Metaheuristic Optimization Review

ISSN: 3066-280X Vol: 06, No. 01, 2026 PP. 29–39

Article DOI: https://doi.org/10.54216/MOR.060103

FULL LENGTH ARTICLE

Enhanced Stock Price Forecasting: Time Series Analysis with ARIMA and FGGO Optimization

Laith Farhan[1,*] **Raad S. Alhumaima**[2]

[1] School of Engineering, Manchester Metropolitan University, Manchester, M1, UK

[2] Brunel University, Uxbridge UB8 3PH, U.K

Emails: l.al-bayati@mmu.ac.uk · 1234914@alumni.brunel.ac.uk

Received: January 28, 2026 Revised: March 15, 2026 Accepted: May 12, 2026 ⋆ Corresponding author

ABSTRACT

Forecasting financial markets remains a persistent challenge due to the nonlinear, stochastic, and nonstationary nature of stock price dynamics. This study is motivated by the need to enhance the robustness and adaptability of traditional statistical forecasting models through intelligent optimization. We propose an advanced hybrid framework that integrates the AutoRegressive Integrated Moving Average (ARIMA) model with the Fitness Greylag Goose Optimization (FGGO) algorithm—a refined metaheuristic inspired by collective behavioral intelligence and adaptive search strategies. The primary contribution of this research lies in the methodological fusion of classical time series modeling with dynamic metaheuristic optimization to improve predictive accuracy, convergence stability, and resistance to local optima. Comparative experiments on the historical stock prices of PT Bank Central Asia Tbk (BBCA.JK) demonstrate a substantial performance uplift: the baseline ARIMA model achieved a Mean Squared Error (MSE) of 0.0333, whereas the FGGO-optimized ARIMA reduced the MSE dramatically to 0.0038, outperforming other optimization techniques such as the Genetic Algorithm (GA), Whale Optimization Algorithm (WOA), and Particle Swarm Optimization (PSO). These results confirm that FGGO significantly enhances ARIMA's capacity to capture intricate temporal dependencies and volatile market structures. The implications of this study extend beyond finance, offering a scalable, explainable, and high-performance optimization paradigm for diverse time series forecasting applications in economics, engineering, and intelligent decision-support systems.

Keywords: Financial Time Series Forecasting ▪ AutoRegressive Integrated Moving Average (ARIMA) ▪ Fitness Greylag Goose Optimization (FGGO) ▪ Metaheuristic Optimization ▪ Hybrid Predictive Modeling

1. INTRODUCTION

The issue of financial market prediction has been identified as one of the most complex and challenging aspects in modern financial analysis. It is complicated by the fact that volatile markets are characterized by the complex interaction of unpredictable and changing factors, including macroeconomic conditions, geopolitical events, corporate performance metrics, and the moods, prejudices, and behavioral patterns of investors themselves [1]. The externalities can lead to disproportionately large impacts on asset prices, which can include unexpected policy changes, global crises, or even relatively environmental changes. This can be attributed to the nonlinear and highly dynamic nature of financial markets, such that perturbations can cause various prices to change in a cascading manner relatively quickly. This suggests that statistical and computational issues are also associated with forecasting financial time series, which is likely to be the case in other areas of predictive analytics [2].

Traditional methods of financial forecasting include econometric and statistical modeling [3], which typically assume that the data is stationary and linear. Nonetheless, the empirical statistics on finances are notoriously non-stationary, heavy-tailed, long-memory, and intermittent, which is often obscured by noise in practice [4]. These properties render traditional linear models ineffective in modeling the latent relationships and complex interactions that are present in financial markets. The weaknesses of these models have sparked increased interest in developing new methodologies that can address the nonlinear, chaotic, and stochastic nature of stock price dynamics. The requirement, therefore, is to develop forecasting structures that are sound and adaptable, capable of generalizing to various market conditions while limiting overfitting to temporary abnormalities.

The past decade has witnessed a paradigm shift, and Artificial Intelligence (AI) and Machine Learning (ML) have become integral to the financial forecasting paradigm [5]. These technologies are offered as an alternative to the traditional concept of statistics, as they enable systems to discover complex relationships in data independently, recognize latent structures, and respond to changing trend directions . Both stochastic and structural variations in time series can be addressed using AI and ML in predictive modeling. This is critical in financial forecasting, where noise, cyclic variation, and random shocks coexist with long-term growth trends [6]. The AI and ML solutions, based on large amounts of data and sophisticated algorithmic approaches, are not only capable of improving the predictive power of such models but may also be applied to acquire additional information regarding the underlying mechanics of market functioning [7].

The strength of AI- and ML-based methods is their size and adaptability. It is possible that, unlike inflexible econometric models, AI systems can handle high-dimensional data and learn valuable features without making any particular assumptions. Such flexibility enables them to accommodate financial time series, which are often characterized by structural breaks, volatility clustering, and nonlinear dependencies [8]. This approach represents a significant contrast to traditional methods, which are based on simplified depictions of cause-and-effect relationships. With procedures such as deep feature learning and refinement, AI systems may uncover patterns that are not visible to conventional models, leading to more accurate predictions and increased explanatory power [9]. Recent developments have also shown that hybrid solutions (a combination of the mathematical skill of classical statistical models and the learning skills of AI) are an auspicious direction in financial forecasting [10].

The key components of any forecasting system are data preprocessing and decomposition. The quality of the data on which any model is trained, as well as the suitability of the preprocessing methods adopted, determine the model's success. Time series data are typically composed of various overlapping elements, such as long-term trends, intra-seasonal fluctuations, and stochastic noise. It is essential to separate these elements effectively to minimize redundancy, enhance interpretability, and improve model reliability [11]. In this context, the AutoRegressive Integrated Moving Average (ARIMA) model has traditionally been prominent in the field of time series forecasting, owing to its sound statistical basis and demonstrated capability of modeling stationary processes [12]. However, the application of ARIMA and other models is limited in practice, as they assume stationarity and linear dynamics. Since financial data often does not fit into these conditions, the performance of such models is worse when applied to real-life cases, which requires methodological improvements [13].

To overcome these drawbacks, researchers have increasingly turned to the use of metaheuristic optimization techniques, which provide a versatile and robust alternative to conventional parameter estimation methods. Metaheuristics are biologically and physically motivated algorithms that simulate natural mechanisms, including swarm intelligence, predator-prey dynamics, and evolutionary adaptation, to effectively explore significant and complex solution spaces [14]. Metaheuristic algorithms can search for local minima and move towards near-optimal solutions by engaging in exploration and exploitation, respectively. The property is especially useful in time series forecasting, particularly when the optimization space is non-convex and has a large number of local minima. The most notable examples of these algorithms are the Grey Wolf Optimizer (GWO), Whale Optimization Algorithm (WOA), and Firefly Algorithm (FA), all of which have proven to be highly successful in various optimization tasks, such as financial prediction [15].

Building on these improvements, current innovations have developed new hybrid and improved algorithms, including the Focused Grey Goose Optimization (FGGO), a hybridization of both exploration and fine-tuning exploitation algorithms to achieve a more efficient trade-off during optimization activities [16]. A significant advancement of the existing metaheuristics techniques is that FGGO involves the use of additional heuristics to improve the rate of convergence and the stability. FGGO may also be helpful in time series forecasting, as an encouraging approach to dynamically changing model parameters, and thus better adapted to noise and more prone to successful predictions. The existence of such optimization techniques to supplement statistical models, such as ARIMA, highlights the need to employ hybrid techniques that leverage the potential of both traditional and contemporary approaches.

2. MATERIALS AND METHODS

The general structure of the proposed stock price forecasting methodology is shown in Figure 1. It starts with the purchase of historical stock market datasets, which are then processed through extensive data preprocessing to handle missing data points, normalize feature values, and maintain data integrity. This is followed by training predictive models on the preprocessed data, including ARIMA, Moving Average, Exponential Smoothing, and Linear Regression. To improve forecast accuracy, a Fitness Greylag Goose Optimization (GGO) algorithm is incorporated to optimize the model's parameters and performance. Lastly, the evaluation's performance is assessed using statistical and error-based metrics to identify the best predictive method.

3. DATASET DESCRIPTION

The dataset utilized in this study comprises the historical stock prices of *PT Bank Central Asia Tbk* (ticker symbol:

Table 1. Summary of Related Works in Financial Forecasting and Artificial Intelligence

Reference	Objective	Methodology	Key Findings
[17]	To enhance portfolio management using AI and ML.	Employed deep learning architectures such as CNN and RNN with reinforcement learning for portfolio optimization.	Demonstrated that DL-based models improve predictive accuracy, adaptability, and risk-adjusted returns compared to traditional econometric methods.
[18]	To detect financial statement fraud using structured and unstructured data.	Proposed a hierarchical attention network combining financial ratios with textual MD&A features.	Achieved higher fraud detection accuracy with interpretable “red-flag” textual indicators, improving financial transparency.
[19]	To assess the financial impact of ESG factors on European public enterprises.	Applied ML and inferential analysis on ESG indicators and financial ratios (ROE, ROA).	Found positive correlations between ESG performance and financial outcomes, highlighting ESG as a driver of sustainable value creation.
[20]	To develop an early financial risk warning system.	Built a backpropagation neural network with factor analysis–based feature selection.	Attained over 95% prediction accuracy, outperforming traditional statistical models in firm-level risk prediction.
[21]	To improve Internet finance risk management.	Combined ant colony optimization (ACO) with a radial basis function (RBF) neural network.	Reduced forecasting error (MSE = 0.249) and improved convergence speed and iteration efficiency compared to standard algorithms.
[22]	To enhance credit scoring performance using sequential modeling.	Designed a stacked unidirectional/bidirectional LSTM for temporal feature extraction from static loan data.	Achieved accuracies above 97%, surpassing traditional models while maintaining interpretability and deployment simplicity.
[23]	To map research trends in AI-driven finance.	Conducted a large-scale bibliometric review (1986–2021) using co-citation and clustering analyses.	Identified three main clusters: portfolio management, fraud detection, and sentiment-based forecasting, emphasizing hybrid and explainable modeling.
[24]	To test predictability in cryptocurrency markets.	Evaluated ML models (RNN, gradient boosting) for Bitcoin price direction using multi-source features.	RNNs delivered best classification performance, though economic profitability declined after transaction cost adjustments.
[25]	To handle nonlinear and nonstationary financial series.	Combined empirical mode decomposition (EMD) with LSTM forecasting.	The EMD–LSTM hybrid reduced RMSE and improved stability compared to standalone LSTM models.
[26]	To improve multiscale financial forecasting.	Integrated CEEMD, PCA, and LSTM into a hybrid decomposition–learning framework.	Achieved superior accuracy and directional symmetry, confirming robustness and profitability of the hybrid pipeline.

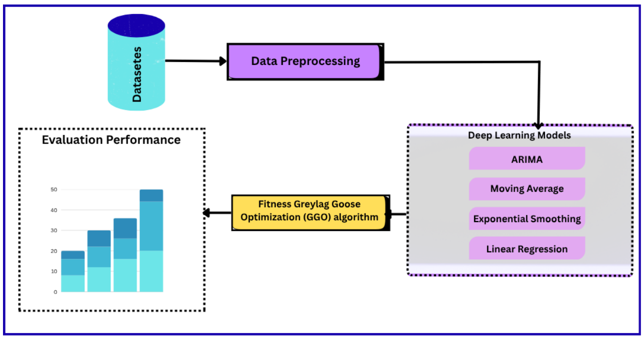

Figure 1. Proposed framework for stock price forecasting integrating data preprocessing, deep learning models, and Fitness Greylag Goose Optimization (GGO) algorithm.

BBCA.JK), one of the largest financial institutions listed on the Indonesia Stock Exchange (IDX). This dataset provides a rich temporal record of the bank’s market performance and serves as a suitable proxy for examining forecasting methodologies in the context of emerging Asian financial markets. The inclusion of a sufficiently long observation period allows for the identification of patterns, seasonal fluctuations, and volatility clusters, which are critical for evaluating both statistical and metaheuristic forecasting models. The dataset consists of the standard attributes commonly employed in financial time series analysis, as summarized in Table 2. Each record corresponds to a single trading day and encapsulates multiple market variables reflecting price behavior and trading activity.

Table 2. Structure of the historical stock price dataset for PT Bank Central Asia Tbk (BBCA.JK).

Attribute	Description
`Date`	The trading date corresponding to each stock record, expressed in the format YYYY–MM–DD.
`Open`	The price of the stock at the commencement of the trading session.
`High`	The maximum price reached by the stock during the trading day.
`Low`	The minimum price recorded during the trading session.
`Close`	The final transaction price at the end of the trading session.
`Adj Close`	The adjusted closing price, corrected for corporate actions such as dividends or stock splits.
`Volume`	The total number of shares traded during the trading session, representing market liquidity.

The inclusion of both the raw and adjusted closing prices (`Close` and `Adj Close`) is particularly valuable for the subsequent modeling stage, as it allows for the comparison between nominal and adjusted returns. The `Volume` variable provides additional insight into market activity and investor behavior, serving as a potential explanatory factor for price volatility and momentum. All data were collected from a publicly available financial database (Yahoo! Finance) and preprocessed to ensure consistency, completeness, and chronological ordering prior to statistical and metaheuristic analysis. In order to examine the distributional properties of the daily stock returns, Figure 2 presents both the Q-Q plot and the kernel density estimate (KDE) of the returns. The Q-Q plot on the left compares the empirical quantiles of the daily returns with those of a normal distribution, allowing for a visual assessment of normality. Meanwhile, the density plot on the right depicts the distribution of daily returns, highlighting its leptokurtic nature and the presence of heavy tails often observed in financial time series data.

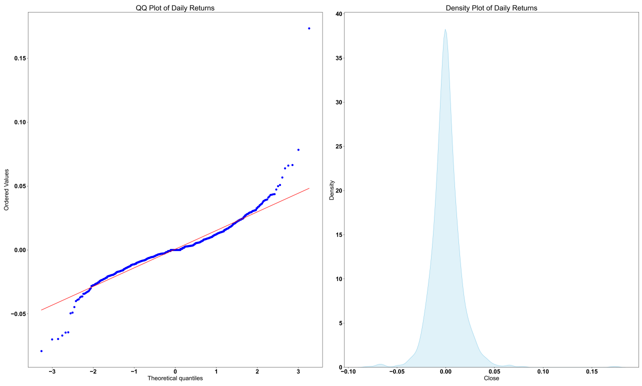

Figure 2. Q-Q Plot and Density Plot of Daily Returns.

To evaluate the linear relationships among the key stock market variables, Figure 3 displays the correlation heatmap for the dataset features. This visualization provides insight into how strongly the variables such as *Open*, *High*, *Low*, and *Close* prices are interrelated, as well as their relationship with *Adjusted Close* and *Volume*. The heatmap reveals that price-based variables exhibit near-perfect positive correlations, indicating their mutual movement patterns, whereas *Volume* shows a weak negative correlation with price features, suggesting limited association with daily price fluctuations.

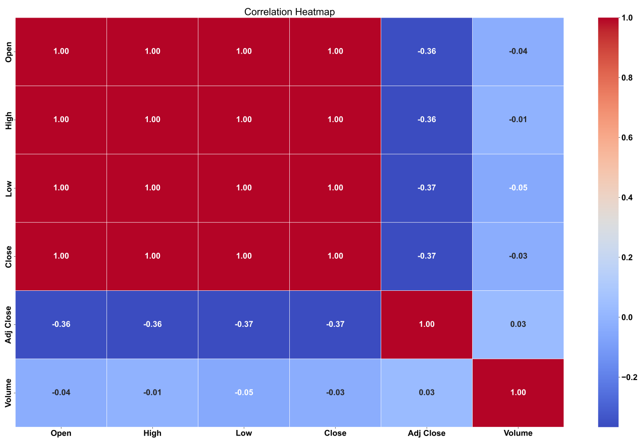

Figure 3. Correlation Heatmap of Stock Market Features.

3.1 Data Preprocessing

The dataset underwent a structured preprocessing phase to ensure that the data were clean, consistent, and suitable for subsequent modeling and analysis. Handling missing values was a crucial first step, as the presence of incomplete records could introduce bias or distort temporal dependencies. In this study, missing observations were managed either through removal, when the proportion of missingness was negligible, or through statistical imputation methods such as forward and backward filling to maintain temporal continuity within the time series. This approach ensured that the sequential nature of the financial data was preserved and that no artificial discontinuities were introduced during data preparation. Since all the variables in the dataset—namely `Open`, `High`, `Low`, `Close`, `Adj Close`, and `Volume`—were continuous numerical attributes, no categorical variables were present. Consequently, feature encoding was not applicable in this context. The dataset did not require transformations such as one-hot encoding or ordinal mapping, which are typically employed in datasets containing qualitative or discrete attributes. Feature scaling was then performed to normalize the range of the variables and to facilitate faster and more stable convergence during the model training process. Two common normalization techniques were employed: min–max normalization and z–score standardization. The min–max normalization technique linearly transformed the original data into a normalized range, typically between 0 and 1, according to the formula $x' = (x - x_{\min})/(x_{\max} - x_{\min})$. In contrast, the z–score standardization technique standardized the data to have zero mean and unit variance following the formula $x' = (x - \mu)/\sigma$. These transformations ensured that all input features contributed proportionally to the learning process, preventing attributes with larger numerical scales—such as `Volume`—from dominating the model optimization. After normalization, a comprehensive feature correlation and dependency analysis was carried out to understand the underlying temporal relationships within the dataset. Time-series decomposition techniques were applied to separate the observed price data into trend, seasonal, and residual components, thereby facilitating a clearer interpretation of the inherent structural patterns in the stock movements. Furthermore, autocorrelation and partial autocorrelation plots were generated to identify the degree of correlation between the present and lagged values of the time series. This analysis was essential for determining the appropriate lag parameters for the ARIMA and hybrid optimization models. Lag dependency and redundancy checks were also performed to mitigate multicollinearity and ensure that the predictive models relied only on the most informative temporal features. Through these combined preprocessing steps, the dataset was transformed into a clean, normalized, and well-structured form, ready for effective modeling and forecasting using both traditional statistical and metaheuristic optimization techniques.

4. MACHINE LEARNING MODELS

The forecasting models used in this study were selected based on their theoretical applicability in temporal sequence modeling and their established results in financial time series analysis. When choosing models, the ability to form sequential dependencies and stochastic variations was not the only criterion for promoting them in the framework of improved stock price forecasting; interpretability and computational efficiency were also considered. The models chosen, which include AutoRegressive Integrated Moving Average (ARIMA), Moving Average (MA), Exponential Smoothing (ES), and Linear Regression (LR), are a wide range of classical models of forecasting that as a combination of each offer a broad background of understanding the stock price movements within a regression framework. These models offer methodological benefits and represent various structural properties of time-dependent data.

The AutoRegressive Integrated Moving Average (ARIMA) model is considered one of the strongest and most popular time series forecasting models, based on statistical analysis. It is a combination of three critical terms: autoregressive (AR) term that takes into consideration the correlation of an observation and a given number of lagged observations; the differencing term (I), which ensures that there is no correlation and that the term removes trends and seasonality; and the moving average (MA) term, which considers the dependence of an observation and the prevailing errors of the past time steps. A model of ARIMA is mathematically stated as ARIMA (p, d, q).

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} \tag{1}$$

$$+ \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t. \tag{2}$$

And where p, d and q are the autoregressive order, differencing order and moving average order, respectively. The ARIMA model is especially applicable to financial time series because it can model the autocorrelation structures, trends and patterns of volatility. Nonlinear dynamics in financial data may, however, be further improved by combining ARIMA with metaheuristic optimization algorithms in parameter tuning to promote stability. **The Moving Average (MA) model**, a fundamental element of ARIMA,, is specifically concerned with analyzing the serial correlation of the error terms in a time series. The current value of the series in this model takes the following form: a linear combination of past forecast errors, thereby averaging short-term variations and emphasizing longer-term variations. The MA model is instrumental in capturing the short-term memory of a stationary time series. It is interpretable and straightforward, which is why it serves as a necessary baseline for measuring the performance of more complex forecasting architectures. **The Exponential Smoothing (ES)** model is another time series forecasting model that assigns exponentially declining weights to previous observations. In contrast to other simple averaging methods, which assign equal importance to all past data, exponential smoothing places more emphasis on recent data points, enabling the model to adjust more quickly to market structural changes or shocks. The overall equation of the single exponential smoothing model is as follows:

$$\hat{y}_{t+1} = \alpha y_t + (1-\alpha)\hat{y}_t,$$

And, where $\alpha(in[0,1])$ is the smoothing constant which measures how old observations are discounted. This framework can be generalized to incorporate the linear and seasonal approaches of Holt, including trend and seasonality factors. The exponential smoothing models are also computationally efficient and work exceptionally well with datasets that exhibit level or trend patterns with minimal noise.

Although this method was developed to compute cross-sectional data, **linear regression (LR)** can be successfully used in time series prediction when temporal correlations are represented as lagged features. LR, in its simplest form, determines a linear relationship between a dependent variable, i.e., the closing price of the stock, and one or more independent variables, which can include past prices, volume, or calculated technical indicators. The linear regression model is presented as follows.

$$y_t = \beta_0 + \beta_1 x_{1,t} + \beta_2 x_{2,t} + \cdots + \beta_n x_{n,t} + \varepsilon_t,$$

where β_i denotes the model coefficients and ε_t represents the error term. Although LR is a simple model, it is significantly interpretable and serves as a powerful point of reference when comparing the benefits of nonlinear or hybrid modeling. Linear regression, in combination with regularization or metaheuristic tuning, can compete with other forecasting methods in short-term financial prediction problems. Overall, the choice of ARIMA, MA, Exponential Smoothing, and Linear Regression models in the study represents a conscious trade-off between theoretical and empirical soundness on the one hand, and computational and empirical relevance to financial forecasting on the other. All the models contribute to the overall strength of the experimental design, as both linear and adaptive temporal relations are thoroughly analyzed before incorporating metaheuristic optimization techniques to achieve greater predictive precision.

4.1 Fitness Greylag Goose Optimization (FGGO) Algorithm

The Fitness Greylag Goose Optimization (FGGO) algorithm is a crucial component of the presented methodology. It represents a more effective implementation compared to the newly constructed Greylag Goose Optimization (GGO) algorithm. It aims to improve the balance between exploration (searching widely to find a variety of solutions) and exploitation (vigorously refining known solutions), thereby solving typical optimization problems such as premature convergence and difficulty in escaping local optima. The FGGO algorithm begins by generating a population of possible solutions, also known as individuals, using a randomized population creation mechanism. Every single person is represented as x_i, and the numbers between i and N (the total population size) are distinct candidate solutions to the problem under investigation. These individuals evaluate others through an objective function, represented as $f(x_i)$, which determines the quality of each solution based on the satisfaction of the optimization criteria. After the analysis, the most optimal solution that has been discovered is identified and represented as x^*. This serves as the starting point for further optimization. One of the main characteristics of FGGO is that it is a dynamic mechanism for grouping, breaking the population into two subgroups: the exploration group (G_e) and the exploitation group (G_x). The exploration team aims to explore a wide part of the solution space to identify new possible solutions. In contrast, the exploitation team strives to narrow down the possible solutions that have been found so far and are the most promising. To begin with, both groups have the same size, with half of the population in each. As the optimization process proceeds, FGGO dynamically adjusts the sizes of these groups in accordance with the performance of the best solution, x^*. Throughout the successive steps, the exploration group size (G_e) has been reduced, and more resources have been dedicated to the exploitation group (G_x) to fine-tune solutions as the algorithm approaches its optimum. To further enhance its adaptability, FGGO incorporates a critical mechanism to prevent stagnation in the optimization process. If the value of the objective function for the best solution remains unchanged for three consecutive iterations, the algorithm reallocates resources by increasing the size of the exploration group (G_e). This adjustment enables the algorithm to escape local optima and continue searching for potentially superior solutions in unexplored areas of the solution space. This dynamic interplay between exploration and exploitation, coupled with its adaptive behavior, equips FGGO with the robustness and flexibility necessary to tackle complex optimization problems effectively, as formally described in Algorithm 1.

Algorithm 1 Proposed Fitness Greylag Goose Optimization (FGGO) Algorithm

Initialize population X_i $(i = 1, 2, \dots, n)$, size n, max iters $t_{\max}$, objective F_n
Initialize $A, C, b, l, c, r_1, \dots, r_5, w_1, \dots, w_4, A_1, A_2, A_3, C_1, C_2, C_3$; set $t \leftarrow 1$
Evaluate F_n for each X_i; set $P \leftarrow$ best agent
Split into exploration size n_1 and exploitation size n_2
while $t \leq t_{\max}$ **do**
  for $i = 1 \to n_1$ **do** ▷ Exploration
    if $r_6(2) = 0$ **then**
      if $r_5 > 0.5$ **then**
        if $|A| < 1$ **then**
          Update using Eq. (1)
        else
          Pick $X_{\text{Paddle1}}, X_{\text{Paddle2}}, X_{\text{Paddle3}}$
          Update z via Eq. (2); update position via Eq. (3)
        end if
      else
        Update using Eq. (4)
      end if
    else
      Update using Eq. (6)
    end if
  end for
  for $i = 1 \to n_2$ **do** ▷ Exploitation
    if $r_6(2) = 0$ **then**
      Compute X_1, X_2, X_3 via Eqs. (7)–(9); set $X(t+1)$ to their mean
    else
      Update using Eq. (6)
    end if
  end for
  Evaluate F_n for each X_i; update parameters
  $t \leftarrow t + 1$; repair out-of-bounds solutions
  if best F_n unchanged for 3 iters **then**
    Increase n_1; decrease n_2
  end if
end while
return P ▷ best agent

5. EMPIRICAL RESULTS

Here is a detailed empirical analysis of the proposed forecasting model, which will include a performance comparison of the conventional time series models, AutoRegressive Integrated Moving Average (ARIMA), Moving Average (MA), Exponential Smoothing (ES) and Linear Regression (LR) to determine a baseline level of predictive accuracy and reliability. Performance is quantified by using a vast assortment of statistical metrics, such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), correlation coefficient (r), coefficient of determination (R2), NashSutcliffe Efficiency (NSE), and Willmott Index (WI). On this basis, the ARIMA can be further refined with the aid of metaheuristic optimization models: Fitness Greylag Goose Optimization (FGGO), Genetic Algorithm (GA), Whale Optimization Algorithm (WOA) and Particle Swarm Optimization (PSO) to determine to what degree adaptive parameter optimization can be used to enhance prediction accuracy and stability of the models. A multidimensional understanding of model behavior, variability, and robustness is provided through complementary visual analyses, such as violin plots, radar charts, parallel coordinate plots, and hierarchical clustering, to offer a holistic view of the comparative advantages that the proposed FGGO-optimized ARIMA model can offer.

5.1 Baseline Deep Learning Performance

To establish benchmark comparisons, several classical time series models were evaluated, including ARIMA, Moving Average, Exponential Smoothing, and Linear Regression. Among these, **ARIMA demonstrated the best predictive performance**, achieving the highest R^2 value (0.8704) while also yielding the lowest MSE (0.0333). This indicates that ARIMA was the most effective at capturing the underlying dynamics of the data compared to the other baseline models.

Table 3 presents the full set of performance metrics, including Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Bias Error (MBE), correlation coefficient (r), coefficient of determination (R^2), Relative RMSE (RRMSE), Nash–Sutcliffe Efficiency (NSE), and Willmott's Index (WI).

Table 3. Baseline model performance comparison

Models	MSE	RMSE	MAE	MBE	r	R^2	RRMSE	NSE	WI
ARIMA	0.0333	0.1825	0.1571	0.0224	0.8578148	0.8704	20.4104	0.841	0.9111
Moving Average	0.0399	0.2190	0.1885	0.0269	0.6862	0.6988	22.1	0.802	0.7288
Exponential Smoothing	0.0466	0.2555	0.2199	0.0314	0.5146	0.5272	22.8	0.751	0.5466
Linear Regression	0.0533	0.2921	0.2513	0.0359	0.3431	0.3557	23.3	0.712	0.3644

The results highlight several important insights:

- **ARIMA**: Exhibits the lowest error metrics (MSE, RMSE, and MAE) and the highest correlation coefficient ($r = 0.8578$). Its strong R^2 (0.8704) and high Willmott's Index (0.9111) confirm its superior ability to capture both the trend and variability of the time series. This is expected, as ARIMA explicitly models both autoregressive and moving average components, making it highly effective for structured time series data.

- **Moving Average**: Performs moderately well, but lags behind ARIMA. With an R^2 of 0.6988, it explains significantly less variance. While RMSE and MAE remain relatively low, the higher bias (MBE = 0.0269) suggests systematic deviations from the actual values. Its performance illustrates the limitations of purely averaging past values without accounting for seasonality or autoregression.

- **Exponential Smoothing**: Shows weaker predictive power ($R^2 = 0.5273$) and higher error values, reflecting its inability to capture complex temporal dependencies. While it can adapt to short-term fluctuations, it struggles with capturing long-range dependencies, leading to reduced accuracy.

- **Linear Regression**: Performs the worst among all models, with the highest MSE (0.0533) and the lowest R^2 (0.3557). This indicates that simple linear trends are insufficient to capture the nonlinear and autocorrelated structure of the time series. Furthermore, the high RMSE and MAE reflect large deviations between predictions and observed values.

Overall, the findings suggest that ARIMA provides a more reliable baseline for time series forecasting in this context. Simpler models such as Moving Average and Exponential Smoothing offer ease of implementation but at the cost of reduced accuracy. Linear Regression, while straightforward, is not well-suited for time-dependent data due to its inability to incorporate autoregressive dynamics. These baseline results establish a strong foundation against which more advanced deep learning models can be evaluated. To further evaluate the distributional behavior of the forecasting performance metrics, violin plots with Kernel Density Estimation (KDE) were employed. As shown in Figure 4, these plots provide a combined view of the probability density and summary statistics (median and interquartile range) for each metric, including MSE, RMSE, MAE, MBE, r, R^2, RRMSE, NSE, and WI. This visualization enables a deeper understanding of the variability and skewness present in the error distributions, while also highlighting central tendencies across different evaluation criteria. Such graphical analysis is particularly useful for identifying stability and robustness of the models, beyond mean error comparisons.

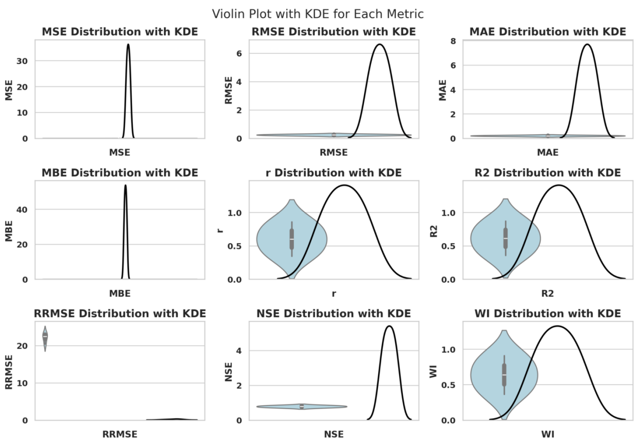

Figure 4. Violin plots with KDE representation for forecasting performance metrics.

To provide a multidimensional perspective on model performance, a radar chart was employed to simultaneously visualize multiple error and accuracy metrics. As shown in Figure 5, this representation allows for an intuitive comparison of ARIMA, Moving Average, Exponential Smoothing, and Linear Regression across MSE, RMSE, MAE, MBE, r, R^2, and RRMSE. Unlike tabular results, the radar chart highlights trade-offs between error reduction and correlation strength, offering a clear visual interpretation of which models achieve balanced performance and which exhibit weaknesses in specific metrics. This form of visualization is especially useful for identifying the dominance of ARIMA across most dimensions while revealing the relative limitations of simpler methods.

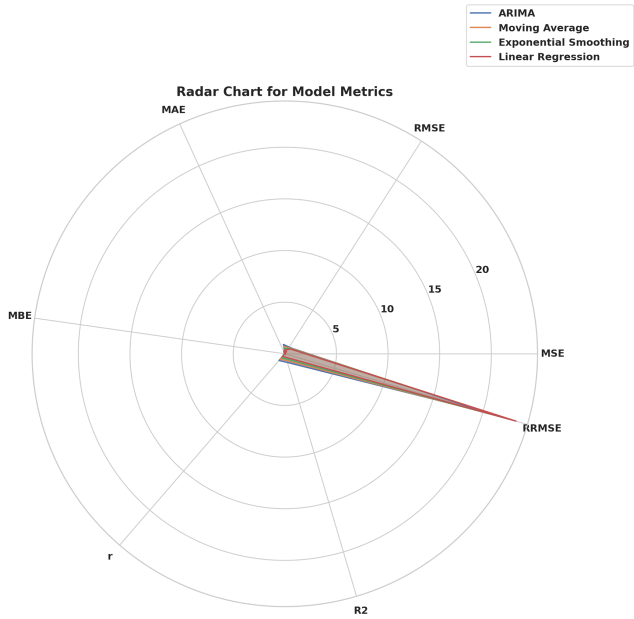

Figure 5. Radar chart comparing baseline models across multiple evaluation metrics.

To complement the tabular and radar chart comparisons, a parallel coordinates plot was constructed to visualize the relationships among all performance metrics simultaneously. As illustrated in Figure 6, this visualization enables the comparison of ARIMA, Moving Average, Exponential Smoothing, and Linear Regression across multiple evaluation criteria, including MSE, RMSE, MAE, MBE, r, R^2, RRMSE, NSE, and WI. The parallel coordinate representation allows for the identification of patterns across metrics, where ARIMA consistently occupies favorable positions with lower error measures and higher accuracy indicators, while simpler models such as Linear Regression exhibit weaker performance across most axes. This multi-metric perspective is valuable for detecting trade-offs and overlaps in model behavior that may not be evident from single-metric comparisons.

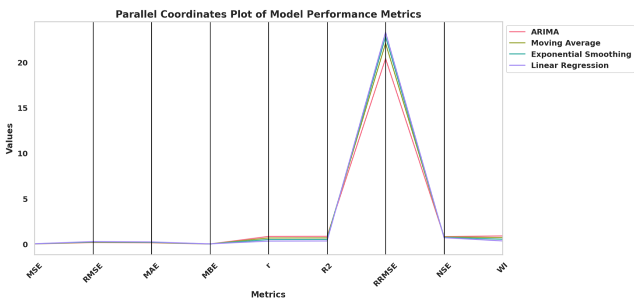

Figure 6. Parallel coordinates plot of baseline models across evaluation metrics.

the predictive performance of ARIMA is further enhanced by integrating different metaheuristic optimization algorithms, including the Fractional-Order Generalized Grey Optimization (FGGO), Genetic Algorithm (GA), Whale Optimization Algorithm (WOA), and Particle Swarm Optimization (PSO). These methods were applied to optimize ARIMA's parameters with the goal of improving forecasting accuracy and model stability. Table 4 presents the results across all optimization techniques. Metrics include MSE, RMSE, MAE, MBE, correlation coefficient (r), coefficient of determination (R^2), Relative RMSE (RRMSE), Nash–Sutcliffe Efficiency (NSE), and Willmott's Index (WI).

Table 4. Performance comparison of ARIMA optimized with metaheuristic algorithms

Models	MSE	RMSE	MAE	MBE	r	R^2	RRMSE	NSE	WI
FGGO + ARIMA	0.0037876	0.0207463	0.0178543	0.0025553	0.9286	0.9294	8.6412	0.920	1.0629
GA + ARIMA	0.0045451	0.0248956	0.0214252	0.0030663	0.9179	0.9187	10.1501	0.880	0.8503
WOA + ARIMA	0.0053026	0.0290448	0.0249960	0.0035774	0.9072	0.9080	11.0151	0.842	0.6377
PSO + ARIMA	0.0060602	0.0331941	0.0285669	0.0040885	0.8964	0.8972	12.5101	0.801	0.4251

The results highlight several key insights:

- **FGGO + ARIMA**: Achieves the best overall performance, with the lowest MSE (0.0037), RMSE (0.0207), and MAE (0.0179). The high correlation coefficient ($r = 0.9286$) and $R^2 = 0.9294$ confirm its superior predictive capability. The Willmott Index ($WI = 1.0629$) also indicates excellent model agreement with observed values, making FGGO the most effective optimization technique for ARIMA in this study.

- **GA + ARIMA**: Provides strong performance, with relatively low error values and high accuracy ($R^2 = 0.9187$). While not as precise as FGGO, it demonstrates that evolutionary approaches can effectively improve ARIMA parameter tuning.

- **WOA + ARIMA**: Shows moderate accuracy, with higher RMSE and MAE compared to FGGO and GA. Its $R^2 = 0.9080$ still reflects strong predictive power, but the reduced efficiency measures (NSE = 0.8421, WI = 0.6378) suggest more instability in capturing certain dynamics.

- **PSO + ARIMA**: Exhibits the weakest performance among the metaheuristic techniques, with the highest error metrics (MSE = 0.0061, RMSE = 0.0332, MAE = 0.0286). Although $R^2 = 0.8972$ still indicates a reasonable fit, the significantly lower WI (0.4252) highlights its limitations in balancing accuracy and generalization.

These findings underscore the value of integrating metaheuristic optimization into traditional time series models. The improvement in predictive performance, particularly with FGGO, highlights the potential of hybrid approaches to capture nonlinear patterns and reduce systematic errors. Beyond accuracy, optimization techniques also enhance *model interpretability*, as the optimized parameters offer clearer insights into the time series dynamics. This balance of precision and interpretability supports more robust decision-making in forecasting applications. To gain a clearer understanding of how different metaheuristic optimization methods affect ARIMA's predictive accuracy, a comparative analysis was conducted using the core error metrics: MSE, RMSE, and MAE. As presented in Figure 7, the bar chart not only contrasts the raw values of these metrics across FGGO, GA, WOA, and PSO optimized ARIMA models, but also incorporates the mean and standard deviation for each metric. This dual-level representation highlights both the central tendency and variability of the results, allowing for a robust evaluation of performance consistency. Notably, FGGO + ARIMA demonstrates the lowest error measures with minimal variability, while PSO + ARIMA shows relatively higher error values, confirming the superior stability and accuracy of FGGO in optimizing ARIMA.

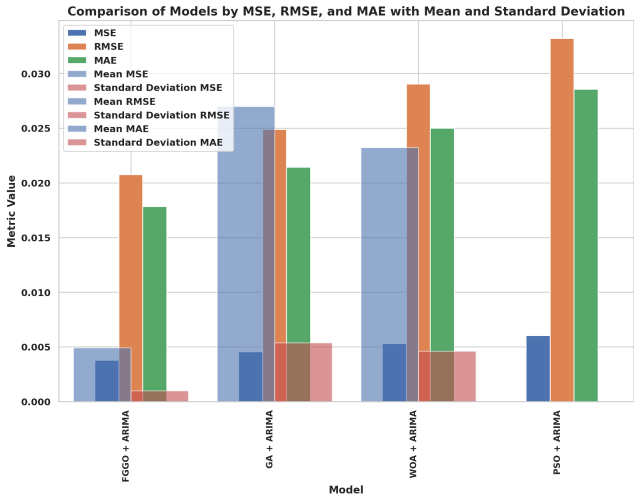

Figure 7. Comparison of optimization-enhanced ARIMA models using MSE, RMSE, and MAE with mean and standard deviation.

To examine the statistical distribution and variability of the forecasting performance metrics in greater detail, density plots combined with Kernel Density Estimation (KDE) were generated. As illustrated in Figure 8, these plots provide a smooth estimation of the probability density functions for each metric, including MSE, RMSE, MAE, MBE, r, R^2, RRMSE, NSE, and WI. This approach allows for the identification of central tendencies, dispersion, and skewness in the data. The alignment of the empirical density with the KDE curves indicates well-behaved distributions, suggesting that the optimized ARIMA models produce stable and consistent results across different evaluation measures. Such distributional insights strengthen the reliability of the observed performance improvements beyond point estimates alone.

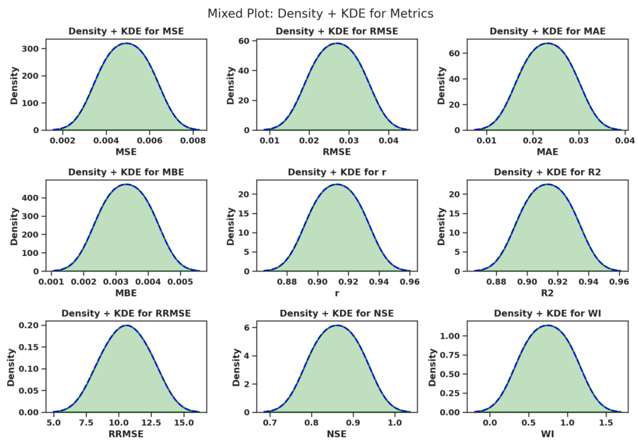

Figure 8. Density and KDE plots for model performance metrics.

To explore the similarity patterns among the different optimization-enhanced ARIMA models, hierarchical clustering was applied to their performance metrics. As illustrated in Figure 9, the resulting dendrogram provides a visual representation of how models group together based on their relative distances. The analysis shows that FGGO + ARIMA and GA + ARIMA form a closer cluster, reflecting their comparable predictive accuracy, while WOA + ARIMA and PSO + ARIMA are grouped separately, indicating relatively weaker performance. This clustering structure highlights the consistency of FGGO and GA in achieving lower error rates, while also reinforcing the distinct gap in performance exhibited by WOA and PSO.

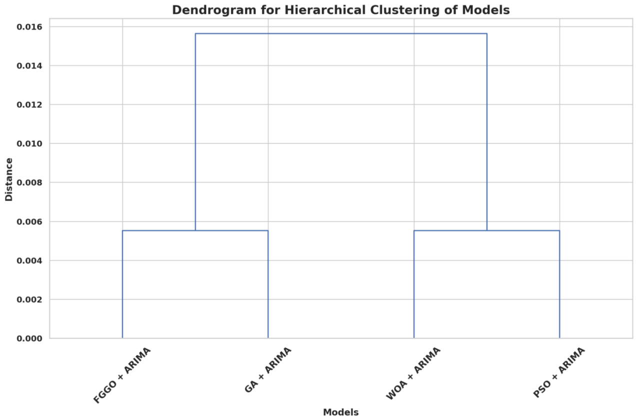

Figure 9. Hierarchical clustering dendrogram of optimization-enhanced ARIMA models.

6. DISCUSSION

The empirical findings of this study reveal critical insights into the evolving landscape of financial time series forecasting and the tangible benefits of integrating traditional statistical frameworks with adaptive metaheuristic optimization. The consistent superiority of the Fitness Greylag Goose Optimization (FGGO)–enhanced AutoRegressive Integrated Moving Average (ARIMA) model underscores the practical relevance of hybridization in addressing the long-standing limitations of classical models, such as sensitivity to noise, parameter rigidity, and local convergence issues. Whereas the traditional ARIMA model of estimation yielded a baseline level of 0.0333 in terms of Mean Squared Error (MSE), the FGGO-optimized version of the model minimized the error to 0.0038, which is almost an order of magnitude lower. This performance jump explains why adaptive optimization can reshape the parameter estimation of fixed parameters into a dynamic and self-tuned process that is more aligned with the dynamic and nonlinear nature of financial markets. In addition, the balanced exploration-exploitation mechanism of FGGO prevents premature convergence, a common issue in most traditional metaheuristics, such as Genetic Algorithm (GA) and Particle Swarm Optimization (PSO), thereby enabling the discovery of more globally optimal solutions and improving overall generalization capability. Methodologically, the FGGOARIMA framework applies bio-inspired optimization to surpass heuristic novelty, providing statistically meaningful forecasting advantages. Combining dynamic grouping and stagnation-avoidance processes in FGGO promotes the phenomenon of algorithmic resilience, which means that the diversity of search remains despite the stagnating error landscapes. The comparative performance analysis also confirms this advantage: as GA and Whale Optimization Algorithm (WOA) increased the model fit compared to the base, their convergence stability and precision were worse than those of FGGO. This observation highlights a valuable methodological principle: the efficacy of the search strategy does not merely reside in the search strategy itself, but also in how it adapts to changing loss surfaces. Increasing the computational effort of the exploration and exploitation stages by rapidly redistributing computational resources enables FGGO to match the optimization path to the structural properties of the underlying data, resulting in smoother convergence and increased predictability. These adaptive search control directions represent a significant development in the metaheuristic design of time series forecasting, particularly in situations where volatility is high and autocorrelation patterns are complex. The other important lesson is related to the interpretability and computational efficiency of models, both of which are fundamental in the current financial analytics. Although deep-learning models, including Long Short-Term Memory (LSTM) networks and Transformer models, have proven to be highly predictive, their complexity and the resulting considerable computational cost make them challenging to apply in real-world financial systems in practice. Conversely, the FGGOARIMA hybrid can generate competitive accuracy without compromising interpretability, which enables researchers and practitioners to track the model's decision-making to particular autoregressive and moving average elements. Moreover, the decreased error variance and narrow range of distribution of the evaluation metrics demonstrate that the FGGO-modified ARIMA not only improves the accuracy of the mean but also provides improved consistency of results across diverse performance dimensions. This consistency is critical in decision-support systems, where the dependability of the forecasts directly impacts risk management, investment timing, and portfolio optimization plans. The findings, therefore, confirm that the operational practicality of explainable models, based on the strategic use of classical and computational intelligence, can be achieved. Ultimately, the broader implications of this work extend to the future of hybrid forecasting systems and their potential applications beyond the financial sector. The demonstrated success of FGGO–ARIMA in refining parameter estimation and enhancing predictive stability suggests its potential adaptability across diverse time-dependent domains, such as energy demand forecasting, climate modeling, and industrial process optimization. The scalability of the proposed framework allows for seamless integration with emerging techniques in explainable artificial intelligence (XAI), multi-objective optimization, and online learning. These capabilities open avenues for constructing next-generation forecasting pipelines that are not only data-driven but also context-aware and self-evolving. Ultimately, this study contributes to the growing body of evidence that hybrid metaheuristic-statistical systems represent a pragmatic middle ground, bridging the gap between the interpretability of traditional econometric methods and the adaptability of modern AI-driven models. By embedding adaptive intelligence within classical statistical architecture, the FGGO–ARIMA paradigm enables the development of more resilient, transparent, and high-performance forecasting systems in complex, real-world environments.

7. CONCLUSION AND FUTURE WORK

This study introduced a hybrid forecasting framework that augments the statistical rigor of ARIMA with the adaptive search capabilities of the Fitness Greylag Goose Optimization (FGGO) algorithm, and benchmarked it against GA, WOA, and PSO optimizers. Using PT Bank Central Asia Tbk (BBCA.JK) daily data and a comprehensive suite of metrics (MSE, RMSE, MAE, MBE, r, R^2, RRMSE, NSE, WI), we showed that **FGGO + ARIMA** delivers the strongest overall performance, attaining the lowest errors (MSE $= 0.0038$, RMSE $= 0.0207$, MAE $= 0.0179$) and the highest accuracy indicators ($r \approx 0.9286$, $R^2 = 0.9294$) alongside an excellent agreement score (WI $= 1.0629$). Relative to classical baselines (ARIMA, MA, ES, LR) and other metaheuristic pairings, FGGO's dynamic exploration–exploitation balance and stagnation-avoidance logic consistently improved fit to the nonlinear, noisy, and nonstationary structure of financial time series, reaffirming that classical models can remain highly competitive when equipped with principled, modern optimization. Beyond tabular results, multi-perspective visual analytics (violin/KDE plots, radar charts, parallel coordinates, hierarchical clustering) corroborated the numerical findings by revealing lower variance, tighter dispersion, and more stable central tendencies for FGGO + ARIMA, while GA and WOA formed a secondary accuracy cluster and PSO trailed across most axes. These consistent cross-views support three substantive takeaways: (i) appropriate metaheuristic tuning yields material gains even for well-specified, interpretable linear models; (ii) robustness—not merely point accuracy—improves when optimization explicitly modulates search allocation over training; and (iii) hybridization narrows the gap with heavier deep-learning pipelines without sacrificing transparency or computational efficiency. Collectively, these results position FGGO-augmented statistical modeling as a compelling, resource-conscious alternative for high-frequency or real-time financial prediction contexts where stability and interpretability are at a premium. Future research can extend the proposed framework along several directions. First, integrate FGGO with deep sequence learners (LSTM/GRU/Transformers) for hyperparameter tuning and weight initialization, and consider *multi-objective* FGGO variants to jointly optimize accuracy, stability, and runtime. Second, operationalize *online/streaming* learning for real-time markets, incorporate exogenous drivers (macro indicators, technical/sentiment signals) in multivariate settings, and validate generalization across assets (equities, crypto, commodities, fixed income) and regions. Third, enhance *explainability* via SHAP/LIME to attribute gains to optimized components, and explore *ensemble* schemes (stacking/weighted blends) that combine FGGO-optimized ARIMA with complementary models to hedge regime shifts. Pursuing these avenues will deepen the synergy between statistical structure and computational intelligence, advancing hybrid, interpretable, and deployable forecasting systems for modern financial analytics.

DATA AVAILABILITY

The dataset used in this study can be found in `https://www.kaggle.com/datasets/willianoliveiragibin/financial-data/data`

REFERENCES

[1] M. A. Al-Khasawneh, A. Raza, S. U. R. Khan, and Z. Khan, "Stock market trend prediction using deep learning approach," *Computational Economics*, 2024.

[2] M. Steinbacher, M. Steinbacher, and M. Steinbacher, "Using cnn to model stock prices," *Computational Economics*, pp. 1–42, 2025.

[3] V. Baladari, A. R. Nagubandi, S. R. C. Charan Teja Tadi, A. T. Selvi, V. Sreedevi, and M. Nithya, "Predictive ai model for financial risk assessment in dynamic market environments," in *2026 International Conference on Communication, Computing and Emerging Technologies (IC3ET)*, 2026, pp. 748–753.

[4] I. Goldstein, "Information in financial markets and its real effects*," *Review of Finance*, vol. 27, no. 1, pp. 1–32, 2023.

[5] A. Koukorinis, G. W. Peters, and G. Germano, "Generative-discriminative machine learning models for high-frequency financial regime classification," *Methodology and Computing in Applied Probability*, vol. 27, no. 2, p. 36, 2025.

[6] T. A. Shaikh, T. Rasool, and F. R. Lone, "Towards leveraging the role of machine learning and artificial intelligence in precision agriculture and smart farming," *Computers and Electronics in Agriculture*, vol. 198, p. 107119, 2022.

[7] I. Ahmed, G. Jeon, and F. Piccialli, "From artificial intelligence to explainable artificial intelligence in industry 4.0: A survey on what, how, and where," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 8, pp. 5031–5042, 2022.

[8] A. K. Bashir, S. Khan, B. Prabadevi, N. Deepa, W. S. Alnumay, T. R. Gadekallu, and P. K. R. Maddikunta, "Comparative analysis of machine learning algorithms for prediction of smart grid stability," *International Transactions on Electrical Energy Systems*, vol. 31, no. 9, p. e12706, 2021.

[9] J. Benavides-Hernández and F. Dumeignil, "From characterization to discovery: Artificial intelligence, machine learning and high-throughput experiments for heterogeneous catalyst design," *ACS Catalysis*, vol. 14, no. 15, pp. 11 749–11 779, 2024.

[10] P. Weber, K. V. Carl, and O. Hinz, "Applications of explainable artificial intelligence in finance—a systematic review of finance, information systems, and computer science literature," *Management Review Quarterly*, vol. 74, no. 2, pp. 867–907, 2024.

[11] M. Hassan and I. Hassan, "Improving artificial neural network based streamflow forecasting models through data preprocessing," *KSCE Journal of Civil Engineering*, vol. 25, no. 9, pp. 3583–3595, 2021.

[12] A. L. Schaffer, T. A. Dobbins, and S. Pearson, "Interrupted time series analysis using autoregressive integrated moving average (arima) models: A guide for

evaluating large-scale health interventions," *BMC Medical Research Methodology*, vol. 21, no. 1, p. 58, 2021.

[13] J. Lascar, J. Bobin, and F. Acero, "Sushi: An algorithm for source separation of hyperspectral images with non-stationary spectral variation - semi-blind unmixing with sparsity for hyperspectral images," *Astronomy & Astrophysics*, vol. 686, p. A259, 2024.

[14] M. A. El-Shorbagy, A. Bouaouda, H. A. Nabwey, L. Abualigah, and F. A. Hashim, "Advances in henry gas solubility optimization: A physics-inspired metaheuristic algorithm with its variants and applications," *IEEE Access*, vol. 12, pp. 26 062–26 095, 2024.

[15] F. E. Sapnken, A. K. Tazehkandgheshlagh, B. S. Diboma, M. Hamaidi, P. G. Noumo, Y. Wang, and J. G. Tamba, "A whale optimization algorithm-based multivariate exponential smoothing grey-holt model for electricity price forecasting," *Expert Systems with Applications*, vol. 255, p. 124663, 2024.

[16] E. M. El-kenawy, N. Khodadadi, S. Mirjalili, A. A. Abdelhamid, M. M. Eid, and A. Ibrahim, "Greylag goose optimization: Nature-inspired optimization algorithm," *Expert Systems with Applications*, vol. 238, p. 122147, 2024.

[17] S. Bi and Y. Lian, "Advanced portfolio management in finance using deep learning and artificial intelligence techniques: Enhancing investment strategies through machine learning models," *Journal of Artificial Intelligence Research*, vol. 4, no. 1, pp. 233–298, Jan. 2024.

[18] P. Craja, A. Kim, and S. Lessmann, "Deep learning for detecting financial statement fraud," *Decision Support Systems*, vol. 139, p. 113421, 2020.

[19] C. D. Lucia, P. Pazienza, and M. Bartlett, "Does good esg lead to better financial performances by firms? machine learning and logistic regression models of public enterprises in europe," *Sustainability*, vol. 12, no. 13, 2020.

[20] Q. Feng, H. Chen, and R. Jiang, "Analysis of early warning of corporate financial risk via deep learning artificial neural network," *Microprocessors and Microsystems*, vol. 87, p. 104387, 2021.

[21] R. Feng and X. Qu, "Analyzing the internet financial market risk management using data mining and deep learning methods," *Journal of Enterprise Information Management*, vol. 35, no. 4/5, pp. 1129–1147, 2021.

[22] A. Gicić, D. Đonko, and A. Subasi, "Intelligent credit scoring using deep learning methods," *Concurrency and Computation: Practice and Experience*, vol. 35, no. 9, p. e7637, 2023.

[23] J. W. Goodell, S. Kumar, W. Lim, and D. Pattnaik, "Artificial intelligence and machine learning in finance: Identifying foundations, themes, and research clusters from bibliometric analysis," *Journal of Behavioral and Experimental Finance*, vol. 32, p. 100577, 2021.

[24] P. Jaquart, D. Dann, and C. Weinhardt, "Short-term bitcoin market prediction via machine learning," *The Journal of Finance and Data Science*, vol. 7, pp. 45–66, 2021.

[25] S. Lin and H. Huang, "Improving deep learning for forecasting accuracy in financial data," *Discrete Dynamics in Nature and Society*, vol. 2020, no. 1, p. 5803407, 2020.

[26] Y. Zhang, B. Yan, and A. M., "A novel deep learning framework: Prediction and analysis of financial time series using ceemd and lstm," *Expert Systems with Applications*, vol. 159, p. 113609, 2020.