Diagnosis classification of dengue fever based on Neural Networks and Genetic algorithms

Martselani Adias Sabara¹, Oman Somantri²*, Heru Nurcahyo³, Nanang Kurnia Achmadi³, Ulfatul Latifah⁴ and Harsono⁵

¹Department of Electrical Engineering, Politeknik Harapan Bersama Tegal, Indonesia
²Department of Informatics, Politeknik Harapan Bersama Tegal, Indonesia
³Department of Pharmacy, Politeknik Harapan Bersama Tegal, Indonesia
⁴Department of Midwifery, Politeknik Harapan Bersama Tegal, Indonesia
⁵Department of Computer Engineering, Politeknik Harapan Bersama Tegal, Indonesia
oman.somantri@poltektegal.ac.id

Abstract. To classify the diagnosis of dengue fever is not easy, an accurate accuracy is needed for doctors and health workers in making a decision related to the diagnosis of dengue fever. In the classification of diagnosis of dengue fever, the Neural Network is used as the model applied. To implement the Neural Network there are several parameters that we must determine such as learning rate and momentum, the problem is the absence of standard guidelines in determining the parameters to be used in this method whom the experimental method is used. For this reason, a method that can solve the problem is needed, with the parameters obtained can be optimized. The solution that can be applied with applying the Genetic Algorithm (GA) on the Neural Network, in order to optimize the learning rate parameter value and momentum. The results obtained are apparently the application of optimization techniques with Genetic Algorithms which can make it easier to find parameter values optimally and can increase the accuracy value in the Neural Network algorithm, thus the model obtained can be used for doctors and health workers in determining the classification of dengue fever diagnosis.

1. Introduction
Infection of dengue virus belonging to the genus of the flavivirus is caused by Dengue Hemorrhagic Fever. Spread of dengue virus through the main cause of the aedes aegypti mosquito. Protein E interacts with the receptor with cellular so as to initiate the process of the virus, the amino acid sequence determines the activity of neutralizing antibodies which classify the dengue virus [1]. According to the World Health Organization an estimated 500-100 million cases of Dengue Fever are recorded worldwide, two-fifths of the world's population is at risk of developing this type of disease and more than one hundred countries are affected by dengue fever. Since 1950, more than 500,000 hospitals have dengue fever cases, it is estimated that around 70,000 deaths in children - children due to dengue fever and about 64 out of every 1,000 children - are infected with dengue fever [2]. Infection mostly occurs in endemic areas, including children who have been infected at least once at the beginning of the decade of life. The symptoms of primary infection are largely unclear, although some develop into a non-typical fever with other symptoms such as headache, retro orbital pain,
muscle aches and sometimes bleeding. Dengue Fever is spread through vectors, mainly aedes aegypti mosquitoes. The Dengue virus genome encodes 10 gene products: C (capsid), pRNA (RNA), E (envelope), and on-structural proteins including NS-1, NS-2A, NS-2B, NS-3, NS-4A, NS-4B, and NS-5. Protein E interacts with cellular receptors as initiate the process of virus entry, the amino acid sequence determines the neutralizing activity of antibodies that classify the Dengue virus (DEN) into 4 serotypes: DEN-1, 2, 3 and 4.

In previous studies on the detection of previous dengue fever, Zuriani and Yuhanis (2011) conducted a study on Prediction of blood Fever, from his research concluded that the accuracy rate was 91.07% [3]. Another study by Suhaeri, S et al in 2014 examined the decision support system for the diagnosis of DHF by using the Back Propagation Method to produce an accuracy of 92.85% [4]. Another study Supriyadi, D et al in 2011 examined the Information System for Dissemination of dengue fever using backpropagation neural network method and produced an accuracy of 88.23% [5].

Neural Network (NN) is an attempt to mimic the function of the human brain. The human brain is believed to consist of millions of small processing units, called neurons, which work in parallel. Neurons are connected to each other through neuron connections. Each individual neuron takes input from a set of neurons. This then processes the input and passes the output to a set of neurons. Output is collected by other neurons for further processing. The human brain is a complex network of neurons in which connections remain breaking and forming. Many models similar to the human brain have been proposed[6]. The advantage of Neural Network is its ability to do learning based on data used for training, be able to do self organization or make representations of the information it receives and has a real time operation in the sense that Artificial Neural Network can perform calculations in parallel and have high tolerant fault [7]. Besides having the advantages of neural networks have disadvantages, including requiring very large data for training and having slow convergence so that the backpropagation algorithm is very dependent on initial parameters such as number of inputs, hidden nodes, output, learning rate and connection weights in the network. In implementing the Neural Network there are several parameters that we must determine. So far there are no standard guidelines in determining the Neural Network parameters, so the method used is an experimental method. Therefore we need a method that can solve these problems, so that the implementation of the Neural Network can be more efficient.

Genetic Algorithms have several advantages such as being able to use for discrete or continuous variables, can be used for large variables, the final results are some optimum variables, not just one solution, optimization is done by encoding variables and can be used in numerical data, experimental data, or analytic function [8][9]. Genetic Algorithm is a reliable optimization method so that it can be used to determine the optimal control parameter values for a particular process [10]. With the implementation of genetic algorithms on Neural Networks, it is expected to accelerate the process of obtaining appropriate and optimal parameter values so that it can increase the level of accuracy in the classification of dengue fever diagnoses.

Based on the background above, the problems that occur in this study are still difficult in determining the diagnosis of patients whether they contract dengue disease or not based on the indication they experienced. Selection of learning rate parameter values and momentum for classification of dengue fever diagnoses with Neural Network parameter selection is still not optimal. The purpose of this study is to obtain the best model to apply Genetic Algorithm (GA) on Neural Network (NN) so that it can optimize the learning rate parameter and momentum values for the classification of dengue fever diagnosis so that the resulting accuracy level becomes better.

2. Method
2.1. Data Collection
This study uses data and sources from the Hospital and doctor Nurhayati as general practitioner in 2010 s.d 2015. The data used in this study include; number of experimental data 170 data with attributes such as gender, body temperature, hemoglobin, platelets, hematocrit, diagnosis.
2.2. Preprocessing Data
At this stage, data normalization and selection of variables relevant to the research are carried out. Data obtained for this study were 170 records of examination of Dengue Fever either infected or not infected. The data used in this study is the dataset obtained from Kardinah Hospital in Tegal City by using several attributes.

| Gender | Temperature | Trombosites | Hematocrites | Diagnosis |
|--------|-------------|-------------|--------------|-----------|
| 2      | 36          | 18000       | 30           | Sakit     |
| 1      | 35.2        | 21000       | 36           | Sakit     |
| 1      | 36.2        | 22000       | 28           | Sakit     |
| 2      | 37          | 32000       | 37           | Sakit     |
| 1      | 37          | 33000       | 24           | Sakit     |
| 2      | 36          | 38000       | 42           | Sakit     |
| 1      | 36          | 39000       | 49           | Sakit     |
| 2      | 36.5        | 39000       | 32           | Sakit     |
| 2      | 37          | 40000       | 36           | Sakit     |
| 2      | 36          | 42000       | 40           | Sakit     |

2.3. Proposed Method
The method used is the Neural Network method based on Genetic algorithms to improve accuracy in detecting dengue fever. Each algorithm will use 170 data data, some of them are used as 80% training data as many as 153 data and some as 20% testing as many as 17 data. Calculations with each algorithm will be repeated several times to get the best parameter magnitude.

![Figure 1. Research design](image-url)
Evaluation is done by comparing the level of accuracy of neural networks classically with neural networks that have been optimized for parameter values. Validation is done using Cross Validation so that a measure of the level of accuracy is obtained. This study uses a type of experimental research. The proposed method is the use of Neural Network (NN) Backpropagation to predict the classification of diagnosis of dengue fever. To increase the level of accuracy of calcification then apply a genetic algorithm (GA).

3. Result and Discussion
3.1. Experimental Results and Model Testing
The results of this study are in the form of data processing collected in accordance with the proposed model, namely Neural Network and Genetic Algorithm. The proposed model is in the classification process of detection of dengue fever as figure 3.

![Figure 2. The Neural Network Model Architecture](image)

In this research, the value of training cycles is determined by testing the value of the range 100 to 1000 for the Learning Rate parameter value and the momentum is given a value of 0.3 for Learning Rate and 0.2 for Momentum.

| Training Cycles | Learning Rate | Momentum | Akurasi   |
|-----------------|---------------|----------|-----------|
| 200             | 0.2           | 0.1      | 97.06%    |
| 200             | 0.2           | 0.2      | 97.06%    |
| 200             | 0.2           | 0.3      | 97.06%    |
| 200             | 0.2           | 0.4      | 97.06%    |
| 200             | 0.2           | 0.5      | 97.06%    |
| 200             | 0.2           | 0.6      | 96.47%    |
| 200             | 0.2           | 0.7      | 96.47%    |
| 200             | 0.2           | 0.8      | 95.88%    |
| 200             | 0.2           | 0.9      | 96.47%    |
| 200             | 0.2           | 1.0      | 48.82%    |

Based on the above experiment, the number of neurons in Hidden Layer 4, 200 for the Neural Network parameter is selected for the values of Training Cycles, 0.2 Learning Rates and 0.1 Momentum and Accuracy which are produced: 97.06% +/- 3.95%.

Some text.
3.2. Results of Proposed Method

After obtaining the best neural network model that produces 97.06% accuracy rate, the next step is to optimize the neural network model proposed by using the Genetic algorithm. By using Rapid Miner tools, a number of combinations are produced from the parameters of the training cycle, the learning rate and the momentum generated. After the experiment was obtained 484 populations. From a combination of neural network parameter optimization with genetic algorithms, the best parameter combination is measured with the highest level of drainage or with the highest fitness value. Based on the experiments that have been done that is optimizing the best parameter values of the neural network by using genetic algorithms to get the results of the predicted value of dengue fever that is equal to 98.24%. To show a comparison of the level of accuracy of the proposed model, namely the Neural Network and genetic algorithms, then compared with neural network algorithm models without genetic algorithms the results are as in the table 3.

| Model                        | Accuracy  |
|------------------------------|-----------|
| Neural Network               | 97.06%    |
| Neural Network + Genetic Algorithm | 98.24%    |

Figure 3. Evaluation model

Based on the table obtained for the level of accuracy in the classification of the determination of whether or not dengue fever patients get the accuracy level of Neural Network = 97.06%. and Neural Network and Genetic Algorithm = 98.24% From these results it can be concluded that the proposed model of the Neural Network accuracy level is better compared to Neural Networks and genetic algorithms.

4. Conclusion

This research was carried out using Neural Network algorithm and genetic algorithms of patient data taken from hospitals that did not exist or did not have Dengue Fever. Parameters generated to get values from Accuracy, Precision and Recall that can be used using Neural Networks and algorithms can be accepted. Accuracy is 98.24% which originates at 97.06%, thus increasing accuracy. By using the Genetic algorithm, optimization of the best neural network parameter values is the level of learning.
and momentum can be obtained from the results that occur. Although the Neural Network algorithm model and the algorithms that are used provide good results, there are some things that can be used for future research, which allows more with more data, but it can be used as part of the health work.

References
[1] Srikiatkhachorn A, 2009 Plasma leakage in dengue hemorrhagic fever Thrombosis and haemostasis 102.
[2] WHO, 1997, Dengue Hemorrhagic Fever: diagnosis, treatment, prevention and control. 2nd ed.
[3] Yusof Y and Mustaffa Z, 2011 Dengue outbreak prediction: A least squares support vector machines approach Int. J. Comput. Theory Eng. 3, 4 p. 489.
[4] Suhaeri S Tundjungsari V Qomariyah Q and Pamuji S, 2014 Sistem Pendukung Keputusan Untuk Diagnosis Penyakit DBD Menggunakan Metode Back Propagation Jaringan Syaraf Tiruan in Seminar Nasional Informatika Medis (SNIMed) p. 27–37.
[5] Supriyadi D Kusworo A and Sarwoko E., 2012, Sistem Informasi Penyebaran Penyakit Demam Berdarah Menggunakan Metode Jaringan Syaraf Tiruan Backpropagation, Universitas Diponegoro.
[6] Fausett L V., 1994 Fundamentals of Neural Networks: Architectures, Algorithms, and Applications p. 461.
[7] Liao G-C, 2011 An Improved Fuzzy Neural Networks Approach for Short-Term Electrical Load Forecasting Control Conf. (ASCC), 2011 8th Asian p. 596–601.
[8] Karimi H and Yousefi F, 2012 Application of artificial neural network-genetic algorithm (ANN-GA) to correlation of density in nanofluids Fluid Phase Equilib. 336 p. 79–83.
[9] LIU G WANG L CHEN G and HUA S, 2011 Parameters Optimization of Plasma Hardening Process Using Genetic Algorithm and Neural Network J. Iron Steel Res. Int. 18, 12 p. 57–64.
[10] Melanie M, 1996 An introduction to genetic algorithms Cambridge, Massachusetts London, England, .. p. 162.