Optimizing Decision Tree Hyperparameters via Random Search for Accurate Heart Failure Risk Prediction

Suyahman*[1]

[1]Department of Computer Science, Universitas Sugeng Hartono, Sukoharjo, Indonesia
Email: [1]suyahman@sugenghartono.ac.id

Abstract

Heart failure remains one of the leading causes of mortality worldwide, highlighting the need for reliable early-detection models to support clinical decision-making. This study investigates the effect of Random Search–based hyperparameter optimization on a Decision Tree model for heart failure risk prediction using a clinical dataset comprising 918 samples and 11 demographic and cardiovascular features. Rather than introducing a novel optimization algorithm, this work focuses on analyzing model performance sensitivity to hyperparameter tuning in a real-world medical dataset. The baseline Decision Tree achieved an accuracy of 0.80. After Random Search optimization, accuracy improved to 0.84, while recall for the positive class increased from 0.83 to 0.90, indicating a notable reduction in false-negative predictions. The optimized configuration, characterized by a shallow tree depth and increased minimum samples per leaf, suggests improved generalization and reduced overfitting. Compared with related studies employing ensemble-based models and genetic optimization, the proposed approach achieves competitive performance using a simpler and more interpretable classifier. These findings demonstrate that systematic hyperparameter tuning can substantially enhance the clinical utility of conventional machine learning models. Practically, the improved recall supports the use of the optimized Decision Tree as a screening-oriented decision support tool, enabling earlier identification of high-risk patients while maintaining model transparency. This study highlights the importance of dataset-specific optimization and provides a foundation for future work involving ensemble methods and advanced optimization strategies to develop robust and clinically applicable heart failure prediction systems.

Keywords: *Decision Tree, Heart Failure Prediction, Hyperparameter Optimization, Machine Learning, Random Search.*

1. INTRODUCTION

Early detection of heart failure risk plays a crucial role in preventive care and timely medical intervention, particularly as cardiovascular diseases continue to account for a substantial proportion of global mortality worldwide [1]. In recent years, machine learning and deep learning techniques have been increasingly adopted to support medical decision making by enabling predictive modeling based on clinical and demographic data [2],[3],[4],[5]. These approaches aim to assist clinicians in identifying high risk patients at an early stage, thereby reducing diagnostic delays and improving treatment outcomes.

Several prior studies have explored the application of machine learning and deep learning models for heart disease and cardiovascular risk prediction. Gangadhar et al. investigated multiple machine learning and deep learning algorithms, including Decision Tree, Naive Bayes, and Artificial Neural Networks, and reported that deep learning models generally outperform traditional classifiers in coronary heart disease prediction tasks [6]. Similarly, Garcia Ordas et al. proposed deep learning models combined with feature augmentation techniques and demonstrated notable improvements in predictive precision, highlighting the potential of representation enrichment for cardiovascular risk assessment [7]. While these studies achieved high predictive performance, they primarily focused on deep learning architectures, which often suffer from limited interpretability and higher computational complexity, posing challenges for deployment in routine clinical settings.

In contrast, interpretable machine learning models, particularly Decision Tree based approaches, remain attractive for healthcare applications due to their transparent decision logic. Ozcan and Peker employed the Classification and Regression Tree algorithm on electronic health record data and achieved an accuracy of 87 percent, while also extracting IF THEN decision rules that can be directly interpreted by clinicians [8]. Likewise, Dalal et al. conducted a large scale comparative study using multiple machine learning models on a cardiovascular disease dataset containing approximately 70,000 records, reporting very high predictive accuracy [9]. However, these studies largely emphasize overall performance metrics or rule extraction, with limited investigation into how hyperparameter configurations influence model stability, recall, and false negative rates in clinically sensitive contexts.

Despite the widespread use of Decision Trees in heart disease prediction, many existing studies rely on default parameter settings or manual tuning strategies, which may result in suboptimal model performance. Prior empirical work has shown that exhaustive tuning strategies such as Grid Search become computationally inefficient as the hyperparameter space grows and may still fail to identify optimal configurations in practice [10]. Consequently, the impact of systematic yet efficient hyperparameter optimization methods on Decision Tree performance, particularly for heart failure risk prediction, remains underexplored.

Motivated by these gaps, this study investigates the application of Random Search based hyperparameter optimization to improve Decision Tree performance for heart failure risk prediction using a real world clinical dataset. Rather than proposing a novel learning algorithm, this work focuses on analyzing the sensitivity of Decision Tree performance to hyperparameter configurations and evaluating improvements in clinically relevant metrics, especially recall and false negative reduction [11]. By comparing a baseline model trained with default parameters to an optimized model obtained through Random Search, this study provides empirical evidence on how dataset specific optimization can enhance the clinical utility of interpretable machine learning models.

The main contribution of this research lies in offering an empirically validated hyperparameter configuration for Decision Tree based heart failure prediction, along with a comprehensive analysis of performance changes before and after optimization. This contribution complements existing deep learning oriented studies by demonstrating that carefully optimized, interpretable models can achieve meaningful performance gains while remaining suitable for screening oriented clinical decision support systems.

2. METHOD

This study employs a quantitative machine learning approach to develop a predictive model for assessing heart failure risk using the Decision Tree algorithm. The methodological framework consists of several sequential stages designed to ensure rigorous data preparation, model construction, and performance evaluation. The first stage involves selecting and understanding the dataset, including an exploratory assessment of feature distributions, missing values, and the clinical relevance of each variable. The second stage is data preprocessing, which includes handling missing data, normalizing numerical features when necessary, encoding categorical variables, and partitioning the dataset into training and testing subsets. The third stage focuses on building a baseline Decision Tree model using default hyperparameters to serve as a reference point for subsequent optimization. The fourth stage implements hyperparameter optimization through the Random Search technique, which enables probabilistic exploration of the hyperparameter space and increases the likelihood of identifying optimal configurations while reducing computational time compared to exhaustive search methods. Key hyperparameters evaluated include max_depth, min_samples_split, min_samples_leaf, and criterion. The fifth stage consists of model evaluation using multiple performance metrics, including accuracy, precision, recall, F1-score, and confusion matrix analysis to assess predictive capability and classification robustness. The overall methodological workflow is illustrated in Figure 1.
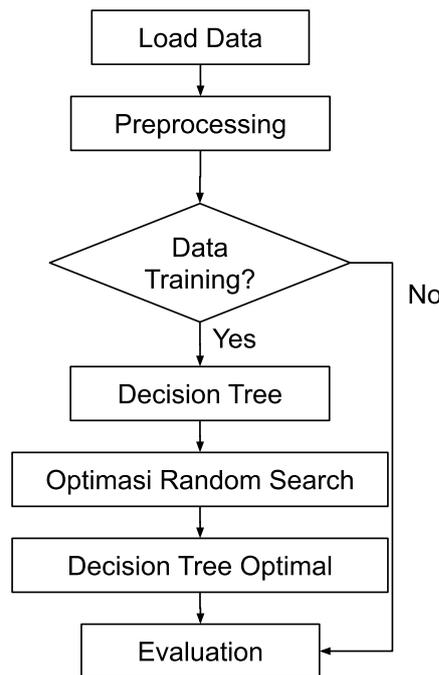
Figure 1. Workflow of Decision Tree Hyperparameter Optimization Using Random Search

2.1. Dataset

This study utilizes the Heart Failure Dataset available on the Kaggle platform (https://www.kaggle.com/datasets/tan5577/heart-failure-dataset). Heart failure represents a common manifestation of cardiovascular disease (CVD), and early detection of its associated risk factors is crucial for both prevention and timely medical intervention. After duplicate removal, the dataset consists of 918 patient records, making it one of the largest consolidated heart disease datasets for predictive modeling [12],[13].

The dataset contains 11 clinical and demographic features used to predict the likelihood of heart disease in patients. Specifically, the features include: patient age (Age), biological sex (Sex), type of chest pain (Chest Pain Type), resting blood pressure (Resting BP), cholesterol level (Cholesterol), fasting blood sugar (Fasting BS), results of the resting electrocardiogram (Resting ECG), maximum heart rate achieved (Max HR), presence of exercise-induced angina (Exercise Angina), ST depression value (Old Peak), and slope of the ST segment during exercise (ST Slope).

The output label consists of a binary variable, Heart Disease, indicating whether heart disease is present (1) or absent (0). The class distribution shows a moderate imbalance, with heart disease cases comprising approximately 55% of the samples and non-heart disease cases approximately 45% [14],[15].

Based on its structure and feature completeness, this dataset is highly suitable for machine learning development, particularly for the Decision Tree algorithm, which performs effectively on tabular data with clinically interpretable variables [16]. Moreover, the predictors included in this dataset have been widely used in medical literature for evaluating patient risk factors, enabling this study to produce a predictive model that supports the early identification of individuals at elevated risk.

2.2. Preprocessing

The preprocessing stage was conducted to prepare the dataset for constructing the heart failure risk prediction model. The initial step involved examining the data structure to identify variable types, confirm the number of observations, and verify the absence of missing values or anomalous records [17],[18]. All attributes were found to be complete, eliminating the need for data imputation or removal.

Categorical variables, including sex, chest pain type, resting electrocardiogram condition, exercise-induced angina, and ST segment slope, were transformed into numerical representations using label encoding. This transformation was required because the Decision Tree algorithm operates on

numerical inputs for impurity-based split selection. Label encoding was preferred over one-hot encoding to maintain feature compactness and avoid unnecessary tree fragmentation.

The dataset was then separated into predictor features and the binary target variable representing the presence or absence of heart disease. To ensure reliable and unbiased model evaluation, the dataset was divided into training and testing subsets using an 80:20 stratified sampling strategy, preserving the original class distribution in both subsets [19]. This approach is particularly important given the moderate class imbalance observed in the dataset.

Overall, the preprocessing procedures ensured that the dataset was clean, structured, and appropriately balanced for model training and evaluation, thereby supporting the development of a reliable and generalizable heart failure risk prediction model.

2.3. Decision Tree

The baseline model in this study utilizes the Decision Tree Classifier without any hyperparameter optimization, serving as a benchmark to evaluate the effectiveness of Random Search in improving predictive performance [20],[21],[22]. The Decision Tree algorithm was selected due to its interpretability, its ability to handle both categorical and numerical variables, and its independence from assumptions regarding data distribution [23],[24]. In its implementation, the model was constructed using the default parameters provided by the scikit-learn library, which typically include Gini impurity as the splitting criterion, the best splitter for attribute selection, and an unrestricted tree depth. Such default configurations often lead to overfitting, particularly in medical datasets characterized by high inter-patient variability and interacting clinical features.

The decision tree is built using a top-down recursive partitioning strategy. At each node, the algorithm selects a feature and split point that minimize node impurity [25]. The Gini impurity of a dataset D is defined in Equation (1) as

$$G(D) = 1 - \sum_{K=1}^{K} p_k^2 \qquad (1)$$

The optimal split S^* is determined by minimizing this value, as expressed in Equation (2)

$$S^* = \arg min_s G_{split} \ (D, S) \qquad (2)$$

This recursive process continues until a stopping condition is reached, such as all samples in a node belonging to the same class or no further meaningful splits being available [26].

2.4. Random Search

Following the evaluation of the baseline model, hyperparameter optimization was performed using the Random Search approach [27]. This method was selected due to its computational efficiency compared to exhaustive Grid Search, particularly when exploring multi-dimensional hyperparameter spaces [28]. Random Search enables probabilistic exploration of diverse parameter combinations, allowing effective identification of high-performing configurations with reduced computational cost [29].

In this study, the hyperparameter search space was defined to include key parameters that directly influence the structure and predictive performance of the Decision Tree. These parameters include maximum tree depth, which controls model complexity, and the minimum number of samples required to split an internal node, which serves as a regularization mechanism to mitigate overfitting. The splitting criterion was also varied between commonly used impurity measures to examine its impact on classification performance.

The Random Search procedure was conducted by evaluating 50 randomly sampled hyperparameter configurations, referred to as search iterations, under a K-Fold cross-validation scheme. The number of iterations was selected to ensure sufficient exploration of the hyperparameter space while maintaining computational efficiency. Cross-validation was employed to provide stable performance estimation and reduce bias associated with a single data partition. To ensure reproducibility, a consistent random initialization was applied throughout the optimization process.

Each iteration produced a Decision Tree model trained with a unique hyperparameter configuration, and its performance was assessed based on average cross-validation metrics. The

configuration yielding the highest validation performance was selected as the optimal model, which was subsequently evaluated on the independent test dataset. This optimized model demonstrates improved generalization capability compared to the baseline Decision Tree, particularly in terms of recall and overall predictive robustness.

2.5. Evaluation

Model evaluation was conducted to assess the performance of the Decision Tree both before and after hyperparameter optimization using Random Search. The evaluation process was performed using the test dataset, which is separate from the training data, ensuring that the performance metrics reflect the model's ability to generalize to unseen data. To measure classification quality, this study employs several commonly used metrics in machine learning, particularly those relevant for disease prediction tasks where clinical consequences are significant.

The first metric used is the confusion matrix, a two-dimensional representation that compares the model's predictions with the actual outcomes [30]. The confusion matrix provides detailed information regarding the number of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). In the context of heart failure risk prediction, the FN value is especially critical, as it represents high-risk patients who are not correctly identified by the model. Therefore, the confusion matrix is an essential tool for understanding the model's error patterns.

In addition, the model's performance was evaluated using several metrics derived from the confusion matrix. Accuracy measures the proportion of correct predictions relative to the total number of observations. However, because accuracy may be insensitive to class imbalance, this study also analyzes precision and recall. Precision reflects the model's ability to correctly identify positive cases among all predicted positives, while recall measures the model's capability to detect all actual positive cases. In medical applications, recall is particularly important because it indicates how effectively the model identifies at-risk patients [31].

To balance precision and recall, the F1-score, defined as the harmonic mean of both metrics is used. The F1-score provides a more comprehensive assessment, especially in datasets with uneven class distributions. Through the combined use of accuracy, precision, recall, F1-score, and confusion matrix analysis, this study evaluates the model thoroughly and compares the performance improvement between the baseline model and the optimized model. This evaluation serves as a crucial basis for determining the effectiveness of Random Search in producing a more accurate and reliable Decision Tree model for predicting heart failure risk.

3. RESULT

The experimental evaluation began with the baseline Decision Tree model trained using default hyperparameters. The baseline achieved an overall accuracy of 0.8098 on the test dataset, demonstrating a reasonable initial performance for heart failure risk prediction. Detailed classification metrics are summarized in Table 1, where precision, recall, and F1-score all reached 0.83, indicating balanced predictive capability across classes.

Table 1. Performance of the Baseline Decision Tree Model

Metric	Value
Accuracy	0.8098
Precision	0.83
Recall	0.83
F1-score	0.83

Visualization of the learned Decision Tree reveals that ST_Slope emerged as the root node, indicating its dominant role in separating heart failure and non-heart failure cases. Subsequent splits involved clinically relevant attributes such as ChestPainType, ExerciseAngina, and Oldpeak, reflecting their contribution to risk stratification. The shallow tree structure suggests that the baseline model relies on a small number of highly informative features to make predictions, providing an interpretable

representation of decision pathways and highlighting the medical relevance of electrocardiographic and exercise-induced indicators. Details of the Decision Tree visualization can be seen in Figure 2.
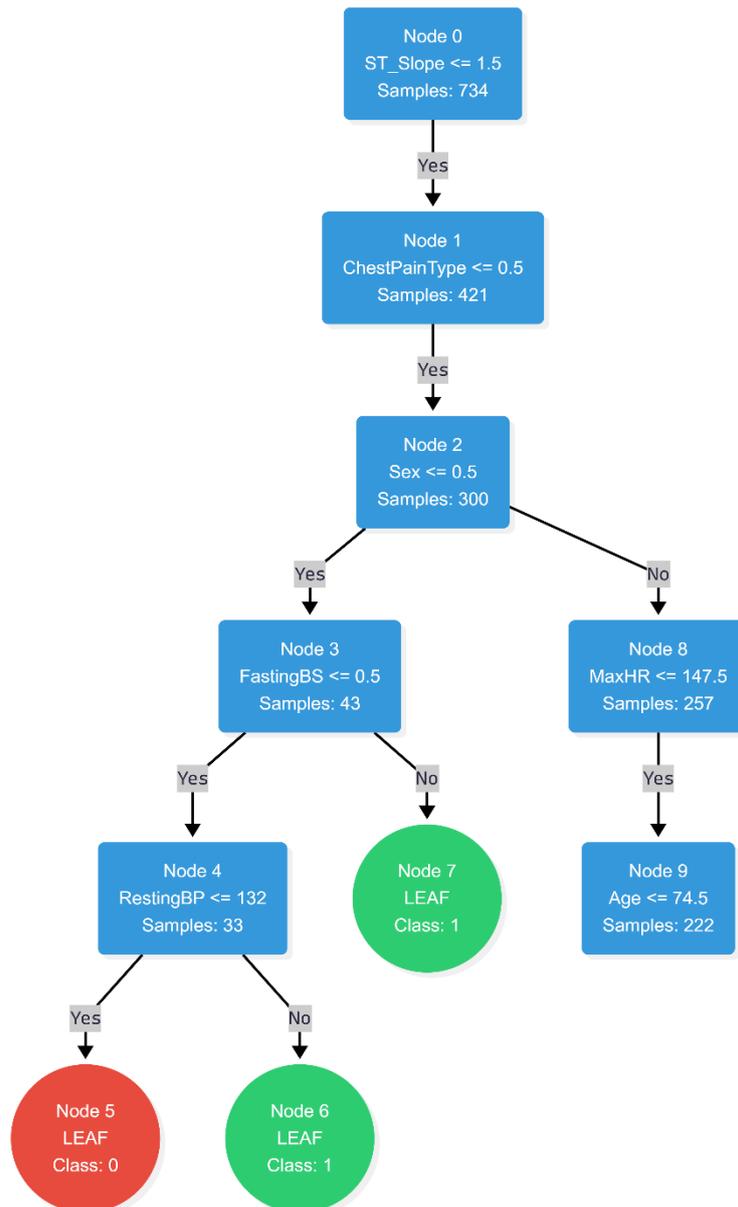

Figure 2. Decision Tree Visualization

Further insight into baseline behavior was obtained through confusion matrix analysis, as shown in Figure 3. The baseline model correctly identified 85 positive cases (true positives) and 18 negative cases (true negatives), while producing 64 false positives and 17 false negatives. Although the model showed acceptable sensitivity, the presence of 17 false negatives indicates that a notable number of heart failure cases remained undetected, which is undesirable in clinical decision-support contexts.

Confusion Matrix

Figure 3. Confusion Matrix of the Baseline Model

To improve predictive performance, Random Search was applied to optimize the Decision Tree hyperparameters. The search process explored multiple parameter combinations and selected entropy as the splitting criterion, a maximum tree depth of 3, a minimum of 10 samples per leaf, a minimum split size of 2, and log2 as the maximum feature selection strategy. These optimized parameters are presented in Table 2.

Table 2. Best Hyperparameters Selected by Random Search

Hyperparameter	Value
criterion	entropy
max_depth	3
min_samples_leaf	10
min_samples_split	2
max_features	log2

After applying the selected hyperparameter configuration, the optimized Decision Tree model demonstrated a substantial improvement in performance. As reported in Table 4, the overall accuracy increased to 0.8478, representing a gain of approximately 3.8 percentage points over the baseline. Precision for the positive class reached 0.84, while recall increased notably to 0.90, resulting in an F1-score of 0.87.

Table 3. Performance After Random Search Optimization

Metric	Value
Accuracy	0.8478
Precision (Class 1)	0.84
Recall (Class 1)	0.90
F1-score (Class 1)	0.87

The optimized confusion matrix further illustrates these improvements. The number of true positives increased from 85 to 92, while false negatives decreased from 17 to 10. True negatives remained constant at 18, and false positives stayed at 64. This shift indicates that Random Search optimization primarily enhanced the model's ability to correctly identify positive heart failure cases, significantly reducing missed diagnoses. The confusion matrix resulting from the optimization can be seen in Figure 4.

Figure 4. Confusion Matrix After Random Search Optimization

A comparative analysis with prior studies is presented in Table 4. Togatoropa et al. achieved accuracies ranging from 81.67% using Random Search to 85.83% using Genetic Algorithm–optimized Random Forest, while Rimal et al. reported baseline Random Forest accuracy of 78.00%, improving to 82.50% with Genetic Algorithm tuning [32],[33]. In contrast, the proposed Decision Tree combined with Random Search attained an accuracy of 84.78%, demonstrating competitive performance despite employing a simpler single-tree architecture.

Table 4. Performance After Random Search Optimization

Model	Accuracy
Random Forest (Togatoropa et al.)	82.50%
RF + Random Search (Togatoropa et al.)	81.67%
RF + Grid Search (Togatoropa et al.)	83.33%
RF + Genetic Algorithm (Togatoropa et al.)	85.83%
Random Forest (Rimal et al.)	78.00%
RF + Manual Tuning (Rimal et al.)	81.10%
RF + Randomized Search (Rimal et al.)	81.10%
RF + Genetic Algorithm (Rimal et al.)	82.50%
Decision Tree (This Research)	80.98%
Decision Tree + Random S. (This Research)	**84.78%**

These results highlight that although ensemble-based methods such as Random Forest with Genetic Algorithm optimization achieve slightly higher peak accuracy, the proposed Decision Tree with Random Search delivers comparable performance with reduced model complexity. This finding suggests that effective hyperparameter optimization can significantly enhance lightweight models, narrowing the performance gap with more computationally intensive approaches.

Overall, the experimental outcomes confirm that Random Search is an effective strategy for improving Decision Tree classification in heart failure prediction. The optimized model not only increases overall accuracy but also substantially enhances recall for the positive class, which is critical in medical applications. These results establish a strong empirical foundation for further discussion regarding model interpretability, stability, and practical deployment in clinical environments.

4.   DISCUSSIONS

The experimental findings indicate that Random Search based hyperparameter optimization substantially enhances the predictive capability of the Decision Tree model for heart failure risk assessment. While the increase in overall accuracy demonstrates improved classification performance, the more critical outcome lies in the rise of recall for the positive class, which reflects a reduced number of undetected heart failure cases. In medical screening applications, this improvement is particularly significant, as minimizing false negative predictions directly supports earlier intervention and more reliable clinical decision making.

Analysis of the confusion matrices further clarifies the nature of this improvement. The optimized model increased the number of true positive predictions while reducing false negatives, whereas false positives and true negatives remained unchanged. This pattern suggests that Random Search primarily strengthened the model sensitivity to heart failure cases without degrading its ability to recognize non heart failure samples. Such behavior indicates a balanced adjustment rather than an aggressive shift toward positive prediction, preserving stability while enhancing clinical relevance.

The selected hyperparameters provide insight into why this behavior emerged. Employing entropy as the splitting criterion promotes more informative partitions, while restricting tree depth to three imposes a strong regularization effect that limits overfitting. Moreover, enforcing a minimum of ten samples per leaf discourages overly specific decision rules that may capture noise rather than meaningful clinical patterns. These constraints collectively indicate that the observed performance gains stem from improved generalization rather than increased model complexity.

Interpretation of the learned tree structure also supports this conclusion. The prominence of ST_Slope as the root node, followed by features such as ChestPainType, ExerciseAngina, and Oldpeak, highlights the importance of electrocardiographic and exercise related indicators in stratifying heart failure risk. The emergence of these clinically meaningful attributes suggests that the optimized model aligns with established medical knowledge, while its shallow structure maintains interpretability and reduces the risk of unstable decision boundaries.

When placed in the context of related work, the proposed Decision Tree with Random Search achieves competitive results compared with Random Forest models optimized using Genetic Algorithms. Although ensemble based approaches report slightly higher peak accuracies, they require greater computational resources and introduce additional architectural complexity. The present study demonstrates that a carefully tuned single tree model can approach the performance of such ensembles, offering a simpler and more transparent alternative that may be preferable in resource constrained or explainability sensitive healthcare environments.

his comparison highlights a broader trade off between predictive power and model simplicity. Random Forest benefits from ensemble diversity but sacrifices interpretability, making it difficult to trace individual decision paths. In contrast, the optimized Decision Tree enables direct inspection of decision rules and feature contributions, which supports clinician trust and facilitates integration into clinical workflows. Such transparency is increasingly recognized as a critical requirement for medical artificial intelligence systems.

The effectiveness of Random Search further reinforces its suitability for practical deployment. Unlike Grid Search, which exhaustively evaluates predefined parameter combinations, or Genetic Algorithms, which involve iterative evolutionary processes, Random Search efficiently samples the hyperparameter space with relatively low computational overhead. The results indicate that this strategy can uncover high performing configurations without extensive search budgets, making it attractive for real world medical applications with limited infrastructure.

espite these encouraging outcomes, several limitations should be acknowledged. The evaluation relied on a single dataset, which may restrict generalizability across populations with different demographic or clinical characteristics. In addition, external validation was not conducted, and only Random Search was explored as the optimization technique. Future research should therefore investigate cross dataset validation, compare multiple optimization strategies such as Bayesian or evolutionary methods, and explore hybrid or ensemble extensions to further improve robustness and clinical applicability.

Overall, the proposed Random Search optimized Decision Tree represents a practical compromise between accuracy, interpretability, and computational efficiency. By improving sensitivity to heart failure cases while maintaining a simple and transparent structure, the approach offers a promising foundation for developing accessible clinical decision support systems and motivates continued exploration of lightweight yet effective predictive models.

5.   CONCLUSION

This study explored the use of Random Search for optimizing Decision Tree hyperparameters in heart failure risk prediction. Rather than focusing solely on numerical improvements, the findings highlight how appropriate hyperparameter selection can enhance model generalization while preserving

interpretability. The optimized Decision Tree demonstrated improved sensitivity to heart failure cases, indicating that lightweight models, when properly tuned, can provide clinically meaningful performance without relying on complex ensemble architectures. Beyond predictive accuracy, this work emphasizes the practical value of model transparency. The resulting shallow tree structure enables direct inspection of decision pathways and feature contributions, which is essential for fostering clinical trust and supporting explainable decision making. Compared with ensemble based approaches reported in previous studies, the proposed method offers a favorable balance between performance, computational efficiency, and interpretability, making it suitable for deployment in resource constrained healthcare environments. Nevertheless, several limitations must be acknowledged. The experiments were conducted using a single publicly available dataset, which may introduce dataset dependency and limit generalizability to broader patient populations. In addition, external validation on independent cohorts was not performed, and although structural regularization was applied through constrained tree depth and minimum leaf size, the risk of overfitting cannot be entirely excluded. These factors should be considered when interpreting the reported outcomes. Future work will focus on validating the proposed approach across multiple datasets and clinical settings to assess robustness and transferability. Further extensions may include comparative studies with advanced optimization techniques such as Bayesian or evolutionary methods, as well as hybrid or ensemble frameworks. Incorporating external validation and domain specific feature analysis is also expected to strengthen clinical applicability. In summary, this study demonstrates that Random Search optimized Decision Trees provide an effective, interpretable, and computationally efficient framework for heart failure prediction. By balancing predictive performance with transparency and practical feasibility, the proposed approach offers a promising foundation for developing scalable clinical decision support systems and motivates continued investigation into lightweight yet reliable machine learning models.

CONFLICT OF INTEREST

The authors declares that there is no conflict of interest between the authors or with research object in this paper.

ACKNOWLEDGEMENT

REFERENCES

[1]   N. L. K. A. Arsani, N. P. D. S. Wahyuni, N. N. M. Agustini, and M. Budiawan, "Deteksi dini dan pencegahan penyakit kardiovaskular," Proceeding Senadimas Undiksha, vol. 1, no. 1, pp. 663–668, 2022.

[2]   A. Čartolovni, A. Tomičić, and E. L. Mosler, "Ethical, legal, and social considerations of AI-based medical decision-support tools: A scoping review," International Journal of Medical Informatics, vol. 161, Art. no. 104738, 2022.

[3]   Q. Xu, W. Xie, B. Liao, C. Hu, L. Qin, Z. Yang, et al., "Interpretability of clinical decision support systems based on artificial intelligence from technological and medical perspective: A systematic review," Journal of Healthcare Engineering, vol. 2023, no. 1, Art. no. 9919269, 2023.

[4]   Q. An, S. Rahman, J. Zhou, and J. J. Kang, "A comprehensive review on machine learning in healthcare industry: classification, restrictions, opportunities and challenges," Sensors, vol. 23, no. 9, Art. no. 4178, 2023.

[5]   Badawy, M., Ramadan, N., & Hefny, H. A. (2023). Healthcare predictive analytics using machine learning and deep learning techniques: a survey. Journal of Electrical Systems and Information Technology, 10(1), 40.

[6]   M. S. Gangadhar, K. V. S. Sai, S. H. S. Kumar, K. A. Kumar, M. Kavitha, and S. S. Aravinth,

"Machine learning and deep learning techniques on accurate risk prediction of coronary heart disease," in Proc. 7th Int. Conf. Computing Methodologies and Communication (ICCMC), Erode, India, 2023, pp. 227–232, doi: 10.1109/ICCMC56507.2023.10083756.

[7]   M. T. García-Ordás, M. Bayón-Gutiérrez, C. Benavides, et al., "Heart disease risk prediction using deep learning techniques with feature augmentation," Multimedia Tools and Applications, vol. 82, pp. 31759–31773, 2023, doi: 10.1007/s11042-023-14817-z.

[8]   M. Ozcan and S. Peker, "A classification and regression tree algorithm for heart disease modeling and prediction," Healthcare Analytics, vol. 3, p. 100130, 2023, doi: 10.1016/j.health.2022.100130.

[9]   S. Dalal, P. Goel, E. M. Onyema, A. Alharbi, A. Mahmoud, M. A. Algarni, and H. Awal, "Application of machine learning for cardiovascular disease risk prediction," Computational Intelligence and Neuroscience, Art. no. 9418666, 12 pp., 2023, doi: 10.1155/2023/9418666.

[10]   R. Azzaz, M. Jahazi, S. E. Kahou, and E. Moosavi-Khoonsari, "Prediction of final phosphorus content of steel in a scrap-based electric arc furnace using artificial neural networks," Metals, vol. 15, no. 1, Art. no. 62, 2025, doi: 10.3390/met15010062.

[11]   RL, M., & Mishra, A. K. (2022). Measuring financial performance of Indian manufacturing firms: application of decision tree algorithms. Measuring Business Excellence, 26(3), 288-307.

[12]   Soni, T., Gupta, D., & Uppal, M. (2024, November). Optimizing Heart Disease Prediction with Random Forest: Insights from the Kaggle Dataset. In 2024 4th International Conference on Advancement in Electronics & Communication Engineering (AECE) (pp. 741-744). IEEE.

[13]   Noh, Y. D., & Cho, K. C. (2022). Heart Disease Prediction Using Decision Tree With Kaggle Dataset. Journal of The Korea Society of Computer and Information, 27(5), 21-28.

[14]   T. Wongvorachan, S. He, and O. Bulut, "A comparison of undersampling, oversampling, and SMOTE methods for dealing with imbalanced classification in educational data mining," Information, vol. 14, no. 1, Art. no. 54, 2023.

[15]   M. Salmi, D. Atif, D. Oliva, A. Abraham, and S. Ventura, "Handling imbalanced medical datasets: Review of a decade of research," Artificial Intelligence Review, vol. 57, no. 10, Art. no. 273, 2024.

[16]   D. Rajput, W. J. Wang, and C. C. Chen, "Evaluation of a decided sample size in machine learning applications," BMC Bioinformatics, vol. 24, no. 1, Art. no. 48, 2023.

[17]   E. Widad, E. Saida, and Y. Gahi, "Quality anomaly detection using predictive techniques: An extensive big data quality framework for reliable data analysis," IEEE Access, vol. 11, pp. 103306–103318, 2023.

[18]   G. R. Hemanth and S. C. Raja, "Proposing suitable data imputation methods by adopting a stage-wise approach for various classes of smart meters missing data—Practical approach," Expert Systems with Applications, vol. 187, Art. no. 115911, 2022.

[19]   V. R. Joseph and A. Vakayil, "SPlit: An optimal method for data splitting," Technometrics, vol. 64, no. 2, pp. 166–176, 2022.

[20]   J. S. Kushwah, A. Kumar, S. Patel, R. Soni, A. Gawande, and S. Gupta, "Comparative study of regressor and classifier with decision tree using modern tools," Materials Today: Proceedings, vol. 56, pp. 3571–3576, 2022.

[21]   Tariq, A., Yan, J., Gagnon, A. S., Riaz Khan, M., & Mumtaz, F. (2023). Mapping of cropland, cropping patterns and crop types by combining optical remote sensing images with decision tree classifier and random forest. Geo-Spatial Information Science, 26(3), 302-320.

[22]   Colledani, D., Anselmi, P., & Robusto, E. (2023). Machine learning-decision tree classifiers in psychiatric assessment: An application to the diagnosis of major depressive disorder. Psychiatry Research, 322, 115127.

[23]   Jo, N., Aghaei, S., Benson, J., Gomez, A., & Vayanos, P. (2023, August). Learning optimal fair decision trees: Trade-offs between interpretability, fairness, and accuracy. In Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society (pp. 181-192).

[24]   Custode, L. L., & Iacca, G. (2023). Evolutionary learning of interpretable decision trees. IEEE Access, 11, 6169-6184.

[25]   Liu, Y. (2022). bsnsing: A decision tree induction method based on recursive optimal boolean rule composition. INFORMS Journal on Computing, 34(6), 2908-2929.

[26]   Blockeel, H., Devos, L., Frénay, B., Nanfack, G., & Nijssen, S. (2023). Decision trees: from efficient prediction to responsible AI. Frontiers in artificial intelligence, 6, 1124553.

[27]   X. Jiang, J. Zhang, X. Shi, and J. Cheng, "Learning the policy for mixed electric platoon control of automated and human-driven vehicles at signalized intersection: A random search approach," IEEE Transactions on Intelligent Transportation Systems, vol. 24, no. 5, pp. 5131–5143, 2023.

[28]   N. Subaşı, "Comprehensive analysis of grid and randomized search on dataset performance," European Journal of Engineering and Applied Sciences, vol. 7, no. 2, pp. 77–83, 2024.

[29]   K. Deligkaris, "Particle swarm optimization and random search for convolutional neural architecture search," IEEE Access, vol. 12, pp. 91229–91241, 2024.

[30]   S. Suyahman and A. Hapsari, "VGG-based feature extraction for classifying traditional batik motifs using machine learning models," Preservation, Digital Technology & Culture, 2025.

[31]   M. Azhari, Z. Situmorang, and R. Rosnelly, "Perbandingan akurasi, recall, dan presisi klasifikasi pada algoritma C4.5, random forest, SVM, dan naive Bayes," Jurnal Media Informatika Budidarma, vol. 5, no. 2, p. 640, 2021.

[32]   P. R. Togatoropa, M. Sianturia, D. Simamoraa, and D. Silaena, "Optimizing random forest using genetic algorithm for heart disease classification," Machine Learning, vol. 2, no. 3, Art. no. 4, 2022.

[33]   Y. Rimal, N. Sharma, and A. Alsadoon, "The accuracy of machine learning models relies on hyperparameter tuning: Student result classification using random forest, randomized search, grid search, Bayesian, genetic, and Optuna algorithms," Multimedia Tools and Applications, vol. 83, no. 30, pp. 74349–74364, 2024.