Presenting a Hybrid Method to Overcome the Challenges of Determining the Uncertainty of Future Stock Price Identification

Zhiqiong Zou[1], Guangyu Xiao[2]*
Jingchu University of Technology, Jingmen 448000, Hubei, China[1]
DongEui University, Busan 47340, Busan, Republic of Korea[2]

***Abstract*—A particular location, framework, or forum where buyers and sellers congregate to trade products, services, or assets is referred to as an economic market. While the future is unpredictable and unknowable, it is still possible to make informed predictions about the course of events. Predicting stock market movements using artificial intelligence and machine learning is one such potential. Even if the stock market is volatile, it is still feasible and wise to use artificial intelligence to create well-informed forecasts before making an investment. The current work suggests a novel approach to increase stock price forecast accuracy by integrating the Radical basis function with Particle Swarm Optimization, Slime Mold Algorithm, and Moth Flame Optimization. The objective of the study is to improve stock price forecast accuracy while accounting for the complexity and volatility of financial markets. The efficacy of the proposed strategy has been tested in the real world using historical stock price statistics. Results demonstrate considerable accuracy improvements over traditional RBF models. The combined strength of RBF and the optimization technique enhances the model's ability to adapt to changing market conditions in addition to increasing prediction accuracy. Results were 0.984, 0.990, 0.991, and 0.994 for RBF, PSO-RBF, SMA, and MFO-RBF, respectively. The performance of MFO-RBF in comparison to RBF shows how combining with the optimizer can enhance the performance of the given model. By contrasting the outcomes of various optimizers, the most accurate optimization has been determined as the main optimizer of the model.**

Keywords—Stock market prediction movement; prediction models; Radical basis function; optimization approaches

I. INTRODUCTION

A crucial component of finance is the stock market. Accurate stock price predictions are essential for investors' risk management and profit-making. It is notable that information on stock prices that is accurately and scientifically forecasted can give regulators crucial help when creating appropriate financial market rules [1]. However, a number of factors, including macroeconomic policies, stock market choices, and the capital flow of significant corporations, and ownership changes, can have an impact on stock values. Unpredictable traits such as non-stationarity, non-linearity, aggregated fluctuation, and stochastic noise are present in the pattern of price movements. Maintaining the stock market's stability and security is vital since it is an integral component of national economies [2] [3]. Analyzing the behavior and performance of stock markets has emerged as a crucial field of research due to the possible hazards involved in [4]. Predicting the movement of stock prices is one of the most significant responsibilities in this respect since it helps investors make educated decisions and avoid dangers, as well as regulators, in stabilizing the financial markets. Uncertain prediction procedures and inaccurate prediction outcomes, however, can result in serious dangers [5]. Therefore, it is essential to create a solid and persuasive prediction model in order to reduce any potential dangers.

The econometric models are not sufficient for all jobs as research issues and application situations get more complicated. Time series analysis' newest favored technique is machine learning, which can be easily deployed, lacks rigid assumptions and considerable prior knowledge, yet has excellent non-linear mapping capabilities [6] [7]. Various methods are used for forecasting; a statistical model is a mathematical framework for analyzing and understanding data patterns. These models, which are a core component of statistics, are used to draw conclusions and forecasts about a population from a sample of data. Simple statistical models like linear regression or complicated ones like hierarchical linear models can be used [8]. However, statistical models also have some limitation that makes the prediction careless. When a statistical model is overly intricate and catches noise in the data rather than underlying patterns, overfitting occurs. As a result, new data may not generalize well, and model interpretability may suffer [9]. Predictive models are created using machine learning methods, including decision trees [10], random forests [11], neural networks [12], and support vector machines [13]. These models can handle non-linear connections and complicated patterns in data. The most potent technology nowadays is ML, which uses a variety of algorithms to enhance its performance on a particular case study. It is a widely held opinion that ML has a substantial capacity for identifying reliable data and seeing patterns in datasets [14]. When it comes to problem-solving, machine learning has proven to be a highly effective approach. Compared to conventional methods, machine learning offers a number of advantages that make it a popular choice in various fields. However, the decision between machine learning and other approaches ultimately depends on the specific problem at hand, the dataset being used, and any relevant restrictions [15].

The radial basis function (RBF) is the model used in this research, and the RBF is a versatile mathematical function that is widely used in numerous fields, including mathematics, machine learning, and data analysis [16]. One of the key features of RBFs is their radial symmetry, which means that their value depends solely on the distance from a central point or center. This makes RBFs particularly useful for a variety of applications, such as interpolation, function approximation, clustering, and more [16]. Due to their flexibility and applicability, RBFs have become an essential tool for researchers and practitioners in many areas of science and engineering [17]. Like other neural networks, RBF has the ability to learn the relationship between dependent and independent variables using several instances from recent datasets. The parallel units that make up the RBF are neurons.

Model optimization is an important step in the development of the presented model. Different methods and techniques are used to optimize model hypermeters, which in this article, Particle swarm optimization [18], slime mold algorithm [19] and moth flame optimization [20] [21] are used to optimize the hyperparameters of the model.

PSO is an optimization algorithm that is inspired by natural phenomena. It has been widely adopted for solving optimization and search problems. PSO is modeled on the social behavior of birds flocking or fish schooling and was developed by James Kennedy and Russell Eberhart in 1995 [18]. This heuristic algorithm is particularly useful for addressing optimization problems that involve complex and high-dimensional search spaces. Another optimization method used in this paper is SMA [22], which gives a fresh method based on the natural mucosal mold's oscillating characteristic. The SMA has some new characteristics thanks to a new mathematical method that applies adaptive weights for the procedure's simulation to produce positive and negative feedback of a biological wave-based mucosal mold emission wave to the best path for connecting food with the capacity to discover and offer high exploitation [23]. Another method for optimizing the hyperparameter of the model is Moth flame optimization; the optimization of the model was carried out using the Moth Flame Optimizer, a nature-inspired approach based on the behavior of butterflies at night. This optimizer takes inspiration from the way in which butterflies navigate towards the moon, which is a proven strategy for long journeys. However, it also recognizes the potential pitfalls of being drawn towards artificial light sources, which can lead to circular movements and a lack of progress. By formalizing this behavior, the MFO has been successfully applied in a range of optimization problems across diverse fields, such as power and energy systems, economic dispatch, engineering design, image processing, and medical applications [24]. Different criteria have been used to evaluate the results of the model, which are chosen depending on the type of model and the data that is used, Root Mean Square Error ($RMSE$), Mean squared error (MSE), Mean absolute error (MAE) and Coefficient of determination (R^2). Several models were used in this project to process a sizable dataset. The time period covered by the dataset was from 2015 to June 2023. The RBF algorithm was carefully developed to take into account a wide variety of input factors in order to guarantee that the final outputs were accurate and trustworthy. The daily transaction volume, high and low prices, and opening and closing prices where criteria has been used. The model was then put through a thorough testing process utilizing these same parameters to evaluate the accuracy of the model outputs. A model that can give traders and investors useful market insights that can aid them in making decisions that result in profitable investments is the final result of this rigorous training and testing procedure. The Google firm owns the stock from which the variable data was received.

The main contributions of the study are as follows:

The research paper presents an innovative methodology for enhancing the precision of stock price predictions through the integration of the RBF with optimization techniques, including PSO, SMA, and MFO. Through the integration of these techniques, the model attains substantial enhancements in precision when compared to conventional RBF models.

The effectiveness of the suggested approach is validated via empirical investigations employing historical stock price data. The findings demonstrate significant enhancements in accuracy, with the MFO-RBF model attaining the highest level of precision among the iterations that were evaluated.

Through a comparison of the results obtained from different optimizers, the research establishes the MFO method as the most precise optimizer for the given model. This emphasizes the significance of choosing the appropriate optimization method in order to maximize the accuracy of predictions.

The study imparts significant knowledge to institutional and individual investors alike through the provision of a dependable approach to forecasting stock prices. Through the utilization of algorithms and historical data, investors are able to execute informed and economical investment decisions, thereby substantially enhancing their prospects of attaining favorable financial results.

II. LITERATURE REVIEW

The use of machine learning algorithms to predict stock market trends has been more popular recently. The goal of this approach is to take advantage of impending price swings and increase investor profits. Agrawal [25] introduced a stock market forecasting system that utilizes deep learning-based nonlinear regression techniques. Agrawal shows that the suggested method performs better than traditional machine learning techniques by doing experiments on a variety of datasets, including data from the New York Stock Exchange and ten years' worth of Tesla stock price data [25]. This topic of study was significantly advanced by the methodology for media and entertainment company stock price forecasting that Petchiappan et al. [26] developed. Through the utilization of machine-learning techniques, specifically logistic and linear regression, they are able to build a robust prediction system that is customized for the industry. By carefully examining stock price information from reliable media outlets, their approach offers investors important insights into maximizing profits and minimizing losses. Petchiappan et al. [26] perform comprehensive studies to demonstrate the efficacy of their system, emphasizing its advantages over traditional ways. Because stock prices are dynamic and have many facets, forecasting stock market movements is still a difficult and challenging task in the finance industry. To overcome this challenge, Sathyabama et al. [27] use machine learning techniques to predict stock market transactions. The effect that news and other external factors have on stock market patterns is heavily stressed in the authors' research. This further highlights how important accurate prediction models are to effectively managing market volatility. Sathyabama et al. [27] include a better learning-based method that incorporates a Naïve Bayes classifier, adding to the body of information already in existence. Menaka et al. [28] conducted a thorough analysis of machine learning algorithms used in stock price prediction on multiple stock exchanges, which contributed to the field of study in this area. Menaka et al. [28] highlighted how different machine-learning techniques can be tailored to create prediction models that are accurate. These techniques included random forests, ensemble approaches, support vector machines, and boosted decision trees. In order to address the particular challenges posed by sudden and erratic market swings, Demirel et al. [29] focused their analysis on the firms that make up the Istanbul Stock Exchange National 100 Index. Employing daily data collected over a nine-year period, the prediction performance of Long Short-Term Memory, Multilayer Perceptrons, and Support Vector Machines was evaluated [29]. Stock market predictions are still the subject of much research because of the wide-ranging implications they have for global financial markets, investors, and businesses. Tembhurney et al. [30] conducted a comparative analysis of machine learning algorithms' performance in projecting the Nifty 50 stock market index in order to address this challenge. Tembhurney et al. [30] implemented the Random Forest and Support Vector Machine techniques using the Python programming language in order to train models using historical stock market data.

The literature evaluation demonstrates the superiority of machine learning algorithms over conventional methods in forecasting stock market trends. Nonetheless, there are still some significant flaws that exist. Feature engineering is neglected, there is a lack of external validation, the interpretation of models is not thoroughly examined, and the assessment of dynamic market situations is insufficient. Moreover, the evaluation of model hazards is inadequate, and there are few comparisons across various market conditions. Improving the dependability and relevance of machine learning-based stock market prediction models requires addressing these shortcomings. Thus, further research is required to concentrate on developing models that are clear, reliable, and flexible, integrating thorough risk assessment frameworks and able to adjust to shifting market situations. In order to address the shortcomings noted in the literature review, this paper focuses on applying novel techniques, specifically the combination of the moth flame optimization and the radial basis function methodology, to improve the accuracy of stock market predictions. This work aims to create a more flexible, robust, and intuitive stock market prediction model in order to increase the reliability and usefulness of machine learning-driven financial market forecasting.

III. Methodology

A. Radial Basis Function

The use of RBF, a type of mathematical operation, is widespread in numerous fields, such as physics, mathematics, and artificial intelligence. When used as an activation function in artificial neural networks, RBF is commonly applied in machine learning, specifically in the radial basis function networks. By using samples from recent datasets, RBF can learn the correlation between dependent and independent variables similar to other neural networks. The parallel components of RBF consist of neurons, and the network is composed of a single buried layer with numerous neurons. The input layer of the neural network receives independent variables, and the nerve cells in the hidden layer compute the input variables to produce the desired output. The RBF network demonstrates satisfactory generalization capacity when compared to new data sets. As long as it has sufficient neurons, the RBF network is capable of estimating any complex function with the necessary accuracy, making it a reliable estimation function. Due to its fast-processing speed, RBF learning is a viable alternative to Multilayer Perceptron learning. Fig. 1 describes the performance of the RBF.

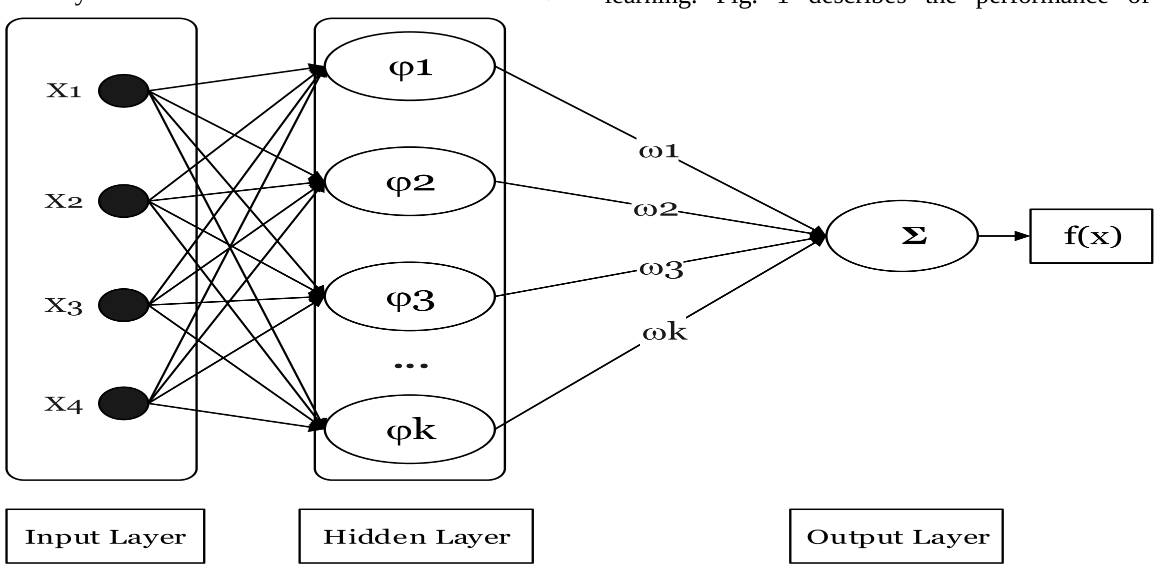

Fig. 1. Description of the performance of the RBF.

Each hidden layer cell is controlled by a non-linear activation equation (φ). The bias component is denoted by constant vector 1 and (φ_0) facilitates the training phase convergence and the restricted reach of the RBF neural network. The research in [31] states that any input vector x may be used to calculate the RBF neural network's output:

$$Y = W^T \Phi = \sum_{j=1}^{L_2} w_{ij} \phi(\|x - c_i\|) \quad (1)$$

When L_2 is the total number of neurons in the hidden layer, c_i is the prototype centers of those neurons, and φ is the Gaussian function, the value of w_{ijv} may be calculated using the following equation:

$$\phi_i(x) = \exp\left(-\frac{\|x-c_i\|^2}{\sigma_i^2}\right) \quad (2)$$

where, σ is the spread parameter. During the training phase of the RBF training, one of the clustering approaches is used to establish the ideal values for the c_i centers, which are initially chosen at random. RBFs are trained and optimized utilizing the Mean Squared Error (MSE) objective function:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(y_d - y_p\right)^2 \quad (3)$$

B. Particle Swarm Optimization

The first inspiration for particle swarm optimization came from studying the social behavior of fish and birds. In continuous and multidimensional environments, this heuristic technique has been successful in tackling optimization and search issues. In the 1990s, James Kennedy and Russell Eberhart developed the PSO technique [18]. Each technique's location inside a D-dimensional search space is considered a potential solution in this method. In accordance with the effect of the determined ideal position and the location of the best-performing particle, each particle modifies its position. The following equation is used by the PSO algorithm to control particle speeds:

$$v_{id}^{t+1} = v_{id}^{t} + C_1 r_1^t \left(Pbest_{id}^t - x_{id}^t\right) + C_2 r_2^t \left(Gbest_{id}^t - x_{id}^t\right) \quad (4)$$

In a d-dimensional search space, v_{id}^k represents the speed of the ith particle at a certain time iteration. For ith individual and iteration t, the ideal particle and location are shown in $Pbest_{id}^t$ and $Gbest_{id}^t$, respectively. The parameters C_1 and C_2 are used to alter particle speed, whilst the numbers r_1^t and r_2^t are arbitrary values between 0 and 1. Additionally, the PSO algorithm's particles travel according to Eq. (5):

$$x_{id}^{t+1} = x_{id}^{t} + v_{id}^{t+1} \quad (5)$$

In this case, x_{id}^t denotes the position of the ith particle in iteration t and in the d dimensional search space.

C. Slime Mold Algorithm

In 2020, Li et al. [19] introduced the SMA, which primarily replicates the behavior and morphological changes of the Physarum polycephalum during foraging. Weights in SMA were used at the same time to imitate the positive and negative feedback created during the slime mold foraging process, resulting in the formation of three distinct morphological forms of slime mold. Slime mold is a eukaryotic creature that lives in a chilly, damp environment. Plasmodium is its major food source. The organic material of slime mold searches for food during the active feeding phase surrounds it and secretes enzymes to break it down. In order to facilitate cytoplasmic flow within, the leading edge of the migration cell moves in sectors, and the trailing end is a network of linked veins. Using a range of food sources, they may concurrently construct linked venous networks based on the characteristics of slime mold. The formula utilized to describe this behavior of the slime mold serves as the foundation for the SMA approach. This strategy may be used in a variety of different sectors.

$$\overrightarrow{X(t+1)} = \begin{cases} \overrightarrow{X_b(t)} + \overrightarrow{v_b} \cdot \left(\overrightarrow{W} \cdot \overrightarrow{X_A(t)} - \overrightarrow{X_B(t)}\right) & r < p \\ \overrightarrow{v_c} \cdot \overrightarrow{X(t)} & r \geq p \end{cases} \quad (6)$$

whereas $X(t)$ and $X(t+1)$ are the locations of the slime mold in repetitions t and $t+1$, respectively, and $X_b(t)$ represents the area of the slime mold with the highest concentration of odor at this specific instant. $X_A(t)$ and X_B display two randomly chosen spots for slime mold and v_b is a variable that changes over time [- a , a]($a = \operatorname{arctanh}(-(\frac{t}{\max\_t}) + 1)$), if v_c is a linearly lowering if r is a random integer between 0 and 1, v_c is a parameter that decreases linearly from 0 to 1, then p is defined as follows:

$$p = tanh|S(i) - DF| \quad i = 1, 2, \dots, n \quad (7)$$

$S(i)$ denotes the fitness of $\overrightarrow{X}$ and DF denotes the iteration that is overall the fittest. The following is a description of the weight W equation:

$$\overrightarrow{W(smell\ index(l))} = \begin{cases} 1 + r.\log\left(\frac{bF-S(i)}{bF-wF} + 1\right), condition \\ 1 - r.\log\left(\frac{bF-S(i)}{bF-wF} + 1\right), others \end{cases} \quad (8)$$

$$smell\ index = sort(S) \quad (9)$$

In this equation, $S(i)$ stands for the first half of the population, bF for best fitness, wF for worst fitness, and the values of the sorted fitness are represented by the scent index. The following equation is used to change the location of the slime mold:

$$\overrightarrow{X^*} = \begin{cases} rand(UB - LB) + LB & rand < z \\ \overrightarrow{X_b(t)} + \overrightarrow{v_b} \cdot \left(\overrightarrow{W} \cdot \overrightarrow{X_A(t)} - \overrightarrow{X_B(t)}\right) & r < p \\ \overrightarrow{v_c} \cdot \overrightarrow{X(t)} & r \geq p \end{cases} \quad (10)$$

where, z is a number between 0 and 0.1 and LB and UB stand for the bottom and upper bounds of the finding interval, respectively.

D. Moth Flame Optimization

Performance improvements for numerous models have been achieved with great success by utilizing the ground-breaking Moth Flame Optimizer, which Fig. 3 demonstrate the general process of this optimizer. This optimizer is motivated by the nighttime behavior of butterflies, which are known to travel towards the direction of light sources. Although this strategy is excellent for traversing large distances, butterflies are in danger of becoming caught in traps as they continuously circle the light source. The MFO method formalizes this movement into a mathematical formula that may be used to solve a wide variety of optimization issues in a variety of industries, including power and energy systems, economic dispatch, engineering design, image processing, and medical applications. Moths use transverse orientation, a unique kind of navigation, to fly directly toward the moon, which allows scientists to examine this behavior. Numerous optimization issues have been successfully handled using this approach. However, MFO struggles with the issue of inadequate exploration [21]. Fig. 2 shows the general organization of the MFO function and Fig. 3 shows the flowchart of the MFO algorithm,

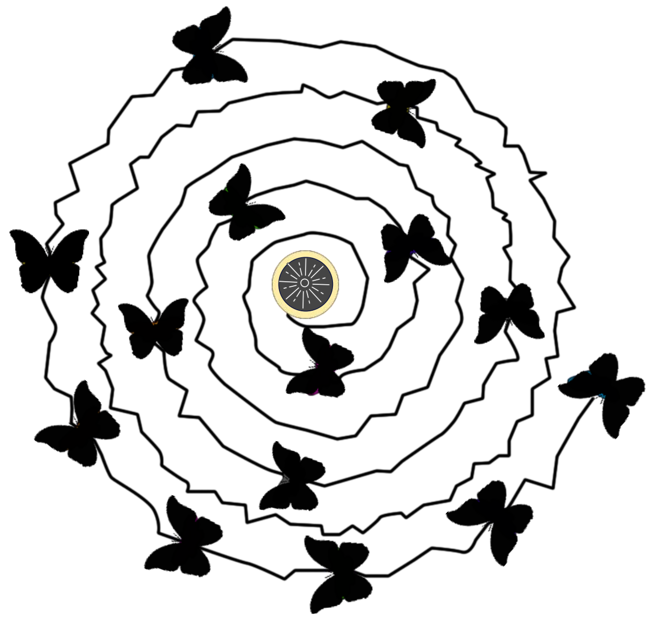

Fig. 2. Flying in a spiral pattern to avoid nearby light sources.

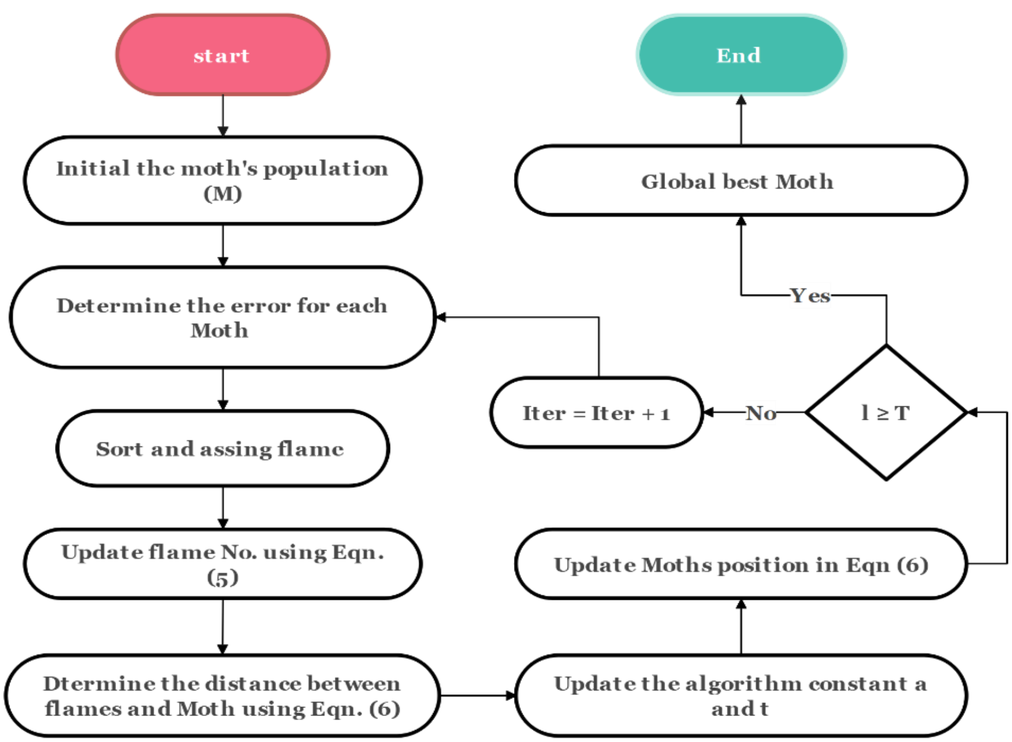

Fig. 3. Flowchart of the MFO algorithm.

Moths are the investigation's possible answers, and the problem aspects are their geographic distributions. By altering their position vectors, moths may fly in 1D, 2D, 3D, or hyper-dimensional space. The suggested method ensures convergence, and MFO is dependable and computationally effective. MFO may be written as follows:

$$M = \begin{bmatrix} \mathrm{CO}_{1,1} & \mathrm{CO}_{1,2} & \cdots & \cdots & \mathrm{CO}_{1,h} \\ \mathrm{CO}_{2,1} & \mathrm{CO}_{2,2} & \cdots & \cdots & \mathrm{CO}_{2,h} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathrm{CO}_{a,1} & \mathrm{CO}_{a,2} & \cdots & \cdots & \mathrm{CO}_{n,h} \end{bmatrix} \quad (11)$$

where h is the number of dimensions and a is the number of moths.

$$S = \begin{bmatrix} S_{1,1} & S_{1,2} & \cdots & \cdots & S_{1,h} \\ S_{2,1} & S_{2,2} & \cdots & \cdots & S_{2,h} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ S_{a,1} & S_{a,2} & \cdots & \cdots & S_{2,h} \end{bmatrix} \quad (12)$$

Worldwide optimization is carried out using the three-step MFO method.

$$MFO = (I, F, T) \quad (13)$$

Where I is a function, F is the flight of the moth in search of space, and T is the stopping criteria.

$$X_i = t(C_i, S_j) \quad (14)$$

S_j indicates the number of jth flames, where C_i is the number of ith moths, where t is the twisting function, which has the following expression:

$$S(C_i, S_j) = Z_i \cdot e^{bt} \cdot \cos(2\pi t) + S_j \quad (15)$$

Where $Z_i =$ separating the moth from the flame, $b =$ constant value, and $t =$ random number between $[-1,1]$.

$$\mathrm{Zi} = \left|\mathrm{S_j} - \mathrm{X_i}\right| \quad (16)$$

E. Dataset Description

The dataset used in this study is intended to make it possible to forecast Google's stock share values over a wide time range, from January 1, 2015, to mid-2023. For investors, financial experts, and decision-makers in the finance sector, accurate stock price forecasting is crucial. The historical stock price information and associated attributes required for conducting predictive studies are provided in this dataset. The dataset's primary sources of financial market data include stock exchanges and financial news outlets. Google's (Alphabet Inc.) historical daily stock share values for the specified time period were compiled. For each trading day between January 1, 2015, and mid-2023, there are several pieces of information available, which are the variables of this paper's dataset, about Google's stock shares. These include the date, the opening price at the start of the trading day, the closing price at the end of the trading day, the highest price the shares reached during the day, the lowest price the shares reached during the day, and the trading volume which represents the total number of shares traded during the day. To guarantee data quality and consistency, stringent data pretreatment processes were used before performing any predictive analyses. Data normalization was also carried out to help with accurate modeling and forecasting. Through the process of data normalization, numerical variables are scaled to a common range, usually between 0 and 1, or with a mean of 0 and a standard deviation of 1. In analytical or modeling activities, this guarantees that variables with varied units or magnitudes are treated equally. The size of the input variables affects the performance of many machines learning techniques, including support vector machines and k-nearest neighbors. These algorithms' performance and convergence may be enhanced by normalizing the data.

Feature scaling, often called Min-Max normalization or data preparation, is the process of rescaling numerical properties in a dataset to a specific range, frequently from zero to one. The objective is to maintain the relative relationships between the values while bringing all the features to a similar scale. In machine learning algorithms that are sensitive to the quantity of input features, this can be particularly crucial. The data normalization approach's formula is as follows:

$$\text{XScaled} = \frac{(X - Xmin)}{(Xmax - Xmin)} \quad (17)$$

A common method for evaluating how well a machine learning model can handle new and untested data is through data splitting. By training the model on a portion of the data and testing it on another subset, the assessment of its performance on real-world data can be revealed by splitting data to train and test. This approach enables us to determine if the model has truly learned from the data and identified patterns or if it is simply recalling information from the training set. Fig. 4 shows that the data set was divided into two parts: the train set and the test set.

The statistical results of the obtained data are shown in Table I. When describing the features of a data collection, statistical measurements like mean, median, skewness, standard deviation, maximum, and minimum are utilized.

By adding up all the values in a data collection and dividing by the total number of values, the mean, sometimes referred to as the average, is determined. When data collection is arranged from lowest to highest, the median is the midway value. The median is the average of the two middle values when there is an even number of values. Comparing to the mean, the median is less impacted by outliers or extreme numbers. A measure of the asymmetry in the data distribution is called skewness. It shows whether the data is roughly symmetric, positively skewed to the right, or negatively skewed to the left. A symmetric distribution is indicated by a skewness value of 0. The distribution or dispersion of data points around the mean is measured by the standard deviation. A higher standard deviation indicates that the data are more variable. It is described mathematically as the square root of the variance. The maximum value among all the data points is simply the maximum value in a data collection. The minimum value among all the data points is the minimum value in a data collection.

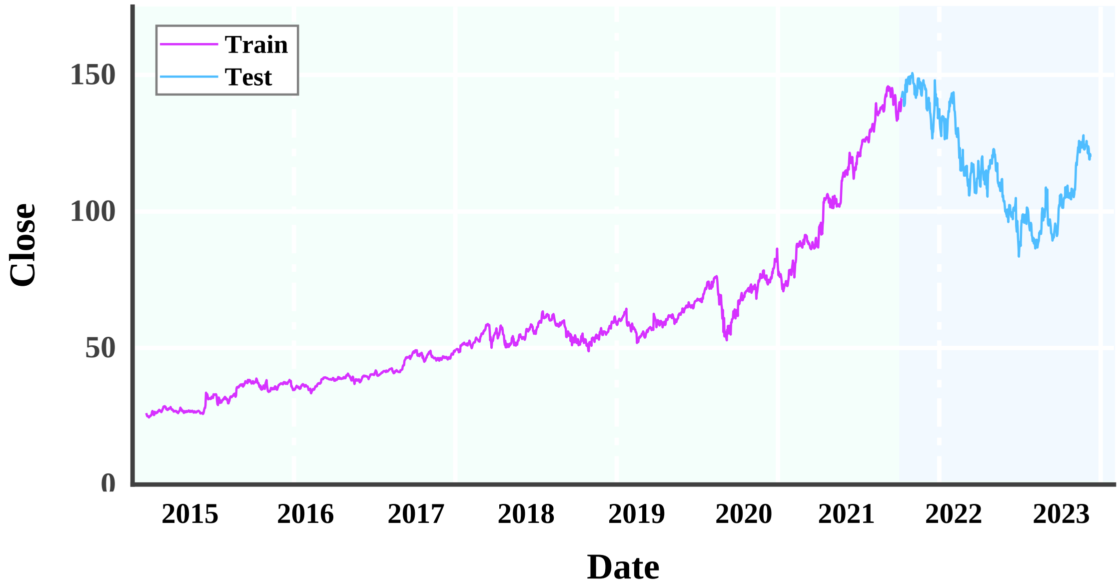

Fig. 4. The overall illustration of the dataset during the training and test.

TABLE I. STATISTICAL RESULTS OF THE PRESENTED MODELS FOR OHCLV

	Open	High	Low	Volume	Close
count	2137	2137	2137	2137	2137
mean	70.05219	70.81457	69.3428	32.59751	70.09629
Std.	34.54605	34.97686	34.14654	15.6062	34.55914
min	24.66478	24.7309	24.31125	6.936	24.56007
25%	41.0205	41.22	40.851	23.248	41.046
75%	96.77	98.94	95.38	37.066	96.73
max	151.8635	152.1	149.8875	223.298	150.709
skew	0.746243	0.736992	0.747426	2.879365	0.741179

F. Evaluation Metrics

For evaluating the performance and efficacy of models, algorithms, and data-driven solutions across a variety of areas, from machine learning and data science to business analytics, evaluation metrics are crucial tools. These metrics offer measurable indicators of how successfully a model or approach completes the goal for which it was designed. The criteria used in presenting this research are MAE, $RMSE$, MSE and R^2. The average absolute difference between anticipated and actual values is calculated using MAE. It offers a simple way to assess prediction errors. The average squared difference between expected and actual values is determined by MSE. More so than MAE, it penalizes significant mistakes. The square root of MSE, or RMSE, offers a measure that can be understood and is expressed in the same units as the objective variable. R-squared measures the percentage of the target variable's variation that the model is responsible for explaining. From 0 (no explanation) to 1 (excellent explanation), it has a range.

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

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

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

$$MSE = \frac{1}{N}\sum_{k=0}^{n}\binom{n}{k}(Fi - Yi)b^2 \quad (21)$$

IV. RESULT AND DISCUSSION

A. Comparative Analysis

In order to successfully forecast the Alphabet stock, an identical dataset was applied to each model. The results of each model were thoroughly analyzed and evaluated for this article in order to present a thorough and instructive comparison of their performance. To establish an accurate and fair comparison, it is essential to define the performance metrics that were applied to assess the models. Evaluating the models using a variety of crucial criteria, as explained in the method section. It is possible to thoroughly assess the performance of each model using a variety of metrics before determining which one best meets the requirements. A thorough Table II with the results displays all the various nuances of how each model performed.

Prior to choosing the RBF model, the obtained result was taken into consideration. After a thorough study of the data, the RBF model was selected because of its higher performance. The Alphabet Inc. index data underwent the process of selecting relevant data and normalizing it from the beginning of 2015 to the middle of 2023. Through this rigorous method, valuable insights will be extracted that will aid in the decision-making process. Due to the problematic optimizer developments, the assessment result for RBF alone is now 0.985 in R^2, as indicated in Table II. The R^2 criteria values for the PSO, SMA, and MFO are 0990, 0.991, and 0.995, respectively, indicating that the optimum course of action may be picked. When compared to other optimizers, the MFO optimizer produces better results. The RMSE model findings shown in Table II further highlight the MFO optimizer's superiority. The $RMSEs$ for RBF, PSO-RBF, SMA-RBF, and MFO-RBF are 2.238, 1.809, 1.710, and 1.253, respectively.

In Fig. 5 through Fig. 6, the experiment's findings are shown, and they show a significant connection between the model and the real data. The MFO-RBF model performed better than the individual RBF, PSO-RBF, and SMA-RBF models among the evaluated models. Notably, the performance of the RBF model was greatly enhanced by the use of the optimizer approach. Fig. 7 and Fig. 8 provide a thorough study of the four models, demonstrating that the chosen model is capable of yielding the best outcomes. These results indicate that the MFO-RBF model is a potentially useful method for precisely forecasting the intended outcomes in the context.

The research presented in this paper demonstrates a higher level of predictive accuracy compared to the studies cited previously [32] [33] , as evidenced by the R^2 value of 0.995 provided in Table III.

TABLE II. THE RESULTS OF EVALUATION CRITERIA FOR THE OPTIMIZED MODEL

MODEL/Metrics	TRAIN SET				TEST SET			
	R2	*RMSE*	*MAE*	*MSE*	*R2*	*RMSE*	*MAE*	*MSE*
RBF	0.988	3.002	1.367	9.012	0.985	2.238	1.899	5.010
PSO-RBF	0.993	2.242	1.642	5.027	0.990	1.809	1.560	3.273
SMA-RBF	0.995	1.865	1.424	3.477	0.991	1.710	1.295	2.923
MFO-RBF	0.998	1.237	0.919	1.530	0.995	1.295	1.009	1.676

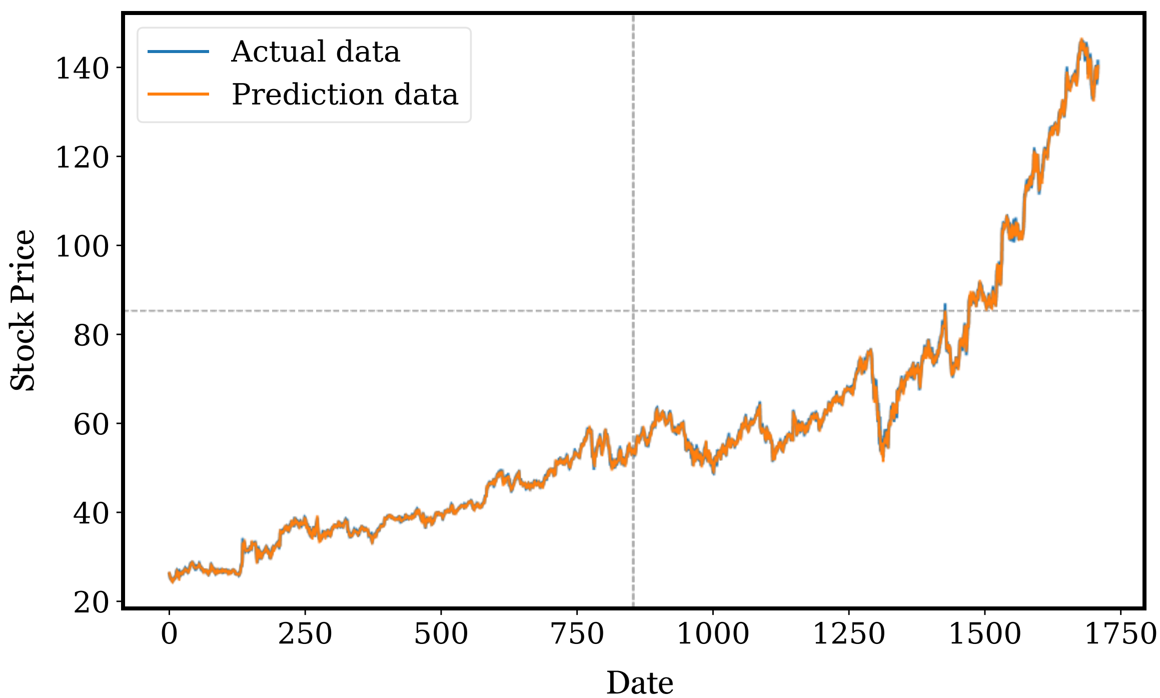

Fig. 5. Assessment of the suggested model's performance in comparison to other models during training.

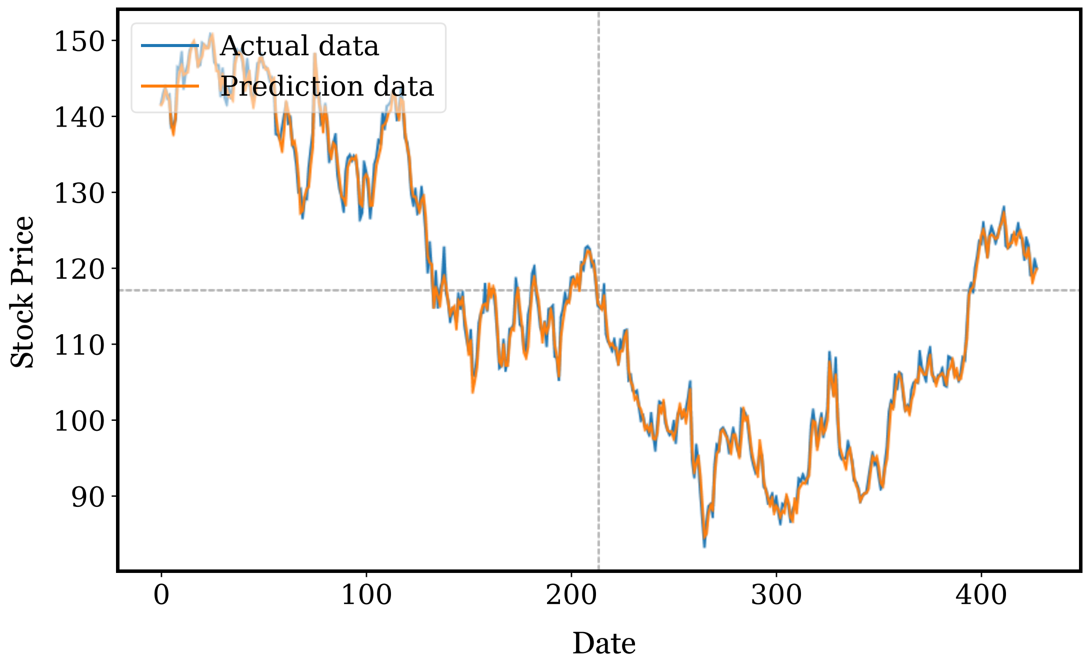

Fig. 6. Assessment of the suggested model's performance in comparison to other models during testing.

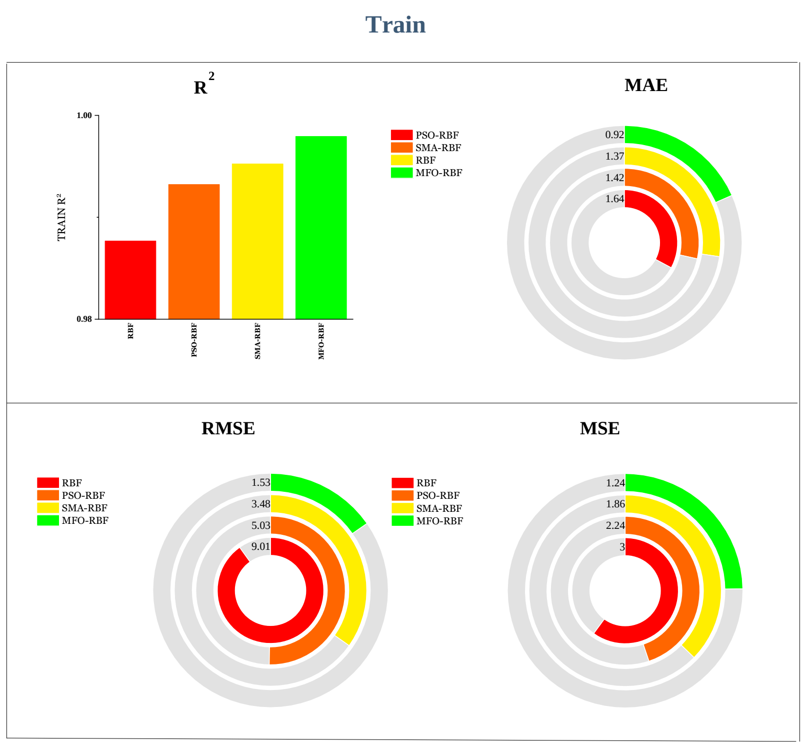

Fig. 7. Result of the Evaluation metrics for the presented models during training.

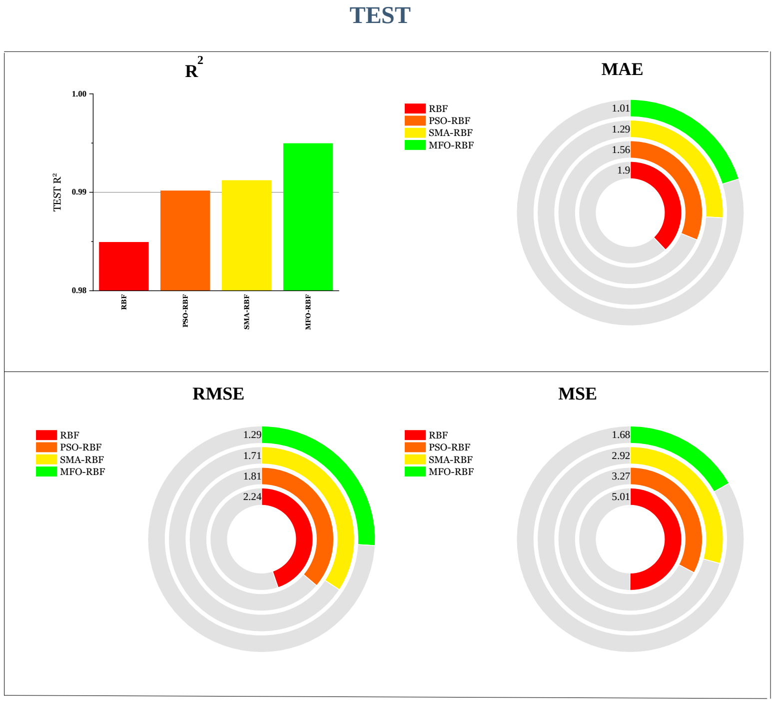

Fig. 8. Result of the Evaluation metrics for the presented models during the test.

TABLE III. AN ASSESSMENT OF THE MODEL IS PROVIDED IN RELATION TO PREVIOUS INVESTIGATIONS

References	Methods	R^2
[32]	DNN and LSTM	0.972
[33]	LSTM	0.981
Present invistigation		0.995

V. CONCLUSION

By leveraging stock prediction techniques to evaluate asset values and identify prevailing market trends, both individual and institutional investors have the opportunity to gain a significant competitive advantage. This allows investors to make well-informed decisions on whether to buy, sell, or hold stocks, utilizing historical data and advanced algorithms. Such a strategy is vital for investors committed to making prudent investment choices, as it mitigates risks and increases the likelihood of achieving profitable outcomes. This research employed various predictive algorithms and data sources to delve into the complex and ever-changing realm of stock prediction. These findings suggest that a combination of models or an ensemble approach may offer more accurate forecasts. Importantly, the development and evaluation of these prediction models underscore the importance of relying on data-driven insights to make reliable decisions. This underscores the benefits of a data-centric approach in today's rapidly evolving business landscape and the broad applicability of predictive analytics across various industries. The primary objective of this study was to create models that could better predict stock prices, enabling interested traders and investors to use these algorithms to make well-timed and cost-effective purchases.

These conclusions were reached in this paper:

First, the data preparation and normalization process were finished, which could have an impact on how the prediction model is displayed. The steps that the selected model would take to examine the data were then prepared for use.

To increase the effectiveness of the model that has been presented, the suitable model should be chosen, the results evaluated, and then the hyperparameters of the model should be adjusted.

By contrasting the outcomes of various optimizers, the most accurate optimization has been determined as the main optimizer of the model. The MFO approach yields the best results when compared to RBF, PSO-RBF, and SMA-RBF, whose results for R^2 evaluation criteria are 0.985, 0.990, and 0.991, respectively.

For the purpose of training and validating the model, the suggested method heavily depends on historical stock price data. The model's ability to predict future market behaviors may be intrinsically limited by its dependence on historical trends, even though it offers a solid foundation. This is especially true in situations where there are unanticipated events or market disruptions. When faced with unusual market dynamics that are not represented in historical data, or during times of increased volatility, the model's effectiveness may be called into question even with the use of sophisticated optimization techniques designed to strengthen the model's flexibility to changing market conditions. Moreover, the model's complexity is increased by combining various optimization methods with the Radical Basis Function. Due to this increased complexity, it is possible that scalability and practical implementation in real-time trading environments will be hampered during the training and evaluation phases, which will require significant computational resources. Additionally, it's important to recognize that the efficacy of the suggested methodology might vary among various financial markets or asset classes, going beyond the parameters of the research. The extent to which the model's predictions can be applied to other markets is largely dependent on variables like investor behavior, regulatory frameworks, and market structure. Furthermore, because the model is complex and combines a variety of optimization methods, there is a greater chance that the training data will be overfit or that the test set will be accidentally incorporated into the model. To reduce these inherent risks and guarantee the validity and reliability of the model, it is therefore essential to use appropriate regularization techniques and strong cross-validation strategies.

Creating methods to make complicated models easier to understand has the potential to reveal important information about the fundamental causes of stock price forecasts. Investor decision-making can be made more informed by providing clear and understandable explanations of model predictions. This builds trust. Furthermore, researching ways to dynamically modify the model's architecture or parameters in response to current market conditions could greatly improve the model's accuracy and robustness, especially in unstable or changing market environments. This project might involve investigating ensemble methods or adaptive learning algorithms that can modify model structures or weights in response to changing market conditions. Additionally, there is a chance to improve the model's predictive power and strengthen its resistance to market swings by investigating non-conventional data sources like news articles, social media sentiment, and macroeconomic indicators. The suggested approach's long-term performance and stability across different market cycles would be assessed through longitudinal research, which would be crucial in revealing important details about its dependability and efficacy as a forecasting tool for investors. These kinds of studies would provide a thorough grasp of the model's long-term performance, illuminating its effectiveness in various market scenarios and its potential as a long-term investment tool.

Funding

YB202215 Key Scientific Research Project of Jingchu University of Technology.

JX2021-005 Key Teaching and Research Project at the School Level of Jingchu University of Technology.

School-level Research Platform of Jingchu University of Technology:Data Analysis Science Laboratory.

HX20220171 Horizontal Scientific Research Project of Jingchu University of Technology.

References

[1] Y.-H. Wang, C.-H. Yeh, H.-W. V. Young, K. Hu, and M.-T. Lo, "On the computational complexity of the empirical mode decomposition algorithm," Physica A: Statistical Mechanics and its Applications, vol. 400, pp. 159–167, 2014, doi: https://doi.org/10.1016/j.physa.2014.01.020.

[2] S. Claessens, J. Frost, G. Turner, and F. Zhu, "Fintech credit markets around the world: size, drivers and policy issues," BIS Quarterly Review September, 2018.

[3] W. Li et al., "The nexus between COVID-19 fear and stock market volatility," Economic research-Ekonomska istraživanja, vol. 35, no. 1, pp. 1765–1785, 2022.

[4] Z. Wang et al., "Measuring systemic risk contribution of global stock markets: A dynamic tail risk network approach," International Review of Financial Analysis, vol. 84, p. 102361, 2022.

[5] Z. Li, W. Cheng, Y. Chen, H. Chen, and W. Wang, "Interpretable click-through rate prediction through hierarchical attention," in Proceedings of the 13th International Conference on Web Search and Data Mining, 2020, pp. 313–321.

[6] R. Bisoi, P. K. Dash, and A. K. Parida, "Hybrid Variational Mode Decomposition and evolutionary robust kernel extreme learning machine for stock price and movement prediction on daily basis," Appl Soft Comput, vol. 74, pp. 652–678, 2019, doi: https://doi.org/10.1016/j.asoc.2018.11.008.

[7] M. Zounemat-kermani, O. Kisi, and T. Rajaee, "Performance of radial basis and LM-feed forward artificial neural networks for predicting daily watershed runoff," Appl Soft Comput, vol. 13, no. 12, pp. 4633–4644, 2013, doi: https://doi.org/10.1016/j.asoc.2013.07.007.

[8] P. McCullagh, "What is a statistical model?," The Annals of Statistics, vol. 30, no. 5, pp. 1225–1310, 2002.

[9] E. Chollet Ramampiandra, A. Scheidegger, J. Wydler, and N. Schuwirth, "A comparison of machine learning and statistical species distribution models: Quantifying overfitting supports model interpretation," Ecol Modell, vol. 481, no. February, 2023, doi: 10.1016/j.ecolmodel.2023.110353.

[10] S. B. Kotsiantis, "Decision trees: a recent overview," Artif Intell Rev, vol. 39, pp. 261–283, 2013.

[11] L. Breiman, "Random forests," Mach Learn, vol. 45, pp. 5–32, 2001.

[12] S. Haykin, Neural networks and learning machines, 3/E. Pearson Education India, 2009.

[13] M. A. Hearst, S. T. Dumais, E. Osuna, J. Platt, and B. Scholkopf, "Support vector machines," IEEE Intelligent Systems and their applications, vol. 13, no. 4, pp. 18–28, 1998.

[14] E. S. Olivas, J. D. M. Guerrero, M. Martinez-Sober, J. R. Magdalena-Benedito, and L. Serrano, Handbook of research on machine learning applications and trends: Algorithms, methods, and techniques: Algorithms, methods, and techniques. IGI global, 2009.

[15] B. Mahesh, "Machine learning algorithms-a review," International Journal of Science and Research (IJSR).[Internet], vol. 9, no. 1, pp. 381–386, 2020.

[16] M. D. Buhmann, "Radial basis functions," Acta Numerica, vol. 9, pp. 1–38, 2000, doi: 10.1017/S0962492900000015.

[17] G. S. Fesaghandis, A. Pooya, M. Kazemi, and Z. N. Azimi, "Comparison of multilayer perceptron and radial basis function in predicting success of new product development," Eng. Technol. Appl. Sci. Res., vol. 7, 2017.

[18] J. Kennedy and R. Eberhart, "Particle swarm optimization," in Proceedings of ICNN'95-international conference on neural networks, IEEE, 1995, pp. 1942–1948.

[19] 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, doi: https://doi.org/10.1016/j.future.2020.03.055.

[20] S. Mirjalili, "Moth-flame optimization algorithm: A novel nature-inspired heuristic paradigm," Knowl Based Syst, vol. 89, pp. 228–249, 2015, doi: https://doi.org/10.1016/j.knosys.2015.07.006.

[21] K. Kaur, U. Singh, and R. Salgotra, "An enhanced moth flame optimization," Neural Comput Appl, vol. 32, no. 7, pp. 2315–2349, 2020, doi: 10.1007/s00521-018-3821-6.

[22] O. Avatefipour et al., "An intelligent secured framework for cyberattack detection in electric vehicles' CAN bus using machine learning," IEEE Access, vol. 7, pp. 127580–127592, 2019.

[23] F. Mirzapour, M. Lakzaei, G. Varamini, M. Teimourian, and N. Ghadimi, "A new prediction model of battery and wind-solar output in hybrid power system," J Ambient Intell Humaniz Comput, vol. 10, no. 1, pp. 77–87, 2019, doi: 10.1007/s12652-017-0600-7.

[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, no. 14, pp. 9859–9884, 2020, doi: 10.1007/s00521-019-04570-6.

[25] S. C. Agrawal, "Deep learning based non-linear regression for Stock Prediction," IOP Conference Series: Materials Science and Engineering ; volume 1116, issue 1, page 012189 ; ISSN 1757-8981 1757-899X, 2021, doi: 10.1088/1757-899x/1116/1/012189.

[26] M. Petchiappan and J. Aravindhen, "Comparative Study of Machine Learning Algorithms towards Predictive Analytics," Recent Advances in Computer Science and Communications ; volume 16, issue 6 ; ISSN 2666-2558, 2023, doi: 10.2174/2666255816666220623160821.

[27] S. Sathyabama, S. C. Stemina, T. SumithraDevi, and N. Yasini, "Intelligent Monitoring and Forecasting Using Machine Learning Techniques," Journal of Physics: Conference Series ; volume 1916, issue 1, page 012175 ; ISSN 1742-6588 1742-6596, 2021, doi: 10.1088/1742-6596/1916/1/012175.

[28] A. Menaka, V. Raghu, B. J. Dhanush, M. Devaraju, and M. A. Kumar, "Stock Market Trend Prediction Using Hybrid Machine Learning Algorithms," International Journal of Recent Advances in Multidisciplinary Topics; Vol. 2 No. 4 (2021); 82-84 ; 2582-7839, Feb. 2021, [Online]. Available: https://journals.ijramt.com/index.php/ijramt/article/view/643

[29] U. Demirel, H. Cam, and R. Unlu, "Predicting Stock Prices Using Machine Learning Methods and Deep Learning Algorithms: The Sample of the Istanbul Stock Exchange," 2021, [Online]. Available: https://hdl.handle.net/20.500.12440/3191

[30] P. M. Tembhurney and S. Pise, "Stack Market Prediction Using Machine Learning (ML) Algorithms," International Journal for Indian Science and Research Volume-1(Issue -1) 08, Feb. 2022, [Online]. Available: https://zenodo.org/record/6787069

[31] M. Taki, A. Rohani, F. Soheili-Fard, and A. Abdeshahi, "Assessment of energy consumption and modeling of output energy for wheat production by neural network (MLP and RBF) and Gaussian process regression (GPR) models," J Clean Prod, vol. 172, pp. 3028–3041, 2018, doi: https://doi.org/10.1016/j.jclepro.2017.11.107.

[32] 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.

[33] 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.