Grenze International Journal of Engineering and Technology, January Issue

Survey on Machine Learning Classifiers for Detecting Rheumatic Heart Disease (RHD) using the UCI Debrecen Dataset

S. Devika[1] and Dr. S. Mangayarkarasi[2]
[1]Research scholar, School of Computing Sciences, VISTAS
Email: devikasresearch@gmail.com
[2]Professor, School of Computing Sciences, VISTAS
Email: drmangaiprabhu@gmail.com

***Abstract*— Rheumatic Heart Disease (RHD) remains a major cause of preventable cardiovascular morbidity and mortality, particularly among children and young adults in low- and middle-income countries. Early detection of valvular damage is essential for preventing disease progression; however, conventional echocardiographic diagnosis is resource-intensive and highly reliant on expert interpretation.**

Recent advances in artificial intelligence (AI) and machine learning (ML) offer promising opportunities to automate RHD screening and support equitable access to care. Clinical Decision Support Systems (CDSS) can assist clinicians in timely and accurate diagnosis, though selecting an effective ML classifier remains a key challenge.

This study evaluates five ML algorithms—Logistic Regression, Gradient Boosting, Random Forest, K-Nearest Neighbors (KNN), and Support Vector Machine—using the UCI Debrecen dataset. Models were assessed using precision, accuracy, F1-score, and recall. Random Forest achieved the best performance, with an F1-score of 72.17% and accuracy of 72.29%, indicating strong overall classification capability. KNN showed the poorest performance, with 60.17% accuracy and the lowest recall (54.47%), highlighting its limited ability to capture complex data patterns.

***Keywords*— Rheumatic Heart Disease (RHD), Logistic Regression (LR), Random Forest (RF), Gradient Boosting (GB), K-Nearest Neighbors (KNN), Support Vector Machine (SVM).**

I. INTRODUCTION

Rheumatic heart disease (RHD) remains a major global health burden, disproportionately affecting children and young adults in low- and middle-income countries. According to recent World Health Organization estimates, millions live with RHD worldwide, and hundreds of thousands die annually from complications such as valvular stenosis, regurgitation, and heart failure [1].

Echo cardiography is the diagnostic gold standard, providing detailed evaluation of valvular structure and function [2]. Recent advances in machine learning (ML) and artificial intelligence (AI) have shown strong potential in cardiovascular imaging.

Grenze ID: 01.GIJET.12.1.557_1

Convolutional neural networks (CNNs) and transformer-based models have been applied to echocardiography for tasks including view classification, chamber segmentation, and lesion detection, achieving performance comparable to human experts [3], [4]. Multimodal approaches integrating electrocardiography (ECG), phonocardiography (PCG), and electronic health records (EHR) can further enhance predictive accuracy [5].

Heart valve diseases (HVD) such as aortic stenosis, mitral regurgitation, and tricuspid abnormalities impair normal blood flow. Early and accurate diagnosis is critical to prevent serious complications like heart failure, arrhythmia, and sudden cardiac arrest. Conventional diagnosis depends on expert interpretation of imaging and clinical tests, which can be time-consuming and subject to inter-observer variability.

Supervised ML methods have emerged as a promising tool for detecting and classifying HVD by learning from labelled patient data (e.g., heart sounds, echocardiograms, clinical parameters). Common algorithms include Support Vector Machines (SVM), Random Forests, Gradient Boosting (e.g., XGBoost), and deep learning models (CNNs, RNNs, Transformers). These techniques support early screening, improve diagnostic accuracy, and assist clinical decision-making.

II. Literature Review

Ahkil et al., in their study "Multi-center retrospective cohort study applying deep learning to electrocardiograms to identify left heart valvular dysfunction," collected transthoracic echocardiography data from five Mount Sinai hospitals in New York City, representing a diverse patient population. Ground-truth valvular status was obtained using a Natural Language Processing pipeline integrated with electrocardiogram (ECG) data. Deep learning models were developed and externally evaluated for valvular disease detection under real-world clinical conditions in 2023 [9].

Daniel Perk et al., in "The Use of Artificial Intelligence Guidance for Rheumatic Heart Disease Screening by Novices," introduced an AI-based method that provides real-time assistance for echocardiographic image acquisition. This approach enables non-experts to obtain diagnostic-quality images using color Doppler, potentially expanding access to RHD screening. Their 2023 study evaluated the effectiveness of AI guidance for novice users [8].

Kelsey Brown et al., in "Using Artificial Intelligence for Rheumatic Heart Disease Detection by Echocardiography: Focus on Mitral Regurgitation," analyzed color Doppler images from 511 pediatric echocardiograms, of which 282 had RHD and 229 were normal. An expert panel independently reviewed all studies. Their automated method employed deep learning with attention mechanisms and mitral regurgitation jet analysis to detect RHD, along with convolutional neural networks to align echocardiographic views and identify the left atrium during systole (2024) [10].

Xinyu Li et al., in "A multi-task deep learning approach for real-time view classification and quality assessment of echocardiographic images," evaluated their method by comparing archived image quality among echocardiographers with varying experience levels—three juniors (1–2 years), three seniors (4–5 years), and three experts (>10 years). Device distribution was similar across groups. Their findings suggest that expert-level images demonstrate higher quality, supporting the feasibility of the proposed technique (2024) [11].

Gino E. Janson et al., in "Automated echocardiography view classification and quality assessment with recognition of unknown views," proposed a 2D ResNet-50 model that processes echocardiogram videos frame-by-frame. Frame-level logits are averaged to generate video-based view predictions, while the maximum logit serves as an anomaly score to accept or reject a view. Frame-based feature vectors are also averaged and linearly regressed to produce a quality score. Only a small subset of labeled studies was used to train the quality assessment component (2024) [12].

Sagnik Ghosh, et al in "Know your Heart: A multimodal pipeline for effective detection and inference of heart diseases" [13] described We have a training dataset, N is the number of modality pairs of Doppler pictures (v), EEG (e), and ECG (c) in our training dataset. Samples of individual modalities for the ith instance are shown here. Every pair of multimodal data has a label (yi) attached to it. To complete the classification objective, the majority of contemporary multimodal techniques use a multi-branch network Cm and accept input from several modalities.[13] at year 2025.

Daniel J. Chung et al., in "Echocardiogram Vector Embeddings Via R3D Transformer for the Advancement of Automated Echocardiography," repurposed an R3D transformer to classify left ventricular ejection fraction (EF) as above or below 50%. The model was trained, validated, and tested on 10,030 echocardiograms from the EchoNet dataset, generating vector embeddings for each study (2024) [14].

Tao Tu and Keyue Chen et al., in "Enhancing Cardiac Disease Detection via a Fusion of Machine Learning and Medical Imaging," described a preprocessing and segmentation workflow for cardiac ultrasound images. Raw

images, containing noise and background artifacts, were first denoised using a Gaussian filter and normalized. Deep learning models such as U-Net and fully convolutional networks (FCN) were then applied to segment key cardiac structures (e.g., ventricles) and extract more precise features, improving diagnostic accuracy for cardiac disease (2025) [15].

Pooja Shah et al., in "Predicting Cardiovascular Risk with Hybrid Ensemble Learning and Explainable AI," utilized three publicly available datasets—the Cleveland Heart Disease, IEEE Dataport Cardiovascular Disease, and Hungarian datasets—to develop a binary classification model (2025) [16]. These datasets include key risk factors such as age, sex, blood pressure, BMI, cholesterol, and glucose. The study describes the preprocessing steps used to prepare the data for model development.

Syed Ali Jafer Zaidi et al., in "HeartEnsembleNet: An Innovative Hybrid Ensemble Learning Approach for Cardiovascular Risk Prediction," used a Kaggle dataset containing clinical records from 70,000 individuals (2025) [17]. The dataset includes 12 key clinical indicators across diverse age groups, enabling comprehensive analysis of cardiovascular risk factors. Its origin from a Kaggle competition highlights its relevance and quality for research applications.

III. Materials and Methods

The dataset used in this work was taken from the UCI Machine Learning Repository. The detailed algorithm for RHD detection is shown in Figure 1. Before training the model, the dataset was preprocessed to ensure data quality and consistency. Using a stratified train-test split, the preprocessed dataset has been split into training and testing sets, with 20% going to testing and 80% going to training. All models were implemented using the scikit-learn library in Python. Default hyperparameters were used for each model. The following measures were used to assess each model's performance:

- Accuracy: a percentage of cases that are properly classified.
  Accuracy=TP+TN/(FP+FN+TP+TN)
  Where:
  TP = True Positives, TN = True Negatives ,FP = False Positives ,FN = False Negatives
- Precision: Precision is the percentage of accurately anticipated positive cases among all positive cases
  Precision=TP/(FP+TP)
- Recall: The percentage of all actual positive instances that were precisely predicted.
  Recall=TP/(TP+FN)
- F1-score: the harmonic mean of recall and precision, which offers a fair assessment of performance.
  F1-score=2×(Precision+Recall/Precision×Recall)

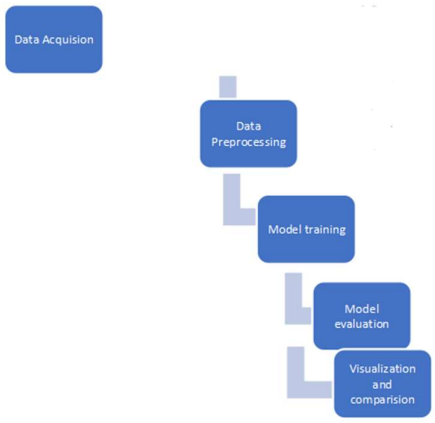

Fig 1. RHD detection algorithm

IV. Results

The performance of five machine learning classifiers—Random Forest, Gradient Boosting, Logistic Regression, Support Vector Machine (SVM), and K-Nearest Neighbors (KNN)—was evaluated for HVD detection using the UCI Debrecen dataset. The results, summarized in Table 1, highlight the variability in performance across the models.

TABLE I: MODEL PERFORMANCE METRICS

Model	Accuracy	Precision	Recall	F1-score
Random Forest	0.7229	0.7757	0.6748	0.7217
Gradient Boosting	0.6840	0.7232	0.6585	0.6894
Logistic Regression	0.7273	0.8333	0.6098	0.7042
Support Vector Machine	0.6797	0.7379	0.6179	0.6726
K-Nearest Neighbors	0.6017	0.6505	0.5447	0.5929

As shown in Table 1, Logistic Regression achieved the highest accuracy (0.7273) and precision (0.8333), indicating its ability to minimize false positives. However, it exhibited a lower recall (0.6098), suggesting a tendency to misclassify positive cases. Random Forest demonstrated a balanced performance, with a high accuracy (0.7229) and F1-score (0.7217), reflecting a good trade-off between precision and recall. Gradient Boosting and SVM showed moderate performance, with accuracy scores of 0.6840 and 0.6797, respectively. KNN performed the least effectively, with the lowest accuracy (0.6017) and recall (0.5447). Figure 2. Shows comparison of accuracies.. The accuracy bar chart visually confirmed the comparative performance of the models, with Logistic Regression and Random Forest showing the highest bars.

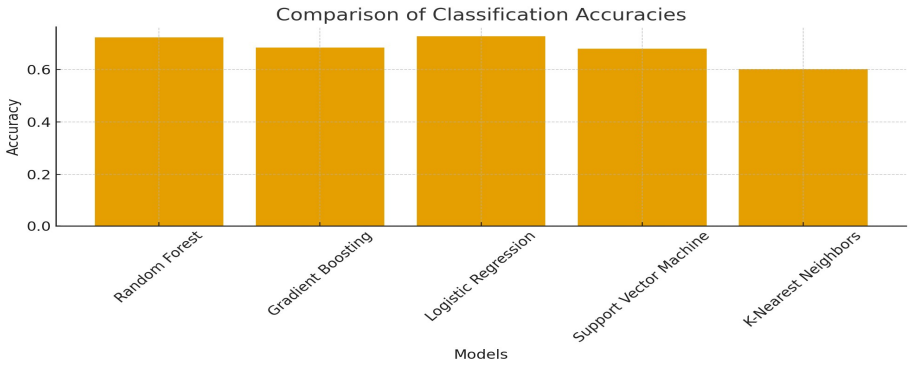

Fig 2. Comparision of accuracies

V. DISCUSSION

The results show that Random Forest and Logistic Regression are the most effective models for RHD detection using the UCI Debrecen dataset. Logistic Regression demonstrated high precision, making it suitable for applications where minimizing false positives is essential, though its lower recall limits its value in screening contexts where missed cases are critical.

Random Forest achieved balanced performance, indicating robustness for general heart valve disease detection due to its resistance to overfitting and ability to handle complex feature interactions. Gradient Boosting and SVM performed moderately, suggesting the need for further optimization, such as hyperparameter tuning or feature engineering. KNN exhibited poor performance, likely due to its inability to capture complex, high-dimensional, and nonlinear patterns.Overall, these findings emphasize the importance of appropriate model selection and evaluation when developing clinical decision support systems for RHD.

VI. CONCLUSION AND FUTURE ENHANCEMENTS

Rheumatic heart valve disease (RHVD) is a chronic complication of rheumatic fever, an autoimmune inflammatory response to Group A streptococcal infection. The resulting valvular scarring and dysfunction can lead to breathlessness, fatigue, chest pain, and, ultimately, heart failure. RHVD remains a major public health concern, especially in low- and middle-income countries.

Clinical Decision Support Systems (CDSS) are gaining importance for early detection and management of RHVD. This study compares five machine learning algorithms—Logistic Regression, Gradient Boosting, Random Forest, K-Nearest Neighbors (KNN), and Support Vector Machine—using the UCI Debrecen dataset. Data were standardized and evaluated using an 80–20 train-test split, with performance assessed by precision, recall, F1-score, and accuracy.

Random Forest achieved the best results, with 72.29% accuracy and a 72.17% F1-score, indicating balanced performance. Logistic Regression showed the highest precision (83.33%) but low recall (60.98%), suggesting

missed positive cases. Gradient Boosting and SVM demonstrated moderate accuracy (68.40% and 67.97%), while KNN performed poorest (60.17% accuracy; 54.47% recall), highlighting its difficulty in modeling complex data relationships.

Future work should investigate deep learning and hybrid approaches to enhance predictive performance. These findings support ongoing efforts to develop reliable automated screening tools for RHVD to enable earlier intervention and improved patient outcomes.

REFERENCES

[1] World Health Organization, Rheumatic Heart Disease – Fact Sheet, 2025.

[2] J. Williamson et al., "Echocardiographic screening for RHD; WHF staging," J. Am. Heart Assoc.,2024.

[3] N. Raissi-Dehkordi et al., "Contemporary applications of AI and ML in echocardiography," Cardiovascular Health,2025.

[4] C. Krittanawong et al., "Deep Learning for Echocardiography: State-of-the-Art Review,"2023.

[5] J.T. Soto et al., "Multimodal deep learning of ECG and Echo time series for LVH detection," Circ.Imaging,2022.

[6] P. Shah et al., "Hybrid ensemble for cardiovascular risk prediction with explainable AI," Sci. Rep.,2025.

[7] S.A.J. Zaidi et al., "HeartEnsembleNet: Hybrid ensemble learning for cardiovascular disease risk," Sci. Rep., 2025

[8] D. Peck, J. Rwebembera, D. Nakagaayi, et al., "Can non-experts screen rheumatic heart disease with AI guidance on handheld echocardiography," J. Am. Soc. Echocardiogr., vol. 36, no. 6, pp. 667–678, 2023, doi: 10.1016/j.echo.2023.02.007.

[9] A. Vaid, et al., "Using deep learning to detect valvular pathology from the electrocardiogram: a multi - centre validation study," Nat. Commun. Med., vol. 4, pp. 12, 2023, doi: 10.1038/s43856-023-00240-w.

[10] K. Brown, P. Roshanitabrizi, C. Sable, et al., "Feasibility and accuracy of artificial intelligence for rheumatic heart disease screening by non-physicians," J. Am. Heart Assoc., vol. 13, no. 2, p. e031245, 2024,

[11] X. Li, Y. Zhang, H. Chen, et al., "Real-time echocardiographic view classification and quality assessment for rheumatic heart disease AI stabilization," Sci. Rep., vol. 14, p. 10234, 2024

[12] P. Jansen, R. Muller, K. Singh, et al., "Automated echocardiographic view selection and out-of-distribution detection using CNN-Transformer pipelines," J. Med. Imaging, vol. 11, no. 3, p. 034501, 2024,

[13] A. Ghosh, R. Banerjee, S. Patel, et al., "Multimodal deep learning using ECG and Doppler echocardiography for cardiac disease detection," in Proc. ACM Conf. Bioinformatics, 2024, pp. 155–164.

[14] J. Chung, H. Lee, L. Wang, et al., "Echocardiographic video representation learning with a 3D Transformer for ejection fraction prediction," iScience, vol. 27, no. 5, p. 108451, 2024,

[15] H. Yu, R. Kumar, L. Zhao, et al., "Hybrid imaging and machine learning for enhanced detection of cardiac disease," Sci. Rep., vol. 15, p. 2211, 2025,

[16] M. Shah, P. Verma, A. Agarwal, et al., "Explainable hybrid ensemble models for cardiovascular prediction with echocardiography transferability," Sci. Rep., vol. 15, p. 3105, 2025.

[17] S. Zaidi, O. Khan, M. Ali, et al., "HeartEnsembleNet: calibrated hybrid ensembles for cardiovascular risk detection,"Diagnostics,Vol.15,no.4,p.722,2025.