

# Enhancing Heart Disease Diagnosis with Hybrid Machine Learning Models: A Case Study

**Ahmed Yousif Falih Saedi**

Computer Engineering Department, College of Engineering, Mustansiriyah University, Baghdad, Iraq  
ahmed.yousif@uomustansiriyah.edu.iq (corresponding author)

**Raof Talal Hussein**

Computer Engineering Department, College of Engineering, Mustansiriyah University, Baghdad, Iraq  
raoofatalal@uomustansiriyah.edu.iq

**Dhari Ali Mahmood Ghrairi**

Department of Computer Engineering, University of Technology Baghdad, Baghdad, Iraq  
dhari.a.mahmood@uotechnology.edu.iq

Received: 19 May 2025 | Revised: 10 June 2025, 28 June 2025, and 30 June 2025 | Accepted: 3 July 2025

Licensed under a CC-BY 4.0 license | Copyright (c) by the authors | DOI: <https://doi.org/10.48084/etasr.12260>

## ABSTRACT

Heart disease remains one of the leading causes of morbidity and mortality worldwide, prompting extensive research into accurate and early detection methods. Recent advancements have highlighted the critical role of Artificial Intelligence (AI) and Machine Learning (ML) in enhancing diagnostic precision. This study explores the effectiveness of hybrid ML models in diagnosing heart disease, specifically focusing on two novel combinations: Random Forest (RF) integrated with Sequential Minimal Optimization (SMO) and J48 decision trees augmented with Logistic Regression (LR). Using a comprehensive heart disease dataset, the models were evaluated based on their classification accuracy, class separability, and risk prediction capability. Both models incorporated advanced preprocessing, cross-validation, and hyperparameter optimization techniques. The RF-SMO model achieved an accuracy of 97% and a Receiver Operating Characteristic (ROC) area of 0.97, while the J48-LR model attained 92% accuracy with a ROC area of 0.96. These findings underscore the potential of hybrid ML approaches to enhance cardiac diagnostics, offering valuable tools for clinical decision support and the advancement of personalized healthcare.

**Keywords-random forest; sequential minimal optimization; machine learning; heart disease; J48; logistic regression**

## I. INTRODUCTION

In recent years, the field of cardiology has undergone a significant transformation through the integration of Machine Learning (ML) methodologies, revolutionizing both diagnostic and treatment strategies [1]. ML techniques enable the analysis of complex, high-dimensional datasets and uncover patterns that are not readily noticeable through traditional methods, thereby supporting more accurate diagnoses and personalized treatment strategies [2–4]. Given that cardiovascular diseases remain among the leading causes of global mortality, the application of ML in cardiology represents a substantial advancement toward improving patient healthcare through data-driven, intelligent decision-making frameworks [5, 6].

A primary focus of ML-based predictive analytics in cardiology is the early detection and prevention of heart disease

[7, 8]. By leveraging historical health records, patient-specific attributes, and real-time physiological metrics, predictive models can forecast the onset of cardiovascular conditions with high accuracy [9, 10]. Moreover, the incorporation of predictive ML models into clinical workflows has been shown to streamline operations, enhance diagnostic precision, and improve patient management [11–13].

Recent advances in ML research have emphasized Deep Learning (DL) and ensemble-based strategies, which exploit the representational power of Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) for analyzing temporal and image-based cardiac data, such as Electrocardiograms (ECGs) and echocardiograms [14–18]. These models have enabled improved feature extraction and classification of subtle cardiac anomalies. In parallel, unsupervised learning approaches have been explored to

identify hidden patterns or clusters in unlabeled cardiac datasets, revealing previously unrecognized subtypes or atypical cases of heart disease [19–21].

Emerging research has also investigated hybrid modeling frameworks, which combine classical ML techniques with DL architectures to enhance model generalization and robustness [22, 23]. For example, studies have proposed integrating deep neural networks with algorithms such as k-nearest neighbors for post-classification refinement, resulting in heightened sensitivity and specificity in detecting early-stage heart abnormalities [24, 25].

Despite significant progress, most existing studies have focused on individual algorithms, with limited exploration of hybrid models that effectively integrate the complementary strengths of distinct techniques. Such hybridization is particularly valuable for mitigating challenges posed by heterogeneous datasets, such as missing values, class imbalance, or noise, which often degrade the performance of single-model approaches.

This study aims to address these limitations by developing and evaluating two novel hybrid ML algorithms for heart disease prediction: Random Forest (RF) combined with Sequential Minimal Optimization (SMO) and J48 decision tree augmented with Logistic Regression (LR). These hybrid configurations are designed to exploit the ensemble strength of tree-based classifiers and the precision of statistical learning methods, with the goal of improving prediction accuracy, model robustness, and clinical applicability.

#### A. Overview of the Dataset

The dataset employed in this research was specifically curated to encompass a broad spectrum of variables associated with cardiovascular diseases. It was sourced from two prominent medical institutions: Ibn Al-Bitar Hospital and Ibn Al-Nafis Hospital. The dataset includes records from over 200 patients across diverse demographic groups, collected over a one-year period.

Data collection was conducted through field visits to the respective hospitals. Due to institutional privacy policies and the complexity of obtaining written records, oral interviews were utilized as the primary data acquisition method. During these interviews, the purpose and scope of the study were clearly explained to each participant, emphasizing their right to refuse participation at any time. Oral informed consent was obtained before initiating each interview, and participants were assured of full confidentiality and data privacy. Interviews were conducted in a secure and comfortable environment to ensure participant well-being. All data collection and handling procedures adhered strictly to ethical research standards and the principles of scientific integrity.

#### B. Characteristics of the Dataset

The dataset is composed of a rich mixture of clinical, demographic, lifestyle, and genetic variables, as outlined below:

- Clinical data: Blood pressure, cholesterol levels, fasting blood glucose, and ECG outcomes.

- Lifestyle factors: Smoking status, alcohol use, physical activity.
- Demographic information: Age, gender.
- Genetic markers: Presence of known gene variants associated with cardiovascular risk.

#### C. Dataset Features

Recent advancements in ML for cardiovascular prediction have demonstrated the importance of incorporating non-traditional features such as lifestyle behaviors and genetic predispositions into consideration [26]. Additionally, advanced signal processing applied to ECG and imaging data, such as heart rate variability and wavelet-based transformations, is also a valuable feature for predictive modeling [27, 28]. The features utilized in this study are summarized in Table I, which outlines each attribute along with its description and possible values. Table II lists some of the samples used in this study.

TABLE I. ATTRIBUTE RULES

Attribute	Description	Value
Age	Patient age (years)	Continuous
Sex	Gender	0 = male, 1 = female
CP	Chest pain	1 = typical burning sensation in heart, 2 = acute stabbing, 3 = burning sensation, 4 = acute crushing pain in heart
BP	Blood Pressure	Continuous
Chol	Cholesterol	0 = normal, 1 = abnormal
Fbs	Fasting blood sugar	1 = (>120mg/del), 0 =<120 mg/del
ECG	Electrocardiograms	0 = normal, 1 = ST-T wave abnormality, 2 = ventricular abnormality
HR	Maximum heart rate	Continuous
Exang	Exercise-induced angina	No = 0, Yes = 1
FH	Family history	No = 0, Yes = 1
SM	Smoking	No = 0, Yes = 1
HYP	Hypo Kinesis	No = 0, Yes = 1
PREANG	Previous attack of angina	No = 0, Yes = 1
Class	Classification type	Coronary/Heart Disease, Arrhythmias, Normal, Angina Pectoris, Congestive Heart Failure (CHF)

## II. METHODOLOGY

#### A. Preprocessing Steps

Given the complexity and heterogeneity of the collected data, several preprocessing steps were applied to ensure the dataset was suitable for ML analysis:

- Data cleaning: Missing values were imputed based on the nature of each feature. For continuous variables (e.g., blood pressure, cholesterol), missing values were replaced with the median value. For categorical variables (e.g., gender, ECG categories), mode imputation was used to preserve class balance.
- Normalization: Continuous features were normalized using min-max scaling to ensure uniform value ranges.

### B. ML Techniques and Algorithms

To enhance prediction accuracy and leverage the complementary strengths of different ML paradigms, two hybrid models were developed, RF-SMO and J48-LR.

RF is an ensemble learning method known for its high classification accuracy, robustness to overfitting, and effectiveness in handling large datasets with multiple input

features. It operates by constructing multiple decision trees and aggregating their outputs via majority voting. SMO is a fast and efficient training algorithm for Support Vector Machines (SVMs). In the proposed hybrid approach, SMO is applied at the decision nodes of the RF structure to optimize the class separation boundaries, enhancing the model's capability to handle nonlinear and complex decision spaces [29].

TABLE II. SAMPLE OF FEATURES USED FOR PREDICTION

Age	Sex	CP	BP	Chol	Fbs	ECG	HR	Exang	FH	SM	HYP	PREANG
48	0	2	120	1	1	1	90	1	1	0	1	1
76	1	1	120	0	1	1	80	1	0	0	1	1
43	0	1	100	1	0	1	74	1	0	0	0	0
68	1	1	140	0	0	1	66	1	1	0	0	0
58	0	3	130	1	1	1	80	0	0	1	1	0
25	0	3	105	0	0	1	100	0	0	0	1	0
51	1	3	110	0	0	1	82	0	1	0	1	0
16	0	3	120	1	0	1	30	0	1	0	0	0

J48 is the open-source Java implementation of the C4.5 decision tree algorithm. It generates a decision tree with pruned or unpruned branches and is widely used for its interpretability and efficient rule-based classification of structured data. LR is integrated at the leaf nodes of the J48 tree to provide probabilistic interpretation of the classifications. This improves the model's handling of borderline cases and increases predictive accuracy for linearly separable patterns [29]. These two hybrid approaches, RF-SMO and J48-LR, were chosen based on their capacity to model both nonlinear and linear relationships, respectively, enabling more comprehensive diagnostic analysis.

### C. Model Training, Validation, and Testing Processes

To reduce dimensionality and enhance model efficiency, two techniques were applied: i) Principal Component Analysis (PCA) was used to capture the most informative components in the dataset, and ii) feature importance rankings derived from an initial RF model helped identify the most influential variables for heart disease prediction. The dataset was partitioned into 70% for training and 30% for testing.

A 10-fold cross-validation strategy was employed to robustly assess the models' generalization performance. This iterative approach mitigates the risk of overfitting and ensures model robustness across different data subsets. To optimize model performance, a grid search technique was implemented within the cross-validation loop. This exhaustive search over a defined parameter space enabled the identification of the most effective hyperparameter configurations for both the ensemble and statistical components

Model performance was assessed using a comprehensive suite of evaluation metrics:

- Accuracy: Measures the proportion of correctly classified instances.

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN} \quad (1)$$

- Sensitivity, recall, or True Positive Rate (TPR): Captures the model's ability to correctly identify positive (disease) cases.

$$\text{Sensitivity} = \frac{TP}{TP+FN} \quad (2)$$

- Specificity or True Negative Rate (TNR): Assesses the model's capability to correctly identify negative (non-disease) cases.

$$\text{Specificity} = \frac{TN}{TN+FP} \quad (3)$$

- Kappa Statistic: Measures the agreement between predicted and actual classifications, adjusted for chance.

$$\text{Kappa} = \frac{P_o - P_e}{1 - P_e} \quad (4)$$

where  $P_o$  is the observed agreement (i.e., overall accuracy),  $P_e$  is the expected agreement by chance, and  $TP$ ,  $TN$ ,  $FP$ , and  $FN$  denote True Positive, True Negative, False Positive, and False Negative. The ROC curve is plotted using pairs of TPR and FPR. Thresholds ranging from 0 to 1 in increments of 0.1 were used to generate the ROC curve.

## III. RESULTS AND DISCUSSION

To validate the effectiveness of the proposed hybrid models, we conducted a comparative analysis with recent studies that utilized traditional single-algorithm approaches for heart disease prediction. These prior studies employed standalone methods such as RF [30], SMO [31], J48 decision trees [32], and LR [33]. Figure 1 presents the diagnostic accuracy achieved by our models compared to those reported in the selected literature.

The RF-SMO model achieved a diagnostic accuracy of 97%, substantially outperforming recent studies, which reported accuracies of 86% [31], 88% [33], 89% [32], and 92% [30] using standalone algorithms. Similarly, the J48-LR model attained a 92% accuracy, aligning with or slightly improving upon the best-performing conventional model. These improvements highlight the efficacy of integrating ML techniques into a hybrid framework, which can better capture complex patterns within clinical datasets.

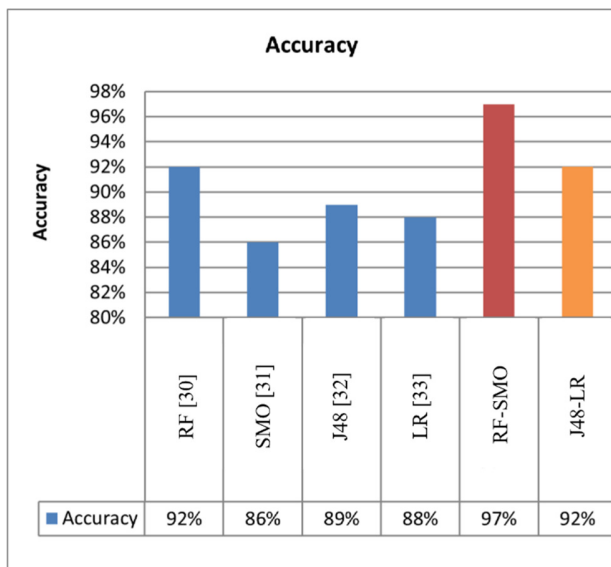


Fig. 1. Differences in accuracy between recent studies.

A. Model Performance Metrics on Collected Dataset

To validate the effectiveness of the proposed hybrid models, their performance was compared with that of traditional standalone algorithms on the same dataset. The RF-SMO model achieved the highest classification accuracy at 97%, followed by the J48-LR model at 92%. In contrast, the traditional models yielded lower accuracies; RF reached 91%, SMO 90%, J48 87%, and LR 85%, as shown in Figure 2

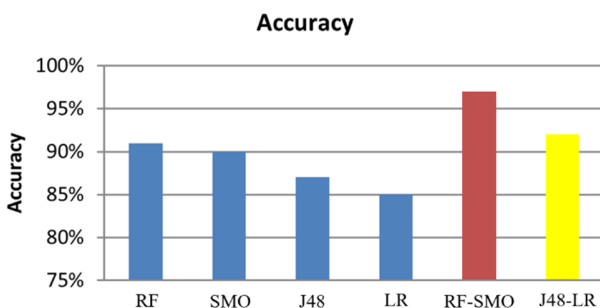


Fig. 2. Accuracy performance on the collected dataset.

Sensitivity and specificity were also examined (Figure 3), given their relevance in clinical diagnostics for identifying true positive and true negative cases. The RF-SMO model recorded a sensitivity of 0.967 and a specificity of 0.995, while the J48-LR model achieved a sensitivity of 0.917 and a specificity of 0.990.

Further evidence of the models' robustness was provided by their ROC-area scores (Figure 4). RF-SMO attained a score of 0.975, and J48-LR achieved 0.963, indicating strong discriminative performance across threshold values. To assess the degree of agreement beyond chance, Cohen's kappa statistic was calculated (Figure 5). The RF-SMO model scored 0.930, and the J48-LR model scored 0.873, confirming

substantial agreement between predicted and actual classifications and reinforcing the reliability of the proposed models.

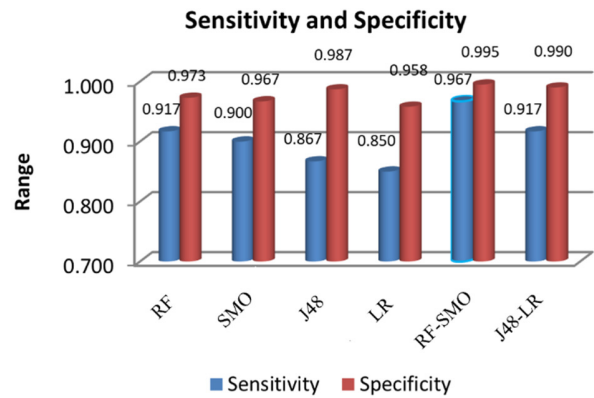


Fig. 3. Sensitivity and specificity performance on the collected dataset.

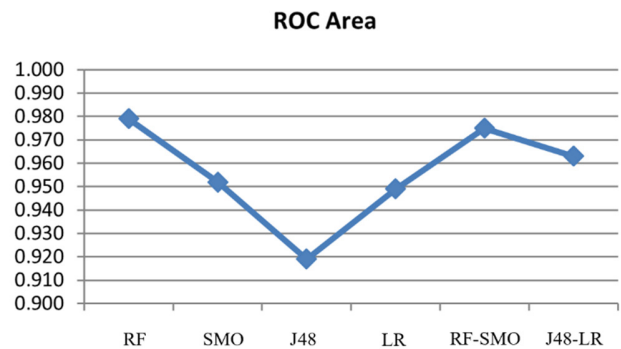


Fig. 4. ROC-area score on the collected dataset.

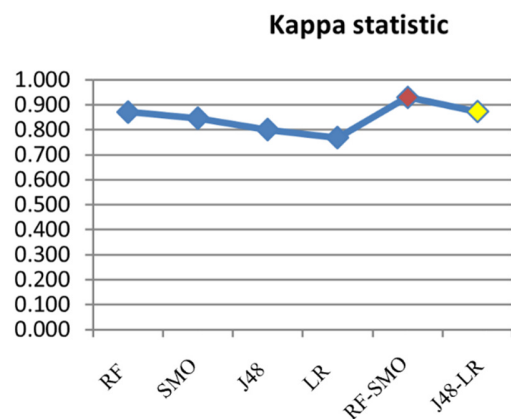


Fig. 5. Kappa statistic on the collected dataset.

The confusion matrices for the proposed models are presented in Tables III and IV. In the case of RF-SMO, almost all instances were correctly classified across all disease categories, with only minor misclassifications. J48-LR showed similarly strong performance, although with slightly more frequent errors in distinguishing between related cardiovascular

conditions. Nonetheless, both models outperformed their base algorithms in nearly every aspect.

The results clearly show that the integration of ensemble and statistical learning techniques significantly improves predictive accuracy and diagnostic reliability. By reducing false positives and false negatives, these hybrid models enhance clinical decision-making processes and present promising avenues for future applications in medical diagnostics.

TABLE III. CONFUSION MATRIX FOR RF-SMO ALGORITHM

a	b	c	d	e	Classified as
13	0	0	1	0	a = Coronary / Heart disease
0	11	0	0	0	b = Angina pectoris
0	0	3	0	0	c = Congestive / heart failure
1	0	0	0	0	d = Arrhythmias
0	0	0	0	31	e = Normal

TABLE IV. CONFUSION MATRIX FOR J48-LR ALGORITHM

a	b	c	d	e	Classified as
13	1	0	1	0	a = Coronary / Heart disease
0	9	0	2	0	b = Angina pectoris
0	0	3	0	0	c = Congestive / heart failure
1	0	0	0	0	d = Arrhythmias
0	0	0	0	30	e = Normal

#### IV. CONCLUSION

This study has demonstrated the efficacy of hybrid Machine Learning (ML) models, Random Forest combined with Sequential Minimal Optimization (RF-SMO), and J48 combined with Logistic Regression (J48-LR), in the prediction of heart disease. The results indicate that these models not only outperform conventional ML algorithms in terms of accuracy but also offer improved probability estimation and class separation, both of which are critical in clinical diagnostics.

Both models exhibited high performance across key evaluation metrics, including accuracy, sensitivity, and specificity, thereby reducing the likelihood of false positives and false negatives, an essential consideration in medical contexts where diagnostic errors can have severe consequences. The RF-SMO model, in particular, achieved superior performance, setting a higher benchmark in all measured aspects. The integration of LR into the decision tree architecture enhanced the model's interpretability and classification precision, providing clinicians with a more reliable tool for informed decision-making.

The successful implementation of these hybrid models represents a meaningful advancement in cardiovascular disease management, improving diagnostic accuracy and enabling more personalized patient care. As ML continues to evolve, its application in healthcare holds the potential to transform disease prediction, monitoring, and treatment. This study contributes to that trajectory by establishing a robust foundation for future developments in clinical predictive modeling.

#### ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to Mustansiriya University ([www.uomustansiriyah.edu.iq](http://www.uomustansiriyah.edu.iq)) and the University of Technology (<https://uotechnology.edu.iq/>), Baghdad, Iraq, for their continued support in this work, as well as to Ibn Al-Bitar Hospital and Ibn Al-Nafis Hospital for their valuable collaboration. All experiments were conducted in accordance with approved guidelines.

#### REFERENCES

- [1] E. Stamate *et al.*, "Revolutionizing Cardiology through Artificial Intelligence—Big Data from Proactive Prevention to Precise Diagnostics and Cutting-Edge Treatment—A Comprehensive Review of the Past 5 Years," *Diagnostics*, vol. 14, no. 11, May 2024, Art. no. 1103, <https://doi.org/10.3390/diagnostics14111103>.
- [2] C. R. Olsen, R. J. Mentz, K. J. Anstrom, D. Page, and P. A. Patel, "Clinical applications of machine learning in the diagnosis, classification, and prediction of heart failure," *American Heart Journal*, vol. 229, pp. 1–17, Nov. 2020, <https://doi.org/10.1016/j.ahj.2020.07.009>.
- [3] C. J. Ejiyi *et al.*, "Enhanced Cardiovascular Disease Prediction Modelling using Machine Learning Techniques: A Focus on CardioVitalnet," *Network: Computation in Neural Systems*, vol. 36, no. 3, pp. 716–748, Jul. 2025, <https://doi.org/10.1080/0954898X.2024.2343341>.
- [4] M. Ahsan, A. Khan, K. R. Khan, B. B. Sinha, and A. Sharma, "Advancements in medical diagnosis and treatment through machine learning: A review," *Expert Systems*, vol. 41, no. 3, Mar. 2024, Art. no. e13499, <https://doi.org/10.1111/exsy.13499>.
- [5] T. Rastogi and N. Girerd, "Enhancing machine learning-based survival prediction models for patients with cardiovascular diseases," *International Journal of Cardiology*, vol. 410, Sep. 2024, Art. no. 132195, <https://doi.org/10.1016/j.ijcard.2024.132195>.
- [6] I. U. Haq, K. Chhatwal, K. Sanaka, and B. Xu, "Artificial Intelligence in Cardiovascular Medicine: Current Insights and Future Prospects," *Vascular Health and Risk Management*, vol. Volume 18, pp. 517–528, Jul. 2022, <https://doi.org/10.2147/VHRM.S279337>.
- [7] M. A. Islam, M. Z. H. Majumder, M. S. Miah, and S. Jannaty, "Precision healthcare: A deep dive into machine learning algorithms and feature selection strategies for accurate heart disease prediction," *Computers in Biology and Medicine*, vol. 176, Jun. 2024, Art. no. 108432, <https://doi.org/10.1016/j.compbiomed.2024.108432>.
- [8] P. Theerthagiri, "Predictive analysis of cardiovascular disease using gradient boosting based learning and recursive feature elimination technique," *Intelligent Systems with Applications*, vol. 16, Nov. 2022, Art. no. 200121, <https://doi.org/10.1016/j.iswa.2022.200121>.
- [9] P. K. Reddy, P. M. Vamsi, C. R. Kumar, K. V. Y. Kumar, P. J. Reddy, and K. L. Nisha, "Predictive Analysis from Patient Health Records Using Machine Learning," in *2023 4th International Conference for Emerging Technology (INCET)*, Belgaum, India, May 2023, pp. 1–6, <https://doi.org/10.1109/INCET57972.2023.10170221>.
- [10] B. Saleh, A. Saedi, A. Al-Aqbi, and L. Salman, "Analysis of Weka Data Mining Techniques for Heart Disease Prediction System," *International Journal of Medical Reviews*, vol. 7, no. 1, Jan. 2020, <https://doi.org/10.30491/ijmr.2020.221474.1078>.
- [11] S. Karthikeyini, G. Vidhya, T. Vetriselvi, and K. Deepa, "Heart Disease Prognosis Using D-GRU with Logistic Chaos Honey Badger Optimization in IoMT Framework," *Information Technology and Control*, vol. 52, no. 2, pp. 367–380, Jul. 2023, <https://doi.org/10.5755/j01.itc.52.2.32899>.
- [12] S. Umamaheswari, A. B. F. Khan, R. Jayababu, and A. A. Ali, "Fuzzy rules-based Data Analytics and Machine Learning for Prognosis and Early Diagnosis of Coronary Heart Disease," *Journal of Information and Organizational Sciences*, vol. 48, no. 1, pp. 167–181, Jun. 2024, <https://doi.org/10.31341/jios.48.1.9>.
- [13] L. B. Elvas, M. Nunes, J. C. Ferreira, M. S. Dias, and L. B. Rosário, "AI-Driven Decision Support for Early Detection of Cardiac Events: Unveiling Patterns and Predicting Myocardial Ischemia," *Journal of*

- Personalized Medicine*, vol. 13, no. 9, Sep. 2023, Art. no. 1421, <https://doi.org/10.3390/jpm13091421>.
- [14] C. M. M. Mansoor, S. K. Chettri, and H. M. M. Naleer, "Development of an efficient novel method for coronary artery disease prediction using machine learning and deep learning techniques," *Technology and Health Care*, vol. 32, no. 6, pp. 4545–4569, Nov. 2024, <https://doi.org/10.3233/THC-240740>.
- [15] F. Li, Y. Chen, and H. Xu, "Coronary heart disease prediction based on hybrid deep learning," *Review of Scientific Instruments*, vol. 95, no. 1, Jan. 2024, Art. no. 015115, <https://doi.org/10.1063/5.0172368>.
- [16] N. S. Alharbi, H. Jahanshahi, Q. Yao, S. Bekiros, and I. Moroz, "Enhanced Classification of Heartbeat Electrocardiogram Signals Using a Long Short-Term Memory–Convolutional Neural Network Ensemble: Paving the Way for Preventive Healthcare," *Mathematics*, vol. 11, no. 18, Sep. 2023, Art. no. 3942, <https://doi.org/10.3390/math11183942>.
- [17] H. Sutanto, "Transforming clinical cardiology through neural networks and deep learning: A guide for clinicians," *Current Problems in Cardiology*, vol. 49, no. 4, Apr. 2024, Art. no. 102454, <https://doi.org/10.1016/j.cpcardiol.2024.102454>.
- [18] M. M. Suhail and T. A. Razak, "Cardiac disease classification from ECG signals using hybrid recurrent neural network method," *Advances in Engineering Software*, vol. 174, Dec. 2022, Art. no. 103298, <https://doi.org/10.1016/j.advengsoft.2022.103298>.
- [19] T. Mohammadi *et al.*, "Unsupervised Machine Learning with Cluster Analysis in Patients Discharged after an Acute Coronary Syndrome: Insights from a 23,270-Patient Study," *The American Journal of Cardiology*, vol. 193, pp. 44–51, Apr. 2023, <https://doi.org/10.1016/j.amjcard.2023.01.048>.
- [20] M. C. Williams *et al.*, "Unsupervised learning to characterize patients with known coronary artery disease undergoing myocardial perfusion imaging," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 50, no. 9, pp. 2656–2668, Jul. 2023, <https://doi.org/10.1007/s00259-023-06218-z>.
- [21] A. M. S.Omar, M. C. Lancaster, S. Narula, A. Baiomi, J. Narula, and P. Sengupta, "Computational Unsupervised Clustering of Echocardiographic Variables for the Assessment of Diastolic Dysfunction Severity," *Journal of the American College of Cardiology*, vol. 71, no. 11, Mar. 2018, Art. no. A1519, [https://doi.org/10.1016/S0735-1097\(18\)32060-6](https://doi.org/10.1016/S0735-1097(18)32060-6).
- [22] S. Anika, M. Islam, and A. Palit, "Early Prediction of Coronary Heart Disease Using Hybrid Machine Learning Models," in *Asia Pacific Advanced Network*, 2024, vol. 1995, pp. 63–75, [https://doi.org/10.1007/978-3-031-51135-6\\_6](https://doi.org/10.1007/978-3-031-51135-6_6).
- [23] P. Dhaka, R. Sehrawat, and P. Bhutani, "An Innovative Approach to Cardiovascular Disease Prediction: A Hybrid Deep Learning Model," *Engineering, Technology & Applied Science Research*, vol. 13, no. 6, pp. 12396–12403, Dec. 2023, <https://doi.org/10.48084/etasr.6503>.
- [24] A. Almulih *et al.*, "Ensemble Learning Based on Hybrid Deep Learning Model for Heart Disease Early Prediction," *Diagnostics*, vol. 12, no. 12, Dec. 2022, Art. no. 3215, <https://doi.org/10.3390/diagnostics12123215>.
- [25] Y. Wu and M. Liu, "Hybrid Deep Learning Model Based on Sparse Recurrent Architecture," *Journal of Circuits, Systems and Computers*, vol. 33, no. 07, May 2024, Art. no. 2450120, <https://doi.org/10.1142/S0218126624501202>.
- [26] C. A. U. Hassan *et al.*, "Effectively Predicting the Presence of Coronary Heart Disease Using Machine Learning Classifiers," *Sensors*, vol. 22, no. 19, Sep. 2022, Art. no. 7227, <https://doi.org/10.3390/s22197227>.
- [27] Y. Zhao, E. P. Wood, N. Mirin, S. H. Cook, and R. Chunara, "Social Determinants in Machine Learning Cardiovascular Disease Prediction Models: A Systematic Review," *American Journal of Preventive Medicine*, vol. 61, no. 4, pp. 596–605, Oct. 2021, <https://doi.org/10.1016/j.amepre.2021.04.016>.
- [28] V. A. Ardeti, V. R. Kolluru, G. T. Varghese, and R. K. Patjoshi, "An overview on state-of-the-art electrocardiogram signal processing methods: Traditional to AI-based approaches," *Expert Systems with Applications*, vol. 217, May 2023, Art. no. 119561, <https://doi.org/10.1016/j.eswa.2023.119561>.
- [29] B. J. Saleh, A. Y. F. Saedi, A. T. Q. Al-Aqbi, and L. A. Salman, "A Review Paper: Analysis of Weka Data Mining Techniques Forheart Disease Prediction System," *Library Philosophy and Practice*, vol. 4032, Aug. 2020.
- [30] A. A. Mohsen, K. Naoufel, T. Alrashahy, and S. Noaman, "Classification and Diagnosis of Heart Disease Using Machine Learning." In Review, Feb. 27, 2024, <https://doi.org/10.21203/rs.3.rs-3985932/v1>.
- [31] I. O. Lawal and O. R. Vincent, "Heart Disease Diagnosis Using Data Mining Techniques and a Decision Support System," in *2022 5th Information Technology for Education and Development (ITED)*, Abuja, Nigeria, Nov. 2022, pp. 1–7, <https://doi.org/10.1109/ITED56637.2022.10051286>.
- [32] O. Gold and A. Iorshase, "Heart failure prediction framework using random forest and J48 with Adaboost algorithms," *Science World Journal*, vol. 18, no. 2, pp. 165–175, Oct. 2023, <https://doi.org/10.4314/swj.v18i2.1>.
- [33] D. Shah, S. Patel, and S. K. Bharti, "Heart Disease Prediction using Machine Learning Techniques," *SN Computer Science*, vol. 1, no. 6, Nov. 2020, Art. no. 345, <https://doi.org/10.1007/s42979-020-00365-y>.