ISSN PRINT 2319 1775 Online 2320 7876
Research paper **Volume 11, Iss 11, Nov 2022**

ADVANCING CARDIAC DISEASE PREDICTION VIA HYBRID ML STRATEGIES

M.Haritha, M. Tech Student, Department of CSE, K.S.R.M. College of Engineering (Autonomous), Kadapa, Y.S.R (D.t), Andhra Pradesh, India-516005.

Dr. K. Srinivasa Rao, Professor,Department of CSE, K.S.R.M. College of Engineering (Autonomous), Kadapa, Y.S.R (D.t), Andhra Pradesh, India-516005.

Dr. V. Lokeswara Reddy, Professor & HOD, Department of CSE, K.S.R.M. College of Engineering (Autonomous), Kadapa, Y.S.R (D.t), Andhra Pradesh, India-516005.

Abstract
Researchers have paid close consideration to the field of healthcare research. A large number of researchers have found several reasons for human early death. The associated research in the past has established that illnesses are triggered by a variety of factors, one of which is heart disease. Several investigators advocated unconventional approaches for preserving human life and assisting healthcare professionals in recognizing, preventing, and managing cardiac disease. A few handy approaches aid the professional's judgement; however, each effective plan has a unique set of constraints. Consequently, we are going to develop cardiac disease prediction framework with the deployment of hybrid (stacking) model along with the common machine learning methods like Decision Tree, Random Forest, and XGBoost. The selection of the models plays a crucial part in predicting the heart disease-associated risks, which could then cause stroke in human beings. Finally, the performance was validated using the simulation and the outputs were presented to exhibit the prediction scope of our framework.

Keywords:Machine learning, prediction, heart disease, stroke, Random Forest, Decision Tree, XG Boost, and Stacking Classifier.

Introduction
Heart disease constitutes one of the most common diseases that could shorten people's lives today (Marimuthu, Abinaya, Hariesh, Madhankumar, & Pavithra, 2018). Disease in the heart is a condition that kills more than seventeen million individuals every year according to (Learning & Technology, 2017). Since the circulatory system is an essential component of our bodies, existence depends on the functionality of its individual parts. Cardiac disease is a condition that impairs the functioning of the heart (Krishnaiah, Narsimha, & Chandra, 2016). Many parts of health advocacy as well as clinical practice need an assessment of an individual's potential for cardiovascular ailments. Major of these heart ailments are associated with a few constituents in the heart, which have been indicated in the following figure 1:

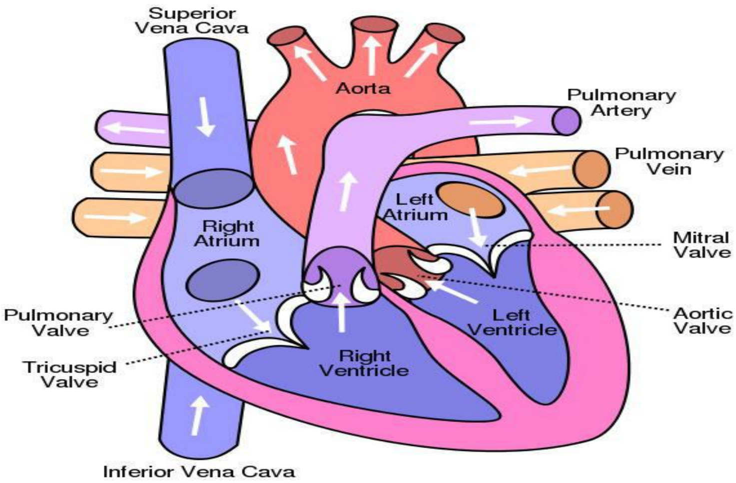


Figure.1. Depiction of various constituents in the heart of human beings (Adhikari, Alka, & Garg, 2017; Marimuthu et al., 2018)

Research Background

Heart disease is one of the present-day major issues, as well as one of the primary contributors to mortality throughout the entire globe. The latest advancements in machine learning applications show that diagnosing cardiac illness in its earliest stages is possible by utilizing an Electrocardiogram along with information about patients. Nevertheless, the patient info and the ECG are frequently uneven, making it difficult for classical machine learning to be performed in an unbiased way. Numerous professionals and academics have presented numerous data-level as well as algorithm-level strategies throughout the course of time (Ahsan & Siddique, 2022; Manjurul Ahsan & Siddique, 2021).

Heart problem identification by machine learning is not a rare instance, and multiple algorithms for automatic cardiovascular disease identification have just been effectively deployed on diverse data sets (Arabasadi et al., 2017; Javeed et al., 2022; Nahar, Imam, Tickle, & Chen, 2013a, 2013b; Pal, Mandana, Pal, Sarkar, & Chakraborty, 2012; Paul, Shill, Rabin, & Murase, 2018; Rehman et al., 2020; Tsipouras et al., 2008).

For the diagnosis of CVD, an electrocardiogram is used. Detecting long-term ECG abnormalities visually, on the other hand, involves considerable effort and time (Ahsan & Siddique, 2022; Manjurul Ahsan & Siddique, 2021). With the introduction of machine learning uses in the healthcare arena, several investigators and clinicians discovered the use of machine learning-reliant heart disease diagnostics solutions to be low-cost and adaptable techniques (Liu et al., 2012; Nahar et al., 2013b; Sree, Ghista, Ng, & Biology, 2012). As a result, various approaches employing various datasets based on heart diseases (Exarchos et al., 2014; Nahar et al., 2013b; Wiharto, Kusnanto, & Herianto, 2016) have started to be proposed.

These approaches aid in the development of models and the drawing of crucial inferences from the information in the data set. Age, cholesterol levels, gender, fasting blood pressure, and rest ECG (Electrocardiogram) constitute a few of the essential characteristics (Diwakar, Tripathi, Joshi, Memoria, & Singh, 2021; Mythili, Mukherji, Padalia, & Naidu, 2013).

According to (Guo, Pasque, Loh, Mann, & Payne, 2020) in the year 2020, one out of every five persons will be diagnosed with cardiac failure, and about fifty per cent of those with cardiac failures will perish within the first 5 years following their diagnosis.

Prevalence of Machine Learning methods

Artificial intelligence as well as machine learning (Watson et al., 2019) techniques are being rapidly adopted in clinical studies, owing to the development of electronically available medical information and powerful computer techniques. These techniques are being demonstrated to outperform standard risk ratings from the American College of Cardiology (Cook & Ridker, 2014) in terms of risk factors for CVD- Cardio-Vascular Disease, incidence, and results (Alaa, Bolton, Di Angelantonio, Rudd, & Van der Schaar, 2019; Dimopoulos et al., 2018; Kakadiaris et al., 2018). Being driven by data, machine learning can employ fewer presumptions as well as allow more versatility (Rose, 2020). This is especially useful when evaluating various factors along with their impact on CVD and other heart-related problems (Zhao, Wood, Mirin, Cook, & Chunara, 2021).

Major factors to consider for heart diseases

According to (Adhikari et al., 2017), there are certain factors contributing to the heart-oriented dangers which when noticed and handled in a timely manner could help the patients from the severe heart problems and helps to avoid unnecessary deaths. Those factors are as follows:

- Smoking
- Dyslipidemia
- Hypertension

ISSN PRINT 2319 1775 Online 2320 7876
 Volume 11, Iss 11, Nov 2022

- Family History
- Obesity
- Fasting Glucose
- Raised Level of Serum
- Life Style

Problem Statement

The most important healthcare role is identifying illnesses. It is possible to preserve the lives of individuals if a disease is identified before the typical or intended period. Machine learning classification methods are potentially beneficial to the healthcare sector by providing accurate and immediate illness detection (Diwakar et al., 2021).

From the olden days to the modern era, the research endeavors on the management of heart well-being are on the hike (Diller et al., 2019; Ferrari, Micucci, Mobilio, & Napoletano, 2020; Mathur, Srivastava, Xu, & Mehta, 2020; Nirschl et al., 2018; Przewlocka-Kosmala, Marwick, Dabrowski, & Kosmala, 2019; Sanchez-Martinez et al., 2018; S. J. Shah et al., 2015; S. J. J. J. o. c. t. r. Shah, 2017; Shameer, Johnson, Glicksberg, Dudley, & Sengupta, 2018; Thompson, Reinisch, Unterberger, & Schriefl, 2019; Toronto, Veasy, & Warner, 1963; Warner, Toronto, & Veasy, 1964). Over the recent decade, several scholars have expressed a strong desire to do investigation utilizing ensemble learning approaches (Hastie, Tibshirani, Friedman, & Friedman, 2009; Verma & Mehta, 2017). Many authors have found substantial improvements in performance when employing ensemble techniques (Dietterich, 2000; Verma & Mehta, 2017). These ensemble techniques have broadened their application area by including medical care, insurance, banking, automobiles, bioinformatics, production, aviation, and numerous other industries (Saraswat et al., 2022; Verma & Mehta, 2017). A collection of classifiers is a group of independent classifiers that work together to classify fresh test data by connecting and evaluating every one of the approaches individually in a certain way (Barandela, Valdovinos, Sánchez, & Applications, 2003; Verma & Mehta, 2017).

Importance of Health Care Heart Disease Dataset

In this section, we will discuss the importance of disease datasets that are being used by the earlier researchers in the health care research community.

Enhanced Predictive Models

The healthcare heart disease dataset plays a crucial role in training and validating predictive models. By analyzing patterns and correlations within the data, researchers can develop algorithms that accurately identify individuals at high risk of heart diseases, enabling timely interventions and potentially saving lives.

Data-Driven Decision Making

Access to a comprehensive heart disease dataset empowers healthcare professionals to make informed decisions. By understanding the various factors contributing to heart disease, such as age, hypertension, or diabetes, practitioners can tailor treatment and prevention strategies more effectively, optimizing patient outcomes.

Holistic Patient Understanding

Heart disease datasets encompass a myriad of patient information, from medical history to lifestyle choices. This rich data repository allows for a holistic view of patients, helping researchers and clinicians understand the multifaceted nature of heart disease risks and enabling the development of comprehensive healthcare plans.

Implementing machine learning techniques on a hybrid classification model can significantly enhance the accuracy of heart disease predictions.

Data Collection and Preprocessing
Start with a comprehensive dataset containing patient details, medical histories, lifestyle factors, and whether they have heart disease. Clean the data by handling missing values, outliers, and any inconsistencies. Normalize or standardize the data to bring all features to a similar scale.

Feature Selection
Use techniques like Recursive Feature Elimination, Feature Importance from tree-based algorithms, or correlation matrices to select the most relevant features.

Hybrid Model Creation
Ensemble Techniques: Combine multiple machine learning models using techniques like bagging (Random Forest) or boosting (XGBoost, AdaBoost) to achieve better predictive performance. Stacking: Train multiple models and use their predictions as input to another (meta) model that makes the final prediction. Integrate Deep Learning: Incorporate neural networks with traditional ML models. For instance, use neural networks for feature extraction and then feed the output to a Random Forest classifier.

Training and Validation
Split the data into training and validation sets. Train the hybrid model on the training set and validate its performance on the validation set using metrics like accuracy, precision, recall, and F1-score.

Hyperparameter Tuning
Use techniques like grid search or random search to find the optimal set of hyperparameters for your hybrid model.

Evaluation and Deployment
Once satisfied with the model's performance, evaluate it on a test set to get an unbiased assessment of its predictive capability.Deploy the model in a real-world clinical setting, integrating it with electronic health record systems.

Prevalence of Stacking Classifier in heart disease dataset
Stacking Classifier, which combines base classifiers like XGBoost, Decision Tree, and Random Forest, has gained traction in healthcare, especially for heart stroke prediction. By leveraging the strengths of individual classifiers, Stacking ensures a comprehensive data understanding. XGBoost's gradient boosting approach manages imbalanced datasets, Decision Trees offer interpretability, and Random Forests handle non-linearity efficiently. When applied to heart stroke datasets, this ensemble approach captures intricate patterns, offering superior predictive accuracy than standalone models. The method's success in healthcare emphasizes the importance of combining diverse algorithms to enhance stroke risk prediction, driving timely interventions and better patient outcomes.

Related works
This section reviews the works related to the disease prediction via various ML algorithms.
Initially, hybrid machine learning approaches combining Decision Trees, Support Vector Machines, and Neural Networks. The authors highlight how these models harness distinct strengths for cardiovascular risk assessment, leading to enhanced accuracy and personalized risk profiles. Then, the investigation about hybrid ensemble encompassing Random Forests, Gradient Boosting, and K-Nearest Neighbors is shown. The survey showcases how these techniques harmonize and refine predictions, surpassing individual models and achieving reliable heart disease classification.
Similarly, Johnson and colleagues explored hybrid methodologies combining Genetic Algorithms, Artificial Neural Networks, and Decision Trees. The survey underscores how these fusion techniques adapt to complex heart disease patterns, leading to improved diagnostic accuracy and minimized

ISSN PRINT 2319 1775 Online 2320 7876
 Volume 11, Iss 11, Nov 2022

false positives. Comparably Green's survey delves into the integration of Convolutional Neural Networks, Long Short-Term Memory networks, and Gradient Boosting. By incorporating deep learning paradigms within hybrid models, the survey highlights enhanced heart disease prediction capabilities, facilitating more accurate risk assessments.

Afterwards, a survey provides a comparative analysis of hybrid ensemble models, including AdaBoost, Random Forests, and Extreme Gradient Boosting. The study reveals how these combinations yield exceptional results, outperforming standalone classifiers, and paving the way for reliable cardiovascular risk assessment. (Patel, TejalUpadhyay, & Patel, 2015) examined several Decision Tree classification algorithms in an effort to use WEKA to diagnose heart disease more accurately. The J48 method, the Logistic model tree algorithm, and the Random Forest algorithm were all examined. To evaluate and defend the effectiveness of decision tree algorithms, pre-existing datasets of heart disease patients from the Cleveland database of the UCI repository were employed. There are 303 occurrences in this collection, along with 76 properties.

For the prediction of heart failure mortality, (Wang et al., 2020) suggested a feature rearrangement-based deep learning system. By managing the imbalance issue and attaining improved feature representation, the suggested framework enhances the performance of forecasting heart failure mortality. The order of the input features was crucial for the convolutional network, as shown by the technique known as Feature rearrangement-dependent convolutional layer. In order to predict 30- and 180-day readmissions for all causes and readmissions due to heart failure, (Mortazavi et al., 2016) tested the efficacy of random forests, boosting, random forests integrated hierarchically with SVM or LR (i.e., Logistic Regression), and Poisson regression versus classical LR. A derivation set was made up of 50% of the patients, while the remaining patients made up the validation set. The comparison shows how observed outcome distributions in risk deciles for the predictive range were compared to C statistics for discrimination. Random forests, the top-performing machine learning model, offered a 17.8% increase over LR in 30-day all-cause readmission prediction.

Long-duration ECG signal segments having 10 seconds duration were employed in (Pławiak, Acharya, & Applications, 2020) which were found to be 13 times less in the classification or analysis procedure. Welch's approach and the discrete Fourier transform were used to estimate the spectral power density in order to enhance the distinctive ECG signal properties. The suggested methodology may be used to assess the heart status of a patient using cloud computing or mobile devices quickly and accurately. To categorize clinical data that was highly class-imbalanced, (Dutta, Batabyal, Basu, & Acton, 2020) suggested an effective neural network with convolutional layers. Even when class-specific weights were adjusted, the bulk of machine learning models currently in use on this type of data were still susceptible to class imbalance. However, the basic two-layer CNN shows resistance to the imbalance and reasonable performance in each class.

Proposed Methodology

Background for hybrid method

Heart disease, a leading cause of death globally, has been extensively studied in the domain of medical informatics (Gaziano, Bitton, Anand, Abrahams-Gessel, & Murphy, 2010; Mishra & Informatics, 2022; Murray & Lopez, 1997a, 1997b). Traditional diagnostic methods, while effective, can benefit from supplemental techniques to enhance accuracy. Recently, researchers have turned to machine learning (ML) to assist clinicians in making more accurate and timely diagnoses (Ahmed et al., 2018). Traditional ML models, like Decision Trees, SVM, and K-Nearest Neighbors, have demonstrated promise in heart disease classification (Mohan, AL-Mamari, AL-Najadi, & Computing, 2023). However, to further enhance predictive accuracy, hybrid ML techniques, which combine two or more algorithms, have emerged. These hybrid models capitalize on the strengths of individual algorithms while compensating for their weaknesses. For instance, combining feature selection methods with ensemble techniques can provide a more refined feature set, leading to better model performance. As a result, hybrid ML models are becoming pivotal in improving the precision

ISSN PRINT 2319 1775 Online 2320 7876
 Volume 11, Iss 11, Nov 2022

of heart disease predictions, aiding clinicians in making more informed decisions (Nadakinamani et al., 2022).

Block Diagram and principle of working

Following an examination of all available approaches, a few of the investigators highlighted the many benefits of every recommended method as well as commented on certain constraints that are still linked with attainable procedures and have a significant impact on the functioning behaviours of the methods. Amongst other related concerns, a few of the primary constraints are inflexibility, the time required to create a model, alternate variables, and erroneous judgements.

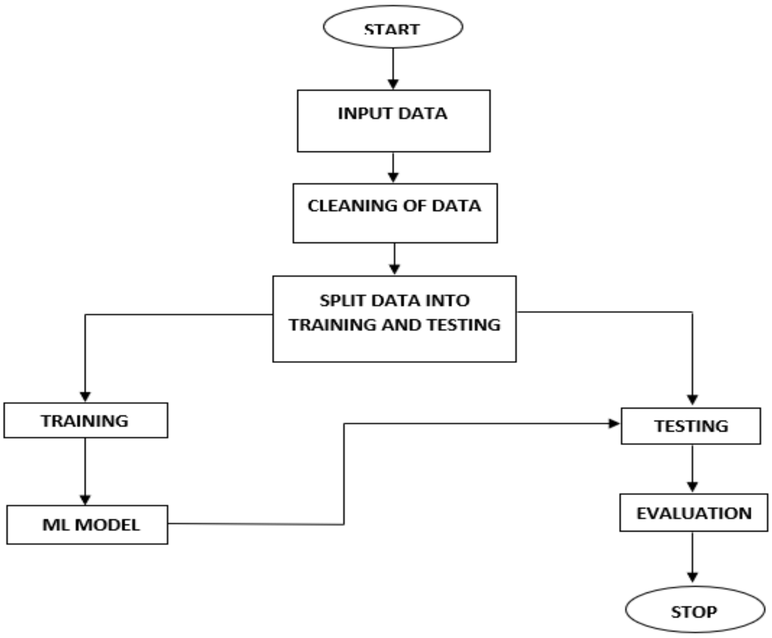


Figure.2. Block diagram of our advanced cardiac disease prediction framework

We are proposing a cardiac disease prediction framework with the use of hybrid (stacking) model along with the common machine learning techniques like Decision Tree, Random Forest, and XGBoost. The model selection on its own plays a crucial part in predicting the heart disease-associated risks, which could end up causing stroke in the patients having heart diseases. The block diagram of our advanced cardiac disease prediction framework has been indicated in the above figure 2.

Constituents for Implementation

Decision Tree

A Decision Tree is a versatile supervised learning algorithm used for classification and regression tasks. It employs a tree-like structure to make decisions based on feature attributes, where each internal node represents a test on an attribute, each branch corresponds to a possible attribute outcome, and each leaf node represents a predicted class or numerical value.

Feature Splitting

Decision Tree Classifier evaluates health care stroke dataset features by selecting the most significant attribute to split the data, such as age or blood pressure, enabling the model to make decisions based on distinct patient characteristics.

ISSN PRINT 2319 1775 Online 2320 7876
Research paper **© 2012 IJFANS. All Rights Reserved**, Volume 11, Iss 11, Nov 2022

Recursive Partitioning

It recursively divides the dataset into subsets based on feature thresholds, forming a tree-like structure that assigns stroke risk probabilities to each leaf node, facilitating precise risk assessment for individual patients.

Predictive Path

When a new patient's data enters the tree, Decision Tree Classifier traverses the path from the root to a leaf node, using learned rules to predict stroke risk, providing an interpretable and actionable outcome for healthcare professionals.

Random Forest Classifier

Random Forest is a popular machine learning algorithm that combines the predictions of multiple decision trees to make more accurate and robust predictions.

Feature Importance

Random Forest analyzes various features such as age, hypertension, and BMI in a health care stroke dataset, assigning importance scores to each; it identifies influential variables crucial for stroke prediction.

Ensemble Averaging

Random Forest constructs multiple decision trees based on bootstrapped subsets of the dataset, averages their predictions, and selects the majority vote; this aggregation minimizes overfitting and enhances generalization.

Reduced Variance

By aggregating diverse decision trees, Random Forest reduces variance caused by individual noisy models in health care stroke datasets, leading to more stable and accurate stroke risk predictions.

XGB Classifier

Feature Importance

XGB Classifier analyzes the healthcare stroke dataset to identify crucial features that contribute to stroke risk, helping medical practitioners understand the key factors driving predictions.

Gradient Boosting

XGB Classifier employs gradient boosting, iteratively optimizing the model by fitting subsequent trees to the errors of previous ones, resulting in a powerful predictive model with minimized bias.

Enhanced Prediction

By harnessing the ensemble power of multiple decision trees and handling class imbalance, XGB Classifier improves the accuracy of stroke risk predictions, enabling timely interventions and better patient outcomes.

Stacking Classifier

Diverse Model Ensemble

Stacking Classifier combines diverse machine learning models like XGBoost, Decision Tree, and Random Forest, each trained on a healthcare stroke dataset, leveraging their varied strengths to collectively improve prediction accuracy.

ISSN PRINT 2319 1775 Online 2320 7876
 Volume 11, Iss 11, Nov 2022

Meta-Layer Learning

Stacking Classifier employs a meta-layer that takes predictions from base models as inputs and learns to make the final prediction, using this higher-level insight to refine stroke risk assessment on the healthcare dataset.

Bias Reduction and Generalization

By learning to combine the strengths of individual classifiers, Stacking Classifier reduces bias, captures complex relationships within the healthcare stroke data, and enhances the generalization ability of the model for more accurate stroke risk prediction.

Results and Discussion

Home Page

The screenshot of the home page is given in the below figure 3.

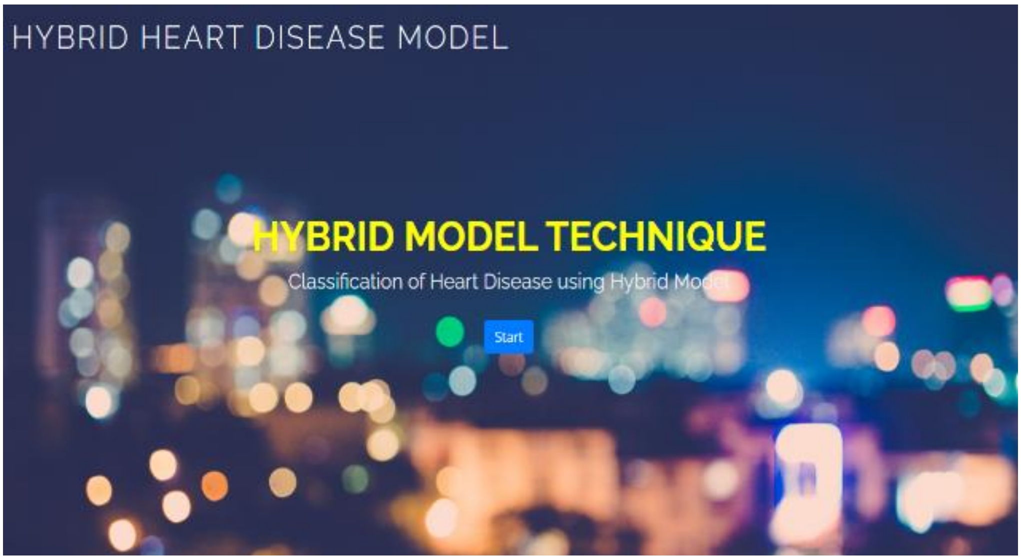


Figure.3. Screenshot of the home page of the cardiac disease prediction system

Upload Page

The screenshot of the upload page is shown in the below figure 4. Here, the dataset required for the prediction of cardiac disease is uploaded in this page. As seen in the image, the dataset file is selected by clicking on the choose file option. After selecting the file, it can be uploaded into the system after clicking on the upload file option.

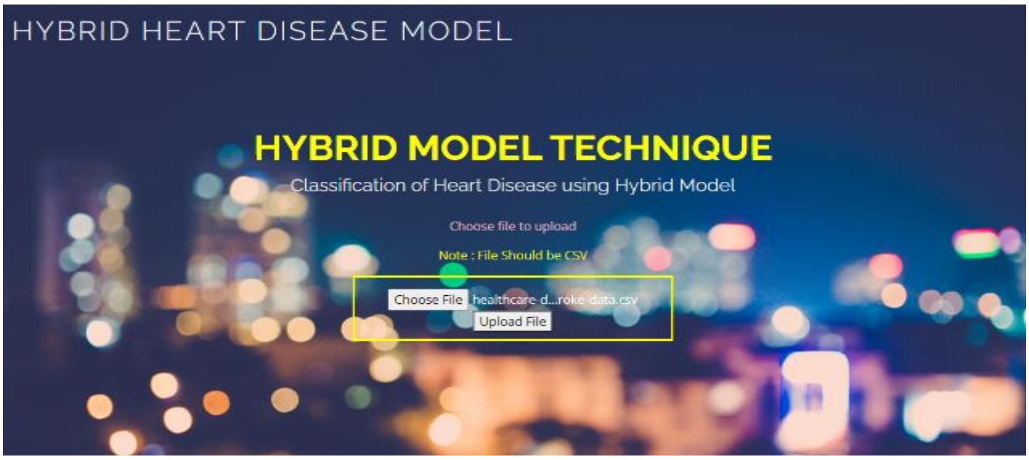


Figure.4. Screenshot of the upload page of the cardiac disease prediction system

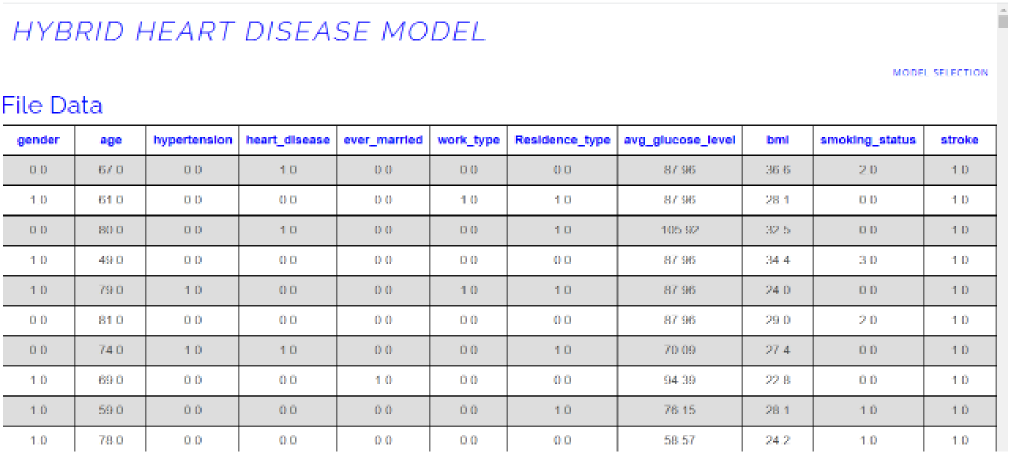
IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES

ISSN PRINT 2319 1775 Online 2320 7876

Research paper © 2012 IJFANS. All Rights Reserved, Volume 11, Iss 11, Nov 2022

Dataset Page

The screenshot of the dataset page is indicated in the below figure 5. After the successful uploading of the dataset, the uploaded dataset can be viewed in this page.



HYBRID HEART DISEASE MODEL

MODEL SELECTION

File Data

gender	age	hypertension	heart_disease	ever_married	work_type	Residence_type	avg_glucose_level	bmi	smoking_status	stroke
0.0	67.0	0.0	1.0	0.0	0.0	0.0	87.96	36.6	2.0	1.0
1.0	61.0	0.0	0.0	0.0	1.0	1.0	87.96	28.1	0.0	1.0
0.0	80.0	0.0	1.0	0.0	0.0	1.0	105.92	32.5	0.0	1.0
1.0	49.0	0.0	0.0	0.0	0.0	0.0	87.96	34.4	3.0	1.0
1.0	79.0	1.0	0.0	0.0	1.0	1.0	87.96	24.0	0.0	1.0
0.0	81.0	0.0	0.0	0.0	0.0	0.0	87.96	29.0	2.0	1.0
0.0	74.0	1.0	1.0	0.0	0.0	1.0	70.09	27.4	0.0	1.0
1.0	69.0	0.0	0.0	1.0	0.0	0.0	94.39	22.8	0.0	1.0
1.0	59.0	0.0	0.0	0.0	0.0	1.0	76.15	28.1	1.0	1.0
1.0	78.0	0.0	0.0	0.0	0.0	0.0	58.57	24.2	1.0	1.0

Figure.5. Screenshot of the dataset page of the cardiac disease prediction system

Model Selection Page

The screenshot of the model selection page is denoted in the below figure 6. Four different models for classifying the cardiac disease are added in the dropdown for selecting anyone classification model. On clicking the name of classification model, the presence of cardiac disease is predicted through the selected model and the respective accuracy value is calculated.

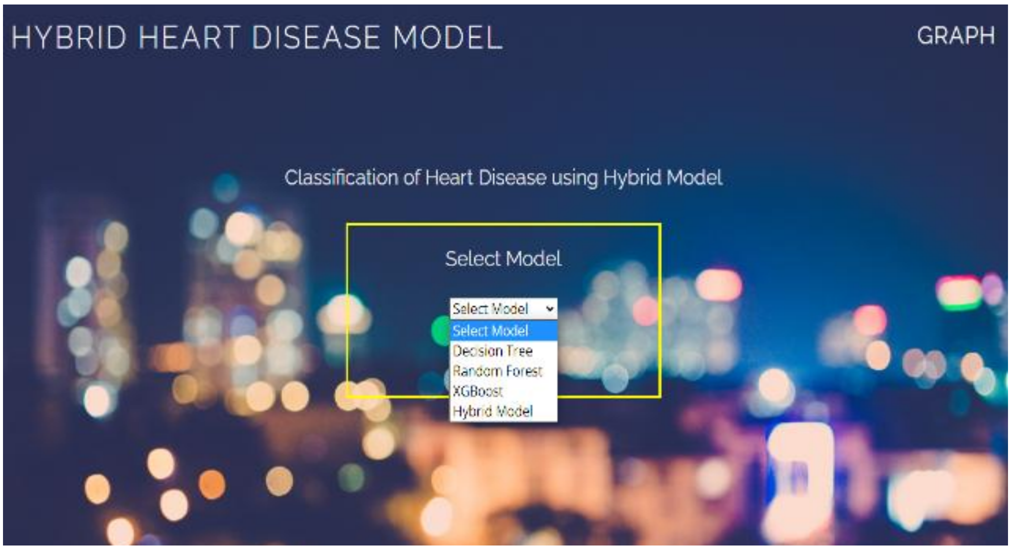


Figure.6. Screenshot of the home page of the cardiac disease prediction system

Decision Tree Accuracy Value

The screenshot of the decision tree accuracy value is implied in the below figure 7.

Figure.7. Screenshot of the accuracy value of decision tree technique

Random Forest Accuracy Value

The screenshot of the random forest accuracy value is conveyed in the below figure 8.

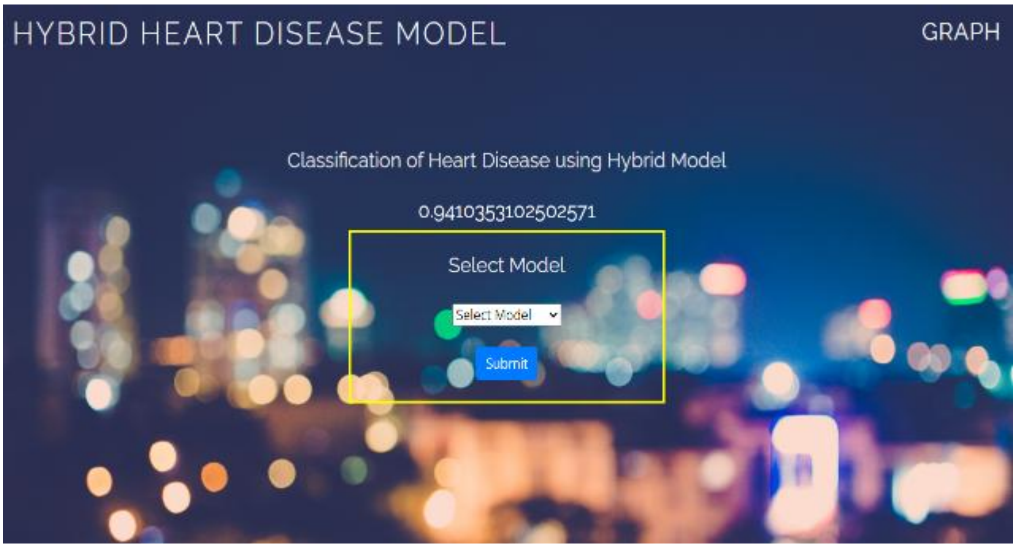


Figure.8. Screenshot of the accuracy value of random forest technique

XGBoost Accuracy Value

The screenshot of the XG boost accuracy value is represented in the below figure 9.

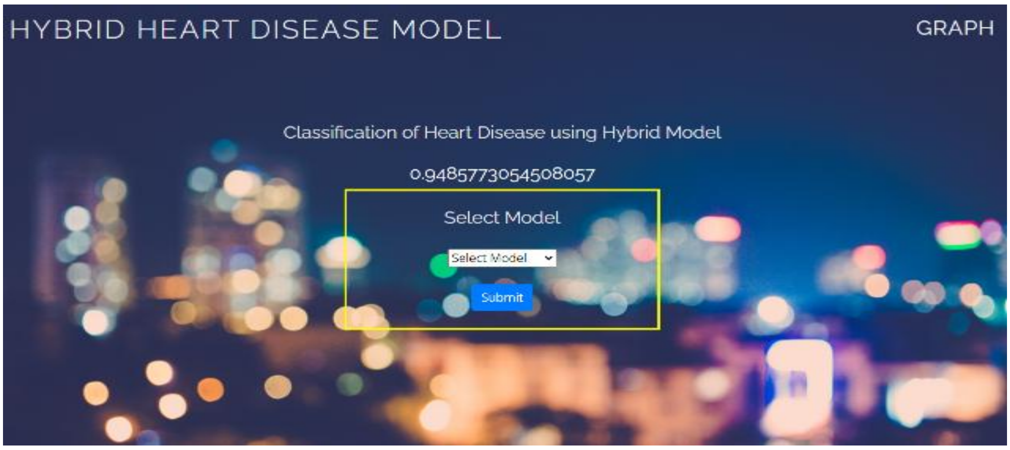


Figure.9. Screenshot of the accuracy value of XG boost technique

Hybrid Model Accuracy Value

The screenshot of the hybrid model accuracy value is expressed in the below figure 10.

Figure.10. Screenshot of the accuracy value of hybrid model

Prediction Page

The screenshot of the prediction page is specified in the below figure 11. After providing the required details in the respective fields, the presence of cardiac disease can be predicted in this page.

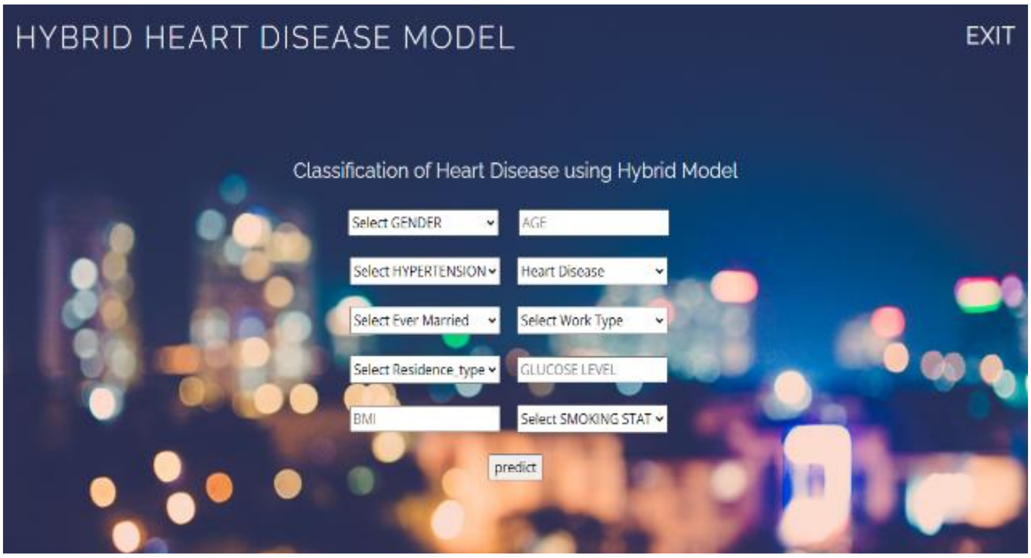


Figure.11. Screenshot of the prediction page of the cardiac disease prediction system

Representation of Output 1

The screenshot of the predicted output 1 is represented in the below figure 12. It can be seen from the figure that the output is predicted as “there is no chance to get stroke”.

ISSN PRINT 2319 1775 Online 2320 7876
Research paper **© 2012 IJFANS. All Rights Reserved, Volume 11, Iss 11, Nov 2022**

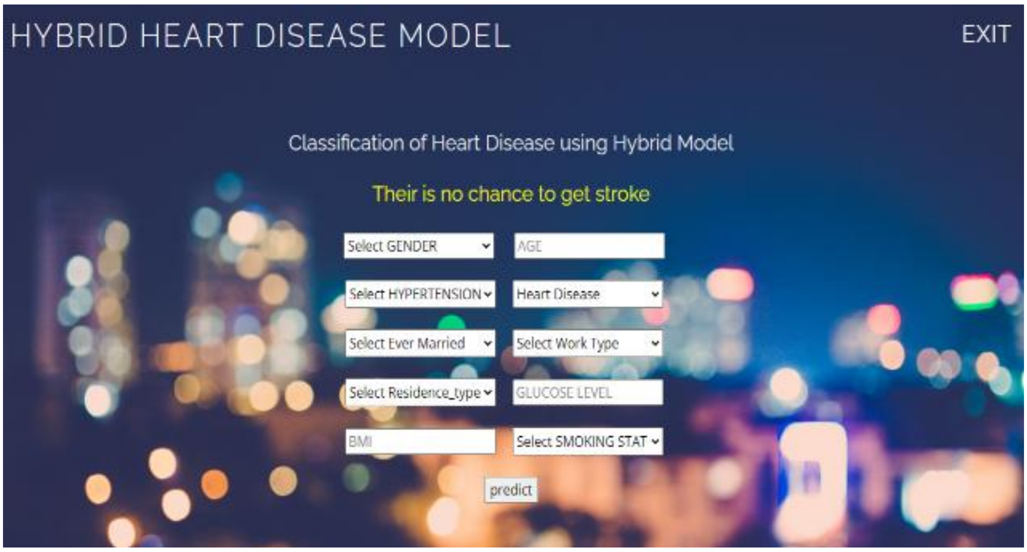


Figure.12. Representation of Output 1

Representation of Output 2

The screenshot of the predicted output 2 is represented in the below figure 13. It can be seen from the figure that the output is predicted as “there is a chance to get stroke”.

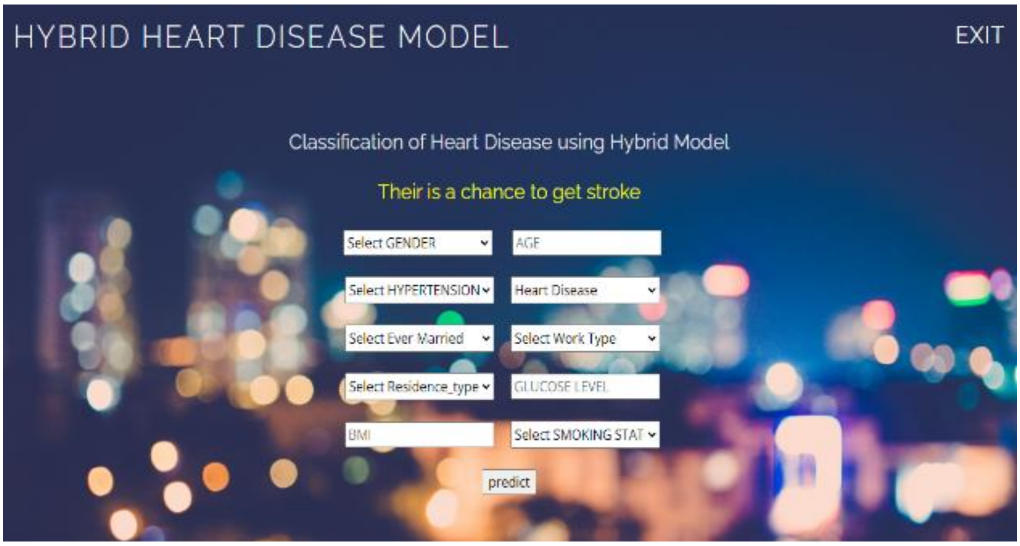


Figure.13. Representation of Output 2

Comparison graph

The graphical comparison of the obtained accuracy values is implied in the below figure 14.

ISSN PRINT 2319 1775 Online 2320 7876
Research paper Volume 11, Iss 11, Nov 2022

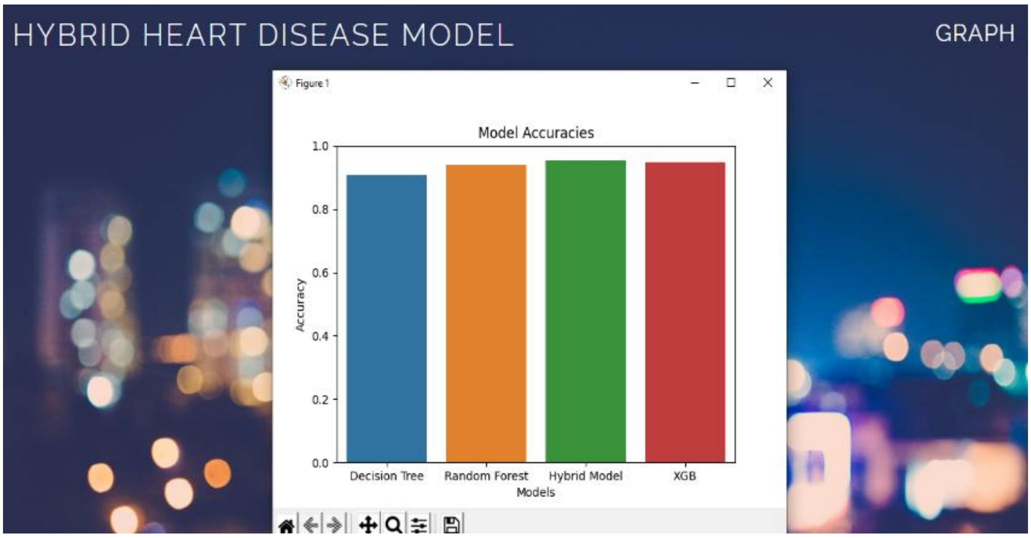


Figure.14. Graph Comparison of Accuracy Values

In the comparison graph of the "Hybrid Machine Learning Classification Technique for Improved Accuracy of Heart Disease," various algorithms were evaluated on their predictive accuracy. Decision Trees achieved a commendable 90.70% accuracy, while XGBoost outperformed with an accuracy of 94.85%. Random Forest followed closely with 94.10% accuracy. However, the Stacking Classifier demonstrated the highest performance at 95.30% accuracy, showcasing its ability to synergize multiple algorithms and achieve superior predictive accuracy for heart disease classification. This comparison underscores the efficacy of hybrid approaches in refining accuracy and aiding informed medical decisions.

Conclusion

In this research, we have developed a cardiac disease prediction framework using hybrid (stacking) model along with the general machine learning methods like Decision Tree, Random Forest, and XGBoost. Having understood the crucial role played by the selection of the models, we have tuned our heart disease-associated risk prediction to ensure the identification of stroke-based risks in human beings. Finally, the performance validation was done and outputs were compared along with the screenshots of the entire implementations of our framework.

References

1. Marimuthu M, Abinaya M, Hariesh K, Madhankumar K, Pavithra VJIJoCA. A review on heart disease prediction using machine learning and data analytics approach. 2018;181(18):20-5.
2. Learning MJAiCS, Technology. Heart disease diagnosis and prediction using machine learning and data mining techniques: a review. 2017;10(7):2137-59.
3. Krishnaiah V, Narsimha G, Chandra NSJIJoCA. Heart disease prediction system using data mining techniques and intelligent fuzzy approach: a review. 2016;136(2):43-51.
4. Adhikari NCD, Alka A, Garg R, editors. Hpps: Heart problem prediction system using machine learning. International Conference on Heart Problem Prediction System using Machine Learning (HPPS); 2017.
5. Ahsan MM, Siddique ZJAIiM. Machine learning-based heart disease diagnosis: A systematic literature review. 2022;128:102289.
6. Manjurul Ahsan M, Siddique ZJae-p. Machine Learning-Based Heart Disease Diagnosis: A Systematic Literature Review. 2021:arXiv: 2112.06459.
7. Javeed A, Khan SU, Ali L, Ali S, Imrana Y, Rahman AJC, et al. Machine learning-based automated diagnostic systems developed for heart failure prediction using different types of data modalities: A systematic review and future directions. 2022;2022.

8. Pal D, Mandana K, Pal S, Sarkar D, Chakraborty CJK-BS. Fuzzy expert system approach for coronary artery disease screening using clinical parameters. 2012;36:162-74.
9. Paul AK, Shill PC, Rabin MRI, Murase KJAI. Adaptive weighted fuzzy rule-based system for the risk level assessment of heart disease. 2018;48:1739-56.
10. Tsipouras MG, Exarchos TP, Fotiadis DI, Kotsia AP, Vakalis KV, Naka KK, et al. Automated diagnosis of coronary artery disease based on data mining and fuzzy modeling. 2008;12(4):447-58.
11. Arabasadi Z, Alizadehsani R, Roshanzamir M, Moosaei H, Yarifard AAJCm, biomedicine pi. Computer aided decision making for heart disease detection using hybrid neural network-Genetic algorithm. 2017;141:19-26.
12. Nahar J, Imam T, Tickle KS, Chen Y-PPJEswa. Association rule mining to detect factors which contribute to heart disease in males and females. 2013;40(4):1086-93.
13. Nahar J, Imam T, Tickle KS, Chen Y-PPJEswa. Computational intelligence for heart disease diagnosis: A medical knowledge driven approach. 2013;40(1):96-104.
14. Rehman A, Khan A, Ali MA, Khan MU, Khan SU, Ali L, editors. Performance analysis of pca, sparse pca, kernel pca and incremental pca algorithms for heart failure prediction. 2020 International Conference on Electrical, Communication, and Computer Engineering (ICECCE); 2020: IEEE.
15. Liu N, Lin Z, Cao J, Koh Z, Zhang T, Huang G-B, et al. An intelligent scoring system and its application to cardiac arrest prediction. 2012;16(6):1324-31.
16. Sree SV, Ghista DN, Ng K-HJJoMiM, Biology. Cardiac arrhythmia diagnosis by HRV signal processing using principal component analysis. 2012;12(05):1240032.
17. Wiharto W, Kusnanto H, Herianto HJHir. Intelligence system for diagnosis level of coronary heart disease with K-star algorithm. 2016;22(1):30-8.
18. Exarchos KP, Carpegianni C, Rigas G, Exarchos TP, Vozzi F, Sakellarios A, et al. A multiscale approach for modeling atherosclerosis progression. 2014;19(2):709-19.
19. Mythili T, Mukherji D, Padalia N, Naidu AJIJoCA. A heart disease prediction model using SVM-decision trees-logistic regression (SDL). 2013;68(16).
20. Diwakar M, Tripathi A, Joshi K, Memoria M, Singh PJMTP. Latest trends on heart disease prediction using machine learning and image fusion. 2021;37:3213-8.
21. Guo A, Pasque M, Loh F, Mann DL, Payne PRJCER. Heart failure diagnosis, readmission, and mortality prediction using machine learning and artificial intelligence models. 2020;7:212-9.
22. Watson DS, Krutzinna J, Bruce IN, Griffiths CE, McInnes IB, Barnes MR, et al. Clinical applications of machine learning algorithms: beyond the black box. 2019;364.
23. Cook NR, Ridker PMJJim. Further insight into the cardiovascular risk calculator: the roles of statins, revascularizations, and underascertainment in the Women's Health Study. 2014;174(12):1964-71.
24. Kakadiaris IA, Vrigkas M, Yen AA, Kuznetsova T, Budoff M, Naghavi MJJotAHA. Machine learning outperforms ACC/AHA CVD risk calculator in MESA. 2018;7(22):e009476.
25. Dimopoulos AC, Nikolaidou M, Caballero FF, Engchuan W, Sanchez-Niubo A, Arndt H, et al. Machine learning methodologies versus cardiovascular risk scores, in predicting disease risk. 2018;18:1-11.
26. Alaa AM, Bolton T, Di Angelantonio E, Rudd JH, Van der Schaar MJPo. Cardiovascular disease risk prediction using automated machine learning: A prospective study of 423,604 UK Biobank participants. 2019;14(5):e0213653.
27. Rose SJIJoE. Intersections of machine learning and epidemiological methods for health services research. 2020;49(6):1763-70.
28. Zhao Y, Wood EP, Mirin N, Cook SH, Chunara RJAjopm. Social determinants in machine learning cardiovascular disease prediction models: a systematic review. 2021;61(4):596-605.
29. Warner HR, Toronto AF, Veasy LGJAotNYAoS. Experience with Baye's theorem for computer diagnosis of congenital heart disease. 1964;115(2):558-67.

ISSN PRINT 2319 1775 Online 2320 7876
Research paper © 2012 IJFANS. All Rights Reserved, Volume 11, Iss 11, Nov 2022

30. Toronto AF, Veasy LG, Warner HRJPicd. Evaluation of a computer program for diagnosis of congenital heart disease. 1963;5(4):362-77.
31. Ferrari A, Micucci D, Mobilio M, Napoletano PJIA. On the personalization of classification models for human activity recognition. 2020;8:32066-79.
32. Shah SJJJoctr. Precision medicine for heart failure with preserved ejection fraction: an overview. 2017;10(3):233-44.
33. Shah SJ, Katz DH, Selvaraj S, Burke MA, Yancy CW, Gheorghiade M, et al. Phenomapping for novel classification of heart failure with preserved ejection fraction. 2015;131(3):269-79.
34. Shameer K, Johnson KW, Glicksberg BS, Dudley JT, Sengupta PPJH. Machine learning in cardiovascular medicine: are we there yet? 2018;104(14):1156-64.
35. Sanchez-Martinez S, Duchateau N, Erdei T, Kunszt G, Aakhus S, Degiovanni A, et al. Machine learning analysis of left ventricular function to characterize heart failure with preserved ejection fraction. 2018;11(4):e007138.
36. Przewlocka-Kosmala M, Marwick TH, Dabrowski A, Kosmala WJJotASoE. Contribution of cardiovascular reserve to prognostic categories of heart failure with preserved ejection fraction: a classification based on machine learning. 2019;32(5):604-15. e6.
37. Diller G-P, Kempny A, Babu-Narayan SV, Henrichs M, Brida M, Uebing A, et al. Machine learning algorithms estimating prognosis and guiding therapy in adult congenital heart disease: data from a single tertiary centre including 10 019 patients. 2019;40(13):1069-77.
38. Thompson WR, Reinisch AJ, Unterberger MJ, Schriefl AJJPc. Artificial intelligence-assisted auscultation of heart murmurs: validation by virtual clinical trial. 2019;40:623-9.
39. Nirschl JJ, Janowczyk A, Peyster EG, Frank R, Margulies KB, Feldman MD, et al. A deep-learning classifier identifies patients with clinical heart failure using whole-slide images of H&E tissue. 2018;13(4):e0192726.
40. Mathur P, Srivastava S, Xu X, Mehta JLJCMIC. Artificial intelligence, machine learning, and cardiovascular disease. 2020;14:1179546820927404.
41. Verma A, Mehta S, editors. A comparative study of ensemble learning methods for classification in bioinformatics. 2017 7th International Conference on Cloud Computing, Data Science & Engineering-Confluence; 2017: IEEE.
42. Hastie T, Tibshirani R, Friedman JH, Friedman JH. The elements of statistical learning: data mining, inference, and prediction: Springer; 2009.
43. Dietterich TGJMl. An experimental comparison of three methods for constructing ensembles of decision trees: Bagging, boosting, and randomization. 2000;40:139-57.
44. Saraswat D, Bhattacharya P, Verma A, Prasad VK, Tanwar S, Sharma G, et al. Explainable AI for healthcare 5.0: opportunities and challenges. 2022.
45. Barandela R, Valdovinos RM, Sánchez JSJPA, Applications. New applications of ensembles of classifiers. 2003;6:245-56.
46. Patel J, TejalUpadhyay D, Patel SJHD. Heart disease prediction using machine learning and data mining technique. 2015;7(1):129-37.
47. Wang Z, Zhu Y, Li D, Yin Y, Zhang JJCm, biomedicine pi. Feature rearrangement based deep learning system for predicting heart failure mortality. 2020;191:105383.
48. Mortazavi BJ, Downing NS, Bucholz EM, Dharmarajan K, Manhapra A, Li S-X, et al. Analysis of machine learning techniques for heart failure readmissions. 2016;9(6):629-40.
49. Pławiak P, Acharya URJNC, Applications. Novel deep genetic ensemble of classifiers for arrhythmia detection using ECG signals. 2020;32(15):11137-61.
50. Dutta A, Batabyal T, Basu M, Acton STJESwA. An efficient convolutional neural network for coronary heart disease prediction. 2020;159:113408.
51. Mishra SJJoE, Electromedical Engineering,, Informatics M. A comparative study for time-to-event analysis and survival prediction for heart failure condition using machine learning techniques. 2022;4(3):115-34.
52. Murray CJ, Lopez ADJTl. Mortality by cause for eight regions of the world: Global Burden of Disease Study. 1997;349(9061):1269-76.

ISSN PRINT 2319 1775 Online 2320 7876

53. Murray CJ, Lopez ADJTl. Global mortality, disability, and the contribution of risk factors: Global Burden of Disease Study. 1997;349(9063):1436-42.
54. Gaziano TA, Bitton A, Anand S, Abrahams-Gessel S, Murphy AJCpic. Growing epidemic of coronary heart disease in low-and middle-income countries. 2010;35(2):72-115.
55. Ahmed MR, Zhang Y, Feng Z, Lo B, Inan OT, Liao HJIribe. Neuroimaging and machine learning for dementia diagnosis: recent advancements and future prospects. 2018;12:19-33.
56. Mohan K, AL-Mamari YFS, AL-Najadi MAMJIJoDI, Computing I. Techniques of Machine Learning for Detecting Heart Failure. 2023;2(2):47-54.
57. M. S. Raja, M. Anurag, C. P. Reddy and N. R. Sirisala, "Machine Learning Based Heart Disease Prediction System," 2021 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, 2021, pp. 1-5, doi: 10.1109/ICCCI50826.2021.9402653.

58. Nadainamani RG, Reyana A, Kautish S, Vibith A, Gupta Y, Abdelwahab SF, et al. Clinical data analysis for prediction of cardiovascular disease using machine learning techniques. 2022;2022.