ARPN Journal of Engineering and Applied Sciences

WCBP: A NEW WATER CYCLE BASED BACK PROPAGATION ALGORITHM FOR DATA CLASSIFICATION

Nazri Mohd. Nawi[1], Abdullah Khan[1], Naim Firdaus[1], M. Z. Rehman[1] and Insaf Ali Siming[2]
[1]Software and Multimedia Centre, Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia, Parit Raja, Batu Pahat, Johor, Malaysia
[2]Faculty of Science, Computer and Human Development, Universiti Tun Hussein Onn Malaysia, Parit Raja, Batu Pahat, Johor, Malaysia
E-Mail: nazri@uthm.edu.my

ABSTRACT

Water Cycle algorithm is a modern nature inspired meta-heuristic algorithm to provide derivative-free solution to optimize complex problems. The back-propagation neural network (BPNN) algorithm performs well on many complex data types but it possess the problem of network stagnancy and local minima. Therefore, this paper proposed the use of WC algorithm in combination with Back-Propagation neural network (BPNN) algorithm to solve the local minima problem in gradient descent trajectory. The performance of the proposed Water Cycle based Back-Propagation (WCBP) algorithm is compared with the conventional BPNN, ABC-BP and ABC-LM algorithms on selected benchmark classification problems from UCI Machine Learning Repository. The simulation results show that the BPNN training process is highly enhanced when combined with WC algorithm.

Keywords: back propagation neural network, water cycle search, local minima, meta-heuristic optimization, hybrid neural networks.

INTRODUCTION

The research on neural network is very popular nowadays and has been made considerable progress in recent years. Neural network has complex nonlinear mapping ability, function approximation and large-scale parallel distributed computing ability. Back-propagation neural network (BPNN) algorithm [1] is the most common example application of neural network because of its rigorous and high universality. The BPNN learns by calculating the errors of the output layer to find the errors in the hidden layers [2]. Because the BPNN learning algorithm uses gradient descent during training therefore it usually takes a long time to converge, network stagnancy and will encounter the local minimum problem [3, 4].

To overcome these limitations, many new hybrid meta-heuristic algorithms are introduced such as cuckoo search with back-Propagation [5, 6], bat based back propagation [7, 8], artificial bee colony with back-propagation [9, 10], and wolf search based back-propagation [11, 12]. Water cycle (WC) algorithm [13] is a novel meta-heuristic algorithm for optimizing constrained functions and engineering problems. The WC algorithm is inspired from the nature and based on the natural water cycle in the ecological system. So, for the sake of accuracy and to avoid local minima in BPNN, this paper proposed a combination of WC and BPNN algorithm. The resultant Water Cycle based Back Propagation (WCBP) algorithm is used to find the optimal weights for BPNN during initialization and update process. The proposed WCBP algorithm is trained on datasets taken from the UCI machine learning repository.

The remaining paper is organized as follows: Section 2 explains Water Cycle (WC) algorithm. Section 3 discusses the weight update process in the proposed WCBP algorithm and the simulation results are discussed in Section 4. And finally the paper is concluded in the Section 5.

THE PROPOSED WCBP ALGORITHM

The Water Cycle algorithm is inspired from nature and based on the observation of water cycle and how rivers and streams flow downhill towards the sea in the real world [13]. A river, or a stream, is formed whenever water moves downhill from one place to another. This means that most rivers are formed high up in the mountains, where snow or ancient glaciers melt. The rivers always flow downhill. On their downhill journey and eventually ending up to a sea, water is collected from rain and other streams. In the proposed WCBP algorithm, the first epoch, the best weights and biases are initialized with WC and then these weights are passed to the BPNN. The weights in BPNN are calculated and compared with the best solution in the backward direction. In the next cycle, WC will update the weights with the best possible solution and WC will continue searching the best weights until the last cycle/epoch of the network is reached or either the MSE is achieved. The pseudo code of the proposed WCBP algorithm is;

Step 1: *WC initializes and passes the best weights to BPNN*
Step 2: *Load the training data*
Step 3: **While** *MSE < Stopping Criteria*
Step 3(a): *WC finds the best weight and pass it on to the network, the weights,* w_{ij} *and biases,* b_i *in BPNN are then adjusted using the following formulae;*

$$w_{ij}(k+1) = (w_{ij}k + \mu\partial_j y_i)$$
$$\mathrm{b}_{\mathrm{i}}(k+1) = b_i k + \mu\partial_j$$

Step 3(b): *Feed forward neural network runs using the weights from WCA*
Step 3(c): *Calculate the backward error*
Step 3(d): *WCA keeps on calculating the best possible weight at each epoch until the Network is converged.*
End While

VOL. 11, NO. 24, DECEMBER 2016 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

RESULTS AND DISCUSSIONS

In this section, the performance of the proposed algorithm known as Water Cycle based Back Propagation (WCBP) algorithm is compared with Artificial Bee Colony Back-Propagation (ABC-BP), Artificial Bee Colony Levenberg-Marquardt (ABC-LM) and conventional BPNN algorithms. The datasets used to verify the accuracy of the proposed algorithm are 7-Bit Parity, Iris, Breast Cancer (Wisconsin), Diabetes and Australian Credit Card Approval. For the experimentation purposes, the Workstation used for carrying out the experimentation comes equipped with a 2.30GHz Core i5 Intel Processor, and 8-GB RAM while the operating system used is Windows 7 Ultimate. The software used to carry-out simulations is MATLAB 2012. During the comparison assessment, the networks parameters such as number of hidden layers, and learning rate are kept same. Three layer back-propagation neural networks is used for training the models, hidden layer is kept fixed to 10 nodes while output and input layers nodes vary according to the dataset given. A total of 20 trials are run to verify the algorithm. The network results are stored in the result file for each trial. Mean Squared Error (MSE), the number of failures are recorded in separate files for each independent trial on selected classification problems.

7-Bit parity dataset

The parity problem is one of the most popular initial testing task and a very demanding classification dataset for neural network to solve. In parity problem, if a give input vectors contains an odd number of one, the corresponding target value is 1, and otherwise the target value is 0. The N-bit parity training set consist of 2N training pairs, with each training pairs comprising of an N-length input vector and a single binary target value. The 2N input vector represents all possible combinations of the N binary numbers.

Table-1. CPU time, Epochs, MSE, Accuracy for 7-bit parity dataset.

Algorithm	BPNN	ABC-BP	ABC-LM	WCBP
CPU Time	22.07	183.3	134.88	925.82
Epochs	1000	1000	1000	1000
MSE	0.263	0.217	0.083	0.0093
Accuracy (%)	85.12	82.12	69.137	99.073

From Table-1, we can see that the proposed WCBP algorithm performs well on 7-Bit Parity dataset. The WCBP converges to global minima in 925.82 seconds of CPU time with an average accuracy of 99.073 percent and achieves a MSE of 0.0093. While, the other algorithm show less CPU time converge to global minima but the accuracy achieved is not very well. The Table-1 shows that the WCBP performs well and converges to global minima within 1000 epochs.

Iris classification dataset

IRIS flower data set classification problem is one of the novel multivariate dataset created by Sir Ronald Aylmer Fisher [14] in 1936. IRIS dataset consists of 150 samples from Iris setosa, Iris virginica and Iris versicolor. Length and width of sepal and petals is measured from each sample of three selected species of Iris flower. The feed forward network is set to 4-5-2.

Table-2. CPU time, Epochs, MSE, accuracy for IRIS dataset.

Algorithm	BPNN	ABC-BP	ABC-LM	WCBP
CPU Time	28.47	156.43	171.52	892.47
Epochs	1000	1000	1000	1000
MSE	0.312	0.155	0.058	0.0087
Accuracy (%)	87.19	86.88	79.559	99.13

From the Table-2, we can see that the proposed WCBP algorithm performs well on Iris Classification dataset. The WCBP converges to global minima in 892.47 seconds of CPU time with an average accuracy of 99.13 percent and achieves a MSE of 0.0087. While, the other algorithm show less CPU time converge to global minima but the accuracy achieved is not very well. The Table-2 shows that the WCBP performs well and converges to global minima within 1000 epochs.

Breast cancer classification dataset

Breast Cancer Dataset was taken from UCI Machine Learning Repository databases. The dataset was created on the information gathered by Dr. William H. Wolberg [15] during the Microscopic study of breast tissue samples selected for the diagnosis of breast cancer. This problem deals with the classification of breast cancer as benign or malignant. The selected feed forward neural network architecture used for this classification problem is 9- 5-2.

Table-3. CPU time, Epochs, MSE, accuracy for breast cancer dataset.

Algorithm	BPNN	ABC-BP	ABC-LM	WCBP
CPU Time	95.46	1482.9	1880.64	1161.35
Epochs	1000	1000	1000	1000
MSE	0.271	0.184	0.0139	0.0085
Accuracy (%)	90.71	92.02	93.831	99.16

From the Table-3, we can see that the proposed WCBP algorithm performs well on breast cancer dataset. The WCBP converges to global minima in 1161.35 seconds of CPU time with an average accuracy of 99.16 percent and achieves a MSE of 0.0085. While, the other algorithm show less CPU time except for ABC-BP AND ABC-LM which are 1482.9 and 1880.64 respectively, converge to global minima but the accuracy achieved is not very well.

Pima Indian diabetes dataset

Pima Indians Diabetes dataset consists of 768 instances, with 8 inputs and 2 outputs. This dataset contains all the information of the chemical changes in a female body whose imbalance can cause diabetes [16]. The feed-forward network architecture for this classification problem is set to 8-5-2.

Table-4. CPU time, Epochs, MSE, accuracy for diabetes dataset.

Algorithm	BPNN	ABC-BP	ABC-LM	WCBP
CPU Time	57.05	4257	2805.09	1388.90
Epochs	1000	1000	1000	1000
MSE	0.269	0.201	0.141	0.0081
Accuracy (%)	84.94	91.468	65.098	99.19

From the Table-4, we can see that the proposed WCBP algorithm performs well on Diabetes dataset. The WCBP converges to global minima in 1388.90 seconds of CPU time with an average accuracy of 99.19 percent and achieves a MSE of 0.0081. While, the other algorithm show less CPU time except for ABC-BP and ABC-LM which are 4257 and 2805.09 respectively, to converge to global minima but the accuracy achieved is not very well.

CONCLUSIONS

BPNN is one of the most widely used and a popular algorithm to train Artificial Neural Networks. Conventional BPNN algorithm has some drawbacks, such as getting stuck in local minima and slow speed of convergence. Nature inspired meta-heuristic algorithms provide derivative-free solution to optimize complex problems. In this paper, a new meta-heuristic search algorithm, called Water Cycle algorithm is proposed to get optimal weights for training BPNN in order to achieve fast convergence rate and accuracy. The performance of the proposed WCBP algorithm is compared with the conventional BPNN, ABC-BP and ABC-LM algorithms. The performance of the proposed WCBP is verified by means of simulations on 7-bit parity, Iris, Diabetes, and Breast Cancer dataset. The simulation results show that the BPNN algorithm effectively avoids the shallow local minima and converge to global minima when coupled with WC algorithm.

ACKNOWLEDGEMENTS

The Authors would like to thank Office of Research, Innovation, Commercialization and Consultancy (ORICC), Universiti Tun Hussein Onn Malaysia (UTHM) and Ministry of Higher Education (MOHE) Malaysia for financially supporting this Research under Fundamental Research Grant Scheme (FRGS) vote no. 1236. This research is also supported by Gates IT Solution Sdn. Bhd under its publication scheme.

REFERENCES

[1] Rumelhart D. E., Hinton G. E., Williams R. J. 1986. Learning Internal Representations by error Propagation, J. Parallel Distributed Processing: Explorations in the Microstructure of Cognition. 1.

[2] Hecht-Nielsen R. 1989. Theory of the Back propagation Neural Network, Department of Electrical and Computer Engineering. University of California, San Diego.

[3] Rehman M. Z., Nawi N. M. 2012. Improving the Accuracy of Gradient Descent Back Propagation Algorithm (GDAM) on Classification Problems, Int. J. of New Computer Architectures and their Applications (IJNCAA). 1(4). pp. 838-847.

[4] Rehman M. Z., Nawi N. M. 2012. Studying the Effect of adaptive momentum in improving the accuracy of gradient descent back propagation algorithm on classification problems. Int. Journal of Modern Physics. 9. pp. 432-439.

[5] YangX. S., and Deb S. 2010. Engineering Optimisation by Cuckoo Search, International Journal of Mathematical Modelling and Numerical Optimisation. 1(4). pp.330-343.

[6] Nawi N. M., Khan A., and Rehman M.Z. 2013. A New Back-propagation Neural Network optimized with Cuckoo Search Algorithm. ICCSA-2013. pp. 413-426.

[7] YangX. S. 2010. A new metaheuristic bat-inspired algorithm, Nature Inspired Cooperative Strategies for Optimization2010. pp. 65-74.

[8] NawiN. M., RehmanM. Z. and Khan A. 2014. A New Bat based Back-Propagation (Bat-BP) algorithm, Advances in Systems Science: Advances in Intelligent systems and Computing. Springer International Publishing.

[9] Karaboga D., Artificial Bee Colony algorithm. 2005. Journal of Scholarpedia. 5(3).

[10] Karaboga D., Basturk B., and Ozturk C. 2007. Artificial bee colony (ABC) optimization algorithm for training feed-forward neural networks, LNCS: Modeling Decisions for Artificial Intelligence. 4617. pp.318-319.

[11] Fong S., Yang X. S. and Deb S. 2012. Wolf Search Algorithm with Ephemeral Memory, Seventh International Conference on Digital Information Management (ICDIM), Macau.

[12] Nawi N. M., Rehman M. Z. and Khan A. 2013. WS-BP: An Efficient Wolf Search Based Back-Propagation Algorithm, International Conference on Mathematics, Engineering, and Industrial Applications (ICOMEIA-2014).

[13] Eskandar H., Sadollah A., Bahreininejad A., and Hamdi M. 2012. Water cycle algorithm - A novel metaheuristic optimization method for solving constrained engineering optimization problems. Computers & Structures. pp. 151-166.

[14] Fisher R. A. 1936. The use of multiple measurements in taxonomic problems, Annual Eugenics. 7. pp.179-188.

[15] Wolberg W. H. and Mangasarian O. L. 1990.Multisurface method of pattern separation for medical diagnosis applied to breast cytology, National Academy of Sciences. 87. pp. 9193-9196.

[16] Smith J. W., Everhart J. E., Dickson W. C., Knowler W. C. and Johannes R. S. 1988. Using the ADAP learning algorithm to forecast the onset of diabetes mellitus, The Symposium on Computer Applications and Medical Care, IEEE Computer Society Press. pp. 261-265.