A Method for Assessing Financial Market Price Behavior: An Analysis of the Shanghai Stock Exchange Index

Zhi Huang[1], Jiansheng Li[2]*
Faculty of Economics and Management, Science and Technology College of Nanchang Hangkong University, Jiujiang 332020, Jiangxi, China[1]
School of Business Administration, Dongbei University of Finance and Economics, Dalian 116025, Liaoning, China[2]

***Abstract*—A stock market is a venue where the shares of publicly traded companies are available for purchase and sale by individuals. The financial markets exert a substantial influence on various domains, including technology, employment, and business. Given the substantial rewards and risks associated with stock trading, investors are exceedingly concerned with the precision of future stock value forecasts. They modify their investment strategies in an effort to achieve even greater returns. Accurate stock price forecasting can be challenging in the securities industry due to the complex nature of the problem and the requirement for a comprehensive understanding of various interconnected factors. The stock market is influenced by a variety of factors, including politics, society, and economics. A multitude of interrelated factors contribute to these behaviors, and stock price fluctuations are capricious. In order to tackle a range of these difficulties, the present investigation proposes an innovative framework that integrates a Grasshopper optimization method with the gated recurrent unit model, a machine-learning approach. The research used data from the Shang Hai Stock Exchange Index for the period of 2015–2023. The proposed hybrid model was also tested on the 2013–2022 S&P 500 and Nikkei 225. The proposed model demonstrated optimal performance, exhibiting a minimal error rate and exceptional effectiveness. The study's findings demonstrate that the proposed model is more suitable for the volatile stock market and surpasses other existing strategies in terms of predictive accuracy.**

Keywords—Financial market; shanghai stock exchange price; gated recurrent unit; grasshopper optimization algorithm

I. INTRODUCTION

One of the most fascinating technological developments of the day is the financial markets [1]. It provides market analysts, investors, and researchers from other fields with various chances [2]. Individuals may have different viewpoints on market involvement, such as understanding market behavior, identifying important elements, trading stocks, and forecasting future events [3]. The market trend, suggesting assets for portfolio management, etc., but a lack of understanding of basic economic concepts and financial literacy may have a significant impact on the returns on investments [4]. Stock forecasting is a complex task requiring a thorough understanding of many interconnected factors [5]. Nevertheless, several factors, such as political, and economic dynamics, impact the stock market [6], [7]. A broad variety of factors, such as changes in the unemployment rate, immigration regulations, public health issues, immigration policies, and monetary policies impacting various nations, might be contributing elements [8]. As a result of a careful examination of the market, everyone involved in the stock market wants to maximize earnings and reduce risks [9], [10]. Consequently, there is an increased demand for market valuations and various forms of analytical assessments to examine market behavior [11], [12]. Fundamental analysis and technical analysis are the two distinct categories into which contemporary approaches to financial forecasting fall [13]. To generate long-term forecasts, fundamental analysis entails the examination of prevalent stock market elements based on knowledge and expertise. As opposed to this, a technical analysis integrates insights obtained from past stock price information [14]. Technical analysis is the systematic examination of past pricing data in conjunction with the application of technical indicators to predict forthcoming trends in financial time series [15], [16]. Conventional methods may improve forecasting precision, but they also add to computing complexity, increasing the risk of prediction mistakes [17]. To effectively use artificial intelligence technology for financial market forecasting, a strong and uncomplicated model is necessary for profitable development [18]. Choosing a suitable methodology is crucial as it relies on the characteristics of the dataset and the desired application. Researchers may encounter datasets that vary with time (time-dependent) or datasets that do not vary with time (time-independent). Each kind of dataset presents distinct issues [19]. The main reason for this is that time series analysis is characterized by consistent price fluctuations occurring at regular periods [20]. Investors must compile voluminous amounts of information and analyze deterministic, non-linear, and non-parametric chaotic systems in order to construct an exact model that can forecast future returns. Due to their nonlinear dynamics, determinism, and absence of well-defined parameters, these systems are unique [21]. It is important to note that these models may sometimes encounter the problem of being trapped in a local minimum. A proposed remedy for this problem is the gated recurrent unit (GRU) model [22]. Using the GRU method, a sophisticated machine-learning model was developed to forecast currency exchange rates. The development of a stock index movement prediction algorithm has been enabled by combining a new technique, the improved online sequential gated recurrent unit, with the grasshopper optimizer algorithm. These strategies use probabilistic concepts that are better suited for sets of responses rather than individual ones.

These algorithms use the principles of natural selection to imitate the most efficient behaviors seen in the natural realm. Slime mold algorithm (SMA)[23], Moth-flame optimizer (MFO) [24], and Grasshopper optimization algorithm (GOA) [25], which is a new and intriguing swarm intelligence system that emulates the natural swarming and foraging habits of grasshoppers. Grasshoppers are a well-known class of insects that pose a threat to agriculture and agricultural production. The two stages of its life cycle are referred to as nymph and maturity. The adult phase is marked by long-range, sharp movements, whereas the nymph phase is characterized by short steps and gradual motions. Given the ever-changing and consequential character of financial markets, specifically the stock market, the significance of precise and dependable stock price prediction is emphasized. To maximize returns and optimize investment strategies, investors are perpetually in search of innovative predictive models. The advent of the GOA-GRU model signifies a critical juncture in the realm of stock market prediction, presenting investors with an exceptional prospect to improve their investment tactics in the face of market instability. By merging Grasshopper optimization and the gated recurrent unit model, this novel approach not only could guarantee enhanced predictive precision but also offer significant insights into the intricate relationship between micro-level market dynamics and macroeconomic factors. The provision of dependable forecasts by the GOA-GRU model not only facilitates technological advancements but also enhances comprehension of market dynamics, thereby it can promote more effective capital allocation and the ongoing development of financial modeling methodologies. The main contributions of the study are as follows:

- The grasshopper optimization algorithm and the gated recurrent unit model have been integrated in a manner that has substantially enhanced predictive accuracy. By capitalizing on this novel framework, investors are able to enhance their decision-making process regarding stock trading, thereby optimizing returns while mitigating risks.
- By subjecting alternative models, including GOA, SMA-GRU, and MFO-GRU, to rigorous evaluation, the GOA-GRU model consistently demonstrated superior performance. The model's superior predictive capabilities are demonstrated by its ability to attain high efficiency and low error.
- The GOA-GRU model provides a pragmatic resolution to the complexities associated with predicting stock prices, as evidenced by its consistent integrity and accuracy. The utilization of this technology has the potential to enable financial analysts, investors, and institutions to make decisions based on data, thereby enhancing the efficacy and knowledge of capital allocation within the financial markets.

The subsequent content of this paper is organized as follows: The literature review is provided in Section II. In addition to the materials and methodology utilized, Section III provides a concise overview of the optimizer techniques and GRU algorithm. The results and discussion are provided in Section IV. At last, the conclusions that have been drawn from the assessments and findings of the review are presented in Section V.

II. Literature Review

A. Related Works

Over the course of the past few decades, there has been a significant amount of potential for the application of machine learning algorithms to the prediction of the future stock market price. In an effort to improve the precision of trend prediction in the context of stock market fluctuations, Nabipour et al. [26] undertook an investigation that utilized deep learning and machine learning algorithms. They conducted a comparative analysis of the performance of different prediction models with respect to four distinct stock market groups that are listed on the Tehran Stock Exchange: diversified financials, petroleum, non-metallic minerals, and basic metals [26]. The outcomes demonstrated that the Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) exhibited superior performance compared to alternative prediction models when applied to continuous data. This underscores the efficacy of these models in capturing intricate temporal dependencies present in the data [26]. In their research, Khan et al. [27] examined the impact of political events and public sentiment on stock market trends, encompassing both the performance of specific companies and the broader market environment [27]. Their objective was to determine whether political situations and public sentiment on a particular day could influence seven-day stock market trends. In pursuit of this objective, a machine learning model was enhanced with sentiment and political situation features in order to examine their impact on the accuracy of predictions [27]. The experimental results indicated that the incorporation of sentiment features resulted in a slight enhancement of 0–3% in the accuracy of predictions. However, the inclusion of the political situation features substantially improved the accuracy of predictions by around 20% [27].

Yuan et al. [28] present an alternative approach to the traditional linear multi-factor stock selection model, which takes into account the dynamic and chaotic characteristics of the stock market. They conducted a thorough feature selection process utilizing a variety of feature selection algorithms in their research. They further refine the parameters of stock price trend prediction models based on machine learning using time-sliding window cross-validation [28]. They utilized a comprehensive eight-year dataset pertaining to the Chinese A-share market in order to ascertain the most efficient integrated models for forecasting stock price trends [28]. By conducting an extensive examination and assessment of various integrated models, their research demonstrates that the random forest algorithm exhibits exceptional efficacy in predicting stock price trends and selecting features [28]. Moghar and Hamiche endeavor to improve the accuracy of inventory value forecasts by leveraging the capabilities of RNNs, with a particular emphasis on LSTM [29]. In their study, Vijh et al. [30] utilized Random Forest and Artificial Neural Networks (ANN) to predict the closing prices of five companies operating in various sectors. They employed financial data that included the opening, closing, high, and low prices of stocks in order to generate novel variables that function as inputs for the predictive models [30]. By employing ANN and Random Forest techniques, they aimed to predict the closing prices of equities on the following business day. The evaluation of the model's effectiveness is conducted by examining the performance of these metrics; lower values signify increased predictive accuracy [30].

In their investigation, Parray et al. [31] examined the feasibility of utilizing three machine learning algorithms—Support Vector Machine (SVM), Perceptron, and Logistic Regression—to predict the trajectory of stock prices for the following day [31]. The experiments are conducted by the researchers using historical stock data from January 1, 2013 to December 31, 2018. The dataset consists of around fifty stocks selected from the NIFTY 50 index of the Indian National Stock Exchange. In addition to calculated technical indicators, the data is utilized for the analysis. The findings suggest that the SVM model attains an average accuracy of 87.35% in its predictions, with Logistic Regression following closely at 86.98% and Perceptron at 75.88% [31]. To predict the closing price of the S&P 500 index the following day, Bhandari et al. [32] employ LSTM, a specialized neural network architecture. A comprehensive analysis of the stock market's behavior is achieved through the formulation of a well-curated ensemble of nine predictors, which includes technical metrics, macroeconomic indicators, and fundamental market data [32]. Moving forward, the chosen input variables are employed to construct both single-layer and multilayer LSTM models, which are subsequently assessed using well-established evaluation metrics [32]. By combining machine learning and deep learning methodologies, Mehtab et al. [33] developed a hybrid modeling strategy for predicting stock prices. The data for this analysis is derived from the NIFTY 50 index values published by the National Stock Exchange (NSE) of India [33]. The period covered by this data is from December 29, 2014, to July 31, 2020. To forecast the open values of the NIFTY 50 index from December 31, 2018, to July 31, 2020, eight regression models are developed utilizing training data spanning from December 29, 2014, to December 28, 2018 [33]. Additionally, four deep learning-based regression models utilizing LSTM networks are implemented to enhance the predictive capability of the framework [33]. Liu and Long introduced a framework for predicting stock closing prices by capitalizing on the capabilities of deep learning, specifically the LSTM network, which excels at processing intricate financial time series [34]. In contrast to conventional models, their framework utilized empirical wavelet transform (EWT) for data preprocessing and an outlier-robust extreme learning machine (ORELM) model for post-processing. The primary constituent, an LSTM network-based deep learning predictor, was optimized by employing the dropout technique and the particle swarm optimization (PSO) algorithm [34]. Combining machine learning techniques with technical analysis indicators, Ayala et al. [35] proposed a hybrid method for generating trading signals in stock market prediction. The simplicity and effectiveness of their approach, which combines machine learning with a technical indicator to inform trading decisions, might be applicable to additional technical indicators in the future [35]. In order to determine the most appropriate machine learning technique, they assessed the performance of Four Neural Networks, a Linear Model, Support Vector Regression, and a Random Forest. As technical trading strategies, they evaluated their approach using daily trading data from major indices such as the DAX and Dow Jones Industrial Average in conjunction with the Triple Exponential Moving Average and Moving Average Convergence/Divergence [35].

B. Challenges and Fulfillment

The exploration of integrating optimization methods with machine learning models is a notable gap in current research on stock market prediction. The proposed framework addresses this deficiency by integrating the GRU model with the SMA, MFO, and GOA, enabling a more refined examination of interrelated variables. Concerns regarding the representativeness and quality of the findings arise due to the utilization of obsolete or irrelevant datasets, which constitutes another deficiency. We ensure the pertinence and contemporaneity of this research findings by addressing this gap with recent data from the Shang Hai Stock Exchange Index spanning the years 2015 to 2023 along with S&P 500 and Nikkei 225. In the realm of stock market prediction research, a divide exists between conceptual progress and tangible implementation. By exhibiting its practical applicability and real-world effectiveness, our demonstrated superior performance—distinguishable from other models—not only verifies that our proposed model is appropriate for volatile markets but also bridges this gap.

III. Method and Materials

A. Slime Mold Algorithm

In 2020, Li et al. introduced SMA, an innovative methodology that was inspired by the natural slime mold activity [23]. The slime mold uses olfaction to perceive and discern the volatile food aromas present in the atmosphere, enabling it to effectively travel toward its prey. The behavior of the slime mold may be formally characterized by the following equation:

$$\overrightarrow{X(t+1)} = \begin{cases} \overrightarrow{X_b(t)} + \overrightarrow{v_b} \cdot \left(\overrightarrow{W} \cdot \overrightarrow{X_A(t)} - \overrightarrow{X_B(t)}\right) & r < p \\ \overrightarrow{v_c} \cdot \overrightarrow{X(t)} & r \geq p \end{cases} \quad (1)$$

The variable $X_b(t)$ reflects the precise region of the slime mold that now displays the greatest concentration of odor. The variables $X(t)$ and $\mathrm{X(t+1)}$ represents the locations of the slime mold in the t-th and $t+1$-th iterations, respectively. $X_A(t)$ and X_B represent two arbitrarily chosen locations of the slime mold. The variable v_b experiences temporal fluctuations within the interval$[-\mathrm{a}, \mathrm{a}]$, where r is a random integer ranging from 0 to 1. The parameter p is defined as the inverse hyperbolic tangent of the negative ratio of t to the maximum value of $= arctanh(-(\frac{t}{max\_t}) + 1)$. The parameter v_c is a linearly decreasing parameter that varies between 0 and 1.

$$p = tanh\,|S(i) - DF| \quad i = 1, 2, \dots, n \quad (2)$$

The symbol DF denotes the iteration with the greatest fitness value, whereas $S(i)$ denotes the fitness of the vector $\overrightarrow{X}$.The equation provided below offers a precise and formal definition of the weight, represented by the symbol W :

$$\overrightarrow{W(smell\ index(l))} = \begin{cases} 1 + r \cdot log\left(\frac{bF - S(i)}{bF - wF} + 1\right), condition \\ 1 - r \cdot log\left(\frac{bF - S(i)}{bF - wF} + 1\right), others \end{cases} \quad (3)$$

$$smell\ index = sort(S) \quad (4)$$

The variable $S(i)$ denotes the first half of the population in the provided equation. The sign bF indicates the maximum fitness value, whereas wF shows the minimum fitness value. Furthermore, the scent index pertains to the arranged values of physical fitness. The spatial coordinates of the slime mold are updated by using the provided formula.

$$\overrightarrow{X^*} = \begin{cases} rand(UB - LB) + LB\ , \quad rand < z \\ \overrightarrow{X_b(t)} + \overrightarrow{v_b}.\left(\overrightarrow{W}.\overrightarrow{X_A(t)} - \overrightarrow{X_B(t)}\right), r < p \\ \overrightarrow{v_c}.\overrightarrow{X(t)}\ ,\ r \geq p \end{cases} \tag{5}$$

Within this particular context, the variable represented by the symbol Z is limited to a numerical interval spanning from 0 to 0.1. The words LB and UB denote the bottom and upper boundaries of the search interval, respectively.

B. Moth-Flame Optimizer

The Moth Flame Optimizer is a clever device that has been shown to significantly increase the performance of many models. The concept for it comes from the way that nocturnal butterflies respond to a source of light at night. When these insects fly toward the moon, they have been shown to be able to navigate over very long distances with success. They might quickly get entangled, however, if they continue to circle the light. Having been well studied, this particular movement may be used as an incredibly effective optimizer in many fields, such as medical applications, business management, image processing, architectural design, electrical and energy systems, and design. Classified as a metaheuristic technique, the MFO algorithm has considerable potential in solving a variety of optimization issues [24]. The challenging components of the inquiry include the geographical distributions of moths, which are viable solutions. Moths may fly in one, two, three, or hyper-dimensional space by changing their position vectors. The suggested method ensures convergence, and MFO has a high computational efficiency and reliability. The following is one way to express the MFO:

$$MFO = (R, M, P) \tag{6}$$

The following equation illustrates how the flames control the rearranging of the moth locations in the movement function M :

$$A_i = s(A_i, B_J) = C_i.e^{br}.\cos(2\pi t) + B_J \tag{7}$$

The variables S, B_j , and A_i are used to represent the spiral function. The moth's index is I, while the flame's index is j. c_i is the symbol for the distance between the i-th moth and the j-th flame. Using the constant symbol b , the geometric characteristics of the spiral are determined. Randomly produced values between -1 and 1 reflect the number denoted by the variable r. To calculate the distance c_i , use the formula below:

$$C_i = |B_j - A_i| \tag{8}$$

Preservation of the exploration phase of the search space is achieved through the implementation of an adaptive technique designed to minimize the frequency of flames. The subsequent procedures are executed to accomplish this:

$$N = round\left(N_{MAX} - \alpha \times \frac{N_{MAX-1}}{\xi}\right) \tag{9}$$

The constants N , N_{MAX} , ξ , and α reflect the number of flames, total number of flames, iterations, and iterations in process, respectively.

C. Grasshopper Optimization Algorithm

Originating from natural processes, the grasshopper optimization technique is a well-known metaheuristic algorithm [36]. The primary objective is to identify optimal solutions that provide the maximum outcome by using randomization to prevent being trapped in suboptimal alternatives. The algorithm's rapid convergence and exceptional exploration abilities have shown its amazing success and efficiency in optimization. GOA has outperformed many other approaches in test scenarios, demonstrating its excellence and promise in practical applications. As it is perceivable from the illustration and framework in Fig. 1 and Fig. 2. Moreover, GOA is adaptable, effectively managing the trade-off between exploring new possibilities and using known solutions to ensure optimal results are achieved. GOA is an excellent choice for research applications because of its unique attributes. Saremi et al. [36] introduced the GOA algorithm, which falls under the category of swarm intelligence algorithms. Each grasshopper's placement in the swarm represents a possible solution, mimicking the behavior of grasshoppers that often gather in swarms.

$$X_i = S_i + G_i + A_i \tag{10}$$

s_i represents social interaction, G_i represents gravitational force, and A_i represents wind advection.

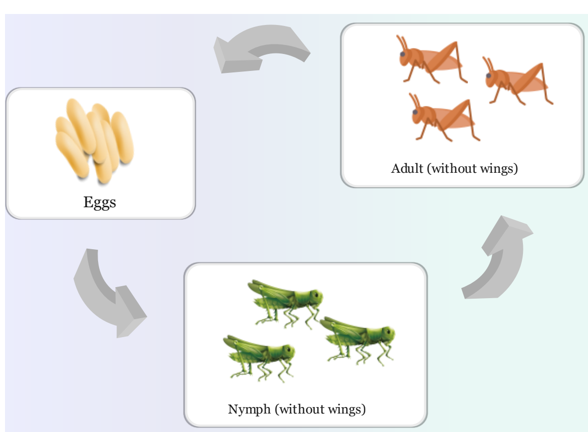

Fig. 1. The illustration of (GOA).

The equation for N grasshopper optimization is expressed as follows, excluding the gravity component and assuming that the wind direction is directed towards the goal [36].

$$X_i^d = c\left(\sum_{\substack{j=1 \\ j\neq i}}^{N} \frac{ub_d - lb_d}{2} s\left(\left|x_j^d - x_i^d\right|\right)\frac{x_j - x_i}{d_{ij}}\right) + \widehat{T_d} \tag{11}$$

The distance between the i-th and j-th grasshoppers is indicated by d_{ij}. Whereas I stands for the beauty scale and f for the power of attraction, function S reflects the strength of social forces. The formulae below are used to compute these values [36]:

$$d_{ij} = \left|d_j - d_i\right|$$

$$s(r) = fe^{\frac{-r}{T}} - e^{-r} \tag{12}$$

The equation is used to calculate the coefficient c, which decreases the comfort zone in proportion to the number of iterations [36].

$$c = c_{max} - l\frac{c_{max} - c_{min}}{L} \tag{13}$$

l represents the current iteration, c_{max} signifies the maximum value, c_{min} the minimum value, and L the iterations' highest number[36].

D. Gated Recurrent Unit

To ensure network correctness, the GRU network is derived by reducing the complicated gate structure of the LSTM. This results in a reduction in the number of network training parameters and an increase in computing efficiency [22]. Within the GRU, the primary function of the gate unit r_t is to regulate the merging of correlation between the current input of the network and the memory from the past instant. Meanwhile, the update gate z_t governs the extent to which memory information is preserved from the historical moment [22]. Fig. 3 illustrates the interior arrangement.

The update gate z in the GRU model is computed at time t using the following formula.

$$z_t = \sigma(W^z x_t + U^z h_{t-1}) \tag{14}$$

In the above formula, σ denotes the sigmoid function, whereas W^z and U^zshow the updated coefficients for the gate weights. The expression for the reset gate r can be formulated as:

$$r_t = \sigma(W^r x_t + U^r h_{t-1}) \tag{15}$$

The cellular memory at time t , denoted as $\widetilde{h_t}$, may be mathematically represented as:

$$\widetilde{h_t} = \tanh(r \times Uh_{t_1} + Wx_t) \tag{16}$$

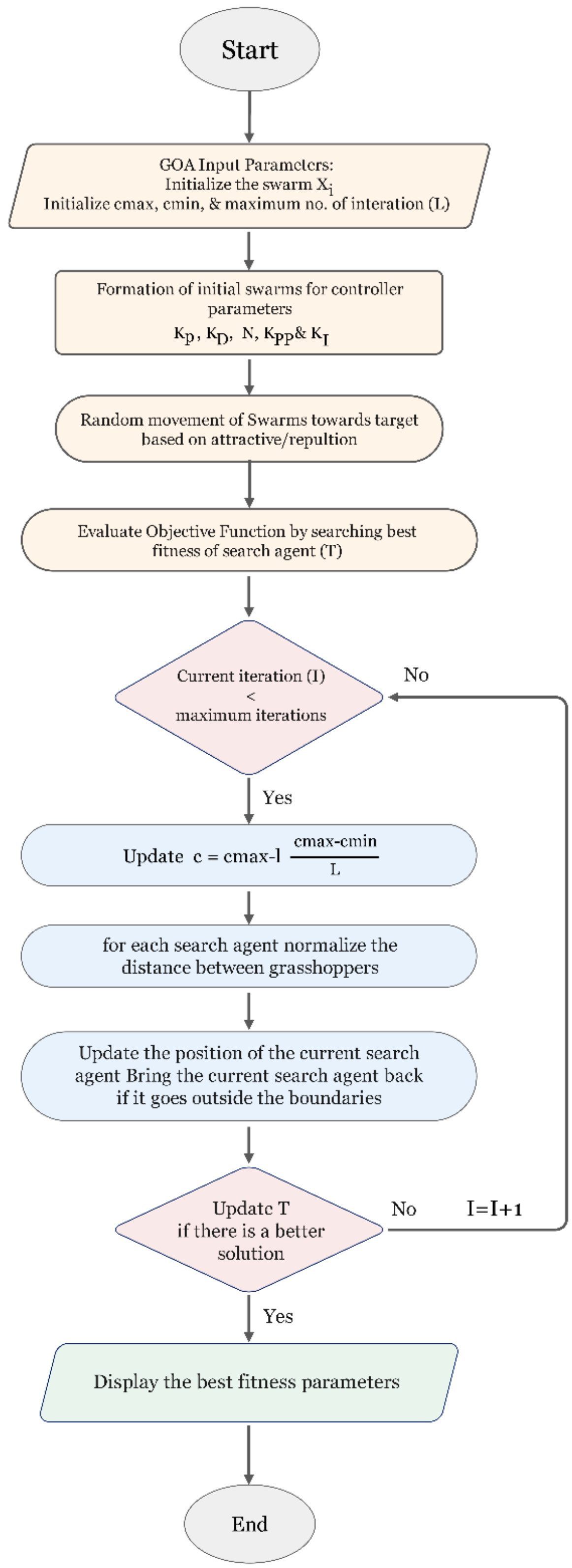

Fig. 2. The framework of (GOA).

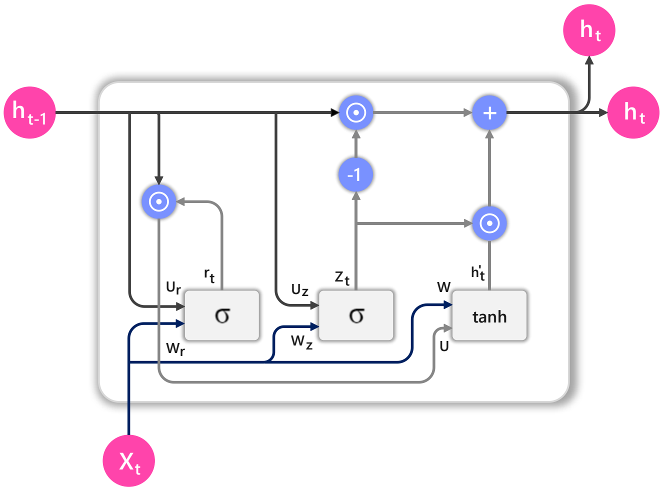

Fig. 3. The illustration of the GRU.

E. Data Collection and Preprocessing

The candlestick chart was devised by Homma, a renowned rice merchant from Sakata City, Japan [37]. Subsequently, several Japanese merchants used it as a means of forecasting forward pricing in rice futures contracts [38]. A candlestick chart is a hybrid chart that combines the features of a line chart and a bar chart. The price changes over a certain time period are frequently shown using it. The three components of a candlestick chart are an authentic body, a lower shadow, and an upper shadow. Each bar on the candlestick chart represents the opening, closing, lowest, and highest prices for a single trading day. The different beginning and closing price gaps are shown in the candlestick chart's body. The many colors of the candlestick also represent different meanings. Red will actually be the color of the body if the opening price is higher than the closing price. As an alternative, the entire body will be heavily green-hued. Near the conclusion of the real body, the higher and lower lines represent the shadows in the upper and lower regions, respectively. Candlestick charts' upper and lower shadows, respectively, display the highest and lowest price ranges over a certain period of time.

A thorough examination of the data was an essential component of the preliminary phase to identify any irregularities, uncommon observations, or inconsistencies that may undermine the credibility of the results. To optimize the performance of the models, two distinct preprocessed data sets were generated. When conducting a thorough analysis, it is important to take into account many elements, such as the trading volume and the OHLC prices over a certain timeframe. The data on the performance of the Shang Hai stock market index from 2015 to 2023 was obtained for this research. The research used a partitioning strategy in which 80% of the dataset was assigned for training. In contrast, the remaining 20% was allocated to conduct tests. The primary goal of this division was to achieve a harmonious equilibrium between the need for a substantial volume of data to train the model and the necessity for a substantial and novel dataset to carry out thorough testing and validation, as seen in Fig. 4. Furthermore, in order to validate the performance of the hybrid model under consideration, data from the Nikkei 225 and the S&P 500 spanning the years 2013 to 2022 were utilized.

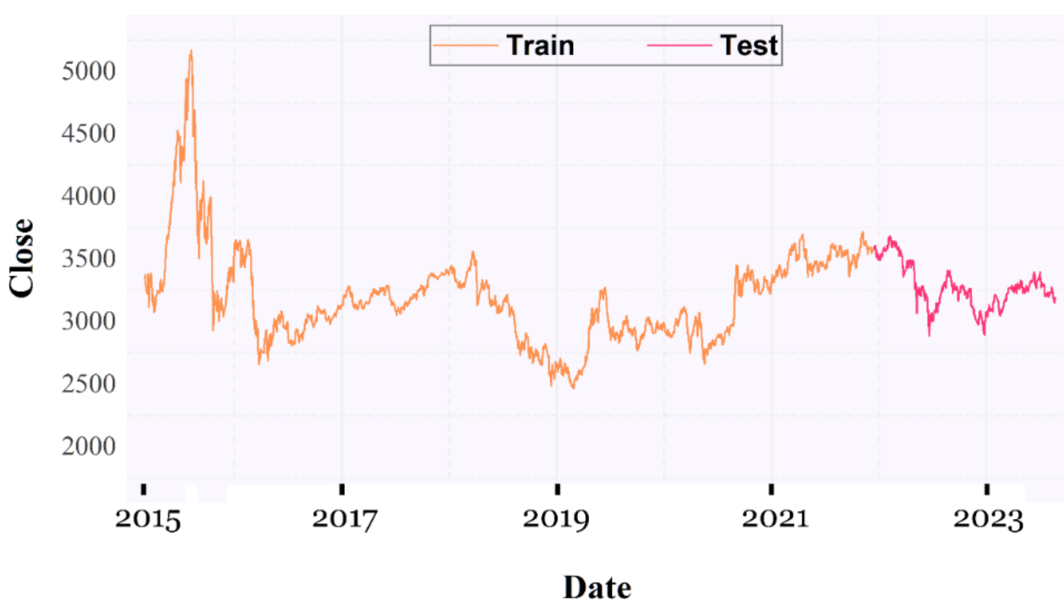

Fig. 4. The process of separating a provided dataset into separate test and train sets.

F. Evaluation Metrics

To evaluate the accuracy of the next projection, many performance criteria were used. Within the field of statistical analysis, four frequently used evaluation criteria are applied to measure the accuracy and effectiveness of a model. The coefficient of determination (R^2), mean absolute error (MAE), mean absolute percentage error (MAPE), and root mean square error (RMSE).

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{17}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \tag{18}$$

$$MAE = \frac{\sum_{i=1}^{n}|y_i - \hat{y}_i|}{n} \tag{19}$$

$$MAPE = \left(\frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right|\right) \times 100 \tag{20}$$

IV. RESULT AND DISCUSSION

A. Statistic Values

Table I is an essential part of the inquiry since it presents a comprehensive summary of the statistical data obtained from the dataset. The table's clarity and comprehensibility are enhanced by the inclusion of OHLC price and volume data. In order to conduct a thorough and precise examination of the data, statistical measures such the mean, variance, minimum (min), maximum (max), standard deviation (Std.), 25%, 50%, and 75% should be employed.

B. Compare and Analyses

Finding and assessing the hybrid algorithm that yields the most accurate stock price forecasts is the primary goal of this research. The main goals of this research effort are to understand the many factors that influence stock market trends and to create prediction models. Giving analysts and investors relevant information to help them make informed and wise investment decisions is the main goal. The performance is comprehensively evaluated in Table II and Fig. 5 and Fig. 6. This report offers a full examination of the effectiveness of each method.

An evaluation of the GRU model has been conducted, both with and without the use of an optimizer. Various assessment metrics have been used, including MAE, RMSE, R^2, and MAPE. By using this approach, users may enhance their comprehension of the model's functioning and form well-informed judgments. Upon analyzing the test and training sets, it was seen that the GRU model produced low results for this technique when the optimizer was not present. According to the presented data, compared to the given metrics, the MFO-GRU model shows better results than the GRU. Upon comparing the findings, it was discovered that the use of the MFO-GRU model led to a decrease in the RMSE test score, resulting in a value of 17.90. The comparative study investigation demonstrated that, as shown in Table II, the SMA-GRU model exhibited superior performance compared to the MFO-GRU model. The MAE value of the SMA-GRU model was determined to be 11.43 for the testing set. The GOA-GRU model demonstrates superior effectiveness compared to the SMA-GRU model. The testing yielded a notable result of 0.9934, indicating the success of the GOA-GRU model. More evidence for the great degree of accuracy and reliability of the GOA-GRU model may be found in the empirical results already cited. The aforementioned results further confirm the model's effectiveness as a useful instrument for the particular purpose of stock price prediction.

TABLE I. A STATISTICAL SUMMARY OF THE GIVEN DATASET

	Open	**High**	**Low**	**Volume**	**Close**
mean	3215.85	3239.99	3191.98	26.68	3219.09
std.	358.75	364.71	349.13	12.25	358.57
min	2446.02	2488.48	2440.91	0.01	2464.36
25%	2987.06	3009.20	2968.36	16.87	2987.97
50%	3206.16	3230.08	3188.54	24.37	3210.37
75%	3386.34	3409.64	3364.42	33.44	3386.00
max	5174.42	5178.19	5103.40	85.71	5166.35
variance	128699.71	133016.59	121892.72	149.94	128573.90

TABLE II. THE ANTICIPATED EVALUATION RESULTS OF THE MODELS

Models/Metrics	**Train Set**				**Test Set**			
	R^2	***RMSE***	***MAPE***	***MAE***	***R^2***	***RMSE***	***MAPE***	***MAE***
GRU	0.9874	44.07	0.87	28.51	0.9854	20.90	0.48	15.71
MFO-GRU	0.9904	38.20	0.79	25.69	0.9893	17.90	0.42	13.65
SMA-GRU	0.9947	28.40	0.59	19.23	0.9928	14.72	0.35	11.43
GOA-GRU	0.9964	23.50	0.47	15.30	0.9934	14.06	0.33	10.61

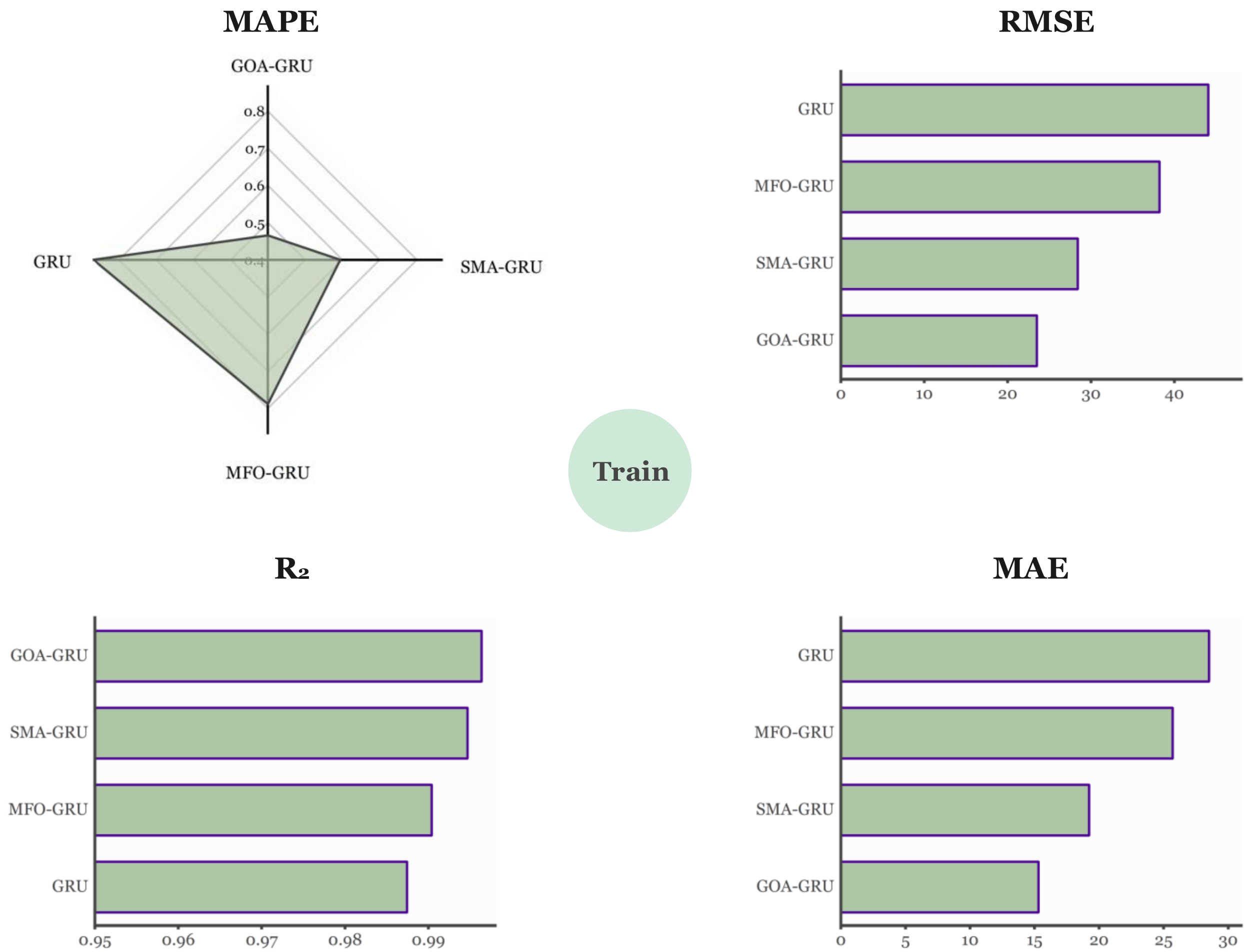

Fig. 5. The results for R^2, MAPE, MAE, and RMSE over the training phase.

The accuracy and consistency of the GOA-GRU model have been validated by the study's findings, which show that it can accurately predict stock prices. One may compare the Hang Seng index curves with the corresponding curves shown in Fig. 7 and Fig. 8 in order to assess the model's effectiveness. The accuracy of the stock value forecasts made by the COA-GRU model is greater than that of the GRU, MFO-GRU, and SMA-GRU models. Finally, the GOA-GRU model is a robust stock price forecasting tool since it consistently demonstrates very high levels of accuracy, reliability, and inference power when applied to historical data. It has been shown that this suggested framework is an excellent tool for precise stock market forecasting. Furthermore, in addition to the SSE index, we also utilized data from the Nikkei 225 and the S&P 500. The daily data for these indices, collected between 2013 and 2022, consist of the identical OHLC and trading volume. The results of the GOA-GRU model applied to this dataset are displayed in Table III. The GOA-GRU model exhibits robust generalizability when applied to the Nikkei 225 and S&P 500 datasets, as indicated by its low error metrics and high R^2 values. This implies that the model has successfully acquired knowledge of the latent patterns present in the data and is capable of generating precise forecasts on fresh data. The robustness and adaptability of the model's architecture and learning mechanisms are demonstrated by the consistency in performance observed across various datasets, which corresponds to distinct market conditions and characteristics.

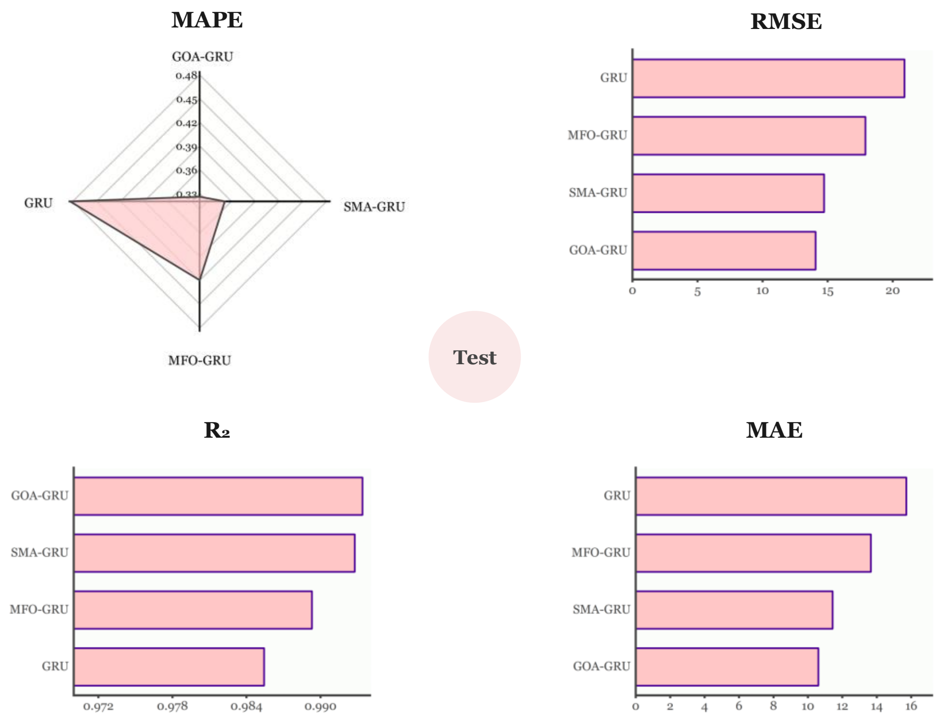

Fig. 6. The results for R^2, MAPE, MAE, and RMSE over the testing phase.

TABLE III. OBTAINED RESULTS OF THE GOA-GRU MODEL FOR S&P 500 AND NIKKEI 225

Train\test	Metrics	S&P 500	Nikkei 225
Train Set	R^2	0.9970	0.9982
	RMSE	28.77	144.47
	MAPE	0.72	0.58
	MAE	17.51	107.77
Test Set	R^2	0.9944	0.9958
	RMSE	22.57	74.29
	MAPE	0.42	0.21
	MAE	17.42	57.85

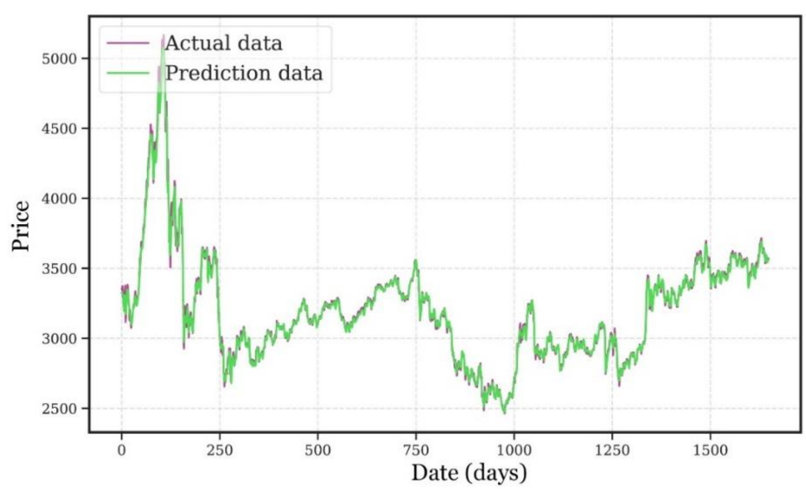

Fig. 7. Throughout the training procedure, the prediction curve was generated using the GOA-GRU approach.

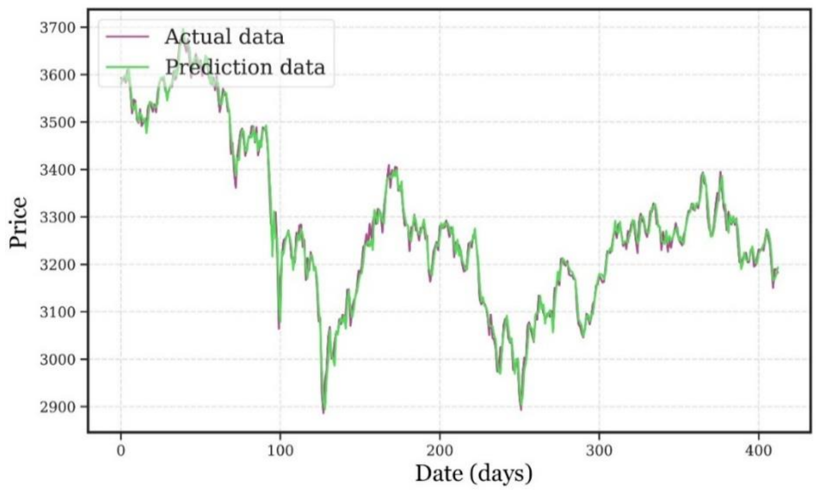

Fig. 8. Throughout the testing procedure, the prediction curve was generated using the GOA-GRU approach.

TABLE IV. A COMPARISON OF THE ASSESSMENT TO PREVIOUS RESEARCH

References	Context	Frameworks	R^2
[39]	Stock market prediciton	Linear regression	0.735
		SVM	0.931
		MLS-LSTM	0.95
[40]	Stock market prediction	LSTM	0.981
[41]	Stock future pdice prediction	LSTM	0.689
		EMD-LSTM	0.87
		CEEMDAN-LSTM	0.903
		SC-LSTM	0.687
		EMD-SC-LSTM	0.911
		CEEMDAN-SC-LSTM	0.92
[42]	Stock price prediction	DNN and LSTM	0.972
Present study	Stock market prediction	GOA-GRU	0.9934

The study's GOA-GRU framework demonstrated superior predictive accuracy compared to alternative methods by effectively mitigating their individual limitations, as indicated in Table IV. Conventional algorithms such as Linear Regression and Support Vector Machines (SVM) exhibit limitations when it comes to reproducing the complex nonlinear associations that are intrinsic in stock market data. LSTM and its variants, although proficient at capturing temporal dependencies, frequently encounter obstacles pertaining to parameter optimization and generalization, thereby impeding their overall performance. Although the integration of deep neural networks (DNN) and LSTM shows potential, it also brings about intricacies and computational demands. On the other hand, the GOA-GRU framework efficiently utilizes the optimization functionalities of the GOA to refine the GRU model's parameters, thereby reducing the impact of overfitting and inadequate parameter configurations. By incorporating this integration, GOA-GRU is capable of capturing temporal dependencies and nonlinear relationships. As a result, it achieves remarkable predictive accuracy and robustness, surpassing other models in the domain of stock market prediction.

Both the quality and quantity of data accessible for training significantly impact the efficacy of the GOA-GRU model. The utilization of a model in practical situations where the quality of data may differ can be hindered by suboptimal or biased predictions resulting from insufficient or biased data. Although the model may exhibit high accuracy when applied to the training dataset, its capacity to extrapolate to novel market conditions or uncover previously unseen data remains a matter of concern. The model's reliability in dynamic market environments may be compromised due to fluctuations in market dynamics or unanticipated occurrences, which can have an impact on its predictive performance. Implementing and interpreting machine learning models is further complicated by the incorporation of intricate optimization techniques, such as Grasshopper optimization. User accessibility and usability of the model may be compromised due to the difficulties that may arise for individuals lacking a comprehensive comprehension of both methodologies. The GOA-GRU model, similar to other predictive models, functions on specific assumptions concerning the fundamental connections between variables and market dynamics. Inaccurate predictions may ensue when the model fails to incorporate changes in market behavior or deviations from these assumptions, thereby reducing its resilience to evolving market conditions. Significant computational resources may be necessary for the training and deployment of the GOA-GRU model, especially when dealing with large-scale datasets or real-time forecasting applications. Practical limitations on scalability and real-time performance may consequently impede the widespread implementation of this technology in environments with restricted computational resources. When developing a predictive model for the financial sector, it is crucial to give utmost importance to ethical considerations such as data privacy, fairness, and regulatory compliance. Ethical guidelines and regulatory requirements must be followed when implementing the GOA-GRU model in order to mitigate potential risks, including misuse or bias, which could erode confidence in the model and its results.

V. Conclusion

The financial markets are a captivating and innovative advancement currently. The financial markets have a substantial influence on several areas, including business, employment, and technology. Predicting stock prices is an intricate undertaking that requires a comprehensive comprehension of several interrelated aspects. The stock market is susceptible to the impact of several factors, such as politics, society, and the economy. The system in issue is characterized by its dynamic and inherent complexity. In order to provide precise forecasts about future stock prices, it is crucial to take into account an extensive array of financial documents, earnings statements, market patterns, and other pertinent data. Furthermore, It is critical to recognize that the behavior of the stock market is significantly influenced by macroeconomic factors, including inflation, interest rates, and world economic situations. Building accurate and dependable prediction models for stock price forecasting can be challenging due to the numerous intricate components involved. One has to have a solid grasp of the intricate and erratic nature of the market to make reliable forecasts. The GOA-GRU model has shown a notable degree of accuracy and reliability and provides a workable solution to these problems. The effectiveness of many models, including GOA, SMA-GRU, and MFO-GRU, to forecast stock prices was assessed in the current research. Volume and Open, High, Low, and Close (OHLC) prices for the Shang Hai Stock Exchange Index spanning the years 2015 to June 2023 were incorporated into the dataset used in this research. Moreover, to verify the effectiveness of the hybrid model being analyzed, data from the Nikkei 225 and the S&P 500 covering the period from 2013 to 2022 were employed. The outcomes of the experiment reveal that the GOA-GRU model accurately predicts stock prices with a high degree of dependability and precision. An evaluation of the predictive capabilities and accuracy of the GOA-GRU model about multiple alternative models was undertaken as an integral component of the research procedure.

- The obtained data consistently demonstrated the higher performance of the GOA-GRU model over the other models. The computations show that the R^2 value of 0.9934 shows a high level of accuracy in the prediction models. As indicated by the observed RMSE score of 14.06 and the MAE value of 10.61, the testing outcomes revealed that the model's predictions exhibited a notable accuracy degree. The model's constant accuracy was shown by its MAPE score of 0.33.

Funding

This work was supported by the 2022 Humanities and Social Sciences Research Projects in Jiangxi Universities and Colleges (JC22219). Project Name: Study on the Incentive Mechanism and Realisation Path of Residents' Consumption under the Carbon Peaking and Carbon Neutrality Goals.

References

[1] R. E. Bailey, *The economics of financial markets*. Cambridge University Press, 2005.

[2] L. N. Mintarya, J. N. M. Halim, C. Angie, S. Achmad, and A. Kurniawan, “Machine learning approaches in stock market prediction: A systematic literature review,” *Procedia Comput Sci*, vol. 216, pp. 96–102, 2023, doi: 10.1016/j.procs.2022.12.115.

[3] K. Pardeshi, S. S. Gill, and A. M. Abdelmoniem, “Stock Market Price Prediction: A Hybrid LSTM and Sequential Self-Attention based Approach,” 2023. doi: 10.48550/arxiv.2308.04419.

[4] M. A. Abdullah and R. Chong, “Financial literacy: An exploratory review of the literature and future research,” *Journal of Emerging Economies and Islamic Research*, vol. 2, no. 3, pp. 32–41, 2014.

[5] S. Mukherjee, B. Sadhukhan, N. Sarkar, D. Roy, and S. De, “Stock market prediction using deep learning algorithms,” *CAAI Trans Intell Technol*, vol. 8, no. 1, pp. 82–94, 2023, doi: 10.1049/cit2.12059.

[6] W. M. Fong and S. K. Koh, “The political economy of volatility dynamics in the Hong Kong stock market,” *Asia-Pacific financial markets*, vol. 9, pp. 259–282, 2002.

[7] A. V Malyshko and O. S. Tykhomyrova, “Political factor’s impact on the ukrainian stock market dynamics,” *ЕКОНОМІЧНИЙ*, 2011.

[8] J. Coppel, J.-C. Dumont, and I. Visco, “Trends in immigration and economic consequences,” 2001.

[9] M. J. Iqbal and S. Z. A. Shah, “Determinants of systematic risk,” *The Journal of Commerce*, vol. 4, no. 1, p. 47, 2012.

[10] D. Siyuan, “How to Maximize Earnings by Using Financial Risk Management in the United States During Coronavirus Pandemic,” in *2021 6th International Conference on Social Sciences and Economic Development (ICSSED 2021)*, Atlantis Press, 2021, pp. 643–652.

[11] C. S. Katsikeas, N. A. Morgan, L. C. Leonidou, and G. T. M. Hult, “Assessing performance outcomes in marketing,” *J Mark*, vol. 80, no. 2, pp. 1–20, 2016.

[12] X. Pang, Y. Zhou, P. Wang, W. Lin, and V. Chang, “An innovative neural network approach for stock market prediction,” *J Supercomput*, vol. 76, no. 3, pp. 2098–2118, 2020, doi: 10.1007/s11227-017-2228-y.

[13] I. K. Nti, A. F. Adekoya, and B. A. Weyori, “A systematic review of fundamental and technical analysis of stock market predictions,” *Artif Intell Rev*, vol. 53, no. 4, pp. 3007–3057, 2020.

[14] J. Behera, A. K. Pasayat, H. Behera, and P. Kumar, “Prediction based mean-value-at-risk portfolio optimization using machine learning regression algorithms for multi-national stock markets,” *Eng Appl Artif Intell*, vol. 120, no. December 2022, p. 105843, 2023, doi: 10.1016/j.engappai.2023.105843.

[15] M. Bansal, A. Goyal, and A. Choudhary, “Stock Market Prediction with High Accuracy using Machine Learning Techniques,” *Procedia Comput Sci*, vol. 215, no. 2022, pp. 247–265, 2022, doi: 10.1016/j.procs.2022.12.028.

[16] D. Shah, H. Isah, and F. Zulkernine, “Stock market analysis: A review and taxonomy of prediction techniques,” *International Journal of Financial Studies*, vol. 7, no. 2, 2019, doi: 10.3390/ijfs7020026.

[17] T. A. Gerds, T. Cai, and M. Schumacher, “The performance of risk prediction models,” *Biometrical Journal: Journal of Mathematical Methods in Biosciences*, vol. 50, no. 4, pp. 457–479, 2008.

[18] R. Chopra and G. D. Sharma, “Application of artificial intelligence in stock market forecasting: a critique, review, and research agenda,” *Journal of risk and financial management*, vol. 14, no. 11, p. 526, 2021.

[19] Y. Li, N. Du, and S. Bengio, “Time-dependent representation for neural event sequence prediction,” *arXiv preprint arXiv:1708.00065*, 2017.

[20] J. D. Hamilton, *Time series analysis*. Princeton university press, 2020.

[21] M. M. Akhtar, A. S. Zamani, S. Khan, A. S. A. Shatat, S. Dilshad, and F. Samdani, “Stock market prediction based on statistical data using machine learning algorithms,” *J King Saud Univ Sci*, vol. 34, no. 4, p. 101940, 2022, doi: 10.1016/j.jksus.2022.101940.

[22] R. Dey and F. M. Salem, “Gate-variants of gated recurrent unit (GRU) neural networks,” in *2017 IEEE 60th international midwest symposium on circuits and systems (MWSCAS)*, IEEE, 2017, pp. 1597–1600.

[23] S. Li, H. Chen, M. Wang, A. A. Heidari, and S. Mirjalili, “Slime mould algorithm: A new method for stochastic optimization,” *Future Generation Computer Systems*, vol. 111, pp. 300–323, 2020.

[24] M. Shehab, L. Abualigah, H. Al Hamad, H. Alabool, M. Alshinwan, and A. M. Khasawneh, “Moth–flame optimization algorithm: variants and applications,” *Neural Comput Appl*, vol. 32, pp. 9859–9884, 2020.

[25] Y. Meraihi, A. B. Gabis, S. Mirjalili, and A. Ramdane-Cherif, “Grasshopper optimization algorithm: theory, variants, and applications,” *Ieee Access*, vol. 9, pp. 50001–50024, 2021.

[26] M. Nabipour, P. Nayyeri, H. Jabani, S. Shahab, and A. Mosavi, “Predicting stock market trends using machine learning and deep learning algorithms via continuous and binary data; a comparative analysis,” *Ieee Access*, vol. 8, pp. 150199–150212, 2020.

[27] W. Khan, U. Malik, M. A. Ghazanfar, M. A. Azam, K. H. Alyoubi, and A. S. Alfakeeh, “Predicting stock market trends using machine learning algorithms via public sentiment and political situation analysis,” *Soft comput*, vol. 24, no. 15, pp. 11019–11043, 2020, doi: 10.1007/s00500-019-04347-y.

[28] X. Yuan, J. Yuan, T. Jiang, and Q. U. Ain, “Integrated Long-Term Stock Selection Models Based on Feature Selection and Machine Learning Algorithms for China Stock Market,” *IEEE Access*, vol. 8, pp. 22672–22685, 2020, doi: 10.1109/ACCESS.2020.2969293.

[29] A. Moghar and M. Hamiche, “Stock Market Prediction Using LSTM Recurrent Neural Network,” *Procedia Comput Sci*, vol. 170, pp. 1168–1173, 2020, doi: https://doi.org/10.1016/j.procs.2020.03.049.

[30] M. Vijh, D. Chandola, V. A. Tikkiwal, and A. Kumar, “Stock Closing Price Prediction using Machine Learning Techniques,” *Procedia Comput Sci*, vol. 167, pp. 599–606, 2020, doi: https://doi.org/10.1016/j.procs.2020.03.326.

[31] I. R. Parray, S. S. Khurana, M. Kumar, and A. A. Altalbe, “Time series data analysis of stock price movement using machine learning techniques,” *Soft comput*, vol. 24, no. 21, pp. 16509–16517, 2020, doi: 10.1007/s00500-020-04957-x.

[32] H. N. Bhandari, B. Rimal, N. R. Pokhrel, R. Rimal, K. R. Dahal, and R. K. C. Khatri, “Predicting stock market index using LSTM,” *Machine Learning with Applications*, vol. 9, p. 100320, 2022, doi: https://doi.org/10.1016/j.mlwa.2022.100320.

[33] J. and D. A. Mehtab Sidra and Sen, “Stock Price Prediction Using Machine Learning and LSTM-Based Deep Learning Models,” in *Machine Learning and Metaheuristics Algorithms, and Applications*, S. and L. K.-C. and B. S. and W. M. and S. D. Thampi Sabu M. and Piramuthu, Ed., Singapore: Springer Singapore, 2021, pp. 88–106.

[34] H. Liu and Z. Long, “An improved deep learning model for predicting stock market price time series,” *Digit Signal Process*, vol. 102, p. 102741, 2020, doi: https://doi.org/10.1016/j.dsp.2020.102741.

[35] J. Ayala, M. García-Torres, J. L. V. Noguera, F. Gómez-Vela, and F. Divina, “Technical analysis strategy optimization using a machine learning approach in stock market indices,” *Knowl Based Syst*, vol. 225, p. 107119, 2021, doi: https://doi.org/10.1016/j.knosys.2021.107119.

[36] S. Saremi, S. Mirjalili, and A. Lewis, “Grasshopper optimisation algorithm: theory and application,” *Advances in engineering software*, vol. 105, pp. 30–47, 2017.

[37] S. Maldonado, C. Vairetti, A. Fernandez, and F. Herrera, “FW-SMOTE: A feature-weighted oversampling approach for imbalanced classification,” *Pattern Recognit*, vol. 124, Apr. 2022, doi: 10.1016/j.patcog.2021.108511.

[38] S. Nison, *Japanese candlestick charting techniques: a contemporary guide to the ancient investment techniques of the Far East*. Penguin, 2001.

[39] A. Q. Md *et al.*, “Novel optimization approach for stock price forecasting using multi-layered sequential LSTM,” *Appl Soft Comput*, vol. 134, p. 109830, 2023, doi: https://doi.org/10.1016/j.asoc.2022.109830.

[40] Z. Jin, Y. Yang, and Y. Liu, “Stock closing price prediction based on sentiment analysis and LSTM,” *Neural Comput Appl*, vol. 32, pp. 9713–9729, 2020.

[41] R. Zhu, G.-Y. Zhong, and J.-C. Li, “Forecasting price in a new hybrid neural network model with machine learning,” *Expert Syst Appl*, vol. 249, p. 123697, 2024, doi: https://doi.org/10.1016/j.eswa.2024.123697.

[42] A. C. Nayak and A. Sharma, *PRICAI 2019: Trends in Artificial Intelligence: 16th Pacific Rim International Conference on Artificial Intelligence, Cuvu, Yanuca Island, Fiji, August 26–30, 2019, Proceedings, Part II*, vol. 11671. Springer Nature, 2019.