Medical Diagnosis System in Healthcare Industry: A Fuzzy Approach

N Sevani1, A Setiawan1, F Saputra1, R. K. Sali1 and O Sunardi2
1Department of Informatics, Faculty of Engineering and Computer Science, Krida Wacana Christian University, Jakarta 11470, Indonesia
2Department of Industrial Engineering, Faculty of Engineering and Computer Science Krida Wacana Christian University, Jakarta 11470, Indonesia

*nina.sevani@ukrida.ac.id

Abstract. Salmonella bacterial infection often cause uncertainty during medical diagnostic phase. Two most common diseases caused by salmonella bacteria are typhus and diarrhea. This study aims to apply fuzzy inference system within medical diagnostic system so that the uncertainty of diagnostic process can be minimized. At first, a knowledge-based system was developed based on physician experience, containing 13 symptoms and 11 rules. Secondly, a web-based platform was designed as a media for physician and or patient to perform diagnostic process. Thirdly, an evaluation of the proposed system was conducted by using black box testing, white box testing, and error measurement via confusion matrix. This study found that by applying triangular membership function, Mamdani inference engine, and defuzzification centroid, the system was able to differentiate between typhus and diarrhea. Furthermore, the web-based medical diagnostic system showed an error rate of 0.3. In other words, the proposed fuzzy-based system was in line with the diagnostic result proposed by the physician.

1. Introduction
Medical diagnosis system is one form of information technology implementation in the healthcare industry. The ability to provide consistent and objective result, the scalability, and the capability to help its user in making clinical decision to improve the quality of the healthcare are some of the advantages offered by the system [1],[2],[3]. Other advantages are the possibility to share experience and also knowledge from many physician [4] and provide personalised diagnose and treatment that can empower the patients to actively participate in their health [5]. The medical diagnosis system is especially useful in an environment with limited resources such as limited physician or limited knowledge of some diseases [3].

There are several ways to diagnose diseases, such as from patient description, physical examination, and laboratory test [3]. With the rapid growth of information technology, the medical diagnosis system can facilitate these ways in one platform. The scientific advancement also promote the birth of various forms of advance medical diagnostic system. Fuzzy Inference System (FIS) is one of the scientific method in artificial intelligent that can be used in medical diagnostic system. This method is fit for an uncertainty condition that often happens in medical diagnostic process [4]. The example of uncertainty in medical diagnose arise due to the similarity of symptoms from several diseases. This condition can lead to differences in diagnostic result form different physicians [1]. Some medical diagnosis system which
using fuzzy system are medical diagnose system for kidney [1] and for typhoid [6],[7],[8],[9]. FIS also already used in diagnostic the disease in plantations [10].

The literatures show that most of the medical diagnosis system was intended for use by physician, nurse, or medical staff only. The domain of the diseases also dominated by certain diseases that can cause death, prevalent disease, or disability disease [5]. There are very limited literatures explaining the medical diagnosis system for infectious diseases caused by salmonella bacteria. There are some diseases caused by salmonella bacterial infection like typhus and diarrhea. These diseases have similar symptoms, such as gastrointestinal disorder which can often lead to misdiagnosis.

Therefore this research aim is to develop a medical diagnosis system that using FIS approach for salmonella bacterial infection. The system was build based on website to reach the wider community. The existence of this system is expected to improve the quality of healthcare services and also empower the community to care about their health. The system enables the community to store their medical record and make self assessment about their health condition. At the end, the system should be able to transform the healthcare delivery for patients and improve healthcare services.

These advantages of medical diagnosis system will lead to efficiency, safety, and better quality in healthcare industry. It is also in line with the implementation of six goals in the conceptual framework for healthcare innovation [3]. The system is expected to overcome challenges and complexity in healthcare industry.

2. Method

The medical diagnostic system was develop using fuzzy inference system (FIS). Figure 1 represents the stages in developing the system using fuzzy mamdani inference (Mamdani FIS). Mamdani FIS provides reasoning stages based on the physician knowledge to get to the conclusion [11]. The input value becomes the crisp value representing the symptoms of salmonella bacterial infection that should be answered by the user. The output would be the kind of disease caused by salmonella bacterial infection. There are two kinds of disease as the output of the system: typhus and diarrhea. The crisp output is also displayed in the form of percentage, with range from 20% to 60%. This range was determined by the physician, due to the output of the system generated only based on the clinical symptoms. Ideally the patients should take laboratory test to complete and get better diagnostic result.

![Figure 1. The stages of the development of medical diagnosis system using FIS](image)

2.1. Membership Function

Tringular membership function will be used to change the crisp value into fuzzy value. The membership range value for the input and the output of the system obtained from the result of the interview with the physician as the human expert. Equation (1) to equation (3) are the formula of membership function for the system input. Meanwhile, Equation (4) and equation (5) are the formula of membership function for the system output.
\[ \mu(x) \text{Not Sure} = \begin{cases} 
0; & \text{if } x \leq 1 \text{ or } x \geq 4 \\
\frac{x - 1}{2.5 - 1}; & \text{if } 1 < x < 2.5 \\
1; & \text{if } x = 5.5 \\
\frac{4 - x}{4 - 2.5}; & \text{if } 2.5 < x < 4,
\end{cases} \]  
\tag{1}

\[ \mu(x) \text{Maybe} = \begin{cases} 
0; & \text{if } x \leq 3 \text{ or } x \geq 8 \\
\frac{x - 3}{5.5 - 3}; & \text{if } 3 < x < 5.5 \\
1; & \text{if } x = 5.5 \\
\frac{8 - x}{8 - 5.5}; & \text{if } 5.5 < x < 8,
\end{cases} \]  
\tag{2}

\[ \mu(x) \text{Definitely} = \begin{cases} 
0; & \text{if } x \leq 7 \\
\frac{x - 7}{10 - 7}; & \text{if } 7 < x < 10 \\
1; & \text{if } x \geq 10 \\
\frac{8 - x}{8 - 5.5}; & \text{if } 5.5 < x < 8,
\end{cases} \]  
\tag{3}

\[ \mu(x) \text{Not Detected} = \begin{cases} 
0; & \text{if } x \leq 2 \text{ or } x \geq 5 \\
\frac{x - 2}{3.5 - 2}; & \text{if } 2 < x < 3.5 \\
1; & \text{if } x = 3.5 \\
\frac{5 - x}{5 - 3.5}; & \text{if } 3.5 < x < 5,
\end{cases} \]  
\tag{4}

\[ \mu(x) \text{Detected} = \begin{cases} 
0; & \text{if } x \leq 4 \text{ or } x \geq 7 \\
\frac{x - 5.5}{5.5 - 4}; & \text{if } 4 < x < 5.5 \\
1; & \text{if } x = 5.5 \\
\frac{7 - x}{7 - 5.5}; & \text{if } 5.5 < x < 7,
\end{cases} \]  
\tag{5}

where \( x \) is the crisp value or value input by the user, and \( \mu(x) \) is the membership value or fuzzy value.

2.2. Defuzzification

Defuzzification is the last step in FIS to convert the output from the inference engine into crisp value. The crisp value will be displayed to the user as the final output of the system. This research applied centroid formula for the defuzzification phase. The centroid formula is commonly used due to the prediction results are often similar with the detection result from the physician [12]. Equation (6) shows the centroid formula.

\[
\mu_{\text{Defuzzification}} = \frac{\sum \int \mu(x)(z).dx}{\sum \int \mu(x).dx},
\]  
\tag{6}

where \( \sum \int \mu(x)(z).dz \) is the moment and \( \sum \int \mu(x).dz \) is the aggregate yield area.
2.3. System Evaluation
The goal of system evaluation is to find out system failure so that it can be fixed. Three kinds of system evaluation method will be carried out in this research, namely black box testing, white box testing, and classification performance using error rate (ERR). The black box testing or input output driven testing tests the system without reference to the internal structure. It can be done by a tester without knowledge to the computer programming and its specification [13]. There were 30 respondents involved and randomly selected from the community. They were requested to try the system and answer some questions in the online questionnaire. The questions covers the ease of use of the features in the system, the system interface, and their opinion about overall system evaluation.

As the opposite, the white box testing tests the internal structure of the system to validate whether the code implementation is in line with the intended design. The tester of the white box testing must have good knowledge in the computer programming. Some factors included in the white box testing were control flow of the data and the information, the exception, and the error handling [13]. The developer of the system acted as the tester in this white box testing.

ERR were conducted to calculate the missclassification error [14]. The error calculated based on the confusion matrix. Equation (7) is the formula used for calculate the ERR.

\[
ERR = 1 - \frac{(TP + TN)}{(TP + TN + FP + FN)}
\]

where TP is True Positive, TN is True Negative, FP is False Positive, FN is False Negative. TP means that both of the system and the actual prediction from expert provide the same positive results. TN is the opposite of the TP, which mean that both of the system and the expert provide the same negative result. Positive result means that users detected with certain disease. Meanwhile, negative result means that user is not detected with certain disease. FP means that the system provides negative result but the expert provides positive result. FN occurs when the system provides positive result and the expert provides negative result.

3. Results and discussion
3.1. Knowledge base
The knowledge-base of the the medical diagnostic system for salmonella bacterial infection contained the representation of the acquired knowledge in the form of 11 rules and 13 symptoms. Table 1 shows the example of some rules stored in the knowledge-base of the system.

| No. | Rules |
|-----|-------|
| 1 | IF Bloatded stomach (Maybe) Or Bloatded stomach (Definetely) Then Diarrhea ( Not Detected ) |
| 2 | IF Frequency of defecate more than 3 times in a day (Maybe) Or Frequency of defecate more than 3 times in a day (Definetely) Then Diarrhea (Detected) |
| 3 | IF Slimy feces (Maybe) Or Slimy feces (Definetely) Then Diarrhea (Detected) |
| 4 | IF Nausea (Maybe) Or Nausea (Definetely) AND Spit up (Maybe) Or Spit up (Definetely) AND Bloatded stomach (Maybe) Or Bloatded stomach (Definetely) AND White tongue (Maybe) Or White tongue (Definetely) AND Fever (Maybe) Or Fever (Definetely) AND Heartburn (Maybe) Or Heartburn (Definetely) AND Frequency of defecate more than 3 times in a day (Not Sure) AND Lethargic (Maybe) Or Lethargic (Definetely) Then Thypus (Detected) |

3.2. Interface of the medical diagnosis system
The medical diagnosis system was build with a website platform. Figure (2) and Figure (3) show partial interface of the website (in Bahasa Indonesia). The system were also designed to provide detailed input symptoms from the users that caused by the disease and the early treatment for the disease.
3.3. System performance evaluation

The result of the black box testing showed that the website is easy to use by the user. Overall user assessed that the system is good enough and can help them to get the early detection of the infection diseases, although the interface is too ordinary.

The error rate was conduct by using 30 cases that obtained randomly with the confusion matrix of the system can be shown in Table (2). Based on the value in the confusion matrix and using Equation (7), the ERR of the system was found at 0.3.

| Table 2. Confusion matrix of the system |
|----------------------------------------|
| The actual result (from physician)    | The medical diagnosis system |
| Positive                              | Positive           | Negative       |
| Positive                              | \( TP = 21 \) cases| \( FP = 0 \)   |
| Negative                              | \( FN = 9 \) cases | \( TN = 0 \)   |
4. Conclusion
The medical diagnosis system for salmonella bacterial infection was one innovation in healthcare industry. The advantages offered by the system may be in the form of scalability and the ability to provide objective and consistent result that could lead to the improvement of services in healthcare industry. The use of FIS method could encourage the system to be capable to handle uncertainty that might emerge because of the lack of knowledge from the user of their own health condition. The evaluation result showed that the system can be used to get early diagnosis and treatment by the community who experienced salmonella bacterial infection like typhus and diarrhea. By the error rate of 0.3, the proposed system showed that it would be able to provide the diagnostic result in accordance with the result from the physician.

5. References
[1] B. Malmir, M. Amini, and S. I. Chang, “A medical decision support system for disease diagnosis under uncertainty,” Expert Syst. Appl., vol. 88, pp. 95–108, 2017.
[2] K. Farooq, B. S. Khan, M. A. Niazi, S. J. Leslie, and A. Hussain, “Clinical decision support systems: A visual survey,” Inform., vol. 42, no. 4, pp. 485–505, 2018.
[3] R. J. Conejar and H.-K. Kim, “A Medical Decision Support System (DSS) for Ubiquitous Healthcare Diagnosis System,” Int. J. Softw. Eng. Its Appl., vol. 8, no. 10, pp. 237–244, 2014.
[4] Y. Jiang, B. Qiu, C. Xu, and C. Li, “The Research of Clinical Decision Support System Based on Three-Layer Knowledge Base Model,” J. Healthc. Eng., vol. 2017, 2017.
[5] B. Martínez-Pérez, I. De La Torre-Díez, M. López-Coronado, B. Sainz-De-Abajo, M. Robles, and J. M. García-Gómez, “Mobile clinical decision support systems and applications: A literature and commercial review,” J. Med. Syst., vol. 38, no. 1, 2014.
[6] O. W. Samuel, M. O. Omosire, and B. A. Ojokoh, “A web based decision support system driven by fuzzy logic for the diagnosis of typhoid fever,” Expert Syst. Appl., vol. 40, no. 10, pp. 4164–4171, 2013.
[7] S. O. Williams and O. M. Olatunji, “Genetic-Neuro-Fuzzy System for the Diagnosis of Typhoid Fever,” no. January, pp. 1–13, 2012.
[8] O. A., A. O.B., and O. K.A., “Computer Aided Diagnostic Systems for Managing Typhoid Fever: A Review of Diagnosis Techniques,” Int. J. Comput. Appl., vol. 126, no. 6, pp. 24–29, 2015.
[9] O. N. N. Celestine, U. F. I. Abdulrahman, and I. Akpabio, “Fuzzy Expert Framework for Diagnosis of Typhoid,” Circ. Comput. Sci., vol. 2, no. 9, pp. 13–16, 2017.
[10] C. P. Diseases and E. I. Survival, VOLUME 15, vol. 15, no. 5.
[11] P. Noviayanti, C. Suhery, and D. M. Midyanti, “Jurnal Coding Sistem Komputer Untan SISTEM PAKAR UNTUK DIAGNOSA PENYAKIT RABIES PADA ANJING MENGGUNAKAN METODE Kata kunci: Sistem pakar , rabies pada anjing , Fuzzy Mamdani . Jurnal Coding Sistem Komputer Untan ISSN : 2338-493X,” vol. 05, no. 2, pp. 77–86, 2017.
[12] L. K. Wardhani and E. Haerani, “Analisis Pengaruh Pemilihan Fuzzy Membership Function Terhadap Output Sebuah Sistem Fuzzy Logic,” in Nthiki Iii 2011, 2011, pp. 326–333.
[13] M. Kumar, A. Professor, S. Kumar Singh, R. K. Dwivedi, and A. Professor, “A Comparative Study of Black Box Testing and White Box Testing Techniques,” Int. J. Adv. Res. Comput. Sci. Manag. Stud., vol. 3, no. 10, pp. 32–44, 2015.
[14] A. Tharwat, “Classification assessment methods,” Appl. Comput. Informatics, 2018.