Hybrid Model of Artificial Neural Networks and Flower Pollination Algorithm for Stock Price Prediction

Lia Farhatuaini
Informatics
IAIN Syekh Nurjati
Cirebon, Indonesia
farhatuaini@syekhnurjati.ac.id

Heru Purnomo Kurniawan
Informatics
IAIN Syekh Nurjati
Cirebon, Indonesia
herukurniawan@syekhnurjati.ac.id

Isnawati Muslihah
Visual Communication Design
ISI Surakarta
Surakarta, Indonesia
isnawati.muslihah@isi-ska.ac.id

Abstract— Predicting the future behavior of the stock market is a difficult task due to its complex and ever-changing nature. This study focuses on predicting BBRI stock prices using an Artificial Neural Network (ANN) improved with the Flower Pollination Algorithm (FPA). We found that the model works well with a 9-100-1 setup, achieving accurate predictions with a Root Mean Square Error (RMSE) of 0.127579154. While FPA effectively reduces errors in the initial 10 iterations, it faces challenges in further improvement, especially in responding to sudden changes in stock prices. Despite performing well overall, the model tends to have a wider margin during unexpected market shifts, indicating a need for additional fine-tuning. This research provides valuable insights into stock price prediction, highlighting the importance of refining models to handle unexpected market changes.

Keywords— Stock price forecasting, Artificial Neural Network, Flower Pollination Algorithm, Model optimization

I. INTRODUCTION

Predicting the future behavior of the stock market is a difficult task due to its complex and ever-changing nature [1]. However, a number of machine learning techniques, including artificial neural networks (ANNs), have shown useful in stock market prediction [2], [3], [4]. ANNs are a type of machine learning algorithm capable of learning intricate relationships between input and output data. They consist of interconnected neurons arranged in layers. Each neuron in a layer receives input from the neurons in the preceding layer and produces an output signal, which is then transmitted to the next layer. During the training process, the weights of connections between neurons are adjusted to enable the ANN to learn how to map input data to the desired output. Typically, supervised learning algorithms like backpropagation are used for this purpose [5].

In recent years, the utilization of ANNs in stock market prediction has gained significant interest. Several research studies have explored the application of ANNs in stock market prediction, demonstrating their potential for improving forecasting accuracy. The ANN model predicts stock prices in stock exchange markets, trained and validated using datasets from the Nairobi Securities Exchange and the New York Stock Exchange. The results indicate a mean absolute percentage error (MAPE) of 0.71% to 2.77%, demonstrating its capability for predicting stock prices in typical markets [6]. The ANN has also been compared with SVM for stock forecasting, and the study concludes that both methods are effective, with ANN performing slightly better. Additionally, the study suggests that incorporating financial news sentiments with stock features can significantly enhance LSTM and GRU models for stock price forecasting [7].

A stock price prediction is also achieved through the utilization of a deep neural network (DNN) model. This model incorporates both technical and fundamental features and undergoes training and validation with historical stock market data. The outcomes of the study indicate that the DNN model exhibits superior performance compared to alternative models in terms of accuracy and robustness [8]. Two models are used for predicting stock market movement: a deep long short-term memory neural network (LSTM) with an embedded layer and an LSTM with an automated encoder. With accuracies of 57.2% and 56.9% for the Shanghai A-shares composite index and 52.4% and 52.5% for individual stocks, respectively, experimental findings demonstrate the improved performance of the deep LSTM with embedded layer [9].

In this paper, we propose a method for optimizing ANNs for stock market prediction using Flower pollination algorithm (ANN-FPA). Flower Pollination Algorithm (FPA) is a metaheuristic algorithm proposed by Yang [10] that has been shown to be effective in optimizing a variety of problems. Several studies have explored the application of the Flower Pollination Algorithm (FPA) in various domains to optimize different problems, including optimization [11], machine learning [12], financial prediction [13], [14], and scheduling problems [15]. The utilization of FPA has demonstrated improvements in solution quality and performance compared to traditional optimization techniques.

A Combination of ANN with FPA proposed in [16] for predicting burr height and burr thickness. Comparing the developed ANN-FPA prediction model to the single ANN model, it is more accurate. Moreover, in [17] ANN models are also used for predicting residual stress, combined with two distinct metaheuristic optimization algorithms, particle swarm optimization (PSO) and FPA. The ANN-FPA had the best prediction accuracy followed by ANN-PSO. To assess the effectiveness of our proposed method, we conduct experiments using a dataset comprising historical stock prices of BBRI, a prominent Indonesian bank [18].

The selection of BBRI stock for this research is driven by its prominent role in the Indonesian banking industry, marked by a solid market presence and an extensive customer base, enhancing its appeal as a stable investment option. Furthermore, BBRI stock's consistent and robust financial performance, including profitability and reliable dividend payouts, underscores its potential to deliver favorable returns to investors. Beyond its financial merits, the inclusion of BBRI stock in the study provides valuable insights into the broader dynamics of the Indonesian stock market, offering a reflection of trends and sentiments within the banking and financial sectors.

The objectives of this research are twofold: Firstly, to develop an accurate and robust stock price prediction model for Bank BBRI by harnessing the capabilities of ANN and the optimization power of the FPA. Secondly, to compare the performance of the proposed model with traditional ANN models and highlight its superiority in terms of prediction accuracy.

To accomplish these objectives, we will leverage historical stock price data of Bank BBRI and preprocess the dataset to remove noise and outliers. The ANN model will be trained on a subset of the data, while the FPA will optimize the model's parameters. The performance evaluation of the proposed model will employ Root Mean Square Error (RMSE).

The findings of this research have implications for investors and financial institutions, providing more accurate stock price predictions for Bank BBRI. Moreover, the integration of the Flower Pollination Algorithm into the optimization of ANN models for stock price prediction contributes to the growing field of computational intelligence in finance.

II. PROPOSED METHOD

A. Artificial Neural Network

Artificial Neural Networks (ANNs) have emerged as powerful tools for solving complex problems and modeling intricate relationships in various domains. ANNs are mathematical models inspired by the structure and functionality of biological neural networks [19]. They consist of interconnected artificial neurons, organized in layers, and capable of learning from data. The fundamental architecture of an ANN involves several key components working in tandem. Neurons, the basic computational units, receive inputs, apply mathematical operations, and generate outputs. These neurons are organized into layers: the input layer for initial data reception, hidden layers for complex computations, and the output layer for final predictions or classifications. The connections between neurons are governed by weights, which signify the importance of specific inputs. The form of the implementation of artificial neural network is shown in Figure 1.

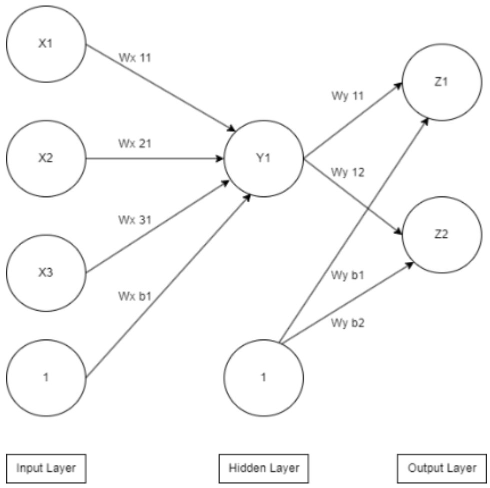

Figure 1. Artificial Neural Network

The fundamental building block of an artificial neuron is the activation function, which introduces non-linearity into the model. The activation function computes the output of the neuron based on the weighted sum of inputs and a bias term. One commonly used activation function is the sigmoid function, denoted as $\sigma(z)$, which is defined in Equation 1.

$$\sigma(z) = \frac{1}{1+e^{-z}} \quad (1)$$

where z represents the weighted sum of inputs. The output of an artificial neuron is then used as input for subsequent neurons in the network. This process of passing information through the network is known as forward propagation. The final layer of neurons in the network provides the output, or prediction.

The weights and biases of the neurons in an ANN are adjusted during a training process, which involves minimizing a defined loss or error function. Backpropagation, a widely used algorithm for training ANNs, iteratively updates the weights and biases by calculating gradients based on the error between predicted and actual outputs.

B. Flower Pollination Algorithm

The Flower Pollination Algorithm (FPA) is an optimization algorithm inspired by the pollination behavior of flowers. It has gained recognition for its efficacy in addressing diverse optimization problems. By imitating the process of flower pollination, FPA efficiently explores the search space to find optimal solutions [10]. In FPA, the population is represented by a collection of flowers, with each flower corresponding to a potential solution for the given optimization problem. The quality of each flower is evaluated based on an objective function. The algorithm initiates with an initial population of flowers and proceeds iteratively to explore and refine the solutions.

The search process in FPA encompasses three fundamental operators: global pollination, local pollination, and flower update. Global pollination facilitates the exchange of information among flowers with higher quality (greater nectar amount), enabling exploration of the solution space. Local pollination encourages information exchange between neighboring flowers, promoting the exploitation of promising regions within the solution space. The flower update operator introduces diversity and exploration by adjusting the positions of flowers according to the best solutions discovered thus far.

Mathematically, FPA updates the position of each flower using Equation 2.

$$x_i^{t+1} = x_i^t + \beta \cdot (x_{best} - x_i^t) + \alpha \cdot (x_r - x_i^t) \qquad (2)$$

where x_i^{t+1} is the new position of i-th flower of iteration $t+1$, x_i^t is its current position at iteration t, x_{best} represents the position of the best flower in the population, x_r is the position of a randomly selected flower, and α and β respectively are control parameters that govern the impact of global and local pollination.

C. Dataset

The dataset used in this research focuses on the stock prices of Bank BRI (BBRI) consists of 1237 data points which is presented in Figure 2, covering a time period of the last five years from October 7th, 2018, until July 7th, 2023.

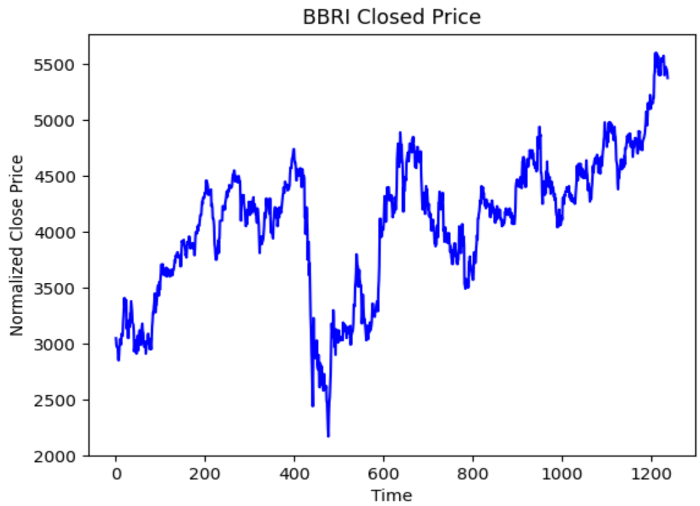

Figure 2. BBRI Closed Price

We utilize the close price for each day in this study. To obtain the dataset's statistical information, we perform exploratory data analysis. The outcome is displayed in Table 1.

Table 1. Exploratory Data Analysis of BBRI Closed Price

Metric	**Score**
Count	1237
Mean	4080.15
Median	4200
STD	627.38
Max	5600
Min	2170

From the exploratory data analysis, The dataset under consideration comprises 1237 observations of Close Price of BBRI. The mean or average close price is approximately

Rp. 4080.15, with a standard deviation of approximately Rp. 627.38, indicating a moderate degree of variability in the close prices around the mean. The range of close prices spans from a minimum of Rp. 2170 to a maximum of Rp. 5600, and the median is Rp. 4200.

D. Pre-Processing

The preprocessing of data is a crucial step in preparing the dataset for analysis and model training. The main objective of data normalization is to bring the data into a standardized range or scale, ensuring that all features contribute equally to the analysis and model training processes. In this paper, the MinMax normalization method is employed to achieve this goal. MinMax normalization transforms the data to a specific range, typically between 0 and 1, while preserving the relative relationships between the values. The formula for MinMax normalization is given by dividing the difference between each data point and the minimum value in the dataset by the range of the dataset, as shown in the Equation 3.

$$X_{\text{normalized}} = \frac{X - X_{min}}{X_{max} - X_{min}} \qquad (3)$$

where $X_{\text{normalized}}$ represents the normalized value of a data point, X is the original value, $X_{\min}$ is the minimum value in the dataset, and $X_{\max}$ is the maximum value in the dataset. By applying this normalization technique, all features in the dataset will be scaled to a common range, eliminating any biases caused by variations in magnitude.

The sliding window technique is employed to create a time-series representation of the data. This technique involves dividing the sequential data into smaller, overlapping windows. Each window consists of a fixed number of previous time steps (input features) and a subsequent time step (target output). By using the sliding window approach, the model can capture temporal dependencies and patterns present in the time series data, enabling more accurate analysis and forecasting.

To evaluate the performance of the prediction model, the preprocessed dataset is split into training and testing sets. The data is divided into an 80:20 ratio, with 80% allocated for training the model and 20% for evaluating its performance. This division ensures that the model is trained on a significant portion of the data while having a separate set of unseen data to assess its ability to generalize and make accurate predictions. The training set is utilized to train the model parameters, while the testing set is used to assess the model's performance on unseen data.

E. ANN-FPA

Following the data splitting, our model is predicted using an ANN that has been tuned via FPA (ANN-FPA). FPA is used to find the best weight in an ANN. The proposed method of ANN-FPA in this paper is written in pseudocode in Algorithm 1 and can be visualized in a flowchart as shown in Figure 3.

Our research aims to obtain the best ANN weights (g_{best}) to predict the close stock price of BBRI. In Algorithm 1, we determine the architecture of the ANN in order to obtain the number of weights (d-dimension). The first population is created by producing d random values from an n set of x, which are then used as ANN weights. The next step is to evaluate each solution (x_i) through the forward propagation process of an ANN, resulting in prediction values $\hat{y}$. From the obtained prediction values, the error is then calculated using RMSE (Root Mean Square Error). RMSE is written in Equation 4, where n is equal to the length of the data and y_i is the target value.

$$\text{RMSE}(x) = \sqrt{\sum_{i=1}^{n} (y_i - \hat{y_i})^2} \qquad (4)$$

Algorithm 1: ANN-FPA Algorithm

Determining the architecture of ANN d
Objective min ANN$(x), x = (x_1, x_2, ..., x_d)$
Initialize a population of n flowers generated with random solutions
Calculate error of solution
Find the g_{best} solution
while *(iter $\leq$ epoch)* **do**
  for $i = 1 : n$ *(all n flowers in the population)* **do**
    if *(rand $< \alpha$)* **then**
      Draw a (d-dimensional) step vector L which obeys a Lévy distribution
      Global pollination $x_i^{iter+1} = x_i^{iter} + L(g_{best} - x_i^{iter})$
    end
    else
      Draw β from a uniform distribution in $0, 1$
      Randomly choose j and k among all the solutions
      Do local pollination $x_i^{iter+1} = x_i^{iter} + \beta(x_j^{iter} - x_i^{iter})$
    end
    Evaluate new solutions
    If new solutions are better, update the population
  end
  Find the current best solution g_{best}
end

Algorithm 1. ANN-FPA Algorithm 1

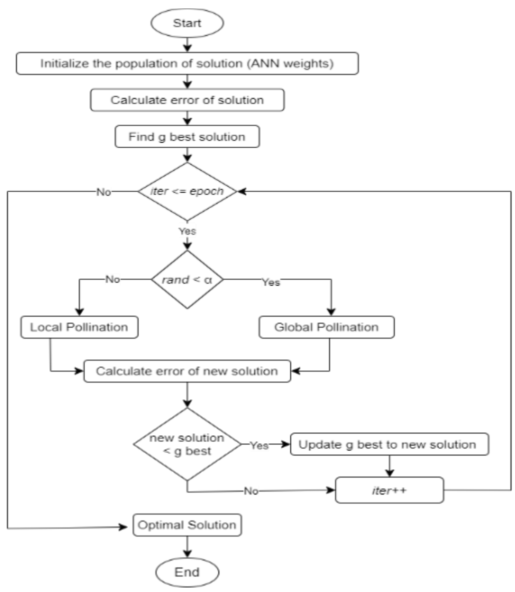

Figure 3. Flowchart of ANN-FPA

For each iter in Algorithm 1, we proposed n solutions as population x. From the initial population, the best solution will be stated as g_{best}. In each iteration *iter* where $iter = (1, 2, \ldots, epoch)$, each solution x_i will be processed based on FPA algorithms. A random value rand will be generated for every solution x_i to determine global or local pollination. If the random value rand is less than α, the global pollination process will be applied. Conversely, if it is greater than or equal to α, the local pollination process will be applied. The pollination process will generate a new solution, and its error value will be recalculated. If the new solution has a smaller error value than g_{best}, then it will become the new g_{best}. This process will continue until the iteration *iter* reaches the maximum *epoch*.

III. RESULT AND DISCUSSION

In this experiment, we conducted forecasting for BBRI stock close prices using an Artificial Neural Network (ANN) optimized with the Flower Pollination Algorithm (FPA). We performed 13 experiments with consistent FPA parameters, including an alpha value of 0.8, beta value of 0.5, and a fixed number of epochs set at 100. From the results of these experiments depicted in Table 2, the optimal outcome was obtained through the evaluation of the ANN-FPA model with an architecture of 9-100-1, where 9 represents the number of nodes in the input layer, 100 denotes the quantity of hidden layer nodes, and 1 signifies the output layer representing the predicted stock price. The Root Mean Square Error (RMSE) for this configuration was recorded as 0.127579154, indicating the accuracy of the forecasting model.

Table 2. RMSE of Testing Data

Input	Hidden Layers	Number of Flowers	Epoch Size	$\boldsymbol{\alpha}$	$\boldsymbol{\beta}$	RMSE
2	100	50	100	0,8	0,5	0,208387
3	100	50	100	0,8	0,5	0,217355
4	100	50	100	0,8	0,5	0,105633
5	100	50	100	0,8	0,5	0,203076
6	100	50	100	0,8	0,5	0,170476
7	100	50	100	0,8	0,5	0,177044
8	100	50	100	0,8	0,5	0,161923
9	100	50	100	0,8	0,5	0,127579
10	100	50	100	0,8	0,5	0,159058
11	100	50	100	0,8	0,5	0,187137
12	100	50	100	0,8	0,5	0,171265
13	100	50	100	0,8	0,5	0,170713
14	100	50	100	0,8	0,5	0,243697

The convergence process of the g_{best} in the 9-100-1 architecture depicted in Figure 4 significantly reduced the error rate in the initial 10 iterations. However, after that point, the Flower Pollination Algorithm (FPA) was unable to further decrease the error rate. The predictions with the corresponding error magnitudes are visualized in Figure 5.

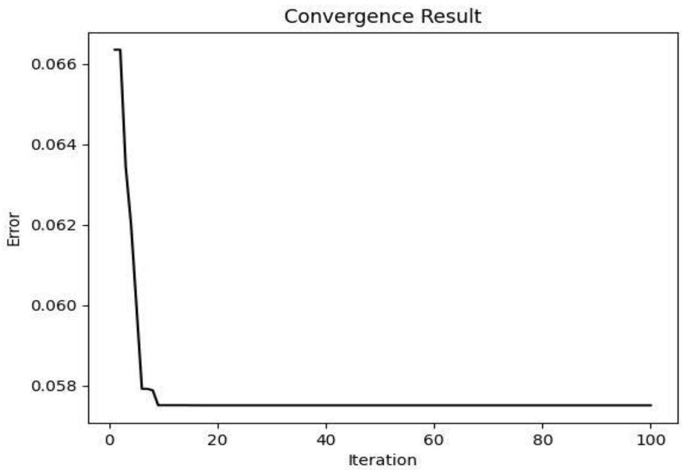

Figure 4. RMSE of gbest in the 9-100-1 ANN Architecture

It seems that the prediction trend may typically follow the pattern of the real data trend when comparing the trend of ANN-FPA forecasts to the trend of the actual data. On the other hand, the predicted trend shows a wider margin than the actual data trend in situations where the trend undergoes dramatic jumps and decreases. Larger mistakes in the prediction model are a result of this difference, especially when there are sudden and sharp swings in the data.

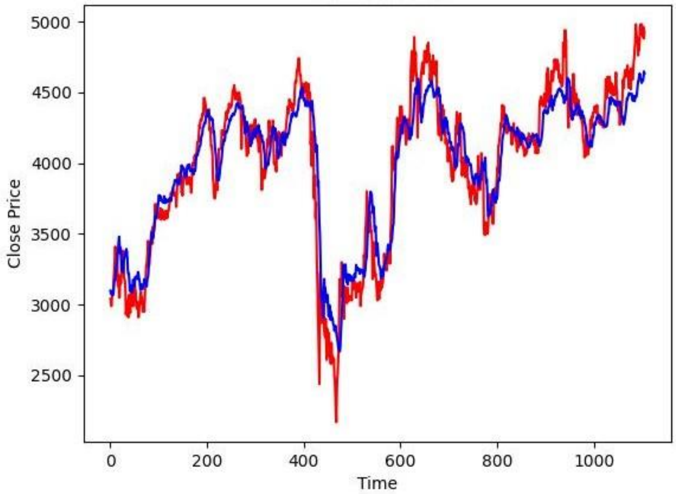

Figure 5. The prediction result of BBRI Stock Price with ANN-FPA

IV. CONCLUSION

The experimentation involving the forecasting of BBRI stock close prices using the Artificial Neural Network (ANN) optimized with the Flower Pollination Algorithm (FPA) provided valuable insights into the model's performance. The identification of an optimal configuration, specifically the 9-100-1 architecture, showcased the model's capability to achieve accurate predictions with a Root Mean Square Error (RMSE) of 0.127579154. The convergence process within the initial 10 iterations highlighted the effectiveness of the FPA in substantially reducing the error rate.

However, during the experiment, there was a noticeable obstacle or limitation in continuing to decrease the error rate of the forecasting model after a certain point. The implication is that the model encountered difficulties when attempting to adapt to sudden and significant changes in the trend of stock prices.

Despite the model's overall ability to capture the general trend, the wider margin observed during periods of sudden market changes indicates a need for further refinement. Future endeavors may involve exploring additional optimization strategies or fine-tuning parameters to enhance the model's adaptability to volatile market conditions. Overall, this study contributes valuable findings to the field of stock price prediction, emphasizing the importance of model robustness in handling unexpected fluctuations for more reliable forecasting.

REFERENCES

[1] E. F. Fama, "Efficient Capital Markets: A Review of Theory and Empirical Work," *J Finance*, vol. 25, no. 2, pp. 383–417, 1970, [Online]. Available: http://www.jstor.org/stable/2325486

[2] M. Vijh, D. Chandola, V. A. Tikkiwal, and A. Kumar, "Stock closing price prediction using machine learning techniques," *Procedia Comput Sci*, vol. 167, pp. 599–606, 2020.

[3] S. Mehtab and J. Sen, "A time series analysis-based stock price prediction using machine learning and deep learning models," *International Journal of Business Forecasting and Marketing Intelligence*, vol. 6, no. 4, pp. 272–335, 2020.

[4] W. Chen, H. Zhang, M. K. Mehlawat, and L. Jia, "Mean–variance portfolio optimization using machine learning-based stock price prediction," *Appl Soft Comput*, vol. 100, p. 106943, 2021.

[5] Y. Wu and J. Feng, "Development and application of artificial neural network," *Wirel Pers Commun*, vol. 102, pp. 1645–1656, 2018.

[6] B. W. Wanjawa and L. Muchemi, "ANN model to predict stock prices at stock exchange markets," *arXiv preprint arXiv:1502.06434,* 2014.

[7] A. Kurani, P. Doshi, A. Vakharia, and M. Shah, "A comprehensive comparative study of artificial neural network (ANN) and support vector machines (SVM) on stock forecasting," *Annals of Data Science*, vol. 10, no. 1, pp. 183–208, 2023.

[8] P. Yu and X. Yan, "Stock price prediction based on deep neural networks," *Neural Comput Appl*, vol. 32, pp. 1609–1628, 2020.

[9] X. Pang, Y. Zhou, P. Wang, W. Lin, and V. Chang, "An innovative neural network approach for stock market prediction," *J Supercomput*, vol. 76, pp. 2098–2118, 2020.

[10] X.-S. Yang, "Flower pollination algorithm for global optimization," in *International conference on unconventional computing and natural computation*, 2012, pp. 240–249.

[11] Afiahayati *et al.*, "Forecasting the Cumulative COVID-19 Cases in Indonesia Using Flower Pollination Algorithm," *Computation*, vol. 10, no. 12, 2022, doi: 10.3390/computation10120214.

[12] M. A. Salam, L. Ibrahim, and D. S. Abdelminaam, "Earthquake prediction using hybrid machine learning techniques," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 5, pp. 654–6652021, 2021.

[13] S. Mohanty and R. Dash, "A flower pollination algorithm based Chebyshev polynomial neural network for net asset value prediction," *Evol Intell*, pp. 1–17, 2021.

[14] I. N. P. Trisna, A. Afiahayati, and M. Auzan, "Flower Pollination Inspired Algorithm on Exchange Rates Prediction Case," *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, vol. 17, no. 3.

[15] P. Singh, G. Bathla, D. Panwar, A. Aggarwal, and S. Gaba, "Performance Evaluation of Genetic Algorithm and Flower Pollination Algorithm for Scheduling Tasks in Cloud Computing," in *International Conference on Signal Processing and Integrated Networks*, 2022, pp. 139–154.

[16] N. Mondal, S. Mandal, M. C. Mandal, S. Das, and B. Haldar, "ANN-FPA Based Modelling and Optimization of Drilling Burrs Using RSM and GA," in *Global Congress on Manufacturing and Management*, 2021, pp. 180–195.

[17] A. B. Khoshaim, A. H. Elsheikh, E. B. Moustafa, M. Basha, and A. O. Mosleh, "Prediction of residual stresses in turning of pure iron using artificial intelligence-based methods," *Journal of Materials Research and Technology*, vol. 11, pp. 2181–2194, 2021.

[18] R. N. Choudhury, "Economic Implications of BRI," in *China's BRI in Different Regions of the World*, Routledge, 2023, pp. 34–45.

[19] J. Zupan, "Introduction to artificial neural network (ANN) methods: what they are and how to use them," *Acta Chim Slov*, vol. 41, p. 327, 1994.