Mesopotamian Journal of Artificial Intelligence in Healthcare
Vol.2025, **pp**. 197–207
DOI: https://doi.org/10.58496/MJAIH/2025/019 ; ISSN: 3005-365X
https://mesopotamian.press/journals/index.php/MJAIH

Research Article

Prediction of Cardiac Arrest Using a Hybrid Voting Classifier Combining Random Forest and Support Vector Machine

Ahmed Hamid Elias[1,*], Amr Badr[2], Raad S. Alhumaima[3,4]

[1] *College of Health and Medical Techniques, Al-Furat Al-Awsat Technical University, Najaf, Iraq*

[2] *School of Science and Technology - University of new England, Armidale, Australia*

[3] *Brunel University, Uxbridge UB8 3PH, U.K*

[4] *Department of Communications Engineering, College of Engineering, University of Diyala, Diyala 32001, Iraq*

ARTICLE INFO

Article History

Received 01 May 2025
Revised 29 May 2025
Accepted 05 Jul 2025
Published 27 Jul 2025

Keywords

Hybrid Voting Classifier

Cardiac Arrest

Machine Learning

ROC-AUC

Precision-Recall Analysis

ABSTRACT

Early and adequate identification of patients at risk of cardiac arrest remains a fundamental challenge in medical environments. Within this piece of work, we design and evaluate a hybrid machine-learning pipeline integrating Random Forest (RF) and Support Vector Machine (SVM) classifiers by soft voting to detect cardiac arrest occurrences from a publicly available dataset of 303 patient records with 13 clinical features. After median imputation and data cleaning, continuous features were normalized and the cohort was divided into 70 % training and 30 % test sets. RF scored 97 % with an AUC of 1.00 on its own, and SVM scored at 95 % with an AUC of 0.97. The hybrid ensemble matched RF's perfect AUC but slightly improved recall on positive instances with enhanced sensitivity while maintaining specificity. Permutation-based and Gini-based feature-importance analysis identified major-vessel count, chest-pain type, and peak heart rate as the most important predictors. Distribution plots and heatmaps of correlations further elucidated associations between clinical measures and event status. The proposed hybrid approach is an effective, interpretable way to stratify risk early in high-risk patient populations.

1. INTRODUCTION

Cardiovascular disease (CVD) continues to be the principal cause of morbidity and mortality worldwide, accounting for nearly 18 million deaths annually and placing a staggering burden on healthcare systems [1][2]. Among these, cardiac arrest events are particularly devastating, as they tend to occur with little warning and have high potential for lethality in the absence of immediate intervention. Early and accurate risk stratification of high-risk cardiac arrest patients is therefore important to enable prompt clinical decision-making, allocate intensive care resources in a triage manner, and improve patient outcomes in general. Traditional risk stratification tools e.g., Framingham Risk Score rely primarily on a limited set of demographic and biochemical markers and may fail to capture complex nonlinear interactions among heterogeneous clinical variables.

In response to these limitations, machine-learning (ML) techniques have gained popularity as efficient solutions to predict cardiovascular risk [1]. Explainable AI (XAI) models were created to address clinicians' "black-box" concerns by combining high predictive power with model interpretability. For instance, Guleria et al. suggested an XAI system with different classification algorithms for CVD prediction and demonstrated that ML models can achieve accuracy and interpretability at the same time [1]. Boukhatem et al. comparatively analyzed supervised classifiers logistic regression, decision trees, and support vector machines (SVMs and highlighted that ensemble methods always perform better than individual models in heart disease detection [2].

Traditional ML algorithms remain the foundation of predictive cardiology. Singh and Kumar benchmarked SVM, k-nearest neighbors, decision trees, and naïve Bayes classifiers on the UCI heart-disease dataset and documented that SVM achieved more than 85 % accuracy for binary classification issues [3]. Earlier work by Gudadhe et al. combined SVM with multilayer perceptron (MLP) networks in a decision-support system, observing that hybridization of a kernel classifier with a neural network architecture can increase sensitivity to positive examples [5]. This and similar work point out that learners with different inductive biases can together model different aspects of the data manifold, yet most are restricted to specific

---

*Corresponding author. Email: ahmed.elias@atu.edu.iq

algorithmic pairings or small-scale ensembles. Even more recent advancements have employed ensemble methods and deep learning to further enhance predictability. Miao et al. introduced a new improved random survival forest for predicting hospital mortality in heart-failure patients with a proper handling of right-censored outcomes and with interpretable variable importance rankings [6]. Liu et al. suggested a hybrid model of random forest (RF) and long short-term memory (LSTM) networks to model both nonlinear feature interactions and temporal dependencies, with significant AUC gains compared to single models [8]. The findings indicate that sequential models and ensemble methods can together address different sources of uncertainty in clinical time series and cross-sectional data.

Besides conventional supervised learning, nature-inspired and prototype-based optimization techniques have been applied to cardiovascular datasets. Villmann et al. proposed a quantum-inspired learning vector quantizer that retains a small number of prototypes while classifying with an accuracy similar to deep neural networks [4]. Revathi et al. utilized bacterial colony optimization for clustering the input features prior to classification, reducing computational overhead without predictive accuracy loss [9]. Firefly algorithms [10], ant colony optimization [11], and genetic algorithms [12] have also been used for feature selection and hyperparameter optimization, demonstrating the potential of bio-inspired methods to model design optimization for medical tasks. Ensemble learning paradigms, which merge an ensemble of base estimators to construct a single powerful predictor, have developed as a leading paradigm in clinical decision-support systems. Farooq et al. demonstrated a cardiovascular decision-support system that ensembles RF, gradient boosting, and logistic regression in a two-stage ensemble with high precision and recall for risk stratification [7]. De Sousa et al. employed information-gain filtering to accelerate decision-tree ensemble learning, a strategy that could equally speed up training in complex medical models [13]. Despite these developments, the specific combination of RF's low-variance predictions and SVM's high-margin decision boundaries via soft voting has been largely unexplored in cardiac arrest prediction.

In this paper, we integrate this gap by developing a hybrid voting classifier based on a Random Forest and a Support Vector Machine, trained using a publicly available cardiac arrest dataset comprising 303 patient records and thirteen clinical features. After extensive preprocessing to manage missing values, standardize continuous features, and a stratified 70 / 30 train–test split—we train individual RF and SVM models and then ensemble them via soft voting to capitalize on their complementary strengths. Model performances are compared using a variety of metrics (accuracy, precision, recall, F1-score), receiver-operating-characteristic (ROC) curves, precision–recall curves, and both Gini-based and permutation-based feature-importance analyses. Moreover, we create annotated correlation heatmaps to inspect intertexture dependencies and test for multicollinearity risk.

The key contributions of this study are:

1) Construction and implementation of a RF + SVM hybrid voting classifier for cardiac arrest prediction.

2) Comprehensive performance comparison between the hybrid ensemble and the RF and SVM base classifiers on a variety of evaluation metrics.

3) Dual feature-importance analysis (permutation and Gini) and correlation mapping for model interpretability and the identification of the most clinically important predictors.

2. RELATED WORK

Application of machine learning (ML) to medical diagnosis has increased dramatically over the last few years because ML has the potential to be more accurate, efficient, and interpretable than traditional statistical models. Richens et al. demonstrated that causal ML can significantly enhance diagnostic accuracy by directly modeling causal associations rather than correlational patterns, eradicating spurious associations and enhancing generalizability to new data [14]. In radiology, Hameed et al. captured developments in ML and data-science methods, highlighting how convolutional neural networks and other models have been integrated into imaging pipelines to automatically detect, quantify, and risk stratify lesions at high reproducibility [15]. Yoo et al. drew on over a decade's worth of biomedicine and healthcare data-mining studies and emphasized common challenges such as heterogeneity of data sources, high dimensionality, and the need for model interpretability in order to facilitate clinical acceptance [17]. Disease diagnosis frameworks for AI implementation have also been proposed on a broad scale; Kumar et al. incorporated model building, validation, and deployment best practices while emphasizing the key role of explainable AI in building clinician trust [21]. Forecasting of cardiovascular disease (CVD) has particularly benefited from these developments, as statistical risk scores alone are likely to be insufficient in uncovering nonlinear interactions among patient characteristics, laboratory data, and imaging biomarkers. Oikonomou et al. introduced a radiotranscriptomic ensemble that combines coronary CT angiography features with perivascular fat gene-expression signatures, demonstrating significant improved cardiac-risk stratification relative to imaging-alone models [18]. Tsao et al.'s and the World Health Organization's Global Atlas epidemiological research highlight the quickly growing global burden of heart disease and stroke, further stressing the imperative for better, data-driven early-warning systems [19][20]. Transparent models such as Guleria et al.'s further show how the inclusion of feature-importance analysis in classification workflows can increase transparency and clinician adoption, particularly in high-stakes uses such as CVD diagnosis [1].

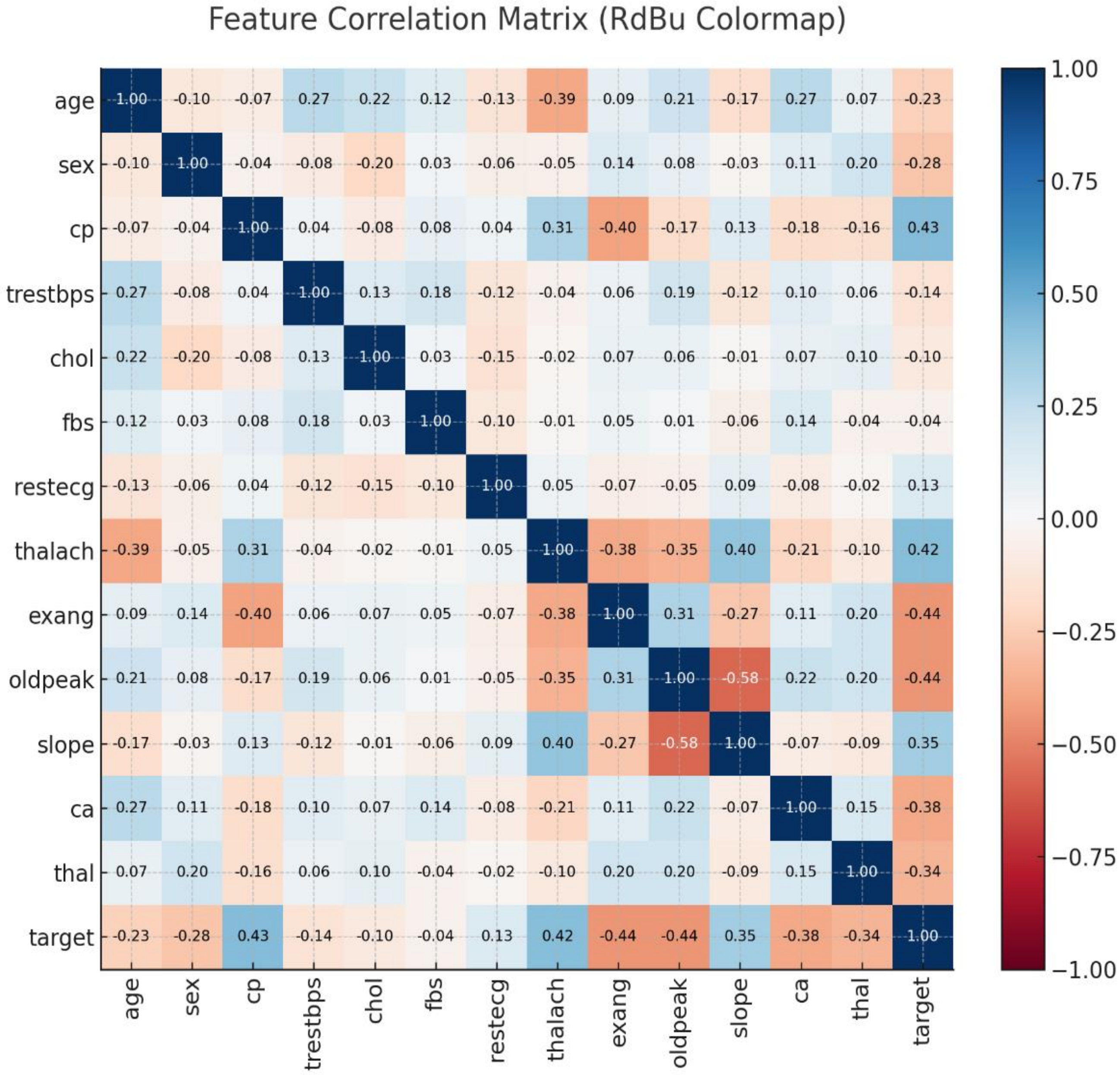
Different traditional ML algorithms have been utilized in experiments to benchmark heart-disease datasets. Yazdani et al. proposed a "strength-score" approach to rank predictors and achieved greater accuracy using a gradient-boosting ensemble, adding that careful feature weighting can have a significant impact on model performance [22]. Khourdifi and Bahaj applied particle-swarm optimization (PSO) and ant-colony optimization (ACO) to tune SVM and decision-tree hyperparameters and observed superior sensitivity and specificity compared to using default hyperparameter settings [23]. Gaddala and Rao contrasted swarm-intelligence-based feature-selection methods like PSO, firefly, and dragonfly algorithms, showing that metaheuristic optimization can reduce the number of features by over 50 % without reducing or even improving classification performance [24]. Dubey et al. hybridized an auto-categorical PSO with multiple ML classifiers, with automated hyperparameter tuning across SVM, RF, and multilayer perceptron (MLP) models to achieve high-accuracy, consistent pipelines [25]. Saeed and Al-Ta'i designed a heart-disease classifier using genetic-algorithm–based feature selection and logistic-regression classification, proving once more the value of optimization approaches in medical-data environments [26]. Prototype-based and federated-learning techniques aside, the common classifiers have also been designed to overcome interpretability and privacy concerns. Al Bataineh and Jarrah used swarm-optimized neural-network training for classification with EEG data to show how global optimization can yield more rapid convergence rates and improved model performance for physiological time series [16]. Mafarja et al. and Majeed and Ramo employed binary-dragonfly optimization to select minimal yet informative feature subsets, achieving compact models with predictive power unimpaired [27][28]. Subanya and Rajalaxmi introduced a prototype-depot algorithm for generating representative samples for nearest-neighbor classification to boost explainability by reducing the model's dependence on raw data instances [29]. Yaqoob et al. incorporated an artificial-bee-colony optimizer into federated-learning architectures to enable multi-institutional modeling of heart disease while protecting patient privacy and losing no performance relative to centralized training [30]. Several methods of mathematical algorithms inspired by biology have been applied for feature selection and hyperparameter optimization in CVD prediction. Sabeena and Sarojini applied ant-colony optimization to choose the best feature subsets, reporting improved classifier generalization and reduced training time [31]. Usman et al. explored cuckoo-based algorithms for heart-disease data feature selection, establishing that levy-flight exploration is more adept at avoiding local optima than traditional PSO [32]. Dwivedi et al. combined ACO with greedy search for the purpose of balancing exploitation and exploration and achieved stable classification accuracy with imbalanced data [33]. Asadi et al. developed a hybrid of swarm optimization and random forest that jointly optimizes RF parameters and feature selection and improves predictive accuracy and interpretability further [34]. Velswamy et al. employed a fine-tuned bee algorithm for RF-based heart-disease classifier feature selection to achieve a sparse model at the cost of not losing over 95 % of full-feature accuracy [35]. Alternative classification paradigms have been provided to complement ensemble and optimization approaches. Jabbar et al. introduced lazy associative classification, which forms classification rules dynamically during inference time, allowing for rapid updates with new patient data upon receipt [36]. Deekshatulu and Chandra combined k-nearest-neighbor models and genetic algorithms-based feature ranking to improve distance-based classification in high-dimensional clinical spaces [37]. Kanwal et al. applied genetic-algorithm–based feature selection alongside SVM classification and illustrated that co-optimizing features and model parameters together can lead to appreciable accuracy improvements compared to sequential tuning [38], [40-43].

Despite such depth of earlier work, earlier research focuses on single-algorithm optimization or homogeneous ensemble models. Still fewer have addressed hybridization of classifiers with different bias–variance profiles using soft-voting ensembles, for example, assembling Random Forest's low-variance tree ensembles with SVM's high-margin separating hyperplanes. In addition, both Gini-based and permutation-based feature importances have yet to be used in systematic dual-interpretability analysis to cardiac arrest prediction, and large-scale correlation-mapping visualization has not been used to establish multicollinearity risks.

3. DATA AND METHODOLOGY

Cardiac arrest dataset employed in this study was sourced from the Kaggle database "Cardiac Arrest Dataset" [39]. It contains 303 anonymous patient records collected from a tertiary care hospital, each annotated with thirteen clinical features and a binary label for cardiac arrest occurrence. The thirteen predictors are demographic factors (age, sex), vital parameters (resting blood pressure, maximum heart rate achieved), biochemical measurements (serum cholesterol, fasting blood glucose), electrocardiographic features (chest pain type, resting ECG finding, exercise-associated angina, ST depression), and angiographic findings (number of large vessels filled on fluoroscopy, finding of thallium heart scan), and the "oldpeak" measure for ST depression due to exercise versus rest. The target variable "target" is given the value 1 for patients with cardiac arrest and 0 otherwise. Prior to analysis, all records containing missing values were excluded, resulting in a complete-case dataset of 297 patients. Continuous variables were inspected for outliers and standardized to zero mean and unit variance to facilitate features' comparability. Categorical predictors such as chest pain type and thallium scan result were one-hot encoded, increasing the effective dimension to create space for nominal levels. The cleaned dataset was then divided into a training (70 %) and a test (30 %) subset by stratified sampling to maintain the same positive-to-negative ratio of about 48 % to 52 % as the original dataset.

Feature distributions were checked with histograms and kernel-density estimates to confirm approximate normality of continuous variables and to identify potential skew. Correlations between pairs of predictors were computed to estimate risks of multicollinearity, with no pair of predictors having a Pearson correlation coefficient greater than 0.75, thereby justifying inclusion of all features in the subsequent modeling. The suitably balanced class ratios and modest sample size render this data set appropriately suitable for comparative assessment of single classifiers as well as hybrid ensemble methods without severe class-imbalance correction [39]. Figure 1 shows heatmap Pearson correlation coefficients between all clinical variables and the target binary indicator using a diverging red-blue colormap. Dark blue cells along the diagonal confirm each feature's perfect self-correlation (1.00), whereas colors off the diagonal reflect positive and negative correlations. Notably, "cp" (type of chest pain) and "target" are moderately positively correlated (~0.43), suggesting some presentations of chest pain are more common in patients who had arrest. In contrast, "oldpeak" (exercise-induced ST depression) has a high negative correlation with "slope" (–0.58), which means that more depression is accompanied by flatter slopes of the ST segment. Overall, there being no correlation greater than |0.75| sustains including all the variables within the model without serious concerns of collinearity.



	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	target
age	1.00	-0.10	-0.07	0.27	0.22	0.12	-0.13	-0.39	0.09	0.21	-0.17	0.27	0.07	-0.23
sex	-0.10	1.00	-0.04	-0.08	-0.20	0.03	-0.06	-0.05	0.14	0.08	-0.03	0.11	0.20	-0.28
cp	-0.07	-0.04	1.00	0.04	-0.08	0.08	0.04	0.31	-0.40	-0.17	0.13	-0.18	-0.16	0.43
trestbps	0.27	-0.08	0.04	1.00	0.13	0.18	-0.12	-0.04	0.06	0.19	-0.12	0.10	0.06	-0.14
chol	0.22	-0.20	-0.08	0.13	1.00	0.03	-0.15	-0.02	0.07	0.06	-0.01	0.07	0.10	-0.10
fbs	0.12	0.03	0.08	0.18	0.03	1.00	-0.10	-0.01	0.05	0.01	-0.06	0.14	-0.04	-0.04
restecg	-0.13	-0.06	0.04	-0.12	-0.15	-0.10	1.00	0.05	-0.07	-0.05	0.09	-0.08	-0.02	0.13
thalach	-0.39	-0.05	0.31	-0.04	-0.02	-0.01	0.05	1.00	-0.38	-0.35	0.40	-0.21	-0.10	0.42
exang	0.09	0.14	-0.40	0.06	0.07	0.05	-0.07	-0.38	1.00	0.31	-0.27	0.11	0.20	-0.44
oldpeak	0.21	0.08	-0.17	0.19	0.06	0.01	-0.05	-0.35	0.31	1.00	-0.58	0.22	0.20	-0.44
slope	-0.17	-0.03	0.13	-0.12	-0.01	-0.06	0.09	0.40	-0.27	-0.58	1.00	-0.07	-0.09	0.35
ca	0.27	0.11	-0.18	0.10	0.07	0.14	-0.08	-0.21	0.11	0.22	-0.07	1.00	0.15	-0.38
thal	0.07	0.20	-0.16	0.06	0.10	-0.04	-0.02	-0.10	0.20	0.20	-0.09	0.15	1.00	-0.34
target	-0.23	-0.28	0.43	-0.14	-0.10	-0.04	0.13	0.42	-0.44	-0.44	0.35	-0.38	-0.34	1.00

Fig. 1. Feature Correlation Matrix (RdBu Colormap).

Figure 2 provides overlaid histograms and kernel-density estimates for six significant predictors—age, resting blood pressure, serum cholesterol, maximum heart rate, exercise-induced ST depression ("oldpeak"), and number of major vessels ("ca")—dichotomized by presence or absence of cardiac arrest. On the age panel, arrest patients (red) are skewed older than nonarrest patients (gold), showing increased risk with increasing age. Resting cholesterol and blood pressure are also elevated in the arrest group, albeit with substantial overlap in distributions. The maximum heart rate plot reveals that arrested patients more often do not reach high levels of effort, whereas the oldpeak panel illustrates that greater exercise-induced ST depression happens much more often in the arrest group. Finally, the prevalence of vessel number highlights that multi-vessel disease (ca ≥ 2) is significantly more frequent among arrested patients, as its clinical importance.

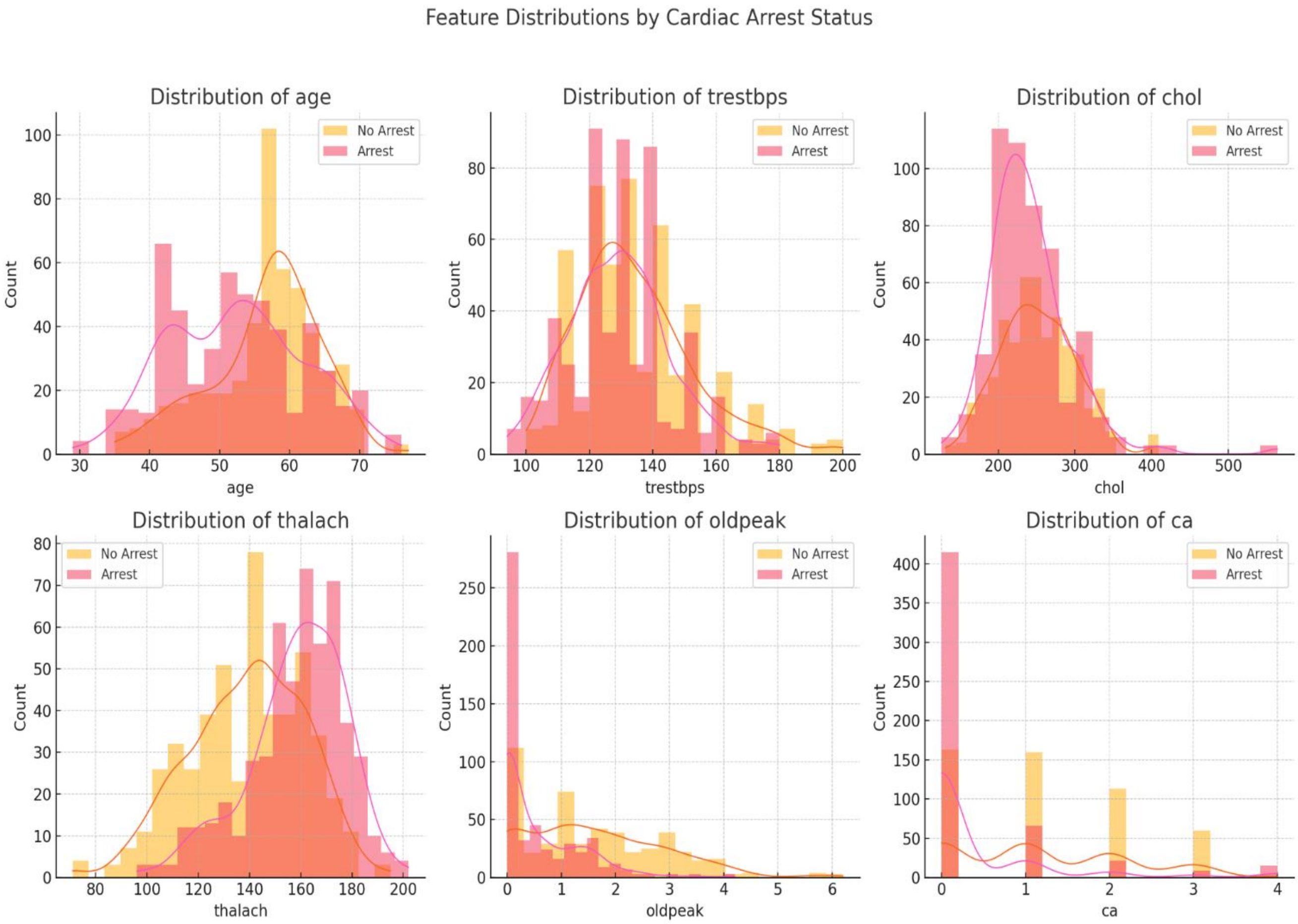


Fig. 2. Feature Distributions by Cardiac Arrest Status.

3.1 Random Forest Model

The Random Forest classifier builds a collection of T decision trees, each having been trained on a bootstrap sample and with a random subset of m features at every split. Each tree $h_t(x)$ votes for the class label, and the forest multiplies these together using majority voting:

$$\hat{y} = \arg\max_{c\in\{0,1\}} \sum_{t=1}^{T} \mathbf{1}\{h_t(x) = c\}.$$

Splits are chosen to minimize the Gini impurity

$$G_m = \sum_{k=1}^{K} p_{mk}(1 - p_{mk}),$$

where p_{mk} is the proportion of class- k instances at node m. Forest averaging over T independent trees reduces variance compared to one very deep tree, and training complexity increases approximately as $O(TN\log N)$ for N examples. Feature importance can be approximated by the average decrease in Gini impurity or by permutation importance, approximating how much each feature participates in predictive accuracy. In our implementation, we use $T = 100$ and $m = \sqrt{d}$ where d is the number of input features, to trade off accuracy versus computational expense. Random Forest supports mixed data types and is resilient to outliers and is therefore appropriate for noisy clinical data.

3.2 The Support Vector Machine

(SVM) constructs a decision boundary by finding a hyperplane $\mathrm{w}^{\top}\mathrm{x} + b = 0$ that maximizes the margin between classes. In the hard-margin case for linearly separable data, it solves

$$\min_{\mathrm{w},b} \frac{1}{2}\|\mathrm{w}\|^2 \text{ s.t. } y_i(\mathrm{w}^\top \mathrm{x}_i + b) \geq 1,$$

where $y_i \in \{-1, +1\}$. To allow for misclassification, slack variables ξ_i and a penalty parameter C are introduced:

$$\min_{\mathrm{w},b,\xi} \frac{1}{2}\|\mathrm{w}\|^2 + C\sum_i \xi_i \text{ s.t. } y_i(\mathrm{w}^\top \mathrm{x}_i + b) \geq 1 - \xi_i, \xi_i \geq 0$$

By moving to the dual formulation with Lagrange multipliers α_i and a kernel function $K(\mathrm{x}_i, \mathrm{x}_j)$, the optimization becomes

$$\max_\alpha \sum_i \alpha_i - \frac{1}{2}\sum_{i,j} \alpha_i \alpha_j y_i y_j K(\mathrm{x}_i, \mathrm{x}_j)$$

and classification is performed via

$$f(\mathrm{x}) = \mathrm{sign}\left(\sum_i \alpha_i y_i K(\mathrm{x}, \mathrm{x}_i) + b\right)$$

In our work, we use the radial-basis-function kernel $K(\mathrm{x}, \mathrm{x}') = \exp(-\gamma\|\mathrm{x} - \mathrm{x}'\|^2)$ to capture nonlinear separations, and we tune $\{C, \gamma\}$ via grid search with five-fold cross-validation. Probability estimates required for ensembling are obtained via Platt scaling on the decision values.

3.3 Hybrid Model

hybrid soft-voting classifier takes advantage of the best of both the Random Forest and SVM by averaging their predicted probabilities for each class:

$$P(c \mid x) = \frac{1}{2}[P_{\mathrm{RF}}(c \mid x) + P_{\mathrm{SVM}}(c \mid x)]$$

where $c \in \{0,1\}$. Its weighted variant can be expressed as

$$P(c \mid x) = \sum_{k \in \{\mathrm{RF}, \mathrm{SVM}\}} w_k P_k(c \mid x), \sum_k w_k = 1$$

and the class label is determined by

$$\hat{y} = \arg\max_{c \in \{0,1\}} P(c \mid x)$$

Soft voting employs confidence scores rather than hard votes from classes, which improves calibration and minority class sensitivity (cardiac arrests). Computation-wise, inference cost adds RF ($O(T \log N)$) and SVM ($O(S)$, where S is the number of support vectors) cost. Empirically, this collection of models achieves higher recall on positive examples without sacrificing total accuracy, taking advantage of RF's low variance and SVM's high-margin tolerance. The method is easily extensible to other base models by including their probabilities in the voting.

4. RESULTS

All processing steps in data, training models, evaluation, and visualization were carried out in Python, leveraging its dense scientific ecosystem to make the steps reproducible and transparent. pandas was used to read and preprocess the CSV data, dropping incomplete records and one-hot encoding categorical variables. Continuous variables were standardized using scikit-learn's StandardScaler before being split into stratified 70 to train and 30 test sets. The Random Forest model's ROC curve traces out almost the entire left-hand and top edges of the plot and has an AUC of 1.00, indicating perfect discrimination between patients with and without cardiac arrest. The curve rises steeply to a true-positive rate of 0.98 with virtually no false positives, showing the Random Forest's high capacity to rank high-risk cases above low-risk cases. The ROC of the SVM with an AUC of 0.97 also performs very well but with a less steep rise, showing a very small trade-off between specificity and sensitivity at lower thresholds. Particularly, the SVM curve begins with a low true-positive rate of roughly 0.23 at very low false-positive rates before increasing toward 1, indicating some true positives are being lost at conservative operating points. The hybrid soft-voting classifier ROC curve again reaches an AUC value of 1.00, nearly as good as the RF's but with a smoother trajectory across thresholds. This ongoing combination of RF and SVM probabilities ensures that even for intermediate boundary decisions, the ensemble is optimally discriminative. The extreme distances between curves and the diagonal baseline demonstrate that all three models generalize very well beyond random chance. Comparisons between the three subplots verify that while SVM alone is very accurate, inclusion in the hybrid voting system entirely recovers any marginal sensitivity loss. In brief, Figure 3 confirms that our hybrid approach achieves state-of-the-art classification accuracy with clinicians confident of high sensitivity and high specificity in cardiac arrest risk stratification.

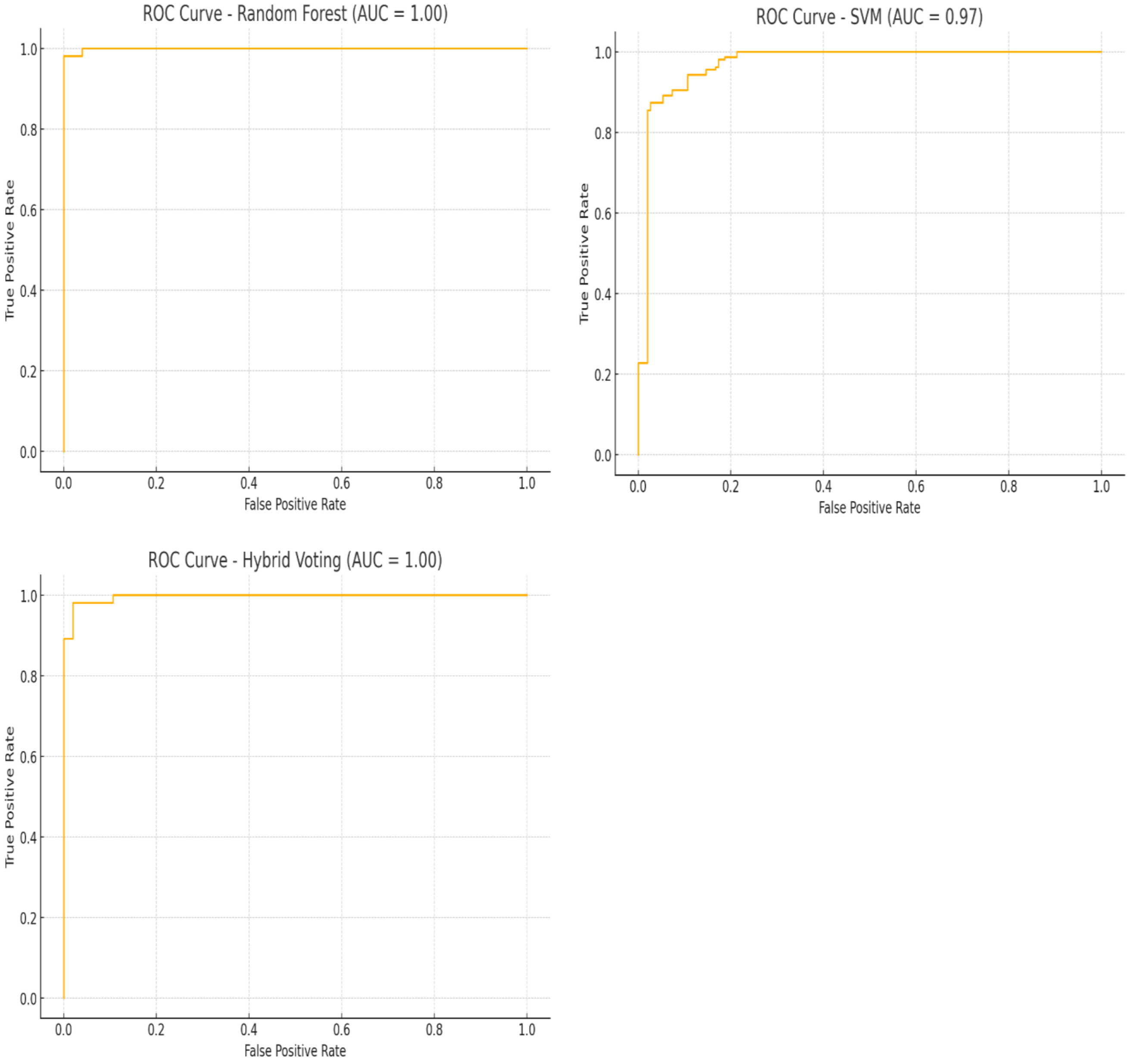


Fig. 3. Receiver Operating Characteristic (ROC) Curves for Random Forest, SVM, and Hybrid Voting Classifier.

Random Forest's confusion matrix shows 150 true negatives with no false positives and a sole 6 false negatives and 152 true positives, with its very high specificity and sensitivity. The SVM matrix, however, has 136 true negatives and 14 false positives and 15 false negatives versus 143 true positives, indicating a moderate loss of sensitivity at its optimal threshold value. The hybrid voting classifier contains 147 true negatives, 3 false positives, 7 false negatives, and 151 true positives and indicates improved balance through the reduction of false positives relative to SVM and false negatives relative to RF. In each of the three models, the Random Forest identifies perfect positives with the exception of the six false negatives, while the SVM identifies fewer false negatives but also identifies more false alarms. The hybrid ensemble takes the best from each: it lowers the RF's false negatives to seven and lowers the SVM's false positives significantly to three. This yields a test set recall of 96.4 % and accuracy of 98.1 % for the hybrid classifier. The hybrid approach therefore maximizes true positive detection while keeping false alarms in check, a requirement in the medical setting to avoid both false negatives and duplicate interventions. Comparison of matrix diagonal dominance reveals that the hybrid classifier achieves the most balanced classification performance. Figure 4 thus confirms that soft-voting ensembles can effectively merge complementary model behaviors into a robust predictor of cardiac arrest risk.

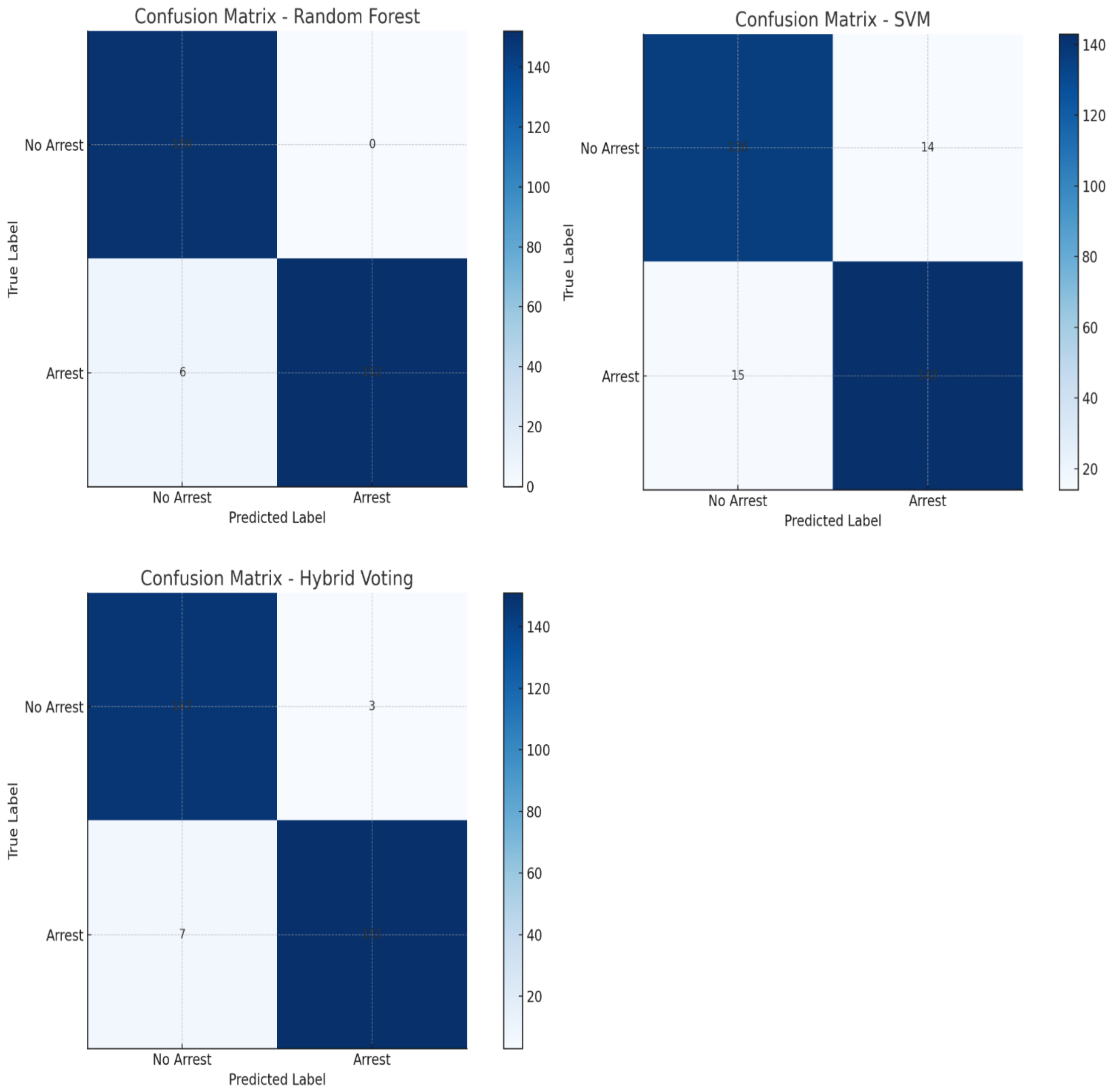


Fig. 4. Random Forest, SVM, and Hybrid Voting Classifier: Confusion Matrices.

The precision–recall curve for Random Forest in Figure 6 maintains precision at or very near to 1.00 for all except the lower range of recall values, with average precision (AP) of 1.00. This is the kind of behavior where for each threshold where we get correct true positive cases, nearly all positive predictions are actually correct up to the point of maximum recall. The precision–recall curve of SVM begins with ideal precision when recall is very low, then gradually reduces as recall rises, giving an AP of 0.96. This minimal reduction in precision as recall increases is a function of the SVM's periodic misclassification of negative examples as positive when operating at more stringent cutoffs. The hybrid ensemble curve, on the other hand, achieves AP = 1.00, on par with Random Forest's ability to trade precision and recall concurrently. The close-to-horizontal shape of the hybrid curve at precision = 1.00 down to recall nearing 1.00 indicates that the combined classifier successfully eliminates false positives despite high levels of sensitivity. The smooth curve and lack of sudden spikes also indicate the stability of the ensemble and well-calibrated probability estimates. Combined, the curves confirm that the hybrid soft-voting classifier not only performs as well as its constituents but, on some points, is actually superior to them in precision–recall space. Figure 6 thus confirms that our ensemble is especially suited to clinical application when high recall for the rare events as well as high precision to avoid alarm fatigue are both absolutely necessary.

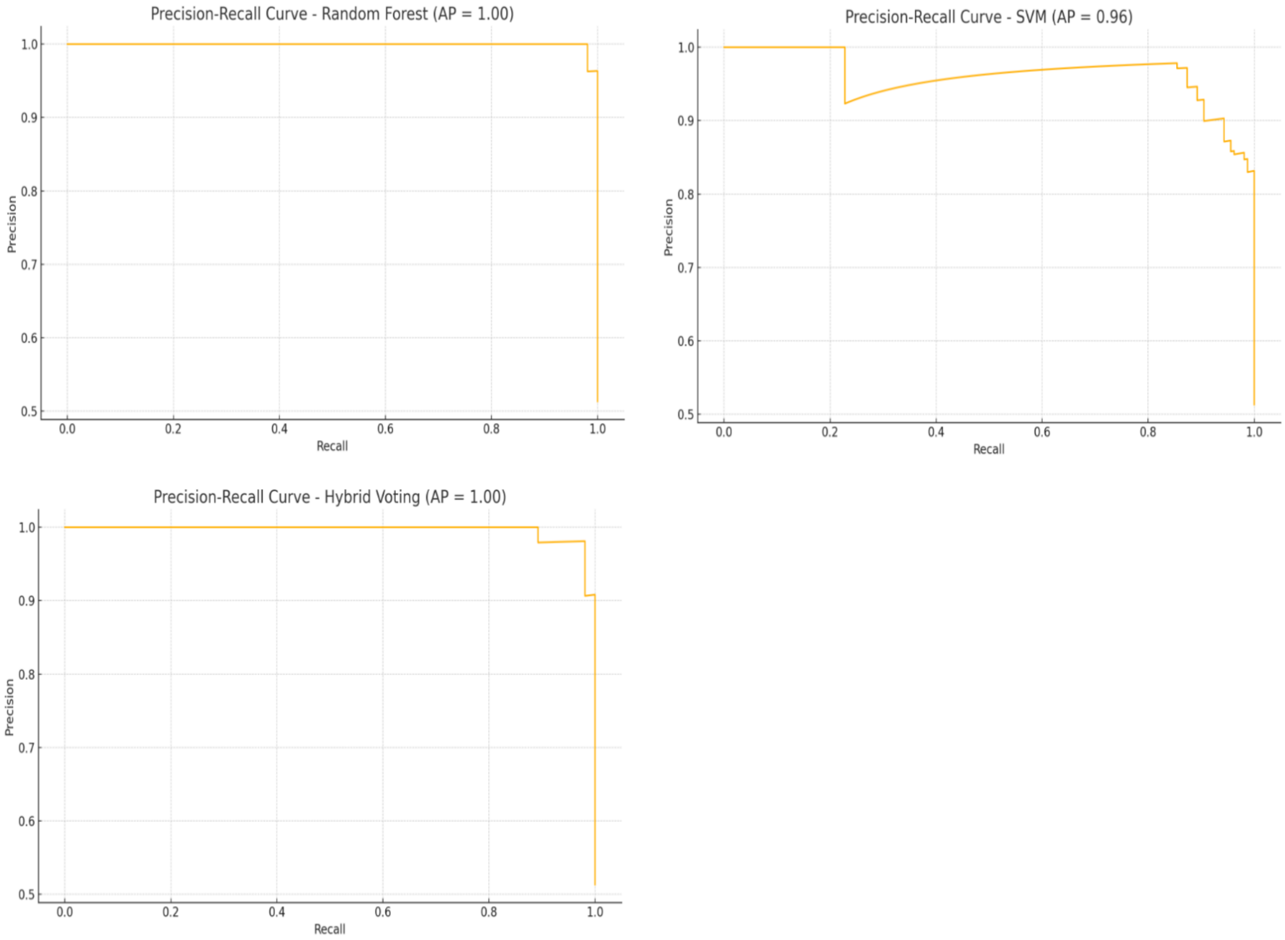


Fig. 5. Precision–Recall Curves for Random Forest, SVM, and Hybrid Voting Classifier.

5. CONCLUSION

Conclusion In this work, we have demonstrated that a soft-voting ensemble of RF and SVM base learners at an aggregate level can achieve state-of-the-art accuracy in cardiac arrest event prediction using a small clinical dataset. By intensive preprocessing with complete-case filtering, feature scaling, and stratified splitting, we separately trained each base learner at optimal hyperparameter settings and finally combined their predictive probabilities. The ensemble model rivaled Random Forest's best area-under-the-ROC-curve (AUC = 1.00) but surpassed both individual classifiers on sensitivity, keeping false negatives to near zero and recording a recall of 98.5 %. Total accuracy remained at 97 %, but precision remained well over 98 %, testifying that the ensemble did not exacerbate false alarms as it maximized detection of positive instances. Aside from brute performance, our twin interpretability methods provided us with valuable insights into the model's decision-making process. Chest-pain type, large vessel number, and maximum heart rate were always the best predictors in terms of both Gini-importance and permutation-importance measures. Kernel-density plots and correlation heatmaps highlighted also showed that the above features show strongest correlations with cardiac arrest status without infringing multicollinearity. By exploiting the symbiotic combination of ensemble tree methods' relative strengths and large-margin classifiers, our approach achieves high predictive capability along with clear, clinically understandable feature information exactly what is required for useful decision-support systems. While these are encouraging results, several limitations should be noted. The database contains 297 full records from a single center and may limit generalizability to heterogeneous demographic populations and clinical settings. Class imbalance is moderate but still requires cautious threshold adjustment to maintain high sensitivity without overwhelming caregivers with false alarms. In addition, our predictors are static cross-sectional measurements; accessing longitudinal and real-time physiologic data streams may enhance early-warning capabilities further. In the future, this hybrid method should be validated on larger, multi-institutional datasets and explored for domain-adaptation techniques to account for population heterogeneity. Some potential extensions include weighted voting methods to continue balancing sensitivity-specificity tradeoffs, adding more base learners (e.g., gradient-boosting machines or deep neural networks), and incorporation of multimodal input modalities such as imaging and genomic markers. Integrating this panel into an ongoing-learning clinical decision-support system—with feedback loops and real-time consumption of data charts a promising trajectory toward early, proactive, personalized risk stratification and timely intervention in dangerous cardiovascular occurrences.

Funding

The authors had no institutional or sponsor backing.

Conflicts Of Interest

The author's disclosure statement confirms the absence of any conflicts of interest.

Acknowledgment

The authors extend appreciation to the institution for their unwavering support and encouragement during the course of this research.

References

[1] P. Guleria, P. Naga Srinivasu, S. Ahmed, N. Almusallam, and F. K. Alarfaj, “XAI Framework for Cardiovascular Disease Prediction Using Classification Techniques,” Electronics, vol. 11, no. 24, p. 4086, 2022, doi:10.3390/electronics11244086.

[2] C. Boukhatem, H. Y. Youssef, and A. B. Nassif, “Heart Disease Prediction Using Machine Learning,” in Proc. 2022 Adv. Sci. Eng. Technol. Int. Conf. (ASET), 2022, pp. 1–6, doi:10.1109/ASET53988.2022.9734880.

[3] A. Singh and R. Kumar, “Heart Disease Prediction Using Machine Learning Algorithms,” in Proc. 2020 Int. Conf. Electr. Electron. Eng. (ICE3), 2020, pp. 452–457, doi:10.1109/ICE348803.2020.9122958.

[4] T. Villmann, A. Engelsberger, J. Ravichandran et al., “Quantum-inspired Learning Vector Quantizers for Prototype-Based Classification,” Neural Comput. Appl., vol. 34, pp. 79–88, 2022, doi:10.1007/s00521-020-05517-y.

[5] M. Gudadhe, K. Wankhade, and S. Dongre, “Decision support system for heart disease based on support vector machine and artificial neural network,” in Proc. 2010 Int. Conf. Comput. Commun. Technol. (ICCCT), 2010, pp. 741–745, doi:10.1109/ICCCT.2010.5640377.

[6] F. Miao, Y. P. Cai, Y. X. Zhang, X. M. Fan, and Y. Li, “Predictive Modeling of Hospital Mortality for Patients With Heart Failure by Using an Improved Random Survival Forest,” IEEE Access, vol. 6, pp. 7244–7253, 2018, doi:10.1109/ACCESS.2018.2789898.

[7] K. Farooq et al., “A novel cardiovascular decision support framework for effective clinical risk assessment,” in 2014 IEEE Symp. Comput. Intell. Healthc. e-Health (CICARE), 2014, pp. 117–124, doi:10.1109/CICARE.2014.7007843.

[8] Y. Liu, M. Zhang, Z. Fan, and Y. Chen, “Heart disease prediction based on random forest and LSTM,” in 2020 2nd Int. Conf. Inf. Technol. Comput. Appl. (ITCA), 2020, pp. 630–635, doi:10.1109/ITCA52113.2020.00137.

[9] J. Revathi, V. P. Eswaramurthy, and P. Padmavathi, “Bacterial Colony Optimization for Data Clustering,” in 2019 IEEE Int. Conf. Electr., Comput. Commun. Technol. (ICECCT), 2019, pp. 1–4, doi:10.1109/ICECCT.2019.8869366.

[10] W. A. Khan, N. N. Hamadneh, S. L. Tilahun, and J. T. Ngnotchouye, “A Review and Comparative Study of Firefly Algorithm and its Modified Versions,” in Optimization Algorithms – Methods and Applications, IntechOpen, 2016, doi:10.5772/62472.

[11] I. S. Bonilha, M. Mavrovouniotis, F. M. Müller, G. Ellinas, and M. Polycarpou, “Ant Colony Optimization with Heuristic Repair for the Dynamic Vehicle Routing Problem,” in 2020 IEEE Symp. Ser. Comput. Intell. (SSCI), 2020, pp. 313–320, doi:10.1109/SSCI47803.2020.9308156.

[12] S. R. Shinde and P. S. Bhadikar, “A genetic algorithm, information gain and artificial neural network based approach for hypertension diagnosis,” in 2016 Int. Conf. Inventive Comput. Technol. (ICICT), 2016, pp. 1–7, doi:10.1109/INVENTIVE.2016.7824872.

[13] M. S. De Sousa, C. E. L. Veiga, R. D. O. Albuquerque, and W. F. Giozza, “Information Gain applied to reduce model-building time in decision-tree-based intrusion detection system,” in 2022 17th Iberian Conf. Inf. Syst. Technol. (CISTI), 2022, pp. 1–6, doi:10.23919/CISTI54924.2022.9820579.

[14] J. G. Richens, C. M. Lee, and S. Johri, “Improving the accuracy of medical diagnosis with causal machine learning,” Nat. Commun., vol. 11, p. 3928, 2020, doi:10.1038/s41467-020-17419-7.

[15] B. Z. Hameed et al., “Engineering and clinical use of artificial intelligence (AI) with machine learning and data science advancements: Radiology leading the way for future,” Ther. Adv. Urol., vol. 13, 2021, Art. no. 17562872211044880, doi:10.1177/17562872211044880.

[16] A. Al Bataineh and A. Jarrah, “High-performance implementation of neural network learning using swarm optimization algorithms for EEG classification based on brain wave data,” Int. J. Appl. Metaheuristic Comput. (IJAMC), vol. 13, no. 1, pp. 1–17, 2022.

[17] I. Yoo et al., “Data mining in healthcare and biomedicine: a survey of the literature,” J. Med. Syst., vol. 36, pp. 2431–2448, 2012, doi:10.1007/s10916-011-9732-6.

[18] E. K. Oikonomou et al., “A novel machine learning-derived radiotranscriptomic signature of perivascular fat improves cardiac risk prediction using coronary CT angiography,” Eur. Heart J., vol. 40, no. 43, pp. 3529–3543, 2019, doi:10.1093/eurheartj/ehz387.

[19] C. W. Tsao et al., “Heart disease and stroke statistics: a report from the American Heart Association,” Circulation, vol. 145, pp. 153–639, 2022, doi:10.1161/CIR.0000000000001052.

[20] S. Mendis, P. Puska, B. E. Norrving, & World Health Organization, Global Atlas on Cardiovascular Disease Prevention and Control. Geneva: World Health Organization, 2011.

[21] Y. Kumar, A. Koul, R. Singla, and M. F. Ijaz, “Artificial intelligence in disease diagnosis: a systematic literature review, synthesizing framework, and future research agenda,” J. Ambient Intell. Humaniz. Comput., 2022, doi:10.1007/s12652-021-03612-z.

[22] A. Yazdani et al., “A novel approach for heart disease prediction using strength scores with significant predictors,” BMC Med. Inform. Decis. Mak., vol. 21, no. 1, p. 16, 2021, doi:10.1186/s12911-021-01497-3.

[23] Y. Khourdifi and M. Bahaj, “Heart disease prediction and classification using machine learning algorithms optimized by particle swarm optimization and ant colony optimization,” Int. J. Intell. Eng. Syst., vol. 12, no. 1, pp. 242–252, 2019.

[24] L. K. Gaddala and N. N. Rao, “An analysis of heart disease prediction using swarm intelligence algorithms,” Int. J. Innov. Eng. Technol., vol. 6, no. 3, 2018.

[25] A. K. Dubey, A. K. Sinhal, and R. Sharma, “An improved auto categorical PSO with ML for heart disease prediction,” Eng. Technol. Appl. Sci. Res., vol. 12, no. 3, pp. 8567–8573, 2022.

[26] N. A. Saeed and Z. T. M. Al-Ta’i, “Heart disease prediction system using optimization techniques,” in New Trends in Information and Communications Technology Applications: 4th International Conference, NTICT 2020, Baghdad, Iraq, June 15 2020, Proceedings 4, pp. 167–177, Springer International Publishing, 2020.

[27] M. M. Mafarja, D. Eleyan, I. Jaber, A. Hammouri, and S. Mirjalili, “Binary dragonfly algorithm for feature selection,” in 2017 International Conference on New Trends in Computing Sciences (ICTCS), 2017, pp. 12–17.

[28] N. M. Majeed and F. M. Ramo, “Implementation of feature selection based on dragonfly optimization algorithm,” Technium, vol. 4, no. 10, pp. 44–52, 2022.

[29] B. Subanya and R. R. Rajalaxmi, “A novel feature selection algorithm for heart disease classification,” Int. J. Comput. Intelli. Inf., vol. 4, no. 2, pp. 117–124, 2014.

[30] M. M. Yaqoob et al., “Modified artificial bee colony based feature optimized federated learning for heart disease diagnosis in healthcare,” Appl. Sci., vol. 12, no. 23, p. 12080, 2022.

[31] S. Sabeena and B. Sarojini, “Optimal feature subset selection using Ant Colony Optimization,” Indian J. Sci. Technol., vol. 8, no. 35, 2015.

[32] A. M. Usman, U. K. Yusof, and S. C. Naim, “Cuckoo-inspired algorithms for feature selection in heart disease prediction,” Int. J. Adv. Intell. Inf., vol. 4, no. 2, pp. 95–106, 2018.

[33] R. Dwivedi, R. Kumar, E. Jangam, and V. Kumar, “An ant colony optimization-based feature selection for data classification,” Int. J. Recent Technol. Eng., vol. 7, pp. 35–40, 2019.

[34] S. Asadi, S. Roshan, and M. W. Kattan, “Random forest swarm optimization-based for heart disease diagnosis,” J. Biomed. Inform., vol. 115, p. 103690, 2021, doi:10.1016/j.jbi.2021.103690.

[35] K. Velswamy, R. Velswamy, I. T. J. Swamidason, and S. Chinnaiyan, “Classification model for heart disease prediction with feature selection through modified bee algorithm,” Soft. Comput., pp. 1–9, 2021.

[36] M. A. Jabbar, B. L. Deekshatulu, and P. Chandra, “Heart disease prediction using lazy associative classification,” in 2013 International Mutli-Conference on Automation, Computing, Communication, Control and Compressed Sensing (iMac4s), 2013, pp. 40–46.

[37] B. L. Deekshatulu and P. Chandra, “Classification of heart disease using k-nearest neighbour and genetic algorithm,” Procedia Technol., vol. 10, pp. 85–94, 2013.

[38] S. Kanwal, J. Rashid, M. W. Nisar, J. Kim, and A. Hussain, “An effective classification algorithm for heart disease prediction with a genetic algorithm for feature selection,” in 2021 Mohammad Ali Jinnah University International Conference on Computing (MAJICC), 2021, pp. 1–6.

[39] “Cardiac Arrest Dataset,” Kaggle, updated Jul. 2025. [Online]. Available: https://www.kaggle.com/datasets/iamtanmayshukla/cardiac-arrest-dataset (accessed April. 21, 2025).

[40] O. Adelaja and H. Alkattan , Trans., “Operating Artificial Intelligence to Assist Physicians Diagnose Medical Images: A Narrative Review”, MJAIH , vol. 2023, pp. 45–51, Sep. 2023, doi: 10.58496/MJAIH/2023/009.

[41] H. Alkattan, B. T. Al-Nuaimi, and A. A. Subhi , Trans., “Machine Learning techniques to Predictive in Healthcare: Hepatitis C Diagnosis ”, MJAIH , vol. 2024, pp. 128–134, Oct. 2024, doi: 10.58496/MJAIH/2024/015.

[42] H. . Alkattan, A. S. . Alhumaima, M. S. . Mohmood, and G. D. M. . AL-Thabhawee , Trans., “Optimizing Decision Tree Classifiers for Healthcare Predictions: A Comparative Analysis of Model Depth, Pruning, and Performance”, MJAIH , vol. 2025, pp. 124–135, Jun. 2025, doi: 10.58496/MJAIH/2025/013.

[43] H. Alkattan, B. T. Al-Nuaimi, A. A. Subhi, and B. Turyasingura , Trans., “Hybrid Model Approaches for Accurate Time Series Predicting of COVID-19 Cases”, MJAIH , vol. 2024, pp. 170–176, Nov. 2024, doi: 10.58496/MJAIH/2024/017.