Journal of Biomedical Engineering and Medical Imaging - Vol. 8, No. 6
Publication Date: December 25, 2021
DOI:10.14738/jbemi.86.11004.
Igodan, E. C., Ukaoha, K. C., & Oliomogbe, S. O. P. (2021). Towards Global Optimization of Neural Network: A Comparative Analysis Using Genetic and Whale Optimization Algorithms. Journal of Biomedical Engineering and Medical Imaging, 8(6). 89-101.

Towards Global Optimization of Neural Network: A Comparative Analysis Using Genetic and Whale Optimization Algorithms

Igodan, E. C.
Department of Computer Science, University of Benin, Benin City, Nigeria

Ukaoha, K. C.
Department of Computer Science, University of Benin, Benin City, Nigeria

Oliomogbe, S. O. P.
Department of Computer Science, University of Benin, Benin City, Nigeria

ABSTRACT

The intelligence and adaptability features of the neural network has made it a technique that is widely used to solve problems in diverse areas such as; detection, monitoring, prediction, diagnostics, data mining, classification, recognition, robotics, biomedicine, etc. However, determination of the optimal number of hidden layers of neural network and other parameters are still a difficult task. Usually, these parameters are decided by trial-and-error which increases the computational complexity and it is human dependent in obtaining the optimal model and parameters alike for any particular task. Optimization has received enormous attention in recent years, primarily because of the rapid progress in computer technology, including the development and availability of user-friendly software, high-speed and parallel processors, and artificial neural networks. This research work is to propose a neuro-evolutionary model using the computational intelligence techniques by combining ANN, GA and WOA for binary classification problems. The proposed optimized ANN-GA and WOA models is to circumvent the problem that is characterized in the trade-off between smoothness and accuracies in selecting the models and optimal parameters of neural network.

Keywords: Artificial Neural Networks; Genetic Algorithms; Optimization; Whale Optimization Algorithm

INTRODUCTION

Evolutionary Artificial Neural Networks (EANNs) are the most common techniques that are used in Machine Learning (ML) which involves the evolutionary algorithms and Artificial Neural Network (ANN) or both combined. While the ANN is used for learning and classification purpose, the evolutionary algorithms are used for optimization purposes. The ANN are based on the analogy of the human brain consisting of a tightly inter-connected group of nerve cells, or basic information processing units called neurons. The simplicity and naturalness with which the biological neural network can learn results in the simulation of artificial biological neural network on a computer (Negnevitsky, 2005). The most peculiar characteristics of ANN are intelligence and self-adaptability (Qin and Tang, 2009). These characteristics are widely applied in domains like fraud detection, health monitoring, stock price forecasting, diagnostics,

and so on (Azzini, 2006). Other areas are data mining, classification and regression, pattern matching and recognition, robotics, biomedicine, etc. The intelligence and adaptability of the neural network comes from the learning process of the network through adjusting the weight coefficients towards reducing the mean square error (or bias). The area of optimization has received enormous attention in recent years, primarily because of the rapid progress in computer technology, including the development and availability of user-friendly software, high-speed and parallel processors, and artificial neural networks. This is because optimization theory and methods deals with selecting the best alternative in the sense of the given objective function (Adhirai et al. 2018).

In principle, neural networks are universal approximators (Svozil et al., 1997), (Qin and Tang, 2009) having similar principles but different types and structures, which are categorized into feed forward, radial basis function, Kohonen self-organizing maps, convolution, recurrent networks and other variants. However, some of the non-trivial limitation of ANN have to do with the optimization of network architecture (or topologies) and model parameters selection. Usually, these parameters are decided by trial-and-error which increases the computational complexity and it is human dependent in obtaining the optimal model and parameters alike for any particular task. As the tasks differ so also are the model and parameters during selection of the best. According to Adhirai et al. (2018) to achieve desired results in a timely and economical fashion in any field, the designated tools must be used.

The need to reduce the dependency and complexity is an interesting and ongoing research in the literature. It is ongoing because there is no mathematical model to use in solving all problems and different problems (datasets) comes with its peculiarities with regards to the model structure and parameters. Also, in addressing this type of problem, optimization theory is central to any problem involving decision making regardless of the field being applied. Some of the optimization algorithms used in finding the best solution from among the set of feasible solutions are Genetic Algorithms (GA), Particle Swarm Optimization (PSO) algorithms, and the Whale Optimization Algorithm (WOA) and so on. Typically, meta-heuristics optimization algorithms have become more popular in different application areas like engineering, biomedical, business, finance, and so on. Nature-inspired meta-heuristic algorithms solve optimization problems by mimicking biological or physical phenomena. The meta-heuristic optimization algorithms can be classified into: the evolutionary, physical-based, swarm-based and human-based algorithms (Mirjalili and Lewis, 2016 and Adhirai et al. 2018). Whilst the GA is evolution-based algorithm proposed by Goldberg (Goldberg, 1989) and others, the WOA is a nature-inspired optimization algorithm proposed by Mirjalili and Lewis (Mirjalili and Lewis, 2016).

Genetic algorithms are stochastic search algorithms based on natural selection processes and human genetics (Goldberg, 1989). The search process starts with a randomly generated population which is evolved over subsequent generations. The strength point of these methods is that the best individuals are always combined together to form the next generation of individuals. This allows the population to be optimized over the course of generations. The most evolutional-inspired genetic algorithms simulate the Darwinian evolution. Other popular algorithms are Evolution Strategy (ES), Probability-Based Incremental Learning (PBIL), Genetic Programming (GP), and Biogeography-Based Optimizer (BBO) (Mirjalili and Lewis, 2016). These algorithms incorporate randomness as the main mechanism to move from a local search

to a global search and therefore, they are more suitable for global optimization (Yang 2014 and Aljarah et al. 2019). This study intends to combine the advantages of ANN and GA to address some of these problems as they are considered the most reliable and promising computational intelligence technique.

The WOA is a nature-inspired meta-heuristic optimization algorithm mimicking the social (hunting) behaviour of humpback whales. WOA has so much attention that it has been tested with 29 mathematical optimization problems and six structural design problems (Mirjalili and Lewis 2016, Adhirai *et al.,* 2018, Nasiri and Khiyabani 2018, Salgotra et al., 2019 and Hussien et al., 2019). To the best of our knowledge, there is no previous study on this use of WOA to optimize the Multi-Layer Perceptron (MLP) models and parameters in the optimization literatures aside the work done by Aljarah *et al.,* 2016) in optimizing the connection weights and biases in neural network. The most interesting thing about the humpback whales, a species of whales, is their special hunting method. This foraging behaviour is called bubble-net feeding method. Humpback whales create the typical bubbles along a circle path or '9'-shaped path while encircling prey during hunting of school of krill or small fishes close to the surface. Simply, bubble-net hunting behaviour could be describe such that humpback whales dive down approximation 12m and then create the bubble in a spiral shape around the prey and then swim upward the surface following the bubbles. The latter maneuver includes three different stages: coral loop, lob-tail and capture loop. However, the spiral bubble-net feeding maneuver is mathematically modelled in order to perform optimization (Mirjalili et al. 2018, Mirjalili and Lewis 2016 and Nasiri and Khiyabani 2018). The optimization results demonstrate that WOA is very competitive compared to the state-of-the-art optimization methods (Mirjalili and Lewis, 2016).

RELATED WORKS

The most widely used ANN is the MLP feed forward neural network, which is trained using back propagation training algorithm (Eng et al., 2008) as one of the most learning algorithms. The MLP comprised of multiple hidden layers. Typically, the network consists of an input layer, at least one hidden layer, and an output layer made up of computational neurons.

In this study, a multilayer perceptron is adopted and can be referred to as the layered feed forward neural network or multilayer feed forward neural network. In the MLP, each processing element makes its computation based upon a weighted sum of its inputs. A threshold function is sometimes used to qualify the output of a neuron in the output layer (Abhishek et al., 2012). The size of the network determines its computational complexity. To address this complexity, various mathematical models have been developed, however, Azzini (2006) and Thomas et al., (2015) posit that there is no mathematically-based theoretical method for predicting the number of hidden nodes to solve the problem.

Although a typical back propagation ANN with one hidden layer has been shown to approximate any function if appropriate free parameters are given (Zurada, 1992, Eng *et al.,* 2008, Jabin, 2014, Nienhold et al., 2015); however, network topology is still a major factor in the performance of the neural network (Erkaymaz *et al.,* 2012). The most challenging aspect is in the choosing of the optimum number of hidden nodes (Thomas *et al.,* 2015). In literature, to the most of our knowledge there appears to be limited attempts in applying GA to determining an ANN model parameter within a wide search space i.e. finding optimal ANN topologies,

optimization minimization algorithm from six different types, and its impact on the training time, performance and complexity (Abdalla *et al.,* 2014). While there is no evidence on the use of the WOA to the best of our knowledge, there is limited comparison on the two optimization algorithm to obtain the best for any particular problem.

Azzini (2006) and Aljarah et al. (2016) show that the architecture of ANN should be strengthened by considering it in tandem with evolutionary algorithms, since neural networks and nature-inspired algorithms techniques can be synergistically combined. Such different types of evolutionary techniques are becoming very useful in complex problems of optimization; and several literatures show that these algorithms represent more efficient ways of optimizing classification systems because they do not require expert knowledge of the problem. In literatures, the determination of the optimal number of hidden units and other parameters are still a difficult task carried out using the rule-of-thumb method which is a trial-and-error method that is time consuming (Ferentinos, 2005 and Panchal et al. 2011) and does not guarantee high classification performance. Batchis (2013) opined that if the ANNs are to be used to approximate complex functions the Back Propagation (BP) algorithm shows several potential weaknesses and so the use of nature-inspired algorithms:

a.) No opportunity for learning algorithm to find a better architecture for the problem since the architecture is constrained
b.) If there are too many nodes or layers, BP becomes impractical to implement, and
c.) BP converges to local maxima that is not global maximum

These approaches adopted in literature shows that the sizes of the networks in width and breath were either pre-determined manually (Ferentinos, 2005 Abdalla et al. 2014 and Sewsynker-Sukai et al. 2017) or obtained with the use of a trial-and-error method (Fisher and Leung 1998 and Sewsynker-Sukai et al. 2017) where there is poor performance due to the inability to solve the problems associated with the trade-off between smoothness and accuracies (Kokko, 2013). The learning process of an ANN, considered as one of the most difficult challenges in machine learning, has attracted a lot of works because of the disadvantages of the conventional training algorithms. These conventional algorithms are local optimal stagnation and slow convergence speed. Although, the MLPs can approximate any function efficiency than shallow ones, and can model complex functions using exponentially fewer neurons allowing them to obtain higher performance with same amount of training datasets (Geron 2019), the evolution of the architecture in existing literatures is all applied on fixed network architectures that are pre-determined. The main difficulty of training the MLP is the nonlinear nature and the unknown best set of main controlling parameters (Aljarah et al. 2016). There is no attempt in evolving two to three hidden layers with a wide search spaces, finding the minimization algorithm with six (6) different types and the impact on training time, performance and complexity (Abdalla et al. 2014). The genetic algorithm and WOA, if applied on a deep network dynamically can improve the selection of the optimal parameters, the minimization of the predictive error and generalization capabilities of the networks (Kokko 2013, Sewsynker-Sukai et al. 2017, Aljarah et al. 2016 and Geron 2019). This study proposed an optimized ANN-GA and WOA models to circumvent the problem that is characterized in the trade-off between smoothness and accuracies in selecting the models and optimal parameters which is demonstrated in building of inspiration from experience, knowledge and skills in the application of the two models. These challenges make the nature-inspired optimization algorithms reliable alternatives as there is no superior optimization algorithm in all

optimization problems according to the 'No-Free Lunch theorem (NFL) (Aljarah et al. 2016). Motivated by these reasons, the GA and WOA are proposed for training multi hidden layer neural nature. The aim of this research work is to proposed and develop a neuro-evolutionary model using the computational intelligence techniques by combining Artificial Neural Network (ANN), Genetic Algorithm (GA) and Whale Optimization algorithm (WOA) for binary classification problems. The objectives of this study are:

a.) To design an optimized MLP models using Genetic algorithm and Whale optimization algorithm
b.) To simulate the models using Python programming language, and
c.) To evaluate and compare both models using performance evaluate matrices

METHODOLOGY

The architecture of the model is shown in Fig 1. The architecture of the models developed shows the different components and the functionalities of each components are explained below:

Datasets (Breast, Cervical cancers and Hepatitis C Virus (HCV)

The datasets used in the research are obtained from the UCI Machine Learning Repository website: archive.ics.uci.edu/ml/index.php. The characteristics of the datasets are shown in Table 1.

Table 1: Dataset characteristics

Datasets	Ration of Training set/Testing set	Attributes	Missing Values	Total
Breast	75/25	10	Yes	699
Cervical	75/25	36	Yes	858
Hepatitis C Virus	75/25	29	N/A	1386

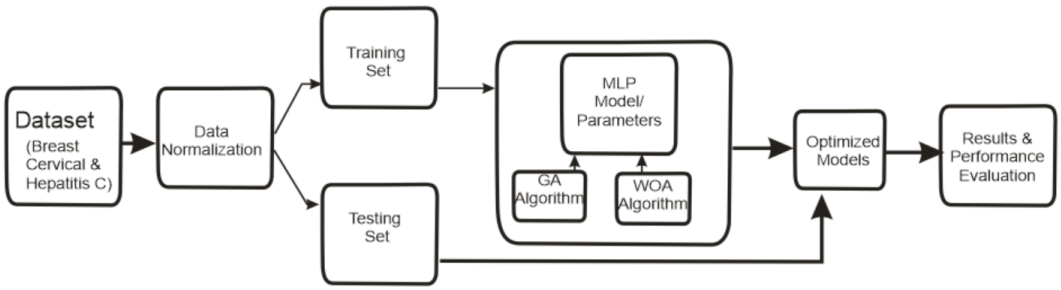

Figure 1: Architecture of the Proposed Model

Data Training and Testing

The normalized data was divided into training and testing sets using a 75% - 25% split ratio as shown in table 1. The training set was used to train the classifier while the testing set was used to evaluate the optimized models for accuracy.

The multilayer perceptron (MLP) is used as the classification algorithm and the Genetic algorithm and Whale Optimization algorithm was used to optimize the model parameters for optimal performance. The associated tasks in this research are binary classification, i.e. predicting the presence of cancer or non-cancer and presence of hepatitis or absence of it.
Given training set as $D = \{(x_i, y_i)\}_i^n = 1$, where D is the training dataset, x_i is an input for the ith instance described by set of attributes given as $\{a_1, a_2, a_i \dots, a_m\} y_i \in \{0,1\}$, an n represents the total number of data samples.

Data Normalization

The obtained data was normalized using the Z-score normalization in order to make training less sensitive to the scale of features and machine learning ready. Z-score converts data into [0,1] distribution using equation 1.

$$x_i' = \frac{x_i - \mu}{\sigma} \dots (1)$$

Where x_i' is the data value, x_i is the data value to be normalized, μ represents the mean of data values in the feature category.

Genetic Algorithm

Genetic Algorithms (Gas) (Goldberg, 1989) are global search stochastic algorithms based on natural selection and natural genetics (evolution) mechanics. GAs are techniques of gradient-free, parallel, efficient search and optimization based on natural selection and genetics laws (Palit and Popovic, 2005 and Deepa and Sivanandam, 2008). Gas are procedures for solving problems by using principles inspired by natural genetics. Starting with an initial population, the GA exploits the information contained in the present population and explores new individuals by generating offspring using genetic operators which can then replace members of the old generation. The commonly used genetic operators are reproduction, crossover, and mutation. After several generations, the algorithm converges on the best solution, which hopefully represents the optimum or near-optimal solution (Devaraj and Roselyn, 2007 and Devaraj and Yegnanarayana, 2000).

$$p(c_i) = \frac{f(c_i)}{p \sum_{i=1}^{N} c_i} \qquad \dots \qquad 2$$

The GA execution in this study follows the following steps (Bouktir *et al.,* 2004 and Goldberg, 1989):

1. Generate randomized initial population $X'(t) = (x_1, x_2, \dots, x_N)$ with N individuals at t=0
2. Compare fitness $f(x_j)$ of each individual x_j in the current population $X(t)$ using Equation 2
3. Check if termination condition is met
    a. If Yes, best individual is selected and stop the search process
    b. If No, a new population $X(t+1)$ is created with N new individuals
        i. Apply reproduction, crossover and mutation operators
        ii. Replacement by elitism

A more complete discussion of Gas can be found in books by Goldberg, (1989) and Rothlauf, 2006) and in articles by Delima et al., (2019)

Whale Optimization Algorithm

The WOA, stochastic optimization algorithm, utilizes a population of search agents to determine the global optimum for optimization problems. It is similar to other population-based algorithms in that the search process starts with creating random solutions for any given problem. The main differences in WOA algorithm compared with others are: the rules improve the candidate solutions, it shows flexibility, gradient-free mechanism, and high local optimal avoidance mechanism. The mathematical equation proposed in the WOA algorithm is as follows:

$$X(t+1) = \begin{cases} X^*(t) - AD & p < 0.5 \\ D' e^{bl} \cos(2\pi t) + X^*(t) & p \geq 0.5 \end{cases} \quad \ldots 3$$

Where p is a random number in [0,1],
$D^* = |X^*(t) - X(t)|$ and indicates the distance of the ith whale to the prey (best solution obtained so far), b is a constant for defining the shape of the logarithmic spiral, and l is a random number in $[-1,1]$, t shows the current iteration, $D = |CX^*(t) - X(t)|$, $A - 2ar - a$, $C = 2r$, a linearly decreasing number from 2 to 0 over the course of iterations (in both exploitation and exploration phases), and r is a random vector in [0,1]. In this research study the number of Whales/Agents used is 30 with a Maximum Cycles = 50.

The first component of this equation simulates the encircling mechanism, whereas the second mimics the bubble-net techniques. The variable p switches between these two components with an equal probability. The exploration and exploitation are the two main phases of optimization using population-based algorithms. They are both guaranteed in WOA (as well as GA algorithms) by adaptively tuning the parameters a *and* c in the main equation.

The WOA algorithm for optimizing MLP parameters: in this section, WOA is applied to optimize the MLP model and parameters with multiple hidden layers. Two major aspects are taken into consideration when the approach is designed: the representation of the search agents in the WOA and the selection of the fitness function. In the WOA-optimized MLP approach, each search agent is encoded as a one-dimensional vector to represent a candidate neural network and its parameters as indicated in table 1.1. That is each vector contains all items in table 1.1. The length of each vector is equal the sum of the individual length of item in table 1.1.

To increase the fitness value of the generated WOA agents, the Mean Square Error (MSE) is adopted. The MSE is the difference between the actual and predicted values by the generated agent (MLP) for all the training samples. The MSE is given in Equation 4 where y is the actual value, $\hat{y}$ is the predicted value, and n is the number of instances in the training dataset.

$$MSE = \frac{1}{2}\sum_{i=1}^{n}(y - \hat{y})^2 \qquad \ldots 4$$

The following steps demonstrate the workflow of the WOA-based approach applied in this study:

1. Initialization
2. Fitness evaluation
3. Update the position of the search agents
4. Repeat steps 2 to 3 until end of iterations.

Multilayer Perceptron

The neural network is built based on the structures of the benchmarks and its parameters obtained from the genetic algorithm and WOA techniques phase following these stages:

1. Initialization of parameters and random weights selection,
2. Feed forwarding using rectified unit and sigmoid functions
3. Error back propagation
4. Updates of the weights and bias
5. Continue steps 1 to 5 until
    a. A pre-defined, arbitrary error level (threshold) is reached, or
    b. The maximum number of calculations.

The MLP is used as the classifier which accepts input vector x_i multiplied by a weight vector w_i, and added to a bias b to produce an output y_i using the following equation 5:

$$y_i = f\left(\sum_{i=1}^{n} w_i x_i + b\right) \ldots (5)$$

Where n is the number of input-output pairs, and f is a non-linear activation function presented in Equation 6.

$f = \frac{1}{1+e^{-x_i}} \ldots (6)$

To determine the prediction error of MLP, the Mean Square Error (MSE) function is applied as shown in Equation 4.

$$E(\hat{y}, y) = \frac{1}{2}\sum_{i=1}^{n} (\hat{y} - y)^2 \ldots (7)$$

Performance Measure

The performance of the proposed classification approach is evaluated using different standard measures based on accuracy, F-score, sensitivity and specificity. Equations 7 to 10 shows the different metrics:

Accuracy: measures the proportion of the total number of predictions that are correct and given as: $\frac{TP+TN}{(TP+TN+FP+FN)} \ldots 7.$

F-Measure (or F-score): measures the weighted average of the true positive rate (recall) and precision. The F-score reaches its best value at 1 (perfect precision and recall), and worst at 0. It is a measure of a test's accuracy and given as: $2 * \frac{Precision*Recall}{Precision+Recall}$ …8.

Sensitivity (True Positive Rate or Recall): means positive classes correctly classified as positive classes and represented as: $\frac{TP}{(TP+FN)} \ldots 9.$

Specificity (True Negative Rate): measures the negative instances truly labeled as negative by classifier. Specificity should be high and given as: $\frac{TN}{(TN+FN)} \ldots 10.$

Where TN means True Negative, TP is True Positive, FN is False Negative, FP is the False Negative.

Our model was implemented using Python 3.7 Spyder Python editor via Anaconda distribution, Google Colab platform, excel spreadsheet, Intel™ i5-3437U CPU @ 1.90GHz 2.40GHz, 64-bit Operating system, x64-based process and HP-Elite-Book.

RESULTS AND DISCUSSION

The MLP models was trained and testing with and without the GA and WOA optimization algorithms and the results are shown below. In Table 2 MLP with GA and WOA algorithms are used to train Hepatitis C Virus dataset. The table indicates that the MLP without optimization algorithm has the lowest accuracy compared to MLP with GA which is higher in accuracy than MLP with WOA. Also MLP with GA shows higher values in precision, recall, F1 score and loss function than using MLP with WOA. MLP with GA has the lowest loss function than others, which indicates improvement in generalization error. In this study 20% - 50% dropout was used to prevent overfitting.

Table 2: Comparison of MLP with GA and WOA algorithms for Hepatitis C Virus

Algorithms	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)	Loss
MLP	94.80	96.53	92.98	94.67	16.48
MLP (GA)	99.35	100.00	95.82	97.81	0.29
MLP (WOA)	97.40	100.00	91.58	95.59	0.56

However, the time complexity of MLP is the highest to MLP with GA and MLP with WOA algorithm but the memory space consumed using MLP with WOA and MLP without optimization algorithm is same and higher than MLP with GA optimization algorithm.

Table 3: Time and Space Complexity for Hepatitis C Virus

Algorithms	Time	RAM Used
MLP	3600s	1.54 GB - 1.54 GB = 0.00 GB
MLP (GA)	1530.977s	1.34 GB - 1.33 GB = 0.01 GB
MLP (WOA)	137.443s	1.54 GB - 1.53 GB = 0.01 GB

Table 4 shows that the accuracy rate, precision, recall and F1 score for MLP with GA are lesser than using MLP or using MLP with WOA algorithm. The Loss function of MLP and GA is far lesser than MLP and MLP with WOA algorithm. However, the time complexity of MLP with WOA is far lesser than MLP and MLP with GA used however, MLP and MLP with WOA consumes more space than using MLP with GA as shown in Table 5.

Table 4: Comparison of MLP with GA and WOA algorithms for Cervical cancer

Algorithms	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)	Loss
MLP	99.06	100.00	98.48	99.22	9.04
MLP (GA)	100.00	100.00	100.00	100.00	0.01
MLP (WOA)	99.07	99.10	98.70	98.90	0.25

Table 5: Time and space complexity for Cervical cancer

Algorithms	Time	RAM Used
MLP	3500s	1.54 GB - 1.54 GB = 0.00 GB
MLP (GA)	1530.977s	1.34 GB - 1.33 GB = 0.01 GB
MLP (WOA)	137.443s	1.54 GB - 1.53 GB = 0.01 GB

Table 6 shows that the accuracy rate, precision, recall and F1 score for MLP with GA are lesser than using MLP and MLP with WOA algorithm. The Loss function of MLP with GA is far lesser than MLP and MLP with WOA algorithm. However, the time complexity of MLP and WOA is far lesser than MLP and MLP with GA which comes next, however, MLP consumes more space compared to other as shown in Table 7.

Table 6: Comparison of MLP with GA and WOA algorithms for Breast cancer

Algorithms	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)	Loss
MLP	91.60	92.50	92.50	92.50	20.28
MLP (GA)	98.60	98.73	98.12	98.42	0.74
MLP (WOA)	96.50	97.88	89.21	93.31	0.91

Table 7: Time and space complexity for Breast Cancer

Algorithms	Time	RAM Used
MLP	3567s	1.54 GB - 1.54 GB = 0.00 GB
MLP (GA)	1468.313s	1.53 GB - 1.52 GB = 0.01 GB
MLP (WOA)	183.514s	1.33 GB - 1.33 GB = 0.00 GB

From the table above, GA takes more space and time to converge while WOA do not. The analysis shows that there is trade trade-off between time and space complexity when comparing GA and WOA. From indication, GA can perform better in reducing the time complexity if the size of the population used and the number of each chromosome are optimize and also if the system configuration is improved upon. The various tables are visualized in the following evaluation metrics in figures 1 to 4 below. From the figures shown, it is indicative that the best generalization error (loss function) achieved is shown in the cervical cancer dataset, followed by hepatitis c virus and lastly breast cancer. Figure 4 indicates that the MLP without optimization takes a longer time to converge.

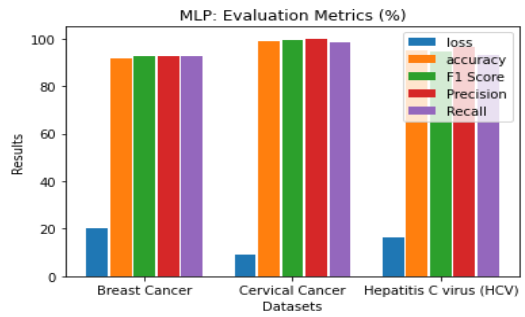

Figure 1: Evaluation for MLP

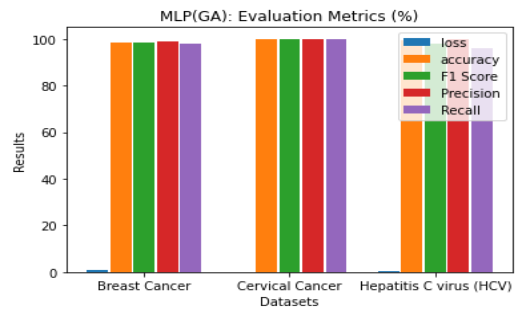

Figure 2: Evaluation for MLP (GA)

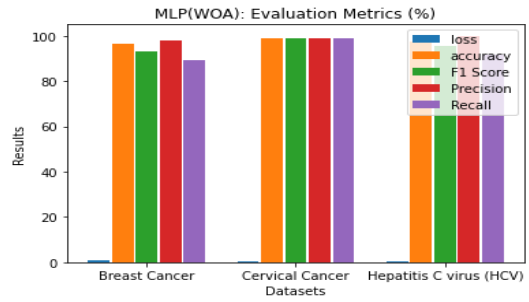

Figure 3: Evaluation for MLP (WOA)

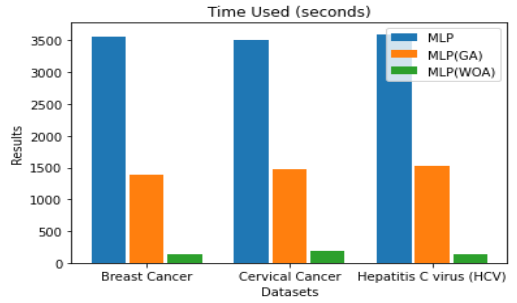

Figure 4: Time Comparison for MLP (with WOA and GA).

SIGNIFICANCE OF THE STUDY

This study seeks to create an optimized MLP model for classification problems specifically for the binary problems using GA and WOA. The optimized models will benefit researchers in the field of ML in addressing simple or complex binary classification problems using MLP with improved accuracies and better performance. However, the scope of the study is restricted to the use of neural network, whale optimization algorithm and genetic algorithm to evolve and select the network optimal parameters for classification problems. The model is applied for training, validation and prediction of breast, cervical and hepatitis c Virus datasets from the UCI machine learning repository website.

CONCLUSION

This study shows that neural network when combined with optimization algorithms like the genetic algorithm and whale optimization algorithm can perform better than when it is used without optimization algorithm. The proposed model reveals the effectiveness in the detection of a binary classification problem through the use of the various metrics. The MLP and GA show improvement than MLP and WOA in accuracy, specificity, sensitivity, F1 score however, the MLP combined with the WOA can also show improvement if the number of whale is further optimized either by reduction or increment, and also by selecting optimal feature subsets for learning. Even the MLP without can be improved upon is feature selection methods (Filter, Wrapper or Embedded) are adopted. The proposed model shows that there is an obvious trade-off in time and space used up.

In the future, we hope to further improve on this study by introducing feature selection techniques so as to reduce the number of features for the cervical cancer and hepatitis c virus.

This is believed that the space and the time consumed will further reduce and by extension improve on the classification accuracy. Also other optimization algorithm may be introduced and compare the best for use. Lastly, the need to also introduce the ensemble methods will be of advantage in the world of machine learning.

References

Yang X-S (ed.) (2014). Random Walks and Optimization. In: Nature inspired Optimization Algorithms, Chap 3. Elsevier, Oxford. Pp. 45-65.

Aljarah I., Faris, H. and Mirjalili, S. (2016). Optimizing Connection Weights in Neural Networks Using the Whale Optimization Algorithm. Soft Comput. Springer.

Nasiri, J. and Khiyabani, M. (2018). A Whale Optimization Algorithm (WOA) Approach for Clustering.Cogent Mathematics & Statistics. pp. 1-13.

Mirjalili, S. and Lewis, A. (2016). The Whale Optimization Algorithm. Advances in Engineering Software. 51-67.

Adhirai, S., Mahapatra, R.P. and Singh, P. (2018). The Whale Optimization Algorithm and Its Implementation in MATLAB. International Scholarly and Scientific Research & Innovation. Vol. 12, No: 10.

Devaraj, D., and Roselyn, J.P. (2007). Improved Genetic Algorithm for Voltage Security Constrained Optimal Power Flow Problem. Int. J. Energy Technology and Policy, Vol. 5, No. 4. Pp475-488.

Goldberg, D.E. (1989). Genetic Algorithm in Search, Optimization and Machine Learning, New York: Addison-Wesley.

Bouktir, T., Slimani, L. and Belkacemi, M. (2004). A Genetic Algorithm for Solving the Optimal Power Flow Problem. Leonardo Journal of Sciences. Issue 4, pp.44-58.

Devaraj, D. and Yagnanarayana, B. (2000). A Combined Genetic Algorithm Approach for Optimal Power Flow. Proc. Of 11th National Conference on Power Systems. pp524-528

Rothlauf, F. (2006). Representations of Genetic and Evolutionary Algorithms. Springer.

Delima, A.J.P., Sison, A.M. and Medina, R.P. (2019). A Modified Genetic Algorithm with a New Crossover Mating Scheme. Indonesian Journal of Electrical Engineering and Informatics (IJEEI). Vol. 7, No. 2, pp. 165-181.

Negnevitsky, M. (2005). Artificial Intelligence: A Guide to Intelligent Systems. Addison-Wesley, Harlow.

Azzini, A (2006), A New Genetic Approach for Neural Network Design and Optimization, Ph.D. Thesis.

Nienhold, D., Schwab, K., and Hanne, T. (2015). Effects of Weight Initialization in a Feedforward Neural Network for Classification Using a Modified Genetic Algorithm. 3rd International Symposium on Computational and Business Intelligence. Pp 6-12.

Abdalla, O.A., Osman, A., and Mohammed, Y. (2014). Optimizing the multilayer Feed-Forward Artificial Neural Networks Architecture and Training Parameters Using Genetic Algorithm. International Journal of Computer Applications. Vol.99, No.10, 42-48.

Ferentinos, K.P. (2005). Biological Engineering Applications of Feedforward Neural Networks Designed and Parameterized by Genetic Algorithms. Elsevier. Pp. 934-950.

Batchis, P. (2013). An Evolutionary Algorithm for Neural Network Learning Using Direct Encoding. Resource 53, Chinese Digital Library, Available Online: www.es.rutgets.edu/~mlittman/courses/ml03/iCML03/.../batchis.pdf. Accessed 25th August, 2018.

Sewsynker-Sukai, Y., Faloye, F., and Gueguim, E.B. (2017). Artificial Neural Networks: An Efficient Tool for Modeling and Optimization of Biofuel Production (a mini review).

Biotechnology and Biotechnological Equipment, 31:2, 221-235.

Kokko, T. (2013). Neural Networks for Computationally Expensive Problems. Master's Thesis. Negnevitsky, M. (2005). Artificial Intelligence: A Guide to Intelligent Systems. Addison-Wesley, Harlow.

Geron, A. (2019). Hands-On Machine Learning with Scikit-Leaming, Keras, and Tensor Flow: Concepts, Tools and Techniques to Build Intelligent Systems. 2nd Edition. O'reilly

Palit, A.K., and Popovic, D. (2005) Computational Intelligence in Time Series Forecasting: Theory and Engineering Applications. Springer.

Deepa, S.N and Sivanandam, S.N. (2008). Introduction to Genetic Algorithms. Springer. New York.

Zurada, J.M. (1992). Introduction to Artificial Neural Systems. West, United States.

Salgotra, R., Singh, U and Saha, S. (2019). On Some Improved Versions of Whale Optimization Algorithm. Arbian Journal for Science and Engineering. http://doi.org/10.1007/s13369-019-04016-0.

Hussien, A.G., Hassanien, A.E., Houssein, E.H., Hoissien, E.H., Amin, M and Azar, T.A. (2019). New Binary Whale Optimization Algorithm for Discrete Optimization Problems. Engineering Optimization. http://doi.org/10.1080/0305215x.2019.1624740.

Svozil, D., Kvasnicka, V. and Pospichal, J. (1997). Introduction to Multi-Layer Feed-forward Neural Networks. Chemometrics and Intelligent Laboratory Systems, 39, 43-62.

Eng, M.H., Li, Y and Wang, Q-G. (2008). Forecast Foresx With ANN Using Fundamental Data. International Conference on Information Management, Innovation Management and Industrial Engineering, 279-282.

Qin, H and Tang, S. (2009). A Solution to Dimensionality Curse of BP Network in Pattern Recognition Based on RS Theory. International Joint Conference on Computational Sciences and Optimization, 636-638.