set to forecast the test dataset. Meanwhile, the time series data is preprocessed to construct input sequences that harmonize with the LSTM model. Subsequently, the architecture of the LSTM model is defined, including the determination of the number of LSTM layers, the number of hidden units, the learning rate, and other key parameters. The model is then compiled with an appropriate loss function and optimizer.

LSTM Model Construction. With the LSTM model architecture in place, the model is trained using the training dataset, and the trained model is set to make predictions on the test dataset. At this juncture, both models have cast their forecasts for the test dataset. The ensuing step involves assigning weights to the two models based on their prediction performance, with the better-performing model donning a higher weight. The prediction results from both the ARIMA and LSTM models are blended using their corresponding weights, and the results are integrated via a weighted average approach. If necessary, additional mathematical alchemy is applied to the weighted average results to optimize the final prediction output.

Hybrid Model Integration and Optimization. Subsequently, the test dataset is utilized to gauge the hybrid model's overall performance. To further escalate predicting accuracy, model parameters are tweaked, or weights are redistributed, in accordance with evaluation outcomes. This process may require several iterations until the model performance ascends to a satisfactory plateau. Through this ensemble approach, the ARIMA-LSTM hybrid model effectively weaves together the advantages of both models, significantly enhancing the accuracy and robustness of time series forecasting.

3 Case Study

3.1 Case Description

Kobiela et al. embarked on a comparative voyage of the ARIMA and LSTM models on the NASDAQ stock exchange data, assessing both models using MSE and MAPE (Mean Absolute Percentage Error) metrics [6]. Their findings unfurled that when predicting based solely on historical price values, the ARIMA model outshines the LSTM model across multiple time horizons. Specifically, in predicting the average price over 30, 90, and 270 days, the ARIMA model is 3.4 times, 1.8 times, and 2.1 times more accurate than the LSTM model, respectively [6]. This suggests that the ARIMA model is a stalwart in short- to medium-term predictions.

He Jie et al. applied ARIMA, ARIMA-LSTM, and ARIMA-LSTM hybrid models to predict the opening price of the Shanghai Composite Index for 10-day, 50-day, and 116-day periods, comparing the R^2, MAE, and RMSE (Root Mean Squared Error) for each model [10]. The results revealed that in the 10-day forecast, the ARIMA model performed the best, while in the 50-day and 116-day forecasts, the ARIMA-LSTM model and the hybrid model outperformed the ARIMA model. These findings suggest that hybrid models have the upper hand in capturing nonlinear features and long-term dependencies.

3.2 Predictive Performance

After a meticulous analysis of two studies on the application of ARIMA, LSTM, and their hybrid models in stock price prediction, this study offers a comprehensive evaluation of the performance of these models and discusses their respective strengths and limitations.

The ARIMA model is favored for its simplicity and efficiency in handling linear time series data. It is particularly well-suited for short-term stock price forecasting, especially when the data exhibit linear characteristics. However, a key limitation of the ARIMA model is its inability to capture nonlinear patterns and long-term dependencies, which limits its effectiveness in modeling complex market dynamics and long-term forecasting needs.

The LSTM model is distinguished by its capacity to capture temporal dependencies over extended periods, especially when dealing with time series data exhibiting intricate dynamic patterns. The network architecture of LSTM enables it to effectively learn long-term dependencies, which is crucial for predicting long-term trends in the stock market. Additionally, LSTM can integrate multiple variables, enhancing its adaptability in complex financial environments. However, the complexity of the LSTM model also presents challenges, as it requires substantial amounts of data and computational resources for effective training. This can become a limiting factor in situations with scarce data or limited computational capacity.

The hybrid model combines the strengths of both ARIMA and LSTM, aiming to improve predictive accuracy. By integrating the linear analytical capabilities of

ARIMA and the nonlinear processing abilities of LSTM, the hybrid model demonstrates excellent performance in medium- to long-term forecasting. This model is better equipped to adapt to market changes, particularly when forecasting over longer time horizons. However, the use of hybrid models also introduces additional challenges in terms of data preprocessing and model tuning, which can increase the complexity and cost of model deployment. The advantages and limitations of each model and their applicable scenarios are summarized in Table 1.

Table 1. Comparison of Different Models

Model Type	Applicable Scenarios	Advantages	Limitations
ARIMA	Short-term forecasting Simple data characteristics Resource-constrained environment	High computational efficiency Simple structure Suitable for linear time series data	Difficult to capture nonlinear features Insufficient handling of long-term dependencies
LSTM	Long-term forecasting Complex market dynamics Multivariate analysis	Able to capture long-term dependencies Suitable for handling complex dynamic characteristics Capable of processing multivariate inputs	High model complexity Requires large amounts of data and computational resources
Hybrid Model	Mid- to long-term forecasting Requires both linear and nonlinear analysis Strong data preprocessing and model tuning capabilities	Combines linear and nonlinear analysis capabilities Enhances prediction accuracy	Requires more data preprocessing Model tuning is complex

To be more precise, the model selection should be determined by the particular demands of the prediction task, the properties of the data, the computational resources at hand, and the required degree of prediction accuracy. For short-term forecasting that requires quick responses and the handling of linear data, the ARIMA model may be a more suitable choice. On the other hand, for medium- to long-term forecasting that demands in-depth analysis of long-term trends and complex dynamics, the LSTM model or hybrid models may offer more optimal solutions.

3.3 Discussion

Due to differences in data scale and sample size, a direct horizontal comparison between the two studies is not feasible. This limitation implies that although both case studies selected in this paper provide valuable insights, their results cannot be directly generalized from one context to another, as they are based on different datasets and

market environments. Such differences may influence the evaluation of model performance, as distinct datasets may exhibit varying levels of volatility, trends, and seasonality, all of which can significantly impact the forecasting ability of the models.

In conclusion, when applying these findings to other markets or datasets, careful consideration must be given to the potential impact of data characteristics on model performance. Future research should be conducted on larger datasets to validate the universality and robustness of these models and explore the specific effects of different data characteristics on model prediction capabilities. This approach will enable more accurate assessments and comparisons of model performance, providing more targeted guidance for practical applications.

4 Conclusion

This review thoroughly explores the application of ARIMA, LSTM, and their hybrid models in stock price prediction, providing a comparative analysis of their performance. The study reveals that the ARIMA model performs exceptionally well in short-term forecasting and linear time series analysis due to its computational efficiency and simplicity. However, it is constrained in handling extended dependencies and complex nonlinear dynamics. The LSTM model, on the other hand, demonstrates strong potential in capturing long-term dependencies and addressing complex market dynamics and multivariate analysis, although it comes with higher model complexity and requires substantial data and computational resources. The hybrid model, which combines ARIMA's linear analytical capability and LSTM's nonlinear processing ability, shows superior performance in medium- and long-term forecasting, albeit at the cost of increased data preprocessing and model tuning efforts.

Based on these findings, the following recommendations are made: When choosing a model, the intrinsic properties of the data set, available computational resources, and the desired level of prediction accuracy should all be taken into consideration. Data preprocessing is a critical step to enhance model performance, including data cleaning, standardization, normalization, and feature engineering. For LSTM and hybrid models, meticulous hyperparameter tuning is essential to optimize model performance. Future research should be conducted on more extensive datasets to validate the generalizability and robustness of these models, as well as to explore the specific impact of different data characteristics on predictive capabilities. Additionally, integrating ensemble learning methods, combining the predictive results of ARIMA and LSTM models, should be considered to improve prediction accuracy and robustness. These recommendations aim to enable researchers and practitioners to more effectively apply these models in stock price prediction and provide a scientific basis for decision-making in financial markets.

References

1. Mondal, P., Shit, L., Goswami, S.: Study of effectiveness of time series modeling (ARIMA) in forecasting stock prices. International Journal of Computer Science, Engineering and Applications 4(2), 13 (2014).
2. Jiang, W.: Applications of deep learning in stock market prediction: recent progress. Expert Systems with Applications 184, 115537 (2021).
3. Yan, B., Aasma, M.: A novel deep learning framework: Prediction and analysis of financial time series using CEEMD and LSTM. Expert Systems with Applications 159, 113609 (2020).
4. Mehtab, S., Sen, J., Dutta, A.: Stock price prediction using machine learning and LSTM-based deep learning models. In: Machine Learning and Metaheuristics Algorithms, and Applications: Second Symposium, SoMMA 2020, pp. 88–106. Springer Singapore (2021).
5. Sunny, M. A. I., Maswood, M. M. S., Alharbi, A. G.: Deep learning-based stock price prediction using LSTM and bi-directional LSTM model. In: 2020 2nd Novel Intelligent and Leading Emerging Sciences Conference (NILES), pp. 87–92. IEEE (2020).
6. Kobiela, D., Krefta, D., Król, W., et al.: ARIMA vs LSTM on NASDAQ stock exchange data. Procedia Computer Science 207, 3836–3845 (2022).
7. Zhang, Y.: Research and application of time series analysis based on ARIMA-LSTM hybrid model. Master's thesis, Yangtze University (2023).
8. Guan, X. Y.: Stock price prediction based on ARIMA-RNN hybrid model. Journal of Harbin University of Commerce (Natural Sciences Edition) 40(02), 250–256 (2024).
9. Su, S. Q.: A comparative study of stock price prediction models based on LSTM and ARIMA. Master's thesis, Chongqing University (2023).
10. He, J., Li, S. P., He, Y. Y., et al.: Stock analysis based on ARIMA and LSTM models. Modern Information Technology 8(21), 41–45 (2024).

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.