International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Enhanced Knowledge Based System for Cardiovascular Disease Prediction using Advanced Fuzzy TOPSIS

Huma Parveen[1]*, Syed Wajahat Abbas Rizvi[2], Raja Sarath Kumar Boddu[3]

Submitted: 02/11/2023 **Revised:** 23/12/2023 **Accepted:** 04/01/2024

Abstract*:* Cases of cardiovascular diseases have risen over the last decade, making them the leading cause of mortality. Early detection of heart diseases helps doctors provide better treatment to patients. In this research, a knowledge-based hybrid heart disease prediction model using advanced fuzzy techniques and artificial neural networks (ANN) is proposed. The ANN and advanced fuzzy Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) techniques are implemented in the proposed methodology for risk prediction of disease and disease classification, respectively. The Analytic Hierarchy Process (AHP) method's attribute weights help make effective prediction of diseases. The proposed model (ANN+ fuzzy TOPSIS) has been measured over various performance-measuring metrics and then compared to other traditional techniques to determine its efficiency. Numerical analysis of the proposed model shows that it performed better than other conventional methods in terms of accuracy (0.99), precision (0.98), specificity (0.978), F-measure (0.981), sensitivity (0.996), and many more. The goal of this research is to improve knowledge-based systems' efficacy through the application of fuzzy logic and ANN for cardiovascular disease prediction and classification.

Keywords: *Cardiovascular Disease, advanced fuzzy TOPSIS, Knowledge based system, Fuzzy Logic, Disease Risk Prediction system*

1. Introduction

There has been a noticeable rise in the prevalence of heart disease, and it has now overtaken all other forms of mortality as the top reason for death for individuals in the majority of nations all over the globe [1]. Many different aspects of cardiovascular disease (CVD) might harm either the anatomy or functionality of the heart [2]. It might be challenging for medical professionals to make a rapid and accurate diagnosis of certain conditions [3, 4].

The treatment of cardiac disease is being performed by several systems, many of which depend on methods of soft computing that have been developed [5]. In particular, the combination of multiple different forms of soft computing to construct hybrid models has been examined as a means of producing results that are superior to those produced by a single kind of computational model [6]. In most cases, these models included two distinct states. In the first stage, approaches for selecting features are employed to pick a subset of those characteristics [7]. After that, the produced subset of characteristics is utilized as data for the categorization procedures that are employed in the second state [8]. Figure 1 illustrates a chart showing various types of CVD. It shows about the coronary heart disease, Myocarditis, congenital, cardiomyopathy etc.

1 Computer science and Engineering, Amity University, Uttar Pradesh, India. huma.parveen@student.amity.edu
2 Computer science and Engineering, Amity University, Uttar Pradesh , India. swarizvi@lko.amity.edu
3 Computer science and Engineering, Lenora College of engineering, Rampachodavaram, Andhra Pradesh, India. iamsarath@gmail.com

Coronary heart disease (CHD), alternatively referred to as CAD (coronary artery disease) or ischemic heart disease, is a distinct form of cardiovascular ailment that predominantly impacts the coronary arteries responsible for delivering oxygenated blood to the cardiac muscle. RHD (rheumatic heart disease) is a pathological condition that may arise because of untreated, or sub optimally treated streptococcal pharyngitis, with a particular predilection for children and young adults. Ischemic heart disease, alternatively referred to as CHD or CAD, is a cardiovascular ailment characterized by diminished or obstructed blood flow to the myocardium owing to atherosclerosis, a condition marked by the constriction of the coronary arteries. Ischemia is characterized by inadequate blood supply and oxygenation to the myocardium. Myocarditis is a cardiac pathology distinguished by myocardial inflammation, which refers to the irritation of the heart muscle. In most cases, a bacterial or viral illness is to blame for this condition, but other causes are possible as well. Congenital heart disease is an umbrella term for a variety of congenital heart defects. Heart and vascular anomalies that are close to the heart are included in this category. Heart disease can range from being completely asymptomatic to being so severe that it could kill the patient. This specific birth defect is extremely common and can have serious consequences for people of all ages. Cardiomyopathy encompasses a range of cardiac disorders characterized by fragility or malfunction of the myocardium, leading to reduced cardiac contractility. The compromised contractile ability of the heart subsequently results in a decrease in its pumping capacity. This particular medical condition possesses the capacity to hinder the ability of the heart to effectively distribute blood to the circulatory system, thereby giving rise to a variety of signs and complications. Angina pectoris, widely known as angina, manifests as a form of thoracic pain or discomfort that arises from an inadequate supply of blood to the myocardium. Angina frequently manifests as a clinical manifestation of an underlying cardiac ailment, commonly known as coronary artery disease (CAD), which arises due to the constriction or obstruction of the coronary arteries. The manifestation of insufficient oxygen supply to the heart, commonly observed during physical exertion or episodes of emotional strain, serves as an indicator of concern.

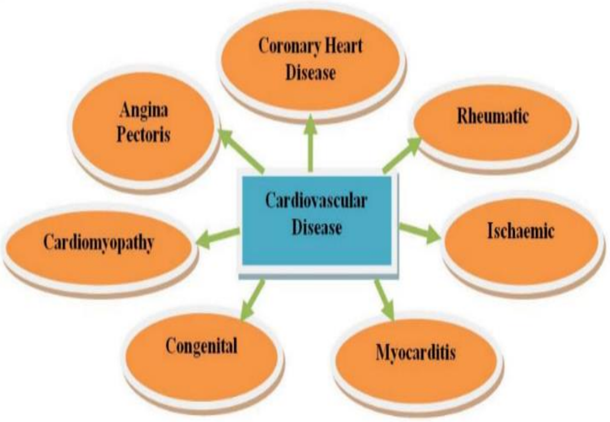


Fig 1 Types of cardiovascular disease (CVD) [9]

The traditional techniques for diagnosing cardiovascular disease (CVD) essentially make use of the patient's healthcare history, diagnostic test reports, and clinical specialist analysis of the pertinent symptoms. All medical professionals foresee the existence of illness based on information they have gained via their training and experience. Heart disease risk is influenced by a variety of factors, including excessive cholesterol, irregular blood pressure, and inactivity. Because of human error, this might result in a wrong diagnosis and possibly a delay in the prediction of the course and severity of the illness. Such erroneous judgments can prevent timely medical care from being provided or perhaps result in fatalities. As a result, a smart medical decision support system (MDSS) enters the scene and plays a crucial role in healthcare by using patient medical data to forecast illness. In this study, an MDSS is offered as a means of accurately predicting and diagnosing illnesses using a patient's computerized health information. This research includes about the various fuzzy based techniques used in heart disease identification.

1.1 Fuzzy technique in the healthcare sector

A set is said to be fuzzy if it permits its components to have varying levels of inclusion in the range [0, 1] [10]. The fuzzy categorization system provides an alternative crisp logic by analyzing data sets according to the individuals' membership in each category [11]. The concept of fuzzy membership is predicated on the idea that a person's participation in a particular group can vary from full membership (100%) to non-participation (0%), and it recognizes the possibility that a dataset might be divided into partial participation in two or more groups [12]. Figure 2 shows an illustration of fuzzy logic. It represents a block diagram of a fuzzy logic system that involves an input interface (also called fuzzification), fuzzy inference system, output variable (also called defuzzication), and membership function. The fuzzy system receives an input in the form of a crisp or a numerical value and converts this into a fuzzy value with the help of the fuzzification process. This fuzzy input is turned into a linguistic input with the use of the membership function. Then the interface engine uses a knowledge base database to generate linguistic output. This fuzzy output is turned back into a crisp value with the help of the defuzzification interface.

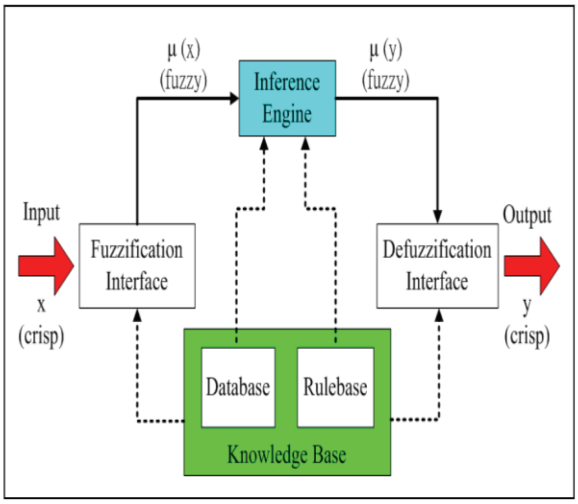


Fig 2 Block diagram of fuzzy logic [13]

To represent the level of membership, fuzzy logic employs truth levels that range from 0.0 to 1.0 [14]. The values of the attributes are changed to fuzzy values. As an example, revenue is projected into the discrete classifications "low, medium, and high," and then fuzzy values are determined for each category. It is possible that more than one fuzzy number would be relevant to a certain fresh sample. Each criterion that is relevant casts a vote about membership in the respective groups. Typically, one would begin by adding up the truth values of each anticipated category [15].

1.2 Fuzzy Technique used for heart disease detection.

In India and other nations, coronary heart disorders (CHD) constitute a leading cause of death. In the Indian population, hypertension, serum cholesterol, diabetes, smoking, hypercholesterolemia, and a high body mass index (BMI) are the primary risk factors for CHD [16]. This study will now cover the fuzzy logic-based method utilized to identify cardiac disease.

1.2.1 Fuzzy C-Mean Clustering (FCM) technique

The process of grouping data points, objects, or instances into various clusters is known as clustering. Data instances in one cluster are more similar to one another than instances in another cluster, yet they differ from one another [17]. In 1973, Duda and Hart originally presented the fuzzy C-mean (FCM) approach, which under ambiguous conditions delivers a more precise clustering than traditional clustering techniques like K-Means and KMedoids [18].

The centroid and domain of each subcluster are chosen iteratively by the FCM clustering algorithm to minimize the provided cost function. This strategy's major objective was to distribute data across clusters in order to reduce variation between clusters. The FCM algorithm is also used for medical image processing, such as picture enhancement or image segmentation. A data set's pixels are classified as a distance-dependent cluster by the FCM method [19].

1.2.2 Hybrid Fuzzy Logic Technique

There exist several fuzzy hybrid categories employed for the detection of heart illness, such as the fuzzy logic-based machine learning model, the fuzzy logic with genetic algorithm approach, the fuzzy logic decision tree model, and the fuzzy clustering with machine learning method, among others.

a) Fuzzy Decision Tree (FDT) technique

Fuzzy Decision Trees (FDTs) can be classified into two primary categories based on the method employed to generate child nodes from a parent node: multi-way split trees and binary (or two-way) split trees. Binary split trees exhibit a defining characteristic of recursively dividing the attribute space into two distinct subspaces. In contrast, multi-way split trees employ a partitioning strategy that divides the space into multiple subspaces, resulting in parent nodes generating, on average, more than two offspring nodes [20].

The architecture of the fuzzy-based decision tree technique is shown in Figure 3. It demonstrates how the fuzzification model is fed the heart disease dataset along with input parameters or attributes. To enhance the model's functionality, the architecture applies the fuzzification approach to a decision tree machine learning model. The result of the Membership function's classification form, which categorizes cardiac disease, is then produced.

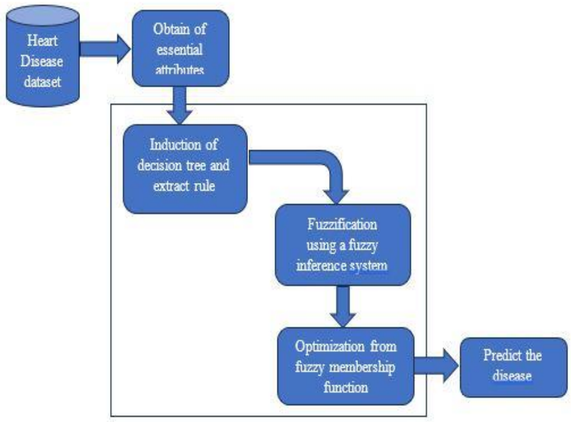


Fig 3. Architecture of the Fuzzy based decision tree model [21].

b) Genetic Fuzzy Model

Genetic Fuzzy technique is a method of stochastic searching those aids in finding the best answer to an optimization problem. The number of attributes in the dataset is decreased by the use of GA for feature selection, which in turn helps to focus the search. The attributes in the dataset are chosen using GA in the proposed study, and the fuzzy inference system then conducts classification and prediction [22]. There are two stages to the genetic fuzzy model: (1) a method that generates weighted fuzzy rules automatically, and (2) the creation of a genetic algorithm-based fuzzy model for coronary artery disease risk level forecasting. At now, the construction of the fuzzy framework has been carried out based on the utilization of weighted fuzzy criteria and the selection of cases of higher quality [23]. Figure 4 depicts the concept under consideration, wherein the genetic process acquires knowledge or adjusts various constituents of a fuzzy rulebased system (FRBS).

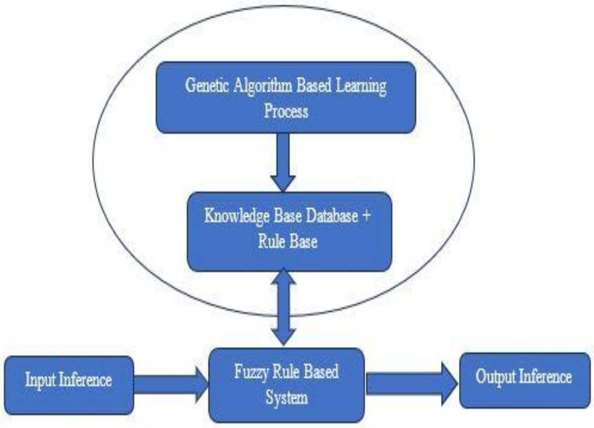


Fig 4. Genetic Fuzzy Systems [24]

1.2.3 Fuzzy Interface System

FISs are well-liked rule-based techniques for modeling ambiguous and imprecise information. Language-based descriptions of patients' symptoms are used to generate suggestions, and fuzzy inference techniques are also employed. The knowledge base contains IF-THEN fuzzy rules that represent the domain knowledge. The "interpretability" of these systems the capacity to clearly convey the rationale underlying the rules in a manner that is understandable to humans is their main strength [25]. Among the most effective categorization and prediction modeling systems used to establish precise relationships between input and output parameters in various processes is the ANFIS (adaptive neuro-fuzzy inference system) [26]. The fundamental component of the inference process, which is used to identify cardiac disease, is shown in Figure 5. It is demonstrated that the fuzzy expressions or patient's parameters give numerical (crisp) input that is transmitted, transformed to fuzzy, used with the fuzzy set to produce fuzzy output (inference engine), and then converted back to crisp output (defuzzification) from fuzzy output.

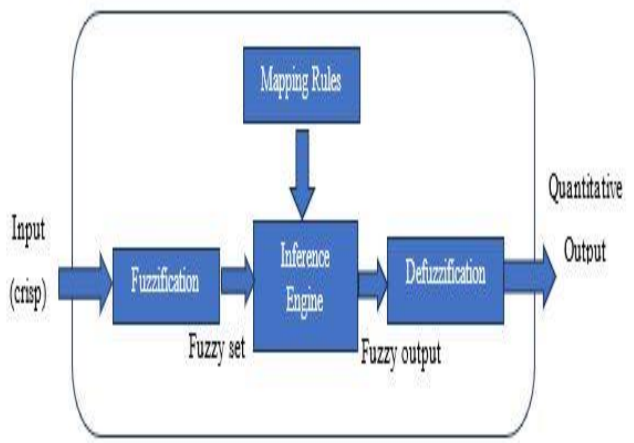


Fig 5. Design of a fuzzy inference system [27]

2. Literature review

The complex world of knowledge-based systems is explored in this literature review, with an emphasis on the application of sophisticated fuzzy techniques to evaluate and enhance system performance. In this review, we specifically focus on advanced fuzzy techniques and examine the state of research, methodologies, and emerging trends in the design and performance evaluation of knowledge-based systems. This section provides a thorough examination of the current body of literature concerning the topic of "Enhanced Cardiovascular Disease Detection using Advanced Fuzzy TOPSIS and Knowledge-Based Systems." The objective of this review is to address existing gaps in the field, provide a comprehensive comprehension of the subject matter, and lay the foundation for future advancements in knowledge-based systems.

Wojcik et al., (2023) [28] developed a health expert system for determining the severity of coronary artery tumors in individuals who suffer from coronary artery disease using fuzzy sets as the underlying data structure. The use of actual data in testing the intelligent system. In the end, it was found that the level of structural abnormality of the coronary artery in individuals with different kinds of coronary disease was 95%, according to the opinion of specialists.

Taylan et al., (2023) [29] discovered that the timely and accurate detection of heart disease and stroke is crucial in order to minimize the risk of experiencing a myocardial infarction. As a result, an approach that makes use of the adaptive neuro-fuzzy inference system (ANFIS) technology has been presented as a means of predicting, classifying, and enhancing the diagnostic efficiency of CVDs. According to the findings of the numerical study, the level of accuracy of prediction offered by ANFIS throughout the training phase is 96.56%.

Kharya et al., (2023) [30] initiated a novel idea known as the fuzzy-weighted Bayesian belief network (FWBBN), which was used to construct and create a healthcare diagnosis support tool based on the BBN. The fuzzy concept is being used for characteristics to cope with real-life circumstances to eliminate sharp boundary concerns that exist in the medical field. In conclusion, it was determined that FWBBN, in comparison to the traditional Bayesian model, is capable of being applied in a manner that is both highly effective and precisely precise in terms of high efficiency and low time intricacy.

Seslier and Karakus (2023) [31] investigated that About 46% of the deaths of people in the world, excluding communicable diseases and accidents, are because of CVDs. In this analysis, various machine learning methods are used to determine heart disease. At last, it concluded that among various classifiers such as Logistics regression, SVM, Naïve Bayes, and Random Forest (RF), the SVM technique achieved the best accuracy outcomes at 87.91%.

Nadakinamani et al., (2022) [32] evaluated a wide variety of cutting-edge machine learning techniques to develop a CVD forecasting method that is very reliable. To determine which machine learning approach is the most appropriate, the efficiency of the suggested CDPS was measured across several different criteria. The DT approach performed very well, with a maximum accuracy of 100%, when it came to forecasting individuals who would be diagnosed with CVD.

Cenitta et al., (2022) [33] recently introduced ischemic heart disease innovative missing value imputation techniques (IHDMIT) using fuzzy-rough sets and their expansions. The novel IHDMIT with RF classification against the cutting-edge technique of expectation maximization, fuzzy C means, and fuzzy roughest is evaluated. According to the findings, the suggested IHDMIT RF classifier achieves a higher accuracy of 93%.

Doppala et al., (2022) [34] discovered that artificial intelligence can transform unprocessed medical data into a useful knowledge base for decision-making and forecasting. The findings of this study offer a robust ensemble model. As a consequence, the suggested model was shown to have an overall success rate of 96.75% on the CVD dataset.

Yilmaz et al., (2022) [35] developed three distinct models for classifying coronary heart disease using RF, logistic regression, and SVM methods respectively. Accuracy served as the criterion for determining how well the models performed. At last, it concluded that the RF classifier had the greatest accuracy of 92.9% among all of the classifiers.

Thukral et al., (2019) [36] discussed about the various applications of fuzzy-based applications in the medical field. The study examined eight widespread medical conditions, including diabetes, cholera, liver, cholera, Parkinson's disease, breast cancer, bronchitis, and asthma. It was based on many fuzzy applications in the medical area. The major goal is to investigate and apply fuzzy logic in future fields, including those outside of current medicine, based on these various medical applications.

Kasbe et al., (2017) [37] introduced a fuzzy-based approach for heart disease identification. The objective of their study is to create a fuzzy expert system that can detect people who are at risk for heart disease. The analysis of the patient's cardiac illness must take into account several variables, which makes the doctor's job challenging. However, scientists need a precise instrument that takes into account and pinpoints risk variables based on given data. Three main steps are involved in the proposed fuzzy expert system. namely, rule basis, defuzzification, and fuzzification.

H. Parveen et al, (2023) [38] observed that healthcare systems have constantly improved because to many sorts of healthcare information. This information is derived from a variety of fresh sources, including digital files, mobile devices, and medical monitoring equipment. From vast healthcare data sets, deep learning and data fusion techniques can be used to create projections that are more exact, accurate, and dependable. This study develops a four-stage paradigm for estimating risk exposure. To predict health risks from a patient's e-health record, a novel Ensemble Classifier (EC) with fuzzy method, neural network and Deep Belief Network (DBN) is described. The neural network is trained using the retrieved characteristics, whilst the fuzzy logic is trained using the returned knowledge sources. A DBN carries out the risk calculation for the disorder using the outputs of both neural networks and fuzzy logic. A cutting-edge hybrid method termed rain-leveraged dynamic butterfly optimization has been used to change the weight that DBN carries in order to improve the efficacy of risk factor forecasting.

3. Research objectives

To have a working knowledge of the many types of cardiac disease and the information mining strategies that can be used to prevent and predict them.

The goal of this project is to create a model for predicting cardiac disease based on knowledge-based systems and fuzzy approaches.

To demonstrate the resilience of the suggested model by comparing it with another conventional model in terms of accuracy and other performance evaluation parameters.

4. Research Methodology

In the framework of research methodology, the idea of designed architecture is investigated. Artificial neural networks and advanced fuzzy TOPSIS method are used to predict and classify the risk. Further, the suggested model is contrasted with the conventional technique to prove the robustness and efficiency of the model.

4.1 Technique used.

The proposed methodology employs two techniques. The following techniques are provided below:

- **Artificial Neural Network (ANN)**

The term ANN refers to a kind of computer model that was developed to imitate the way a human brain operates. Applications involving prediction, categorization, and pattern identification are the primary areas in which ANNs are used. ANN models can understand and recognize similar patterns via training, and then they can recognize the outcome for input data that are presented to the network [39-41]. Figure 6 shows the schematic structure of an ANN where input data is fed to the input.

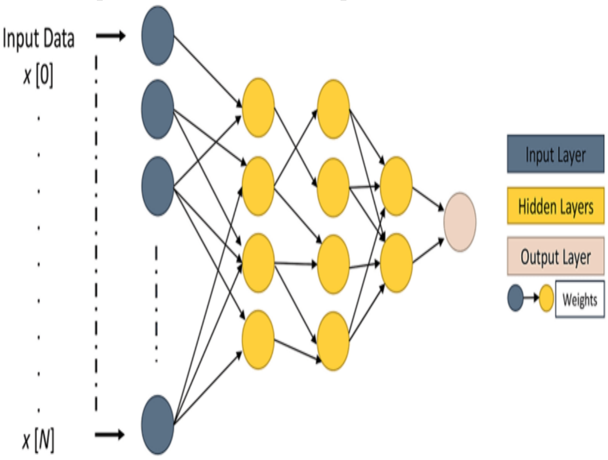


Fig 6 Structure of ANN [42]

- **Fuzzy Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS)**

The fuzzy TOPSIS was first developed by Yoon and Hwang. The method now holds the highest level of recognition for addressing decision-making challenges that involve the utilization of many criteria. This strategy is predicated on the theory that the selected option ought to have the shortest route to the Positive Ideal Solution, while also having the greatest length to the Negative Ideal Solution [43].

The procedure for implementing fuzzy TOPSIS involves the following steps:

Step 1: Ratings should include conditions and options.

It initially worked under the presumption that there are K members in the decision-making group. The weight of the criteria C_j is represented by $\widetilde{w}_j^k = \left(w_{j1}^k, b_{j2}^k, c_{j3}^k\right)$, and $\tilde{x}_{ij}^k = \left(a_{ij}^k, b_{ij}^k, c_{ij}^k\right)$ is the definition of the fuzzy score of the K^{th} selection maker on substitution A_i regarding condition C_j.

Step 2: Determine the cumulative fuzzy ratings for the given alternatives, as well as the aggregated fuzzy weights for each criterion.

The calculation of the weighted average fuzzy rating for the i^{th} choice with respect to the j^{th} criteria is determined using the following formula: $\tilde{x}_{ij} = (a_{ij}, b_{ij,}c_{ij})$.

$$a_{ij} = \frac{min}{k}\{a_{ij}^k\}, b_{ij} = \frac{1}{K}\sum_{k=1}^{k} b_{ij}^k, c_{ij} = \frac{max}{k}\{c_{ij}^k\}. \quad (1)$$

Formulae are employed to calculate the aggregate fuzzy weight $\tilde{w}_j = (w_{j1}, w_{j2}, w_{j3})$ for each criteria C_j.

$$w_{j1} = \frac{min}{k}\{w_{j1}^k\}, w_{j2} = \frac{1}{K}\sum_{k=1}^{k} w_{j2}^k, w_{j3} = \frac{max}{k}\{w_{j3}^k\} \quad (2)$$

Step 3 Perform the computation on the normalized fuzzy decision matrix.

$\tilde{R} = [\tilde{r}_{ij}]$ is the formula for the normalized fuzzy decision matrix, where,

R˜ denotes the normalized fuzzy decision matrix, $\tilde{R} = [\tilde{r}_{ij}]$, where.

$$\tilde{r}_{ij} = \left(\frac{a_{ij}}{c_j^*}, \frac{a_{ij}}{c_j^*}, \frac{a_{ij}}{c_j^*}\right) \text{ and } c_j^* = {}^{max}_{i}\{c_{ij}\} \quad (3)$$

$$\tilde{r}_{ij} = \left(\frac{a_j^-}{c_{ij}}, \frac{a_j^-}{b_{ij}}, \frac{a_j^-}{a_{ij}}\right) \quad \text{and} \quad c_j^- = {}^{min}_{i}\{a_{ij}\} \quad (4)$$

Step 4: Perform the computations necessary to produce the weighted normalized fuzzy decision matrix.

$\tilde{V}$ denote the weighted normalized fuzzy decision matrix, $\tilde{V} = (\tilde{v}_{ij})$, where $\tilde{v}_{ij} = \tilde{r}_{ij} \times w_j$.

Step 5: Determine both the fuzzy negative ideal solution (FNIS) and the fuzzy positive ideal solution (FPIS). These are the FPIS and FNIS calculations:

$$A^* = (\tilde{v}_1^*, \tilde{v}_1^*, \cdots \tilde{v}_n^*), \text{ where } \tilde{v}_j^* = {}^{max}_{i}\{v_{ij3}\} \quad (5)$$

$$A^- = (\tilde{v}_1^-, \tilde{v}_1^-, \cdots \tilde{v}_n^-), \text{ where } \tilde{v}_j^- = {}^{min}_{i}\{v_{ij1}\}. \quad (6)$$

Step 6: Calculate the length between each option and the FPIS and FNIS.

The distances between each alternative A_i and the First Preference Ideal Solution (FPIS) and the First Non-Preference Ideal Solution (FNIS) should be provided:

$$d_i^* = \sum_{j=1}^{n} d(\tilde{v}_{ij}, \tilde{v}_j^-) \quad (7)$$

Step 7: Calculate the CC_i similarity coefficient for each option.

The proximity coefficient CC_i is determined for each option A_i using the formula below:

$$CC_i = \frac{d_i^-}{d_i^- + d_i^*} \quad (8)$$

Step 8: Rank all alternatives.

The choice with the highest proximity coefficient is the most favorable one [33].

- **Proposed Algorithm**

Start

1. Read the Electronic health record of the patient as input variable = I
2. Perform data_ pre-processing (Data cleaning, Tokenization) on → I
3. Weight determination using AHP
4. Select features (I) → using Pearson correlation
5. For → KE, FE; where, KE=knowledge_ extraction, FE= feature_ extraction
 divide the selected features in group as →g1, g2.
6. Perform→ KE (Ontology-based, Improved semantic similarity) → on g1
7. Perform → FE (statistical features, Information entropy) → using PCA → on g2
8. G1 ← g1+g2
9. Optimize →G1
10. For prediction risk → using→ ANN
11. Classify risk level → using →fuzzy TOPSIS
12. Output

End

4.2 Proposed methodology.

Figure 7 displays the proposed methodology's flowchart. In the below methodology determines the sequence of operations involves in this research methodology. In this methodology we are using an electronic health record (EHR) dataset which is divided in 2 parts. 30% of the dataset is used in testing purpose and remaining 70% used for training. Firstly, the dataset is preprocessed which involving the data cleaning and the tokenization. The data cleaning process eliminate the duplicate, incorrect format data, or an incomplete data. After that this data is breakdown into small chunks or object called tokenization. This helps to increase the speed and performance of the model. After the data preprocessing now the weight is determine using the analytic hierarchy process (AHP). The calculated weight helps to determine the importance of priority or preference for different criteria using a structured manner. After assigning the preference of the dataset. It goes to the feature selection process. In the Pearson correlation method, the correlation coefficient varies between the -1 to +1. If the value closer to 0 represents the weak correlation if the value closer to -1 represents the strong negative correlation and the value closer to +1 shows a positive correlation. The feature is extracted using a principal component analysis (PCA) technique. It reduces the number of input dimension of the input dataset. Using this method, linear combinations of the original features are generated; however, the relationship between these linear combinations and the original features may not be as straightforward or easily interpretable as one might expect. Traditional feature selection techniques, such as filter, wrapper, or embedded methods, may be more suitable for your needs if you require a method for feature selection that keeps the interpretability of individual features intact. The knowledge is extracted from dataset using an ontology-based knowledge extraction which helps to organize the large volume of our dataset in structured and organized manner. After the feature extraction the optimization of feature is done which makes the input dataset ready for an early prediction of cardiovascular disease. In this research we use an artificial neural network (ANN) model for the early risk prediction of the cardiovascular disease. The risk is predicted by an ANN network and the risk level is classified by Fuzzy topsis technique. This model (ANN + Fuzzy topsis) predict the cardiovascular disease with a model accuracy, specificity, sensitivity, FPR rate and various other parameters.

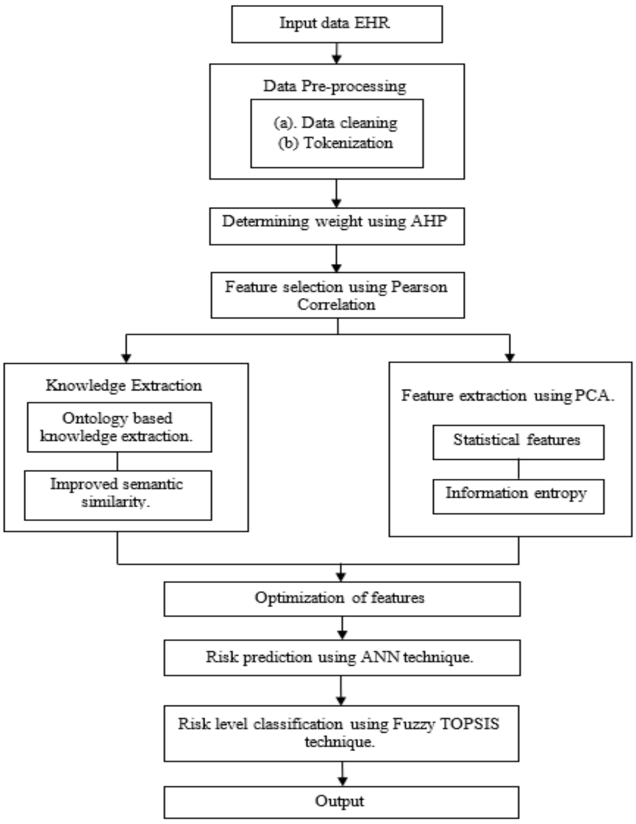


Fig 7. Flowchart of the suggested methodology

5. Results and discussion

In this section, the outcomes that were attained through the application of the suggested methodology are thoroughly addressed. The dataset is explained in detail. Performance-measured parameters are calculated and finally, the proposed model is compared with the conventional technique.

5.1 Dataset

The UCI Cleveland dataset is used in this research. It is a collection of EHR of the patients. Information such as age, gender, chest pain type, blood pressure, ECG., etc. is stored in the records. The dataset is split into two parts. 30% of the data is utilized to test the model after it has been trained on 70% of the data. The code implementation is done using python language.

Sensitivity analysis of fuzzy Topsis ANN based model.

In this analysis, the sensitivity of the suggested model is calculated on different learning percentages such as 60%, 70%, 80%, and 90%. It is seen in Table 1 and Figure 8 that as the learning percentage of the dataset increases the sensitivity of the suggested model is also increased. The sensitivity of the suggested model can be calculated by the following formulae:

$$Sensitivity = \frac{true\ positive}{true\ positive + false\ negative} \qquad (10)$$

Table 1. Sensitivity at different learning percentages

Parameter	Learning percentage			
	60%	**70%**	**80%**	**90%**
Sensitivity	0.992	0.994	0.995	0.996

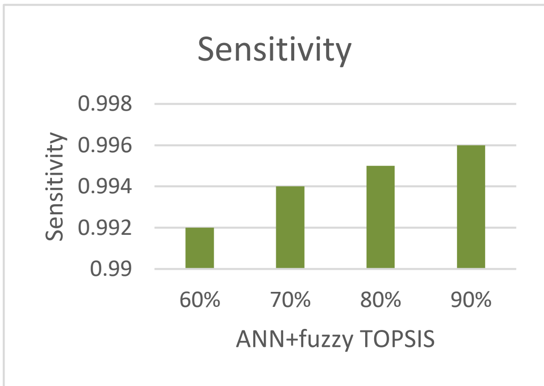


Fig 7 Graph showing the Sensitivity of the proposed model.

Specificity analysis of ANN and Fuzzy topsis method.

In this analysis, the specificity of the suggested model is calculated on different learning percentages such as 60%, 70%, 80%, and 90%. It is seen in Table 2 and Figure 9 that as the learning percentage of the dataset increases the specificity of the suggested model is also increased. The specificity of the suggested model can be calculated by the following formulae:

$$Specificity = \frac{Total\ negative}{total\ negative + false\ positive} \qquad (11)$$

Table 2. Specificity at different learning percentages

Parameter	Learning percentage			
	60%	**70%**	**80%**	**90%**
Specificity	0.975	0.976	0.978	0.978

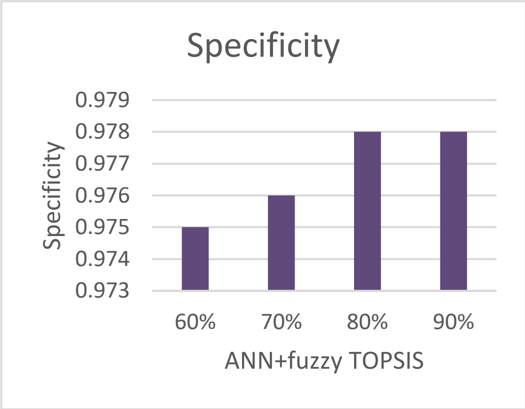


Fig 8 Graph showing the Sensitivity of the proposed model.

Accuracy determination of fuzzy Topsis ANN based model.

In this analysis, the accuracy of the suggested model is calculated on different learning percentages such as 60%, 70%, 80%, and 90%. It is seen in Table 3 and Figure 10 that as the learning percentage of the dataset increases the accuracy of the proposed model is also increased. The accuracy of the proposed model can be calculated by the following formulae:

$$Accuracy = \frac{true\ negative + true\ positive}{total\ no.of\ results} \quad (12)$$

Table 3 Accuracy at different learning percentages

Parameter	Learning percentage			
	60%	**70%**	**80%**	**90%**
Accuracy	0.981	0.983	0.986	0.99

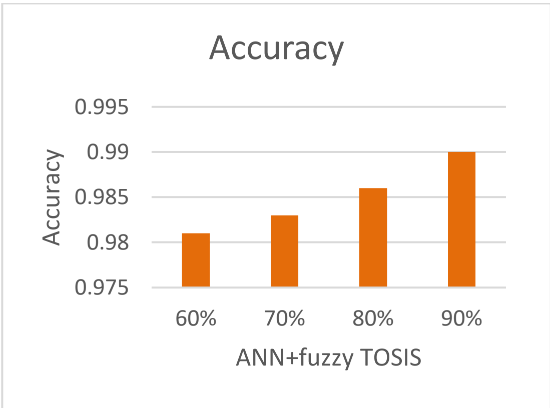


Fig 9 Graph showing the Accuracy of the proposed model.

Precision analysis of fuzzy Topsis ANN based model.

In this analysis, the precision of the suggested model is calculated on different learning percentages such as 60%, 70%, 80%, and 90%. It is seen in table 4 and figure 11 that as the learning percentage of the dataset increases the precision of the suggested model is also increased. The precision of the suggested model can be calculated by the following formulae:

$$Precision = \frac{True\ positive}{Total\ positive\ values} \quad (13)$$

Table 4. Precision at different learning percentages

Parameter	Learning percentage			
	60%	**70%**	**80%**	**90%**
Precision	0.962	0.970	0.974	0.98

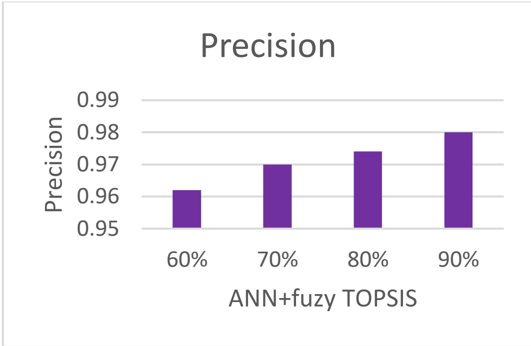


Fig 10 Graph showing the Precision of the proposed model.

F-measure analysis of fuzzy Topsis ANN based models.

In this analysis, the F-measure of the proposed model is calculated on different learning percentages such as 60%, 70%, 80%, and 90%. It is seen in table 5 and figure 12 that as the learning percentage of the dataset increases the sensitivity of the proposed model is also increased. The F-measure of the proposed model can be calculated by the following formulae:

$$F - measure = \frac{Precision \times recall}{precision + recall} \quad (14)$$

Table 5. F-measure at different learning percentage

Parameter	Learning percentage			
	60%	**70%**	**80%**	**90%**
F-measure	0.912	0.981	0.982	0.981

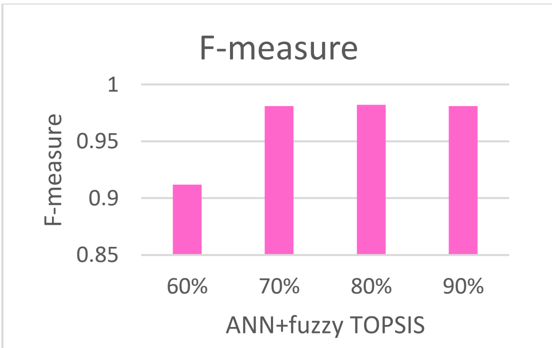


Fig 11 Graph showing the Sensitivity of the proposed model.

MCC analysis of fuzzy Topsis ANN based model.

In this analysis, the Mathews Correlation Coefficient (MCC) of the proposed model is calculated on different learning percentages such as 60%, 70%, 80%, and 90%. It is seen in table 6 and figure 13 that as the learning percentage of the dataset increases the MCC of the proposed model is also increased.

Table 6 MCC at different learning percentage

Parameter	Learning percentage			
	60%	**70%**	**80%**	**90%**
MCC	0.96	0.963	0.971	0.98

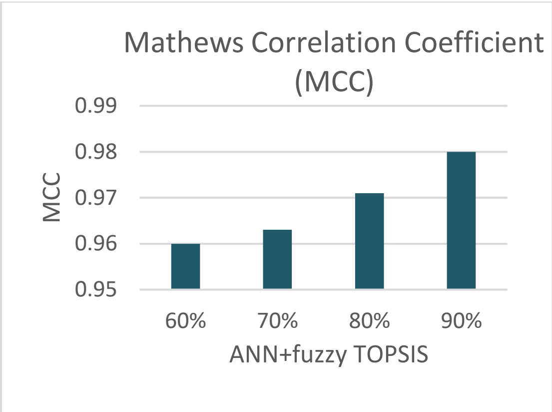


Fig 12 Graph showing MCC of the proposed model.

Result 7:

In this analysis, the Negative Predictive Value (NPV) of the proposed model is calculated on different learning percentages such as 60%, 70%, 80%, and 90%. It is seen in table 7 and figure 14 that as the learning percentage of the dataset increases the NPV of the proposed model is also increased.

Table 7. NPV at different learning percentages

Parameter	Learning percentage			
	60%	**70%**	**80%**	**90%**
NPV	0.994	0.995	0.998	1.0

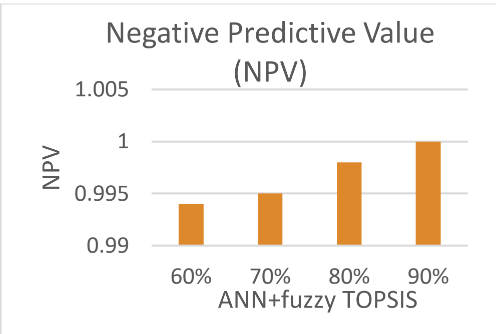


Fig 13 Graph showing NPV of the proposed model.

FPR analysis of fuzzy Topsis ANN based model.

In this analysis, the False Positive Rate (FPR) of the proposed model is calculated on different learning percentages such as 60%, 70%, 80%, and 90%. It is seen in table 8 and figure 15 that as the learning percentage of the dataset increases the FPR of the proposed model is decreased.

Table 8. FPR at different learning percentages

Parameter	Learning percentage			
	60%	**70%**	**80%**	**90%**
FPR	0.024	0.020	0.016	0.015

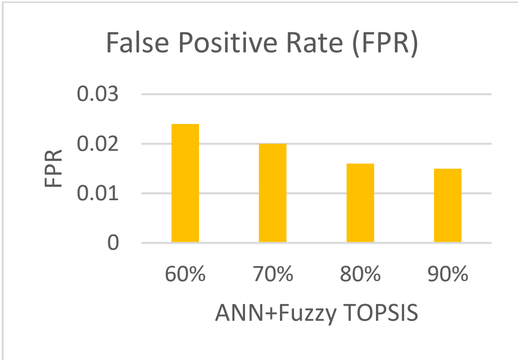


Fig 14 Graph showing the FPR of the proposed model.

FNR analysis of fuzzy Topsis ANN based model.

In this analysis, the False Negative Rate (FNR) of the proposed model is calculated on different learning percentages such as 60%, 70%, 80%, and 90%. It is seen in table 9 and figure 16 that as the learning percentage of the dataset increases the FNR of the proposed model is also increased.

Table 9. FNR at different learning percentages

Parameter	Learning percentage			
	60%	**70%**	**80%**	**90%**
FNR	0.008	0.006	0.004	0.003

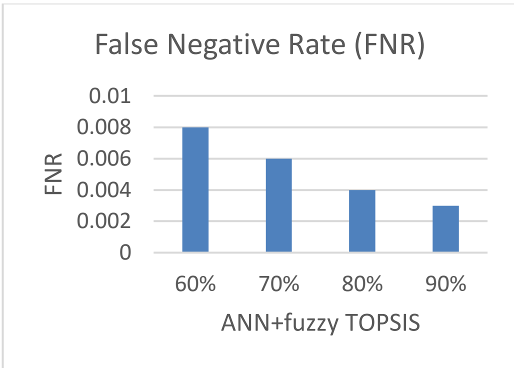


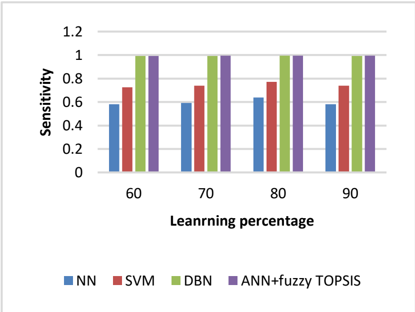
Fig 15 Graph showing FNR of the proposed model.

- **Comparative analysis**

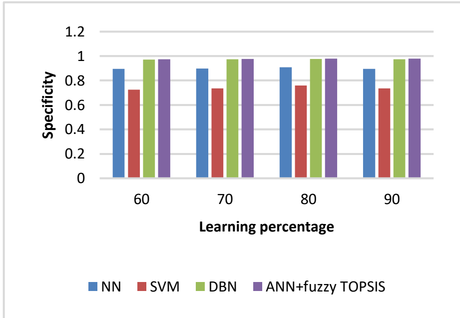
In this section, the proposed paradigm is contrasted with existing traditional approaches such as Neural network (NN), SVM, and DBN. It is compared based on positive metrics parameters such as sensitivity, specificity, accuracy, and precision. Figure 17 (a) shows the comparison of the conventional technique with the proposed model based on sensitivity, and it is seen the sensitivity of the suggested model is higher among all the methods. Figure 17 (b) shows the comparison of the conventional technique with the proposed model based on specificity, and it is seen the specificity of the suggested model is higher among all the methods. Figure 17 (c) shows the comparison of the conventional technique with the proposed model based on accuracy, and it is seen that the precision of the suggested model is higher among all the methods. Figure 17 (d) shows the comparison of the conventional technique with the proposed model based on precision, and it is seen the precision of the suggested model is higher among all the methods. Table 10 shows the overall comparison of the suggested model with other conventional methods in terms of sensitivity, specificity, accuracy, and precision.

Table 10. Comparison table

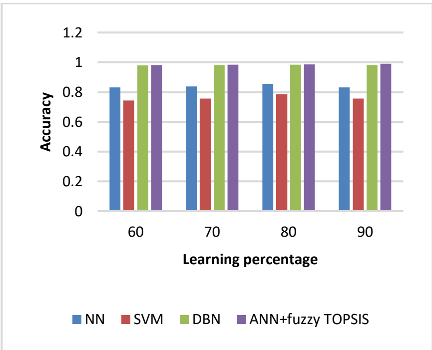
Model	Sensitivity				Specificity				Accuracy				Precision			
	60	**70**	**80**	**90**	**60**	**70**	**80**	**90**	**60**	**70**	**80**	**90**	**60**	**70**	**80**	**90**
NN	0.581	0.593	0.639	0.580	0.895	0.898	0.909	0.895	0.832	0.837	0.855	0.832	0.581	0.593	0.639	0.580
SVM	0.726	0.740	0.772	0.739	0.725	0.735	0.759	0.735	0.744	0.757	0.786	0.757	0.491	0.516	0.573	0.516
DBN	0.991	0.992	0.994	0.992	0.971	0.973	0.976	0.973	0.979	0.981	0.984	0.981	0.960	0.963	0.971	0.963
ANN+TOPSIS	0.992	0.994	0.995	0.996	0.975	0.976	0.978	0.978	0.981	0.983	0.986	0.99	0.962	0.97	0.974	0.98



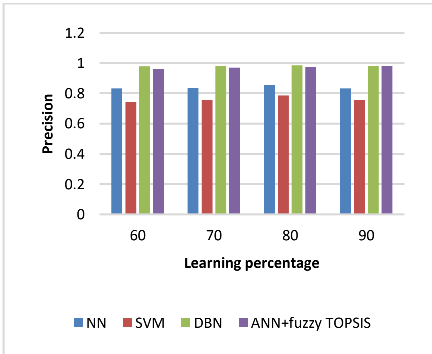
(a)



(b)



(c)



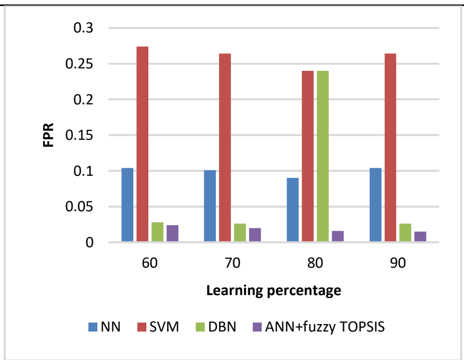
(d)

Fig 16 Comparison of the proposed work's performance to that of similar current schemes in terms of (a) Sensitivity (b) specificity (c) accuracy (d) precision

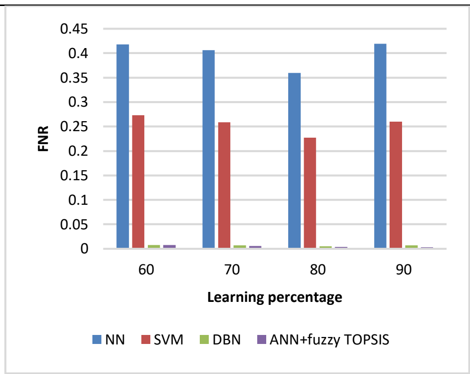
In Figures 18 (a) and (b), a comparison of the proposed model with other conventional techniques is shown based on FPR and FNR. Figure 18 (a) demonstrates that the proposed model's total FPR is quite low when compared to other methods and figure 18 (b) demonstrates that, in comparison to previous methods, the suggested model's total FNR is quite low. Table 11 shows the overall comparison of the proposed model with other conventional techniques in terms of FPR and FNR.

Table 11. Comparison table

Model	FPR				FNR			
	60	**70**	**80**	**90**	**60**	**70**	**80**	**90**
NN	0.104	0.101	0.090	0.104	0.418	0.406	0.360	0.419
SVM	0.274	0.264	0.240	0.264	0.273	0.259	0.227	0.260
DBN	0.028	0.026	0.023	0.026	0.008	0.007	0.005	0.007
ANN+TOPSIS	0.024	0.02	0.016	0.015	0.008	0.006	0.004	0.003



(a)



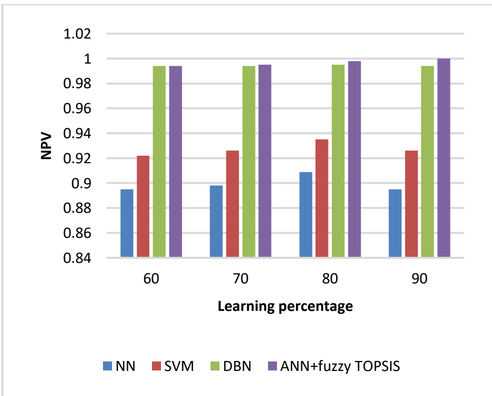
(b)

Fig 17 Comparison of the proposed work's performance to that of similar current schemes in terms of (a) FPR (b) FNR

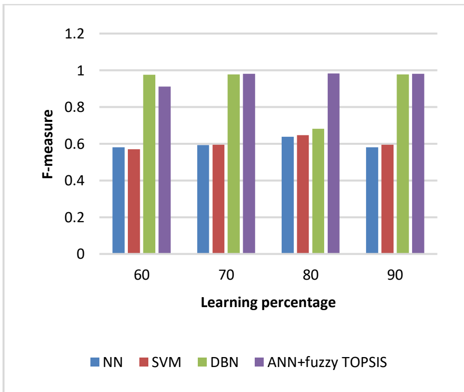
In figure 19 (a), (b), and (c) a comparison of the proposed model with other conventional techniques is shown based on NPV, F-measure, and MCC respectively. Figure 19 (a) shows that the overall NPV of the proposed model is quite low as compared to other techniques. Figure 19 (b) shows that the overall F-measure of the suggested model is quite low as compared to other methods. Figure 19 (c) demonstrates that the suggested model's total MCC is fairly low when compared to other methods. Table 12 shows the overall comparison of the proposed model with other conventional techniques in terms of F-measure, MCC, and NPV.

TABLE 12. Comparison table

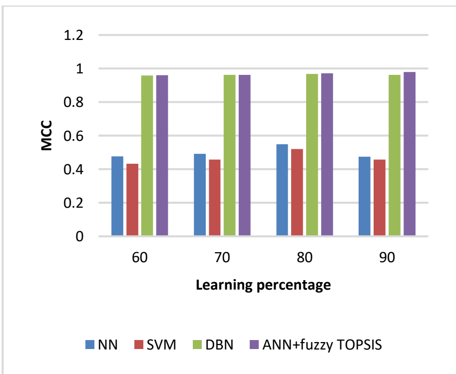
Model	F-measure				MCC				NPV			
	60	70	80	90	60	70	80	90	60	70	80	90
NN	0.581	0.593	0.639	0.580	0.477	0.491	0.549	0.475	0.895	0.898	0.909	0.895
SVM	0.571	0.594	0.647	0.594	0.432	0.458	0.520	0.458	0.922	0.926	0.935	0.926
DBN	0.975	0.978	0.682	0.977	0.959	0.962	0.968	0.962	0.994	0.994	0.995	0.994
ANN+TOPSIS	0.912	0.981	0.982	0.981	0.96	0.963	0.971	0.98	0.994	0.995	0.998	1



(a)



(b)



(c)

Fig 18 Comparison of the proposed work's performance to that of similar current schemes in terms of (a) NPV, (b) F-measure, and (c) MCC

6. Conclusion

In most nations, heart disease has surpassed all others as the top cause of mortality among adults during the last decade. Heart disease is a common health problem, and early identification helps physicians treat these patients more effectively. The proposed model achieved encouraging outcomes in predicting and classifying CAD risk. The AHP method's attribute weights can help in making sound decisions about illness diagnosis. By reducing the need for invasive biopsies and clinical procedures, the system protects users from the risks associated with CVD diagnosis. The system provides a second perspective to the doctor to confirm the diagnosis of sickness. According to numerical analysis, the suggested model outperformed standard techniques in many respects, including accuracy (0.99), precision (0.98), specificity (0.978), F-measure (0.981), sensitivity (0.996), and many more. In future work, the researcher will carry on their investigation into and development of effective heuristic methods for attribute reduction strategies, to manage enormous quantities of features and big numbers of records. The utilization of a knowledge-based system within the framework of fuzzy logic has resulted in notable progress within the realm of AI and knowledge-based systems. The following inquiries may examine the amalgamation of sophisticated fuzzy methodologies with machine learning algorithms, including deep learning and neural networks. Knowledge-based systems may become more adaptive and capable of learning with this combination. Fuzzy logic knowledge-based systems can be improved for real-time decision assistance in vital applications like finance, healthcare, and autonomous systems. The point of view of specialists in a variety of disciplines, such as psychology, sociology, and ethics, to make sure that the creation of knowledge-based systems using cutting-edge fuzzy techniques is in line with societal demands and human values Future research on this topic will focus on improving knowledge-based systems' dependability, interpretability, and ethics in addition to their performance. The incorporation of sophisticated fuzzy techniques will remain essential in solving challenging real-world issues in a variety of fields as technology develops.

7. Acknowledgments

The author would like to express her sincere gratitude to the co-author for their invaluable contributions to this research work. Without their support, this study would not have been possible. The author would like to state that this research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. Funding for this research was covered solely by the author(s) of this article.

References

[1] P. Bhatnagar et al., “The epidemiology of cardiovascular disease in the UK 2014,” Heart, vol. 101, no. 15, pp. 1182-1189, 2015 [doi:10.1136/heartjnl-2015-307516].

[2] M. T. Maeder et al., “A clinical approach to obstructive sleep apnea as a risk factor for cardiovascular disease,” Vasc. Health Risk Manag., vol. 12, pp. 85-103, 2016 [doi:10.2147/VHRM.S74703].

[3] R. Alugubelli, “Exploratory study of artificial intelligence in healthcare,” Int. J. Innov. Eng. Res. Technol., vol. 3, no. 1, pp. 1-10, 2016.

[4] N. C. Long et al., “A highly accurate firefly based algorithm for heart disease prediction,” Expert Syst. Appl., vol. 42, no. 21, pp. 8221-8231, 2015 [doi:10.1016/j.eswa.2015.06.024].

[5] Chitra and V. Seenivasagam, “Heart disease prediction system using supervised learning classifier,” BIJSESC, vol. 3, no. 1, p. 1-7, 2013 [doi:10.9756/BIJSESC.4336].

[6] S. Jurado et al., “Hybrid methodologies for electricity load forecasting: Entropy-based feature selection with machine learning and soft computing techniques,” Energy, vol. 86, pp. 276-291, 2015 [doi:10.1016/j.energy.2015.04.039].

[7] M. Rostami et al., “Review of swarm intelligence-based feature selection methods,” Eng. Appl. Artif. Intell., vol. 100, p. 104210, 2021 [doi:10.1016/j.engappai.2021.104210].

[8] S. Bouktif et al., “Optimal deep learning LSTM model for electric load forecasting using feature selection and genetic algorithm: Comparison with machine learning approaches,” Energies, vol. 11, no. 7, p. 1636, 2018 [doi:10.3390/en11071636].

[9] M. Thiyagaraj and G. Suseendran, “Survey on heart disease prediction system based on data mining techniques,” Indian J. Innov. Dev., vol. 6, no. 1, pp. 1-9, 2017.

[10] H. B. Sola et al., “Interval type-2 fuzzy sets are generalization of interval-valued fuzzy sets: Toward a wider view on their relationship,” IEEE Trans. Fuzzy Syst., vol. 23, no. 5, pp. 1876-1882, 2014 [doi:10.1109/TFUZZ.2014.2362149].

[11] S. A. Mokeddem, “A fuzzy classification model for myocardial infarction risk assessment,” Appl. Intell., vol. 48, no. 5, pp. 1233-1250, 2018 [doi:10.1007/s10489-017-1102-1].

[12] P. A. Burrough and A. Frank, Geographic Objects with Indeterminate Boundaries. CRC Press, 2020.

[13] B.-J. Wen, “A biomedical application by using optimal fuzzy sliding-mode control” in Sliding Mode Control. IntechOpen, 2011.

[14] R. Dhanalakshmi et al., “RETRACTED ARTICLE: Association rule generation and classification with

fuzzy influence rule based on information mass value," J. Ambient Intell. Hum. Comput., vol. 12, no. 6, pp. 6613-6620, 2021 [doi:10.1007/s12652-020-02280-9].

[15] V. Krishnaiah et al., "Heart disease prediction system using data mining techniques and intelligent fuzzy approach: A review," Int. J. Comput. Appl., vol. 136, no. 2, pp. 43-51, 2016 [doi:10.5120/ijca2016908409].

[16] Bashir et al., "Improving heart disease prediction using feature selection approaches" in 16th international bhurban conference on applied sciences and technology (IBCAST), vol. 2019. IEEE, 2019, pp. 619-623 [doi:10.1109/IBCAST.2019.8667106].

[17] J. Han et al., Data Mining: Concepts and Techniques. Morgan Kaufmann, 2022.

[18] R. O. Duda and P. E. Hart, Pattern Classification and Scene Analysis. New York: Wiley, 1973.

[19] V. R. Elangovan et al., "Heart atherosclerosis detection using FCM+ kMeans Algorithm" in 2nd International Conference on Computation, Automation and Knowledge Management (ICCAKM), vol. 2021. IEEE, 2021, pp. 102-106 [doi:10.1109/ICCAKM50778.2021.9357719].

[20] Segatori et al., "On distributed fuzzy decision trees for big data," IEEE Trans. Fuzzy Syst., vol. 26, no. 1, pp. 174-192, 2017 [doi:10.1109/TFUZZ.2016.2646746].

[21] Pamela et al., 2013, "'A Fuzzy Optimization Technique for the Prediction of Coronary Heart Disease Using Decision Tree,' semantic scholar". Available at: https://www.semanticscholar.org/paper/A-Fuzzy-Optimization-Technique-for-the-Prediction-Pamela-Gayathri/aad05b5d65d39eb4ca2fa897566e6d4f619c849b.

[22] T. Santhanam and E. P. Ephzibah, "Heart disease prediction using hybrid genetic fuzzy model," Indian J. Sci. Technol., vol. 8, no. 9, p. 797, 2015 [doi:10.17485/ijst/2015/v8i9/52930].

[23] P. Sharma and K. Saxena, "Application of fuzzy logic and genetic algorithm in heart disease risk level prediction," Int. J. Syst. Assur. Eng. Manag., vol. 8, no. S2, pp. 1109-1125, 2017 [doi:10.1007/s13198-017-0578-8].

[24] "Genetic Fuzzy Systems and Multi-Objective Evolutionary Fuzzy Systems: Taxonomy, Current Research Trends and Prospects | Soft Computing and Intelligent Information Systems. Available at: sci2s.ugr.es. Available at: https://sci2s.ugr.es/gfs.

[25] Casalino et al., "A hierarchical fuzzy system for risk assessment of cardiovascular disease" in IEEE Conference on Evolving and Adaptive Intelligent Systems (EAIS), vol. 2020. IEEE, 2020, pp. 1-7 [doi:10.1109/EAIS48028.2020.9122750].

[26] A. M. Sagir and S. Sathasivam, "A novel adaptive neuro fuzzy inference system based classification model for heart disease prediction," Pertanika J. Sci. Technol., vol. 25, no. 1, 2017.

[27] Rizvi et al., "A fuzzy inference system (FIS) to evaluate the security readiness of cloud service providers," J. Cloud Comput., vol. 9, pp. 1-17, 2020.

[28] W. Wójcik et al., "Medical fuzzy-expert system for assessment of the degree of anatomical lesion of coronary arteries," Int. J. Environ. Res. Public Health, vol. 20, no. 2, p. 979, 2023 [doi:10.3390/ijerph20020979].

[29] O. Taylan et al., "Early prediction in classification of cardiovascular diseases with machine learning, neuro-fuzzy and statistical methods," Biology, vol. 12, no. 1, p. 117, 2023 [doi:10.3390/biology12010117].

[30] S. Kharya et al., "Fuzzy weighted Bayesian belief network: A medical knowledge-driven Bayesian model using fuzzy weighted rules," Int. J. Inf. Technol., vol. 15, no. 2, pp. 1117-1125, 2023 [doi:10.1007/s41870-022-01153-y].

[31] T. Seslier and M. Ö. Karakuş, In healthcare applications of machine learning algorithms for prediction of heart attacks Journal of Scientific Reports-A, vol. 051, pp. 358-370, 2022.

[32] R. G. Nadakinamani et al., "Clinical data analysis for prediction of cardiovascular disease using machine learning techniques," Comp. Intell. Neurosci., vol. 2022, 2973324, 2022 [doi:10.1155/2022/2973324].

[33] D. Cenitta et al., "Ischemic heart disease multiple imputation technique using machine learning algorithm," Eng. Sci., vol. 19, pp. 262-272, 2022 [doi:10.30919/es8d681].

[34] P. Doppala et al., "A reliable machine intelligence model for accurate identification of cardiovascular diseases using ensemble techniques," J. Healthc. Eng., vol. 2022, 2585235, 2022 [doi:10.1155/2022/2585235].

[35] R. Yilmaz and F. H. Yağin, "Early detection of coronary heart disease based on machine learning methods," Med. Rec., vol. 4, no. 1, pp. 1-6, 2022 [doi:10.37990/medr.1011924].

[36] S. Thukral and J. S. Bal, "Medical applications on fuzzy logic inference system: A review," Int. J. Adv. Netw. Appl., vol. 10, no. 4, pp. 3944-3950, 2019 [doi:10.35444/IJANA.2019.10046].

[37] T. Kasbe and R. S. Pippal, "Design of heart disease diagnosis system using fuzzy logic" in International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), vol. 2017. IEEE, 2017, pp. 3183-3187 [doi:10.1109/ICECDS.2017.8390044].

[38] H. Parveen et al., "Disease risk level prediction based on knowledge driven optimized deep ensemble framework," Biomed. Signal Process. Control, vol. 79, p. 103991, 2023 [doi:10.1016/j.bspc.2022.103991].

[39] Y. Atkov et al., “Coronary heart disease diagnosis by artificial neural networks including genetic polymorphisms and clinical parameters,” J. Cardiol., vol. 59, no. 2, pp. 190-194, 2012 [doi:10.1016/j.jjcc.2011.11.005].

[40] Knok et al., “Implementation of intelligent model for pneumonia detection,” Tehn. Glas., vol. 13, no. 4, pp. 315-322, 2019 [doi:10.31803/tg-20191023102807].

[41] A. H. N. Kishore and V. E. Jayanthi, “Neuro-fuzzy based medical decision support system for coronary artery disease diagnosis and risk level prediction,” J. Comp. Theor. Nanosci., vol. 15, no. 3, pp. 1027-1037, 2018 [doi:10.1166/jctn.2018.7198].

[42] “Artificial neural network modelling,” Ebrary. Available at: https://ebrary.net/132934/engineering/artificial_neural_network_modelling.

[43] Nădăban et al., “Fuzzy TOPSIS: A general view,” Procedia Comput. Sci., vol. 91, pp. 823-831, 2016 [doi:10.1016/j.procs.2016.07.088].