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

HMLT: Hybrid Machine Learning Technique for Prediction Heart Disease based on Internet of Things

Miss. Rina S. Patil[1] Dr. Tripti Arjariya[2] Prof. (Dr.) Mohit Gangwar[3]

Submitted: 26/01/2023 **Revised:** 22/03/2023 **Accepted:** 16/04/2023

Abstract: It is very vital and alarming to be able to predict heart diseases before it manifest, as the number of people diagnosed with cardiac diseases is rising at an exponential rate day by day. Since determining this diagnosis is a challenging process, it is essential that it be carried out in an accurate and timely manner. The proposed study effort centres mostly on the issue of which patients are more prone to suffer from heart disease depending on a variety of different medical parameters. By using patient's past medical history, a method called the Hybrid Machine Learning Technique (HMLT) that is based on the Internet of Things (IOT) has been presented as an approach for determining whether or not a patient is likely to be identified with a cardiovascular disorders. The performance of the proposed HMLT is compared to that of a number of traditional methods, including Decision Tree (DT), Support Vector Machine (SVM), Naive Bayes (NB), X-GBOOST, Artificial Neural Network (ANN) and Random Forest (RF). The results of the experiments demonstrate that the effectiveness of the HMLT-based cardiac disease prediction system that has been developed is superior to that of other techniques. This conclusion was reached as a result of the findings of the studies. The newly developed methodology indicates that HMLT has a performance accuracy of 96%, which is superior to that of the traditional classification algorithms that are presently in use.

Keywords: *Heart Disease Prediction, Hybrid Machine Learning Technique, Internet of Things, Machine Learning.*

1. Introduction

According to the research that was conducted by the World Health Organization in the year 2016, cardiovascular diseases play the biggest role in the top 10 global leading causes of death, with ischemic heart disease contributing the most and stroke ranking as the second most prevalent cause of death. It is generally agreed upon that adequate health care services ought to be made available so that individuals can undergo routine examinations of their health. Cardiovascular disease is the leading cause of death across the globe, accounting for close to 31% of all deaths [14]. Because of a shortage of diagnostic centres, skilled medical professionals, and other factors that influence the precise prognosis of heart disease, early identification [15, 16, 17] and therapy of a number of different cardiac illnesses can be extremely difficult. This is especially true in developing nations. In light of this issue, current developments in computer technology and methodologies for machine learning have been used to the development of medical aid program as a support system for advance diagnosis of cardiac disease. If a heart-related ailment is diagnosed in its early stages, the risk of dying from it is significantly reduced. Several machine learning strategies are applied to medical data in order to comprehend the pattern of the data and derive a forecast from it. The data associated with healthcare are typically very extensive and difficult to understand. The ability of ML techniques to manage large amounts of data and mine it for relevant information is a significant advantage [18, 19]. The machine learning algorithms utilize previous data and base their predictions on current data. This kind of machine learning approach for predicting cardiac disease can motivate cardiologists to take speedier actions, which in turn saves a significant number of people's lives. This is because more individuals can acquire drugs within a shorter time span.

Machine learning is a subfield of artificial intelligence study [2] that has quickly become one of the most well-known aspects of data science. The techniques that are used in machine learning are meant to be able to carry out a wide variety of tasks, including prediction, categorization, decision - making process, and so on. ML models were tested and verified on a set of real-time test datasets that they had not seen before. After that, the overall accuracy of the anticipated outcome is validated by comparing the model's final efficiency with the actual value.

Therefore, finding a quick and accurate diagnosis for cardiovascular disease is a crucial concern. There have previously been a significant number of attempts made to forecast cardiac disease by using machine

[1]*Ph.D. Research Scholar, Department of Computer Science and Engineering, Bhabha University, Bhopal, Madhya Pradesh, India*
[2]*Head, Department of Computer Science and Engineering, Bhabha University, Bhopal, Madhya Pradesh, India*
[3]*Director (Alumni Cell), B. N. College of Engineering and Technology, Lucknow*
[3]*mohitgangwar@gmail.com*

learning algorithms [20, 21, 22, 23], but in the study that is being presented, IOT healthcare datasets are being compared to common ML algorithms in order to determine which ML method is the most accurate.

Since cardiovascular disease is the leading cause of death in adults, the proposed investigation is being carried out to determine, with the use of a patient's medical information, which detects individuals who are most likely to be diagnosed with a heart disease [6]. The proposed HMLT determines patients who are experiencing any heart disease-related symptoms such as chest discomfort or high blood pressure and therefore can help in diagnosing disorder with fewer diagnostic exams and effective care, so that they may be cured appropriately. For example, chest pain or high blood pressure are two examples of these symptoms.

The purpose of this study is to determine whether or not it is likely that the patient will be detected with any cardiac heart diseases due to their medical factors like gender, age, heart palpitations, fasting insulin levels, etc. The findings of this study will be used to inform future research. An Internet of Things healthcare dataset is utilised in order to acquire the information about the patient from the sensor. With the use of this dataset, it is possible to determine the extent to which the patient suffers from a heart condition. Whether or not a patient is likely to even have heart disease was determined by classifying the patient according to 13 medical characteristics. The following machine learning methods namely Decision Tree, Support Vector Machine, Naive Bayes, X-GBOOST, Artificial Neural Network and Random Forest are used to train the dataset attributes. The HMLT method is the most effective algorithm with an accuracy rate of 96 %, which is significantly higher than that of other classifiers. And finally, patients are classified according to whether or not they are at danger of developing a heart disease using this proposed procedure, which is also completely cost effective.

The structure of this paper is as organized in 5 sections. In Section 2, related work of heart disease prediction approach using machine learning techniques performed by researchers is discussed in detail. In section 3, proposed hybrid machine learning heart disease prediction approach is discussed. In section 4, experiments using various attributes of IOT dataset is discussed in detail. Finally, in section 5, conclusion of proposed research work is discussed.

2. Literature Survey

In the environment that we live in today, heart disease is a serious problem that needs to be addressed. It is getting more difficult because of random characteristics like blood pressure and cholesterol, for example. Because of these factors, the researchers have focused their attention on Data Mining or Machine Learning as the current methodologies for forecasting the outcomes of an illness. These are considered to be contemporary methods.

Data mining and machine learning are both helpful tools for a variety of problem-solving arrangements. Utilizing this method results in a reliable strategy based on the assessments of free components. Because healthcare is an important application for data mining, the management of the healthcare system must personally deal with an enormous number of information assets. Heart disease has indeed been identified as one of the leading causes of death in nations around the world [24, 25].

Among the many Data Mining strategies, a variety of methods have been developed specifically for the purpose of predicting cardiovascular diseases. There are some issues in medical care that can be resolved by employing strategies from the fields of machine learning and data mining. These strategies have a broad range, which is one of their most powerful advantages. In comparison to other algorithms that have been developed in the past, such as clustering and so on, machine learning approaches play a significantly more effective role in the prediction of heart disease. The existing literature available plays an important part in finding solutions to all of the issues, as well as in making decisions that are both reasonable and successful.

S. Ouyang et al. [1] provides assistance in the process of conducting cardiac disease prediction commencing from a variety of heart disease forms (coronary heart disease) as well as data sets, summarising the currently adopted ML detection and prediction approach, emphasising the features and distinctions of these methods, and analysing the difficulties and potential developments. The findings indicate that methods of machine learning have a wide range of uses in the study of cardiovascular disorders. Moreover, due to the non-uniform nature of medical data, every technique for machine learning has to be applied to a particular domain. The likelihood of developing heart disease is briefly discussed in the final section of the article.

K. S. K. Reddy et al. [2] intends to make a prediction about the Innovative Smart model for the prediction of Heart Problems utilizing Dynamic KNN. It compares this model to Support Vector Machine. Predictions of heart illness are made using dynamic KNN (N=92) as well as SVM (N=92), both of which are examples of machine learning approaches. The prediction of disease is made using a straightforward method called dynamic KNN. The dataset pertaining to heart disease is utilised for disease prediction. There are

20 samples obtained from each group, and then the total is split into training and testing datasets. According to the findings of the tests, the accuracy of the dynamic KNN is 84.44%, while the accuracy of the SVM is 67.21%. There is a critically important distinction to be made among Dynamic KNN as well as SVM in terms of analysis. Therefore, the conclusion that can be drawn is that dynamic KNN looks to perform substantially better than SVM when it comes to predicting novel heart disease.

T. Xue et al. [3] extract the data of individuals who have cardiac disease and have a physical index. Particle swarm optimization (PSO) is employed to optimise, and a classification prediction approach of cardiovascular disease utilising PSO, SVM is developed. This addresses the issue that the optimised parameters in the conventional SVM framework are hard to find. The findings of the experiments indicate that, in comparison to the conventional support vector machine model, the optimised model enhances the predictive performance by 1.33% and also reduces the time required for model training; both of these factors contribute to an improvement in the effectiveness with which heart disease can be diagnosed.

The goal of G. S. Reddy Thummala et al. [4] is to use the uneven forest approach to anticipate cardiovascular problems and to attain improved accuracy in their forecasts by utilizing machine learning techniques and assessing the effectiveness of those algorithms to that of KNN. Both the Decision Tree and the K-Nearest Neighbor algorithms are looked at closely in this study. A dataset of 1700 records was utilised in order to put the tactics to the test and evaluate their effectiveness. As aspect of the programming investigation, N=20 iterations on every technique are done. The objective of this study is to find distinct grades of model performance using a reliability of repeated measures that is 80%. After the completion of the tests, the mean effectiveness of the decision tree approach and the k-nearest neighbour method, correspondingly, for the disease forecasting was found to be 86.75 % and 82.55 %, respectively. It is possible, through the use of t-tests on independent samples, to demonstrate that there is a statistically significant difference between the two algorithms' levels of precision (p 0.05). The findings of the analysis show that the RF method operates considerably better than the KNN classifier.

D. K. Chohan et al. [5] attempted to predict heart disease using a variety of different algorithms, including LR, DT, SVM, NB, RF, and KNN, with the intention of establishing which technique is the most accurate. The DT came out on top, achieving an accuracy of 98.53% overall.

Machine learning algorithms as well as deep learning algorithms were utilised in the process of predicting cardiac disease by K. Vayadande et al. [6]. The dataset that was used came from Kaggle and contained 303 rows with 14 different attributes. LR, Naive Bayes, Artificial Neural Network, K-NN, Support vector machine, Multi-Layer Perceptrons, DT, RF, XG Boost, as well as Cat Boost are the algorithms that are utilised in the model.

X. Yuan et al. [7] offer a prediction model that is based on machine learning in order to accomplish binary and multiple categorization for the purpose of heart disease prediction. Firstly, a Fuzzy-GBDT technique is constructed by integrating fuzzy logic together gradient boosting decision tree. This is done with the goal of lowering the level of data complexity and raising the degree to which binary classification forecast is generalised. After that, overfitting is presented by combining bagging with the Fuzzy-Gradient Boost Decision Tree (Fuzzy-GBDT) technique. Additional categorization of the degree of heart disease is provided by the Bagging-Fuzzy-GBDT for predicting multiple categorization. According to the findings of the evaluation, the Bagging-Fuzzy-GBDT possesses superior accuracy as well as stability when making binary and multiple classification predictions.

Dual illness prediction methodology is a user interactive-based method that was proposed by D. Sharathchandra et al. [8]. The suggested technique for predicting cardiovascular disease and diabetes uses inputs from the final user along with data that is representative of the real world. In the work that has been provided, the forecasting of diseases is done with the help of the model of logistic regression and the Support vector machine framework. When it comes to accurately predicting cardiovascular disease and diabetes, the proposed method achieves a success rate of 85% and 78%, correspondingly.

Gnaneswari G [9] utilised a variety of different techniques for the early diagnosis of the disease, including SVC, KNN, RF and NB method. The new aspect of this research is the ability to forecast whether or not a patient who already has cardiovascular disease will experience a cardiac event in the future. On the other hand, the majority of researchers are exclusively concerned with attempting to forecast the presence of cardiac disease. The forecasting of heart attacks in people who already have cardiac disease is the primary subject of this article.

E. Maraj et al. [10] demonstrate a fuzzy logic model that can predict coronary heart diseases. This method was developed with seven input variables and one output variable, and it was tested on 30 subjects in Albania. This framework is constructed with the help of the fuzzy

logic toolbox, which is used in this investigation. Inputs into the fuzzy model that are taken into account include blood pressure, lipid, activity level, diabetes, age, body mass index, and smoke, while the disease categorization is determined by the output. It has been determined that fuzzy sets as well as membership functions are the best options. Defuzzification is accomplished through the use of the centroid approach.

Data mining techniques are utilised, as stated by V. Kannagi et al. [11], in order to construct the Heart Disease Forecasting Scheme. Several different kinds of data formats, such as photos, text, graphs, and figures, are made use of in different information management systems. In order to get an early diagnosis of cardiac disease, risk variables such as system problems were investigated. Using a new method known as Intelligent Learning Assisted Support Vector, it is able to rapidly detect heart disease by specifying several traits and risk factors. On the basis of these criteria, mining concepts are utilised to find factors that pose a high risk for coronary disease. The utilisation of data mining techniques will make it possible to provide predictions of sickness that are both quick and accurate.

Using a dataset that was provided by Kaggle, A. Basak et al. [12] applied a variety of machine learning approaches, including SVM, RF, NB, KNN. The findings presented by the proposed ML were among the most accurate, with a level of accuracy of 93%. Methods of model validation are also utilised in this study in order to develop the model that is going to be the most applicable to the circumstances that now exist.

G. Kumar Sahoo et al. [13] creates a model for individualised care to combat the risk of heart disease by employing a system that may be used at home. LR, NB, RF, DT, SVM, GBDT, KNN are the examples of machine learning models that are utilised in the process of predicting heart disease. The early and accurate diagnosis of cardiovascular illness is an essential component of quality medical care. It is crucial to recognise coronary heart disease at its early stages, seek a specialist physician before the disease's seriousness, and begin treatment with medication as soon as possible. The Cleveland Heart Disease database that is located in the UCI ML Repository was utilised in order to evaluate the effectiveness of the model that was suggested. The Random Forest method demonstrates superior performance accuracy, scoring 90.16 %, in comparison to all other machine learning methods. It's possible that the most effective model involves evaluating patients' fitness levels rather than their usual hospital visits. The work that is suggested will lighten the load on hospitals and make it easier for them to treat only the most urgent patients.

Alperen Erdogan et al. [30] Came to the conclusion that heart disease, which is the condition that results in the majority of patients' deaths, was one of the most serious ailments. The accurate diagnosis of heart conditions presents significant challenges for medical professionals. People who have various ailments that display common symptoms, such as chest tightness, a briefness of breath, tremors, and nausea, may incorrectly assume that they have a heart ailment even though cardiac disorders may be objectively diagnosed. Due to this, it is challenging for doctors and other medical experts to diagnose conditions related to the heart. In order to determine whether or not heart conditions do in fact exist, researchers turned to several methods of machine learning. It was determined how much of an impact each piece of information had on the total success rate, and then that data was analysed and weighted accordingly. Throughout the course of this investigation, a method for determining the weight coefficient has been proposed for consideration. The findings of the approach that was suggested resulted in a success rate of 86.90%, and 13 separate characteristics were obtained from the patients to use in the analysis.

3. Research Methodology

Both men and women face an equal risk of dying from heart disease as a leading cause of death. According to the available data, heart disease is responsible for around one third of all the deaths that occur throughout the entire world. A significant amount of research has been done in an effort to improve the diagnostic accuracy of heart disorders. When there is a large volume of data, those working in the healthcare industry have a difficult time diagnosing diseases. Data Mining and Machine Learning both take massive amounts of data and turn them into data that can be used to make accurate predictions and choose appropriate actions. Because of the complexities involved, the prediction of heart disease depends on detailed information in the majority of cases. As a result, there is an increased level of optimism among the researchers over the prognosis of diseases, particularly those pertaining to the heart.

This study presents the development of a framework for the prediction of heart illness. The framework, which makes use of 13 characteristics of the benchmark UCI heart disease dataset, is designed to assist medical practitioners in the early detection of heart disease. The procedure of the suggested approach is divided into a number of stages, each of which is discussed below.

Consider the following flowchart of proposed heart disease prediction model depicted in figure 1

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

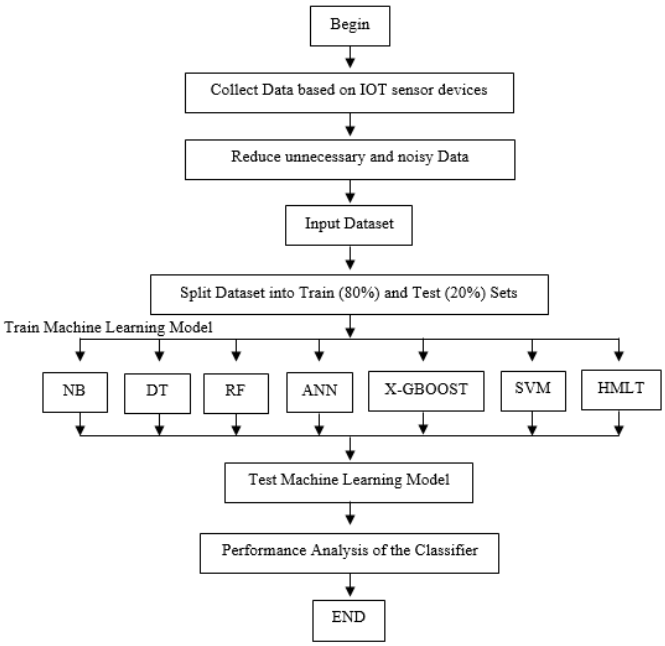

Figure 1: Flowchart of proposed framework for predicting heart disease

Collecting of Data

It is possible to acquire patient data from patients who are spread out over multiple locations using Internet of Things devices that can be worn or implanted. In this first stage, the data is gathered by the utilisation of IoT devices that are implanted into the human body, data gathered from benchmark datasets, and health records.

For the purpose of predicting heart disease, the data contained in the 304 records that belong up the standard UCI heart disease dataset is taken into consideration. The following table 1, describes the attributes of the dataset

Table 1: Description of Attributes of Dataset

1	Age (in years)	Continuous
2	Gender (male or female)	1- Male,0- Female
3	Trestbps (resting blood pressure)	Continuous value
4	Cp (type of chest pain)	Chest pain type, values are taken between 1 to 4 1- typical angina 2- atypical angina 3- non-anginal pain 4- asymptomatic
5	fbs (fasting blood sugar)	Measured in mg/dL <=100 - normal 110 to 125 pre- diabetes >=126 - diabetes
6	restecg (ECG result)	0-normal 1-having ST- T 2-hypertrophy
7	thal (Heart rate of patient)	It takes following values 3-normal 6- fixed defect 7-reversible defect
8	chol (serum cholesterol)	Continuous value
9	exang (exercised induced angina)	0 - no 1 - yes
10	thalach (max heartbeat rate)	Continuous value
11	Oldpeak (ST depression)	Continuous value
12	Ca (major vessels number colored by fluoroscopy)	Number of major vessels from 0-3
13	Slope (peak slope)	Takes values from 1 to 3

Data Pre-Processing:

After the acquisition of a variety of records, the data on heart disease are then pre-processed. The 300 patient records were there in the dataset, out of which 5 records contain missing values. This 5 records were eliminated from the dataset, and the pre-processing was performed on the remaining 257 patient records. Both multi-class variable and a binary classifier were used the features of the dataset. Evaluating for the presence or absence of heart disease involves the utilisation of the multi-class variable. If the patient is found to have heart disease, the value is assigned to 1, but if the patient is found not to have heart disease, the value is assigned to 0, indicating that the patient does not have heart disease. The transformation of medical records into diagnostic values is a crucial step in the process of pre-processing data. According to the findings of the initial processing of the data for 295 patient records, 130 of the records display a value of 1 which indicates the existence of heart disease, while the other 165 records indicate a value of 0 showing the absence of cardiovascular disease.

Feature Selection:

Out of the total of 13 features that make up the data set, the two attributes having to do with the patient's age and gender are the ones that are used to determine their personal details. The remaining 12 qualities are significant since they contain essential medical record. When it comes to diagnosis and determining the degree of cardiac disease, clinical records are absolutely necessary. In this research, many machine learning (ML) approaches, including DT, NB, SVM, ANN, RF, KNN, Xg-Boost, and HMLT, are implemented. The experiment was carried out using each of the ML approaches and 14 attributes.

Classification:

This research makes use of a wide variety of classification models derived from machine learning. The accuracy of each model is computed for the provided dataset, and then that result is compared to the HMLT. The description of each conventional classifier is discussed as follows:

- **Decision Tree:** In the field of ML, the Decision Tree method is utilised to create classification models [26]. The structure of a tree serves as the foundation for this hierarchical classification approach. This belongs to the class of supervised learning, which describes situations in which the desired final outcome is developed previously. The decision tree method may accommodate both category and numerical data sets. The root, branches, as well as leaf nodes make up the components of a decision tree. The data is assessed based on the path that it took to get from the root node to one of the leaf nodes. A number of 283 tuples were evaluated as part of the dataset's decision-making process using the decision tree. It was possible that a favourable or negative evaluation would be reached about the heart disease prediction. The correctness, selectivity, and sensitivities of the model were evaluated by comparing them to the specific parameters.
- **Random Forest:** The random forest model is made up of a number of independent decision trees that work together as a collection; more specifically,

every tree in the random forest has its own unique way of predicting classes [28]. By employing the bagging approach to the feature space, the Random Forest technique is able to generate a greater degree of randomness and variation. Rather than searching aggressively for the best predictors to generate branches, it randomly samples components of the predictor area, which results in the addition of more variety and a reduction in the variance of the trees, but at the expense of an equivalent or higher level of bias. The procedure, "feature bagging," is the powerful method that ultimately results in a model that is more robust.

- **Naive Bayes:** The statistical classifier known as Naive Bayes makes the assumption that there is no connection between the attributes being considered. It assumes an attribute value on a particular class but determines its independence based on how that value compares to the values of other attributes. This is known as conditional independence [28].
- **Support Vector Machine**: SVM is a sort of supervised learning that accomplishes a task that is comparable to that of the C4.5 method, with the exception that it does not make any use of decision trees [28]. The support vector machine makes an effort to reduce the likelihood of incorrect classification.
- **K nearest neighbor:** The KNN technique is a supervised technique which is used to classify data [28]. The samples' distance function as well as the results of the majority of the k-nearest neighbours are utilised as the basis for this approach, which is then employed to retrieve the knowledge. It examines the entirety of the dataset to locate the k examples that are spatially closest to the new one, and then it outputs the mode for a classification issue. In some situations, the KNN algorithm does not work very well, which means that it achieves a poor level of accuracy in comparison to the other algorithms.
- **Xg-Boost:** This approach makes use of a framework known as gradient boosting [27]. This technique uses the maximum gradient instead of the methods that use conventional gradients, so the error gets minimised with each iteration. This is how it outperforms the other ways. The Xg- Boost models achieve the best outcomes possible in the laboratory tests in every scenario.
- **Artificial neural network:** Artificial neural networks were primarily developed for computational purposes. The key idea behind this model is to complete a task in a shorter amount of time compared to the conventional model [29]. The biological pattern of neurons in the human brain can be compared to this model's representation of that structure. The same manner that neurons link to one another in the brain, here also neurons (nodes) will link to one another. This model is made up of a very large number of neurons that are all connected to one another and work together to carry out a certain function. A neural network with a single layer is referred to as a perceptron, as it only produces a single output.

Evaluation Indicators: Comparing the efficacy of the various algorithms, which is discussed as follows, requires the utilisation of four distinct classification performance evaluation metrics:

- **Accuracy:** It is defined as the percentage of samples that are correctly categorised in relation to the overall number of samples.
- **Precision:** It refers to the proportion of true positives among samples that have been projected to be positive by a model compared to the total number of samples that have been expected to be positive.
- **Recall:** It is the proportion of instances that actually are positive that are projected to be positive out of the total number of instances that actually are positive.
- **F1 score:** It is the harmonic mean of precision and recall rate, with 1 being the highest possible score and 0 being the lowest possible score.

4. Result and Discussion

Experiments are carried out in order to determine whether or not the patient suffers from heart disease. These experiments take into account several characteristics, including the patient's age, gender, chol, restecg, thal, trestbps, fps, cp, exang , slope, ca, oldpeak and thalach readings. In order to carry out this experiment, a dataset consisting of 300 records that are individually characterised by 14 distinct attributes is trained.

Figure 2 is a representation of a graph showing the effects of heart symptoms on patients who have different characteristics. If a person has a job that causes them to be under a great deal of stress and has a poor eating personality, it is almost certain that they will become compromised with heart syndrome within a small period of time, and he/she is one of the people who could suffer a serious coronary strike.

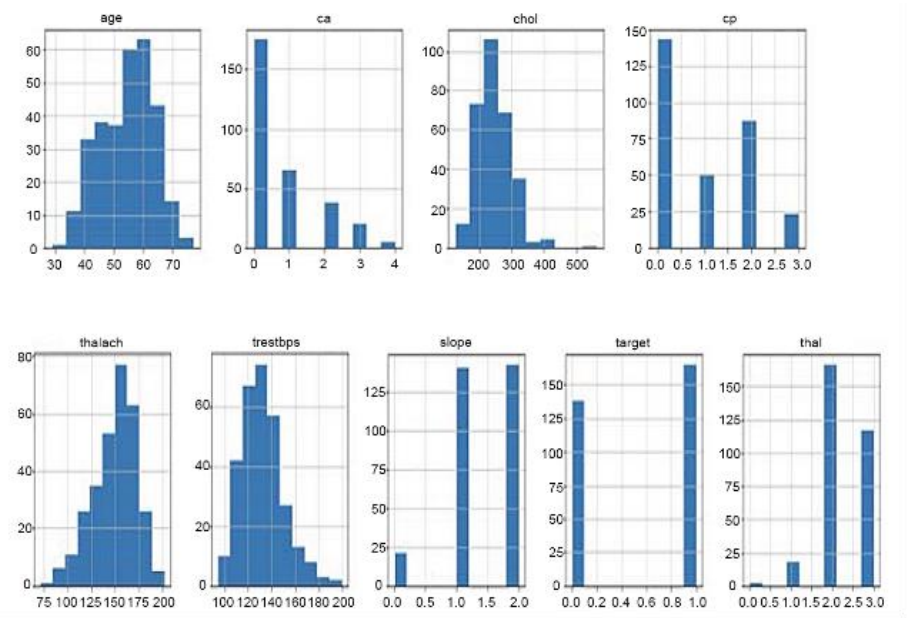

Fig. 2. Histogram of the distribution of target on features of dataset

On the same data set, seven different machine learning classification algorithms were tested, and the results of those tests were compared with those produced by the proposed HMLT as shown in table 2.

Table 2: Performance measurements of seven ML classifiers and proposed HMLT

Classifier	Accuracy	Precision	Recall	F-Measure
RF	0.92	0.94	0.97	0.98
SVM	0.96	0.97	0.98	0.98
X-GBOOST	0.95	0.96	0.97	0.98
NB	0.84	0.89	0.96	0.92
ANN	0.94	0.96	0.97	0.97
HMLT	0.96	0.98	0.99	0.99
DT	0.92	0.95	0.96	0.97

Consider the following figure 3, which depicts the performance measurements of various conventional classifiers and proposed HMLT.

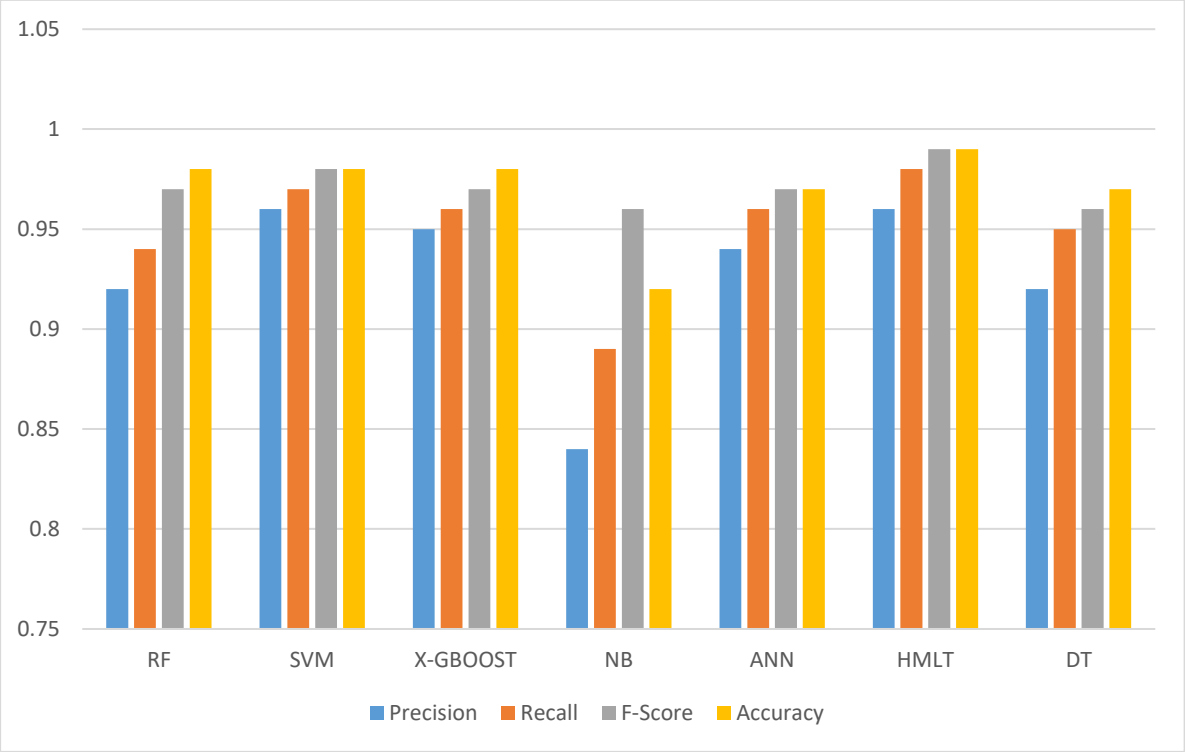

Fig. 3: Performance measurements of seven ML classifiers and proposed HMLT

It is observed from the above figure 3, proposed hybrid machine learning technique outperforms than other conventional technique.

The curve seen in Figure 4 represents the receiver operating characteristic (ROC) of the HMLT which is used to predict cardiac disease.

The correlation between specificity and sensitivity is reflected on the ROC curve. Specificity is measured along the abscissa of X-axis; when it is nearer it is to zero, accuracy will be enhanced. Sensitivity is measured along the Y-axis of the ordinate; when it is closer to one, the accuracy is enhanced. Figure 4 demonstrates that out of the seven algorithms, HMLT has the maximum prediction accuracy.

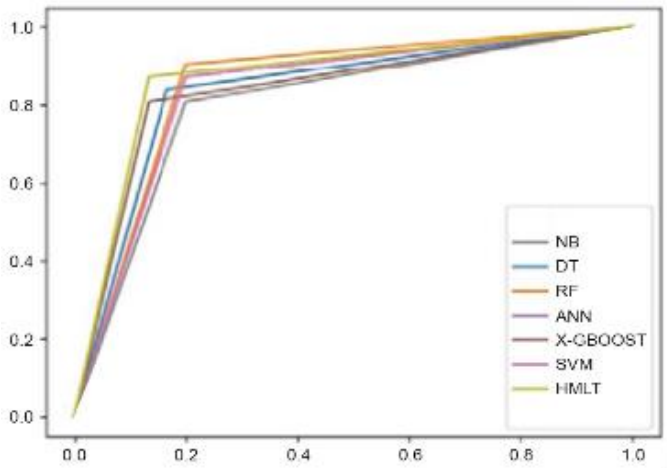

Fig. 4. ROC curve of seven algorithms in predicting heart disease

Thus, the data have been analyzed using seven different conventional classifiers in addition to the HML approach that was proposed. The performance of the proposed HMLT is compared to the effectiveness of 7 traditional classifiers through the utilisation of several evaluation metrics. It has been determined that the suggested HMLT has a classification accuracy of 96%, which is higher than that of the other classifications models.

5. Conclusion

In the long term, it is possible to save more lives of patients with heart disease and get early diagnosis of

heart irregularities with the assistance of the original patient records that are recognised by the proposed HMLT. If the condition can be treated and diagnosed at an earlier stage, it is possible to avoid the development of heart disease and bring down the overall death rate, both of which would be very promising information to have. When compared to other conventional ML algorithms, the outcomes obtained from this research effort using hybrid machine learning technology are superior. Proposed hybrid machine learning technique has obtained a classification accuracy of 96%, which is better than Decision Tree, Support Vector Machine, Naive Bayes, X-GBOOST, Artificial Neural Network and Random Forest. The resumption of this research could make use of additional combinations of various machine learning methodologies in the future to obtain better predictive performance. This indicates that there is a significant amount of room for development in the implementation of ensemble learning is an area of medicine. On the other hand, the effectiveness of prediction of heart disease can be improved by developing new techniques for choosing features to acquire a more comprehensive understanding of important characteristics.

References

[1] S. Ouyang, "Research of Heart Disease Prediction Based on Machine Learning," 2022 5th International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE), Wuhan, China, 2022, pp. 315-319, doi: 10.1109/AEMCSE55572.2022.00071.

[2] K. S. K. Reddy and K. V. Kanimozhi, "Novel Intelligent Model for Heart Disease Prediction using Dynamic KNN (DKNN) with improved accuracy over SVM," 2022 International Conference on Business Analytics for Technology and Security (ICBATS), Dubai, United Arab Emirates, 2022, pp. 1-5, doi: 10.1109/ICBATS54253.2022.9758996.

[3] T. Xue and Z. Jieru, "Application of Support Vector Machine Based on Particle Swarm Optimization in Classification and Prediction of Heart Disease," 2022 7th International Conference on Intelligent Computing and Signal Processing (ICSP), Xi'an, China, 2022, pp. 857-860, doi: 10.1109/ICSP54964.2022.9778616.

[4] G. S. Reddy Thummala and R. Baskar, "Prediction of Heart Disease using Decision Tree in Comparison with KNN to Improve Accuracy," 2022 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES), Chennai, India, 2022, pp. 1-5, doi: 10.1109/ICSES55317.2022.9914044.

[5] D. K. Chohan and D. C. Dobhal, "A Comparison Based Study of Supervised Machine Learning Algorithms for Prediction of Heart Disease," 2022 International Conference on Computational Intelligence and Sustainable Engineering Solutions (CISES), Greater Noida, India, 2022, pp. 372-375, doi: 10.1109/CISES54857.2022.9844328.

[6] K. Vayadande et al., "Heart Disease Prediction using Machine Learning and Deep Learning Algorithms," 2022 International Conference on Computational Intelligence and Sustainable Engineering Solutions (CISES), Greater Noida, India, 2022, pp. 393-401, doi: 10.1109/CISES54857.2022.9844406.

[7] X. Yuan, J. Chen, K. Zhang, Y. Wu and T. Yang, "A Stable AI-Based Binary and Multiple Class Heart Disease Prediction Model for IoMT," in IEEE Transactions on Industrial Informatics, vol. 18, no. 3, pp. 2032-2040, March 2022, doi: 10.1109/TII.2021.3098306.

[8] D. Sharathchandra and M. R. Ram, "ML Based Interactive Disease Prediction Model," 2022 IEEE Delhi Section Conference (DELCON), New Delhi, India, 2022, pp. 1-5, doi: 10.1109/DELCON54057.2022.9752947.

[9] G. G, "Analysis of The Diagnostic Parameters of Heart Diseases and Prediction of Heart Attacks," 2022 IEEE 3rd Global Conference for Advancement in Technology (GCAT), Bangalore, India, 2022, pp. 1-6, doi: 10.1109/GCAT55367.2022.9972211.

[10] E. Maraj and S. Kuka, "Prediction of Coronary Heart Disease Using Fuzzy Logic: Case Study in Albania," 2022 International Conference on Electrical, Computer and Energy Technologies (ICECET), Prague, Czech Republic, 2022, pp. 1-6, doi: 10.1109/ICECET55527.2022.9872569.

[11] V. Kannagi, M. Rajkumar, I. Chandra, K. Sangeethalakshmi and V. Mohanavel, "Logical Mining Assisted Heart Disease Prediction Scheme in Association with Deep Learning Principles," 2022 International Conference on Electronics and Renewable Systems (ICEARS), Tuticorin, India, 2022, pp. 1409-1415, doi: 10.1109/ICEARS53579.2022.9751820.

[12] A. Basak, M. S. Rahman and M. Rahman, "Prediction of Heart Disease Using an Approach Based on Machine Learning," 2022 13th International Conference on Computing Communication and Networking Technologies

(ICCCNT), Kharagpur, India, 2022, pp. 1-5, doi: 10.1109/ICCCNT54827.2022.9984555.

[13] G. Kumar Sahoo, K. Kanike, S. K. Das and P. Singh, "Machine Learning-Based Heart Disease Prediction: A Study for Home Personalized Care," 2022 IEEE 32nd International Workshop on Machine Learning for Signal Processing (MLSP), Xi'an, China, 2022, pp. 01-06, doi: 10.1109/MLSP55214.2022.9943373.

[14] Z. Lin, Y. Wang, Z. Xiao and L. Xie, "A Survey of Heart Disease Risk Computer Intelligent Prediction Model," 2022 IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC), Dalian, China, 2022, pp. 1433-1436, doi: 10.1109/IPEC54454.2022.9777302.

[15] T. Kavitha, S. Hemalatha, S. Mownika, G. M. Priya and J. K. Vigashini, "Review on Cardio Vascular Disease Prediction Using Machine Learning," 2022 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, 2022, pp. 1-4, doi: 10.1109/ICCCI54379.2022.9740953.

[16] S. K. L. Sameer and P. Sriramya, "Improving the Accuracy for Prediction of Heart Disease by Novel Feature Selection Scheme using Decision tree comparing with Naive-Bayes Classifier Algorithms," 2022 International Conference on Business Analytics for Technology and Security (ICBATS), Dubai, United Arab Emirates, 2022, pp. 1-8, doi: 10.1109/ICBATS54253.2022.9758926.

[17] M. S. Babu and V. Karthick, "Improving Accuracy in Intelligent Coronary Heart Disease Diagnosis Prediction Model using Support Vector Clustering Technique Compared over Random Forest Classifier Algorithm," 2022 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES), Chennai, India, 2022, pp. 1-5, doi: 10.1109/ICSES55317.2022.9914283.

[18] R. Vijayarangan., M. Hemalatha., K. Umadevi. and R. Satishkumar., "Prediction of Heart Disease at an Early Phase Using Machine Learning Techniques," 2022 IEEE International Conference on Data Science and Information System (ICDSIS), Hassan, India, 2022, pp. 1-5, doi: 10.1109/ICDSIS55133.2022.9915987.

[19] R. Jani, M. Shariful Islam Shanto, M. Mohsin Kabir, M. Saifur Rahman and M. F. Mridha, "Heart Disease Prediction and Analysis Using Ensemble Architecture," 2022 International Conference on Decision Aid Sciences and Applications (DASA), Chiangrai, Thailand, 2022, pp. 1386-1390, doi: 10.1109/DASA54658.2022.9765237.

[20] A. Srivastava and A. k. Singh, "Heart Disease Prediction using Machine Learning," 2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), Greater Noida, India, 2022, pp. 2633-2635, doi: 10.1109/ICACITE53722.2022.9823584.

[21] L. Riyaz, M. A. Butt and M. Zaman, "Ensemble Learning for Coronary Heart Disease Prediction," 2022 2nd International Conference on Intelligent Technologies (CONIT), Hubli, India, 2022, pp. 1-9, doi: 10.1109/CONIT55038.2022.9848292.

[22] M. O. Butt, A. Ur Rehman, S. Javaid, T. M. Ali and A. Nawaz, "An Application of Artificial Intelligence for an Early and Effective Prediction of Heart Failure," 2022 Third International Conference on Latest trends in Electrical Engineering and Computing Technologies (INTELLECT), Karachi, Pakistan, 2022, pp. 1-6, doi: 10.1109/INTELLECT55495.2022.9969182.

[23] G. N. Ahmad, S. Ullah, A. Algethami, H. Fatima and S. M. H. Akhter, "Comparative Study of Optimum Medical Diagnosis of Human Heart Disease Using Machine Learning Technique With and Without Sequential Feature Selection," in IEEE Access, vol. 10, pp. 23808-23828, 2022, doi: 10.1109/ACCESS.2022.3153047.

[24] K. Shah, T. Shetty, S. Sawant and J. Sisodia, "Molles Corde: A Neural Network-based Multi-model System for Detecting Heart Diseases," 2022 2nd International Conference on Intelligent Technologies (CONIT), Hubli, India, 2022, pp. 1-5, doi: 10.1109/CONIT55038.2022.9848141.

[25] E. Mun and J. Cho, "Review of IoT-based AI analysis method for heart monitoring and heart disease prediction," 2022 13th International Conference on Information and Communication Technology Convergence (ICTC), Jeju Island, Korea, Republic of, 2022, pp. 2249-2251, doi: 10.1109/ICTC55196.2022.9952501.

[26] Vijeta Sharma, Shrinkhala Yadav and Manjari Gupta. “Heart Disease Prediction using Machine Learning Techniques”, 2020, 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN), IEEE.

[27] Shaik Farzana and Duggineni Veeraiah. “Dynamic Heart Disease Prediction using Multi-Machine Learning Techniq”, 2020, IEEE.

[28] Noor Basha, Ashok Kumar P S, Gopal Krishna C and Venkatesh P. "Early Detection of Heart Syndrome Using Machine Learning Technique", 2019, 4th International Conference on Electrical, Electronics, Communication, Computer Technologies and Optimization Techniques (ICEECCOT), IEEE.

[29] Rahul Katarya and Polipireddy Srinivas. "Predicting Heart Disease at Early Stages using Machine Learning: A Survey", 2020, International Conference on Electronics and Sustainable Communication Systems (ICESC), IEEE.

[30] Alperen ERDOĞAN and Selda GÜNEY. "Heart Disease Prediction by Using Machine Learning Algorithms", 2021, IEEE.