Clinical Healthcare Applications: Efficient Techniques for Heart Failure Prediction Using Novel Ensemble Model

Ashwitha Anni Poojari *

*Corresponding author, Department of Information Technology, Manipal Institute of Technology Bengaluru, Manipal Academy of Higher Education, Manipal, Karnataka, India. E-mail: ashwitha.a@manipal.edu

Uppala Sivaji

Department of IT, Institute of Aeronautical Engineering, Dundigal, Hyderabad. E-mail: sivaji.u117@gmail.com

Bhavani Rupa Devi

Department of CSE, Annamacharya Institute of Technology and Sciences, Tirupati. E-mail: rupadevi.aitt@annamacharyagroup.org

Gandela Sujatha

Department of Electronics and Communication, Sri Venkateswara College of Engineering Karakambadi road, Tirupati, Andrapradesh, India. E-mail: sujatha.g@svcolleges.edu.in

Konduru Reddy Madhavi

School of Computing, Mohan Babu University, Tirupati, A.P., India. E-mail: kreddymadhavi@gmail.com

Tammineni Ravi Kumar

Department of CSE, Aditya Institute of Technology and Management, Tekkali, A.P. E-mail: ravi.4u.kumar@gmail.com

Abstract

Heart failure is a severe medical ailment that significantly impacts patients' well-being and the healthcare system. For improved results, early detection and immediate treatment are essential. This work aims to develop and evaluate predictive models by applying sophisticated ensemble learning techniques. In order to forecast heart failure, we used a clinical dataset from Kaggle. We used the well-known ensemble techniques of bagging and random forest

(RF) to create our models. With a predicted accuracy of 82.74%, the RF technique, renowned for its versatility and capacity to handle complex data linkages, fared well. The bagging technique, which employs several models and bootstrapped samples, also demonstrated a noteworthy accuracy of 83.98%. The proposed model achieved an accuracy of 90.54%. These results emphasize the value of group learning in predicting cardiac failure. The area under the ROC curve (AUC) was another metric to assess the model's discriminative ability, and our model achieved 94% AUC. This study dramatically improves the prognostic modeling for heart failure. The findings have extensive implications for clinical practice and healthcare systems and offer a valuable tool for early detection and intervention in cases of heart failure.

Introduction

Cardiovascular disease is an ailment that occurs in the heart or blood vessels. It is one of the primary causes of death, which includes other heart-related issues. According to World Health Organization (WHO) data, cardiovascular diseases, such as heart attacks and strokes, affect approximately 17 million people, leading to fatalities (Ashraf, 2021). According to the Centers for Disease Control (CDC), around 610,000 individuals die annually due to heart illness (Dangare, 2012). Modern cardiovascular diseases such as coronary artery disease, heart failure, arrhythmias, and stroke pose significant global health challenges (Austin, 2013). Several risk factors, such as hypertension, high cholesterol, diabetes, obesity, and lifestyle choices, influence these diseases. It is crucial to manage the time after getting a heart attack because the patient's chances of survival decrease if they don't receive proper attention promptly. Therefore, CVD prevention and management are essential. It is better to be proactive than reactive. These are critical findings for health professionals and researchers. Numerous research studies have been carried out to forecast the likelihood of cardiovascular disease, with the primary objective being to prevent these diseases, which contribute significantly to global mortality (Zhang, 2020). Cardiology has transformed, focusing more on prevention, early intervention, and lifestyle adjustments. Cardiology now incorporates sophisticated diagnostic tools, minimally invasive procedures, and innovative surgical techniques. Digital technologies have become essential in preventing CVD, offering several benefits. These technologies encourage the adoption of healthier lifestyles, help individuals maintain these changes, identify CVD risk factors early, and provide personalized care and

management. Additionally, healthcare professionals receive valuable clinical decision support from these technologies (Islam, 2021).

We represented the cardiovascular disease prediction and the death event, a probability measurement of whether the death risk is affirmative or negative. The UCI dataset is the most common dataset utilized in different research. Ensemble Model, the proposed model of this research, we compared the precision, recall, and F1 score evaluation with the Random Forest (RF), Bagging model. We choose the ensemble model regarding the machine learning model because the expected performance may not be reflected with the traditional machine learning methods. The data preprocessing technique has been applied to improve the data quality for data training. Our proposed model achieved the highest precision, explained in the Result section.

Research Objectives and Motivation

1. This objective involves utilizing sophisticated statistical and machine learning techniques to build a robust predictive model for heart failure. The model will be trained on a comprehensive dataset that includes relevant clinical parameters, biomarkers, and patient information.

2. Exploring correlations between variables, identifying potential outliers, and understanding distributions will inform feature selection and model development.

3. One of the primary goals is to provide a tool with direct clinical relevance, especially in the early detection of heart failure. By leveraging the predictive model, healthcare professionals can identify individuals at risk, enabling them to intervene promptly and potentially prevent or mitigate the severity of heart failure.

The motivation behind this study lies in the compelling need to address the global health burden caused by lung cancer. Early detection of lung cancer can guide to more satisfactory treatment and ultimately reduce survival.

Contribution

This research study creates several notable contributions:

• Applying the Random Forest ensemble model demonstrates a thorough understanding of advanced machine learning techniques. This contribution showcases the effectiveness of RF in predicting heart failure, highlighting its potential as a powerful tool in clinical decision support systems.

• The utilization of Bagging, a technique that leverages multiple models trained on bootstrapped data samples, represents a significant contribution. This approach not only diversifies the models but also enhances predictive accuracy. Our research thus introduces an additional dimension to ensemble learning methods, comprehensively evaluating their effectiveness in heart failure prediction.

- Developing a novel ensemble model represents a pioneering contribution to the field. This innovation signifies our commitment to pushing the boundaries of predictive modeling. Introducing a new ensemble technique has demonstrated a keen ability to think critically and creatively, potentially offering a unique solution to heart failure prediction.

Literature Review

A comprehensive analysis of different machine learning and deep learning techniques was conducted to formulate predictive models for heart disease are reviewed in this section.

A proposed "UCI Heart Disease" application utilized ensemble machine learning to assess 14 characteristics from 76 attributes. The author claims that the proposed model, XGBoost, achieved an accuracy of 81%. However, it should be noted that no comparison was made with traditional classifiers (Sai Shekhar, 2020). Another study from various experiments in this area compared Logistic regression, Naive Bayes, Decision Tree, Random Forest, Artificial Neural Networks, Multi-Layer Perception, and Deep Neural Networks. Though Ml is convenient in Heart disease prediction, the significance is the comparison between different machine learning classifiers. But Logistic regression achieved 93%, which is comparatively the highest accuracy (Katarya, 2021). Researchers conducted a study to develop a method for detecting heart disease early. They analyzed various risk factors associated with cardiac patients and used multiple learning classifiers to process the dataset. Among the different classifiers, the SVM had an impressive accuracy rate of 84.33% (Babu, 2017).

Another research introduces an innovative healthcare system designed to predict heart disease. This system employs ensemble deep learning techniques and feature fusion, effectively combining sensor data with electronic medical records to create comprehensive healthcare; the Cleveland dataset was previously employed individually, utilizing a limited set of only 16 features. Using an information gain technique, it identifies and selects crucial elements and refines the model by calculating specific feature weights for each class through a conditional probability approach. Notably, the results demonstrate that this novel system outperforms existing methods, achieving higher accuracy in heart disease prediction through a hybrid deep learning method (Ali, 2020). Based on a well-structured dataset containing information on 304 patients with a history of heart problems and other medical conditions. The dataset includes age, resting blood pressure, and fasting sugar levels. The study highlights the effectiveness of machine learning algorithms, especially K-nearest neighbors (KNN) and logistic regression. These algorithms have shown superior results in predicting and classifying patients with heart disease compared to other algorithms used in previous research. The study achieved an impressive maximum accuracy of 88.5% by applying KNN and logistic regression (LK, 2021). A separate study utilized a supervised learning approach, achieving an impressive 99% accuracy rate using the Random Forest (RF) algorithm. This was accomplished through a rigorous 10-fold cross-validation process. This study used a heart

disease dataset from Kaggle. However, the dataset may not represent all heart disease cases in the population. (Ali M. M., 2021).

A new model called MARKER-HF has been developed. It utilizes a boosted decision tree algorithm to distinguish between high and low-risk patients. The model has been assessed for its ability to predict the risk of death in different patient groups, including the UCSD, UCSF, and BIOSTAT-CHF populations. The results indicate that the proposed model has a high potential for accurate risk prediction, scoring 88% validity for mortality risk (Ahmed, 2017). Maruf et al proposed the bagging model in research work (Al Maruf, 2023). The dataset used in their research is imbalanced. Louridi et al. (2019) and Dwivedi et Al. (2023) both utilized the K-Nearest Neighbors method on the AFIC dataset. The results were, respectively, 90.16% and 80%. 46% of the 450 statistics from Kita Hospital Jakarta Enriko were accurate. Using AdaBoost, the Cleveland dataset improved the outcome by 56.13%. Shetty (Chicco, 2020) obtained 89% accuracy using the Starlog 270 data, and Chaurasia produced 75.9% accuracy using the same technique and a hybrid approach. The accuracy of an artificial neural network model was 82.5%, and the accuracy of a hybrid norm was 78.88%, according to Cheng and Chaurasia an ATTICA dataset was chosen as the subject of this inquiry. Depending on the type of classifier and training dataset, the Hellenic Score showed 20% specificity, 58% NPV, 85% accuracy, 87% PPV, and 97% sensitivity. Although the ML algorithms' specificity, accuracy, sensitivity, PPV, and NPV, respectively, ranged from 65% to 84%, 46% to 56%, 67% to 89%, and 24% to 45%.

The Hoeffding classifier used ensemble models to predict the risk of heart disease and had a CVD prediction accuracy of 88.56%. The Framingham and Cleveland tests showed 91% and 93% classification accuracies, respectively (Kavitha, 2021). The accuracy, sensitivity, and specificity of the three different ensemble approaches were 93.44%, 89.28%, and 96.96%, respectively (Al Maruf, 2023) . A novel hybrid classifier selection technique dubbed HRFLC that combines random forest, AdaBoost, and linear correlation was first demonstrated using 280 instances of the UCI dataset (Dinesh, 2018). Several researchers used an ensemble model in their investigation (Raza, 2019; Kumar, 2022; Shorewala, 2021). From the above information in this section, we have reviewed the most recent research work that has been conducted on the relevant topic (Nagarajan, 2021; Rajendran, 2021; Subasi, 2021). From these research findings, we proposed our system and showed the evaluation in the next section.

Methodology

Throughout our research, we followed a structured approach consisting of several steps, including data collection, preprocessing, model building, and analysis of the model outcomes. We proposed a new model based on our findings, which we evaluated using various metrics. To visualize our research flow, shown in Figure 1.

Dataset description

The Kaggle cardiovascular disease dataset has been used in our proposed system. The dataset includes the patient's age, anemia, creatinine phosphokinase, diabetes, ejection fraction, high pressure, platelets, serum creatinine, Serum Sodium, Sex, Smoker, and Time. The dataset has several columns, and the target class death event is binary classification 0 and 1. The dataset encompasses 299 entries, each representing an individual's pertinent medical information about heart health. Key features include age, presence of anemia, creatinine phosphokinase levels, diabetes status, ejection fraction percentage, high blood pressure indication, platelet count, serum creatinine levels, serum sodium levels, gender.
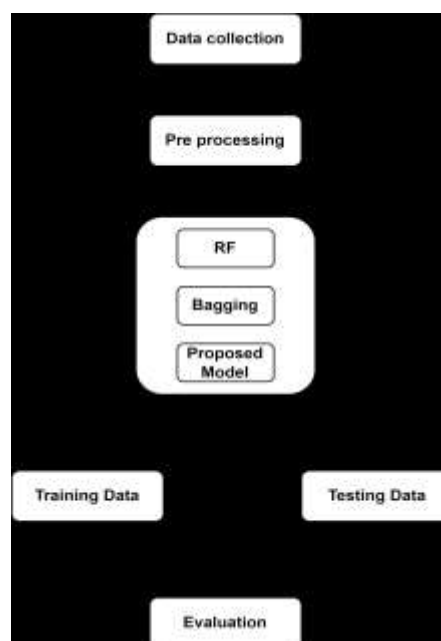

Figure 1. Flow of our research smoking habits, and the duration of observation.

Conducting a thorough statistical analysis reveals noteworthy insights. The age range spans 40 to 95 years, with a mean of approximately 60.83 years. About 43.14% of individual's exhibit anemia, while roughly 41.81% have diabetes. Creatinine phosphokinase levels range widely, from 23 to 7861, with a mean of approximately 581.84. Ejection fraction percentages range from 14 to 80%, with a mean of 38.08%. Notably, around 35.12% of individuals have high blood pressure, and approximately 32.11% are smokers. Platelet counts vary significantly, ranging from 25100 to 850000. Serum creatinine levels range from 0.5 to 9.4, with a mean of approximately 1.39. Serum sodium levels range from 113 to 148, with a mean of 136.63. The dataset's gender distribution is slightly skewed towards males, constituting about 64.88%, while females account for 35.12%. Approximately 32.11% of individuals experienced heart-related events leading to death. The observation period ranges from 4 to 285 units, with a mean duration of approximately 130.26. This comprehensive statistical analysis provides valuable insights into the distribution and characteristics of the

dataset's features, offering a deeper understanding of cardiovascular health factors and their potential associations with heart-related events.

The dataset attributes are:

- **Anaemia:** Anemia is a medical condition that arises due to a drop in healthy red blood cells or a drop in haemoglobin situations, leading to an impairment in the blood's capability to carry oxygen. Polychrome causes of anemia include iron or vitamin scarcities, habitual conditions, or inheritable factors. Symptoms of anemia include fatigue, weakness, and conciseness of breath, and treating anemia depends on relating and addressing the bearing cause. Proper opinion and operation of anemia are essential to ensure that the body's tissues and organs admit optimal oxygen delivery, promoting overall health and well-being.

- **Creatine phosphokinase:** It is a pivotal enzyme in varied tissues, including muscles and the brain. CPK has different isoenzymes in specific tissues, similar to CK- MM in cadaverous strength, CK- MB in the heart muscle, and CK- BB in the brain and nervous system. Measuring CPK situations in the blood is essential for diagnosing and covering muscle or heart damage conditions, including heart attacks and muscular diseases.

- **Ejection Fraction:** EF is a crucial cardiac parameter that measures the percentage of blood pumped out by the left ventricle during each contraction. EF is an essential diagnostic tool for heart conditions. A reduced EF can lead to heart dysfunction, especially heart failure, while a high EF may indicate specific cardiac issues.

- **Platelets:** Platelets are thrombocytes, which are tiny blood cells essential for blood clotting. Produced in the bone marrow, they play a vital role in stopping bleeding by forming clots at the site of blood vessel injuries.

- **Serum Creatinine:** This blood test measures the amount of creatinine present in the blood. Creatinine is a waste product produced by the muscles and removed by the kidneys. Monitoring the level of serum creatinine is essential for evaluating kidney function since higher levels can indicate kidney issues.

- **Serum Sodium:** A blood test that checks the serum sodium levels. Sodium is an essential electrolyte that is crucial in maintaining fluid balance and nerve function. Any abnormality in the serum sodium levels can indicate various medical conditions such as dehydration, kidney issues, or other health problems.

Classification

Before applying the proposed ensemble model, we applied two baseline ensemble models. We used Random Forest and Bagging Ensemble.

Random Forest (RF): Random Forest is a powerful ensemble learning technique frequently employed in binary classification tasks, such as predicting heart failure. It operates

by constructing numerous decision trees, each trained on a bootstrapped dataset sample, thereby introducing diversity. At each split in the trees, only a random subset of features is considered, enhancing robustness [23]. In the case of binary classification, the final prediction is determined by a majority vote among the individual trees. Mathematically, let T represent the number of trees, and pt (1|xi) be the predicted probability of the positive class by tree t. The final prediction for a given sample xi is calculated as:

$$\hat{y}_i = \begin{cases} 1, & if \ \frac{1}{T}\sum_t^T p_t(1|x_i) > 0.5, \\ 0, & otherwise \end{cases}$$

In the context of heart failure prediction, for a new patient's features, the Random Forest model will consider the opinions of all its constituent trees and predict whether the patient is at risk of heart failure based on the majority decision. This technique is known for its accuracy and robustness, making it a popular choice in medical prediction tasks.

Bagging: Bagging, or Bootstrap Aggregating, is an ensemble learning method often employed for predicting heart failure. It entails building multiple base models, typically decision trees, on different subsets of the training data.

In the context of heart failure prediction, let T represent the number of base models. For classification, the final class prediction $\hat{y}_i$ for a given patient's features xi is calculated as:

$$\hat{y}_i = arg \max_c \frac{1}{T} \sum_{t=1}^T \mathcal{P}_t(c|x_i)$$

Where pt (c|xi) is the predicted probability of class c by base model t. This equation signifies that the class determines the final prediction with the highest average predicted probability across all base models.

In the case of regression, the final prediction $\hat{y}_i$ is computed as:

$$\hat{y}_i = \frac{1}{T}\sum_{t=1}^T y_t(x_i)$$

Where yt (xi) represents the predicted value by base model t. This formula indicates that the final regression prediction is the average of the predicted values from all base models.

By applying bagging, especially in heart failure prediction, one can leverage the combined knowledge of multiple models to make more accurate and robust predictions, which is crucial in critical healthcare applications.

Proposed model

In this section, we have presented our proposed model. We denote the dataset as D, where D = {(xk,yk), k ∈ N}. The free inconstant is xk = [xk1,xk2,xk3,...,xkp] (Equation 1),

D = {(xk,yk)},k = 1,2,...,N   (1)

While yk signifies the dependent variable, where yk takes values in {0,1}. Each instance or row's independent variable is expressed as xj = [x1j,x2j,x3j,...,xNj] .

xk = [xk1,xk2,xk3,...,xkp] (2)

The weighted mean of the variable xj is computed as

$$x_j = \frac{1}{N}\sum_{k=1}^{N} \mathcal{W'}_k x_{kj,}$$ (3)

Where wk′ ≥ 0 and ▮▮▮▮ = 1. Notably, higher weighted data elements exert more influence on the weighted mean compared to lower weights. Additionally, it is worth noting that the weights are non-negative. We have applied randomization for the proposed data split algorithm. This means that in every divided step, it arbitrarily selects from the set {N}[Here,N =Natural number]. This variation makes generating a variable number of trees possible, reducing variation and improving efficiency.

We begin with the original dataset D, representing the entire trial. It is divided between distinct homogeneous sets. At each step, single data is randomly chosen from {N} and labeled as j. The chosen data is then taken out of the set to make room for more choices. D is then split into two partitions, as shown in Equation (4).

$$D = \begin{cases} D_{11,} & if\ x_{ij} < x_j, \\ D_{12,} & if\ x_{ij} \geq x_j, \end{cases}$$ (4)

Considering each child node independently, namely viewing D11 and D12 as root nodes and applying Equation (4) recursively, we obtain D21, D22 from D11 and D23, D24 from D12. Until a halting rule is satisfied, this partitioning process continues. To prevent over-partitioning of the dataset, we used two stopping rules. Maximum tree height is used as first, denoted as Hmax. This stops the method from running endlessly or from excessively splitting the input dataset. Thus, first stops growing when H = Hmax. Tree height is defined as H = 0 for main node, and H = 1 for D11 and D12, and so on.

Second, the tree stops expanding if the partition shrinks sufficiently regarding the number of instances it contains. As N (Dkl) is how many instances there are in a node Dkl, the tree-building process stops when N (Dkl) ≤ Nmin, where Nmin is the anticipated bare minimum of examples in the split shown in equation 5.

H = Hmax   or   N (Dkl) ≤ Nmin (5)

Once the dataset is partitioned, we apply the Classification task (CART) method to model each split. It calculates the likelihood of incorrectly classifying a given variable when a random selection uses Gini Impurity.

$$Gini = 1 - \sum_{i=1}^{J} \mathcal{P}_i^2$$ (6)

We now possess a forest of Tmax trees, all fitted with the CART algorithm. We employ the AB-WAE to calculate an ensemble from these CART models. This ensemble technique modifies the weighted ageing classifier ensemble [?]. Given Pa (Ψk) as the correct prediction

frequency of classifier Ψk and itter(ψk)total iterations Ψk has spent in the ensemble, the classifier's weight $\omega(\Psi k)$ is determined as shown in Equation (7).

$$w(\Psi_i) = \begin{cases} \frac{P_a(\Psi_i)}{P_\Pi} & if\ P_a(\Psi_i) > P_{\Pi a} \\ \frac{P_a(\Psi_i)}{\overline{itter(\psi_\iota)}} & else \end{cases} \qquad (7)$$

The final predictions of the model Ψ is obtained using Equation (8),

$$\Psi(x) = if\ \sum_{t=1}^{T_{max}} w\,(\Psi_t)F(i)_t(x) = \max_{j\in\{1,2,....,j\}} \sum_{t=1}^{T_{max}} w(\Psi_t)F\,(j)_t\,(x) \qquad (8)$$

Where F (k)t(x) and F(j)t(x) represent the predictions of the t-th tree for class k and j respectively. According to their accuracy and the amount of time they have been in the ensemble, the different classifiers are given weights in this ensemble technique. Classifiers whose weights fall below a specified threshold are removed from the ensemble [?]. The ensemble is guaranteed to produce the best outcomes when accuracy is used as the optimization criterion. The proposed method is depicted in Figure 2.

Results and Discussion

In this research, we harnessed the power of ensemble learning techniques in machine learning, amplifying predictive performance by combining multiple models. Leveraging Google Colab and Python as the programming language, we employed various ensemble methods in the Scikit-learn library. This included Random Forests, a robust decision tree-based ensemble, and Bagging, a technique that creates multiple subsets of the dataset for training. We experimented with the powerful and flexible library to further enhance predictive accuracy, known for its speed and performance. By utilizing ensemble learning, we aimed to create robust models capable of handling complex datasets and delivering superior predictive outcomes. We used Python programming language for this.

In this section, we commence with an in-depth analysis of the heart failure dataset, focusing on discerning influential factors contributing to heart failure. Subsequently, we will expound upon the outcomes of employing our predictive model for heart failure prognosis. In a dedicated subsection, we will provide a concise overview of the classifier's performance, evaluating its effectiveness using various assessment metrics alongside the proposed model.
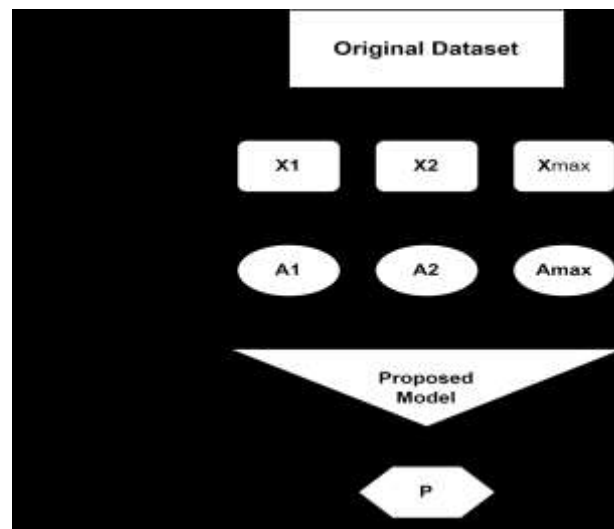
Figure 2. Architecture of the proposed model

Exploratory data analysis

Firstly, we will draw the histogram for all columns. Creating histograms for each column is a crucial step as it enables us to represent the distribution patterns within the data visually. This visualization aids in promptly detecting any potential outliers or irregularities. Furthermore, it grants us a comprehensive understanding of the general structure of the data distribution. Moreover, histograms furnish valuable insights into the dataset's range, dispersion, and central tendencies - pivotal metrics for various analytical endeavors. Hence, generating histograms for all columns is an indispensable component of exploratory data analysis, empowering us to make well-informed decisions rooted in the data's characteristics.

A deficiency in the quantity or quality of red blood cells characterizes anemia. Creatine phosphokinase (CPK), an essential enzyme, is crucial in various bodily functions. When CPK levels are significantly elevated, it often indicates injury or stress to muscle tissue, the heart, or the brain. From Figure 3, the typical range for CPK levels varies, with males falling between 39 to 308 U/L and females between 26 to 192 U/L. However, CPK levels can surge as high as 5,000 U/L in certain circumstances. Ejection fraction, another vital metric, gauges the percentage of blood expelled from the heart during each contraction. An average ejection fraction typically falls between 50% and 75%, while a borderline ejection fraction can range between 41% and 50%. Platelets, also known as thrombocytes, are blood cells crucial for clotting. A standard platelet count ranges from 150,000 to 450,000 platelets per microliter of blood. Serum creatinine level, determined through a blood test, assesses kidney function.

The typical range for serum creatinine is 0.74 to 1.35 mg/dL for adult men and 0.59 to 1.04 mg/dL (52.2 to 91.9 micromoles/L) for adult women. Elevated levels exceeding 1.2 for women and 1.4 for men may signify early kidney dysfunction. Serum sodium, a critical electrolyte, usually falls between 135 and 145 milliequivalents per liter (mEq/L) in a healthy individual. Deviations from this range, either too high or too low, may indicate underlying

kidney issues, dehydration, or other medical conditions. Understanding these parameters is pivotal in assessing and monitoring overall health. As evident from the visual representation in Figure 4, it is apparent that there is a striking similarity in the distribution between the two entities. This observation strongly suggests a high degree of mutual information associated with this feature. The resemblance in distribution patterns indicates a significant level of.
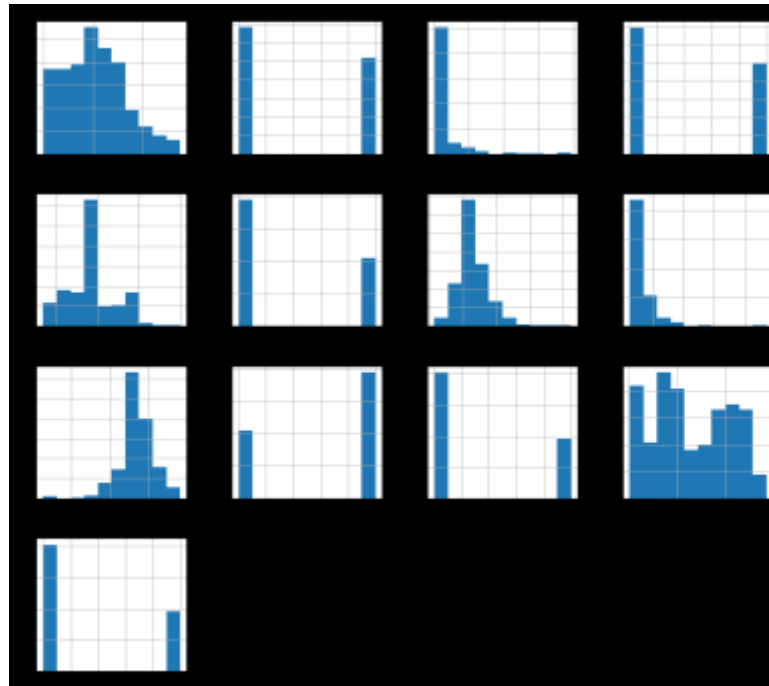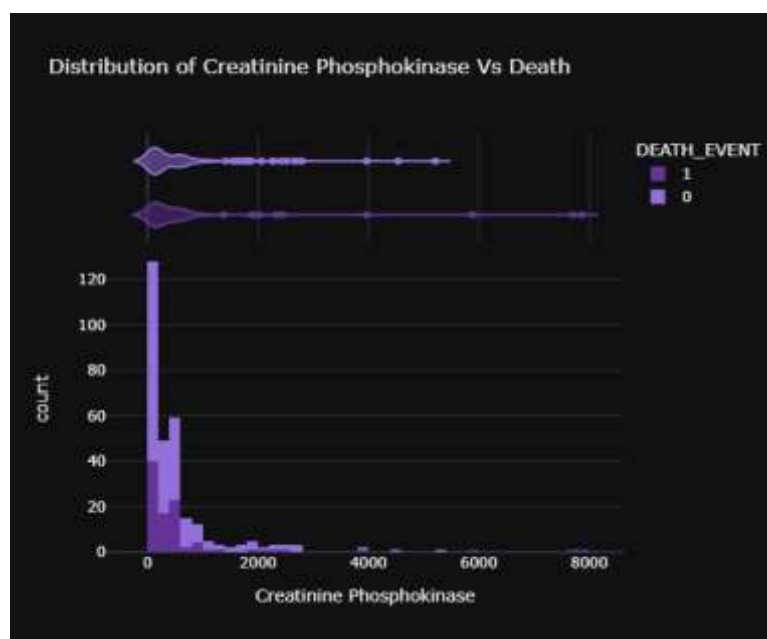

Figure 3. Histogram of all columns

Figure 4. Distribution of creatinine phosphokinase vs death interdependence or shared information between the two variables.
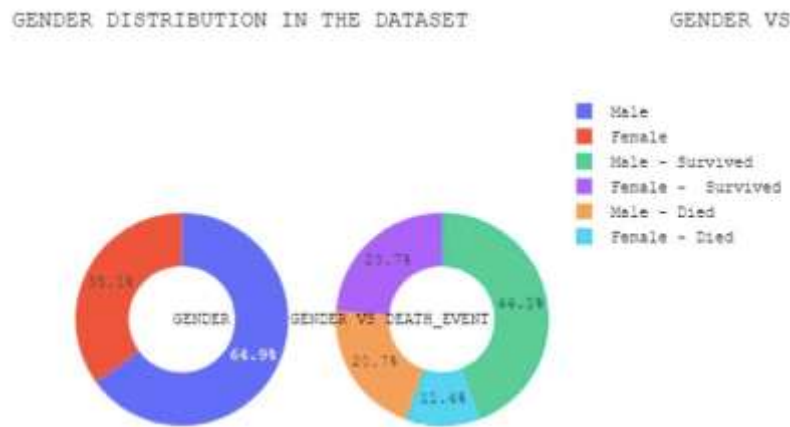
Figure 5. Distribution of gender in the dataset

The presented plot provides a comprehensive view of the dataset's gender distribution and its relationship with the occurrence of death events. The first pie chart in Figure 5 illustrates the proportion of males to females in the dataset, indicating a balanced representation of both genders. Moving to the second pie chart, it delves deeper into the data by categorizing each gender based on their survival status. Specifically, it distinguishes between males and females who have survived and those who have unfortunately experienced a death event. This visualization aids in understanding the gender-specific trends in survival outcomes within the dataset. The doughnut-like structure of the pie charts adds a layer of visual clarity, allowing for easy comparison of proportions. Furthermore, the annotations strategically placed at the center of each chart provide clear labels, enhancing interpretability. Overall, this plot offers valuable insights into the interplay between gender distribution and survival outcomes, which is critical in understanding potential patterns and making informed decisions in medical analyses. From the figure, females are 35.1%, and males are 65.9%. The male survival rate is 44.1%, and the female survival rate is 23.7%. From the analysis, males died at 20.7%, and females died at 11.4%.

Result of the proposed model

After exploratory data analysis, we will present our classifier's performance. Firstly, we have calculated all the classifier's precision, recall and f-score. In Table 1, we have presented the result of our proposed ensemble model and another two models. In this comprehensive model evaluation shown in Table 1.

Table 1. Performance of the proposed model

Model	Precision	Recall	F1-Score
RF Model	83.24%	82.54%	82.89%
Bagging Model	84.54%	86.27%	85.39%
Proposed Model	88.71%	88.34%	88.53%

Table 1, we scrutinized three distinct models - the RF, Bagging, and Proposed Ensemble. The RF model demonstrated a precision of 83.24%, indicating its proficiency in correctly identifying positive instances. Furthermore, with a recall of 82.54%, it displayed a commendable ability to capture a substantial proportion of positive cases. The resultant F1-Score of 82.89% exemplified a harmonious balance of precision and recall. The Bagging Model showcased a slightly improved precision at 84.54%. This precision indicated a heightened accuracy in the identification of actual positive instances. The model's recall, however, was particularly noteworthy at 86.27%. The F1-Score of 85.39% emphasized the model's effectiveness, showcasing its proficiency in balancing positive and negative.

Enter the Proposed Model, which outshone both counterparts with an outstanding precision of 88.71%. This signifies an exceptional ability to pinpoint positive cases accurately. The recall of 88.34% further underlined its prowess in identifying positive instances. This culminated in an exemplary F1-Score of 88.53%, affirming a robust equilibrium between precision and recall.

This comprehensive analysis elucidates that the Proposed Model emerges as the frontrunner in this evaluation. Its superior performance is evident in its ability to identify positive instances while minimizing false positives and negatives. This positions the Proposed Model as the most promising candidate among the models examined, demonstrating its potential as an effective tool in this problem domain. Three different models were evaluated based on their accuracy scores in the context of heart failure.
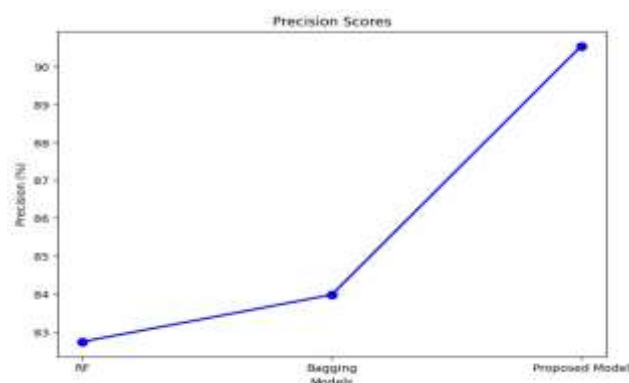

Figure 6. Accuracy of the proposed model

Prediction. From Figure 6, The RF model achieved an accuracy of 82.74%, indicating that it correctly classified approximately 82.74% of all cases. The Bagging model performed better, with an accuracy of 83.98%, suggesting that it could accurately classify about 83.98% of the cases. However, the Proposed Model stood out with an impressive accuracy of 90.54%. This indicates that the Proposed Model excelled in accurately identifying cases of heart failure, surpassing the other models' overall accuracy. It demonstrates the potential effectiveness of this model in heart failure prediction based on the available data.

In heart failure prediction, the Receiver Operating Characteristic (ROC) curve is a critical tool for assessing the performance of classifiers. It visually represents the trade-off between positive and false favorable rates, offering insights into a model's ability to distinguish between positive and negative cases. We have shown graphically the best ROC value in Figure 8. The Random Forest model exhibits a respectable ROC value of 91%, demonstrating a solid discriminatory capacity. The Bagging model elevates this performance further, achieving a notable ROC value of 92%. However, the Proposed Model truly shines, boasting the highest ROC value at 94%. This exceptional result signifies its ability to differentiate between positive and negative cases of heart failure accurately. These promising ROC values underscore the potential clinical utility of our models in predicting heart failure with a high degree of accuracy.
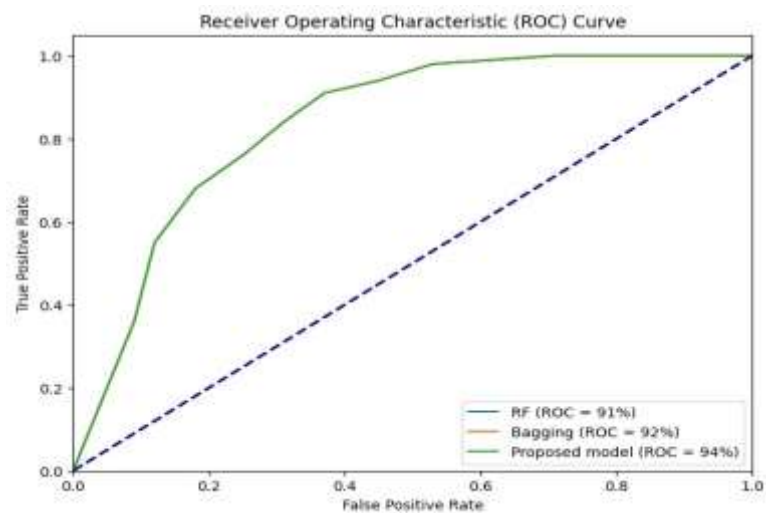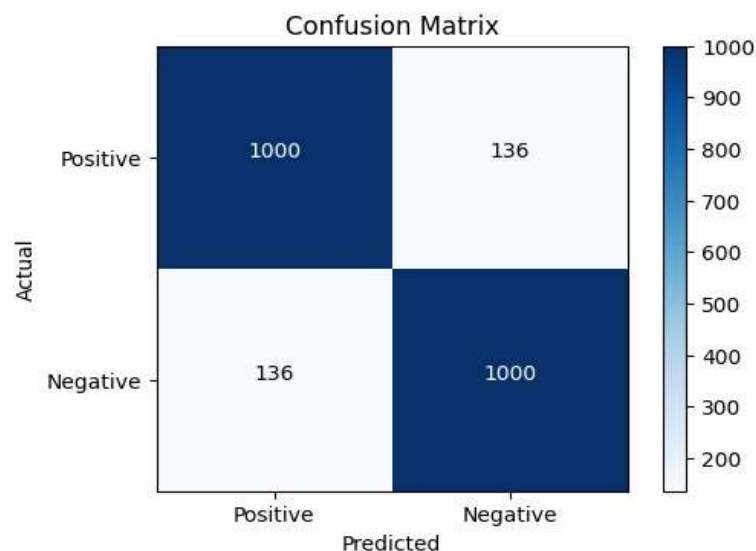

Figure 7. ROC value of the proposed model

Figure 8. Confusion matrix of the proposed model

Conclusion

Detecting and managing cardiovascular complaints is a significant health challenge worldwide. This exploration proposes an innovative approach to enhance heart complaint prediction. The technique involves an ensemble model with a delicacy-grounded weighted aging classifier ensemble, outperforming conventional models like RF and Bagging. The proposed model achieved an average of 88 across three essential evaluation criteria: perfection, recall, and F1- score. This system can efficiently prognosticate heart complaint threats and give precious clinical advice. Combining machine learning technology and clinical perceptivity can enhance patient issues and address this significant global health concern. The study involved a comprehensive analysis of ROC angles, which validated the model's superior performance in identifying heart disease. Detecting and managing cardiovascular complaints is a significant health challenge worldwide. This exploration proposes an innovative approach to support heart complications. The technique involves an ensemble model with a delicacy-grounded weighted aging classifier ensemble, outperforming conventional models like RF and Bagging. The proposed model achieved a normal of 88 across three essential evaluation criteria: precision, recall, and F1- score. This system can efficiently prognosticate heart complaint threats and give precious clinical advice. Combining machine learning technology and clinical insights can enhance patient issues and address this significant global health concern. The study involved a comprehensive analysis of ROC curves, which validated the model's superior performance in identifying heart disease.

Conflict of interest

The authors declare no potential conflict of interest regarding the publication of this work. In addition, the ethical issues including plagiarism, informed consent, misconduct, data fabrication and, or falsification, double publication and, or submission, and redundancy have been completely witnessed by the authors.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

Ahmed, S. T. (2017). A study on multi objective optimal clustering techniques for medical datasets. In *2017 international conference on intelligent computing and control systems (ICICCS)* (pp. 174-177).

Ahmed, S. T. (2023). AITel: eHealth Augmented Intelligence based Telemedicine Resource Recommendation Framework for IoT devices in Smart cities. *IEEE Internet of Things Journal*.

Al Maruf, A. a. (2023). Prediction of Heart Disease and Heart Failure Using Ensemble Machine Learning Models. In *International Conference on Machine Learning, IoT and Big Data* (pp. 481-492).

Ali, F. a.-S.-S. (2020). A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion. *Information Fusion*, 208-222.

Ali, M. M. (2021). Heart disease prediction using supervised machine learning algorithms: Performance analysis and comparison. *Computers in Biology and Medicine*, 104672.

Ashraf, M. a. (2021). *Prediction of cardiovascular disease through cutting-edge deep learning technologies: an empirical study based on TENSORFLOW, PYTORCH and KERAS.* Springer.

Austin, P. C. (2013). Using methods from the data-mining and machine-learning literature for disease classification and prediction: a case study examining classification of heart failure subtypes. *Journal of clinical epidemiology*, 398-407.

Babu, S. a. (2017). Heart disease diagnosis using data mining technique. In *2017 international conference of electronics, communication and aerospace technology (ICECA)* (pp. 750-753).

Chicco, D. a. (2020). Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone. *BMC medical informatics and decision making*, 1--16.

Dangare, C. S. (2012). Improved study of heart disease prediction system using data mining classification techniques. *International Journal of Computer Applications*, 44-48.

Dinesh, K. G. (2018). Prediction of cardiovascular disease using machine learning algorithms. In *2018 International Conference on Current Trends towards Converging Technologies (ICCTCT)* (pp. 1-7).

Islam, S. M. (2021). Digital health approaches for cardiovascular diseases prevention and management: lessons from preliminary studies. *Mhealth*.

Katarya, R. a. (2021). Machine learning techniques for heart disease prediction: a comparative study and analysis. *Health and Technology*, 87--97.

Kavitha, M. a. (2021). Heart disease prediction using hybrid machine learning model. In *2021 6th international conference on inventive computation technologies (ICICT)* (pp. 1329-1333).

Kumar, S. S. (2022). Unstructured Oncological Image Cluster Identification Using Improved Unsupervised Clustering Techniques. *Computers, Materials \& Continua*.

LK, S. S. (2021). COVID-19 outbreak based coronary heart diseases (CHD) prediction using SVM and risk factor validation. In *2021 Innovations in Power and Advanced Computing Technologies (i-PACT)* (pp. 1-5).

Louridi, N. a. (2019). Identification of cardiovascular diseases using machine learning. In *2019 7th mediterranean congress of telecommunications (CMT)* (pp. 1-6).

Nagarajan, S. M. (2021). Feature selection model for healthcare analysis and classification using classifier ensemble technique. *International Journal of System Assurance Engineering and Management*, 1-12.

Rajendran, N. A. (2021). Heart disease prediction system using ensemble of machine learning algorithms. *Recent Patents on Engineering*, 130-139.

Raza, K. (2019). Improving the prediction accuracy of heart disease with ensemble learning and majority voting rule. In *U-Healthcare Monitoring Systems* (pp. 179-196).

Sai Shekhar, M. a. (2020). Heart Disease Prediction Using Machine Learning. In *International Conference on Emerging Trends and Advances in Electrical Engineering and Renewable Energy* (pp. 603-609).

Shorewala, V. (2021). Early detection of coronary heart disease using ensemble techniques. *Informatics in Medicine Unlocked*, 100655.

Subasi, A. a. (2021). Intrusion detection in smart healthcare using bagging ensemble classifier. In *International Conference on Medical and Biological Engineering* (pp. 164-171).

Zhang, Z. a. (2020). Enhanced character-level deep convolutional neural networks for cardiovascular disease prediction. *BMC medical informatics and decision making*, 1-10

Bibliographic information of this paper for citing: