Optimizing Coronary Artery Disease Prognosis with a Dual-Class Boosted Decision Trees Approach

Mohd Arbaz Mazharuddin
PG Scholar, Department of Computer Science and Engineering, Shadan College of Engineering and Technology, Hyderabad, Telangana,
India – 500086. Email: arbazmohd.ac@gmail.com

Subramanian K.M
Professor, Department of Computer Science and Engineering, Shadan College of Engineering and Technology, Hyderabad, Telangana,
India – 500086. Email: kmsubbu.phd@gmail.com

Md. Ateeq Ur Rahman
Professor, Department of Computer Science and Engineering, Shadan College of Engineering and Technology, Hyderabad, Telangana,
India – 500086. Email: mail_to_ateeq@yahoo.com

Abstract: The better lifespan and growth of the population has resulted in an increased prevalence of stable CAD which has placed a great burden on the healthcare systems. The introduction of ML models has been identified as a tool which can play an important role in predicting individual risk factors and allow early interventions and lighten the burden for patients and their families. These algorithms can be used to forecast the number of hospitalizations, surveillance of people at risk and to help optimising treatment programmes. In this paper, we present a robust dual-class BDT approach towards CAD prognosis, where we use advanced ML tools to analyze clinical data, medical imaging, genetic markers and lifestyle factors. We proposed a feature selection method using a RFECV with the random forest to reduce the data set. All models were trained on the Heart Disease dataset, and the Voting Classifier (Bagging Classifier with Random Forest and Decision Tree) was able to achieve an accuracy rate of 100%. Ensemble methods can be used to improve the prediction of the onset and progression of CHD, as shown here. By leveraging this approach, AI can significantly contribute to healthcare advancements and showcase the potential of ML models in delivering high-performing solutions for personalized patient care and early disease detection.

***"Index Terms -** Health issue, coronary heart disease, two-class LR, two-class NN, two-class DJ, two-class BDT ervices."*

I. INTRODUCTION

One of the most important organs of the body is the human heart because it circulates blood around a body and enables it to carry out vital physiological functions. It has four chambers and coronary arteries are directly connected to the main artery of the body, which is the aorta, that originates from the left ventricle. An aortic valve leads into the coronary arteries, which branch into right and left coronary arteries that carry oxygenated blood to the myocardium. CVDs like a heart attack are caused by the accumulation of plaque in blood vessels which causes the narrowing of the arteries, which in turn affects blood flow all over the body [1, 2]. Of these ailments, CHD is the most common and is one of the major causes of death in the world. Cardiovascular diseases kill > 17 million people each year in the world, and CHD alone is responsible for > 240,000 deaths in men and 76,000 deaths in women in Europe prior to 65 years of age [3]. An estimated 3.8 million men and 3.4 million women die of CHD each year worldwide [4]. Although the economic costs of heart disease have declined in recent years, the impact was still substantial: $219.6 billion in 2017, comprising healthcare services, drugs and lost productivity [5].

Recently, AI has shown great potential in a wide range of areas including healthcare, autonomous driving and gaming [6]. ML algorithms are at the heart of AI applications, including SVMs, ANNs, DTs, NB, and KNN [7]. Each method has its own pros and limitations that can be utilized in different medical activities such as illness screening, risk assessment, prediction and decision support [8]. There is a growing focus on cardiovascular illnesses, one of the number one causes of death globally as identified by the WHO, resulting to a large volume of research. The accurate diagnostic techniques and preventive methods have tremendous potential for reducing mortality from CAD. The need of early detection is emphasised as conventional diagnostic tools such as coronary angiography are intrusive and have associated hazards [9]. The challenge has been met by ML, which offers the potential to develop non-invasive, efficient and accurate tools that allow for early CAD identification, risk classification and treatment planning. The shift is indicative of the growing importance of AI in combating cardiovascular diseases on a global scale.

II. RELATED WORK

ML has become an indispensable tool in the healthcare industry, especially in the prediction and diagnosis of CAD. In recent years, many studies have investigated several ML methods to predict CAD with better accuracy and efficiency and this indicates an increasing interest in utilizing AI-driven solutions in healthcare. These approaches employ multiple ML algorithms, multiple feature selection methods, and multiple ensemble methods that further contribute to better CAD diagnosis and risk assessment.

Hassan et al. explored the effective prediction of CAD using various ML algorithms, in conjunction with feature selection algorithms, in [10]. They proved the implementation of feature selection enhances the prediction performance by decreasing overfitting and raising the interpretability of the model. They demonstrated that the application of feature selection methods like RFE significantly improves the performance of the classifiers, like RF and SVM and makes them more suitable for CAD prediction. They emphasized the importance of optimizing the features, as including irrelevant features might decrease the accuracy of the model and complicate it. Finally, the study has proven that ML models, if well calibrated and adapted with efficient feature selection strategies, can play a significant role in prediction of CAD.

Devi et al. [11] also made a significant contribution in the prediction of CAD by applying the ML approaches for predicting the coronary artery disease in a clinical environment. The researchers employed multiple ML models such as DT, KNN, and SVM to predict CAD using patient data. The research demonstrated the flexibility of ML algorithms and their ability to handle diverse data types and highlighted their potential in healthcare applications in general. They tried different training sets on various traits such as demographic and clinical history and lifestyle traits to improve prediction. Furthermore, their study pointed to the importance of integrating multiple data sources and using complex algorithms to greatly improve diagnostic performance, especially in CAD patients, where early intervention is critical.

Gonsalves et al., in 2019 [12] conducted an experimental study for predicting coronary heart disease using various ML techniques. The authors have done a comparative assessment of common ML approaches like ANNs, DTs and NB. The study revealed that each of the algorithms had its own strengths, but the results of the ensemble methods were found to be more accurate and robust. Ensemble models like boosting and bagging were shown to be effective for tackling complex medical data, as they reduce the variance and bias, leading to improved predictive performance [13]. The authors highlighted that the choice of the model and the data quality are crucial for the performance of ML models and require the use of high-quality data. Their results highlighted that a good model with the correct collection of variables and data pre-processing procedures can beat traditional diagnostic methods in predicting CAD risk.

To extend the use of ensemble learning for CAD prediction, Puneet et al. [14] applied Voting Classifier. In this study, we employed the ensemble of multiple base classifiers (DT and RF) in order to predict the CHD. The results showed that the ensemble model had better performance than individual classifiers with higher accuracy and generalization capabilities. The Voting Classifier leverages the combined strength of different models to limit the chances of overfitting and to improve the overall prediction performance. The technique was able to obtain the prediction accuracy and provided a more solid construction of a real-world application, where each model often suffers from problems related to the presence of different data sources and poor datasets. The authors of the study determined that ensemble learning methods, such as Voting Classifiers, are extremely effective for predicting complex diseases, such as CAD, for which some classifiers may not be able to discover all the patterns in the data [15].

A more sophisticated LightGBM model to predict CAD was assessed by Yang et al. [16] and compared with other popular ML models. The results of the study showed that gradient boosting models such as LightGBM outperform the traditional classifiers in terms of computing efficiency and accuracy of the predictions. The authors tested this model with a large number of CAD patients and demonstrate that this model can be used to forecast the likelihood of the CAD onset with high accuracy. The results showed that the advanced boosting techniques like LightGBM could be used effectively for large-scale, high-dimensional datasets, making it a valuable tool for the CAD prediction. Furthermore, the study emphasized on the significance of hyperparameter tuning in enhancing the performance of the model, with the fine tuning process helping the model to learn the peculiarities of the data and deliver better results. Yang et al. contribute to an increasing amount of research that shows the strength of advanced boosting methods to handle the difficulties of complicated medical data.

Dutta et al. [17] proposed the use of CNNs for the prediction of CHD is another important addition.

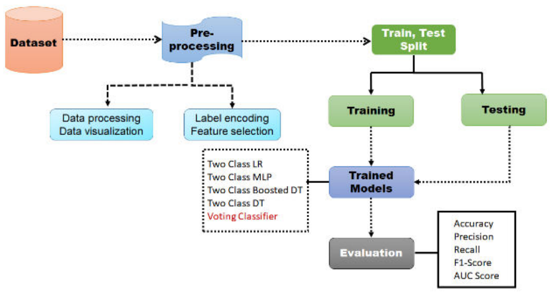
Structured medical data was employed for CAD prediction by CNNs which are widely used for image processing. The study revealed that CNNs are able to learn from raw data efficiently and identify complex patterns that are typically difficult to learn using conventional ML techniques. The authors have trained their model with a huge amount of data and found that the CNN model has significantly outperformed other models in terms of prediction accuracy. This was very helpful when dealing with the interactions between the attributes of the data, which, in the case of medical data, is of high demand. The use of CNNs for CAD prediction opens new ways to use DL methods in the field of healthcare for diagnosis and risk stratification.

In the study of Al-Ssulami et al. [18] they evaluated the impact of new data augmentation methods on Coronary Heart Disease prediction with ML. They used data augmentation together with conventional ML methods to increase the diversity and amount of training data, which is important in healthcare where labeled data is often scarce. They were able to construct a more robust model that could handle small or imbalanced samples by producing synthetic data and utilizing augmentation techniques including rotation, scaling and translation. The method overcame the issue of limited access to medical data and offered a technique for improving the generalizability and accuracy of CAD prediction models . The study highlighted the necessity of data augmentation to deal with data scarcity, a typical problem in medical research, particularly in the case of uncommon disorders like CAD.

In terms of research in CAD prediction applications, ML has come a long way, developing new algorithms and approaches with the years. From the traditional classifiers such as DT and Naïve Bayes classifiers, the range of CAD prediction solutions has expanded to include more sophisticated approaches like ensemble methods and DL techniques. The future of CAD diagnostics lies in the potential of ML to provide tools that can process large and complex sets of data, identify trends that are not apparent in the data, and produce more accurate and timely predictions [19]. Future work in this field will likely continue to focus on optimization of existing models, inclusion of new characteristics and improvement of data quality to ensure that ML models are able to provide solid predictions in clinical contexts. In addition, the incorporation of AI and ML in health care systems can contribute to lowering health care expenditure, enhancing patient results, and aiding health care staff in making more effective choices, which can result in optimized health care system management and prevention of coronary heart disease [20].

III. MATERIALS AND METHODS

The approach proposed will lie on enhancing the prognosis of CAD through the development of state-of-the-art ML models, based on real world clinical data. This will involve incorporating several data sources such as medical imaging, genetic markers, lifestyle and environmental factors to create a comprehensive risk assessment framework. The 4 new models will be implemented: two-class Logistic Regression, two-class MLP, two-class DT and two-class Boosted Decision Trees. An ensemble Voting Classifier will also be implemented to boost the accuracy of the prediction, by utilizing the strength of different algorithms (Bagging Classifier with RF + DT). The device will aid early detection of CAD and early monitoring of high risk patients, thereby optimizing treatment strategies that are patient-specific and based on individualized risk profiles. The system intends to provide a reliable decision-making support system for the healthcare providers to improve the patients' management and minimize the hospitalization rate and consequently the quality of care of the CAD patients.



"Fig.1 Proposed Architecture"

Figure 1 is a representation of the ML process for a two-class classification problem. It begins with data processing and visualization then label encoding and feature selection. The data is then divided into two parts: training and testing. A set of classifiers, such as LR, MLP, Boosted Decision Tree and DT are trained on the training data. A voting model of classifiers is used to combine the predictions from these individual classifiers. The trained model is evaluated against performance metrics such as accuracy, precision, recall, F1 score and AUC score on the testing set.

A) Dataset Collection:

For this study, the Heart Disease dataset [17] used is the entire dataset with 1190 cases and 14 characteristics. Age, gender, type of chest pain (cp), resting blood pressure (trestbps), cholesterol (chol), fasting blood sugar (fbs), resting electrocardiographic results (restecg), maximum heart rate (thalach), exercise induced angina (exang), oldpeak, slope of peak exercise ST segment (slope), number of major vessels colored by fluoroscopy (ca), thalassemia (thal), and the target variable indicating the presence or absence of CHD. Dataset is publicly available and have been used to analyze features and to visualize data distribution and correlations in CAD patients.

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	targe
0	63	1	0	145	233	1	2	150	0	2.3	2	0	2	
1	67	1	3	160	286	0	2	108	1	1.5	1	3	1	
2	67	1	3	120	229	0	2	129	1	2.6	1	2	3	
3	37	1	2	130	250	0	0	187	0	3.5	2	0	1	
4	41	0	1	130	204	0	2	172	0	1.4	0	0	1	

“Fig.2 Dataset Collection”

B) Processing:

Pre-processing is a key step in the process of preparing the data set for model training. This includes data cleaning, visualization of the features, encoding the labels, selecting the best features and data, and ensuring that the model does well and makes the correct predictions.

a) Data Processing: Data processing is a crucial procedure that helps in arranging the data set for analysis. Data cleaning is carried out to address missing data and inconsistencies, first. For simplicity, we remove irrelevant columns that are not useful for the analysis from the dataset. Then the data of interest for model training is picked by separating the independent variables (X) and the target variable (y). The aim of these processes is to ensure the accurate modeling of the dataset while storing only meaningful information to make predictions.

b) Data Visualization: Data visualization can be used to understand how and where the features in the data relate to each other. A correlation matrix is utilized to explore correlations among several variables, to find out if any variables are strongly correlated that may impact model performance. Sample results are also displayed to give a better idea of the distribution of the target variable and to detect if there are any imbalances and trends within the data. The intuitive insight of the data given by visualization approaches is crucial for informed decisions and optimization of models.

c) Label Encoding: The data are processed into a numerical format by encoding the categorical data with labels, to allow the ML models to process the data efficiently. Categorical variables such as ‘sex’, ‘cp’, ‘restecg’, ‘exang’ and ‘thal’ in this dataset are converted into numeric labels. Each category is assigned a distinct number in this way to enable the models to interpret the data accurately. Label encoding is crucial if you want to map non-numeric columns to a format that can be processed by ML models without any mistakes or misinterpretations.

d) Feature Selection: It is crucial to perform a feature selection to get the best out of the model with the most relevant features. For this study, RFECV (Recursive Feature Elimination with Cross-Validation) was used with RF (Random Forest) to do feature selection. The aim of RFECV is to eliminate one characteristic at a time and test the performance of the model after the elimination to obtain the optimal subset of features that are most important for predicting CHD. With this approach it is possible to eliminate overfitting, increase the quality of the model and optimize the computing efficiency by selecting only the most important variables.

C) Training and Testing:

One of the key steps to assessing model performance is to partition the data into training and test sets. In this research, the data set is divided into two parts, 80% is used for training set and 20% is used for test set. The model is trained with the training set, and it assists the model to learn the patterns and relationships that are contained in the data. Generalization capability of the model is evaluated by testing it with the test set which the model has not encountered before. The purpose of this split is to ensure that the model is tested on data it has not seen before and for a suitable measure.

D) Algorithms:

Two-class Logistic Regression [18] is a statistical model for binary classification. It is used to estimate the probability of the provided input to be a member of a particular class by using a logistic function. The technique has been applied to our CAD prognosis research to identify risk factors associated with the coronary artery disease in order to classify patients based on their risk of developing the disease. Easy to use and easy to understand, and can be used as a good initial evaluation and risk classification instrument.

Two-Class Multi-Layer Perceptron (MLP) is a neural network that can be used for binary classification problems. It consists of an input layer, one or more hidden layers and an output layer. In our work, [20] MLP is used to learn non-linear relationships between features, by identifying complex patterns in the dataset. It is crucial for accurate forecasting of the likelihood of coronary artery disease, and thus early intervention opportunities and patient risk assessment based on patient profile.

All of the units in the MLP are sigmoid function as Activation Function

f (x) = 1 1 + exp (- x) (1)

and is the default activation function that is used in feedforward neural networks.

Boosted Decision Trees (2-class) A combination of several DT is used to improve the prediction. One tree corrects the errors of the preceding tree. This ensemble method has shown to perform better on accuracy and robustness for classification problems. As part of our [19] CAD prognosis research, boosted decision trees are considering a wide variety of patient data variables to achieve a more powerful model for the prediction of the onset and progression of coronary artery disease without sacrificing efficiency. Flexibility and performance play important roles in their treatment in high-risk patients and in optimizing treatment strategies.

Two-Class Decision Tree is a decision tree used for splitting the data into two classes based on the values of the features. This technique is useful in our project to understand and illustrate the impact of many risk factors for CAD. DT are explainable and therefore useful for doctors to identify key factors which impact patient health. They are also used as a reference model for comparison with better algorithms to predict CAD.

The expected value of each outcome can be computed if you know the cost of each outcome, and the probability that it will occur in the following way:

“Expected value (EV) = (First possible outcome x Likelihood of outcome) + (Second possible outcome x Likelihood of outcome) – Cost (2)”

For obtaining high classification accuracy, Voting Classifier is a combination of several models including RF and DT. An ensemble of both algorithms can make use of the strength of both algorithms and yield a more accurate estimation of the risk of coronary artery disease. In our project, the Voting Classifier helps to improve the performance of each model by combining their outputs which in turn reduces the chances of overfitting and gives good predictions to assist in patient evaluation and treatment decisions.

IV. RESULTS AND DISCUSSION

Accuracy: The level of accuracy of a test is related to how well it can distinguish between patient and healthy cases. Test accuracy can be measured by the percentage of positive and negative test results that are correct in all test instances. Mathematically:

$$Accuracy = \frac{\text{TP + TN}}{\text{TP + FP + TN + FN}} \quad (3)$$

Precision: Precision indicates the percentage of positive instances (or samples) that are accurately classified. The formula for computing precision is:

$$Precision = \frac{\text{True Positive}}{\text{True Positive } + \text{ False Positive}} \quad (4)$$

Recall: Recall in ML is the percentage of the relevant instances of a class that is correctly identified by the model. It is a measure of the accuracy of a model that predicts a positive instance of a class by comparing the number of positive instances correctly predicted by the model with the number of observed positive instances.

$$Recall = \frac{\text{TP}}{\text{TP + FN}} \quad (5)$$

F1-Score: The F1 score is used to measure the accuracy of the ML model. Combination of precision and recall scores of model. Accuracy statistic: measures the percentage of times a model is correct for the data.

$$F1\ Score = 2 * \frac{Recall \text{ X } Precision}{Recall + Precision} * 100 \quad (6)$$

AUC_Score: AUC-ROC Curve is a performance measure of a classification problem at different threshold levels. ROC is a plot of the TPR against the FPR. Overall ability of the model to discriminate a class is given by the AUC. The higher the AUC , the better the model is doing .

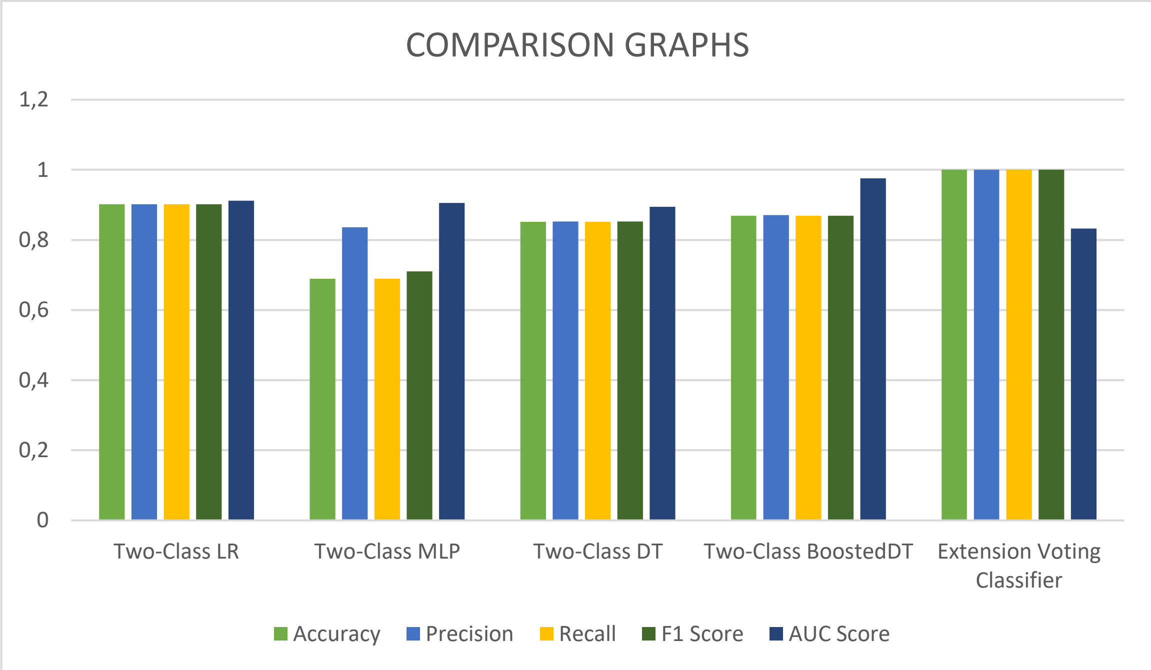
$$AUC = \sum_{i=1}^{n-1} (FPR_{i+1} - FPR_i) \cdot \frac{TPR_{i+1} + TPR_i}{2} \quad (7)$$

The accuracy, precision, recall, F1 score and AUC score of each method are illustrated in Table 1. Voting Classifier shows the best results having an accuracy of 100%. In addition, metrics of other algorithms are included in the table below for comparison.

“Table.1 Performance Evaluation Table”

Model	Accuracy	Precision	Recall	F1 Score	AUC Score
Two-Class LR	0.902	0.902	0.902	0.902	0.912
Two-Class MLP	0.689	0.836	0.689	0.710	0.906
Two-Class DT	0.852	0.853	0.852	0.853	0.895

Two-Class BoostedDT	0.869	0.871	0.869	0.869	0.976
Extension Voting Classifier	**1.000**	**1.000**	**1.000**	**1.000**	**0.833**



"Fig.3 Comparison Graphs – Classification"

The light green line in Fig 2 represents accuracy, light yellow recall, blue precision, green F1-score and dark blue AUC_Score. The Voting Classifier has the highest values for all measures when compared to the other algorithms and thus out performs the other algorithms. These facts are clearly illustrated by the above graph.

V. CONCLUSION

Last but not least, the suggested system is a great advance in the prognosis and treatment of CAD. We aim to provide healthcare practitioners with powerful tools for the early diagnosis and risk assessment with modern ML techniques. A combination of various data sources enables medical imaging, genetics and lifestyle to form a complete picture of individual risk profiles. Results of the models evaluation showed that Voting Classifier had the highest accuracy percentage of 100%. This amazing feat demonstrates the power of ensemble modeling techniques in enhancing prediction capabilities in healthcare settings. The suggested approach is designed to enhance patient outcomes and meets the growing demand on healthcare systems by enabling quick interventions and individualized treatment techniques. The prevalence of CAD is increasing and a powerful machine learning framework like this could be used to more effectively allocate resources and enhance the treatment of heart disease in general.

VI. Future Work

In future work, the research will be further enhanced by using sophisticated approaches such as DL, transfer learning and hybrid models. Also we will talk about using natural language processing for evaluating unstructured clinical notes and reinforcement learning for optimizing therapy suggestions. Also, the possibility to expand the research to other areas such as predictive analytics in chronic diseases and customized medicine will be explored, increasing its application and effectiveness in different healthcare settings.

REFERENCES

[1] T. G. Richardson, E. Sanderson, T. M. Palmer, M. Ala-Korpela, B. A. Ference, G. D. Smith, and M. V. Holmes, ''Evaluating the relationship between circulating lipoprotein lipids and apolipoproteins with risk of coronary heart disease: A multivariable Mendelian randomisation analysis,'' PLOS Med., vol. 17, no. 3, Mar. 2020, Art. no. e1003062.

[2] C.-Y. Chou, D.-Y. Hsu, and C.-H. Chou, ''Predicting the onset of diabetes with machine learning methods,'' J. Personalized Med., vol. 13, no. 3, p. 406, Feb. 2023.

[3] E. Wilkins, L. Wilson, K. Wickramasinghe, P. Bhatnagar, J. Leal, R. Luengo-Fernandez, R. Burns, M. Rayner, and N. Townsend, European Cardiovascular Disease Statistics 2017. Brussels, Belgium: European Heart Network, 2017.

[4] J. Mackay and G. A. Mensah, The Atlas of Heart Disease and Stroke. Geneva, Switzerland: World Health Organization, 2004.

[5] S. S. Virani et al., ''Heart disease and stroke statistics—2021 update: A report from the American Heart Association,'' Circulation, vol. 143, no. 8, pp. 535–552, 2021.

[6] Wanyonyi, M., Morris, Z. N., Musyoka, F. M., & Kitavi, D. M. (2025). Enhanced machine learning and hybrid ensemble

approaches for Coronary Heart Disease prediction. PLoS One, 20(12), e0328338.

[7] Wanyonyi, M., Morris, Z. N., Musyoka, F. M., & Kitavi, D. M. (2025). Enhanced machine learning and hybrid ensemble approaches for Coronary Heart Disease prediction. PLoS One, 20(12), e0328338.

[8] Aswini, K., & Arya, K. (2025). Optimizing Heart Disease Prediction: A Comparative Analysis of Tree-Based Ensembles With Feature Expansion and Selection. IEEE Access.

[9] Madhumathi[1], A., Karthikeyan, P., & Saravanakumar, P. (2026, May). Optimized Bi-GRU Models for Enhanced Heart Disease Prediction Through Data Mining. In Evolution in Signal Processing and Telecommunication Networks: Proceedings of Tenth International Conference on Microelectronics Electromagnetics and Telecommunications (ICMEET 2025), Volume 2 (Vol. 2, p. 358). Springer Nature.

[10] Manjurathi, B., Madhumathi, A., Karthikeyan, P., Saravanakumar, P., & Kannadhasan, S. (2025, October). Optimized Bi-GRU Models for Enhanced Heart Disease Prediction Through Data Mining Techniques. In International Conference on Microelectronics, Electromagnetics and Telecommunication (pp. 358-368). Cham: Springer Nature Switzerland.

[11] Rathinam, V., Valarmathi, K., Madhumathi, A., & Lalitha, S. D. (2026). A Hybrid Quantum-Inspired Deep Learning Framework with Bio-Inspired Optimization for Cardiovascular Disease Prediction. Journal of Electronics, Electromedical Engineering, and Medical Informatics, 8(2), 806-824.

[12] Aggarwal, A., Taneja, P., Arora, V., & Rao, S. (2026, February). HeartSense: Machine-Augmented Intelligence for Predicting Cardiovascular Disease from Structured Clinical Data with Patient Diagnostic Profiles. In 2026 IEEE Applied Sensing Conference (APSCON) (pp. 1-4). IEEE.

[13] Aggarwal, A., Taneja, P., Arora, V., & Rao, S. (2026, February). HeartSense: Machine-Augmented Intelligence for Predicting Cardiovascular Disease from Structured Clinical Data with Patient Diagnostic Profiles. In 2026 IEEE Applied Sensing Conference (APSCON) (pp. 1-4). IEEE.

[14] Rathinam, V., Valarmathi, K., Madhumathi, A., & Lalitha, S. D. (2026). A Hybrid Quantum-Inspired Deep Learning Framework with Bio-Inspired Optimization for Cardiovascular Disease Prediction. Journal of Electronics, Electromedical Engineering, and Medical Informatics, 8(2), 806-824.

[15] Hu, W., Wang, G., Ren, L., Hu, J., Wu, X., Zhuang, W., ... & Hu, W. (2026). Current trends and future directions of artificial intelligence in lung cancer diagnosis. Chinese Journal of Cancer Research, 38(2), 148.

[16] Sourav, M. M. I., Aman, N., Ali, M. I., Rana, M. J., Hasan, M. M., & Roy, A. (2025, July). An Ensemble Machine Learning Approach for Cardiovascular Disease Prediction Using Feature Selection and Gradient Boosting. In 2025 International Conference on Quantum Photonics, Artificial Intelligence, and Networking (QPAIN) (pp. 1-6). IEEE.

[17] Jaradat, M., & Awad, M. (2026). Explainable AI in cardiology diagnostics: a systematic review of machine learning, meta-heuristic optimization, and clinical text mining for coronary artery disease. International Journal of Medical Informatics, 106321.

[18] Rabcan, J., Levashenko, V., Zaitseva, E., & Kvassay, M. (2025). Advancing ECG signal classification with a fuzzy classifier approach. IEEE Access.

[19] Rathinam, V., Valarmathi, K., Madhumathi, A., & Lalitha, S. D. (2026). A Hybrid Quantum-Inspired Deep Learning Framework with Bio-Inspired Optimization for Cardiovascular Disease Prediction. Journal of Electronics, Electromedical Engineering, and Medical Informatics, 8(2), 806-824.

[20] Ali, M., Hashmi, M. U., Ahmad, Z., Kazmi, N. U. A., Ittfaq, A., & Ashraf, A. (2025). Hybrid Supervised Machine Learning Models for Enhanced Alzheimer's Disease Classification. Proceedings of the Pakistan Academy of Sciences: A. Physical and Computational Sciences, 62(4), 323-336.