Heart Disease Diagnosis Using Tsukamoto Fuzzy Method

M. Faris Al Hakim[1*], Nafa Fajriati[1], Rizka Nur Pratama[1]

[1] Department of Computer Science, Universitas Negeri Semarang, Semarang, Indonesia
*Corresponding author: farisalhakim@mail.unnes.ac.id

ARTICLE INFO

Article history
Received 21 March 2023
Revised 31 March 2023
Accepted 1 April 2023

Keywords
Fuzzy Logic
Fuzzy Tsukamoto
Fuzzy Inference System
Heart Disease
Tsukamoto Method

ABSTRACT

As one of the leading causes of death in the world, heart disease needs special attention. Heart disease often causes sudden death because the signs of a heart attack are not easy to detect. However, early detection efforts can still be pursued and continue to be carried out, especially using information technology. This study aims to diagnose the risk level of heart disease using Tsukamoto method and involving 11 input variables such as cholesterol, blood pressure, ECG, and others. At the same time, the output variables include healthy, small, medium, large, and very large. The stages of the method consist of four main processes, namely literature review, fuzzy inference system design, applying of Tsukamoto fuzzy, and evaluation. The research concluded that the fuzzy logic of the Tsukamoto method can be used to diagnose the risk level of heart disease, although the model performance is still limited to an accuracy value of 58%.

1 Introduction

The heart is one of the organs that are vital and very important for the human body. The main role of the heart is to pump blood throughout the body and collect it after being cleaned by the lungs. Heart disease is often called the “silent killer” because, in many cases, a person is unaware of having this disease until they show signs of a heart attack or heart failure (Athiyah, 2021). Heart disease is currently listed as one of the leading causes of death in the world (Ahsan & Siddique, 2022; Gooding et al. 2020; Mohan et. al, 2019). It is not an easy matter to diagnose heart disease. This can be caused by several risk factors, such as high blood pressure, high cholesterol, diabetes, abnormal pulse rate, and various other influencing factors (Mohan et. al, 2019). World Health Organization (WHO) has determined that approximately 17.9 million deaths occur yearly due to heart disease (WHO, 2021). Therefore, early heart disease prediction is important so people with the disease can be treated before a heart attack or stroke occurs (Bahani et. al, 2021; Paul et. al, 2018).

Several previous studies regarding the diagnosis of heart disease have been carried out by (Athiyah, 2021; Mohan et. al, 2019; Putra et. al, 2019; Fiano & Purnomo, 2017). Heart disease diagnosis was done using four input variables, namely cholesterol, blood pressure, blood sugar, and body mass index (Athiyah, 2021). However, the expert knowledge base system has drawbacks. It could be more effective as a decision-making tool for doctors in providing an accurate diagnosis and treatment of heart disease due to unclear information and inaccuracy and uncertainty in decision making. Therefore, this research proposes the design, implementation, and evaluation of a fuzzy inference system using the Tsukamoto method for predicting the risk of heart disease to overcome the shortcomings of previous studies.

Along with the development of increasingly advanced technology, fuzzy logic is used for diagnosis because it has many advantages, including mathematical concepts that form the basis of reasoning, is simple but easy to understand, flexible, has tolerance for imprecise data, and can model complicated non-linear functions (Pamela et. al, 2013). A fuzzy inference system is a system for concluding some fuzzy rules (Fuzzy Rule Based) in the form of if-then. The method that can be used to get the output from the if-then rule is the Tsukamoto method. In the Tsukamoto method, Membership functions are used to represent fuzzy sets as a consequence of the if-then rule. As a result, the output of inference results from each rule is presented in the form of a crisp set based on α-predicate (fire strength). The final result is obtained by using a weighted average (Satria & Sibarani, 2020; Berlian et. al, 2020; Napitupulu et. al, 2019; Falatehan et al., 2018).

Based on this description, this study aims to help predict the level of risk of heart disease using the Fuzzy Inference System (FIS) with the Tsukamoto fuzzy method. The fuzzy rules used in this paper consist of 44 rules and 11 input variables. The fuzzy rules are obtained from an article by (Iancu, 2018). The use of fuzzy logic was chosen because it has been widely used in disease diagnosis. In addition, the Tsukamoto fuzzy method is used because it is a method that can predict and provide tolerance for fluctuating and flexible data.

2 Basic Theory

2.1 Fuzzy Inference System (FIS)

The set of reasoning and fuzzy in the form of IF-THEN is based on a computational framework called the Fuzzy Inference System (FIS) (Setyono & Aeni, 2018). Fuzzy sets are based on extending the range of feature functions so that the fuzzy set will contain real numbers in the interval [0,1]. The membership value indicates that the item in the chat universe is not just 0 or 1 but somewhere in between. In other words, the truth value of an item is not just true or false. 0 means wrong, 1 means right, and has a value between true and false (Fiano & Purnomo, 2017).

Application programs that use specific reasoning methods to produce fuzzy systems are called fuzzy inference system applications (Napitupulu et. al, 2019). Fuzzy Inference System, also called fuzzy inference engine, is a system that can evaluate all rules simultaneously to produce conclusions, and the order of the rules can be arbitrary (Sari et al, 2021). The Fuzzy Inference System can be carried out using the Tsukamoto method (Reynaldi et. al, 2021; Permadi & Alamsyah, 2020), the Mamdani method (Damayanti et al., 2022), and the Sugeno method (Sari et al., 2021; Reynaldi et. al, 2021). In this research, the fuzzy inference system method is the Tsukamoto method.

2.2 Fuzzy Tsukamoto

In the Tsukamoto method, each rule is represented by a fuzzy set with a monotonous membership function. This method is called defuzzification to determine the output value by changing the input (Zaidatunni'mah et. al, 2021). To get output in Tsukamoto fuzzy, 4 stages are needed, namely: Fuzzification, Fuzzy Rule Base, Inference Engine, and Defuzzification. Fuzzification converts system inputs with the firm or crisp values into fuzzy sets and determines the degree of membership in the fuzzy sets. The process for building Rules that will be used in the form of IF–THEN is stored in the fuzzy membership function. Converting fuzzy input into fuzzy output is fuzzifying each predefined Rule (IF-THEN Rules). MIN implication function is used to get the alpha-predicate value for each rule. After that, output of each rule (z value) is calculated using alpha-predicate value obtained. Defuzzification is giving firm or crisp values by transforming back the fuzzy output obtained from the inference engine. The final results are obtained using the average weighting equation using the average Weight Average method (Satria & Sibarani, 2020; Berlian et. al, 2020; Napitupulu et. al, 2019; Falatehan et al., 2018). Formula used to calculate defuzzification is described in Equation 1 as follows.

$$Z^* = \frac{\sum \alpha_r z_r}{\sum \alpha_r} \tag{1}$$

3 Method

3.1 Research Stage

This study used the Tsukamoto method to detect the level of risk of heart disease based on 11 disease symptoms as input variables. In this study, the steps taken were divided into four stages. The research work is described in Figure 1 as follows.

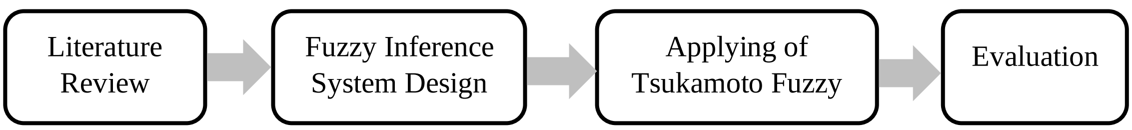

Figure 1. Research Stage

3.2 Fuzzy Inference System

This study used the Tsukamoto fuzzy method to design a *fuzzy inference system*. The steps of the Tsukamoto method used in this research consisted of four main processes that was fuzzification, determination of fuzzy rules, application of implication functions, and defuzzification. The inference process in fuzzy is described clearly by (Iancu, 2018) and shown in Figure 2 below.

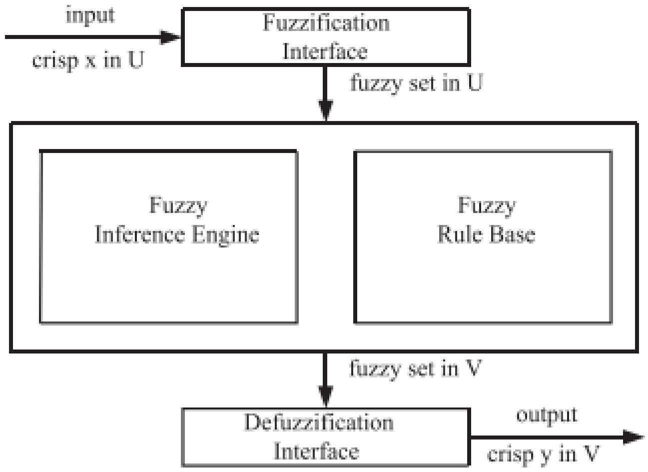

Figure 2. Fuzzy Logic System

The data were needed in the fuzzy inference system to discover the risk level of heart disease with the Tsukamoto fuzzy method include:

1) Input Variable

Input variables used in this research was also adopted from (Iancu, 2018). They were the symptoms of heart disease, including chest pain, blood pressure, cholesterol, blood sugar, ECG, maximum heart rate, exercise, old peak, gender, thallium, and age. The values of these variables were in the form of numerical and linguistic values.

2) Output Variable

The output variable was the variable that will be used as a diagnosis based on the input variable. The output variable was an integer value between 0 and 4. The greater the value of the number, the greater the risk of heart disease. The research was done by (Fiano & Purnomo, 2017), only used three risk levels. The levels are Small, Medium, and Large. The risk levels defined in this study were Healthy, Small, Medium, Large, and Very Large Levels. The output variable has several value limits, as explained in Table 1 below.

Table 1. Output variable range

Risk Level	Score
Healthy	≤ 1
Small	0-2
Medium	1-3
Large	2-4
Very Large	≥ 3

The rule base is the central part of the fuzzy inference system. The rule base used in this study was obtained from (Iancu, 2018), which consisted of 44 rules. Each rule has one input and one output. Several samples of fuzzy rule used in this research are presented in Table 2 below.

Table 2. Fuzzy Rule

Rule	Risk Level
R_1	If Chest Pain is Typical Angina then Risk is Healthy
R_5	If chest pain is without symptoms then Risk is very large
R_6	If Gender is Female then Risk is Small

R_7	If Gender is Female then Risk is Small
R_8	If Blood Pressure is Low then Healthy
R_{12}	If Blood Pressure is Very High then Risk is Very Large
R_{13}	If Cholesterol is Low then Healthy
R_{17}	If Cholesterol is Very High then Risk is Very Large
R_{18}	If Blood Sugar is True then Risk is Moderate
R_{19}	If ECG is Normal then Healthy
R_{23}	If ECG is Hypertrophy then Risk is Very Large
R_{24}	If Maximum Heart Rate is Low then Healthy
R_{28}	If the Maximum Heart Rate is High Risk is Very Large
R_{29}	If Exercise is True then Moderate Risk Level
R_{30}	If Old Peak is Low then Healthy
R_{34}	If Old Peak is Risky then Risk is Very Large
R_{35}	If Thallium is Normal then Healthy
R_{39}	If Thallium is a Reversible Defect then Risk is Very Great
R_{40}	If Age is Young then Healthy
R_{44}	If Age is Very Old then Risk is Very Large

3.3 Membership Function

The membership function of each variable was described in detail in this section.

1) Chest pain. The variable has 4 values represented by numeric value. The members were 1 for typical angina, 2 for atypical angina, 3 for non-angina, and 4 for without symptoms.
2) Blood pressure. It has 4 linguistic values, namely low, medium, high, and very high. The membership function of the blood pressure variable can be seen in Figure 3.

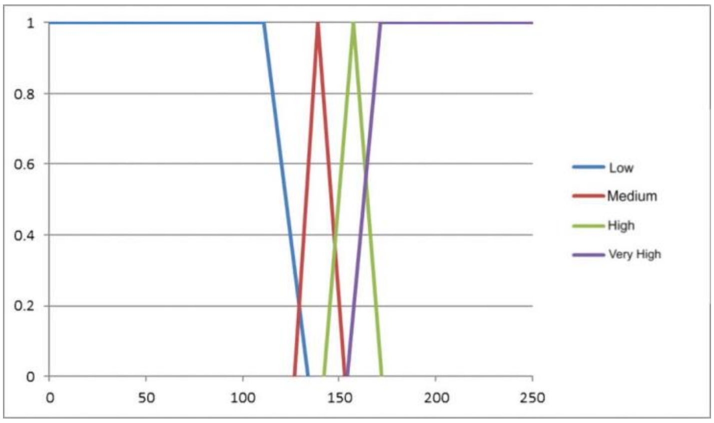

Figure 3. Membership function of blood pressure variable

3) Cholesterol. The variable has 4 linguistic values, namely low, medium, high, and very high. The membership function of the cholesterol variable can be seen in Figure 4.

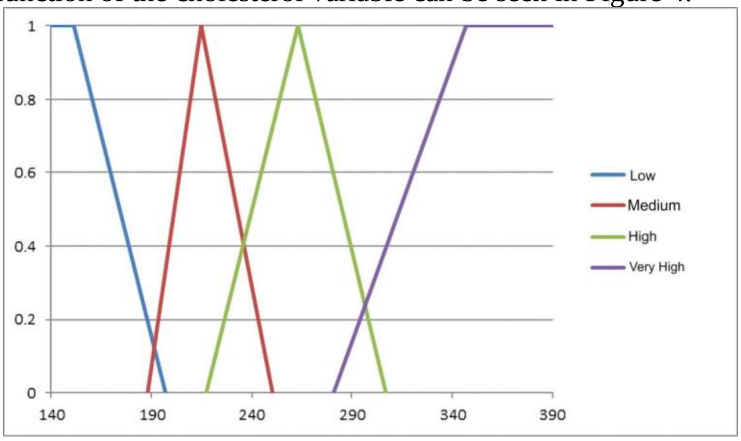

Figure 4. Membership function of cholesterol variable

4) Blood sugar. The input variable has only one linguistic value, it is True. The membership function of the blood sugar variable can be seen in Figure 5.

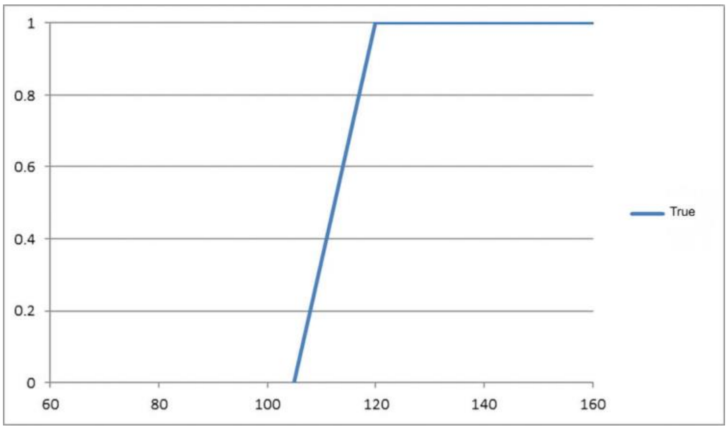

Figure 5. Membership function of blood sugar variable

5) Electrocardiography (ECG). The variable has 3 linguistic values, namely Normal, ST-T Abnormal, and Hypertrophy. The membership function of the ECG variable can be seen in Figure 6.

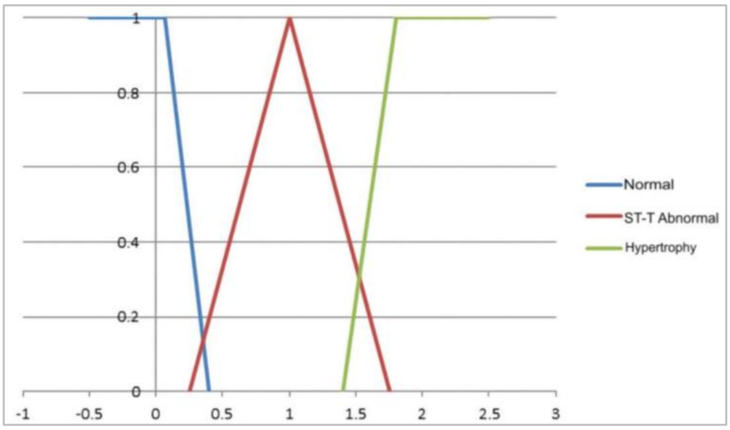

Figure 6. Membership function of ECG variable

6) Maximum heart rate. The variable has 3 linguistic values, namely low, medium, and high. The membership function of the maximum heart rate variable can be seen in Figure 7.

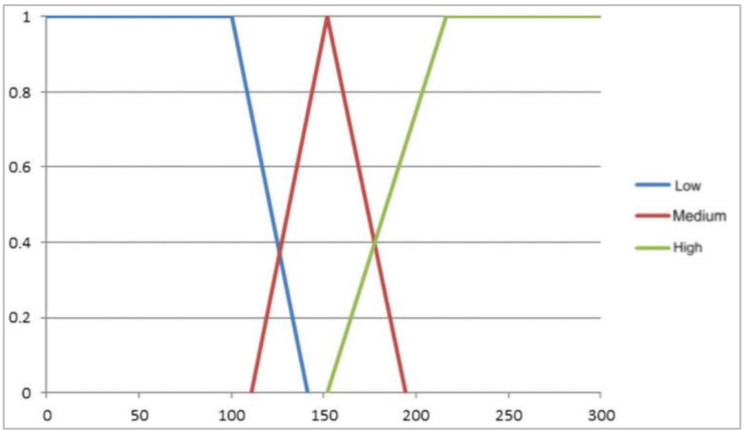

Figure 7. Membership function of maximum heart rate variable

7) Exercise. The variable has 2 values, namely 1 and 0. One (1) represents value that exercise is determined for the patient and Zero (0) represents otherwise.
8) Old peak. The variable has 3 linguistic values, namely low, risk, and horrible. The membership function of the old peak variable can be seen in Figure 8.

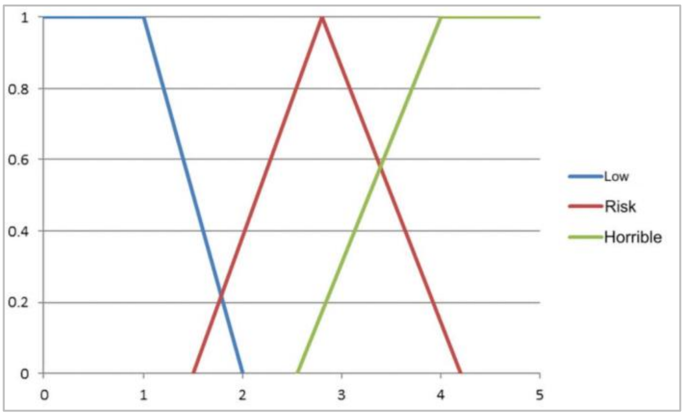

Figure 8. Membership function of old peak variable

9) Thallium. The variable has 3 numerical values, namely 3 for normal, 6 for permanent disability, and 7 for reversible disability.
10) Gender. The variable has 2 values, namely 0 for female and 1 for male.
11) Age. The variable has 4 linguistic values, namely young, midlife, old, and very old. The membership function of the age variable can be seen in Figure 9.

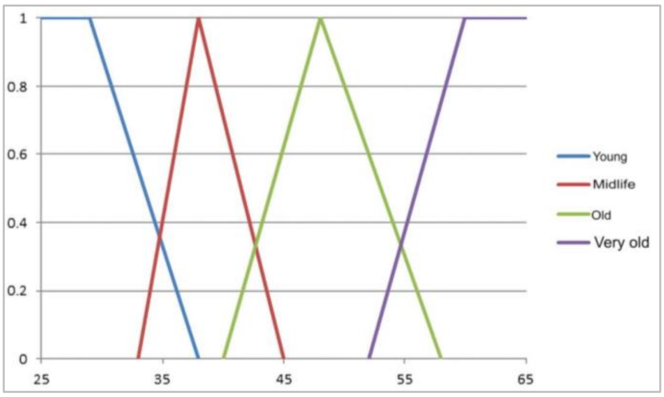

Figure 9. Membership function of age variable

12) Output variable. As mentioned before, the variable has 5 values, namely Healthy, Small, Medium, Large, and Very Large. The membership function of the output variable can be seen in Figure 10.

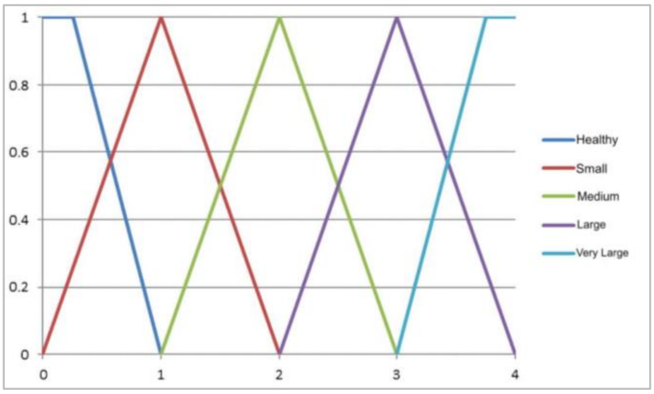

Figure 10. Membership function of output variable

4 Results & Discussion

In this stage, the method was used to calculate a sample of the patient's condition and was evaluated using more data.

4.1 Calculation of Tsukamoto Method

Calculation process was started from a sample of patient's condition as input variables and would be inferenced using Tsukamoto Method to obtain the risk level of the disease. The sample of patient's condition is given in Table 3 below.

Table 3. A sample of input variables

Input Variable	Value
Chest pain	1
Gender	1
Blood pressure	120 mmHg
Cholesterol	200 mg/dl
Blood sugar	140 mmHg
ECG	0.4
Maximum Heart Rate	150
Exercise	1
Old peak	1.5
Thallium	3
Age	36 years old

All of the input variables are applied for fuzzification. The crisp value of the input variables will be converted to fuzzy number using the membership functions of each variable. The result of calculation at each variable in fuzzification is presented in Table 4 as follows.

Table 4. Fuzzification result

Input Variable	Fuzzification
Chest pain	1
Gender	1
Blood pressure	0.6
Cholesterol	0.4
Blood sugar	1
ECG	0 (Normal) 0.2 (ST-T Abnormal)
Maximum Heart Rate	0.9
Exercise	1
Old peak	0.5 (Low) 0 (Risk)
Thallium	1
Age	0.2 (Young) 0.6 (Midlife)

The next step is finding z score for each variable by substituting α-predicate in the membership function of output variable. Based on the results in Table 4, several rules used in continuous step are R_1, R_8, R_{19}, R_{30}, R_{35}, and R_{40} for Healthy, R_{14}, R_{20}, R_{25}, R_{36}, and R_{41} for Low Risk, R_7, R_{18}, R_{21}, R_{26}, and R_{29} for Medium Risk, and R_{34} for Very Large Risk. The z score obtained from each rule are presented in Table 5.

Table 5. The result of z score

Rule	z score
R_1	0.25
R_7	2
R_8	0.55
R_{14}	0.4

R_{18}	2
R_{19}	1
R_{20}	0
R_{21}	1.2
R_{25}	0.9
R_{26}	1.9
R_{29}	2
R_{30}	0.625
R_{34}	3
R_{35}	0.25
R_{36}	1
R_{40}	0.85
R_{41}	0.6

The process is continued by defuzzification step. Equation 1 is used to calculate the result. The calculation process of defuzzification is presented in Table 6.

Table 6. Defuzzification process

Rule	α-predicate	α-predicate*z score
R_1	1	0.25
R_7	1	2
R_8	0.6	0.33
R_{14}	0.4	0.16
R_{18}	1	2
R_{19}	0	0
R_{20}	0	0
R_{21}	0.2	0.24
R_{25}	0.9	0.81
R_{26}	0.9	1.71
R_{29}	1	2
R_{30}	0.5	0.3125
R_{34}	0	0
R_{35}	1	0.25
R_{36}	1	1
R_{40}	0.2	0.17
R_{41}	0.6	0.36
∑	**10.3**	**11.5925**

The final calculation is to divide the sum of multiplication between z score and α-predicate by alpha. The result of dividing 11.5925 over 10.3 is 1.1255. The result shows that the patient has a risk level of heart disease of 1.125485. The score is included in the small risk category with membership degree 0.87 and in the medium risk category with membership degree 0.12.

4.2 Evaluation

In this stage, the method was evaluated using small number of data. The results generated would be compared with the real output of the dataset to calculate model performance. The dataset for testing is presented in Table 7.

Table 7. Dataset for Testing

Data	Input Variable											Output
	A[a]	G[b]	CP[c]	BP[d]	Ch[e]	BS[f]	ECG[g]	MHR[h]	E[i]	OP[j]	T[k]	
1	67	1	4	160	286	105	1.9	108	1	1.5	3	Medium
2	56	1	3	130	256	135	2.4	142	1	0.6	6	Medium

3	58	1	3	132	224	100	2	173	0	3.2	7	Large
4	50	0	3	120	224	90	0	158	0	1.6	3	Healthy
5	60	1	4	117	230	130	0	160	1	1.4	7	Medium
6	42	1	4	140	226	103	0.05	178	0	0	3	Healthy
7	43	1	4	120	177	95	1.9	120	1	2.5	7	Large
8	65	0	4	150	225	98	2.2	114	0	1	7	Very large
9	58	1	3	112	230	104	2.5	165	0	2.5	7	Very large
10	50	1	3	140	233	95	0	163	0	0.6	7	Small
11	60	1	3	140	185	96	2.1	155	0	3	3	Small
12	67	1	4	125	254	140	0.02	163	0	0.2	7	Large

[a]Age; [b]Gender; [c]Chest pain; [d]Blood pressure; [e]Cholesterol; [f]Blood sugar; [g]ECG; [h]Maximum Heart Rate; [i]Exercise; [j]Old peak; [k]Thallium;

By using the dataset for testing, the model performance can be seen in Table 8 below.

Table 8. Dataset for Testing

Data	Input Variable		Validation
	Dataset	Tsukamoto	
1	Medium	Medium	Yes
2	Medium	Medium	Yes
3	Large	Large	Yes
4	Healthy	Small	Not
5	Medium	Medium	Yes
6	Healthy	Medium	Not
7	Large	Large	Yes
8	Very large	Large	Not
9	Very large	Large	Not
10	Small	Small	Yes
11	Small	Medium	Not
12	Large	Large	Yes

Based on the validation results, then the calculation is carried out accuracy using a 5x5 multiclass confusion matrix where the results of the output will be classified by category. The confusion matrix is presented in Table 9.

Table 9. Confusion matrix result

Output		Risk Level				
		Healthy	Small	Medium	Large	Very large
Dataset	Healthy	0	1	1	0	0
	Small	0	1	1	0	0
	Medium	0	0	3	0	0
	Large	0	0	0	3	0
	Very Large	0	0	0	2	0

Based on Table 8, the data predicted right (true positive) by the model was 7 of 12 test data given. The model failed to predict the 4^{th}, 6^{th}, 8^{th}, 9^{th} , and 11^{th} data. The 4^{th} data which is actually healthy is predicted to be small risk. The 6^{th} data which is actually healthy is predicted to be medium risk. The 8^{th} data which is actually very large risk is predicted to be large risk. The 9^{th} data which is actually vary large risk is predicted to be large risk. The last, the 11^{th} data which is actually small risk is predicted to be medium risk. The model showed that it could diagnose heart disease using Tsukamoto method with an accuracy value of 58%.

5 Conclusion

Based on the results of the study, it was concluded that the fuzzy logic of the Tsukamoto method can be used to diagnose the risk level of heart disease by using the same components, namely 44 fuzzy rules, 11 input variables, and 5 output variables. Although, the model is capable of performing the diagnostic tasks, the model performance is still limited to an accuracy value 58%. From the testing data given, 7 data can be predicted correctly, and 5 other data failed to be predicted correctly. For further research, it is suggested that the Tsukamoto method can be improved so that it has better performance for diagnosing heart disease. Another suggestion is that input variables can be taken from other heart disease factors.

6 References

Ahsan, M.M. & Siddique, Z. (2022). Machine Learning-Based Heart Disease Diagnosis: A Systematic Literature Review. *Artificial Intelligence in Medicine, 128*, 102289.

Athiyah, U., Rosyadi, F.C.D.P., Saputra, R.A., Hekmatyar, H.D., Satrio, T.A., & Perdana, A.I. (2021). Diagnosa Resiko Penyakit Jantung Menggunakan Logika Fuzzy Metode Tsukamoto. INFOKES: Jurnal Ilmiah Rekam Medis dan Informatika Kesehatan, 11(1), 31-40.

Bahani, K., Moujabbir, M., & Ramdani, M. (2021). An Accurate Fuzzy Rule-Based Classification Systems for Heart Disease Diagnosis. Scientific African, 14, 1019.

Berlian, H.R., Hasbi, M., & Kustanto, K. (2020). Optimasi Stok Ayam Potong Menggunakan Metode Fuzzy Tsukamoto di Rumah Makan Boyolali. Jurnal TIKomSiN, 8(1), 74-84.

Damayanti, D.R., Wicaksono, S., Hakim, M.F.A., Jumanto, J., Subhan, S., & Ifriza, Y.N. (2022). Rainfall Prediction in Blora Regency Using Mamdani's Fuzzy Inference System, Journal of Soft Computing Exploration, 3(1), 62-69.

Falatehan A. I., Hidayat, N., & Brata, K. C,. (2018). Sistem Pakar Diagnosis Penyakit Hati Menggunakan Metode Fuzzy Tsukamoto Berbasis Android. Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer (JPTIIK), 2(8), 2373– 2381.

Fiano, D. S. I., & Purnomo, A. S. (2017). Sistem Pakar Untuk Mendeteksi Tingkat Resiko Penyakit Jantung Dengan Fuzzy Inferensi (Mamdani). INFORMAL: Informatics Journal, 2(2), 64-78.

Gooding, H.C., Brown, C.A., Revette, A.C., Vaccarino, V., Liu, J., Patterson, S., …, Ferranti, S.D. (2020). Young Women's Perceptions of Heart Disease Risk. Journal of Adolescent Health, 67(5), 1-6.

Iancu, I. (2018). Heart Disease Diagnosis Based on Mediative Fuzzy Logic. Artificial Intelligence in Medicine, 89, 51–60.

Mohan, S., Thirumalai, C., & Srivastava, G. (2019). Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques. IEEE Access, 7, 81542-81554.

Napitupulu, F.R., Irwansyah, M.A., & Priyanto, H. (2019). Sistem Informasi Jual Beli Rumah dengan Fitur Rekomendasi Harga Menggunakan Logika Fuzzy Tsukamoto. JEPIN (Jurnal Edukasi dan Penelitian Informatika), 5(3), 308-315.

Pamela, P.I., Gayatri, P. & Jaisankar, N. (2013). A Fuzzy Optimization Technique for the Prediction of Coronary Heart Disease Using a Decision Tree. International Journal of Engineering and Technology (IJET), 5(3), 2506-2514.

Paul, A.K., Shill, P.C., Rabin, M.R.I., & Murase, K. (2018). Adaptive Weighted Fuzzy Rule-Based System for The Risk Level Assessment of Heart Disease. Applied Intelligence volume, 48, 1739–1756.

Permadi, D.B.S. & Alamsyah, A., (2020). Application of Fuzzy Algorithms and Analytical Hierarchy Process Modification in Decision SupportSystems for Lazis Scholarship UNNES. Scientific Journal of Informatics, 7(1), 87-98.

Putra, I.G.Y.P, Khrisne, D.C., & Suyadnya, I.M.A. (2019). Expert System for Early Diagnosis of Heart Disease Using Random Forest Method. Journal of Electrical, Electronics and Informatics, 3(1), 15-18.

Reynaldi, R., Syafrizal, W., & Hakim, M.F.A. (2021). Analisis Perbandingan Akurasi Metode Fuzzy Tsukamoto dan Fuzzy Sugeno dalam Prediksi Penentuan Harga Mobil Bekas, Indonesian Journal of Mathematics and Natural Sciences, 44(2), 73-80.

Sari, F., Desyanti, D., Radillah, T., Nurjannah, S., Julimar, J., & Pakpahan, J.Y. (2021). Examining Child Obesity Risk Level Using Fuzzy Inference System. International Journal of Public Health Science (IJPHS), 10(3), 679-687.

Satria, F. & Sibarani, A.J.P. (2020). Penerapan Metode Fuzzy Tsukamoto untuk Pemilihan Karyawan Terbaik Berbasis Java Desktop. Jurnal Teknologi Informasi & Komunikasi Digital Zone, 11(1), 130-143.

Setyono, A. & Aeni, S.N. (2018). Development of Decision Support System for Ordering Goods using Fuzzy Tsukamoto. International Journal of Electrical and Computer Engineering (IJECE), 8(2), 1182-1193.

World Health Organization. (2021). Fact sheet: Cardiovascular Diseases (CVDs). Retrieved from https://www.who.int/health-topics/cardiovascular-diseases/

Zaidatunni'mah, U., Dewi, M.T., & Hakim, M.F.A. (2021). Implementation of Fuzzy Tsukamoto in Employee Performance Assessment. Journal of Soft Computing Exploration, 2(2), 143-152.