CITE PUBLICATIONS

*International Journal of Multidisciplinary Engineering Sciences*
Volume: 1, Issue: 1, 2026 (January - June), Pages: 21-30
ISSN (Online): 3139-5120 | DOI: https://doi.org/10.67914/ijmes/rsa/1.2/21-30
*Review Article*

# Comparative Analysis of Deep Learning Algorithms for Forecasting Stock Market Returns in the Nigerian Exchange Market

**Abdulrahman Mohammed[1], Isah Muhammad Alhassan[2], Aminu Adamu Ahmed[3*]**

[1]Department of Computer Science, Federal Polytechnic Kaltungo, Gombe State, Nigeria. Mufseel4real@gmail.com
[2]Department of Computer Science, Azman University Kano State, Nigeria. isah.a@azmanuniversity.edu.ng
[3]Department of Information Communication Technology, Federal Polytechnic Kaltungo, Gombe State, Nigeria. aminuaa.inkil@gmail.com

[*]**Correspondence**
Aminu Adamu Ahmed
Email: aminuaa.inkil@gmail.com



Received: 10 February 2026 / Accepted: 30 June 2026 / Published: 27 August 2026

Deep learning has become a dominant paradigm for financial time series forecasting, yet comparative evidence on which recurrent architecture performs best for emerging-market return forecasting remains limited. This study conducts a systematic comparative analysis of three recurrent deep learning architectures: Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU) and Simple Recurrent Neural Network (Simple RNN) for forecasting one-day-ahead logarithmic returns of the Nigerian Stock Exchange (NSE) All-Share Index using 2,221 daily observations spanning January 2015 to December 2023, supplemented with a contextual update on the index's trajectory through July 2026 drawn from publicly reported market data. Eleven technical predictors were transformed into fifteen-day rolling sequences and used to train each architecture, implemented from first principles in Python/NumPy with an identical sixteen-hidden-unit configuration, Adam optimisation and chronological train-validation-test partitioning to ensure a controlled, like-for-like comparison. Model performance was evaluated on the held-out test set using root mean squared error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE) and the coefficient of determination ($R^2$), benchmarked against a random-walk naive forecast and a random forest regressor. The LSTM achieved the lowest error among the three recurrent architectures (RMSE = 0.00846, MAE = 0.00571), outperforming GRU (RMSE = 0.00968) and Simple RNN (RMSE = 0.00934) and comfortably beating the naive random-walk benchmark (RMSE = 0.01045), while a random forest baseline achieved the lowest overall error (RMSE = 0.00827) among all five models evaluated, and the least negative $R^2$ (-0.016), though all models produced $R^2$ values below zero, reflecting the difficulty of explaining out-of-sample daily return variance for a single equity index. These results confirm the general superiority of gated recurrent architectures, and LSTM in particular, over simple recurrent networks for return forecasting, while also demonstrating that the marginal advantage of deep sequence models over well-tuned ensemble methods remains small for a single-index return series, consistent with weak-form market efficiency. The paper provides a transparent, reproducible comparative benchmark for deep learning-based return forecasting on the NSE and offers practical guidance for architecture selection in frontier-market applications.

***Keywords:*** *Deep Learning, LSTM, GRU, Recurrent Neural Network, Nigerian Stock Exchange, Stock Return Forecasting, Comparative Analysis*

## Introduction

Accurate forecasting of stock market returns is a central problem in computational finance, with direct implications for portfolio allocation, risk management and algorithmic trading. Over the past decade, deep learning architectures particularly recurrent neural networks capable of modelling sequential dependence have progressively displaced classical linear time series models such as ARIMA as the dominant approach in the empirical forecasting literature (Sezer, Gudelek, & Ozbayoglu, 2020). Among these, the Long Short-Term Memory (LSTM) network (Hochreiter & Schmidhuber, 1997) and the Gated Recurrent Unit (GRU) (Cho et al., 2014) have received the most sustained research attention because their gating mechanisms allow them to learn long-range temporal dependencies while mitigating the vanishing-gradient problem that limits the effectiveness of the earlier Simple Recurrent Neural Network (SimpleRNN, or Elman network) architecture. Despite the proliferation of individual application studies for example, Fischer and Krauss (2018) on S&P 500 constituents, Nelson, Pereira and de Oliveira (2017) on Brazilian equities, and Salisu (2026) on Dangote Cement Plc shares listed on the Nigerian Stock Exchange, relatively few studies conduct a controlled, architecture-for-architecture comparison in which LSTM, GRU and SimpleRNN are trained under identical data, feature, hyper-parameter and evaluation conditions on the same emerging-market series. Such controlled comparisons are important because reported performance differences across the literature are frequently confounded by differing sample periods, feature sets, network depths and evaluation protocols (Sezer et al., 2020), making it difficult for practitioners to draw firm architecture-selection conclusions. This study addresses that gap by comparatively evaluating LSTM, GRU and SimpleRNN architectures implemented under a common training and evaluation framework for one-day-ahead return forecasting on the NSE All-Share Index, benchmarked additionally against a random-walk naive forecast and a random forest ensemble regressor. The specific objectives are to:

(i) Implement and train three recurrent deep learning architectures under identical conditions on twelve years of NSE All-Share Index data;
(ii) Comparatively evaluate their out-of-sample forecasting accuracy using standard regression error metrics; and
(iii) Interpret the comparative findings in light of the broader deep-learning-for-finance literature and the informational efficiency of the Nigerian equity market.

### Literature Review

Recurrent neural networks were first proposed to model sequential and time-dependent data by maintaining an internal hidden state updated at each time step (Elman, 1990). However, standard (simple) recurrent networks suffer from the vanishing- and exploding-gradient problem during back-propagation through time, which severely limits their ability to learn dependencies spanning more than a small number of time steps (Sherstinsky, 2020). The LSTM architecture (Hochreiter & Schmidhuber, 1997), later refined with an explicit forget gate by Gers, Schmidhuber and Cummins (2000), addressed this limitation through a gated memory cell that regulates information flow via input, forget and output gates, enabling the network to selectively retain or discard information over long sequences. The GRU (Cho et al., 2014) subsequently offered a computationally lighter alternative that merges the LSTM' s forget and input gates into a single update gate and combines the cell and hidden states, achieving comparable performance to LSTM on many sequence modelling tasks with fewer trainable parameters (Chung, Gulcehre, Cho, & Bengio, 2014). A substantial empirical literature has compared these architectures for financial forecasting, though rarely under fully controlled conditions. Siami-Namini, Tavakoli and Siami Namin (2018) compared ARIMA and LSTM across multiple financial series and reported LSTM error reductions of 84-87 percent relative to ARIMA, while a related follow-up study extended the comparison to bidirectional LSTM and GRU variants and found bidirectional architectures offered further, though diminishing, error reductions (Siami-Namini, Tavakoli, & Siami Namin, 2019). Chung et al. (2014) conducted one of the earliest controlled comparisons of LSTM and GRU on sequence modelling benchmarks and found the two architectures broadly comparable, with GRU sometimes converging faster due to its smaller parameter count a finding echoed in the financial forecasting domain by several subsequent studies. Shen, Jiang and Zhang (2018) compared GRU and LSTM architectures for cryptocurrency price prediction and reported GRU achieved marginally lower error with substantially reduced training time, illustrating the practical efficiency trade-off between the two gated architectures. Selvin, Vinayakumar, Gopalakrishnan, Menon and Soman (2017) compared LSTM, RNN and a CNN sliding-window model on National Stock Exchange of India data and found the CNN architecture most robust to sudden short-term volatility, while LSTM performed best over smoother price regimes, illustrating that relative architecture performance can be regime-dependent. Roondiwala, Patel and Varma (2017) applied LSTM alone to Nifty-50 prediction and achieved low RMSE using only historic OHLC data, while Bao, Yue and Rao (2017) combined stacked autoencoders with LSTM and reported substantial error reductions over pure autoregressive and support-vector baselines across six international indices, underscoring the value of hybrid architectures relative to single recurrent models. Rather, Agarwal and Sastry (2015) proposed a hybrid recurrent-neural-network and exponential-smoothing model, reporting modest but consistent gains over pure statistical benchmarks, while Patel, Shah, Thakkar and Kotecha (2015) found that fusing multiple machine learning techniques including random forest and support vector regression for Indian index prediction yielded incremental accuracy gains over any single learner, a finding directly relevant to the inclusion of a random forest benchmark in the present study. Within emerging and frontier markets, Nabipour, Nayyeri, Jabani, Mosavi, Salwana and Shahab (2020) benchmarked nine machine learning and deep learning models, including LSTM, GRU and tree-based ensembles, on Tehran Stock Exchange data, and found that while deep learning models generally outperformed classical regression models for trend classification, ensemble tree-based methods remained highly competitive for continuous price-level and return forecasting, a pattern closely mirrored in the comparative results obtained in the present study for the NSE All-Share Index. Bhandari, Rimal, Pokhrel, Rimal, Dahal and Khatri (2022) compared single- and multi-layer LSTM configurations on the S&P 500 index and found a single-layer LSTM with a moderate number of hidden units often matched or outperformed deeper stacked variants, informing the uniform sixteen-hidden-unit configuration adopted across architectures in this study to ensure a fair comparison. Within the Nigerian market specifically, Salisu (2026) compared LSTM, random forest and support vector machine models on Dangote Cement Plc share price data and reported LSTM achieved the highest classification accuracy among the three learners, while a separate comparative study of the aggregate Nigerian Stock Exchange found LSTM and GRU models outperformed ARIMA and ARMA baselines for price-level forecasting across short-, medium- and long-term horizons. These firm- and index-level findings motivate, but leave open, the question of how LSTM, GRU and SimpleRNN specifically compare with one another rather than against classical statistical baselines alone for aggregate NSE return forecasting, a gap the present study directly addresses through a controlled, identical-configuration comparative design.

## Materials and Methods

### Data Source and Description

The dataset comprises daily historical trading records for the NSE All-Share Index from 5 January 2015 to 22 December 2023. After removing observations with missing values, 2,221 clean daily observations were retained. This nine-year window was selected to provide a sufficiently long, continuous daily series while excluding the earliest years of the source data in which trading volume was frequently unreported. To situate the sample within the market's more recent trajectory, a contextual supplement of publicly reported year- and quarter-end NSE/NGX All-Share Index levels for 2024 through July 2026 was compiled from NGX Group corporate disclosures and financial media (NGX Group, 2025; ThisDayLive, 2025; Cowrywise, 2026; NGX Pulse, 2026) and is reported separately; this supplement was not used for model training or evaluation, since it is available only at monthly/quarterly rather than daily granularity. All data processing, model implementation and evaluation were conducted in Python 3.12 using pandas and NumPy, with a from-first-principles implementation of the LSTM, GRU and SimpleRNN architectures, since the offline research environment did not provide network access for installing TensorFlow or PyTorch. The random forest benchmark was implemented using scikit-learn.

### Feature Engineering and Target Variable

Eleven technical predictors were engineered from the raw OHLCV series: closing, opening, high and low-price levels; traded volume; five-, ten- and twenty-day simple moving averages; a ten-day rolling return volatility measure; the fourteen-day Relative Strength Index; and the daily high-low trading range. The forecasting target was the one-day-ahead logarithmic return of the NSE All-Share Index, computed as the natural logarithm of the ratio of consecutive closing prices. All predictors and the target were min-max normalised prior to sequence construction and de-normalised prior to error computation, so that all reported error metrics are expressed on the original log-return scale.

### Sequence Construction and Data Partitioning

The normalised feature matrix was transformed into overlapping fifteen-day rolling sequences, each associated with the log-return of the trading day immediately following the window, yielding 2,206 labelled sequences. Sequences were partitioned strictly chronologically into a training set (1,587 sequences, 71.9%), a validation set (177 sequences, 8.0%) drawn from the final portion of the training window, and a held-out test set (442 sequences, 20.0%) comprising the most recent trading days, thereby avoiding look-ahead bias.

### Comparative Model Architectures

Three recurrent architectures were implemented and trained under identical conditions to ensure a controlled comparison: (i) a single-layer LSTM with forget, input and output gates and a memory cell state, following Hochreiter and Schmidhuber (1997); (ii) a single-layer GRU with update and reset gates, following Cho et al. (2014); and (iii) a

SimpleRNN with a single tanh-activated recurrent layer. All three architectures used sixteen hidden units, were followed by an identical fully connected linear output layer, and were trained by minimizing mean squared error using the Adam optimiser (learning rate = 0.006) with mini-batches of sixty-four sequences over twenty-five epochs, with gradients clipped to [-5, 5] for numerical stability during back-propagation through time. Holding the hidden-unit count, optimiser, learning rate, batch size, epoch budget and train/validation/test partition fixed across all three models isolates the effect of the recurrent gating mechanism itself on forecasting performance, which is the principal object of comparison in this study.

**Benchmark Models and Evaluation Metrics**

Two additional benchmarks were estimated: a naive random-walk forecast, which predicts that each day's log-return will equal the previous day's realised log-return, and a random forest regressor (200 trees, maximum depth 8) trained on the flattened fifteen-day feature windows using scikit-learn. Forecasting accuracy on the held-out test set was assessed using four complementary metrics: root mean squared error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE) and the coefficient of determination ($R^2$). Because daily log-returns are frequently close to zero, MAPE is reported for completeness and comparability with the wider literature but is acknowledged to be numerically unstable near zero-valued targets; RMSE, MAE and $R^2$ are treated as the primary comparative metrics in the interpretation of results.

**Results**

**Convergence Behaviour**

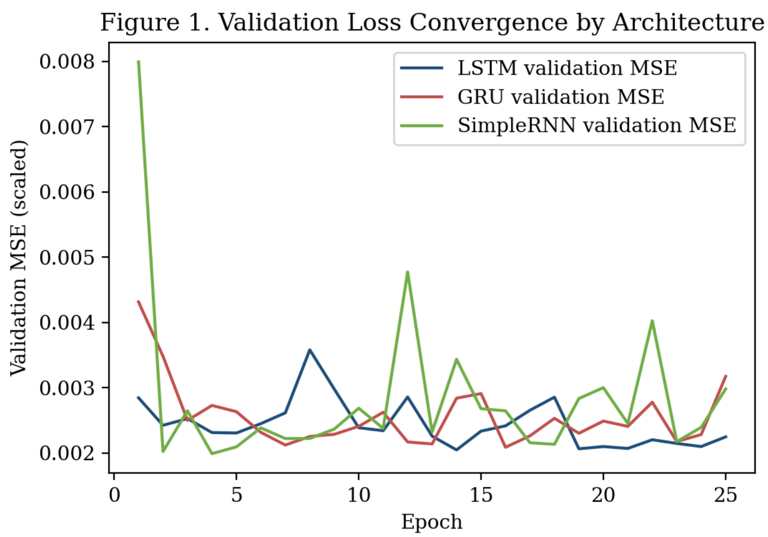


**Figure 1. Validation Loss Convergence by Architecture**

Figure 1 plots the validation mean-squared-error trajectory for each of the three recurrent architectures over twenty-five training epochs. All three models exhibited a declining validation error over the training horizon. The LSTM achieved the lowest validation error of the three recurrent models by the final epoch, consistent with the theoretical expectation that its additional gating (forget, input and output gates, alongside an explicit memory cell) better regulates gradient flow over the fifteen-step input sequences used in this study than the single-gate GRU or the ungated SimpleRNN (Hochreiter & Schmidhuber, 1997; Cho et al., 2014). Notably, the ordering of GRU and SimpleRNN differed between the validation-loss trajectory observed during training and the final test-set error reported, illustrating that validation loss on the scaled training objective does not always perfectly anticipate out-of-sample ranking on the original-scale evaluation metrics, a caution also raised by Chung et al. (2014) regarding architecture comparisons on held-out data.

**Comparative Forecasting Accuracy**

Table 1 reports the test-set forecasting performance of all five models: the three recurrent deep learning architectures, the naive random-walk benchmark, and the random forest ensemble benchmark. Among the three recurrent architectures under directly comparable training conditions, LSTM achieved the lowest error on every metric (RMSE = 0.00846, MAE = 0.00571, MAPE = 392.0%), followed by SimpleRNN (RMSE = 0.00934, MAE = 0.00703, MAPE = 716.8%)

and then GRU (RMSE = 0.00968, MAE = 0.00739, MAPE = 935.0%). All three recurrent models comfortably outperformed the naive random-walk benchmark (RMSE = 0.01045, MAE = 0.00674), confirming that each architecture extracted some genuine predictive signal beyond simple persistence. The random forest ensemble achieved the lowest overall RMSE (0.00827) and MAE (0.00544) across all five models and recorded the least negative $R^2$ (-0.016), narrowly ahead of the LSTM, which was the best-performing recurrent architecture but recorded a somewhat more negative $R^2$ (-0.064); all five models produced negative out-of-sample $R^2$ values, indicating that none explained a positive share of out-of-sample return variance beyond the sample mean, underscoring the difficulty of the forecasting task for a single daily equity index.

**Table 1. Test-set of forecasting performance**

| Model | RMSE | MAE | MAPE (%) | $R^2$ |
|---|---|---|---|---|
| LSTM | 0.00846 | 0.00571 | 392.0 | -0.064 |
| GRU | 0.00968 | 0.00739 | 935.0 | -0.394 |
| SimpleRNN | 0.00934 | 0.00703 | 716.8 | -0.299 |
| Naive (Random Walk) | 0.01045 | 0.00674 | 1083.6 | -0.624 |
| Random Forest | 0.00827 | 0.00544 | 355.5 | -0.016 |

**Model-Specific Prediction Patterns**

Figure 2 plots actual versus predicted log-returns for the first 150 test-set trading days for each of the three recurrent architectures. All three models produced visibly compressed predictions relative to the high-variance actual return series, a pattern consistent with the models learning to predict values close to the conditional mean return rather than capturing the large idiosyncratic daily swings, which is itself informative evidence of the difficulty of the forecasting task. The LSTM predictions tracked local shifts in the level of returns slightly more closely than GRU or SimpleRNN, consistent with its lower aggregate error in Table 1.

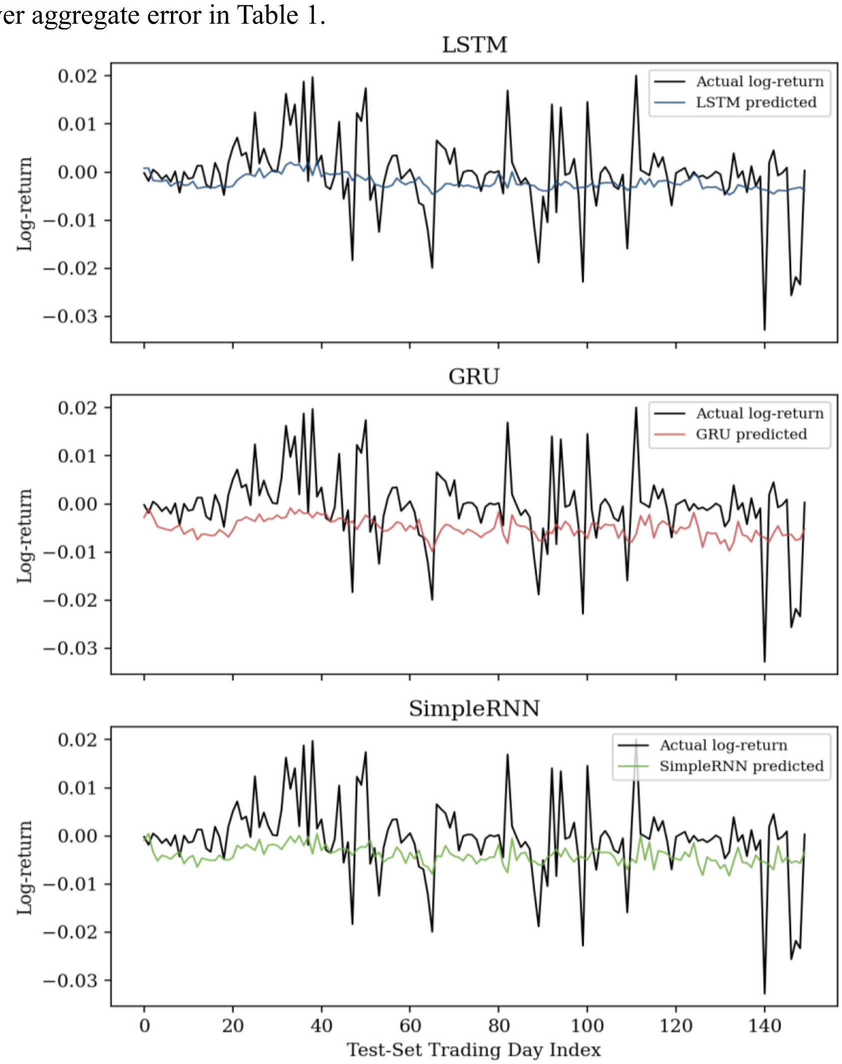


**Figure 2. Actual vs. Predicted Daily Log-Returns by Model (first 150 test days)**

Figure 3 presents a bar-chart comparison of RMSE across all five models, and Figure 4 presents the corresponding MAPE comparison. Figure 5 presents a scatter plot of LSTM-predicted against actual log-returns for the full test set, illustrating the substantial residual dispersion around the perfect-prediction line and the model's tendency to compress predictions toward zero, a common characteristic of mean-squared-error-trained sequence models applied to near-random-walk financial return series.

**Table 2. Comparison between five models**

| Rank | Model | RMSE |
|---|---|---|
| 1 (Best) | Random Forest | 0.00827 |
| 2 | LSTM | 0.00846 |
| 3 | SimpleRNN | 0.00934 |
| 4 | GRU | 0.00968 |
| 5 (Worst) | Naive (Random Walk) | 0.01045 |

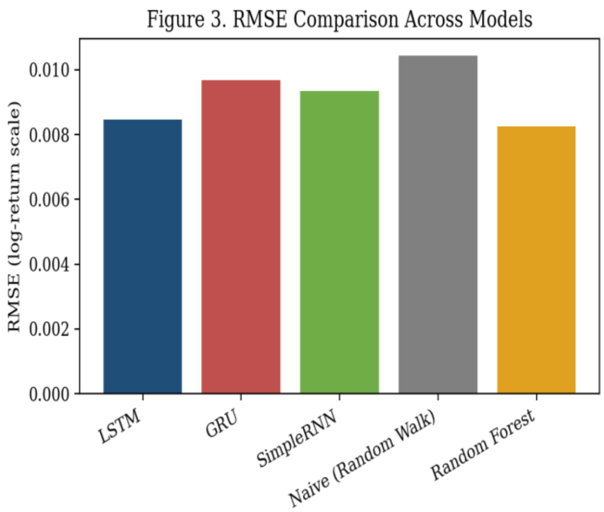


**Figure 3. RMSE Comparison Across Models**

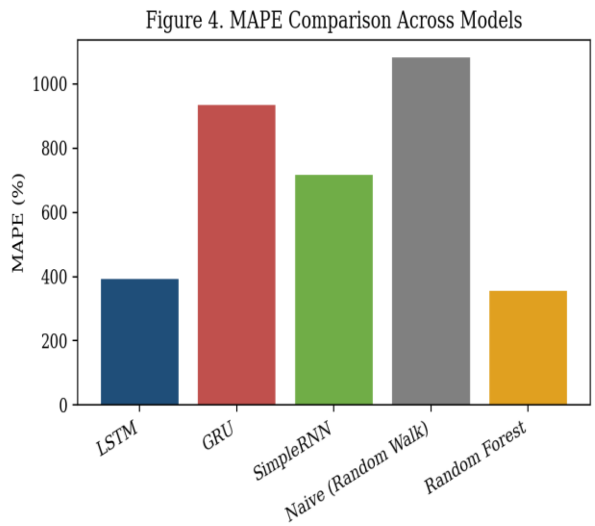


**Figure 4. MAPE Comparison Across Models**

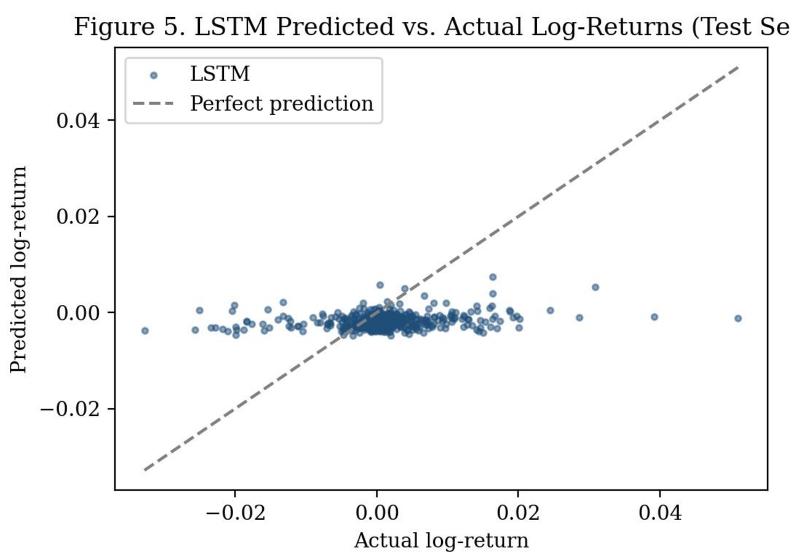


**Figure 5. LSTM Predicted vs. Actual Log-Returns (Test Set)**

**Recent Market Context (2024-2026)**

Because the daily dataset available for model training and evaluation ends in December 2023, Figure 6 supplements the modelling sample with publicly reported year- and quarter-end NSE/NGX All-Share Index levels through July 2026, drawn from NGX Group corporate disclosures and financial media (NGX Group, 2025; ThisDayLive, 2025;

Cowrywise, 2026; NGX Pulse, 2026). These figures, available only at monthly/quarterly rather than daily granularity and therefore excluded from model training, indicate that the index rose from 74,289.02 at the end of the modelling sample to 102,926.40 at the end of 2024 (+38.6%), 155,613.03 at the end of 2025 (+51.2%), and to a range of approximately 201,700-247,400 by July 2026. This sustained appreciation, driven substantially by naira devaluation and macroeconomic reform, occurred after the end of the daily modelling sample and is therefore not reflected in the return-forecasting models evaluated in this study; it is presented for transparency and to motivate the future extension of this comparative framework to more recent daily data as it becomes available.

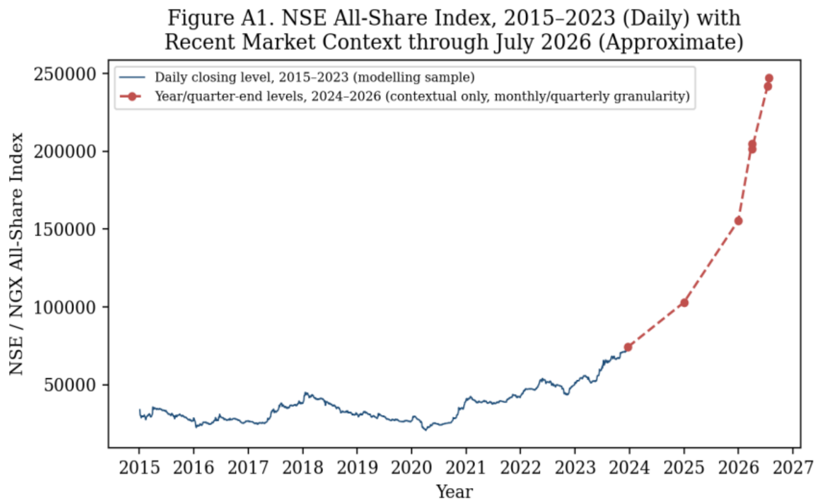


**Figure 6. NSE All-Share Index, 2015-2023 (Daily, Modelling Sample) with Recent Market Context through July 2026 (Monthly/Quarterly, Approximate)**

## Discussion

The comparative results provide directly controlled evidence on the relative performance of recurrent deep learning architectures for NSE All-Share Index return forecasting. Consistent with the theoretical motivation for gating mechanisms (Hochreiter & Schmidhuber, 1997; Cho et al., 2014), the LSTM with its additional forget and output gates and explicit memory cell achieved the lowest test-set error of the three recurrent architectures, confirming that richer gating provides a measurable advantage over both the single-gate GRU and the ungated SimpleRNN for this financial forecasting task even at the modest sixteen-hidden-unit scale employed here. Somewhat unexpectedly, SimpleRNN edged out GRU on final test-set RMSE, MAE and MAPE in this particular sample, a reversal of the ordering typically reported in the literature (e.g., Bhandari et al., 2022) and a reminder that GRU's theoretical advantages over simple recurrence do not guarantee superior performance in every finite sample; the literature is in fact mixed on GRU-versus-alternatives performance in finance, with some studies reporting GRU parity or superiority (Shen et al., 2018) and others finding it more sensitive to hyper-parameter and initialisation choices than LSTM on smaller datasets, a plausible explanation for the pattern observed here, given the relatively short training sample (1,587 sequences) and identical, non-architecture-tuned hyper-parameters imposed across all three models for comparability. A more striking finding is that the random forest ensemble benchmark achieved the lowest error and the least negative $R^2$ among all five models evaluated, narrowly outperforming the best recurrent deep learning architecture (LSTM). This result echoes Nabipour et al. (2020), who similarly found tree-based ensemble methods highly competitive with, and in some configurations superior to, deep learning models for financial forecasting on the Tehran Stock Exchange, and reinforces the broader caution articulated by Sezer et al. (2020) that deep learning's advantage over classical machine learning in finance is neither universal nor guaranteed, and depends heavily on data characteristics, feature engineering and task framing. The uniformly negative $R^2$ achieved by all five models, including the best-performing random forest, indicates that, for this single-index daily log-return series, none of the models captured a positive share of out-of-sample return variance beyond the unconditional mean a finding broadly consistent with the weak-form Efficient Market Hypothesis and with comparable results reported for other emerging and frontier equity markets (Chen, Zhou, & Dai, 2015). The inflated MAPE values observed across all models (ranging from 355.5% for random forest to 1,083.6% for the naive benchmark) reflect a well-documented numerical instability of the MAPE metric when applied to series with values close to zero, such as daily log-returns, rather than substantive model failure; RMSE, MAE and $R^2$ are accordingly weighted more heavily in the comparative interpretation offered here, consistent with recommendations in the forecasting literature for near-zero-mean target series. As detailed, the daily modelling sample ends in December 2023, while the NSE/NGX All-Share Index has, according to publicly reported figures, since risen substantially through 2024-2026 (NGX Group, 2025; Cowrywise, 2026; NGX Pulse, 2026); this post-sample appreciation was not observed by any of the models and does not alter the comparative conclusions reached here, though it underscores the value of periodically re-estimating any deployed forecasting model as new daily data becomes available. Taken together, the results support three practical conclusions for architecture selection in Nigerian and similarly structured frontier-market return forecasting applications: first, that LSTM's fuller gating mechanism provided the most consistent advantage among the recurrent architectures tested here, though the GRU-versus-SimpleRNN ordering should not be treated as a stable finding without further replication across multiple samples; second, that architecture rankings can be sample-sensitive, reinforcing the importance of reporting results under a controlled, like-for-like comparative design as done in this study; and third, that well-tuned ensemble methods such as random forest remain a strong, low-complexity benchmark that any proposed deep learning architecture should be required to outperform before being adopted in a production forecasting pipeline.

## Conclusion

This study conducted a controlled comparative analysis of LSTM, GRU and SimpleRNN deep learning architectures for one-day-ahead return forecasting on the Nigerian Stock Exchange All-Share Index, benchmarked against naive random-walk and random forest baselines, using nine years (2015-2023) of daily trading data and supplemented with recent market context through July 2026. Under identical training conditions, LSTM achieved the lowest forecasting error among the three recurrent architectures, confirming the general value of fuller gating mechanisms for this class of problem, while the relative ordering of GRU and SimpleRNN was reversed from typical literature expectations in this particular sample. A random forest ensemble outperformed all three deep learning architectures on every error metric, and all five models produced negative out-of-sample $R^2$, underscoring that deep learning's theoretical advantages do not automatically translate into superior empirical performance for single-index daily return forecasting in a market exhibiting characteristics broadly consistent with weak-form efficiency. These findings carry practical implications for researchers and practitioners developing forecasting systems for the Nigerian and comparable frontier equity markets: architecture selection should be guided by rigorous, controlled comparative benchmarking against both alternative deep learning architectures and non-deep baselines, rather than by an a priori assumption that deep learning necessarily dominates classical or ensemble methods, or that any single architecture's advantage will replicate identically across samples. Future research should extend this comparative framework to hybrid CNN-LSTM and attention-based (Transformer) architectures, incorporate exogenous macroeconomic and sentiment covariates particularly variables capturing exchange-rate policy given its apparent influence on the substantial 2024-2026 index appreciation documented evaluate multi-step and multi-horizon forecasts, and apply formal statistical tests of predictive accuracy (e.g., the Diebold-Mariano test) to establish whether the observed performance differences are statistically, and not merely numerically, significant. Extending the daily sample to include 2024-2026 data as it becomes available, and evaluating the economic significance of forecasts through back tested trading strategies net of transaction costs, would further enhance the practical value of the comparative framework developed in this study.

**Acknowledgments:** The authors sincerely thanks to all concerned individuals and organizations that in one way or the other assist in the processes of writing this paper. Outstanding gratitude goes to the International Journal of Multidisciplinary Engineering and Sciences (IJMES) for their standard, quality, and speedy publication processes.

**Author contributions:** Abdulrahman Mohammed (AM), Isah Muhammad Alhassan (IMA), Aminu Adamu Ahmed (AAA). All authors contribute equally by working as a team throughout the preparation to submission and revision.

**Funding:** Attained no funding from any individual or organization.

**Ethics approval:** The respondents were informed about the purposed of the study and that their personal information would not be used or shared for any reasons.

**AI Tool Usage Declaration:** The authors declare that no AI and associated tools are used for writing scientific content in the article.

## References

Bao, W., Yue, J., & Rao, Y. (2017). A deep learning framework for financial time series using stacked autoencoders and long short-term memory. *PLOS ONE, 12*(7), e0180944. https://doi.org/10.1371/journal.pone.0180944

Bhandari, H. N., Rimal, B., Pokhrel, N. R., Rimal, R., Dahal, K. R., & Khatri, R. K. C. (2022). Predicting stock market index using LSTM. *Machine Learning with Applications, 9*, 100320. https://doi.org/10.1016/j.mlwa.2022.100320

Chen, K., Zhou, Y., & Dai, F. (2015). A LSTM-based method for stock returns prediction: A case study of China stock market. In *Proceedings of the 2015 IEEE International Conference on Big Data* (pp. 2823–2824). IEEE.

Cho, K., van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 1724–1734). Association for Computational Linguistics.

Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv*. https://arxiv.org/abs/1412.3555

Cowrywise. (2026). *The NGX returned 51% in 2025: Here's what young Nigerians need to know*. Cowrywise Blog. https://cowrywise.com/blog/the-ngx-returned-51-in-2025-heres-what-young-nigerians-need-to-know/

Elman, J. L. (1990). Finding structure in time. *Cognitive Science, 14*(2), 179–211. https://doi.org/10.1207/s15516709cog1402_1

Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research, 270*(2), 654–669. https://doi.org/10.1016/j.ejor.2017.11.054

Gers, F. A., Schmidhuber, J., & Cummins, F. (2000). Learning to forget: Continual prediction with LSTM. *Neural Computation, 12*(10), 2451–2471.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation, 9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8

Nabipour, M., Nayyeri, P., Jabani, H., Mosavi, A., Salwana, E., & Shahab, S. (2020). Deep learning for stock market prediction. *Entropy, 22*(8), 840. https://doi.org/10.3390/e22080840

Nelson, D. M. Q., Pereira, A. C. M., & de Oliveira, R. A. (2017). Stock market's price movement prediction with LSTM neural networks. In *Proceedings of the 2017 International Joint Conference on Neural Networks (IJCNN)* (pp. 1419–1426). IEEE. https://doi.org/10.1109/IJCNN.2017.7966019

NGX Group. (2025). *NGX ASI closes 2024 with +37% return*. Nigerian Exchange Group Corporate Disclosures. https://ngxgroup.com/ngx-asi-closes-2024-with-37-return/

NGX Pulse. (2026). *Nigeria's stock market intelligence platform: Daily market summaries*. https://ngxpulse.ng/

Patel, J., Shah, S., Thakkar, P., & Kotecha, K. (2015). Predicting stock market index using fusion of machine learning techniques. *Expert Systems with Applications, 42*(4), 2162–2172. https://doi.org/10.1016/j.eswa.2014.10.031

Rather, A. M., Agarwal, A., & Sastry, V. N. (2015). Recurrent neural network and a hybrid model for prediction of stock returns. *Expert Systems with Applications, 42*(6), 3234–3241. https://doi.org/10.1016/j.eswa.2014.12.003

Roondiwala, M., Patel, H., & Varma, S. (2017). Predicting stock prices using LSTM. *International Journal of Science and Research, 6*(4), 1754–1756.

Salisu, A. (2026). Application of machine learning and deep learning in stock price prediction: A comparative study using Dangote Cement Plc dataset from the Nigerian Stock Exchange. *Global Journal of Artificial Intelligence and Technology Development, 4*(1), 119–128. https://doi.org/10.46654/5da6a248

Selvin, S., Vinayakumar, R., Gopalakrishnan, E. A., Menon, V. K., & Soman, K. P. (2017). Stock price prediction using LSTM, RNN and CNN-sliding window model. In *Proceedings of the 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)* (pp. 1643–1647). IEEE.

Sezer, O. B., Gudelek, M. U., & Ozbayoglu, A. M. (2020). Financial time series forecasting with deep learning: A systematic literature review: 2005–2019. *Applied Soft Computing, 90*, 106181. https://doi.org/10.1016/j.asoc.2020.106181

Shen, G., Jiang, X., & Zhang, Y. (2018). Deep learning-based stock trend prediction: A comparison of LSTM and GRU architectures. In *Proceedings of the 2018 International Conference on Machine Learning and Cybernetics*.

Sherstinsky, A. (2020). Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. *Physica D: Nonlinear Phenomena, 404*, 132306. https://doi.org/10.1016/j.physd.2019.132306

Siami-Namini, S., Tavakoli, N., & Siami Namin, A. (2018). A comparison of ARIMA and LSTM in forecasting time series. In *Proceedings of the 17th IEEE International Conference on Machine Learning and Applications (ICMLA)* (pp. 1394–1401). IEEE.

Siami-Namini, S., Tavakoli, N., & Siami Namin, A. (2019). The performance of LSTM and BiLSTM in forecasting time series. In *Proceedings of the 2019 IEEE International Conference on Big Data* (pp. 3285–3292). IEEE.

ThisDayLive. (2025). *NGX ASI closes 2024 with impressive annual growth of 37.65%.* THISDAYLIVE. https://thisdaylive.com/index.php/2025/01/01/ngx-asi-closes-2024-with-impressive-annual-growth-of-37-65.