

# **PERFORMANCE EVALUATION OF THE NEURAL NETWORK DIAGNOSTIC SYSTEM FOR THE RE-EMERGING ARBOVIRAL INFECTION -DENGUE**

A.ShameemFathima<sup>1</sup>, Dr.D.Manimegalai<sup>2</sup>

**Abstract-** The prevalence of emerging infectious diseases - Dengue and Chikungunya in humans has increased within the recent past and also are of great public health concern in India. This study aims to design and develop the neural network model that finds relationships and aims to classify arboviral infection amongst patients. It also investigates the effects of clinical manifestations and examines the performance of the four training algorithms of Levenberg-marquardt back propagation attempting to minimize error in classifying the patients as dengue positive or negative. The learning algorithm is applied on the proposed neural network model to minimize the error by continuously training the network until the optimum efficiency level is obtained.

**Keywords –** Artificial Neural Networks; Levenberg-Marquardt Algorithm; Arboviral Classification; Multi Layer Perceptron Networks

## **I. INTRODUCTION**

The prevalence of emerging infectious diseases in humans has increased within the recent past and also threatens to increase in the near future. Dengue, a Flavivirus and Chikungunya, an Alphavirus, transmitted by Aedes mosquitoes, [1] are a cause of great concern to public health in India, especially in delta regions. Because of a number of factors, including urbanization, globalization of travel, and lack of efficient chemical pesticide-based vector control interventions, these infections have re-emerged as a significant international community health problem. Chikungunya fever resembles Dengue fever, and is characterized by severe, sometimes persistent, joint pain (arthritis), as well as fever and rash [2]. Diagnosis of Dengue starts with a clinical suspicion, prompted by the recognition of a collection of presenting symptoms and signs. In the early acute febrile phase of illness, Dengue patients often present with a history of sudden onset fever, which is often accompanied by nausea, aches and pains. As, the clinical manifestations of Chikungunya fever resemble those of Dengue fever it has to be distinguished from Dengue fever. [3] Co-occurrence of both fevers has been observed in Tamil Nadu state of India thus highlighting the importance of strong clinical suspicion and efficient laboratory support. It has been postulated that many cases of Dengue virus infection are misdiagnosed and that the incidence of CHIKV infection is much higher than reported [4]. Therefore, the present study was undertaken to diagnose Chikungunya infection in clinically suspected Dengue patients and the vice-versa.

Most of the research studies done on Dengue, involves weather parameters, considering the significance of weather variables in spreading DF. However, it is difficult to find a research that is using the same Dengue data to do the classification, but there are researches related with this technique that is using different Dengue data to choose the best method for knowledge acquisition for Dengue dataset. In medical decision making, a variety of neural networks are used for decision accuracy. A neural network with enough elements (called neurons) can classify any data with arbitrary accuracy. They are particularly well suited for complex decision boundary problems over many variables. Therefore we choose neural networks as a good candidate for solving the Arboviral disease classification problem which contains the numerical data such as our dataset. Considering all the above research studies, first, we perform a standard presentation of results that promote and facilitate future comparisons to recognize important trends and patterns which, in turn, will inform the development of opinion and medical practice and patients affected by

---

<sup>1</sup> Department of Computer Science and Engineering, Manonmanium Sundaranar University, Tirunelveli, Tamilnadu, India

<sup>2</sup> Department of Information Technology, National Engineering College, Kovilpatti, Tamilnadu, India

Arbovirus. The Rest of this paper elaborates on the methodology adapted in carrying out the study to develop an ANN model to predict Dengue disease, the findings and conclusions drawn from those outcomes.

## II. METHODOLOGY

### A. Data Collection and Selecting Input Parameters–

The medical dataset we are classifying includes real records of patients suffering from viral infection from several hospitals, King Institute of preventive Medicine and laboratory diagnostic centers in Tamil Nadu, India. Blood samples were collected from patients, who were clinically diagnosed to have acute viral infection, at the time of hospital admission and the time of discharge. The relationships between objects are represented by numerical relationships between variables..

In our study of interest, Datasets that are available for Dengue describe information about the patients suffering with Dengue disease and without Dengue disease along with their symptoms .Thirty one independent variables factors/symptoms causes for infection (Age, Gender, Fever, Chills, Coryza, Systolic pressure, Diastolic pressure, Shock, Myalgia, Malaise, Arthralgia , Hallucinations,Confusion, Altered consciousness, Unconscious, Convulsion, Neck rigidity, Motorweakness, Paralysis, Lymphadenopathy, Skin rash, Hemorrhagic symptoms, Pleural effusion, Hemoglobin, Red Blood Count and Platelet count at admission, Platelet count during treatment and Platelet count at discharge) are the scaled numeric variables and Infection is the categorical variable. The data undergoes an extensive data preprocessing task such as data selection, cleaning, reduction, and discretization.

### B. Designing and Training the ANN–

The focus of this study is only on supervised learning, in which the output from the neural network is compared with a set of targets and the error signal is used to update the weights in the neural network.

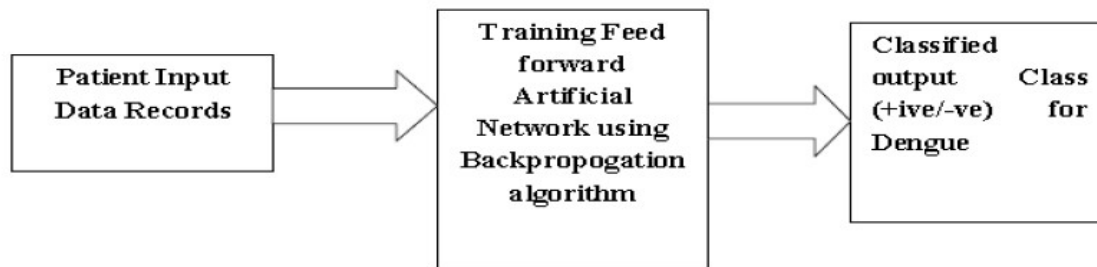


Figure 1. Neural Network Classification System Diagram

The ANN used to predict Dengue cases is designed as a Multi Layer Perceptron (MLP) network with three initial layers, namely – Input layer, Output layer and one hidden layer. The input layer consists of the input parameters identified in the previous section. The hidden layer is designed with default ten neurons with the Levenberg-Marquardt back propagation as the learning algorithm and Output layer consists of a single parameter which is the prediction of Dengue cases. Tan-sigmoid was used as the activation function for the input, while pure-linear was used for the output.

The entire dataset that is read as numeric data in double array from the Microsoft Excel spreadsheet file and is split in to two datasets based on the output of the target variable - Dengue positive or Dengue negative. This dataset is then again divided in 80% for training and 20% for testing purpose. At the next stage, both 80% training set is concatenated in to training and 20% of the testing set in both dataset is concatenated together to achieve a testing set for 20%.This is done to achieve best classification accuracy results.

	1	2	3	4	5	6	7	8	9
1	736	19	2	1	3	1	1	130	
2	737	50	2	2	7	1	1	120	
3	738	33	2	2	7	1	1	120	
4	739	14	1	2	7	0.5000	1	110	
5	740	15	1	1	3	0.5000	1	100	
6	741	53	2	1	13	0.5000	1	100	
7	742	17	2	1	3	1	1	100	
8	743	50	2	2	2	1	1	140	
9	744	70	2	2	4	0.5000	0.5000	120	
10	745	62	2	1	15	1	1	110	
11	746	44	2	1	23	1	0.5000	120	

Figure 2. Training Data

As the classification is done using MATLAB, which does not give convincing results when input layer contains 0 values all the input values that has 0 for all the attributes are converted to 0.5.[ 5]The target variable is separated from the dataset and the remaining 31 attributes are applied as input with completing all the preprocessing. This can be viewed in the Figure2 and 3

	8	9	10	11	12	13	14	15	16
1	90	70	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000
2	120	80	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000
3	120	90	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000
4	110	90	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000
5	110	90	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000
6	90	70	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000
7	120	80	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000
8	120	90	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000
9	120	80	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000
10	120	90	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000
11	120	80	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000

Figure 3. Testing Data

### III. EXPERIMENTS AND FINDINGS

After converting the raw data, we put in the data into MATLAB workspace, ready to continue the main process. The Steps involved in this next task are defined below :

- 1) Define one sample: inputs and outputs
- 2) Define and custom network
- 3) Define topology and transfer function
- 4) Configure network
- 5) Train net and calculate neuron output .

The model we designed for the basic classification performance analysis is shown in Figure 4 below.

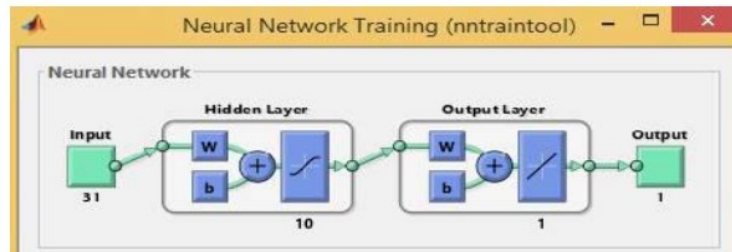


Figure 4. Architecture of Proposed Neural Network

The following step after creating the architecture of neural network is doing training session that matches measurements that each individual has with their classes (Dengue positive, Dengue Negative).After calling the „train“ function, there will be an interface as shown below,

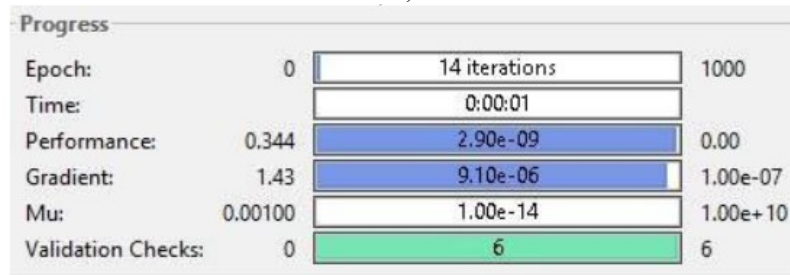


Figure 5. Performance Measures of Proposed Neural Network

The first experiment is done using the `trainlm`, often the fastest Back Propagation algorithm, and is highly recommended as a first-choice supervised algorithm. The trained network is used to make a prediction on the test data and confusion matrix is generated for comparison with other techniques. The network performance with the different parameters will be determined by the mean of the squared errors on the test set since those test data will not be used for training. The mean squared error (MSE) is the average squared difference between outputs and targets. Lower values are better while zero means no error. [6][7]

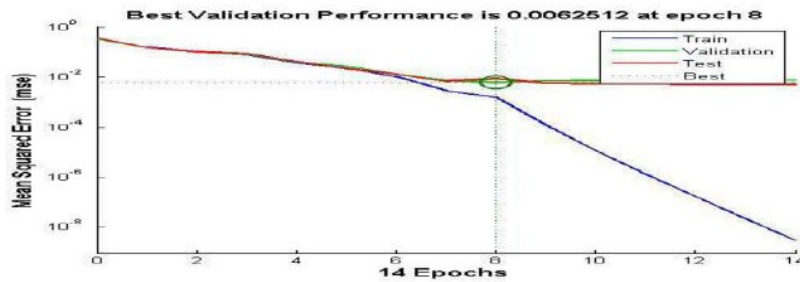


Figure 6. Plot showing the best validation Performance of the Levenberg-Marquardt (`trainlm`)

The second experiment is done using the Artificial Neural Network with OneHidden Unit and default ten neurons in the hidden layer with the Bayesian regulation back propagation. `trainbr` is a network training function that updates the weight and bias values according to Levenberg-Marquardt optimization. [8] It minimizes a combination of squared errors and weights, and then determines the correct combination so as to produce a network that generalizes well.

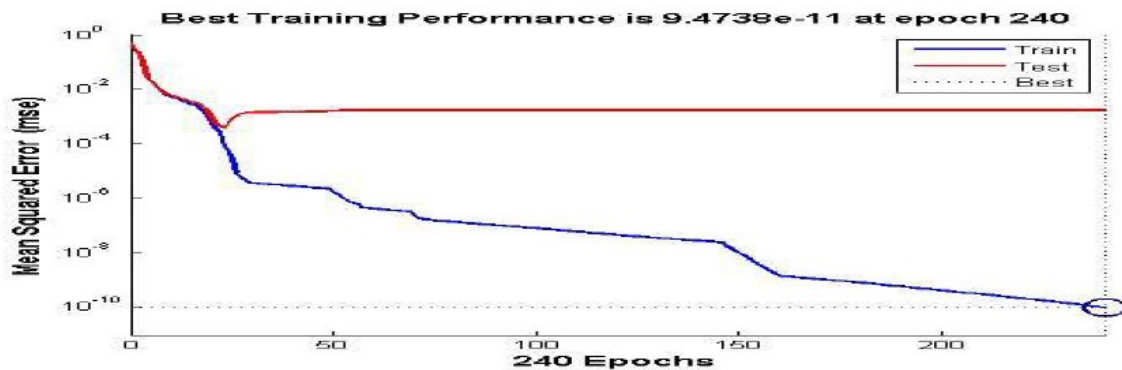


Figure 7. Plot showing the best validation Performance of the Bayesian regulation Back Propagation (`trainbr`)

The third experiment is done using with the Scaled conjugate gradient Back propagation. `trainsg` is a network training function that updates weight and bias values according to the scaled conjugate gradient method.[9]

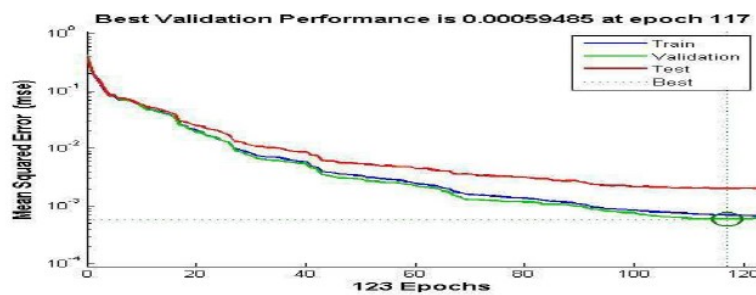


Figure 8. Plot showing the best validation Performance of the Scaled Conjugate Gradient Back propagation. (`trainsg`)

Finally, the experiment is done using with the Resilient Back Propagation. trainrp is a network training function that updates weight and bias values according to the resilient Back Propagation algorithm (Rprop).[10]

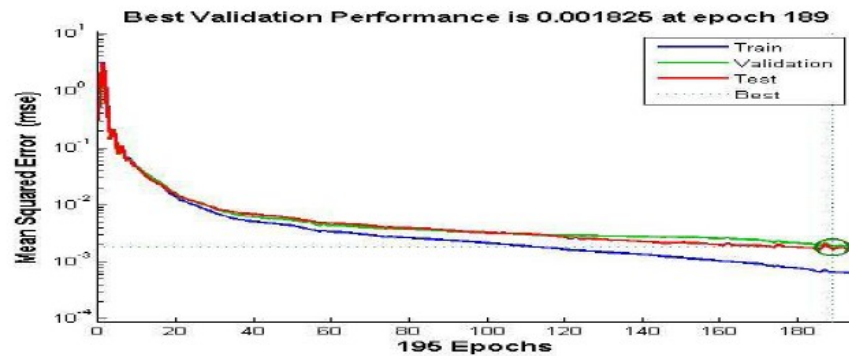


Figure 9. Plot showing the best validation Performance of the Resilient Back Propagation (trainrp)

The confusion matrix of the classifiers - Levenberg-Marquardt Back Propagation (trainlm), Bayesian regulation Back Propagation (trainbr) Scaled conjugate gradient backpropagation (trainscg), Resilient back propagation (trainrp) are shown in the Table – 1.

Table -1 Evaluation Measure of the Proposed Neural Network Model

Training algorithms used for Classification	Evaluation Measure			
	TP	TN	FP	FN
Lavenberg-Marquardt back propagation	843	112	36	0
Bayesian regulation back propagation	843	117	31	0
Scaled conjugate gradient back propagation	843	115	33	0
Resilient back propagation	843	108	40	0

The performance measures of the compared classifiers are calculated as sensitivity, specificity and accuracy are tabulated as shown in the Table 2.

Table -2 Performance Measure of the Proposed Neural Network Model

Methods	Performance Measurements		
	Sensitivity	Specificity	Accuracy
Lm	1	0.756757	0.963673
Br	1	0.790541	0.968718
Scg	1	0.777027	0.9667
Rp	1	0.72973	0.959637

#### IV. CONCLUSION

The experiments revealed that best accuracy is achieved by Bayesian Regularization Back Propagation when the number of input layers is 31. Using all of these previously described dataset, the experiment determines how well neural network structures perform in predicting the disease-Dengue. As the research was performed on a single dataset, it is recommended that the LM algorithm be applied to additional real world datasets to determine its proficiency in dealing with different conditions and phenomena. Even with the massive amount of neural network research that has taken place, there is still a large amount of controversy about the best way to determine the number of hidden layers and hidden layer neurons. It is the scope of this research, to explore how best the hidden layers and hidden layer neurons can be evolved through the use of genetic algorithms and relevant inputs can be determined while simultaneously searching for a global solution.

#### V. ACKNOWLEDGEMENTS

We acknowledge the help and guidance in terms of medical expertise provided by the esteemed doctors and microbiologists at the virology department of King Institute of Preventive Medicine, Chennai, Tamilnadu who provided us with a cosmic amount of viral data needed for our research study.

#### REFERENCES

- [1] Kaushik RM, Varma A, Kaushik R, Gaur, KJBS. "Concurrent dengue and malaria due to Plasmodium falciparum and P. vivax". *Trans R Soc Trop Med Hyg*2007; 101:1048-1050.
- [2] Srinivasan.V, Iyem Perumal.A, Rajagopalan.S.P, Dheva Rajan., "A Surveillance of Dengue Fever: A Review Article ", *American Journal of Sustainable Cities and Society*, Issue 3, Vol. 1 January 2014
- [3] Deen JL et al., "The WHO dengue classification and case definitions: time for a reassessment". *Lancet*, 2006,368:170–173.
- [4] Dengue. Guidelines for diagnosis, treatment prevention and control, Geneva, World Health Organization, 2009, WHO/HTM/NTD/DEN/2009
- [5] [https://www.mathworks.com/help/pdf\\_doc/nnet/nnet\\_ug.pdf](https://www.mathworks.com/help/pdf_doc/nnet/nnet_ug.pdf)
- [6] Hornik, K., Stinchomb, M. and White, H. (1989), "Multilayer Feedforward Neural Networks are Universal Approximators," *Neural Networks*, vol. 2, pp. 359-366.
- [7] P. Lippmann, R. P. (1987), "An Introduction to Computing with Neural Nets," *IEEE ASSP Magazine*, pp. 4-22.
- [8] Tickle, A., Andrews, R. and Diederich, J. (1998), "The Truth will Come to Light: Directions and Challenges in Extracting the Knowledge Embedded within Trained Artificial Neural Networks," *IEEE Transactions on Neural Networks*, vol. 9, pp. 1057-1068.
- [9] Patil, S. &Kumaraswamy, Y., "Intelligent and effective heart attack prediction system using data mining and artificial neural network", *European Journal of Science Research*. Vol 31, No. 4.2009.
- [10] Simon Haykin, (2001) "Neural Networks – A Comprehensive Foundation", PearsonEducation.