A Novel Proposal for Improving Economic Decision-Making Through Stock Price Index Forecasting

Xu Yao[1], Weikang Zeng[2], Lei Zhu[3], Xiaoxiao Wu[4], Di Li[5]
School of Business Administration, Xi'an Eurasia University, Xi'an 710065, Shaanxi, China[1, 3, 5]
School of Economics and Management, Northwest University, Xi'an 710127, Shaanxi, China[2]
School of Public Administration, Xi'an University of Architecture and Technology, Xi'an 710054, Shaanxi, China[3]
School of International Education, Southeast Asian University of Thailand, Bangkok 10160, Thailand[4]

***Abstract*—The non-stationary, non-linear, and extremely noisy nature of stock price time series data, which are created from economic factors and systematic and unsystematic risks, makes it difficult to make reliable predictions of stock prices in the securities market. Conventional methods may improve forecasting accuracy, but they can additionally complicate the computations involved, increasing the likelihood of prediction errors. To address these issues, a novel hybrid model that combines recurrent neural networks and grey wolf optimization was introduced in the current study. The suggested model outperformed other models in the study with high efficacy, minimal error, and peak performance. Utilizing data from Alphabet stock spanning from June 29, 2023, to January 1, 2015, the effectiveness of the hybrid model was assessed. The gathered information comprised daily prices and trading volume. The outcomes showed that the suggested model is a reliable and effective method for analyzing and forecasting the time series of the financial market. The suggested model is additionally particularly well-suited to the volatile stock market and outperforms other recent strategies in terms of forecasting accuracy.**

Keywords*—Hybrid model; recurrent neural networks; grey wolf optimization; stock price prediction

I. Introduction

The finance market is a fascinating and intricate structure that profoundly affects some domains, including business, employment, and technology [1], [2]. It gives investors numerous chances to invest money and generate returns with minimal risk [3], [4]. Fama conducted one such study [5]. The current understanding of stock price behavior has been greatly influenced by this subject, which is still an important area of research in the field of finance. Fundamental analysis and technical analysis are the two main stock market decision-making methodologies used by investors. While technical analysis looks at historical market data, trends, and patterns to forecast future market movements, fundamental analysis looks at a company's financial health and its possibilities for future growth [6], [7]. Stock market forecasting has been increasingly important among knowledgeable analysts and investors in recent years. However, because of the chaotic nature of the market, it is quite challenging to analyze stock market movements and price actions [8]. Global economic conditions, market news, and quarterly earnings reports are just a few of the factors that have a significant impact on the stock market. These elements make it difficult to predict stock prices with any degree of accuracy. The market capitalization of the listed companies is used to create the stock market indices. As a result, stock market indexes' values reflect the total value of the underlying stocks. In this market environment that is continuously changing, it is difficult to make accurate stock market predictions. Utilizing various statistical tools, researchers and market analysts have been eager to develop and test stock market behavior. These methods, which offer insights into the intricate and dynamic character of the stock market, include clustering and autoregressive integrated moving averages [9]. However, given that price changes tend to be chaotic, noisy, nonparametric, nonlinear, and nonstationary, it is challenging for analysts to accurately assess and forecast price changes [10]. These characteristics imply that traditional statistical methods may not be sufficient for effective stock market analysis.

Data analysis technologies including spreadsheets, automated data collection, and prediction models started to appear in the 1980s. With the development of technology, researchers started to employ deep learning and artificial neural networks to create models that could learn complicated relationships and extract crucial information more quickly than they could with earlier technologies. Many different architectures have been designed to address various issues and handle the complicated structure of datasets as deep learning has become more popular in recent years. Information only moves forward in a basic feedforward neural network architecture. There is no memory of past inputs; each input is handled independently. Because previous events are essential for forecasting future events, these models are not appropriate for sequential data. Recurrent neural networks (RNNs) have been created to handle similar tasks. Loops built into RNN architecture enable the persistence of pertinent data across time. Internal information is transferred within the network from one time-step to the next. With this architecture, RNNs are more suited for time series applications including stock market forecasting, language translation, and signal processing as well as sequential data modeling. Many other tasks, such as speech recognition, image captioning, and natural language processing, have been carried out using RNNs. Additionally, they have been applied in the area of finance, where they have demonstrated excellent potential for predicting stock values. RNNs are excellent for simulating complex and dynamic systems due to their capacity to retain information and develop context over time. When Hu et al. [11] used an ensemble RNN technique to forecast stock market movements, their opposing model was able to outperform the competition. To assess the significance of features, individual data points, and particular cells in each architecture, Freeborough et al. [12] applied four well-known techniques to the RNN, long short-term memory (LSTM), and a gated recurrent unit (GRU) network trained on S&P 500 stocks data. These techniques are ablation, permutation, added noise, and integrated gradients.

More and more, real-world engineering design issues are solved optimally using stochastic operators-based metaheuristic algorithms [13]. Deterministic algorithms are trustworthy, but they are less efficient at locating global optima because they can become stuck in local optima [14]. Randomness is a technique used by stochastic optimization algorithms, such as evolutionary ones, to avoid local solutions and locate global optima in search spaces [15]. Although each run of these techniques results in a different solution, they outperform deterministic algorithms in terms of avoiding local solutions. Take for example, ant lion optimization (ALO) [16], Biogeography-based optimization (BBO) [17], Aquila optimizer (AO) [18], grey wolf optimization (GWO) [19], and so on. GWO is a recently developed meta-heuristic optimization algorithm that draws inspiration from the communal foraging behavior of grey wolves in the wild. Mirjalili et al. [19] made the initial pitch in 2014. Alpha, beta, delta, and omega wolves are used by the GWO algorithm to replicate the leadership structure and hunting strategy of grey wolves. Certain frameworks were proposed by Rajput et al. [20] for stock price forecasting, including ARIMA (Auto Regressive-Integrated-Moving Average), FLANN (Functional Link Artificial Neural Network), ELM (Extreme Learning Machine) models, and Grey Wolf optimizer. Kumar Chandra. [21] employed Elman neural network (ENN) and GWO algorithm to optimize the parameters of ENN for forecasting the stock market. A model using an artificial neural network (ANN) optimized by the GWO method was given by Sahoo et al. [22] and the Bombay Stock Exchange (BSE) was used as the dataset in their essay.

The article introduces the GWO-RNN hybrid model, a highly reliable stock price forecasting tool. By contrasting it with many other models, including RNN, BBO-RNN, and AO-RNN, the study evaluated its accuracy. A method involving several analytical steps was used to achieve this. The principal contributions of the investigation are as follows:

A novel hybrid model is presented in this study, which integrates GWO and RNNs to tackle the intricacies associated with stock price prediction. The purpose of this hybrid model is to address the drawbacks of traditional methods by providing a more efficient forecasting method, with reduced computation complexity and prediction error probability.

The analysis utilizes Alphabet stock data spanning from January 1, 2015, to June 29, 2023, to assess the effectiveness of the hybrid model that has been proposed. In terms of high efficacy, minimal error, and optimum performance, the results demonstrate that the proposed model outperforms alternative models, thereby establishing its dependability for forecasting and analyzing financial market time series data.

The research paper validates the reliability of the GWO-RNN model as an instrument that generates exceptionally precise forecasts of stock prices. The text underscores the model's capacity to rapidly analyze and interpret substantial amounts of data by combining grey wolf optimization-based optimization with recurrent neural networks. This functionality offers investors significant insights of market trends as well as prospective investment prospects.

The proposed model exhibits exceptional suitability for chaotic stock markets, outperforming other contemporary strategies in terms of the accuracy of its forecasts. This indicates that the GWO-RNN model is capable of efficiently managing the difficulties presented by market volatility, rendering it a dependable instrument for investors in search of precise forecasts in ever-changing financial landscapes.

Section II of this research contains the literature review. Section III offers a thorough examination of the data source and all of its relevant components. The data was analyzed using a variety of methodologies, including the RNN model, evaluation metrics, and the GWO optimizer. The experimental findings are presented in Section IV. Next, they are contrasted and discussed with those from other approaches in Section V. The research's findings are finally summarized in Section VI.

II. Literature Review

In the last decade, the application of machine learning (ML) algorithms to predict stock markets has increased substantially. Christanto et al. [23] suggested an examination of methodologies employed in the capital market to predict stock prices through a comparative analysis of ML, technical analysis, and fundamental analysis. They utilized Support Vector Regression (SVR) and Support Vector Machine (SVM) as ML methodologies to forecast stock prices. The assessment includes three parameter groups: technical-only (TEC), financial statement-only (FIN), and a combination of the two (COM). Experiments revealed that the integration of financial statements had no impact on SVR forecasts yet had a beneficial impact on SVM forecasts. 83 percent was the accuracy rate attained by the model in the conducted investigation. In their study, Chen et al. [24] investigated the historical context of economic recessions, emphasizing the abrupt and disastrous outcomes that can be observed in instances like the 2008 financial crisis, which was marked by a significant decline in the S&P 500. Motivated by the potential benefits of prompt crisis detection, they applied advanced ML techniques, such as Extreme Gradient Boosting and Random Forest, to predict potential market declines in the United States. By comparing the efficacy of these methodologies, their investigation aims to determine which model exhibits superior predictive capability for US stock market collapses. An analysis was conducted on market indicators utilized in crisis forecasts. This involved the utilization of daily financial market data and 75 explanatory variables, including both broad US stock market indexes and sector indexes. Through the utilization of specific classification metrics, they arrived at conclusions regarding the effectiveness of their predictive models. Tsai et al. [25] examined the interest of investors in stock forecasting, focusing on the recent application of ML to enhance precision. They deliberated on fiscal year-end selection and the impact that misaligned reporting periods have on investment decisions and comparability. They utilized ML models for fundamental analysis to predict Taiwan's (TW) stock market returns with an emphasis on synchronized fiscal years. Using models such as Financial Graph Attention Network (FinGAT), Feedforward Neural Network (FNN), Random Forest (RF), and Gated Recurrent Unit (GRU), they constructed stock portfolios with higher anticipated returns. Their research indicates that in terms of returns and portfolio scores, these portfolios surpassed the benchmarks of the TW50 index. They assert that ML models proved advantageous in the domains of investment decision-making and stock market analysis. Ardakani et al. [26] presented a federated learning framework utilizing Random Forest, Support Vector Machine, and Linear Regression models for stock market forecast. To determine the optimal strategy, they contrasted federated learning with centralized and decentralized frameworks, and the strategies of learning frameworks for stock market prediction were elucidated by their results. Mamluatul et al. [27] developed an innovative approach for forecasting stock prices by integrating ML, stock price data, technical indicators, and Google trends. To forecast stock prices, SVR, Multilayer Perceptron (MLP), and Multiple Linear Regression were implemented. With a Mean Absolute Percentage Error (MAPE) of 0.50%, SVR outperforms MLP and Multiple Linear Regression in forecasting Indonesian stock prices. They discovered that SVR accurately forecasts stock prices, enabling investors to make informed judgments regarding the stock market. Juare et al. [28] emphasized the significance of stock market analysis in determining financial market profits. Random Forest (RF), SVM, K-nearest neighbors (KNN), and Logistic Regression were utilized in their research to forecast stock market trends. The evaluation criteria for these algorithms are accuracy, recall, precision, and F-Score. Locating the optimal algorithm for stock market prediction was the primary objective. The value that investors and stock exchanges can derive from accurate forecasts underscores the significance of predictive models in the realm of financial decision-making. The importance of investors utilizing forecasting stock prices (SPP) models for profit in the global financial market was underscored by Swathi et al. [29]. SPP models from the past employed statistical and ML techniques. They presented SCODL-SPP, a method for predicting stock prices that integrates deep learning and Sine Cosine Optimization (SCO), in their investigation. The SCODL-SPP model employs deep learning and a stacked long short-term memory (SLSTM) model to predict the closing prices of stocks. The SCO algorithm is employed to optimize the hyperparameters of the SLSTM model after the min.

The literature review on stock market prediction closes several identified research gaps effectively. By utilizing the Alphabet stock as a central metric, this study offers specific insights about Alphabet stock, thus expanding the purview of research in this field. In addition, it addresses a void in the literature concerning the evaluation of data quality and preprocessing methodologies by providing clear and transparent explanations of data preprocessing procedures, thereby guaranteeing data quality and reproducibility. Furthermore, the incorporation of domain-specific insights into the GWO-RNN model improves the precision of predictions, thereby surmounting the drawback associated with the inadequate integration of domain knowledge. Furthermore, conducting a comparative analysis between the GWO-RNN model and other hybrid methods provides significant contributions to the understanding of the effectiveness of ensemble methods, thereby addressing a gap in the current body of research on this subject. Through an assessment of the GWO-RNN model's ability to on Alphabet stock market data encompassing the period from January 1, 2015, to June 29, 2023, this research ultimately establishes the model's dependability and efficacy in the domain of financial time series analysis and prediction. The aforementioned contributions symbolize noteworthy progressions in the field of stock market forecasting, resulting in enhanced predictive models' resilience, precision, and practicality within financial markets.

III. Research Methods

A. Recurrent Neural Network

In the complicated field of machine learning, different algorithms and techniques are used to interpret and analyze data. The recurrent neural network (RNN), which is made to handle sequential input, is one of the most crucial tools in this field and it is displayed in Fig. 1 and Fig. 2. RNNs can incorporate prior knowledge and produce outputs based on prior learning, in contrast to conventional feedforward neural networks, which only take into account the current input. The network's loop structure, which maintains a memory of previous inputs and outputs, enables this. The activation layer of each RNN unit uses a hyperbolic tangent function to process input and convert it into a format that the rest of the network can understand. The network updates its internal state as the input is analyzed, enabling it to take into account the context of earlier inputs when generating outputs. Because RNNs can take into account temporal context, which results in more effective and efficient learning, they are more accurate and useful than regular neural networks [30].

$$\mathrm{h_t} = \sigma(\mathrm{W_{xt}} + \mathrm{U_{ht-1}} + \mathrm{b}) \tag{1}$$

in the equation, b stands for the bias of the neuron, and W and U are weight matrices that represent the input to the current cell and the recurrent input, respectively. The input and hidden states of the cell at time t are represented by the values of $\mathrm{x_t}$ and $\mathrm{h_t}$. The symbol σ designates the sigmoid function of the neuron.

B. Biogeography-based Optimization

The foundation of biogeography-based optimization (BBO) is the movement of species according to the appropriateness of their habitat. Thus, a solution is like a habitat for an optimization issue. A crowded environment, where conditions for living species are better than in other habitats, is a better option for the population. The habitat in which living things struggle is the worst answer for the population. By sharing their traits, the superior solutions draw in the inferior ones. The following operators are used in processing this feature sharing. The operator of Migration: Migration is the process by which, following emigration and immigration rates, a better habitat replaces a worse option. The pace at which a species leaves its environment is known as its emigration rate. A better solution will have a greater emigration rate than a bad one. Conversely, the rate of immigration represents the amount by which a species departs from its natural environment. As a result, the bad option will have a larger immigration rate than the better one. The fundamental form of BBO has been represented by straight lines, as shown in Fig. 3. We have for the straight lines

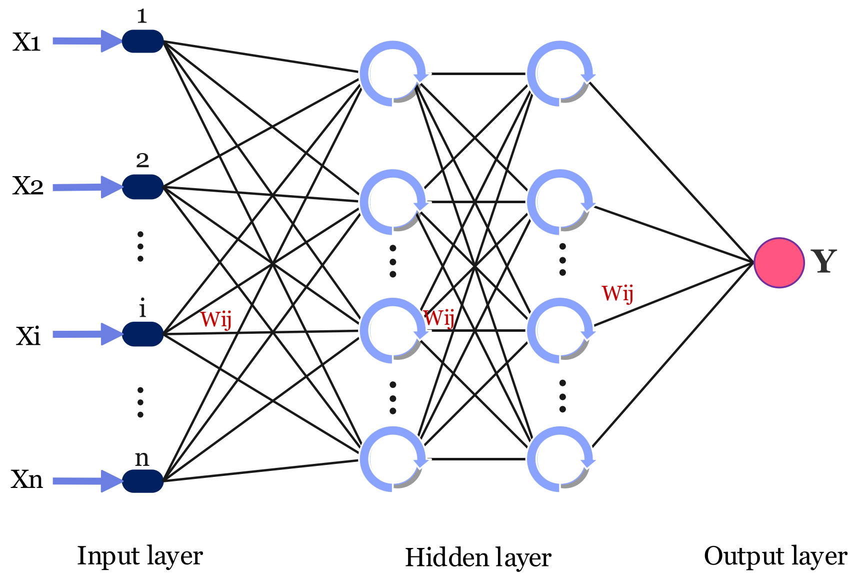

Fig. 1. RNN structure.

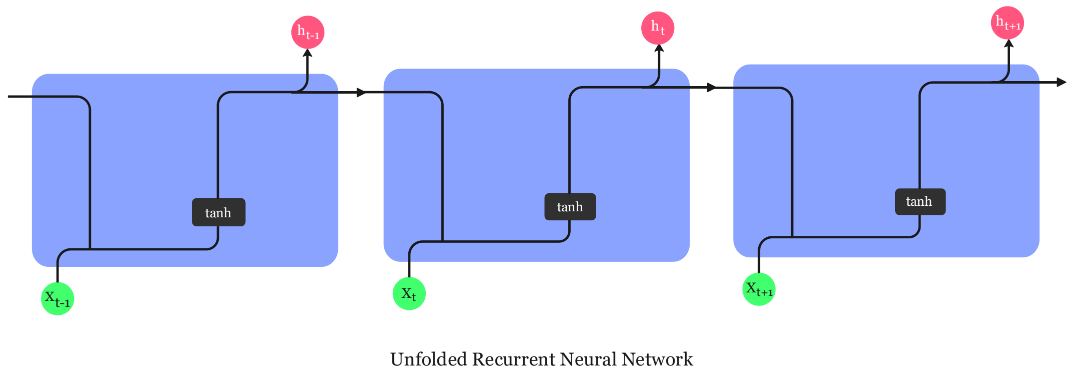

Fig. 2. The process of transforming data in RNN nodes.

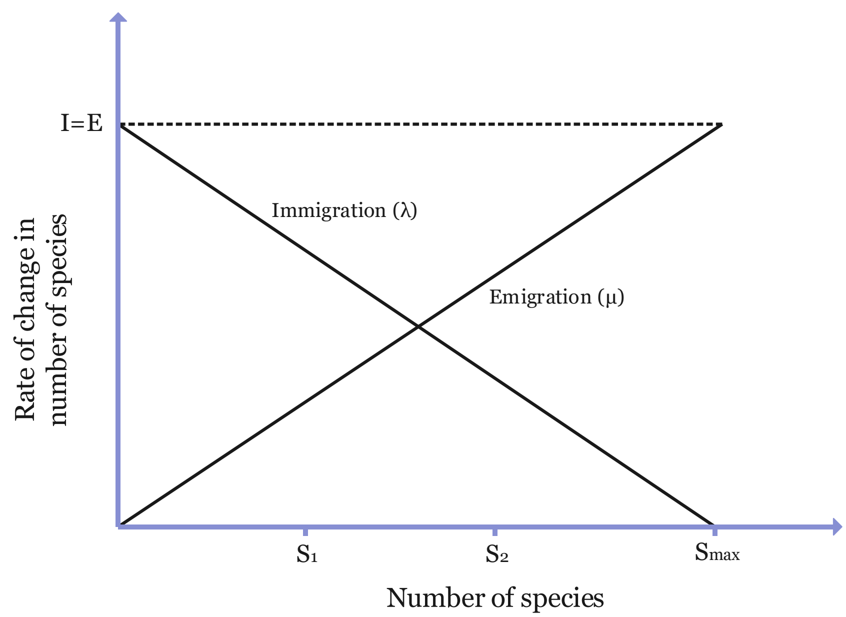

Fig. 3. The diagram of two habitats.

$$\mu_k = \frac{E \times k}{n} \lambda_k = I\left(1 - \frac{k}{n}\right) \tag{2}$$

where,

μ_k : The k^{th} habitat's emigration rate.

λ_k : The k^{th} habitat's emigration rate.

I : The highest rate of immigration.

E : The highest rate of emigration.

$\mathbf{n = S_{max}}$: The most species that a habitat can sustain.

k : The total number of species.

An abundance of species is indicated by a high HSI, which is characterized by a high rate of emigration and a low rate of immigration to nearby habitats. An increase in species leads to a decrease in immigration rate. The rate of emigration, however, rises in tandem with the number of species. Solutions S1 and S2 are the two possible options. For the most part, S2 is a good response, while S1 is a bad one. S1 has, on average, higher immigration rates than S2. Comparing S1 with S2 emigration rates, the former will be lower. The Fig. 3 represents the process of two habitats.

Mutation: A BBO mutation is comparable to an abrupt shift in environmental circumstances brought on by other events such as a tornado, volcanic eruption, or natural disaster. The species migrates to a new habitat when its old one becomes unsuitable for survival, as shown by the random change in the solution.

C. Aquila Optimization

The AO algorithm, a novel one, was released in 2021 [18]. The four categories in this algorithm were inspired by the hunting tactics of the raptor bird Aquila. In the first category, birds of prey are tracked down and pursued while flying high in the air. The swift attack of prey at low altitudes close to the ground is carried out by the second category, which glides. With a slow descent and low-altitude flight, the third group gradually attacks its prey. Using diving to catch terrestrial prey is the fourth category. Quick acceleration, convergence, and stability are all made possible by the AO algorithm's potent optimization capabilities [18]. The reliability and consistency of it are also very high. The act of a vertical dive is what an eagle does when it spots a potential prey area. The bird quickly determines the ideal hunting location on the ground by flying at great altitudes. The most efficient course of action is determined using an equation that takes the search area into account. Fig. 4 is an example of the one of hunting strategies of this bird.

$$\begin{cases} Z_1(t+1) = Z_{best}(t) \times \left(1 - \frac{t}{T}\right) \\ \quad +\left(Z_M(t) - Z_{best}(t) \times \text{ rand }\right) \\ \qquad Z_M(t) = \frac{1}{N}\sum_{i=1}^{N} Z_i(t), \\ \qquad \forall j = 1,2,\dots\dots, Dim \end{cases} \tag{3}$$

where, (t) is the best course of action, indicating the location of the closest target prey, and $Z(t+1)$ is the solution of generation $t+1$, produced by the search method $Z_1 \cdot Z_{best}(t)$. This iteration's t is the number. T is the maximum number of iterations that can be done. $Z(t)$ is a visual representation of the current solution's position mean at the t-th iteration. The name of a random integer between 0 and 1 is Rand. The following rapid gliding attack: The eagle soars to a height to identify the prey region in order to reduce the hunting zone or the search space for the most effective response according to the equation that follows:

$$\begin{cases} Z_2(t+1) = Z_{best}(t) \times Z(D) \\ \quad +Z_R(t) + (y - z) \times \text{ rand} \\ \qquad L(D) = s \times \frac{\mu \times \sigma}{|\nu|^{\frac{1}{\beta}}} \end{cases} \tag{4}$$

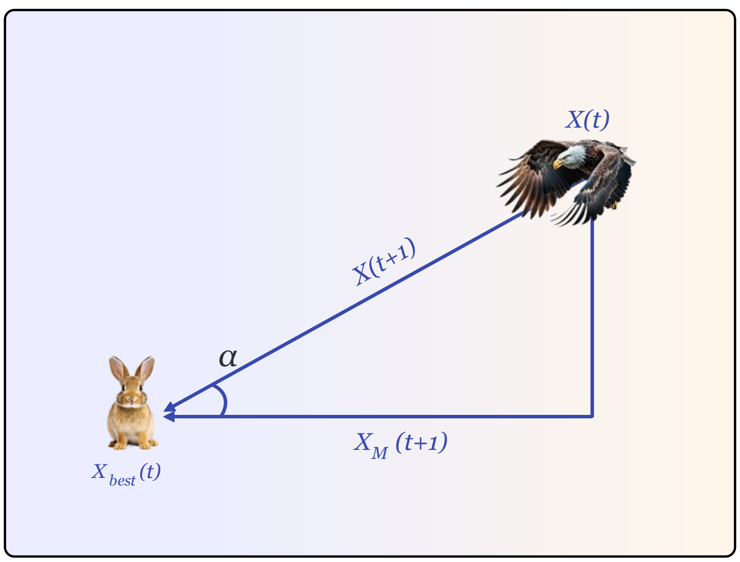

Fig. 4. The illustration of Aquila optimization.

where, $L(D)$ stands for the hunting flight distribution function, D for the dimensional space, and $Z_R(t)$ for the random solution between $[1, N]$. Once the prey region has been precisely identified and the Aquila is prepared to land and strike, it switches to the low-flying, slow-falling assault mode at the chosen target position. The third low-altitude flight pattern is this one. By employing this strategy, the bird may see how its prey would react and slowly move in its direction as in the following formula:

$$Z_3(t+1) = \left(Z_{best}(t) - Z_M(T)\right) \times \alpha - \text{rand} + \left((U_b - L_b) \times \text{rand} + L_b\right) \times \delta \quad (5)$$

where, α and δ are the modifying parameters. The upper limit of the given problem is U_b. The bottom limit of the problem is denoted by L_b. The fourth method of walking capture is the eagle striking the target from above while performing quick convergence using the following equation:

$$\begin{cases} Z_4(t+1) = Q_F \times Z_{best}(t) - (G_1 \times Z(t) \times \text{rand}) - G_2 \times L(D) + \text{rand} \times G_1 \\ Q_F(t) = \dfrac{2 \times \text{rand} - 1}{t^{(1-T)2}} \\ G_1 = 2 \times \text{rand} - 1 \\ G_2 = 2 \times \left(1 - \dfrac{t}{T}\right) \end{cases} \quad (6)$$

the quality function, or Q_F, and the search technique are balanced. Aquila's actions while pursuing its prey are fully visible on G_1. Aquila's flying slope when hunting is represented by G_2. $Z(t)$ is the answer for the current iteration.

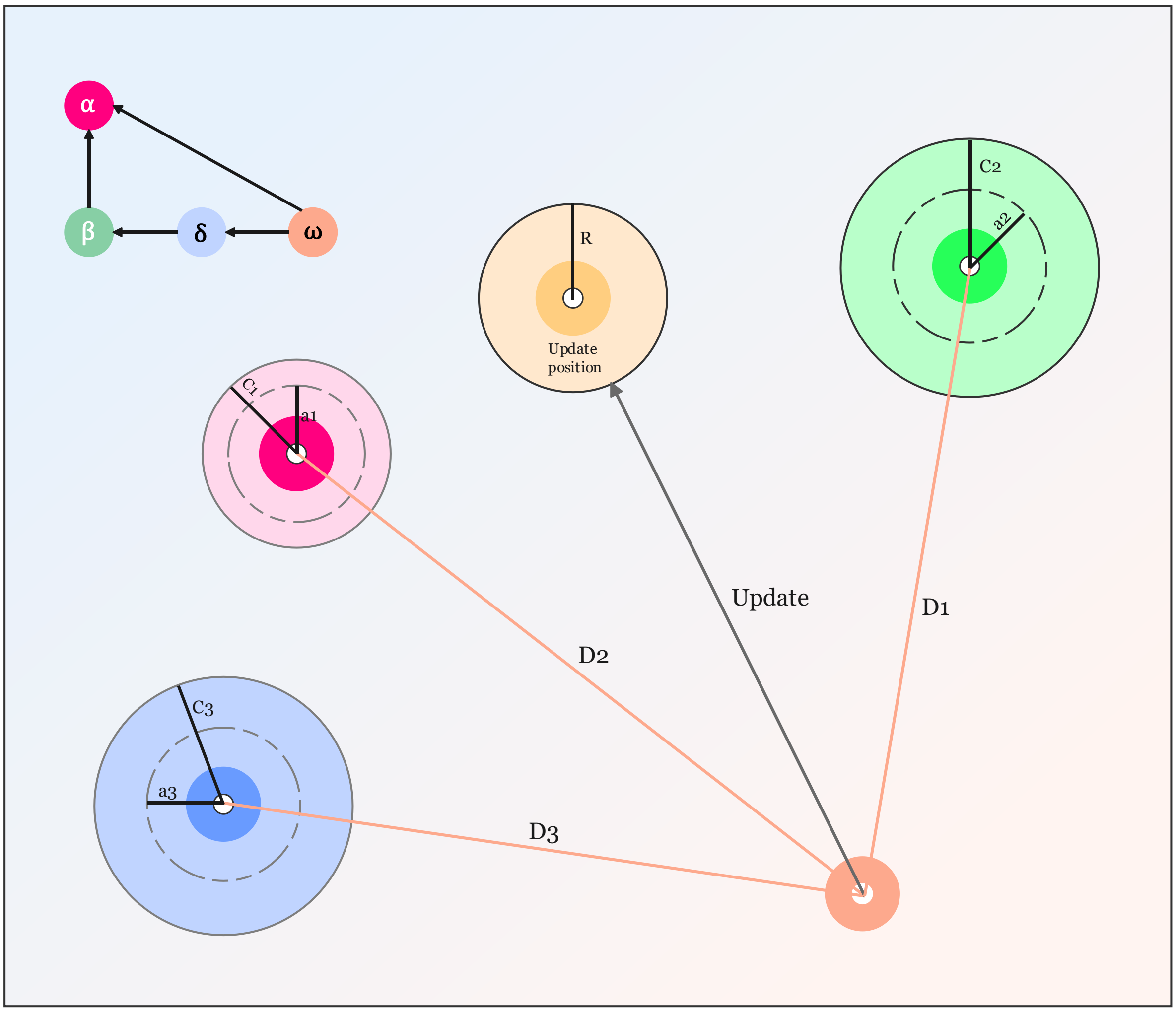

Fig. 5. The structure of Gray wolf optimizer.

D. Gray wolf optimizer

Utilizing a meta-heuristic approach, the Gray Wolf Optimizer, a unique optimization method, has been created. Mirjalili et al. [19] proposed a method, that mimics the gray wolf social structure and hunting methods. According to Fig. 5, the hierarchy of leadership contains four options: Alpha, Beta, Delta, and Omega, with Omega being the final challenger. Alpha is the best alternative and Fig. 6 represents how this optimizer works. The three main hunting methods used in the approach include chasing, encircling, and attacking prey in an effort to replicate wolf behavior. The equation illustrated below was used to simulate the movement of gray wolves during nature hunting:

$$\vec{D} = |\vec{C} \cdot \vec{X}_p(t) - \vec{X}(t)| \\ \vec{X} \mid (t+1) = \vec{X}_p(t) - \vec{A} \times \vec{D} \tag{7}$$

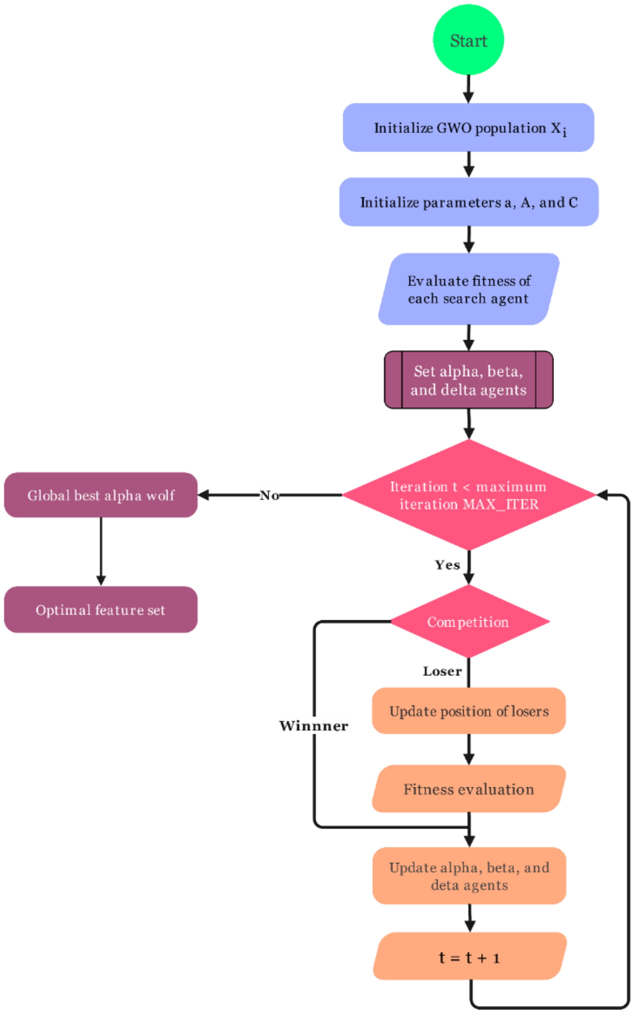

Fig. 6. Diagram of the Gray wolf optimizer.

where t is the current iteration, $\vec{D}$ stands for movement, $\vec{X}_p$ for prey location, $\vec{A}$ and $\vec{C}$ for coefficient vectors, and $\vec{X}$ stands for a gray wolf's position. The coefficient vectors($\vec{A}$ and $\vec{C}$) are built using the relationships shown below:

$$\vec{A} = 2\vec{a} \times \vec{r}_1 - \vec{a} \\ \vec{C} = 2 \times \vec{r}_2 \tag{8}$$

Using information from alpha, beta, and delta, the position of new search reps that include omegas is modified accordingly:

$$\vec{D}_a = |\vec{C}_1 \cdot \vec{X}_a - \vec{X}|, \vec{D}_\beta = |\vec{C}_2 \cdot \vec{X}_\beta - \vec{X}|, \vec{D}_\delta = |\vec{C}_3 \cdot \vec{X}_s - \vec{X}| \tag{9}$$

$$\vec{X}_1 = \vec{X}_a - \vec{A}_1 \cdot \vec{D}_u, \vec{X}_2 = \vec{X}_\beta - \vec{A}_2 \cdot \vec{D}_\beta, \vec{X}_3 = \vec{X}_b - \vec{A}_3 \cdot \vec{D}_\delta \tag{10}$$

$$\vec{X}(t+1) = \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3} \tag{11}$$

when the wolves undertake a last assault to accomplish the mission, as shown by the subscripts α, β, and δ. An is a random variable with a value between $-2\vec{a}$ and $2\vec{a}$, whereas an is used to simulate the previous assault by altering $\vec{a}$ value from 2 to 0. Therefore, decreasing a would also cause a decrease in $\vec{A}$. $|\vec{A}| < 1$ forced the wolves into sticking to their prey. Gray wolves hunt in groups and follow the alpha wolf, dispersing to forage and reassembling to attack. Whenever $|\vec{A}|$ has a random value bigger than unity, the wolves may split apart in pursuit of prey. The two setup parameters for the GWO method that are most important are the wolf count and generation number. Each generation is a wolf's final deed, and the quantity of wolves accurately reflects function assessments across time. This means that the total number of objective function evaluations will be equal to the wolf population times the size of the generation, or,

$$OFEs = N_W \times N_G \tag{12}$$

E. Data source and Preparation

When conducting a thorough analysis of a stock, it's important to take into account some factors, such as the trading volume and the Open, High, Low, and Close (OHLC) prices over a specific period. For this specific study, information on Alphabet Inc. began to be gathered in 2015. For each day during the specified period, this data included information on both the trading volume and the OHLC prices. Examining the data landscape carefully to spot any anomalies, outliers, or discrepancies that might have an impact on the accuracy of the findings was the first step. Following this analysis, the dataset underwent several cleaning and preparation steps, using various techniques like scaling and normalization to reduce errors and guarantee consistency in training results via using the following equation:

$$Xscaled = \frac{(X - Xmin)}{(Xmax - Xmin)} \tag{13}$$

This methodical approach aimed to raise the general level of data quality that served as the basis for the forecasting models. To further improve the models, the prepared data was split into two subsets, with 80% of the data used for training and the remaining 20% for testing and validation according to Fig. 7. The goal of this division was to strike a balance between the requirement for a sizable amount of data for model training and the requirement for a varied and untested set for exhaustive testing and validation.

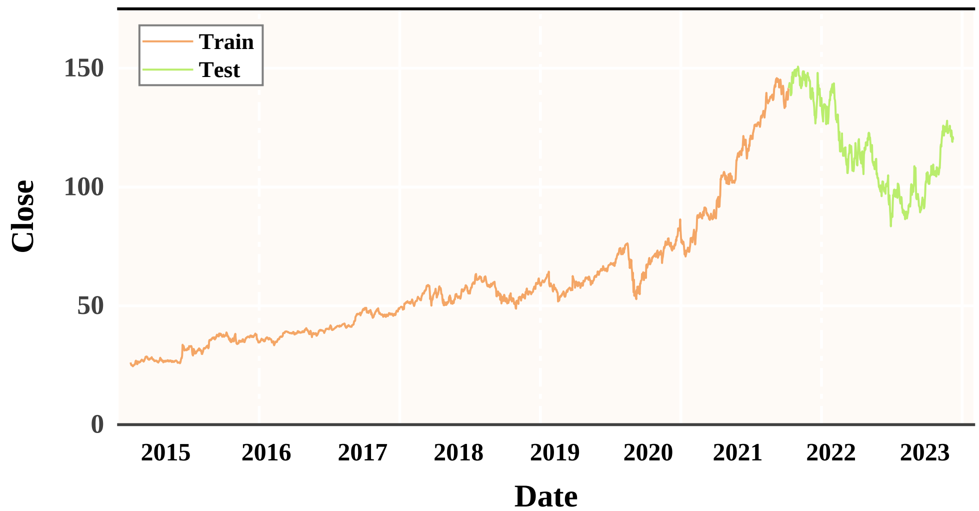

Fig. 7. Dividing the data into training and tests.

F. Assessment Criteria

In order to ensure the validity and accuracy of future projections, a diverse range of performance indicators were employed. These metrics were meticulously selected to provide a comprehensive evaluation of the prognostications. Throughout the evaluation process, several metrics, including the mean absolute error (MAE), which calculates the average absolute difference between the predicted and actual values, the root mean square error (RMSE), which measures the root mean square of the errors between the predicted and actual values, and the coefficient of determination (R^2), which gauges the proportion of variance in the dependent variable that is predictable from the independent variable was considered. These methodologies are highly valuable for assessing the accuracy of forecasting models and can facilitate more informed decisions.

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

$$MAE = \frac{\sum_{i=1}^{n}|y_i - \hat{y}_i|}{n} \tag{15}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{16}$$

IV. Experimental Results

A. Statistical Values

The study's component includes a table (see Table I) that presents comprehensive statistical information about the dataset. The table displays OHLC pricing and volume statistics, which help to clarify the data. Statistical metrics such as count, mean, minimum (Min), standard deviation (Std.), 25%, 75%, variance, and maximum (Max) values are also provided, enabling a more detailed and precise analysis of the data.

B. Results of the Each Model

The primary goal of this study is to identify and assess the top hybrid algorithm for stock price forecasting. To do this, the research created forecasting models and looked at intricate factors that affect stock market patterns. The objective is to deliver analytical information that can help analysts and investors make wise investment decisions. Each model's performance is comprehensively evaluated, along with a thorough study of its efficacy, in Table II, Fig. 8, and Fig. 9.

TABLE I. Statistical Summary of the Data Set

	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
Variance	1193.43	1223.381	1165.986	243.5536	1194.334

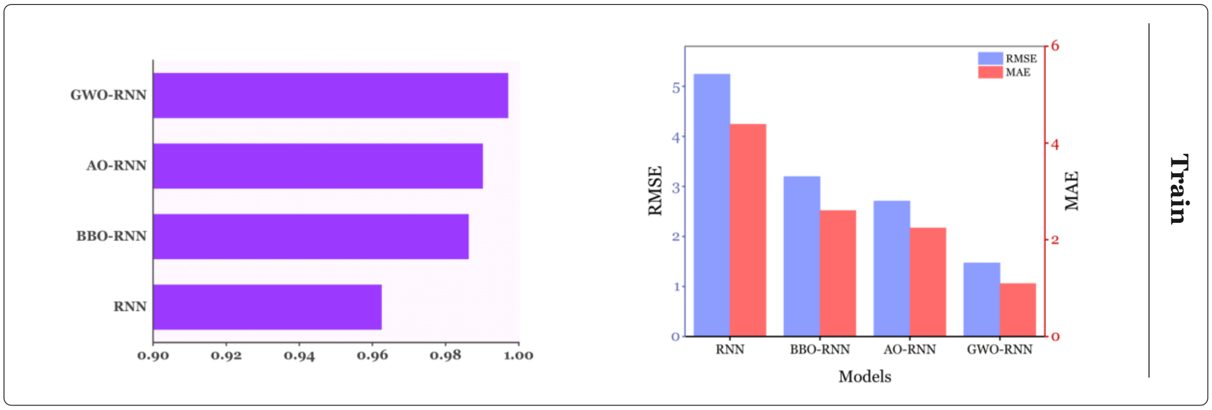

Fig. 8. Values for each model's training-related assessment.

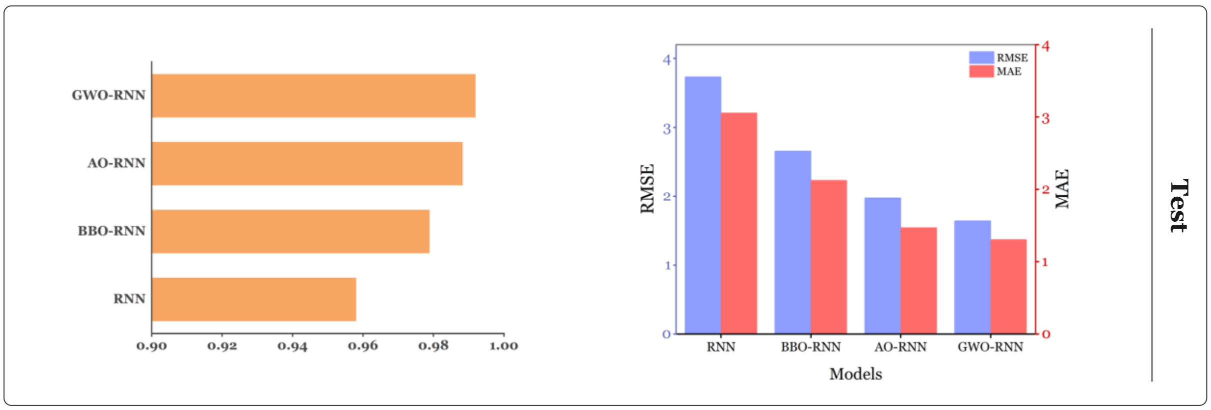

Fig. 9. Values for each model's testing-related assessment.

TABLE II. THE RESULTS OF THE MODELS FORECAST EVALUATION

	TRAIN SET			TEST SET		
	R^2	***RMSE***	***MAE***	***R^2***	***RMSE***	***MAE***
RNN	0.962	5.241	4.389	0.958	3.733	3.058
BBO-RNN	0.986	3.193	2.604	0.979	2.654	2.126
AO-RNN	0.990	2.705	2.244	0.988	1.976	1.472
GWO-RNN	0.997	1.467	1.094	0.992	1.643	1.306

Three widely accepted metrics—RMSE, MAE, and R^2—were used to carry out a detailed review of the data analysis. These indicators are well known for being able to offer a precise assessment of the analysis's dependability, correctness, and overall effectiveness. With and without an optimizer, the RNN model's performance was evaluated using the R^2, RMSE, and MAE criteria. This evaluation improved comprehension of the model's performance and aided in making decisions based on the outcomes.

V. DISCUSSION

A. Analysis of the Models

This study's principal objective is to identify and evaluate the most effective hybrid algorithm for stock price prediction. In order to accomplish this, the study developed forecasting models and examined complex variables that influence stock market patterns. The primary aim is to provide analysts and investors with valuable analytical data that can assist them in making informed investment choices. The performance of each model is assessed exhaustively, and its effectiveness is thoroughly examined in Table II, Fig. 8, and Fig. 9. The R^2, RMSE, and MAE criteria were used to assess the performance of the RNN model both with and without the optimizer. This enabled wise decisions by giving a thorough grasp of the model's performance. Analysis of the training and test sets revealed that, without the optimizer, the RNN model generated R^2 values for training and testing of 0.962 and 0.958, respectively. The RMSE values for training and testing were 5.241 and 3.733, respectively, despite the MAE values being 4.389 and 3.058. By integrating optimizers, the RNN model performs significantly better. For instance, when the BBO optimizer was used, the R^2 value for the tests increased to 0.979. Furthermore, the RMSE and MAE values for testing, which came in at 2.654 and 2.126, respectively, were lower than those for training. The AO-RNN model performed better than the BBO-RNN model and produced better results. Specifically, the results for the training and testing R^2 were

0.990 and 0.988, respectively. It's crucial to remember that the MAE and RMSE figures fell to 1.472 and 1.976 during the testing. This shows increased precision. The GWO-RNN model has demonstrated remarkable accuracy and reliability in regression analysis. The model achieved the highest R^2 scores of 0.997 and 0.992 for the training and testing datasets, respectively. This indicates that the model has a high predictive power and can explain almost all the variability in the data. Furthermore, the model's performance is exceptional, as evidenced by the low MAE and RMSE training values of 1.094 and 1.306, and testing values of 1.467 and 1.643, respectively. These values represent the amount of error between the actual and predicted values, with lower values indicating higher accuracy. Therefore, the GWO-RNN model has shown exceptional accuracy in both the training and testing data sets.

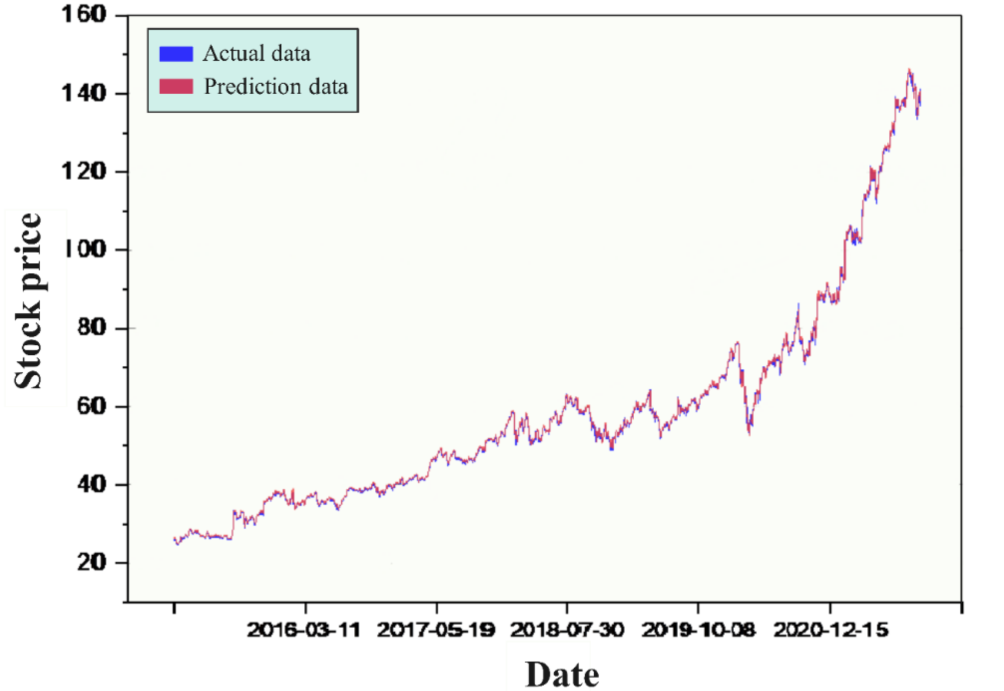

Fig. 10. The prediction curve is produced by training with GWO-RNN.

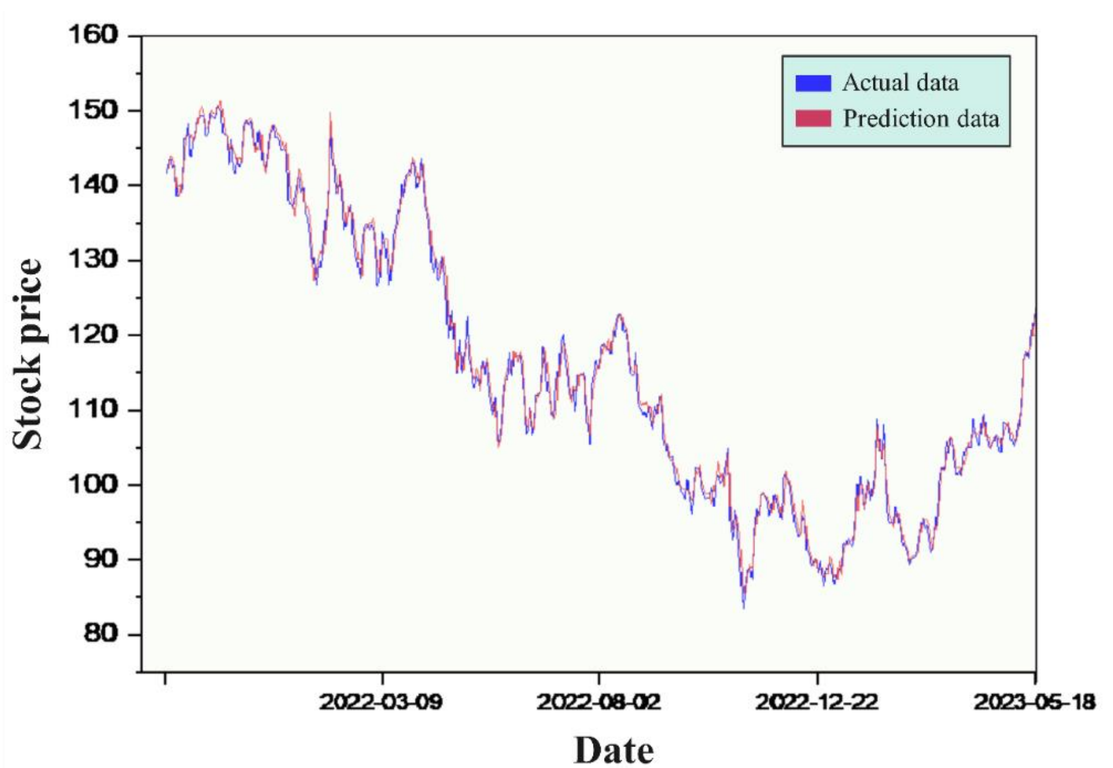

Fig. 11. The prediction curve is produced by testing with GWO-RNN.

For creating accurate stock price predictions, the GWO-RNN model is a very trustworthy instrument. As illustrated in Fig. 10 and Fig. 11, this model is successful in forecasting the Alphabet stock curves. Due to the RNN method's ability to lower price fluctuations, streamline trend prediction, and increase model precision, the GWO-RNN model performs better than other models at forecasting stock prices. The capacity of the GWO-RNN model to learn from prior data sets is one of its distinguishing characteristics. Learning from prior data sets is essential for a model to produce accurate stock value predictions and adapt to shifting market trends. In summary, the GWO-RNN model is an effective and useful tool for forecasting stock prices. Because of its accuracy, precision, and adaptability, it is highly advised for anyone wishing to make profitable transactions in the stock market. It stands out from other models due to its use of the RNN algorithm and GWO optimizer, making it the greatest option for anyone looking to make informed investing decisions.

B. Comparison with Recent Works

Validation measures and comparisons with prior relevant literature are essential elements in evaluating the credibility and significance of a research investigation. In addition to guaranteeing the dependability and accuracy of the study's findings, they aid in contextualizing the research within a wider framework. The present evaluation assesses, which is illustrated in Table III, the predictive capabilities of different models, such as the GWO-RNN model employed in this research, concerning the behavior of the stock market. Significantly, the GWO-RNN model, which was developed and evaluated using Alphabet stock data, attains a remarkable coefficient of determination (R^2 = 0.992), outperforming alternative algorithms including Linear Regression, SVM, and several iterations of LSTM. The remarkable accuracy of this forecast highlights the consistency and consistency of the GWO-RNN model in capturing the intricate dynamics of stock prices. Through the utilization of the Grey Wolf Optimization-Recurrent Neural Network architecture, the model adeptly exploits past stock price data to deliver resilient predictions, thereby showcasing its consistency in adjusting to fluctuations in the market. Incorporating Alphabet stock data into the evaluation process enhances coherence and ensures that the model's performance is in line with actual market conditions. As a result, the GWO-RNN model can be established as a logical and consistent methodology for forecasting the stock market, providing significant contributions to financial decision-making and risk mitigation.

TABLE III. AN EVALUATION OF THE MODEL IN COMPARISON TO PRIOR RESEARCH

Authors	**Abdul et al.** [31]			**Zhu et al.** [32]						**Present work**
Algorithms	Linear regression	SVM	MLS-LSTM	LSTM	EMD-LSTM	CEEMDAN-LSTM	SC-LSTM	EMD-SC-LSTM	CEEMDAN-SC-LSTM	GWO-RNN
R^2	0.73	0.93	0.95	0.6896	0.8703	0.9031	0.6871	0.9111	0.9206	0.992

C. Limitations and Future Works

The efficacy of the GWO-RNN model was assessed using a distinct dataset comprising Alphabet stock from January 1, 2015, to June 29, 2023. The efficacy of the model could potentially be compromised by the attributes of the given dataset, thereby restricting its applicability to alternative equities, industries, or periods. Hybrid models frequently depend on precise parameter configurations, particularly when optimization algorithms such as grey wolf optimization are integrated. The extent to which the GWO-RNN model can withstand variations in these parameters may have been inadequately investigated in the research. The exclusive dependence of the study on historical stock price data may result in the omission of critical information, including geopolitical developments, external economic events, and market sentiment. The extent to which the GWO-RNN model captures and incorporates such external factors into its predictions is not exhaustively investigated, which may have implications for its overall predictive capabilities. Although the study showcases the superior performance of the GWO-RNN model in comparison to other models, a more extensive evaluation comparing it to a broader spectrum of established and cutting-edge models would offer a more precise assessment of its comparative merits and drawbacks.

Enhancing the range of data sources beyond daily prices and trading volume may yield significant insights that can be applied to the prediction of stock market trends. The incorporation of supplementary data streams, including sentiment analysis from social media platforms, economic indicators, or news articles, may bolster the predictive capacities of models through the inclusion of supplementary market dynamics. The incorporation of external variables into predictive models, including macroeconomic indicators, geopolitical events, and regulatory changes, may increase their predictive capability. Gaining insight into how these extraneous variables impact the conduct of the stock market and integrating them into prognostic models may enhance their precision and dependability. Undertaking thorough assessments of predictive models under various market conditions, encompassing periods of stability and volatility, would yield valuable insights regarding the models' ability to withstand and apply to a wide range of situations. Conducting performance tests on models across diverse economic conditions may unveil their merits and drawbacks, thereby providing valuable insights for their implementation in practical situations. By expanding the utilization of predictive models to encompass financial methods other than stock prices, including commodities, currencies, and cryptocurrencies, their applicability and significance could be significantly enhanced. An examination of the adaptability of predictive models to various asset classes and market segments would yield valuable insights regarding their efficacy and versatility in a wide range of financial markets.

VI. CONCLUSIONS

The task of predicting stock prices is difficult and complex because it requires examining many different aspects of society, the economy, and politics. Since the stock market is dynamic and ever-evolving, it is important to take financial statements, earnings reports, market trends, and other factors into account when forecasting future stock values. The stock market's behavior can be significantly impacted by macroeconomic factors like interest rates, inflation, and global market conditions. Due to the complexity and numerous variables involved in predicting stock values, it can be challenging to develop accurate and trustworthy prediction models. Making accurate predictions requires an understanding of the market's erratic and non-linear characteristics. In this study, the RNN, BBO-RNN, and AO-RNN stock price prediction models were evaluated for their performance.

The GWO-RNN model performed better than other models in the tests that were conducted. R^2, RMSE and MAE values for the model were 0.992, 1.643, and 1.306, respectively. These findings show that the GWO-RNN model is highly predictively accurate and that it can be trusted to deliver accurate and dependable outputs.

The GWO-RNN model has been demonstrated to be a trustworthy tool for making highly accurate stock price predictions. This model can analyze and interpret large amounts of data in real-time using a combination of recurrent neural networks and grey wolf optimization-based optimization, giving investors useful insights into market trends and potential investment opportunities.

ACKNOWLEDGMENT

This work was supported by the Key Courses Construction Project of Xi'an Eurasia University (2019KC026); the Project of Shaanxi Provincial Sports Bureau (2022307); The Innovation Team of Eurasia University (2021XJTD07); Xi'an Social Science Foundation Project (23JX116); Xi'an Social Science Foundation Project (23JX118).

REFERENCES

[1] B. Qian and K. Rasheed, “Stock market prediction with multiple classifiers,” Applied Intelligence, vol. 26, no. 1, pp. 25–33, 2007, doi: 10.1007/s10489-006-0001-7.

[2] D. Kumar, P. K. Sarangi, and R. Verma, “A systematic review of stock market prediction using machine learning and statistical techniques,” Mater Today Proc, vol. 49, no. September, pp. 3187–3191, 2020, doi: 10.1016/j.matpr.2020.11.399.

[3] C. Zhang, J. Ding, J. Zhan, and D. Li, “Incomplete three-way multi-attribute group decision making based on adjustable multigranulation Pythagorean fuzzy probabilistic rough sets,” International Journal of Approximate Reasoning, vol. 147, pp. 40–59, 2022, doi: https://doi.org/10.1016/j.ijar.2022.05.004.

[4] Y. Chen, P. Zhao, Z. Zhang, J. Bai, and Y. Guo, “A Stock Price Forecasting Model Integrating Complementary Ensemble Empirical Mode Decomposition and Independent Component Analysis,” International Journal of Computational Intelligence Systems, vol. 15, no. 1, 2022, doi: 10.1007/s44196-022-00140-2.

[5] E. F. Fama, “Random walks in stock market prices,” Financial analysts journal, vol. 51, no. 1, pp. 75–80, 1995.

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

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

[8] 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.

[9] J. V Hansen, J. B. McDonald, and R. D. Nelson, “Time series prediction with Genetic-Algorithm designed neural networks: An empirical comparison with modern statistical models,” Comput Intell, vol. 15, no. 3, pp. 171–184, 1999.

[10] Y. S. Abu-Mostafa and A. F. Atiya, “Introduction to financial forecasting,” Applied Intelligence, vol. 6, no. 3, pp. 205–213, 1996, doi: 10.1007/BF00126626.

[11] R. Chiong, Z. Fan, Z. Hu, and S. Dhakal, “A Novel Ensemble Learning Approach for Stock Market Prediction Based on Sentiment Analysis and the Sliding Window Method,” IEEE Trans Comput Soc Syst, vol. 10, no. 5, pp. 2613–2623, 2023, doi: 10.1109/TCSS.2022.3182375.

[12] W. Freeborough and T. van Zyl, “Investigating Explainability Methods in Recurrent Neural Network Architectures for Financial Time Series Data,” Applied Sciences (Switzerland), vol. 12, no. 3, 2022, doi: 10.3390/app12031427.

[13] I. Boussaïd, J. Lepagnot, and P. Siarry, “A survey on optimization metaheuristics,” Inf Sci (N Y), vol. 237, pp. 82–117, 2013.

[14] A. R. Simpson, G. C. Dandy, and L. J. Murphy, “Genetic algorithms compared to other techniques for pipe optimization,” J Water Resour Plan Manag, vol. 120, no. 4, pp. 423–443, 1994.

[15] T. Back, Evolutionary algorithms in theory and practice: evolution strategies, evolutionary programming, genetic algorithms. Oxford university press, 1996.

[16] S. Mirjalili, “The ant lion optimizer,” Advances in Engineering Software, vol. 83, pp. 80–98, 2015, doi: 10.1016/j.advengsoft.2015.01.010.

[17] D. Simon, “Biogeography-based optimization,” IEEE transactions on evolutionary computation, vol. 12, no. 6, pp. 702–713, 2008.

[18] L. Abualigah, D. Yousri, M. Abd Elaziz, A. A. Ewees, M. A. A. Al-qaness, and A. H. Gandomi, “Aquila Optimizer: A novel meta-heuristic optimization algorithm,” Comput Ind Eng, vol. 157, p. 107250, 2021, doi: https://doi.org/10.1016/j.cie.2021.107250.

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

[20] S. Agarwal, P. Rajput, and A. K. Jena, “A Hybrid Evolutionary model for Stock Price Prediction Using Grey Wolf Optimizer,” in 2022 OITS International Conference on Information Technology (OCIT), 2022, pp. 1–6. doi: 10.1109/OCIT56763.2022.00062.

[21] S. Kumar Chandar, “Grey Wolf optimization-Elman neural network model for stock price prediction,” Soft comput, vol. 25, no. 1, pp. 649–658, 2021, doi: 10.1007/s00500-020-05174-2.

[22] S. Sahoo and M. N. Mohanty, “Stock market price prediction employing artificial neural network optimized by gray wolf optimization,” in New Paradigm in Decision Science and Management: Proceedings of ICDSM 2018, Springer, 2020, pp. 77–87.

[23] F. W. Christanto, V. G. Utomo, R. Prathivi, and C. Dewi, “The Impact of Financial Statement Integration in Machine Learning for Stock Price Prediction,” International Journal of Information Technology and Computer Science ; volume 16, issue 1, page 35-42 ; ISSN 2074-9007 2074-9015, 2024, doi: 10.5815/ijitcs.2024.01.04.

[24] Y. Chen, X. Andrew, and S. Supasanya, “CRISIS ALERT:Forecasting Stock Market Crisis Events Using Machine Learning Methods,” 2024. doi: 10.48550/arxiv.2401.06172.

[25] P.-F. Tsai, C.-H. Gao, and S.-M. Yuan, “Stock Selection Using Machine Learning Based on Financial Ratios,” Mathematics, Vol 11, Iss 23, p 4758 (2023), Mar. 2023, doi: 10.3390/math11234758.

[26] S. Pourroostaei Ardakani, N. Du, C. Lin, J. Yang, Z. Bi, and L. Chen, “A federated learning-enabled predictive analysis to forecast stock market trends,” Mar. 2023, [Online]. Available: https://eprints.lincoln.ac.uk/id/eprint/53623/

[27] M. Hani'ah, M. Z. Abdullah, W. I. Sabilla, S. Akbar, and D. R. Shafara, "Google Trends and Technical Indicator based Machine Learning for Stock Market Prediction," MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer; Vol 22 No 2 (2023); 271-284 ; 2476-9843 ; 1858-4144 ; 10.30812/matrik.v22i2, Mar. 2023, [Online]. Available: https://journal.universitasbumigora.ac.id/index.php/matrik/article/view/2287

[28] K. Juare and A. Kulkarni, "Machine Learning Algorithms for Stock Market Prediction," International Journal of Innovative Science and Research Technology 7(12) 2193-2199, Mar. 2023, [Online]. Available: https://zenodo.org/record/7698476

[29] T. Swathi, N. Kasiviswanath, and A. A. Rao, "A Novel Sine Cosine Optimization with Stacked Long Short-term Memory-enabled Stock Price Prediction," Recent Advances in Computer Science and Communications ; volume 16 ; ISSN 2666-2558, 2023, doi: 10.2174/0126662558236061230922074642.

[30] Z. C. Lipton, J. Berkowitz, and C. Elkan, "A critical review of recurrent neural networks for sequence learning," arXiv preprint arXiv:1506.00019, 2015.

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

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