Journal of Positive Sciences (JPS) **eISSN: 2582-9351, https://positive-sciences.com** **JPS. 2021 Sep; 1(6): 1-5**

Article info **Research Article**
Received on: 20.08.2021
Accepted on: 28.09.2021
Published on: 30.09.2021
doi: https://doi.org/10.52688/ASP52535

Facebook metrics classifications using whale optimization algorithm and artificial neural network

Mohammed Al-Bayati [1, *]

[1] Department of Information Technology, Altinbaş University, Istanbul, Turkey
[*] mohammed.mt2016@gmail.com

ABSTRACT

This paper explores the official hybridization of two Artificial intelligence approaches for the 'whale' 'optimization' 'algorithm' ('WOA') and 'artificial' 'neural' 'network' ('ANN') on the Facebook metrics Estimate. The Whale optimization algorithm is a global search technique, while a local search method is the method of gradient backpropagation. In order to reduce the chances that an ANN slowly mixed with the local limit and converging to the global ideal, ANN has calculated the relationship between inputs and output variables, using the stochastic search power. The results of this study show that in contrast to the traditional neural back-propagation network (BPNN), the proposed model increase the convergence rate as well as predictable accuracy for the hybrid model (WOA-ANN).

Keywords: Whale Optimization Algorithm (WHO), Artificial Neural Network (ANN), backpropagation, multilayer perceptron, classification

INTRODUCTION

The vast amount of data (registers) is considered to be the most important resources for the collection and interpretation of knowledge which allows physicians to assume that they are right and that increase the probability of survival. In the field of health, scientists use (Backpropagation) BPNN to create unstructured problems because they are complex and nonlinear in relation to input and output variables. The reference algorithm for propagating the weight and distortion of a downward gradient is the regional search algorithm. Reduce fitness due to the root error average between the current output and the anticipated ANN function. The downside of the back-propagation algorithm is quickly stuck in a local minimum and slow convergence. [1] used a network of nerve back spread to predict neonatal disease diagnoses. In various datasets of neonatal diseases the researchers applied normal back propagation. The new model's efficiency is 75 percent more reliable. A number of hybrid prediction models, including heart regression (LR), multivarian MARS, artificial neural network (ANN), and raw range (RRS) have been suggested by [2]. The model's efficiency is higher than the neural network artificial. The hybrid network of neural artificial (ANN) and neural fuzzy networks (FNN) was developed by [3] to prevent diabetes and heart diseases. In the proposed model 84.24% and 86.8% were accepted for the data set for Pima Indian diabetes and for Cleveland cardiovascular disease. Milan and Sunila have researched a range of algorithms for the extraction of cardiovascular data, including the artificial neural web (ANN, SVM, Decision-making tree, and *Ripper). SVM learned algorithms for data mining. The accuracies of the support vector machines (80.06%), ANN (84.12%), the |Decision| Tree (79.05%) , and Ripper is (81.08%). The accuracy of cardiovascular disease can be estimated by SVM. The effects of class disparity in education data using the Medical Diagnostic Neural Network have been analyzed by [4]. Standard back-propagation (BP) and particle swarm optimization PSO have been used for two algorithms. There are two algorithms. The results show, for large and small data sets, that BP exceeds PSO performance. [5] have used SAS fundamental technology 9.1.3 to predict heart disease. The model proposed is built on the basis of neural networks. The accuracy obtained is 89.01%, which is not optimal but considered to be decent enough. [6] developed a new algorithm for hybrid predictions with a missing imputation value. A model based on classification of K-means and the perceptron was developed by the researchers. Three clinical sets of reference data, namely diabetes of Pima, UCI machine learning databases of Wisconsin breast cancer, and hepatitis, have evaluated the performance of the proposed algorithm. The findings are very good and the proposed model would suit very well if other missing values are introduced into the DataSet. The Bajaj et al. have introduced three algorithms for data classification, such as the |decision| |tree|, the split |validation| and the prototype [7]. The model suggested eliminates medical errors and makes a forecast with high accuracy. (CAPSO) method is proposed by [8]. The method proposed is used to improve the performance and reliability of ANN training. The quality of none standard medical data sets has been tested and the findings are positive. More analysis on [9] is available for interested readers. Back-propagation algorithm used to produce various weights and preconditions each re-executed during the training phase. This results in different prediction

***Corresponding author**
Mohammed Al-Bayati,
Department of Information Technology, Altinbaş University, Istanbul, Turkey
e-mail: mohammed.mt2016@gmail.com

outcomes and convergence speeds in each sprint. The downside of back propagation algorithms was solved by using the Gray Wolf Optimizer (GWO) so that the weight and preferences for back propagation can be found at the end. GWO is one of the most recent organic algorithms for automating gray animals hunting behavior. GWO seeks optimal solutions for reduced local capture and speed mitigation. GWO is looking in various directions for the best solutions.

MATERIALS AND METHODS

In this section, we are going to discuss the methods that used in this work which these methods are |Whale| |Optimization| |Algorithm|, |Artificial| |Neural| |Network|, and Backpropagation.

WHALE OPTIMIZATION ALGORITHM (WOA)

The whales are wonderful animals, world's biggest mammals. An adult whale is 180 meters long and weights up to 30 tones. A major mammal with seven different main species has: Minke, Sei, Right, Finback and Black, Humpback. Most of the time, whales were regarded as predatory. You never rest, because the ground of the oceans are breathing. Half the brain sleeps simply just. Ironically, the whales are known to be highly intelligent emotionally pets [11]. Buckling whales may identify and encircle the location of the prey. Sinine, or nearly ideal, the WOA algorithm considers the target probe to be the most effective solution, as at first sight the optimal location of the model in the search area is not understood. If the best search officer is set up, other search companies try to make the best search officer update their positions. The following equations explain this:

$$D^{\rightarrow} = |C.X^{\rightarrow}(T) - X^{\rightarrow}(T)|| \quad (1)$$

$$X^{\rightarrow}(T) = X^{*\rightarrow}|(T)| - A^{\rightarrow}.D^{\rightarrow} \quad (2)$$

T is present, vector-coefficient are A and C, X-a the best way to achieve the variable. Location X-a of the vector. Absolute value. Absolute value. Absolute value. Absolute value. It is important to note that X Tem must be modified in every iteration if a better solution is found. For vectors A and C, the following was established:

$$A = 2a \cdot r - a \quad (3)$$

$$C = 2 \cdot r \quad (4)$$

In both exploratory and development stages, a linear decrease of 2-0 is observed and r in [0,1] is a random variable. Illustration, the reasoning behind Eq is shown in 1(a). (2) a 2D problem. The quest agent's location (X, Y) can be changed to the current best record (Xjusqu'au Y jusqu'au). Different positions in the best agent can be reached by changing the value of A and C vectors in relation to the current position. Figure (1.1) indicates the possible change in 3D space of a search agent (b). It is important to note that it may be possible to achieve any location between the keypoints shown in the Figure (1) by specifying the random vector (r). So Eq. So. (2) allows every search officer to change their location in the current best solution neighborhood and to calculate the circumference of the prey. The concept can be implemented in a n-dimensional search room and search agents transfer the best possible solution. The bumps attack the prey with the |bubble| net technique, as defined in the preceding paragraph [12].

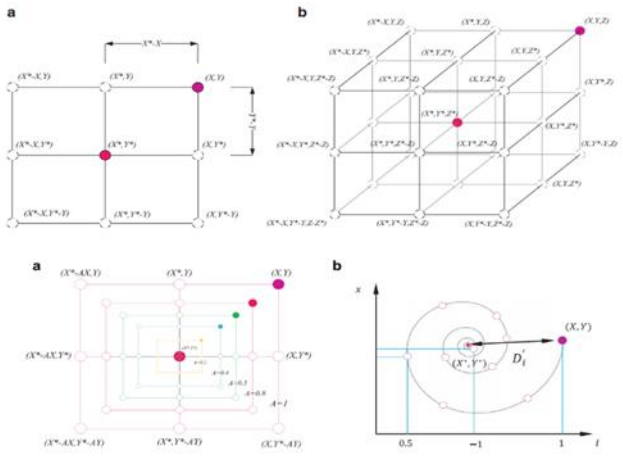

Figure 1: WOA-based bubble-net search mechanism.

The psedue code of this algorithm is:

*Corresponding author
Mohammed Al-Bayati,
Department of Information Technology, Altinbaş University, Istanbul, Turkey
e-mail: mohammed.mt2016@gmail.com

return X*

```
Population initialization Xi (i = 1, 2, ..., n)
Fitness Calculation
The best search agent is X*
As long as [t < maximum iterations number]
        for each search agent
                Updating a, A, C, l, and p
                If (1) (p<0.5)
                        if (2) (|A| < 1)
                                Updating the current search agent position
                        else if (2) (|A|
                                Random search agent selection ( )
                                Updating the current search agent position
                        end if (2)
                        else if (1) (p 0.5)
                        Updating the current search position
                end if (1)
        end for
        Checking any search agent in case it goes beyond the search space and amend it
        Calculating each search agent's fitness
        Updating X* in case there is a better solution
        t=t+1
end of iteration loop
```

ARTIFICIAL NEURAL NETWORK

One of the main machine learning methods is artificial neural networks. They are structures that stimulate the brain which repeat how the neural part of its name implies. Elements which turn the incoming layer into the neural network layer, and in most cases, layers of units occurring Input and output layers for Identifications patterns, there are appropriate tools, it's very complex or numerous for the human to learn or extract and recognize the machine [12].

RESULTS AND DISCUSSIONS

DATASET STRUCTURE AND DESCRIPTION

In this study, we will apply prediction with "ANN-BP" and ANN-WOA' on the dataset that we took from |UCI| |Machine| |Learning| |Repository| using exactly the Facebook metrics dataset, we execute the Normalization Method as a kind of Preprocessing Method on the Dataset. Facebook page manager created the last column. the senior marketing managers suggested this categorization because it is related to the promotions and advertising that cosmetic businesses execute the last column providing a manual categorization based on the campaign that linked with the post content. the categorization checked by another social media professional into the organization, all of 790 posts to reduce the likelihood of classification error because of error in written considered as a manual process. This Dataset contents are 500 instances, 15 features, and 4 result attributes (total interactions, likes, comments, and shares), which during this proposed study the target output is the total interaction. If the instance's target is less than 300 interactions the output is 0 else the output is 1.

IMPLEMENTATION AND RESULTS

The 'ANN'-BP and 'ANN'-WOA classification accuracy were collected at different times after the Hidden Layer 7 with 1000 iterations. In addition, (MES) mean square error in the hidden| layer with value is lesser than the other (MES) mean square error in the hidden layers with values as showing in comparing. The matrix contains the outcomes of the group real (Rows) and (Columns). 'TP' "(True Positive)" mention to the expected case, 'TN' "(True Negative)" mention to the normally expected case, 'FP' "(False Positive)" mention to the normal event predicted, and 'FN' "(False Negative)" mention to the real expected as a normal occurrence. Sensitivity shown in Equation. (5), and Specificity are shown in Equation. (6), The accuracy that's shown in Equation. (7), the general indicator of how much the classifier is right. Accuracy tests how much it's right when the classifier predicted an attack. Note, Specificity tests how much it is expected correctly when an campaign actually occurs.

*Corresponding author
Mohammed Al-Bayati,
Department of Information Technology, Altinbaş University, Istanbul, Turkey
e-mail: mohammed.mt2016@gmail.com

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+FN+TN} \quad (5)$$

$$\text{Sensitivity} = \frac{TP}{TP+FN} \quad (6)$$

$$\text{Specificity} = \frac{TN}{TN+FP} \quad (7)$$

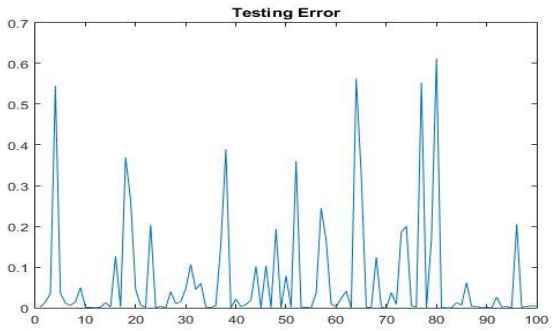

Figure 2. The gradient calculation of ANN-WOA

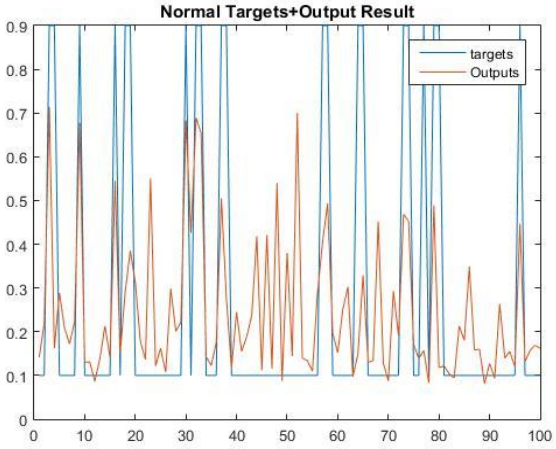

Figure 3. The general comparison between the neural network output and target from the dataset

Table 1. (WOA-ANN) Model Performance

(MSE) Mean Squared Error	0.0200458
Training Accuracy	97.99542
Testing Accuracy	99.875312
Specificity	98.5309
Sensitivity	100
True Positive	19
False Positive	2
True Negative	79
False Negative	0

CONCLUSION

Social media data are increasing day by day since the number of subscribers are increased. Hence, data mining technology is proposed for addressing the issues raised due to big data. Machine learning algorithms can perform the mining tasks efficiently as compare to traditional coding lines. Artificial neural networks are made for performing learning tasks using multi-layers perception. Data are made to pass through consecutive layers where particular process are being applied on the data in each layer resulting the data in good approximation. Artificial neural network is structured in similar form as human neural system. Data are

*Corresponding author
Mohammed Al-Bayati,
Department of Information Technology, Altinbaş University, Istanbul, Turkey
e-mail: mohammed.mt2016@gmail.com

travelling from one layer into other through the so-called weights. The quality of results in neural network is depending on the method used for setting the weight coefficients. Training process are made for setting up the model weight coefficients using various training schemes over this project. Optimization approaches such as whale optimization algorithm and gradient descent algorithm are used during the training stage of the proposed model. Feed forward neural network is proposed for classifying a Facebook data and classification performance was measured using variety of performance metrics. Mean square error, root mean square error, classification accuracy and confusion matrix were used for determining the feed forward neural network performance. Results revealed that whale optimization algorithm is outperformed in producing the optimum classification results over feed forward neural network classifier. An optimum accuracy of 99.875 % was realized from this model while other proposed optimization model (e.g. gradient descent algorithm) has produced classification accuracy of 86 %. The results were compared with other optimization approach that listed in the literature (e.g. support vector machine algorithm), the same was yielded a classification accuracy of 97.97 %.

REFERENCES

[1] Dilip, R.C., Mridula, C. and Samanta, R.K. (2011) An Artificial Neural Network Model for Neonatal Disease Diagnosis. International Journal of Artificial Intelligence and Expert Systems, 2, 96-106.
[2] Shao, Y.E., Hou, C.-D. and Chiu, C.-C. (2014) Hybrid Intelligent Modeling Schemes for Heart Disease Classification. Applied Soft Computing, 14, 47-52. http://dx.doi.org/10.1016/j.asoc.2013.09.020.
[3] Kahramanli, H. and Allahverdi, N.(2008) Design of a Hybrid System for the Diabetes and Heart Diseases. Expert Systems with Applications, 35, 82-89. http://dx.doi.org/10.1016/j.eswa.2007.06.004
[4] Kumari, M. and Godara, S. (2011) Comparative Study of Data Mining Classification Methods in Cardiovascular Disease Prediction 1. International Journal of Computer Science and Technology, 2, 304-308.
[5] Mazurowski, M.A., Habas, P.A., Zurada, J.M., Lo, J.Y., Baker, J.A. and Tourassi, G.D. (2008) Training Neural Network Classifiers for Medical Decision Making: The Effects of Imbalanced Datasets on Classification Performance. Neural Networks, 21, 427-436. http://dx.doi.org/10.1016/j.neunet.2007.12.031.
[6] Das, R., Turkoglu, I. and Sengur, A. (2009) Effective Diagnosis of Heart Disease through Neural Networks Ensembles. Expert Systems with Applications, 36, 7675-7680. http://dx.doi.org/10.1016/j.eswa.2008.09.013.
[7] Purwar, A. and Singh, S.K. (2015) Hybrid Prediction Model with Missing Value Imputation for Medical Data. Expert Systems with Applications, 42, 5621-5631. http://dx.doi.org/10.1016/j.eswa.2015.02.050.
[8] Bajaj, P., Choudhary, K. and Chauhan, R. (2015) Prediction of Occurrence of Heart Disease and Its Dependability on RCT Using Data Mining Techniques. Advances in Intelligent Systems and Computing, Springer India, 340, 851-858.
[9] Beheshti, Z., Shamsuddin, S.M., Beheshti, E. and Yuhaniz, S.S. (2014) Enhancement of Artificial Neural Network Learning Using Centripetal Accelerated Particle Swarm Optimization for Medical Diseases Diagnosis. Soft Computing, 18, 2253-2270. http://dx.doi.org/10.1007/s00500-013-1198-0.
[10] Seera, M. and Lim, C.P. (2014) A Hybrid Intelligent System for Medical Data Classification. Expert Systems with Applications, 41, 2239-2249. http://dx.doi.org/10.1016/j.eswa.2013.09.022.
[11] Woźniak, M., Graña, M. and Corchado, E. (2014) A Survey of Multiple Classifier Systems as Hybrid Systems. Information Fusion, 16, 3-17. http://dx.doi.org/10.1016/j.inffus.2013.04.006.
[12] Nishara Banu, M.A. and Gomathy, B. (2014) Disease Forecasting System Using Data Mining Methods. International Conference on Intelligent Computing Applications (ICICA), Coimbatore, 6-7 March 2014, 130-133

*Corresponding author
Mohammed Al-Bayati,
Department of Information Technology, Altinbaş University, Istanbul, Turkey
e-mail: mohammed.mt2016@gmail.com