Introducing an Innovative Approach to Mitigate Investment Risk in Financial Markets: A Case Study of Nikkei 225

Xiao Duan
Taizhou Vocational College of Science and Technology, Taizhou Zhejiang, 318020, China

***Abstract*—When the value of an investor's stock portfolio rises during a period of great market performance, investors often experience a gain in wealth. Spending may increase when people feel more at ease and confident about their financial circumstances. On the other hand, during a market crisis, a fall in wealth could lead to lower consumer spending, which could impede economic growth. Stock market trend prediction is thought to be a more important and fruitful endeavor. Stock prices will, therefore, provide significant returns from prudent investing decisions. Because of the outdated and erratic data, stock market forecasts pose a serious challenge to investors. As a result, stock market forecasting is among the main challenges faced by investors trying to optimize their return on investment. The goal of this research is to provide an accurate hybrid stock price forecasting model using Nikkei 225 index data from 2013 to 2022. The construction of the support vector regression involves the integration of multiple optimization approaches, including moth flame optimization, artificial bee colony, and genetic algorithms. Moth flame optimization is proven to produce the best results out of all of these optimization techniques. The evaluation criteria used in this research are MAE, MAPE, MSE, and RMSE. The results obtained for MFO-SVR, which is 0.70 for criterion MAPE, show the high accuracy of this model for estimating the price of Nikkei 225.**

***Keywords*—NIKKEI 225 index; artificial bee colony; stock price; financial markets; support vector regression**

I. Introduction

A. Background Knowledge

The global stock market is a burgeoning industry in every nation. This industry directly affects a large number of individuals. Therefore, those individuals must learn about the current market trend. With the growth of the stock market, people's interest in stock price forecasting has increased. Trend forecasting has become a crucial subject for investors, shareholders, and other authorities involved in the stock market industry. Stock price prediction is thought to be an arduous endeavor [1]. Due to the fact that stock markets are fundamentally a noisy, non-parametric, non-linear, and in deterministically chaotic system [2][3]. Market trends are affected by a multitude of variables, including equities, liquid funds, consumer behavior, and stock market news. Collectively, these factors govern how stock market trends behave. Tools for technology and parametric pricing approaches, or a mix of these, can be used to study trend behavior [4][5]. To lessen any possible risks, it is crucial to develop a strong and convincing prediction model. There are several hypotheses explaining why stock markets are unpredictable. Conventional approaches to trend prediction are based on patterns that do not change over time. This method ignores the stock market's volatility, which makes predicting stock prices difficult given the myriad factors at play. However, the development of machine learning (ML) [6][7] has offered a solid remedy that uses a variety of algorithms to improve performance in certain situations. It's an exciting breakthrough that might completely change the way we make stock market forecasts. Many people believe that ML is capable of identifying trustworthy data and identifying patterns within the dataset [8]. The majority of traditional time series prediction techniques rely on static patterns, which makes predicting stock prices difficult by nature. Furthermore, forecasting the price of stocks is a challenging endeavor in and of itself due to the sheer number of influencing factors. Longer-term market behavior is more like a weighing machine than a voting machine, making it possible to predict changes in the market over extended periods [9].

Artificial neural networks (ANNs) are a frequently used and beneficial model for many different sectors, with applications ranging from classification and grouping to pattern recognition and prediction. The overall usefulness of an ANN may be evaluated by utilizing metrics related to data analysis, including volume, scalability, convergence, fault tolerance, accuracy, processing speed, latency, and performance. [10][11]. One of the artificial neural networks' main potentials is its ability to process data rapidly in a massively parallel implementation; this has generated interest and raised demand for study in the subject [12]. Natural language processing, picture recognition, and other uses can be facilitated by the development and deployment of ANNs. This method was used to investigate breast cancer detection by Mahan et al. [13]. Qihao Weng et al. [14] compute impervious surfaces using a medium-sized geographic dataset by applying this technique. Ecological modeling is another area in which this technique is utilized. Lek, Sovan, and others, 2012 [15]. Support vector regression (SVR) is a well-known method in the machine learning field and has been regarded as a reliable substitute for outlier detection and a means of reducing overfitting in the setting of linear regression. The approach employed in this work is called SVR; it is a potent supervised learning strategy that lowers the confidence range of training samples while simultaneously minimizing structural and empirical risk. Solving complicated nonlinear issues, especially with limited sample sizes, is made much easier with this method. SVR contributes to the ability to anticipate and evaluate future samples with accuracy, increasing decision-making processes and offering significant insights by optimizing generalization performance while lowering risk [16]. For success and best outcomes, it is vital to comprehend the concepts and advantages of SVR, regardless of your field of expertise—data science, machine learning, or another similar one [17]. The procedure is called optimization determining which potential solution, if any, is optimum for a given issue.

The need for novel optimization procedures has become more evident over the past several decades as issues have grown more complex. Prior to the advent of heuristic optimization methods, the sole tools for problem optimization were mathematical optimization methods. The main problem with most mathematical optimization methods that are deterministic is the trapping of regional optimum. A few of them also need the derivation of the search space, such as gradient-based algorithms. As such, they are utterly useless for fixing real problems. Recent work has introduced Moth Flame Optimization (MFO) [18] to resolve worldwide optimization issues and practical applications. How well MFO works in terms of convergence rate and population variety has previously been shown while dealing with difficult challenges. Based on the MFO, this study suggests a novel method for data clustering. The MFO has several competitive advantages, which are utilized in this work. These benefits include its ability to avoid local optima and its rapid convergence to the global optimal solution. Our main objective is to employ MFO to identify the data items into clusters more precisely and to cover the search space more thoroughly than the existing methods can.

B. *Research Gaps, Contributions, and Novelties*

Research gaps include a lack of comprehension of the dataset characteristics that contribute to the outperformance of the algorithms, an examination of the generalizability of sector-specific methodologies, an investigation into the most effective methods for integrating external variables, an assessment of algorithm performance across different market conditions, and broader geographic comparisons to determine the efficacy of the algorithms. By integrating multiple optimization approaches—including moth flame optimization, artificial bee colony, and genetic algorithms—the MFO-SVR model investigated the efficacy of hybrid methodologies for improving prediction accuracy, thereby filling a number of gaps in the literature on stock market prediction. By employing Nikkei 225 index data spanning from 2013 to 2022, the issue of dealing with obsolete and inconsistent data is effectively resolved, thereby guaranteeing the prediction model's pertinence and precision. Moreover, the implementation of comprehensive evaluation metrics offers a uniform evaluation of model performance, thereby augmenting the body of literature's demand for standardized assessment techniques. The efficacy of the MFO-SVR model in forecasting Nikkei 225 prices is evidenced by its reported high accuracy. This provides investors with valuable insights that can assist them in maximizing their investment returns and bridging the gap in accurate stock market prediction. The study presents an innovative hybrid model for predicting stock prices that combines SVR optimized by MFO with additional optimization techniques, including GA and ABC. By utilizing this hybrid methodology, the Nikkei 225 index stock price forecasts are rendered more precise, thereby mitigating the issue of obsolete and inconsistent data that is commonly encountered in stock market prediction. This research makes a valuable contribution to the extensive empirical implementation of SVR networks in the prediction of financial time series. Through an examination and comparison of several optimization methodologies (GA, ABC, and MFO), the study establishes the efficacy of SVR in forecasting financial markets spanning a substantial duration from 2013 to 2022. This empirical analysis contributes significantly to the body of knowledge regarding financial forecasting. The study employs stock price data from the Nikkei 225 index, which covers the time period from January 1, 2013 to January 1, 2023. This extensive dataset is utilized to train and test the forecasting models. Implementing standardized evaluation criteria guarantees a rigorous evaluation of the performance of the model. Section II of the study represents the literature review. Methodology is given in Section III. Result and discussion are demonstrated in the Section IV. The conclusion is given in Section V.

II. Literature Review

There has been an increasing inclination in recent times to utilize machine learning algorithms for the purpose of forecasting stock market trends, with the objective of leveraging forthcoming price fluctuations and augmenting investor profitability. Agrawal proposed of a stock market prediction system that employs non-linear regression techniques based on deep learning [19]. By conducting experiments on a variety of datasets, such as ten years' worth of Tesla stock price and New York Stock Exchange data, Agrawal establishes that the proposed method outperforms conventional machine learning techniques [19]. Petchiappan et al. [20] made a substantial contribution to this discussion through their novel methodology for forecasting the stock prices of media and entertainment companies. By utilizing machine-learning methodologies, particularly logistic and linear regression, they construct a resilient prediction system that is customized to the needs of this industry. Through the examination of media stock price data, their model provides investors with valuable insights on how to optimize profits and mitigate losses. By means of meticulous experimentation, Petchiappan et al. [20] establish the effectiveness of their methodology, emphasizing its superiority in comparison to conventional approaches. Predicting stock market movements remains an enduring and complex task within the field of finance, owing to the ever-changing and multifaceted characteristics of stock prices. Sathyabama et al. [21] tackle this obstacle by employing machine learning algorithms for the purpose of forecasting stock market transactions. The research conducted by the authors highlights the importance of external variables, including news, in shaping stock market trends. Additionally, it emphasizes the criticality of precise prediction models in order to successfully navigate market volatility. In their contribution to this discussion, Sathyabama et al. [21] present an improved learning-based approach that makes use of a Naïve Bayes classifier. Menaka et al. [22] made a scholarly contribution to this domain through the provision of an exhaustive examination of machine learning algorithms that are employed to forecast stock prices on a variety of stock exchanges. Menaka et al. [22] emphasized the adaptability of various machine learning methodologies—such as ensemble methods, support vector machines, random forests, and boosted decision trees—when constructing accurate prediction models. Demirel et al. [23] tackled the distinct obstacles presented by abrupt and uncertain market conditions by concentrating on companies that are included in the Istanbul Stock Exchange National 100 Index. Utilizing nine years of daily data, their study assessed the performance of Multilayer Perceptrons, Support Vector Machines, and Long Short-Term Memory models in predicting opening and closing stock prices [23]. Tembhurney et al. [24] tackled this obstacle by comparing the performance of machine learning algorithms in the context of Nifty 50 stock market index forecasting. Tembhurney et al. [24] employed the Python programming language to execute the Support Vector Machine and Random Forest algorithms for the purpose of training models with historical stock market data. The literature review offers a thorough examination of diverse approaches utilized in the prediction of stock market trends. It emphasizes the significance of employing machine learning algorithms to anticipate patterns and optimize investment choices. However, there are a number of gaps that can be identified. To begin with, although the review examines the efficacy of various algorithms and methodologies, it does not present a cohesive framework or conduct a comparative analysis of these approaches across diverse datasets or market conditions. Furthermore, there is a dearth of discourse regarding the integration of extraneous variables, including geopolitical events, economic indicators, and news sentiment, into predictive models. Such an expansion would substantially bolster the precision and resilience of such models. Moreover, greater emphasis must be placed on the performance and implementation of these predictive models in the real world, as well as their influence on tangible investment strategies and results. Ultimately, insufficient emphasis is placed on mitigating the difficulties presented by obsolete and inconsistent data, a critical factor in establishing the dependability and efficacy of stock market predictions. This study presents an innovative hybrid stock price prediction model that incorporates various optimization techniques—including moth flame optimization, artificial bee colony, and genetic algorithms—and utilizes Nikkei 225 index data. Through the integration of these optimization methodologies, the objective of this model is to enhance the precision of stock market forecasts, thus mitigating the issue presented by obsolete and volatile data. Furthermore, this research underscores the significance of practical implementation through the assessment of the MFO-SVR model's performance using evaluation metrics including MAE, MAPE, MSE, and RMSE. By adopting this methodology, one can guarantee that the predictive model not only possesses sound theoretical foundations but also effectively directs investment decisions in practice.

III. Methodology

A. Dataset Description

A thorough dataset analysis takes into account the volume of transactions in addition to the open, high, low, and closing (OHLC) prices during a certain period. In order to make this analysis easier, information from 2013 to 2022 was gathered. A thorough data-cleaning process was used to guarantee the accuracy and consistency of the forecasting models. This multi-phase approach was put into place with the intention of protecting the integrity of the dataset and reducing the possibility of problems due to incomplete or erroneous data. A great deal of work was put into carefully examining the data to look for unusual trends, high or low numbers, or discrepancies that might undermine the reliability of the conclusions. The data was then put through a number of procedures to make sure it was clean and ready for processing. Methods like normalization were used to reduce gradient mistakes and encourage reliable training outcomes. As shown in Fig. 1, the dataset was then split into two subgroups, with 80% designated for training and the remaining 20% for testing.

Fig. 1. Illustration of dataset and separation to train and test.

For a comprehensive analysis, it is necessary to include the number of transactions as well as the OHLC prices for a given time period. A type of financial chart called a candlestick chart is used to show price changes over time. The Japanese candlestick chart was created by rice trader Munehisa Hooma and is known as the Japanese candlestick chart [25]. A candlestick chart is similar to a combined line and bar chart. Four important pieces of information for a trading day are represented by each bar, which are the open, close, low, and high prices. There are usually three parts to a candlestick: the actual body, the lower shadow, and the upper shadow. If the initial price exceeds the closing price, the actual body will be filled in red. Otherwise, the actual body will just be green filler.

In a given time frame, the high and low-price ranges are shown with a high and low shadow. However, not every geranium has a shadow. A visual aid to decision-making in the stock market is the candlestick chart. When using a candlestick chart, it will be easier for the trader to communicate the highs and lows, as well as the open and close. As a result, a trader can identify the trend of stock market fluctuations in a certain period by examining candlestick patterns [26]. When the close exceeds the open, the candlestick is referred to as bullish. If not, it's referred to as a bearish candlestick. Fig. 2 explains a candlestick plot.

A common data preparation technique in statistics and machine learning is min-max normalization, also known as feature scaling or min-max scaling. Scaling a feature's values to a predefined range, typically between 0 and 1, is the primary objective of Min-Max normalization. The formula for Min-Max normalization is as follows [1]:

$$\text{Xscaled} = \frac{(X - Xmin)}{(Xmax - Xmin)} \quad (1)$$

B. Data Analysis and Model Preparation

Apart from the exhaustive preprocessing and analysis of the datasets that were previously delineated, it is critical to recognize the possible constraints and prejudices that are intrinsic to the process of training the models and data. Notwithstanding the diligence we expend in data cleansing and normalization, specific elements might introduce biases or compromise the dependability of our forecasting models. A possible constraint pertains to the exhaustiveness and precision of the dataset itself. Notwithstanding stringent data cleansing protocols, inherent biases or errors may persist and potentially compromise the efficacy of our models. Furthermore, potential biases may arise due to the selection of features and the level of detail in the data, especially if specific variables are disproportionately represented or absent. Moreover, the utilization of candlestick charts and pattern recognition in our predictive processes introduces an element of subjectivity. Although these methodologies provide valuable insights into market trends, they are not devoid of constraints. The assessment of candlestick patterns is inherently subjective and subject to analyst variation, which may result in the omission of significant factors that impact market behavior or the formation of biased conclusions. In order to address these constraints, we have incorporated rigorous validation methods and conducted sensitivity analyses to evaluate the models' robustness and applicability. Furthermore, continuous monitoring and improvement of our methodologies are crucial in order to rectify any emerging biases or constraints that may arise during the course of our analysis. By recognizing these possible constraints and prejudices, our objective is to offer a more equitable and clear analysis of our findings, thereby cultivating trust in the dependability of our predictive models.

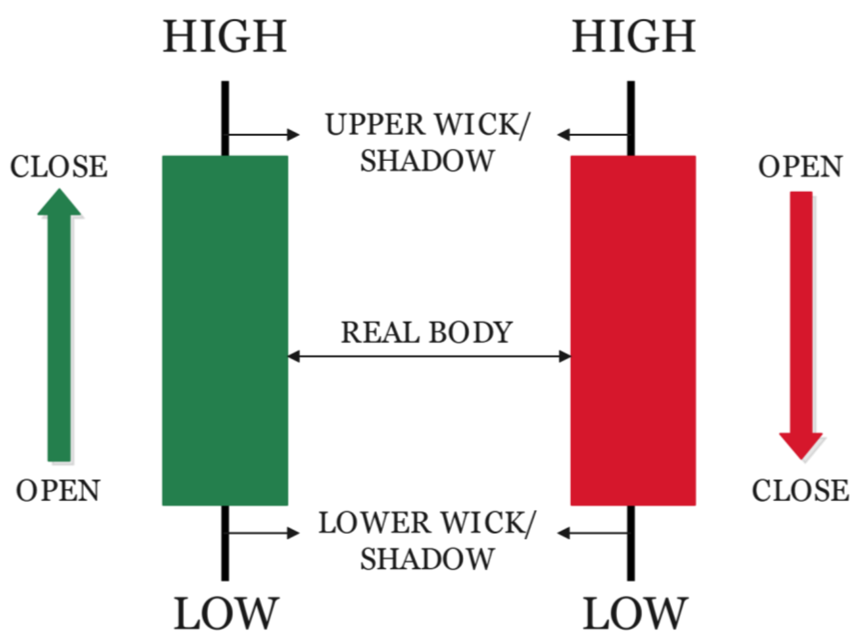

Fig. 2. Brief overview of candlestick chart.

C. Support Vector Regression

The algorithm's basic idea is as follows: given a training vector, $x_i \in \mathbb{R}^p, i = 1,2, \ldots, n$, and a vector $y \in \mathbb{R}^p$, our goal is to find $\omega \in \mathbb{R}^p$ and $b \in \mathbb{R}^p$ such that the prediction gives by sign $\omega^T \varphi(x_i) + b$ is correct for most samples. The optimization problem that must be addressed in order to estimate ω and b is indicated by the minimal value of the equation that follows:

$$\operatorname{argmin}\left(\frac{1}{2}\parallel \omega \parallel^2\right) + C\sum_{i=1}^{n} \xi_i + \xi_i^*$$
$$\text{s.t.} \begin{cases} y_i - (\omega^T\varphi(x_i) + b) \leq \varepsilon + \xi_i \\ (\omega^T\varphi(x_i) + b) - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* > 0, i = 1,2 \ldots, n \end{cases} \tag{2}$$

In this context, C represents the penalty parameter, ξ_i and ξ_i^* denote the slack variables, and ε stands for the insensitive loss function. The inclusion of ε enhances the resilience of the estimation. To address the issues above, the duality theory is commonly employed to convert it into a convex quadratic programming problem. Through the application of Lagrange transformation to Eq. (2), we can derive:

$$(\omega, b, \xi, \xi^*, \beta, \beta^*, \mu, \mu^*) = \left(\frac{1}{2}\parallel \omega \parallel^2\right)$$
$$+C\sum_{i=1}^{n}(\xi_i + \xi_i^*) - \sum_{i=1}^{n}\beta_i[\varepsilon + \xi_i - y_i + (\omega^T\varphi(x_i) + b)]$$
$$-\sum_{i=1}^{n}\beta_i^*[\varepsilon + \xi_i^* + y_i - (\omega^T\varphi(x_i) + b)]$$
$$-\sum_{i=1}^{n}\mu_i\xi_i - \sum_{i=1}^{n}\mu_i^*\xi_i^*, \ \text{S.T. } \beta_i \geq 0, \beta_i^* \geq 0, \mu_i \geq 0, \mu_i^* \geq 0 \tag{3}$$

$\beta_i, \beta_i^*, \mu_i, \mu_i^*$ are Lagrange coefficients. Using a partial derivative of Lagrange function concerning variables ω, b, ξ, ξ^* are equal to 0. Using a partial derivative of the Lagrange function with respect to the variables ω, b, ξ, ξ^* are equal to 0. Upon importing the Lagrangian operator and the optimization restriction expression, the decision function of Eq. (3) becomes the following form:

$$f(x) = \sum_{i=1}^{n}(\beta_i - \beta_i^*)K(x_i, x) + b \tag{4}$$

In Eq. (4), $\beta_i, \beta_i^* \geq 0$, and $K(x_i, x)$ is a kernel function. The overall structure of the SVR methods is demonstrated in Fig. 3.

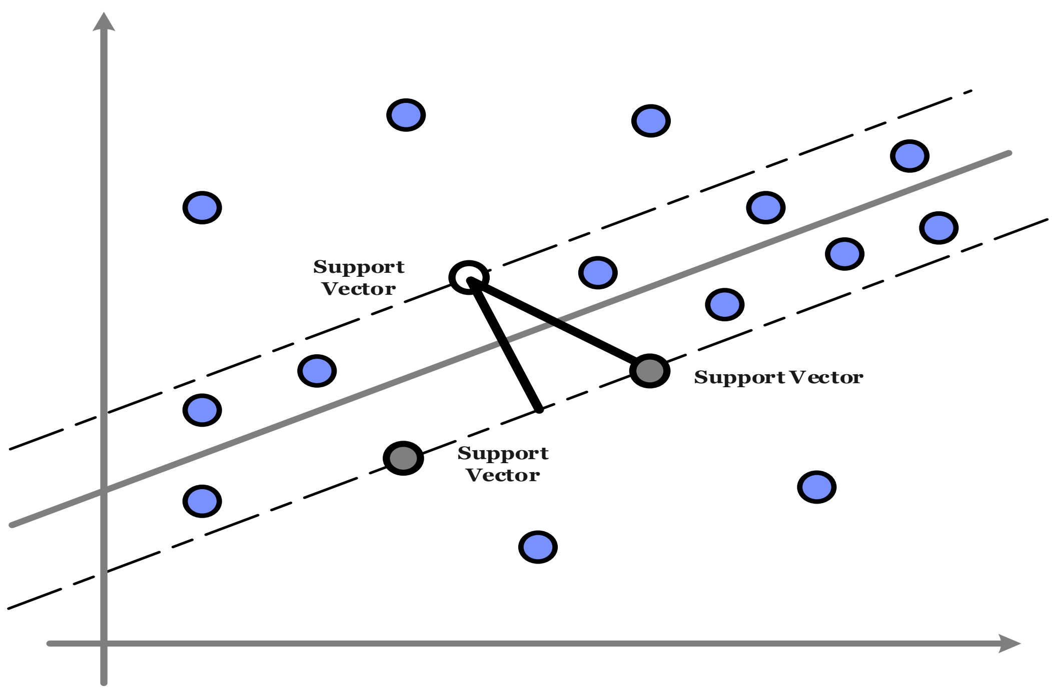

Fig. 3. Illustration of SVR.

D. Genetic Algorithm

Natural selection is simulated by the GA method, which is used to solve optimization and search problems. Fundamentally, it involves using genetic operators, such as crossover, consistently (recombination), mutation, and selection, to a population of candidate solutions, or persons, to generate new individuals. Next, the role of fitness, which evaluates the solution's quality, is applied to evaluate the new individuals. This procedure is carried out across a number of generations until a workable answer is discovered [27]. GA is made up of three essential components [28]: An individual's chromosome is a sequence of characters or numbers. Which encoding is used depends on the specific problem being addressed. Assessing each person's contribution to the solution's quality is done using the fitness function. Considering the present problem, the fitness component was developed. From already-existing individuals, it can be made new ones by using evolutionary operators. Operators that are used most frequently include crossover, mutation, and selection. Selection is the process of identifying which persons are most fertile. Crossover is the process of combining the chromosomes of two people to combine a third person. The purpose of the mutation is to cause small, haphazard changes to a person's chromosomes. It's critical to keep in mind that heuristic optimization is what GA does; while it can provide a decent answer at a reasonable processing cost, it cannot be trusted to find the optimal solution overall. For large-scale issues, however, it might be computationally demanding and time-consuming, particularly if the dataset is huge and the training procedure is drawn out [29]. The optimal values for the hyperparameters of the SVR, as determined by GA, are presented in Table I.

E. Artificial Bee Colony

Because the ABC algorithm can find excellent solutions with very little processing overhead, it has been chosen as the best tool. Previous studies [30][31] have optimized multi-dimensional numerical problems using the ABC technique. After being published by Basturk and Karaboga [32], Karaboga et al. [30][31] developed a unique population-based meta-heuristic approach known as the ABC algorithm. The ABC algorithm was inspired by the ingenious foraging strategies employed by swarms of honeybees. Bees that forage come in three varieties: employed bees, onlooker bees, and scout bees. Every bee that is actively searching for food is classified as employed. The ABC algorithm's specifics are as follows. The first solutions are created at random and utilized by the bee agents as their food supply locations. Following initiation, the bee agents go through three main cycles of iterative changes: choosing viable solutions, upgrading the workable solutions and steering clear of less-than-ideal solutions. Every hired bee chooses a new potential food supply status should be updated the workable solutions. Their decision is influenced by the area around the food source they had previously chosen. Eq. (5) is used to determine where the new food supply is located.

$$v_{ij} = x_{ij} + \phi_{ij}\left(x_{ij} - x_{kj}\right) \tag{5}$$

where, v_{ij} is a fresh, workable answer that has been altered from its initial solution value $\left(x_{ij}\right)$ according to a comparison with a place chosen at random from its nearby solution $\left(x_{kj}\right), \phi_{ij}$ is a random number between $[-1,1]$ that is used to randomly in the following iteration, modify the previous answer to become a new one, and $k \in \{1,2,3, \ldots, SN\}$ and $k \neq i, j \in \{1,2,3, \ldots, D\}$ are arbitrary index selections. What distinguishes between x_{ij} and x_{kj} is a shift in location within a certain dimension. Suppose a new candidate's food source position has a higher fitness value than the previous one. The old food source position will be replaced in the employed bee's memory. The working bees will go back to their colony and share the fitness with the other bees benefits of their new food sources. The fitness value that the working bees provide determines which of the suggested food sources each observer bee chooses in the following stage. Eq. (6) provides the likelihood that a suggested food source will be chosen.

$$P_i = \frac{fit_i}{\sum_{i=1}^{SN} fit_i} \tag{6}$$

where, the food source's fitness value is represented by fit $_i$ There are i possible food sources, and their sum is SN. Table I contains the optimal values for the hyperparameters of the SVR that have been determined by ABC.

F. Moth Flame Optimization

Mirjalili [18] proposed the MFO Algorithm. It builds an efficient swarm-based optimization technique by taking into account the intricate flying patterns of phototactic moths and modeling their movement around a flame analytically. Like other nocturnal animals, moths navigate by using celestial bodies. They commonly use transverse orientation navigation, which uses the moon as a navigational aid. To continue producing fruit, a moth travels at a constant angle to the moon. The moth's minuscule movement about its distance from the moon is what makes this navigational method effective.

On the other hand, man-made light sources frequently veer off into a lethal spiral around a light source. Rather than the moth and moon's separation, this occurs due to the light source's close closeness. In this instance, the moth enters the light source spirally rather than in a straight path, as would be required by keeping the transverse orientation. Fig. 4 and a thorough explanation of this phenomenon found in [18].

A haphazard population of moths is first formed in the search space. They are updated in a spiral pattern concerning the flame, keeping in mind that the moth's movement shouldn't go beyond the search space. Fig. 5 suggests that the moths are circling the flame in a hyperelliptical pattern, traveling in all directions. Because the moths migrate towards the flame, the algorithm gets confined to limited optimal states, and each moth's location is updated concerning its matching flame. This reduces the possibility of local optima stagnation because each month will circle different flames. Furthermore, the flame position is modified every iteration concerning the best answer, improving the algorithm's opportunity for investigation.

Moth movement limits the ability to use new flame positions in search space while also increasing the degree of exploration. The primary objective of any optimization algorithm is to create equilibrium between the periods of exploration and exploitation. An approach that is adaptable to determine the number of flames is proposed to increase the algorithm's exploitation. During the iteration, the number of flames steadily drops. In the most recent round of retries, it ensures that moth modifies their location to match the most advantageous updated flare. The best positions that the moths have so far managed to achieve are also displayed in a flame matrix, and an array indicates the matching fitness of these places. Moths look for the optimum outcome by updating their locations and searching around their associated flame; as a consequence, they never lose their optimal position. All moths' positions are updated concerning the respective flames, as indicated by Eq. (7).

$$M_i = S\left(M_i, F_j\right) \tag{7}$$

where, M_i and F_j stand for the i^{th} moth and the j^{th} flame, respectively, and S is the spiral function. Eq. (8) defines an exponential spiral, which serves as the primary updating mechanism.

$$S\left(M_i, F_j\right) = D_i \cdot e^{bt} \cdot \cos(2\pi t) + F_j \tag{8}$$

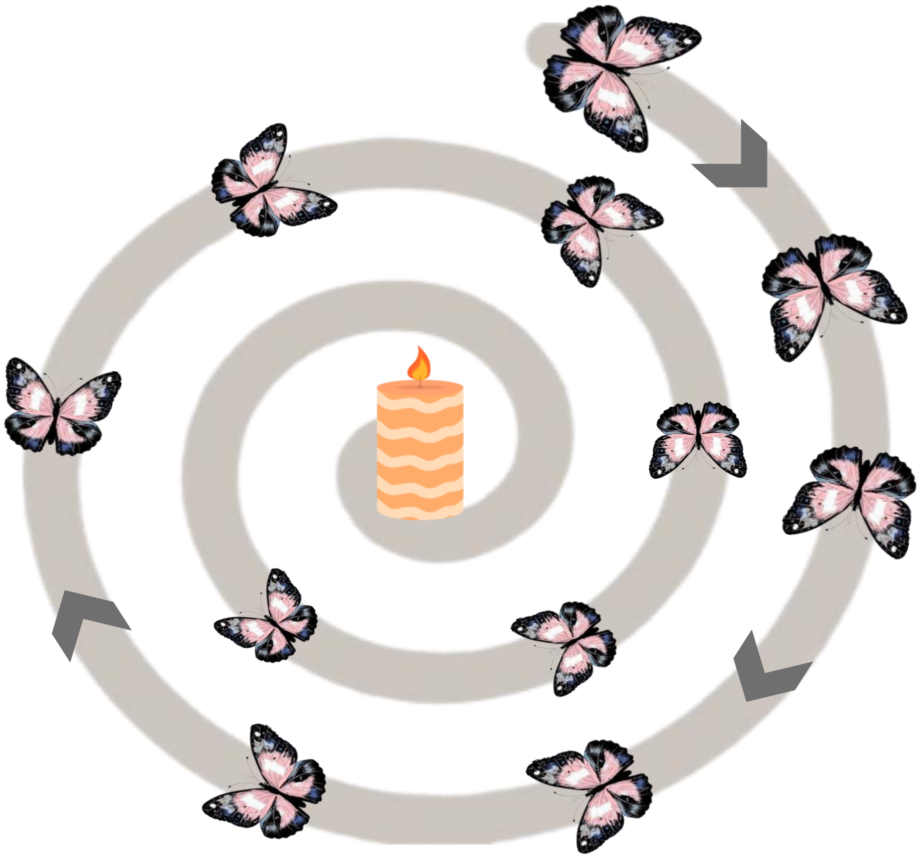

Fig. 4. Visualizing the moth flame optimization.

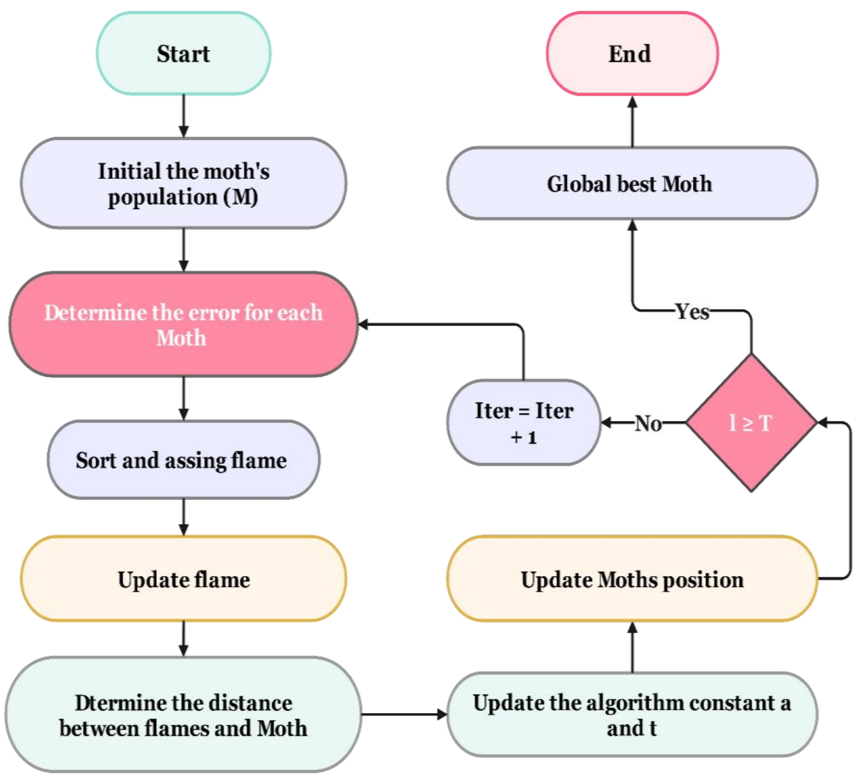

Fig. 5. MFO flowchart.

Eq. (9) computes the distance D_i, which is the separation between the i^{th} moth and the j-th flame. The constant value b is used to define the form of the logarithmic spiral. Over the duration of the iterations, the parameter t is a random number in the interval $[r, 1]$, where r is a factor of convergence and falls linearly from -1 to -2.

$$D_i = \left|M_i - F_j\right| \tag{9}$$

Every moth updates its position with one flame to prevent becoming trapped in local optima. Every time, the flames list is refreshed and arranged according to their fitness values. The first moth modifies its location based on the optimal flame, whereas the final moth modifies its position based on the least optimal flame. Additionally, an adaptive mechanism reduces the number of flames between iterations to improve the exploitation of the most promising solutions. Eq. (10) illustrates this technique in action.

$$\text{flame}_{\text{No}} = round(N - \text{ iter } \cdot (N-1)/ \text{ MaxIter }) \tag{10}$$

The maximum number of moths is denoted by N, while the current and maximum number of iterations are represented by iter and Maxlter, respectively. The optimal values that have been found for the SVR's hyperparameters by MFO are presented in Table I.

TABLE I. SETTING OF THE SVR HYPERPARAMETERS

SVR		GA	ABC	MFO
kernal	[Linear, RBF, Poly, and Sigmoid]	linear	linear	linear
gamma	[1, 0.5, 0.1, 0.01, 0.001]	0.5	0.1	0.5
C	[0.1, 1, 10, 20, 50, 100]	10	20	10
epsilon	[0.01, 0.05, 0.1, 0.5]	0.05	0.05	0.1

IV. RESULT AND DISCUSSION

A. Evaluation Metrics

Regression models' accuracy and efficacy when forecasting the values of the output by using the input data are assessed using evaluation criteria. The discrepancy between the expected and actual numbers is measured by the Mean Squared Error (MSE). It is computed by first computing the square of the discrepancy between what was anticipated and what was observed and then averaging all of these squared variations. The model's correctness is determined by this value; the lower the MSE, the more accurate the model.

$$MSE = \frac{1}{N}\sum_{k=0}^{n} \binom{n}{k} (Fi - Yi)b^2 \tag{11}$$

The disparity between the expected and actual values is also measured by the MAE. To compute it, take the total amount that differs between the actual and anticipated numbers, then average the whole difference. The lower the MAE, the better the model; this number is also used to evaluate the model's correctness.

$$MAE = \frac{\sum_{i=1}^{n}|y_i - \hat{y}_i|}{n} \tag{12}$$

MAPE is a percentage-based metric used to assess a model's accuracy. The calculation involves splitting the whole amount of the discrepancy between the actual and anticipated values by the real amount and then averaging all of these percentages. The lower the MAPE, the better the model; this number is used to evaluate the model's accuracy.

$$MAPE = \left(\frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right|\right) \times 100 \tag{13}$$

Root means square error (RMSE) is another indicator that provides significant support in evaluating the precision of forecasting models.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \tag{14}$$

B. Statistical Values

A long analysis of the dataset is provided by the study report, which is displayed in Table II. The table offers a comprehensive statistical representation of the OHLC volume and price data. This enables a more thorough comprehension of the information. The variance, kurtosis, skewness, mean, standard deviation (Std.), minimum (Min), and maximum (Max) values are among the several statistical measures displayed in the table. These measures offer an accurate and comprehensive data analysis. Central tendency, variability, and dispersion of the data are only a few of the many aspects of the data about which each of these measures provides insightful information.

TABLE II. STATISTICAL RESULTS OF THE PRESENTED MODELS FOR OHCLV

	Count	Mean	Std.	Min	Median	Max	Skew	Kurtosis
Open	2442	20813.83	4765.013	10405.67	20538.9	30606.15	0.171314	-0.80564
High	2442	20926.76	4777.106	10602.12	20632.72	30795.78	0.177934	-0.80271
Low	2442	20690.79	4748.074	10398.61	20451.26	30504.81	0.16562	-0.81171
Volume	2442	3730.003	1985.287	0	3180	19840	1.912733	6.522738
Close	2442	20812.22	4763.784	10486.99	20559.85	30670.1	0.170538	-0.80791

C. Analyze and Discussion

The major objective of this work is to identify and assess the optimal hybrid algorithm for stock price forecasting. To do this, researchers have created prediction models and evaluated a wide range of intricate variables that affect stock market patterns. The major objective is to give analysts and investors pertinent information so they can make well-informed investment decisions. Together with a comprehensive analysis of each model's effectiveness, Table III, Fig. 6, and Fig. 7 offer a detailed assessment of each model's performance.

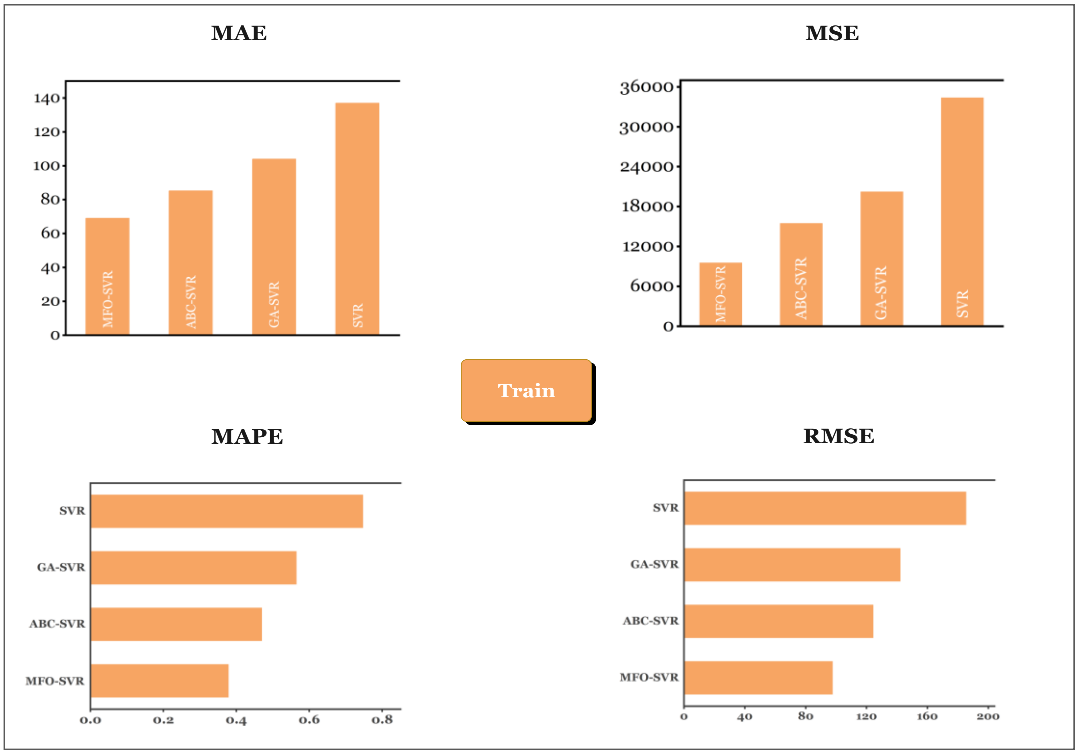

Fig. 6. Result of the evaluation metrics for the presented models during training.

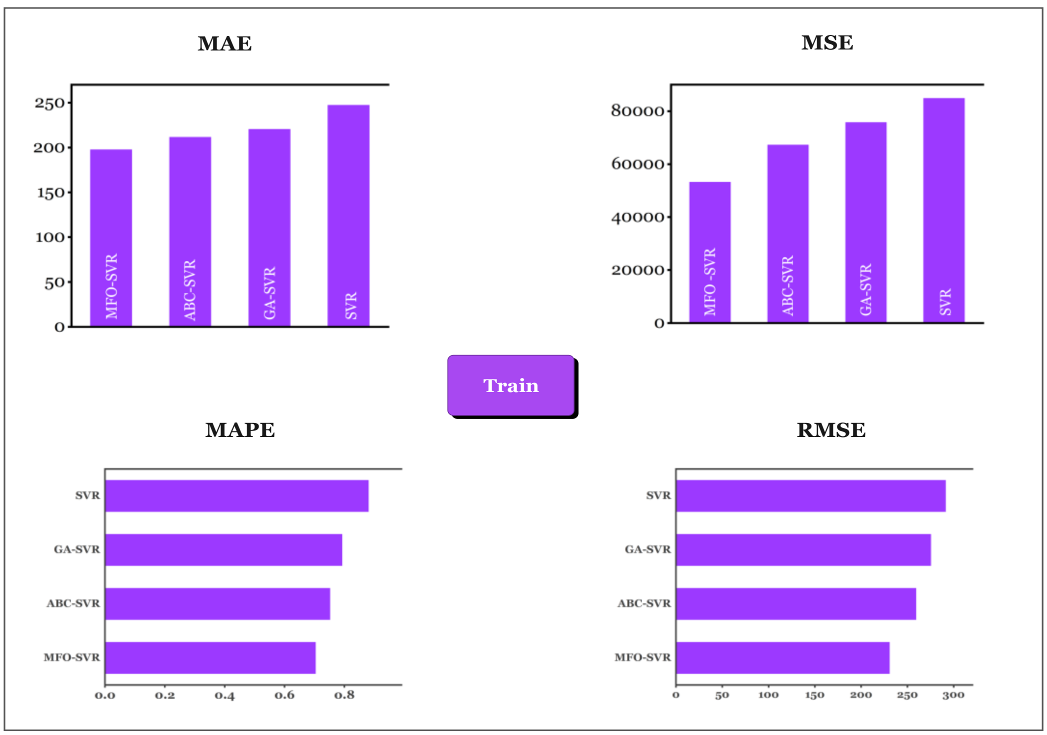

Fig. 7. Result of the evaluation metrics for the presented models during the test.

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

MODEL/Metrics	TRAIN				TEST			
	RMSE	MAPE	MAE	MSE	RMSE	MAPE	MAE	MSE
ARIMA	255.34	1.01	184.44	65199.22	348.19	1.09	307.16	121236.11
MLP	226.20	0.91	168.52	51164.83	314.20	1.02	287.83	98721.97
SVR	185.35	0.75	136.94	34352.80	291.21	0.88	247.12	84803.10
GA-SVR	142.10	0.56	103.95	20193.55	275.18	0.79	220.27	75724.66
ABC-SVR	124.29	0.47	85.23	15448.45	259.20	0.75	211.43	67182.42
MFO-SVR	97.55	0.38	68.98	9516.14	230.60	0.70	197.53	53175.49

A thorough study of the data analysis was conducted using four well-known metrics: RMSE, MAE, MAPE, and MSE. These indicators are well renowned for providing a precise evaluation of the overall accuracy, reliability, and efficacy of the analysis. The performance of the SVR model, both with and without an optimizer, was assessed using the RMSE, MAPE, MSE, and MAE criteria. This assessment enhanced the capacity to understand the model's performance and make decisions in light of the findings. This provided a comprehensive understanding of the model's performance, enabling well-informed decision-making. During training and testing, SVR's RMSE values were 185.35 and 291.21, respectively, while the MAE values were 136.94 and 247.12, respectively. The values of MAPE were 0.75 and 0.88. MSE values for SVR during train and test were 34352.80 and 84803.10, respectively. The GA-SVR model performs well when optimizers are included. Additionally, compared to the training values, the testing RMSE, MAPE, MAE, and MSE values for GA-SVR were reduced at 275.18, 0.79, 220.27, and 75724.66, respectively. From a production standpoint, the ABC-SVR model outperformed the GA-SVR model. Additionally, to prove the ability of the MFO-SVR model two other benchmark models were utilized; these models are Autoregressive integrated moving average (ARIMA) and Multilayer perception (MLP). The obtained results of the ARIMA during the testing phase for RMSE, MAPE, MAE, and MSE were 348.19, 1.09, 307.16, and 121236.11. Likewise, these results for the MLP models were 314.20, 1.02, 287.83, and 98721.97, respectively. Having compared these models with MFO-SVR it can be concluded that the proposed model is more effective than these models.

In the training and testing data sets, the MFO-SVR model has demonstrated remarkable accuracy as a result. The MFO-SVR model is a superb resource for very accurate stock price prediction. How accurately our model predicts the paths of the Nikkei 225 index stocks is shown in Fig. 8 and Fig. 9. The SVR approach makes the MFO-SVR model different from other models in stock price forecasting because it can reduce price fluctuations, simplify trend prediction, and boost model precision. Among its distinctive features is the MFO-SVR model's ability to learn from previous data sets. In order for a model to accurately anticipate stock values and adjust to changing market trends, it must be trained on past data sets.

The potential real-world applications of the identified MFO-SVR hybrid algorithm for stock price forecasting are substantial throughout the financial industry. The precise forecasts it generates have the potential to form the basis of investment decision support systems, assisting analysts and investors in making well-informed decisions. Furthermore, the incorporation of this technology into algorithmic trading systems empowers the implementation of automated trading tactics that take advantage of anticipated fluctuations in stock prices. Furthermore, the capacity of the model to mitigate price volatility and offer valuable perspectives on market sentiment contributes to the improvement of risk management tactics and portfolio optimization endeavors. Furthermore, its predictions can be utilized by individuals for the purpose of financial planning, and by quantitative analysts to construct and validate trading strategies in the past. In general, the MFO-SVR model demonstrates its versatility as a tool that can be applied in a variety of contexts, providing stakeholders with informative insights into the dynamics of the stock market and enabling them to optimize their financial activities and accomplish their investment objectives.

Although the MFO-SVR hybrid algorithm exhibits potential applications in stock price forecasting, it is imperative to recognize specific constraints and avenues that warrant further investigation. A potential drawback of the model is its dependence on historical data, which might not comprehensively capture abrupt market fluctuations or unanticipated occurrences. As a result, forecasts made during periods of market volatility could be rendered less precise. Furthermore, the intricacy of market dynamics could potentially impede the model's capacity to extrapolate findings to diverse asset classes and market conditions. Further research may be dedicated to improving the model's resilience through the integration of real-time data streams and external factors, including news sentiment analysis and macroeconomic indicators, in order to enhance the accuracy of predictions. Additionally, further research endeavors may investigate alternative hybrid algorithms or machine learning methodologies in order to augment the performance of forecasting and tackle the concern of model interpretability. This would guarantee that stakeholders are able to comprehend and place confidence in the insights delivered. In addition, conducting an examination of the potential ramifications of transaction costs and liquidity limitations on the efficacy of the model in practical trading situations may yield significant knowledge for its application. In general, the ongoing progress and enhancement of stock price forecasting algorithms will be aided by the resolution of these constraints and the exploration of additional research directions. This will ultimately be to the advantage of investors and financial practitioners.

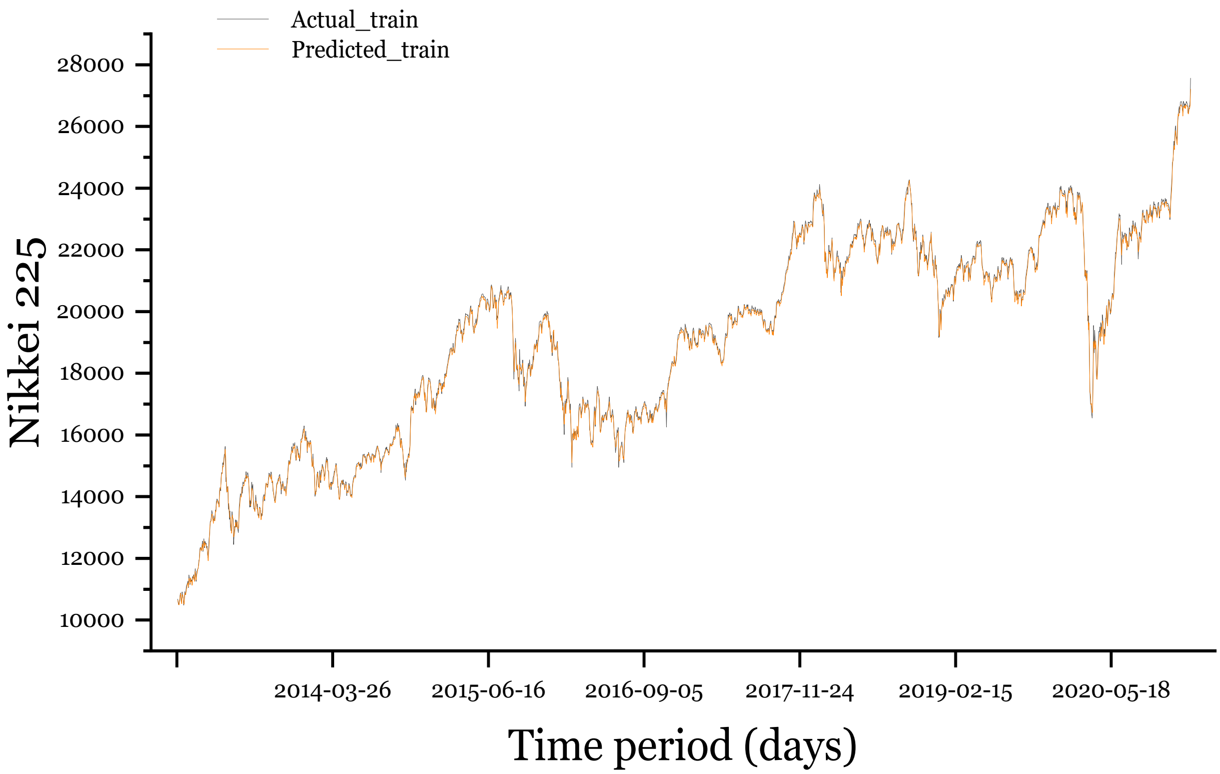

Fig. 8. Evaluation of the performance of the proposed model in comparison to real data during training.

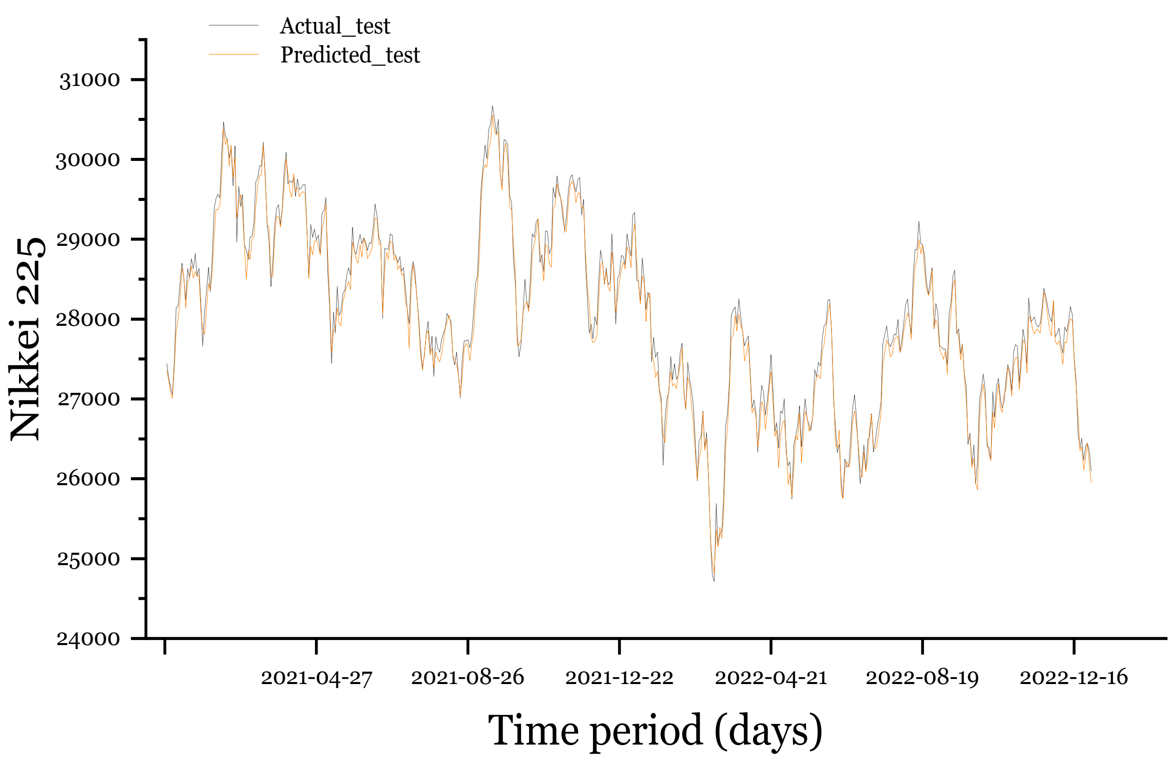

Fig. 9. Evaluation of the performance of the proposed model in comparison to real data during testing.

V. Conclusion

The financial market is a realm that captivates investors, market analysts, and academics, providing an abundance of opportunities for investigation. By employing stock prediction methodologies, both individual and institutional investors can potentially attain a competitive advantage in identifying market trends and assessing assets. By leveraging historical data and sophisticated algorithms, investors are empowered to render well-informed decisions pertaining to stock transactions, encompassing purchases, sales, and holdings. The present study utilized support vector regression networks, which were optimized using the MFO approach, in order to predict the values of stocks. The objective of the MFO-SVR model presented in this study is to forecast trends in the stock market. Through the application of Nikkei 225 index data encompassing the period from January 1, 2013, to January 1, 2023, this research has established a foundation for subsequent inquiries. The dataset, which is composed of 20% test data and 80% training data, provides a solid foundation for subsequent analysis. Anticipating the future, numerous pathways exist for further investigation. To commence, the generalizability of predictive models could be improved by broadening the scope of analysis to include supplementary datasets sourced from various financial markets. Furthermore, an examination of alternative optimization methodologies or the integration of ensemble techniques may enhance the precision and resilience of forecasts. In addition, ongoing research into real-time prediction models may provide valuable insights for timely decision-making, given the dynamic nature of financial markets. Through the seamless integration of these prospective research concepts into our current conclusions, we establish a foundation for ongoing progress and improvement in the domain of financial market forecasting. Conducting research and exploration in an iterative manner is critical for remaining informed about the ever-changing dynamics of the market and guaranteeing that predictive models remain practical and applicable.

References

[1] K. Kim, “Financial time series forecasting using support vector machines,” Neurocomputing, vol. 55, no. 1–2, pp. 307–319, 2003.

[2] H. Ince and T. B. Trafalis, “Kernel principal component analysis and support vector machines for stock price prediction,” Iie Transactions, vol. 39, no. 6, pp. 629–637, 2007.

[3] R. K. Dash, T. N. Nguyen, K. Cengiz, and A. Sharma, “Fine-tuned support vector regression model for stock predictions,” Neural Comput Appl, no. July, 2021, doi: 10.1007/s00521-021-05842-w.

[4] C.-J. Lu, C.-H. Chang, C.-Y. Chen, C.-C. Chiu, and T.-S. Lee, “Stock index prediction: A comparison of MARS, BPN and SVR in an emerging market,” in 2009 IEEE International conference on Industrial Engineering and Engineering Management, IEEE, 2009, pp. 2343–2347.

[5] K. S. Kannan, P. S. Sekar, M. M. Sathik, and P. Arumugam, “Financial stock market forecast using data mining techniques,” in Proceedings of the International Multiconference of Engineers and computer scientists, 2010.

[6] M. Bansal, A. Goyal, and A. Choudhary, “A comparative analysis of K-Nearest Neighbor, Genetic, Support Vector Machine, Decision Tree, and Long Short Term Memory algorithms in machine learning,” Decision Analytics Journal, vol. 3, no. November 2021, p. 100071, 2022, doi: 10.1016/j.dajour.2022.100071.

[7] S. Ray, “A quick review of machine learning algorithms,” in 2019 International conference on machine learning, big data, cloud and parallel computing (COMITCon), IEEE, 2019, pp. 35–39.

[8] 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.

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

[10] R. A. Khurma, H. Alsawalqah, I. Aljarah, M. A. Elaziz, and R. Damaševičius, “An enhanced evolutionary software defect prediction method using island moth flame optimization,” Mathematics, vol. 9, no. 15, pp. 1–20, 2021, doi: 10.3390/math9151722.

[11] A. Mozaffari, M. Emami, and A. Fathi, “A comprehensive investigation into the performance, robustness, scalability and convergence of chaos-enhanced evolutionary algorithms with boundary constraints,” Artif Intell Rev, vol. 52, pp. 2319–2380, 2019.

[12] N. Izeboudjen, C. Larbes, and A. Farah, “A new classification approach for neural networks hardware: from standards chips to embedded systems on chip,” Artif Intell Rev, vol. 41, pp. 491–534, 2014.

[13] M. Desai and M. Shah, “An anatomization on breast cancer detection and diagnosis employing multi-layer perceptron neural network (MLP) and Convolutional neural network (CNN),” Clinical eHealth, vol. 4, pp. 1–11, 2021.

[14] X. Hu and Q. Weng, “Estimating impervious surfaces from medium spatial resolution imagery using the self-organizing map and multi-layer perceptron neural networks,” Remote Sens Environ, vol. 113, no. 10, pp. 2089–2102, 2009.

[15] Y.-S. Park and S. Lek, “Artificial neural networks: Multilayer perceptron for ecological modeling,” in Developments in environmental modelling, vol. 28, Elsevier, 2016, pp. 123–140.

[16] W. S. Noble, “What is a support vector machine?,” Nat Biotechnol, vol. 24, no. 12, pp. 1565–1567, 2006.

[17] L. N. Mintarya, J. N. M. Halim, C. Angie, S. Achmad, and A. Kurniawan, “Machine learning approaches in stock market prediction: A systematic literature review,” Procedia Comput Sci, vol. 216, pp. 96–102, 2022, doi: 10.1016/j.procs.2022.12.115.

[18] 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.

[19] 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.

[20] 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.

[21] 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.

[22] 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

[23] 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

[24] 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

[25] T. Logan, Getting started in candlestick charting, vol. 73. John Wiley & Sons, 2008.

[26] T.-H. Lu, Y.-M. Shiu, and T.-C. Liu, "Profitable candlestick trading strategies—The evidence from a new perspective," Review of Financial Economics, vol. 21, no. 2, pp. 63–68, 2012.

[27] B. Mohan and J. Badra, "A novel automated SuperLearner using a genetic algorithm-based hyperparameter optimization," Advances in Engineering Software, vol. 175, no. September 2022, p. 103358, 2023, doi: 10.1016/j.advengsoft.2022.103358.

[28] 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.

[29] D. M. Rocke and Z. Michalewicz, "Genetic algorithms+ data structures= evolution programs," J Am Stat Assoc, vol. 95, no. 449, p. 347, 2000.

[30] D. Karaboga and B. Basturk, "A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm," Journal of global optimization, vol. 39, pp. 459–471, 2007.

[31] D. Karaboga and B. Akay, "A comparative study of artificial bee colony algorithm," Appl Math Comput, vol. 214, no. 1, pp. 108–132, 2009.

[32] B. Basturk, "An artificial bee colony (ABC) algorithm for numeric function optimization," in IEEE Swarm Intelligence Symosium, Indianapolis, IN, USA, 2006, 2006, p. 12.