Design Fuzzy Expert System And Certainty Factor In Early Detection Of Stroke Disease

1st Laurentinus
Faculty of Information Technology
ISB Atma Luhur
Pangkalpinang, Indonesia
laurentinus@atmaluhur.ac.id

2nd Kiswanto
Faculty of Information Technology
ISB Atma Luhur
Pangkalpinang, Indonesia
kiswanto@atmaluhur.ac.id

3rd Rahmat Sulaiman
Faculty of Information Technology
ISB Atma Luhur
Pangkalpinang, Indonesia
rahmatsulaiman@atmaluhur.ac.id

4th Fransiskus Panca Juniawan
Faculty of Information Technology
ISB Atma Luhur
Pangkalpinang, Indonesia
fransiskus.pj@atmaluhur.ac.id

5th Dwi Yuny Sylfania
Faculty of Information Technology
ISB Atma Luhur
Pangkalpinang, Indonesia
dysylfania@atmaluhur.ac.id

6th Putri Kurniawan
Faculty of Information Technology
ISB Atma Luhur
Pangkalpinang, Indonesia
putri.kyowa@gmail.co.id

7th Harrizki Arie Pradana
Faculty of Information Technology
ISB Atma Luhur
Pangkalpinang, Indonesia
harrizkiariep@atmaluhur.ac.id

***Abstract*— With the rapid development of the world, science and technology have progressed a lot, where all require computerized work, including technology in the health sector. Health problems are one of the social issues that must be carried out, with the cost of medical treatment and doctor consultation is very expensive. It is difficult to get information and treat the disease properly and then to prevent strokes. To see someone who is at risk of strokes, of course, requires a diagnosis from a doctor or an expert in stroke disease. But unfortunately, not all of the society can support themselves by doctors or experts seeing that many strokes sufferers have to help quickly. Based on this problem, an expert system needs to diagnose or make a decision based on expert sources in a stroke to be more effective, accurate, and efficient. The system design uses the Certainty Factor Method to measure the certainty of 20 stroke symptoms and the Fuzzy Logic as logical logic in processing nine patient's medical history. In this study's discussion, getting the expert system's F-measure (F1 score) was 0.857, which indicates a good performance expert system.**

Keywords*— *expert system, certainty factor, fuzzy Tsukamoto logic, early detection, strokes

I. INTRODUCTION

Stroke is a disease characterized by brain function loss signs due to the cessation of blood supply to the brain[1]. This stroke is one of the leading causes of death and also the primary neurological disability in Indonesia. The types of risk of strokes used in this study are blood pressure, diabetes, family history, smoking, physical activity, cholesterol, height, weight, and a history of atrial fibrillation. For the detection process, the risk level is divided into high, medium, and low. Stroke is one of the third leading causes of death in the United States, with an annual incidence of 130,000 cases. In Indonesia alone, the number of stroke sufferers is increasing every year, based on the results of primary health research (Riskesdas), in 2018, stroke sufferers increased by 3.9% from 7% in 2013 to 10.9% in 2018. Lack of information and early prevention and action taken must be done on strokes patients themselves is an obstacle for people who lack awareness about health, where it is a necessary action or treatment that is fast and appropriate for stroke patients themselves[1]. To find out someone is at risk of strokes certainly requires the diagnosis of a doctor or expert on strokes, but not all can be handled alone by the doctor or expert seeing the number of strokes patients who must be dealt with quickly, it requires an expert system application that helps the work they diagnose or make a decision based on expert sources on a stroke to be more effective, accurate and efficient. The expert system is the development of artificial intelligence in the form of practical applications.

II. RELATED WORK

A literature study in this research is taken from books and journals related to strokes and prevention and rehabilitation. Some of them are research on a fuzzy expert system for heart disease diagnosis conducted [2] designed a Fuzzy Expert System for the diagnosis of heart disease. The system is designed based on the V.A. Database of Medical Centers, Long Beach, and the Cleveland Clinic Foundation. The system has 13 input fields and one output field. Input fields are types of chest pain, blood pressure, cholesterol, resting blood sugar, maximum heart rate, resting electrocardiography (ECG), exercise, prolonged peak (ST depression caused by exercise relative to rest), thallium scanning, gender, and age. The output field refers to the presence of heart disease in a patient. The results obtained from the designed system are compared with the data in the database. The results of the designed system are correct at 94%.

Research on the Classification of Strokes Risk Levels Using the Tsukamoto GA-Fuzzy Method was conducted [3] about estimating the risk of strokes using Fuzzy Logic inference, from 15 test data obtaining an accuracy of 60%. So to optimize these results, Fuzzy Tsukamoto's inference is used to optimize the membership degree function using Genetic Algorithms. The chromosome representation used in this study is the real code, which on each chromosome initializes the boundaries of all fuzzy variables. The crossover method using one cut point, then the mutation method used is random mutation, and the selection method used is elitism selection. It is known that the accuracy of the system using the Fuzzy Tsukamoto-GA method of optimizing the membership function limit is 86.66%. With the best parameters with optimal results, namely the number of pop size of 500, then many generations of 1000, and a combination of Cr = 0.5 and Mr = 0.6.

Research on Heart disease diagnosis based on mediative fuzzy logic with the aim of designing an expert system based on the type of fuzzy logic to diagnose possible heart disease for patients. The proposed system is an extension of Mamdani's fuzzy logic controller and contains 44 single input-output type rules. The system works with 11 variables as input and one variable as the output[4].

Research by [4] entitled An intuitionistic fuzzy analytics for stroke disease proposes a Fuzzy Intuitionistic Decision Tree to diagnose various types of strokes. Mapping observation data implement this approach to the Intuitionistic Fuzzy Set. This result leads to membership functions, non-membership functions, and degrees of the doubt for each record, calculated using Hamming Distance as the main requirement for Intuitionistic Fuzzy Entropy. Hamming distance calculates the difference between the values on the same variable. This approach's main advantage is that we can find out the variables affected by strokes using information obtained from Intutionistics Entropy.

Research in the development of expert systems using fuzzy has been carried out in classification of hypertension and differential symptoms based diagnosis [5] [6], predict chronic kidney disease [7] [8], diabetes [9] [10], groundwater quality[11], blood pressure [12], Rheumatic-Musculoskeletal diagnosis [13], depression [14] [15].

However, the above research only focuses on one algorithm so that it allows for wrong predictions or inaccuracy because there are so many influencing factors such as symptoms to the life history of patients and their families where this aspect requires a different theory in its calculations. From this problem, we need a system that can early diagnose strokes with different functions. The certainty factor method is a theory of certainty in calculating the patient's error symptoms and fuzzy logic, a cryptic logic to make a decision-making process based on the rules based on the patient's medical history with the results of the calculation of the certainty factor method.

III. Certainty Factor and Fuzzy Logic

A. Certainty Factor

The certainty factor method is used when facing a problem whose answer is uncertain. This uncertainty can be a probability. Short life Buchanan introduced this method in the 1970s. He uses this method when diagnosing and treating meningitis and blood infections. This method's development team noted that doctors often analyze the information provided with phrases such as 'maybe', 'almost certain'. Uncertainty is represented by the degree of confidence; while the difference is in the fuzzy logic when calculating for a rule with more than one premise. Fuzzy logic does not have a confidence value for the rule so that the calculation only looks at the smallest value for the AND operator or the most massive value for the rule. The OR operator of each premise on the rule is different from the certainty factor; that is, each rule has its belief value, not only the premises with a belief value. Certainty factor shows a measure of the certainty of a fact or rule.

$$CF[h,e] = MB[h,e] - MD[h,e] \tag{1}$$

Information :

CF[h,e] = certainty factor

MB[h,e] = measure of belief,

A measure of confidence or level of confidence in the hypothesis (h), if given evidence (e) between 0 and 1 MD [h, e] = measure of disbelief, a measure of distrust or level of confidence in the hypothesis (h), if given evidence (e) between 0 and 1. As for some combinations of certainty factors on certain premises:

1. Certainty factors with one premise.

   CF [h, e] = CF [e] * CF [rule] = CF [user] * CF [expert]

2. Certainty factors with more than one premise.

   CF [A ∧ B] = Min (CF [a], CF [b]) * CF [rule]

   CF [A ∨ B] = Max (CF [a], CF [b]) * CF [rule]

3. Certainty factors with similar conclusions. Combined CF [CF1, CF2] = CF1 + CF2 * (1 - CF1)

The advantage of this method is that it is suitable for use in expert systems that measure certain or uncertain things such as diagnosing diseases, and the calculation of this method is only valid for one calculation, and can only process two data so that its accuracy is maintained.

B. Fuzzy Logic

Fuzzy logic is a branch of the artificial intelligence system (Artificial Intelligence) that emulates the human ability to think in the form of an algorithm executed by a machine. This algorithm is used in various data processing applications that cannot be represented in binary form. Fuzzy logic interprets cryptic statements in a logical sense. Fuzzy is the language defined as fuzzy or vague, so the value can be true or false simultaneously. Fuzzy logic is a logic that has a value of ambiguity or confusion between true or false. In fuzzy logic theory, a value can be true or false simultaneously. However, how much truth and error depends on the weight of the membership it has. Fuzzy logic has a degree of membership in the range 0 (zero) to 1 (one), and fuzzy logic shows how a value is correct and the extent to which a value is false. Fuzzy logic is an appropriate way to map an input space into an output space and have a continuous value. Fuzzy is expressed in degrees of membership and degrees of truth. Therefore something can be said to be partly right and partly wrong at the same time.

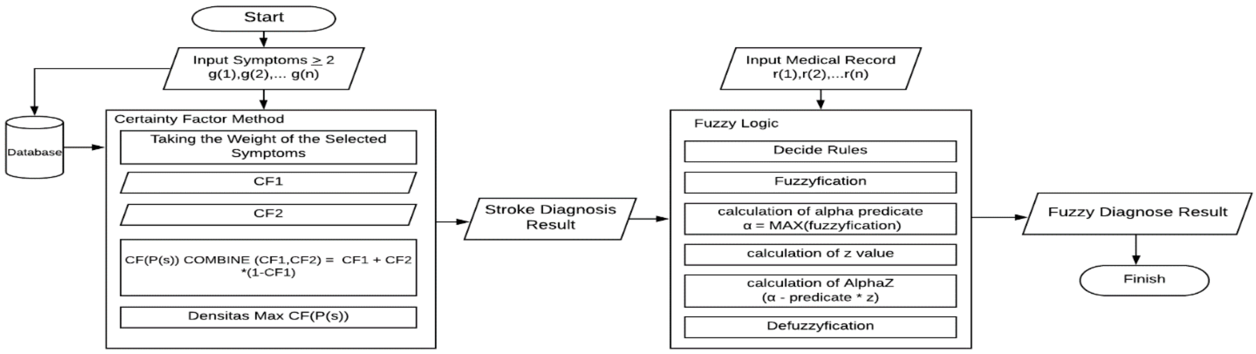

fig 1. design of Certainty factor and fuzzy logic

IV. Result and Discussion

In this study, data collection was conducted at the Bhakti Timah Hospital Pangkalpinang with one of the Neurologists in the hospital. This study's variables are the types of strokes and their factors and the symptoms of a stroke. Data obtained from experts/doctors consists of data on symptoms and diseases and their aspects. Here is a table of disease relations and signs in the expert system of strokes diagnosis:

TABLE I. RELATIONSHIP BETWEEN STROKE AND SYMPTOMS

Symptoms	**Disease**			
	1	**2**	**3**	**4**
1	✓			✓
2	✓	✓	✓	
3	✓	✓	✓	✓
4	✓		✓	
5	✓			
6	✓			
7	✓			✓
8	✓	✓		
9	✓	✓		
10	✓	✓	✓	
11	✓			
12	✓			✓
13	✓			
14	✓			
15	✓			
16	✓	✓	✓	✓
17		✓		✓
18				✓
19				✓
20		✓		

TABLE II. SYMPTOMS DATA

ID Symptoms	**Symptoms Name**
1	Hard to swallow
2	Loss of Concentration
3	Severe headache
4	Balance disorders
5	Weakness or paralysis of limbs
6	Speak vaguely
7	Numbness
8	Tingling On The Side Of The Body
9	Blurred vision / visual disturbances
10	Convulsions
11	Movement Difficult to Coordinate
12	Fainting
13	Not able to speak/write
14	Not Connecting While Talking
15	Smile Asymmetric / Turning To One Side
16	nausea and vomiting
17	Dizzy
18	Choking
19	Restless
20	Loss of consciousness

TABLE III. DISEASE DATA AND FACTORS

ID Disease	**Disease**
1	Stroke
2	High Blood Pressure / Hypertension
3	Fever
4	Heart Disease

TABLE IV. BELIEF WEIGHT VALUE SYMPTOMS OF DISEASE

Disease Name	**Symptoms**	**Belief**
Stroke	Hard to swallow	0.3
	Loss of Concentration	0.3
	Severe headache	0.5
	Balance disorders	0.3
	Weakness or paralysis of limbs	0.5
	Speak vaguely	0.5
	Numbness	0.3
	Tingling On The Side Of The Body	0.4
	Blurred vision / visual disturbances	0.3
	Movement Difficult to Coordinate	0.2
	Fainting	0.6
	Not able to speak/write	0.5
	Not Connecting While Talking	0.2
	Smile Asymmetric / Turning To One Side	0.5
	nausea and vomiting	0.5
High Blood Pressure / Hypertension	Sakit Kepala Parah	0.5
	Dizzy	0.3
	Blurred vision / visual disturbances	0.3
	nausea and vomiting	0.4
	Convulsions	0.4
	Loss of Concentration	0.3
	Tingling On The Side Of The Body	0.4
	Loss of Concentration	0.2
Fever	Severe headache	0.5
	Convulsions	0.4
	nausea and vomiting	0.4
	Balance disorders	0.3

	Loss of Concentration	0.3
Heart Disease	Restless	0.2
	Hard to swallow	0.3
	Numbness	0.3
	Dizzy	0.3
	Fainting	0.6
	Severe headache	0.5
	nausea and vomiting	0.4
	Choking	0.3

TABLE V. DATA HISTORY

No.	Name History / Other Factors	Value
1	Hypertension	Nothing
		exist
2	Diabetes Mellitus	Nothing
		exist
3	Cholesterol	Nothing
		exist
4	Weight	normal
		fat
		obesity
5	Genetic Factors	Nothing
		1 - 2
		> 3
6	Gender	Female
		Male
7	Age	< 45
		45-55
		> 55
8	Hearth Problem	Nothing
		exist
9	Respiratory disorders	Nothing
		exist
10	Addiction to addictive substances (alcohol)	Nothing
		exist

A. *Analysis of Certainty Factor Methods*

At this stage of analysis, data determination of disease, symptoms, certainty values, and calculation of certain factors will be carried out according to the data from the book and the results of interviews conducted with neurologists. Then the calculation is carried out based on the selected symptoms. The user-selected symptom has the cf value, which is used to determine the combined cf value of the 2 symptoms selected, namely convulsions and dizziness.

B. *Fuzzy Method Analysis*

Fuzzy theory is strict logic, which has true or false values. In fuzzy logic theory, a value is considered valid and false simultaneously, depending on its membership [6]. In a fuzzy set, the membership value is between 0 and 1. If x has a value of 0, it means x is not a member of set A. Likewise, if x has a value of 1, x is a full member of set A.

In this research process, the eligibility criteria for giving linguistic value are represented as input variables. At the same time, the output variable in this process is in the form of linguistic value.

1) *Variable Input*

This study's input variable is the Certainty Factor value based on weights that have a supportive value or symptoms and a list of the patient's history or medical records.

2) *Output variable*

This study's output variable is a linguistic value where the results of the linguistic value are obtained from the specified input variable.

The fuzzy membership value calculation process results are then referenced to the fuzzy rules (rules). There are two input variables (t) that need to be implemented for fuzzy rules. In determining the linguistic value given, there are four rules used :

[R1] If the diagnostic value is small and the historical value is small, then the linguistic variable is "Small".

[R2] If the diagnostic value is large and the historical value is small, then the linguistic variable is "doubtful".

[R3] If the diagnostic value is small and the historical value is large, then the linguistic variable is "Doubtful"

[R4] If the diagnostic_-value is large and the historical value is large, then the linguistic variable is "Large".

The function of variable membership in this study is divided into 2, the input variable, and the output variable. The input variable consists of symptoms and history, while the output variable is the linguistic value.

a. Symptom Variables

The Symptom variable consists of 2 fuzzy sets: SMALL and LARGE

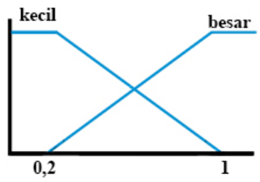

Fig 2. Symptom Variables

b. History Variable

The historical variable consists of 2 fuzzy sets: SMALL and LARGE.

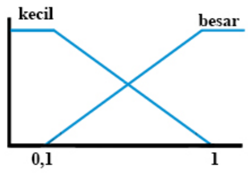

Fig 3. History Variable

c. Linguistic Value Variables

The Symptom variable consists of 2 fuzzy sets: SMALL and LARGE.

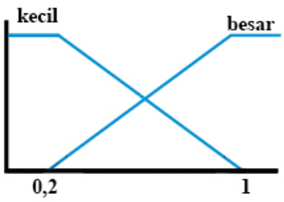

Fig 4. Linguistic Value Variables

[R1] If the diagnostic value is small and the historical value is small, then the linguistic variable is "Small"

α-predicate1=min (μ diagnosed SMALL, μ history SMALL)

z1 = (100-stroke weight) / (100-20)

[R2] If the diagnostic value is large and the historical value is small, then the linguistic variable is "Small"

α-predicate2 = min (μdigosedLARGE, μmall history)

z2 = 1 - little stroke weight

[R3] If the diagnostic value is small and the historical value is large, then the linguistic variable is "Small"

α-predicate3 = min (μdiagnosed SMALL, μMaster history)

z3 = (100-total history) / (100-10);

[R4] If the diagnostic value is large and the historical value is large, then the linguistic variable is "Small"

α-predicate 4 = min (μBig diagnosis, μBig history)

z4 =1-little stroke weight

defuzzification = αpred1 * z1 + αpred2 * z2 + αpred3 * z3 + αpred4 * z4 / (αpred1 + αpred2 + αpred3 + αpred4)

C. *User Interface*

Displays the home page, which is the main menu page that appears after the splash screen. There are four menus to choose from, there is consultation, about applications, symptom data, and disease information.

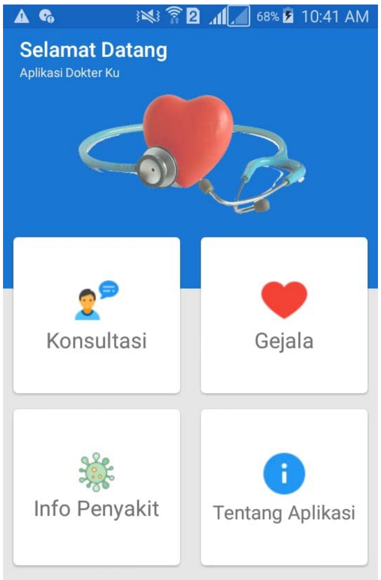

Fig 5. *Home* Screen Display

The Symptom Screen Display lists the symptoms a patient may experience.

Fig 6. Symptom Screen Display

The Diseases Screen displays a list of diseases.

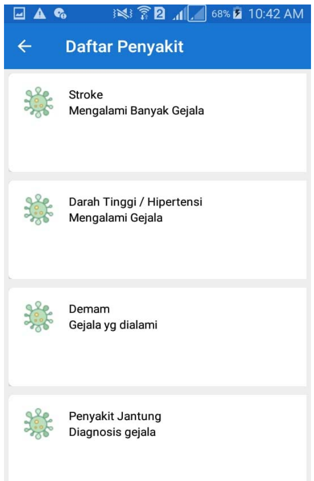

Fig 7. Diseases Result Screen Display

Displays a consultation where there is a list of symptoms to be selected by the community.

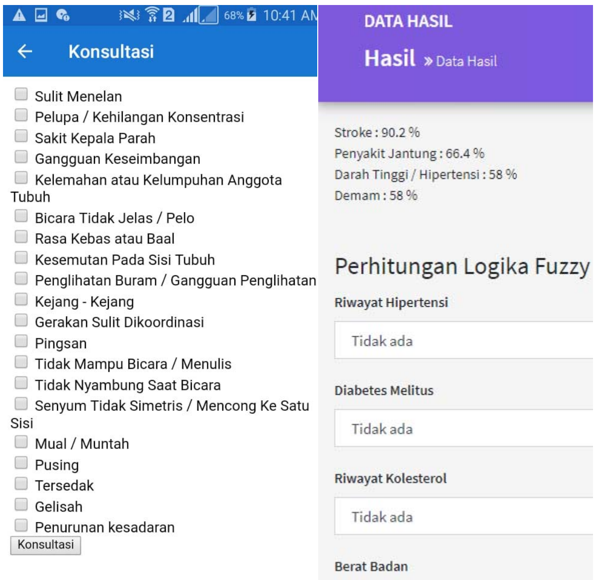

Fig 8. Consultation Display Screen

Displays the Certainty Factor Method and a Fuzzy algorithm's results page when we select a symptom and press the disease check button on the consultation page.

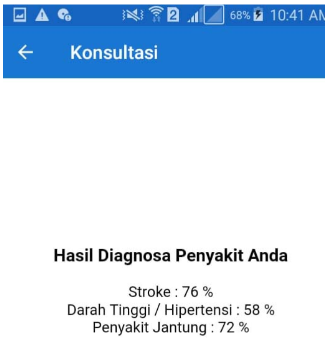

Fig 9. Consultation Result Displays

System testing is done by comparing the results with expert judgment. the results of the expert's assessment will be matched with the expert system so that the results are obtained.

The design of fuzzy logic we proposed system was evaluated using a confusion matrix that contains the real case information of the patient diagnosed by the doctor who expert in the area and the diagnosis predicted by a fuzzy expert system. A two-class classifier confusion matrix is shown in Table 1 and the result is presented in Table 2.

TABLE VI. CONFUSION MATRIX

Confusion Matrix		Predicted	
		Negative	**Positive**
Actual	Negative	TN: is the number of diagnosed cases of society don't have a disease, the result of expert system predicted correctly don't have the disease	FN: is the number of diagnosed cases of society don't have a disease, the result of expert system predicted incorrectly they do have the disease
	Positive	FP: is the number of diagnosed cases of society have a disease, the result of expert system predicted incorrectly the do not have a disease	TP: is the number of diagnosed cases of society have a disease, the result of expert system predicted correctly the do have a disease

TABLE VII. THE PROPOSED SYSTEM SYSTEM CONFUSION MATRIX

Proposed System Confusion Matrix		Predicted		Total
		Incorrect by the system	**Correct by the system**	
Actual	Incorrectly diagnosis cases	4	2	6
	Correctly diagnosis cases	2	12	14

The performance of our proposed system was generally rated using the data in the matrix. Some metrics, including accuracy, precision, sensitivity (recall), F-measure (F1 score), and specificity were applied as the criteria to implement this evaluation. In (2) to (6) show the formulas for these metrics [24] :

$$\text{Accuracy} = (TN + TP) / (TN+FP+FN+TP) \quad (2)$$

$$\text{Precision} = TP / (TP + FP) \quad (3)$$

$$\text{Recall} = TP / (TP+FN) \quad (4)$$

$$\text{Specificity} = TN / (TN+FP) \quad (5)$$

$$\text{F1 - Score} =(\text{precision x recall}) / (\text{precision} + \text{recall}) \quad (6)$$

TABLE VIII. THE ACCURACY OF THE PROPOSED SYSTEM

Variable	Result
Accuracy (A)	80%
Precision (C)	85.71%
Recall (R)	85.71%
Specificity (S)	66.67%
$F1 - Score$ (F)	0.857

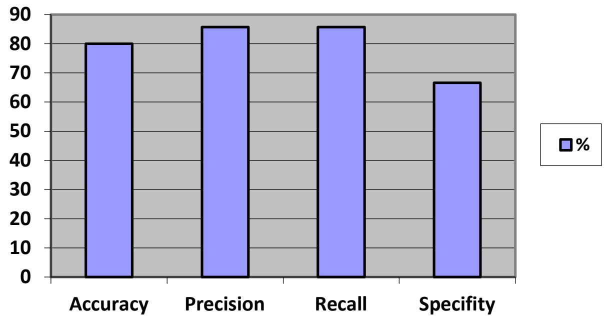

Fig 10. Persentage of expert system performance

The best value of this test is one and the worst is zero. As shown in Table VIII, the F1 score of expert system was 0.857 which indicates a good performance expert system.

V. CONCLUSION

Certainty Factor method is a theory of certainty which in its application calculates the symptoms of patients with stroke indications based on 20 symptoms and 4 diseases. The application of fuzzy logic, which is the logic of doing calculations on certainty factor results with the patient history that has 9 input variables and 1 output result. The results of the expert system test for early detection of strokes obtained results with 80% accuracy, 85.71% precision, 85.71% Recall, and 66.67% specificity, with an F1 score of 0.8571 which indicates a good accuracy of the expert system.

ACKNOWLEDGMENT

This research is being conducted and was supported by RISTEK-BRIN and research institutions and community service of ISB Atma Luhur, Thank you for the help of RISTEK-BRIN so that the research with contract number 830/SP2H/LT/MONO/LL2/2020 can be completed and published.

REFERENCES

[1] K. L. Furie *et al.*, "Guidelines for the prevention of stroke in patients with stroke or transient ischemic attack: A guideline for healthcare professionals from the American Heart Association/American stroke association," *Stroke*, 2011.

[2] L. B. Goldstein *et al.*, "Guidelines for the primary prevention of stroke: A Guideline for Healthcare Professionals from the American Heart Association/American Stroke Association," *Stroke*, 2011.

[3] W. J. Powers *et al.*, "Guidelines for the early management of patients with acute ischemic stroke: 2019 update to the 2018

guidelines for the early management of acute ischemic stroke a guideline for healthcare professionals from the American Heart Association/American Stroke Association," *Stroke*. 2019.

[4] W. J. Powers *et al.*, "2015 American Heart Association/American stroke association focused update of the 2013 guidelines for the early management of patients with acute ischemic stroke regarding endovascular treatment: A guideline for healthcare professionals from the American Heart Association/American stroke association," *Stroke*. 2015.

[5] P. Melin, I. Miramontes, and G. Prado-Arechiga, "A hybrid model based on modular neural networks and fuzzy systems for classification of blood pressure and hypertension risk diagnosis," *Expert Syst. Appl.*, 2018.

[6] M. I. Fale and Y. G. Abdulsalam, "Dr. Flynxz – A First Aid Mamdani-Sugeno-type fuzzy expert system for differential symptoms-based diagnosis," *J. King Saud Univ. - Comput. Inf. Sci.*, 2020.

[7] F. Hamedan, A. Orooji, H. Sanadgol, and A. Sheikhtaheri, "Clinical decision support system to predict chronic kidney disease: A fuzzy expert system approach," *Int. J. Med. Inform.*, 2020.

[8] S. Thukral and V. Rana, "Versatility of fuzzy logic in chronic diseases: A review," *Med. Hypotheses*, 2019.

[9] P. Rani, A. R. Mishra, and A. Mardani, "An extended Pythagorean fuzzy complex proportional assessment approach with new entropy and score function: Application in pharmacological therapy selection for type 2 diabetes," *Appl. Soft Comput. J.*, 2020.

[10] S. S. G. and M. K., "Diagnosis of diabetes diseases using optimized fuzzy rule set by grey wolf optimization," *Pattern Recognit. Lett.*, 2019.

[11] I. S. Babiker, M. A. A. Mohamed, and T. Hiyama, "Assessing groundwater quality using GIS," *Water Resour. Manag.*, 2007.

[12] R. Sharma, K. K. Deepak, P. Gaur, and D. Joshi, "An optimal interval type-2 fuzzy logic control based closed-loop drug administration to regulate the mean arterial blood pressure," *Comput. Methods Programs Biomed.*, 2020.

[13] B. A. Akinnuwesi, B. A. Adegbite, F. Adelowo, U. Ima-Edomwonyi, G. Fashoto, and O. T. Amumeji, "Decision support system for diagnosing Rheumatic-Musculoskeletal Disease using fuzzy cognitive map technique," *Informatics Med. Unlocked*, 2020.

[14] M. H. Fazel Zarandi, S. Soltanzadeh, A. Mohammadi, and O. Castillo, "Designing a general type-2 fuzzy expert system for diagnosis of depression," *Appl. Soft Comput. J.*, 2019.

[15] G. Zheng, K. Li, W. Bu, and Y. Wang, "Fuzzy comprehensive evaluation of human physiological state in indoor high temperature environments," *Build. Environ.*, 2019.

[16] L. Laurentinus and S. Rinaldi, "Implementasi Metode Analytical Hierarchy Process dan Simple Additive Weighting untuk Pemilihan Dosen Terbaik Studi Kasus STMIK Atma Luhur," *J. Teknol. Inf. dan Ilmu Komput.*, 2019.

[17] T. Mantoro, Laurentinus, N. Agani, and M. A. Ayu, "Improving the security guarantees, authenticity and confidentiality in short message service of mobile applications," in *Proceedings of 2016 4th International Conference on Cyber and IT Service Management, CITSM 2016*, 2016.

[18] C.- Kirana, H. A. Pradana, and R.- Sulaiman, "Intestine Disease Diagnosis System Using Certainty Factor Method," *Sci. J. Informatics*, 2019.