Enhanced Heart Disease Diagnosis using Fuzzy Expert System

Priyanka Kumari[1], Manish Sahu[2]

[1]*MTech Scholar,* [2]*Assistant Professor*
Department of Computer Science and Engineering
Bhabha College of Engineering, RKDF University, *Bhopal, India*

* Corresponding Author: Priyanka Kumari
Manuscript Received: Manuscript Accepted:

Abstract

Medical diagnosis systems have been widely applied to diagnosing the symptoms of diseases such as cancer and diabetes. However, the analysis tools and methods are insufficient for identifying hidden relationships in the symptoms of heart attack. Furthermore, the diagnosis of heart attack is a complex task that requires precision and effectiveness. Many alternative methods have been suggested for medical diagnosis in the healthcare domain. However, evaluating the functionality of heart attack diagnosis systems remains challenging. Therefore, this study aims to develop a system that diagnoses heart attack via fuzzy logic and evaluate the functionality of the proposed diagnostic heart attack system. This study contributes to the healthcare domain as the developed system can assist doctors in accurately diagnosing when heart attack symptoms have an ambiguous relationship. Therefore, the developed system will decrease doctor's workloads during consultations.

Keywords: Diagnosis, Expert System, Fuzzy, Heart Disease, Matlab, Patient

I. INTRODUCTION

At present, the use of traditional methods of collecting information from patients at medical centres is slowly decreasing. Meanwhile, the use of computer technology in the fields of medical diagnosis and treatment is significantly increasing [1]. In addition, data sizes are gradually increasing. Addressing these large amounts of data is a challenging task [2]. A regular database is based on Boolean logic, in which the information is either completely true or completely false [3]. Thus, the development of artificial intelligence methodology has been recognised as an important requirement in complex problem-solving situations. Medical diagnosis is a particularly effective example because of the complexity of the human mind and body, which results in a limited and vague knowledge of their functions [4]. Furthermore, medical diagnosis is a process of defining disease or explaining disorders with respect to symptoms and signs [5]. Diagnosis in the medical context is frequently regarded as implicit. The appropriate data for diagnosis are primarily obtained from the medical history and a physical examination of a patient.

Medical diagnosis is a complex task that requires precise and effective work. In addition, addressing ambiguous information in classical database management systems is challenging.

Finding convenient approaches for storing and managing human perception-based data, which is often vague and uncertain, is important in regular database systems [2]. Technology helps address problems of data inaccuracy, redundancy and loss. Furthermore, technology improves data retrieval speed [3]. Technology also helps in identifying various diseases via medical diagnosis. According to the World Health Organization [2], 12 million deaths occur annually due to heart disease. This ailment is the primary cause of death of adults. In the United States, 50% of deaths occur due to heart disease, one person dies every 34 seconds due to this health condition. Similarly, in other developed countries, heart disease is one of the main causes behind death in adults [2]. To decrease the mortality rate of heart disease, medical practitioners must diagnose it at an early stage [6]. An ambiguous data management experience in a database is important in storing ambiguous data and avoiding missing ambiguous information in the database. Ignoring ambiguous data management results in the risk of losing important information, which may be beneficial for some applications. A database that supports ambiguous, imprecise and uncertain information is called a fuzzy database and based on fuzzy logic and fuzzy set theory, which was introduced by Zadeh in the mid-1960s [7]. Fuzzy logic, which refers to a group of many-valued logic, is used to solve miscellaneous challenges, including clinical diagnostics. Fuzzy logic focuses on approximate values in place of fixed and accurate values [2].

According to Zadeh [7], fuzzy logic is an expansion of classical logic, which is based on Boolean logic. A variable in fuzzy logic can use an accurate value between 0 and 1 rather than using true or false in traditional binary sets. Furthermore, fuzzy logic delivers a method to make exact decisions on the basis of imprecise and ambiguous input data. Generally, fuzzy logic is used for applications in control systems because it closely resembles how a human makes decisions but in a faster manner [3]. This paper is structured as follows: section II discusses related work concerning heart disease diagnosis; section III reveals our proposed algorithm. Section IV demonstrates implementation of designing of fuzzy expert system along with result analysis. Finally, section V concludes the paper.

II. Related Work

Healthcare issues are a major concern. Thus, researchers have conducted many studies to propose a system especially related to diagnosing heart disease. Lavanya et al. [8] designed a fuzzy rule-based inference system for the detection and diagnosis of lung cancer. The dataset was derived from domain experts and referred to symptoms, stages and treatment facilities to provide a proficient and simple method for diagnosing lung cancer. Five variables were selected, namely, weight loss, shortness of breath, chest pain, persistent cough and blood in sputum. The diagnosis system was implemented via the software JADE and MATLAB. Kadhim et al. [9] emphasised the design and implementation of a fuzzy expert system for back pain diagnosis. The rules were established by experts, and the decision sequence was illustrated by a decision tree. This system was developed using Visual Prolog 7.1. The variables used as input for this fuzzy expert system were body mass index, age, gender and clinical observation symptoms.

Neshat et al. [10] implemented a fuzzy expert system for diagnosing liver disorders. The authors studied two cases, namely, people with healthy livers and those with unhealthy livers, and determined disease risk intensity measure. The fuzzy inference system was developed in MATLAB. These approaches are developed systems that use fuzzy logic in diagnosing disease. Most of them are based on a fuzzy inference engine via MATLAB, and one of them uses Visual Prolog 7.1 and JADE. Anbasari et al. [11] proposed a genetic algorithm that determines the attributes that contribute to the diagnosis of heart ailments, thereby indirectly reducing the number of tests needed to be taken by a patient. Thirteen attributes were reduced to six through genetic search. However, three classifiers, namely, naive Bayes, classification by clustering and decision tree were used in predicting the diagnosis of patients with the same accuracy as that with 13 attributes. The classifier (data mining) was the main method in heart disease prediction. The researchers incorporated genetic search to enhance the prediction of classifiers and used Weka 3.6.0 tools to conduct their experiment.

Taneja [12] described a classification model that sought to predict heart disease cases on the basis of patterns generated from the International Cardiovascular Hospital database. This study was conducted through two approaches. First, the data used for this study were collected from transthoracic echocardiography reports of patients; second, the classification models were developed on a sizable dataset. The researcher observed the system in real time to understand the business process of the hospital, and a Microsoft Word file was used for separate records of each patient, thereby integrating the data. A database with variables of interest was created, and the values of each variable were recorded into the new database. Hussain et al. [13] proposed a fuzzy logic-based home healthcare system for chronic heart disease patients (in stable conditions) for out-of-hospital follow-up and monitoring. The study focused on health monitoring that provided 24/7 service. The authors efficiently used a wearable battery-powered sensor unit and designed a fuzzy scheme model to diagnose the health parameters of the patient.

Krishnaiah et al. [14] reviewed data mining algorithms used for predicting heart disease. The study aided future work in finding the optimal model by providing a fast and simple analysis of different prediction models in data mining. Their work indicated that data mining tools can answer trade questions that conventionally use much time in deciding, whereas fuzzy intelligent techniques increase the accuracy of heart disease prediction systems. However, the analysis showed that different technologies use different numbers of attributes in obtaining results. Different levels of accuracy depend on the tools used for implementation.

Recently, Paul et al. [15] proposed a fuzzy decision support system for the prediction of heart disease's risk level using GAs. Using the preprocessed datasets, the model finds out the knowledge in the form of decision rules using GA. The simulation results depict that better accuracy and considered to be more efficient in the heart diseases prediction is obtained by the FDSS. For different selection methods, the system's accuracy is almost 80% for different datasets. Various studies give only a glimpse into predicting heart disease with ML techniques. Diverse data mining approaches and prediction methods, such as KNN, LR, SVM, NN, and Vote have been rather popular lately to identify and predict heart disease [20]. The novel method Vote in conjunction with a hybrid approach using LR and NB is proposed in this paper. The UCI dataset is used for conducting the experiments of the proposed method, which resulted in 87.4% accuracies in the prediction of heart disease.

In 2019, Mohan et al. [21] propose a novel method that aims at finding significant features by applying machine learning techniques resulting in improving the accuracy in the prediction of cardiovascular disease. The prediction model is introduced with different combinations of features and several known classification techniques. The authors produce an enhanced performance level with an accuracy level of 88.7% through the prediction model for heart disease with the hybrid random forest with a linear model (HRFLM).

III. PROPSED ALGORITHIM

The proposed Fuzzy Logic Controller is designed using Matlab fuzzy logic tool for Heart Disease Diagnosis which consists of 8 Linguistic Inputs and produces 1 output:

A. Input and Output Parameters

Linguistic Input parameters

I_1 Blood Pressure

I_2 Serum Cholesterol

I_3 Maximum Heart Rate

I_4 Chest Pain Type

I_5 Fasting Blood Sugar

I_6 Resting Electrocardiographic Results

I_7 Gender

I_8 Age

Output parameters

O Diagnosis of Heart Disease

B. Proposed Algorithm

INPUT

Input the fuzzy set for I_1, I_2, I_3, I_4, I_5, I_6, I_7 and I_8.

OUTPUT

Output the fuzzy set for O.

METHOD

Begin.

Step1: Create input fuzzy set I_1 (I_{11}, I_{12}, I_{13}, I_{14}), I_2 (I_{21}, I_{22}, I_{23}, I_{24}), I_3 (I_{31}, I_{32}, I_{33}), I_4 (I_{41}, I_{42}, I_{43}, I_{44}), I_5 (I_{51}, I_{52}), I_6 (I_{61}, I_{62}, I_{63}), I_7 (I_{71}, I_{72}), I_8 (I_{81}, I_{82}, I_{83}, I_{84}) and output fuzzy set O (O_1, O_2, O_3, O_4).

Step2: Built the fuzzy numbers for input fuzzy set I_1, I_2, I_3, I_4, I_5, I_6, I_7, I_8 and output fuzzy set O.

Step3: Set the triangular membership function for the fuzzy sets.

Step4: Execute the fuzzy inference by Mamdani's method.

Step5: Apply the merged rules {Rule 1,2...679}.

Step6: Compute the degree of similarity among all rules

Step7: Defuzzify it using centroid method.

End

IV. IMPLEMENTATION AND RESULT ANALYSIS

Fuzzy logic approach has been adopted in this study to estimate the health conditions of patients from their symptoms. There are four stages and seven steps involved in the implementation of fuzzy logic approach. The stages are fuzzification, fuzzy rule base, fuzzy inference engine and defuzzification. In this study, the data are obtained from UCI Machine Learning Repository: Heart Disease Data Set [16]. Eight input parameters are considered as the patient's symptoms, namely, blood pressure, serum cholesterol, maximum heart rate, chest pain type, fasting blood sugar, resting electrocardiographic results, gender and age as shown in Table I.

TABLE I: Representation of Fuzzy Variables and Fuzzy Numbers

Fuzzy Variables	Representation of Fuzzy Variables	Fuzzy Numbers	Representation of Fuzzy Numbers	Fuzzy Triangular/ Trapezoidal Numbers
Blood Pressure	I_1	Low	I_{11}	**[100 100 111 134]**
		Medium	I_{12}	**[127 139 153]**
		High	I_{13}	**[142 157 172]**
		Very High	I_{14}	**[154 171 320 320]**
Serum Cholesterol	I_2	Low	I_{21}	**[40 40 151 197]**
		Medium	I_{22}	**[188 215 250]**
		High	I_{23}	**[217 263 307]**
		Very High	I_{24}	**[281 347 681 681]**
Maximum Heart Rate	I_3	Low	I_{31}	**[70 70 100 141]**
		Medium	I_{32}	**[111 152 194]**
		High	I_{33}	**[152 216 400 400]**
Chest Pain Type	I_4	TAntiga	I_{41}	**[0 1 2]**
		ATAntiga	I_{42}	**[1 2 3]**
		NonAntiga	I_{43}	**[2 3 4]**
		Asynpt	I_{44}	**[3 4 5]**
Fasting Blood Sugar	I_5	False	I_{51}	**[-1 0 1]**
		True	I_{52}	**[0 1 2]**
Resting Electrocardiographic Results	I_6	Normal	I_{61}	**[-0.5 -0.5 0.1 0.4]**
		St-T abnormal	I_{62}	**[0.25 1 1.8]**
		Hypertrophy	I_{63}	**[1.4 2 2.5 2.5]**
Gender	I_7	Female	I_{71}	**[-1 0 1]**
		Male	I_{72}	**[0 1 2]**
Age	I_8	Young	I_{81}	**[20 20 29 38]**
		Medium	I_{82}	**[33 38 45]**
		Old	I_{83}	**[40 48 58]**
		Very Old	I_{84}	**[52 60 100 100]**
Diagnosis of Heart Disease	O	Healthy	O_1	**[-1 0 1]**
		Low Risk	O_2	**[1 1.5 2]**
		Risk	O_3	**[2 2.5 3]**
		High Risk	O_4	**[3 3.5 4]**
		Very High Risk	O_5	**[4 4.5 5]**

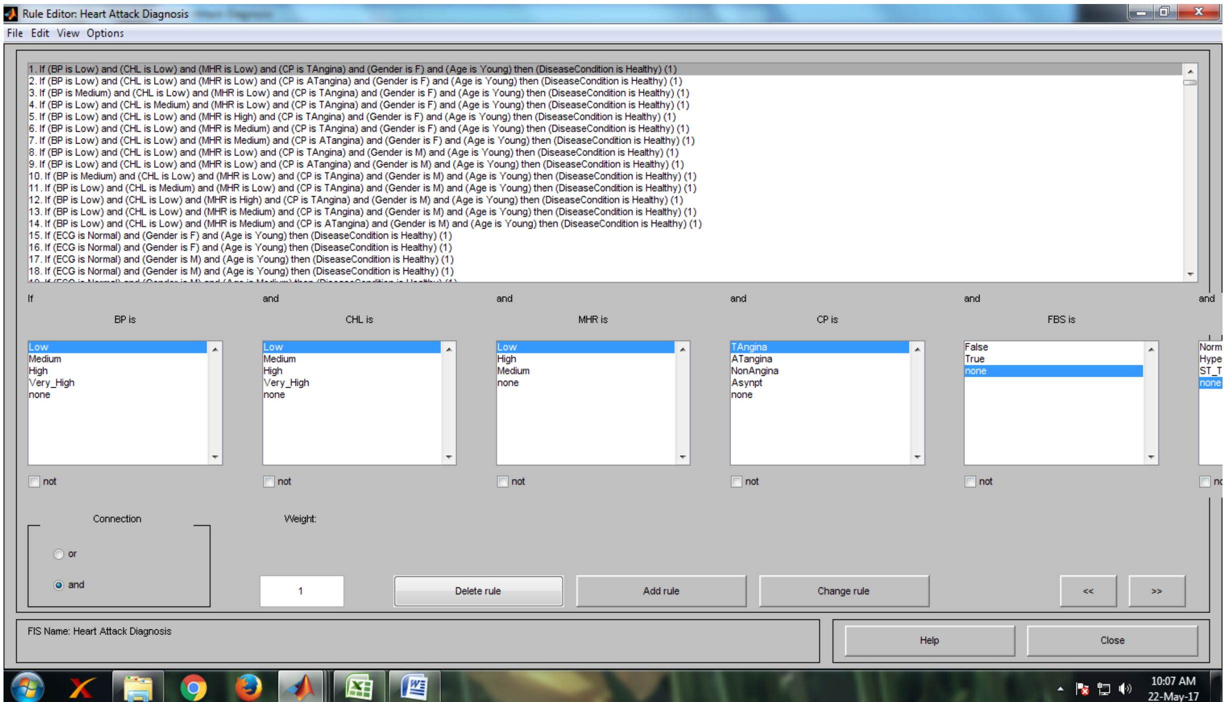

Figure 1. Rule Base Editor

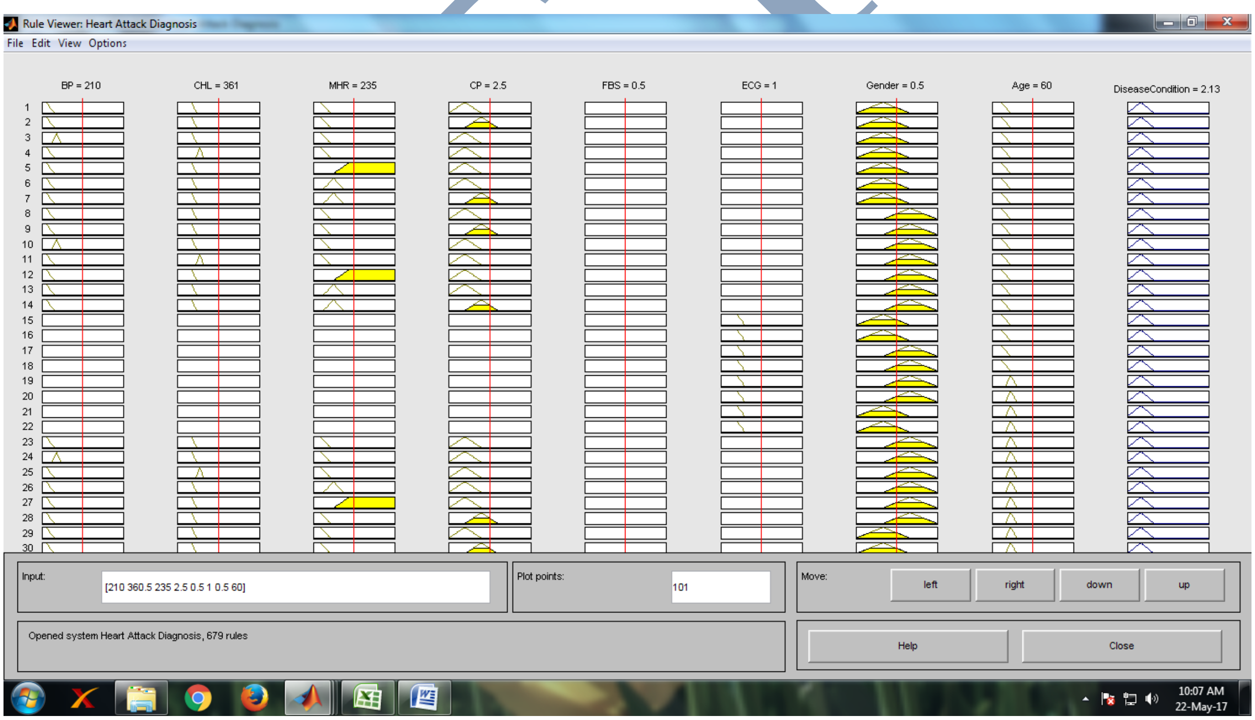

Figure 2. Fuzzy Rule Viewer

Performance Assessment Statement can be assessed based on the accuracy level. The True Positive (TP) and the True Negative (TN) denote the correct classification. False Positive (FP) is the outcome when the predicted class is yes (or positive) and actual class is no (or negative).

Still, a False Negative (FN) is the outcome when the predicted class is no (or negative) and actual class is yes (or positive). Table II lists the various outcomes of a two-class prediction. Accuracy is the proportion of the total number of predictions that were correct as shown by following formula:

$$Accuracy = \frac{TN + TP}{TN + TP + FN + FP} \times 100\%$$

TABLE II: Different Outcomes of a Two-Class Prediction

Actual Class	Predicted Class	
	YES	**NO**
YES	True Positive (TP)	False Negative (FN)
NO	False Positive (FP)	True Negative (TN)

This system tested on 200 patient databases randomly [16] and 89% accuracy has been achieved. Our experimental results are compared with Paul et al.'s algorithm in Table III and shows that the proposed system is more accurate as compare to other existing expert systems.

TABLE III: Comparison of Accuracy of Proposed Method with Previous Research

S. No.	Authors	Accuracy
1	Paul et al. [15]	**80%**
2	Bache and M. Lichman [17]	**84.5%**
3	Duch et al. [18]	**84%**
4	Duch et al. [19]	**82.8%**
5	Rao et al. [20]	**87.4%**
6	Mohan et al. [21]	**88.7%**
7	Proposed Method using Fuzzy Logic	**89%**

V. CONCLUSION

In order to diagnose any disease, the expert system developed by human may be a cheering way out to reduce time, cost, medical error and human efforts. In this paper, we have proposed Fuzzy Logic based Expert System for the diagnosis of Heart Disease, most severe disease in the present time. It makes use of the well known UCI database [16]. Through our implementation, 89% accuracy is achieved. Mamdani FIS in MatLab Fuzzy logic tool is used to design our expert system.

REFERENCES

1. A. Adeli and M. Neshat, "A fuzzy expert system for heart disease diagnosis," in Proceedings of International Multi Conference of Engineers and Computer Scientists, Hong Kong, 2010.
2. R. Parvin and A. Abhari, "Fuzzy database for heart disease diagnosis," in Proceedings of Medical Processes Modeling and Simulation (MPMS) of the 2012 Autumn Simulation Multi-Conference (SCS/AutumnSim'12), 2012.
3. R. Parvin, "Fuzzy Database for Medical Diagnosis," Department of Computer Science, Ryerson University, 2004.
4. P. R. Innocent, R. I. John, and J. M. Garibaldi, "Fuzzy methods for medical diagnosis," Applied Artificial Intelligence, vol. 19, pp. 69-98, 2004.
5. Farlex, "Medical Diagnosis," in Medical Dictionary, ed, 2009.
6. Info Kesihatan (2014, 12 November). Penyakit Jantung Pembunuh Nombor Satu. Available: http://infokesihatan4u.blogspot.my/2014/09/penyakitjantung-pembunuh-nombor-satu.html
7. L. A. Zadeh, "Fuzzy Sets," Information and Control, vol. 8, pp. 338-353, June 1965 1965.
8. K. Lavanya, M. S. Durai, and N. C. S. N. Iyengar, "Fuzzy rule based inference system for detection and diagnosis of lung cancer," International Journal of Latest Trends in Computing (EISSN: 2045-5364) Volume, vol. 2, 2011.
9. M. A. Kadhim, M. A. Alam, and H. Kaur, "Design and implementation of fuzzy expert system for back pain diagnosis," International Journal Of Innovative Technology & Creative Engineering, vol. 1, pp. 16-22, 2011.
10. M. Neshat, M. Yaghobi, M. Naghibi, and A. Esmaelzadeh, "Fuzzy expert system design for diagnosis of liver disorders," in Knowledge Acquisition and Modeling, 2008. KAM'08. International Symposium on, 2008, pp. 252-256.
11. M. Anbarasi, E. Anupriya, and N. Iyengar, "Enhanced prediction of heart disease with feature subset selection using genetic algorithm," International Journal of Engineering Science and Technology, vol. 2, pp. 53705376, 2010.
12. A. Taneja, "Heart disease prediction system using data mining techniques," Orient. J. Comput. Sci. Technol, 2013.
13. A. Hussain, R. Wenbi, Z. Xiaosong, W. Hongyang, and A. L. da Silva, "Personal home healthcare system for the cardiac patient of smart city using fuzzy logic," Journal of Advances in Information Technology Vol, vol. 7, 2016.
14. V. Krishnaiah, G. Narsimha, and N. S. Chandra, "Heart Disease Prediction System using Data Mining Techniques and Intelligent Fuzzy Approach: A Review," Heart Disease, vol. 136, 2016.
15. Animesh Kumar Paul, Pintu Chandra Shill, Md. Rafiqul Islam Rabin, M. A. H. Akhand, "Genetic Algorithm Based Fuzzy Decision Support System for the Diagnosis of Heart Disease," 5th International Conference on Informatics, Electronics and Vision, pp. 145-150, 2016.
16. K. Bache and M. Lichman, "UCI Machine Learning Repository [http://archive. ics. uci. edu/ml]. University of California, School of Information and Computer Science," Irvine, CA, 2013.
17. B. Ster and A. Dobnikar, ."Neural networks in medical diagnosis: Comparison with other methods". In: A. Bulsari et al., editor, Proceedings of the International Conference EANN '96, pages 427-430,1996.
18. Duch W, Adamczak R, Grąbczewski K, Żal G, " A hybrid method for extraction of logical rules from data", Second Polish Conference on Theory and Applications of Artificial Intelligence, pp. 61-82,1998.
19. Duch W, Grudzinski K and Diercksen G.H.F ," Minimal distance neural methods", World Congress of Computational Intelligence, Anchorage, Alaska, IEEE IJCNN'98 Proceedings, pp. 1299-1304,1998.
20. S. N. Rao, P. Shenoy M, M. Gopalakrishnan and A. Kiran B, "Applicability of the Cleveland clinic scoring system for the risk prediction of acute kidney injury after cardiac surgery in a South Asian cohort," Indian Heart J., vol. 70, no. 4, pp. 533_537, 2018.
21. Senthilkumar Mohan, Chandrasegar Thirumalai, Gautam Srivastava, "Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques," IEEE Access, pp. 81542-81554, 2019.