

Prediction of Zika Virus by Multilayer Perceptron Neural Network (MLPNN) Using Cloud

B. Mahalakshmi and G. Suseendran

Abstract--- Zika virus a mosquito borne flavivirus disease, which is spreading hastily across all over the world. Nearly 95 countries are infected with Zika, *Aedes aegypti* Mosquitoes is the source of spreading the virus. Microcephaly, myelitis, Guillain – Barre Syndrome and neuropathy are the causes of ZVD. Miscarriages and preterm birth also possible also occur during the time of infection. To overcome an early prediction system is used for detecting the virus using symptoms. The zika dataset is stored in cloud and in our proposed work a Multilayer Perceptron Neural Network classifier used for predicting the Zika virus. The classifier produces accuracy level of 97% the highest accuracy level. Based on the symptoms ZVD is predicted at earlier stage, if they found as infected RNA test will be taken for the concerned person.

Keywords--- Zika Virus, MLPNN, Cloud Computing, Microcephaly, Backpropagation.

I. INTRODUCTION

In 1947, Uganda forest first Zika virus was identified in Rhesus monkeys blood. It is a mosquito-borne flavivirus the spreader was *Aedes* mosquitoes and it is the major reason behind ZIKA virus. In 1952 in Uganda and united republic of Tanzania it was identified in humans. In Africa, America, Asia and the Pacific are the places where the outbreak of ZIKA found [1]. During the period of 1960s to 1980s, intermittent cases of human infections were found across Africa and Asia, usually come with mild illness.

In 2009, outbreak of first ZIKA was found in the Federated States of Micronesia an Island of Yap. In 2013, a hefty outbreak of ZIKA infection was found in French Polynesia followed by other countries and some territories in pacific [2]. In Brazil a report said outsized outbreak of ZIKA found by symptoms of rashes and in July 2015 the Guillain –Barre syndrome was identified. During October 2015, Zika Virus with microcephaly was found. Outbreaks and confirmation of diffusion appeared all through the America, Africa and other parts of the world. From the last update around 95 countries and territories have a report of Zika transmission [3].

These types of mosquitoes primarily transmit the virus and cause the infections during day time. It is similar to Dengue, Chikungunya, malaria etc where the symptoms are same. Symptoms such as mild fever, rash, muscle pain, joint pain, rashes, headache usually occurs for 2-7 days [4]. For most people the symptoms do not grow for ZIKA virus. If a

woman is pregnant it should be taken more care for examine whether ZIKA infected or not. Because ZVD during pregnant affect the infants with a disease called microcephaly i.e. a child born with small head, not only this and also congenital Zika syndrome [5]. The ZVD makes pregnancy period a complicated time, where miscarriages and preterm birth also occur. ZVD cause neurological complications for adults and children such as GB Syndrome-Guillain –Barre Syndrome, neuropathy and myelitis.

II. RELATED REVIEW

Devendra T. Mourya et.al, 2016[6] have proposed that ZiV, flavivirus like dengue and Chikungunya were spread among various countries include Brazil and declared it as a global health emergency. Various geographical areas were infected because of this virus. Because of this various neurological symptoms and infections were revealed. So precautions should be taken to prevent from Zika as it is very infective for the pregnant women and infants, adults everyone.

Sumit Bhardwaj et.al, 2017[7] confirmed about the presence of Zika virus in India. On screening four of the zika infected cases were found based on the samples. Here prevalence of low level infection in India was discussed that is the threshold level and gradient levels of the mosquitoes are low compared with other country mosquitoes. So it may be a great issue in future. Comparing with Chikungunya Zika is going to be a big issue in future.

Petersen et al.[8] represent about Zika outbreak and the risk level of Zika among pregnant women's and the infections of infants were discussed. The microcephaly is a major problem of the infants and some syndromes related to neurological issues were discussed.

Amadin, F. I. and Bello, M. E. 2018[9], have proposed that MPNN algorithm is considered to be a better algorithm for psychological symptoms diagnosis and provide and accuracy of 88%. This algorithm provides a high level of accuracy and backpropagation method was used for predicting the accurate result.

Kadri SM et.al 2016[10] have proposed that Zika virus declared as a global public emergency by WHO in 2016. It was spreading from Africa to Asia and captured and exceptional attention among globally. So infographic knowledge was provided to the countries that are in a position of Zika virus risk at low level. A preventive measure should be taken.

Manuscript received September 16, 2019.

B. Mahalakshmi, Ph.D. Research Scholar, Department of Computer Science, School of Computing Science, Vels Institute of Science, Technology & Advanced Studies (VISTAS), Chennai, Tamil Nadu, India. (e-mail: maha.karthik921@gmail.com)

G. Suseendran, Department of Information Technology, School of Computing Science, Vels Institute of Science, Technology & Advanced Studies (VISTAS), Chennai, Tamil Nadu, India. (e-mail: suseendar_1234@yahoo.co.in)

III. PROPOSED METHOD

In our proposed system the data collection of Zika is purely based on the symptom based so a complete analysis taken through internet and collected the symptoms. A dataset is created based on the information collected. So ANN is considered to be best way to predict the virus. The entire dataset is stored in cloud, and then a classifier is used for predicting the system. In e-healthcare system cloud plays a major role in maintaining the patient’s information. The dataset used here for predicting the zika are synthetic data. The symptoms are collected thoroughly through internet. Based on the information collected the dataset is prepared. By analyzing Zika virus and its symptoms we created nearly 530 instances and nearly 15 attributes are created. Out of that the pregnant women are to be given a top priority because of the microcephaly disease for the infants.

3.1. Artificial Neural Network

An ANN- Artificial Neural Network is a model for imitating as human brain for the functionality and structure of neurons. Artificial neuron is the basic component of ANN [11] [12]. Basically it has three layers, such as input layer, hidden and output layer. In the hidden layer there may be two, three and up to n layers available. Every value in the input layer has its own value and the connection through hidden layer having the neuron weight. Each has to perform a set of operation such as multiplication, summation and activation in that order. Each neuron is weighted and it is multiplied with the input value and in hidden layer neuron summation value from the input layer generates the output by final activation function from the hidden layer. The below figure 1 shows the Multilayer perceptron network.

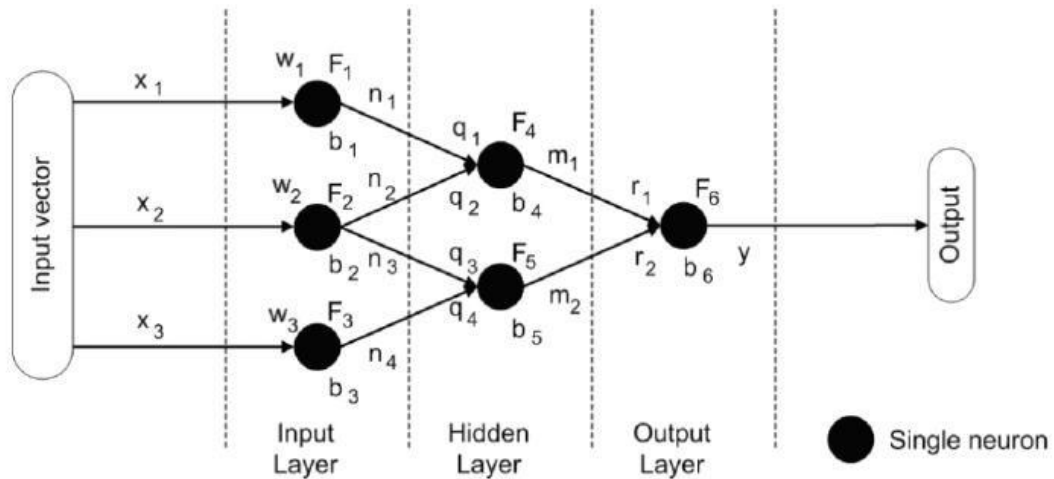


Figure 1: Multilayer Perceptron Neural Network

The purpose is to construct network with perfect number of nodes which are from hidden nodes which gives a good quality solution for the problem. The figure 2 below shows the good fit data result.

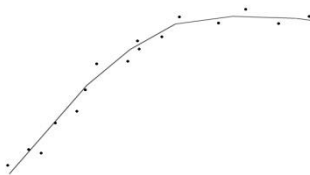


Figure 2: Hidden data with good fit

3.2. Feedforward Method

For classifying the instances back propagation method is used in the multilayer perceptron classifier. The MLP algorithm created the network, which is scrutinized and customized at the training time. The network is with sigmoid nodes for the threshold value and except for numeric class [13] [14]. For producing complex output the back propagation method is necessary for the elements. It performs a feed forward method for inputs in the network. While doing the iterative process the set of weights are taken for predicting the class label. The feed forward algorithm having one input layer, one or more hidden and

one output layer. While at the time of classification if error occurs the backpropagation method is used. To improve the accuracy of classification the number of inputs and the training time is reduced. [15]

$$Net_j = \sum w_{ij} x_j + \theta_j \quad (1)$$

Where w is the weight of each node, x is the input value from i to j. Next the sigmoid function nothing but the activation function are calculated using the equation

$$O_j = x_k = \frac{1}{1 + e^{-Net_j}} \quad (2)$$

Using the above equation if any error occurs during the feedforwarding method i.e. while forwarding the input value it passes through the hidden layer and output is received. If the required output is not attained not the backpropagation method is used. In this the sigmoid function is calculated by giving the output back to the hidden layer then the input again processed using feedforwarding concept[16].

3.3. Backpropagation Method

3.3.1. Correcting Error and Adjustment of Weights

The output node k, having the activation value as and for node k the expected target value is t_k , and then the difference between the expected and actual is given as

$$\Delta_k = t_k - O_k \quad (3)$$

For node k the error signal is calculated as

$$\delta_k = \Delta_k O_k (1 - O_k) \quad (4)$$

Where $O_k (1 - O_k)$ is the sigmoid function derivative.

Using the delta rule, the weight of input node j and output node k is proportional to error at node k produced by node j activation value. For adjusting the weight $w_{j,k}$, the formula is used between the output node k and j is follows,

$$\Delta w_{j,k} = I_r \delta_k X_k \quad (5)$$

$$w_{j,k} = w_{j,k} + \Delta w_{j,k} \quad (6)$$

Where $\Delta w_{j,k}$ weight change between nodes j and k and I_r is learning rate which is relatively small and constant that indicates the adjustment in weight. Based on the learning rate the network is identified. If it is too low learning of network is very slow, if it too high the network will move back and forth around minimum values. Changing the learning rate from large to small during backpropagation method provides some advantages. Assume a state that a network initiates far from the set of optimal weights the training will be rapid initially.

While at the process of learning if the learning rate decreases it is said that it reaches a minimum optimal point. If the learning processes get slowed the network will join a solution as the overshooting possibility reduced [21].

When the learning process close to optimal point initially, the system may oscillate but when the learning rate reduced the effect also decrease with time.

So as to improve the adjustment in the weights, changes have to be made with equation (5) as follows.

$$\Delta w_{j,k}^n = I_r \delta_k X_k + \Delta w_{j,k}^{(n-1)} \mu \quad (7)$$

Here the iteration is established with a momentum which is multiplied to the (n-1) iteration. The term momentum is used speed up the learning process by changing the weight by continuing in same path with better steps.

3.3.2. Hidden Layer

Hidden Layer error signal for node j is calculated as follows

$$\delta_k = (t_k - O_k) O_k \sum (w_{j,k} \delta_k) \quad (8)$$

Finally the purpose of back propagation is achieved by minimizing the error on the output nodes in the neural network. The global error is calculated for all type of neural network patterns

$$E = \frac{1}{2} \sum (\sum (t_k - O_k)^2) \quad (9)$$

When the neural network correctly trained then the value of error function should be zero. The below the pseudo code for the multilayer perceptron algorithm is explained.

3.3. Algorithm

Initializing all the Input X_i and weight w_i to the i^{th} node for each layer

Repeat the procedure for every training data Net_{ij} ,

Apply the input to the network

Feed forward

Giving input to every node and every layer in the network

Calculate the sum of weight of each inputs to the node adding threshold t_k value to the total

Finally activation for every node O_j is inverse of e^{-Net_j}

end

end

Backpropagation

In output every node is to be verified

If there is any error in the output layer then, calculate

$$w_{j,k} = w_{j,k} + \Delta w_{j,k}$$

end

Hidden Layer

For every hidden layer and node in the layer

The nodes error signal δ_k is calculated as $(t_k - O_k)$

Then each nodes weight w_i is updated $O_k \sum (w_{j,k} \delta_k)$

end

end

Error Calculation

Error Function is calculated

End

While ((max No of iterations < than precise) AND (Error Function > than precise))

End.

IV. EXPERIMENT AND RESULT

The Weka 3.8 version is used for the prediction of zika virus. It is software for data mining which contains a collection of machine learning algorithms. It uses the ARFF file format for analyzing the data as a default one, then CSV also a supported file format for Weka. In this the explorer option is used for training set data and testing dataset the explorer tab is used. The synthetic data was created with all possible combinations with our assumptions. It is considered to be a rare infection in India it is difficult to collect the data. The dataset which provided in above section is used in weka tool. All the combinations of the symptoms are used. And nearly 500 instances and 15 attributes are created. Based on the information the person is classified as infected or uninfected. Here the MLP classifier is used for classifying the instances.

The MLP classifier provides an accuracy of 97 percent. The instances are classified with its true positive and false positive rate as infected and uninfected. The TP and FP values are 97 and 1.4 respectively for the given dataset. The sensitivity and the specificity of the instances are measured using the formula.

$$\text{Sensitivity} = \text{TP} / (\text{TP} + \text{FN})$$

The Sensitivity of the instances is found out using the True Positive by the total no of True positive and false negative values.

$$\text{Specificity} = \text{TN} / (\text{TN} + \text{FP})$$

The Specificity is calculated using the total no of True Negative by its total no of True Negative and False Positive values.

$$\text{Accuracy} = (\text{TN} + \text{TP}) / (\text{TN} + \text{TP} + \text{FN} + \text{FP})$$

The accuracy is the proportion of the total number of true negative that is correctly found instances by total number of all the instances values. The accuracy, specificity and sensitivity result analysis is shown in figure 3.



PREDICTION OF ZIKA VIRUS BY MULTILAYER PERCEPTRON NEURAL NETWORK (MLPNN) USING CLOUD

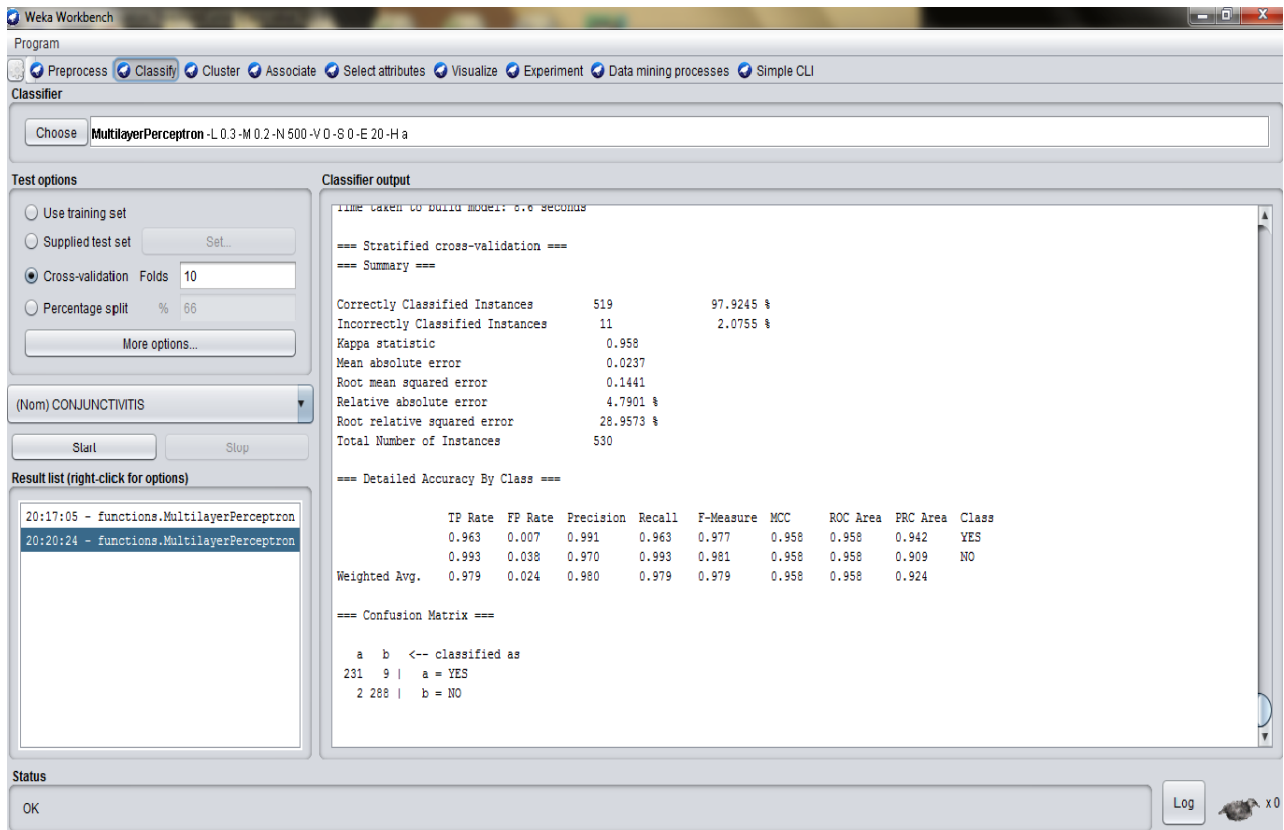


Figure 3: Result analysis using Weka

The threshold value for each class probability varies. So a classifier returns a threshold for MLP is described. In the X axis the sample size is shown in the Y axis shows the True

positive rate is given. The graph is plotted based on the instance class value. The below figure 4 shows the threshold curve graph for the class value YES using the MLP.

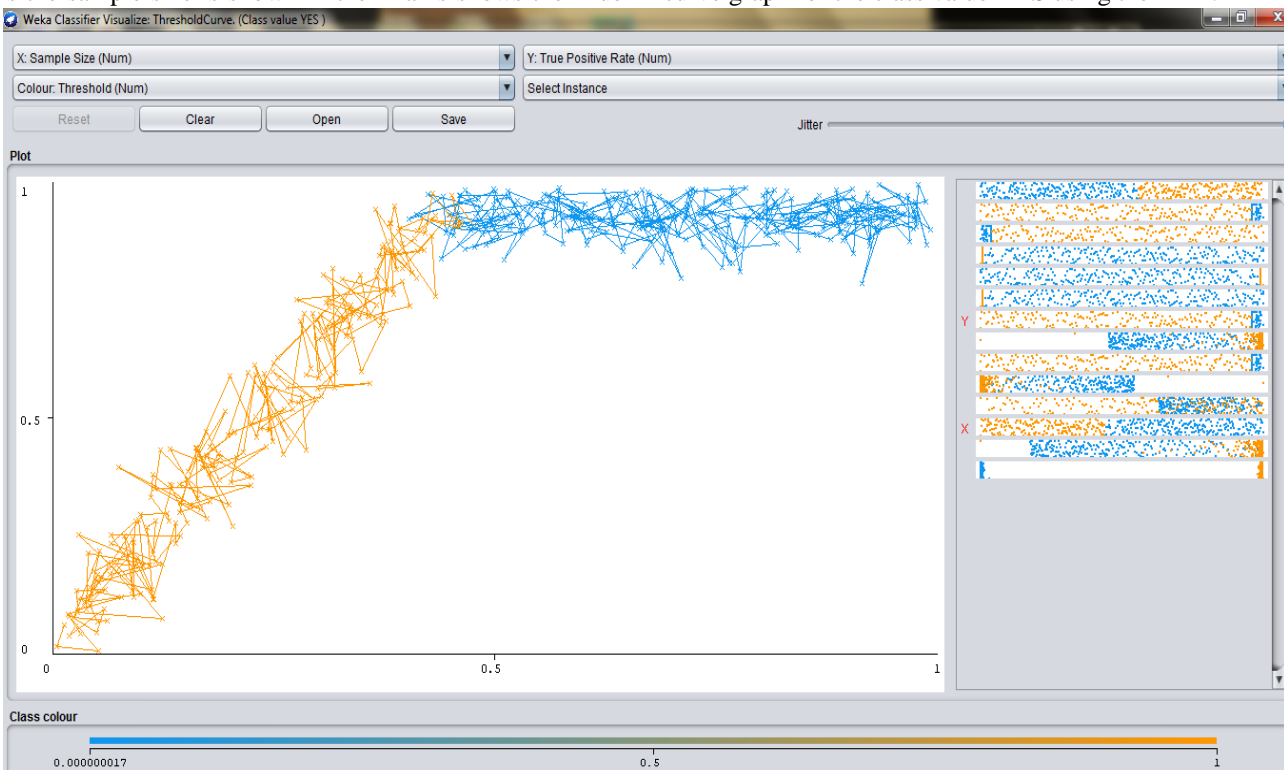


Figure 4: Threshold curve graph

Let's, in the above equations sensitivity, specificity and accuracy are calculated. Based on these values the diagnostic test is correctly identified. The true positive rates of the sensitivity represent the correctly identified pin points of a test and shows how goodness of the test. Next the specificity shows the normal diagnostic that is negative

condition of the test. The comparative analysis of various classifier algorithms are explained in below table 1.

Table 1: Comparative analysis of classifiers

AUTHOR	METHODS	ACCURACY	SENSITIVITY	SPECIFICITY
Koc et al. 2014	LR	65	68	69
Hagan et al.1996	NN	83	84	85
Saren sanjay 2017	NBN	87	80	73
Proposed Method	MLP	97.9	98	97

Whereas accuracy is to identify the true result of the entire population either true positive or true negative. It deals the actual degree of a test diagnosis condition. A high sensitivity test is for capturing all the possible positive conditions of a test. Thus sensitivity is used for screening of a disease. [17][18]

The symptoms are normal comparing with other mosquito borne disease. The incubation period of the virus is nearly 2 to 7 days. Once a mosquito bites a person it takes around 5 days to develop a virus in human body. So if the symptoms persist more than 7 days the person has to check for early prediction method that is using our proposed method. If they found to be infected that person has to go for a RNA test in which it is completely analyzed [19] [20]. Otherwise no need to go for any other test. In the above graph overall probability of threshold value is given. Figure 5 represent the graphical representation of accuracy, sensitivity and specificity of classifiers.

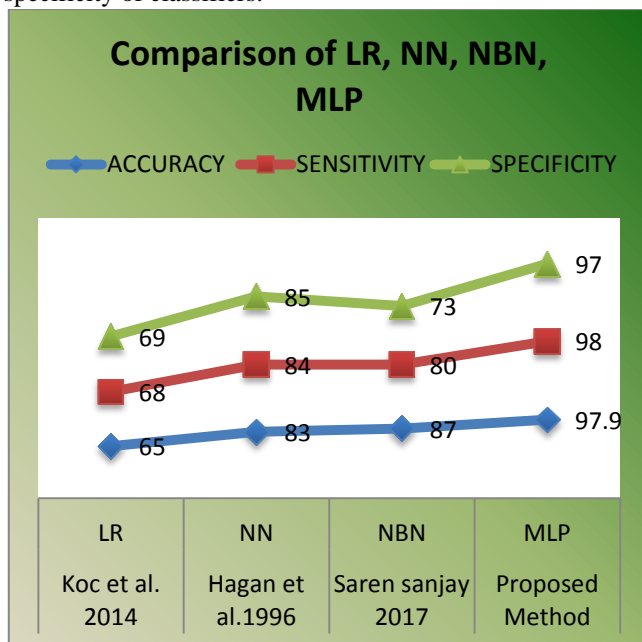


Figure 5: Comparison of LR, NN, NBN, MLP

The accuracy, sensitivity, specificity are compared with different classifiers. Using the LR- Linear Regression the specificity 69%, Sensitivity is 68, accuracy is 65[21]. For NN the specificity 85%, Sensitivity is 84, accuracy is 83 [22] and for NBN the specificity 73%, Sensitivity is 80, accuracy is 87[20]. Finally the MLP our proposed classifier gives the result of specificity 97%, Sensitivity is 98, and accuracy is 97.9.

V. CONCLUSION AND FUTURE ENHANCEMENT

Zika virus is deemed as a globalized emergency, so every country affected with Zika is to take a preventive measure from avoiding other people to get infected. Here a multilayer perceptron algorithm is used which provides a maximum of 97% of accuracy comparing with other. A synthetic dataset is applied in Weka tool and using MLP classifier algorithm the accuracy level is high. If they travelled to risk prone area where Zika infection found the person has to check whether infected or not. Then the RNA test is taken for the infected for the person to know the level of infection. In future the attributes of the disease should be improved in order to get more accurate prediction result. The prediction system is not only for Zika virus but also for other diseases for predicting at its early stage. Disease such as heart disease, arthritis and other chronic disease every thing are in need of early prediction stage so it is better to apply this methodology in a better way to get accurate result.

REFERENCES

1. Sarmiento-Ospina A, Vsquez-Serna H, Jimenez-Canizales CE, Villamil-Gmez WE, Rodriguez-Morales AJ. Zika virus associated deaths in Colombia. *Lancet Infect Dis.* 2016; 16: 523-524.
2. World Health Organization, Zika outbreak: WHO's global emergency response plan. 2016.
3. Musso D, Roche C, Robin E, Nhan T, Teissier A, Cao-Lormeau VM. Potential sexual transmission of Zika virus. *Emerg Infect Dis.* 2015;21:359- 361.
4. Paixao ES, Barreto F, Teixeira GM, Costa CM, Rodrigues L. History, epidemiology, and clinical manifestations of Zika: A systematic review. *Am J Public Health.* 2016;106:606-612.
5. Nishiura H, Mizumoto K, Rock KS, Yasuda Y, Kinoshita R, Miyamatsu Y. A theoretical estimate of the risk of microcephaly during pregnancy with Zika virus infection. *Epidemics.* 2016;15:66-70.
6. Mourya, Devendra T., et al. "Zika virus: Indian perspectives." *The Indian journal of medical research* 143.5 (2016): 553.
7. Bhardwaj, Sumit, Mangesh D. Gokhale, and Devendra T. Mourya. "Zika virus: Current concerns in India." *The Indian journal of medical research* 146.5 (2017): 572
8. Petersen E, Wilson ME, Touch S, et al. Rapid spread of Zika virus in the Americas - Implications for public health preparedness for mass gatherings at the 2016 Brazil Olympic Games. *Int J Infect Dis.* 2016;44:11-15.
9. Amadin, F. I., and M. E. Bello. "Prediction of yellow fever using multilayer perceptron neural network classifier." *Journal of Emerging Trends in Engineering and Applied Sciences* 9.6 (2018): 282-286.
10. Kadri, S. M., and Kausar Rehana Saleem-ur-Rehman. "Zika virus disease (ZIKV) Infographics, language-tailored: Innovations and Good Practices in Integrated Disease Surveillance Programme Kashmir, India 2016." *J Mol Pathol Epidemiol* 1 (2016): 1.
11. Pal, Sankar K., and Sushmita Mitra. "Multilayer perceptron, fuzzy sets, and classification." *IEEE Transactions on neural networks* 3.5 (1992): 683-697.
12. Keller, James M., and Hossein Tahani. "Backpropagation

PREDICTION OF ZIKA VIRUS BY MULTILAYER PERCEPTRON NEURAL NETWORK (MLPNN) USING CLOUD

- neural networks for fuzzy logic." *Information Sciences* 62.3 (1992): 205-221.
13. Mansourkhaki, Ali, Mohammadjavad Berangi, and Majid Haghiri. "Comparative Application of Radial Basis Function and Multilayer Perceptron Neural Networks to Predict Traffic Noise Pollution in Tehran Roads." *Journal of Ecological Engineering* 19.1 (2018).
 14. Memarian, Hadi, and Siva Kumar Balasundram. "Comparison between multi-layer perceptron and radial basis function networks for sediment load estimation in a tropical watershed." *Journal of Water Resource and Protection* 4.10 (2012): 870.
 15. Koskela, Timo, et al. "Time series prediction with multilayer perceptron, FIR and Elman neural networks." *Proceedings of the World Congress on Neural Networks*. INNS Press San Diego, USA, 1996.
 16. Tang, Jiexiong, Chenwei Deng, and Guang-Bin Huang. "Extreme learning machine for multilayer perceptron." *IEEE transactions on neural networks and learning systems* 27.4 (2016): 809-821.
 17. Mahalakshmi, B., and G. Suseendran. "An Analysis of Cloud Computing Issues on Data Integrity, Privacy and Its Current Solutions", *Advances in Intelligent Systems and Computing* 839, 2019, pp.467-482.
 18. Mahalakshmi, B., and G. Suseendran. "Effectuation of Secure Authorized Deduplication in Hybrid Cloud." *Indian Journal of Science and Technology*, Vol.9(25),(2016), pp.1-7.
 19. Ko, M., and A. Barkana. 2014. "Application of Linear Regression Classification to Low-Dimensional Datasets." *Neurocomputing* 131:pp. 331–335.
 20. Sareen, Sanjay, Sandeep K. Sood, and Sunil Kumar Gupta. "Secure internet of things-based cloud framework to control zika virus outbreak." *International journal of technology assessment in health care* 33.1 (2017): 11-18.
 21. Ko, M., and A. Barkana. 2014. "Application of Linear Regression Classification to Low-Dimensional Datasets." *Neurocomputing* 131: 331–335.
 22. Hagan, M. T., H. B. Demuth, and M. H. Beale. 1996. *Neural Network Design*. Boston: Pws Pub.
 23. T. Nathiya, G. Suseendran , 2018."An Effective Way of Cloud Intrusion Detection System Using Decision tree, Support Vector Machine and Naïve Bayes Algorithm "International Journal of Recent Technology and Engineering", Vol.7(4S2), 2018, pp.38-42..
 24. Mahalakshmi, B., and G. Suseendran. "ZIKA Virus: A Secure System USING NBN Classifier for Predicting and Preventing ZIKA in Cloud", *International Journal of Recent Technology and Engineering*, Vol.7(4S2), 2018, pp.28-32.