International Journal of Innovative Research in Advanced Engineering (IJIRAE) ISSN: 2349-2163
Issue 03, Volume 4 (March 2017) www.ijirae.com

PREDICTION THE NUMBER OF PATIENTS AT DENGUE HEMORRHAGIC FEVER CASES USING ADAPTIVE NEURAL FUZZY INFERENCE SYSTEM (ANFIS)

Basuki Rachmat [a)], Oky Dwi Nurhayati [b)], Suryono [c)]
Master of Information Systems Diponegoro University [a)]
Departement of Computer Engineering Diponegoro University [b)]
Department of Physics, Faculty of Science and Mathematics, Diponegoro University [c)]

Abstract*— Dengue Hemorrhagic Fever is one of the dangerous infectious diseases that can cause death within a short time and often cause epidemic. The spread of dengue fever outbreaks globally with frequency levels tend to be higher during the period of last 50 years gave rise to an idea that systematic prevention. The purpose of this paper was to design an application to predict the number of dengue hemorrhagic fever patients with ANFIS method. Weather factors such as air humidity, air temperature, rainfall and number of rain days is used as the factors that influence the incidence of dengue hemorrhagic fever. In this paper using three methods for establishment of FIS: Grid Partition, Substractive Clustering and Fuzzy C Means. By simulating three methods for maximum predicted results, it was found that the ANFIS method with Grid Partition as the establishment of FIS is the best model to generate value with the smallest RMSE testing is 0.71. It indicates That ANFIS models is well proven to be used in predicting The cases of dengue fever.

***Keywords* — dengue hemorrhagic fever, weather, ANFIS, predictition, RMSE**

I. INTRODUCTION

Various methods have been used to predict the incidence of dengue fever. These methods include Artificial Neural Network (ANN) [1], fuzzy systems [2], entropy algorithm [3] and many more. Artificial Neural Network (ANN) has an advantage in predictive nonlinear, strong in parallel processing and the ability to tolerate errors, but an Artificial Neural Network (ANN) also has a number of limitations, including a lack of ability to perform an operation - numeric operation with high precision, the operation algorithm arithmetic, logic operations, and symbolic operation and duration of the training process that can sometimes take a very long time for a substantial amount of data [1].

Neuro Fuzzy is a combination of the two systems, namely the system of fuzzy logic and neural network. Neuro fuzzy system based on fuzzy inference systems that are trained using learning algorithm derived from the neural network system. Thus, neuro fuzzy system has all the advantages possessed by the fuzzy inference system and neural network systems. Fuzzy inference system can translate the knowledge of experts in the form of rules, but it usually takes a long time to set membership functions. Therefore required learning techniques of artificial neural networks to automate the process so as to reduce the search time. From the ability to learn the neuro fuzzy systems are often referred to as ANFIS (Adaptive Neuro Fuzzy Inference Systems) [4]. The use of ANFIS one of which is very useful in terms of stock market predictions, in Istanbul Turkey ANFIS models used to predict the return on the stock price index by using six macroeconomic variables and the three index as input parameters. The experimental results show that the model successfully predicted earnings ANFIS stock market in Turkey Istanbul Stock Exchange with 98.3% accuracy rate [5].

ANFIS method used to predict the temperature of the tool CNC (Computer Numerical Control). Improper temperature can have a significant effect on the accuracy of CNC machine tools. Mistakes come from the deformation temperature and machine elements caused by heat sources in the engine structure or of changes in ambient temperature. The results showed that the method is superior in terms of accuracy ANFIS prediction capability with fewer rules. This method can improve the accuracy and robustness of the system error compensation temperature affects the accuracy of CNC tool [6]. The world's total population of approximately 7.5 billion people by the amount of 2/5 of the population who are in tropical and subtropical regions at risk of disease transmission Dengue Hemorrhagic Fever [3].

Dengue Hemorrhagic Fever (DHF) endemic in more than 100 countries in the African region, the Americas, the Eastern Mediterranean, Southeast Asia and the Western Pacific. Southeast Asia and the Western Pacific are the worst affected areas due to the impact of infectious diseases [7]. Dengue viruses are transmitted by female Aedes mosquitoes through bloodfeeding from human hosts. After the incubation of 4–10 days, dengue infection can produce a wide spectrum of illness which range from asymptomatic or subclinical to severe hemorrhagic manifestations, plasma leakage, and severe organ impairment with potentially lethal complication. Since there are no specific antiviral medicines treating dengue or vaccines preventing dengue, the most effective way to manage this disease is through preventive medicine and control of vector populations [8]. Changes in demographics, urbanization, water supplies are inadequate, the migration, and the introduction of new territory through international travel led to a significant increase of the outbreaks of dengue and about 3.6 billion people today have the potential risk of outbreaks, in addition to the climate change factor also affects a region [9].

II. ADAPTIVE NEURO FUZZY INFERENCE SYSTEM (ANFIS)

Adaptive neuro fuzzy inference system (ANFIS) model is an architecture that consists of input–output variables and a fuzzy rule base of the Takagi–Sugeno type. For simplification, it is assumed that the framework of ANFIS has two inputs x, y and one output z. Then, the corresponding rule set with two fuzzy if–then rules for a first-order Sugeno fuzzy model can be expressed as shown in Eq. (1). Entries are evaluated by linguistic variables (A1, B1). A linear combination of the input values with a constant term (r) is used to obtain each rule result.

$$\begin{aligned} &Rule1 : \text{If } x \text{ is } A_1 \text{ and } y \text{ is } B_1 \text{ then } z_1 = p_1x+q_1y+r_1 \\ &Rule2 : \text{If } x \text{ is } A_2 \text{ and } y \text{ is } B_2 \text{ then } z_2 = p_2x+q_2y+r_2 \end{aligned} \tag{1}$$

Fig. 1 illustrates the ANFIS architecture, which contains five layers with different functions. The function of each layer is described as follows. Layer 1: The main purpose of layer 1 is to map input variables (x and y) into fuzzy sets, say A = $\{A_1, A_2, B_1, B_2\}$ through the process of fuzzification. Each node in this layer is a square node with anode functions for generating membership grades. A and B are linguistis labels (such as ‘‘low” and ‘‘high”) characterized by different membership functions such as generalized bell, sigmoid or triangular. Layer 2: In this layer, firing strength will be used after combining the fuzzy sets of each input. The Π-norm operator performing The fuzzy conjunction (‘‘and”), is used to obtain the output. Layer 3: The main purpose of this layer is to calculate the ratio of ith rute is firing strength to the sum of all firing strength. Layer 4: In The layer, the output from the previous layer is multiplied with the function of Sugeno fuzzy rule. Layer 5: There is only one node in this layer. This single node computes the sum of all outputs of each rule from the previous layer. Then, the weighted averaged method is used to perform the process of defuzzification, which transforms the fuzzy result into a crisp output.

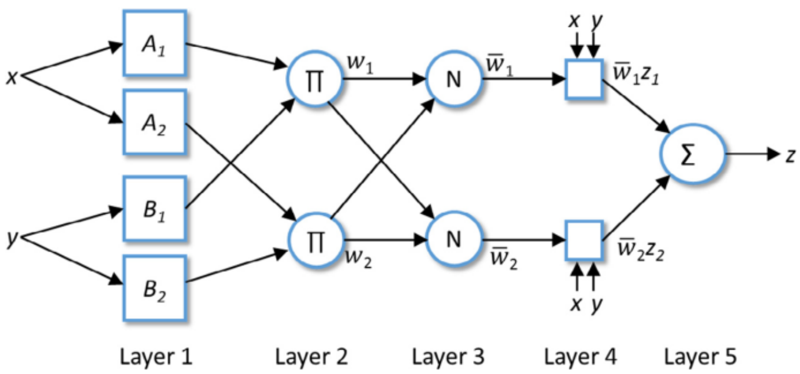

Fig. 1 The ANFIS structure

The parameters in ANFIS can be classified in two types: nonlinear parameters in the premise part and linear parameters in the consequent part. There are several methods to optimize these parameters; such as gradient descent method or steepest descent method [10]. However, in the ANFIS, there is a more efficient method which is hybrid learning method. In the forward-pass part, the output of node will transfer forward to layer 4 and the method of least squares estimate is applied to adjust the parameters of consequent part. In the backward-pass part, the error signals are transferred backward to layer 1 and the method of gradient descent si used to update the parameters of the premise part. The process needs iteration and each iteration denotes an epoch. The training data set is utilized in the network learning. To minimize the eror between the desired and the real output, there is a critical adjustment of the weights of the connections during the learning process [4].

III. METHOD

The data in this study using secondary data, the number of patients with dengue hemorrhagic fever cases are cases that have been clinically tested clinically in health laboratories that have been reported on a monthly basis in Pekalongan District Health Office from 2010 to 2015. At the same time interval is also collected Pekalongan weather data from the Meteorology and Geophysics Semarang, weather data were used to predict the number of cases of dengue hemorrhagic fever is a data humidity, temperature, rainfall and number of rainy days. Data on the predicted number of cases of dengue hemorrhagic fever is divided into two, namely the training data (training) and test data (testing). To obtain optimal results to vary the ratio between the training data and testing the data, in this study the variation of data is done by dividing the training data ranging from 60% to 95%.

The data processing in this study conducted FIS forming process using method 3 as the comparison is Grid Pratition, substractive Clustering, Fuzzy C-Means and the prediction process using Adaptive Neuro Fuzzy Inference System (ANFIS). FIS modeling method produces structures Grid Partitioning FIS Sugeno types of training data sets. Grid Partitioning involves eight types of membership functions (trimf, trapmf, gbellmf, gaussmf, gauss2mf, pimf, dsigmf, psigmf). In modeling Substructive Clustering (SC) to obtain the optimal structure of ANFIS to predict the number of dengue cases was obtained by varying the parameters manually grouping. Various grades Range Influence (RI) changes between 0.1 and 0.3 with the difference value 0.01. Fuzzy C-Means used as an alternative modeling ANFIS FIS to generate optimal structure, by varying the number of clusters and iterations of the FCM itself. Number of cluster piloted ranging from 2 to 40, while the maximum iteration is set at 100. For the exponent is set unchanged at number 2, because the change has no effect exponent.

The stages of the predicted number of cases of dengue hemorrhagic fever using ANFIS method in this study are as follows:

1. *Enter data monthly number of cases of dengue hemorrhagic fever, temperature, rainfall and humidity at the same time as input parameters,*
2. *Do the training process and testing data has been normalized by the percentage of training data that produces an optimal RMSE values,*
3. *Predicting the number of patients with dengue hemorrhagic fever cases using FIS formation model that has the smallest RMSE value,*
4. *Then the result predicted number of cases of dengue fever patients.*

Fig. 2 Flowchart prediction using ANFIS

IV. RESULT

This study resulted in a predictive modeling cases of dengue hemorrhagic fever with data obtained from Pekalongan District Health Office and the Meteorology and Geophysics Semarang. Prediction of cases of dengue hemorrhagic fever using Adaptive Neuro Fuzzy Inference System based on the number of cases of dengue fever, the average temperature, rainfall, number of days of rain and humidity. From the results of the predicted number of cases of dengue fever will be obtained endemic level. Predictive modeling of dengue fever cases this is done by a monthly time intervals.

ANFIS parameters are not the same for each condition. So to predictions, necessary to find appropriate ANFIS parameters and produce the best output. In general, the models predicted number of cases of dengue fever using ANFIS method consists of two processes, namely the processes of training and testing process.

A. ANFIS Model Building

In the process of modeling ANFIS, functions associated with ANFIS in Matlab software is used to generate a model program. In the process of model building, ANFIS model parameters determined in the learning phase at a particular epoch for the smallest error checking. The test data set is used to verify the accuracy and effectiveness of the trained model.

1. Grid Partitioning

Grid Partitioning method produces structures FIS Sugeno type of training data sets. Grid Partitioning method involves eight types of membership functions (trimf, trapmf, gbellmf, gaussmf, gauss2mf, pimf, dsigmf, psigmf). The amount of the membership function can be determined in relation to each input. Because the structure is used Sugeno FIS, only one output can be used. Modeling results are shown in Table 1.

TABLE I - THE FEATURES AND PERFORMANCE RESULTS OF ALTERNATIVE ANFIS-GP MODELS

MODEL NAME	% TRAINING DATA	NUMBE OF MF	TYPE MF	TYPE OF OUTPUT MF	EPOCH	MSE		RMSE	
						TRAINING	TESTING	TRAINING	TESTING
GP 1	95	3	trimf	constant	20	1.85	1.5	1.36	1.22
GP 2	95	3	gbellmf	constant	20	0.91	0.50	0.95	0.71
GP 3	95	3	dsigmf	constant	50	0.65	2.25	0.81	1.5
GP 4	95	3	psigmf	constant	50	0.65	2.25	0.81	1.5
GP 5	95	3	pimf	constant	50	2.13	3.75	1.46	1.93
GP 6	95	3	trapmf	constant	50	2.29	4.74	1.51	1.93
GP 7	90	3	trimf	constant	10	1.95	4.57	1.39	2.1
GP 8	80	2	trapmf	constant	10	4.63	8.43	2.15	2.9
GP 9	70	3	trapmf	constant	10	2.02	8.19	1.42	2.86
GP 10	60	3	gbellmf	constant	20	0.02	133.64	0.15	11.56

To search for the best model performed a number of experiments by modifying existing variables in the model Grid Partitioning. In this study, the table displays the test results only 10 models that have the MSE and RMSE values are small. Data is divided into two, or training and the training data and test or testing. After testing, the percentage of training data that produces the best model is 95 percent of the data. The greater the percentage of training data given outcome tends to be better. From the test results that have been made optimal results obtained are models of GP 2, where the model of GP 2 value testing the smallest RMSE is 0.71.

2. SUB-CLUSTERING METHOD

ANFIS structure optimum based methods substractive Clustering grouping is achieved by varying the parameters manually. Various influences (RI) changes from 0.10 up to 0.3 with the difference in value of 0.01. Table 2 shows the results of the simulation of alternative models, it can be concluded from the table that the optimal subclustering is predictive models with SC1 and SC2 models for these two models have the value of testing the most minimal RMSE is 2.56. Modeling results are shown in table 2.

TABLE 2- THE FEATURES AND PERFORMANCE RESULTS OF ALTERNATIVE ANFIS-SC MODELS

MODEL NAME	% TRAINING DATA	IR	EPOCH	MSE		RMSE	
				TRAINING	TESTING	TRAINING	TESTING
SC1	90	0.10	10	0	6.57	0	2.56
SC2	90	0.11	30	0	6.57	0	2.56
SC3	90	0.12	30	0	6.71	0	2.59
SC4	90	0.13	40	0	6.71	0	2.59
SC5	90	0.14	40	0	6.71	0	2.59
SC6	90	0.15	61	0	6.71	0	2.59
SC7	90	0.16	83	0	6.71	0	2.59
SC8	90	0.17	205	0	6.71	0	2.59

Based on table 2, when the value of IR (Influence Radius) rises 0.01 after 0.11 it can be observed that the same testing RMSE values of 2.59 though IR values changed by adjusting the value of the Epoch.

3. FUZZY C-MEANS METHOD

Simulation and performance results for alternative models are done to find the optimum FIS structure is given in table 3. As can be seen in this table, it cannot be concluded that the result of better performance is obtained by the number of clusters to more. This can be seen FCM1 and FCM3 models have the lowest value of RMSE testing despite its cluster number is not the highest.

TABLE 3 - THE FEATURES AND PERFORMANCE RESULTS OF ALTERNATIVE ANFIS-FCM MODELS

MODEL NAME	% TRAINING DATA	NUMBER OF CLUSTER	MAX ITERATION	EPOCH	MSE		RMSE	
					TRAINING	TESTING	TRAINING	TESTING
FCM1	90	4	30	14	2.35	6.71	1.53	2.59
FCM2	95	3	20	10	3.22	7.5	1.79	2.73
FCM3	95	6	100	48	0.89	6.5	0.94	2.54
FCM4	90	7	100	15	0.93	7.28	0.96	2.69

Based on some of the experiments that have been carried out by three methods FIS generation of weather data of dengue cases, the model generation method FIS Grid Partitioning with GP2 models have the smallest error rate to be able to predict the number of patients with dengue cases. This is evidenced by the value of the smallest RMSE testingnya is 0.71.

B. Prediction by Using ANFIS Method

In this research, the prediction for August to December 2015. The prediction is done comparing the data actually exists. From the results obtained prediction RMSE value of 0.632455532. As for the comparison of actual data with the prediction data for August to December 2015 may be seen in Table 4.

TABLE 4 - COMPARISON OF ACTUAL DATA WITH THE DATA PROCESSED ANFIS PREDICTIONS

Month	Actual Data	Prediction	RMSE
August	1	1	0.63
September	6	6	
October	3	3	
November	3	2	
December	0	1	

In table 4 shows a comparison of actual data with the data processed ANFIS predictions. The actual data presented from August - December 2015 with a line of blue. While the prediction data depicted in red. From the graph comparison of predicted results with actual data can be seen that outline the results predicted by ANFIS method already approached the actual value of the number of cases of dengue hemorrhagic fever. Basically, the prediction of time series data with ANFIS method works pretty well.

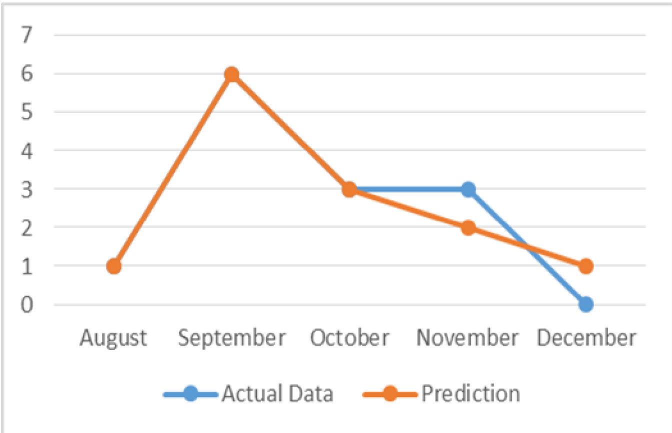

Fig. 3 Comparison of actual data with the data processed ANFIS predictions

V. CONCLUSIONS

From the study predicted the number of patients with dengue fever cases with the method of Adaptive Neural Fuzzy Inference System (ANFIS) can be taken some conclusions as follows:

1. *Methods Adaptive Neuro Fuzzy Inference System (ANFIS) can be applied to predict the number of cases of dengue fever sufferers, it can be proven that the average error or difference between the predicted value to the actual data value number of patients with dengue fever cases is relatively small.*
2. *FIS formation of the three methods used in this research is that Grid Partitioning, substractive Clustering and Fuzzy C-Means, Partition Grid method provides the best model, evidenced by the value of testing the smallest RMSE is 0.71.*

REFERENCES

[1] Aburas, H.M., Cetiner B.G., Sari M., 2009, Dengue Confirmed-Cases Prediction: A Neural Network Model, Expert Systems with Applications, 37, 4256–4260.
[2] Torres, C., Barguil S., Melgarejo M. dan Olarte A., 2014, Fuzzy Model Identification Of Dengue Epidemic In Colombia Based On, *Artificial Intelligence in Medicine.* 60 (2014) 41– 51.
[3] Chen, C, & Chang H., 2013, Predicting Dengue Outbreaks Using Approximate Entropy Algorithm and Pattern Recognition. *The Journal of Infection*, 67(1), 65–71.
[4] Jang, J., & Sun, C., 1995, Neuro-fuzzy Modeling and Control, *Proceedings of IEEE,* 83,378–406.

[5] Boyacioglu, M.A., dan Avci D., 2010, An Adaptive Network-Based Fuzzy Inference System (ANFIS) for the prediction of stock market return: The case of the Istanbul Stock Exchange, Expert Systems with Applications Volume 37, Issue 12, December 2010, Pages 7908–7912.

[6] Abdulshahed, A.M., Longstaff, A.P., Fletcher, S., 2014, The application of ANFIS prediction models for thermal error compensation on CNC machine tools, Applied Soft Computing, (2015) 158–168

[7] World Health Organization (WHO), 2013, Comprehensive Guidelines for Prevention and Control of Dengue Haemorraghic Fever. Regional Office for South East Asia.

[8] Al-Muhandis, N., Hunter, P.R., 2011. The value of educational messages embedded in a community-based approach to combat dengue fever: a systematic review and meta-regression analysis. PLoS Negl. Trop. Dis. 5, e1278

[9] Wilder-Smith A, Gubler DJ. (2008). Geographic expansion of dengue: the impact of international travel. Med Clin North Am. 92:1377–1390.

[10] Jang, J.-S.R., 1993. ANFIS: Adaptive Network-based Fuzzy Inference System. *IEEE Transaction on System, Man and Cybernetics, 655-685*.