A Stable AI-Based Binary and Multiple Class Heart Disease Prediction Model for IoMT

Xiaoming Yuan ⬚, Jiahui Chen, Kuan Zhang ⬚, Yuan Wu ⬚, *Senior Member, IEEE,*
and Tingting Yang ⬚, *Member, IEEE*

Abstract—**Heart disease seriously threatens human life due to high morbidity and mortality. Accurate prediction and diagnosis become more critical for early prevention, detection, and treatment. The Internet of Medical Things and artificial intelligence support healthcare services in heart disease monitoring, prediction, and diagnosis. However, most prediction models only predict whether people are sick, and rarely further determine the severity of the disease. In this article, we propose a machine learning based prediction model to achieve binary and multiple classification heart disease prediction simultaneously. We first design a Fuzzy-GBDT algorithm combining fuzzy logic and gradient boosting decision tree (GBDT) to reduce data complexity and increase the generalization of binary classification prediction. Then, we integrate Fuzzy-GBDT with bagging to avoid overfitting. The Bagging-Fuzzy-GBDT for multiclassification prediction further classify the severity of heart disease. Evaluation results demonstrate the Bagging-Fuzzy-GBDT has excellent accuracy and stability in both binary and multiple classification predictions.**

Index Terms—**Fuzzy logic, gradient boosting decision tree (GBDT), heart disease predication and diagnosis, Internet of Medical Things (IoMT), machine learning.**

Xiaoming Yuan and Jiahui Chen are with Qinhuangdao Branch Campus, Northeastern University, Qinhuangdao 066004, China (e-mail: yuanxiaoming@neuq.edu.cn; 798283635@163.com).

Kuan Zhang is with the Department of Electrical and Computer Engineering, University of Nebraska–Lincoln, Omaha, NE 68182 USA (e-mail: kuan.zhang@unl.edu).

Yuan Wu is with the State Key Laboratory of Internet of Things for Smart City and the Department of Computer and Information Science, University of Macau, Taipa, Macau SAR, China (e-mail: ywucism@gmail.com).

Tingting Yang is with Navigation College, Dalian Maritime University, Dalian 116026, China, and also with Pengcheng Laboratory, Shenzhen 518055, China (e-mail: yangtingting820523@163.com).

I. INTRODUCTION

HEART disease with high morbidity and mortality has been viewed as one of the most complicated and life deadliest human diseases [1]. It seriously affects people's quality of life and brings huge economic losses in monitoring and treatment. Artificial intelligence (AI) could be used for early prediction, detection, and diagnosis of health problems [2]. It help patients obtain effective health guidance, intervention, and treatment from doctors and could alleviate the serious influence of heart disease. AI learning methods could effectively process mass Internet of Medical Things (IoMT) data and could give real-time heart disease prediction and diagnosis results in high efficiency and accuracy for e-healthcare system [3]–[6]. Moreover, it significantly reduces the economic and administrative burdens of intelligent chronic diseases prevention, monitoring, and treatment systems. However, critical problems on how to guarantee the high accuracy, generalization, stability of prediction algorithms, and models based on machine learning are the bottleneck to be solved.

Many researchers have paid attention to improving the accuracy of heart disease prediction models. Machine learning classifiers, including logistic regression (LR), K-nearest neighbors (K-NN), artificial neural network (ANN), support vector machine (SVM), naive Bayes (NB), XGBoost, and decision tree (DT), are used in the designing of the classification system [7], [8]. Particularly, gradient boosting decision tree (GBDT) is more efficient and applicable as a heart disease classification prediction algorithm [9], [10]. It can process various types of data, including continuous and discrete values [11].

However, most current prediction models or algorithms only solve the binary classification problem of heart disease prediction without considering the risk level of heart disease. According to the angiographic results, the severity of heart disease is classified into five categories, i.e., from zero (no presence) to four [12]. Thomas *et al.* [13] predicted the risk of each person. Nevertheless, there is still room for improvement in the accuracy of the original multiclassification algorithm for classifying different risk levels of heart disease. In machine learning, improving the accuracy of algorithms is usually achieved by reducing variance and deviation. The learning model [14] obtains higher area under the curve (AUC) and reduces variance by adding bootstrap aggregating (bagging) algorithm. Nevertheless, previous studies on heart disease prediction have high data complexity and the accuracy of each type of heart disease prediction is not given.

In this article, we propose a stable and high accuracy prediction approach to achieve binary classification and multiple classification of heart disease prediction at the same time. On the one hand, fuzzy logic reduces the model deviation and increases the model generalization by reducing the complexity of data. So, we introduce the fuzzy logic to GBDT algorithm to improve the accuracy. On the other hand, the bagging algorithm reduces the variance of the model through multiple random sampling. Thus, we further introduce the bagging algorithm to improve the stability of the model. In particular, our proposed algorithm can not only predict whether people are sick but also predict the severity of heart disease compared with traditional algorithms. Our contributions can be summarized as follows.

1) We propose a Fuzzy-GBDT-based binary-classification prediction algorithm for heart disease diagnosis, which reduces the complexity of heart disease data and increases the generalization ability of GBDT.
2) We integrate the bootstrap aggregating algorithm with Fuzzy-GBDT to avert the overfitting problem. The modified Bagging-Fuzzy-GBDT improves the prediction accuracy and stability by reducing the variance and deviation of the prediction model.
3) We develop the Bagging-Fuzzy-GBDT for multiple classification prediction to classify what severity of the disease is, which makes the diagnosis and treatment more accurate for different disease types.
4) Finally, we employ the grid search method to search the best parameters combination of the Bagging-Fuzzy-GBDT algorithm and evaluate the performance of the proposed model. The evaluation results demonstrate the model is highly accurate and stable for heart disease prediction.

The rest of this article is organized as follows. Section II presents the related work. Section III introduces our proposed system architecture. Section IV describes the details of the proposed algorithm integrated Bagging-Fuzzy-GBDT prediction. Section V discusses and analyzes the simulation results. Finally, Section VI concludes this article.

II. RELATED WORK

Heart disease prediction has attracted much attention over the last decades. Many prediction algorithms have been proposed for accurate heart disease diagnosis [5]–[10]. Tao *et al.* [7] used XGBoost classifier and extracted all 164 features from averaged magnetocardiography recordings for heart disease detection. The heart disease data complexity is relatively high. The application of data mining methods to predict heart disease has certain limitations. It could be made simpler and better by reducing the set of attributes. Features selection algorithms could be used to reduce the number of input features in a classifier. Four features selection algorithms and six classification algorithms [8] were employed simultaneously for the heart disease diagnosis, which is too complicated. Jabbar *et al.* [15] presented an available associative classification based on genetic algorithm. According to the actual data size, using genetic algorithm in disease prediction could obtain the best attribute set from a large dataset.
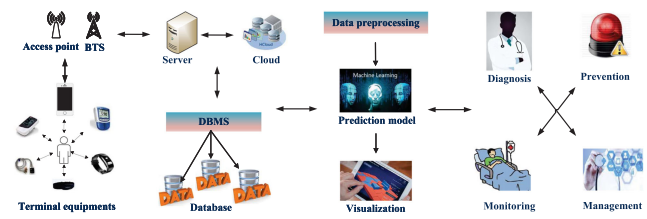

Fig. 1. IoT-based system architecture.

Decision tree is a celebrated classification algorithm learning technique. People have proposed many different decision tree algorithms for different applications, such as GBDT [11], C4.5 [16], and classification and regression tree [17]. Soni *et al.* [18] made a comparison among decision tree, neural network, and naive Bayes algorithms. The results showed decision tree performed best in the accuracy of heart disease prediction. But decision tree has difficulty to process continuous data. Hybrid learning procedures are used to diagnosis the heart disease [19], [20]. Seven different data mining algorithms [19] were introduced to predict heart disease, including SVM, NB, DT, KNN, RBF, multilayer perceptron, and single conjunctive rule learner. The comparison results indicated that SVM using the boosting technique is the best identifier and the recognition accuracy of retained samples is 90.5%. Fuzzy logic method is introduced to reduce the data complexity and improve the accuracy of prediction model [21]. ANN with fuzzy was employed to predict heart disease risks [22]. The global weights of the attributes based on risk contributions are calculated by fuzzy analytic hierarchy process [23]. But the overfitting problem in disease prediction should be noticed.

The binary classification results of heart disease prediction can only get the result if people have a disease. The risk level of heart disease is not given, which is ineffective for doctors in aided diagnosis and treatment. The prediction of multiple classification heart disease is also necessary. Polonsky *et al.* [24] evaluated the extent to which adding coronary artery calcium score to a model based on traditional risk factors. This correctly reclassified participants in the Multi-Ethnic Study of Atherosclerosis cohort in terms of risk of future heart disease events. The evaluation results of the model improved the classification of risk.

III. SYSTEM MODEL

Our IoMT and AI-based healthcare monitoring system architecture is shown in Fig. 1. The IoMT system consists of a few sensors, terminal devices, wireless links, and a local or cloud server. The sensors and devices attached to or around human body, such as ECG sensors or blood glucose meters, collect vital signs data of the human body. The collected data are transmitted over wireless links to edge or cloud server for real-time or historical data processing with the database management system. The health-related monitoring data could be accessed by doctors or authorized personnel at any time and any place. Moreover, health data can be further processed or trained for disease prediction, diagnosis, and management based on AI
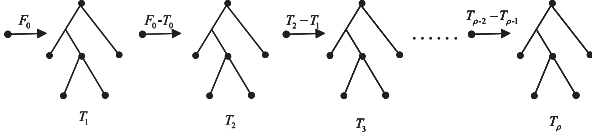
Fig. 2.   Schematic diagram of the GBDT algorithm.

Fig. 3.   Effect of different values α in membership.

technologies, such as machine learning. The prediction accuracy and stability of prediction model could be highly improved.

We employ the popular machine learning based GBDT algorithm to generate the prediction model. GBDT is a concrete example of gradient boosting, but its weak classifier is the decision tree, which are trained in sequence. It builds a new decision tree in each iteration in the direction of reducing residuals. GBDT algorithm uses the negative gradient value of the loss function to approximate the residual. The schematic diagram is shown in Fig. 2, where F_0 is the given initial input value and T is the traversed decision tree.

GBDT produces a prediction model in the form of an ensemble of weak prediction models. GBDT uses the value of the negative gradient of the loss function in the current model to simulate the approximate value of the residual in the regression problem. The loss function is an expression to measure loss and error, which reflects the credibility of the model. The smaller loss function is, the higher accuracy of model is. The expression is described as (1), where parameters $a, b, c, \ldots, \rho$ are the weight coefficients. It sums up the conclusions of all trees to make the conclusion. GBDT algorithm not only solves regression problems but also solves classification problems by adding the sigmoid function

$$F(x) = aF_0(x) + bF_1(x) + cF_2(x) + \cdots + \rho F_\rho(x). \quad (1)$$

IV. BINARY AND MULTIPLE CLASSIFICATION METHODS FOR HEART DISEASE PREDICTION

GBDT is one of the best algorithms to fit real distribution in traditional machine learning algorithms. The GBDT algorithm can process various types of data and could be used for heart disease prediction. In this section, we implement binary classification and multiple classification algorithms for heart disease prediction simultaneously. At first, we use the fuzzy logic method to reduce the complexity of heart disease data. The Fuzzy-GBDT algorithm is proposed to realize binary classification for heart disease prediction. Moreover, bagging algorithm is also integrated to avoid overfitting problem and achieve multiple classification prediction by Bagging-Fuzzy-GBDT algorithm.

A. Fuzzy-GBDT: Fuzzy Logic Integrates GBDT Algorithm

1) Data Fuzzification: Prediction and diagnosis results are usually affected by data complexity. The diagnosis results of patients are sometimes the same, and even their health data have great value differences under the same attribute. That is owing to the high complexity of health-related data. Then, fuzzy logic is introduced to reduce the data complexity and improve
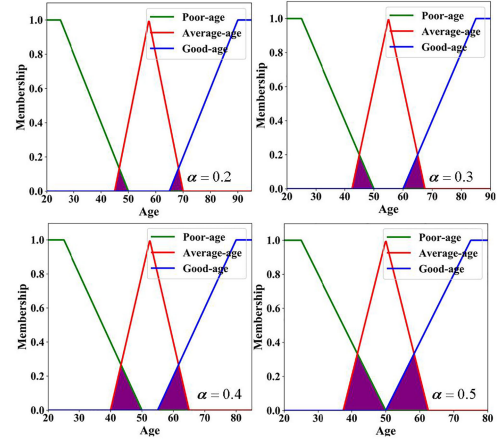
the accuracy of prediction model, as well as the generalization ability.

Data fuzzification is the progress that we employ a fuzzy set to describe the exact data. A fuzzy set is a hierarchical data object that describes the degree to which a data object belongs to an attribute. The complexity of a fuzzy set is greatly reduced by the membership function [21]. The key issue of data fuzzification is the function of membership degree. The membership function is based on the combination of actual life experience and mathematical theory knowledge. The degree of membership function is between 0 and 1.

In this article, we process the vital clinical records for diagnosis and learning the severity of heart disease. At first, we employ the triangular function to define the subordinated degree to complete data fuzzification process. Given μ is the membership degree, a and b are the minimum and maximum value of the data interval, respectively. If μ is converted into three membership degrees (μ_1, μ_2, μ_3), we define $\sigma = \frac{b-a}{3-2\alpha}(0 < \alpha < 1)$, where α is the degree of coincidence of subintervals. Let $\theta = a + (1-\alpha)\sigma$, the corresponding equations are shown as follows:

$$\mu_1 = \begin{cases} 0, & x > a + \sigma \\ \frac{a+\sigma-x}{\sigma}, & a <= x <= a + \sigma \\ 1, & x < a \end{cases}$$

$$\mu_2 = \begin{cases} 0, & x < \theta \ \text{ or } \ x > a + (2-\alpha)\sigma \\ \frac{x-\theta}{\sigma/2}, & \theta \leq x \leq \theta + \sigma/2 \\ \frac{a+(2-\alpha)\sigma-x}{\sigma/2}, & \theta + \sigma/2 < x < \theta + \sigma \end{cases} \quad (2)$$

$$\mu_3 = \begin{cases} 0, & x < a + (2-2\alpha)\sigma \\ \frac{x-\sigma}{\sigma}, & a + (2-2\alpha)\sigma \leq x \leq b \ . \\ 1, & x > b \end{cases}$$

The effect of different values of α on the results of data fuzzification is shown in Fig. 3. Considering attribute age ranges from 0 to 90, age is fuzzed according to the regularity of the data samples. The coincidence degree of age subintervals is positively correlated with the value of α. There are 14 attributes about the heart disease dataset. The detailed characteristics of data distribution are shown in Section V. Trestbps, chol, and thalach attributes are also fuzzed into three levels.

Algorithm 1: Fuzzy-GBDT Algorithm for Binary Classification.

Input: Training set D, the maximum number of iterations T, loss function L, fuzzy attribute A_i, indicator function I

Output: Strong learner $f_T(x)$

1:    Let $B_i = \max(\mu_{i1}, \mu_{i2}, \mu_{i3}) \rightarrow (\alpha_i, \beta_i, \gamma_i)$
2:    Let $A_i = B_i$
3:    Initialize $f_0(x)$
4:    **for** each $t \in [1, T]$ **do**
5:      **for** each $i \in [1, N]$ **do** (calculate negative gradient $\tilde{y}_i$)
6:        $\tilde{y}_i = -[\frac{\partial L(y_i, f(x_i))}{\partial f(x_i)}]_{f(x)=f_{t-1}(x)}$
7:      **end for**
8:      Fit a regression tree to the target c_{tj} giving terminal regions $R_{tj}, j = 1, 2, \ldots, J$
9:      **for** each $j \in [1, J]$ **do** (calculate the best fit c_{tj} for the leaf j)
10:       $c_{tj} = \underbrace{\arg\min}_{c} \sum_{x_i \in R_{tj}} \log(1 + \exp(-y_i(f_{t-1}(x_i) + c)))$
11:      **end for**
12:      Update $f_t(x) = f_{t-1}(x) + \sum_{j=1}^{J} c_{tj}I(x \in R_{tj})$
13:    **end for**

---

2) Fuzzy-GBDT Algorithm Description: The fuzzified data are combined with GBDT to be the Fuzzy-GBDT algorithm, as shown in Algorithm 1. Parameters α_i, β_i, and γ_i are the representative values of the three intervals of ith attribute A_i. They could be mapped to their membership functions $(\mu_{i1}, \mu_{i2}, \mu_{i3})$, respectively.

Given the training set $D=\{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$, we initialize weak Fuzzy-GBDT $f_0(x)$ in step 3 as $f_0(x) = 0.5 * \log(\frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n}(1-y_i)})$. In step 6, we adopt the cross-entropy loss function as L, where $p_i = \frac{1}{1+e^{-f_t(x_i)}}$, so

$$L(y_i, f_t(x)) = -\{y_i \ln p_i + (1 - y_i) \ln(1 - p_i)\} \quad (3)$$

$$\tilde{y}_i = -\left[\frac{\partial L(y_i, f(x_i))}{\partial f(x_i)}\right]_{f(x)=f_{t-1}(x)}$$
$$= y_i - \frac{1}{1 + e^{(-f_{t-1}(x_i))}} \quad (4)$$

where J is the number of leaf nodes and the indicator function $I \in [0, 1]$. Finally, we can get the strong learner

$$f_T(x) = f_0(x) + \sum_{t=1}^{T} \sum_{j=1}^{J} c_{tj}I(x \in R_{tj}). \quad (5)$$

B. Bagging-Fuzzy-GBDT Algorithm for Binary Classification

The Fuzzy-GBDT algorithm could improve the generalization ability. However, GBDT is sensitive and could not process mass medical data efficiently. Bagging is a well-known group learning algorithm of parallel integrated learning. Through multiple

self-sampling, bagging eliminates the interference of a single sensitive point. It can effectively avoid overfitting problem and reduce the variance of classifier results. In order to improve the accuracy and stability of Fuzzy-GBDT prediction model and decrease the learning time, the parallel Bagging-Fuzzy-GBDT algorithm is proposed.

1) Bagging-Fuzzy-GBDT Algorithm: The proposed Bagging-Fuzzy-GBDT algorithm could realize the binary classification for heart disease prediction. It reduces the data complexity and outputs standard available data. The bagging method generates several parallel week classifiers. The integrated Bagging-Fuzzy-GBDT algorithm has high adaptability even in the situation that data have small fluctuation. The accuracy and stability for heart disease prediction and diagnosis are guaranteed. The detailed sequences are operated as follows.

1) Given a training set T, and the size of which is N.
2) The subsets $T_i(i = 1, 2, \ldots, m)$ is replaced with the training set, and the size is $n'(n' < N)$.
3) Generate Fuzzy-GBDT classifier, and train them with sampled subset T_i.
4) Repeat step 3) to generate m Fuzzy-GBDT classifiers.
5) The strong Bagging-Fuzzy-GBDT classifier is voted from m weak learners.

2) Algorithm Complexity: The definition of the tree $f(x)$ is defined as $f_t(x) = \varpi_{q(x)}, \quad \varpi \in R^N, \quad q: R^d \rightarrow \{1, 2, \ldots, N\}$, where N is the number of leaves, ϖ is the vector of scores on leaves, q is a function that assigns each data to the corresponding leaf node, and d is the data dimension. The time complexity of the GBDT algorithm [9] is determined by the regularization term $\Omega(f_t)$. It could be obtained by the following:

$$\Omega(f_t) = \gamma N + \frac{1}{2}\lambda \sum_{j=1}^{N} \varpi_j^2. \quad (6)$$

Generally speaking, we cannot enumerate all possible tree structures. So, we employ a greedy algorithm instead. It starts with a leaf and then iteratively adds branches to the tree. The time complexity of the sample sorting of a single tree is $O(n \log n)$. The complexity of GBDT is determined by the depth of the tree. The fuzzy logic reduces the complexity of data by the triangular membership function, which will not increase the algorithm of GBDT. The complexity of Fuzzy-GBDT is the same with GBDT. Bagging-Fuzzy-GBDT generate several weak Fuzzy-GBDT classifiers in parallel and voting the strongest classifier in binary classification. Therefore, the time complexity of the proposed Bagging-Fuzzy-GBDT algorithm for binary classification is $O(dn \log n)$.

C. Bagging-Fuzzy-GBDT Algorithm for Multiple Classification

Using binary classification to predict heart disease, we can only predict whether a patient is ill or not. However, we cannot predict which type of heart disease a patient is suffering from. In order to diagnose diseases more accurately and improve the efficiency of diagnosis, we classify the types of heart diseases

into five types according to the angiographies disease status values in the data of cardiac patients. The five heart disease types are represented in training data by 0, 1, 2, 3, and 4, respectively, of which only 0 represents no heart disease, and the larger the number, the more serious the disease is.

We improve the Bagging-Fuzzy-GBDT algorithm so that it can be used for not only binary classification but also multi-classification prediction of heart disease. A detailed description of the algorithm is shown in Algorithm 2. From Algorithm 2, we can know that M is the number of extraction times in bagging algorithm, and step 2 to step 5 are the data fuzzification progress and convert fuzzified data into formats that can be used by GBDT algorithm. In step 6, k (0,1,...,4) stands for the type of heart disease. $\tilde{y}_{ik}$ is negative gradient, and it is calculated by loss function, where loss function is: $L(\{y_k, F_k(x)\}_1^K) = -\sum_{k=1}^{K} y_k \ln p_k(x)$. In step 19, I stands for learning rate, which is a regularization method for controlling the contribution of a single decision tree. It can prevent overfitting of Bagging-Fuzzy-GBDT. Finally, the final output $F(x)$ is obtained by the voting method from M training results.

Comparing Algorithm 1 with Algorithm 2, we can see that the multiple classification algorithm is roughly the same as the process of the binary classification algorithm, but Algorithm 2 adds a for-loop about K, where K represents the category of multiple classifications. We parallelize split searching at the entire level rather than a single node. The benefit of Bagging-Fuzzy-GBDT parallel computing is enough to cover the overhead. For Bagging-Fuzzy-GBDT multiclassification algorithms, we need to pay attention to two points. The first point is to assume that there are K classes, each of which trains M classifiers. After training, there will be a total of $M * K$ trees. The second point is that the order of loops cannot be changed. K categories must fit the first tree before they begin to fit the second tree. It is not allowed to learn all categories of a tree. The time complexity of Bagging-Fuzzy-GBDT for multiclassification is changed as $O(dMKn \log n)$.

V. PERFORMANCE EVALUATION

In this section, we make the performance comparisons between our Bagging-Fuzzy-GBDT and some other classification models. The evaluation results demonstrate the validity of the proposed algorithm on binary and multiple prediction of heart disease.

A. Dataset

In this article, we employ the University of California, Irvine (UCI)'s open-source heart disease dataset [25]. The data are from four different medical databases, including Cleveland, Hungary, Switzerland, and VA Long Beach. This database contains 14 important attributes and has a total of 836 data. The detailed description of UCI dataset with attributes and their ranges used is shown in Table I.

The binary classification and multiple classification variables are introduced to the attributes of the UCI dataset. Sex and age are used to identify the personal information of the patient. The remaining 12 attributes are vital clinical records that are trained

---

Algorithm 2: Bagging-Fuzzy-GBDT Algorithm for Multiple XClassification.

Input: Training set $T=\{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$
Output: $F(x)$ by voting method
1:    Set number of sample M
2:    **for** each $m \in [1, M]$ **do**
3:      Fuzzy attribute Ai
4:      Let $B_i = \max(\omega_{i1}, \omega_{i2}, \omega_{i3}) \to (\alpha_i, \beta_i, \gamma_i)$
5:      Let $A_i = B_i$
6:      **for** each $k \in [1, K]$ **do**
7:        Initialize $f_{k0}(x)$
8:      **end for**
9:      **for** each $t \in [1, T]$ **do**
10:       Calculate p_x for each category
         $p_k(x) = \frac{e^{f_k(x)}}{\sum_{l=1}^{K} e^{f_l(x)}}, k \in [1, K]$
11:       **for** each $k \in [1, K]$ **do**
12:         **for** each $i \in [1, N]$ **do**
13:           Calculate negative gradient
           $\tilde{y}_{ik} = y_{ik} - p_k(\mathbf{x}_i)$
14:         **end for**
15:         Fit a regression tree to the target c_{jkt} giving terminal regions $R_{jkt}, j = 1, 2, \ldots, J$
16:         **for** each $j \in [1, J]$
          do $c_{jkt} = \frac{K-1}{K} \frac{\sum_{\mathbf{x}_i \in c_{jkt}} \tilde{y}_{ik}}{\sum_{\mathbf{x}_i \in c_{jkt}} |\tilde{y}_{ik}|(1-|\tilde{y}_{ik}|)}$
17:         **end for**
18:         Update
19:         $f_{kt}(x) = f_{k,t-1}(x) + \sum_{j=1}^{J} c_{jkt}I \quad (x \in c_{jkt})$
20:       **end for**
21:      **end for**
22:    **end for**

---

and learned for heart disease prediction and diagnosis. The heart disease data needs to be preprocessed before training. We need complete data normalization, data conversion, filling in missing values, and removing outliers and null values.

The filling of missing values is mainly for databases Switzerland and VA Long Beach. The missing fields are filled with the average values of other complete databases. In this article, the dataset is divided into training set and test set according to the ratio of 7 to 3.

B. Evaluation Metrics

Five metrics are used to evaluate the performance of the proposed model, including the receiver operation characteristic (ROC) curve, AUC, accuracy, precision, and F1. ROC is a receiver operating characteristic curve, in which false positive ratio is horizontal axes, and true positive ratio is vertical axes [26]. It describes the tradeoff between the classifier's ability to identify true positive cases and the number of false positive cases misclassified. AUC is the area under the ROC curve and the encirclement of the coordinate axis. Specially, accuracy is the ratio of correctly classified samples to the total number of given

TABLE I
DATA SET DESCRIPTION

Attributes	Description	Ranges
Sex	Sex of subject	[0, 1]
Age	Age in year	[29, 77]
Cp	Chest pain	[1, 4]
Trestbps	Resting blood pressure	[94, 200]
Chol	Serum cholesterol	[126, 465]
Fbs	Fasting blooding sugar	[0, 1]
Restecg	Resting electrocardiographic result	[0, 2]
Thalach	Maximum 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]

samples. The calculation of accuracy is described as follows:

$$\text{Accuracy} = \frac{TS + TN}{TS + TN + FS + FN} \qquad (7)$$

where TS is the number of people who are predicted to have heart disease and actually sick, true positive; FS refers to the number of people who are predicted to be sick but not sick, false positive; FN refers to the number of people who are predicted to have no heart disease but are actually sick, false negative; TN is the number of people who do not have heart disease and are actually not sick, true negative.

Precision is defined as the ratio of correctly detected sick samples to all detected sick samples, and it shows the specificity. Recall is defined as the ratio of correctly detected sick samples to all actual sick samples, and it shows the sensitivity. F1 comprehensively describes the precision and recall and obtains quantitative results. The equations are shown as follows:

$$\text{Precision} = \frac{TS}{TS + FS} \qquad (8)$$

$$\text{Recall} = \frac{TS}{TS + FN} \qquad (9)$$

$$\text{F1} = 2 * \frac{\text{Precision} + \text{Recall}}{\text{Precision} * \text{Recall}}. \qquad (10)$$

C. Parameter Determination of Bagging-Fuzzy-GBDT Algorithm

In Bagging-Fuzzy-GBDT algorithm, there are six parameters [11] that need to be determined. They are shown as following. The values of these parameters affect the stability and accuracy of the proposed prediction model. Therefore, how to find the optimal values is a key problem.

1) The number of decision trees is M. In the growth process of Bagging-Fuzzy-GBDT, each iteration generates a decision tree. The prediction accuracy of training sets is improved with the increasing number of decision trees. However, improper M value may lead to overfitting problem. M is selected by testing.
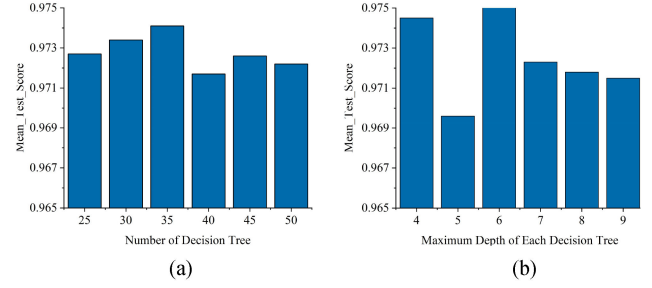

Fig. 4. Parameter determination diagram. (a) *mean_test_score* values of different numbers of trees. (b) *mean_test_score* values of different maximum depth of each decision tree.

2) The maximum depth of each decision tree is MD. The value of MD cannot be too large or too small. If MD is too large, the training time of each tree is extended, and the algorithm is too long. If MD is too small, a single tree cannot generate residuals to fit well, so the generated algorithm is ill-fitting.

3) The minimum number of samples required to split an internal node is MS. The value of MS is determined by two situations. If the value is int, MS is the minimum number. If the value is float, then MS is a fraction. The minimum number of samples changes to (MS * MD) for each node.

4) The minimum number of samples required to be at a leaf node is ML. A splitting point of any depth will only be considered if at least ML training samples are left in each of its left and right branches. This makes the model fit smooth, especially in regression.

5) The number of sampling for bagging is m. m represents the number of samples that need to be replaced from the original set to get m subdatasets. Appropriate m value is helpful to get more accurate prediction results.

6) Learning rate I. Learning rate $I(0 < I < 1)$ is a regularization method used to control the contribution of a single decision tree in the model, which can prevent overfitting of Bagging-Fuzzy-GBDT.

In the Bagging-Fuzzy-GBDT algorithm, grid search is employed to find the best parameters, which not only improves the performance of the model, but also greatly optimizes the stability of the model. Grid search uses the method of cross-validation to exhaustively search for the best parameter combination of cross-validation scores. *mean_test_score* represents the highest score obtained by the testing set in the process of finding the optimal parameters. So, we use *mean_test_score* as the evaluation criteria for finding the optimal parameters.

First, in order to determine the number of decision trees M, we let the other parameters unchanged by default, and then use the grid search to find the optimal value of M, as shown in Fig. 4(a). From the graph, we can see that when the number of decision trees M takes 35, and the *mean_test_score* reaches its maximum value. Second, MD is considered as a variable to be optimized by grid search, where M equals 35 fixed and the remaining value is the default value. The *mean_test_score* of various maximum depth of each decision tree is shown in Fig. 4(b).
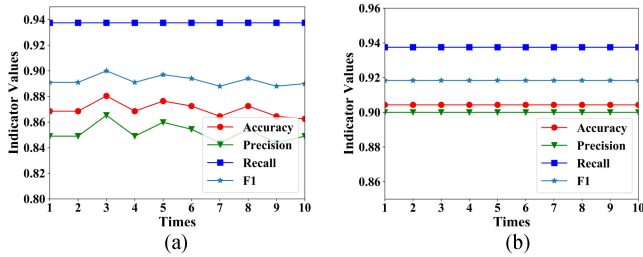
Fig. 5. Metric values before and after grid search. (a) Values before grid search. (b) Values after grid search.

TABLE II
THE INDICATOR VALUES OF DIFFERENT MODELS

Models	Accuracy	Precision	Recall	F1
Decision Tree	0.82	0.79	0.93	0.85
GBDT	0.84	0.80	0.97	0.88
Bagging-GBDT	0.86	0.81	0.98	0.89
Fuzzy-GBDT	0.88	0.86	0.93	0.90
Bagging-Fuzzy-GBDT	0.90	0.90	0.94	0.92

TABLE III
PERFORMANCE COMPARISON RESULTS OF SEVERAL STATE-OF-THE-ART WORKS

Authors	Accuracy	Precision	Recall	F1
Haq et al. [18]	0.89	-	0.77	-
Mohan et al. [19]	0.8847	0.875	0.928	0.90
Mienye et al. [26]	0.90	0.89	0.91	0.90
Proposed Bagging-Fuzzy-GBDT	0.90	0.90	0.94	0.92

Fig. 6. Experimental results of ROC and AUC. (a) ROC of different models. (b) AUC of different models.

When MD reaches 6, the $mean\_test\_score$ reaches the highest point. When the MD value is greater than 6, $mean\_test\_score$ value decreases continuously. So, 6 is the most appropriate value. In addition, through analogy, the remaining parameters are also optimized through grid search. The results are as follows: MS is 20, ML is 3, I is 0.1, and m is 20.

Fig. 5(a) is a change chart of evaluation indicators for ten experiments without grid search. From the figure, we can see that the accuracy and precision are constantly changing, and only the recall rate is stable. However, there is no change in accuracy, precision, recall, or F1 of ten experiments in Fig. 5(b). It fully demonstrates that the stability of the model has been greatly improved after the optimization.

D. Performance Evaluation of the Algorithm

1) Performance Evaluation of Binary Classification: We constructed five prediction models to evaluate the rationality and effectiveness of the prediction models. The performance of Bagging-Fuzzy-GBDT is compared with decision tree, GBDT, Bagging-GBDT, Fuzzy-GBDT, and models on accuracy, precision, recall, and F1. The proposed Bagging-Fuzzy-GBDT model achieves the best prediction results after determining the parameters through grid search, we list the corresponding metrics value of five models for predicting heart disease in Table II.

The Bagging-Fuzzy-GBDT approach has a significant advantage in accuracy and precision. First, GBDT has better heart disease prediction performance than traditional decision tree prediction model. Second, from the comparison results of Bagging-GBDT and Fuzzy-GBDT with GBDT, we can see that the improvement of GBDT algorithm by using bagging algorithm and fuzzy logic is effective. Finally, the Bagging-Fuzzy-GBDT approach has similar performance in accuracy and precision, whereas the other four models have different levels. The

value of recall is determined by the inaccurate prediction values. That is, the values of precision affect the values of recall. The high value of recall is not always better. Therefore, we introduce F1 to evaluate the prediction model more fair and objective. The higher value of F1 is, the better performance of the prediction model is. The highest F1 value of Bagging-Fuzzy-GBDT in Table II proves Bagging-Fuzzy-GBDT has the best effectiveness and stability. Moreover, the proposed Bagging-Fuzzy-GBDT approach is compared with some recent scholarly works. The comparison results shown in Table III indicate that the proposed Bagging-Fuzzy-GBDT obtains better performance than those introduced in the literature.

The evaluation results of ROC and AUC of different modes are described in Fig. 6(a) and (b). The Bagging-Fuzzy-GBDT algorithm has the best performance, whereas the performances of decision tree and GBDT are weaker. Their AUC values corresponding to ROC are 0.8 and 0.82, respectively. The AUC of our proposed Bagging-Fuzzy-GBDT approach is 0.9 in Fig. 6(b), which is the highest of the five models. This further shows that its comprehensive performance of the model is the best. When the bagging algorithm and fuzzy logic are added, respectively, the AUC value tends to increase. This again shows that we apply these two methods to improve the correctness of the GBDT algorithm. Therefore, Bagging-Fuzzy-GBDT has high stability and accuracy in both theory and experiment for heart disease prediction.

2) Performance Evaluation of Multiple Classification: In order to further determine a certain type of disease, we make a multiclassification prediction of the type 1, type 2, type 3, and type 4 of the disease according to angiography. We calculate the average accuracy of each type of prediction through ten experiments. The specific description is shown in Fig. 7, where C-P stands for the correct number of predictions, and I-P for the
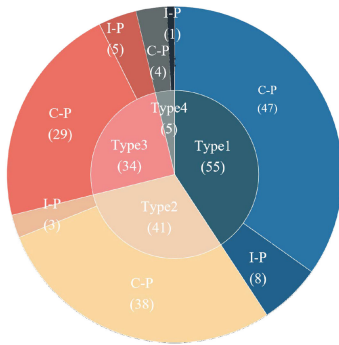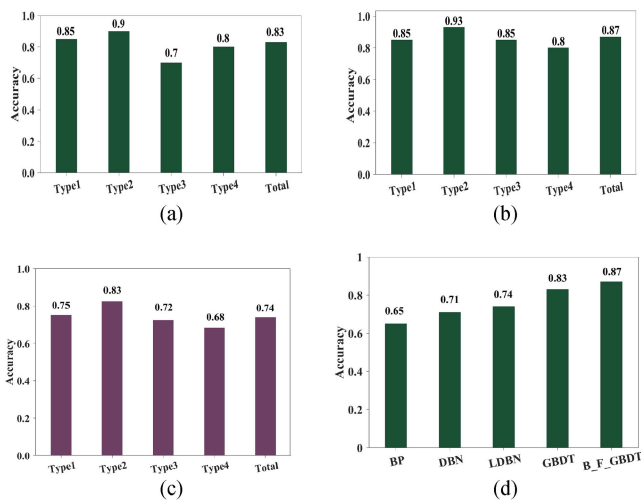
Fig. 7. Indicator values of proposed model.

Fig. 8. Accuracy of different types of heart disease. (a) GBDT. (b) Bagging-Fuzzy-GBDT. (c) LDBN. (d) Comparisons.

wrong number of predictions. The number of accurate predictions for each type accounts for the vast majority. There are 135 sets for testing, among which there are 55 sets for type 1 and 41 sets for type 2. However, due to the small amount of data in type 4, there are only five sets for type 4 in the testing set after randomly splitting the dataset. As a result, the accuracy of type 4 data is more sensitive.

We calculate the accuracy of the four categories of heart disease according to the single category. The accuracy is calculated by $MAccuracy = \frac{C\text{-}P}{C\text{-}P + I\text{-}P}$. The comparison results of the multiclassification prediction accuracy of each type are shown in Fig. 8(b). The predictive accuracy of each category is from 80 to 95% in the experiment, indicating that the Bagging-Fuzzy-GBDT algorithm has excellent performance on multiclassification implemented. Among them, type 2 has the best prediction effect with an average accuracy of 93%. The predictive performances of type 1 and type 3 are the same, reaching 85%. The predictive performance of each type is basically the same, which makes the diagnosis more accurately and quickly, and gives patients different treatments according to different types. However, it is unreasonable to compare the accuracy of each individual classification for multiclassification.

The overall accuracy should be evaluated, which will make more sense for evaluating the multiclassification model. There are 135 testing data of four types in Fig. 7, of which 118 are correctly predicted. We can calculate that the overall accuracy of Bagging-Fuzzy-GBDT for multiclassification prediction of heart disease is 87.4%.

Fig. 8(a) is the result of multiclassification prediction of heart disease based on a simple GBDT algorithm. The accuracy of type 1 is 0.9, whereas the accuracy of type 3 is only 0.7. Moreover, we also evaluate the prediction performance of long deep belief network (LDBN) for multiple classification of heart disease. The total prediction accuracy of LDBN is only 0.74 in Fig. 8(c), which is much lower than GBDT and Bagging-Fuzzy-GBDT. The similar prediction methods, such as deep belief network and back propagation neural network, are also valuated on heart disease prediction. Comparing the results in Fig. 8(d), the Bagging-Fuzzy-GBDT multiple classification model has higher prediction accuracy.

From the above, the illness prediction could transform into the binary classification problem, and the determination of heart disease severity could transform into the multiclassification problem. Moreover, the application of Bagging-fuzzy-GBDT algorithm in binary and multiclassification could achieve high accuracy and high stability.

VI. CONCLUSION

In this article, we proposed a stable and accurate Bagging-Fuzzy-GBDT approach for heart disease prediction and diagnosis in IoMT. The proposed Bagging-Fuzzy-GBDT approach achieved binary and multiple classification prediction of heart disease. We introduced the fuzzy logic and bagging algorithm into GBDT algorithm to reduce the complexity of data and avoid overfitting. Specifically, the stability of the model was greatly improved after the parameters were determined by the grid search. The evaluation results showed that the proposed model achieves good performance in terms of accuracy, stability, AUC, and other indicators compared with other traditional algorithms. The Bagging-Fuzzy-GBDT algorithm can not only predict the disease in high accuracy but also further determine the certain type of disease. It can be used in e-healthcare field to better diagnose and manage patients' health. In the future, we will optimize the proposed model and cooperate with local hospitals to further develop and evaluate the performance on both open and real data.

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," in *Proc. IEEE Int. Conf. Comput. Commun.*, Honolulu, HI, USA, 2018, pp. 2087–2095.

[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," in *Proc. Int. Conf. Comput. Commun. Informat.*, Coimbatore, India, 2018, pp. 1–5.

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

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

[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.

[16] V. A. S. Hernndez *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," in *Proc. IEEE Symp. Comput. Commun.*, Heraklion, Greece, 2017, pp. 204–207.

[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*. Accessed: Feb. 10, 2021. [Online]. Available: http://archive.ics.uci.edu/ml/datasets/Heart Disease

[26] X. Yuan *et al.*, "A high accuracy integrated bagging-fuzzy-GBDT prediction algorithm for heart disease diagnosis," in *Proc. IEEE/CIC Int. Conf. Commun. China*, Changchun, China, 2019, 467–471.

[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, no. 2020, pp. 100307–100311, Mar. 2020.

Xiaoming Yuan received the B.E. degree in electronics and information engineering from Henan Polytechnic University, Jiaozuo, China, in 2012, and the Ph.D. degree in communication and information system from Xidian University, Xi'an, China, in 2018.

She has been an Assistant Professor with Qinhuangdao Branch Campus, Northeastern University, Qinhuangdao, China, since 2018. Her research interests include cloud/edge computing, artificial intelligence, medium access control, and performance analysis for Internet of Vehicles and the Internet of Medical Things.

Jiahui Chen received the B.E. degree in Internet of Things engineering from the Lanzhou University of Lanzhou, China, in 2020. He is currently working toward the master's degree in computer science and technology with Qinhuangdao Branch Campus, Northeastern University, Shenyang, China.

His research interests include cloud/edge computing and Internet of Vehicles.

Kuan Zhang received the B.Sc. degree in communication engineering and the M.Sc. degree in computer applied technology from Northeastern University, Shenyang, China, in 2009 and 2011, respectively, and the Ph.D. degree in electrical and computer engineering from the University of Waterloo, Waterloo, ON, Canada, in 2016.

He has been an Assistant Professor with the Department of Electrical and Computer Engineering, University of Nebraska–Lincoln, Lincoln, NE, USA, since September 2017. He was also a Postdoctoral Fellow with the Broadband Communications Research Group, Department of Electrical and Computer Engineering, University of Waterloo, from 2016 to 2017. His research interests include security and privacy for mobile social networks, e-healthcare systems, cloud/edge computing, and cyber-physical systems.

Yuan Wu (Senior Member, IEEE) received the Ph.D. degree in electronic and computer engineering from The Hong Kong University of Science and Technology, Hong Kong, in 2010.

He is currently an Associate Professor with the State Key Laboratory of Internet of Things for Smart City, and also with the Department of Computer and Information Science, University of Macau, Macau. Prior to that, he was a Full Professor with the College of Information Engineering, Zhejiang University of Technology, Hangzhou, China. During 2016–2017, he was a Visiting Scholar with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada. His research interests include resource management for wireless networks, green communications and computing, mobile edge computing, and smart grids.

Tingting Yang (Member, IEEE) received the B.Sc. and Ph.D. degrees from Dalian Maritime University, Dalian, China, in electronic information engineering in 2004 and in communication and information system in 2010, respectively.

She is currently a Professor with Dalian Maritime University, and also with Pengcheng Laboratory, Shenzhen, China. Her research interests include space-air-ground-sea integrated networks, network AI, and edge intelligence.