FUDMA Journal of Engineering and Technology
ISSN: 3092-9385
https://fjet.fudutsinma.edu.ng/index.php/fjet/
Volume 2, Issue 1, 2026, 1056-1062

A Hybrid CNN–BiGRU Model with Grey Wolf Optimization and LightGBM for Stock Price Prediction

Fatima A. MUSA[1*], Abdulmajid B. UMAR[2], Abba M. BALA[3]

[1*]School of Science Aminu Kano College of Islamic and Legal Studies, Kano, Kano State, Nigeria
[1,2,3]Faculty of Computing, Northwest University, Kano, Kano State, Nigeria

[1*]fatima4ado@gmail.com, [2]abumar@yumsuk.edu.ng, [3]abbambala01@gmail.com

Abstract
Accurate stock price prediction remains a challenging task due to the non-linearity, volatility, and noisy nature of financial time-series data. Traditional statistical models such as ARIMA are limited in capturing complex market dynamics, while standalone deep learning models often suffer from overfitting and suboptimal hyperparameter selection . This paper proposes a hybrid stock price prediction framework that integrates Convolutional Neural Networks (CNN) for feature extraction and Bidirectional Gated Recurrent Units (BiGRU) for temporal dependency modeling. The Grey Wolf Optimizer (GWO) is employed for automated hyperparameter optimization, while Light Gradient Boosting Machine (LightGBM) is used as the final regression layer to enhance generalization. ARIMA is applied as a preprocessing technique to remove linear components and ensure stationarity. Experimental results show that the proposed model significantly outperforms the baseline ARIMA + Attention-CNN-BiLSTM + XGBoost model, achieving a high explanatory power with an ***R^2 value of 0.93903****, alongside notable reductions in MAE, MSE, and RMSE. These results demonstrate the effectiveness of combining deep learning architectures, metaheuristic optimization, and ensemble learning for accurate and reliable stock price prediction.*

Keywords: *Stock Price Prediction, CNN, BiGRU, Grey Wolf Optimizer, LightGBM, Time-Series Forecasting.*

1 Introduction

Stock markets play a critical role in global economic development by facilitating capital formation and investment decision-making. Accurate stock price prediction is essential for investors and financial institutions to manage risk and maximize returns. However, stock price movements are highly volatile, non-linear, and influenced by numerous internal and external factors, making forecasting extremely challenging [1,2]. Time series forecasting is the process of estimating future values by examining previously recorded data points that are ordered. Despite their widespread use, conventional time-series models like the Autoregressive Integrated Moving Average (ARIMA) frequently fail to reflect the intricate temporal patterns and nonlinear correlations seen in financial time series. Assume linearity and stationarity, which limits their performance in real-world financial markets [3].

Hybrid deep learning frameworks combined with optimization and ensemble learning enhance generalization and robustness in challenging prediction challenges [4]. Advances in machine learning and deep learning have demonstrated superior capabilities in capturing complex temporal dependencies. Models such as LSTM, BiLSTM, and GRU have shown promising results in financial forecasting [5,6]. Nevertheless, these models are sensitive to hyperparameter settings and are prone to overfitting.

To address these challenges, this study proposes a hybrid CNN–BiGRU model optimized using Grey Wolf Optimizer (GWO) and enhanced with LightGBM for improved prediction accuracy and robustness.

2 Related Work

Numerous studies have explored stock price prediction using statistical and machine learning techniques. ARIMA-based models have been widely applied due to their simplicity and interpretability but perform poorly on non-linear financial data [7]. Convolutional Neural Networks (CNN) have been adopted to extract local temporal features, while recurrent architectures such as LSTM and GRU capture long-term dependencies [8]. The time series of stock prices in financial markets is noisy, non-linear, and extremely volatile. Because time series frequently exhibit characteristics like non-stationarity, seasonality, trend, and evident noise, forecasting in financial markets can be challenging. For instance, because stock values are affected by both internal market dynamics and external macroeconomic factors, they are quite volatile and difficult to model [9].

Historically, stock values have been predicted using statistical models, but because financial data is complicated and dynamic, these models frequently fall short. The availability of large financial datasets is increasing, and advancements in machine learning and deep learning approaches have created new opportunities to increase forecasting accuracy [10]. [11] Mean Absolute Error (MAE) was used as the main metric to assess the stock market forecast accuracy of CNN, LSTM, and CNN–LSTM hybrid models.

Received: 29/12/2025; Revised: 09/01/2026; Accepted: 19/01/2026; Published: 27/06/2026

Although the study emphasizes the potential of these models, it is devoid of information on the dataset and might benefit from additional measures such as RMSE, insights into model training, and evaluation under different market conditions. Its relevance could be further increased by incorporating useful applications in actual trading systems. Ensemble techniques being used, such as LightGBM, that help in improving the final prediction has also been investigated in an increasing number of research. For stock price prediction, [12] suggested a hybrid ARIMA–CNN–LSTM model in conjunction with LightGBM. In contrast to solo CNN–LSTM models (0.902), their results demonstrated that LightGBM could learn from the residuals of the base learners and obtained a higher R2 value.

Combinations of different deep learning that is Hybrid models with optimization algorithms have gained increasing attention. Grey Wolf Optimizer (GWO) has been used for hyperparameter tuning in neural networks, demonstrating improved convergence and prediction accuracy compared to GA and PSO [13]. Despite the fact that ensemble learning methods like LightGBM have been demonstrated to enhance generalization and lower variance in financial forecasting tasks, there is still a lack of research on how these methods can be integrated with CNN–BiGRU architectures that are tuned using metaheuristic algorithms. This research addresses this gap by incorporating LightGBM as the final prediction layer within a GWO-optimized hybrid deep learning framework [14].

3 Methodology

3.1 Overall Framework

A deep learning architecture comprising Convolutional Neural Networks (CNN), Bidirectional Gated Recurrent Units (BiGRU), and an Attention Mechanism (Attention - Based CNN – BiGRU) is incorporated into the suggested hybrid model. The fact that each part of this design tackles a certain time-series forecasting limitation serves as its inspiration: Grey Wolf Optimizer optimizes model hyperparameters, CNN collects local temporal information, BiGRU records long-term relationships, and LightGBM improves prediction accuracy and generalization. This combination helps in minimizing errors, as measured by evaluation metrics like MAE, MSE, RMSE, and R^2 also ensures more reliable and precise market price predictions, making it a suitable choice for forecasting stock or market trends [15].

Figure 1 illustrates the proposed hybrid CNN–BiGRU–GWO–LightGBM framework. ARIMA preprocessing is first applied to remove linear trends. CNN layers extract local features, followed by BiGRU layers with an attention mechanism to model temporal dependencies. GWO optimizes hyperparameters, and LightGBM produces the final prediction.

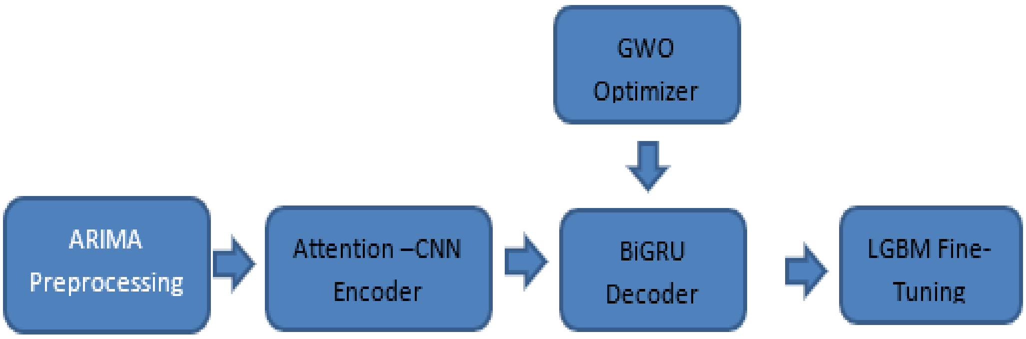

Figure 1: Proposed CNN–BiGRU–GWO–LightGBM Framework

3.2 CNN Feature Extraction

The CNN layers apply one-dimensional convolutional filters to historical stock price sequences to extract short-term temporal patterns and suppress noise. Given an input sequence

$X = [X_1 \, X_2, ..., X_T]$ the convolution operation is defined as:

$$\mathcal{Y}_t = f\left(\sum_{i=0}^{k-1} w_i \cdot x_{t+i} + b\right) \quad (1)$$

where k is the kernel size, w_i denotes the convolutional weights, b is the bias term, and $f(\cdot)$ represents the activation function. In this study, multiple convolutional layers were employed with filter sizes of k=3 and k=5, followed by max-pooling layers to reduce dimensionality and enhance feature robustness before feeding the extracted features into the BiGRU layer

3.3 BiGRU with Attention

The BiGRU processes input sequences in both directions that is forward and backward directions to capture temporal dependencies from past and future contexts. Let h_t and h_t denote the hidden states of the forward and backward GRU at time step t, respectively. The combined BiGRU output is given by;

$$h = [h_t \cdot h_t] \tag{2}$$

3.4 Hyperparameter Optimization Using GWO

The Grey Wolf Optimizer (GWO) was utilized to optimize essential hyperparameters of the suggested model, such as the learning rate, quantity of convolutional filters, number of BiGRU hidden units, dropout rate, and time-window length. The fitness function was defined as the validation Mean Squared Error (MSE), which GWO sought to reduce. The optimizer was setup with a population of 20 wolves and operated for 50 iterations. In each iteration, candidate solutions were assessed by training the CNN–BiGRU model and calculating the validation loss. For the final model training, the ideal hyperparameter set that matched the lowest fitness value was chosen.

<u>Algorithm 1: GWO-based Hyperparameter Optimization</u>
Input: Hyperparameter search space, population size N, maximum iterations T
Output: Optimal hyperparameter set θ*
1: Initialize wolf population {θ1, θ2, ..., θN}
2: Evaluate fitness (validation MSE) of each wolf
3: Identify α, β, and δ wolves
4:   for t = 1 to T do
5:     Update positions of wolves using GWO equations
6:     Evaluate fitness of updated wolves
7:     Update α, β, and δ
8: end for
9: Return θ* corresponding to α wolf

4 Experimental Setup

4.1 Dataset Description

Tushare database is the dataset for the Chinese stock market was used to gather historical stock price data. Specifically, daily trade data of the chosen stock or stocks listed on the Shanghai and Shenzhen Stock Exchanges were used (e.g., [insert exact stock name or ticker, such as China Merchants Bank – 600036.SH]). The dataset, which spans January 1, 2007 to March 31, 2022, contains Open, High, Low, Close, and Volume attributes. All input features were normalized using Min–Max normalization, which was defined as:

$$X_{scaled} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{3}$$

where X_{min} and X_{max} denote the minimum and maximum values of each feature, respectively. Normalization prevents dominance of features with larger scales and improves learning efficiency in the deep learning model Dataset description incomplete.

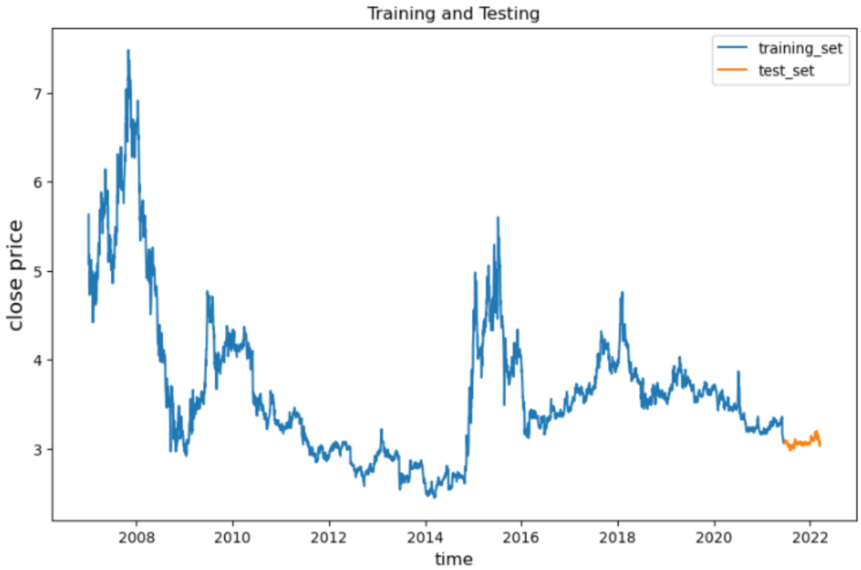

Figure 2: Training and Testing

4.2 Evaluation Metrics

Model performance was evaluated using Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and the coefficient of determination (R2).

RMSE is a measure of the average squared difference between actual stock prices and expected stock prices. It is prone to outliers because it penalizes large errors more harshly than minor ones due to squaring. A lower RMSE indicates a better model.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \qquad (4)$$

where:

- y_i = Actual stock price at time
- y_i= Predicted stock price at time
- n= Total number of observations

Interpretation:

- RMSE close to zero indicates accurate predictions.
- High RMSE suggests significant prediction errors.

Mean Absolute Error (MAE): The MAE calculates the average absolute difference between forecast and actual stock prices. Because it treats all errors equally without squaring them, it is simpler to read than RMSE.

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \qquad (5)$$

Interpretation:

- A smaller MAE means better model performance.
- Less sensitive to large errors compared to RMSE

R-Square (R2) – Coefficient of Determination: R2 measures how well the model explains stock price volatility. It has a range of 0 to 1, where 1 represents a perfect prognosis.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \qquad (6)$$

where:

- $\hat{y}$ = Mean of actual stock prices

Interpretation:

- $R^2 = 1$: Model perfectly fits the data.
- $R^2 > 0.8$: Strong predictive capability.

$R^2 < 0.5$: Poor fit

Mean Square Error (MSE): The average squared difference between the expected and actual values is measured by the MSE.

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \qquad (7)$$

Where:

- y_i = Actual stock price value
- y_i= Predicted stock price a
- n= Total number of observations

Interpretation

- Lower MSE = better prediction (closer to actual values).
- Higher MSE = poor prediction (model has large errors).

5 Results and Discussion

5.1 Performance Comparison

Compared with the baseline ACNN-BiLSTM model, the proposed CNN-BiGRU-GWO framework reduces MAE, MSE, and RMSE by approximately 29.55%, 45.00%, and 27.67%, respectively. In addition, the coefficient of determination (R^2) improves from 0.88342 to 0.93903, representing a 5.56% increase in explanatory power. These research results get demonstrate that the proposed model achieves significantly lower prediction error while providing more accurate and robust stock price forecasting, as shown in Table 1 below:

Table 1: Performance Comparison of Different Models (Comparison with baseline model)

METRICS	CNN-BiGRU-GWO	ACNN-BiLSTM	IMPROVEMENT
MAE	0.00789	0.01120	Improved
MSE	0.00011	0.00020	Slightly Improved
RMSE	0.01030	0.01424	Improved
R^2	0.93903	0.88342	Improved

5.2 Training and Validation Analysis

Figures 3 and 4 show the training and validation loss curves for the BiGRU and proposed hybrid model, respectively.

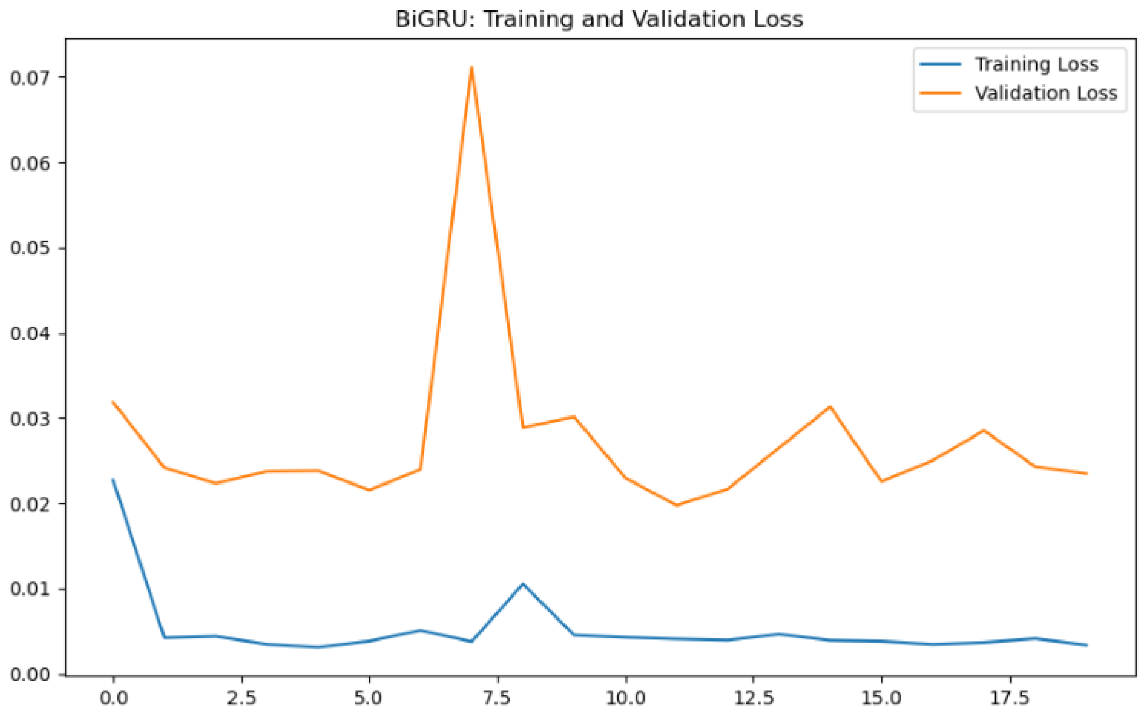

Figure 3: Training and Validation Loss for BiGRU Model

Received: 29/12/2025; Revised: 09/01/2026; Accepted: 19/01/2026; Published: 27/06/2026

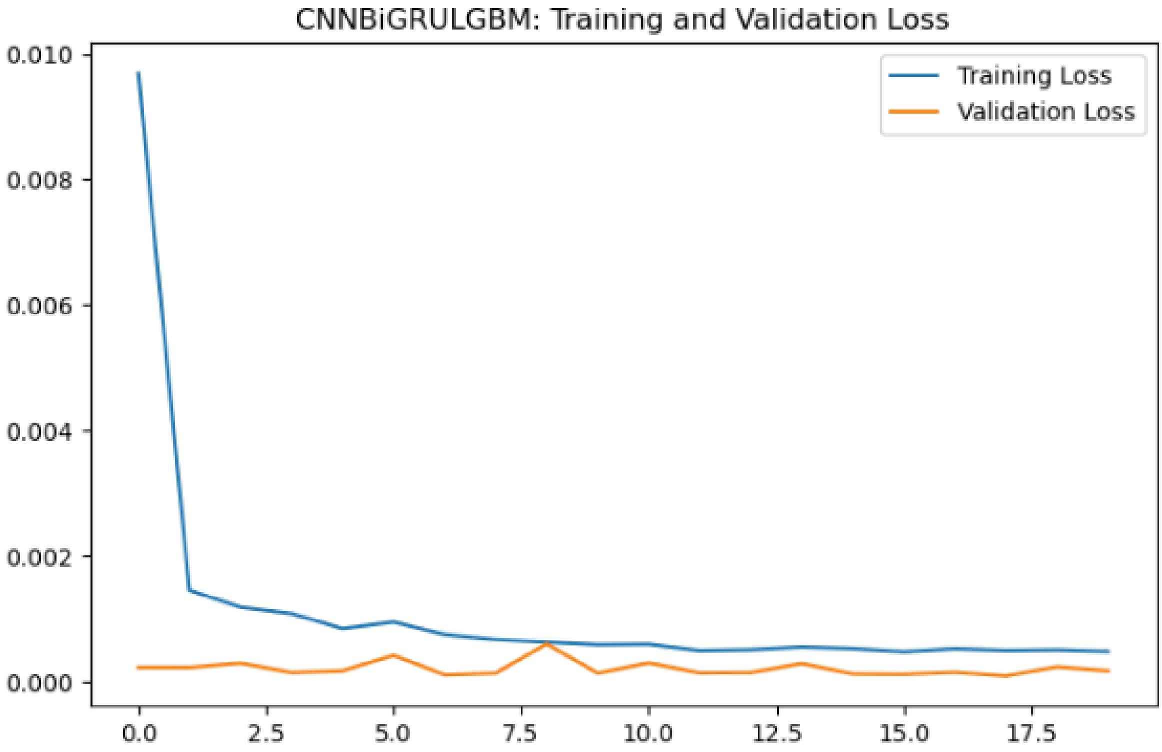

Figure 4: Training and Validation Loss for Proposed Model

The results indicate that integrating CNN feature extraction, BiGRU temporal modeling, GWO optimization, and LightGBM regression significantly improves prediction accuracy.

6 Conclusion

This paper proposed a hybrid CNN–BiGRU–GWO–LightGBM model for stock price prediction. Experimental results confirm superior performance over statistical models and baseline deep learning approaches. The integration of meta heuristic optimization and ensemble learning enhances robustness and generalization, making the model suitable for real-world financial forecasting.

References

[1] Singh, R., & Srivastava, S. (2017). Stock market forecasting using machine learning techniques. Procedia Computer Science, 122, 1491–1498.
[2] Gupta, G. S., Gupta, P., Gupta, S. K., & Jadon, R. S. (2020). Financial time-series forecasting: Challenges and solutions. Expert Systems with Applications.
[3] Vargas, M. (2017). Time-series modeling for stock price prediction. Journal of Financial Analytics.
[4] Sonkavde, G., Dharrao, D. S., et al. (2023). Forecasting stock market prices using machine learning and deep learning models: A systematic review. International Journal of Financial Studies, 11(3), 94.
[5] Wu, Y., & Zhang, H. (2017). Deep learning for stock price prediction. Neurocomputing
[6] Wang, S. (2023). Comparison of LSTM and BiGRU for time-series forecasting. Applied Soft Computing.
[7] Dadhich, R., Rezaei, H., Faaljou, H., & Mansourfar, G. (2021). Predictive models for stock market index using stochastic time series. Journal of Economics.
[8] Livieris, I., Li, Z., Zhu, W., Zhu, B., Wang, B., & Wang, Q. (2020). Deep learning models for stock market forecasting. AlgorithmsI.
[9] Dadhich, M., Pahwa, M. S., Jain, V., & Doshi, R. (2021). Predictive models for stock market index using stochastic time series ARIMA modeling in emerging economy. Lecture Notes in Mechanical Engineering, 281–290. https://doi.org/10.1007/978-981-16-0942-8_26
[10] Daul, S., Jaisson, T., & Nagy, A. (2022). Performance attribution of machine learning methods for stock returns prediction. Journal of Finance and Data Science, 8, 86–104. https://doi.org/10.1016/j.jfds.2022.04.002
[11] Wang, S. (2023). A stock price prediction method based on BiLSTM and improved Transformer. IEEE Access, 11, 104211–104223. https://doi.org/10.1109/ACCESS.2023.3296308
[12] Mekelburg, E., & Strauss, J. (2024). Pooling and winsorizing machine learning forecasts to predict stock returns with high-dimensional data. Journal of Empirical Finance, 79. https://doi.org/10.1016/j.jempfin.2024.101538.

[13] Abotaleb, M., Lim, W. H., Mishra, P., Qenawy, A.-S. T., & Almetwally, E. (2024). Enhancing stock price prediction accuracy using ARIMA and advanced Greylag Goose Optimizer algorithm. Journal of Artificial Intelligence in Engineering Practice, 1(1), 49–65. https://doi.org/10.21608/jaiep.2024.355004

[14] Ke, G., Su, Y., & Jiang, X. (2017). LightGBM: A highly efficient gradient boosting decision tree. Advances in Neural Information Processing Systems (NeurIPS).

[15] Dai, H., Wang, W., Cao, J., & Wu, H. (2021). A deep neural network for stock price prediction. Journal of Physics: Conference Series, 1994(1). https://doi.org/10.1088/1742-6596/1994/1/012029