International Journal of Engineering & Technology, 7 (4.19) (2018) 458-463

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Fuzzy Rules Base System for Early Self-Diagnosis of Dengue Symptoms

Nor Azura Husin[1], Akram Saleh Naseer Al-Harogi[1], Norwati Mustapha[1], Hazlina Hamdan[1], UmmiAmalina Husin[2]

[1]Intelligent Computing Research Group,
Faculty of Computer Science and Information Technology, University Putra Malaysia
[2]Faculty of Medicine, University Kebangsaan Malaysia
**Corresponding Author Email: [1]n_azura@upm.edu.my*

Abstract

Dengue has become rapidly expanding and significant public health problem in tropical and subtropical regions. In severe cases, people infected with dengue may experience severe bleeding, shock and death. Thus, increasing dengue fever (DF) can be very serious, potentially life threatening and becoming global threat. Therefore, this research aimed to develop an accurate model that could better detect early signs and symptoms of dengue fever and develop a practical system for self-notification of the disease. Two techniques were applied to provide early self-notification to the patients namely the fuzzy expert system and data mining technique. The rules of dengue diagnosis are developed based on an interview with a medical doctor and those rules will be applied in an expert system using a fuzzy logic. However, before applying the extracted rules, the accuracy of rules was tested by data mining tool. This research applies the methodology to dengue related-data from a hospital and compares the rules to the training dataset by Multilayer Perceptron network. Furthermore, the finding showed that the accuracy of result for self-diagnosis of dengue symptoms produce a reliable result.

Keywords: *Dengue; data mining; early detection, Fuzzy logic.*

1. Introduction

According to the World Health Organization (WHO) [1] incident rate of dengue infection has increased 30-fold over the last five decades with approximately 100 million infections occurring worldwide annually, making almost half of the world's population at risk of dengue. As eloquently stated by Health Minister, Datuk Seri Dr. S. Subramaniam (2017) DF is the most prevalent disease in Malaysia with a ratio of 328.3 cases per 100,000 populations compared to other infectious disease like hand, foot, and mouth disease (HFMD), leptospirosis, hepatitis B, and HIV [2].

DF is a disease caused by a family of viruses, known as Flaviviridae. It is mosquito-borne viral infection that is being rapidly transmitted by the Aedes mosquitoes. These dengue viruses are transferred on to humans through the bites of an infective female Aedes mosquito, which mainly acquires the virus during feeding on the blood of an infected humans. Once infected, dengue virus can cause a flu-like illness and develop symptoms typically 2-7 days after the first symptoms simultaneously with a decrease in temperature, below 38°C and included other symptoms such as abdominal pain, persistent vomiting, rapid breathing, bleeding gums, fatigue, restlessness and vomiting blood [3]. Sometimes dengue infection might cause fatal outcome in severe cases. Severe dengue occurs less often than DF but has a more serious clinical presentation [4]. As stated by World Health Organization (2017) severe dengue, previously known as dengue haemorrhagic fever (DHF) is potentially life threatening to the patient due to severe plasma leaking, severe fluid accumulation that cause shortness of breath, severe bleeding, or organ impairment [5].

There is no specific treatment for DF or severe dengue, other than supportive management and fluid therapy to maintain patient's body fluid volume [6].

Life threatening complication can occur rapidly in delay of proper medical care. Vicente et, al. in her study revealed that delay in medical treatment is found to be significantly associated with complications leading to severe dengue [7]. In addition, this DF can be difficult to distinguish from other viral fever states especially during the early phase of illness, where non-specific clinical symptoms and signs accompany the febrile illness [8]. Furthermore, patient with early sign and symptoms will only appear when the patient's state is critically ill [9]. Thus, it is very crucial to recognize early signs and symptoms of DF, so that proper treatment can be administered to prevent complications and reduce dengue mortality. World Health Organization indicated that early detection of DF and appropriate treatment will decrease fatality rates from more than 20% to less than 1% [5].

Based on our observation, despite several recent studies on diagnosis of DF some of these models still have constraint in terms of accuracy and limited practical use for medical doctors. Hence, it is crucial to develop an accurate model that could better detect early signs and symptoms of DF and self-notification of the disease by the laymen. This study proposesan accurate system that could provide early self-notification to the patients whether they are

Suspected to have DF or not by using data mining tool and fuzzy logic techniques. The model enables infected patients to determine whether they need to seek medical care immediately before the symptoms get worse.

Copyright © 2018 Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

To aid discussion, this paper is structured as follows: The next section deliberates on previous studies related to methods of notification of DF.
Section 3 explains the dataset focusing on training and testing dataset. The methodology and data implementation will be discussed in section 4. Section 5 presents the result and finding from the system process. Finally, section 6 concludes the work in this paper.

2. Related Works

Initial observation reveals that there are several recent studies on early diagnosis of DF symptoms. This section summarizes literature and previous findings on the research subject.
Razak T. et al. [10] proposed a fuzzy logic system for enhanced early diagnosis of dengue disease. This research seeks to provide a more convenient way for patient to get consultancy with an expert and facilitate the doctor's role in diagnosing DF with a fuzzy logic system. The strength in this research is the ability for analyzing the unclear symptoms because it contains 'no', 'yes' and 'maybe' values where the maybe value is used in unclear symptoms. However, this research is limited to the symptom based and the researchers do not consider other significant factors such as the duration of symptoms as part for analyzing the disease.
This research coincides with another observation made by V. Pabbi et al. [11], in the application of fuzzy expert system for early diagnosis of DF. In term of accuracy, this current fuzzy system has an equivalent ability with doctors in diagnostic and assessing the risk of disease. The result output in this research reveals almost 100% accuracy in predicting the type of DF. However, the researchers only focus on specific types of dengue which are DF, DHF and dengue shock syndrome. Besides, this research does not consult with domain expert as this is very important to ensure the accuracy and correctness of the information.
Salman A. et al. [12] in their study revealed that fuzzy logic system has more benefit compare with others methods of expert system such as neural network. This is because fuzzy logic system could build a very complex linear function and can apply experts' experience without prior training process. This fuzzy expert system application facilitates the user in diagnosing dengue hemorrhaging fever via Android mobile device based on four clinical symptoms of DF. Those four criteria are fever, skin rash, spontaneous bleeding and amount of petechial appeared after tourniquet test. Users need to answer some questions provided by the application to get data for processing. In addition, rules in this study is built according to the expert researches. This expert system application capable to produce diagnosis of DHF earlier than laboratory test that require longer time to produce result. However, this research had use insufficient clinical symptoms as others clinical symptoms such as abdominal pain, headache, vomiting, and fluid accumulation are not included. It is only focused on detection of DHF which is late presentation of dengue. Late recognition of dengue diagnosis could progress to severe condition and may cause mortality.
Meanwhile, Faisal T. et al. [13] developed a non-invasive medical diagnostic system to classify the risk in dengue patients. Statistical analysis was used to determine the system's output. Two multilayer perceptron neural network models were applied to train the dataset which are Levenberg-Marquardt algorithm (LMA) or Scaled Conjugate Gradient algorithm (SCGA). Clinical symptoms, signs and Bioelectrical Impedance Analysis measurements of the dengue patients were included as the predictors for constructing the diagnostic system. As a result, SCGA contributes 75% prediction accuracy, while 70.7% prediction accuracy achieved by using LMA.
This research assists the doctors for classifying the level of risk in dengue patients. However, this research has weakness in term of its prediction accuracy which is only up to 75% and the researchers suggest to improve the systems' accuracy by using adaptive neuro-fuzzy model.
Farooqi W. et al. [14] in their study have used different algorithms of data mining technique for the efficient classification of the DF type.This is due to the increase in number of people infected by DF in Pakistan in recent years. The authors proposed to estimate the Weka tool used in classifying the DF and DHF datasets. These algorithms included Naive Bayes classifier, Decision Tree, K-nearest neighbor algorithm, and Multilayered Perceptron algorithm. These algorithms are measured based on five criteria which are accuracy, precision, sensitivity, specificity and false negative rate. The strength in this paper is the researchers provide new approach to assess the best algorithm for datasets classification in different criteria. Nevertheless, this study cannot employ the best algorithm as a practical method for predicting the dengue disease either by doctor or patient.
J. Matthews et al. [15] have concerned with usage of android technique for early diagnosis. This is due to a long distance of patients' house to the nearest medical centers, especially rural areas which may cause the patients' condition to get worse and leads to complication before the patients able to access proper medical care. The authors proposed Dengue Detector Mobile Application (DDMA) technique through a smart phone for image processing and then user can send the image to the doctor for diagnosing the disease in early stage. The creation of smartphone apps allowing patient to engage with doctors and this method can be accessed at any time on a smartphone. Although this research enables patient to access early diagnosis, there is a disadvantage for users in the rural areas because of limited internet coverage.

3. Dataset

We have obtained the dataset and done interviews with doctors who have experiences in dengue domain. Based on these interviews, we have extracted 20 rules about DF and DHF. Those rules have been converted to dataset format as CSV file and it be considered as testing dataset. The dataset consists of 20 columns representing DF symptoms and DHF symptoms. Waikato Environment for Knowledge Analysis (WEKA) software version 3.7 software has been used for training and testing two datasets where the first one represents hospital dataset and another represents doctors' interviews rules and comparing the accuracy of testing dataset and to predict accurate rules.

3.1. Training Dataset

The training dataset is a collection from 1055 instances and 18 attributes represented the diagnosis of 1054 patients during 2015 where the attributes represent the patient symptoms and the period of infection by dengue disease. The classification algorithm was selected for training the data where our dataset does not need clustering process as the datasets instances are similar in attributes. In our training, we have used Multilayer Perceptron algorithm that is a part of neural networks algorithms that distinguish from other algorithms in high accuracy. Figure 1 below shows a training dataset.

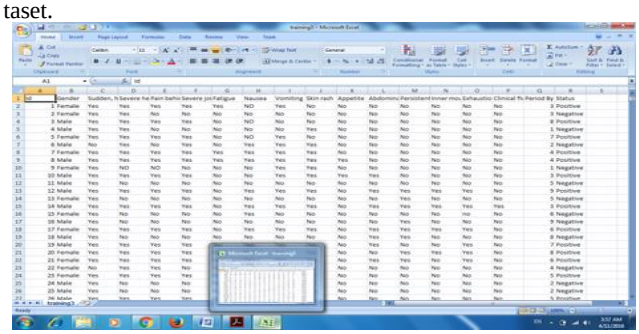


Fig. 1: Training Dataset

3.2. Testing Dataset

The testing dataset is a collection from 29 instances and 18 attributes representing the rules that have been extracted from doctors at interview sessions where the attributes represent the patient symptoms and the period of infection by dengue disease. The testing dataset has been examined byMultilayer Perceptron algorithm to produce the results. Figure 2 below shows a testing dataset used in the testing process.

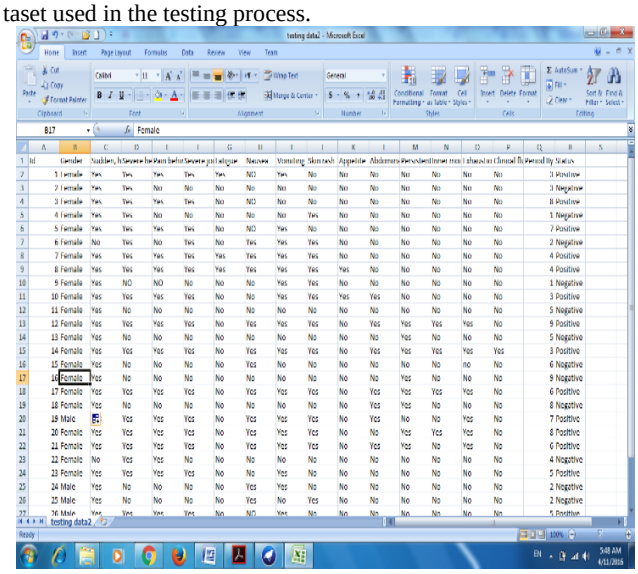


Fig. 2: Testing Dataset

4. Implementation Model

The main objective of this research is to develop an accurate system that will notify patients whether they are suspected to have DF or not by using two approaches which are fuzzy logic and data mining techniques. The system flow chart represents the general processes of this system. This flow chart provides a guideline for the dengue patients' diagnostic system and notify whether they need to seek for medical attention. It consists of several steps in this study as illustrated in Figure 3.

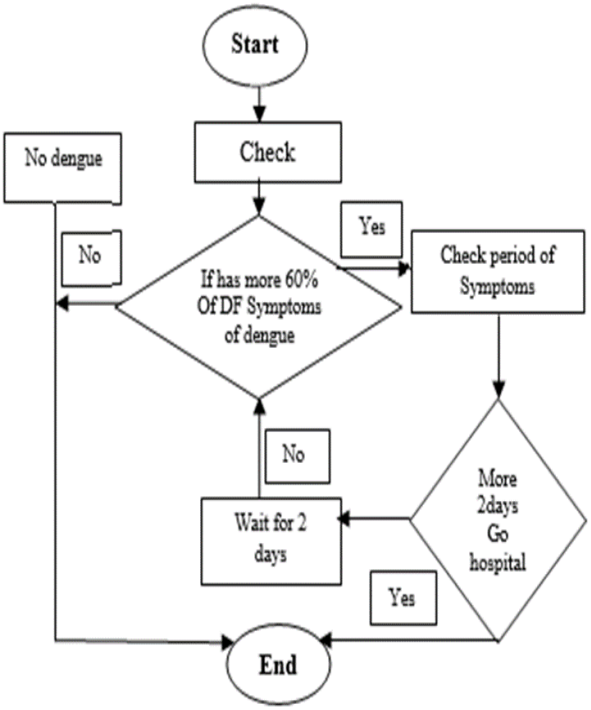


Fig. 3: System Flow Chart

There are two main algorithms in this system:

1. Dengue fever (DF) diagnosis algorithm
2. Dengue hemorrhagic fever (DHF) diagnosis algorithm

According to Figure 4 and Figure 5 below, these algorithms detail how this system analyzes the input for assuring whether these symptoms conform to dengue fever or dengue hemorrhagic fever symptoms or not.

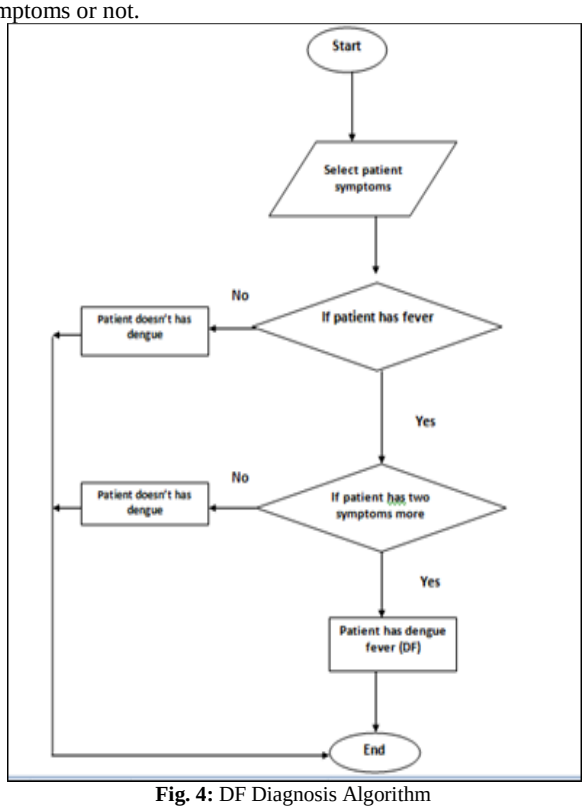


Fig. 4: DF Diagnosis Algorithm

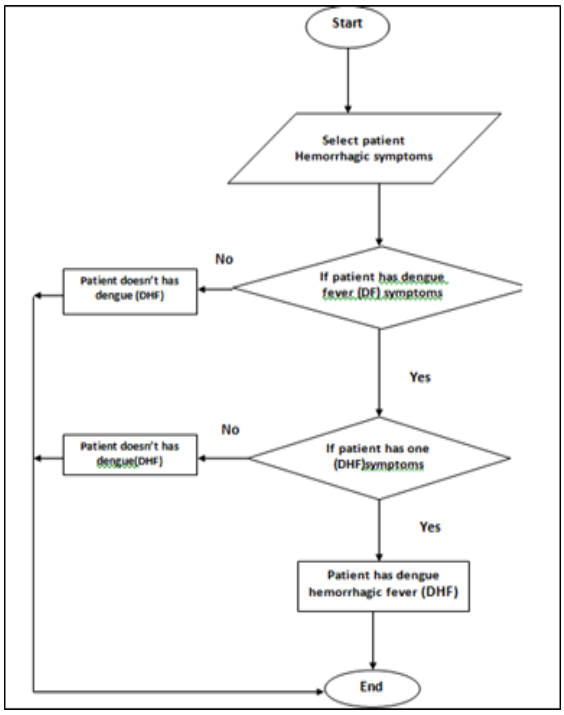


Fig. 5: DHF Diagnosis Algorithm

Figure 6 details on the system architecture in this research which consists of the following elements:

1. **Fuzzifier:** is the interface for receiving the input data from the user where this data represents patient's symptoms.

2. **Rules:** are the stored knowledge in this system as a result for doctors' experience.
3. **Inference:** is algorithm for comparing the input data with system stored rules.
4. **Defuzzifier:** is the algorithm to display the results of the comparison between input data and system rules.

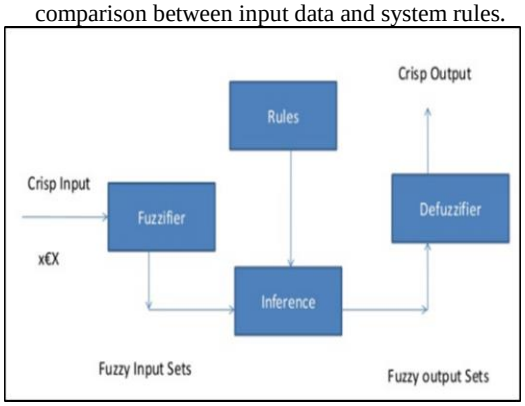


Fig. 6: System architecture

4.1. Knowledge Base

In this phase, a data mining technique is used for extracting accurate rules by converting the rules to dataset to examine data mining tool that has been trained by dengue dataset. The expert domain has been interviewed for getting the correct information about the dengue disease and how to provide the right suggestions for dengue disease. Two experienced doctors from University Malaya (UM) will be interviewed and all the information from the interview session must be converted to rule based and stored in the knowledge based. From this interview with doctors, we obtain that the diagnosis of DF should be considered in anyone who develop a fever and two others of the following symptoms which is as elaborated in Table 1 below:

Table 1: System of DF and DHF

Id	DF	Id	DHF
1	Sudden, high fever	1	Abdominal pain
2	Severe joint and muscle pain	2	Persistent vomiting
3	Vomiting	3	Clinical fluid accumulation
4	Skin rash	4	Inner mouth bleeding
5	Pain behind the eyes	5	Exhaustion
6	Fatigue		
7	Severe headaches		
8	Nausea		
9	Loss of appetite		

Moreover, there are several significant factors that should be considered when we diagnose DF. These include the incubation period of the infection and previous history of dengue infection. Then, doctors evaluate the percentage of each symptoms of DF. From this interview session, we determine the symptoms and early warning sign of DF for each stage of dengue.

Besides that, we learn that DF can be difficult to differentiate from other viral infection because they share many similar symptoms. In fact, there are instances which some of the dengue notification received by doctors at vector unit is not accurate. Thus, the doctors provided us more information on how to differentiate between those symptoms. After conducting this interview session, we have extracted 20 rules about DF and DHF. Those rules have been converted to dataset format as CSV file and it will be considered as testing dataset.

4.2. Rules

In this phase, we will obtain the information from the doctors using an expert system technique. As mentioned earlier, two experienced doctors have been interviewed to obtain the correct knowledge about dengue. Then, the general rule for dengue diagnosis was built and 20 rules about dengue were extracted. Table 2 and table 3 below describes the general rule and extracted rules to diagnose patient with DF.

Table 2: General Rule for Dengue Diagnosis

General rule:
High fever symptom + more than 1 other symptoms + Period of infection more 2 days = Positive
Else = Negative

Table 3: Extract Rules

Id	Symptoms	Status
1	High fever+ headaches+ Pain behind the eyes+ Severe joint and muscle pain+ Fatigue+ Nausea+ Vomiting+ Skin rash+ not Appetite+ Abdominal+ Inner mouth bleeding+ Exhaustion+ Period of infection more 2 days	Positive
2	High fever+ Severe joint and muscle pain+ Fatigue+ Nausea+ Vomiting+ Skin rash +Period of infection less 2 days	Positive
3	High fever+ headaches+ Pain behind the eyes + Period of infection more 2 days	Positive
4	High fever+ headaches+ Pain behind the eyes Period of infection more 2 days	Positive
5	High fever+ headaches+ Pain behind the eyes+ Severe joint + Period of infection more 2 days	Positive
6	High fever+ headaches+ Severe joint and muscle pain+ Period of infection more 2 days	Positive
7	High fever+ headaches Vomiting+ Period of infection more 2 days	Positive
8	Fever+ Vomiting+ Skin rash+ Period of infection more 2 days	Positive
9	High fever + headaches + not Appetite+ Abdominal+ Inner mouth bleeding+ Exhaustion+ Period of infection more 2 days	Positive
10	High fever+ Abdominal + Exhaustion+ Period of infection more 2 days	Positive
11	Headaches Vomiting+ Period of infection more 2 days	Negative
12	Vomiting+ Skin rash+ Period of infection more 2 days	Negative
13	Headaches + not Appetite+ Abdominal+ Inner mouth bleeding+ Exhaustion+ Period of infection more 2 days	Negative
14	Abdominal + Exhaustion+ Period of infection more 2 days	Negative
15	High fever+ Vomiting+ Skin rash+ Period of infection less 2 days	Negative
16	High fever+ headaches + Skin rash+ Period of infection less 2 days	Negative
17	High fever+ Inner mouth bleeding + Skin rash+ Period of infection less 2 days	Negative
18	Inner mouth bleeding + Skin rash+ Period of infection less 2 days	Negative
19	Inner mouth bleeding + Skin rash+ Period of infection more 2 days	Negative
20	Headaches+ Pain behind the eyes+ Severe joint and muscle pain+ Fatigue+ Nausea+ Vomiting+ Skin rash+ not Appetite+ Abdominal+ Inner mouth bleeding+ Exhaustion+ Period of infection more 2 days	Negative

4.3. Convert Rules

In this phase, the doctors evaluated each symptom in percentage values as table 4 to assess the system whether it meets the requirement for accurate diagnosis of dengue disease. In addition, it contains the evaluation of the period of the dengue infection and

the assurance whether patient has dengue in the past or not. The other significant factors that increase the infection probability by dengue are the period of infection and the previous infection of dengue. When the period of infection more than two days, the percentage of infection will increase by 10%. Besides, patient who has previous history of DF will increase percentage of infection by 5%. Those symptoms mentioned in Table 4 represent the normal DF, the other symptoms are representing the DHF. However, these DHF symptoms do not have any percentage values because they are extensions of DF symptoms.

Table 4: Symptoms in Percentage Values

Id	Symptom	Percentage value
1	Sudden, high fever	41%
2	Severe headaches	3%
3	Pain behind the eyes	5%
4	Severe joint and muscle pain	10%
5	Fatigue	3%
6	Nausea	3%
7	Vomiting	5%
8	Skin rash	5%
9	Loss of appetite	3%

At the next phase, we will convert the rules that have been extracted in the previous stage to database and we also design the suitable algorithms for dengue diagnosis. This phase consists of two tasks. The first task is system flow chart as we mentioned earlier and the second task is database design.

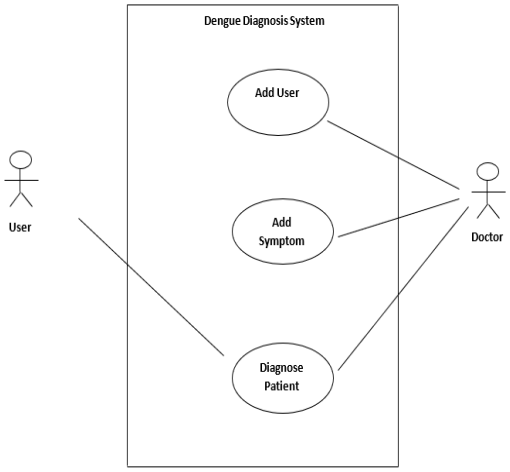


Fig. 7: Use Case Diagram

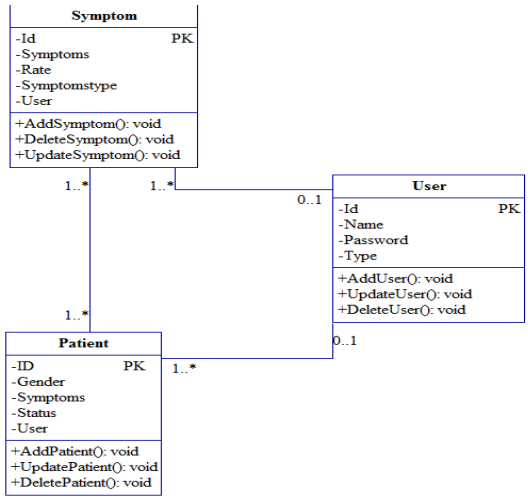


Fig. 8: Class Diagram

Database design consist of use case diagram details the system functionalities and explains the responsibilities of each user level (Figure 7) and class diagram explains the different objects of the system and the relationship between them (Figure 8).

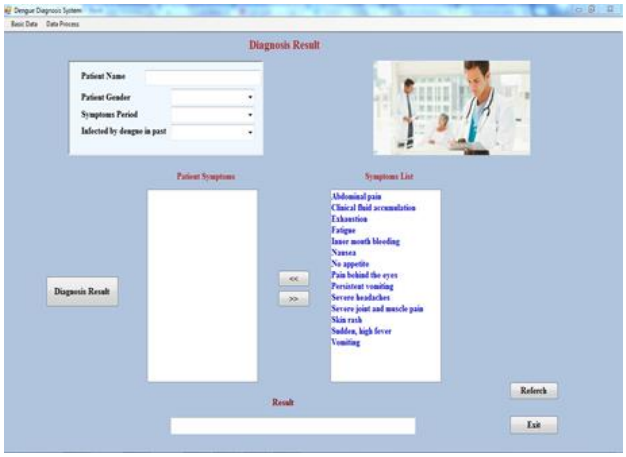


Fig. 9. Patient diagnosis interface

Figure 9 shows one of the interfaces in this project. Patient diagnosis interface is the interface where both user and doctor can diagnose the patient's condition if they suffer from dengue fever or not by choosing the symptoms in lift list to right list and determine the period of infection and determine if the patients infected by dengue in past or not from the above combo boxes. The result of the diagnosis will appear in a textbox situated at the bottom of the screen.

5. Experimental Results and Discussion

In this section, we will concern with collecting the data and examining the accuracy for the collected data. The data in this research is divided into training dataset and testing dataset, where the training dataset represents the dengue dataset that has been obtained from hospital and the testing dataset represents the extracted rules that have been obtained from doctor's interviews. Therefore, we will examine the accuracy in both datasets and the conformance between them.

Table 5: Training and Testing Result

Evaluation training dataset	Training rate	Testing rate
Correctly Classified Instances	1055 (100%)	29 (100%)
Incorrectly Classified Instances	0 (0%)	0 (0%)
Kappa statistic	1	1
Mean absolute error	0.0009	0.0009
Root mean squared error	0.0011	0.0011
Coverage of cases (0.95 level)	100%	100%
Mean rel. region size (0.95 level)	50%	50%
Total number of Instances	1055	29

Table 5 above explains the evaluation of training and testing results, where these results appear the classification of dataset close to 100% for training which means that dataset is clean and the machine has detected each attribute of the dataset. From this experiment also shows the classification of dataset for testing the doctor's rules equal to 100%, which means the result assure that our dataset is highly correct and conformable to training dataset. In addition, these results also tell us that rules that have been extracted from doctors' interviews have high accuracy.

In addition, this table explains the results in various type of error measurement such as Mean Absolute Error (MAE) with 0.0009 and the Root-Mean-Square Deviation (RMSD) with 0.0011 for both training and testing. Mean Absolute Error (MAE) is a quantity used to measure how close forecasts or predictions are to the eventual outcomes. The root-mean-square deviation (RMSD) is a frequently used measure of the differences between values (sample and population values) predicted by a model or an estimator and the values observed. The following Table 6explains the relevance between the records in training and testing dataset.

Table 6: The Relevance between the Records in Training and Testing Dataset

	TP	FP	Precision	Recall	F	ROC	Class
	1	1	1	1	1	1	+ve
	1	1	1	1	1	1	-ve
W.avg.	1	1	1	1	1	1	

W. avg. = Weight Average
+ve = Positive class
-ve =Negative class
F = F-measure

Table 6 above shows the relevant and irrelevant records that are retrieved to the total number of the records in the training and testing dataset, where these results mean all the records are relevant. The Confusion Matrix is classified as A= Positive=509 and B= Negative=546 for training and A= Positive=14 and B= Negative=15 for testing data set. In view of interpretation of the expert system, the result shows that the expert system is reliable and it is working without faults. Besides, in terms of correctness, the system could yield correct outputs as per its objective of development. The expert system is also easy to use without specific training requirements. As a conclusion, results of both training and testing datasets verify the accuracy of the rules that have been extracted from doctors' interview. Thus, these rules become reliable to be used as knowledge base in the expert system for this research.

6. Conclusion

This system has many benefits in collecting and analyzing the data by using various resources for the data. In addition, this system has used fuzzy logic neural network model based on authoritative information provided by qualified doctors for getting accurate results. Our system assists doctors to make their work easier and it provides a practical platform for infected patient to get the information whether they need to seek medical care immediately before the symptoms get worse. Furthermore, the user will find the system useful because it helps them to save their time and cost while experiencing as though they are directly getting consultation with the experts.

7. Future Work

For future research, we need to convert the expert system in this research into application in android mobile deviceto improve this research and make it more useful. This method will assist to spread the application of this service in a wider range.

Acknowledgement

The authors would like to thank Dr.AbdualkareemAlazee and Dr.AbalazeezAlrymy from Universiti Malaya for their guided and helped on evaluating the system. Recognition and gratitude is also due to University Putra Malaysia for their funding under Geran Putra IPM and other facilities provided to complete this research.

References

[1] World Health Organization, Dengue Guidelines for Diagnosis, Treatment, Prevention and Control. World Health Organization-New Edition 2009. WHO: Geneva 2009.
[2] https://www.nst.com.my/news/nation/2017/08/263438/health-minister-various-methods-adopted-tackle-spread-infectious-diseases
[3] World Health Organization, Dengue Control. http://www.who.int/denguecontrol/human/en/ [Accessed 20 Nov 2017].
[4] Suzanne Moore Shepherd, Medscape. https://emedicine.medscape.com/article/215840-overview#a3 [Accessed on 1 December 2017]
[5] World Health Organization, Dengue and Severe Dengue, 2013, http://www.who.int/mediacentre/factsheets/fs117/en/ (updated April 2017).
[6] National Guidelines for Clinical Management of DF, National Vector Bourne Disease Control Programme, 2015, http://www.nvbdcp.gov.in/Doc/Clinical%20Guidelines.pdf.
[7] C.R. Vicente, J.C.Lauar, B.S. Santos, V.M. Cobe, and C. Jr. Cerutti, Factors related to severe dengue during an epidemic in Vitória, State of Espírito Santo, Brazil, 2011, Rev Soc Bras Med Trop, 2013,Vol. 46(5), p. 629–32.
[8] Clinical Practices Guidelines Management of Dengue Infection in Adult, 2015, 3rd Edition.
[9] World Health Organization,Dengue and Dengue Haemorrhagic Fevers.WHO Fact Sheet 117,2002, http://www.who.int/inffs/en/fact117.html.
[10] T. Razak, M. Ramli, Helmi, and R. Wahab, Dengue Notification System using Fuzzy Logic, International Conference on Computer, Control, Informatics and Its Application, (IEEE, 2013), pp. 231-235.
[11] V. Pabbi, Fuzzy Expert System for Medical Diagnosis. International Journal of Scientific and Research Publication. 2015, Vol. 5, pp. 1-3.
[12] A. Salman, Y. Lina,and C. Simon,Computational Intelligence Method for Early Diagnosis DHF Using Fuzzy on Mobile Device, (European Physical Journal Conferences, 2014), Vol. 68, p. 00003.
[13] T. Faisal,M.N. Taib, and F. Ibrahim,Neural network diagnostic system for dengue patients risk classification. Journal of Medical Systems, 2012. Vol. 36, p. 661-676.
[14] W. Farooqi,and S. Ali,A Critical Study of Selected Classification Algorithms for DF and DHF, Proceedings of the 2013 11th International Conference on Frontiers of Information Technology, (IEEE, 2013), p. 140-145.
[15] J. Matthews, M. Chang,Z. Feng, R. Srinivas andM.Gerla, PowerSense: Power Aware Dengue Diagnosis on Mobile Phones, Proceedings of the First ACM Workshop on Mobile Systems, Applications and Services for Healthcare, p. 6.