J. Integr. Sci. Technol. 2025, 13(1), 1016

Journal of Integrated
SCIENCE & TECHNOLOGY

Enhancing heart disease diagnosis: Leveraging classification and ensemble machine learning techniques in healthcare decision-making

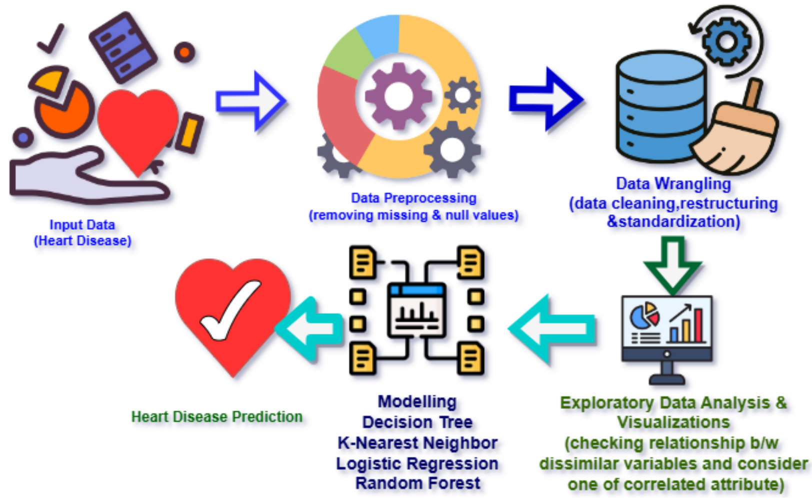
Narendra Kumar Sharma[1,2], Alok Singh Chauhan[3*], Shahnaz Fatima[2], Swati Saxena[4]

[1]Department of Computer Application, Pranveer Singh Institute of Technology, Kanpur, Uttar Pradesh, India. [2]Amity Institute of Information Technology, Amity University, Lucknow, Uttar Pradesh, India. [3]School of Computer Applications and Technology, Galgotias University, Greater Noida, Uttar Pradesh, India. [4]Department of Computer Application, Maharana Pratap Engineering College, Kanpur, Uttar Pradesh, India

Submitted on:12-Aug-2024, Accepted and Published on: 16-Nov-2024

Article

ABSTRACT



Cardiovascular disease is one of the main reasons for the demise of people in the world today, whether it is a developed country or a developing country. It is not only affecting the people living in the urban areas, but it has also affected the people of rural areas. If we know it at the primary stage, then its side effects can be avoided by reducing the chances of heart disease. So, the correct prediction of heart disease is an imperative task to assist doctors and medical experts to take decisions and make effective treatment policies to save the lives of people. In this paper, we use and combine multiple classification methods of data mining and machine learning to perk up the precision of the classifier. For this, we have used the ensemble machine learning method, which combines multiple models into a single predictive model that utilizes the advantages of multiple base models, usually called weak learners, to compensate for each model's weakness. We intend an iterative ensemble approach to integrate various low-performance classifiers to form a strong classifier with high precision. We took a dataset from the IEEE data port for its implementation, which contains around 1190 instances with 11 features of heart disease. We examine on the basis of initial symptoms whether the patient has heart disease or not. We explore the application of classification and ensemble machine learning techniques to augment healthcare decision-making for heart disease. By bridging the gap between data-driven insights and clinical decision-making, these techniques pave the means for a more proactive and patient-centric approach to cardiovascular health management.

Keywords: Healthcare, Heart Disease, Decision Making, Data Mining, Machine Learning, Ensemble Classifier.

INTRODUCTION

Heart disease remains solitary of the foremost reasons for mortality globally, posing significant challenges to healthcare systems and clinicians. With the ever-growing volume of patient data and the complexity of heart disease diagnosis and treatment, there is a pressing need for advanced computational methods to aid in decision making. Machine learning (ML) techniques have emerged as powerful tools capable of extracting valuable insights from large datasets, thus revolutionizing healthcare decision-making.

According to the World Heart Report 20.5 million deaths worldwide in 2021 were attributed to cardiovascular disease[1]. Heart disease is one of the leading reasons for the demise of non-communicable diseases in human kind today. As stated by the WHO, about 17.9 million people died of heart disease in 2019,

*Corresponding Author: Alok Singh Chauhan, Galgotias University, Greater Noida, India and Narendra Kumar Sharma, Pranveer Singh Institute of Technology, Kanpur, India
Email: alokchauhan.1983@gmail.com, narendrasharmauim@gmail.com

Cite as: *J. Integr. Sci. Technol., 2025, 13(1), 1016.*
URN:NBN:sciencein.jist.2025.v13.1016
DOI:10.62110/sciencein.jist.2025.v13.1016

which accounts for a possible 32% of all deaths worldwide. It is anticipated that 24% of the casualties owed to incommunicable ailments in India are due to heart disease.

Cardiovascular disease prediction is a significant confront in the field of medicinal science and healthcare decision-making. Machine learning is one such medium that can effectively facilitate healthcare forecasting and decision-making from the huge amount of data formed by the health sector[2]. Data mining or machine learning is an unearthing method to analyze and encapsulate large data from a mixed approach to discover important decisions[3,4]. Data mining provides a wide range of disciplines including computational theory, pattern recognition, artificial neural networks, genetic algorithms, statistics, data analytics, and probabilistic modeling.[5,6,7,8]

Heart disease persists to be a significant global health challenge,[9,10] demanding innovative solutions to improve diagnostic accuracy and treatment outcomes.[11] In this work, we introduce a novel methodology to augment healthcare decision-making for heart disease by leveraging four distinct classification algorithms: decision trees, k-nearest neighbors, logistic regression, and random forests. Additionally, we propose an iterative ensemble technique designed to boost precision and reliability in risk prediction and diagnosis.

The prevalence of heart disease underscores the urgency for more effective diagnostic tools and treatment strategies. Despite advances in medical technology, accurately assessing individual risk factors and determining optimal interventions remain complex tasks[12,13]. Machine learning algorithms offer a promising avenue for extracting actionable insights from diverse patient datasets, potentially revolutionizing clinical decision-making in cardiovascular healthcare. Our research focuses on harnessing the strengths of four classification algorithms: decision trees, k-nearest neighbors (KNN), logistic regression, and random forests. Decision trees provide intuitive interpretability, allowing clinicians to understand the underlying decision-making process. K-Nearest neighbors excel in identifying patterns within data and are particularly useful for clustering similar cases[14]. Logistic regression tends to be a straightforward probabilistic scaffold for binary classification tasks. Random forests, as an ensemble of decision trees, combine the predictive supremacy of manifold models to perk up accuracy and sturdiness[15,16].

In addition to these individual classifiers, we introduce a novel iterative ensemble approach tailored specifically for heart disease diagnosis. This iterative ensemble method iteratively combines predictions from multiple base classifiers, iteratively refining the ensemble to optimize precision while maintaining high recall. By dynamically adjusting the contribution of each classifier based on its performance, our approach adapts to the complexity and variability of heart disease data, enhancing diagnostic accuracy and clinical utility. Throughout this paper, we detail the implementation of each classification algorithm and describe our iterative ensemble approach in depth. We present experimental results demonstrating the efficacy of our method compared to traditional standalone classifiers and existing ensemble techniques. Moreover, we discuss the practical implications of our approach for real-world healthcare decision-making, emphasizing its impending to get better patient outcomes and diminish healthcare costs.

A. Data Mining and Machine Learning in Healthcare Decision Making

Data mining (DM) and machine learning (ML) are very useful techniques in healthcare and the medical field. These techniques are very helpful in accurate prediction and quick diagnosis of diseases[17,18,19]. These are very useful to predict and analyze the particular health condition of patients and guide them to medical professionals and decision makers to make effective decisions regarding the particular medical condition of the patient and their illnesses[20,21]. Hence, these techniques save a lot of time of doctors and medical experts and can be helpful in correct treatment by working out individual imprecision.[22]

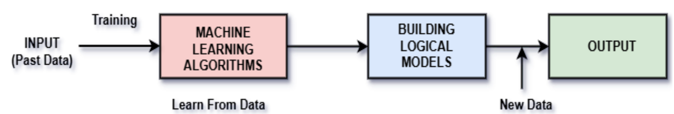


Figure 1. Machine Learning Process

B. Ensemble Learning in Decision Forecasting

Ensemble learning is a machine learning technique that mingles multiple models into a single predictive model. Ensemble learning utilizes the advantages of multiple base models usually called weak learners to compensate for each model's weakness. The foremost belief behind ensemble learning is to group weak learners mutually to form a strong learner that achieves enhanced recital than any individual weak learner.[23] This approach permits a better predictive ability in contrast to a single model.

Eventually, Ensemble methods are meta-algorithms that mingle multiple ML techniques into one predictive model to amplify the performance of the model. The advanced ensemble learning method possibly will use bagging to diminish change, boost predisposition, and stacking approach to improve prediction. We can categorize ensemble learning models into two categories based on the choice of weak learners. Homogeneous ensemble models utilize single-based learning and heterogeneous ensemble models use different base learning algorithms.[9]

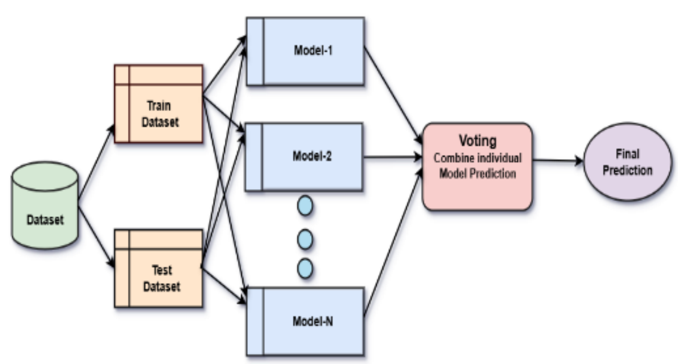


Figure 2. Ensemble Learning

Problem Definition and Motivation

There is an invaluable endowment of a healthy life, but in today's busy life, we are all affected by it. There are many such diseases but among them, heart disease and diabetes are mainly found in today's people. These diseases are not only affecting the people living in the urban areas but they have also affected the people of rural areas. In such a situation, we need a kind of health system which can reduce the risks of heart disease by making accurate and correct predictions. Thus, accurate prediction of heart disease is an imperative task to assist doctors and medical experts in making effective decisions and valuable treatment policies to save the lives of people.

Related Work

H.S. Niranjana Murthy *et al.*[24] explore and employ different data mining approaches along with machine learning techniques on diabetes and analyze the performance of the ML technique. The study is concerned with the SVM model for diabetes prediction and identification. Satish Ranbhise*et al.*[25] introduce a wireless-based health monitoring system using data mining and IoT that can monitor the health status of patients and predict heart disease based on symptoms. Basma Saleh *et al.*[26] worked on cardiovascular disease and predicted heart disease using data extraction algorithms and data mining to identify the main risk factors involved in heart disease. They explore the integration of different data extraction algorithms through data frameworks to predict heart disease. H Benjamin Fredrick David *et al.*[27] present a comparative analysis of ensemble learning algorithms that explores the future of medications and the state of the patient and identifies the crucial risk involved in heart disease.

Ahmad Mousa Altamimi *et al.*[28] worked on a children's diabetic dataset and predicted the diabetic condition in children using different data mining algorithms to analyze the performance and make early predictions on children's diabetic condition. Sinkon Nayak *et al.*[29]emphasized the heart disease diagnosis at the premature phase and used data mining classification and frequent item mining approach to filter the attribute and determine the heart disease at the primitive stage. Anjan Nikhil Repaka*et al.*[30] carry out heart disease forecasting by using Naïve Bayesian and complex encryption systems to find the risk factors involved in heart disease and design a smart heart disease prediction based on mobile health technology. Rahma Atallah *et al.*[23] use an ensemble technique to mingle several machine learning methods on the way to envisage heart disease. They determine the majority voting accomplished by ensemble method and predict the probable heart disease in people.

Chandrasekhar and Peddakrishna[31] demonstrate the efficacy of leveraging machine learning techniques and optimization strategies to enhance the accuracy of heart disease prediction. By utilizing datasets from IEEE data port, they have shown the potential for significant advancements in medical diagnostics. Their study emphasizes the importance of integrating data driven approaches into healthcare practices, paving the way for more effective and timely diagnoses, and ultimately improving patient outcomes. M. Alshraideh et al. explore how ensemble classification techniques can enhance heart disease risk prediction accuracy, offering a tailored framework integrating diverse features like demographic data and medical history. They experiment with various machine learning algorithms and ensemble strategies, such as bagging and boosting, to identify the most effective approach.[32] Through their study, they contribute valuable insights for advancing intelligent systems in healthcare.

Designed Approach

In the designed approach, manifold classifiers are promoted to increase the precision of the classifier. We intend an iterative ensemble approach to integrate various low performance classifiers to form a strong classifier with high precision. The main insight of this proposal is to situate classifier weights and prepare a test of the data at every iteration that produces correct predictions from heterogeneous interpretation. Any ML approach is often used as a base classifier as elongated because it accepts the instruction set weights. The designed approach is anticipated to meet the subsequent conditions:

- The classifier training should be carried out interactively with different weighted learning examples.
- At every iteration, the classifier attempts to render the finest for those examples at the same time as minimizing getting-to-know errors.

A. Working Steps

- Primarily, a training separation is selected randomly.
- Iteratively train the designed ML model by deciding on the training set mainly pedestal on the accurate predictions of final training.
- Give higher significance to missorted instances in order to the ones instances are much more probable to be categorized in the next recurrence.
- Based on the accuracy of the classifier, additionally allot weights to the classifier trained on each iteration. The higher classifier receives greater weight.
- We replicate the same process until the whole training data is matched with no miscalculation otherwise until the maximum number of evaluators individually is achieved.
- Toward rank, you 'vote' for each of the learning algorithms that you create.

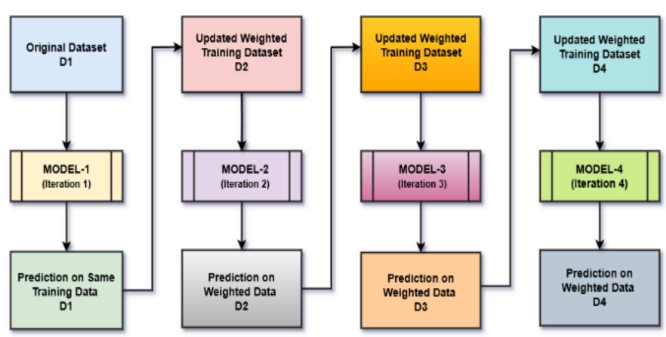


Figure 3. Designed Iterative Approach

Algorithm of designed approach is shown below:

Input

Dataset D = { (x_1, y_1), (x_2, y_2)…., (x_M, y_M) };
L = {Decision Tree (DT), Random Forest (RF), Logistic Regression (LR), KNN}
T = 4 (#Number of learning round)

Algorithm

Step 1: Initialize the weight distribution

$$D_1 = \frac{1}{M}$$

Step 2: for t= 1, 2..….T

- Train a base learner ht from D using Dt
  Ht = L (D, Dt)
- Calculate Error which is nothing, but the summation of all the sample weights of misclassified data points.

$$e_t = \sum_{i=1}^{M} D_t(i)[h_t(x_i) \neq y_i)]$$

- Determine the weight of h_t

$$\alpha_t = \frac{1}{2} ln \frac{1 - e_t}{e_t}$$

- Calculate normalization factor that enables D_{t+1} to be a distribution

$$Z_t = \sum_{i=1}^{m} D_t(i) \times \begin{cases} e^{-\alpha t} if\, h_t(x_i) = y_i \\ e^{\alpha t} if\, h_t(x_i) \neq y_i \end{cases}$$

- Update the distribution

$$D_{t+1}(i) = \frac{D_t(i)}{Z_t} \times \begin{cases} e^{-\alpha t} if\, h_t(x_i) = y_i \\ e^{\alpha t} if\, h_t(x_i) \neq y_i \end{cases}$$

Step 3: End

Output

$$H(x) = sign \sum_{t=1}^{T} \alpha_t\, h_t(x)$$

Designed method transforms weak individual models into a strong ensemble of models. We need to update weights because if the same weights are applied to the next model, the output would be the same. Records with higher weights have higher chances to be selected in the next model.

B. Model Hyperparameters

Model hyperparameters are properties that take care of the entire training of an algorithm. While an algorithm learns model parameters from data, hyperparameters are used to control the behavior of the algorithm. These parameters must be initialized before the algorithm is trained. To initiate the approach, first, we tend to put a few necessary hyperparameters, like:

- Base_estimator: weak learners trained the models and the decision tree classifier is applied as the default weed learner.
- N_estimators: add up to weak learners who train iteratively.
- Learning_rate: endorse weight obtained for weak learners. Use one as default.

C. Parameter Tuning

Hyperparameter tuning is the procedure of deciding the correct combination of hyperparameters that allows the model to maximize the recital of the model. Grid search and Bayesian optimization algorithms are usually utilized to learn when to tune hyperparameters within a system. When hyperparameter changes are primed, the grid search algorithm might also additionally seem like a nearby optimization in preference to an inclusive optimization, however Bayesian optimization set of rules is able to offer global optimizations. Grid search was utilized in this study.

METHODOLOGY

The focal intention of this research work is to successfully envisage probable heart disease problems from the medical dataset. A model is established using a prediction technique to access the heart disease facility by certain characteristics. Ensemble learning is used to construct class prediction based on selected features in this work to combine multiple classification approaches to perk up the precision of the classifier. In this research work, the dataset is obtained from the IEEE data port which contains around 1190 instances with 11 features of heart disease. We implemented a designed approach in Python with various libraries like pandas, numpy, matplotlib, seaborn, scikit-learn etc. The work execution is represented within the following flow diagram 4.

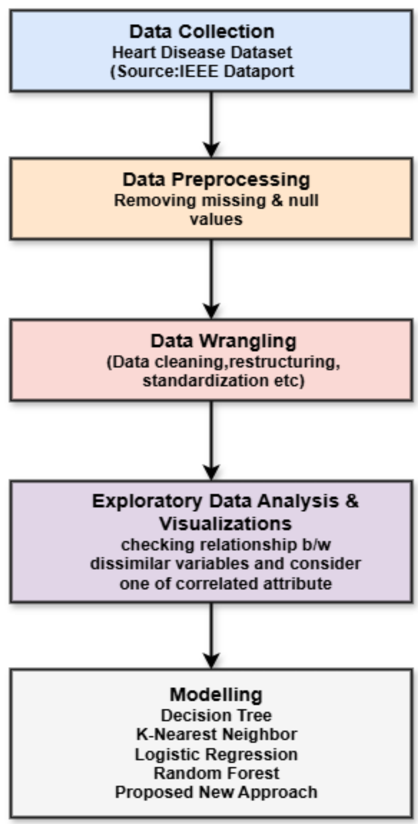


Figure 4. Flow Diagram of Execution

A. Data Set

In this research work, the dataset is obtained from the IEEE data port (https://ieee-dataport.org/open- access/heart-disease-dataset-comprehensive) which contains around 1190 instances with 11 features of heart disease patient and their health condition. The dataset is combined with five popular heart disease datasets resembling Cleveland, Hungarian, Switzerland, Long Beach VA, and Statlog dataset which make it a prevalent and comprehensive heart disease dataset. The attribute included in the dataset are

- *age*
- *sex*
- *chest_pain_type*
- *resting_blood_pressure*
- *serum_cholesterol*
- *fasting_blood_sugar*
- *resting_electrocardiogram_results*
- *max_heart_ rate_achieved*
- *exercise_induced_angina*
- *oldpeak _S T*
- *slope_of_the_peak_excersise_ST_segment*

B. Data Preprocessing

In this process, we have to check whether the dataset which has any missing or null values is available or not. This is a significant step when building a machine learning model. After the preprocessing of the dataset, no null value was found.

C. Data Wrangling

Data wrangling is a process that provides the facility to restructure, clean, enrich, discover, publish, and validate the available raw data into a more standard format. It is also known as data munging. We perform data wrangling in Python with different libraries numpy, pandas, matplotlib, plotly etc.

D. Exploratory Data Analysis

We carry out exploratory data analysis to obtain vital information about the data. To understand the data examination and the relationship between different variables, we ensure that there is a strong relationship between two variables or not, and after that, we can consider one of them. After examining the dataset, we find that there is no predicament of multi-alignment.

E. Modelling

A model represents what has been learned by a machine learning algorithm. In our experiment, we employ a Decision Tree, K-nearest neighbor, Logistic Regression, and Random Forest and propose a new approach.

We considered and put into subsequent features of the dataset:

- Sex
  - 1 = male
  - 0 = female

  Chest Pain Type
  - Value 1: typical angina
  - Value 2: atypical angina
  - Value 3: non-anginal pain
  - Value 4: asymptomatic

  Fasting Blood sugar (fasting blood sugar > 120*mg*/*dl*)
  - 1 = true
  - 0 = false

  Resting electrocardiogram results
  - Value 0: Normal
  - Value 1: having ST-T wave abnormality (T wave inversions and/or ST elevation or depression of > 0.05mV)
  - Value 2: showing probable or definite left ventricular hypertrophy by Estes' criteria

  Exercise-induced angina
  - 1 = yes
  - 0 = no

- the slope of the peak exercise ST segment
  - Value 1: upsloping
  - Value 2: flat
  - Value 3: downsloping and class
    - 1 = heart disease
    - 0 = Normal

The method was trained to recognize the target class from heart disease features and evaluate the performance metrics of the model as accuracy, precision, recall, f1-score, support, macro and weighted average. The performance metrics used for the research work is described as:

Accuracy: In this, we predict how many times the classifier is correct in total and determined as:

$$Accuracy = \frac{Number\_of\_Correct\_\Pr edictions}{Total\_\Pr edictions}$$

Accuracy can be computed as TP, TN, FP, and FN[33]

$$Accuracy = \frac{TN + TP}{TP + FP + TN + FN}$$

Where,

TP: These are the cases in which we predicted yes who have the heart disease

TN: In this we predicted whether not, and they do not have this disease.

FP: In this we predicted yes, but they do not actually have heart disease.

FN: We predicted that they did not, but they actually have the disease.

Precision: In this, we predict yes and check how many times it is correct.

$$\Pr ecision = \frac{TP}{FP + TP}$$

Recall: It is defined out of the total number of positive classes that our model envisaged correctly. The recall should be as high as possible.

$$Recall = \frac{TP}{FN + TP}$$

F1 Score: In this, recall and precision average weighted are calculated as:

$$\text{F1 score} = \frac{2}{\frac{1}{\text{recall}} + \frac{1}{\text{precision}}} = 2 \times \frac{\text{recall} \times \text{precision}}{\text{recall} + \text{precision}}$$

Macro-average: It is analyzed for each performance and calculated separately.

Weighted-Average: This is the weighted average and estimated as F1 score as follows:

$$\text{weighted average (F1 Score)} = \frac{\sum_{i=1}^{n} F_i \times N_i}{N}$$

Where,
F_i = F1 Score for i^{th} class,
N_i =calculation of instances in i^{th} class,
N = entire calculation of occurrences and n is calculation of classes in the dataset.

EXPERIMENTAL EVALUATION

Here we evaluate the performance of manifold machine learning models based on the chosen input dataset. The method of evaluation mainly focuses on the accuracy of the model during predicting the final results. To employing ML and the designed approach, we must sure that data is suited as follows:

1. First perform cleaning on data,
2. Converting null values in integer value by 0,
3. Replace entity data by unknown,
4. Substitute 0 and unknown,
5. Perform encoding on data,
6. Leave unrelated feature,
7. Perform normalization on data,
8. Split data into training and validation datasets,
9. The training data contain 80% and validation contains the remaining 20% of data.

Analyses of machine learning method are performed by Python tool with various libraries numpy, pandas, scikit- learn, seaborn, matplotlib, plotly etc. and explored categorical variable.

RESULT ANALYSIS

We can examine the existing approach and obtain the result of the designed approach on the basis of its correctness. The performance of the current approach and designed can be tested through its accuracy, precision, recall, f1-score and support. The results are revealed as follows:

The performance of all classifiers is graphically represented by the ROC (Receiver Operating Characteristics) curve in figure 6. An ROC curve is a graphical illustration that indicates the presentation of a classification model on every categorization doorstep[28]. These curve plots are on the basis of two parameters one is TP rate and other is FP rate[34]:

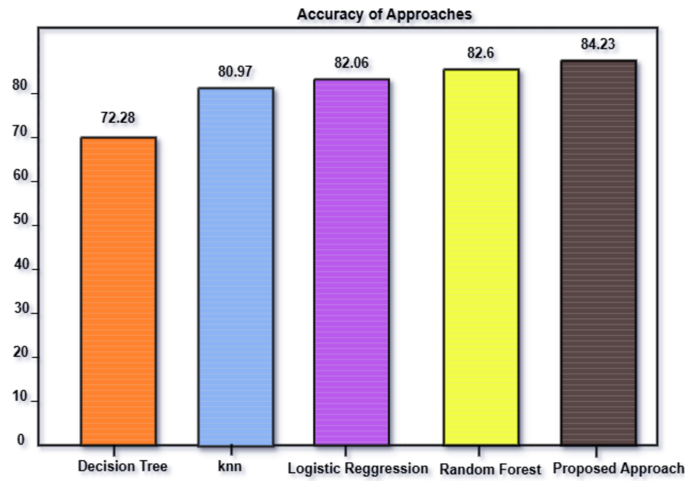


Figure 5. Accuracy of Approaches

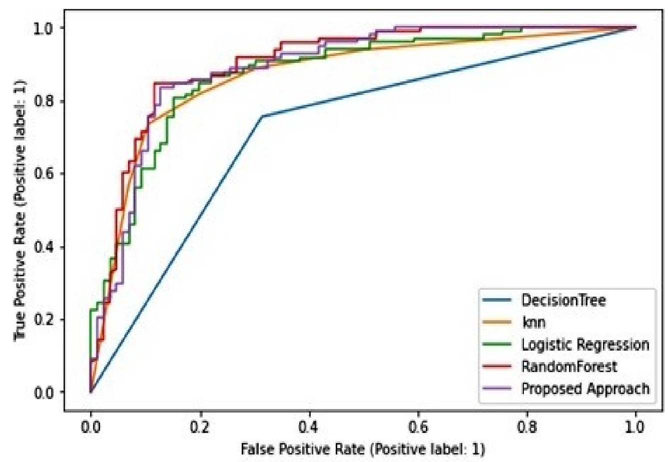


Figure 6. ROC curve of designed and other approaches

The comparison of accuracy of existing approach and designed approach is graphically represented in figure 5. It clearly shows that the obtained accuracy after performing the experiment using machine learning tool python on heart disease. In our case and scenario, the designed approach obtained the highest 84.23% accuracy as compared to other approaches like random forest get 82.60%, logistic regression get 82.06%, K-nearest neighbor get 80.97% and decision tree get 72.28% of accuracy.

Thus, the end result shows that the designed approach will be capable of assisting and forecasting the possibility involved in heart disease and effectively taking decisions to prevent heart disease. After the experiment we found that 55.3% have heart disease and 44.7% do not have heart disease.

DISCUSSION

The results demonstrate the efficacy of the designed iterative ensemble approach, which outperformed individual classifiers and traditional ensemble techniques in terms of predictive accuracy. Several key points emerge from these findings:

Precision of Predictive Models: The improved performance of the designed approach highlights the importance of leveraging ensemble techniques to enhance the exactitude of heart disease prediction. By iteratively refining the ensemble based on the performance of individual classifiers, our approach achieves higher accuracy levels compared to standalone models.

Clinical Utility: The high accuracy of predictive models is of paramount importance in clinical decision making. Healthcare practitioners can use these models to identify individuals at elevated risk of heart disease and intervene early with preventive measures or targeted treatments. The ability to accurately predict heart disease can lead to improved patient outcomes and reduced healthcare costs.

Comparative Analysis: The relative investigations afford precious imminent keenness on the strengths and weaknesses of dissimilar classification approaches. While random forest, logistic regression, and k- nearest neighbor methods demonstrate respectable accuracy, the designed iterative ensemble approach surpasses them all. Understanding the relative performance of these methods can guide researchers and clinicians in selecting appropriate algorithms for heart disease prediction tasks.

Dataset Distribution: The distribution of individuals diagnosed with heart disease versus those without heart disease highlights the prevalence of the condition within the dataset. This distribution reflects the real-world scenario where heart disease remains a significant public health concern. Predictive models trained on such datasets need to accurately capture the complexities of heart disease risk factors and manifestations to yield clinically relevant predictions.

Future Directions: Future research endeavors may focus on further optimizing predictive models, exploring alternative ensemble techniques, and incorporating additional data sources to improve accuracy and generalizability. Moreover, efforts to validate predictive models on diverse patient populations and evaluate their impact on clinical outcomes are essential for real-world deployment.

Conclusion

In this paper, we have attempted to correctly assess the future health challenges taking into relation the severity of heart disease. For which we have attempted to combine multiple data mining and machine learning methods to create a more correct and accurate predictive model, so that the probability of heart disease can be forecast correctly. The designed approach uses ensemble learning techniques that mingle multiple machine learning classifiers to produce a strong classifier with high accuracy. The foremost proposal of this approach is to put the classifier weights and train the sample data on every iteration thereby the accurate prediction can be achieved. After evaluating the accuracy with different present approaches, we can say that the precision of the designed approach is improved compared to others, which can be useful in prediction and diagnosis of heart disease. In summary, the utilization of classification and ensemble machine learning techniques represents a promising avenue for enhancing healthcare decision-making in the realm of heart disease. Through the comprehensive analysis of diverse datasets and the deployment of advanced algorithms, our research has demonstrated the potential of these methods to improve diagnostic accuracy and prognostic prediction. It can be further helpful for clinicians and medical experts to make decisions and make effective treatment policies to save the lives of people.

Conflict of Interest Statement

The authors declare that there is no conflict of interest regarding the publication of this paper.

References

1. W.H. Organization. World Heart Report 2023: Full Report - World Heart Federation. *World Hear. Fed. Accessed Oct* **2023**, 15.
2. S. Nashif, M.R. Raihan, M.R. Islam, M.H. Imam. Heart Disease Detection by Using Machine Learning Algorithms and a Real-Time Cardiovascular Health Monitoring System. *World J. Eng. Technol.* **2018**, 06 (04), 854–873.
3. A.S. Chauhan, U.K. Lilhore, A.K. Gupta, et al. Comparative Analysis of Supervised Machine and Deep Learning Algorithms for Kyphosis Disease Detection. *Appl. Sci.* **2023**, 13 (8), 5012.
4. A.S. Chauhan, J. Singh, S. Kumar, et al. Design and assessment of improved Convolutional Neural Network based brain tumor segmentation and classification system. *J. Integr. Sci. Technol.* **2024**, 12 (4), 793.
5. A. Bin Naeem, B. Senapati, A.S. Chauhan, et al. Deep Learning Models for Cotton Leaf Disease Detection with VGG-16. *Int. J. Intell. Syst. Appl. Eng.* **2023**, 11 (2), 550–556.
6. P.K. Mall, M. Kumar, A. Kumar, et al. Self-Attentive CNN+BERT: An Approach for Analysis of Sentiment on Movie Reviews Using Word Embedding. *Int. J. Intell. Syst. Appl. Eng.* **2024**, 12 (12s), 612–623.
7. M. Novita, A.S. Chauhan, R.M.D. Ujianti, et al. Exploring deep learning and machine learning for novel red phosphor materials. *J. Lumin.* **2024**, 269, 120476.
8. A. Bin Naeem, B. Senapati, A.S. Chauhan, et al. Hypothyroidism Disease Diagnosis by Using Machine Learning Algorithms. *Int. J. Intell. Syst. Appl. Eng.* **2023**, 11 (3), 368–373.
9. N.K. Sharma, S. Fatima, S. Saxena. Improvising healthcare decision making by employing ensemble technique. *J. Math. Comput. Sci.* **2021**, 11 (4), 4008–4022.
10. A.S. Chauhan. Healthcare decision making using k-means clustering technique. *PranjanaThe J. Manag. Aware.* **2020**, 23 (2), 28–33.
11. J. Singh, S. Fatima, A.S. Chauhan. Design and Performance Study of Improved Fuzzy System with Genetic Algorithm. *Int. J. Recent Innov. Trends Comput. Commun.* **2023**, 11, 242–248.
12. J. Singh, S. Fatima, A.S. Chauhan. Multi-Objective Travel Route Optimization Using Non-Dominated Sorting Genetic Algorithm. *Int. J. Intell. Syst. Appl. Eng.* **2023**, 11 (3), 785–794.
13. N.K. Sharma, S. Fatima, A.S. Chauhan. Artificial Intelligence and Internet of Medical Things in the Diagnosis and Prediction of Disease. In *Data Modelling and Analytics for the Internet of Medical Things*; CRC Press, **2023**; pp 91–108.
14. M.F. al. Maternal Health Risk Classification-A Comparison between Algorithm Decision Tree and k-Nearest Neighbor (kNN. In *2023 International Conference on Sustainable Emerging Innovations in Engineering and Technology (ICSEIET*; IEEE; pp 507–511.
15. S. Kumar, R. Dhir, N. Chaurasia. Brain Tumor Detection Analysis Using CNN: A Review. In *Proceedings - International Conference on Artificial Intelligence and Smart Systems, ICAIS 2021*; IEEE, **2021**; pp 1061–1067.
16. A.M. Madni, A. Bin Naeem, A.M. Soomro, B. Senapati, A. Singh. Breast cancer diagnosis comparative machine learning analysis algorithms. *J. Comput. Biomed. Informatics* **2023**, 04 (02), 49–65.
17. D. Anand, O.I. Khalaf, F. Hajjej, et al. Optimized Swarm Enabled Deep learning technique for bone tumor detection using Histopathological Image. *Sinergi (Indonesia)* **2023**, 27 (3), 451–466.
18. S. Mangalampalli, G.R. Karri, M. Kumar, et al. DRLBTSA: Deep reinforcement learning based task-scheduling algorithm in cloud computing. *Multimed. Tools Appl.* **2024**, 83 (3), 8359–8387.

19. G.M. Abdulsaheb, H.J. Hassan, O.I. Khalaf. A Modified Bandwidth Prediction Algorithm for Wireless Sensor Networks. *J. Inf. Sci. Eng.* **2024**, 40 (1), 177–188.
20. X. Xue, G.M. Abdulsahib, O.I. Khalaf, et al. Soft computing approach on estimating the lateral confinement coefficient of CFRP veiled circular columns. *Alexandria Eng. J.* **2023**, 81, 599–619.
21. S. Sadesh, O.I. Khalaf, M. Shorfuzzaman, et al. Automatic Clustering of User Behaviour Profiles for Web Recommendation System. *Intell. Autom. Soft Comput.* **2023**, 35 (3), 3365–3384.
22. M. Anjaneyulu, D.P. Degala, P. Devika, V. Hema. Effective heart disease prediction using hybrid machine learning techniques. *AIP Conf. Proc.* **2023**, 2492, 81542–81554.
23. R. Atallah, A. Al-Mousa. Heart Disease Detection Using Machine Learning Majority Voting Ensemble Method. In *2019 2nd International Conference on New Trends in Computing Sciences, ICTCS 2019 - Proceedings*; IEEE, **2019**; pp 1–6.
24. N.M. H.S. Early Prognosis of Diabetes Using Supervised Learning Techniques: A Comparison of Performance. *Rev. Gestão Inovação e Tecnol.* **2021**, 11 (4), 140–148.
25. R. Rastogi, M. Bansal. Diabetes prediction model using data mining techniques. *Meas. Sensors* **2023**, 25 (4), 100605.
26. B. Saleh, A. Saedi, A. Al-Aqbi, L. Salman. Analysis of weka data mining techniques for heart disease prediction system; IJECE, **2020**; Vol. 7.
27. H.B.F. David. Impact of Ensemble Learning Algorithms Towards Accurate Heart Disease Prediction. *ICTACT J. Soft Comput.* **2020**, 10 (3), 2084–2089.
28. A.M. Altamimi, M. Hijjawi. Performance Analysis of Supervised Classifying Algorithms to Predict Diabetes in Children. *J. Xi'an Univ. Archit. Technol.* **2020**, XII (III), 2010–17.
29. S. Nayak, M.K. Gourisaria, M. Pandey, S.S. Rautaray. Prediction of heart disease by mining frequent items and classification techniques. In *2019 International Conference on Intelligent Computing and Control Systems, ICCS 2019*; IEEE, **2019**; pp 607–611.
30. A.N. Repaka, S.D. Ravikanti, R.G. Franklin. Design and implementing heart disease prediction using naives Bayesian. In *Proceedings of the International Conference on Trends in Electronics and Informatics, ICOEI 2019*; IEEE, **2019**; Vol. 2019-April, pp 292–297.
31. N. Chandrasekhar, S. Peddakrishna. Enhancing Heart Disease Prediction Accuracy through Machine Learning Techniques and Optimization. *Processes* **2023**, 11 (4), 1210.
32. M. Alshraideh, N. Alshraideh, A. Alshraideh, et al. Enhancing Heart Attack Prediction with Machine Learning: A Study at Jordan University Hospital. *Appl. Comput. Intell. Soft Comput.* **2024**, 2024 (1), 5080332.
33. V.M. Patro, M. Ranjan Patra. Augmenting Weighted Average with Confusion Matrix to Enhance Classification Accuracy. *Trans. Mach. Learn. Artif. Intell.* **2014**, 2 (4), 77–91.
34. B. Van Calster, L. Wynants, G.S. Collins, J.Y. Verbakel, E.W. Steyerberg. ROC curves for clinical prediction models part 3. The ROC plot: a picture that needs a 1000 words. *J. Clin. Epidemiol.* **2020**, 126, 220–223.