Iraqi Journal for Computer Science and Mathematics

Journal Homepage: http://journal.esj.edu.iq/index.php/IJCM
e-ISSN: 2788-7421 p-ISSN: 2958-0544

Improved Artificial Neural Networks Based Whale Optimization Algorithm

Royida A.Ibrahem Alhayali[1], Mostafa Khaled Abd Alrahman Aladamey[2], Mohammed Rashid Subhi [3], Mostafa Abdulghafoor Mohammed[4]*, Amiza Amir[5], Zahraa A. Abdalkareem[6]

[1]Department of Architecture Engineering, College of Engineering, University of Diyala, Diyala, Iraq

[2,4,6]Alimam Aladham University College, Baghdad, Iraq

[3]Department of Petroleum System Control Engineering, College of Petroleum Processes Engineering, Tikrit University

[5] Faculty of Electronic Engineering & Technology, Universiti Malaysia Perlis, Malaysia

*Corresponding Author: Mostafa Abdulghafoor Mohammed

DOI: https://doi.org/10.52866/ijcsm.2023.02.03.015
Received April 2023; Accepted July 2023; Available online August 2023

ABSTRACT: Owing to the increasing interest in artificial neural networks (ANNs) across various fields of study, many studies have focused on enhancing their performance through the utilisation of different learning algorithms. This study examines the use of the Whale Optimization Algorithm (WOA) as a training algorithm to improve the classification accuracy of ANNsGetting artificial neural networks (ANNs) to categorize things accurately requires good model design - you need to pick the right structure, training algorithm, and activation functions. For this project, the researchers tested using the Whale Optimization Algorithm (WOA) method to train the ANN models instead of the commonly used backpropagation approach. They gathered 10 real-world datasets from the UCI machine learning repository to train and compare models on. When they compared ANNs trained with WOA against standard backprop-trained ANNs, the WOA ones did better at classifying the data. Using the whale algorithm to optimize the models led to higher accuracy in predicting the categories for the datasets. The results showed that fine-tuning an ANN with WOA outperformed just using regular backpropagation training. The classification accuracy of a WOA-trained ANN was compared with that of a backpropagation-trained ANN, and the results showed that the WOA-trained ANN exhibited superior performance.
Keywords: artificial neural network; backpropagation; whale optimization algorithm; metaheuristics.

1. INTRODUCTION

The backpropagation algorithm is commonly employed for training purposes, Some potential challenges that should be considered are convergence problems and the risk of falling into the trap of local optimality. [1]. The interesting by the researchers of the classifying optimization problems that included with the analyzing of the large datasets with many variables. The number of variables in the optimization problem, the problem's dimensions, is directly related to the overall size of the problem. As the amount of data that needs to be processed increases, so does the number of dimensions that must be taken into account [2].The Traditional mathematical frameworks also encounter the difficulty of depending on derivatives to solve optimization tasks. [3]–[5],[6].

This study attempted to discover the possibility of utilizing the Whale Optimization Algorithm (WOA) to train artificial neural networks (ANNs) with the intent of enhance their accuracy in the classification of the object. Artificial neural networks needs proper design to gain a better classification, and this includes selecting the best architectures, training methods, and efficient functions of the nodes. Our study focused on using the whale optimization algorithm to optimize the set of correlated weights in the network, which may help in finding the ideal weight for each neuron. The idea was that the whale optimization algorithm would improve the weights so that artificial neural networks could classify categories more accurately. So they tested having the algorithm tune the weights as a way to boost classification performance for the ANN models. Hence, the original WOA was improved by enhancing the framework

*Corresponding author: alqaisy86@gmail.com

through the addition of a local search parameter, contraction factors and inertia weight. Therefore, the model used to train the ANN model was based on 10 known datasets sourced from the UCI ML repository for performance improvement [7]. The classification accuracy of the WOA-trained ANN was compared with that of the backpropagation-trained ANN, and the results demonstrated the superior performance of the WOA-trained ANN. The contributions to this work are as follows: i) improvement of the classification accuracy of ANN through the use of WOA for training; ii) enhancement of step size control by incorporating the mean of whales in the algorithm, allowing all whales in a group to contribute to finding the optimal solution for the worst whale based on inertia weight and contraction factors. This article is divided into the following sections: In Section 2, we aim to provide some background information on the proposed system, and Section 3 presents a brief overview of the proposed method. The discussion of the experimental findings can be found in Section 4, while the conclusions are outlined in Section 5.

2. RELATED WORK

Scholars have focused on the issue of high-dimensional classification tasks; the dimension of an optimisation problem is determined by its size, as an increase in problem size also leads to an increase in dimensionality. The conventional mathematical frameworks face the challenge of relying on derivatives to solve optimisation tasks [8]–[10]. To address these issues, scholars have developed nature-inspired frameworks for training ANN models [11]. The study by [12] reported the training of an ANN using the Krill Herd Algorithm. The Adjusted Bat-inspired Algorithm has also been employed by many researchers for training ANNs. Several studies have utilised the modified Bat algorithm to enhance the structure and weights of ANNs, resulting in improved convergence quality [13]. Metaheuristics are categorised into population-based and single-based methods.

The use of a single-based metaheuristic for ANN training begins with one solution and incorporates neighbouring solutions to find the optimum solution [14], [15]. On the other hand, training with population-based metaheuristics begins with multiple solutions and simultaneously searches for neighbouring solutions to achieve a diverse range of solutions. The search for the optimal solution continues until a stopping condition is met. The population-based metaheuristics are subclassified into swarm intelligence and evolutionary algorithms. Among the evolutionary algorithms, the Genetic algorithm has recently found application in improving the performance of ANNs.

The ant colony algorithm has been used by some researchers to train ANNs. Adjusted ACO has also been adopted for learning the structure of feed-forward neural networks [16]. In [17], the multiverse model was used to find a solution to a global optimisation problem. PSO was developed by Kennedy & Eberhart, and since its launch, it has found applications in several areas, including ANN training [18]. The performance of adjusted PSO in finding solutions to simple problems is high, especially in tasks where other approaches may not perform well [19]. The improvement of the structure and weight of FFNN using Gaussian and Fuzzy-based models has also been reported. The use of the Gaussian distribution has been reported to improve the convergence of PSO [20].

A PSO-based approach has been proposed for determining synaptic weights in NN models to achieve low RMSE [21]–[23]. The study by Liu & Yin presented a nature-inspired PSO for ANN training. The adjusted acceleration operator of the PSO and the developed adjusted inertia weight helped optimise the weights and backpropagation thresholds of the ANN model. The efficiency of the approach was evaluated using the multimedia evaluation approach [24]. The present work proposes the use of the WOA for training ANNs. The algorithm is evaluated using 10 problem datasets sourced from the UCI ML repository to assess its performance [25].

By examining the advantages of ANNs, it should be noted that the drawbacks of ANNs are being addressed gradually, and the advantages are improving daily. It is believed that ANNs will become an indispensable part of our lives in the future due to their growing importance.

3. PROPOSED SYSTEM

In a memplex, the best and worst whales are denoted by X_b and X_w, respectively, while the best whale in the population is denoted by X_g. The following equations are utilised to ascertain the worst whale:

$$S_i = rand\,(0,1) * (X_b - X_w) \tag{1}$$

$$X_w = X_w + S_i \qquad -S_{\max} \leq S_i \leq S_{\max} \tag{2}$$

where S_{max} is the maximum feasible speed at which the whale may travel at a given latitude and longitude. i = 1, 2; the "Mgn" symbol represents the passage of time at the whale's current location. Each memplex goes through Mgn iterations, and a submemplex is built to avoid getting stuck in a local optimum. Members of the submemplex are chosen based on the fitness of the amphibians. S_i equals a random number between 0 and 1 multiplied by the difference in quality between the best and poorest whales. Calculating the worst whale's new position with a step size of S_i (Equation 2) gives us the updated position of the worst whale. If the fitness value of the new position is higher than that of the old position, the old whale is replaced. In this case, the optimal global solution X_g is substituted into X_b, as indicated in Equations 3 and 4.

$$S_i = rand\ (0,1) * (X_g - X_w) \tag{3}$$

$$X_w = X_w + S_i \qquad - S_{\max} \leq \quad S_i \ \leq S_{\max} \tag{4}$$

If the fitness value does not increase, the ranking of the previous form will be determined randomly. This will continue until the termination condition is met. Equations 1 and 2 are utilised to adjust the position of the worst whale in the original WOA, and in each memplex, the algorithm attempts to leverage the information from the best whale to improve the performance of the worst whale. In addition, if the regional best whale is unable to perform its job, only then will the global best whale be utilised. However, this is not the best way to disseminate information about global best practices to the general public. The purpose of this research was to enhance the local search capability of the WOA by developing two algorithmic models. The first model was the original WOA. Secondly, the ASFL has been modified to include the contraction factors C1 and C2, and the ranking of the worst whale has been raised by averaging it with the others. The WOA can be enhanced with the use of the following equations:

$$S = C_1 R_1 (X_b - X_w) + C_2 R_2 (mean - X_w) \tag{5}$$

$$S = \begin{Bmatrix} \min(S, smax) & S \geq 0 \\ \max(S, -smax) & S < 0 \end{Bmatrix}$$

$$\text{new } X_w = X_w + S * w \tag{6}$$

where C1 and C2 are non-negative constant acceleration parameters, R1 and R2 are random values between 0 and 1 and w is the inertia weight (0.6). The average number of whales in a given memplex is expressed as a fraction of the total number of whales in that memplex. As seen in Equation 5, C1 and C2 determine the algorithm's path towards the optimal solution. Additionally, C1 and C2 evaluate the value orientation of one of the sections of Equation 5 on the right. If C1 is greater than C2, the algorithm will randomly search for the optimal solution around the Xb value. However, if C1 is less than C2, it will randomly search around the mean value. When C1 = C2, a parallel movement is implemented on both sides to find the optimal solution. Furthermore, a large step size, S, means that the algorithm will search faster with fewer steps, but reducing the step size can affect the chance of finding the best solution. Hence, the inertia weight, w, was introduced for step size adjustment. The flowchart of the standard WOA is shown in Fig. 2.

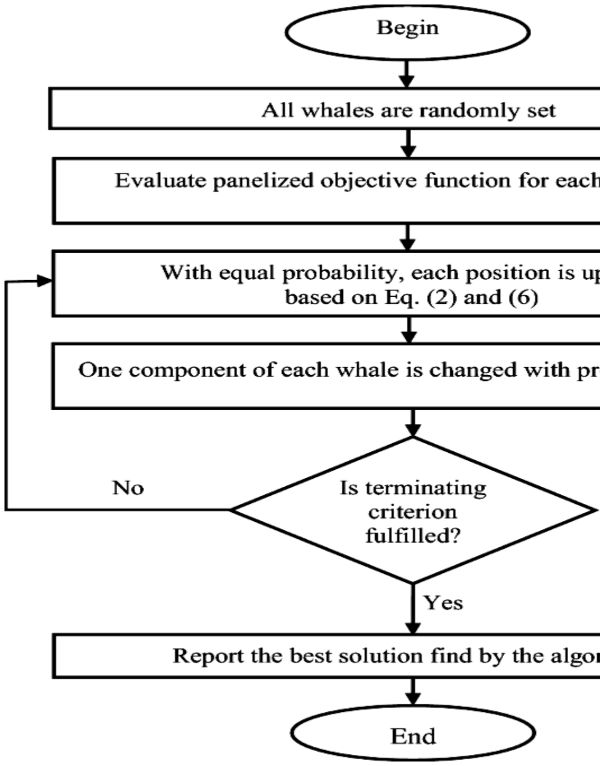

FIGURE 2. - **Flowchart for the Whale Optimization Algorithm (WOA)**

4. RESULTS AND DISCUSSION

To enhance the performance of the WOA, two variations of the algorithm are adopted. Classification problems were addressed using the WOA. The data set shown in Table 1 was used to evaluate the performance of the WOA. Programming of the adjusted algorithm was performed using C# in Visual Studio 2013. The dataset was normalised within the interval [0, 1], while the selection of initial weights for the ANN model was done randomly within the range of [-10, 10]. The hidden layer output was calculated using the Tang-Sigmoid transfer function. The computation of the output of the output layer neurons was done with the log-sigmoid transfer function. The fitness function of the algorithm was based on the mean squared error (MSE). A total of 1000 iterations were implemented for the WOA, while performance accuracy was validated using 10-fold cross-validation. Table 1 shows the parameters of the WOA used in this study.

We set the number of iterations of the WOA's (WOA) w parameter to 0.6. This number was used to adjust the distance between the original and replacement whales' locations. That's why it's useful to examine the inertia weight value when comparing the effectiveness of the algorithm with large and tiny steps. Both C1 and C2 were modified at 2, making their respective impacts on the right-hand side of the step size equation S equal. In addition, the random R1 and R2 values have opposite effects on the left and right sides of the step size equation S. The parameter values, memplex count, total population size and individual whale count were all the same between the two versions of the WOA. We compared the efficiency of our suggested method to. The results of the experiment with the WOA are listed in Table 2.

Table 1. - **Whale Optimization Algorithm (WOA) parameters**

Parameter	Value
Number of memplex	m = 4
Number of whales in each memplex	n = 9
Number of whales in the community as a whole	F = 49
The maximum allowed change in the whale position	$S_{max} = 1.8$
Constant acceleration C1 and C2	C1 = C2 = 3
Random numbers R1 and R2	Random value between 0 and 2
Weight of inertia	w = 0.5

Table 2 explains the methodology employed to determine the particular values for the parameter w, which were derived through a series of iterative tests.

Table 2. - Repeated inertia-weight tests for reliability

Dataset test	**w = 0.2**	**w = 0.6**	**w = 0.7**	**w = 0.8**	**w = 0.9**	**w = 0.1**	**w = 0.2**
Iris	94.22	**95.64**	94.00	94.32	**95.65**	94.31	**94.65**
Haber	**75.16**	73.83	72.55	71.16	71.55	72.51	73.77
Liver	63.93	63.95	73.44	74.56	65.72	75.63	**75.35**
Wine	93.45	96.04	97.88	97.26	98.64	95.66	94.85
Thyroid	94.18	93.75	99.18	90.72	90.70	**96.66**	95.70
E. coli	76.89	78.72	76.22	**87.42**	81.72	77.25	78.12
Heart	71.37	71.22	**76.92**	68.62	66.14	67.17	66.87
Balance	88.71	87.12	87.23	91.21	**90.86**	90.16	90.47
Blood	78.48	77.68	76.66	**79.68**	78.87	76.88	77.45
Diabetes	71.68	75.54	78.63	78.14	77.10	74.21	**77.66**

Table 3 indicates the rankings and shows that a value of w = 0.6 yielded the best results. Consequently, the previously established value for inertia mass was employed in the implementation of the WOA.

Table 3. - Values of inertia weight is ranked

Dataset test	w = 2	w = 0.7	w = 0.6	w = 0.5	w = 0.5	w = 0.4	w = 0.2
Iris	2	2	1	4	5	3	1
Haberman	2	3	3	5	4	5	2
Liver disorder	5	5	3	3	6	2	1
Wine	6	4	4	3.1	7	2	2
Thyroid	3	2	2	3	7	2	3
E. coli	6	3	6	1	8	3	4
Heart	2	4	2	3	8	4	7
Balance	6	6	6	4	4	3	2
Blood transfusion	3	5	7	3	3	3	3
Diabetes	6	7	6	3	2	4	1
Mean	4	4.7	4.7	6.2	**7.8**	3.3	2.2

Table 3 shows that the WOA demonstrated improved performance in both training and testing accuracy at w = 0.6 compared to w = 1 under identical conditions. The small step size of the method, when w is small, allows for a more thorough search of the solution space, leading to improved speed. When applied to the datasets, the WOA showed superior training performance with a value of w = 0.6. Except for the Heart, Haberman and Blood datasets, the suggested algorithm outperformed the other methods. In terms of testing accuracy, the proposed approach outperformed the SFLA on all datasets, except for the Haberman, Heart and Thyroid datasets. At w = 0.6, the WOA exhibited greater performance compared to the alternative technique in relation to training accuracy. However, it only demonstrated superior testing accuracy compared to the algorithm on the Haberman's Survival, Liver Illness, Blood, Iris, Wine, E. coli, Thyroid and Balance datasets. When compared to the SFLA, the results of training and testing favour the WOA.

To achieve optimal results, ANN algorithms necessitate a thorough training period. Except for the Heart dataset, the WOA performed admirably on all the other datasets. For all three data types—continuous, integer and mixed—the optimal value of w for the WOA is 0.6, not 1. The following charts display the results of the accuracy tests conducted on each dataset using various comparison algorithms.

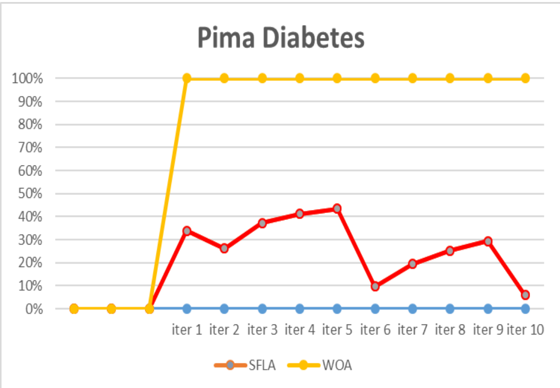

FIGURE 1. - Pima diabetes test results

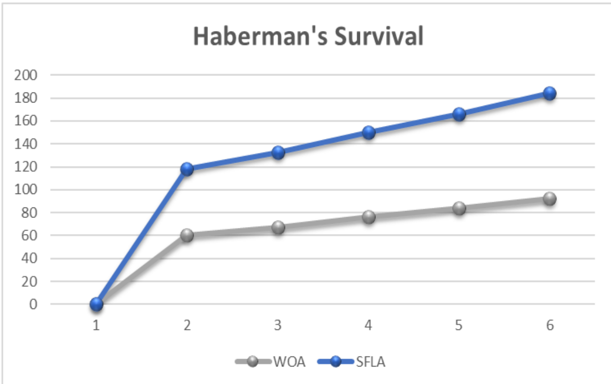

FIGURE 2. - Haberman's survival function test results

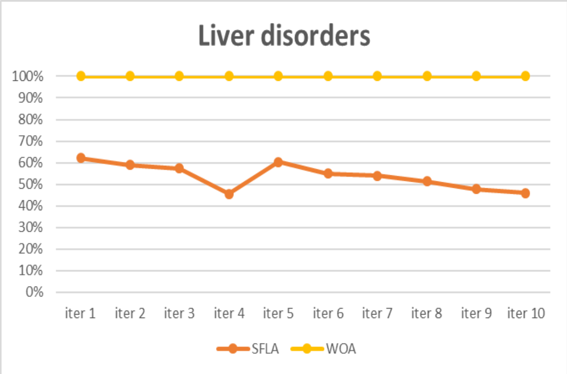

FIGURE 3. - Liver disorder algorithm comparison tests

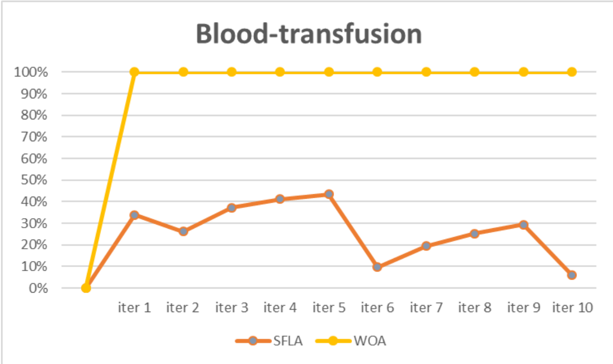

FIGURE 4. - Blood-transfusion algorithm evaluation results

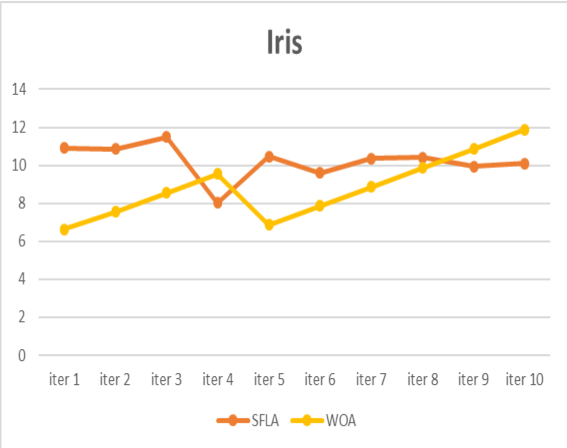

FIGURE 5. - Iris algorithm comparison test outcomes

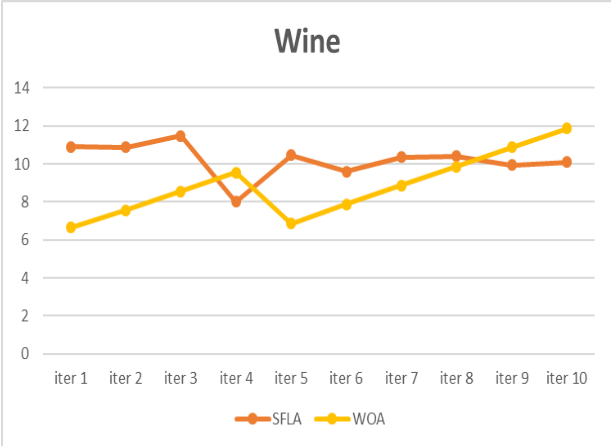

FIGURE 6. - Wine algorithm comparison test results

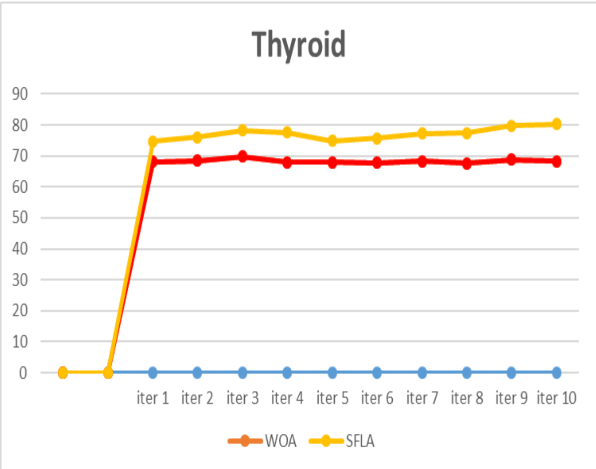

FIGURE 7. - The outcomes of thyroid tests

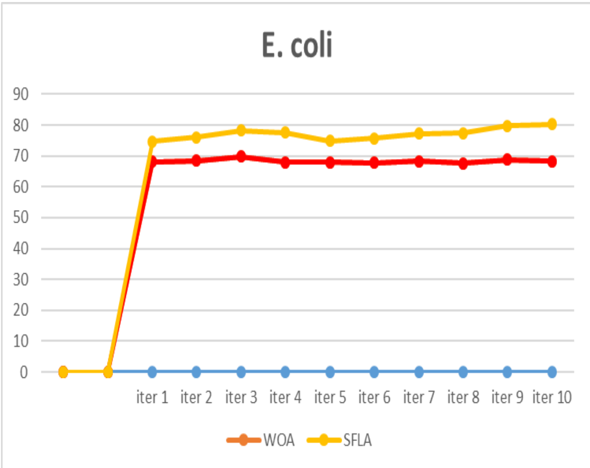

FIGURE 8. - The outcomes of the E. coli algorithm tests

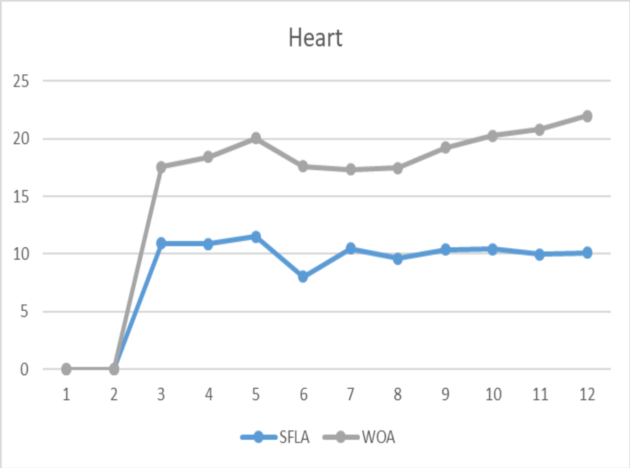

FIGURE 9. - The outcomes of heart failure comparison tests

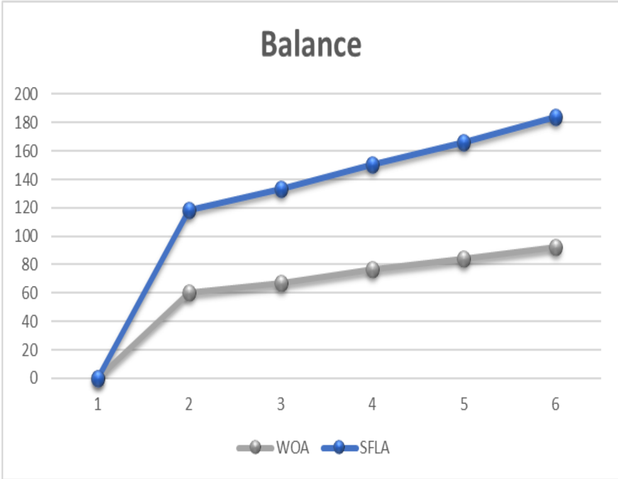

FIGURE 10. - Compared balance algorithm testing results

5. CONCLUSION

A WOA algorithm was proposed in this paper. WOA was able to identify the synaptic weights of the ANN. The efficacy of the suggested WOA-trained ANN model was measured using 10 datasets available through the UCI ML repository. The Min-Max normalisation technique was used to modify the datasets. The effectiveness of the proposed WOA algorithm was evaluated using datasets with different features. When applied to the selected datasets, WOA approaches enhanced accuracy. By incorporating the mean of the whales into the algorithm, every whale in the group was able to contribute to finding the optimal solution for the weakest whale. When compared to the SFLA algorithm, the suggested WOA often produces higher accuracy rates during both training and testing. The MSE value of the WOA is low, indicating that it converges quickly. In comparison to the other SFLA algorithms discussed in our research, the WOA has a shorter execution time. The WOA algorithm will be improved in the future by being hybridised with other optimisation strategies to provide more robust frameworks for training ANNs.

Funding

None

ACKNOWLEDGEMENT

None

CONFLICTS OF INTEREST

The author declares no conflict of interest.

REFERENCES

[1] M. Alhamdoosh and D. Wang, "Fast decorrelated neural network ensembles with random weights," *INS Information Sciences*, vol. 264, pp. 104–117, 2014.

[2] D. Chakraborty, S. Saha, and S. Maity, "Training feedforward neural networks using hybrid flower pollination-gravitational search algorithm," in *2015 International Conference on Futuristic Trends on Computational Analysis and Knowledge Management (ABLAZE)*, 2015, pp. 261–266.

[3] M. Yaghini, M. M. Khoshraftar, and M. Fallahi, "A hybrid algorithm for artificial neural network training," *Engineering Applications of Artificial Intelligence*, vol. 26, no. 1, pp. 293–301, 2013.

[4] P. A. Kowalski and S. Łukasik, "Training neural networks with krill herd algorithm," *Neural Processing Letters*, vol. 44, no. 1, pp. 5–17, 2016.

[5] R. A. I. Alhayali, M. Aljanabi, A. H. Ali, M. A. Mohammed, and T. Sutikno, "Optimized machine learning algorithm for intrusion detection," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 24, no. 1, pp. 590–599, 2021.

[6] M. M. Mijwil, R. Doshi, K. K. Hiran, O. J. Unogwu, and I. Bala, "MobilenetV1-based deep learning model for accurate brain tumor classification," *Mesopotamian Journal of Computer Science*, vol. 2023, pp. 32–41, 2023.

[7] A. H. Ali, R. A. I. Alhayali, M. A. Mohammed, and T. Sutikno, "An effective classification approach for big data with parallel generalized Hebbian algorithm," *Bulletin of Electrical Engineering and Informatics*, vol. 10, no. 6, pp. 3393–3402, 2021.
[8] M. Mijwil, I. E. Salem, and M. M. Ismaeel, "The significance of machine learning and deep learning techniques in cybersecurity: A comprehensive review," *Iraqi Journal for Computer Science and Mathematics*, vol. 4, no. 1, pp. 87–101, 2023.
[9] N. S. Jaddi, S. Abdullah, and A. R. Hamdan, "Optimization of neural network model using modified bat-inspired algorithm," *Applied Soft Computing*, vol. 37, pp. 71–86, 2015.
[10] S. Chalup and F. Maire, "A study on hill climbing algorithms for neural network training," in *Proceedings of the 1999 Congress on Evolutionary Computation-CEC99*, 1999, pp. 2014–2021.
[11] D. Hemanand and R. R. Vallem, "Cyber security system based on machine learning using logistic decision support vector," *Mesopotamian Journal of CyberSecurity*, vol. 2023, pp. 64–72, 2023.
[12] R. Abada, A. M. Abubakar, and M. T. Bilal, "An overview on deep leaning application of big data," *Mesopotamian Journal of Big Data*, vol. 2022, pp. 31–35, 2022.
[13] R. S. Sexton, B. Alidaee, R. E. Dorsey, and J. D. Johnson, "Global optimization for artificial neural networks: A tabu search application," *European Journal of Operational Research*, vol. 106, pp. 570–584, 1998.
[14] K. M. Salama and A. M. Abdelbar, "Learning neural network structures with ant colony algorithms," *Swarm Intelligence*, vol. 9, pp. 229–265, 2015.
[15] N. M. Hussien, Y. M. Mohialden, N. T. Ahmed, M. A. Mohammed, and T. Sutikno, "A smart gas leakage monitoring system for use in hospitals," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 19, pp. 1048–1054, 2020.
[16] Z. F. Hussain and H. R. Ibraheem, "Novel convolutional neural networks based Jaya algorithm approach for accurate deepfake video detection," *Mesopotamian Journal of CyberSecurity*, vol. 2023, pp. 35–39, 2023.
[17] S. Mirjalili, S. M. Mirjalili, and A. Hatamlou, "Multi-verse optimizer: A nature-inspired algorithm for global optimization," *Neural Computing and Applications*, vol. 27, pp. 495–513, 2016.
[18] R. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," in *MHS'95., Proceedings of the Sixth International Symposium on Micro Machine and Human Science*, 1995, pp. 39–43.
[19] M. A. Kadhim and A. M. Radhi, "Heart disease classification using optimized machine learning algorithms," *Iraqi Journal for Computer Science and Mathematics*, vol. 4, no. 2, pp. 31–42, 2023.
[20] V. G. Gudise and G. K. Venayagamoorthy, "Comparison of particle swarm optimization and backpropagation as training algorithms for neural networks," in *Proceedings of the 2003 IEEE Swarm Intelligence Symposium. SIS'03*, 2003, pp. 110–117.
[21] S. Chatterjee, S. Sarkar, S. Hore, N. Dey, A. S. Ashour, and V. E. Balas, "Particle swarm optimization trained neural network for structural failure prediction of multistoried RC buildings," *Neural Computing and Applications*, vol. 28, no. 8, pp. 1–12, 2016.
[22] T. Liu and S. Yin, "An improved particle swarm optimization algorithm used for BP neural network and multimedia course-ware evaluation," *Multimedia Tools and Applications*, vol. 76, pp. 11961–11974, 2017.
[23] M. S. I. Alsumaidaie, K. M. A. Alheeti, and A. K. Alaloosy, "Intelligent detection of distributed denial of service attacks: A supervised machine learning and ensemble approach," *Iraqi Journal for Computer Science and Mathematics*, vol. 4, no. 3, pp. 12–24, 2023.
[24] R. A. Hasan, R. A. I. Alhayali, M. A. Mohammed, and T. Sutikno, "An improved fish swarm algorithm to assign tasks and cut down on latency in cloud computing," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 20, pp. 1103–1108, 2022.
[25] H. Zhao, O. Gallo, I. Frosio, and J. Kautz, "Loss functions for image restoration with neural networks," *IEEE Transactions on Computational Imaging*, vol. 3, pp. 47–57, 2016.