AN INTELLIGENT SYSTEM BASED ON BACK-PROPAGATION NEURAL NETWORK AND PARTICLE SWARM OPTIMIZATION BASED NEURAL NETWORK FOR DIAGNOSING ANEMIA IN PREGNANT LADIES

By

NEHA SHARMA * VIKAS KHULLAR ** ASHISH KUMAR LUHACH ***

** PG Scholar, Department of Computer Science & Engineering, CT Institute of Engineering Management and Technology, Jalandhar, India.*
*** Assistant Professor, Department of Computer Science & Engineering, CT Institute of Engineering Management and Technology, Jalandhar, India.*
**** Vice Principal & Associate Professor, Department of Computer Science, and Engineering, CT Institute of Engineering Management and Technology, Jalandhar, India.*

Date Received: 17/08/2017 Date Revised: 22/08/2017 Date Accepted: 23/08/2017

ABSTRACT

According to WHO According to World Health Organization (WHO) survey, anemia is one of the most commonly encountered medical deficiencies during pregnancy [5]. BP network has been successfully used for anemia diagnosing in pregnant ladies, however BP network's drawbacks, such as long execution time and its easy fall into local optima have restricted its wider applications. Recently proposed stochastic optimization method Particle Swarm Optimization (PSO) is also been discussed. Also the way BP network's initial weights and bias are optimized; Particle swarm optimization is also carefully discussed. In this paper, firstly BP is used to initially train and test the BP network, then the Particle Swarm Optimization Based Back-Propagation (PSO-BP) networks is used to train and diagnose the anemia in pregnant ladies. While concluding the experimental results, it shows variation in the taken parameters, execution time, and accuracy [10, 11], [16].

Keywords: Anemia, Back Propagation, Particle Swarm Optimization, Artificial Neural Network, Pregnant Ladies, Particle Swarm Optimization based Neural Network.

INTRODUCTION

Anemia is one of the most widespread public health problems, especially in developing countries. It impairs cognitive development, reduces physical work capacity and in severe cases cause increased risk of mortality particularly during perinatal period [11]. During pregnancy, approximately 75% of all anemia diagnosed are due to iron deficiency. Furthermore, World Health Organization (WHO) considered that ladies in developing countries may be pregnant or as much as one half of their reproductive lives and therefore are an increased risk of anemia during this time [14, 15].

Globally, anemia has been found to be the most common complication in pregnancy, the World Health Organization (WHO) estimates that more than 50% of the pregnant ladies in developing countries have been affected [7]. About 50% of pregnant ladies suffer anemia, and most of the cases are iron, vitamin B or folic acid deficient [16].

ANN based models for medical parameters such as mortality and of anemia deficiency in pregnant ladies have been studied. The research applies ANN and their associated analysis techniques to anemia deficient pregnant ladies. The ANN has been trained by disclosing it to the set of existing data where the desired, i.e. output following history of pregnant ladies in known ANN have varying complexity and types which has been used in research diagnosis of anemia deficiency. ANN has been shown to be more accurate in diagnosing anemia and obtaining the results from neural network [1].

In this paper, BP and PSO algorithms were applied to train the neural network [3, 4]. These algorithms were established using Anaconda Python simulator. The experiment shows that PSO training model has more ideal results as compared to the BP training neural network model [13].

1. Related Work

Kononenko (2001) [9] explains that artificial machine learning algorithms were very useful to solve medical diagnostic tasks. Qeethara Kadhim Al-Shayea (2011) had investigated the Artificial Neural Network (ANN) accuracy for diagnosing the disease. The author proposed that Feed Forward back propagation neural network was used to distinguish between infected and non-infected patients [1]. Perceptions and its computational ability in single layer feed forward network was discussed by Reosenblatt (1962). Further Rumelhart and William popularized the generalized delta rule for learning by back-propagation, which is the most widely used training algorithm for multi-layer network. Later Carvalho and Ludermir [2] had used PSO and BP neural networks for training. In conclusion, the authors show that the PSO algorithm had enhanced the generalization ability of neural network. Hippert, et al., [6] had also used particle swarm optimization algorithm with inertia weight to train the neural network, the experiment showed that the particle swarm optimization algorithm is the fastest and the easiest way to train the neural network.

2. Algorithms Used

2.1 Back-Propagation

Back Propagation, an abbreviation for "backward propagation of errors", is a common method of training Artificial Neural Networks. From a desired output, the network leans from many inputs, similar to the way a child learns to identify a dog from examples of dogs. Back propagation is a neural network learning algorithm [10]. The feedback forward network structure is the most important ANN structure. Design of a feed forward net for any specific application involves many issues, most of which require problem dependent solutions. Overall approach comprises of two parts: feed forward implementation of learned mapping and training of 3 layer network, where three layers comprises of Input layer, Hidden layer, and Output layer represented as in Figure 1 [7].

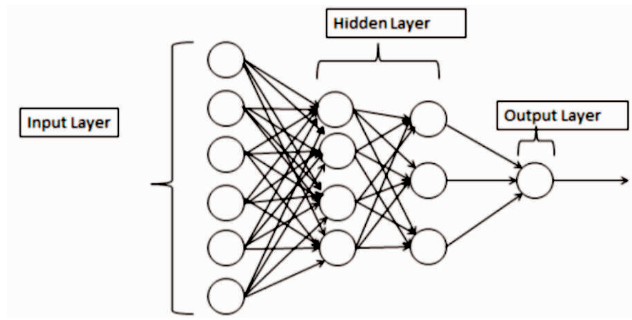

Figure 1. Back-propagation Workflow

2.2 Particle Swarm Optimization (PSO)

Like other evolutionary algorithms, PSO uses a fitness function to search for the best position. Each particle is initialized with a random position and velocity. In every simulation run, the fitness function is evaluated by taking the current position of the particle in the solution space. The particles keep track of two best values. The first one is the best fitness value obtained so far by the particle, the corresponding position being termed as personal best (p_{best}). The other is the best fitness value achieved so far considering all the particles in the swarm. The location of the best fitness value in a whole swarm is called global best (g_{best}). At each run, there is only one g_{best}, and all the particles are attracted toward g_{best} [8, 12].

Each individual (called particle) is characterized by its position $\vec{x_i}$, its velocity $\vec{v_i}$, its personal best position $\vec{p_i}$ and its neighborhood best position $\vec{p_g}$.

The elements of the velocity vector for particle i are updated as;

$$v_{ij} \leftarrow \omega v_{ij} + c_1 q\left(x_{ij}^{pb} - x_{ij}\right) + c_2 \gamma \left(x_j^{sb} - x_{ij}\right) \quad (1)$$

where j = 1,..,n, ω is the inertia weight, x_i^{pb} is the best variable vector encountered so far by particle I, and x^{sb} is the swarm best vector, i.e. the best variable vector found by any particle in the swarm, so far c_1 and c_2 are constants, and q and r are random numbers in the range [0, 1]. Once the velocities have been updated, the variable vector of particle is modified according to;

$$x_{ij} \leftarrow x_{ij} + v_{ij} \quad (2)$$

The cycle of evaluation followed by updates of velocities and positions (and possible update of x_i^{pb} and x^{sb}) is then repeated till a satisfactory solution has been found [5].

3. Overall Methodology

In this work, after preprocessing, the data is trained with two different training algorithms BP and PSO and were compared, in order to find the overall best model for diagnosis. Figure 2 shows the overall methodology of the work performed. All neural networks have been simulated using Python.

3.1 Dataset Used

Live Data records of pregnant ladies were collected from various hospitals. Total of 180 records of pregnant ladies were collected between 1 March 2017 and 30 April 2017. The dataset is classified into two different classes either positive or negative for anemia in pregnant ladies respectively. Results of the final diagnosis after the confirmation with the specialists had been used for study. Consisted records have 8 attributes each while 4 of them are actually used as the major attribute.

Total no. of observations = 1180

Total no. of anemia positive observations = 593

Total no. of anemia negative observations = 588

3.1.1 Description of the Database Entities

Figure 3 represents the dataset samples of few pregnant ladies with the medical parameters included during the study and the attributes used for the research [7] are listed in Table 1.

3.2 Network Model

PSO as a heuristic optimization method is successfully applied to train the model. It is proposed to update network weights by reasons of easy implementation and realization, the small number of parameters to be set, and capable of treatment with real numbers, no derivative information. According to this study, in order to improve the ability of Back-propagation neural network to escape from a local optimum, the PSO algorithm was used to modify the Network accuracy and execution time. Figures 4 and 5 show the workflow of Back propogation and

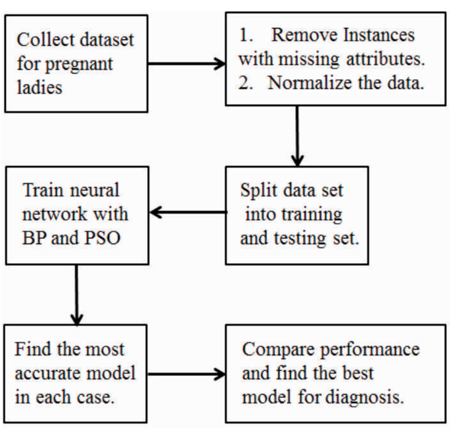

Figure 2. Overall Methodology

	A	B	C	D	E	F	G	H
1	Gender	RBC (4.5-6)	HGB(12-16)g/dl	HCT(36-48)	MCV(80-100)fb/cell	MCH(27-34)	MCHC(31-37)	Output
2	F	4.65	12	36.7	97.4	32.7	35.6	0
3	F	4.2	12	36.8	97.8	33.2	35.7	1
4	F	5.6	12	36.8	98.5	33.6	36.2	0
5	F	5.32	12	38	98.9	34	36.6	1
6	F	5.7	12	39.7	99	34.2	37.4	1
7	F	5.33	12.5	39.6	97.6	28	34.2	0
8	F	5.83	12.8	39.9	97.9	29.5	35.6	1
9	F	5.35	12.9	39.9	98.3	29.8	35.2	0
10	F	5.63	12.9	40.3	98.8	30.8	37.4	0
11	F	4.85	12.9	40.6	98.9	33.8	36.2	0
12	F	5.73	13	39.6	97.7	32.5	35.6	0
13	F	5.3	13	39.9	98.6	32.8	35.2	0
14	F	5	13	40.9	99.2	30.8	35.9	0
15	F	5.93	13	45.9	99.4	32.7	35.7	0
16	F	5.48	13.2	46.4	97.5	31	35.2	0
17	F	4.38	13.2	46.8	98.4	32	34.2	0
18	F	5.3	13.2	47.9	98.6	32	34.7	1
19	F	5.73	13.2	48.6	99.2	33.9	36.2	1
20	F	5.3	13.2	48.9	99.5	31.7	34.5	0
21	F	5.43	13.3	47.5	96.4	32.8	36.2	1
22	F	5.4	13.3	47.9	97.8	32.8	35.6	1
23	F	5.29	13.3	48.3	98.4	33.8	36.4	1
24	F	5.38	13.3	48.6	99.4	33.5	36.2	1

Figure 3. Example of Dataset Representing for few Patients

Attributes	Range
Gender	Female
Hemoglobin (HGB)	(12-16)g/dl
Red Blood Cell (RBC)	4.5-6/ul
Hematocrit (HCT)	(32-36)%
Mean Corpuscular Volume (MCV)	(80-100)fb/cell
Output	Anemia=0 Non-anemia=1

Table 1. Attributes used for Research

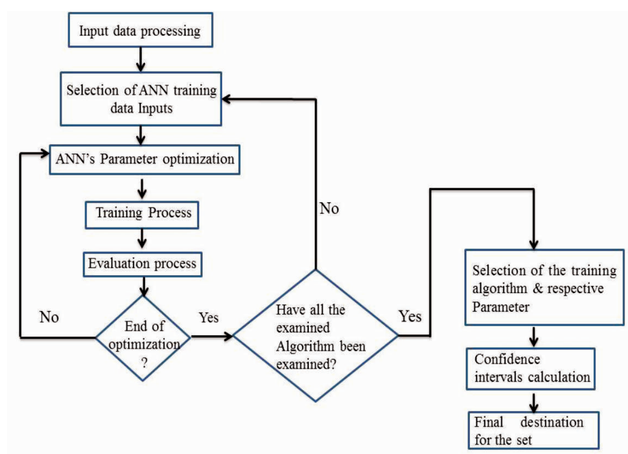

Figure 4. Back-propagation Workflow

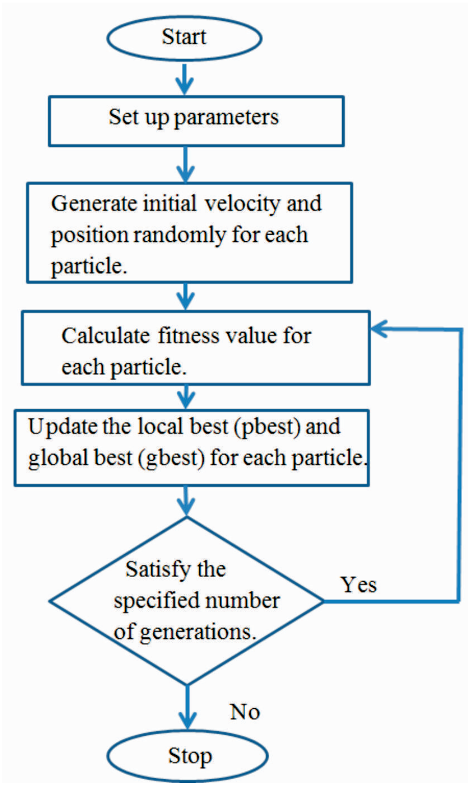

Figure 5. PSO Workflow

particle swarm optimization Algorithm, respectively.

4. Performance Evaluation

The result reflects the research work in two phases. The first phase is obtained from the back-propagation training and testing of the network while the second phase is obtained from the PSO optimized BP algorithms for diagnosing anemia in pregnant ladies. Parameters used are accuracy and execution time.

4.1 Accuracy during Different Data Ratio

According to Table 2 it has been analyzed that during BP, the accuracy increases when data has been divided in the ratio of 80 – 20 as training and testing, respectively while during learning with PSO, the data shows highest

Training/Testing	BP (Accuracy in % age)	PSO (Accuracy in % age)
90/10	65	82
80/20	72	80
70/30	69	89
60/40	53	81
50/50	65	69

Table 2. Accuracy during Different Data Ratio

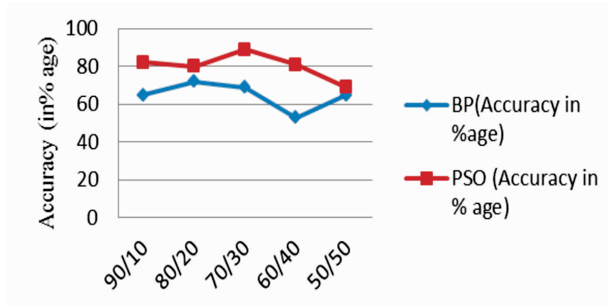

Figure 6. Accuracy during Different Data Ratio

Training/Testing	Execution Time (in seconds)	Execution Time (in seconds)
90/10	3.3806	2.2290
80/20	3.4399	2.2936
70/30	3.4796	2.3505
60/40	3.5038	2.4314
50/50	3.5153	2.4891

Table 3. Execution Time for BP and PSO during Different Data Ratio

accuracy of 70-30. Figure 6 represents the graphical view for the same.

4.2 Execution Time

Execution time is the total real time taken by the algorithm to train, test, and return the output. It can be measured in seconds, minutes, etc., according to the requirement. Back-propagation takes more time to reach the neighborhood of an optimal solution, but then reaches it more precisely. On the other hand, PSO-BP algorithms investigate the entire search space. Hence, they reach faster than the region of optimal solutions, but have difficulties to localize the exact point.

Table 3 represents the analyzed execution time for both the algorithms during different data ratios. Execution time increases as the data splitting ratio increases.

Conclusion

Based on the training and testing instances of BP and PSO-BP networks for the anemia diagnosis in pregnant ladies, some conclusions are abstracted:

- PSO as a stochastic optimization approach trains network using global particle and the accuracy for testing is high as compared to local optimization approach, i.e. Back-propagation.
- PSO takes less time to train and test the model, i.e. it

takes less execution time than BP.

Therefore overall it can be concluded that as per anemia diagnosis in pregnant ladies with back-propagation and particle swarm optimization based neural network on the basis of comparative parameters, such as accuracy and execution time during different data splitting ratio PSO-BP performs far better than network trained with BP.

References

[1]. **Al-Shayea, Q. K. (2011).** Artificial neural networks in medical diagnosis. *International Journal of Computer Science Issues*, 8(2), 150-154.

[2]. **Carvalho, M., & Ludermir, T. B. (2006, October).** An analysis of PSO hybrid algorithms for feed-forward neural networks training. In *Neural Networks, 2006. SBRN'06. Ninth Brazilian Symposium on* (pp. 6-11). IEEE.

[3]. **Dang, T. L., Cao, T., & Hoshino, Y. (2016).** Data Pre-processing for a Neural Network Trained by an Improved Particle Swarm Optimization Algorithm. *International Journal of Computer Applications*, 154(1), 1-8.

[4]. **Gandomi, A. H., Yang, X. S., & Alavi, A. H. (2013).** Cuckoo search algorithm: A metaheuristic approach to solve structural optimization problems. *Engineering with Computers*, 29(1), 17-35.

[5]. **Gandomi, A. H., Yang, X. S., Talatahari, S., & Deb, S. (2012).** Coupled eagle strategy and differential evolution for unconstrained and constrained global optimization. *Computers & Mathematics with Applications*, 63(1), 191-200.

[6]. **Hippert, H. S., Pedreira, C. E., & Souza, R. C. (2001).** Neural networks for short-term load forecasting: A review and evaluation. *IEEE Transactions on Power Systems,* 16(1), 44-55.

[7]. **Kaya, E., Aktan, M. E., Koru, A. T., & Akdoğan, E. (2015).** Diagnosis of anemia in children via artificial neural network. *International Journal of Intelligent Systems and Applications in Engineering*, 3(1), 24-27.

[8]. **Kenndy, J., & Eberhart, R. C. (1995, December).** Particle swarm optimization. In Proceedings of *IEEE International Conference on Neural Networks* (Vol. 4, pp. 1942-1948). IEEE Press.

[9]. **Kononenko, I. (2001).** Machine learning for medical diagnosis: History, state of the art and perspective. *Artificial Intelligence in Medicine*, 23(1), 89-109.

[10]. **Mohaghegi, S., Del Valle, Y., Venayagamoorthy, G. K., & Harley, R. G. (2005, June).** A comparison of PSO and backpropagation for training RBF neural networks for identification of a power system with STATCOM. In *Swarm Intelligence Symposium, 2005. SIS 2005. Proceedings 2005 IEEE* (pp. 381-384). IEEE.

[11]. **Sharma, N., & Khullar, V. (2016).** Comparative Review of Artificial Neural Network Machine Learning for diagnosing Aneamia in Pregnant Ladies. *i-manager's Journal on Information Technology*, 5(4), 33-38.

[12]. **Shi, Y. (2001).** Particle swarm optimization: developments, applications and resources. In *evolutionary computation, 2001. Proceedings of the 2001 Congress on* (Vol. 1, pp. 81-86). IEEE.

[13]. **Suresh, A., Harish, K. V., & Radhika, N. (2015).** Particle swarm optimization over back propagation neural network for length of stay prediction. *Procedia Computer Science*, 46, 268-275.

[14]. The LINKAGES Project, Maternal Nutrition during Pregnancy and Lactation, United State Agency for International Development (USAID), Academy for Educational Development [AED], The CORE Group, 2004.

[15]. **United Nations. World Health Organization. (2001).** *Iron Deficiency Anaemia: Assesment, Prevention and Control: a Guide for Programme Managers*. World Health Organization.

[16]. **Yılmaz, Z., & Bozkurt, M. R. (2012).** Determination of women iron deficiency anemia using neural networks. *Journal of Medical Systems*, 36(5), 2941-2945.

ABOUT THE AUTHORS

Neha Sharma is currently pursuing her Masters in Technology (M.Tech) in the Department of Computer Science and Engineering at CT Group of Institutions, Jalandhar, Punjab, India. Her research area is Artificial Neural Network and Applications.

Vikas Khullar is currently working as an Assistant Professor in the Department of Computer Science and Engineering at CT Group of Institutions, Jalandhar and also a Research Scholar at I.K.G. Punjab Technical University, Jalandhar. His research area is Interventions for Mental Disorders and Big Data Machine Learning.

Dr. Ashish Kumar Luhach is Vice Principal & Associate Professor in the Department of Computer Science and Engineering at CT Group of Institutions, Jalandhar. His research area is E-commerce and Security Frameworks for E-commerce.