



Optimizing Connection Weights in a Long Short-Term Memory (LSTM) Using Whale Optimization Algorithm (WOA): A Review

¹Abuzairu Ahmad, ²Kabiru Ibrahim Musa, ¹Fatima Umar Zambuk, ¹Mustapha Abdulrahman Lawal

¹Department of Mathematical Sciences,
Faculty of Science,

²Department of Management Information Technology,
Abubakar Tafawa Balewa University Bauchi

ABSTRACT

Many academics have recently developed an interest in the learning process of long short-term memory, which is widely accepted as one of the trickiest issues in machine learning. The main challenges of the standard training procedure are usually local optima stagnation and slow convergence. As a result, the whale optimization algorithm can be used to solve long short-term memory local minima. Optimizing connection weights can go a long way toward assisting the authority to properly addressing network problems. A review of the related literatures for optimization weights has been conducted in this paper with the sole purpose of identifying a gap for future research. A keyword search was conducted to locate relevant literatures from various academic databases, and the papers gathered were reviewed to determine their areas of strength and weakness. The term "independent" refers to a person who does not work for the government.

ARTICLE INFO

Article History

Received: June, 2022

Received in revised form: July, 2022

Accepted: September, 2022

Published online: September, 2022

KEYWORDS

Whale optimization algorithm, WOA, Long short term memory, LSTM, Training neural network, evolutionary algorithm

INTRODUCTION

Artificial neural networks (ANNs) are nonparametric, intelligent mathematical models that draw inspiration from the biological nervous system. In the past three decades, ANNs have received a great deal of attention and have been used to solve several classification, pattern recognition, regression, and forecasting issues. (Schmidhuber, 2015; Chatterjee *et al.*, 2016; Braik *et al.*, 2008; Linggard *et al.*, 2012; Huang *et al.*, 2015; Rezaeianzadeh *et al.*, 2014). The mechanism of an ANN's learning has a significant impact on its effectiveness.

A lot of academics have recently become interested in artificial neural network learning because it is one of the hardest problems in machine learning. The nonlinearity and lack of knowledge of the optimal set of primary regulating parameters make training a neural network problematic (weights and biases). The main problems of conventional training methods are delayed convergence and local optima stagnation. As a result, the stochastic optimization algorithm offers a solid choice for addressing these

problems (Aljarah *et al.*, 2016). Deep learning (DL) has recently emerged as the big data analysis trend with the quickest rate of growth. It has been widely and successfully used in a variety of computer applications, including speech recognition, processing of neural language, sequential data, and image classification (Abdel-Hamid, Mohamed, Jiang, & Penn, 2017). Due to its superior performance as compared to that of conventional learning algorithms, Machine learning approaches are becoming more and more significant as the paradigm shift toward data-intensive science occurs. Deep learning, in particular, has demonstrated to be an incredibly potent tool in many domains, constituting a significant advancement in the science.

Statement of the Problem

Single-hidden-layer artificial neural networks (ANNs) are always less accurate than ANNs with several hidden layers (Mirjalili *et al.*, 2018). The issue of local minima is still unresolved despite the literature's deployment and investigation of a broad variety of evolutionary and



swarm-based techniques for training MLP (Aljarah *et al*, 2018). In order to identify issues, the whale optimization algorithm (WOA) makes use of a population of search agents. Making a set of arbitrary solutions (candidate solutions) to a particular problem is the first step in the search process. It then enhances this set till a final criterion is met (Mirjalili and Lewis, 2018).

Aim and Objectives

The aim of this work is to optimize the connection weights in a long short-term memory using whale optimization algorithm. This research has the following objectives to;

- (i) Propose a novel long short-term memory (LSTM) training technique using whale optimization algorithm (WOA).
- (ii) Compare and visualize the results obtained on different datasets in order to find the best classification accuracy, convergent speed, and precision score.
- (iii) Developed WOA-based approach in this research will be evaluated and tested based on six (6) selected cancer classification datasets.

LITERATURE REVIEW

In order to properly grasp how neural networks operate, their merits and shortcomings, their limitations, and how they accomplish their goals, this part provides a quick introduction to computational techniques used in the field of neural network optimization as mentioned in the literature.

Machine Learning

The process of evolution gave rise to the brain's capacity to mimic the world. It didn't emerge because such forward prediction was specifically optimized for; rather, it did so because it helped organisms thrive and survive in their unique settings. The capacity of modules of current world model techniques to anticipate

future states of the environment is frequently directly rewarded, in contrast to the emergent brain representations in nature (Schmidhuber 1990; Ha and Schmidhuber 2018; Hafner *et al*,, 2018; Wayne *et al*,, 2018). While it is obviously advantageous to be able to explicitly motivate a model to anticipate what will happen next, we are interested in the more challenging challenge of agents that should learn to predict what is crucial for their survival without being overtly motivated to do so.

A problem in end-to-end training of complicated neural models without requiring individual training of each component (Schmidhuber *et al*. 2018). The infamous credit assignment problem (Minsky *et al*,, 1961). Deep learning has proven to be generally effective at solving the CAP for deep networks (i.e., figuring out how much each weight contributes to the error of the network), but research suggests that more diverse networks may do better (SchmidHuber *et al*, 2018).

OVERVIEW OF FUNDAMENTAL CONCEPT AND COMPUTATIONAL INTELLIGENCE

The following sub sections provide an overview of some basic ideas that are frequently used in relation to this planned research project.

Artificial Neural Network

An artificial neural network (ANN) is a type of computational intelligence made up of neurons. It draws inspiration from biological neural networks. Since the neurons receive information from their input nodes and aggregate it, it models the complicated relationship between the input and output. Each neuron's activation is calculated, and the response is conveyed across the output node of neurons. As depicted in Figure 1, the input, hidden, and output layers make up the classic ANN's three primary structural layers. These strata are linked together. Numerous neurons with an activation function make up each of these levels.

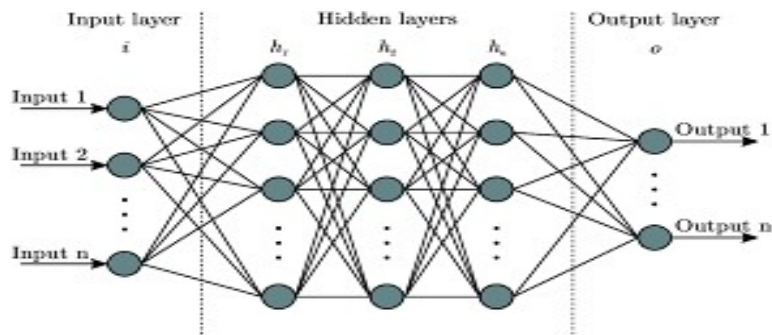


Figure 1: Structure of Artificial neural network (Bre, Gimenez, &Fachinotti, 2018).

Long Short-Term Memory Neural Network

Recurrent neural networks (RNNs) with a particular architecture called long short-term memory were created to simulate temporal sequences. Because of its long-range dependencies, LSTM is more accurate than traditional RNNs. It was first suggested by (Hochreiter et al., 1997). Error backflow is a problem caused by the back propagation technique in RNN architecture. Memory blocks are unique units found in the recurrent hidden layer of LSTM unlike RNN. In addition to specific multiplicative units called gates that regulate the flow of information, the memory blocks also

include memory cells with self-connections that store the network's temporal state. In the original architecture, there were three different gate types for each memory block, including: (1) Input gate: This gate regulates how input activations go to the memory cell. (2) Output gate: This gate regulates how cell activations exit the network and enter the rest of the system. (3) Skip the gate: Output gate controls the output flow of cell activations into the rest of network. (3) **Forget gate**: by scaling the internal state of the cell before adding it as input through the cell's self-recurrent link, the cell's memory is adaptively forgotten or reset.

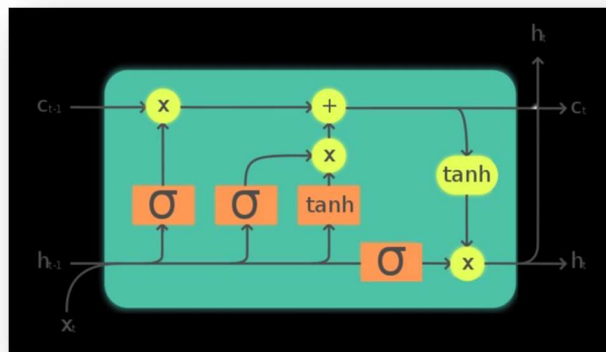


Figure 2: Long short-term memory (LSTM) neural network architecture (Guillaume Chevalier et al, 2018)

REVIEW OF RELATED EMPIRICAL STUDIES

Shallow learning-based and deep learning-based approaches can be used to classify the RUL prediction algorithms. When the performance of the prediction model is superior to

the available computing power, deep learning has become the preferred option. Here, all previous research on the RUL prediction of turbofan engines using the same benchmark dataset is compared fairly and consistently. The approach of



related works is described in a concise synopsis. Brief learning-based approaches are firstly summarized.

The best solutions will be utilized as parents in one evolutionary algorithm called GA. These solutions will be crossed to create an improved version of the parent solutions (Suganda et al, 2019). The regression model's parameters were adjusted using a genetic algorithm (GA). The performance's RMSE was on average 47.63. A novel strategy that included the maximum Rao-Blackwellized Particle filter, kernel two sample tests, and maximum mean discrepancy was developed in place of conventional machine learning algorithms (Cai et al, 2020).

The suggested model is a two-layer, constrained Boltzmann machine-based GA (Ellefsen et al, 2021). It received a 19.8 average RMSE. A single layer vanilla LSTM model and the Adam adaptive learning rate optimization algorithm were used, and the resulting average RMSE was 28.4. (2018). Similar to that, the Adam

adaptive learning rate optimization approach was used to enhance the LSTM model (Wu et al, 2020). Evaluation revealed that 5 layers and 100 neurons per layer produced the best results (average RMSE of 19.1). Execution times are undoubtedly much accelerated when using heuristic optimization strategies (Suganda et al, 2020). For weather forecasting, it was proposed to use a single-layer and multi-layer long short-term memory (LSTM) model using intermediate variables (Salman et al., 2018) Multiple layers LSTM is the most effective LSTM model in this experiment, and pressure variable is the best intermediate data. This model's validation accuracy was 0.8060 and RMSE was 0.0775 when pressure was used as a variable. The best model for predicting water quality is the LSTM after deep learning was used to improve it (Amir et al, 2021). Additionally, they merged an LSTM neural network with the whale optimization technique (Lei et al., 2020), and the RMSE of the LSTM-NN was 0.0041.

Table1. An overview of relevant works organized by proposed technique, contributions, findings, and future research directions

S/N	Author Name	Year	Title	Proposed model	Contributions	Results	Issues/Gap
1.	KT Chui et al,	2021	RNN-LSTM model for remaining usable life was optimized by a genetic algorithm. Turbofan engine forecast.	RNN-LSTM	RUL prediction of turbofan engines under short- and long-term contentions benefits from the combination model RNN-LSTM. The results show that it lowers RMSE by 6.07–14.72% when compared to standalone RNN and standalone LSTM.	According to the results, CEEMD-WPT reduces the RMSE utilizing the individual dataset FD001–FD004 by 5.14–27.15 percent compared to CEEMD, EEMD, EMD, WPT, EEMD–WPT, and EMD–WPT.	Recurrent neural networks (RNN) and LSTM are combined because they have advantages in managing both short-term and long-term RUL forecasts.
2.	Salman et al,	2018	Weather prediction using	LSTM	The relationship	Multiple layers	This effort has failed to

Corresponding author: Ahmad, A. ✉ abuzairuhammad2020@gmail.com ✉ Department of Mathematical Sciences, Faculty of Science, Abubakar Tafawa Balewa University, Bauchi. © 2022. Faculty of Technology Education. ATBU Bauchi. All rights reserved



			a single-layer and multi-layer long short-term memory (LSTM) model with intermediary variables		between the expected and intermediate variables can be formulated or explained by the LSTM.	LSTM is the most effective LSTM model in this experiment, and pressure variable is the best intermediate data.	build the combination of LSTM and convolution neural network after attempting several strategies for smoothing the data input (CNN) models.
3.	Wang <i>et al</i> ,	2020	Prediction of energy usage for 3RRR PPM using LSTM neural network with whale optimization method.	WOA trainer	After gathering training data, the whale optimization algorithm (WOA) reduced the RMSE of the LSTM-NN to 0.00041 without the need for the gradient of the loss function, significantly enhancing the effectiveness of the training LSTM-inability NN's to solve predictions.	According to the findings, the relative error between experimental data and predictions was less than 2.50%.	To achieve the intelligent use of 3-RRR PPMs, the energy consumption forecast based on LSTM-NN will be helpful.
4.	Lei <i>et al</i> ,	2021	A regularized LSTM approach for estimating rolling bearings' remaining useful life	LSTM	In this paper, an improved LSTM algorithm known as E-LSTM is introduced. It was successful in resolving the over-fitting issue during the LSTM	The long-term dependence issue is successfully resolved by the e-LSTM technique, which performs better than RNN.	The investigation of techniques to speed up E-LSTM calculations was unsuccessful in this work.



5.	Y. Lei <i>et al,</i>	2018	A model-based approach for estimating the useful life of machinery	MLE algorithm	model's training process. Our primary goal is to raise the overall average accuracy of RUL forecast outcomes.	The findings indicate that, among the three ways, the proposed model-based method performs the best in terms of RUL prediction of machinery.	The suggested solution still has certain drawbacks. For instance, we want to build a health indicator with a monotonic trend in the indicator development module.
6.	Elsaid <i>et al,</i>	2018	Recurrent neural networks with long- and short-term memory that have been enhanced by ant colonies	LSTM-RNN	Contributions can be translated using some significant metrics that are measured and stored on the flight data recorder. Since vibrations are likely to be the result of a combination of these factors, it is very challenging to pinpoint the exact origin of the excess in vibrations.	The results show that by optimizing the gates within LSTM cells, the ACO method can drastically reduce the number of connections required while enhancing the recurrent neural network's capacity for prediction.	The attempt to individually develop each mesh in the LSTM RNN using ACO failed.
7.	Esaid <i>et al,</i>	2020	Deep recurrent neural networks' neuroevolutionary transfer learning is accomplished by network-aware adaptation.	RNN	It has been demonstrated that N-ASTL can execute transfer learning on takes where it	The results show that N-ASTL approaches produce performance that is considerable	The difficult task of using RNNs to forecast time series has been the main



					was previously.	y better in all investigations.	focus of this work.
8.	Amina <i>et al.</i>	2020	Comparative effectiveness of machine learning techniques for predicting indoor daylight illuminances	ML algorithm	Our study's contribution relates to the best methods for forecasting indoor illuminations.	Our findings indicate that lengthy short-term memory is precise and dependable.	The biggest drawback concerned the type of data used to create and test the proposed models.
9.	S. siami <i>et al.</i> ,	2018	Forecasting time series in the economy and finance.	ARIMA and LSTM	Compare the effectiveness of LSTM with ARIMA in terms of the reduction of prediction error rates.	The RMSE numbers clearly show that LSTM-based models perform significantly better than ARIMA-based models (84%–87% less error rates).	Applying similar techniques to other situations and datasets with various amounts of features was intended to explore the improvement brought about by deep learning.
10.	Suganda <i>et al.</i> ,	2019	LSTM with artificial bee colony optimization for Bit coin price forecast	LSTM	This work's key contribution is the combination of ABC and LSTM to optimize the hyper parameter and produce models with low RMSE.	The outcome shown that the ABC-LSTM model performed better than LSTM models that were not optimized.	The comparison between hyper parameter optimization on LSTM-based models utilizing ABC and GA is the limitation.
11.	Anita <i>et al.</i> ,	2020	LSTM time series prediction optimization for the Indian stock market.	LSTM	One-way ANOVA was used to assess whether this data could be	Due to its greater stability, this suggests that the	The basic LSTM architecture is the only thing that has to be



					statistically validated.	stateless LSTM model is preferred.	tuned. Similar tuning techniques can be used with many LSTM models.
12.	Aljarah <i>et al</i> ,	2018	Utilizing the whale optimization technique to optimize connection weights in neural networks.	MLP	The outcomes demonstrated that on the majority of datasets, the suggested WOA-based training method outperformed the existing algorithms.	The outcomes were more favorable in terms of convergence and accuracy.	Other ANN types were not successfully trained using the WOA in this experiment.

METHODOLOGY

The Proposed System

Optimizing the connection weights in neural networks is the goal of this research. The suggested system uses the Whale Optimization Algorithm to train its long short-term memory (WOA).

Long Short-Term Memory

An artificial recurrent neural network (RNN) architecture called a long short-term memory (LSTM) is utilized in deep learning. LSTM has feedback connections in contrast to traditional feed forward neural networks. It can analyze whole data sequences in addition to single data points (like photos) (such as speech or video). For instance, LSTM can be used for tasks like linked,

unsegment handwriting identification, speech recognition, and network traffic anomaly detection, or IDSs (intrusion detection systems). A cell, an input gate, an output gate, and a forget gate make up a typical LSTM unit. The three gates control the flow of information into and out of the cell, and the cell remembers values across arbitrary time intervals. Onetime series classification, processing, and prediction are all tasks that are well suited to LSTM networks. To solve the vanishing gradient issue that can arise when training conventional RNNs, LSTMs were created. The advantage of LSTM over RNNs, hidden Markov models, and other sequence learning techniques in many applications is their relative insensitivity to gap length.

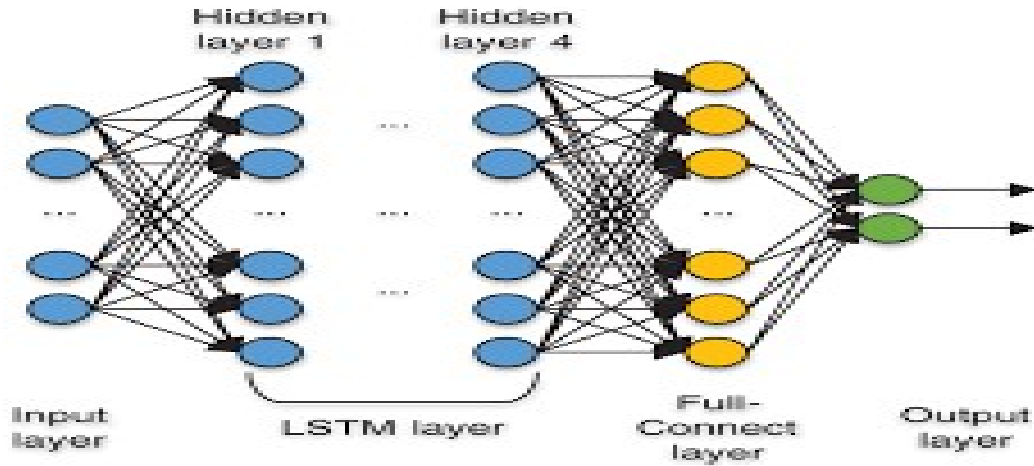


Figure 3: Structure of Long Short-Term Memory (LSTM) neural network (Demeng wu et al, 2021)

A typical LSTM with an input layer, four hidden layers, and an output layer is depicted in the image above. Input neurons in the input layer communicate with the hidden layers. Each neuron has one output, one activation function (which determines the output of a certain input), and weighted inputs (synapses). The movable components known as synapses are what

transform a neural network into a parameterized system. The activation signal is created by the weighted sum of the inputs and is then sent to the activation function to cause the neuron to produce one output. Linear, step, sigmoid, tanh, and ReLU functions are the most often utilized activation functions.

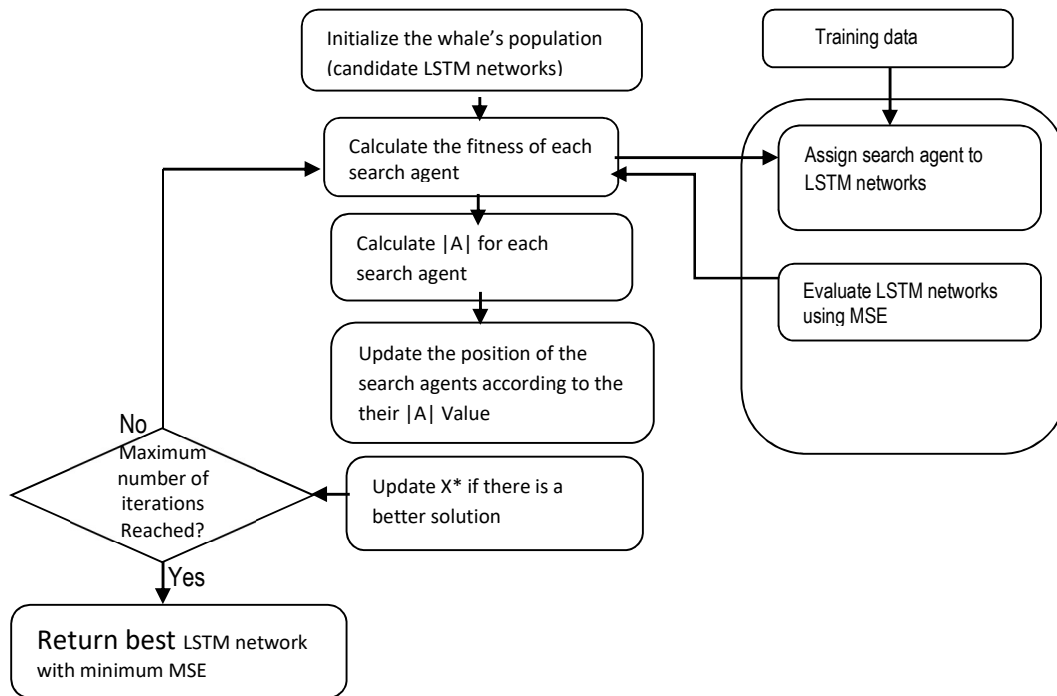


Figure 4: Framework of the proposed system

The Whale Optimization Algorithm

A recently developed stochastic optimization technique is called the whale optimization algorithm (WOA) (Mirjalili and Lewis, 2018). For optimization issues, it uses a population of search agents to identify the global optimum. The search procedure begins by generating a set of random solutions (candidate solutions) for a particular issue, much like other

population-based algorithms. It then enhances this set till a final condition is met. The criteria that enhance the potential solutions at each optimization phase are the key distinction between WOA and other algorithms. In reality, WOA imitates humpback whale hunting behavior by using a technique called bubble-net feeding to locate and attack prey. The following diagram depicts the bubble-net hunting model:

Figure 5: Bubble net hunting behavior (Abdel-Salam et al, 2018)

A humpback whale can be seen setting up a trap in this illustration by travelling in a spiral pattern around its prey and blowing bubbles as it goes. The WOA's primary source of inspiration is this strategy of shrewd foraging. The encircling mechanism is another simulated action of humpback whales in WOA. To begin their bubble-net method of hunting, humpback whales surround their target. It is important to note that humpback whales are the only species known to engage in bubble-net feeding. The spiral bubble-net feeding maneuver is mathematically analyzed and optimized in this work. This algorithm's primary mathematical equation is as follows:

$$X(t+1) = X^*(t) - AD \quad \text{for } p < 0.5 \dots\dots\dots 2$$

$$X(t+1) = D'e^{bl} \cos(2\pi t) + X^*(t) \quad \text{for } p \geq 0.5 \dots\dots 3$$

Where:

P is a random number in [0, 1],

X is a position vector,

X* is the position vector of the best solution obtained so far,

D' = |X*(t) - X(t)| indicates the best solution obtained so far,

b is a constant for defining the shape of the logarithmic spiral,

l is a random number in [-1, 1],

t shows the current iteration,

$$D = |CX^*(t) - X(t)| \dots\dots\dots 4$$

The vectors A and C are calculated as follows:

$$A = 2ar - a \dots\dots\dots 5$$

$$C = 2r \dots\dots\dots 6$$

A linearly decreases from 2 to 0 over the course of iterations (in both exploration and exploitation phases), and r is a random vector in [0, 1].

While equation (2) imitates the bubble-net method, equation (1) simulates the encircling process. The chance of switching

Between these two components is the same for the variable p. The WOA generates a set of random solutions before beginning to optimize a given problem. The best search agent so far or a randomly chosen search agent is used to update search agent positions in each optimization phase. The pivot point to update the position of additional search agents when $|X| > 1$ is the best option to ensure exploration and convergence. In other cases, the best solution found thus far serves as the pivot point (when $|X| \leq 1$). The WOA's pseudo codes are displayed in Algorithm 1.

Algorithm 1 (Pseudocodes of WOA)

```

1. Initialize the whales' population  $X_i$  ( $i = 1, 2, 3, \dots, n$ )
Initialize a, A, and C
Calculate the fitness of each search agent
 $X^*$  = the best search agent
Procedure WOA (Population, a, A, C, MaxIter,)
    t=1
    While t ≤ MaxIter do for each search agent do
        if  $|A| \leq 1$  then
            Update the position of the current search agent
            using eqn. (4)
        else if  $|A| \geq 1$  then
            Select a random search agent  $X_r$  and
            Update the position of the current agent
            using eqn. (2)
        end if
    end for
    Update a, A, and C
    Update  $X^*$  if there is a better solution
    t = t + 1
end while
return  $X^*$ 
  
```



end procedure

Evaluation Parameters and Performance Metrics

The performance of the suggested system will be assessed after it has been put into place. The performance criteria used in this work include accuracy, convergence speed, and precision score. The following parameters are mathematically calculated;

i. Accuracy: This performance metric pertains to the model's accurate prediction, and it may be written as:

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+FN+TN} \dots\dots\dots 7$$

ii. Precision: Precisions show how accurate and precise your model is by showing how many of the positive results that were predicted are actually positive. When the cost of false positives is substantial, precisions is a useful metric to use. In mathematics, it is written as:

$$\text{Precision} = \frac{TP}{TP+FP} \dots\dots\dots 8$$

IMPLEMENTATION OF THE SYSTEM

MATLAB R2021a was used to implement the described algorithms. The system in use is an HP laptop with a Pentium® Core i7 processor, 8GB of RAM, and a 1TB hard drive. We implemented the proposed WOA trainer and other algorithms in MATLAB R2021a for all experiments. In order to preserve the class distribution as much as feasible, the datasets are split into 70% for training and 30% for testing using stratified sampling. Furthermore, all datasets are standardized using min-max normalization to remove the impact of features with various scales. Each experiment is run through 50, 100, 150, 200, and 250 iterations over the course of five different runs. The simulation was conducted for 250 iterations for each of six (6) distinct cancer datasets.

METHOD OF DATA COLLECTION

The primary source of the data used in this research was a journal article written by Aljarah I. et al. (2018). During the data gathering, a large number of other journals and websites were visited. As a result, the study uses six (6) distinct cancer

datasets, including those for the bladder, leukemia, lung, ovary, pancreas, and prostate, respectively.

CONCLUSIONS AND FUTURE RESEARCH DIRECTION

This research work can extend to other variants of deep recurrent neural networks; thus, future work should consider training other variants of RNNs. According to the findings of this work, none of these models were employed when the whale optimization technique was applied to optimize link weights in a long short term memory. It is important to take into account how the WOA-trained MLP might be used to solve engineering categorization issues (Aljarah et al., 2018). It was also suggested that employing the WOA-trained MLP to solve function approximation datasets could be a useful addition (Mirjalili et al., 2018).

REFERENCES

Aljarah, I., Faris, H., & Mirjalili, S. (2018). Optimizing connection weights in neural networks using the whale optimization algorithm. *Soft Computing*, 22(1), 1-15.

Chui, K. T., Gupta, B. B., & Vasant, P. (2021). A Genetic Algorithm Optimized RNN-LSTM Model for Remaining Useful Life Prediction of Turbofan Engine. *Electronics*, 10(3), 285.

ElSaid, A., Jamiy, F. E., Higgins, J., Wild, B., & Desell, T. (2018, July). Using ant colony optimization to optimize long short-term memory recurrent neural networks. *In Proceedings of the Genetic and Evolutionary Computation Conference* (pp. 13-20).

ElSaid, A., Karns, J., Ororbial II, A., Krutz, D., Lyu, Z., & Desell, T. (2020). Neuroevolutionary Transfer Learning of Deep Recurrent Neural Networks through Network-Aware Adaptation. *arXiv preprint arXiv:2006.02655*.

Gao, Y., Chen, K., Gao, H., Zheng, H., Wang, L., & Xiao, P. (2020). Energy consumption prediction for 3-RRR PPM through combining LSTM neural network with whale optimization

Corresponding author: Ahmad, A. ✉ abuzairuhammad2020@gmail.com ✉ Department of Mathematical Sciences, Faculty of Science, Abubakar Tafawa Balewa University, Bauchi. © 2022. Faculty of Technology Education. ATBU Bauchi. All rights reserved



- algorithm. *Mathematical Problems in Engineering*, 2020.
- Lei, Y., Li, N., Gontarz, S., Lin, J., Radkowski, S., & Dybala, J. (2019). A model-based method for remaining useful life prediction of machinery. *IEEE Transactions on reliability*, 65(3), 1314-1326.
- Liu, Z. H., Meng, X. D., Wei, H. L., Chen, L., Lu, B. L., Wang, Z. H., & Chen, L. (2021). A regularized LSTM method for predicting remaining useful life of rolling bearings. *International Journal of Automation and Computing*, 18(4), 581-593.
- Ngarambe, J., Irakoze, A., Yun, G. Y., & Kim, G. (2020). Comparative performance of machine learning algorithms in the prediction of indoor daylight illuminances. *Sustainability*, 12(11), 4471.
- Salman, A. G., Heryadi, Y., Abdurahman, E., & Suparta, W. (2018). Single layer & multi-layer long short-term memory (LSTM) model with intermediate variables for weather forecasting. *Procedia Computer Science*, 135, 89-98.
- Siami-Namini, S., & Namin, A. S. (2018). Forecasting economics and financial time series: ARIMA vs. LSTM. *arXiv preprint arXiv:1803.06386*.
- Yadav, A., Jha, C. K., & Sharan, A. (2020). Optimizing LSTM for time series prediction in Indian stock market. *Procedia Computer Science*, 167, 2091-2100.
- Yuliyono, A. D., & Girsang, A. S. (2019). Artificial bee colony-optimized LSTM for bitcoin price prediction. *Advances in Science, Technology and Engineering Systems Journal*, 4(5), 375-383.