

Expert System of Dengue Disease Using Artificial Neural Network Classifier

Hamdani Hamdani¹, Zainal Arifin², Anindita Septiarini³

^{1,2,3}*Department of Informatics, Faculty of Engineering, Mulawarman University, Samarinda, Indonesia*

¹hamdani@unmul.ac.id, ²zainal.arifin@unmul.ac.id, ³anindita@unmul.ac.id

Abstract – An expert system has been applied to the classification of dengue fever. Dengue is a severe disease and can be fatal if not diagnosed and treated properly. Headache, muscle aches, fever, and rash are the most prevalent symptoms. Dengue fever is an endemic disease in various South Asian and Southeast Asian nations. Dengue fever (DF), dengue hemorrhagic fever (DHF), and dengue shock syndrome are the three types of dengue (DSS). Currently, these diseases may be classified using a machine learning approach with dengue symptoms as the input data. This study proposes implementing an Artificial Neural Network (ANN) with the Backpropagation (BPNN) algorithm as the classifier to categorize dengue types, divided into three categories: DF, DHF, and DSS. There were 21 attributes in the dataset that represent the dengue symptoms. It was gathered from 110 patients. Cross-validation with k-fold 2, 3, 5, and 10 was applied as the evaluation method. In order to evaluate the BPNN classification method, three parameters were obtained: precision, recall, and accuracy. These were used to justify the most optimal performance. Cross-validation using k-fold 3 produced the best evaluation performance with the learning rate values of 0.1, 0.3, and 0.5, obtaining the precision, recall, and accuracy values of 0.969, 0.967, and 0.967, respectively.

Keywords: Dengue, Expert System, Classification, Backpropagation, Cross Validation.

I. INTRODUCTION

Nowadays, computer technology is being used in all areas, one of which is the health care field, specifically in the form of an expert system [1-2]. In recent years, it has been widely developed. Expert systems were taken out of the need to assist doctors and patients in diagnosing and treating disease, even to alert and notify doctors and patients. This system is still being developed further because it may be linked to clinical decision-making to denote disease and aid doctors in diagnosis. It is a computer application that contains information from one or more human specialists about a specific disease. The expert system assists patients more easily in getting diagnostic results based on the symptoms that occur. Furthermore, it can be utilized at any time, making it

cost-effective. The expert system can be applied to detect dengue fever disease [3].

Dengue fever is an arboviral disease caused by infection with one of the four dengue virus serotypes. According to the World Health Organization (WHO), this disease increases continuously every year, reaching 50 million cases and around 2.5 billion people living in dengue-endemic areas [4]. Dengue fever can cause a variety of symptoms, including fever, headache, muscle aches, and a measles-like rash, often known as fracture fever [5]. The rapid spread of the dengue virus has grown increasingly harmful, and resolving this issue should be regarded as an emergency. Based on the statistics data, the national incidence of DHF in Indonesia increased from 50.8 per 100,000 people in 2015 to 78.9 per 100,000 people in 2016 [6]. The clinical diagnosis can range from symptomatic DF to DHF, the most severe variant, and DSS, the most lethal [4]. Furthermore, dengue fever has risen since 2012, according to the Malaysian Ministry of Health. The Malaysian Ministry of Health produced a study in 2015 that recorded 107,079 cases of dengue fever with 293 deaths, compared to 43,000 cases of dengue fever with 92 deaths in 2013 [7].

Computer technology can be used to classify different forms of dengue illnesses. As input data, several indicators were employed, including the patient's hemoglobin and thrombocyte values and the symptoms reported (headache, fever, joint pain, muscle pain, pain behind the eyeball, and other symptoms). The classification method is required to determine the patient's symptoms without having to consult an expert or doctor. It has been implemented using the machine learning approach [8]. Prior studies used the following techniques to implement the machine learning-based classifying process: Logistic Regression [9], Naive Bayes [10], Support Vector Machine (SVM) [5], K-Nearest Neighbor (KNN) [11], and ANN [12-13].

Prediction of diabetes was applied using the ID3 decision tree with the application of attribute selection Correlation-based Feature Selection (CFS) and Information Gain (IG) [14]. The number of attributes used was 13, obtained from the patient's age, gender, and

symptoms. The test data used were in the form of balanced data and unbalanced data. The use of balanced data in this study improves prediction performance using the ID3 algorithm. The use of Correlation-based Feature Selection and Information Gain attribute selection methods can improve prediction performance using the ID3 algorithm. The two look slightly different, but the highest performance was obtained using Correlation-based Feature Selection with five attributes, namely *gpost*, *glun*, *upost*, *urn*, and *actn*, with an average accuracy of 84.77, an average sensitivity of 87.18, and an average specificity of 82.37. The initial diagnosis of gastric disease was implemented using the Dempster Shafer method [15], which focuses on several gastric disorders such as GERD, peptic ulcers, gastroparesis, and dyspepsia. The data used are 100 test data, and 20 samples were taken. They tested the system's accuracy by conducting interviews with internal medicine doctors, which resulted in an accuracy rate of 94%.

The expert system used the Dempster Shafer to assess the uncertainty of 20 stroke symptoms. The findings of the illness slice yield a percentage of the chance of stroke, hypertension / high blood pressure, fever, and heart disease. In addition, Fuzzy Logic was used to process nine patient medical history data. In this case, combining the two methods in providing a stroke diagnosis based on pain symptoms and patient history. Those methods were evaluated using several metrics, including accuracy, precision, sensitivity (recall), F-measure, and specificity, resulting in an expert system value of 0.786, indicating good expert system performance [16]. Malaria detection employed a Saliency-based Convolutional Neural Network [17]. The scenario contrasts saliency approaches, specifically Region Contrast Saliency, Frequency-tuned Saliency, Spectral Residual, and Histogram Contrast. The frequency-tuned saliency approach outperformed conventional saliency methods in identifying malaria, with an accuracy of 90.32% to 62.67% for the contrast saliency region, 50% for spectral residual saliency, and 79.06% for histogram contrast saliency.

The KNN and Random Forest methods were carried out to predict the Vulnerability Level of Dengue Fever Using. Based on the results, it can be concluded that population density, growth rate, population mobility, rainfall, and wind speed are the most influencing factors of the six factors, with the KNN algorithm RMSE value of 29.26. The results of this study are good enough to be implemented in the real world, and the model shows the most influencing factors are population density, growth rate, population mobility, rainfall, and wind speed. By

utilizing the study results, the government can adjust actions to each level of sub-district vulnerability and pay more attention to the factors that most influence dengue fever, according to the study's results [11].

This study aims to classify dengue disease was accomplished using machine learning with ANN algorithms. The disease was divided into three classes: DF, DHF, and DSS. The symptoms caused by the condition are used as input data. The evaluation was carried out through cross-validation with different k-fold values.

II. METHOD

This section discusses the dataset details used and the classification algorithm utilized in this study. It also describes the procedure for assessing the performance of the ANN classification method. A total of 90 dengue patients were included in the dataset given by Dirgahayu Hospital in Samarinda, Indonesia. The dataset was determined into three classes: DF, DHF, and DSS, with 40, 41, and 9 data, respectively. Each patient's patient code (C), name, address, age, thrombocyte (T), and hemoglobin (H) readings are among the attributes acquired. There were also eighteen various symptoms of the patients recorded and the expert's diagnosis results. Headache (S1), fever (S2), joint pain (S3), muscle soreness (S4), petechiae (S5), maculopapular skin rash (S6), shock (S7), bruising (S8), vomiting (S9), anxiety (S10), constipation (S11), heartburn (S12), diarrhea (S13), red eyes (S14), cough (S15), sore throat (S16), lower jaw discomfort (S17), and nasal cavity inflammation (S18) were among the symptoms recorded. The symptoms that each patient experiences can differ, causing the expert to reach various judgments.

There were two phases in the dengue classification approach: training and testing. Before carrying out the process, the pre-processing was carried out in both phases. Subsequently, the ANN method with the backpropagation (BPNN) algorithm is applied in the learning process, followed by weighted selection. In the testing phase, the classification was implemented to justify the testing data into certain classes. A performance evaluation process was required to assess the classifier's performance. The diagnosis from the expert (actual class) and BPNN method (predicted class) is the input to the evaluation process. Fig. 1 presents the process stages involved in the dengue classification method. Meanwhile, the process details are described as follows.

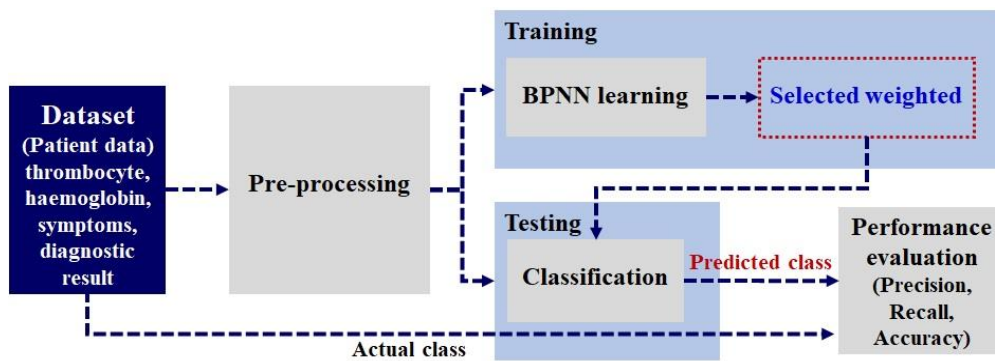


Fig. 1 The process stages of the dengue classification method

A. Pre-processing

This process applied the selection of attributes, which was important in classification and data discretization. The attributes selection was necessary since not all of the qualities of patients play an essential role and are employed as input parameters in the classification process, specifically the patient's name and address. As a result, there were a total of 20 attributes employed, including C, T, H, and S (S1 until S18). Meanwhile, discretization was required to convert text-based symptom data into numeric data that may be used as input in the subsequent process. Those symptoms (S1 until S18) have a value of 1 if the patient has specific symptoms; otherwise, the symptoms value of 0. The result of the pre-processing data is shown in Table I.

B. Learning and Classification

ANN is a computational model that forms based on biological principles. Numerous processing elements are included in its construction (neurons). There is a connection between neurons that allows information received by one neuron to be turned into information received by another neuron. This connection is referred to as a weight relationship. Deboeck and Kohonen present non-linear, multi-layered parallel regression approaches as a collection of mathematical techniques that may be used for signal processing, forecasting, and grouping. These techniques are referred to as non-linear multi-layered parallel regression techniques.

As one of the core components of soft computing, BPNN has been widely used in several sectors of human life for both research and solving technical problems such as forecasting, diagnostics, and pattern recognition [18]. It is widely utilized in different fields of science and engineering nowadays due to its capacity to describe

both linear and nonlinear systems without requiring the use of implicit assumptions, as most classic statistical approaches do. ANN is an information processing system that shares some properties with biological neural networks found in humans. In 1943, Waffen McCulloch and Walter Pitts invented the neural network. Several professionals have built ANN to answer human life difficulties [19]. The design of the BPNN learning algorithm specifies each step connected with the entire process. It consists of several processes, such as initializing the weights and calculating the parameter value at the hidden layer, as illustrated in Fig. 2.

The BPNN algorithm employs a binary sigmoid function with the target output values ranging from 0 to 1. The primary issue with the BPNN is the iteration uncertainty, which must be performed in a relatively long period. These uncertainties have an effect on the value of the epoch procedure used to generate the appropriate iteration conditions. Each researcher has a unique perspective on deciding the value of the parameter. The steps of the BPNN algorithm are identified in [20].

C. Performance Evaluation

This method was designed to evaluate the dengue classification method's performance. Three performance metrics were used to evaluate the classification method's performance: precision, recall, and accuracy based on the confusion matrix multiclass. These parameters are defined using (1)–(3) as follows [21]:

$$precision = \frac{\sum_{k=1}^n N_{ki}}{N_{ii}} \tag{1}$$

$$recall = \frac{\sum_{k=1}^n N_{ik}}{N_{ii}} \tag{2}$$

$$accuracy = \frac{\sum_{i=1}^n N_{ii}}{\sum_{i=1}^n \sum_{j=1}^n N_{ij}} \tag{3}$$

TABEL I
THE EXAMPLE OF PRE-PROCESSED DATA

C	T	H	S1	S2	S3	S4	S5	S6	S7	...	S13	S14	S15	S16	S17	S18	Diagnose
C1	120000	32	1	0	0	1	1	1	0	...	1	1	0	0	1	0	DF
C2	4000	20	1	1	1	1	1	1	1	...	0	0	0	0	1	1	DHF
C3	32000	30	1	0	1	0	1	1	0	...	1	1	0	0	1	0	DSS
C4	166000	38	1	0	1	0	0	1	1	...	0	0	1	0	1	0	DHF
C5	102000	41	1	1	1	0	1	1	1	...	0	0	1	0	0	0	DSS
C6	9500	32	1	1	1	0	0	1	0	...	0	0	0	0	1	1	DF
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
C86	97000	39	1	0	1	1	0	1	0	...	0	0	1	0	1	0	DF
C87	63000	47	1	1	1	1	1	1	0	...	0	0	1	0	0	0	DHF
C88	77000	34	1	1	1	1	0	1	1	...	0	0	1	0	1	0	DSS
C89	43000	49	1	1	1	0	1	1	1	...	1	1	1	1	0	0	DHF
C90	27000	32	1	1	1	1	0	0	0	...	1	1	1	1	0	0	DHF

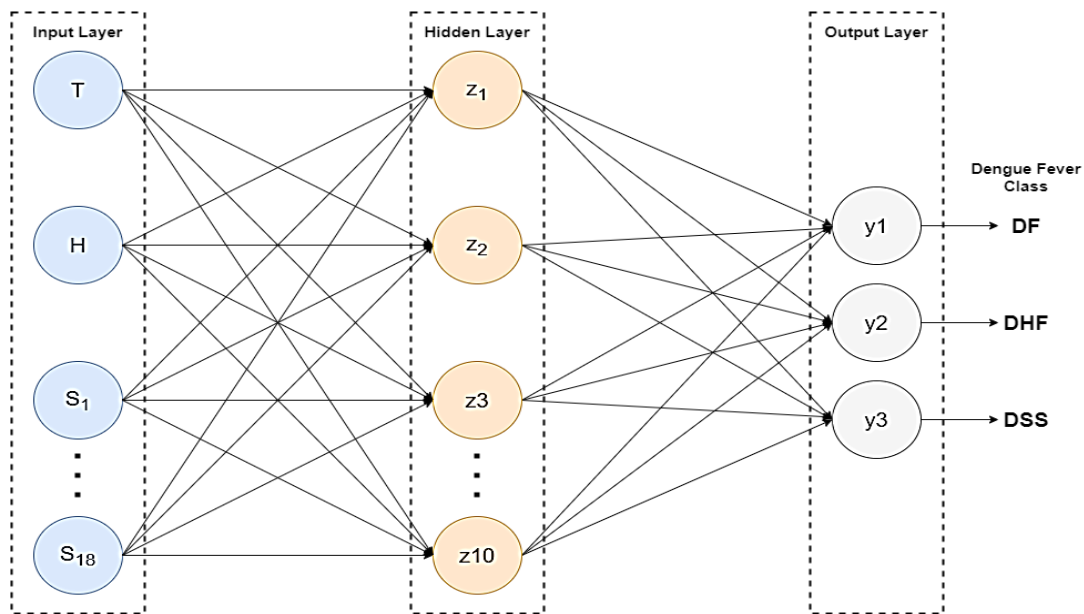


Fig. 2 The BPNN architecture for dengue disease classification

A confusion matrix is a machine learning concept that describes a given classification algorithm's actual and predicted classifications. It has two dimensions: one for the actual class and another for the expected class predicted by the classifier. The basic structure of a confusion matrix for multi-class classification challenges is depicted in Fig. 3. There are n classes, namely $A_1, A_2,$ and A_n . N_{ij} denotes the number of samples belonging to class A_i but incorrectly classified as class A_j in the confusion matrix [22].

Cross-validation with k -fold values of 3, 5, and 10 was utilized to evaluate performance in this study. The dengue patient dataset was subdivided into k subsets based on the type of diagnosis data. Generally, training data is $(k-1)/k$, while testing data is $1/k$. After that, the

procedure is repeated k times. Finally, the final rate estimation was based on the validation result of the mean k -time.

		Output Class		
		A_1	$\dots A_j \dots$	A_n
Target Class	A_1	N_{11}	N_{1j}	N_{1n}
	\vdots		\vdots	
	A_i	N_{i1}	$\dots N_{ij} \dots$	N_{in}
	\vdots			
	A_n	N_{n1}	N_{nj}	N_{nn}

Fig. 3 Confusion Matrix Multiclass

III. RESULT AND DISCUSSION

ANN classification method required pre-processing to perform the discretization, which tried to transform all patient data to a numeric type. There were 18 symptoms (S1–S18) and values for thrombocyte (T) and hemoglobin (H) based on data acquired from dengue patients. As a result, a total of 20 attributes were employed as input data for the subsequent classification process or input layer of the BPNN algorithm. If the patient reports experiencing the symptom, it was assigned a value of 1. Otherwise, if the patient does not perceive it, it was worthless. Dengue diagnosis outcomes were classified into three classes: DF, DHF, and DSS.

The BPNN algorithm was used in the classification procedure. A three-layer Backpropagation Neural Network model is developed, with 20 neurons in the input layer, 10 neurons in the hidden layer, and 3 neurons in the output layer. Following that, the epoch value was 500, and the learning rate consisted of three values, with the magnitude of the change being 0.2, consisting of 0.1, 0.3, or 0.5. This study used a certain threshold value. The training results for each different value of learning rate and class are shown in Table II. The result was indicated using eight parameters, including true positive (TP) rate, false positive (FP) rate, precision, recall, F-score, receiver operating characteristic (ROC), mean absolute error (MAE), and root mean square error (RMSE).

Table II shows the training result with the learning rates of 0.1, 0.3, and 0.5 indicated by the TP rate, FP rate, precision, recall, and F-score obtaining the equal values.

Meanwhile, the parameters of ROC, MAE, and RMSE have no significant difference value for each learning rate. However, the average value of ROC was able to achieve the highest with the lowest value of MAE and RMSE.

Furthermore, the testing step was taken to achieve the method's optimal performance, quantified by three parameters: precision, recall, and accuracy. This number was calculated using a cross-validation technique on a multiclass confusion matrix with three distinct k-fold values of 2, 3, 5, and 10. The testing result performance of the three learning rates obtained with the various k-fold values is summarized in Table III. Table III shows that implementing the BPNN algorithm with a learning rate of 0.1 using k-fold 2, 5, and 10, a learning rate of 0.3 using k-fold 2, 5, and 10, and a learning rate of 0.5 using k-fold 5, and 10 get the lowest performance indicated by the accuracy that achieved the value of 0.956. BPNN with a learning rate of 0.1 using k-fold 3, the learning rate of 0.3 using k-fold 3, and a learning rate of 0.5 using k-fold 2 and 3 obtained the maximum performance of the classification method with the accuracy value 0.967, respectively. The maximum performance was influenced by the highest training results obtained based on the learning rate value of 0.5. The learning rate and k-fold value affect the accuracy, as seen in Table III. Table III represents the number of successfully classified and incorrectly classified data for each method. Meanwhile, Fig. 4–6 show the confusion matrix as the comprehensive study results with different learning rate values and k-fold.

TABEL II
THE TRAINING RESULT OF DENGUE CLASSIFICATION USING DIFFERENT LEARNING RATE VALUES

Learning Rate	Class	Parameter Evaluation (%)							
		TP rate	FP rate	Precision	Recall	F-score	ROC	MAE	RMSE
0,1	DF	1.000	0.060	0.930	1.000	0.964	0.987	0.0554	0.1627
	DHF	0.902	0.000	1.000	0.902	0.949	0.966		
	DSS	1.000	0.012	0.900	1.000	0.947	0.999		
	Average	0.956	0.028	0.959	0.956	0.955	0.978		
0,3	DF	1.000	0.060	0.930	1.000	0.964	0.992	0.0431	0.1608
	DHF	0.902	0.000	1.000	0.902	0.949	0.969		
	DSS	1.000	0.012	0.900	1.000	0.947	1.000		
	Average	0.956	0.028	0.959	0.956	0.955	0.982		
0,5	DF	1.000	0.060	0.930	1.000	0.964	0.992	0.0399	0.1598
	DHF	0.902	0.000	1.000	0.902	0.949	0.972		
	DSS	1.000	0.012	0.900	1.000	0.947	1.000		
	Average	0.956	0.028	0.959	0.956	0.955	0.983		

TABEL III
THE TESTING RESULT OF DENGUE CLASSIFICATION USING DIFFERENT K-FOLDS

Learning Rate	K-Fold	Parameter Evaluation		
		Precision	Recall	Accuracy
0,1	2	0.959	0.956	0.956
	3	0.969	0.967	0.967
	5	0.957	0.956	0.956
	10	0.957	0.956	0.956
0,3	2	0.959	0.956	0.956
	3	0.969	0.967	0.967
	5	0.957	0.956	0.956
	10	0.957	0.956	0.956
0,5	2	0.969	0.967	0.967
	3	0.969	0.967	0.967
	5	0.957	0.956	0.956
	10	0.957	0.956	0.956

Confusion Matrix				
Class		Output Class		
		DF	DHF	DSS
Target Class	DF	40	0	0
	DHF	3	37	1
	DSS	0	0	9

(a)

Confusion Matrix				
Class		Output Class		
		DF	DHF	DSS
Target Class	DF	40	0	0
	DHF	2	38	1
	DSS	0	0	9

(b)

Confusion Matrix				
Class		Output Class		
		DF	DHF	DSS
Target Class	DF	39	1	0
	DHF	2	38	1
	DSS	0	0	9

(c)

Fig. 4 Confusion matrix with learning rate 0.1: (a) k-fold 2, (b) k-fold 3, and (c) k-fold 5&10

Confusion Matrix				
Class		Output Class		
		DF	DHF	DSS
Target Class	DF	40	0	0
	DHF	3	37	1
	DSS	0	0	9

(a)

Confusion Matrix				
Class		Output Class		
		DF	DHF	DSS
Target Class	DF	40	0	0
	DHF	2	38	1
	DSS	0	0	9

(b)

Confusion Matrix				
Class		Output Class		
		DF	DHF	DSS
Target Class	DF	39	1	0
	DHF	2	38	1
	DSS	0	0	9

(c)

Fig. 5 Confusion matrix with learning rate 0.3: (a) k-fold 2, (b) k-fold 3, and (c) k-fold 5&10

Confusion Matrix				
Class		Output Class		
		DF	DHF	DSS
Target Class	DF	40	0	0
	DHF	3	38	0
	DSS	0	0	9

(a)

Confusion Matrix				
Class		Output Class		
		DF	DHF	DSS
Target Class	DF	40	0	0
	DHF	2	38	1
	DSS	0	0	9

(b)

Confusion Matrix				
Class		Output Class		
		DF	DHF	DSS
Target Class	DF	39	1	0
	DHF	2	38	1
	DSS	0	0	9

(c)

Fig. 6 Confusion matrix with learning rate 0.5: (a) k-fold 2, (b) k-fold 3, and (c) k-fold 5&10

Fig. 4–6 illustrate the frequency the DF is misclassified as the DHF. As illustrated in Fig. 4(a), using a learning rate of 0.1 with k-fold 2 results in misclassifying three data from the DF into the DHF. As illustrated in Fig. 4(b), k-fold 3 results in the misclassification of two data in the DF as DHF and one in the DHF as DSS. In contrast, Fig. 4(c) illustrates the implementation of k-fold 5 results in the misclassification of two data in the DF as DHF, one data in the DHF as DF, and one in the DHF as DSS.

Fig. 5(a) shows the implementation of learning rate 0.3 with k-fold 2 causes misclassification of 3 data in the DF, which is classified as DHF, and 1 data in DSS classified as DHF. Fig. 5(b) shows the implementation of k-fold 3 causes misclassification of 2 data in the DF, which is classified as DHF, 1 data in the DHF classified as DSS, and 1 data in the DHF classified as DSS. Meanwhile, Fig. 5(c) shows that the implementation of k-fold 5 causes misclassification of 2 data in the DF, classified as DHF, 1 data in the DHF classified as DF, and 1 data in DHF classified as DSS.

As illustrated in Fig. 6(a), using a learning rate of 0.5 with k-fold 2 results in the misclassification of three data points in the DF, categorized as DHF, and one data point in the DSS is classified as DHF. The implementation of k-fold 3 is depicted in Fig. 6(b). It exhibits the misclassification of two data sets in the DF as DHF, one data set in the DHF as DSS, and one in the DHF as DSS. Furthermore, as illustrated in Fig. 6(c), k-fold 5 results in the misclassification of two data points in the DF as DHF, one data point in the DHF as DF, and one data point in the DHF as DSS. It demonstrates that BPNN with a learning rate of 0.1 and a k-fold value of 2 and 3 may improve precision, recall, and accuracy values and exhibit fewer classification errors.

IV. CONCLUSION

Dengue is a dangerous disease that causes death with different symptoms experienced by patients. There are three varieties of dengue: DF, DHF, and DSS. This disease needs to be detected early to get the appropriate treatment. The expert system for dengue classification was applied using a machine learning approach using BPNN. The input data for the classifier is 20 attributes consisting of 18 symptoms experienced by patients and the values of platelets and hemoglobin. A three-layer BPNN model is constructed, with 20 neurons in the input layer, 10 neurons in the hidden layer, and 3 neurons in the output layer. The afterward epoch value used in this study is 500, with the learning rate consisting of three values, including 0.1, 0.3, and 0.5. Cross-validation was used to evaluate the BPNN classifier with varied k-fold

values of 2, 3, 5, and 10. The evaluation findings indicate that using k-fold 3 with all a learning rate values, the performance of BPNN achieved the highest accuracy value of 0.967. The evaluation shows the learning rate and the k-fold value affect the classification results. This method may be used for larger datasets for future research. The result can improve accuracy and is more helpful in predicting or identifying other diseases.

ACKNOWLEDGEMENT

The authors would like to thank the Faculty of Engineering, Mulawarman University, Samarinda, Indonesia, for supporting this research in 2021.

REFERENCES

- [1] A. Saibene, M. Assale, and M. Giltri, "Expert systems: Definitions, advantages and issues in medical field applications," *Expert Syst. Appl.*, vol. 177, no. November 2020, p. 114900, 2021.
- [2] S. Chatterjee, D. Dey, S. Munshi, and S. Gorai, "Dermatological expert system implementing the ABCD rule of dermoscopy for skin disease identification," *Expert Syst. Appl.*, vol. 167, no. November 2020, 2021.
- [3] A. Salam, Sri Suryani Prasetyowati, and Yuliant Sibaroni, "Prediction Vulnerability Level of Dengue Fever Using KNN and Random Forest," *J. RESTI (Rekayasa Sist. dan Teknol. Informatika)*, vol. 4, no. 3, pp. 531–536, 2020.
- [4] B. Ajlan, M. Alafif, M. Alawi, N. Akbar, E. Aldigs, and T. Madani, "Assessment of the new World Health Organization's dengue classification for predicting severity of illness and level of healthcare required," *bioRxiv*, p. 516229, 2019.
- [5] R. Gangula, L. Thirupathi, R. Parupati, K. Sreeveda, and S. Gattoju, "Ensemble machine learning based prediction of dengue disease with performance and accuracy elevation patterns," *Mater. Today Proc.*, no. xxxx, 2021.
- [6] R. T. Sasmono, A. Sutjiyanto, M. S. Santoso, K. Sriwedari, B. Yohan, E. Mayasanti, R. F. Hayati, and D. Denis, "Molecular epidemiology of dengue in North Kalimantan, a province with the highest incidence rates in Indonesia in 2019," *Infect. Genet. Evol.*, vol. 95, no. May, p. 105036, 2021.
- [7] H. Abid, M. Malik, F. Abid, M. R. Wahiddin, N. Mahmood, and I. Memon, "Global Health Action Nature of Complex Network of Dengue Epidemic As Scale-Free Network," *Healthc. Inform. Res.*, vol. 25, no. 3, pp. 182–192, 2018.
- [8] I. Tougui, A. Jilbab, and J. El Mhamdi, "Impact of the Choice of Cross-Validation Techniques on the Results of Machine Learning-Based Diagnostic Applications," *Healthc. Inform. Res.*, vol. 27, no. 3, pp. 189–199, 2021.
- [9] Z. Rustam, F. Zhafarina, G. S. Saragih, and S. Hartini,

- “Pancreatic cancer classification using logistic regression and random forest,” *IAES Int. J. Artif. Intell.*, vol. 10, no. 2, pp. 476–481, 2021.
- [10] N. A. Mashudi, N. Ahmad, and N. M. Noor, “Classification of adult autistic spectrum disorder using machine learning approach,” *IAES Int. J. Artif. Intell.*, vol. 10, no. 3, pp. 743–751, 2021.
- [11] A. Salam, Sri Suryani Prasetyowati, and Yuliant Sibaroni, “Prediction Vulnerability Level of Dengue Fever Using KNN and Random Forest,” *J. RESTI (Rekayasa Sist. dan Teknol. Informatika)*, vol. 4, no. 3, pp. 531–536, 2020.
- [12] A. Septiarini, D. M. Khairina, A. H. Kridalaksana, and H. Hamdani, “Automatic glaucoma detection method applying a statistical approach to fundus images,” *Healthc. Inform. Res.*, vol. 24, no. 1, pp. 53–60, 2018.
- [13] J. J. Patel and S. K. Hadia, “An enhancement of mammogram images for breast cancer classification using artificial neural networks,” *IAES Int. J. Artif. Intell.*, vol. 10, no. 2, pp. 332–345, 2021.
- [14] M. S. Efendi and H. A. Wibawa, “Prediksi Penyakit Diabetes Menggunakan Algoritma ID3 dengan Pemilihan Atribut Terbaik (Diabetes Prediction using ID3 Algorithm with Best Attribute Selection),” *Juita*, vol. VI, no. 1, pp. 29–35, 2018.
- [15] R. Ardiansyah, F. Fauziah, and A. Ningsih, “Sistem Pakar Diagnosa Awal Penyakit Lambung Menggunakan Metode Dempster-Shafer Berbasis WEB,” vol. 24, no. 3, pp. 182–196, 2018.
- [16] Laurentinus, “Demster Shafer dengan Fuzzy Expert System Pada Deteksi Penyakit Stroke,” *J. Resti*, vol. 1, no. 10, pp. 837–845, 2021.
- [17] Y. Yohannes, S. Devella, and K. Arianto, “Deteksi Penyakit Malaria Menggunakan Convolutional Neural Network Berbasis Saliency,” *JUITA J. Inform.*, vol. 8, no. 1, p. 37–44, 2020.
- [18] H. Mustafidah and S. Suwarsito, “Performance of Levenberg-Marquardt Algorithm in Backpropagation Network Based on the Number of Neurons in Hidden Layers and Learning Rate,” *JUITA J. Inform.*, vol. 8, no. 1, p. 29–35, 2020.
- [19] Y. A. Lesnussa, C. G. Mustamu, F. Kondo Lembang, and M. W. Talakua, “Application of Backpropagation Neural Networks in Predicting Rainfall Data in Ambon City,” *Int. J. Artif. Intell. Res.*, vol. 2, no. 2, 2018.
- [20] A. F. Daru, M. B. Hanif, and E. Widodo, “Improving Neural Network Performance with Feature Selection Using Pearson Correlation Method for Diabetes Disease Detection,” *JUITA J. Inform.*, vol. 9, no. 1, p. 123–130, 2021.
- [21] M. Sokolova, N. Japkowicz, and S. Szpakowicz, “Beyond accuracy, F-score and ROC: A family of discriminant measures for performance evaluation,” *AAAI Work. - Tech. Rep.*, vol. WS-06-06, no. January, pp. 24–29, 2006.
- [22] X. Deng, Q. Liu, Y. Deng, and S. Mahadevan, “An improved method to construct basic probability assignment based on the confusion matrix for classification problem,” *Inf. Sci. (Ny)*, vol. 340–341, pp. 250–261, 2016.