FFANN based Cost Effective Major Infant Disease Management

A.M.Agarkar
Department of Electronics and Telecom. Engineering,
S.S.G.M. College of Engineering Shegaon, 444203, India

Dr. A.A.Ghatol
Ex-Vice Chancellor, Dr. BATU, Lonere
Ex-Director, PIET, Pune, India

ABSTRACT
In India, 30% to 40 % babies are low birth weight babies (LBW) as opposed to about 5% to 7% of newborn in the west. In India, 7 to 10 million LBW infants are born annually. About 10 % to 12% of Indian babies are born preterm (less than 37 completed weeks) as compared with 5% to 7% incidence in the west. These infants are physically immature and therefore their neonatal mortality is high. It is possible to increase the survival of the infants and quality of human life through prompt and adequate disease management of the newborn.

The proposed model of cost effective major infant disease management system based on *artificial intelligence algorithm* is helpful for diagnostic- cum- preventive approach to reduce the immaturity, fragility, vulnerability and dependence of the neonates in the developing countries to reduce neonatal and child mortality. Secondly, a significant proportion of the pediatricians' time, especially in major hospitals, large cities and overpopulated areas is spent on examination and evaluation of apparently healthy babies and detection of minor developmental defects.

In addition to these facts, many third world countries including countries like Pakistan and Bangladesh face the major problem of child health diagnosis and malnutrition. Medical facilities and expertise is either absent or out of reach of these tribal and poor communities, many public health centers (PHCs) lack in advice by experts on immediate basis in case of emergencies. Major hurdles include the lack of medical experts and trained manpower, scarcity of funds and improper budgetary allocation for rural health at state and central government level.

General Terms
Artificial Intelligence, Medical Diagnosis, Disease Management.

Keywords
Artificial Neural Network, Infant Disease Management, Malaria, Typhoid, Dengue, FFANN.

1. INTRODUCTION
Physicians intuitively exercise knowledge obtained from previous patients' symptoms. In everyday practice, the amount of medical knowledge grows steadily, such that it may become difficult for physicians to keep up with all the essential information gained. To quickly and accurately diagnose a patient, there is a critical need in employing computerized technologies to assist in medical diagnosis and access the related information. Computer-assisted technology is certainly helpful for inexperienced physicians in making medical diagnosis as well as for experienced physicians in supporting complex decisions.

Computer-assisted technology has become an attractive tool to help physicians in retrieving the medical information as well as in making decisions in face of today's medical complications. Machine learning techniques with computer-aided medical diagnosis should have good comprehensibility, i.e., the transparency of diagnostic knowledge and the explanation ability. Nowadays, as the computational power increases, the role of automatic visual inspection becomes more important. Image processing and artificial intelligence techniques are introduced that may provide a valuable tool.

Currently in Malaysia, the traditional method for the identification of Malaria parasites requires a trained technologist to manually examine and detect the number of the parasites subsequently by reading the slides. This is a very time consuming process, causes operator fatigue and is prone to human errors and inconsistency. An automated system is therefore needed to complete as much work as possible for the identification of Malaria parasites. The integrated system including soft computing tools has been successfully designed with the capability to improve the quality of the image, analyze and classify the image as well as calculating the number of Malaria parasites [1]. The cost of such system is high enough and hence not recommended for third world countries, especially India which has a dense rural population living below poverty line. Such systems do not consider multiple diseases at the same time for diagnosis.

Medical expert systems in various areas are certain to grow because huge medical data are provided according to increment of performance of medical systems/scanners. In them, the most famous medical expert system would be MYCIN [2]. MYCIN did not use fuzzy logic directly; one of primary components was the use of certainty factors. The medical expert systems using fuzzy logic directly are given in the references [3], [4] and [5]. Fuzzy degree of uncertainty or possibility degree of certain diagnosis is employed, as explained in [3], [4] and [5].

Artificial Intelligence (AI) is the study of mental facilities through the use of computational models. It has produced a number of tools. These tools are of great practical significance in engineering to solve various complex problems normally requiring human intelligence. In general, an artificial neural network is built in two steps, that is, generating component artificial neural networks and then combining their predictions. The powerful tools among these are expert system (knowledge-based system), ANN, Genetic Algorithm based ANN, Neural-Fuzzy, and Support Vector Machines (SVM).

1.1 Expert System
The expert system (ES), also known as knowledge-based systems (KBS), is basically computer programs embodying knowledge about a narrow domain for the solution of problems related to that domain. An ES mainly consists of a knowledge base and an inference mechanism. The knowledge base contains domain knowledge, which may be expressed as any combinations of 'IF-THEN' rules, factual statements, frames, objects, procedures and cases. The inference mechanism manipulates the stored knowledge to produce solutions.

1.2 Fuzzy Logic System

A demerit of an ordinary rule-based ES is that they cannot handle new situations which are not covered explicitly in their knowledge bases. Hence, ESs cannot give any conclusions in these situations. The fuzzy logic systems (FLSs) are based on a set of rules. These rules allow the input to be fuzzy, i.e. more like the natural way that human express knowledge [XS]. The use of fuzzy logic can enable ESs to be more practical. The knowledge in an ES employing fuzzy logic can be expressed as fuzzy rules (or qualitative statements). A reasoning procedure, the compositional rule of inference and conclusions are to be drawn by extrapolation or interpolation from the qualitative information stored in the knowledge base.

1.3 Artificial Neural Network

Artificial neural network (ANN) can capture domain knowledge from examples they can readily handle both continuous and discrete data and have good generalization capability as with fuzzy expert systems. An ANN is a computational model of the human brain. ANNs assume that computation is distributed over several simple units called neurons, which are interconnected and operate in parallel thus known as parallel distributed processing systems or connectionist systems. Implicit knowledge is built into a neural network by training it. Some ANNs can be trained by typical input patterns and the corresponding expected output patterns. The error between the actual and expected outputs is used to strengthen the weights of the connections between the neurons. This type of training is known as supervised training. Some of the ANNs are trained in an unsupervised mode, where only the input patterns are provided during training and the network learns automatically to cluster them in groups with similar features.

1.5 Genetic Algorithm

Genetic algorithm (GA) is a stochastic optimization procedure inspired by natural evolution. It can yield the global optimum solution in a complex multi-model search space without requiring specific knowledge about the problem to be solved. A genetic or evolutionary algorithm operates on a group or population of chromosomes at a time, iteratively applying genetically based operators such as crossover and mutation to produce fitter populations containing better solution chromosomes.

1.6 Support Vector Machine

Support Vector Machines (SVMs) are the methods for creating functions from a set of labeled training data. The function can be a classification function or the function can be a general regression function. For classification, SVMs operate by finding a hyper surface in the space of possible inputs, which will attempt to split the positive examples from the negative examples.

2. DETAILS OF THE ARTIFICIAL INTELLIGENT METHOD EMPLOYED

ANNs are highly interconnected processing units inspired in the human brain and its actual learning process. Interconnections between units have weights that multiply the values which go through them. Also, units normally have a fixed input called bias. Each of these units forms a weighted sum of its inputs, to which the bias is added. This sum is then passed through a transfer function.

Prediction with ANNs involves two steps, one is *training* and the other is *learning*. Training of Feed forward artificial neural networks (FFANNs) is normally performed in a supervised manner. The success of training is greatly affected by proper selection of inputs. In the learning process, a neural network constructs an input-output mapping, adjusting the weights and biases at each iteration based on the minimization or optimization of some error measured between the output produced and the desired output. This process is repeated until an acceptable criterion for convergence is reached. In the back propagation (BP) algorithm, the input is passed layer through layer until the final output is calculated, and it is compared to the real output to find the error. The error is then propagated back to the input adjusting the weights and biases in each layer. In the standard BP learning algorithm the sum of square errors is minimized. In order to accelerate the learning process, two parameters of the BP algorithm are adjusted: the learning rate and the momentum. The learning rate is the proportion of error gradient by which the weights are to be adjusted. Larger values give a faster convergence to the minimum. The momentum determines the proportion of the change of past weights that are used in the calculation of the new weights.

In this paper, the fully-connected feed forward multilayer perceptron network is used and trained. The network consists of an input layer representing the input data to the network, hidden layers and an output layer representing the response of the network. Each layer consists of a certain number of neurons; each neuron is connected to other neurons of the previous layer through adaptable synaptic weights *w* and biases *b*, as shown in Figure 1(a) and Figure 1(b).

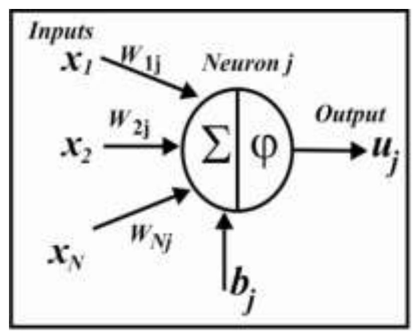

Figure 1(a) Information processing in ANN

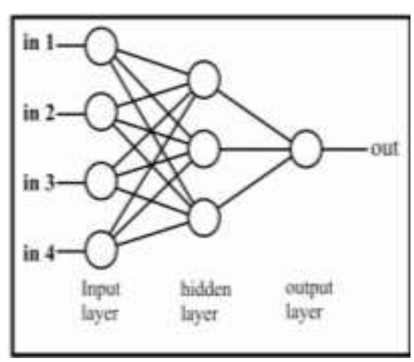

Figure 1(b) Architecture of ANN

If the inputs of neuron j are the variables $x_1, x_2, .. x_i, .., x_N$, the output u_j of neuron *j* is obtained as ,

$$u_j = \varphi(\sum_{i=1}^{N} w_{ij} x_i + b_j)$$

where, w_{ij} is the weight of the connection between neuron *j* and *i*-th input; b_j is the bias of neuron *j* and φ is the transfer (activation) function of neuron *j*.

An ANN of three layers (one hidden layer) is considered with N, M and Q neurons for the input, hidden and output layers, respectively. The input patterns of the ANN represented by a vector of variables $x = (x_1, x_2, \ldots, x_i, \ldots, x_N)$ submitted to the NN by the input layer are transferred to the hidden layer. Using the weight of the connection between the input and the hidden layer and the bias of the hidden layer, the output vector $u = (u_1, u_2, \ldots, u_j, \ldots, u_M)$ of the hidden layer is determined. The output u_j of neuron j is obtained as,

$$u_j = \varphi hid(\sum_{i=1}^{N} w_{ij}^{hid} x_i + b_j^{hid})$$

where, w_{ij}^{hid} is the weight of connection between neuron *j* in the hidden layer and the *i*-th neuron of the input layer, b_j^{hid} represents the bias of neuron *j* and φhid is the activation function of the hidden layer.

The values of the vector *u* of the hidden layer are transferred to the output layer. Using the weight of the connection between the hidden and output layers and the bias of the output layer, the output vector $y = (y_1, y_2, \ldots y_k, \ldots, y_Q)$ of the output layer is determined.

The output y_k of neuron *k* (of the output layer) is obtained as,

$$y_k = \varphi out(\sum_{j=1}^{M} w_{jk}^{out} u_j + b_k^{out})$$

where, w_{jk}^{out} is the weight of the connection between neuron *k* in the output layer and the *j*-th neuron of the hidden layer, b_k^{out} is the bias of neuron *k* and φout is the activation function of the output layer.

The output y_k is compared with the desired output (target value) y_k^d. The error E in the output layer between y_k and y_k^d $(y_k^d - y_k)$ is minimized using the mean square error at the output layer (which is composed of *Q* output neurons), defined by,

$$E = \frac{1}{2}\sum_{k=1}^{Q}(y_k^d - y_k)^2$$

Training is the process of adjusting connection weights *w* and biases *b*. In the first step, the network outputs and the difference between the actual (obtained) output and the desired (target) output (i.e., the error) is calculated for the initialized weights and biases (arbitrary values). In the second stage, the initialized weights in all links and biases in all neurons are adjusted to minimize the error by propagating the error backwards (the BP algorithm). The network outputs and the error are calculated again with the adapted weights and biases, and this training process is repeated at each epoch until a satisfied output y_k is obtained corresponding with minimum error. This is done by adjusting the weights and biases of the BP algorithm to minimize the total mean square error and is computed as,

$$\Delta w = w^{new} - w^{old} = -\eta \frac{\partial E}{\partial w} \quad (1a)$$

$$\Delta b = b^{new} - b^{old} = -\eta \frac{\partial E}{\partial b} \quad (1b)$$

where, η is the learning rate. Equations (1a) and (1b) show the generic rule used by the BP algorithm. Equations (2a) and (2b) illustrate this generic rule of adjusting the weights and biases. For the output layer, we have,

$$\Delta w_{jk}^{new} = \alpha \Delta w_{jk}^{old} + \eta \delta_k y_k \quad (2a)$$

$$\Delta b_k^{new} = \alpha \Delta b_k^{old} + \eta \delta_k \quad (2b)$$

where, α is the momentum factor (a constant between 0 and 1) and $\delta_k = y_k^d - y_k$

For the hidden layer, we get,

$$\Delta w_{ij}^{new} = \alpha \Delta w_{ij}^{old} + \eta \delta_j y_j \quad (3a)$$

$$\Delta b_j^{new} = \alpha \Delta b_j^{old} + \eta \delta_j \quad (3b)$$

where,

$$\delta_j = \sum_k^Q \delta_k w_{jk} \quad \text{and} \quad \delta_k = y_k^d - y_k$$

3. SYMPTOMATIC STUDIES OF SOME DISEASES

3.1 Typhoid

Typhoid is also known as enteric fever or salmonellosis. It is an infectious disease and is a very common cause of persistent high grade fever. Typhoid is a bacterial disease caused by *Salmonella typhi*. A related bacterium called *Salmonella paratyphi* causes paratyphoid fever. The disease is transmitted by contaminated food or water. These bacteria live within the gall bladders of some human beings without causing disease for years. These carriers pass these bacteria in their stools and if the carrier is a food handler, the disease spreads to a large number of people. The illness is also spread by a contaminated water supply. Since the bacteria are passed in the stools of the carriers as well as patients afflicted with acute illness, any contamination of the water supply with sewage spreads the disease in epidemic proportions [10, 12].

3.2 The symptoms of Typhoid

Fever is the main symptom that gradually increases over four to five days. The fever is high grade (up to 40.5ºC or 105ºF) and almost continuous unless some fever-relieving drugs (antipyretics) are taken. The appetite is poor and the patient feels weak. The liver and spleen become enlarged. In serious conditions, perforation of the intestines may occur in a few cases [9, 12].

3.3 Malaria

Malaria is a parasitic disease characterized by high fever, chills and rigors. *Falciparum* malaria, one of four different types, affects a greater proportion of the red blood cells than the other types and is more serious. The disease is a major health problem in India as in most of the tropics and subtropics. Malaria is caused by a parasite (*Plasmodium*) that is transmitted from one human to another by the bite of infected anopheles mosquitoes. The symptoms occur in cycles of 48 to 72 hours. This is the time taken by the parasites to multiply inside the red blood cells, which then rupture, and the parasites infect more red blood cells. Malaria can also be transmitted congenitally (from a mother to her unborn baby) and, rarely, by blood transfusions [10, 12].

3.4 The Symptoms of Malaria

There are sequential chills, fever, and sweating accompanied by headache, nausea and vomiting, muscle pain and anemia. In severe cases, there may also be jaundice, convulsions or coma [6].

3.5 Dengue

The principle vector involved in transmission of the virus is the mosquito, *Aedes aegypti* (dominant in India) and *Aedes albopictus*. Clinical illness begins after a period of 5 to 6 days (variation 3 to 15 days) of the bite preceded a day before by viremia which continues till 4 to 5 days subsequent to the onset of clinical illness [10, 12].

3.6 The Symptoms of Dengue

Symptoms of dengue are sudden onset of moderate to high fever, headache, retro-orbital pain, muscle, bone and joint pain, anorexia, bad taste in mouth and flushing of face. In fair-skinned individuals, a maculopapular rash may be seen over the trunk and upper limbs for 3 to 4 days after the onset of fever, lasting from a few hours to a few days [11, 12].

4. RESULTS AND DISCUSSION

The FFANN is a widely accepted classifier. However, the success of FFANN to distinguish between Malaria and Thyphoid is strongly related to the success in the pre-processing of its input data. The inputs should contain lot of information in order for the network to properly classify the events. In this paper, three-layer FFANN is used and trained with a supervised learning algorithm called back propagation (BP). The FFANN consists of one input layer, one hidden layer and one output layer. The input layer consists of neurons: the inputs to these neurons are various symptoms for Malaria, Thyphoid and Dengue like temperature, abdominal pain, pulse, vomiting, rashes, joint pain and muscle pain etc. The output layer consists of three neurons representing the Malaria, Thyphoid and Dengue. With respect to the hidden layer, it is customary that the number of neurons in the hidden layer is done by trial and error. The same approach is used in the proposed algorithm. Symptoms of three diseases were used as an input to the network. For generalization, the randomized data is fed to the network and is trained for different hidden layers. Various training methods of Conjugate Gradient (CG) BP and Levenberg– Marquardt BP are used for training the network and average minimum MSE on training and testing data is obtained. For all training methods, it is assumed that learning rate LR = 0.7, momentum MM = 0.6, data used for training purpose TR = 10%, for cross validation CV = 10% and for testing purpose TS = 80%. With these assumptions, the variation of average MSE and percent accuracy of classification for Malaria, Thyphoid and Dengue with respect to the number of processing elements in the hidden layer is obtained.

Figure 2 shows the variation of percentage of classification accuracy with respect to the number processing elements in the hidden layer. It is found that in Levenberg–Marquardt BP ('trainlm') method for four processing elements in the hidden layer the minimum MSE (0.00014) is obtained, which the lowest value is obtained by any method and 100% classification between Malaria, Thyphoid and Dengue, which means there is a clear discrimination between Malaria, Thyphoid and Dengue with this method. Hence, this network of Levenberg–Marquardt BP 'trainlm' for learning rate LR = 0.7, momentum MM = 0.6, training data TR = 10%, cross validation CV = 10% and testing data TS = 80% with 04 number of processing elements in the hidden layer is the best-suited network.

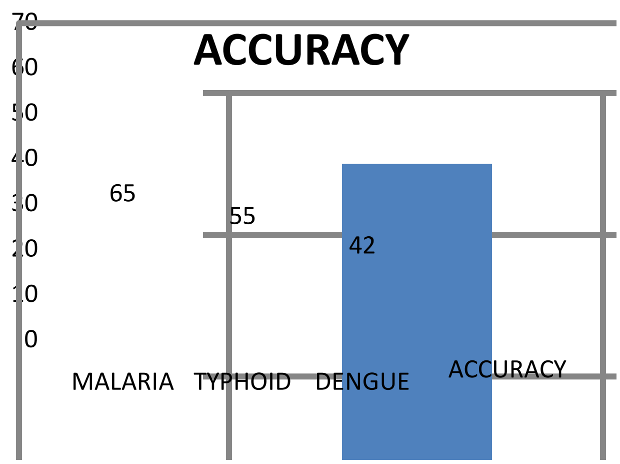

Figure 2 (a) Accuracy for number of processing element 1

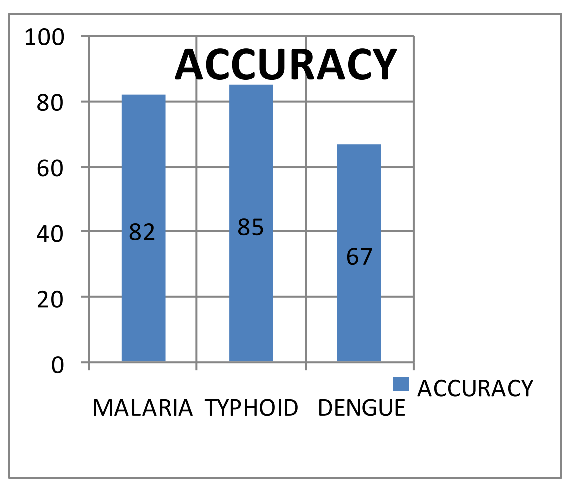

Figure 2(b) Accuracy for number of processing element 2

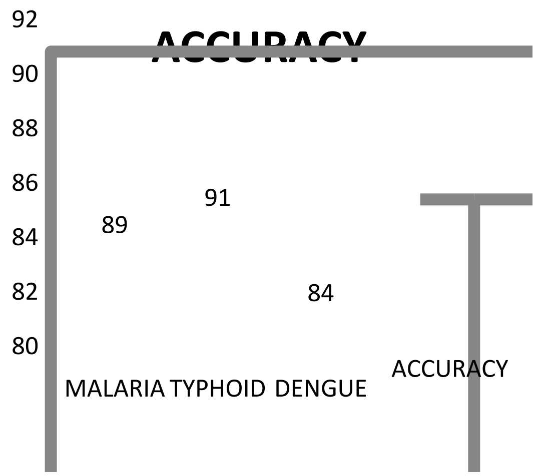

Figure 2(c) Accuracy for number of processing element 3

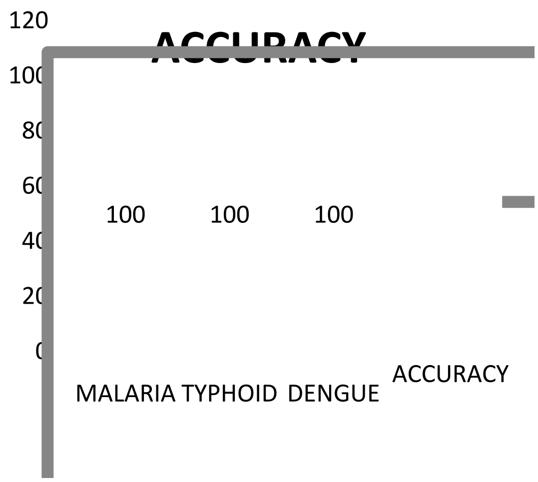

Figure 2(d) Accuracy for number of processing element 4

Figure 2 Variation of percentage accuracy with number of processing elements in hidden layer

5. CONCLUSION

The artificial neural network for the diagnosis provides an efficient way to assist the doctor in the diagnosis of Malaria, Typhoid and Dengue. The feedback error propagation learning allows the programmer to use the doctor's experience in training the network. The assistant of doctors can feed symptoms and the diagnosis given is similar to that of given by the doctor itself.

Also it is a monotony free and presumption free diagnostic system. It also provides a better alternative to process abstract medical data over the conventional programming method.

6. ACKNOWLEDGMENTS

The authors are thankful to Dr. G.M.Dhole, Prof. S.R.Paraskar and Prof. M.A.Beg of Shri.Sant Gajanan Invention and Advanced Research Centre (SGIARC), Shegaon – 444203 (India) for necessary support and testing facilities. Authors are also thankful to the child specialists Dr. Manohar Wankhade, Shegaon (India) and Dr. Hari Wadode, Khamgaon (India) for permission of collecting patient's data.

7. REFERENCES

[1] Tohal S.F. and Ngah U.K., "Computer Aided Medical Diagnosis for the Identification of Malaria Parasites", IEEE - ICSCN 2007, MIT Campus, Anna University, Chennai, India. Feb. 22-24, 2007. pp.521-522.

[2] Shordiffe E.H., "Computer-Based Medical Consultation, MYCIN", Elsevier /North Holland, New York , 1976.

[3] Adlassnig K.P., "A survey on medical diagnosis and fuzzy subsets in M.M. Gupta and E. Sanchez (Eds): Approximate Reasoning in Decision Analysis", North-Holland, pp.203-217, 1982.

[4] Esogbue A.O., "Measurement and valuation of a fuzzy mathematical model for medical diagnosis", Fuzzy sets and Systems, 10, pp.223-242,1983.

[5] Hudson D.L. and Cohen M.E., "Fuzzy Logic in Medical Expert System", IEEE Eng. Med. and Bio., pp. 693-698, 1994.

[6] Haykin Simon, Neural Networks – A Comprehensive Foundation, Pearson Education, Inc., 2001.

[7] Zurada J.M., Introduction to Artificial Neural Systems, West Publishing Co, 2002.

[8] Freeman J.A., Skapura D.M., Neural Networks – Algorithms, Applications, and Programming Techniques, Computation and Neural Systems Series, Pearson Education, Inc., 2007.

[9] Chatterjee C.C., Human Physiology – Volume I, Medical Allied Agency, 2004.

[10] Chatterjee C.C., Human Physiology – Volume II, Medical Allied Agency, 2002.

[11] Santosh Kumar A., Paediatric Clinical Examination, Paras Medical Publisher, 2008.

[12] Suraj Gupte, The Short Textbook of Pediatrics – Incorporating National and International Recommendations (MCI, IAP, NNF, WHO, UNICEF, IPA, ISTP, AAP, etc), Jaypee Brothers Medical Publishers, 2009.