

Early Detection of *Dengue* Disease Using Extreme Learning Machine

Suhaeri[#], Nazri Mohd Naw^{i*}, Muhamad Fathurahman¹

[#] Faculty of Information Technology, Universitas YARSI, Jakarta, Indonesia

^{*} Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia, 86400, Johor, Malaysia

E-mail: nazri@uthm.edu.my

¹ Faculty of Medicine, Universitas YARSI, Jakarta Indonesia

Abstract—Dengue disease is one of the serious and dangerous diseases that cause many mortality and spread in most area in Indonesia. There are about 201,885 cases had been reported in 2016 including 1,585 death cases. The availability of nowadays clinical data of Dengue disease can be used to train machine learning algorithm in order to automatically detect the present of Dengue disease of the patients. This study will use the Extreme Learning Machine (ELM) method to classify the dengue by using the clinical data so that first aid can be given which can decrease some death risk. The back propagation neural network is one of the popular machine learning technique that capable of learning some complex relationship and had been used in many applications. However, back propagation neural network still suffers with some limitations such as slow convergence and easily getting stuck in local minima during training. Therefore, this research proposed an improved algorithm known as ELM which is an extension of Feed Forward Neural Network that utilize the Moore Penrose Pseudoinver matrix that gain the optimal weights of neural network architecture. The proposed ELM prevents several backpropagation issues by reducing the used of many parameters that solves the main drawbacks of Backpropagation algorithm that uses during the training phase of Neural Network. The result shows that the proposed ELM with selected clinical features can produce best generalization performance and can predict accurately with 96.94% accuracy. The proposed algorithm achieves better with faster convergence rate than the existing state-of-the-art hierarchical learning techniques. Therefore, the proposed ELM model can be considered as an alternative algorithm to apply for early detection of Dengue disease.

Keywords— machine learning; artificial neural network; back propagation; dengue disease; classification.

I. INTRODUCTION

There is no doubt that staying healthy in daily life is a priceless treasure anyone in this world are willing to do anything in order to maintain our physical health. Anyone may face health problems, regardless of age, sex, and profession. Anyone can be infected with different kind of diseases that can cause losing life and one of the most dangerous diseases in the world that caused so many fatalities is dengue.

In medical definition, any Dengue infection is caused by dengue germ that can be infected human through female mosquito is known as *Aedes aegypti* and the symptoms of the infections are shown such as feeling headache, joint-pain, retro-orbital pain, muscular pain and some evidence of rash. In other words, the transmission is also known as a bone breaking unwellness [1]. Dengue occurs almost every year

during the monsoon in sub-tropical and tropical regions around the world. At present, there is no specific medicine prescribed for its treatment – the patient must be given complete rest and plenty of fluids to help the body flush out the toxins. Over 2.5 billion populations all around the globe have been endangered by Dengue infection, and there are almost 50 million people suffer from it globally every year [1].

Indonesia also has been a victim of this rapidly growing sickness, and dengue is one of the serious diseases in Indonesia. There are 201,885 cases in 2016 including 1,585 death cases [2]. The doctor recognizes the Dengue disease by analyzing some common signs, symptoms, and risk factors of this disease based on clinical data. However, all those signs, symptoms and the risk factors cannot be measured with 100% certainty because of the occurrence of since various types of uncertainties such as imprecision, vagueness, ambiguity, and ignorance are associated with

them. Hence, physicians faced difficulty in diagnosing the dengue patient accurately since they do not consider the uncertainties as part of critical factors for the disease.

Based on world health organization, according to Rigau [3], dengue can be divided into two types. The first type of dengue is known as classical dengue or commonly known as *Dengue* fever, and the other one type of dengue is known as dengue hemorrhagic fever (DHF). Also, medical had categorized further dengue hemorrhagic fever into four types which are DHF1, DHF2, DHF3, and DHF4. It is easy for anyone to identify symptoms of DHF such as high fever which continues up 2 to 7 days as well as a few numbers of pre-indication like leakage of plasma, feeling stupor and feeble heart rate. However, in the early spotting stage, scientist found out that it still very hard for them to differentiate between dengue febricity from dengue hemorrhagic pyrexia.

Nowadays, there are different techniques of data mining that are available and can be used for Dengue fever classification. This includes NB classifier; KNN Technique, decision tree, SVM and multilayered Technique [3]. Most of those techniques are evaluated based on five common measurements in term of the accuracy, sensitivity, precision, specificity, and number of negative rate.

Some researchers such as Tanner et al. only focused on classifying Dengue (fever) by using Decision tree technique to classify 1200 patient datasets where from the research they found out 6 remarkable features from the datasets, and they also managed to predict 84% accurately [4]. Later in 2010, Self-Organizing MAP (SOM) and ML feed-forward neural networks (MFNN) been used and proposed by Tarig's team to cluster patients into two sets and they managed to classify only 70% correctly [5]. Another researcher from Pakistan used ML perceptron's (MLP), and they managed to achieve better results of 90% accuracy [6]. Later, Daranee et al. [7] suggested in his work to use the decision tree method to classify dengue patients from two different datasets. They managed to gain better results and managed to achieve accurate results of 97.6% and 96.6 % accuracy from first as well as second experiment respectively. However, for unseen data, their method performed badly. Based on the experimental results, it can be concluded that the decision tree approach did not make counterpart and solve the task for dengue classification very well.

Some researchers also used the expert system on the analysis of dengue fever, such as those conducted by Hossain et al., [8] which diagnosing the Dengue disease with forwarding Chaining method through the input of symptoms. Unfortunately, their study showed that the results obtained from the test were uncertain and not accurate.

Another clustering approach was proposed by P. Manivannan et al. [9] where K-Mediod Clustering algorithm was used to predict dengue fever. The proposed method predicted the people who were affected by dengue based on the categorization of age from a given dengue dataset. The simulations result showed that the proposed K-medoid clustering algorithmic rule increased the efficiency of the output and the author mentioned that the proposed method was also the most effective technique to predict the dengue patient with serotypes. However, the paper discovered that the method still facing some problems and still lack in term

of scalability particularly on large datasets as well as facing with high time processing and space complexity.

Tarmizi et al. [10] from Malaysia had done their research by referring to the weather in Thailand, Indonesia, and Malaysia. The main reasons for choosing those countries are because the climates of those three countries are known for their humidity, and therefore that kind of environment attracted a different kind of water-borne diseases such as dengue. They proposed different machine learning techniques such as Data mining (DM) and Artificial neural network (ANN). Moreover, the paper also includes rough set theory (RS) to predict Dengue disease. The dataset was taken from a public hospital from Selangor state. First, the dataset was split into 10 cross-fold validation where two test datasets were used with the simulator WEKA tool. The results showed that the accuracy obtained with ten cross-fold validation from those three methods were almost the same where it showed that DM classified accurately with 99.5% accuracy, ANN performed slightly better with 99.8% accuracy, and RS achieved the best accuracy of 100%. Also the paper showed that by using a split percentage, 99.2% of accuracy was achieved with DM and ANN, whereas 99.72% of accuracy was obtained using RS.

Later, an artificial neural network with multi-layer feed forward neural network was suggested by Ibrahim et al. [6] which used to forecast the defervescence fever in patients of Dengue disease. The data was taken from 252 hospitalized patients, in whom 4 patients were identified having Dengue Fever and 252 patients had Dengue hemorrhagic fever. They used the neural network toolbox from MATLAB and managed to achieve the accuracy of 90%. They found out that the accuracy can be increased by using different methodologies in artificial neural networks.

Since then, Artificial Neural Network (ANN) have been known as one of the popular method in machine learning and have been applied in many applications including research in medical world due to the introduction a new way of analysis and decision making by mimicking the human brain's work, which increase the accuracy of the analysis [11]-[14].

Even though ANN had gained its popularity in most applications, yet ANN still suffer from some limitations such as slow in learning speed, and it has been a significant bottleneck in their applications for past decades. Two main reasons cause the problems: (1) the extensively use of slow first order gradient-based learning algorithms to train neural networks, and (2) most of the parameters of the networks are tuned heuristically by using such learning algorithms. Many improvements on ANN is ongoing research where researchers had introduced a different kind of parameters adjustment as well as introducing higher order method rather than first order method. Besides that, researchers also introduced hybrid methods that combine machine learning techniques with ANN. The main reason for using machine learning is that their popularity and their good performances had attracted many researchers.

In the recent advances. Machine Learning techniques have widely applied in analyzing the clinical data. One of the favorite Machine Learning Techniques is Extreme Learning Machine (ELM). It is an improved model of Neural Network that proposed by Huang [15] to solve some

drawbacks of the *back propagation algorithm* during the training phase. The drawbacks including the training time, stuck in local minima and too many parameters should be tune [15], [16]. Many studies have been done to analyze the performance of ELM biomedical application that uses clinical data such as prediction the mortality of bladder cancer [17], Heart Disease diagnosis [16], and Prediction of protein interaction [18]. Therefore, in this study, we will propose an extreme learning machine (ELM) to build an

early detection system for Dengue disease. Also, this study aims to measure the accuracy of the ELM model in term of classification of Dengue Fever.

The rest of the paper is organized as follows. Section II provides the methodology that uses in this study and the basic idea of ELM. Section III provides the experiment design, result, and discussion. The last section presents the conclusion and future work

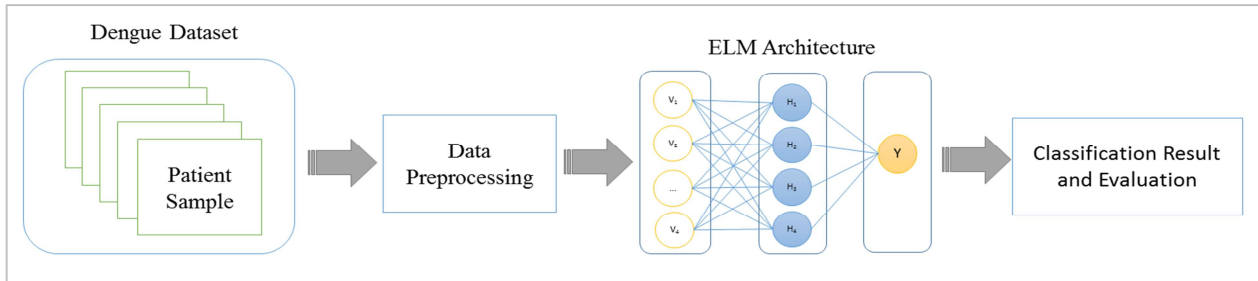


Fig. 1: The proposed approach

II. MATERIAL AND METHOD

In this section, we describe the approach to automatically classify the Dengue disease by using ELM architecture (ELM). Extreme Learning Machine (ELM) is an extension of Feed Forward Neural Network that is proposed by Huang et al. [15]. Instead of using the *back propagation algorithm* for learning schema, the ELM utilizes the Moore Penrose *Pseudo inverse matrix* in order to gain the optimal weights of the neural network.

In this research, we proposed an approach that used ELM for a better classification for detecting Dengue disease. The architecture of the ELM is shown in Figure 1. In the first step, we collect the patient samples from several public health centers in Jakarta, Indonesia. Then we apply data preprocessing techniques before feeding to ELM architecture.

The architecture of ELM is shown in Figure 2. The preprocessing techniques that are used in this study including fill in missing values remove outliers sample, feature selection, and data normalization. The next step is feeding the preprocessed dataset into the ELM architectures. Finally, the evaluation of the classification result is provided.

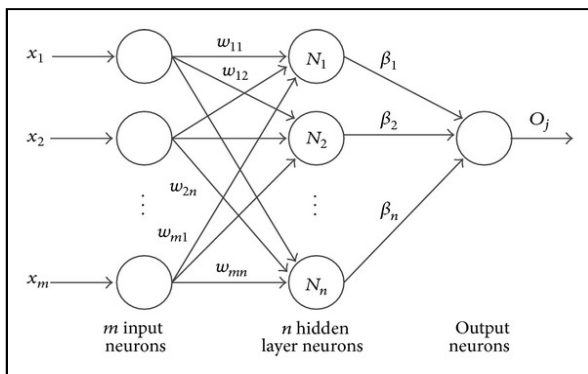


Fig..2 The ELM architecture

A. Dataset

We gain a hundred patient samples from Public Health Center in Jakarta, Indonesia. Each sample contains clinical features. The clinical features that are used in this research including leukocytes, temperature, hemoglobin, thrombocyte, and the total number of the day getting a fever. These clinical features are checked and recorded periodically by the clinician. Doctors can determine the diagnosis of a patient by analyzing the clinical features. In this study, the status of a patient is presented as binary values. The 0 means that the patient is free of Dengue disease and the 1 means that the patient potentially has a Dengue disease.

B. Data Preprocessing

The dataset contains many missing values. Therefore data preprocessing are required. The missing values occur because of every single patient has a different total number of clinical features. Some patients are diagnosed by the doctor only using few clinical features. Otherwise the other require more periodically check of clinical features such as leukocytes hemoglobin, thrombocyte to diagnose the patient. One of the solutions to missing values is asking to the expert, So that, the missing values in the dataset are filled by the doctor. The next step of data preprocessing is data normalization. The dataset is normalized by Equation 1.

$$x_i = \frac{x_i - \min}{\max - \min} \quad (1)$$

After the normalization process, generalized cross-validation (CGV) was used as a feature selection method.

C. Extreme Learning Machine(ELM)

One of the popular learning algorithms for the generalized single hidden layer feed-forward neural networks is Extreme Learning Machine (ELM). The algorithm randomly generates and the output weights by the unknown node parameters and analytically computed. ELM is an extension

of Feed Forward Neural Network that been proposed by Huang et al. [15]. Instead of using the *back propagation algorithm* for the learning schema, the ELM utilizes the *Moore Penrose Pseudo inverse* matrix in order to gain the optimal weights of neural network architecture. ELM has more advantages to the traditional neural network. It prevents several *back propagation* issues such as many parameters should be tuned, long training time and trap in local minima [15]. In contrast to standard feed forward neural network, where this neural network architecture has formed the computation as follows [15];

$$\sum_{i=1}^N \beta_i g(w_i \cdot v_i + b_i) = y_i \quad (2)$$

Where v_i is the input feature, w_i is the weight that connects visible and hidden layer. $g(v)$ is the activation function and β_i is the weight between hidden and output layer. In ELM architecture, the Equation 1 can also be written in the following form [15].

$$\mathbf{H}\beta = \mathbf{Y} \quad (3)$$

Where

$$\mathbf{H} = (\mathbf{w}_1, \dots, \mathbf{w}_N, \mathbf{b}_1, \dots, \mathbf{b}_N, \mathbf{v}_1, \dots, \mathbf{v}_N) = \begin{bmatrix} g(\mathbf{w}_1 \cdot \mathbf{v}_1 + \mathbf{b}_1) & \dots & g(\mathbf{w}_N \cdot \mathbf{v}_1 + \mathbf{b}_N) \\ \dots & \dots & \dots \\ g(\mathbf{w}_1 \cdot \mathbf{v}_N + \mathbf{b}_1) & \dots & g(\mathbf{w}_N \cdot \mathbf{v}_N + \mathbf{b}_N) \end{bmatrix} N \times N, \quad (4)$$

$$\beta = \begin{bmatrix} \beta_1 \\ \dots \\ \beta_N \end{bmatrix} N \times m, \quad \mathbf{Y} = \begin{bmatrix} Y_1 \\ \dots \\ Y_N \end{bmatrix} N \times m \quad (5)$$

Where \mathbf{H} is the matrix for hidden output, x_i refers to the input feature whereas N represents the total number of hidden units. Then the training phase of ELM can be solved by finding the optimal solution of β of the linear system $\mathbf{H}\beta = \mathbf{Y}$ [15]. The optimal solution for that linear system can be gained by using this following equation:

$$\beta = \mathbf{H}^\dagger \mathbf{Y} \quad (6)$$

Where \mathbf{H}^\dagger is the matrix for *moore penrose* generalized invers. Basically, the architecture for ELM consists of three layers, The first layer is in the input layer, followed by hidden layer and the last layer is output layer. ELM avoid the *backpropagation* issues such as, the predefined the learning rate, has a long training time and easily trap in local minima. The basic process of ELM algorithm is described as the following steps[15] :

1. Initially choose the value of weights between the input randomly hidden layer and hidden to the output layer.
2. Calculate the hidden output matrix \mathbf{H} by using Equation (4)
3. Find the weights β by using Equation (5).

The further detail of the ELM algorithm can be seen in some studies [15], [19], [20].

III. RESULTS AND DISCUSSION

In this section, the performance of the proposed ELM learning algorithm is compared with the different activation functions and different classification algorithms such as SVM, decision tree and Naive Bayes. All simulations for all algorithms such as BP and ELM are carried out using MATLAB 6.5 environment which run in Pentium 4 processor with 1.9 GHz CPU. The datasets were taken from the public health center in Jakarta, Indonesia. The datasets contain of 98 dengue fever samples and the total number of clinical features for each sample reaches 54 including the status of the patient. The status of the patient is presented as a binary value. The 0 value refers to the healthy patient with free dengue symptoms and the 1 value refer to the patient with confirmed dengue disease by the doctor.

Before the dataset is fed to the ELM model, all samples are normalized by using Equation (1). Then, the ELM model are built with a different configuration including the number of hidden neurons and the variety of activation function such as sin, radial basis, pure linear and tan-sigmoid. K-Fold cross-validation is used to perform the training and testing phase. In this study, 3-2000 neurons in the hidden layer are analyzed in order to choose the best ELM model. In the next scenario, important feature are selected based on generalized cross-validation (CGV), number of subset models and Residual Sum of Squares (RSS) that is available in earth package in R. Other classification model are built for the comparison purpose including Support Vector Machine (SVM), Naive Bayes (NB) and Decision Tree (DT). Caret Package that is provided by R is utilized in order to build all classification models.

For evaluation purposes, the accuracy of each model are calculated according to the following equation:

$$\text{acc} = \frac{\text{TP} + \text{TN}}{\text{TN} + \text{TP} + \text{FN} + \text{FP}} \quad (7)$$

The experiment results are given in Figure 3 until Figure 5 where all figures show the validation accuracy of ELM model. The number of hidden nodes are essential parameters that affect the performance of ELM models. Also, the type of activation functions in the hidden unit also shows the different mean of accuracy. The pure linear activation function provides the best performance when the number of hidden unit's is approximately reach $\sim 15 - 45$ which is less than the number, of features of the sample data. In contrast, the sin, radial basis, and tan-sigmoid activation function have a better accuracy result when the number of hidden units more significant than the number of features of the sample data.

It clearly shows that the ELM model that uses tan-sigmoid activation function outperforms the performance of ELM model with other activation function. The ELM model had achieved the best accuracy of 93.45% with 1295 hidden unit and tan-sigmoid activation function. Figure 3 shows the classification result based on the selected essential features. Seven selected features are based on CGV and RSS. All of the ELM has slightly improved the classification result. When the number of hidden units exceed 100, the ELM model produces greater than $\sim 90\%$ in accuracy.

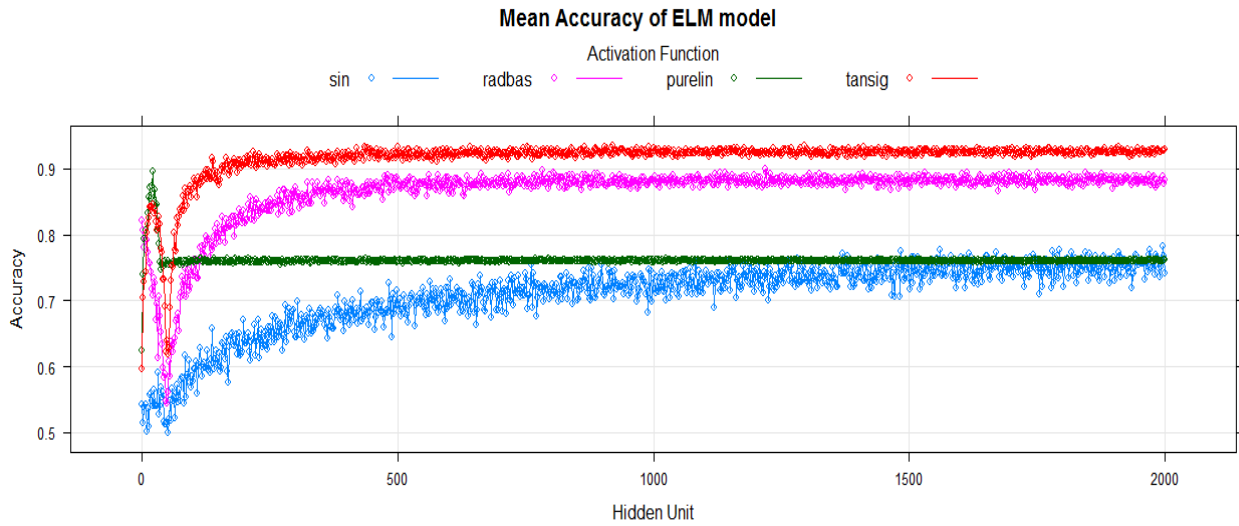


Fig.3: Mean Accuracy of ELM model

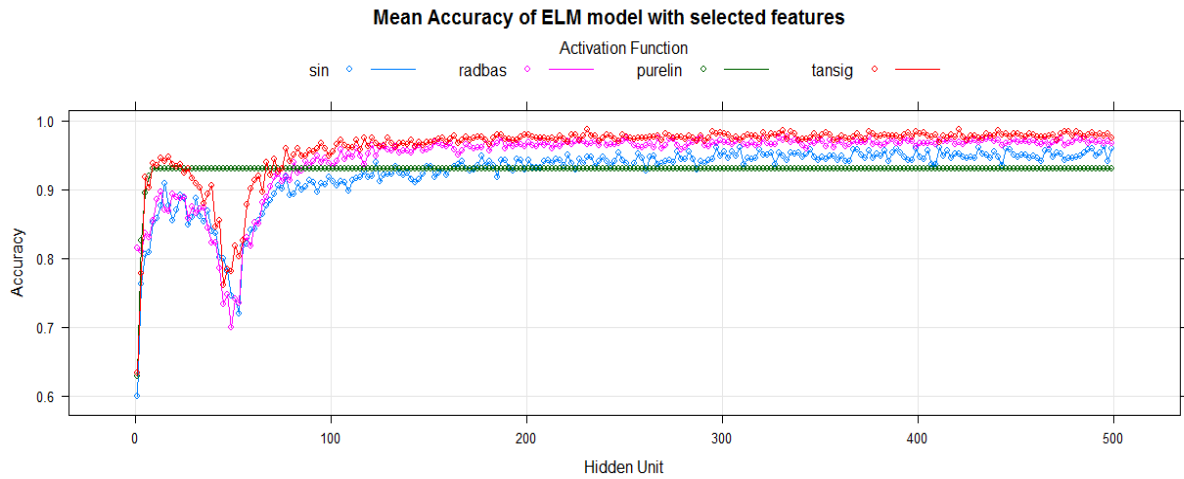


Fig.4: Mean accuracy of ELM model with selected features

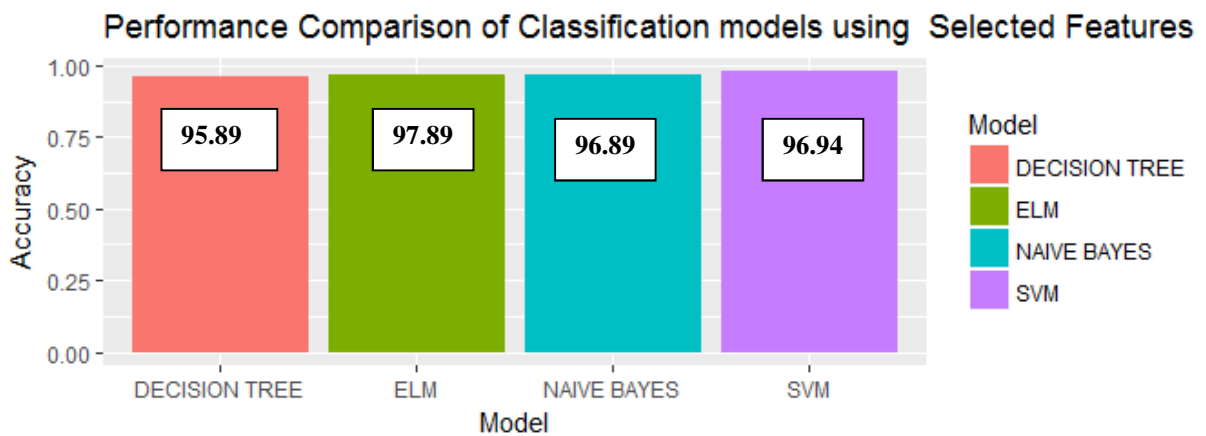


Fig.5: Performance Comparison of Classification Model

In comparison with the other classification models (see Figure 4). ELM has competitive classification result as compared to others. In this experiment, SVM produces the highest % in the mean of testing accuracy, followed by ELM 97.89%, NB 96.89%, and DT 95.89%. It shows that ELM produces good generalization performance in classifying Dengue disease and can learn thousands of times faster than conventional popular learning algorithms for feed-forward neural networks.

The results showed that the ELM performs better with faster learning time than SVM by a factor up to thousands. The results in our simulations show that for classifying Dengue disease, the response speed of trained feed-forward neural networks is much faster than SVM particularly to the external new unknown observations. The main reasons are because of SVM algorithms is more complex and usually generate a much larger number of support vectors (computation units) with better processing whereas, for the traditional feed-forward neural networks, the network requires only very few hidden nodes (computation units) for learning on the same applications. Moreover, it is not easy for SVM models to classify real-time data particular for this application since several hours may be spent for such classification (testing) set, while the ELM demonstrates that the technique is more suitable in applications which require fast classification as well as fast response capability.

IV. CONCLUSION

In conclusion, this research aims to analyze the performance of ELM in classifying the Dengue disease. The result shows that the proposed ELM model can achieve 96.94% in the mean of testing accuracy that is slightly lower than that of SVM (97.89%). This study also shows that the ELM model with tan-sigmoid activation function has better performance as compared to the ELM model with another type of activation functions. The results show that ELM performs better learning efficiency than the other machine learning techniques. Simulations result on the Dengue disease classification dataset show that the proposed algorithm achieves better with faster convergence rate than the existing state-of-the-art hierarchical learning techniques. Therefore, our proposed ELM model can be considered as an alternative algorithm to apply for early detection of Dengue disease.

ACKNOWLEDGMENT

This research was funded by the Internal Research grant from Universitas YARSI Number 101/INT/UM/WR II/UY/VII/2017. The authors also would like to thank Universiti Tun Hussein Onn Malaysia (UTHM) and Ministry of Higher Education (MOHE) Malaysia for financially supporting this Research under Trans-disciplinary Research Grant Scheme (TRGS) vote no. T003.

REFERENCES

[1] R. T. Sasmono, I. Wahid, H. Trimarsanto, B. Yohan, S. Wahyuni, M. Hertanto, I. Yusuf, H. Mubin, I. J. Ganda, R. Latief, P. J. Bifani, P. Y. Shi, and M. J. Schreiber, "Genomic analysis and growth

characteristic of dengue viruses from Makassar, Indonesia," *Infect. Genet. Evol.*, vol. 32, pp. 165–177, 2015.

[2] B. Hardhana, C. S. Budiono, N. Kurniasih, E. V. Manullang, M. I. Susanti, S. Pangribowo, A. Harpini, R. Aprianda, E. Satriani, R. Mardina, E. Susetyoaji, H. A. Habibi, M. Sari, and B. B. Sigit, *Data dan Informasi Profil Kesehatan Indonesia 2016*. 2016.

[3] Rigau-Pérez JG, et.al. (1998) Dengue and dengue haemorrhagic fever. *The Lancet* 19: 971-977

[4] Vong S, et.al. (2010) Dengue incidence in urban and rural Cambodia: results from population-based active fever surveillance, 2006–2008. *PLoS neglected tropical diseases* 4: e903.

[5] Faisal T, Ibrahim F, Taib MN (2010) A noninvasive intelligent approach for predicting the risk in dengue patients. *Expert Systems with Application* 37:2175-2181.

[6] Ibrahim F, Taib MN, Abas WA, Guan CC, Sulaiman S (2005) A novel dengue fever (DF) and dengue haemorrhagic fever (DHF) analysis using artificial neural network (ANN). *Computer Methods and Programs in Biomedicine* 79: 273-281.

[7] Daranee T, PrapatS, Nuanwan S (2012) Data mining of dengue infection using decision tree. *Entropy* 2: 2.

[8] Hossain, Mohammad Shahadat, Habib, Israt Binteh, Andersson, Karl. (2017). A Belief Rule based Expert System to Diagnose Dengue Fever under Uncertainty. *Computing Conference 2017*, 179-186 p.

[9] P.Manivannan, Dr. P. Isakki, "Dengue Fever Prediction using K-Medoid Clustering Algorithm", *International Journal of Innovative Research in Computer and Communication Engineering*, Vol. 5, Special Issue 1, March 2017

[10] Tarmizi, N.D.A., Jamaluddin, F., Abu Bakar, A., Othman, Z.A., Zainudin, S. and Hamdan, A.R. (2013) Malaysia "Dengue Outbreak Detection Using Data Mining Models", *Journal of Next Generation Information Technology (JNIT)*, 4, 96-107.

[11] Chin Kim On, Teo Kein Yau, Rayner Alfred, Jason Teo, Patricia Anthony, Wang Cheng, Backpropagation Neural Ensemble for Localizing and Recognizing Non-Standardized Malaysia's Car Plates. *International Journal on Advanced Science, Engineering and Information Technology*, Vol. 6 (2016) No. 6, pages: 1112-1119.

[12] Yusuf Hendrawan, Dimas Firmanda Al Riza, Machine Vision Optimization using Nature-Inspired Algorithms to Model Sunagoke Moss Water Status. *International Journal on Advanced Science, Engineering and Information Technology*, Vol. 6 (2016) No. 1, pages: 45-57.

[13] Haruna Chiroma, Sameem Abdul-kareem, Abdullah Khan, Nazri Mohd. Nawi, Abdulsalam Ya'u Gital, Liyana Shuib, Adamu I. Abubakar, Muhammad Zubair Rahman, Tutut Herawan, " Global Warming: Predicting OPEC Carbon Dioxide Emissions from Petroleum Consumption Using Neural Network and Hybrid Cuckoo Search Algorithm", Volume 10, Issue 8, 25 (2015)

[14] Habib Shah, Rozaida Ghazali, Nazri Mohd Nawi, " Hybrid Ant Bee Colony Algorithm for Volcano Temperature Prediction", *Communications in Computer and Information Science* Volume 281 CCIS, 2012, Pages 453-465, pp 453-465

[15] G.-B. H. G.-B. Huang, Q.-Y. Z. Q.-Y. Zhu, and C.-K. S. C.-K. Siew, "Extreme learning machine: a new learning scheme of feedforward neural networks," *2004 IEEE Int. Jt. Conf. Neural Networks (IEEE Cat. No. 04CH37541)*, vol. 2, pp. 25–29, 2004.

[16] M. Fathurachman and U. Kalsum, "Heart disease diagnosis using extreme learning based neural networks." ... *Concept, Theory* ..., 2014.

[17] G. Wang, K.-M. Lam, Z. Deng, and K.-S. Choi, "Prediction of mortality after radical cystectomy for bladder cancer by machine learning techniques," *Comput. Biol. Med.*, vol. 63, pp. 124–132, 2015.

[18] Z. You, L. Li, Z. Ji, M. Li, and S. Guo, "Prediction of protein-protein interactions from amino acid sequences using extreme learning machine combined with auto covariance descriptor," *Memetic Comput. (MC)*, 2013 ..., pp. 80–85, 2013.

[19] J. Tang, Chenwei Deng, and G.-B. Huang, "Extreme Learning Machine for Multilayer Perceptron," *IEEE Trans. Neural Networks Learn. Syst.*, pp. 1–13, 2015.

[20] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification.," *IEEE Trans. Syst. Man. Cybern. B. Cybern.*, vol. 42, no. 2, pp. 513–29, 2012.