

Research Article

Neural Fuzzy Hybrid Rule-Based Inference System with Test Cases for Prediction of Heart Attack Probability

Rahul Kumar Jha ¹, **Santosh Kumar Henge** ¹, **Sanjeev Kumar Mandal**,¹ **Amit Sharma** ¹,
Supriya Sharma,¹ **Ashok Sharma**,² and **Afework Aemro Berhanu** ³

¹School of Computer Applications, Lovely Professional University, Phagwara, Punjab, India

²Jammu University, Jammu, India

³Department of Environmental Engineering, College of Biological and Chemical Engineering, Addis Ababa Science and Technology University, Addis Ababa, Ethiopia

Correspondence should be addressed to Rahul Kumar Jha; dr.rahulkumarjha@gmail.com and Afework Aemro Berhanu; afework.aemro@aastu.edu.et

Received 26 April 2022; Revised 6 July 2022; Accepted 21 July 2022; Published 29 September 2022

Academic Editor: Vijay Kumar

Copyright © 2022 Rahul Kumar Jha et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Heart disease has reached to the number one position in last decade in terms of mortality rate, and more wretchedly, heart attack has affected life in 80% of the cases. Cardiac arrest is an incurable incongruity that requires special treatment and cure. It has been a key research area for many years, and the number of researchers across the globe is devoted toward finding the optimal solution to avoid the ill-effect of this disease. Along with predicting heart disease, if focus moves towards prevention of heart attack as well, then this could result in major life saver area for masses. This research work is fully devoted toward finding out the probability of heart attack so that people can take preventive measure before it hit the wall. This research proposed the neural fuzzy inference system (NFIS) to represent the training data formed from the n -dimensions of functions. The NFIS consists of error computing module to improve the learning instructions when the errors have been measured, initially the membership functions are defined, and the parameters of membership functions are activated and learnt through when needed for an operation. The proposed methodology has experimented with sample test cases on Cleveland heart disease dataset from University of California Irvine (UCI) repository with the integration of supporting dependable and nondependable parameters, causing-factors, and data-matrices. This research has integration more than 13000 fuzzification rules to generate best decision-making, normalization process, planting techniques to create the feasibility to compute the heart attack probability and achieved 94 percentage of accuracy. This research can be extendable to build auto-altering and advise system with integration hardware peripheral circuit devices.

1. Introduction

Cardiovascular disease, in actual, dominating all other diseases in terms of mortality rate and heart attack is one of the biggest helpers of it which is taking almost 80% life per year. Cure of heart disease is possible, and there are many methods available to diagnose HD, for example, angiography, but the point is that they are very expensive and therefore is out of reach for masses. To overcome this, an opportunity arises to search for an alternate so that this

could be diminished from root and people can breathe coming out from such an extreme trouble. AI after its emergence and acceptance approached to invent the solution by mean of technology, and many researchers have therefore used this technology and conducted experiment using machine learning and data mining techniques to find the best solutions for such problems.

The cardiovascular disease (CVD), according to World Health Organization (WHO), holds an estimated number around 17.9 million deaths, 31% of global death [1]. There

are many parameters which impact heart and cause disease like high blood pressure, body temperature, blood sugar, cholesterol, long-term smoking, etc., and if they surpass that line, they can trigger a heart attack at any level, attempt to cause one, and often result in a person's death if they already have an illness or are in poor health. Diagnosis of heart disease before time could save millions of people, it has been a significant topic for many researchers around the world, and experimental results are phenomenal. Artificial intelligence (AI) has been vital with machine learning (ML), and with its wide range of algorithms and methods, it has proved itself and made a difference toward providing effective solution, not only for diagnosis/prediction of heart disease but also in prediction of many other critical diseases. In this research work, we tried to create a system which can find the probability of heart attack in a person to conclude if he is a normal or already suffering from heart disease. In the experiment, we used fuzzy inference system to solve this purpose and have succeeded till an extent.

1.1. Neural Networks and Fuzzy Inference System. The neural networks and fuzzy inference system work independently from each other. The co-operative NFIS uses the mechanism to learn all the parameters from fuzzy system [2]. The NFIS system can perform the inference operations based on the system of fuzzy rules with the help of prior predefined available knowledge. The human-like reasoning style of fuzzy system is incorporated by five layers of NFIS with the use of a linguistic model, and fuzzy set contains a huge set of If-Then constructed fuzzy-based rules [3–5].

NFIS can represent the training data formed from the n-dimensions of functions.

- (i) The NFIS consists of error computing module to improve the learning instructions when the errors have been measured, initially the membership functions are defined, and the parameters of membership functions are activated and learnt through when needed for an operation. As shown in Figure 1, the imprecise input values can be converted into neuron-based inputs, and then these inputs are transformed into fuzzy input to form the fuzzy sets [3–7].
- (ii) Then, the fuzzification approach has implemented with fuzzy-based inference rules $M(A)$ for prior collected hug fuzzy set formed of datasets $A(i1-in)$ based on the fuzzy membership functions and fuzzy approached rules and it derives the values of $M(B)$. By using the data in the knowledge base, the process of “fuzzification” transforms a sharp input value into a fuzzy one. The Gaussian, triangular, and trapezoidal MFs are the ones that are most frequently utilized in the fuzzification process, despite the fact that several types of curves can be encountered in literature. Embedded controllers can quickly implement these kinds of mathematical functions (MFs).

- (iii) The MFs are mathematically described using a number of parameters. These characteristics, or the geometry of the MFs, can be altered to fine-tune the performance of a fuzzy logic controller (FLC). Fuzzification provides for the linguistic expression of the system's inputs and outputs, making it possible to apply rules to complicated systems in a straightforward way.

This article scopes with the experiment, and study done so far in finding the best solution can find the probability of heart attack. The subsequent sections of paper are structured as follows: Section 2 covers the literature review of past study of fuzzy inference system and other ML-based techniques, Section 3 contains the methodology, Section 4 contains experiment and dataset details, Section 5 enlightens the result which ends up with Section 6 carrying conclusion and suggestive future work, section 7 with acknowledgment, and section 8 with the references.

2. Related Work

Many researchers in past did fantastic research on finding optimal solution for prediction of heart disease using many traditional and advance ML classification algorithms; few of them are decision tree (DT), k-nearest neighbor (KNN), support vector machine (SVM), deep neural network (DNN), random forest (RF), Naive Bayes (NB), logistic regression (LR), artificial neural network (ANN), genetic algorithms (GA), accelerated greedy additional (AGA) algorithm, network file system (NFS), but we have not found much study in finding the probability of heart attack. On dataset, UCI provided dataset; i.e., Cleveland has been preferred one while doing experiment.

Nashif et al. [8] presented a cloud-based ML system trained using Weka, a Java-based data mining tool for prediction of HD. Author claimed to develop a real-time patient-monitoring system which was capable of doing real-time sensing of body parameters such as temperature, blood pressure (BP), humidity, and heartbeat which get in sync with every 10 sec. Model was trained using UCI HD database and cross-validated using k-fold (10) cross-validation technique. Result was analyzed, and the model (SVM) recorded Accuracy (ACC), Sensitivity (SENS), and specificity (SPEC) of 97.53%, 97.50%, and 94.94%, respectively.

Haq et al. [9] presented a study using KNN, DT, RF, LR, ANN, SVM, and NB classification algorithms enabled with three feature selection methods, i.e., maximum relevance-minimum redundancy (mRMR), least absolute shrinkage and selection operator (LASSO), and relief validating using K-fold validation over popular HD dataset like Cleveland from UCI repository. Different results were obtained with different methods combined with different feature selection methods, and it was observed that applying feature reduction methods on models can reduce the execution time and improve accuracy [10].

Jha et al. [11] presented a study to compare different classification algorithms for prediction of heart disease where many classical methods such as decision tree (DT), k-nearest

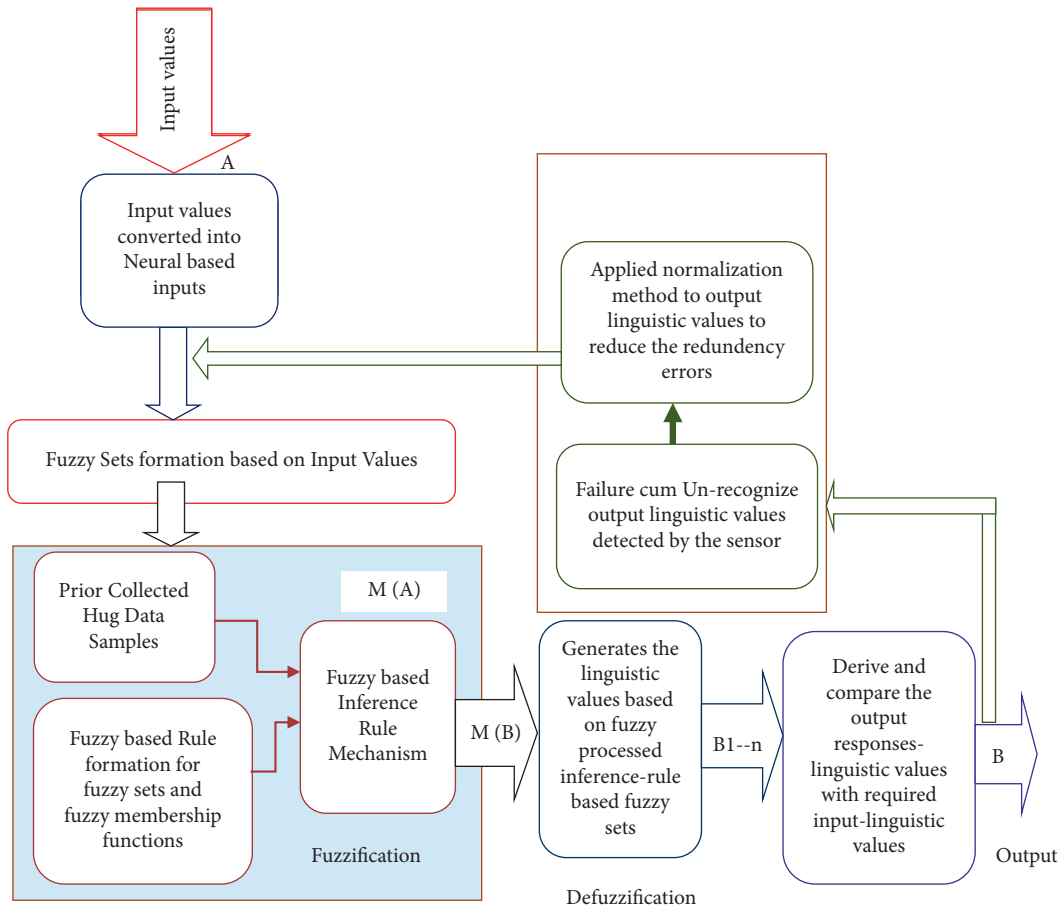


FIGURE 1: Architecture of neural fuzzy inference system.

neighbor (KNN), support vector machine (SVM), deep neural network (DNN), random forest (RF), and Naive Bayes (NB) were used applying feature selection over Rapid Miner tool to train model using Cleveland dataset from UCI repository. Result was compared among all methods experimented, and DNN outperformed among all with recorded ACC, SENS, and SPEC of 93.3%, 91.6%, and 88.4%, respectively.

Davari Dolatabadi et al. proposed a methodology for the automatic diagnosis of normal and coronary artery disease using HRV signal, extracted from ECG. Model was trained using SVM and achieved 99.2% accuracy [12].

Aghamohammadi et al. [13] had proposed a hybrid system comprising GA, ANFIS, and K-fold cross-validation for prediction of heart disease experimented over Cleveland dataset from UCI taking 14 features. Model efficiency was evaluated using ACC, SENS, and SPEC with result as 84.43%, 91.15%, and 79.16%, respectively.

Ziasabounchi and Askerzade [14] proposed an approach a hybrid system combining GA and adaptive neuro-fuzzy inference system (ANFIS) algorithms for the prediction of heart disease. Model was trained using Cleveland dataset, and performance was evaluated on the parameters ACC, SENS, SPEC, and root mean square error (RMSE). An accuracy of 92.30% was recorded with the proposed system.

Samuel et al. [15] proposed a hybrid system (fuzzy-analytical hierarchy process (fuzzy AHP) and ANN) to predict heart failure risk. In the experiment, ANN classifier was used to train the model, wherein weights were computed by fuzzy AHP system obtaining a result with 91.10% accuracy.

Kumar [16] proposed a hybrid system ANFIS having fuzzy inference system and neural network (NN) for predicting HD where training encompassed iterative tuning of parameters of the ANFIS using a hybrid learning procedure trained on UCI Cleveland dataset with MATLAB tool. This model had been segregated into 5 layers; i.e., first layer had the input variables, second, third and fourth layers were responsible for internal computation, and fifth layer resulted output of model. Model showed ACC of 91.83%.

Yazdani and Ramakrishnan [17] developed a clinical decision support system to help the doctors in predicting the risk of heart disease using optimal artificial neural network model.

Arabasadi et al. [18] presented a study with hybrid NN-GA over Z-Alizadeh Sani dataset where after applying feature selection, model with 22, 5, and 1 neurons was put in input, hidden, and output layers, respectively, undergoing experiment using feed-forward structure and showed result having ACC, SENS, and SPEC of 93.85%, 97%, and 92%, respectively. Weights were generated using GA.

Kaan and Ahmet [19] proposed a genetic algorithm-based recurrent fuzzy neural system (GARFNN) to diagnose heart disease using Cleveland heart disease dataset from UCI. Experiment was accomplished by comparing with the ANN-fuzzy AHP system, and GARFNN showed better result with 97.78% accuracy with test set as compared with ANN-fuzzy AHP which ended at 91.1%.

Reddy et al. [20] presented an experiment using hybrid adaptive genetic algorithm empowered by fuzzy logic (AGAFL) system, rough set feature selection, and fuzzy rule-based classification to predict heart disease. Experiment was performed with UCI dataset in multiple steps; i.e., first step covered feature selection from available dataset using rough set logic which was later in consecutive steps refined and trained using traditional model as well as adaptive model proposed in the study. Result shows an ACC, SENS, and SPEC of 90%, 91%, and 90%, respectively, with proposed (AGAFL) model which outperformed as compared to other experimented models.

Akgul et al. [21] presented a hybrid approach of using ANN-GA for diagnosis of HD using UCI Cleveland dataset and showed comparison in result with ANN with and without GA applied to model. Result clearly showed that ANN-GA gives better result having accuracy of 95.82% as compared with ANN without GA having accuracy of 85.02%.

Amma [22] in the conference paper presented an experiment using GA-NN to generate and optimize weight using GA and later train model using NN, calculating fitness, applying crossover, and continuing for n generation till the aim achieved. Experiment was performed on UCI datasets, and result recorded 94.17% accuracy.

Nikam et al. [23] presented a model GA-NFS consisting of GA, NN, and fuzzy set to improvise the prediction capability of the model. Study showed the capability of GA to reduce the error rate and NFS to increase the model performance. In the experiment, fitness value roll was explained to create better offspring in the next generation that could help in achieve the goal to increase in model accuracy.

The author Jabbar et al. [24] presented a study to predict heart disease combining KNN and GA over UCI repository dataset. They used GA to achieve optimal solution performing global search in the dataset and KNN to calculate model accuracy based on GA's input. Result showed that the proposed algorithm outperformed in many scenarios.

The author Jabbar et al. [25] presented a study on Weka tool using GA and association rules showing complex search capability of GA. Data were collected from Andhra Pradesh, India. Another study presented by Abdeldjouad et al. [26] experimented on various ML hybrid techniques including LR, MOEFC, FURIA, GFSLB, and Vote using Keel and Weka software. Vote outperformed with 80.2% accuracy.

The authors Yekkala and Dixit [27] presented an experiment over Z-Alizadeh Sani dataset using RF, XGBoost, and NN to tune with GA to optimize accuracy and error rate for HD prediction solution with 93.85% accuracy.

Farman Ali et al. [28] proposed ensemble deep learning and feature fusion for prediction of heart disease. Feature

extraction was done using fusion method, and model was trained using deep network and achieved 98.5% accuracy.

Demidova et al. [29] presented a self-tuning multi-objective system comprising GA having self-tuning capability embedded with SVM which is more flexible in terms of selecting parameter and can be used in multi-object training purpose and claimed that it could outperform over many models.

Hayashi et al. [30] conducted experiment considering the importance of variations in hemoglobin (Hb) levels; the aim of the study was to explore the upper limit of Hb levels during anemia treatment in predialysis chronic kidney disease (CKD) patients utilizing rule extraction.

Santhanam and Ephzibah [31] concluded a study with objective to diagnose heart disease using computing techniques like genetic algorithm and fuzzy logic using hybrid genetic fuzzy heart disease diagnosis system and achieved satisfactory results.

Abushariah et al. [32] presented the prediction of heart disease with MATLAB which demonstrated the comparison study with different available methods like multilayer perceptron (MLP) structure on the artificial neural network (ANN) and adaptive neuro-fuzzy inference systems (ANFIS) approach. Experiment result showed ANN to come as powerful model over ANFIS with score more than 87%.

Kaur and Khehra [33] presented the study that intends to review several studies on FL and hybrid-based methods for assessing patient risk for heart disease. The current study lists articles from 2010 together with the power, operating system, accuracy rate, and other requirements employed in the diagnosis of heart disease using FL and hybrid-based techniques. This survey encourages researchers to come up with new, creative ideas and to carry on with their work in the relevant sector and also presents the potential concept for direct patient transport from nursing homes to intensive care units via ambulance services.

Feng et al. [34] suggest the use of an ANFIS (adaptive neuro-fuzzy inference system) to identify cardiac problems. Utilizing the genetic method, membership function parameters in ANFIS are optimized. The open UCI heart disease datasets were used for the investigation. According to comparison, the experimental result, which showed 91.25 percent accuracy on the testing set, was deemed satisfactory.

3. Proposed Methodology

The neural fuzzy inference system (NFIS) was suggested in this study as a way to describe training data made up of n -dimensional function space. In order to better learning instructions after mistakes have been measured, the NFIS contains an error computing module. Initially, membership functions are constructed, and their parameters are activated and learnt via as needed for an operation.

To approach the research work, help of genetic algorithm and neural fuzzy inference system has been taken. Figure 2 explains the stage-wise execution or methodology to be used in finding the probability of heart attack. Research work has been carried out using Cleveland dataset.

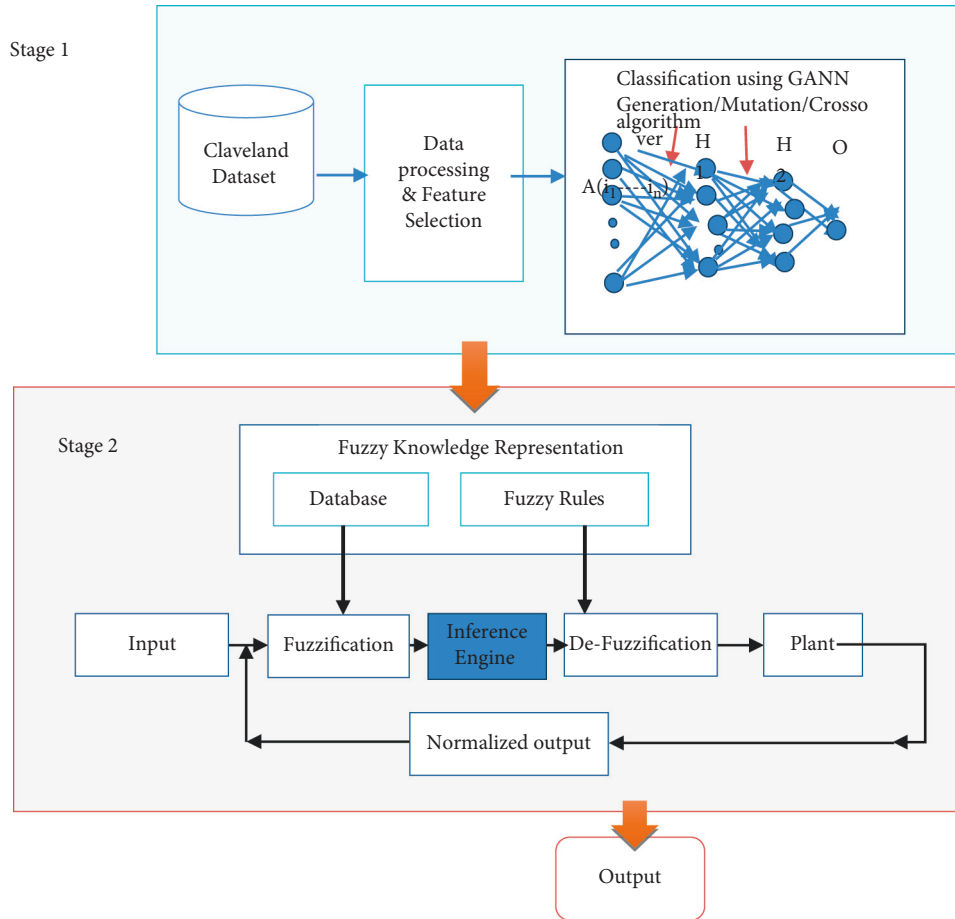


FIGURE 2: Stage-wise execution of genetic algorithm—neural network-based fuzzy inference system.

- (i) total of 17 features from dataset have been used out of which 16 is used as input in different stages and 1 (num) has been used as a target feature.
- (ii) Stage one predicts if patient has heart disease or not.
- (iii) The outcome of stage one has been included as a part of input parameter and has been used as the part of stage two of experiment.
- (iv) Input parameter has been passed to create 126 rule set merged with 13824 predefined rules created with the help of threshold from the medical parameters taken from Cleveland dataset.
- (v) All the rules have been fed in fuzzy system and step in for creating membership function with the help of fuzzy sets. Input underwent through fuzzification process.
- (vi) After training and defuzzification stage, crisp output has been presented as percentage probability of heart attack.

Figure 2 demonstrates the research work in two stages—1 and 2. Stage 1 explains the heart disease prediction, and stage 2 explains the calculation of heart attack probability. Research work demonstrates the hybrid system which comprises neural network powered by genetic algorithm and incorporates the intelligence of fuzzy inference system.

3.1. Stage 1: Predict Heart Disease. In stage one, prediction of heart disease in patient has been achieved. To achieve this goal, input has undergone various stages; first, it passed through data preprocessing stage, and for feature selection mRMR algorithm has been used. In next stage, a neural network has been formed for training purpose, and a weight has been generated using genetic algorithm. Model has been trained, and on every cycle, new offspring has been generated using parent by applying crossover and mutation technique. Best result has been considered and used for prediction of heart disease.

3.2. Stage 2: Heart Attack Probability. In second stage, heart attack probability has been computed and for this intelligence of fuzzy inference system has been added to existing system. Input from previous stage has been passed to fuzzy system where 13000+ fuzzy rules have been fed into the system which would cover both normal and risky cases, and membership functions have been created for each medical parameter passed to the model. In the next step, Mamdani system has been used and input has undergone through fuzzification process where data converted into fuzzy values. In the next step, same converted value has undergone through the defuzzification process to get crisp output (Algorithm 1).

Step 1: Input dataset
 Step 2: Create prerequisite for model, and select number of input, hidden, and output neurons.
 Step 3: Generate feature indices using mRMR feature selection algorithm.
 Step 4: Generate initial population with random uniform values of chromosomes.
 Step 5: Compute weight matrix of population using genetic algorithm.
 Step 6: Train the model
 Step 7: Repeat steps 4 to 6 for different sets of features and record the highest score.
 Step 8: Test data to find if patient has heart disease or not.
 Step 9: Create inference rules using the priority and complexity of medical parameters from the dataset.
 Step 10: Create membership functions for each medical parameter based on the fuzzy set range defined.
 The mapping of a set of real numbers (x_i) onto membership values (u_i), which typically fall between $[0, 1]$, is known as a fuzzy set. A fuzzy set is represented in this fuzzy package by a set of pairs, u_i/x_i , where u_i is the membership value for real number x_i .
 Step 11: Process the input using fuzzification process.
 A point with a low membership value should have less influence on the calculation of the infimum according to the weighting approach (or minimum). As a result, the distance between x and can be expressed as follows: $d(x, \mu) = \inf_{y \in \mathcal{S}} [d_{\mathcal{S}}(x, y) f(\mu(y))]$, where f is a decreasing function of μ (e.g., $f(\mu(y)) = 1/\mu(y)$) such that $f(1) < +\infty$ (in order to guarantee that if x belongs completely to μ , i.e., if $\mu(x) = 1$, the distance is attained for $y = x$), and with the convention $0f(0) = +\infty$. If $\mu(x) = 0$, i.e., if x is completely outside of μ , this definition leads to satisfactory results.
 Step 12: Pass the fuzzy output to defuzzification process to get crisp output.
 Defuzzification is the process of converting a fuzzy set to a precise integer. When you want a precise number as the output from a fuzzy system, defuzzification is necessary. $A = \{(x, \mu_A(x) : x \in X)\}$, where $\mu_A(x)$ is called the membership function of A in (c, d) .
 Step 13: Test with test data to show output.

ALGORITHM 1: Pseudo-code for hybrid neuro-fuzzy inference system.

4. Experimental Setup and Test Case Scenarios

Experiment has been conducted using 16 parameters from Cleveland dataset and divided into two parts: (1) first input was passed to predict the heart disease. (2) Heart attack probability was calculated based on fuzzy inference system.

4.1. System Specification. Experiment has been conducted on machine with configuration: 64-bit Windows 10 operating system, Intel Core i5-5200U CPU@2.20 GHz, and 16 GB RAM. All code has been written using Python libraries using PyCharm IDE and C# using Visual Studio 2019 Community version.

4.2. Dataset. In this research work, Cleveland heart disease dataset has been picked from UCI containing 76 attributes and 303 records for the experiment. This is very popular dataset and has been used by many researchers in various experiments using machine learning [2]. From a total of 76 available attributes, 15 attributes have been selected as an input features and 1 as target feature is shown in Table 1.

Below is the graphical representation of patient statistics shown in Table 2.

Dataset has been categorized into 0–4 numeric status where 0 stands for healthy person and 1–4 indicates person on risk considering 1 is least risky and 4 is the most as shown in Figure 3.

4.2.1. Limitation of Dataset. During data preparation to be used in the research, we found lot of limitation of Cleveland dataset which also exposed the future scope of the study. Few of the limitations are as follows:

TABLE 1: List of attributes from Cleveland dataset to be used in the experiment.

Attributes
Patient age (Age)
Cholesterol (Chol)
Resting ECG result (RestECG)
Blood sugar at fasting (FBS)
ST-depression due to exercise (OldPeak)
Resting heart rate
Patient gender (sex)
Chest pain type (CP)
Heart status (Thal)
Exercise-induced angina (Exang)
Diagnosed heart disease (Num)
Patient family heart disease history
Max heart rate (Thalach)
Resting BP (RBP)
Major vessel count colored by fluoroscopy (CA)
Slope for peak exercise (slope)
Cigarette per day

TABLE 2: Patient health statistics showing sample distribution among healthy and patient at risk.

Patient without heart disease (healthy)	Patient with heart disease
53.87%	46.13%

- (i) Out of 303 records, only 282 records were found correct, and others contain bad data.
- (ii) Data are not divided equally in different age groups, and most of the data where heart disease is positive are from age range 40–75

- (iii) Data are not divided equally between male and female, and the number of male records is much higher than the females.
- (iv) Most of records are for male only, and very few data are related to female patient.
- (v) Out of 282 correct data, only 127 were found risky and other records were of normal patients.

Out of 127, only six records were from age group adult-group1 (18–40) and 1 from old-group2. So, it is clearly noticed that this is not clearly divided correctly into age groups.

4.3. Rule Creation. This study has added more features to a selected processed dataset, such as resting heart rate, cigarette consumption per day, blood sugar, and family history of heart disease. It has also implicated Cleveland dataset dependable, non-dependable attributes, matrices, and supportive computational parameters.

This research had composed with two approaches.

- (i) Out of 303 data, the Cleveland dataset’s data processing revealed 282 excellent data and 21 poor data. Computed 126 rules are based on the dataset record based on num feature which has risk level defined as 1, 2, 3, and 4.
- (ii) A preliminary analysis was conducted to determine the relative contributions of various parameters, such as the maximum heart rate, which was maintained at the normal, medium, high, and critical levels. Similarly, blood sugar can also contribute, but its level cannot affect health in a minute, so we kept its threshold as normal and medium. Same way we picked different features and set its contribution level to determine if a person could face heart attack or not.

A total of 13824 rules have been generated which would cover both normal and risk cases.

4.4. Implications of Membership Functions. A membership function $\mu_A(x)$ enables us to graphically represent a fuzzy set. In this research work, triangular and trapezoidal membership functions have been used. Membership functions have been created for each fuzzy set created for all the medical parameters used in the system. In this research work, triangular and trapezoidal functions have been used to show the fuzzy set graphically.

4.4.1. Triangular Membership Function. The triangular curve is a function of vector x and depends on three scalar parameters a , b , and c ,

$$f(x; a, b, c) = \begin{cases} 0, & x \leq 0, \\ \frac{x-a}{b-a}, & a \leq x \leq b, \\ \frac{c-x}{c-b}, & b \leq x \leq c, \\ 0, & c \leq x. \end{cases} \quad (1)$$

4.4.2. Trapezoidal Membership Function. The trapezoidal curve is a function of a vector, x , and depends on four scalar parameters a , b , c , and d .

$$f(x; a, b, c, d) = \begin{cases} 0, & x \leq 0, \\ \frac{x-a}{b-a}, & a \leq x \leq b, \\ 1, & b \leq x \leq c, \\ \frac{c-x}{c-b}, & c \leq x \leq d, \\ 0, & d \leq x. \end{cases} \quad (2)$$

There are two special cases of a trapezoidal function, which are called R -functions and L -functions having parameters $a = b = -\infty$ and $c = d = +\infty$.

R -functions ($a = b = -\infty$)

$$\mu_A(x) = \begin{cases} 0, & x > d, \\ \frac{d-x}{d-c}, & c \leq x \leq d, \\ 1, & x < c. \end{cases} \quad (3)$$

L -functions ($c = d = +\infty$)

$$\mu_A(x) = \begin{cases} 0, & x < a, \\ \frac{x-a}{b-a}, & a \leq x \leq b, \\ 1, & x > b. \end{cases} \quad (4)$$

4.4.3. Gaussian Function. This is a function of vector x and depends on two parameter a and b represented by

$$f(x; a, b) = e^{-(x-b)^2/2a^2}. \quad (5)$$

4.4.4. Sigmoid Function. This is a function of vector x and depends on two parameters a and b represented by

$$f(x; a, b) = \frac{1}{1 + e^{-a(x-b)}}. \quad (6)$$

4.4.5. Resting Blood Pressure. Input field is divided into 4 fuzzy sets normal, moderate, high, and critical, and their range is shown in Figure 4. Membership functions of “medium” and “high” are trapezoidal, and membership functions of “normal” and “critical” are triangular.

4.4.6. Cholesterol. Input field is divided into 4 fuzzy sets normal, moderate, high, and critical, and their range is

TABLE 3: Continued.

Condition	Attack probability
RestingBP_RBP IS High AND SerumCholesterol_SCH IS Moderate AND FastingBloodSugar_FBS IS Normal AND RestingECGResult_RES IS Normal AND RestingHeartRate_RHR IS Good AND MaxHeartRate_MHR IS Moderate AND CigratePerDay_CPD IS Normal AND HeartDiseaseFamilyHistory IS No AND IsHeartPatient IS No	Normal
RestingBP_RBP IS High AND SerumCholesterol_SCH IS Moderate AND FastingBloodSugar_FBS IS Normal AND RestingECGResult_RES IS Normal AND RestingHeartRate_RHR IS Good AND MaxHeartRate_MHR IS Critical AND CigratePerDay_CPD IS High AND HeartDiseaseFamilyHistory IS Yes AND IsHeartPatient IS Yes	High risk
RestingBP_RBP IS High AND SerumCholesterol_SCH IS Moderate AND FastingBloodSugar_FBS IS Normal AND RestingECGResult_RES IS Normal AND RestingHeartRate_RHR IS High AND MaxHeartRate_MHR IS Normal AND CigratePerDay_CPD IS Normal AND HeartDiseaseFamilyHistory IS No AND IsHeartPatient IS No	Normal
RestingBP_RBP IS Critical AND SerumCholesterol_SCH IS Critical AND FastingBloodSugar_FBS IS High AND RestingECGResult_RES IS Normal AND RestingHeartRate_RHR IS High AND MaxHeartRate_MHR IS High AND CigratePerDay_CPD IS High AND HeartDiseaseFamilyHistory IS Yes AND IsHeartPatient IS No	High risk
RestingBP_RBP IS Critical AND SerumCholesterol_SCH IS Critical AND FastingBloodSugar_FBS IS High AND RestingECGResult_RES IS Normal AND RestingHeartRate_RHR IS High AND MaxHeartRate_MHR IS High AND CigratePerDay_CPD IS High AND HeartDiseaseFamilyHistory IS Yes AND IsHeartPatient IS Yes	Critical risk
RestingBP_RBP IS Critical AND SerumCholesterol_SCH IS Critical AND FastingBloodSugar_FBS IS High AND RestingECGResult_RES IS Risk1 AND RestingHeartRate_RHR IS Good AND MaxHeartRate_MHR IS Normal AND CigratePerDay_CPD IS Moderate AND HeartDiseaseFamilyHistory IS No AND IsHeartPatient IS No	Moderate risk

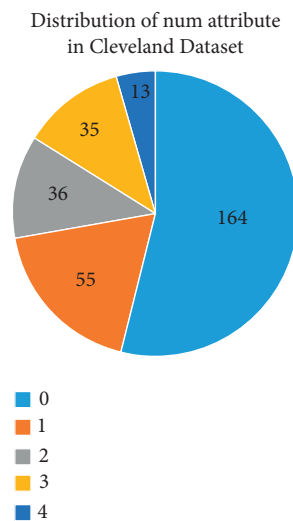


FIGURE 3: Graph showing data distribution among healthy and risky patient.

shown in Figure 5. Membership functions are defined as trapezoidal and triangular.

4.4.7. *Fasting Blood Sugar.* Input field is divided into 2 fuzzy sets normal and high with triangular membership function shown in Figure 6.

4.4.8. *Resting Electrocardiogram (ECG) Result.* Input is divided into 3 fuzzy sets normal, risk1, and risk2 having triangular membership function shown in Figure 7.

4.4.9. *Resting Heart Rate.* Input is divided into 3 fuzzy sets good, average, and high having triangular membership function shown in Figure 8.

4.4.10. *Max Heart Rate.* Input is divided into 4 fuzzy sets normal, moderate, high, and critical, and their range is shown in Figure 9. Membership function is defined as triangular.

4.4.11. *Cigarette per Day.* Input field is divided into 2 fuzzy sets normal and high with triangular membership function. Ranges have been shown in Figure 10.

4.4.12. *Family History for Heart Disease.* Input field is divided into 2 fuzzy sets yes and no with triangular membership function. Ranges have been shown in Figure 11.

4.4.13. *Heart Patient.* Input field is divided into 2 fuzzy sets yes and no with triangular membership function. Ranges have been shown in Figure 12.

4.4.14. *Heart Attack Probability.* This is output parameter and is divided into 4 fuzzy sets normal, moderate risk, high risk, and critical risk, and their ranges have been shown in Figure 13. Membership function is defined as triangular.

4.5. *Inference Rules.* Table 3 depicts few inference rules demonstrating experiment scenarios. It demonstrates the rule set with probability for heart attack.

5. Results and Discussion

Experiment has been performed on Cleveland dataset with many test cases in hand, and few of them are added below for reference. Here, medical parameters have been taken into consideration, fuzzy sets and membership functions have been defined, input underwent fuzzification process applying fuzzy rule sets to get fuzzy values, and later

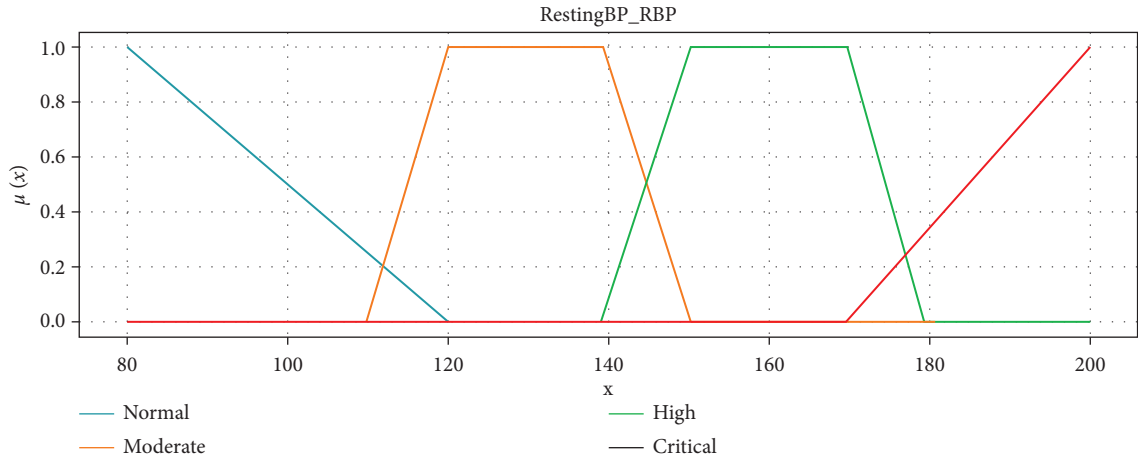


FIGURE 4: Graph showing membership functions for input feature RBP.

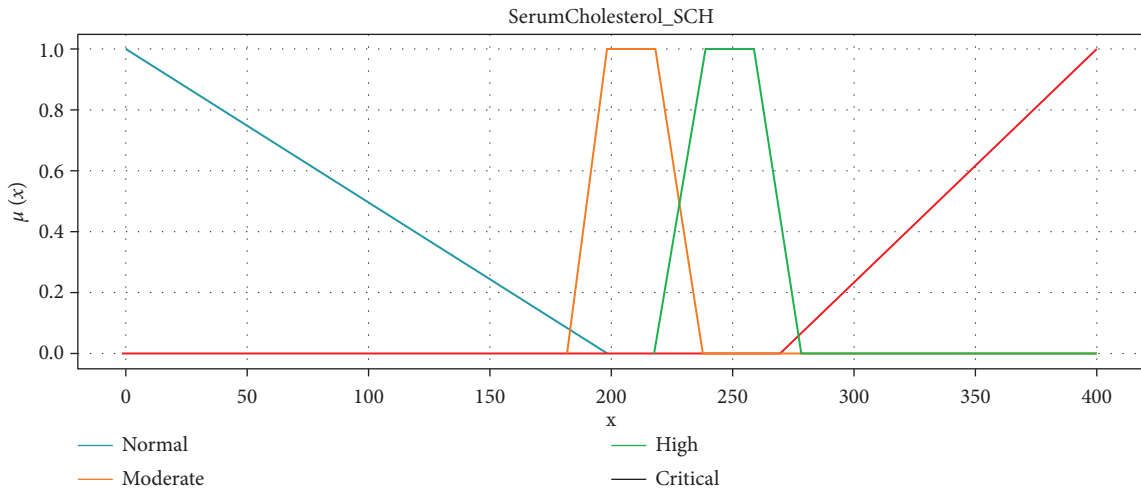


FIGURE 5: Graph showing membership functions for input feature SCH.

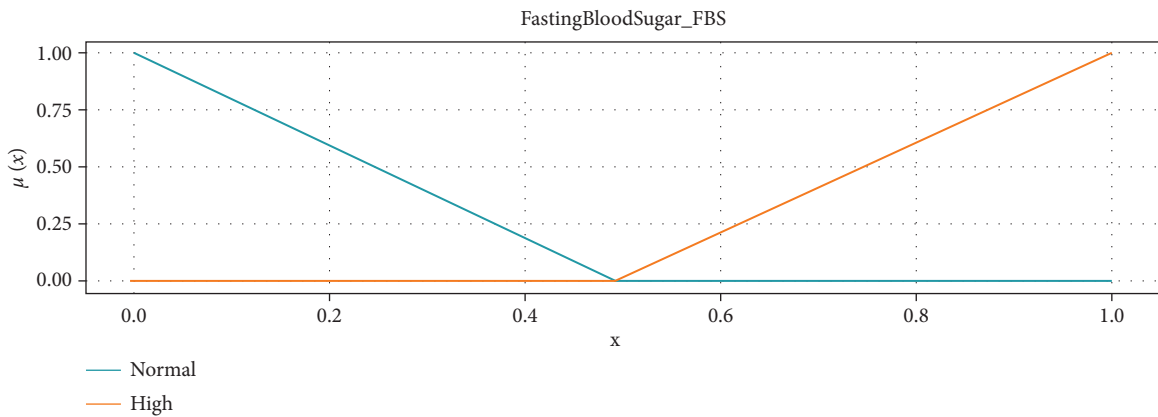


FIGURE 6: Graph showing membership functions for input feature FBS.

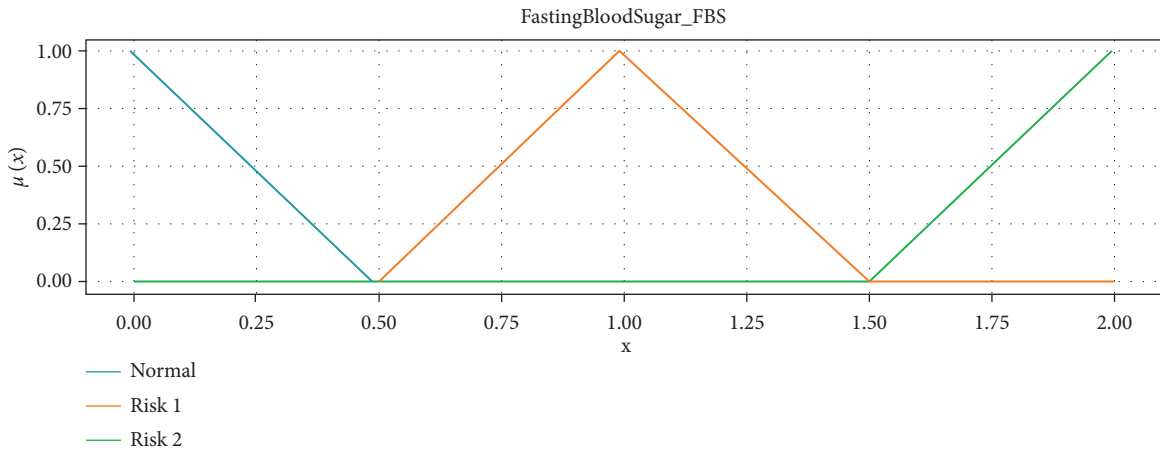


FIGURE 7: Graph showing membership functions for input feature RES.

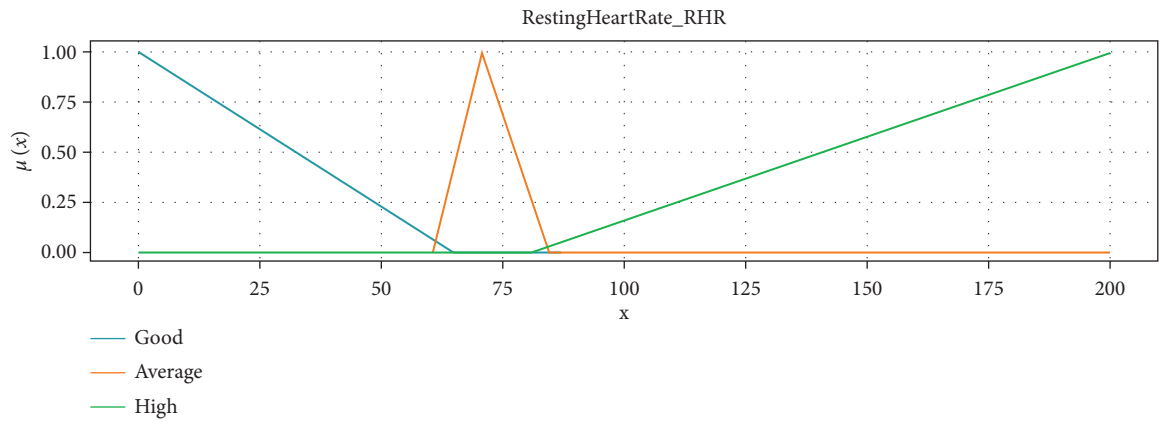


FIGURE 8: Graph showing membership functions for input feature RHR.

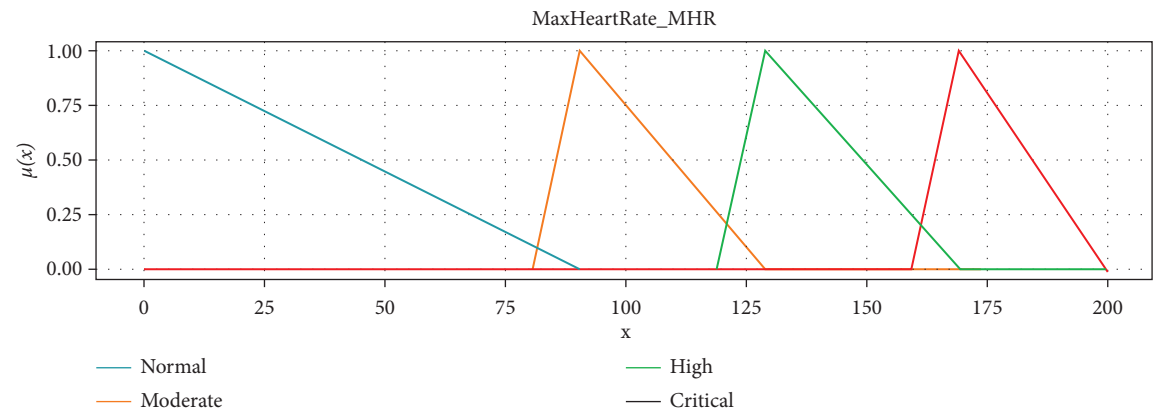


FIGURE 9: Graph showing membership functions for input feature MHR.

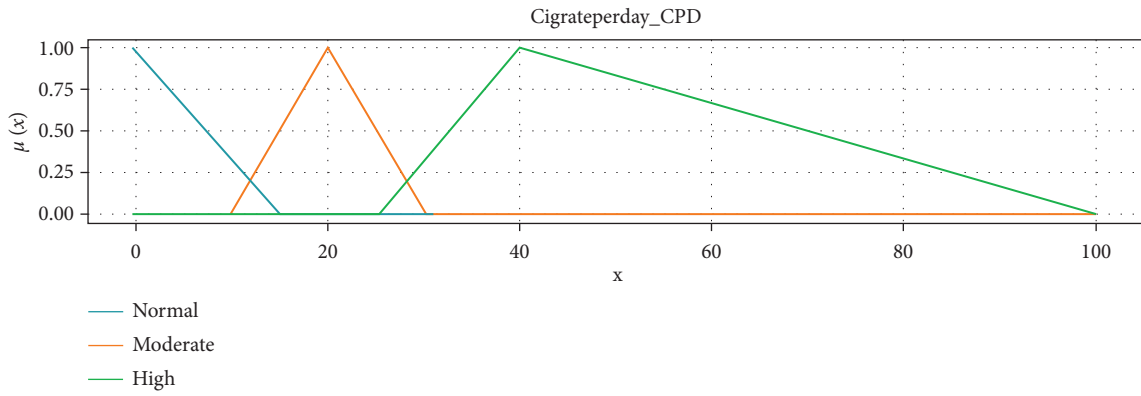


FIGURE 10: Graph showing membership functions for input feature CPD.

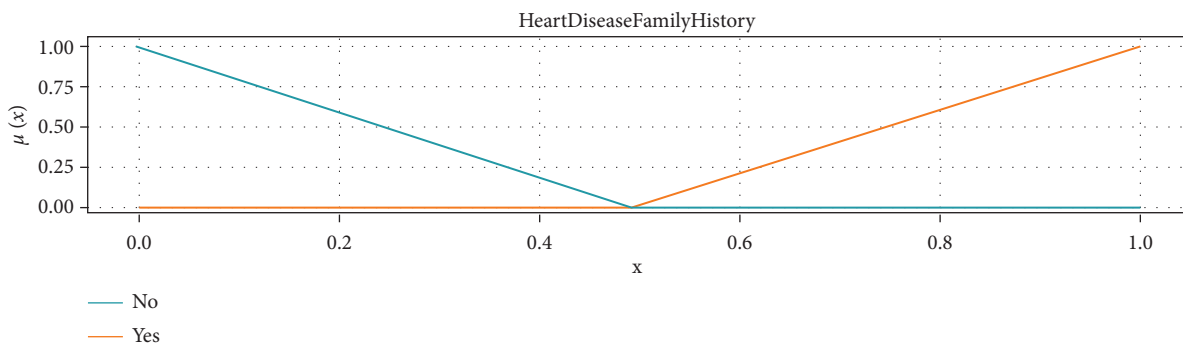


FIGURE 11: Graph showing membership functions for input feature heart disease family history.

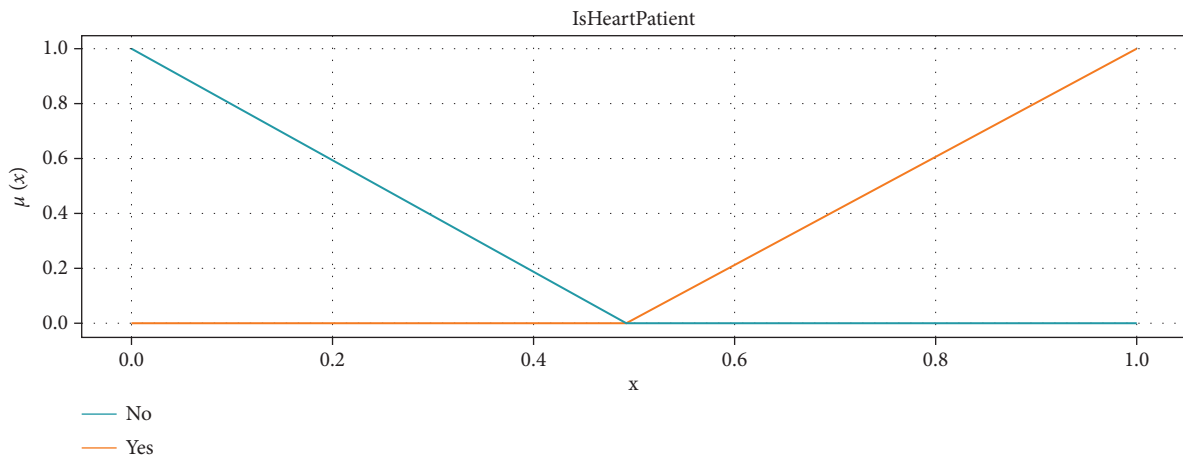


FIGURE 12: Graph showing membership functions for input feature heart patient.

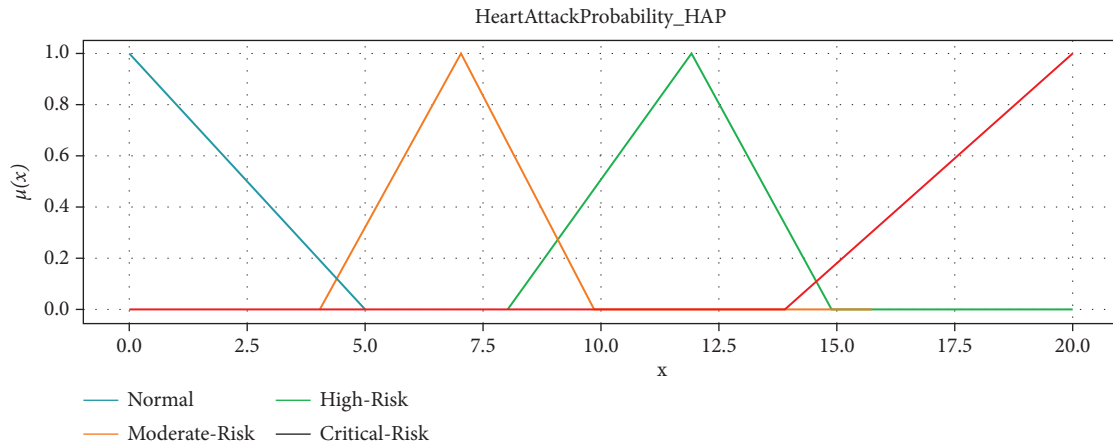


FIGURE 13: Graph showing membership functions for output feature HAP.

TABLE 4: Test cases showing result with inference system.

Test case	Result (shows the probability of heart attack based on the rule mapped)
"RestingBP_RBP:" 200 "SerumCholesterol_SCH:" 280 "FastingBloodSugar_FBS:" 0 "RestingECGResult_RES:" 0 "RestingHeartRate_RHR:" 100 "MaxHeartRate_MHR:" 180 "CigratePerDay_CPD:" 5 "HeartDiseaseFamilyHistory:" 0 "IsHeartPatient:" 0	Outcome (%) = 57.57 Here, it is clear that RBP, SCH, RHR, and MHR and higher than usual but rest features have controlled data so here probability of heart attack is 50+% means person is at risk but not much. This case be measured as moderate-high risk.
"RestingBP_RBP:" 130 "SerumCholesterol_SCH:" 204 "FastingBloodSugar_FBS:" 0 "RestingECGResult_RES:" 2 "RestingHeartRate_RHR:" 71 "MaxHeartRate_MHR:" 172 "CigratePerDay_CPD:" 0 "HeartDiseaseFamilyHistory:" 0 "IsHeartPatient:" 0	Outcome (%) = 34.96 Here, most of the parameters are in good shape but still features like RES, RBP, and MHR are more than usual but in control. So, this case could be considered as low-moderate risk.
"RestingBP_RBP:" 105 "SerumCholesterol_SCH:" 198 "FastingBloodSugar_FBS:" 0 "RestingECGResult_RES:" 0 "RestingHeartRate_RHR:" 92 "MaxHeartRate_MHR:" 168 "CigratePerDay_CPD:" 20 "HeartDiseaseFamilyHistory:" 0 "IsHeartPatient:" 0	Outcome (%) = 34.84 Here, most of the parameters are in good shape but still features like CPD, RHR, and MHR are more than usual but in control. So, this case could be considered as low-moderate risk.
"RestingBP_RBP:" 110 "SerumCholesterol_SCH:" 172 "FastingBloodSugar_FBS:" 0 "RestingECGResult_RES:" 2 "RestingHeartRate_RHR:" 68 "MaxHeartRate_MHR:" 158 "CigratePerDay_CPD:" 50 "HeartDiseaseFamilyHistory:" 1 "IsHeartPatient:" 1	Outcome (%) = 33.35 This is a heart patient but his parameters other than CPD and RES are normal so is at low-moderate risk.

TABLE 4: Continued.

Test case	Result (shows the probability of heart attack based on the rule mapped)
"RestingBP_RBP:" 139 "SerumCholesterol_SCH:" 240 "FastingBloodSugar_FBS:" 0 "RestingECGResult_RES:" 0 "RestingHeartRate_RHR:" 87 "MaxHeartRate_MHR:" 200 "CigratePerDay_CPD:" 40 "HeartDiseaseFamilyHistory:" 1 "IsHeartPatient:" 1	Outcome (%) = 57.57 This is a heart patient, and his few parameters are beyond normal so is at moderate-high risk.
"RestingBP_RBP:" 140 "SerumCholesterol_SCH:" 290 "FastingBloodSugar_FBS:" 0 "RestingECGResult_RES:" 2 "RestingHeartRate_RHR:" 68 "MaxHeartRate_MHR:" 158 "CigratePerDay_CPD:" 50 "HeartDiseaseFamilyHistory:" 1 "IsHeartPatient:" 1	Outcome (%) = 57.57 This is a heart patient, and his few parameters are beyond normal so is at moderate-high risk.
"RestingBP_RBP:" 140 "SerumCholesterol_SCH:" 290 "FastingBloodSugar_FBS:" 1 "RestingECGResult_RES:" 0 "RestingHeartRate_RHR:" 103 "MaxHeartRate_MHR:" 158 "CigratePerDay_CPD:" 50 "HeartDiseaseFamilyHistory:" 1 "IsHeartPatient:" 1	Outcome (%) = 85.35 This is a heart patient, and his most of the parameters are beyond normal so is at critical risk.
"RestingBP_RBP:" 140 "SerumCholesterol_SCH:" 290 "FastingBloodSugar_FBS:" 0 "RestingECGResult_RES:" 0 "RestingHeartRate_RHR:" 103 "MaxHeartRate_MHR:" 158 "CigratePerDay_CPD:" 50 "HeartDiseaseFamilyHistory:" 1 "IsHeartPatient:" 0	Outcome (%) = 57.57 Here, parameters are more or less same as above case but this patient is not a heart patient so will be at moderate-high risk.
"RestingBP_RBP:" 83 "SerumCholesterol_SCH:" 188 "FastingBloodSugar_FBS:" 0 "RestingECGResult_RES:" 0 "RestingHeartRate_RHR:" 56 "MaxHeartRate_MHR:" 110 "CigratePerDay_CPD:" 0 "HeartDiseaseFamilyHistory:" 0 "IsHeartPatient:" 0	Outcome (%) = 0 Here, parameters are normal and also this patient is not a heart patient so will be at normal or low risk.

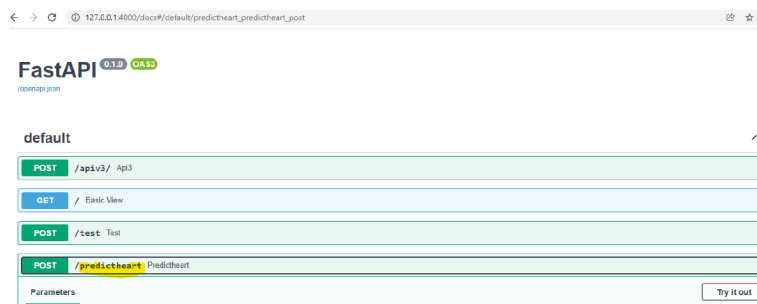


FIGURE 14: The API interface for heart attack probability.

TABLE 5: Model comparison with existing similar research work.

Author name	Proposed approach	Classification technique	Prediction rate in percentage
Samuel et al. [15]	Hybrid system which combines fuzzy analytic hierarchy process (fuzzy AHP) technique and ANN	ANN and fuzzy AHP	Accuracy: 91.1%
Aghamohammadi et al. [13]	Genetic algorithm (GA) and adaptive neural fuzzy inference system (ANFIS)	ANFIS	Accuracy: 84.43%, sensitivity: 91.1504%, specificity: 79.1667%
Li et al. [35]	Combination of FuzzyGMEn-generated DDM and CNN (Inception_v4 model)	CNN	Accuracy: 81.85%
Ziasabounchi and Askerzade [14]	ANFIS-based classification model	ANFIS	Accuracy: 92.30%
Abushariah et al. [32]	Artificial neural network (ANN) and adaptive neuro-fuzzy inference systems (ANFIS)	ANN, ANFIS	Accuracy: 87.04%
Rahul, Henge and Sharma	This study	Hybrid GANN + fuzzy inference system	Accuracy: 94%

defuzzification process is applied to convert fuzzy value into crisp output. Table 4 shows the test cases with result obtained from the experiment.

5.1. Expose Model via Application Programming Interface (API). In the next part of the experiment, fast API form Joblib Python Library has been used to integrate model and exposing to end user for testing and further use. Figures 14–16 demonstrate the API integration of trained model.

The system comprises hybrid neural network which is powered by genetic algorithm and has an intelligence of fuzzy inference system, and this combination boosted the research work in achieving the desired result which as compared to other proposed solutions seems more powerful in terms of prediction and calculating the probability. Result is given in Section 4; Table 4 demonstrates the test scenarios covered during the experiment. The comparative analysis has been done based on the proposed approach, implicated practices, classification technique, and achieved prediction accuracy in the form of percentage as shown in Table 5 and Figure 17.

Result shows that system was effectively calculating the probability of heart attack using which relevant preventive measures could be taken to improvise health parameters.

6. Conclusion

Cardiac arrest is an incurable incongruity that requires special treatment and cure. It has been a key research area for many years, and the number of researchers across the globe is devoted toward finding the optimal solution to avoid the ill-effect of this disease. Along with predicting heart disease, if focus moves towards prevention of heart attack as well, then this could result in major life saver area for masses. This research work is fully devoted toward finding out the probability of heart attack so that people can take preventive measure before it hits the wall. This research proposed the neural fuzzy inference system (NFIS) to represent the training data formed from the n -dimensions of functions. The NFIS consists of error computing module to improve the learning instructions when the errors have been measured, initially the membership functions are defined, and

the parameters of membership functions are activated and learnt through when needed for an operation. The proposed methodology has experimented with sample test cases on Cleveland heart disease dataset from UCI repository with the integration of supporting dependable and nondependable parameters, causing-factors, and data-matrices. This research has integration more than 13000 fuzzification rules to generate best decision-making, normalization process, and planting techniques to create the feasibility to compute the heart attack probability, and achieved 94 percentage of accuracy. Result shows that system was effectively calculating the probability of heart attack using which relevant preventive measures could be taken to improvise health parameters. Further work can be achieved to make this system more robust so that it could help in creating more accurate system. Experiment was carried out on limited age group due to limitation of dataset but that can be improved and more age group might be included in future experiments. The future work can be extendable to build auto-alert and advise system with integration hardware peripheral circuit devices.

Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

The authors declare no conflicts of interest.

Authors' Contributions

R. K. Jha and S. K Henge conceptualized the study; R. K. Jha and S. K Henge developed the methodology; R. K. Jha and S. K Henge. helped with software; S. K Henge validated the study; S. K. Henge carried out formal analysis; R. K. Jha and S. K. Henge investigated the study; R. K. Jha and S. K. Henge collected resources; S. K. Henge curated the data; R. K. Jha and S. K. Henge wrote and prepared the original draft; S. K. Henge wrote, reviewed, and edited the manuscript; S. K. Henge visualized and supervised the study.

References

- [1] Who, "WHO cardiovascular disease," 2022, <https://www.who.int/health-topics/cardiovascular-diseases/>.
- [2] V. A. Medical Center, "Long beach and Cleveland clinic foundation:robert detrano, M.D., ph.D. Hungarian institute of cardiology," 2022, <https://archive.ics.uci.edu/ml/datasets/Heart+Disease>.
- [3] S. K. Henge and B. Rama, "Five layered-neural fuzzy closed loop hybrid control system with compound bayesian decision making process for classification cum identification of mixed connective conjunct consonants and numerals," *Advances in Intelligent Systems and Computing*, vol. 553, pp. 619–629, 2017.
- [4] B. R. Santosh Kumar Henge, "Comprative study with analysis of OCR algorithms and invention analysis of character recognition approched methodologies," in *Proceedings of the IEEE 1st International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES)*, Delhi, India, 2016.
- [5] S. K. Bhatia, K. K. Mishra, S. Tiwari, and V. K. Singh, *Advances in Computer and Computational Sciences* Springer, Berlin, Germany, 2016.
- [6] S. K. Henge and B. Rama, "Neural fuzzy closed loop hybrid system for classification, identification of mixed connective consonants and symbols with layered methodology," *Intelligent Control and Energy Systems (ICPEICES)*, pp. 1–6, 2016.
- [7] B. Singh and S. K. Henge, "Neural fuzzy inference hybrid system with SVM for identification of false singling in stock market prediction for profit estimation," *Advances in Intelligent Systems and Computing*, vol. 1197, pp. 221–227, 2020.
- [8] S. Nashif, R. Rakib Raihan, R. Islam, and H. Imam, "Heart disease detection by using machine learning algorithms and a real-time cardiovascular health monitoring system," *World Journal of Engineering and Technology*, vol. 6, pp. 854–876, 2018.
- [9] A. U. Haq, J. P. Li, M. H. Memon, S. Nazir, and R. Sun, "A hybrid intelligent system framework for the prediction of heart disease using machine learning algorithms," *Mobile Information Systems*, Special Issue, vol. 2018, Article ID 3860146, 21 pages, 2018.
- [10] M. Abdar, W. Książek, U. R. Acharya, R. S. Tan, V. Makarenkov, and P. Pławiak, "A new machine learning technique for an accurate diagnosis of coronary artery disease," *Computer Methods and Programs in Biomedicine*, vol. 179, Article ID 104992, 2019.
- [11] R. K. Jha, S. Henge, and A. Sharma, *Optimal Machine Learning Classifiers for Prediction of Heart Disease Sersc*, 2020.
- [12] A. Davari Dolatabadi, S. E. Z. Khadem, and B. M. Asl, "Automated diagnosis of coronary artery disease (CAD) patients using optimized SVM," *Computer Methods and Programs in Biomedicine*, vol. 138, pp. 117–126, 2017.
- [13] M. Aghamohammadi, M. Madan, J. K. Hong, and I. Watson, "Predicting heart attack through explainable artificial intelligence," *Computational Science – ICCS*, vol. 11537, pp. 633–645, 2019.
- [14] N. Ziasabounchi and I. Askerzade, "ANFIS based classification model for heart disease prediction," *International Journal of Electrical & Computer Sciences*, vol. 14, p. 2, 2014.
- [15] O. W. Samuel, G. M. Asogbon, A. K. Sangaiah, P. Fang, and G. Li, "An integrated decision support system based on ANN and Fuzzy_AHP for heart failure risk prediction," *Expert Systems with Applications*, vol. 68, pp. 163–172, 2017.
- [16] A. S. Kumar, "Diagnosis of heart disease using fuzzy resolution mechanism," *Journal of Artificial Intelligence*, vol. 5, no. 1, pp. 47–55, 2011.
- [17] A. Yazdani and K. Ramakrishnan, "Performance evaluation of artificial neural network models for the prediction of the risk of heart disease," in *Proceedings of the International Conference for Innovation in Biomedical Engineering and Life Sciences*, pp. 179–182, Putrajaya, Malaysia, December 2015.
- [18] Z. Arabasadi, R. Alizadehsani, M. Roshanzamir, H. Moosaei, and A. A. Yarifard, "Computer aided decision making for heart disease detection using hybrid neural network-Genetic algorithm," *Computer Methods and Programs in Biomedicine*, vol. 141, pp. 19–26, 2017.
- [19] K. Uyar and A. İlhan, "Diagnosis of heart disease using genetic algorithm based trained recurrent fuzzy neural networks," *Procedia Computer Science*, vol. 120, pp. 588–593, 2017.
- [20] G. T. Reddy, M. P. K. Reddy, K. Lakshmana, D. S. Rajput, R. Kaluri, and G. Srivastava, "Hybrid genetic algorithm and a fuzzy logic classifier for heart disease diagnosis," *Evol. Intel*, vol. 13, no. 2, pp. 185–196, 2020.
- [21] M. Akgul, O. E. Sönmez, and T. Ozcan, "Diagnosis of heart disease using an intelligent method: a hybrid ANN – GA approach," in *Proceedings of the Intelligent and Fuzzy Techniques in Big Data Analytics and Decision Making*, pp. 1250–1257, Stanbul, Turkey, July 2019.
- [22] N. G. B. Amma, "Cardiovascular disease prediction system using genetic algorithm and neural network," in *Proceedings of the International Conference on Computing, Communication and Applications*, pp. 1–5, Dindigul, India, February 2012.
- [23] S. Nikam, P. Shukla, and M. Shah, "Cardiovascular disease prediction using genetic algorithm and neuro-fuzzy system," *International Journal of Latest Trends in Engineering and Technology*, vol. 8, no. 2, pp. 104–110, 2017.
- [24] M. A. Jabbar, B. L. Deekshatulu, and P. Chandra, "Classification of heart disease using K- nearest neighbor and genetic algorithm," *Procedia Technology*, vol. 10, pp. 85–94, 2013.
- [25] M. A. Jabbar, B. L. Deekshatulu, and P. Chandra, "An evolutionary algorithm for heart disease prediction. Wireless networks and computational intelligence. ICIP 2012," *Communications in Computer and Information Science*, vol. 292, pp. 378–389, 2012.
- [26] F. Z. Abdeldjouad, M. Brahami, and N. Matta, "A hybrid approach for heart disease diagnosis and prediction using machine learning techniques. The impact of digital technologies on public health in developed and developing countries," in *Proceedings of the 18th International Conference*, vol. 12157, pp. 299–306, Hammamet, Tunisia, June, 2020.
- [27] I. Yekkala and S. Dixit, "A novel approach for heart disease prediction using genetic algorithm and ensemble classification," *Advances in Intelligent Systems and Computing*, vol. 1251, pp. 468–489, 2020.
- [28] F. Ali, S. El-Sappagh, S. R. Islam et al., "A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion," *Information Fusion*, vol. 63, pp. 208–222, 2020.
- [29] L. A. Demidova, M. M. Egin, and R. V. Tishkin, "A self-tuning multiobjective genetic algorithm with application in the SVM classification," *Procedia Computer Science*, vol. 150, pp. 503–510, 2019.
- [30] Y. Hayashi, K. Nakajima, and K. Nakajima, "A rule extraction approach to explore the upper limit of hemoglobin during anemia treatment in patients with predialysis chronic kidney disease," *Informatics in Medicine Unlocked*, vol. 17, Article ID 100262, 2019.
- [31] T. Santhanam and E. P. Ephzibah, "Heart disease prediction using hybrid genetic fuzzy model," *Indian Journal of Science and Technology*, vol. 8, no. 9, pp. 797–803, 2015.
- [32] M. A. M. Abushariah, A. A. M. Alqudah, O. Y. Adwan, and R. M. M. Yousef, "Automatic heart disease diagnosis system

- based on artificial neural network (ANN) and adaptive neuro-fuzzy inference systems (ANFIS) approaches,” *Journal of Software Engineering and Applications*, vol. 7, pp. 1055–1064, 2014.
- [33] J. Kaur and B. S. Khehra, “Fuzzy logic and hybrid based approaches for the risk of heart disease detection: state-of-the-art review,” *J. Inst. Eng. India Ser. B*, vol. 103, no. 2, pp. 681–697, 2022.
- [34] J. Feng, Q. Wang, and Na Li, “An intelligent system for heart disease prediction using adaptive neuro-fuzzy inference systems and genetic algorithm,” *Journal of Physics: Conference Series*, no. 1, Article ID 012172, 2021.
- [35] Y. Li, Y. Zhang, L. Zhao et al., “Combining convolutional neural network and distance distribution matrix for identification of congestive heart failure,” *IEEE Access*, vol. 6, Article ID 39734, 2018.