Medical Diagnosis System in Knowledge-based Healthcare Industry: A Fuzzy Approach

by Sevani, Setiawan, Saputra, Sali and Sunardi

Medical Diagnosis System in Knowledge-based Healthcare Industry: A Fuzzy Approach

N Sevani[1]*[1], A Setiawan[1], F Saputra[1], R K Sali[1] and O Sunardi[2]
[1]*Department of Informatics, Faculty of Engineering and Computer Science, Krida Wacana Christian University, Jakarta 11470, Indonesia*
[2]*Department 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 causes uncertainty during the medical diagnostic phase. The two most common diseases caused by salmonella bacteria are typhus and diarrhea. This study aims to apply the fuzzy inference system within the medical diagnostic system so that the uncertainty of the diagnostic process can be minimized. At first, a knowledge-based system was developed based on physician experience, containing 13 symptoms and 11 rules. Second, a web-based platform was designed as a media for physicians and or patients to perform a diagnostic process. Third, an evaluation of the proposed system was conducted by using black-box testing, white box testing, and error measurement via the 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.* **Keywords:** *fuzzy, medical diagnostic, knowledge-based, healthcare industry.*

1. Introduction

The medical diagnosis system is one form of knowledge management platform in the healthcare industry. Its ability to provide a consistent and objective result and the scalability and capability to help its users in making a clinical decision are some of the advantages offered by the system [1],[2],[3]. Other advantages include the possibility to share experience and knowledge from physicians [4] and the capability to provide personalized diagnostics and treatment procedures that enable patients to actively participate in their health management [5]. The medical diagnosis system is helpful when no physician is available and or when patients have no prior knowledge about their current disease [3].

There are several methods 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 methods in one platform. The scientific advancement also promotes the birth of various forms of advanced medical diagnostic system. Fuzzy Inference System (FIS) is one of the scientific methods in artificial intelligence that can be used in the medical diagnostic system. This method is fit for an uncertainty condition that often happens in the medical diagnostic process [4]. An example of uncertainty in medical diagnose arises due to the similarity of symptoms from several diseases. This condition can lead to differences in diagnostic results form different physicians [1]. Some medical diagnosis system which using the fuzzy system is a medical diagnosis system for kidney [1] and typhoid [6],[7],[8],[9].

The literature shows that most of the medical diagnosis system was intended for use by a physician, nurse, or medical staff only. The domain of the diseases also dominated by certain types of diseases that can cause death, permanent, or disability [5]. There is very limited literature explaining the medical diagnosis system for salmonella bacteria infection, such as typhus and diarrhea. These diseases have similar symptoms, such as gastrointestinal disorder, which can often lead to misdiagnosis. Therefore, this research aims to develop a medical diagnosis system using the FIS approach to identify salmonella bacterial infection. The system was built using web-based interaction to reach a wider community. The proposed system is expected to improve the quality of healthcare services, and also to empower the community to care about their health. It enables the community to store their medical records and make a self-assessment on their health condition. In the end, it should be able to improve patients' conscience on their health management.

From the healthcare industry point of view, the advantages of the medical diagnosis system will lead to efficiency, safety, and better quality in healthcare industry services. 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 the healthcare industry.

2. Method

The medical diagnostic system was developed using Mamdani Fuzzy Inference System (Mamdani FIS). Figure 1 represents the stages in developing the system using Mamdani FIS. Mamdani FIS provides reasoning stages based on the physician's knowledge to get to the conclusion [10]. 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 diseases as the output of the system: typhus and diarrhea. The crisp output is also displayed in the form of a percentage, with a 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 tests to complete and get a better diagnostic result.

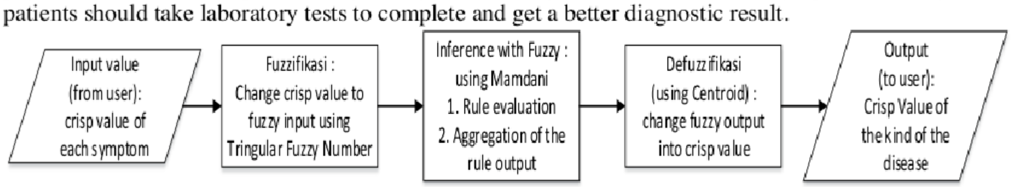

Figure 1. Stages of medical diagnosis system using FIS

2.1. *Membership Function*

Triangular membership function will be used to change the crisp value into a fuzzy value. The membership range value for the input and the output of the system was obtained from the interview with the physicians as the human expert. Equations (1) to (3) are the formulae of membership function for the system input. Meanwhile, Equation (4) and (5) are the formulae of a membership function for the system output.

$$\mu(x)Not\ Sure = \begin{cases} 0;\ if\ x\ \leq 1\ or\ x\ \geq 4 \\ \dfrac{x-1}{2.5-1};\ if\ 1 < x < 2.5 \\ 1;\ if\ x = 5.5 \\ \dfrac{4-x}{4-2.5};\ if\ 2.5 < x < 4, \end{cases} \quad (1)$$

$$\mu(x)Maybe = \begin{cases} 0;\ if\ x \le 3\ or\ x \ge 8 \\ \frac{x-3}{5.5-3};\ if\ 3 < x < 5.5 \\ 1;\ if\ x = 5.5 \\ \frac{8-x}{8-5.5};\ if\ 5.5 < x < 8, \end{cases} \tag{2}$$

$$\mu(x)Definetely = \begin{cases} 0;\ if\ x \le 7 \\ \frac{x-7}{10-7};\ if\ 7 < x < 10 \\ 1;\ if\ x \ge 10 \\ \frac{8-x}{8-5.5};\ if\ 5.5 < x < 8, \end{cases} \tag{3}$$

$$\mu(x)Not\ Detected = \begin{cases} 0;\ if\ x \le 2\ or\ x \ge 5 \\ \frac{x-2}{3.5-2};\ if\ 2 < x < 3.5 \\ 1;\ if\ x = 3.5 \\ \frac{5-x}{5-3.5};\ if\ 3.5 < x < 5, \end{cases} \tag{4}$$

$$\mu(x)Detected = \begin{cases} 0;\ if\ x \le 4\ or\ x \ge 7 \\ \frac{x-5.5}{5.5-4};\ if\ 4 < x < 5.5 \\ 1;\ if\ x = 5.5 \\ \frac{7-x}{7-5.5};\ if\ 5.5 < x < 7, \end{cases} \tag{5}$$

where x is the crisp value or the value input by the user, and μ(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 a crisp value. The crisp value will be displayed to the user as the final output of the system. This research applied a centroid formula for the defuzzification phase. The centroid formula is commonly used due to the prediction results that are often similar to the detection result from the physician [11]. Equation (6) shows the centroid formula.

$$\mu Defuzzyfikasi = \frac{\sum \int \mu x(z).zdz}{\sum \int \mu x(z).dz}, \tag{6}$$

where $\Sigma \int \mu x(z).zdz$ is the moment and $\Sigma \int \mu x(z).dz$ is the aggregate yield area.

2.3. *System Evaluation*

The goal of system evaluation is to identify system failure so that it can be fixed. Three kinds of system evaluation methods 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 the knowledge of computer programming and its specification [12]. 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 cover the ease of use of the features in the system, the system interface, and their opinion about overall system evaluation.

On 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 of computer programming. Some factors included in the white box testing were the control flow of the data and the information, the exception, and the error handling [12]. The developer of the system acted as the tester in this white box testing.

ERR was conducted to calculate the misclassification error [13]. The error calculated based on the confusion matrix. Equation (7) is the formula used to calculate the ERR.

$$ERR = 1 - [(TP + TN)/(TP + TN + FP + FN)], \quad (7)$$

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

3. Results and discussion

3.1. Knowledge base

The knowledge-base of 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 3 shows an example of some rules stored in the knowledge-base of the system.

Table 3. Rules

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

3.2. *The interface of the medical diagnosis system*

The medical diagnosis system was built with a website platform. Figure (2) and Figure (3) show the partial interface of the website (in Bahasa Indonesia). The system was also designed to provide detailed input symptoms from the users caused by the disease and the early treatment for the disease.

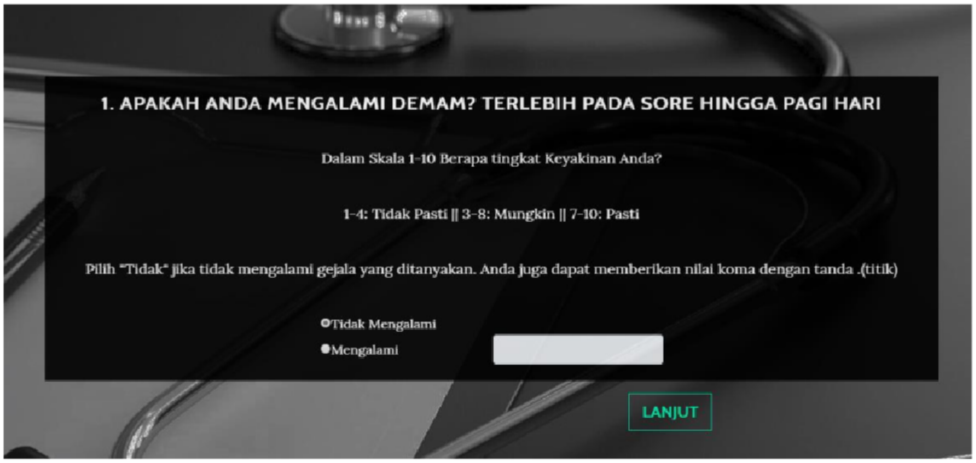

Figure 2. Interface of the diagnose menu

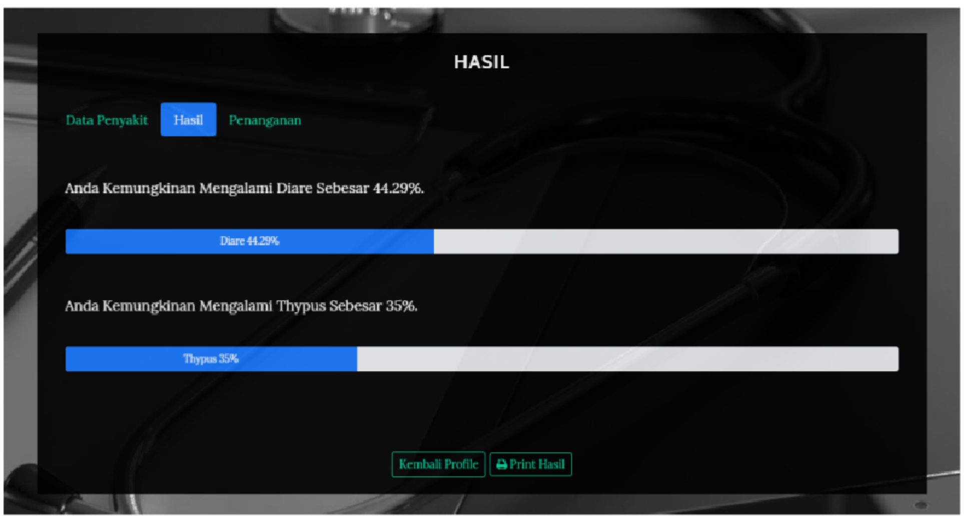

Figure 3. Interface of the diagnose result

3.3. *The system performance evaluation*

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

The error rate was conducted using 30 cases that obtained randomly. The confusion matrix of the system can be shown in Table (4). Based on the value in the confusion matrix and using Equation (7), the ERR of the system was found.

Table 4. Confussion matrix of the system

The actual result (from physicians)	*The medical diagnosis system*	
	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 the healthcare industry. The advantages offered by the system may be in the form of scalability and the ability to provide an objective and consistent result that could lead to the improvement of services in the healthcare industry. The use of the FIS method could encourage the system to be capable to handle the uncertainty that might emerge because of the lack of knowledge from the user of their 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 similar to the result from the physicians.

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, pp. 1-8, 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. Omisore, 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," vol. January, pp. 1–13, 2012.

[8] A. Oguntimilehin, O.B. Abiola, and K.A. Olatunji, "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] P. Noviyanti, C. Suhery, and D. M. Midyanti, "Sistem Pakar untuk Diagnosa Penyakit Rabies pada Anjing Menggunakan Fuzzy Mamdani," Jurnal Coding Sistem Komputer Untan, vol. 5, no. 2, pp. 77–86, 2017.

[11] L. K. Wardhani and E. Haerani, "Analisis Pengaruh Pemilihan Fuzzy Membership Function Terhadap Output Sebuah Sistem Fuzzy Logic," SNTIKI III, pp. 326–333, 2011.

[12] M. Kumar, S. K. Singh, and R. K. Dwivedi, "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.

[13] A. Tharwat, "Classification assessment methods," Appl. Comput. Informatics, article in press, 2018.

Medical Diagnosis System in Knowledge-based Healthcare Industry: A Fuzzy Approach

ORIGINALITY REPORT

5%	5%	5%	0%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

MATCH ALL SOURCES (ONLY SELECTED SOURCE PRINTED)

3%

★ Liaqat Ali, Atiqur Rahman, Aurangzeb Khan, Mingyi Zhou, Ashir Javeed, Javed Ali Khan. "An Automated Diagnostic System for Heart Disease Prediction Based on ${\chi^{2}}$ Statistical Model and Optimally Configured Deep Neural Network", IEEE Access, 2019

Publication

Exclude quotes	Off	Exclude matches	< 1%
Exclude bibliography	Off		