



Exploring Deep Learning on Apple Stock Volatility: Long Short-Term Memory and Gated Recurrent Unit Performance

**Kia Kurniawan^{1*}, Ikhwan Ash-Shiddiqi²,
Rutul Trivedi³, Reylita Putri Ramadhan⁴**

^{1,2}Department of Information System, Faculty of Science and Technology,
Universitas Islam Negeri Sultan Syarif Kasim Riau, Indonesia

³Department of Computer Science and Engineering, Faculty of Computer Science,
Darshan University, India

⁴Department of E-commerce, Faculty of Economics and Management,
Nantong College of Science and Technology, China

E-Mail: ¹kiakurniawan10@gmail.com, ²ikhwanashshiddiqi23@gmail.com,
³rutultrivedi07@gmail.com, ⁴reylitap@gmail.com

Received Jun 12th 2025; Revised Jun 20st 2026; Accepted Jul 26th 2026; Available Online Aug 01st 2026

Corresponding Author: Kia Kurniawan

Copyright © 2026 by Authors, Published by Institute of Research and Publication Indonesia (IRPI)

Abstract

Forecasting Apple Inc. (AAPL) stock prices is challenging due to the nonlinear and dynamic nature of financial data. This study compares the performance of Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) architectures in predicting AAPL stock volatility. Daily stock data from January 1, 2020, to May 5, 2025, were used to evaluate both models under various hyperparameters. Performance was assessed using Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE). The results show that GRU with the AdamW optimizer (learning rate 0.01, batch size 8) achieved the best performance (MAE: 3.4188, RMSE: 4.9261, MAPE: 1.5648). Beyond identifying optimal hyperparameters, this study highlights critical architectural differences. The findings reveal that GRU's streamlined gating mechanism adapts more effectively and more rapidly to sudden price fluctuations than LSTM's complex memory-retention structure. These results demonstrate GRU's distinct advantage for computationally agile and accurate stock forecasting under highly volatile market conditions.

Keywords: Apple Inc. (AAPL), Deep Learning, Gated Recurrent Unit, Long Short-Term Memory, Stock Volatility

1. INTRODUCTION

The increasing complexity, volatility, and nonlinearity of financial markets have intensified the need for accurate forecasting models. Stock prices are affected by historical price movements, firm-specific events, macroeconomic conditions, investor sentiment, and sectoral dynamics. These interconnected factors make stock prediction difficult and highly uncertain [1], [2], [3]. In the technology sector, Apple Inc. (AAPL) is a highly relevant forecasting target because its stock price is strongly influenced by innovation cycles, product announcements, investor expectations, and broader technology market trends [4]. Therefore, forecasting AAPL closing prices is important for both academic research and practical investment decision-making.

Numerous recent studies have comprehensively demonstrated that deep learning architectures significantly outperform traditional statistical methods in capturing the noisy and non-stationary behavior of modern financial markets [5], [6], [7]. Specifically, for forecasting highly volatile technology stocks and financial assets, data-driven historical approaches are continually optimized to minimize investment risk and predict sudden market shifts [8], [9].

Traditional forecasting models, including ARIMA, exponential smoothing, and GARCH, have been widely used in financial time-series analysis due to their interpretability and statistical foundations [10], [11]. However, these methods often depend on assumptions of linearity, stationarity, or predefined volatility structures. Such assumptions are difficult to satisfy in stock-market data, which are noisy, nonlinear, and highly dynamic [12]. As a result, conventional models may fail to capture complex temporal patterns in financial markets. Deep learning has therefore become a promising alternative because it can learn nonlinear relationships and sequential dependencies directly from historical data [13], [14].

Among deep learning models, Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) are widely adopted for stock-price forecasting. LSTM uses memory cells and gating mechanisms to capture long-term dependencies in sequential data [1], [15], [16]. GRU simplifies the recurrent structure through update and reset gates, reducing computational complexity while maintaining strong predictive capability [3], [16], [17]. Previous studies show that both models can outperform traditional forecasting approaches. However, their relative performance varies across datasets, markets, horizons, and feature representations [18]. This indicates that LSTM and GRU may provide complementary predictive strengths, rather than one model consistently being superior.

Recent research has improved stock forecasting through hybrid models, ensemble learning, feature engineering, and hyperparameter optimization. Ensemble deep learning has been shown to improve prediction accuracy by combining multiple recurrent models and information sources [19]. Multi-source information fusion can also enhance prediction performance, although it often increases data complexity [14]. Other studies have combined decomposition methods such as EMD and VMD with LSTM to reduce noise and improve the modeling of nonlinear financial series [11], [12]. Hyperparameter optimization has further improved LSTM forecasting performance [20]. In addition, attention-based, bidirectional, and adaptive learning frameworks have strengthened model stability, generalization, and responsiveness to changing market conditions [15], [21], [22].

Despite these advances, important research gaps remain. Many studies evaluate LSTM and GRU independently or compare them as competing models, rather than integrating them into a unified ensemble framework [1], [4], [10]. Several high-performing models also rely on heterogeneous data, such as financial news, sentiment, Google Trends, or multi-source indicators, which may reduce practical reproducibility [2], [14]. Moreover, studies specifically focused on AAPL closing price prediction using an LSTM-GRU ensemble with historical prices, technical indicators, and systematic tuning remain limited. This gap highlights the need for a practical, efficient, and robust forecasting framework [12], [19].

Therefore, this study proposes an ensemble deep learning framework that integrates LSTM and GRU to predict Apple Inc.'s (AAPL) daily closing price. The framework uses historical stock price data, enriched with adaptive technical indicators and systematic hyperparameter tuning. By combining LSTM and GRU, the proposed model aims to capture both long-term temporal dependencies and efficient recurrent representations. Unlike prior studies that focus mainly on standalone models, statistical-deep learning hybrids, or complex multi-source systems, this study examines the complementary predictive capability of LSTM and GRU within a unified ensemble architecture. Model performance is evaluated using Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE), which are commonly used in stock-price forecasting research [10], [12], [20]. The findings are expected to contribute to practical deep learning-based financial forecasting and provide useful insights for technology-stock prediction.

2. MATERIAL AND METHOD

This study adopts a quantitative approach using deep learning methods to forecast Apple Inc. (AAPL) stock prices through LSTM and GRU models. Processing financial time-series data requires robust memory-architectural mechanisms that mitigate the vanishing-gradient problem inherent in standard recurrent networks [23], [24], [25]. The research procedure begins with the collection of historical stock data, followed by data preparation stages that include preprocessing, normalization, time-step windowing, and dataset division using hold-out validation. After the dataset is prepared, the LSTM and GRU models are constructed and trained separately to learn sequential patterns and temporal dependencies in stock price movements. This workflow is consistent with previous stock forecasting studies that apply LSTM and GRU to financial time series data and compare their performance in terms of prediction accuracy and model efficiency [1], [3], [17]. The Research Methodology is shown in Figure 1.

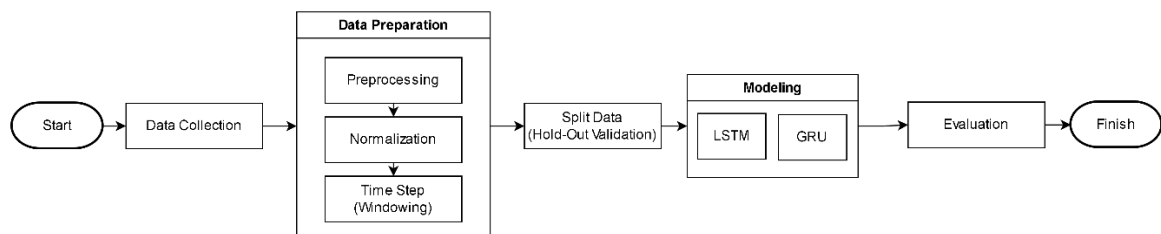


Figure 1. Research Methodology

2.1. Stocks

Stocks are capital market instruments that represent ownership in a company and function as investment assets whose values are strongly affected by market dynamics. In this study, Apple Inc. (AAPL) stock was

selected because its price movement is highly sensitive to company-specific information, technical market conditions, investor sentiment, and broader economic factors. Stock price fluctuations reflect market responses to new information and investor expectations, making stock movement prediction a complex task in financial forecasting [1], [2]. This complexity is also supported by previous studies showing that financial time series are generally nonlinear, volatile, and influenced by multiple economic and non-economic factors, which makes accurate stock prediction challenging [3], [10].

2.2. Data Collection

Data acquisition was conducted using the Yahoo Finance platform via a Python-based API to obtain historical stock market data. The main parameters used in the retrieval process included the stock ticker, date, and close price. In this study, Apple Inc. stock was identified using the correct ticker symbol AAPL, and the dataset covered the period from January 1, 2020, to May 5, 2025. The use of Yahoo Finance as a source of historical stock data is consistent with previous financial forecasting research that retrieved market data from Yahoo Finance for deep learning-based stock price prediction [10].

2.3. Long Short-Term Memory (LSTM)

LSTM is a recurrent neural network architecture introduced by Hochreiter and Schmidhuber, designed to capture long-term dependencies in sequential time series data [26]. In this study, LSTM is applied to predict AAPL stock volatility by learning temporal patterns from historical stock data. LSTM uses memory cells and three main gates, namely the input gate, forget gate, and output gate, to control how information is stored, updated, and passed through time. This gating mechanism enables LSTM to overcome common recurrent neural network limitations, such as vanishing gradients, and makes it suitable for financial time series forecasting [1], [10]. Although LSTM has a more complex structure than GRU, previous studies show that it remains effective for modeling nonlinear and volatile stock market data [11], [22]. The LSTM algorithm equations are presented in Equations 1-6.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (1)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (3)$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \quad (4)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (5)$$

$$h_t = o_t \cdot \tanh(C_t) \quad (6)$$

2.4. Gated Recurrent Unit (GRU)

GRU is a recurrent neural network architecture proposed by Cho et al., designed to capture sequential patterns and long-term dependencies in time series data [27]. In this study, GRU is applied to predict AAPL stock volatility by learning temporal patterns from historical stock data. GRU uses two main gates, namely the update gate and reset gate, to control how past information is retained, discarded, and combined with current input. Compared with LSTM, GRU has a simpler structure and fewer parameters, thereby reducing training complexity while maintaining strong predictive performance in financial time series forecasting [3], [17]. Previous studies also show that GRU is suitable for stock market prediction because it can model dynamic temporal patterns efficiently and, in some cases, outperform LSTM in terms of accuracy and training efficiency [4], [10]. The GRU algorithm equations are presented in Equations 7-10.

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t]) \quad (7)$$

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t]) \quad (8)$$

$$\tilde{h}_t = \tanh(W \cdot [r_t \cdot h_{t-1}, x_t]) \quad (9)$$

$$h_t = (1 - z_t) \cdot h_{t-1} + z_t \cdot \tilde{h}_t \quad (10)$$

2.5. Time Steps

Time steps are an important component of time series processing because they determine how many previous observations are used as input to predict future values. In this study, the time-step approach is applied to transform Apple Inc. (AAPL) stock price data into a supervised learning structure, allowing the LSTM and GRU models to learn sequential dependencies from historical price movements. The selection of an appropriate historical window is important because an overly short window may fail to capture relevant market patterns, while an overly long window may include obsolete information and reduce predictive effectiveness. Therefore, this study evaluates several time-step configurations, ranging from short intervals to a maximum of 30 time steps, to identify the most effective input sequence length for stock price prediction. This approach is consistent with previous studies that emphasize the importance of the historical window size and sequential input design in recurrent neural network-based stock forecasting [1], [2], [17].

2.6. Optimizer

This study uses the Adam optimizer to improve the training efficiency of the LSTM and GRU models by enabling adaptive parameter updates during training. In addition, AdamW is applied as an alternative optimizer because it incorporates decoupled weight decay regularization, which can help improve model generalization and reduce overfitting when modeling volatile AAPL stock data [28], [29]. The use of adaptive optimization is consistent with previous stock forecasting studies that employed Adam-based training to update recurrent neural network parameters and improve prediction performance in nonlinear financial time series [15], [16], [22].

2.7. Evaluation

The evaluation of these predictions using MAE, RMSE, and MAPE provides a standardized, direct comparison between the model's forecasting capability and the actual market trajectory [30], [31]. MAE measures the average squared difference between the actual and predicted stock prices. A lower MAE indicates that the model produces predictions closer to the observed data, reflecting better forecasting precision. This metric is commonly used in deep learning-based stock price prediction studies to evaluate model accuracy and to compare forecasting performance across models [2], [20]. The MAE is calculated using the following Equation 11.

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (11)$$

RMSE is obtained by taking the square root of the MSE, allowing the prediction error to be interpreted in the same scale as the original stock price data. A lower RMSE indicates better forecasting accuracy and smaller deviations between predicted and actual values. RMSE is widely used in stock price forecasting studies to evaluate and compare the predictive performance of deep learning models such as LSTM and GRU [10], [12], [20]. The RMSE is calculated as Equation 12.

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (12)$$

MAPE measures forecasting error as a percentage by calculating the average relative difference between actual and predicted values. This metric is useful because it expresses model error as a percentage, making prediction performance easier to interpret and compare across models. A lower MAPE value indicates better forecasting accuracy and smaller relative deviation from the actual stock price. MAPE is commonly used in stock price prediction studies to evaluate deep learning models, including LSTM, GRU, and hybrid forecasting approaches [1], [10], [17]. The MAPE is determined using the following Equation 13.

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\% \quad (13)$$

3. RESULTS AND DISCUSSION

An experiment was conducted to evaluate and compare the performance of two neural network architectures, LSTM and GRU, in predicting Apple Inc.'s stock price. Model performance was assessed using three evaluation metrics: MAE, RMSE, and MAPE. The experimental scenarios involved several combinations of hyperparameters, including optimizer type, learning rate, and batch size. Specifically, Adam and AdamW

were used as optimizers, with learning rates of 0.1, 0.01, and 0.001 and batch sizes of 8, 16, and 32. The results show that each combination of hyperparameters produced different levels of forecasting performance for both the LSTM and GRU models. A detailed comparison of these experimental results is presented in Table 1.

Table 1. Comparative Performance of LSTM and GRU Models

Index	Model	Optimizer	Batch Size	Learning Rate	MAE	RMSE	MAPE
0	LSTM	Adam	8	0.1	14.514037	17.676572	6.384236
1	LSTM	Adam	16	0.1	14.469635	15.773862	6.244741
2	LSTM	Adam	32	0.1	21.832522	24.805992	9.385906
3	LSTM	AdamW	8	0.1	15.998586	19.290156	7.000371
4	LSTM	AdamW	16	0.1	39.021302	42.028209	16.754512
5	LSTM	AdamW	32	0.1	15.584288	18.531247	6.842859
6	LSTM	Adam	8	0.01	23.840259	26.795441	10.255925
7	LSTM	Adam	16	0.01	16.841537	17.785335	7.372128
8	LSTM	Adam	32	0.01	35.865176	38.996718	15.376011
9	LSTM	AdamW	8	0.01	9.0099450	11.760921	4.077310
10	LSTM	AdamW	16	0.01	5.4877570	7.9576280	2.530864
11	LSTM	AdamW	32	0.01	52.723917	55.232445	22.775774
12	LSTM	Adam	8	0.001	9.6896920	10.835013	4.192072
13	LSTM	Adam	16	0.001	6.168747	8.7173340	2.833250
14	LSTM	Adam	32	0.001	67.624558	69.631561	29.366852
15	LSTM	AdamW	8	0.001	11.304045	12.514036	4.879520
16	LSTM	AdamW	16	0.001	7.2664130	9.7518260	3.306408
17	LSTM	AdamW	32	0.001	77.684678	79.437917	33.820986
18	GRU	Adam	8	0.1	5.0413710	7.0029070	2.344876
19	GRU	Adam	16	0.1	10.538547	13.843793	4.831420
20	GRU	Adam	32	0.1	7.5028890	9.6326610	3.309659
21	GRU	AdamW	8	0.1	7.0939010	8.6483440	3.105711
22	GRU	AdamW	16	0.1	15.901854	17.691440	6.831434
23	GRU	AdamW	32	0.1	26.708578	29.262188	12.231568
24	GRU	Adam	8	0.01	3.7015930	5.1151930	1.680593
25	GRU	Adam	16	0.01	14.367728	15.811853	6.183859
26	GRU	Adam	32	0.01	4.0125170	5.6735040	1.826600
27	GRU	AdamW	8	0.01	3.4188030	4.926152	1.564814
28	GRU	AdamW	16	0.01	13.415980	14.916845	5.776897
29	GRU	AdamW	32	0.01	4.1660190	5.6178860	1.878472
30	GRU	Adam	8	0.001	4.0295100	5.2530170	1.816802
31	GRU	Adam	16	0.001	19.499816	20.850697	8.404973
32	GRU	Adam	32	0.001	6.0213710	7.0077940	2.651956
33	GRU	AdamW	8	0.001	4.2524640	5.4111580	1.909426
34	GRU	AdamW	16	0.001	19.126225	20.4715210	8.243205

Table 1 presents the results of an experiment comparing the LSTM and GRU models in predicting Apple Inc. (AAPL) stock prices. The experiment was conducted with several hyperparameter configurations: the Adam and AdamW optimizers, batch sizes of 8, 16, and 32, and learning rates of 0.1, 0.01, and 0.001. Model performance was evaluated using three primary metrics: MAE, RMSE, and MAPE, with lower values indicating smaller prediction error. Based on the results in Table 1, the GRU model generally demonstrated better predictive performance than the LSTM model. This is evident from the lower MAE, RMSE, and MAPE values for several GRU configurations compared to LSTM.

Meanwhile, the LSTM model tended to produce higher error values than the GRU model in most configurations. The best configuration for the LSTM was achieved using the AdamW optimizer, a batch size of 16, and a learning rate of 0.01. However, its results still fell short of the GRU's best performance. Thus, the experimental results show that hyperparameter selection significantly affects model performance, and the GRU with AdamW, a batch size of 8, and a learning rate of 0.01 is the most optimal configuration in this study.

Table 2. Top Three Best-Performing Models for AAPL Stock Price Prediction

Index	Model	Optimizer	Batch Size	Learning Rate	MAE	RMSE	MAPE
27	GRU	AdamW	8	0.01	3.418803	4.926152	1.564814
24	GRU	Adam	8	0.01	3.701.593	5.115193	1.680593
30	GRU	Adam	8	0.001	4.029.510	5.253017	1.816802

Table 2 presents the top three best-performing models for AAPL stock price prediction based on the lowest MAE, RMSE, and MAPE values. All three selected models use the GRU architecture with a batch size of 8, indicating that GRU provides the most accurate forecasting performance among the tested configurations. The best overall result is achieved by the GRU model with the AdamW optimizer and a learning rate of 0.01, producing the lowest error values across all evaluation metrics. The second- and third-best models are also GRU-based configurations using the Adam optimizer with learning rates of 0.01 and 0.001, respectively.

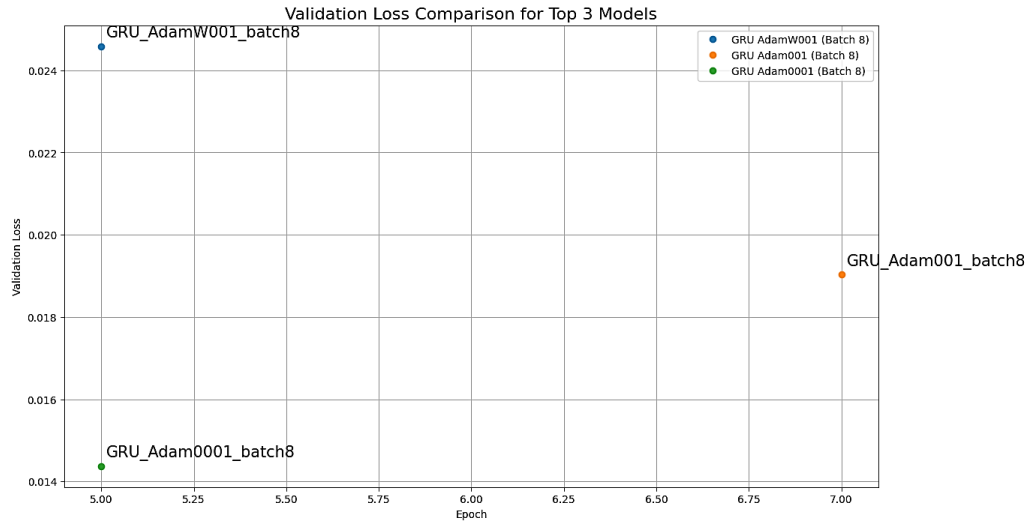


Figure 2. Validation Loss Comparison for Top 3 Models

Figure 2 shows the validation loss comparison of the top three GRU-based models for AAPL stock price prediction. The lowest validation loss is achieved by GRU_Adam0001_batch8, with a value of approximately 0.014, indicating the best validation performance among the three models. The GRU_Adam001_batch8 model records a moderate validation loss of around 0.019, while GRU_AdamW001_batch8 produces the highest validation loss, approximately 0.025. These results indicate that the GRU model using the Adam optimizer with batch size 8 and learning rate 0.001 provides the most stable validation performance compared with the other top model configurations.

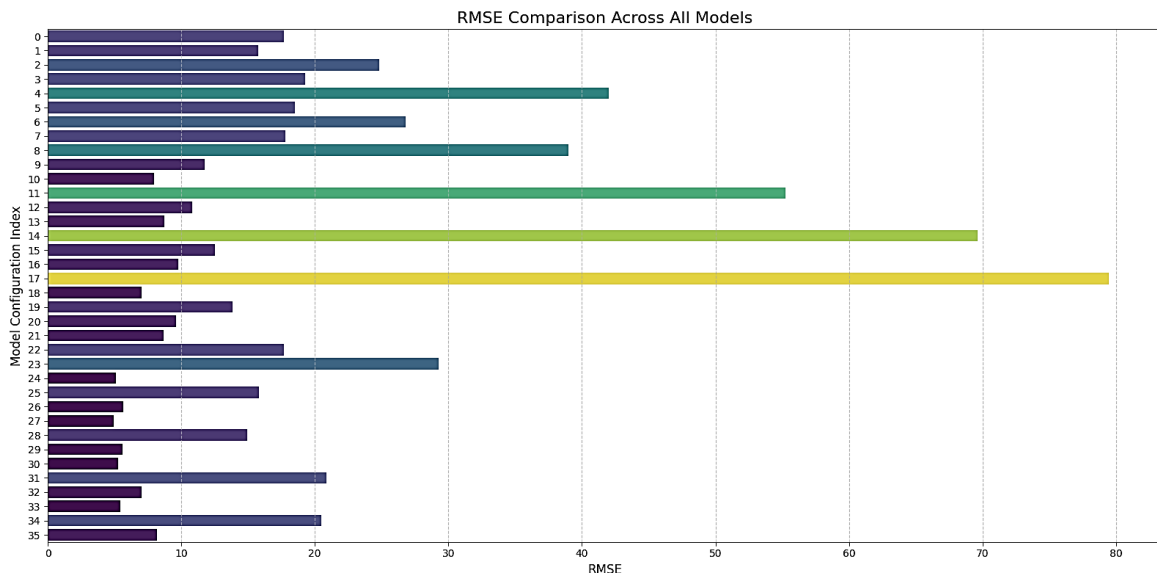


Figure 3. RMSE Comparison Across All Model

Figure 3 presents the RMSE comparison across all LSTM and GRU model configurations, showing the variation in prediction error produced by different optimizer, batch size, and learning rate settings. Lower RMSE values indicate better forecasting accuracy, and the figure shows that several GRU configurations achieve the smallest RMSE values, particularly around model indices 24, 27, and 30. In contrast, some LSTM configurations produce much higher RMSE values, especially at indices 14 and 17, indicating weaker

prediction performance under those hyperparameter settings. Overall, the chart confirms that GRU-based models generally provide more stable and accurate prediction results than LSTM models in this experiment.

Figure 4 compares the actual and predicted AAPL stock prices generated by the three best GRU-based models. Overall, the three models are able to follow the actual price movement reasonably well, especially during upward trends, downward trends, and turning points in the price pattern. The GRU model using AdamW with a batch size of 8 and a learning rate of 0.01 produces the closest predictions to the actual values, as its predicted line closely aligns with the actual price line across most time steps. The GRU models using Adam with learning rates of 0.01 and 0.001 also show similar predictive patterns, although slight deviations appear in several periods, particularly during sharp price fluctuations. Therefore, Figure 4 indicates that all three GRU models demonstrate strong predictive capability, with the AdamW configuration at a learning rate of 0.01 showing the most stable and accurate performance.

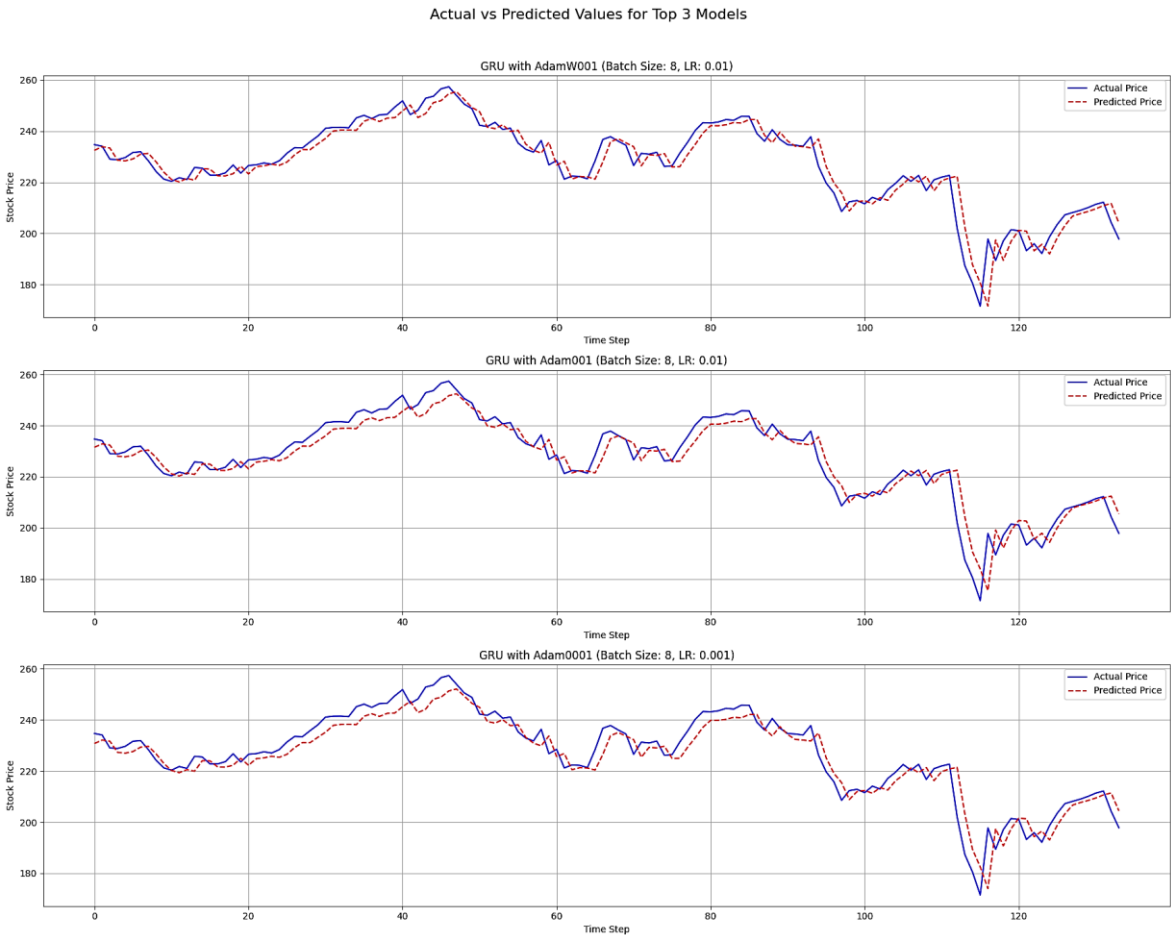


Figure 4. Actual vs Predicted Values for Top 3 Models

3.1. Discussion

The experimental results demonstrate a clear performance advantage of the GRU architecture over LSTM when forecasting Apple Inc. (AAPL) stock volatility. While both models are designed to capture long-term dependencies in sequential data, their architectural differences play a crucial role in how they handle the highly volatile and non-linear nature of technology sector stocks.

LSTM utilizes a complex three-gate mechanism input, forget, and output gates to carefully regulate information flow across time steps [26]. While powerful for sequences requiring deep, long-term memory retention, this structural complexity can lead to slower adaptation when processing sudden, erratic price fluctuations driven by rapid market sentiment shifts or product announcements typical of AAPL. Conversely, GRU employs a more streamlined structure with only two gates: the update and reset gates [27]. The comparative results suggest that GRU's simpler architecture acts as a more agile mechanism in this specific context. It allows the model to quickly discard irrelevant past information and adapt to new, fluctuating price trends without being constrained by an over-reliance on historical states.

The robustness of the GRU architecture in handling sharp data fluctuations aligns with recent findings in predicting other high-risk financial assets, where reduced architectural complexity leads to faster adaptation [32], [33], [34]. The more streamlined gating structure in GRU allows for efficient error propagation and

temporal mapping compared to LSTM, especially when optimized with weight decay strategies like AdamW [35], [36].

Furthermore, the superiority of the AdamW optimizer in the best-performing GRU model highlights the importance of weight decay regularization in preventing the network from overfitting to the noise inherent in daily stock movements. By proving that GRU outperforms LSTM across multiple evaluation metrics (MAE, RMSE, and MAPE), this study establishes that architectural agility and rapid temporal adaptation are more critical than complex memory retention when modeling the volatile dynamics of the Apple stock market.

4. CONCLUSION

This study evaluated the performance of LSTM and GRU models for predicting Apple Inc. (AAPL) stock prices across several hyperparameter configurations, including optimizer types, batch sizes, and learning rates. The best-performing model was a GRU with the AdamW optimizer, a batch size of 8, and a learning rate of 0.01, achieving the lowest MAE of 3.418803, RMSE of 4.926152, and MAPE of 1.564814.

Crucially, the contribution of this research extends beyond empirical hyperparameter tuning, it provides an architectural evaluation of how different recurrent neural networks process volatile financial time series. The findings establish that GRU is fundamentally better suited to capturing the rapid, non-linear fluctuations of technology stocks such as AAPL. GRU's simplified update-and-reset gating structure enables more agile responses to sudden market shifts, offering a distinct predictive advantage over LSTM's more complex memory-retention mechanism. Future research may improve the model by incorporating additional technical indicators, market sentiment data, or ensemble learning approaches to enhance prediction robustness in highly volatile market conditions.

REFERENCES

- [1] Y. Gao, R. Wang, and E. Zhou, "Stock Prediction Based on Optimized LSTM and GRU Models," *Sci. Program.*, vol. 2021, 2021, doi: 10.1155/2021/4055281.
- [2] Y. Li and Y. Pan, "A novel ensemble deep learning model for stock prediction based on stock prices and news," *Int. J. Data Sci. Anal.*, vol. 13, no. 2, pp. 139–149, Mar. 2022, doi: 10.1007/s41060-021-00279-9.
- [3] K. Sako, B. N. Mpinda, and P. C. Rodrigues, "Neural Networks for Financial Time Series Forecasting," *Entropy*, vol. 24, no. 5, May 2022, doi: 10.3390/e24050657.
- [4] V. Chang, Q. A. Xu, A. Chidozie, and H. Wang, "Predicting Economic Trends and Stock Market Prices with Deep Learning and Advanced Machine Learning Techniques," *Electronics (Switzerland)*, vol. 13, no. 17, Sep. 2024, doi: 10.3390/electronics13173396.
- [5] O. B. Sezer, M. U. Gudelek, and A. M. Ozbayoglu, "Financial time series forecasting with deep learning: A systematic literature review: 2005–2019," *Applied Soft Computing Journal*, vol. 90, May 2020, doi: 10.1016/j.asoc.2020.106181.
- [6] W. Jiang, "Applications of deep learning in stock market prediction: Recent progress," *Expert Syst. Appl.*, vol. 184, Dec. 2021, doi: 10.1016/j.eswa.2021.115537.
- [7] B. Lim and S. Zohren, "Time-series forecasting with deep learning: A survey," Apr. 05, 2021, Royal Society Publishing. doi: 10.1098/rsta.2020.0209.
- [8] G. Sonkavde, D. S. Dharrao, A. M. Bongale, S. T. Deokate, D. Doreswamy, and S. K. Bhat, "Forecasting Stock Market Prices Using Machine Learning and Deep Learning Models: A Systematic Review, Performance Analysis and Discussion of Implications," Sep. 01, 2023, Multidisciplinary Digital Publishing Institute (MDPI). doi: 10.3390/ijfs11030094.
- [9] S. Zaheer et al., "A Multi Parameter Forecasting for Stock Time Series Data Using LSTM and Deep Learning Model," *Mathematics*, vol. 11, no. 3, Feb. 2023, doi: 10.3390/math11030590.
- [10] J. K. Mutinda and A. K. Langat, "Stock price prediction using combined GARCH-AI models," *Sci. Afr.*, vol. 26, Dec. 2024, doi: 10.1016/j.sciaf.2024.e02374.
- [11] Z. Zhang, Q. Liu, Y. Hu, and H. Liu, "Multi-feature stock price prediction by LSTM networks based on VMD and TMFG," *J. Big Data*, vol. 12, no. 1, Dec. 2025, doi: 10.1186/s40537-025-01127-4.
- [12] M. Ali, D. M. Khan, H. M. Alshanbari, and A. A. A. H. El-Bagoury, "Prediction of Complex Stock Market Data Using an Improved Hybrid EMD-LSTM Model," *Applied Sciences (Switzerland)*, vol. 13, no. 3, Feb. 2023, doi: 10.3390/app13031429.
- [13] Q. Liu, Z. Tao, Y. Tse, and C. Wang, "Stock market prediction with deep learning: The case of China," *Financ. Res. Lett.*, vol. 46, May 2022, doi: 10.1016/j.frl.2021.102209.
- [14] I. K. Nti, A. F. Adekoya, and B. A. Weyori, "A novel multi-source information-fusion predictive framework based on deep neural networks for accuracy enhancement in stock market prediction," *J. Big Data*, vol. 8, no. 1, Dec. 2021, doi: 10.1186/s40537-020-00400-y.
- [15] S. Sang and L. Li, "A Novel Variant of LSTM Stock Prediction Method Incorporating Attention Mechanism," *Mathematics*, vol. 12, no. 7, Apr. 2024, doi: 10.3390/math12070945.

- [16] Y. Zhang and H. D. Fill, "TS-GRU: A Stock Gated Recurrent Unit Model Driven via Neuro-Inspired Computation," *Electronics (Switzerland)*, vol. 13, no. 23, Dec. 2024, doi: 10.3390/electronics13234659.
- [17] A. Lawi, H. Mesra, and S. Amir, "Implementation of Long Short-Term Memory and Gated Recurrent Units on grouped time-series data to predict stock prices accurately," *J. Big Data*, vol. 9, no. 1, Dec. 2022, doi: 10.1186/s40537-022-00597-0.
- [18] Y. Touzani and K. Douzi, "An LSTM and GRU based trading strategy adapted to the Moroccan market," *J. Big Data*, vol. 8, no. 1, Dec. 2021, doi: 10.1186/s40537-021-00512-z.
- [19] A. D. Kayit and M. T. Ismail, "Advancing stock price prediction through the development of hybrid ensembles: a comprehensive comparative analysis of machine learning approaches," *J. Big Data*, vol. 12, no. 1, Dec. 2025, doi: 10.1186/s40537-025-01185-8.
- [20] B. Gülmez, "Stock price prediction with optimized deep LSTM network with artificial rabbits optimization algorithm," *Expert Syst. Appl.*, vol. 227, Oct. 2023, doi: 10.1016/j.eswa.2023.120346.
- [21] K. Noor and U. Fatima, "Meta Learning Strategies for Comparative and Efficient Adaptation to Financial Datasets," *IEEE Access*, vol. 13, pp. 24158–24170, 2025, doi: 10.1109/ACCESS.2024.3516490.
- [22] S. Sang and L. Li, "A Stock Prediction Method Based on Heterogeneous Bidirectional LSTM," *Applied Sciences (Switzerland)*, vol. 14, no. 20, Oct. 2024, doi: 10.3390/app14209158.
- [23] F. Kamalov, "Forecasting significant stock price changes using neural networks," *Neural Comput. Appl.*, vol. 32, pp. 17655–17667, May 2020, doi: 10.1007/s00521-020-04942-3.
- [24] H. Zouaoui and M. N. Naas, "Option pricing using deep learning approach based on LSTM-GRU neural networks: Case of London stock exchange," *Data Science in Finance and Economics*, vol. 3, no. 3, pp. 267–284, Aug. 2023, doi: 10.3934/DSFE.2023016.
- [25] J. Zhao, D. Zeng, S. Liang, H. Kang, and Q. Liu, "Prediction model for stock price trend based on recurrent neural network," *J. Ambient Intell. Humaniz. Comput.*, vol. 12, no. 1, pp. 745–753, Jan. 2021, doi: 10.1007/s12652-020-02057-0.
- [26] S. Hochreiter and J. "Urgen Schmidhuber, "Long Short-Term Memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.
- [27] K. Cho et al., "Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation," in *Empirical Methods in Natural Language Processing (EMNLP)*, 2014, pp. 1724–1734. doi: 10.3115/v1/D14-1179.
- [28] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," in *International Conference on Learning Representations (ICLR)*, 2015. Accessed: Oct. 18, 2025. [Online]. Available: <https://arxiv.org/abs/1412.6980>
- [29] I. Loshchilov and F. Hutter, "Decoupled Weight Decay Regularization," in *International Conference on Learning Representations (ICLR)*, Jan. 2019. [Online]. Available: <https://dblp.org/rec/conf/iclr/LoshchilovH19.html>
- [30] I. Kasse, J. Kusuma, A. Lawi, and A. Rahim, "Enhancing Stock Price Forecasting Accuracy Through Compositional Learning of Recurrent Architectures: A Multi-Variant RNN Approach," *IEEE Access*, vol. 13, pp. 161073–161102, 2025, doi: 10.1109/ACCESS.2025.3602721.
- [31] J. Sen, S. Mehtab, and A. Dutta, "Stock Price Prediction Using Machine Learning and LSTM-Based Deep Learning Models," Aug. 2021, doi: 10.36227/techrxiv.15103602.v1.
- [32] S. Ji, J. Kim, and H. Im, "A comparative study of bitcoin price prediction using deep learning," *Mathematics*, vol. 7, no. 10, Oct. 2019, doi: 10.3390/math7100898.
- [33] A. Dutta, S. Kumar, and M. Basu, "A Gated Recurrent Unit Approach to Bitcoin Price Prediction," *Journal of Risk and Financial Management*, vol. 13, no. 2, Feb. 2020, doi: 10.3390/jrfm13020023.
- [34] I. Javid, R. Ghazali, I. Syed, M. Zulqarnain, and N. A. Husaini, "Study on the Pakistan stock market using a new stock crisis prediction method," *PLoS One*, vol. 17, no. 10 October, Oct. 2022, doi: 10.1371/journal.pone.0275022.
- [35] H. Song and H. Choi, "Forecasting Stock Market Indices Using the Recurrent Neural Network Based Hybrid Models: CNN-LSTM, GRU-CNN, and Ensemble Models," *Applied Sciences (Switzerland)*, vol. 13, no. 7, Apr. 2023, doi: 10.3390/app13074644.
- [36] J. Xiao, T. Deng, and S. Bi, "Comparative Analysis of LSTM GRU and Transformer Models for Stock Price Prediction," in *Proceedings of International Conference on Digital Economy, Blockchain and Artificial Intelligence, DEBAI 2024*, Association for Computing Machinery, Inc, Dec. 2024, pp. 103–108. doi: 10.1145/3700058.3700075.