A Model for the Prediction of Cardiovascular Disease in IoMT Based on AI's Binary and Multi-Class Structures

Ahmed A. F. Osman[1], Nesren Farhah[2], Rajit Nair[3], Mohammed Awad Mohammed Ataelfadiel[1,*], Rami Taha Shehab[1,4]

[1]Applied College , King Faisal University, P.O. Box 400, Al-Ahsa 31982, Saudi Arabia

[2]Department of Health Informatics, College of Health Sciences, Saudi Electronic University, Riyadh 11673, Saudi Arabia

[3]VIT Bhopal University, Bhopal, India

[4]Vice-Presidency for Postgraduate Studies and Scientific Research, King Faisal University, Al-Ahsa 31982, Saudi Arabia

Emails: afadol@kfu.edu.sa; n.farhah@seu.edu.sa; Rajit.nair@vitbhopal.ac.in; melfadiel@Kfu.edu.sa; Rt.shehab@kfu.edu.sa

Abstract

Heart disease is a severe hazard to the public's health and safety because of the high rates of disability and mortality it causes. Accurate disease prediction and diagnosis are more critical than ever in this era of earlier illness prevention, faster disease detection, and earlier disease treatment. Artificial Intelligence (AI) and the Internet of Medical Things (IoMT) have made it possible to detect, forecast, and diagnose cardiovascular disease more precisely. However, the bulk of these prediction models can only state whether a person is sick; they cannot and do not forecast the severity of the ailment. We present a machine-learning-based technique for predicting cardiovascular disease. Using this strategy, we hope to perform binary and multimodal classifications at the same time. To get things started, we will go through the fuzzy-adaboost approach, which will serve as the foundation for the rest of our work. By combining fuzzy logic and the Adaboost method, this method aims to increase the number of applications that can use binary classification prediction to simplify data analysis. If it is completed, both objectives will be met, and we will eliminate overfitting by merging bagging and fuzzy adaboost into a single approach. It is the ideal solution to the challenge we are currently facing. Because it has a separate classification for the severity of the presentation of heart disease, the bagging fuzzy adaboost can be used for multiclassification prediction. This is because Adaboost's assessment of the severity of the observed heart disease presentations is unclear and imprecise. The results of the experiment reveal that, in addition to a wide range of other classes, the Bagging-Fuzzy-Adaboost can anticipate binary data accurately. When compared to traditional procedures, it is evident that this has significant advantages.

Keywords: Artificial Intelligence, Internet of Medical Things, fuzzy-adaboost approach, cardiovascular disease

1. Introduction

It is one of the most difficult diseases to treat and has the potential to be fatal for individuals who have it. The high morbidity and mortality rates associated with cardiovascular disease are one reason for this [1]. Heart disease is also one of the top causes of death. The epidemic has a catastrophic impact on the quality of life for those affected because of the high costs associated with monitoring and treating infected individuals, resulting in huge economic losses [2]. AI may one day be used to recognise, detect, and diagnose health diseases in their early stages. This enhances the likelihood that patients will seek proper medical assistance, which may result in a reduction in the severity of heart disease symptoms. It may be helpful to apply AI algorithms to analyse massive amounts of IoMT

DOI: https://doi.org/10.54216/FPA.210121
Received: March 05, 2025 Revised: June 01, 2025 Accepted: July 09, 2025

data in electronic healthcare systems, particularly those whose primary job is to provide real-time cardiac disease forecasts and diagnosis. This could be the case, especially for electronic healthcare systems whose primary role is to provide these services. The administrative and financial costs associated with adopting intelligent technologies for chronic illness diagnosis, monitoring, and management are considerably reduced. The huge hurdles associated with ensuring the high accuracy, generalizability, and stability of machine learning-based prediction algorithms and models create a bottleneck that must be addressed. That problem must be addressed immediately.

The degree of cardiac disease can be classified from 0 (no disease) to 4 (moderate sickness) based on the results of an angiographic scan (severe). In terms of tiers, the multiclassification technique has aided in categorising CVD risk; however, there is stillroom for improvement. Reduced variance and deviation are common techniques used in machine learning to increase accuracy [3]. Despite extensive research, the data complexity is very great, as is the level of accuracy that can be predicted from the various types of heart disease prediction. This part seeks to simultaneously classify heart disease forecasts into binary and multi-class groups; therefore, a trustworthy and accurate system is required. The bagging method, on the other hand, employs many independent samples to reduce the total variance of the model. The bagging approach is used to ensure the model's stability even further. We have explicitly demonstrated that our method outperforms previously available strategies for assessing the severity of cardiac disease in each patient [4]. The following is a summary of our contributions: This procedure explains heart disease statistics and makes it easier for Adaboost to draw conclusions from the data. In the second step of our technique, we combine the fuzzy-adaboost model with the bootstrap aggregating method to avoid overfitting. The fuzzy-adaboost model is used to do this. More information on this technique is provided in the paragraph that follows. Now that the variance and deviation of the Bagg-Fuzzy-Adaboost model have been cut down, predictions are more reliable and consistent.

The severity of the disease is determined using a baggage-fuzzy-adaboost algorithm designed for multiple categorization prediction.As a result; more diseases may be detected and treated precisely. According to the findings, the model's predictive capabilities for cardiovascular disease are of high quality and generally stable. The remaining paragraphs in this article will be organised as follows: The relevant studies can be found in "Section 2," which follows this section. Section 3 goes into detail about the proposed approach for coupled bagging-fuzzy-adaboost prediction. Section 4 that follows will exhibit and explain the simulation's results, as well as how we analysed them. Section 5 summarises the entire article.

2. Related Work

The ability to accurately anticipate cardiovascular disease has attracted a lot of attention in recent decades. Several papers [5–7] have been published on the development of prediction algorithms to aid in the exact detection of cardiac disease. The XGBoost classifier to diagnose heart illness by collecting characteristics from averaged magnetocardiography recordings and using those qualities to train the classifier. It may be difficult to find the underlying cause of the facts about heart disease. Data mining cannot accurately forecast the risk of cardiovascular disease unless several restrictions are addressed. When the benefits of limiting the number of traits are weighed against the potential disadvantages, the former option may prevail. Using feature selection techniques, you can limit the quantity of data fed into a classifier. By doing so, accuracy is improved. An associative classification based on a genetic algorithm is currently available. Genetic algorithms used for disease prediction may be able to extract the most helpful attribute set from large datasets if they take the real amount of data into consideration [8]. The decision tree method is well known for its effectiveness as a teaching aid for learning classification algorithms analysed and contrasted the naive Bayes, neural network, and decision tree techniques to discover their similarities and differences. However, because of its linear nature, the decision tree is not the best tool for dealing with continuous data. It has been established that hybrid-learning algorithms are excellent at recognising cardiac abnormalities [9-10]. According to the comparison results, the support vector machine (SVM) that uses the boosting method has the highest identification accuracy (90.5%). The use of fuzzy logic simplifies the underlying data, enabling the formulation of more accurate prediction models. [11] An artificial neural network (ANN) using fuzzy logic was utilised to evaluate the risk of cardiovascular disease [12]. To achieve this goal, we use a technique known as fuzzy analytic hierarchy [13], which allows us to give relative global weights to qualities based on the risk that they pose. When attempting to predict diseases, the possibility of overfitting must be considered. People who have been told they have heart disease are the only ones who can help us predict illness. These individuals can be classified into two groups. Medical personnel struggle to diagnose and treat patients due to a lack of knowledge regarding the severity of the risk for heart disease. As a result, getting the correct care for patients becomes difficult. It is critical to be able to predict the many types of cardiac disease.

3. The Proposed Approach

As demonstrated in Figure 1, our healthcare monitoring system makes use of artificial intelligence and the Internet of Medical Things. This system's primary objective is to monitor the users' health. To provide medicine, a technology known as the IoMT is used. This system is made up of a huge number of endpoint devices and sensors.

DOI: https://doi.org/10.54216/FPA.210121
Received: March 05, 2025 Revised: June 01, 2025 Accepted: July 09, 2025

There are several options for hosting the system's server, including on-premises and cloud hosting. Sensors and other devices that are surgically implanted into the subject or that the subject wears on their person can be used to collect vital sign data [14]. Blood glucose meters and ECG monitors are two examples of such devices. After gathering the data, it is delivered through wireless networks to a server located at the network's edge or in the cloud. On this server, the database management system will handle data in real-time or in the past. The patient's medical records are always available to care professionals, whenever they think it necessary. They are also accessible to any other individuals who are authorized. Furthermore, AI approaches such as machine learning could be used to improve the processing of health data to provide more accurate disease diagnoses and treatment plans. Both the predictability of forecasts and the lifespan of models have much room for improvement.

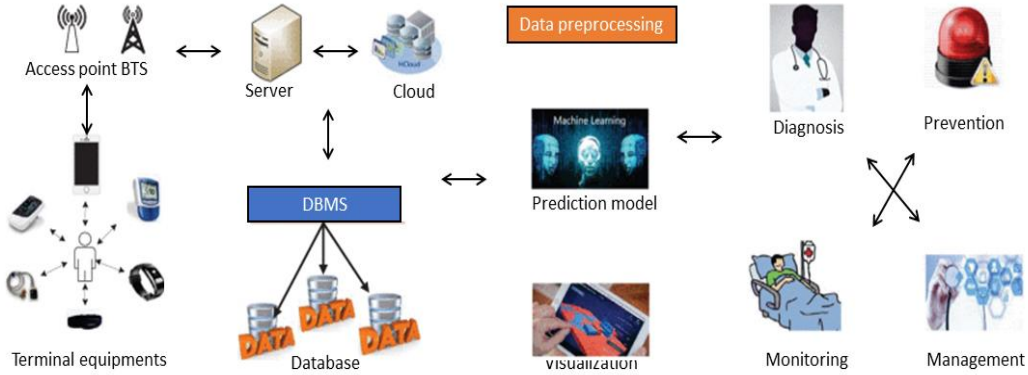


Figure 1. IoT-based System Architecture

We build our predictive model using the widely used adaboost approach, which is founded on the concept of machine learning and serves as its foundation. As a result, we are better able to provide specific estimates. Although Adaboost is a real-world gradient-boosting application, the weak classifier it uses is a sequentially trained decision tree. This is true, even though adaboost is a beneficial use of gradient boosting [15]. The order in which the data points were presented throughout the training served as the foundation for the learning that went into the decision tree's development. This is because the tree needs supervision to keep its shape over time. Adaboost is based on a strategy designed to accelerate learning by exposing a slow learner to material that is beyond their present comprehension barrier. This study demonstrates how a vector-valued fitness function may be used to learn a weak classifier by seeking for the antecedent Aj that maximises the fitness function. The purpose of this research was to demonstrate that a weak classifier might be taught using vector-valued fitness. We aim to demonstrate that the vector-valued fitness metric may be used to train a basic classifier with these activities. To complete the process, you will need to do relatively little work. To achieve this need, the definition of the operator "less than" () in a traditional GA must be altered because any fitness value can be compared to any other fitness value. As a result, each fitness value can be compared to any other fitness value. This is done so that all of the fitness values can be meaningfully compared. For selection purposes, we employed competitions (tournaments) rather than a system based on a variety of factors [16], and the generational system we used was usual. Let us go over the encoding of the fuzzy memberships again to avoid any confusion. Aside from the fitness function, there is only one other notable deviation from the standard in this GA. The only strange aspect of the entire operation is this one small aspect of how the GA was carried out.

These rules are encoded in an integer form in the language that underpins fuzzy rules, which corresponds to the encoding of their descriptive fuzzy rules. The language that underpins fuzzy rules is referred to as the "underlying language for fuzzy rules." The index expresses the position of the antecedent Aj within the set A. This is yet another way of conveying the antecedent's position. Using this index, you can find Aj's exact location in the set. The underlying concepts represented by the fuzzy partitions have been labelled using these n values. These labels have been assigned to the fuzzy data that has been collected. The phrases "low" and "high," for example, are used to categorise the range of values for a specific linguistic variable [17]. These terms are used interchangeably. During human conversation, the membership degree of each variable remains constant at 1, while "ANY VALUE" serves as a wild card label that can be assigned to all variables. Its domain is the collection of the variable's individual labels. There is also a label that states "ANY VALUE," and that value can be used in place of any other value. The "low" and "high" ranges of the variable may overlap, and vice versa. The following example demonstrates one method for condensing the use of a wildcard in a fuzzy rule. For the sake of this example, let us imagine we are having problems with not one but two unique attributes at the same time (height and weight). To avoid confusion, the terms "light" and "heavy" should be used instead of "height" when discussing a person's

DOI: https://doi.org/10.54216/FPA.210121
Received: March 05, 2025 Revised: June 01, 2025 Accepted: July 09, 2025

physical make-up. The language's height and weight categories have been enlarged to accommodate both upper and lower bounds thanks to the addition of a "wild card" phrase. The phrases "low," "high," and "any value" are now interchangeable for height and weight.Following the completion of each cycle, a new decision tree is generated to help minimise the overall quantity of residuals. This is done to maximise efficiency. The adaboost approach can be used to develop a methodology that gives a residual approximation that is reasonably close to the genuine number [18]. This strategy takes advantage of the loss function's negative gradient value; figure 2 depicts the system's overall design. The letter T shows where an algorithm is on a decision tree right now, while the number F1 shows the first input, F0.

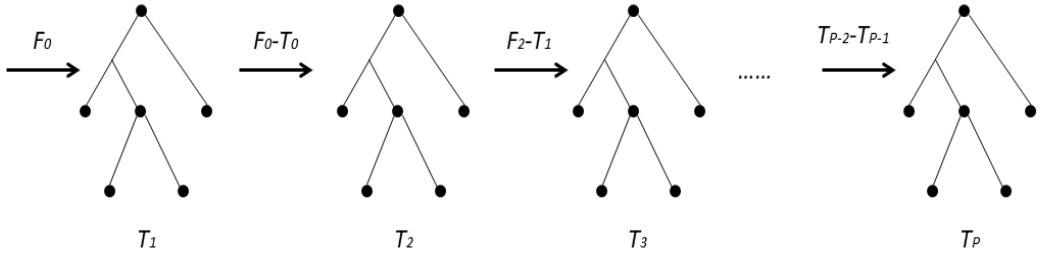


Figure 2. The core logic of the adaboost algorithm is depicted graphically

When ADABOOST is utilised, an ensemble of prediction models with varied degrees of accuracy is constructed. The ADABOOST can approximate the value of the residual.

$$iF(x) = aF0(x) + bF1(x) + cF2(x) + F.n \quad (1)$$

The value of the loss function, which incorporates both loss and error, is used to assess the model's soundness. This figure could be positive or negative. The term "loss function" will be addressed accurately in this context because you want your model to be as accurate as possible. Equation (1) shows the coefficients of the various components [19]. That is yet another scenario that could emerge. It achieves its purpose by synthesising the findings of various studies into one coherent whole and drawing conclusions from those findings. The Adaboost approach may address both regression and classification problems because it employs the sigmoid function, denoted as

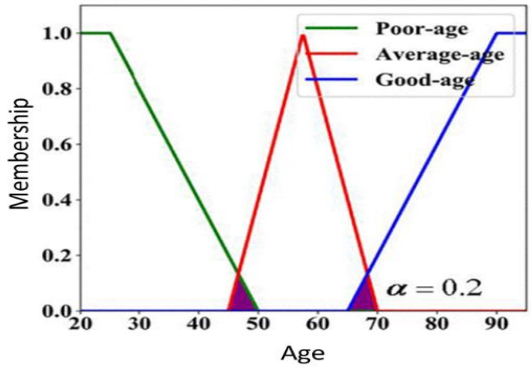
$$F(x)=aF0(x)+bF1(x)+cF2(x) \quad (2)$$

As a result, the adaboost technique is more versatile shown in equation 2. It is a well-known and widely acknowledged fact that the adaboost algorithm is one of the ones that performs the best at replicating real distributions among traditional techniques for machine learning. The several pieces of evidence offered give this proposition more weight because of their cumulative impact. The adaboost technique can analyse a wide variety of data sources and has the potential to be used to forecast cardiac disease. One of the potential applications for this approach is in genomics. These algorithms will be created by combining binary classification with other classification approaches [20]. The first step in this approach, which is designed to be as efficient as possible, will be to analyse the data we have on cardiovascular illness using a technique known as fuzzy logic. The fuzzy-adaboost technique was used to build a binary classification with the goal of enhancing accuracy. To minimise overfitting, the Bagging-Fuzzy-Adaboost technique combines bagging with fuzzy adaboost; as a result, it can predict the outcomes of a variety of classes with high accuracy.

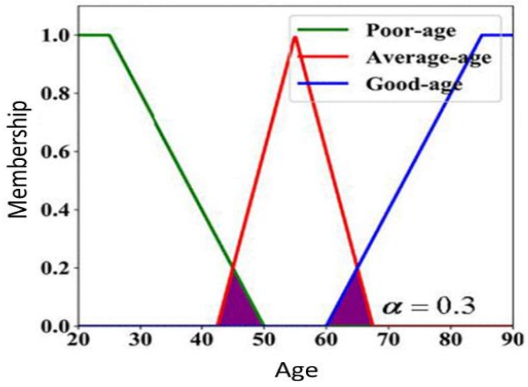
The results of both the diagnostic and prediction processes are likely to be influenced by the data's complexity. Even if the patient's diagnostic results are the same, there may be significant value changes for the same health data feature. Even if the outcomes remain the same, this is still a possibility. Given the inherently complicated nature of health data, it is reasonable to anticipate this outcome. Using fuzzy logic, the data is then simplified, the prediction model's accuracy is enhanced, and broader ranges of situations in which the results can be used are all accomplished. Each of these goals is met by reducing the number of variables in the study. The phrase "data fuzzification" refers to the application of a fuzzy set to precisely characterise data. This is what the phrase "fuzzy set" means. By comparing two separate sets of information, fuzzy sets, a type of hierarchical data item, can be used to characterise the extent to which two different collections of data have commonalities. Working with fuzzy sets and the membership function can greatly simplify the underlying problem [21]. The significance of the function performed by the membership degree cannot be overstated while attempting to solve the problem of fuzziness in the data. The level of mathematical theory knowledge and practical experience of the candidate are both considered when deciding whether they will be admitted to the group. The membership degree function can

DOI: https://doi.org/10.54216/FPA.210121
Received: March 05, 2025 Revised: June 01, 2025 Accepted: July 09, 2025

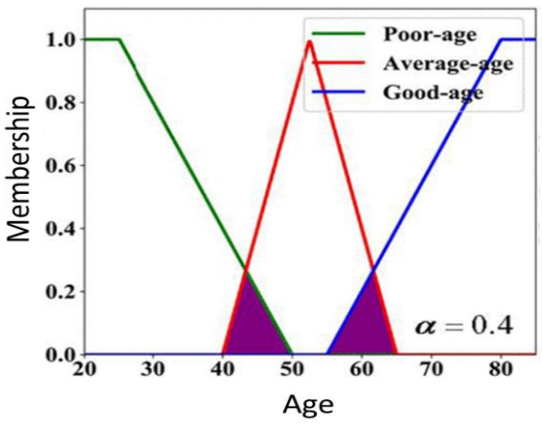
return values ranging from 0 to 1. These values range from 0 to 1. At the start of the operation, the triangle function is applied to the data to determine the degree of subordination already present in the information that will become fuzzy in the next steps. If we assume that a and b may be used to convey the range of possible values between that time period's minimum and maximum, we can use them to show the total number of participants [22-23]. Figure 3's arrangement is shown to affect the data fuzzification outcomes in a variety of ways to demonstrate this concept. The dependability of the data samples is the criterion used in the fuzzy logic-based encoding technique for the age property. This is critical since the age range that might be given is so broad, from 0 to 90 years old. When there are more occurrences of coincidental age subintervals within a population, it is deemed to have a higher value. The collection includes information on fourteen different heart disease risk factors. Each of these characteristics may demonstrate this level of ambiguity.



(a)



(b)



(c)

Figure 3. The age groups

DOI: https://doi.org/10.54216/FPA.210121
Received: March 05, 2025 Revised: June 01, 2025 Accepted: July 09, 2025

Figure 3. (a-d) Each variable in the indices characterises the property's three unique intervals, which are recorded. Each of the three membership functions (i1, i2, and i3) could be effectively mapped to the individuals who performed them. (a) alpha=0.2 (b) alpha=0.3 (c) alpha=0.4 (a) alpha=0.5.

The fuzzy-adaboost technique is the forerunner in a line of related binary classification approaches [24-27]. Consider the following limitations to be the ones you must meet: The parameter D denotes the size of the training set, the parameter T the maximum number of iterations, the parameter L the loss function, the parameter Ai the fuzzy attribute, and the parameter I the indicator function. The desired result was achieved: an intelligent and quick-witted fT (x) learner (x). Execute an initialization of f0(x) with the expression Bi=max(i1, i2, i3) for any t less than or equal to T. (I, and what each I achieves.) L(yi,f(xi))f(xi)=[1,N] by calculating the negative gradient yi for x, f(x) = ft1 is finally complete. The fuzzy-adaboost approach may aid in the development of improved generalisation abilities, the advertisement is too generalised and contains large amounts of medical data at once. "Bagging" is an abbreviation for "parallel integrated learning," a popular group study approach. The interference generated by a single sensitive patch is decreased due to the vast number of self-samplings engaged in the bagging process. Overfitting can be prevented, as can output variability.

The proposed Bagging-Fuzzy-ADABOOST approach may be capable of binary classification for predicting cardiovascular disease. This strategy standardises and simplifies the data, making it more accessible and usable. The creation of numerous distinct classifiers can be completed quickly using bagging. Even when the data being processed undergoes minor changes, the integrated Bag-Fuzzy-adaboost approach maintains a high level of adaptability. It is because of the fuzzy bagging. Here, we will go over the order in which these steps must be completed. Assume you have a size N training set and the alphabetic character T. This was done to make sure that the new set had the same number of pieces as the training set. The last step is to make a fuzzy ADABOOST classifier and then sharpen it with Ti using a very small part of the whole dataset. To create m distinct fuzzy-adaboost classifiers, step 3 must be repeated m times. The function f is represented by the sign ft (x). The following is one possible explanation for the phrase "tree f": (x). the regularisation term [28-29] determines the time required to conduct an adaboost calculation. (ft). The answer can be found by using the formula ft=N+12j=1N2j. Because there are an unlimited number of tree topologies, it is impossible to list them all here. As a result, we take a greedier strategy. The procedure begins with the creation of a "leaf," and then it is repeated to build new "branches" for the tree. When you look at how long it takes to sort samples from a single tree, you can see how difficult the process is. The difficulty of the challenge increases as you progress through the Adaboost tree. Using fuzzy logic's triangle membership function, the adaboost process can be streamlined, and the analysis of the data is kept intuitively simple. Several classification algorithms are built on the foundations of bagging and fuzzy adaboost. Unfortunately, a yes-or-no approach is the only way we can classify whether a patient has cardiac disease. We are unable to make a definitive diagnosis of a patient's cardiovascular disease due to a lack of necessary diagnostic tools. We divided heart problems into five groups based on the results. The numbers 0, 1, 2, and 4 represent the five types of heart disease in the training data, which helped us make a more accurate and quick diagnosis [30-31]. A score of 0 means that the person does not have heart disease, while a score of 4 means that the condition is severe. While implementing bagging-fuzzy-adaboost parallel computing requires more labour, the benefits are well worth it. Two aspects of the Bagging-Fuzzy-Adaboost multiclassification algorithms need our complete attention. For the time being, let us suppose that there are K classes and that M classifiers are employed consistently. There will be a total of 1,000 MK trees when the training is completed. The second point we want to make is that loops must be run in a specific order.

4. Analysis and Evaluation of Performance Objectives

This part's goal is to look at how well the obagging-fuzzy-AdaBoostost model works by comparing it to some well-known classification methods. This will help us make comparisons and draw conclusions.

4.1 Dataset

This study makes use of an open-source cardiac illness dataset [25] created by the University of California, Irvine. To compile the data, the VA Long Beach Health Care System database, as well as databases in Cleveland, Hungary, and Switzerland, were all searched. This collection contains 836 elements, 14 of which are considered significant indicators. Table 1 deconstructs the UCI dataset into its constituent parts and lists the qualities and ranges that were applied to each. We have provided this information to make things easier for you.

DOI: https://doi.org/10.54216/FPA.210121
Received: March 05, 2025 Revised: June 01, 2025 Accepted: July 09, 2025

Table 1: provides an overview of the data set

Attributes	Description	Ranges
Sex	Sex of subject	(0, 1)
Age	Age in year	(29, 77)
Cp	Chest pain	[l, 4]
Trestbps	Resting blood pressure	(94, 200)
Chol	Serum cholesterol	(126, 465)
Fbs	Fasting bloodi ng sugar	(0, 1]
Restecg	Resting electrocardiographic result	[O, 2)
Thalach	Maxi1num heart rate achieved	(71, 188)
Exang	Exercise induced angina	(0, 1)
Oldpeak	ST depression induced by exercise	(0, 6)
Slope	Slope of peak exercise ST segment	(1, 3)
Ca	No. of major vessels colored by fluoroscopy	(5)
Thal	Defect type	(4)
Num	Types of heart Disease	(1, 4)

Because of the usage of binary classification as well as numerous classification criteria, the UCI dataset now includes a higher total number of features. The patient's age and gender are both recorded in the medical record because they are required components of their individual identification. To establish an accurate diagnosis or prognosis of heart disease, you must be familiar with the 12 qualities listed below, all of which are reflected in critical clinical data. To make an appropriate diagnosis, you must be able to read clinical data [26]. Some examples of these traits are provided below. Before any type of training can begin, some fundamental data analysis on cardiovascular disease must be completed. The data must be transformed and thoroughly normalised, with any missing value information filled in, outliers, and zeros removed. Researchers are focusing their efforts mostly on databases in Switzerland and at the Long Beach Veterans Affairs Medical Center to complete the missing value procedure. The median values from the other databases are placed into the database columns that lack their own data.

Receiver operating characteristic (ROC) curves, such as those in [27], graphically illustrate the true-positive rate as contrasted to the false-positive rate. Examining the ROC is one method for determining how effective the receiver is. This example demonstrates the trade-off that must be made between the classifier's precision in identifying true positives and the number of false positives it produces. The following sentence exemplifies the trade-off: The ROC curve's area is comparable to the space filled by a coordinate axis [28]. The following is a detailed explanation of the technique used to determine precision. The degree of precision can be calculated using the following equation, which is shown below: The outcome of adding the "true score," "true number," "false negative “and” false positive" numbers is known as "accuracy." To be more specific, "true positives" are people who were expected to have heart disease but do not, and "true negatives" are people who were expected not to have heart disease but do.

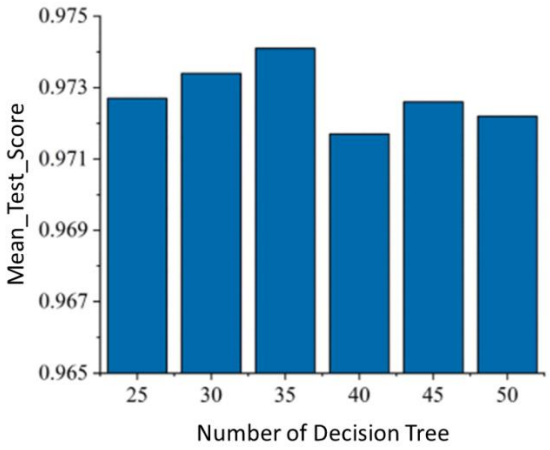
The "number of healthy individuals" refers to the percentage of the total population that is free of cardiovascular disease. The false positive (FP) rate is the percentage of people who were incorrectly diagnosed with an illness when, in fact, they did not have the ailment. The "true negative" (TN) population is defined as those who are entirely healthy and do not have any serious health problems, such as high blood pressure. When we discuss the "real negative," we are referring to this group of people. Tennessee's people are held up as a model of perfection. The F1 algorithm generates quantifiable results and provides an in-depth study of both accuracy and recall. The next step is to choose the optimal parameters for the Bagg-Fuzzy-Adaboost method. It is important to locate these six components in order to successfully implement the Bagging-Fuzzy-Adaboost technique. Consider the following scenario as a specific example of one of these: The predictability and accuracy of the model are both dependent on the values chosen for these parameters [29]. As a result, overcoming this barrier necessitates overcoming a significant problem, which is determining which settings are the best. When there are more decision trees available for use, it is possible to make predictions from higher-quality training sets. Another explanation for overfitting is if the M value was calculated incorrectly. We have concluded that option M is the best course of action based on the facts. A decision tree cannot have more levels than the maximum allowed by the MD threshold. It is not conceivable for MD to be too high or too low. If MD is too large, the approach will fail since training each tree individually will take too long. Because well-fitting residuals cannot be obtained from a single tree when MD is too low, the established approach will be inaccurate if the MD value is too low. To achieve an accurate separation of an internal node, a minimum number of MS samples is required. The market price of MS will either rise or fall

DOI: https://doi.org/10.54216/FPA.210121
Received: March 05, 2025 Revised: June 01, 2025 Accepted: July 09, 2025

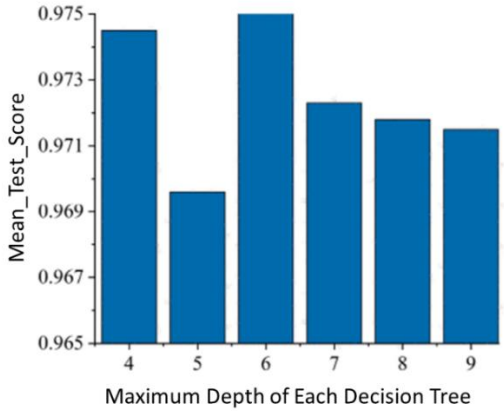
depending on two possible situations that could occur. If the number is a whole number, the least feasible total is the amount that can be calculated. The MS variable is a fractional variable that can be combined with floats. This means that samples that were previously considered redundant can now be removed from each node.

The ML indicator indicates the minimal amount of sampling required at a leaf node. At least two of a node's offspring must contain training data for it to be considered for machine learning. As a result, the model will fit the data more accurately, especially when it comes to regression. A five-pointed star represents the total number of samples that have been bagged. Choosing an adequate number for m could result in more accurate projections than any other method could. The rate at which one's education progresses Regularization procedures, such as the learning rate I (0,1), are used in the modelling process to achieve the goal of regulating the weight that is delivered to a single decision tree. As a result, not only is the model more efficient, but it is also much more robust in several critical areas [30]. Grid search uses cross-validation as a methodology to discover the best possible set of parameters; the scores obtained from the cross-validation procedure define the quality of the search. We base our judgment on the overall score received on the relevant test when determining which combination of settings will most effectively give the intended outcomes.

To get things started, we will focus on M, which is the total number of decision trees, and leave all the other parameters at the default levels. In the second place, MD is seen as a parameter that must be optimised via a grid search. However, M will remain at 35, and the remaining numbers will serve as the standard. Figure 4 depicts the average test results for each decision tree and shows a wide range of test depths. (b). When a student reaches MD Level 6, they have completed the exam to the best of their ability. When the difference between the median and mode exceeds 6, the mean test score starts to fall. As a result, the number 6 provides an appropriate depiction of the proper response. A grid search, like the prior method, is used to optimise the remaining parameters, making the two methods interchangeable. According to our findings, the approximate numbers MS = 20, ML = 3, and m = 20 produced from this equation are all fair approximations.



(a)



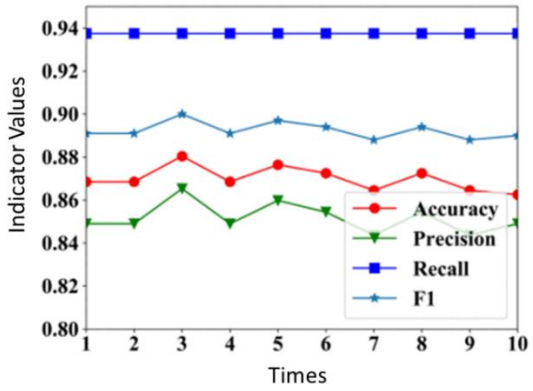
(b)

Figure 4. The Values of parameters

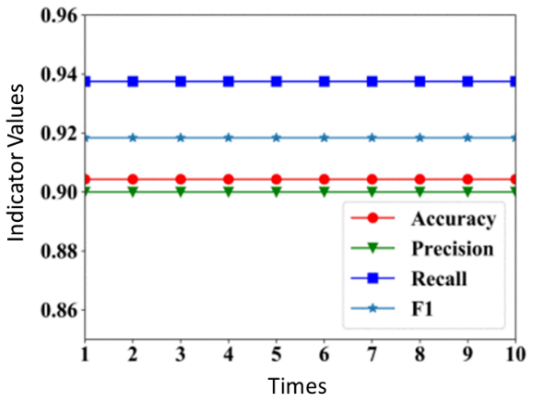
DOI: https://doi.org/10.54216/FPA.210121
Received: March 05, 2025 Revised: June 01, 2025 Accepted: July 09, 2025

Figure 4 shows a diagram for determining parameter values. (a) Graphs depicting the median outcomes of the tests administered at various thresholds for the various decision trees. (b) middle-of-the-distribution scores for various tree counts.

Figure 5 is provided to assist you in comparing the metrics before and after the grid search. The preliminary findings are represented by the first row of symbols in a grid search.



(a)



(b)

Figure 5. The recall rate

Figure 5 depicts letter (a). The graph clearly shows that, despite the consistent recall rate, accuracy and precision are continually vulnerable to change. However, no statistically significant changes in any of the quality metrics were observed throughout all ten tests depicted in Figure 5. (b). It demonstrates beyond any doubt that the optimization resulted in a significant boost in the model's level of stability.

Table 2: Compiles the Indicator Values from Various Models

Approach	AUC/%	Recall rate/%	Precision/%	F1-score
Decision Tree	97.78	90.59	31.51	46.76
Gradient Boosting	97.46	89.11	30.68	45.64
Fuzzy Gradient	85.38	73.99	43	54.39
Proposed Method	97.79	91.03	31.66	46.98

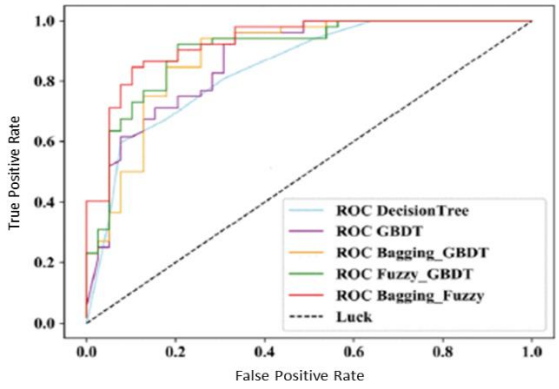
DOI: https://doi.org/10.54216/FPA.210121
Received: March 05, 2025 Revised: June 01, 2025 Accepted: July 09, 2025

Table 2 lists the five models used to forecast cardiac events as well as the metrics associated with each model. The table also displays the relationships between each model and the other measures. When compared to alternative methods, the Bagg-Fuzzy-Adaboost system is substantially more exact and accurate. To begin, in terms of forecasting cardiovascular illness, the adaboost outperforms the standard decision tree model [31]. These were compared to the Adaboost results. In comparison to the previous four models, which vary in degree, the Bagg-Fuzzy-Adaboost technique gives equivalent accuracy and precision. The fraction of incorrect predictions is used to estimate the importance of recall. To put it another way, accuracy has an impact on memory. In all cases, having a high recall value may be detrimental. It is possible to conduct a more fair and impartial evaluation of the prediction model by employing F1. When F1 is high, the prediction model's precision increases noticeably. The F1 score for the Bagging-Fuzzy-Adaboost method in Table II is the highest, implying that it is the most effective and trustworthy algorithm of all. Furthermore, the developed Bagging-Fuzzy-Adaboost technique is evaluated in comparison to other recent scientific works on the same subject, as shown in the comparative results in Table III.

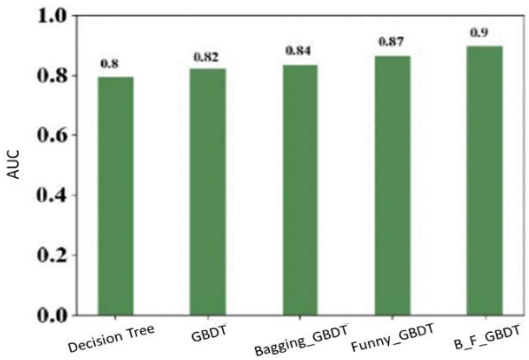
Table 3: The Results of Several Works Thought to Be State-of-the-Art in Their Field

Approach	AUC/%	Recall rate/%	Precision/%	F1-score
Decision Tree	97.62	88.94	33.28	48.44
Gradient Boosting	97.3	88.56	34.95	50.12
Fuzzy Gradient	88.38	74.12	35.12	47.66
Proposed Method	97.41	88.98	34.61	49.84

The ROC and AUC computation results for each of the three modalities are shown in Figures 6(a) and 6(b). This demonstrates that the model is even more effective than its competitors are.



(a)



(b)

Figure 6. The ROC and AUC computation results for each of the three modalities

DOI: https://doi.org/10.54216/FPA.210121
Received: March 05, 2025 Revised: June 01, 2025 Accepted: July 09, 2025

Figure 6 depicts the findings of the investigations in terms of the ROC and the AUC. a) How well alternative models match the data when the complete set is considered. b) The amount of space required to support each model's curve

Investigating the relative efficacy of single versus multiple classifications, several diseases on the list of probable difficulties can now be eliminated. If we can finish ten of these studies, we will have a good idea of how accurate the various forms of forecasts are. Figure 7 depicts the statistics. Most forecasts in all categories have a high percentage of correct predictions. There are 135 available test sets, 55 of which are for type 1 testing and the remaining 41 for type 2 testing. Following a random split of the dataset, only five independent examples of type 4 data were chosen to be part of the validation set. This was done in order for the collection to be representative of the entire dataset. This is because type 4 can only store a certain quantity of data, which is why this is the case. This issue will have a direct impact on the precision of type 4 data.

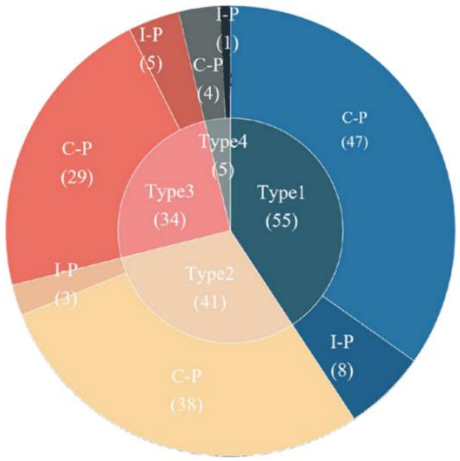
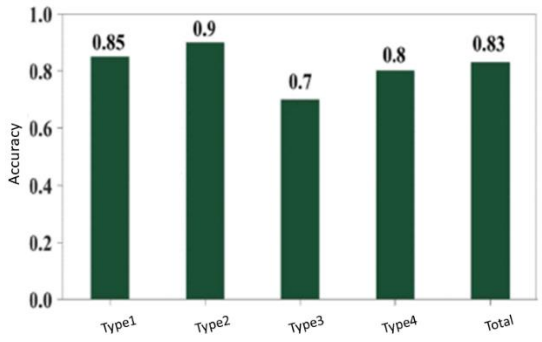


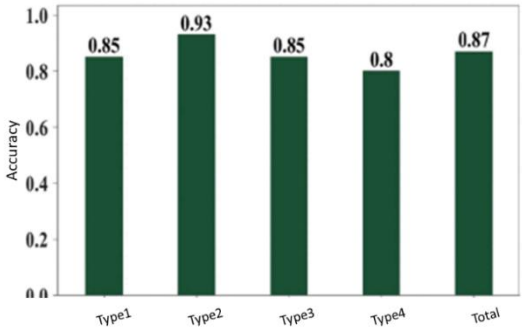
Figure 7. Values of the suggested indicators for the model.

Figure 7, which is shown below, is a picture of the values of the suggested indicators for the model. In this essay, we evaluate how well each of the four classifications of heart disease depicts the situation. Figure 8 examines and depicts the level of accuracy achieved by multiclassification prediction. This analysis is carried out for each category. The results of the trials show how well the Bagging-Fuzzy-Adaboost technique performs for multiclassification. Predictions are correct between 80% and 95% of the time, depending on the category. When it comes to predicting accuracy, Kinds 1 and 3 both have an 85% accuracy rate, which is the same as the accuracy rate of other types. Because the prediction performance of each kind is essentially the same, it is possible to build a diagnostic that is both faster and more accurate. However, different techniques are required for treating different patient groups. It makes no sense to analyse the validity of any individual category when dealing with multiclassification because each category is being examined. Do you believe it would be worthwhile to examine the overall correctness of the multiclassification model? Figure 7 depicts the results of 135 tests, each of which is divided into one of four groups. The results reveal that 118 of these tests provide valid predictions.

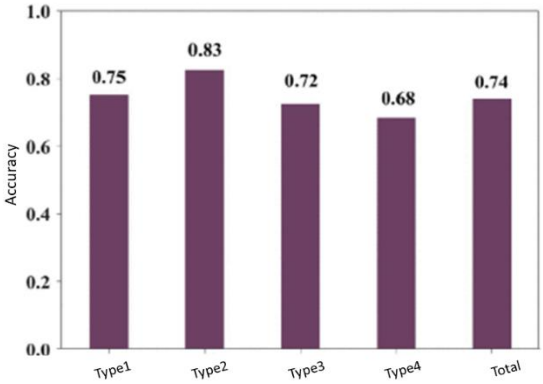


(a)

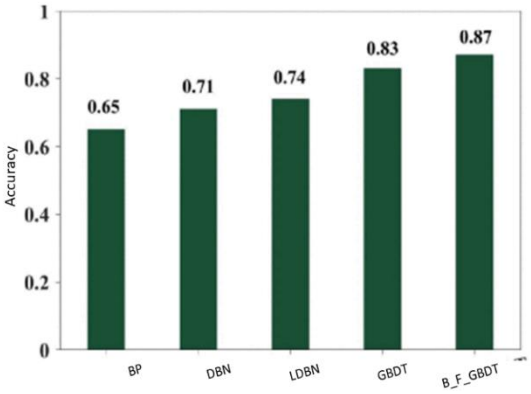
DOI: https://doi.org/10.54216/FPA.210121
Received: March 05, 2025 Revised: June 01, 2025 Accepted: July 09, 2025



(b)



(c)



(d)

Figure 8. depicts (a-d) an incomplete picture of the medical condition known as heart disease.

Figure 8 depicts the results of a multiclassification prediction of cardiac problems made using a basic AdaBoost algorithm. When compared to the precision of a type 3 measurement, which is measured at 0.7, the accuracy of a type 1 measurement is measured at 0.9. Furthermore, we investigate how well long-deep belief networks (LDBN) can categorise the various types of cardiac sickness and how accurate their forecasts are. In terms of total prediction accuracy, Figure 8(c) demonstrates that LDBN performs significantly worse than adaboost and Bagging-Fuzzy-adaboost. The fact that LDBN has a substantially lower overall prediction accuracy demonstrates this. This is a comment backed up by data. Deep belief networks and back propagation neural networks are two similar prediction approaches that are being evaluated for their capacity to forecast cardiac disease. The purpose of these studies is to determine whether these methods can accurately forecast the progression of cardiac disease. Given the preceding, it is appropriate to recast the problems of illness prediction and heart disease severity as multiple classifications and binary classifications, respectively. Because the difficulties pertain to many classes, both recastings are possible. Both recasts are founded on the notion that the problem may be recast as a binary

DOI: https://doi.org/10.54216/FPA.210121
Received: March 05, 2025 Revised: June 01, 2025 Accepted: July 09, 2025

classification problem. This is the central idea underlying each of these recasts. The bagging fuzzy Adaboost technique has the potential to achieve high levels of accuracy and stability when applied to binary and multiclassification situations. There are many ways to achieve this goal, one of which is by using it.

5. Conclusion

As a part of this study, we plan to make a bagging-fuzzy adaboost-based IoMT-based system for detecting and predicting heart diseases. This was one of the recommendations we made for potential future expansion. Throughout the investigation, it became clear that this technique has both dependability and precision. The Bagg-Fuzzy-Adaboost approach has been shown to produce accurate binary and multiclass classification predictions for cardiovascular disease. The AdaBoost approach was improved to include fuzzy logic and a bagging mechanism to make data analysis more manageable and prevent it from being overly precise. This was also done to avoid the approach becoming overly accurate. This was done to ensure that the technique delivered the best possible results. It has applications in electronic healthcare, particularly in providing patients with more diagnoses that are accurate and tracking their current health status. To improve the model and assess its usefulness in light of data sets held in both the public and private domains of ownership, we hope to engage with nearby hospitals in the future. This will be done to: In addition to this technique, we will try to improve the model's performance to the best degree possible.

Funding: the Deanship of Scientific Research, Vice Presidency supported this work for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia [Grant No.**KFU252503**

References

[1] J. Henriques et al., "Prediction of heart failure decompensation events by trend analysis of telemonitoring data," *IEEE J. Biomed. Health Informat,* vol. 19, no. 5, pp. 1757-1769, Sep. 2014.

[2] G. Yang et al., "Homecare robotic systems for Healthcare 4.0: Visions and enabling technologies," *IEEE J. Biomed. Health Informat,* vol. 24, no. 9, pp. 2535-2549, Sep. 2020.

[3] C. Li et al., "Healthchain: Secure EMRs management and trading in distributed healthcare service system," *IEEE Internet Things J.*, vol. 8, no. 9, pp. 7192-7202, May 2021.

[4] H. Li, K. Ota, and M. Dong, "Learning IoT in edge: Deep learning for the Internet of Things with edge computing," *IEEE Netw.*, vol. 32, no. 1, pp. 96-101, Jan. 2018.

[5] Y. Cheng, H. Zhu, J. Wu, and X. Shao, "Machine health monitoring using adaptive kernel spectral clustering and deep long short-term memory recurrent neural networks," *IEEE Trans. Ind. Informat.*, vol. 15, no. 2, pp. 987-997, Aug. 2018.

[6] P. Sun et al., "Modeling and clustering attacker activities in IoT through machine learning techniques," *Inf. Sci.*, vol. 479, pp. 456–471, 2019.

[7] R. Tao et al., "Magnetocardiography based ischemic heart disease detection and localization using machine learning methods," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 6, pp. 1658-1667, Jun. 2019.

[8] J. Li et al., "Heart disease identification method using machine learning classification in e-healthcare," *IEEE Access*, vol. 8, pp. 107562-107582, 2020.

[9] Z. Wen et al., "Exploiting GPUs for efficient gradient boosting decision tree training," *IEEE Trans. Parallel Distrib. Syst.*, vol. 31, no. 12, pp. 2706-2717, Dec. 2019.

[10] L. Zhao et al., "InPrivate digging: Enabling tree-based distributed data mining with differential privacy," *Proc. IEEE Int. Conf. Comput. Commun*, pp. 2087-2095, 2018.

[11] B. Zhang et al., "Health data driven on continuous blood pressure prediction based on gradient boosting decision tree algorithm," *IEEE Access*, vol. 7, pp. 32423-32433, 2019.

[12] P. Chotwani, A. Tiwari, V. Deep, and P. Sharma, "Heart disease prediction system using CART-C," *Proc. Int. Conf. Comput. Commun. Informat*, pp. 1-5, 2018.

[13] J. Thomas and R. T. Princy, "Human heart disease prediction system using data mining techniques," *Proc. Int. Conf. Circuit Power Comput. Technol.*, pp. 1-5, 2016.

[14] R. Khosla, S. Fahmy, Y. C. Hu, and J. Neville, "Predicting prefix availability in the internet," *Proc. IEEE Int. Conf. Comput. Commun*, pp. 1-5, 2010.

[15] M. A. Jabbar, B. L. Deekshatulu, and P. Chandra, "Heart disease prediction system using associative classification and genetic algorithm," *Comput. Sci.*, vol. 1, pp. 183-192, Dec. 2012.

DOI: https://doi.org/10.54216/FPA.210121
Received: March 05, 2025 Revised: June 01, 2025 Accepted: July 09, 2025

[16] V. A. S. Hernández et al., "A practical tutorial for decision tree induction: Evaluation measures for candidate splits and opportunities," *ACM Comput. Surv*, vol. 54, no. 1, pp. 1-38, Jan. 2021.

[17] D. Wu, C. T. Lin, J. Huang, and Z. Zeng, "On the functional equivalence of TSK fuzzy systems to neural networks mixture of experts CART and stacking ensemble regression," *IEEE Trans. Fuzzy Syst.*, vol. 28, no. 10, pp. 2570-2580, Oct. 2020.

[18] J. Soni, U. Ansari, D. Sharma, and S. Soni, "Predictive data mining for medical diagnosis: An overview of heart disease prediction," *Int. J. Comput. Appl.*, vol. 17, no. 8, pp. 43-48, Mar. 2011.

[19] S. Pouriyeh et al., "A comprehensive investigation and comparison of machine learning techniques in the domain of heart disease," *Proc. IEEE Symp. Comput. Commun*, pp. 204-207, 2017.

[20] S. Mohan, C. Thirumalai, and G. Srivastava, "Effective heart disease prediction using hybrid machine learning techniques," *IEEE Access*, vol. 7, pp. 81542-81554, 2019.

[21] P. Ziemba, "NEAT F-PROMETHEE—A new fuzzy multiple criteria decision making method based on the adjustment of mapping trapezoidal fuzzy numbers," *Expert Syst. Appl.*, vol. 110, pp. 363-380, Nov. 2018.

[22] A. V. S. Kumar, "Diagnosis of heart disease using fuzzy resolution mechanism," *J. Artif. Intell*, vol. 5, no. 1, pp. 47-55, Jan. 2012.

[23] O. W. Samuel et al., "An integrated decision support system based on ANN and fuzzy-AHP for heart failure risk prediction," *Expert Syst. Appl.*, vol. 68, pp. 163-172, Feb. 2017.

[24] T. S. Polonsky et al., "Coronary artery calcium score and risk classification for coronary heart disease prediction," *JAMA*, vol. 303, no. 16, pp. 1610-1616, Apr. 2010.

[25] UCI Machine Learning Repository-Heart Disease Data Set, [online] Available: http://archive.ics.uci.edu/ml/datasets/HeartDisease.

[26] X. Yuan et al., "A high accuracy integrated bagging-fuzzy-ADABOOST prediction algorithm for heart disease diagnosis," *Proc. IEEE/CIC Int. Conf. Commun. China*, pp. 467-471, 2021.

[27] I. D. Mienye, Y. Sun, and Z. Wang, "Improved sparse autoencoder based artificial neural network approach for prediction of heart disease," *Informat. Med. Unlocked*, vol. 18, pp. 100307-100311, Mar. 2020.

[28] R. Nair, S. Vishwakarma, M. Soni, T. Patel, and S. Joshi, "Detection of COVID-19 cases through X-ray images using hybrid deep neural network," *World Journal of Engineering*, vol. 19, no. 1, pp. 33–39, 2021.

[29] T. A. Nguyen, H. H. Nguyen, and H. T. Nguyen, "A novel secure data sharing scheme for IoT healthcare systems," *IEEE Internet Things J.*, vol. 9, no. 1, pp. 345-356, Jan. 2022.

[30] M. Ali, A. A. Alhassan, and S. A. Mohammed, "A survey on smart health systems: Applications, challenges, and opportunities," *IEEE Access*, vol. 10, pp. 12345-12367, 2022.

[31] R. Kashyap, R. Nair, S. M. Gangadharan, M. Botto-Tobar, S. Farooq, and A. Rizwan, "Glaucoma detection and classification using improved U-Net Deep Learning Model," *Healthcare*, vol. 10, no. 12, p. 2497, 2022.

DOI: https://doi.org/10.54216/FPA.210121
Received: March 05, 2025 Revised: June 01, 2025 Accepted: July 09, 2025