

# Prediction of Financial Markets Utilizing an Innovatively Optimized Hybrid Model: A Case Study of the Hang Seng Index

Xiaopeng YANG

School of Economics and Management, Weifang University, Weifang 261061, Shandong, China

**Abstract**—Stock trading is a highly consequential and frequently discussed subject in the realm of financial markets. Due to the volatile and unpredictable nature of stock prices, investors are perpetually seeking methods to forecast future trends in order to minimize losses and maximize profits. Nevertheless, despite the ongoing investigation of various approaches to optimize the predictive efficacy of models, it is indisputable that a method for accurately forecasting forthcoming market trends does not yet exist. A multitude of algorithms are currently being employed to forecast stock prices due to significant developments that have occurred in recent years. An innovative algorithm for predicting stock prices are examined in this paper which is a Gated Recurrent Unit combined with the Aquila optimizer. A comprehensive data implementation utilizing the Hang Seng Index stock price was executed as a dataset of this research which was collected between the years of 2015 and the end of June 2023. In the study, several additional methods for predicting stock market movements are also detailed. A comprehensive comparative analysis of the stock price prediction performances of the aforementioned algorithms has also been carried out to offer a more in-depth analysis and then the results are displayed in an understandable tabular and graphical manner. The proposed model obtained the values of 0.9934, 0.71, 143.62, and 36530.58, for  $R^2$ , MAPE, MAE, and MSE, respectively. These results proved the efficiency and accuracy of the suggested method and it was determined that the proposed model algorithm produces results with a high degree of accuracy and performs the best when it comes to forecasting a time series or stock price.

**Keywords**—Financial markets; stock future trend; Hang Seng Index; Gated Recurrent Units; Aquila Optimizer

## I. INTRODUCTION

The stock market, as defined, is a marketplace wherein individuals engage in the buying and selling of stocks associated with certain companies. Over time, the values of these stocks exhibit significant fluctuations. Nevertheless, it would be unwise to disregard the factors contributing to the significant fluctuations in the stock market, which may include political influences, brand perception, and the prevailing global conditions. The aforementioned elements have the ability to significantly influence the perspectives and convictions of prospective investors, hence contributing to fluctuating patterns in the stock market. Hence, while it is crucial to comprehend the potential variables contributing to these fluctuations, it remains insufficient to devise a methodology that can reliably forecast trends in light of perpetual worldwide transformations

and uncertainties [1]. Given the inherent unpredictability and significant market volatility, a considerable number of individuals interested in the stock market want to acquire a tool or method that can dependably forecast market trends, so enabling them to achieve more profitability [1]. Nonetheless, persistent endeavors are being undertaken to construct a model or algorithm that can assist investors in forecasting changes with more precision than previously achieved. The utilization of machine learning (ML) algorithms is a prevalent and well-accepted approach for constructing predictive models [2]. ML is a computational paradigm in which computers acquire information and make predictions based on previous experiences and training, without relying on external programming [2]. A few approaches and algorithms linked to ML have been investigated and addressed in [3-5]. Many advancements in data science and ML have occurred in the last few years, which have led to the creation of certain specific algorithms that are highly effective for predictive analytics across all sectors. Among the models examined in this article is GRU, a subfield of machine learning.

Recurrent neural networks (RNNs), such as Gated Recurrent Units (GRUs) [5], are used to handle sequential data, including time series. It was presented as an abridged form of the long short-term memory (LSTM) architecture. GRU is intended to address the vanishing gradient issue in RNNs and allow the network to store data over extended sequences, much like LSTM. The GRU model has one fewer gate structure than the LSTM; it consists of an update gate and a reset gate. It has been demonstrated that GRU performs comparably to LSTM while processing time series data, therefore this difference in operation impact is not very important. Compared to LSTM, GRU can converge more quickly because of its streamlined structure, which also speeds up training. Applications for the GRU model include speech recognition, video analysis, and natural language processing [5]. GRU was utilized by Ya Gao et al. [6] to predict stocks. GRU Neural Network Based on CEEMDAN-Wavelet was utilized by Chenyang Qi et al. [7] to predict stock prices.

The accuracy of predicting the value of the stock market rose as optimizers were developed and combined with a range of models to provide better outcomes in the predictions. Among the optimizers that were demonstrated were the whale optimization algorithm (WOA) [8], biogeography-based optimization (BBO) [9], genetic algorithm (GA) [10], moth-flame optimization (MFO) [11], ant lion optimization (ALO) [12], grey wolf optimization (GWO) [13], and Aquila

optimizer (AO) [14]. The most recent method to replicate the four distinct stages of Aquila hunting behavior is the AO, which was put out by Abualigah et al. [14]. Aquila employs four main hunting techniques: strolling and capturing prey, contour flying with a brief glide, low flying with a slow drop, and high soar with a vertical stoop [15]. These four core Aquila hunt processes served as the inspiration for the creation of the AO, a nature-inspired optimization algorithm that fundamentally clarifies the actions of each hunt stage. Initialization, Expanded Exploration, Narrowed Exploration, Expanded Exploitation, and Narrowed Exploitation are the five main processes that the traditional AO concentrates on. One of the most important aspects of the algorithm, the current iteration maximum iteration, usually guides the AO algorithm from the exploration to the exploitation stage. The exploration phase will be activated if the condition mentioned is true; otherwise, the exploitation step will be carried out [15].

From the commencement of 2015 until the conclusion of June 2023, daily transaction data from the Hang Seng Index (HSI) was collected, encompassing the following metrics: opening price, closing price, highest price, lowest price, and trading volume. To assess the reliability of each model, the research looked at a variety of models, including GRU, ALO-GRU, GWO-GRU, and AO-GRU. For this post, the AO-GRU model was selected since it has the best performance. The rest of the paper is structured as follows. The literature review is given in the Section II. Numerous analytical methods, including optimizer approaches, and the GRU model along with the dataset were provided in Section III. The study's findings are provided in Section IV and their discussions are demonstrated in Section V. Section VI provides a quick summary of the research's findings.

## II. LITERATURE REVIEW

In recent years, there has been a significant increase in the utilization of machine learning algorithms for the purpose of forecasting the stock market. A comparative study of fundamental analysis, technical analysis, and machine learning (ML) approaches was what Christanto et al. [16] proposed as an investigation into methodologies utilized in the capital market to forecast stock prices. For predicting stock prices, they employed Support Vector Regression (SVR) and Support Vector Machine (SVM) as ML techniques. Technical-only (TEC), financial statement-only (FIN), and a combination of the two (COM) parameter groups are assessed. Financial statement integration had a neutral effect on SVR predictions but a positive effect on SVM predictions, according to their experiments. The model achieved an accuracy rate of 83% in the conducted study. Chen et al. [17] examined the historical backdrop of economic recessions, highlighting the sudden and catastrophic consequences of occurrences such as the 2008 financial crisis, characterized by a substantial decrease in the SP 500. Driven by the prospective advantages of timely crisis detection, they implemented sophisticated machine learning methodologies, including Random Forest and Extreme Gradient Boosting, to forecast possible market downturns in the United States. Comparing the performance of these approaches, their research seeks to ascertain which model is more accurate at predicting US stock market crashes. Market indicators for crisis prediction were analyzed by employing

daily financial market data and 75 explanatory variables, which encompass general US stock market indexes as well as sector indexes. By employing particular classification metrics, they derived conclusions concerning the efficacy of their predictive models. Tsai et al. [18] discussed investors' interest in stock prediction, especially with the recent use of machine learning to improve accuracy. Machine learning works in technical, fundamental, and sentiment analysis, according to prior research. They discussed fiscal year-end selection and how misaligned reporting periods affect comparability and investment decisions. They emphasized synchronized fiscal years and use machine learning models for fundamental analysis to forecast Taiwan (TW) stock market returns. They created stock portfolios with higher predicted returns using Random Forest (RF), Feedforward Neural Network (FNN), Gated Recurrent Unit (GRU), and Financial Graph Attention Network (FinGAT) models. These portfolios outperformed TW50 index benchmarks in returns and portfolio scores, according to their study. Machine learning models were beneficial for stock market analysis and investment decision-making, according to Tsai et al. [18]. Ardakani et al. [19] proposed a federated learning framework for stock market prediction using Random Forest, Support Vector Machine, and Linear Regression models. They compared federated learning to centralized and decentralized frameworks to find the best approach. Federated learning outperformed centralized and decentralized frameworks in Mean Square Error (MSE) using Random Forest (0.021) and Support Vector Machine (37.596). Linear regression model-based centralized learning (MSE = 0.011) outperformed federated and decentralized frameworks. Federated learning had a lower model training delay than benchmarks for Linear Regression (9.7 s) and Random Forest (515 s), while decentralized learning saves time for Support Vector Machine (3847 s). Their findings illuminated stock market prediction learning framework strategies [19]. A novel stock price prediction method by Mamluatul et al. [20] uses machine learning, stock price data, technical indicators, and Google trends. SVR, MLP, and Multiple Linear Regression were used to predict stock prices. SVR predicts Indonesian stock prices better than MLP and Multiple Linear Regression with a MAPE of 0.50%. They found that SVR predicts stock prices accurately, helping investors make informed stock market decisions [20]. The importance of stock market forecasts in financial market profits was stressed by Juare et al. [21]. Their research used Random Forest, Support Vector Machine, KNN, and Logistic Regression to predict stock market trends. These algorithms are evaluated using accuracy, recall, precision, and F-Score. The main goal was to find the best stock market prediction algorithm. Effective forecasts can benefit stock exchanges and investors, highlighting the importance of predictive models in financial decision-making [21]. Swathi et al. [22] emphasized the need for investors to use stock price prediction (SPP) models in the global financial market for profit. Earlier SPP models used statistical and machine learning (ML) methods. In their study, the authors introduced SCODL-SPP, a stock price prediction method using Sine Cosine Optimization (SCO) and deep learning. The SCODL-SPP model forecasts share closing prices using deep learning and a stacked long short-term memory (SLSTM) model. The SLSTM model hyperparameters are optimized

using the SCO algorithm after the min-max normalization of primary data. The SCODL-SPP model outperformed other models in stock price prediction accuracy, according to experiments [22]. Su et al. [23] proposed a stack framework using LGBM to predict the Taiwan stock market index. Their study created a comprehensive feature set to account for political events, economic conditions, investor psychology, and global market trends that affect stock market predictions. They introduced a feature selection algorithm to identify important features and improve training performance. A stacking strategy integrates multiple classifiers to improve prediction accuracy. The proposed model is tested using a 10-year Taiwan Stock Exchange Capitalization Weighted Stock Index dataset. Both the prediction model and feature selection method performed well in experiments, indicating that the proposed approach is effective in stock market index prediction [23]. Ryan Chipwanya's study examined how stock market prediction tools and data have improved, making market predictions possible [24]. Logistic regression, decision trees, and random forest algorithms were compared for predicting Japanese stock market asset movements using machine learning models for time-series forecasting. The models were also compared to feedforward deep neural networks. Overall, all models achieved directional bias forecasting accuracy above 50% [24]. Pardeshi et al. [25] emphasized the importance of stock market prediction for profitable investing. To address complex financial market dynamics, they emphasized deep learning. Geopolitical events and historical price trends affect stock market volatility. They introduced Long Short-Term Memory with a Sequential Self-Attention Mechanism (LSTM-SSAM) to predict stock prices with low error. Their proposed model was tested using SBIN, HDFCBANK, and BANKBARODA stock datasets. Their study showed that LSTM-SSAM improves stock price prediction accuracy through extensive experimentation [25].

The literature review on stock market prediction effectively addresses several identified gaps. The inclusion of the Hang Seng Index as a focal point allows for the provision of market-specific insights pertaining to Hong Kong, thereby expanding the geographical scope of research within this field. Moreover, through the transparent elucidation of data preprocessing procedures, the assurance of data quality and reproducibility is achieved, thereby addressing the existing gap in the literature regarding the evaluation of data quality and preprocessing methodologies. Furthermore, the enhancement of predictive accuracy through domain knowledge is exemplified by the integration of domain-specific insights into the AO-GRU model, thereby addressing the limitation of limited domain knowledge integration. Moreover, the comparison between the AO-GRU model and ensemble techniques offers valuable insights into the efficacy of ensemble methods, thus filling the void in the limited investigation of these methods. Finally, a more thorough evaluation of model performance can be achieved by integrating supplementary evaluation metrics or examining the constraints of current ones, thereby addressing the deficiency in insufficient evaluation metrics. These contributions have made significant advancements in the field of stock market prediction, resulting in improved robustness, accuracy, and applicability of predictive models in financial markets.

### III. METHODS AND MATERIALS

#### A. Data Gathering and Preparation

The Hang Seng Index is a prominent stock market index in Hong Kong that tracks the performance of several notable companies listed on the Hong Kong Stock Exchange. The Hang Seng Index is composed of a diverse group of companies that lead their respective industries in many sectors of the Hong Kong economy. These sectors include, among others, the manufacturing, banking, real estate, and telecommunications sectors. Like other notable stock indices, the Hang Seng Index is weighted by market capitalization. Each company's weight in the index is based on its share market value; larger companies have a greater impact on the index's movements. Many factors, including trading volume and the Open, High, Low, and Close (OHLC) prices during a specific time period, should be included in a comprehensive study. The Hong Kong Stock Exchange provided hundreds of stocks from various industries that were used as the source of stock data for this study. Raw transaction data, comprising the opening price, closing price, highest price, lowest price, and trading volume, was gathered for every day from the start of 2015 to the end of June 2023. The gathered data was divided into two groups in order to maximize the performance of the models. As shown in Fig. 1, a partitioning approach was used in this experiment. In particular, twenty percent of the data was reserved for testing, while the remaining eighty percent was used for training. This division's main objective was to determine the most workable solution that balanced the need for a sizable amount of data for model training with the demands of a large, untested dataset for thorough testing and validation.

#### B. Gated Recurrent Unit

The GRU network was first introduced by Cho et al. [26]. The basic RNN concept [27] is to determine outputs by taking into account inputs and the hidden state, which is determined by squaring up past outputs or hidden states. With an update gate and a reset gate, the GRU offers sophisticated control over the data in a concealed state. In general, it can determine which data adding from the existing inputs is necessary (because it may be crucial to the future) and which data from the past may be eliminated from the hidden state (because it is unrelated to the present state). Compared to long short-term memory [28], which features a unit made up of three gates and a cell structure, the GRU has fewer parameters. A single GRU's construction is seen in Fig. 2,  $h$ , in which  $z$  is the update gate and  $r$  is the reset gate. Several of these units, designated  $h_j$  (along with associated  $r_j$  and  $z_j$ ) in a GRU network, are updated using the following equations:

$$r_j = \sigma \left( [V_r x]_j + [U_r h_{(t-1)}]_j \right) \quad (1)$$

$$z_j = \sigma \left( [V_z x]_j + [U_z h_{(t-1)}]_j \right) \quad (2)$$

$$h_j^{(t)} = z_j h_j^{(t-1)} + (1 - z_j) \tilde{h}_j^{(t)} \quad (3)$$

$$\tilde{h}_j^{(t)} = \phi \left( [V x]_j + [U (r \odot h_{(t-1)})]_j \right) \quad (4)$$

Let  $x^{(t)}$  be the vector that is input at a time  $t$ . The bias parameter is a part of  $V$ , which stands for parameter matrices.

To take the  $j$ -th element of a vector, use the boldface notation  $h$  and  $r$ , which represent the vectors with all of the values  $h_j$  and  $r_j$ . When first introduced,  $\sigma$  represents a hyperbolic tangent function and  $\sigma$  is a logistic function. In the beginning, for every  $j$ ,  $h_j^{(0)} = 0$ .

### C. Ant Lion Optimizer

The predatory behavior of ant lions in the wild, which primarily consists of ants, ant lions, and elite ant lions, served as the primary inspiration for the creation of ALO in 2015 [12]. The structure of the ALO algorithm is as follows: Using the roulette and random walk techniques, the ant colony and ant placements are first established at random. After an ant has completed its journey, their fitness is assessed using a fitness function. If the ant's location outperforms that of the ant lions around it, then it is deemed the best option available at this time. Furthermore, the position of the ant lion becomes the best option if it manages to capture an ant. In every cycle, the elite antlion stands in for the best possible outcome within the ant lion population. In contrast, the elite antlion is updated if the optimal antlion outperforms it; if not, it stays the same until the

end of the iteration, at which point it overcomes the elite antlion [12].

The random walk, which is used to show how ants travel, is stated as follows by Eq. (5):

$$X(t) = [0, cs(2r(t_1) - 1)], \dots, cs(2r(t_n) - 1) \quad (5)$$

where,  $r(t_n)$  is the random walk function of the  $n$ -th iteration and  $cs$  is the cumulative sum.

Moreover, Eq. (6) illustrates how Eq. (5) is further standardized.

$$X_i^{(t)} = \frac{(X_i^{(t)} - a_i) \times (d_i - c_i^{(t)})}{(d_i^{(t)} - a_i)} \quad (6)$$

The  $i$ -th individual's min and max values are denoted by  $a_i$  and  $d_i$ , respectively, while the  $t$ -th iteration of the  $i$ -th variable maximum value is represented by  $c_i^{(t)}$  and  $d_i^{(t)}$ , respectively.

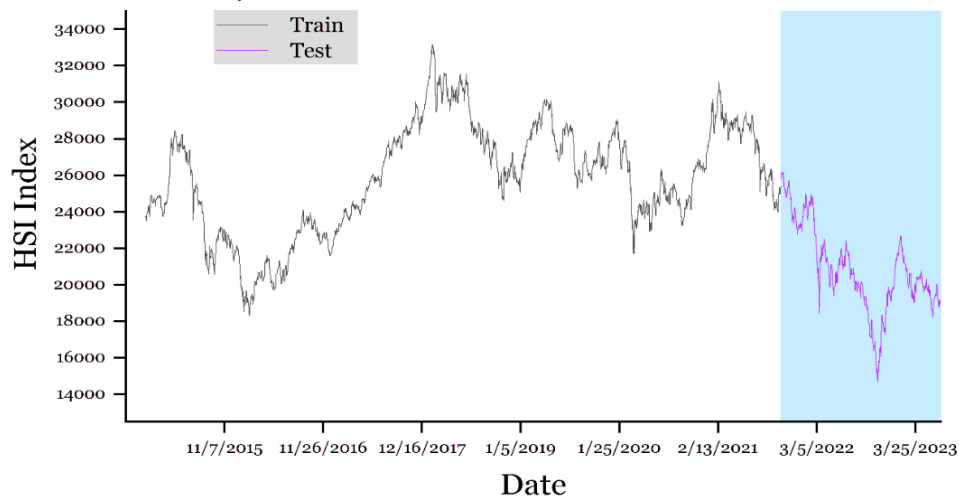


Fig. 1. Creating a separate train and test set of data.

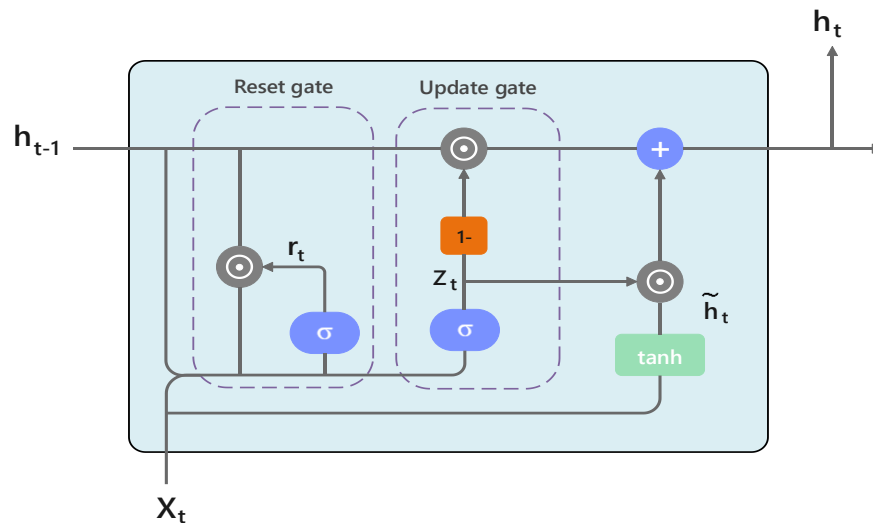


Fig. 2. The GRU model's structure.

Antlion traps have an impact on the random movement of ants, as demonstrated by Eq. (7) and Eq. (8):

$$c_i^{(t)} = \text{Antlion}_j^{(t)} + c^{(t)} \quad (7)$$

$$d_i^{(t)} = \text{Antlion}_j^{(t)} + d^{(t)} \quad (8)$$

In the  $t$ -th iteration, the individual's minimum and maximum values are denoted by  $c^{(t)}$  and  $d^{(t)}$ , respectively, whereas  $\text{Antlion}_j^{(t)}$  signifies the location of the  $j$ -th ant-lion.

Ants move randomly or in a roulette pattern around the antlion, as seen by Eq. (9):

$$\text{Ant}_i^{(t)} = \frac{R_A^{(t)} + R_E^{(t)}}{2} \quad (9)$$

The positions of the  $i$ -th ant at the  $t$ -th iteration are indicated by  $\text{Ant}_i^{(t)}$ , and  $R_A^{(t)}$  and  $R_E^{(t)}$  are the random walks around the elite or the roulette wheel on the second day of the, respectively.

As the number of iterations increases, antlion will get closer to the approximate optimal solution by reducing the boundaries in the manner described by Eq. (10) and Eq. (11):

$$c^i = \frac{c^{(t)}}{I} \quad (10)$$

$$d^i = \frac{d^{(t)}}{I} \quad (11)$$

where the lowest and maximum values of all variables at the  $t$ -th iteration are denoted by  $I$ , and  $c^{(t)}$  and  $d^{(t)}$  accordingly. Eq. (12) displays the location update formula, which the ant lion will use to feed on the ants after the iteration.

$$\text{Antlion}_j^{(t)} = \text{Ant}_i^{(t)}, \text{ if } f(\text{Ant}_i^{(t)}) > f(\text{Ant}_j^{(t)}) \quad (12)$$

Where the locations of the  $i$ -th and  $j$ -th ant-lions of the  $t$ -th iteration are represented by  $\text{Antlion}_j^{(t)}$  and  $\text{Ant}_i^{(t)}$ .

#### D. Grey Wolf Optimization

A novel swarm intelligence algorithm called the GWO makes use of the capabilities that Mirjalili et al. [29] discovered. Accurate stability is achieved between exploration and development, and it is expandable and adaptable. Following the wolf cooperation mechanism, the GWO imitates the actions of a population of grey wolves that are predators. Following natural law and rigid social structures, every wolf in the population has a certain role to fulfill [29]. Wolf populations in a GWO are arranged based on fitness levels.

The wolf with the lowest health,  $\omega$ , is regarded as the lowest-ranking person. A wolf within the wolf pack can be considered a feasible response, and the wolves corresponding to the finest solution, superior answer, and suboptimal answer of the present day may be designated as the  $\alpha$ ,  $\beta$ , and  $\gamma$  Wolf, respectively. The following sums up the grey wolf population's predatory behavior during the search:

$$D = |C \times X_p(t) - X(t)| \quad (13)$$

$$X(t + 1) = X_p(t) - A \times D \quad (14)$$

The distance  $D$  between the wolf and the target is given in Eq. (13). The coordinate transformation of a wolf is represented by Eq. (14), where  $X_p(t)$  represents the target's location in the  $t$ -generation,  $X(t)$  represents the position of a lone wolf inside the  $t$ -generation wolf pack,  $A$  and  $C$  are coefficients, and the calculation formula is as follows:

$$a = 2 - 2 * \frac{\text{iter}}{\text{Max}_{\text{iter}}} \quad (15)$$

$$A = 2a * r_1 \quad (16)$$

$$C = 2r_1 \quad (17)$$

where,  $r_1, r_2 \in [0,1]$ ,  $\text{iter}$  is the population of iterations, and  $\text{Max}_{\text{iter}}$  is the maximum number of iterations. The three different sorts of wolves choose which gray wolf will replace which when the wolf assaults its prey, or when it catches quarry. The paradigm for this decision is as follows:

$$D_i^j(t) = |C \times X_i^j(t) - X^j(t)| \quad (18)$$

$$X_i^j(t + 1) = X_i^j(t) - A \times D_i^j(t) \quad (19)$$

$$X(t + 1) = \frac{1}{3} \times \sum X_m(t + 1) \quad (20)$$

The difference between the  $t$ -generation and  $i(i = \alpha, \beta, \gamma)$  wolves are denoted by  $D_i^j(t)$ . In accordance with  $\alpha, \beta$ , and  $\gamma$  stride length and the wolf's motion direction, in that order, Eq. (18)-(20) determine the  $\omega$  wolf. The new period of grey wolves formed following a location update is represented by Eq. (20).

#### E. Aquila Optimizer

The AO is a recently introduced algorithm that aligns with the inherent hunting behavior of the Aquila species [14]. The hunting process has four distinct stages: an initial phase of extensive exploration achieved via soaring at high altitudes followed by a rapid vertical drop, a subsequent phase of focused exploration accomplished through gliding with precise contour flight, a subsequent phase of extensive exploitation achieved through a low-flying descending attack, and a final phase of focused exploitation accomplished through walking and capturing prey as seen in Fig. 3. The AO algorithm employs a range of characteristics to facilitate the transition from the exploration stage to the exploitation stage [14]. The initial two-thirds of iterations are dedicated to simulating the exploration stage, while the remaining one-third of iterations is allocated for imitating the exploitation stage [14] as the summary of AO optimizer performance is shown in Fig. 4.

The eagle starts the initial form of vertical descent when it identifies a potential area for prey and promptly determines the optimal hunting spot on the planet by ascending to significantly high elevations and identifying the region of investigation where the most efficient approach is determined using the subsequent formula:

$$\begin{cases} Z_1(t+1) = Z_{\text{best}}(t) \times \left(1 - \frac{t}{T}\right) + (Z_M(t) - Z_{\text{best}}(t) \times \text{rand}) \\ Z_M(t) = \frac{1}{N} \sum_{i=1}^N Z_i(t), \forall j = 1, 2, \dots, \text{Dim} \end{cases} \quad (21)$$

The generation  $t + 1$  solution is represented by  $Z(t + 1)$ , which is the result of the search strategy  $Z_1 \cdot Z_{\text{best}}(t)$ , where  $(t)$  is the ideal approach that indicates the position of the closest target prey. This loop has  $t$  iterations remaining.  $T$  is the highest possible number of iterations. The location means of the current solution at the  $t$ -th iteration is denoted by  $Z(t)$ . A random number between 0 and 1 is referred to as a Rand. The subsequent swift-gliding assault: The eagle soars to a height to discover the prey region in order to reduce the hunting area or search space in line with the following equation for the best reaction:

$$\begin{cases} Z_2(t+1) = Z_{\text{best}}(t) \times Z(D) + Z_R(t) + (y - z) \times \text{ra} \\ L(D) = s \times \frac{\mu \times \sigma}{|v|^{\beta}} \end{cases} \quad (22)$$

$$\begin{cases} Z_4(t+1) = Q_F \times Z_{\text{best}}(t) - (G_1 \times Z(t) \times \text{rand}) - G_2 \times L(D) + \text{rand} \times G_1 \\ Q_F(t) = \frac{2 \times \text{rand} - 1}{t^{(1-T)^2}} \\ G_1 = 2 \times \text{rand} - 1 \\ G_2 = 2 \times \left(1 - \frac{t}{T}\right) \end{cases} \quad (24)$$

The quality function, or  $Q_F$ , and the search technique is balanced. The Aquila's motions as it searches for its food are

The Aquila enters the low-flying, slow-falling assault mode at the chosen target point when the prey zone has been carefully located and it is ready to land and attack. This is the third pattern of low-altitude flying; by using this tactic, the bird could see how its prey would respond and gradually approach it, as in the formula below: with  $D$  representing the dimensional space,  $Z_R(t)$  representing the random solution between  $[1, N]$ , and  $L(D)$  representing the hunting flight distribution function.

$$\begin{aligned} Z_3(t+1) = & (Z_{\text{best}}(t) - Z_M(T)) \times \alpha \\ & - \text{rand} + ((U_b - L_b) \times \text{rand} + L_b) \\ & \times \delta \end{aligned} \quad (23)$$

$\alpha$  and  $\delta$  are the two moderating elements in this case.  $U_b$  is upper bound on the issue.  $L_b$  is the lower bound of the issue. Walking capture is the fourth method when the eagle uses the following equation to rapidly converge and attack the target from above.

seen on  $G_1$ . The hunting flying slope of Aquila is represented by  $G_2$ .  $Z(t)$  is the solution for this iteration.

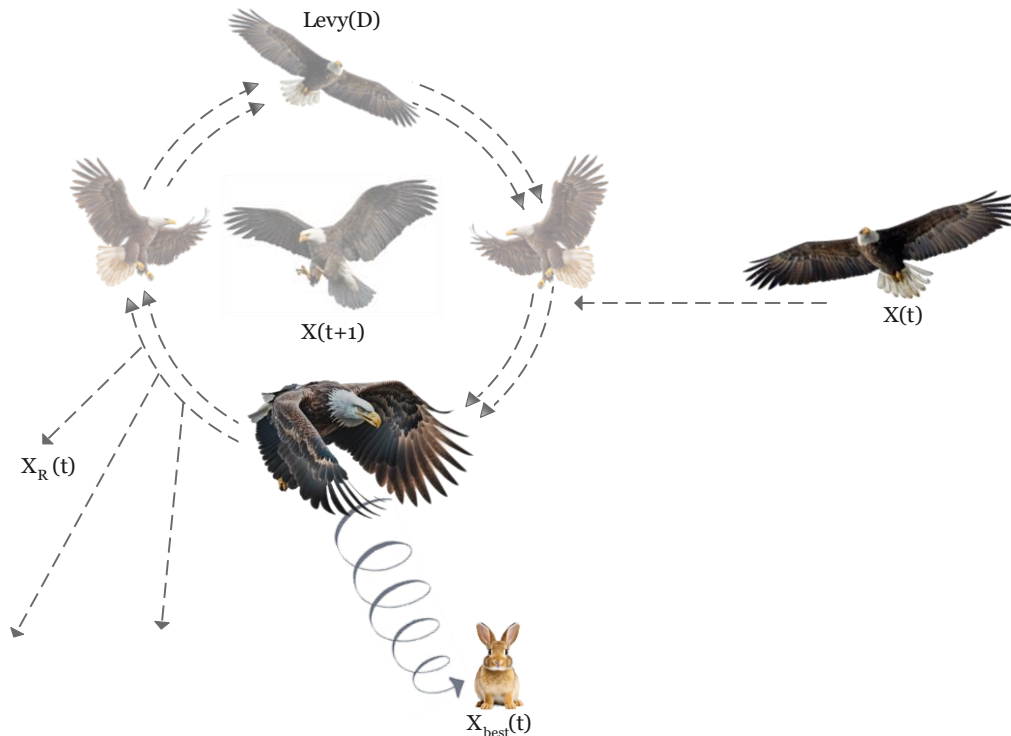


Fig. 3. An illustration of the Aquila hunting.

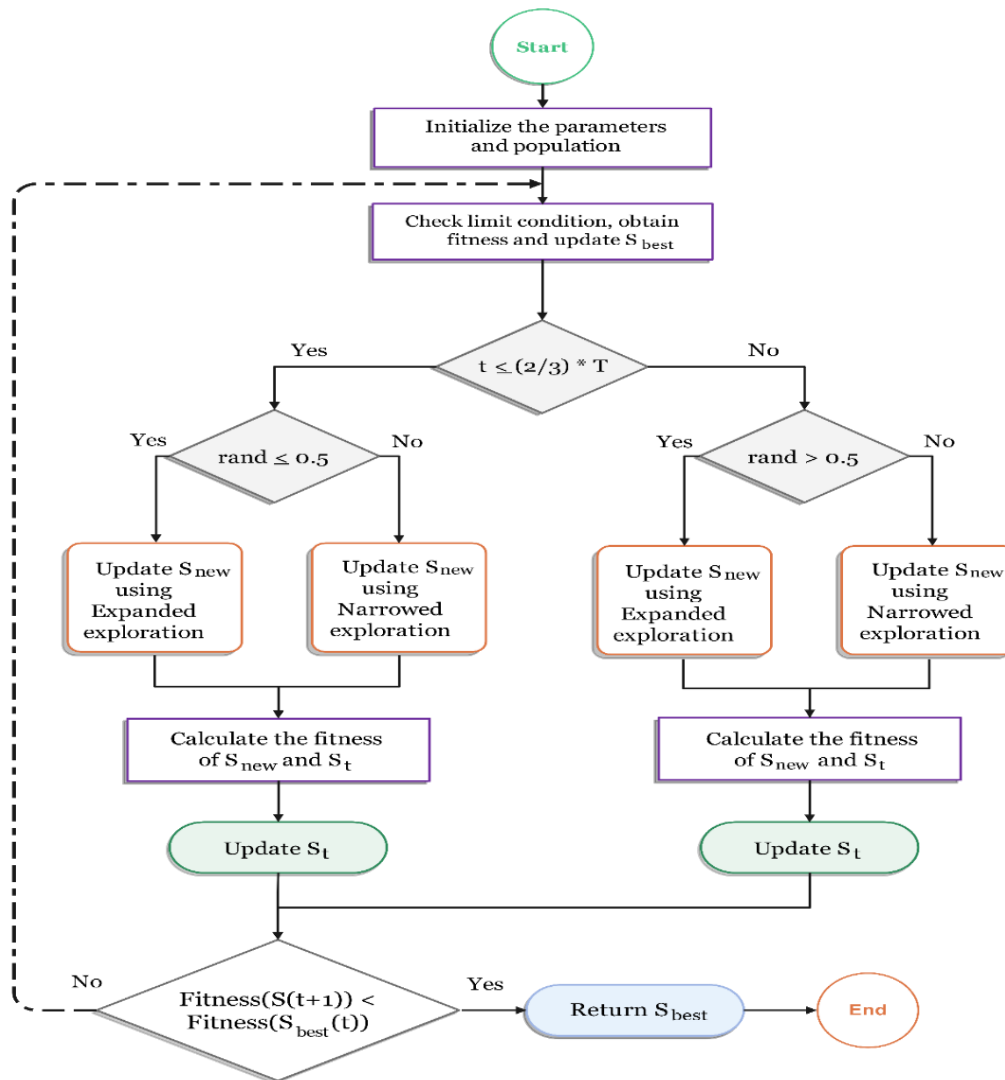


Fig. 4. The selective AO optimizer's graphical flowchart.

#### IV. RESULTS

##### A. Evaluation Metrics

The evaluation of the accuracy of the future forecast was conducted by employing a variety of performance measures. The carefully chosen metrics provide a thorough evaluation of the reliability and precision of the predictions. The assessment criteria employed in this article encompass:

- Coefficient of determination ( $R^2$ ):

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (25)$$

- Mean absolute error (MAE):

$$MAE = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n} \quad (26)$$

- Mean squared error (MSE):

$$MSE = \frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2, \quad (27)$$

- Mean absolute percentage error (MAPE):

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

##### B. Statistic Values

The dataset encompassed a significant temporal range, commencing on January 1, 2015, and concluding at the termination of June 2023. This section presents tabular representations of the project's outcomes subsequent to its effective implementation. The inclusion of OHLC price and volume statistics in Table I enhances the comprehensibility of the information. The utilization of statistical metrics such as count, mean, minimum (min), maximum (max), standard deviation (Std.), 25%, 50%, 75%, and variance enables a thorough and precise examination of the data.

TABLE I. A COLLECTION OF STATISTICAL VALUE SUMMARIES

	Open	High	Low	Volume	Close
count	2090	2090	2090	2090	2090
mean	24877.8	25026.72	24689.52	4013.656	24862.03
Std.	3492.279	3486.289	3484.234	1462.996	3486.437
min	14830.69	15113.15	14597.31	0	14687.02
25%	22194.44	22350.36	21998.24	3068.985	22151.32
50%	25002.49	25118.69	24755.93	3679.685	24973
75%	27716.1	27860.96	27525.43	4594.719	27693.23
max	33335.48	33484.08	32897.04	12025.52	33154.12
variance	12196013	12154214	12139887	2140358	12155245

C. Compare and Analysis

The successful completion of the project has yielded the results, which are displayed in this section as Table II and Fig. 5, 6. For this study, four algorithms were used to predict HSI

stock values: GRU, ALO-GRU, GWO-GRU, and AO-GRU. The dataset was massive, spanning from the beginning of 2015 to the end of June 2023. The models were evaluated using four important performance indicators:  $R^2$ , MAPE, MAE, and MSE.

TABLE II. A PREDICTION OF THE EVALUATION OUTCOMES OF THE MODELS

MODEL/Metrics	TRAIN SET				TEST SET			
	$R^2$	MAPE	MAE	MSE	$R^2$	MAPE	MAE	MSE
GRU	0.9894	1.09	280.91	92243.40	0.9875	1.01	202.16	69134.21
ALO-GRU	0.9923	0.83	213.74	67466.58	0.9906	0.97	198.26	51900.77
GWO-GRU	0.9941	0.77	197.79	51335.10	0.9921	0.78	161.26	43986.23
AO-GRU	0.9952	0.60	154.69	41515.24	0.9934	0.71	143.62	36530.58

V. DISCUSSIONS

The tabulated results for the GRU obtained during the testing of a specific model on the dataset of that firm are shown in Table II. The AO-GRU method has produced the greatest predictive performance out of all four models, with an MSE of 36530.58, the lowest of all four. This is because the optimum MSE value is near zero. The results of the  $R^2$  calculations for several models evaluated on a certain dataset are listed in Table II. The model fit is better when the value of  $R^2$  is nearer 1. The obtained results demonstrate that the AO-GRU has yielded the most encouraging outcomes among the four computational models that were examined. Because its  $R^2$  score of 0.9934 is the closest to 1, which suggests that the accuracy rate of the model is high. This table offers a thorough summary of the MAPE that a given model obtained during testing on the dataset. A lower number in the MAPE indicates greater performance. With a MAPE of 0.71, Table II analysis reveals that AO-GRU performed the best out of the three algorithmic

models examined. Because it produces the most accurate predictions, it can be concluded that the AO-GRU algorithm is the best computational model to apply when working with comparable datasets. Regarding the MAE report, it should be noted that the more accurate the prediction, the closer the reported value is to zero. This is supported by the reported table, which indicates that the reported number for the AO-GRU model's test phase is 143.62, indicating that the model in question is more accurate at forecasting and is the one used in this article. The findings may be used in business and other domains where data analysis is essential for making well-informed choices and forecasts.

The GRU model's combination with the Aquila optimizer as seen in Fig. 7 and Fig. 8, it is clear from the analysis of the HSI index curves and their comparative assessment that the AO-GRU model performs better and is more effective than the other models investigated in this study.

Train

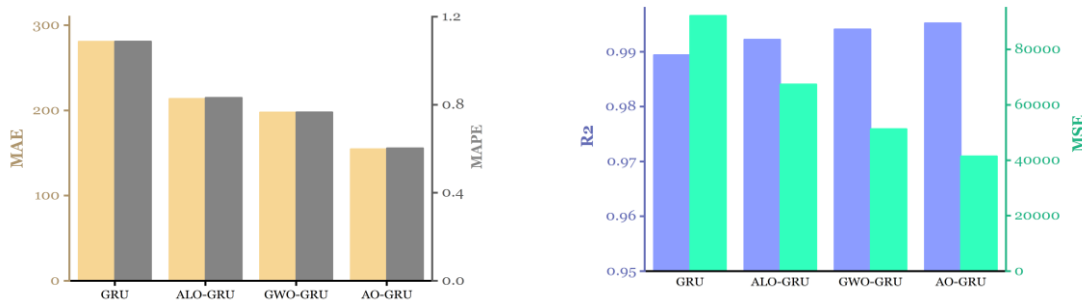


Fig. 5. The methods' training outcomes contain a variety of measures, including  $R^2$ , MAPE, MAE, and MSE.



### Test

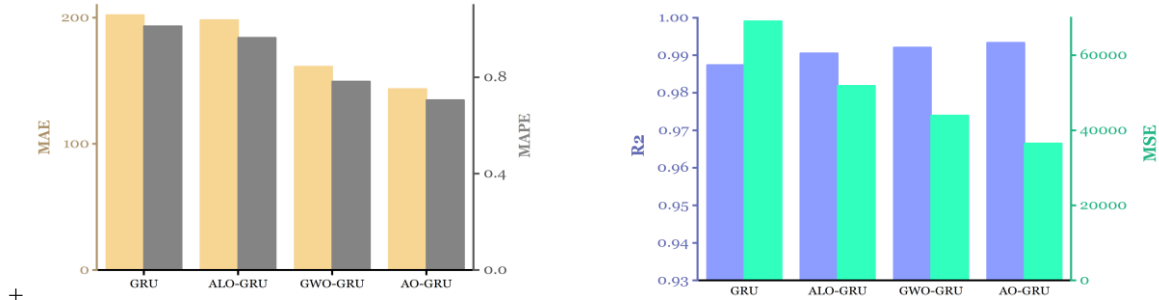


Fig. 6. The methods' testing outcomes contain a variety of measures, including  $R^2$ , MAPE, MAE, and MSE.

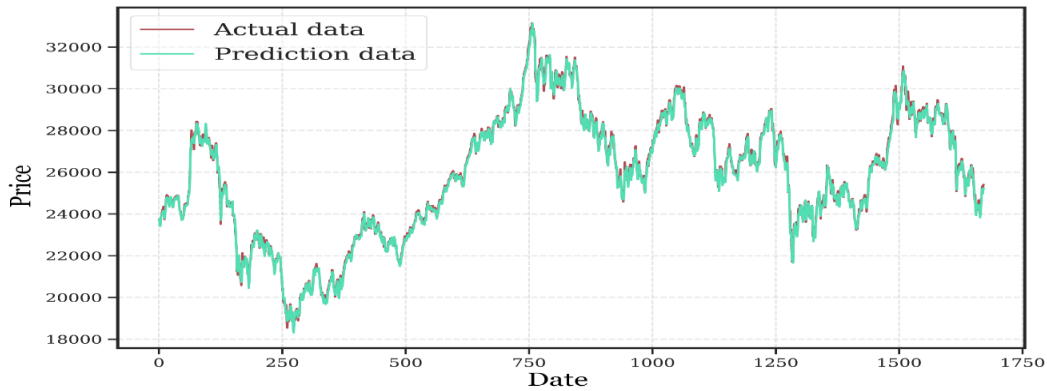


Fig. 7. The prediction curve produced during the training phase by applying the AO-GRU technique.

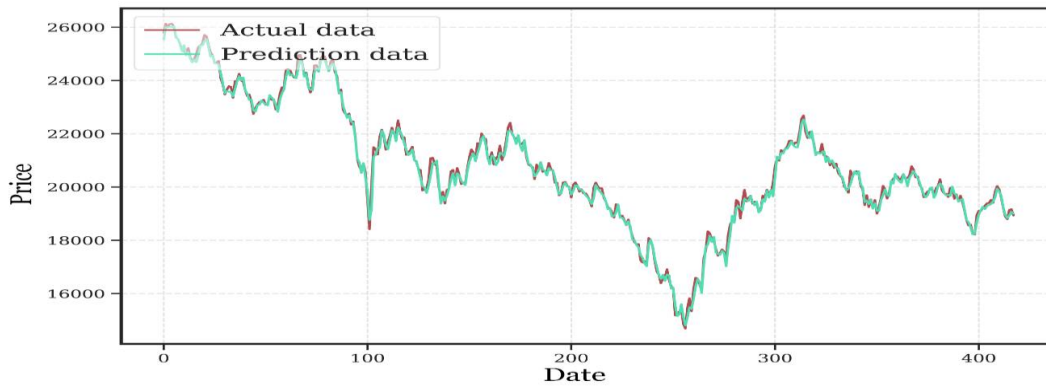


Fig. 8. The prediction curve produced during the testing phase by applying the AO-GRU technique.

TABLE III. A COMPARATIVE ANALYSIS OF THE MODEL IS PROVIDED IN RELATION TO PREVIOUS STUDIES

Authors	Method	$R^2$
Zhu et al. [30]	LSTM	0.6896
	EMD-LSTM	0.8703
	CEEMDAN-LSTM	0.9031
	SC-LSTM	0.6871
	EMD-SC-LSTM	0.9111
	CEEMDAN-SC-LSTM	0.9206
Abdul et al. [31]	Linear regression	0.73
	SVM	0.93
	MLS-LSTM	0.95
Current study		0.9934

Based on the findings presented in Table III, the AO-GRU model demonstrates superior performance when compared to other models evaluated in the domain of stock price prediction. In this study, the  $R^2$  value of 0.9934 was found to be highly impressive, surpassing the performance of other evaluated models such as LSTM, SVM, and MLS-LSTM. This indicates a stronger correlation between the predicted and actual stock prices, suggesting that the model is more effective in capturing the underlying data patterns. By integrating adaptive optimization techniques, the AO-GRU model gains the capability to dynamically modify its parameters during training. This allows it to adapt to the subtleties of the data and improve its capacity to handle variations and complexities in the stock market. The model effectively utilizes the GRU architecture to capture long-range dependencies in sequential data, while also addressing challenges such as the vanishing gradient problem. The gating mechanisms of GRU control the flow of information within the network, enabling more effective learning, especially in situations where there is a scarcity of training data. Significantly, the efficiency of the GRU architecture, which necessitates fewer parameters and computations in comparison to more intricate models such as LSTM, results in expedited training durations and reduced computational expenses. Consequently, the AO-GRU model becomes more viable for real-time or extensive applications. Moreover, the model's architecture, which is relatively uncomplicated, improves its interpretability, offering stakeholders valuable insights into the determinants of stock price fluctuations and the underlying reasoning behind the model's predictions, as opposed to certain intricate black-box models.

## VI. CONCLUSION

This study aimed to develop machine learning models with improved stock price prediction accuracy. By making the appropriate sort of investment at the appropriate moment, traders and investors would be able to take advantage of these strategies and maximize their gains. This project included the effective implementation of four algorithms: GRU, ALO-GRU, GWO-GRU, and AO-GRU. A comprehensive comparative analysis of the algorithms' performances during stock price prediction was conducted after GRU algorithms were utilized to create accurate predictive models for use in the stock price prediction of HSI. The collected stock values 1st of January 2015 to the end of June 2023. The four assessment metrics MSE,  $R^2$ , MAPE, and MAE as well as the data gathered during the model testing are displayed in tabular and graphical form in the research study's results section.

- Following a comprehensive review and analysis of the data, the AO-GRU approach is shown to be the most error-free among the available strategies for time series prediction, with the lowest MSE (36530.58), MAPE (0.71), and MAE (143.62) errors and the highest value of  $R^2$  (0.9934).

The study's findings present compelling evidence that suggests several promising avenues for future research in the field of stock price prediction. In order to enhance predictive accuracy and robustness, it is imperative to investigate alternative optimization techniques beyond Aquila, the

optimizer employed in the AO-GRU model. The incorporation of supplementary datasets not limited to the Hang Seng Index has the potential to yield more profound insights and improve the generalizability of the model. This could involve the inclusion of macroeconomic indicators, industry-specific data, or sentiment analysis derived from diverse sources. Furthermore, the implementation of strategies such as attention mechanisms or feature importance analysis has the potential to improve the interpretability of models, thereby promoting increased trust and usability among various stakeholders. It is imperative to conduct thorough testing in various market conditions and economic cycles in order to evaluate the robustness and flexibility of models. Additionally, investigating the applicability of these models to different asset classes, such as commodities or currencies, may reveal novel opportunities for investment strategies. By directing attention towards these specific areas, forthcoming research endeavors possess the capacity to propel the domain of stock price prediction forward and effectively address the changing demands of investors and financial professionals.

## REFERENCES

- [1] M. Bansal, A. Goyal, and A. Choudhary, "Stock Market Prediction with High Accuracy using Machine Learning Techniques," *Procedia Comput Sci*, vol. 215, pp. 247–265, 2022, doi: <https://doi.org/10.1016/j.procs.2022.12.028>.
- [2] V. U. Kumar, A. Krishna, P. Neelakanteswara, and C. Z. Basha, "Advanced prediction of performance of a student in an university using machine learning techniques," in 2020 international conference on electronics and sustainable communication systems (ICESC), IEEE, 2020, pp. 121–126.
- [3] M. Mohri, A. Rostamizadeh, and A. Talwalkar, *Foundations of machine learning*. MIT press, 2018.
- [4] L.-P. Chen, "Multiclassification to gene expression data with some complex features," *Biostat Biom Open Access J*, vol. 9, no. 1, pp. 1–2, 2018.
- [5] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," *arXiv preprint arXiv:1412.3555*, 2014.
- [6] Y. Gao, R. Wang, and E. Zhou, "Stock Prediction Based on Optimized LSTM and GRU Models," *Scientific Programming*, Vol 2021 (2021), 2021, doi: [10.1155/2021/4055281](https://doi.org/10.1155/2021/4055281).
- [7] C. Qi, J. Ren, and J. Su, "GRU Neural Network Based on CEEMDAN-Wavelet for Stock Price Prediction," *Applied Sciences*, Vol 13, Iss 12, p 7104 (2023), 2023, doi: [10.3390/app13127104](https://doi.org/10.3390/app13127104).
- [8] S. Mirjalili and A. Lewis, "The whale optimization algorithm," *Advances in engineering software*, vol. 95, pp. 51–67, 2016.
- [9] D. Simon, "Biogeography-based optimization," *IEEE transactions on evolutionary computation*, vol. 12, no. 6, pp. 702–713, 2008.
- [10] 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](https://doi.org/10.1016/j.advengsoft.2022.103358).
- [11] S. Mirjalili, "Moth-flame optimization algorithm: A novel nature-inspired heuristic paradigm," *Knowl Based Syst*, vol. 89, pp. 228–249, 2015.
- [12] S. Mirjalili, "The ant lion optimizer," *Advances in engineering software*, vol. 83, pp. 80–98, 2015.
- [13] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Advances in engineering software*, vol. 69, pp. 46–61, 2014.
- [14] 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.
- [15] B. Sasmal, A. G. Hussien, A. Das, and K. G. Dhal, "A Comprehensive Survey on Aquila Optimizer," *Archives of Computational Methods in*

- Engineering, vol. 30, no. 7, pp. 4449–4476, 2023, doi: 10.1007/s11831-023-09945-6.
- [16] 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.
- [17] 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.
- [18] 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.
- [19] 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/>.
- [20] 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.universitاسbumigora.ac.id/index.php/matrik/article/view/2287>.
- [21] 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>.
- [22] 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.
- [23] I. Su, P. L. Lin, Y. Chung, and C. Lee, “Forecasting of Taiwan’s weighted stock Price index based on machine learning,” *Expert Systems*; volume 40, issue 9; ISSN 0266-4720 1468-0394, 2023, doi: 10.1111/exsy.13408.
- [24] R. Chipwanya, “Stock Market Directional Bias Prediction Using ML Algorithms,” 2023, doi: 10.48550/arxiv.2310.16855.
- [25] 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.
- [26] K. Cho, B. Van Merriënboer, D. Bahdanau, and Y. Bengio, “On the properties of neural machine translation: Encoder-decoder approaches,” *arXiv preprint arXiv:1409.1259*, 2014.
- [27] K. Cho et al., “Learning phrase representations using RNN encoder-decoder for statistical machine translation,” *arXiv preprint arXiv:1406.1078*, 2014.
- [28] R. Ying, Y. Shou, and C. Liu, “Prediction Model of Dow Jones Index Based on LSTM-Adaboost,” in *2021 International Conference on Communications, Information System and Computer Engineering (CISCE)*, 2021, pp. 808–812. doi: 10.1109/CISCE52179.2021.9445928.
- [29] S. Mirjalili, S. M. Mirjalili, and A. Lewis, “Grey Wolf Optimizer *Adv Eng Softw* 69: 46–61.” ed, 2014.
- [30] 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>.
- [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>.