Stock Price Forecasting with Optimized Long Short-Term Memory Network with Manta Ray Foraging Optimization

Zhongpo Gao[1], Junwen Jing[2]*
School of Economics and Management, Harbin University, Harbin 150086, Heilongjiang, China[1]
School of Economics, Harbin University of Commerce, Harbin 150028, Heilongjiang, China[1]
School of Mathematical Physics, Xi 'an Jiaotong-Liverpool University, Suzhou 215028, Jiangsu, China[2]

***Abstract*—The stock market is a financial marketplace where investors may participate through the acquisition and sale of stocks in publicly traded companies. Predicting stock prices in the securities sector may be challenging due to the intricate nature of the subject, which necessitates a comprehensive grasp of several interconnected factors. Numerous factors, including politics, society, as well as the economy, have an impact on the stock market. The primary objective of financial market investing is to exploit larger profits. Financial markets provide many opportunities for market analysts, investors, and researchers in several industries due to significant technology advancements. Conventional approaches encounter difficulties in capturing the complex, non-linear connections that exist in market data, which requires the implementation of sophisticated artificial intelligence models. This paper presents a new approach to tackling certain issues by suggesting a unique model. It combines the long short-term memory method and Empirical Mode Decomposition with the Manta Ray Foraging Optimization. When tested in the current study's dynamic stock market, the EMD-MRFO-LSTM model outperformed other models regarding performance and efficiency. The Nasdaq index data from January 2, 2015, to June 29, 2023, were used in this study. The findings demonstrate how the suggested model is capable of making precise stock price predictions. The suggested model offers a workable approach to studying and predicting stock price time series by obtaining values of 0.9973, 91.99, 71.54, and 0.57, for coefficient of determination (R^2), root means square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE), respectively. Compared to other methods currently in use, the proposed model has a higher accuracy in forecasting and is more physically relevant to the dynamic stock market, according to the outcomes of the experiment.**

Keywords*—Stock price; hybrid forecasting method; Manta Ray Foraging Optimization; empirical mode decomposition; Nasdaq index

I. Introduction

Having a thorough understanding of the stock market is essential for anyone working in the finance sector. To reduce risk and maximize returns, investors have to be able to correctly forecast stock prices. Macroeconomic policies, stock options, capital movements of significant firms, and changes in ownership are only a few of the important variables that might affect stock prices. Making precise predictions is difficult due to the unpredictable nature of price fluctuations, which are characterized by non-linear, non-stationary, stochastic noise, and fluctuating [1]. Although econometric and statistical methods could be employed to forecast asset market-related series statistics analysis is constrained by excessive noise and non-linearity.

Artificial intelligence models are preferred over traditional methods for learning complex and non-linear relationships. In the domain of quantitative analysis in finance, their popularity is on the rise because of their ability to extract valuable data from input variables. They are used to detect patterns in historical data and forecast future trends [2], [3]. Deep learning has developed a number of sophisticated structures that address a variety of issues when it comes to varied data sets. A well-known model of neural networks that only processes data in one direction is called feedforward. However, working with sequential data in which earlier occurrences are crucial to projecting future results, can be challenging. As a result, such models have trouble correctly forecasting outcomes in such situations. To handle sequential data more successfully, there are sophisticated neural network models like recurrent neural network (RNN) as well as long short-term memory (LSTM). Data can be conveyed from one step to the next thanks to the RNNs' loop-based structural design, which allows them to retain critical data across time [4]. To train an RNN, a labeled training dataset is utilized to compute the error or cost between the actual and predicted values. The network's biases and weights are then continually adjusted to lower the error until it reaches the lowest practicable level. A gradient is used in the training process to calculate how much each parameter raises the cost [5]. Utilizing the gradient as a guide, backpropagation is used to iteratively change the error surface's parameters. Errors are propagated through this procedure between the intake and output layers. The main issue with this strategy is that to compute the gradient, partial derivatives must be generated for each parameter. Vanishing gradients are a typical problem during neural network training when gradients disappear or get smaller as they go back through the networks [6], [7]. The vanishing gradient problem was addressed through recurrent neural network development, including LSTM. The fundamental advantage of LSTM is that it can sustain long-term memory, which makes it a great option for tasks requiring long-term memory [8]. The vanishing gradient problem makes it impossible for traditional recurrent neural networks to retain long-term dependencies, however, LSTM is made to get around these limitations.

The process of decomposing time series data using empirical mode decomposition (EMD) involves separating the data into interpretable intrinsic mode functions (IMFs) and a residue that represents the trend [9]. Obtaining immediate frequency data from natural signals, which often exhibit nonlinear and nonstationary characteristics, is a technique supported by actual evidence.

As opposed to previous methods, the optimization process has witnessed breakthroughs recently, making it more efficient in handling challenges associated with confined, rigid, or unidentified search areas. The genetic algorithm (GA), a powerful computer-based technique, mimics natural selection to find the best solutions. GA utilizes a set of potential solutions called individuals and genetic operations involving selection, crossover, and mutation to produce new individuals [10]. A set of optimization techniques known as meta-heuristic algorithms was developed to overcome the constraints of mathematical computation, convergence issues, and the need for informed guesses [11]. By continually iterating through a collection of initial random replies, these distinct kinds of optimization strategies seek the best general solutions for specific problems. battle royale optimization (BRO), manta ray foraging optimization (MRFO), and grey wolf optimization (GWO) are three of the most well-known techniques in the field [12], [13], [14]. In response to gray wolves' social foraging behavior, the Gray Wolf Optimizer algorithm, a meta-heuristic optimization technique, was created [13].

The motivation for this research is multifaceted and is rooted in the complexity and non-linearity that are inherent in stock markets. These complex relationships are not adequately captured by conventional linear models, which is why advanced models are necessary to make more precise predictions. The unprecedented opportunity to improve stock price forecasting is presented by the accelerated advancements in artificial intelligence and machine learning, which can be employed to implement sophisticated algorithms such as LSTM networks and optimization techniques like MRFO. For investors, analysts, and financial institutions, accurate predictions are essential for making informed investment decisions, optimizing portfolios, mitigating risks, and maximizing returns. The dynamic and volatile nature of stock markets presents a challenge for conventional methods, underscoring the necessity of innovative approaches that provide reliable performance in real-world scenarios. The powerful combination of EMD and LSTM networks is achieved through the synergy between the two. EMD decomposes complex time series data into simpler components, enabling LSTM to accurately represent them. The efficacy of the model is further improved by the integration of advanced optimization methods, such as MRFO, for hyperparameter tuning. The value of accurate stock price prediction models in trading and investment contexts can be underscored by the EMD-MRFO-LSTM model's superior performance on Nasdaq index data, which can demonstrate their practical application and real-world relevance. Following are the contributions of the investigation:

- A novel hybrid model that integrates EMD, LSTM network, and MRFO is introduced in this research. The accuracy and robustness of stock price predictions are improved by this combination, which capitalizes on the assets of each method.
- This research contributes to a more profound comprehension of market dynamics by effectively capturing the complex, non-linear relationships in stock market data. The model's capacity to analyze and process complex financial data is crucial in identifying the factors that influence stock price fluctuations.
- The research offers a thorough assessment of a variety of stock price forecasting models, such as EMD-LSTM, LSTM, EMD-GA-LSTM, EMD-BRO-LSTM, and EMD-GWO-LSTM. The study provides vital insights into the relative performance of these models and the advantages of the proposed EMD-MRFO-LSTM model by benchmarking them.

The following text comprises the remaining contents of the paper. The background of the study is covered in Section II. Related works are specified in Section III. The materials, data gathering, decomposition, evaluation metrics, and methodology are detailed in Section IV. The experimental results are reported in Section V. The discussions of the results are presented in Section VI. In the concluding section, the study's findings are briefly discussed in Section VII. The prospects and challenges are discussed in Section VIII.

II. BACKGROUND

A multitude of determinants impact the stock market, which is a dynamic and intricate system. These determinants comprise investor sentiment, geopolitical events, and [15], [16], [17]. Accurately forecasting stock prices is critical in order to facilitate well-informed investment decision-making and proficient risk management. Conventional financial models, which heavily depend on technical and fundamental analysis, frequently fail to encompass the intricacies of market dynamics. Conventional approaches frequently encounter challenges in capturing the complex patterns and non-linear associations that are intrinsic to market data. The utilization of artificial intelligence models in finance is becoming more prevalent due to their capacity to extract valuable insights from historical data as well as to discover complex relationships among input variables, as discussed in this study. In addition to price forecasting, trend analysis, and anomaly detection, ML algorithms have been implemented in a variety of stock market prediction domains. These methods assist in discerning parallels and distinctions between equities, identifying market anomalies, and revealing concealed correlations that could potentially impact price fluctuations. The article's primary contribution is the introduction of the GWO-LSTM hybrid model, which integrates the operational characteristics of LSTM and GWO to enhance the accuracy of stock price forecasts [18], [19]. To verify the efficacy of the hybrid model, the research utilizes a stringent methodology that includes data analysis, model evaluation, and comparison with alternative techniques.

III. Related Works

The Shanghai Index's close price for the next day was predicted by Lu et al. [20] using the convolutional neural network (CNN) and LSTM approach. CNN's primary objective was to identify the most valuable features in the data, as well as the closing stock price was predicted using the LSTM approach. The problem stemmed from CNN's inability to identify the optimal feature from the input data. Rezaei et al. [21] introduced two hybrid algorithms called EMD-CNN-LSTM for stock price prediction. On the historical data of the S&P 500, Dow Jones Industrial Average, and Hang Seng Index dataset, Qiu et al. [22] constructed an LSTM-based model. Using the stock trading data from the S&P 500, to forecast the stock price for the ensuing 1, 5, and 10 minutes, Lanbouri et al. [23] employed the LSTM model for the high-frequency. To forecast the close price of the National Stock Exchange and NIFTY50 index, Yadav et al. [24] employed deep learning using the LSTM-based approach. According to the findings, a stateless LSTM model was discovered to be preferable due to its increased stability for time series forecasting problems. For the purpose of stock forecasting using time series data, Dash et al. [25] developed a novel machine learning (ML) technique that makes use of an optimized form of support vector regression. Zhang et al. [26] developed a two-stage prediction methodology that can accurately forecast stock prices. Three machine learning models, a nonlinear ensemble technique, and a decomposition algorithm are all incorporated into this mode. In the first stage, they decomposed stock price time series into sub-series using variational mode decomposition (VMD) as well as then used extreme learning machine (ELM), support vector regression (SVR), and deep neural network (DNN) to forecast each sub-series. Rao et al. [27] addressed the challenge of accurate stock market forecasting by proposing a hybrid machine learning model for stock market prediction. Accounting et al. [28] employed LSTM for predicting the Tehran stock market. The CatBoost algorithm has been utilized to predict financial distress, and the dataset was gathered from the Chinese stock market between 2016 and 2020 by Zhao et al. [29]. Kumar et al. [30] suggested a hybrid deep learning model that combines adaptive particle swarm optimization (PSO) as well as LSTM network.

IV. Materials and Method

A. Data Description

This study employs time series data, which are distinguished by their temporal dependencies. It is crucial to look at the volume of financial data and the open, high, low, and close prices (OHLC) over a specific period to conduct a thorough analysis. Open price is the initial price agreed upon by vendors and purchasers to conduct business following the market's regular trading hours. The open price holds considerable importance as it establishes the security's initial value for the course of the trading day. The term high price refers to the maximum price that fluctuates during a particular trading session for a given security. It represents the highest value point that the security price has reached during that particular period. The high price is indicative of the peak level of investors' demand and enthusiasm for the security throughout the trading session. In the context of a given trading process, a low price denotes the lowest price at which a specific security was transacted. The expression close price denotes the ultimate price at which certain securities were exchanged after a trade. It is the final price at which a transaction occurs just before the market's closing time. Volume denotes the aggregate quantity of shares or contracts that have been traded. As a result, information was gathered between January 2, 2015, and June 29, 2023, from the Yahoo Finance website's Nasdaq index.

The aforementioned information is contained within the dataset used in the investigation. Following the acquisition of the dataset, a comprehensive process of data cleansing was executed in order to preserve the precision and uniformity of the forecasting models. The multi-step procedure was designed to protect the dataset's integrity and avoid any inaccurate or incomplete information from being added that would cause problems. One of the critical stages required a careful analysis of the data to identify any outliers, anomalies, or discrepancies that can potentially undermine the validity of the outcomes. The information was preprocessed and cleaned using a variety of methods to ensure its suitability for use. To prevent gradient errors and inconsistent training results, the data was scaled and normalized. The data were normalized before training using the Min-Max-Scaler method, which helped to guarantee a stable model and avoid having too high weight values. Prices and volume for OHLC were used as the training data, which was fed into the model. High price, low price, open price, as well as volume data were given to the model for testing. The data were divided into 20% for testing and 80% for training.

By preprocessing the data to preserve its accuracy and consistency, it can be guaranteed that the models can accurately learn from historical stock price trends and make precise predictions. The inclusion of OHLC prices and volume data provided a comprehensive view of market activities, allowing the model to capture intricate patterns and dependencies in the stock market data.

B. Empirical Mode Decomposition

The empirical mode decomposition is a technique utilized to break down time series data into two components: a residual component that represents the trend, and a collection of interpretable intrinsic mode functions (IMFs) [31]. It is a technique that is backed by empirical evidence and used to obtain immediate frequency data from natural signals, which frequently display nonlinear and nonstationary properties. An IMF is a mathematical function that has an average value of zero and one extreme value between each time it crosses zero. Fig. 1 to 5 demonstrate the decomposition of a variety of stock market characteristics, including the Open, High, Low, Volume, and Close prices, using EMD. Each figure commences with the original time series data at the top, followed by 11 IMFs that capture various frequency components of the data. The highest frequency is captured in IMF 1, and subsequent IMFs progress to lower frequencies. After extracting the IMFs, the residual component is represented by the bottom plot in each figure, which illustrates the long-term trend. In addition to enhancing prediction models, this decomposition process also facilitates a detailed analysis by isolating various patterns within the stock market data.

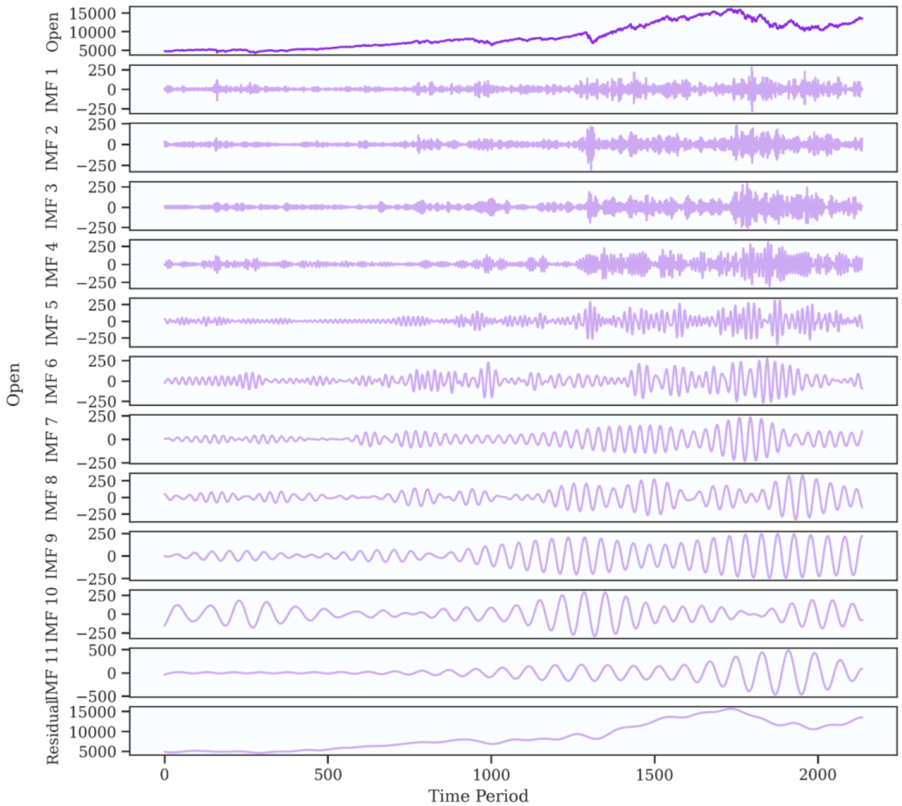

Fig. 1. The decomposition of Open price by EMD.

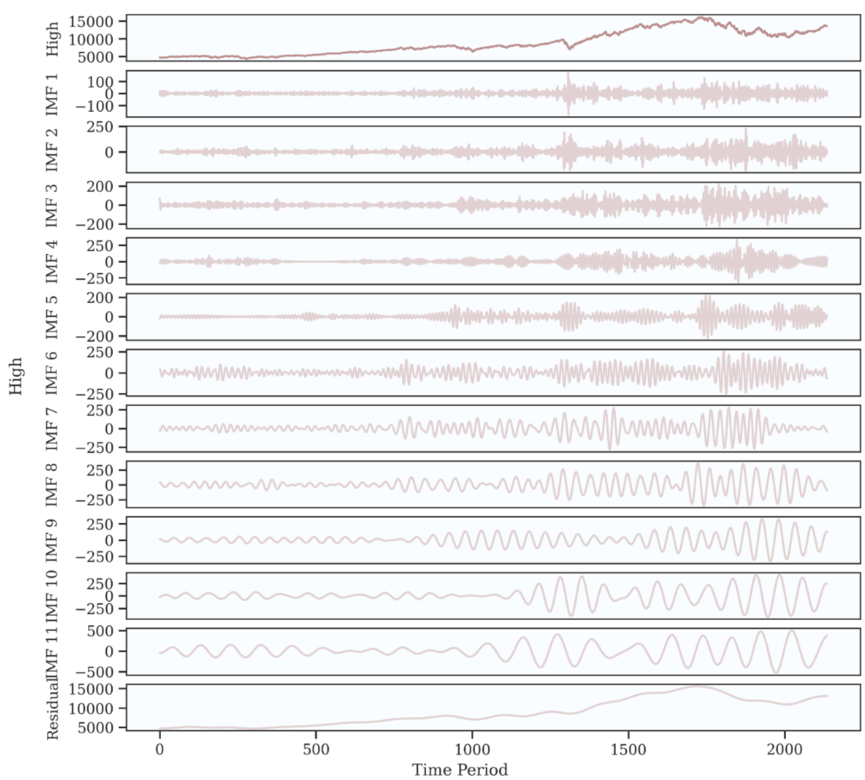

Fig. 2. The breakdown of High price using EMD.

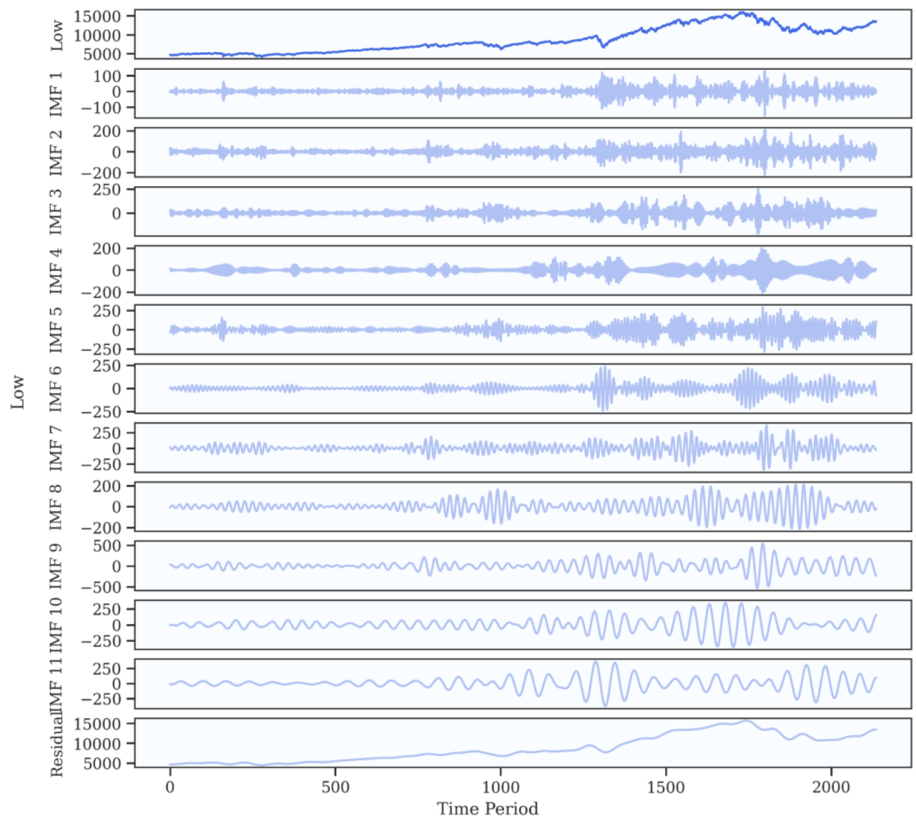

Fig. 3. The breakdown of Low price using EMD.

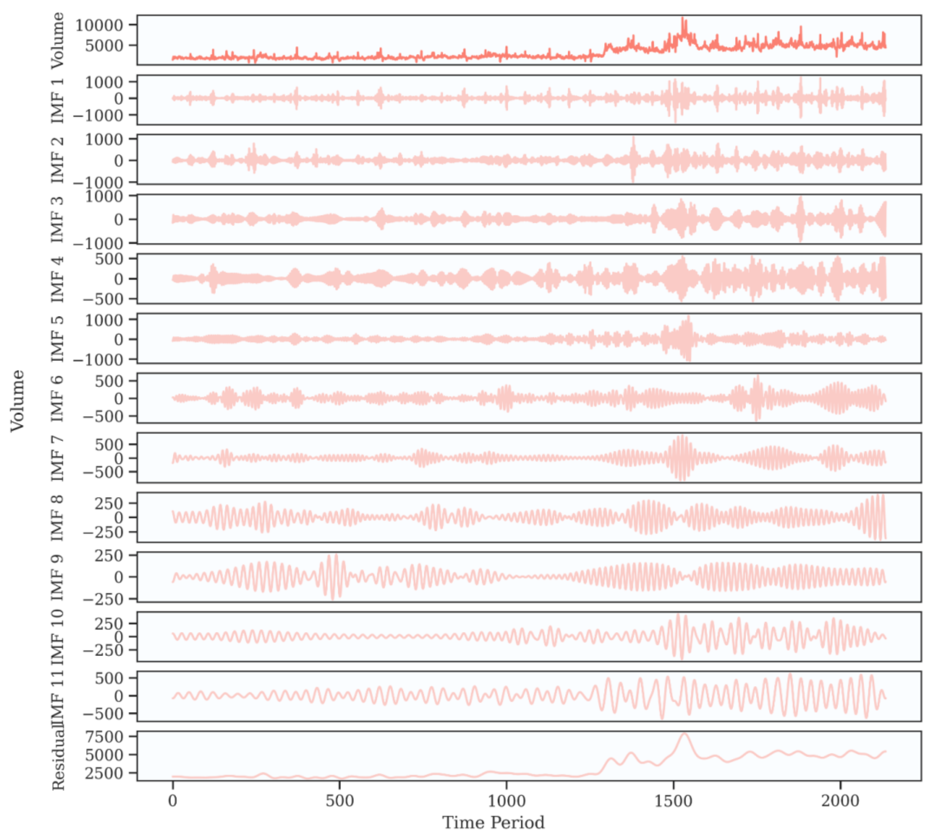

Fig. 4. The decomposition of Volume by EMD.

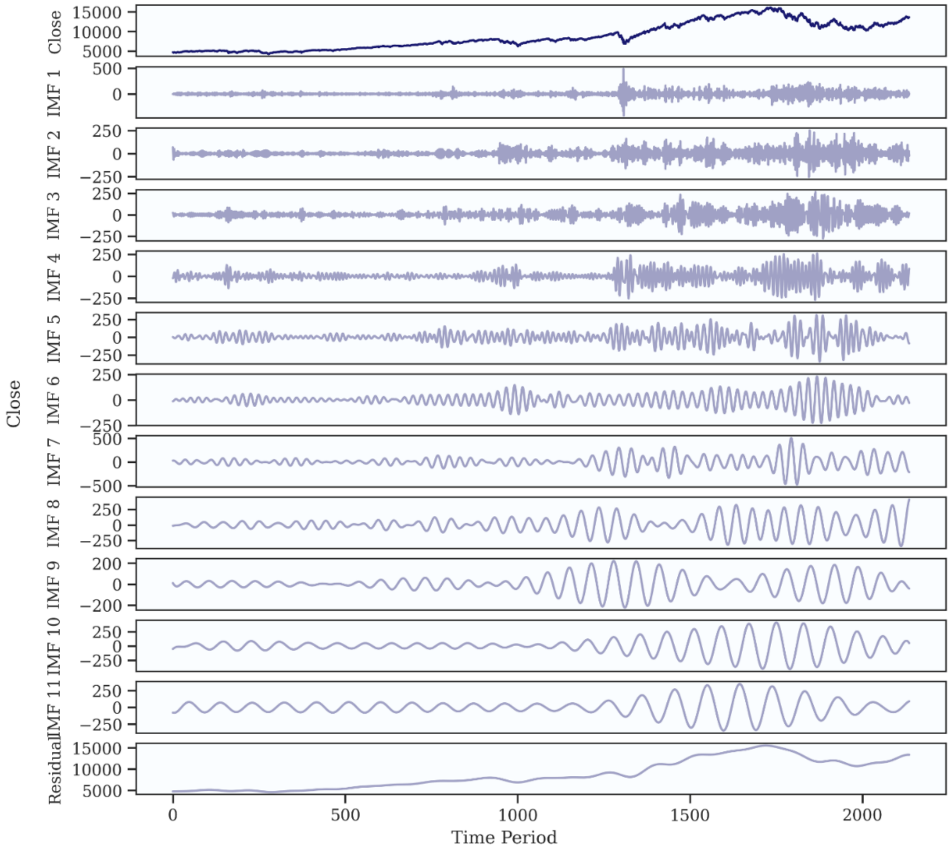

Fig. 5. The decomposition of Close price by EMD.

The method employed by EMD to decompose respiratory motion into IMFs is explained as follows [31].

- $Z_1$is the mean value of the upper and lower boundaries of a time series signal $x(t)$. These boundaries are determined by interpolating the local maximum and minimum points.
- Deducting Z_1 from the original time series $x(t)$yields the initial component P_1, which is defined as $p_1 = x(t) - Z_1$.
- Let P_1 represent the data in which the means of the upper and lower envelopes are Z_{11} during the second shifting process; $p_{11} = p_1 - z_{11}$.
- By the following conditions, the shifting process is terminated k times: (a) z_{1k} approaches zero; (b) the distinction between zero-crossings and the p_{1k} number of extrema does not surpass one, or (c) the maximum number of iterations has been completed. When this occurs, the IMF, denoted as p_{1k}, can be determined by dividing $p_{1k} = p_{1(k-1)} - z_{1k}$.
- The initial IMF (the shortest component of the data), represented by$a_1 = p_{1k}$, is subtracted from the data as $x(t) - a_1 = y_1$. This operation is repeated for each of the following values of $y_2 = y_1 - a_2, \dots y_n = y_{n-1} - a_n$.

Consequently, the initial time series $x(t)$is reduced to the collection of IMF functions shown below:

$$x(t) = (\sum_{i=1}^{n} a_i + y_n \quad (1)$$

C. Manta Ray Foraging Optimization

1) Inspiration: Manta rays are complex organisms despite their menacing appearance. They are among the largest marine organisms known to science [14]. Manta rays are flat-bodied from top to bottom as well as have two pectoral fins; they swim elegantly while birds soar effortlessly. Furthermore, they possess a pair of cephalic appendages that protrude anterior to their enormous, terminal jaws. They funnel prey as well as water into their jaws utilizing horn-shaped cephalic lobes while foraging. Then, using modified gill rakers, the prey is removed from the water. Two distinct species are identified as manta rays. The reef manta ray (Manta alfredi) is one of them that can attain a width of 5.5 meters and inhabits the Indian Ocean, western Pacific, and southern Pacific. The other is the 7-meter-wide giant manta ray (Manta birostris), which inhabits mild temperate, tropical, and subtropical oceans [14]. Their estimated age of existence is five million years. Many do not live to be the average age of 20 years due to the fact that they are pursued by fishermen. The illustration of MFRO is covered in Fig. 6.

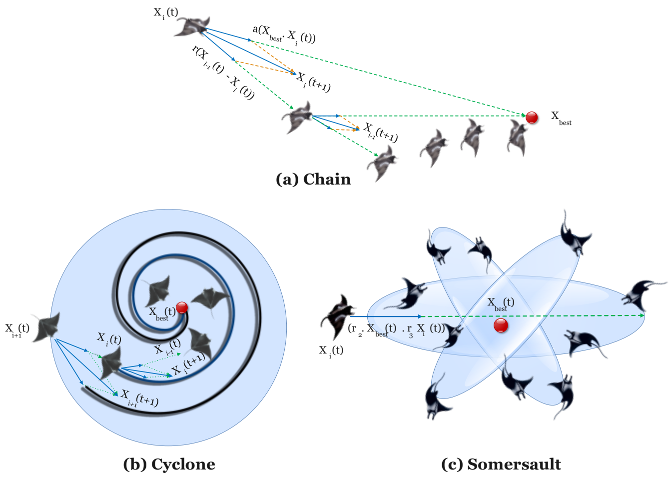

Fig. 6. The illustration of MRFO.

MRFO has been impacted by three distinct foraging behaviors: cyclone foraging, chain foraging, as well as somersault foraging.

2) *Chain foraging:* Manta rays can travel in the direction of plankton they detect using MRFO. An elevated plankton concentration correlates with a more favorable geographical location. While the optimal solution remains unknown, MRFO hypothesizes that the ideal solution thus far is the high-concentration plankton that manta rays desire to consume. A foraging chain is formed when manta rays are arranged from head to tail. At the same moment that individuals approach the food, they also approach the item that is immediately in front of them. Put simply, in each iteration, each individual gets revised with the best possible option that has been identified thus far, in addition to the solution that is currently in front of it. This is a representation of the chain foraging theoretical framework:

$$x_i^d(t+1) = \begin{cases} x_i^d(t) + r.\left(x_{best}^d(t) - x_i^d(t)\right) \\ +\alpha.\left(x_{best}^d(t) - x_i^d(t)\right) \ i = 1 \\ x_i^d(t) + r.\left(x_{i-1}^d(t) - x_i^d(t)\right) \\ +\alpha.\left(x_{best}^d(t) - x_i^d(t)\right) \ i = 2, \ldots N \end{cases} \tag{2}$$

$$\alpha = 2.r.\sqrt{|\log(r)|} \tag{3}$$

where $x_i^d(t)$denotes the location of the i-th individual at time t in the d-th dimension, r signifies an arbitrary vector from 0 to 1, α denotes the value of the ratio, and $x_{best}^d(t)$represents the plankton with the highest concentration. The current status of the i-th individual is established using the situation $x_{i-1}(t)$ as well as the position $i-1$-th of the food at the time $x_{best}(t)$ respectively.

3) *Cyclone foraging:* When a group of manta rays detects a region of deep-water plankton, they will spiral in their pursuit of the food in a continuous foraging chain. In contrast, as part of their cyclone foraging strategy, manta ray clusters swim each individual manta ray in the direction of the one in front of it, as opposed to spiraling towards the food. In other words, manta ray colonies engage in spiral foraging in a helical formation. An individual not only replicates the motion of the one preceding it but also proceeds in a spiral trajectory toward sustenance. The expression in mathematics that characterizes the spiral motion of manta rays in a two-dimensional space is as follows:

$$\begin{cases} X_i(t+1) = X_{best} + r.\left(X_{i-1}(t) - X_i(t)\right) \\ \quad +e^{bw}.cos(2\pi w).\left(X_{best} - X_i(t)\right) \\ Y_i(t+1) = Y_{best} + r.\left(Y_{i-1}(t) - Y_i(t)\right) \\ \quad +e^{bw}.sin(2\pi w).\left(Y_{best} - Y_i(t)\right) \end{cases} \tag{4}$$

where w represents a random number from zero to one.

This behavior of motion is extensible to n_D space. Theoretical representation of cyclone scavenging may be defined succinctly as:

$$x_i^d(t+1) = \begin{cases} x_{best}^d + r.\left(x_{best}^d(t) - x_i^d(t)\right) \\ +\beta.\left(x_{best}^d(t) - x_i^d(t)\right) \ i = 1 \\ x_{best}^d + r.\left(x_{i-1}^d(t) - x_i^d(t)\right) \\ +\beta.\left(x_{best}^d(t) - x_i^d(t)\right) \ i = 2, \ldots, N \end{cases} \tag{5}$$

$$\beta = 2e^{r_1\frac{T-t+1}{T}}.\sin(2\pi r_1) \tag{6}$$

The variables denoted as $[0,1], T$ the maximal number of iterations, β the weight coefficient, and r_1 the rand number.

Each individual conducts the search in a random manner, using the food as their reference position. As a result, the region where the most effective solution has been identified thus far benefits from cyclone foraging. Additionally, this behavior serves to significantly enhance the exploration process. By designating each individual, a reference position that is arbitrary and distinct from the current optimal one, we can compel them to seek out a new position. The mathematical equation for this mechanism, which enables MRFO to conduct an exhaustive global search and is primarily concerned with exploration, is provided below.

$$x_{rand}^d = Lb^d + r.(Ub^d - Lb^d) \tag{7}$$

$$x_i^d(t+1) = \begin{cases} x_{rand}^d + r.\left(x_{rand}^d - x_i^d(t)\right) \\ +\beta.\left(x_{rand}^d - x_i^d(t)\right) \ i = 1 \\ x_{rand}^d + r.\left(x_{i-1}^d - x_i^d(t)\right) \\ +\beta.\left(x_{rand}^d - x_i^d(t)\right) \ i = 2, \ldots, N \end{cases} \tag{8}$$

The variable x_{rand}^d denotes a position generated at random within the search space. Lb^d as well as Ub^d represent, respectively, the d-th dimension's minimum and maximum boundaries.

4) *Somersault foraging:* This behavior is characterized by the food's position being considered a pivot. Every individual undergoes a series of back-and-forth swims that involve a pirouette to a different position. As a result, they continuously adjust their positions in accordance with the most advantageous one discovered thus far. The formulation of the mathematical model is as follows:

$$x_i^d(t+1) = x_i^d(t) + S.\left(r_2.x_{best}^d - r_3.x_i^d(t)\right), i = 1, \dots, N \quad (9)$$

where S is the somersault factor that controls the variety of somersaults that manta rays perform, $S = 2, r_2$ and r_3 are two arbitrary values from the interval $[0,1]$. The framework of MRFO can be displayed in Fig. 7.

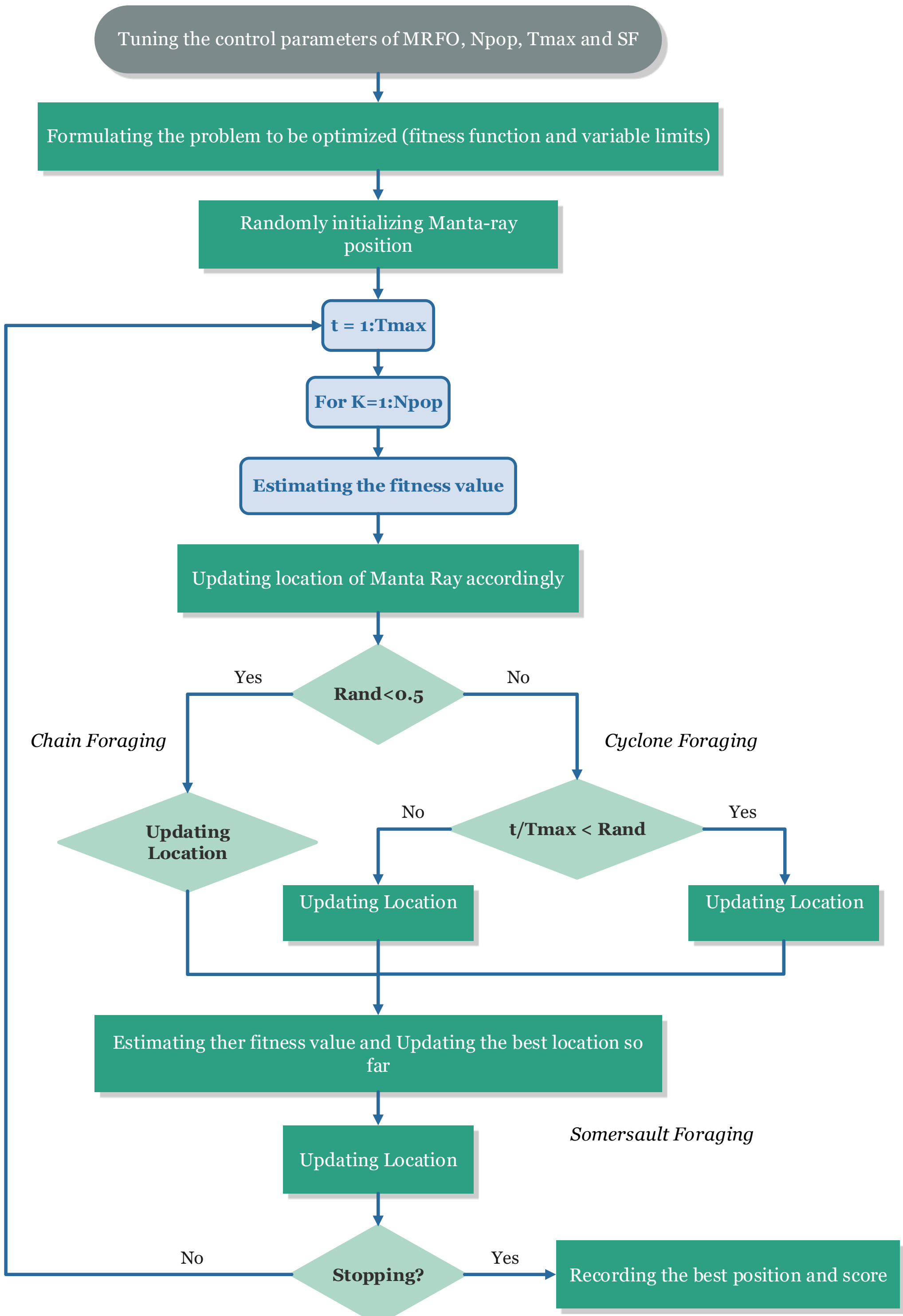

Fig. 7. The framework of MRFO.

D. Long Short-term Memory

The long short-term memory is a very well-liked and successful deep learning method. It is an effective tool for a variety of applications since it is built to handle and process enormous amounts of data [8]. Three a memory unit and gating units are used by the LSTM model to process incoming input. Together, these components control the flow of data, eliminating any extraneous material and generating output that is both brief and pertinent. The forgetting gate removes any potentially present irrelevant information, whereas the input gate handles the processing of incoming data. The function of the output gate is to regulate the flow of data that has been processed and generate a precise and relevant output. The gate formulas are used to sort, process, and store data, while the memory unit stores pertinent information for later use. The LSTM model may exclude any extraneous data by employing these algorithms, ensuring that only essential data is retained. As a result, it is a very effective method for handling vast amounts of data without creating extra clutter. The LSTM method is a significant asset in the field of deep learning since it is a strong and dependable tool for processing complex data sets [32]. The LSTM's operation is shown in the following equations. The forget gate decides whether to preserve or discard the information. A sigmoid layer processes the current input and the prior hidden state. The value that this layer output ranges from 0 to 1. Keep the data if the result value is more closely related to 1. Otherwise, disregard the knowledge.

$$f_t = \sigma\left(W_f \times [h_{t-1}, x_t] + b_f\right) \quad (10)$$

where σ denotes the sigmoid function, W_f is the mass that is linked to the forget gate, the prior hidden state demonstrates as h_{t-1}, the input value is x_t, as well as b_f denotes the bias associated with the forget gate.

The gate that accepts input is responsible for modifying the cell state. An individual sigmoid layer and a tan layer process the present input as well as the prior hidden state first. A data value is transformed by the sigmoid layer into a value that ranges from 0 to 1. Using the tanh layer, a data value is transformed into a value between -1 and 1. The outputs of the sigmoid layer and the tanh layer are multiplied by a point-wise procedure. then computes the new cell state value.

$$i_t = \sigma(W_i \times [h_{t-1}, x_t] + b_i) \quad (11)$$

The weight that is expressed as a symbol for the input gate by W_i and b_i is the component of the input gate that introduces bias.

The below equation is used to calculate the output of the tanh layer:

$$\tilde{C}_t = tanh(W_c[h_{t-1}, x_t] + b_c) \quad (12)$$

Equation below is used to calculate the new cell state:

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \quad (13)$$

The output gate decides what secret state will be shown next. A sigmoid layer processes the input at hand in addition to the preceding hidden state at the beginning. The changed cell state is then transmitted to a tan layer. The outputs of the sigmoid layer and the tanh layer are multiplied point-wise to find the subsequent hidden state. The new cell state as well as the next concealed state are then transferred to the following time step.

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (14)$$

where b_o denotes the bias associated with the output gate and w_o is the weight associated with the output gate.

$$h_t = o_t tanh(C_t) \quad (15)$$

Using the following equation, the subsequent hidden state is determined:

$$h_t = o_t tanh(C_t) \quad (16)$$

E. Genetic Algorithm

The genetic algorithm is a method of computation that simulate natural selection's technique to handle optimization and search issues [10]. With this algorithm, a collection of probable solutions known as people is created. To create new individuals, these individuals are then subjected to genetic processes like mutation, recombination, and selection. The illustration and framework of GA can be seen in Fig. 8 and 9. This evaluation procedure is iterative and is performed over several generations until a workable answer is discovered. As a result, GA is a potent instrument that is frequently employed in many different industries, involving engineering, finance, and science, to name a few [33]. Three components are essential to GA [34]. A chromosome is an encoded string of numbers or characters that is given to each individual by the encoding component. Encoding methodology is determined by the precise problem that must be resolved. Following this, the fitness metric is applied to assess how each individual embodies the solution. The ability to exercise has been deliberately designed to mitigate the present issue. Evolutionary operators utilize the crossing, transformation, and choice procedures. When two people's chromosomes crossover, a new being is created, mutation indiscriminately modifies an individual's chromosomes, and selection is used to determine which individuals are the most fertile.

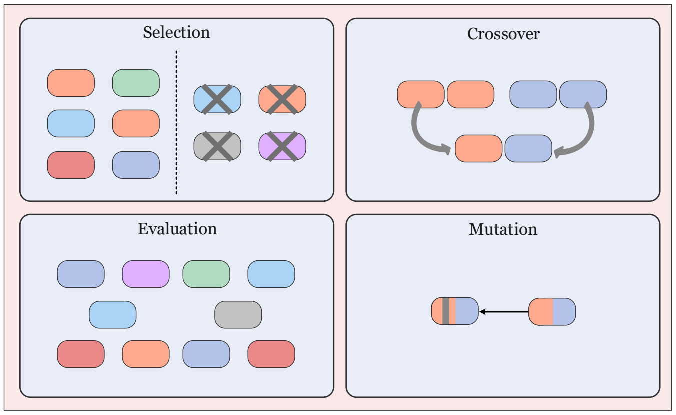

Fig. 8. The illustration of GA.

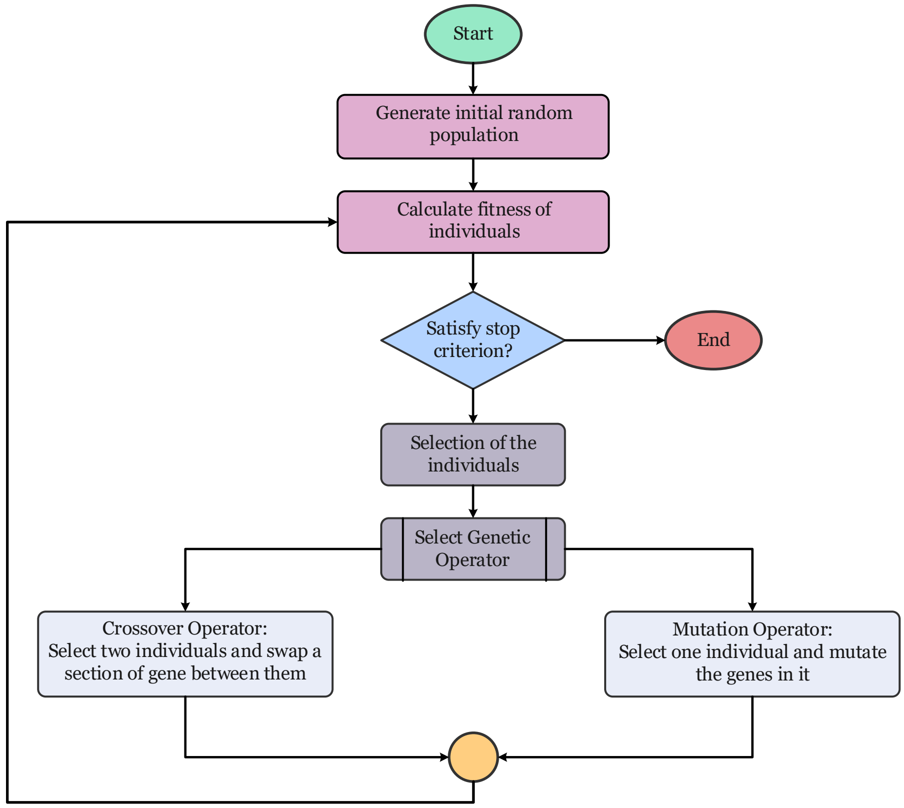

Fig. 9. The framework of GA.

F. *Battle Royale Optimizer*

The battle royale optimizer is the name of a meta-heuristic algorithm that Farshi proposed [12]. The algorithm was inspired by a popular multiplayer online game in which players must eliminate rivals to find a safe haven to survive. Stepping outside the safe zone in the game puts the player at risk of getting hurt or being eliminated [12]. The damage rate of the injured player is calculated using the equation shown below:

$$x_i.damage = x_i.damage + 1 \tag{17}$$

Injured players strive to switch to other positions to confront the enemy. The equation that follows shows the players' newest placements:

$$x_{dam,d} = x_{dam,d} + r\left(x_{best,d} - x_{dam,d}\right) \tag{18}$$

where $x_{dam,d}$ denotes the location of the wounded player in the dimension d, the best solution in dimension d is indicated by the notation $x_{best,d}$, and r is a random number generated from a uniform distribution between 0 and 1. The search agents are distributed at random throughout the problem space and cover it equally.

In a d-dimensional problem space, the upper limit and lower bound are represented by ub_d and lb_d the following equation:

$$x_{dam,d} = r(ub_d - lb_d) + lb_d \tag{19}$$

The best approach is shown in the formula below, and the worst-fitting options are discarded. Considering this, the starting value Δ is $\log_{10}(MaxCicle)$, where MaxCicle is the number of repetitions:

$$\Delta = \Delta + round\,(\frac{\Delta}{2}) \tag{20}$$

G. *Gray Wolf Optimizer*

A meta-heuristic technique has been used to create a novel optimization process known as the gray wolf optimizer. The approach, which mimics the seeking strategies and social organization of gray wolves, was first presented by Mirjalili et al. [13]. The best solution is Alpha, while there are four alternatives in the leadership hierarchy, Omega is the last competitor: Beta, Alpha, Omega, and Delta.

The strategy employs three primary hunting techniques to mimic wolf behavior: pursuing prey, enclosing prey, and attacking prey.

$$\vec{D} = \left|\vec{C} \cdot \vec{X}_p(t) - \vec{X}(t)\right|$$
$$\vec{X}\,|\,(t+1) = \vec{X}_p(t) - \vec{A} \cdot \vec{D} \tag{21}$$

where $\overrightarrow{C}$ and $\vec{A}$ represent coefficient vectors, $\overrightarrow{D}$ signifies motion, t is the current stage of iteration, as well as $\overrightarrow{X}$ represents the whereabouts of a gray wolf. The construction of the parameter variables ($\overrightarrow{A}$ and $\overrightarrow{C}$) is based on the subsequent relationships:

$$\vec{A} = 2\vec{a} \cdot \vec{r}_1 - \vec{a}$$
$$\vec{C} = 2 \cdot \vec{r}_2 \quad (22)$$

The location of new search reps that involve omegas is adjusted utilizing the data from alpha, beta, and delta as follows:

$$\vec{D}_a = |\vec{C}_1 \cdot \vec{X}_a - \vec{X}|, \vec{D}_\beta = |\vec{C}_2 \cdot \vec{X}_\beta - \vec{X}|, \vec{D}_\delta = |\vec{C}_3 \cdot \vec{X}_s - \vec{X}| \quad (23)$$

$$\vec{X}_1 = \vec{X}_a - \vec{A}_1 \cdot \vec{D}_u, \vec{X}_2 = \vec{X}_\beta - \vec{A}_2 \cdot \vec{D}_\beta, \vec{X}_3 = \vec{X}_b - \vec{A}_3 \cdot \vec{D}_\delta \quad (24)$$

$$\vec{X}(t+1) = \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3} \quad (25)$$

where the wolves are indicated by the subscripts $\alpha, \beta,$ and δ to mount a final attack to complete the task. $\vec{a}$ is used to mimic the last attack by changing a value from 2 to 0, whereas an is a random variable between $-2\vec{a}$ and $2\vec{a}$. As a result, decreasing $\vec{a}$ would also cause $\vec{A}$ to decrease. The wolves were forced by $|\vec{A}| < 1$to cling to their prey. After following the leader wolf on a pack search, gray wolves disperse to collect sustenance before reconvening for an assault. In pursuit of prey, wolves may divide into groups if the value of $|\overrightarrow{A}|$ exceeds unity at random. Two of the most critical settings for the algorithm used by GWO are the number of wolves and generation. Generation after generation signifies a wolf's conclusive action. In addition, the number of wolves precisely reflects changes in performance estimates over time. In other words, the generation size multiplied by the wolf population will result in an equivalent quantity of objective function evaluations.

$$OFEs = N_W \times N_G \quad (26)$$

H. Performances Metrics

Several performance metrics were utilized to determine the dependability of future estimates. The root means square error (RMSE), mean absolute error (MAE), coefficient of determination (R^2) , and mean absolute percentage error (MAPE) were some of these. The accuracy of forecasting models can be evaluated using these metrics, which additionally helps to ensure that the estimates are solid and reliable.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \quad (27)$$

$$MAPE = \left(\frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right|\right) \times 100 \quad (28)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (29)$$

$$MAE = \frac{\sum_{i=1}^{n}|y_i - \hat{y}_i|}{n} \quad (30)$$

where $\overline{y}_i$ is the sample mean, $\hat{y}_i$is the predicted value, and y_i is the actual value.

V. Experimental Results

A. Statistical Values

Table I which contains comprehensive statistical data about the dataset is included in this section of the study. The data are made more understandable by the table's inclusion of OHLC prices and volume figures. A robust dataset for analysis is provided by the 2137 observations in this table. This provides a clear indication of the data's scope and the extensive period over which it was collected by presenting the count. The central tendencies of the dataset are represented by the mean values in the table, which provide an average perspective on the market's performance during the analyzed period. For example, the mean closing price of 8745.8210 serves as a foundation for comprehending typical market behavior. In the same way, the minimum and maximum values capture the extremes within the dataset, indicating the lowest and highest market activities observed, such as the minimum volume of 706.880 and the maximum closing price of 16057.440. Additionally, the table contains skewness values, which provide a deeper understanding of the asymmetry in the data distribution. The skewness, which is nearly zero, indicates that the distribution is fairly symmetrical, indicating that the data does not significantly favor one tail. This information is essential for comprehending the fundamental patterns in the data, which can have a substantial impact on modeling endeavors.

TABLE I. SUMMARY STATISTICS FOR THE DATA SET

	Open	**High**	**Low**	**Volume**	**Close**
count	2137	2137	2137	2137	2137
mean	8744.3560	8805.2870	8677.5740	3143.80	8745.8210
minimum	4218.810	4293.220	4209.760	706.880	4266.840
maximum	16120.920	16212.230	16017.230	11621.190	16057.440
skewness	0.4993140	0.493110	0.5027470	1.0283760	0.4977320

B. Models' Outcomes

Using data from Nasdaq Finance, the proposed method was both trained and evaluated. To forecast a numerical value, regression analysis is implemented. A large number of vendors, purchasers, and investors participate in the stock market, an extremely risky investment destination. A share typically signifies ownership in a corporation. Recognizing stock price patterns and making investments at the optimal time and location are the sole prerequisites for potentially generating profits. Determining as well as evaluating the hybrid algorithm that is most efficient in predicting stock pricing is thus the primary objective, provided that an event is accurately predicted at the appropriate moment. Determining and evaluating the most effective hybrid algorithm for investing price forecasts is the principal aim of this study. Elaborate variables that influence stock market patterns have been analyzed in conjunction with the development of forecasting models. The goal was to provide insightful information that would aid investors and analysts in making prudent investment decisions. A detailed assessment of the performance of each mode, as well as an examination of its effectiveness, is presented in Table II. The principal goal of this investigation is to ascertain as well as evaluate the most efficient hybrid method for predicting stock prices. Through the development of predictive models and an understanding of the fundamental factors that influence stock market trends, this research aims to assist analysts and investors in making well-informed investment choices.

TABLE II. EVALUATION FINDINGS OF THE SIX BENCHMARKING ALGORITHMS' STATISTICAL FORECASTS

MODEL/METRICS	**TRAIN SET**				**TEST SET**			
	R^2	***RMSE***	***MAE***	***MAPE (%)***	R^2	***RMSE***	***MAE***	***MAPE (%)***
LSTM	0.9766	449.05	383.63	5.62	0.9609	311.60	259.87	2.06
EMD-LSTM	0.9817	396.55	247.51	2.59	0.9703	271.63	208.55	1.70
EMD-GA-LSTM	0.9890	307.53	193.82	2.16	0.9809	217.40	165.24	1.32
EMD-BRO-LSTM	0.9916	268.93	178.14	2.17	0.9904	154.48	121.27	0.98
EMD-GWO-LSTM	0.9966	170.78	153.74	1.99	0.9951	122.01	93.96	0.75
EMD-MRFO-LSTM	0.9981	127.72	107.76	1.43	0.9973	91.99	71.54	0.57

VI. DISCUSSION

To assess the efficacy of the data analysis, four widely used metrics—RMSE, MAPE, MAE, and R^2—were applied. To fully evaluate the results, a comprehensive assessment of the accuracy of the analysis, precision, as well as overall performance can be done using these metrics. R^2, RMSE and MAPE criteria were assessed for the LSTM model using EMD decomposition both with and without the optimizer. During the evaluation period, LSTM achieved a R^2value of 0.9609, as shown in Table II and Fig. 10 and 11. Frequently, the process of decomposing a problem reveals functions or recurrent patterns that apply to numerous components. The act of reusing modules or components serves to enhance stability, decrease the probability of errors, and accelerate the development process. In light of the data provided in this section, Clearly, it is apparent that the utilization of EMD decomposition decreases the value of MAE to 208.55 and 247.51, respectively, during the testing and training phases. When optimizers are added to the LSTM model, its efficacy is substantially enhanced. By enabling the efficient modification of model parameters, optimizers mitigate performance degradation. In order to achieve a convergent set of parameters, a multitude of optimizers implement unique strategies, including adaptive learning rate, slope descent, momentum, and others. This degree of effectiveness accelerates the convergence process during training. The cited examples demonstrate that the simultaneous implementation of the GA optimizer and EMD decompose produces a more precise outcome and reduces calculation error, as shown in Table II. Moreover, in conjunction with GA, the BRO optimizer exhibited enhanced performance, as indicated by its reduced RMSE value. By attaining an MAE score of 93.96, EMD-GWO-LSTM demonstrated superior efficacy in comparison to EMD-BRO-LSTM. According to regression analysis, the EMD-MRFO-LSTM model is an exceptionally precise and reliable instrument. The respective R^2 score of the model for the testing dataset was 0.9973. The outcomes of this study illustrate the model's robust predictive capability and its ability to explain nearly all of the variability in the data. Decreased values signify greater precision, as they represent the discrepancy between the predicted and realized values. In light of the exceptional accuracy demonstrated by both the training and assessment datasets, the EMD-MRFO-LSTM model has been validated.

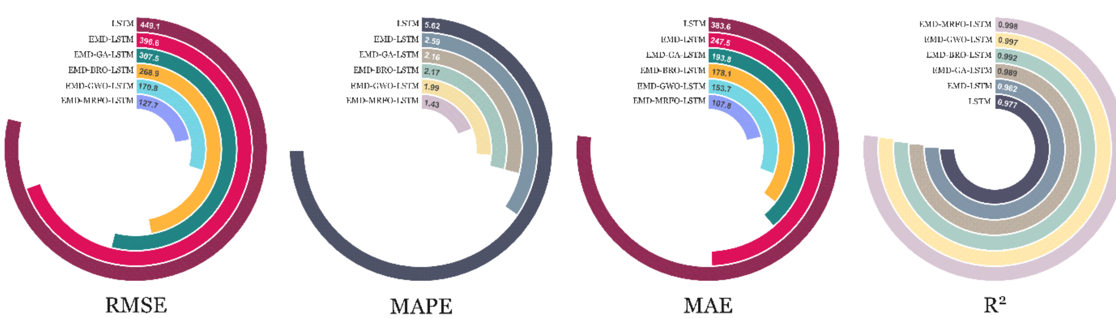

Fig. 10. Evaluation values for each model in the training set.

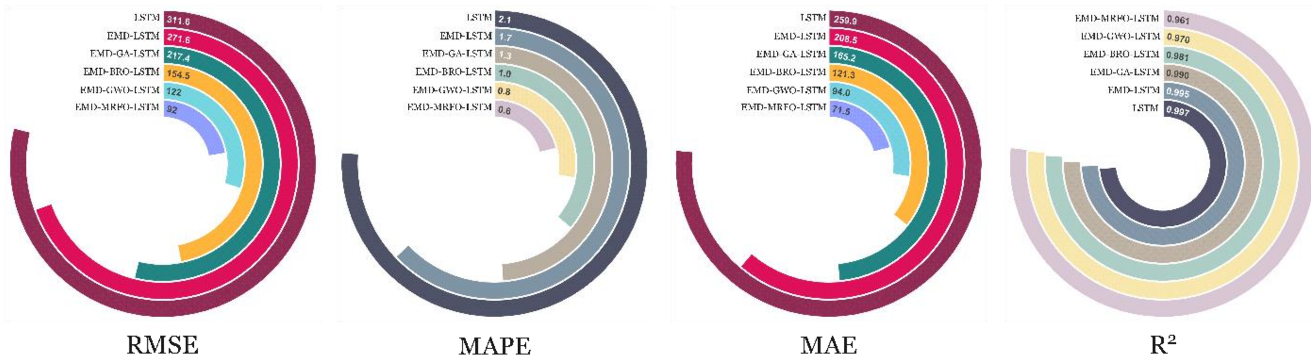

Fig. 11. Evaluation values for each model in the testing set.

The EMD-MRFO-LSTM model's ability to forecast the Nasdaq index during both the training and testing phases is illustrated in Fig. 12 and 13. The black line in Fig. 12, which illustrates the training phase, represents the actual Nasdaq index, while the red line illustrates the predicted values from the model. The model's capacity to learn from historical data is underscored by the close alignment of the two curves. Particularly, the Nasdaq index's reversal points, peaks, and troughs are precisely captured by the model. The predicted values closely follow the actual values at reversal points, where the market changes direction. The model also accurately predicts the market's peaks and troughs, matching them with precision. The model's capacity to learn intricate market dynamics and patterns is illustrated by this precise fit. The EMD-MRFO-LSTM model's robustness and generalization capability are validated by its continued performance in Fig. 13, which represents the testing phase. The model closely aligns the predicted peaks and troughs with the actual market values and maintains its accuracy in predicting reversal points, where the market shifts direction. This consistency in the testing phase, during which the model encounters new, unseen data, emphasizes its reliability and effectiveness in real-world applications. The model's ability to accurately capture critical market prices is demonstrated by the detailed fit between the predicted and actual curves in both phases, rendering it a valuable tool for investors and analysts. The EMD-MRFO-LSTM model's utility in stock market prediction is confirmed by its ability to accurately predict future market trends, which aids in the making of informed investment decisions.

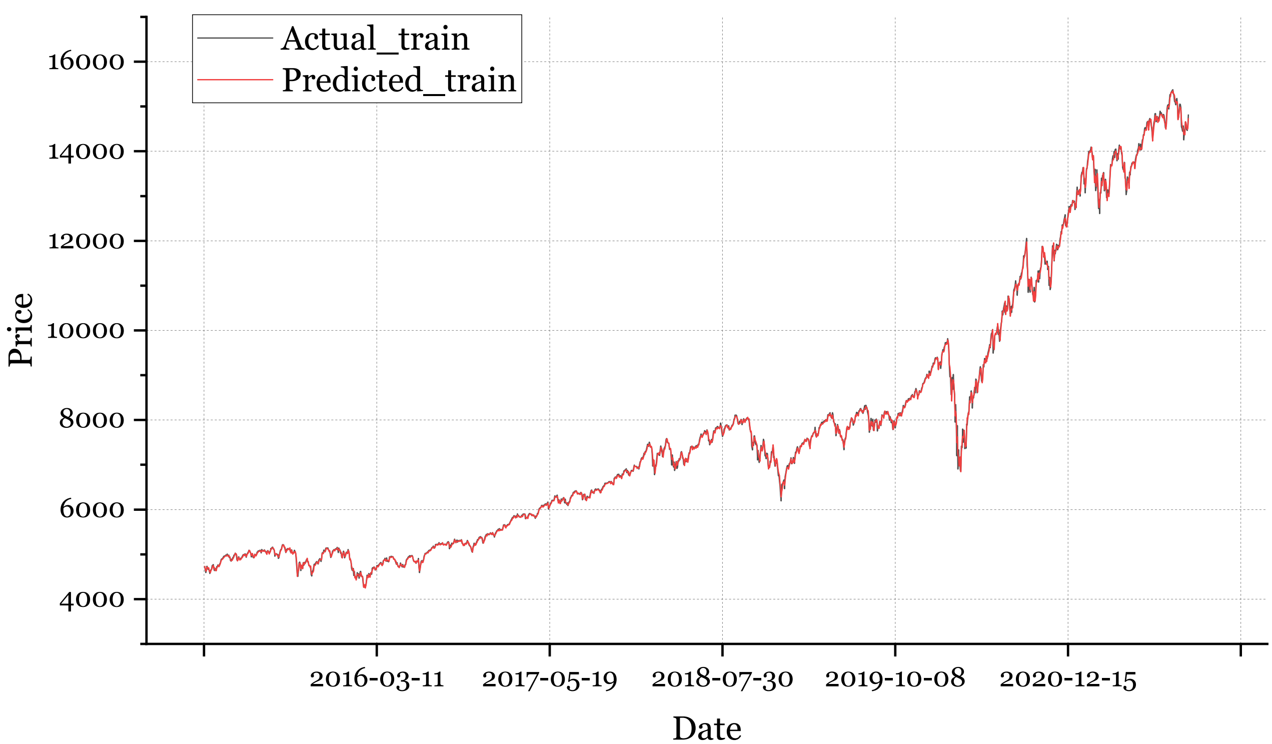

Fig. 12. Training-generated forecasting curve employing EMD-MRFO-LSTM.

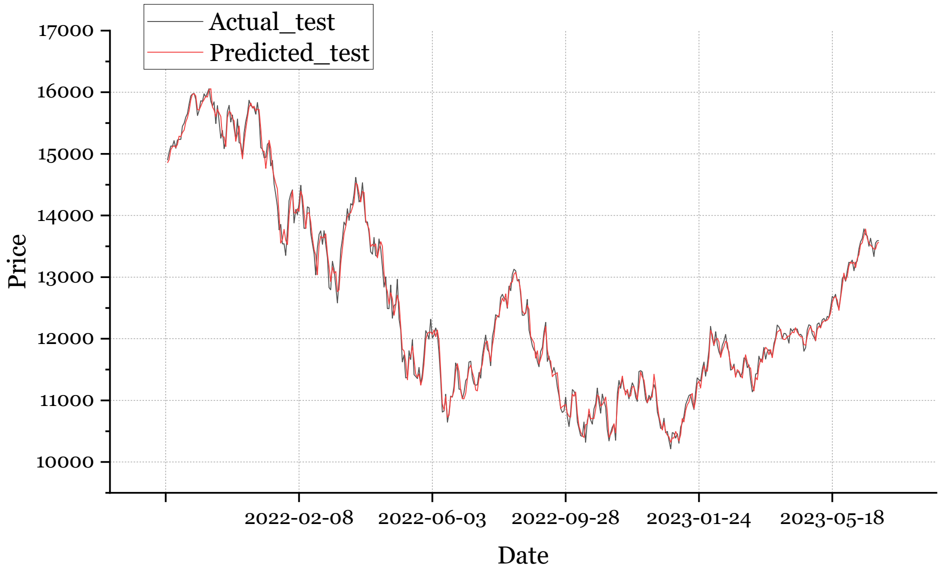

Fig. 13. Testing-generated forecasting curve employing EMD-MRFO-LSTM.

According to Table III, in comparison to both traditional and advanced benchmark models, such as support vector regression (SVR), random forest (RF), multilayer perceptron (MLP), and long short-term memory (LSTM), the proposed EMD-MRFO-LSTM model has exhibited remarkable superiority in predicting stock market prices. In particular, the model attains an R^2 value of 0.9973, which is statistically significantly greater than those of SVR (0.9097), RF (0.9258), MLP (0.9442), and LSTM (0.9609). The EMD-MRFO-LSTM model's high R^2 value suggests that it can account for nearly 99.73% of the variance in the target data, which is a significant improvement over the closest competing model, LSTM, which explains 96.09%. Additionally, the EMD-MRFO-LSTM model demonstrates superior error metrics, with an RMSE of only 91.99, in contrast to 311.60 for LSTM, 329.03 for MLP, 379.37 for RF, and 418.38 for SVR. The MAE exhibits a comparable pattern, as evidenced by its value of 71.54, which is significantly lower than that of LSTM (259.87), MLP (254.90), RF (287.73), and SVR (361.18). Furthermore, the MAPE of the proposed model is a mere 0.57%, which is a negligible error margin in comparison to 2.06% for LSTM, 2.09% for MLP, 2.24% for RF, and 3% for SVR.

TABLE III. PERFORMANCE OF THE PROPOSED MODEL IN COMPARISON WITH BENCHMARK MODELS

MODEL/METRICS	***TEST SET***			
	R^2	***RMSE***	***MAE***	***MAPE (%)***
SVR	0.9097	418.38	361.18	3
RF	0.9258	379.37	287.73	2.24
MLP	0.9442	329.03	254.90	2.09
LSTM	0.9609	311.60	259.87	2.06
EMD-MRFO-LSTM	0.9973	91.99	71.54	0.57

The EMD-MRFO-LSTM model's stability and reliability are further demonstrated by its consistent results across various cross-validation methods, as illustrated in Table IV. The model achieves an R^2 of 0.9967, an RMSE of 92.82, an MAE of 72.33, and a MAPE of 0.58% through 5-fold cross-validation. The model maintains this high-performance level with an R^2 of 0.9970, an RMSE of 92.18, an MAE of 71.91, and a MAPE of 0.57% when a 10-fold cross-validation is implemented. The model's potential as a dependable tool for financial forecasting is underscored by these results, which are capable of delivering precise and stable predictions under varying conditions.

The EMD-MRFO-LSTM method utilized in the current study has the highest R^2 value among other methodologies in the literature, as illustrated in Table V. This model outperforms linear regression, support vector machine (SVM), multi-layer stacked long short-term memory (MLS-LSTM), convolutional neural network-bidirectional long short-term memory with attention mechanism (CNN-BiLSTM-AM), and combinations of long short-term memory and deep neural networks (LSTM and DNN) in terms of stock price prediction accuracy. This implies that the current model can account for nearly all of the variability in the stock market data. The incremental improvements in R^2 in comparison to existing methods underscore the effectiveness of integrating EMD with the MRFO and LSTM networks.

TABLE IV. PERFORMANCE OF THE PROPOSED MODEL WITH DIFFERENT CROSS-VALIDATIONS

K-folds/METRICS	***TEST SET***			
	R^2	***RMSE***	***MAE***	***MAPE (%)***
Without k-fold	0.9973	91.99	71.54	0.57
5-fold	0.9967	92.82	72.33	0.58
10-fold	0.9970	92.18	71.91	0.57

TABLE V. COMPARISON OF THE COEFFICIENT OF DETERMINATION FOR A DIVERSE ARRAY OF STOCK MARKET PREDICTION METHODOLOGIES IN LITERATURE

References	***Frameworks***	***R^2***
[35]	Linear regression	0.735
	SVM	0.931
	MLS-LSTM	0.95
[36]	LSTM	0.981
[37]	CNN-BiLSTM-AM	0.98
[38]	LSTM and DNN	0.972
Present study	EMD-MRFO-LSTM	0.9973

The EMD-MRFO-LSTM model is capable of forecasting stock market trends. This model has the potential to assist investors in making more informed decisions regarding the purchase, sale, or retention of stocks by predicting the future values of stock indices, such as the Nasdaq. Portfolio managers and financial advisors may find the EMD-MRFO-LSTM model advantageous for optimizing their investment portfolios. The model can help diversify portfolios to reduce risk and enhance returns by predicting the potential future performance of a variety of assets. It has the potential to offer valuable insights into the most advantageous times to rebalance portfolios by identifying the optimal moments to adjust the allocation of various assets. The EMD-MRFO-LSTM model may be implemented by financial institutions to evaluate and mitigate risks. The model can assist in the identification of potential periods of high volatility or downturns by forecasting market trends. This allows institutions to employ risk mitigation strategies, such as adjusting their exposure to specific assets or hedging, to potentially safeguard their investments from adverse market movements.

VII. CONCLUSIONS

Forecasting stock prices is a complex and multifaceted process that poses numerous challenges. Social, political, and economic changes, as well as other factors, all have an impact on the stock market, which is an ever-evolving and dynamic system. Future stock prices must be correctly predicted by considering a variety of factors. Constraints and variables abound in the procedure of forecasting stock prices, which can make it difficult to develop accurate and dependable prediction models. It is imperative to comprehend the market's non-linear and unpredictable characteristics to achieve this objective. Thankfully, the EMD-MRFO-LSTM model has proven to be dependable and precise, providing a workable solution to these issues. The current study evaluated the efficacy of various stock price forecasting models, such as EMD-LSTM, LSTM, EMD-GA-LSTM, EMD-BRO-LSTM, as well as EMD-GWO-LSTM. The hyperparameter optimization techniques GA, BRO, MRFO, and GWO were utilized to enhance the LSTM's parameters. Nevertheless, optimal outcomes were achieved when the MRFO optimization method was coupled with LSTM. From January 2, 2015, to June 29, 2023, OHLC pricing and the Nasdaq index's volume comprised the dataset employed in the research. According to the analysis, the EMD-MRFO-LSTM model was highly reliable and accurate at forecasting stock prices. The EMD-MRFO-LSTM model consistently displayed superior accuracy and efficacy in its predictions compared to other models tested during the study by having 0.9973, 91.99, 71.54, and 0.57 values for R^2, RMSE, MAE, and MAPE for the testing, respectively. In general, the EMD-MRFO-LSTM model demonstrates efficacy as a stock price forecasting instrument and the investor supplies astute information to enable prudent investment choices.

VIII. CHALLENGES AND PROSPECTS

A. Challenges

Numerous variables impact stock market data, which is inherently chaotic. Deriving significant patterns and trends from the data is a difficult task due to its intricate nature. Conventional feedforward neural networks encounter difficulties when confronted with sequential data in which past events hold significant importance in forecasting future outcomes. Gradients in models may dissolve or become extremely small during training, making it challenging for the network to effectively learn long-term dependencies. To make accurate predictions, it is vital to identify pertinent characteristics of the supplied data. The process of regard to the optimization for complex models necessitates traversing extensive parameter spaces. The evaluation of various models and techniques necessitates the use of rigorous metrics for assessment and methodologies. When evaluating the precision and efficacy of different methodologies, it is imperative to meticulously contemplate elements such as the origins of the data, the structures of the models, and the optimization strategies.

B. Prospects

Potentially more accurate and dependable stock market forecasts could result from the combination of sophisticated artificial intelligence models LSTM and optimization techniques (gray wolf optimization and meta-heuristic algorithms). These models have the potential to optimize forecasting capabilities by more effectively capturing the intricate relationships and patterns that are intrinsic in stock market data. The finance industry relies heavily on precise stock market forecasts to manage risk effectively. The proposed models have the potential to aid financial institutions and investors in making well-informed decisions related to risk mitigation and return optimization through the provision of more dependable forecasts. Investment firms and financial institutions that implement sophisticated predictive analytics methods are likely to gain a competitive advantage in the marketplace. By investing in modern technologies to improve decision-making procedures, these institutions can strengthen their financial achievements and maintain an excellent market position. The creation and implementation of advanced forecasting models offer educational programs and academic institutions with a vision to offer courses and seminars covering areas such as machine learning in the financial sector, optimization methodology, and quantitative analysis. This can help develop the next generation of professionals who have the expertise to handle complex financial issues.

Acknowledgment

This work was supported by the General Project of the National Social Science Foundation of China: Research on the Impact of the "Chinese Rules" of Digital Trade on the Construction of a Strong Trade Country (Project Approval No. 23BGJ031) (Zhongpo Gao).

References

[1] R. G. Ahangar, M. Yahyazadehfar, and H. Pournaghshband, "The comparison of methods artificial neural network with linear regression using specific variables for prediction stock price in Tehran stock exchange," arXiv preprint arXiv:1003.1457, 2010.

[2] A. Ashta and H. Herrmann, "Artificial intelligence and fintech: An overview of opportunities and risks for banking, investments, and microfinance," Strategic Change, vol. 30, no. 3, pp. 211–222, 2021.

[3] C. Milana and A. Ashta, "Artificial intelligence techniques in finance and financial markets: a survey of the literature," Strategic Change, vol. 30, no. 3, pp. 189–209, 2021.

[4] M. Kaur and A. Mohta, "A review of deep learning with recurrent neural network," in 2019 International Conference on Smart Systems and Inventive Technology (ICSSIT), IEEE, 2019, pp. 460–465.

[5] N. F. Hardy and D. V Buonomano, "Encoding Time in Feedforward Trajectories of a Recurrent Neural Network Model," Neural Comput, vol. 30, no. 2, pp. 378–396, Feb. 2018, doi: 10.1162/neco_a_01041.

[6] S. Hochreiter, "The vanishing gradient problem during learning recurrent neural nets and problem solutions," International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, vol. 6, no. 02, pp. 107–116, 1998.

[7] F. Adeeba and S. Hussain, "Native Language Identification in Very Short Utterances Using Bidirectional Long Short-Term Memory Network," IEEE Access, vol. 7, pp. 17098–17110, 2019, doi: 10.1109/ACCESS.2019.2896453.

[8] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Comput, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.

[9] A. Lotfipoor, S. Patidar, and D. P. Jenkins, "Deep neural network with empirical mode decomposition and Bayesian optimisation for residential load forecasting," Expert Syst Appl, vol. 237, p. 121355, 2024.

[10] M. Mitchell, An introduction to genetic algorithms. MIT press, 1998.

[11] X.-S. Yang, "Metaheuristic optimization," Scholarpedia, vol. 6, no. 8, p. 11472, 2011.

[12] T. Rahkar Farshi, "Battle royale optimization algorithm," Neural Comput Appl, vol. 33, no. 4, pp. 1139–1157, 2021.

[13] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey Wolf Optimizer," Advances in Engineering Software, vol. 69, pp. 46–61, 2014, doi: https://doi.org/10.1016/j.advengsoft.2013.12.007.

[14] W. Zhao, Z. Zhang, and L. Wang, "Manta ray foraging optimization: An effective bio-inspired optimizer for engineering applications," Eng Appl Artif Intell, vol. 87, p. 103300, 2020.

[15] Z. Chen, L. Zhang, and C. Weng, "Does climate policy uncertainty affect Chinese stock market volatility?," International Review of Economics & Finance, vol. 84, pp. 369–381, 2023.

[16] M. Khraiche, J. W. Boudreau, and M. S. R. Chowdhury, "Geopolitical risk and stock market development," Journal of International Financial Markets, Institutions and Money, vol. 88, p. 101847, 2023.

[17] J. W. Goodell, S. Kumar, P. Rao, and S. Verma, "Emotions and stock market anomalies: a systematic review," J Behav Exp Finance, vol. 37, p. 100722, 2023.

[18] M. M. Salamattalab, M. H. Zonoozi, and M. Molavi-Arabshahi, "Innovative approach for predicting biogas production from large-scale anaerobic digester using long-short term memory (LSTM) coupled with genetic algorithm (GA)," Waste Management, vol. 175, pp. 30–41, 2024.

[19] E. Gholamian, S. M. S. Mahmoudi, and S. Balafkandeh, "Techno-economic appraisal and machine learning-based gray wolf optimization of enhanced fuel cell integrated with stirling engine and vanadium-chlorine cycle," Int J Hydrogen Energy, vol. 51, pp. 1227–1241, 2024.

[20] W. Lu, J. Li, Y. Li, A. Sun, and J. Wang, "A CNN-LSTM-Based Model to Forecast Stock Prices," Complexity, vol. 2020, p. 6622927, 2020, doi: 10.1155/2020/6622927.

[21] H. Rezaei, H. Faaljou, and G. Mansourfar, "Stock price prediction using deep learning and frequency decomposition," Expert Syst Appl, vol. 169, p. 114332, 2021, doi: https://doi.org/10.1016/j.eswa.2020.114332.

[22] J. Qiu and B. Wang, "dan Zhou, C. Forecasting Stock Prices with Long-Short Term Memory Neural Network Based on Attention Mechanism," Advanced Design and Intelligent Computing, vol. 15, no. 1, 2020.

[23] Z. Lanbouri and S. Achchab, "Stock Market prediction on High frequency data using Long-Short Term Memory," Procedia Comput Sci, vol. 175, pp. 603–608, 2020, doi: https://doi.org/10.1016/j.procs.2020.07.087.

[24] A. Yadav, C. K. Jha, and A. Sharan, "Optimizing LSTM for time series prediction in Indian stock market," Procedia Comput Sci, vol. 167, pp. 2091–2100, 2020, doi: https://doi.org/10.1016/j.procs.2020.03.257.

[25] R. K. Dash, T. N. Nguyen, K. Cengiz, and A. Sharma, "Fine-tuned support vector regression model for stock predictions," Neural Comput Appl, vol. 35, no. 32, pp. 23295–23309, 2023.

[26] J. Zhang and X. Chen, "A two-stage model for stock price prediction based on variational mode decomposition and ensemble machine learning method," Soft comput, vol. 28, no. 3, pp. 2385–2408, 2024.

[27] K. V. Rao and B. V. Ramana Reddy, "Hm-smf: An efficient strategy optimization using a hybrid machine learning model for stock market prediction," Int J Image Graph, vol. 24, no. 02, p. 2450013, 2024.

[28] M. Accounting, S. Majid, M. Anzahaei, and H. Nikoomaram, "A Comparative Study of the Performance of Stock Trading Strategies Based on LGBM and CatBoost Algorithms Keywords :," International Journal of Finance and Managerial Accounting, vol. 7, no. 26, pp. 63–76, 2022.

[29] S. Zhao, K. Xu, Z. Wang, C. Liang, W. Lu, and B. Chen, "Financial distress prediction by combining sentiment tone features," Econ Model, vol. 106, no. November 2021, 2022, doi: 10.1016/j.econmod.2021.105709.

[30] G. Kumar, U. P. Singh, and S. Jain, "An adaptive particle swarm optimization-based hybrid long short-term memory model for stock price time series forecasting," Soft comput, vol. 26, no. 22, pp. 12115–12135, 2022, doi: 10.1007/s00500-022-07451-8.

[31] A. Zeiler, R. Faltermeier, I. R. Keck, A. M. Tomé, C. G. Puntonet, and E. W. Lang, "Empirical mode decomposition-an introduction," in The 2010 international joint conference on neural networks (IJCNN), IEEE, 2010, pp. 1–8.

[32] P. Suebsombut, A. Sekhari, P. Sureephong, A. Belhi, and A. Bouras, "Field data forecasting using lstm and bi-lstm approaches," Applied Sciences (Switzerland), vol. 11, no. 24, 2021, doi: 10.3390/app112411820.

[33] B. Gülmez and E. Korhan, "COVID-19 vaccine distribution time optimization with Genetic Algorithm," 2022.

[34] E. Alkafaween, A. B. A. Hassanat, and S. Tarawneh, "Improving Initial Population for Genetic Algorithm using the Multi Linear Regression Based Technique (MLRBT)," Communications - Scientific Letters of the University of Zilina, vol. 23, no. 1, pp. E1–E10, 2021, doi: 10.26552/com.C.2021.1.E1-E10.

[35] 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.

[36] 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.

[37] W. Lu, J. Li, J. Wang, and L. Qin, "A CNN-BiLSTM-AM method for stock price prediction," Neural Comput Appl, vol. 33, no. 10, pp. 4741–4753, 2021, doi: 10.1007/s00521-020-05532-z.

[38] 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.