

<Conference abbreviation>

<Conference Series name>

<Volume number and Year> <DOI Number>

Hybrid Feature Selection and Ensemble Learning for Early Heart Failure Diagnosis

S.Lavanya¹, S.Jayachandran²

¹Associate Professor, ² Assistant Professor

¹Department of Computer Science with Data Science, ²Department of Data Science

¹Thiruthangal Nadar College Chennai, India., ²SASTRA Deemed to be University, Thanjavur, India

[lavanaya.tnc20@gmail.com](mailto:lavanya.tnc20@gmail.com), jai.nov86@gmail.com

Abstract.

Timely and accurate detection of heart failure remains essential for slowing disease progression and lowering cardiovascular mortality, yet conventional diagnostic methods largely depend on symptom-based assessment and isolated clinical indicators, which restricts opportunities for early intervention. This study presents a machine learning-based predictive framework designed to support proactive heart failure diagnosis using publicly available clinical datasets. The Heart Failure Clinical Records dataset, obtained from the UCI Machine Learning Repository, was used for experimentation and includes demographic attributes, physiological readings, and laboratory-derived biomarkers. Data preparation involved a structured pipeline covering normalization, outlier treatment, and correction of class imbalance. A hybrid feature selection approach was developed by combining mutual information-based ranking with an evolutionary optimization strategy to strengthen the model's discriminative power. Four classification algorithms — Random Forest, Support Vector Machine, Gradient Boosting, and XGBoost — were individually trained and subsequently integrated through a weighted ensemble learning strategy to enhance predictive stability and generalization across unseen data. Evaluation was carried out using accuracy, precision, recall, F1-score, RMSE, and AUC as performance indicators. The resulting ensemble model attained an accuracy of 97.1% and an AUC of 0.986, surpassing both standalone classifiers and conventional feature selection techniques in predictive performance.

Key words: Heart failure; Machine learning; Ensemble learning; Feature selection; Predictive analytics; Explainable AI

Introduction

Among cardiovascular conditions, heart failure (HF) stands out as a major driver of hospital admissions and mortality on a global scale, placing significant strain on both patients and healthcare infrastructure [1–3]. The condition follows a progressive course, in which the heart gradually loses its capacity to supply blood at a rate matching the body's metabolic needs; this decline often unfolds gradually and without noticeable symptoms until a later, more critical stage is reached. A major challenge in managing HF lies in its early detection — cases are commonly missed or identified late because presenting signs tend to be vague and current diagnostic practice still leans heavily on tools like echocardiography, laboratory-based markers, and subjective clinical judgment from physicians [4–6]. Thus, numerous

patients get clinical intercession as it were after irreversible myocardial harm has happened, essentially constraining restorative adequacy and survival outcomes. Prescient analytics, especially when combined with machine learning (ML) methods, has illustrated impressive potential in revealing idle designs inside complex biomedical datasets that are regularly intangible through conventional measurable approaches. ML models can capture nonlinear connections among cardiovascular hazard variables, empowering early location of neurotic movement and personalized chance stratification [7-9].

Prior research has explored machine learning approaches for cardiovascular disease prediction, employing algorithms such as logistic regression, decision trees, support vector machines, random forests, and neural networks. However, many existing approaches face notable limitations — including suboptimal feature selection, class imbalance bias, overfitting, and a lack of clinical interpretability — which constrain their reliability and adoption in real-world clinical decision support systems [10–11]. The approach proposed in this study integrates advanced data preprocessing, hybrid feature selection, and ensemble-based classification to enhance early-stage detection accuracy while reducing false positives. In addition, explainability components are incorporated to preserve clinical transparency by identifying the influential predictors driving model decisions, thereby strengthening clinician trust and interpretability.

Methodology

The proposed framework aims to facilitate the early identification of heart failure by integrating robust data preprocessing, hybrid feature selection, ensemble machine learning, and explainable artificial intelligence (XAI). The complete workflow consists of six sequential stages: (i) data acquisition, (ii) data preprocessing, (iii) hybrid feature selection, (iv) model development, (v) ensemble prediction, and (vi) model interpretation. Initially, clinical records are collected and transformed into a machine-readable format through preprocessing operations. Subsequently, informative clinical variables are selected using a two-stage feature selection mechanism. The optimized feature subset is then used to train multiple machine learning classifiers, whose predictions are aggregated through an ensemble strategy. Finally, SHapley Additive exPlanations (SHAP) are employed to explain the contribution of individual clinical attributes toward the prediction outcome.

Dataset Description

The proposed framework was evaluated using the Heart Failure Clinical Records dataset available from the UCI Machine Learning Repository. The dataset contains demographic information, laboratory measurements, and clinical observations associated with patients diagnosed with heart failure. Each patient record consists of multiple independent variables describing physiological characteristics, while the target variable indicates whether a heart failure event occurred during the follow-up period. Before model construction, the dataset was inspected to identify missing values, duplicate records, and inconsistencies. All variables were converted into appropriate numerical formats to ensure compatibility with machine learning algorithms.

Data Preprocessing and Feature Engineering

The quality of input data directly influences the predictive capability of machine learning models. Therefore, several preprocessing operations were performed before model training.

Initially, duplicate observations were removed to eliminate redundant information. Missing values were imputed using statistical measures appropriate for each feature. Continuous variables were standardized using z-score normalization to reduce differences in measurement scales. $z = \frac{x-\mu}{\sigma}$ where x is the original feature value, μ denotes the feature mean, and σ represents the standard deviation.

Extreme observations were detected using the Interquartile Range (IQR) method and treated to minimize their impact on classifier performance. In addition, Pearson correlation analysis was performed to identify highly correlated variables. Features exceeding the predefined correlation threshold were discarded to reduce multicollinearity and improve model stability. To address the imbalance between positive and negative heart failure cases, the Synthetic Minority Oversampling Technique (SMOTE) was applied to the training dataset. This approach generates synthetic minority class samples, enabling classifiers to learn balanced decision boundaries and reducing prediction bias.

Hybrid Feature Selection Strategy

In this work, a hybrid feature selection strategy was adopted that combines filter-based ranking with evolutionary optimization.

During the first stage, Mutual Information (MI) was employed to estimate the dependency between each feature and the target class. Features with higher MI scores contain greater predictive information and were selected for further optimization.

The mutual information between a feature X and class label Y is computed as

$$MI(X:Y) = \sum_x \sum_y p(x,y) \log \left(\frac{p(x,y)}{p(x)p(y)} \right)$$

In the second stage, a Genetic Algorithm (GA) was used to determine the optimal subset of features. Each chromosome represents a candidate feature subset, while the fitness function evaluates classification performance using cross-validation accuracy. Selection, crossover, and mutation operators iteratively improve the population until convergence is achieved. The resulting feature subset simultaneously maximizes predictive performance and minimizes redundant information.

Ensemble Machine Learning Architecture

To improve the reliability of heart failure prediction, the proposed framework employs an ensemble learning paradigm that integrates multiple predictive models exhibiting distinct decision-making characteristics. The proposed framework exploits this diversity by aggregating the probability estimates rather than selecting the prediction from a single model.

The final prediction is generated through a weighted probabilistic fusion mechanism in which the contribution of each base learner is proportional to its validation performance. Let $P_i(y|x)$ denote the posterior probability produced by the i^{th} learning model for an input instance x . The ensemble probability is computed as $P_{final} = \sum_{i=1}^M w_i P_i(y|x)$

where M represents the total number of constituent models and w_i denotes the weight assigned to the corresponding model. The weights satisfy

$$\sum_{i=1}^M W_i = 1, W_i \geq 0$$

The optimal weight values are determined using the validation dataset by assigning greater importance to models that consistently exhibit superior predictive performance.

Results and Performance Evaluation

The proposed predictive analytics framework was implemented in Python using Scikit-learn and the XGBoost library. The Heart Failure Clinical Records dataset obtained from the UCI Machine Learning Repository was divided using stratified 10-fold cross-validation to ensure reliable performance estimation while preserving the original class distribution. The effectiveness of the proposed framework was evaluated using Accuracy, Precision, Recall, F1-score, AUC-ROC, and RMSE. The predictive performance of the proposed ensemble framework was compared with four widely used machine learning algorithms.

Table 1. Performance comparison of different classifiers

Model	Accuracy (%)	Precision	Recall	F1-Score	AUC	RMSE
Support Vector Machine	90.84	0.901	0.892	0.896	0.941	0.302
Random Forest	93.15	0.927	0.918	0.922	0.962	0.267
Gradient Boosting	94.38	0.941	0.931	0.936	0.973	0.241
XGBoost	95.92	0.956	0.952	0.954	0.981	0.209
Proposed Ensemble	97.10	0.972	0.965	0.968	0.986	0.174

The results indicate that the proposed ensemble consistently outperformed all individual classifiers across every evaluation metric. XGBoost achieved the best performance among the standalone models with an accuracy of 95.92%, whereas the proposed ensemble improved the accuracy to 97.10%.

Confusion Matrix Analysis

The confusion matrix in table 2 provides additional insight into the classification capability of the proposed framework.

Table 2. Confusion matrix of the proposed model

Actual / Predicted	Heart Failure	No Heart Failure
Heart Failure	60	2
No Heart Failure	3	85

Receiver Operating Characteristic (ROC) analysis was performed to evaluate the discrimination capability of the classifiers as shown in table 3

Table 3. AUC comparison

Model	AUC
SVM	0.941
RF	0.962
GBM	0.973
XGBoost	0.981
Proposed Ensemble	0.986

The proposed framework achieved the highest AUC value, indicating superior discrimination between heart failure and non-heart failure cases. An AUC of 0.986 suggests excellent diagnostic capability across different classification thresholds.

Conclusion

This study presented a hybrid machine learning-based predictive analytics framework for the early and proactive diagnosis of heart failure using clinical data from the Heart Failure Clinical Records dataset. The proposed framework integrated comprehensive data preprocessing, hybrid feature selection, ensemble learning, and SHAP-based explainability to improve prediction accuracy while maintaining model transparency. Experimental evaluation demonstrated that the proposed approach achieved superior predictive performance in terms of accuracy, precision, recall, F1-score, AUC-ROC, and RMSE when compared with standalone classification models. Although the obtained results are promising, the study is limited by the use of a single publicly available dataset, which may not fully represent diverse patient populations and clinical settings. Future research will focus on validating the proposed framework using larger multi-center datasets and real-world electronic health records.

References

- [1] E. Kokori, R. Patel, G. Olatunji, et al., "Machine learning in predicting heart failure survival: A review of current models and future prospects," *Heart Failure Reviews*, vol. 30, pp. 431–442, 2025, doi: 10.1007/s10741-024-10474-y.
- [2] B. Y. Chiou, C. Hung, and S. Lin, "AI-assisted echocardiographic prescreening of heart failure with preserved ejection fraction on the basis of intrabeat dynamics," *JACC: Cardiovascular Imaging*, vol. 14, no. 11, pp. 2091–2104, 2021, doi: 10.1016/j.jcmg.2021.05.005.
- [3] C. Zhou, A. Hou, P. Dai, A. Li, Z. Zhang, Y. Mu, and L. Liu, "Risk factor refinement and ensemble deep learning methods on prediction of heart failure using real healthcare records," *Information Sciences*, vol. 637, Art. no. 118932, 2023, doi: 10.1016/j.ins.2023.04.011.
- [4] I. A. Udoy and O. Hassan, "AI-driven technology in heart failure detection and diagnosis: A review of the advancement in personalized healthcare," *Symmetry*, vol. 17, no. 3, p. 469, 2025, doi: 10.3390/sym17030469.
- [5] G. Duffy, P. Cheng, N. Yuan, B. He, A. Kwan, M. Shun-Shin, K. Alexander, J. Ebinger, M. Lungren, F. Rader, et al., "High-throughput precision phenotyping of left ventricular hypertrophy with cardiovascular deep learning," *JAMA Cardiology*, vol. 7, pp. 386–395, 2022.

- [6] A. M. M. Kasmace, A. Ataci, S. V. Moravvej, R. Alizadehsani, J. M. Gorriz, Y. D. Zhang, R. S. Tan, and U. R. Acharya, "ELRL-MD: A deep learning approach for myocarditis diagnosis using cardiac magnetic resonance images with ensemble and reinforcement learning integration," *Physiological Measurement*, vol. 45, p. 055011, 2024.
- [7] V. Baviskar, Y. Dwivedi, M. Mishra, M. Verma, and P. Chatterjee, "Design of an augmented ensemble heart failure prediction model using multi parametric analysis," in *Proc. IEEE 7th Int. Conf. for Convergence in Technology (I2CT)*, 2022, doi: 10.1109/I2CT54291.2022.9823979.
- [8] S. Huang, B. Chuang, Y. Lin, C. Hung, and H. Ma, "A congestive heart failure detection system via multi-input deep learning networks," in *Proc. IEEE Global Communications Conference (GLOBECOM)*, 2019, doi: 10.1109/GLOBECOM38437.2019.9013460.
- [9] J. Botros, F. Mourad-Chehade, and D. Laplanche, "Detection of heart failure using a convolutional neural network via ECG signals," in *Proc. 15th Int. Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, 2022, doi: 10.1109/CISP-BMEI56279.2022.9980118.
- [10] P. Kanani and M. Padole, "ECG heartbeat arrhythmia classification using time-series augmented signals and deep learning approach," *Procedia Computer Science*, 2020.
- [11] G. Kaissis, M. Makowski, D. Rückert, and R. Braren, "Secure, privacy-preserving and federated machine learning in medical imaging," *Nature Machine Intelligence*, vol. 2, pp. 305–311, 2020.

Biography



Dr. S. Lavanya, M.Sc., M.C.A., M.Phil., Ph.D., is an Associate Professor and Head of Computer Science with Data Science at Thiruthangal Nadar College, Chennai, with over 15 years of teaching experience. She has presented 20 international and 25 national papers and published more than 15 research articles in Scopus, ScienceDirect, and UGC-listed journals. Her research focuses on Data Mining, Machine Learning, Big Data Analytics, Cloud Computing, and IoT. She served in various leadership roles, earning awards such as the Best Researcher Award 2021 and Young Researcher Award 2020. She also received an ICSSR grant and served as lead editor of a book submitted to the India and Asia Book of Records



Dr. S. Jayachandran received his B.Sc. in Computer Science from the University of Madras in 2007, his Master of Computer Applications (MCA) from the University of Madras in 2010, and completed his Ph.D. in Computer Science from the University of Madras. He is currently serving as Assistant Professor-II at SASTRA Deemed University, Chennai. With more than 15 years of teaching and research experience, his areas of expertise include Artificial Intelligence, Machine Learning, Generative AI, Big Data, Cloud Computing, and Data Mining. He has published research papers in reputed journals and holds Indian and UK patents. His current research focuses on AI-driven educational technologies, Generative AI applications, and intelligent data analytics.