Brilliance
Research of Artificial Intelligence

E-ISSN : 2807-9035
Volume 5, **Number** 2, November 2025
https://doi.org/10.47709/brilliance.v5i2.7505

Comparison of XGBoost and LightGBM Algorithms in Predicting Heart Disease

Fionna Caroline[1*], Nur Rachmat[2]
[1,2]Universitas Multi Data Palembang, Indonesia,
[1]fionnacaroline@mhs.mdp.ac.id, [2]nur.rachmat@mdp.ac.id

open access

***Corresponding Author**

Article History:
Submitted: 12-12-2025
Accepted: 18-12-2025
Published: 28-12-2025

Keywords:
Heart Disease Prediction; Machine Learning; XGBoost; LightGBM; Stratified K-Fold Cross Validation; SMOTE

Brilliance: Research of Artificial Intelligence is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).

ABSTRACT

Heart disease remains a leading cause of mortality worldwide, underscoring the need for early and accurate diagnosis to reduce complications and improve patient outcomes. Recent advances in machine learning have enabled the development of predictive models that assist healthcare professionals in disease detection using patient medical records. This study aims to develop and compare the performance of Extreme Gradient Boosting (XGBoost) and Light Gradient Boosting Machine (LightGBM) for heart disease prediction.

The dataset used in this research was obtained from the UCI Machine Learning Repository and consists of 303 patient records with binary class labels indicating the presence or absence of heart disease. Data preprocessing involved feature standardization using StandardScaler and handling class imbalance through the Synthetic Minority Over-sampling Technique (SMOTE). Model evaluation was conducted using Stratified K-Fold Cross Validation with K values of 3, 5, and 7 to ensure robust and unbiased performance assessment. Hyperparameter optimization was carried out using RandomizedSearchCV to efficiently identify optimal model configurations.

Experimental results indicate that both XGBoost and LightGBM achieved strong classification performance, with accuracy exceeding 80% and AUC values above 0.89. LightGBM demonstrated slightly superior performance in terms of average accuracy, F1-score, and stability across folds, while XGBoost achieved higher precision, reflecting better control of false positives. Overall, both algorithms are effective for heart disease prediction, supporting the potential of machine learning in early disease detection and clinical decision-support systems.

INTRODUCTION

The heart is one of the most vital organs in the human body. It works to circulate blood throughout the body via the blood vessels. Therefore, despite its crucial role in delivering oxygen to all tissues and removing metabolic waste, the heart is one of the organs most vulnerable to disease.nerable to disease (Prabowo & Kurniadi, 2023). Heart disease is influenced by multiple factors, including genetic predisposition and unhealthy lifestyle habits such as physical inactivity, smoking, alcohol consumption, high intake of fat and salt, and psychological stress (Naomi et al., 2021). This disease may develop as a result of daily lifestyle behaviors and habits, beginning with atherosclerosis, a process characterized by the accumulation and proliferation of elements within the arterial walls, accompanied by plasma lipoprotein deposition. This process leads to the formation of foam cells and subsequent disruption of blood flow to the heart (Yogianto et al., 2024)

According to data released by the World Health Organization (WHO), in 2021, heart disease caused approximately 17.8 million deaths annually, accounting for one in every three deaths worldwide. The data also reported 21.2 million new cases of heart disease. Furthermore, the prevalence of heart disease was higher among males than females.(Kemenkes, 2024)

The symptoms of heart disease vary widely, ranging from chest pain to the most severe condition, sudden cardiac arrest, which remains one of the leading causes of death globally (Nafisah et al., 2024). In principle, heart disease can be identified at an early stage; however, many individuals fail to recognize the condition in a timely manner due to a lack of awareness. If left untreated, this condition can pose serious long-term health risks. The likelihood of full recovery from heart disease is relatively low, and it is particularly difficult for patients diagnosed with chronic heart conditions to regain normal health. Therefore, early diagnosis and treatment remain the most effective options (Ahmed, 2023).

The application of technological approaches is essential to assist healthcare professionals in achieving more effective prevention and treatment. High-risk patients have a greater chance of avoiding severe complications when diagnosis is performed accurately and promptly (Arjun Vahlevy et al., 2023). Technology-based diagnostic systems are capable of providing accurate and reliable detection within a shorter time frame by leveraging machine learning techniques to generate, process, and analyze large volumes of data. Ultimately, these systems enable healthcare

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Brilliance
Research of Artificial Intelligence
E-ISSN : 2807-9035
Volume 5, Number 2, November 2025
https://doi.org/10.47709/brilliance.v5i2.7505

professionals to make better clinical decisions by identifying patterns and symptoms associated with heart disease (Handayani et al., 2021)

Therefore, this study aims to develop and compare the XGBoost and LightGBM algorithms for predicting heart disease using hyperparameter tuning with the RandomizedSearchCV method and Stratified K-Fold Cross Validation, along with oversampling using SMOTE. XGBoost is a variant of decision tree based algorithms that shares similarities with gradient boosting. It differs primarily in its objective function design, which is optimized to minimize the loss function, and it focuses on classification-based decision trees (Aulia et al., 2025) The architecture of XGBoost, illustrating the workflow from data processing to output generation, is presented in Figure 1.

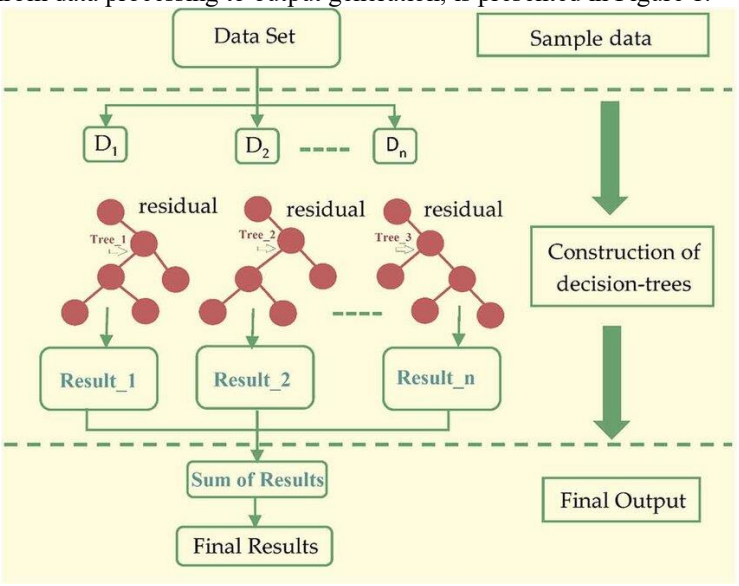

Figure 1. *The architecture of XGBoost*

LightGBM is a Gradient Boosting Decision Tree (GBDT) algorithm optimized for computational speed and memory efficiency. It employs Gradient-based One-Side Sampling (GOSS) to reduce the number of data instances and Exclusive Feature Bundling (EFB) to simplify the feature space. According to the formulation presented by (Wang & Wang, 2020), LightGBM uses the same general objective function and second-order Taylor approximation as XGBoost. However, the two algorithms differ in their tree growth strategies. XGBoost adopts a level-wise tree growth approach, as illustrated in Figure 2, in which all nodes at the same level are split in parallel. In contrast, LightGBM employs a leaf-wise tree growth strategy, as shown in Figure 3, where the leaf with the largest information gain is selecte9d for splitting .

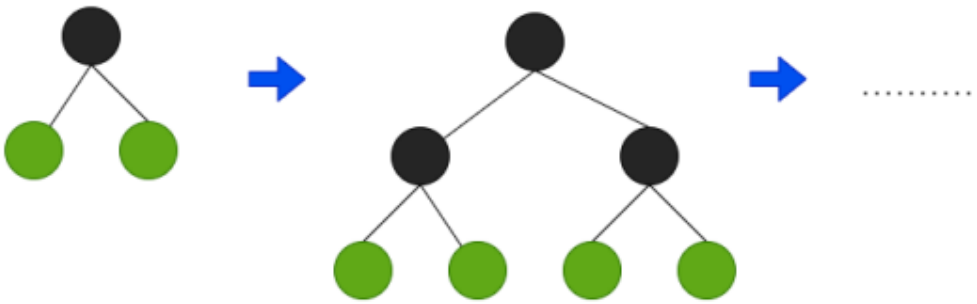

Figure 2. Level-Wise Tree Growth of XGBoost

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Brilliance
Research of Artificial Intelligence

E-ISSN : 2807-9035
Volume 5, Number 2, November 2025
https://doi.org/10.47709/brilliance.v5i2.7505

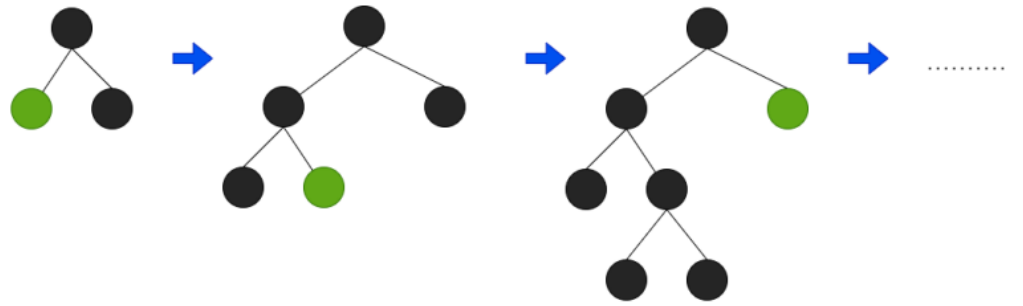

Figure 3. Leaf-Wise Tree Growth of LightGBM

LITERATURE REVIEW

Several previous studies have investigated heart disease detection using various machine learning algorithms. A study conducted by (Pradana et al., 2022) employed an Artificial Neural Network (ANN) for heart disease classification, achieving an accuracy of 73.77%, a precision of 80.43%, a recall of 84.09%, and an F1-score of 82.22%. However, this study did not apply techniques such as Stratified K-Fold Cross Validation, nor did it address data imbalance issues. As a result, the reported performance may still be prone to overfitting or bias toward the majority class.

Subsequently, (Hidayat et al., 2023) utilized the Random Forest Classifier to perform heart disease classification and reported an accuracy of 94%. Nevertheless, the study did not describe the hyperparameter tuning process, making it uncertain whether the model configuration was optimal. In addition, the evaluation strategy did not include cross-validation, and no approach was implemented to handle class imbalance.

Study by (Sausan et al., 2024) compared the Decision Tree Classifier and XGBoost Classifier for heart disease prediction using a dataset from the UCI Machine Learning Repository. The results showed that XGBoost achieved an accuracy of 93%, while Decision Tree achieved 90%. The study included data preprocessing steps such as oversampling using SMOTE to address class imbalance. However, it did not implement cross-validation techniques such as Stratified K-Fold Cross Validation to ensure model generalization, nor did it perform hyperparameter tuning using methods such as RandomizedSearchCV.

Furthermore, (Nurmasani & Pristyanto, 2021) proposed an ensemble stacking approach for heart disease classification under imbalanced data conditions. The study utilized the heart failure clinical records dataset from the UCI Machine Learning Repository, with C4.5 and SVM as base learners and logistic regression as the meta-classifier. The evaluation results indicated that the stacking method outperformed several single classifiers, including K-NN, SVM, Decision Tree, and Neural Network, achieving an accuracy of 81%, a G-Mean of 75%, and an AUC of 87%. However, the study did not explain the hyperparameter tuning process, which may affect optimal model performance. Moreover, despite addressing data imbalance through an ensemble approach, the study did not employ Stratified K-Fold Cross Validation (SKCV) or explicit oversampling techniques such as SMOTE to strengthen minority class distribution in a controlled manner.

Another study by (Ramadhanti et al., 2024) applied the XGBoost algorithm to predict hemoglobin levels. The experimental results showed that the average difference between non-invasive and invasive hemoglobin predictions was 0.76, with an average accuracy of 94.09%. The oxygen saturation sensor exhibited an error rate of 1%, while the heart rate sensor and temperature sensor showed error rates of 5.91% and 2.24%, respectively. Additionally, (Sankar et al., 2022) compared XGBoost with several other machine learning algorithms for thyroid disease prediction and reported that XGBoost achieved the highest accuracy, reaching 98.59%.

In another study, (Kurniadi & Larasati, 2022) compared LightGBM with SVM and Random Forest under two scenarios using feature selection and without feature selection for stroke detection. The results showed that all three methods achieved identical accuracy values of 98% in both scenarios. Subsequently, (R et al., 2025) compared LightGBM and ANN for red wine quality prediction, where LightGBM demonstrated superior performance with an accuracy of 80%. Moreover, (Hakim et al., 2025) compared XGBoost and LightGBM for optimizing diabetes classification models, reporting accuracies of 90.78% for XGBoost and 91.29% for LightGBM.

METHOD

The data used in this study consist of patient medical record data related to heart disease. The dataset was obtained from a publicly available dataset in the UCI Machine Learning Repository known as the “Heart Disease” dataset. This dataset contains 303 patient medical records, indicating whether each patient has a history of heart disease or not. In this dataset, the 14 features used are:

1. age – Age of the patient
2. sex – Gender (1 = male, 0 = female)

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Brilliance
Research of Artificial Intelligence
E-ISSN : 2807-9035
Volume 5, Number 2, November 2025
https://doi.org/10.47709/brilliance.v5i2.7505

3. cp – Chest pain type
4. trestbps – Resting blood pressure
5. chol – Serum cholesterol in mg/dl
6. fbs – Fasting blood sugar > 120 mg/dl
7. restecg – Resting electrocardiographic results
8. thalach – Maximum heart rate achieved
9. exang – Exercise induced angina
10. oldpeak – ST depression (rounded, positive values)
11. slope – Slope of the peak exercise ST segment
12. ca – Number of major vessels (0-3) colored by fluoroscopy
13. thal – Thalassemia test result
14. num target –Diagnosis of heart disease

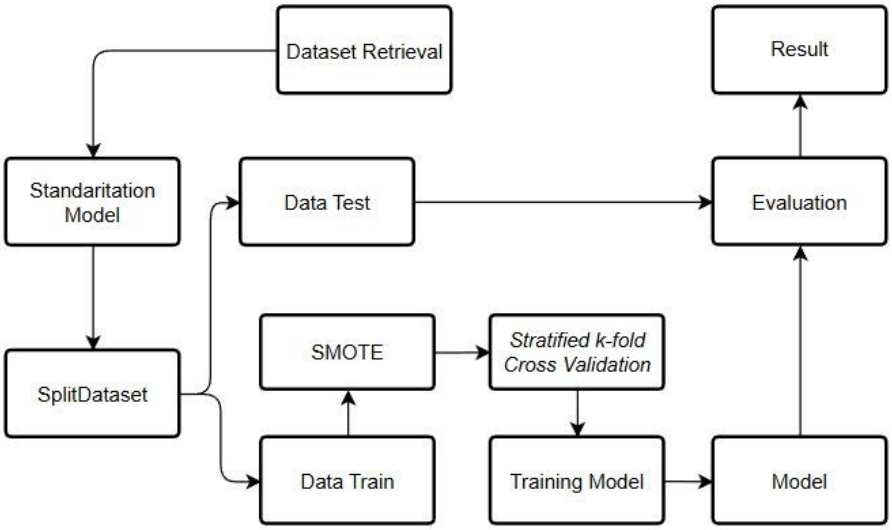

Figure 4. Computational model

The overall computational model is illustrated in Figure 4. The process begins with the input dataset consisting of patient medical records. After the imputation process, all numerical features are standardized using the StandardScaler method. Feature standardization using StandardScaler is an essential preprocessing step that transforms numerical features to have zero mean and unit variance. This technique improves the stability and performance of machine learning models, particularly scale-sensitive algorithms. By equalizing feature scales, StandardScaler enables more efficient model optimization and enhances overall predictive performance (Mathivanan et al., 2025). This standardization also aims to accelerate the training process and After standardization, all numerical and categorical variables are identified to ensure proper processing. Relevant features are retained based on medical domain considerations and their potential contribution to prediction performance. The dataset is subsequently split into training and testing sets with an 80:20 ratio. Before model training, the target class distribution is analyzed. Since heart disease diagnosis data often exhibits class imbalance between positive and negative cases, the Synthetic Minority Over-sampling Technique (SMOTE) technique is applied to balance the dataset. This method generates synthetic samples for the minority class, preventing model bias toward the majority class and enabling better recognition of minority patterns (Effendy et al., 2024). Class imbalance is addressed using SMOTE, which has been proven effective in improving the performance of classification models on imbalanced medical data.

The training data is evaluated using Stratified K-Fold Cross Validation with K values of 3, 5, and 7 to ensure stable and unbiased model assessment across different data subsets. In each iteration, the first fold serves as the testing data, while the remaining folds serve as the training data. This process is repeated by rotating the testing fold, as illustrated in Figure 5.

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Brilliance
Research of Artificial Intelligence
E-ISSN : 2807-9035
Volume 5, Number 2, November 2025
https://doi.org/10.47709/brilliance.v5i2.7505

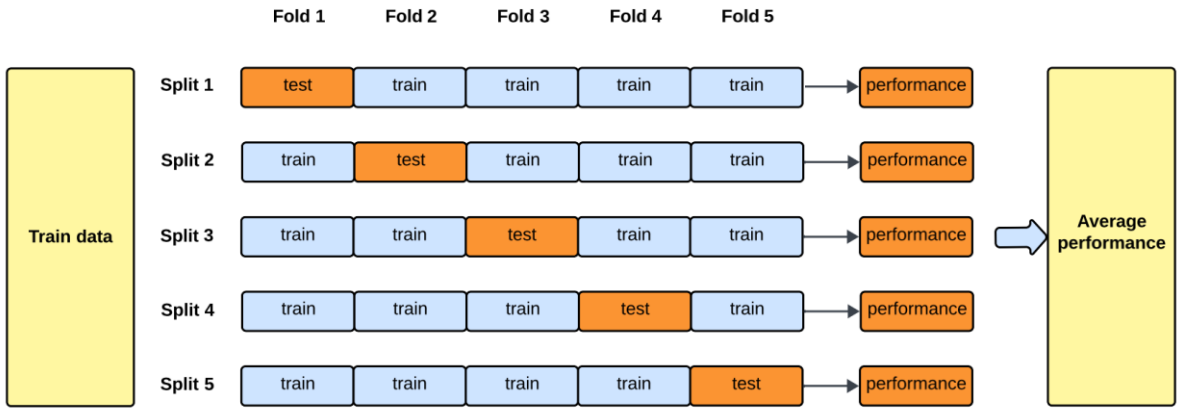

Figure 5. Cross-Validation Fold Partitioning Scheme

The next stage involves training the XGBoost and LightGBM models using the training data. The XGBoost model is configured with 300 decision trees (n_estimators = 300) and a learning rate of 0.05 to ensure that the learning process takes place gradually and steadily, thereby improving the model's generalization ability. The subsample and colsample_bytree parameters are each set to 0.9 to introduce data and feature variation in each tree, which aims to reduce overfitting and improve model robustness. Tree complexity is limited by max_depth = 4, so that the model remains capable of capturing non-linear relationships without becoming too complex. In addition, L2 regularization (reg_lambda = 1.0) is applied to control model weights and improve training stability. The training process is made reproducible by setting random_state = 42, while the logarithmic loss (logloss) evaluation metric is used to optimize probability estimation in binary classification. The histogram tree construction method is chosen to improve computational efficiency and memory usage during the training process.

Hyperparameter tuning is performed using RandomizedSearchCV, which efficiently explores the hyperparameter space by randomly sampling a fixed number of parameter combinations. RandomizedSearchCV is used to optimize hyperparameters by randomly exploring the parameter space and evaluating each combination using cross-validation techniques. This approach has proven to be more computationally efficient than exhaustive search methods such as Grid Search, especially for complex ensemble-based models such as XGBoost and LightGBM, as it is able to find the optimal hyperparameter combination without having to test all possible parameters available (Aribisala et al., 2022). This approach is computationally more efficient than exhaustive grid search, particularly for complex models such as gradient boosting, while still achieving near-optimal performance. Randomized hyperparameter optimization has been shown to be effective in reducing training time and improving generalization performance in machine learning models (Firdaus et al., 2025).

The final models are then evaluated using the testing dataset. The evaluation includes the confusion matrix and classification metrics such as accuracy, precision, recall, and F1-score to assess overall model performance. Additionally, computation time, ROC curves, and AUC values are recorded to compare the efficiency of both algorithms. Model evaluation is conducted using confusion matrix–based metrics such as accuracy, precision, recall, and F1-score, along with ROC curves and AUC values to assess the discriminative ability of the models, while computational time is recorded to compare algorithm efficiency.

RESULT

XGBoost Experimental Results

After conducting experiments using the XGBoost method with 3-fold, 5-fold, and 7-fold cross-validation, the results show that the 5-fold CV configuration performs the best, providing the most balanced combination of accuracy, F1-score, and AUC. However, it requires longer computation time (2.28 seconds) compared to the 3-fold and 7-fold CV. Meanwhile, the 3-fold CV configuration is superior in terms of efficiency, with the fastest computation time (0.17 seconds) and relatively small standard deviation, indicating consistent results. However, its performance metrics remain slightly lower than those of the 5-fold CV. The 7-fold CV configuration shows weaker and less stable performance, as indicated by higher standard deviation values, particularly on recall (8.73%). The detailed experimental results for XGBoost are presented in Table 1.

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Brilliance
Research of Artificial Intelligence

E-ISSN : 2807-9035
Volume 5, Number 2, November 2025
https://doi.org/10.47709/brilliance.v5i2.7505

Table 1. XGBoost Experimental Results

	3-fold CV	5-fold CV	7-fold CV
Accuracy	0.8359 ± 0.0145	0.8398 ± 0.0218	0.8208 ± 0.0441
Precision	0.8521 ± 0.0371	0.8546 ± 0.0518	0.8322 ± 0.0536
Recall	0.8166 ± 0.0205	0.8245 ± 0.0388	0.8083 ± 0.0873
F-1 Score	0.8330 ± 0.0077	0.8374 ± 0.0205	0.8167 ± 0.0527
ROC AUC	0.8897 ± 0.0137	0.9013 ± 0.0264	0.8904 ± 0.0418
Compute Time	0.17 s	2.28 s	0.41 s

LightGBM Experimental Results

After conducting experiments using the LightGBM method with 3-fold, 5-fold, and 7-fold cross-validation, the best results were obtained using the 5-fold CV scheme, which achieved an average accuracy of 0.8285 ± 0.0325, an F1-score of 0.8280 ± 0.0305, and the highest AUC of 0.9037 ± 0.0306. The relatively low standard deviation indicates consistent performance across folds. Meanwhile, the 3-fold CV configuration produced a lower accuracy (0.8206 ± 0.0063), but with stable results due to its small variability. The 7-fold CV configuration yielded accuracy similar to the 3-fold CV (0.8207 ± 0.0488), but with higher variability especially in recall (0.7928 ± 0.0984) indicating lower model stability. The detailed LightGBM experimental results are presented in Table 2.

Table 2. LightGBM Experimental Results

	3-fold CV	5-fold CV	7-fold CV
Accuracy	0.8206 ± 0.0063	0.8285 ± 0.0325	0.8207 ± 0.0488
Precision	0.8312 ± 0.0302	0.8344 ± 0.0520	0.8418 ± 0.0504
Recall	0.8087 ± 0.0492	0.8245 ± 0.0388	0.7928 ± 0.0984
F-1 Score	0.8179 ± 0.0127	0.8280 ± 0.0305	0.8129 ± 0.0578
ROC AUC	0.8889 ± 0.0132	0.9037 ± 0.0306	0.8823 ± 0.0416
Compute Time	0.21 s	2.46 s	0.54 s

DISCUSSION

Based on the experimental results using 3-fold, 5-fold, and 7-fold cross-validation on both XGBoost and LightGBM, the 5-fold CV scheme consistently emerges as the most optimal choice. For XGBoost, the 5-fold CV configuration produces the highest accuracy (0.8398 ± 0.0218), F1-score (0.8374 ± 0.0205), and AUC (0.9013 ± 0.0264). Although it requires more computation time (2.28 seconds), these results indicate a strong balance between accuracy, stability, and generalization capability. Similarly, LightGBM also achieves its best performance with 5-fold CV, yielding an accuracy of 0.8285 ± 0.0325, an F1-score of 0.8280 ± 0.0305, and the highest AUC (0.9037 ± 0.0306). The relatively small standard deviations demonstrate consistent performance across folds. Table 3 presents the comparison of the best experimental results for both LightGBM and XGBoost using the 5-fold CV scheme.

Table 3. Comparison of the Best Experimental Results of LightGBM and XGBoost

	5-fold CV	
	LightGBM	**XGBoost**
Accuracy	0.8285 ± 0.0325	0.8398 ± 0.0218
Precision	0.8344 ± 0.0520	0.8546 ± 0.0518
Recall	0.8245 ± 0.0388	0.8245 ± 0.0388
F-1 Score	0.8280 ± 0.0305	0.8374 ± 0.0205
ROC AUC	0.9037 ± 0.0306	0.9013 ± 0.0264
Compute Time	2.46 s	2.28 s

CONCLUSION

This study successfully developed a heart disease prediction model based on machine learning using the XGBoost and LightGBM algorithms, and conducted a comparative evaluation of their performance through cross-validation and testing on a separate test set. The experimental results show that both models achieved high classification performance, with accuracy levels exceeding 80% and AUC values above 0.89, indicating strong discriminative capability in distinguishing patients with and without heart disease risk.

Overall, LightGBM demonstrated slightly superior performance compared to XGBoost in terms of average accuracy, F1-score, and result stability across folds. This improvement can be attributed to LightGBM's histogram-based learning mechanism and leaf-wise tree growth strategy, which enable more efficient learning and better adaptation to data patterns. As a result, LightGBM produced faster and more responsive models when handling the variation introduced by SMOTE. Therefore, it can be concluded that 5-fold cross-validation is the most optimal validation scheme for the LightGBM model in this dataset, as it provides a balanced combination of accuracy, high

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

E-ISSN : 2807-9035
Volume 5, **Number** 2, November 2025
https://doi.org/10.47709/brilliance.v5i2.7505

AUC, and performance stability.

Meanwhile, XGBoost exhibited an advantage in achieving higher precision, indicating that the model is more conservative when generating positive predictions and tends to produce fewer false positives. This characteristic makes XGBoost more suitable for contexts where minimizing false positives is critical, such as early diagnosis or pre-screening stages. Accordingly, 5-fold cross-validation is also identified as the most optimal validation scheme for the XGBoost model in this study, as it offers the best balance between accuracy, stability, and generalization capability.

Based on the overall evaluation, LightGBM is considered more efficient and stable for implementation in medical decision-support systems due to its ability to detect more positive cases with an acceptable error rate. In conclusion, the findings of this study confirm that both XGBoost and LightGBM are effective algorithms for supporting predictive analytics in the healthcare domain, particularly for early detection of heart disease.

REFERENCES

Ahmed, R. W. (2023). Decision support system for early diagnosis of heart diseases. *VFAST Transaction on Software Engineering*, *11*, 124–130.

Aribisala, J. O., Abdulsalam, R. A., Dweba, Y., Madonsela, K., & Sabiu, S. (2022). Identification of secondary metabolites from Crescentia cujete as promising antibacterial therapeutics targeting type 2A topoisomerases through molecular dynamics simulation. *Computers in Biology and Medicine*, *145*. https://doi.org/10.1016/J.COMPBIOMED.2022.105432

Arjun Vahlevy, D., Levis Putra Zendrato, E., Fadillah, R., Jafar Sidiq, R., & Sahruddin. (2023). Tinjauan Literatur Sistematik pada Sistem Pakar untuk Diagnosa Penyakit Manusia. *Jurnal Artificial Inteligent Dan Sistem Penunjang Keputusan*, *1*(1), 1–8. https://garuda.kemdikbud.go.id/.

Aulia, M. K., Utaminingsih, E., & Prihatin, N. (2025). Model Prediksi Risiko Kesehatan Perkotaan Berbasis Lingkungan dengan XGBoost. *Computer Science (Co-Science)*, *5*, 95–102.

Effendy, F., Sianturi, G. A. S., Silaban, D. F., Aqil, M. F., Simangunsong, E., Sitorus, Y. A. L., & Arnita. (2024). Application of Random Forest for Heart Disease Classification with SMOTE Approach to Balance Data. *Journal of Informatics and Data Science (J-IDS)*, *3*(2). https://doi.org/10.24114/j-ids.xxxxx

Firdaus, E. A., Manurung, J., Saragih, H., & Prabukusumo, M. A. (2025). *Optimization of XGBoost hyperparameters using grid search and random search for credit card default prediction*. *14*, 269–280.

Hakim, L., Zyen, A. K., & Sarwido. (2025). Optimasi Model Klasifikasi Diabetes dengan Stacking pada Algoritma XGBoost dan LightGBM. *Jurnal JUPITER*, *17*, 821–833.

Handayani, F., Kusuma, K. S., Asbudi, H. L., Purnasiwi, R. G., Kusuma, R., Sunyoto, A., & Pardnya, W. M. (2021). Komparasi Support Vector Machine, Logistic Regression Dan Artificial Neural Network Dalam Prediksi Penyakit Jantung. *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, *7*(3), 329–334. https://doi.org/10.26418/jp.v7i3.48053

Hidayat, Sunyoto, A., & Fatta, H. Al. (2023). Klasifikasi Penyakit Jantung Menggunakan Random Forest Clasifier. *Jurnal SISKOM-KB (Sistem Komputer Dan Kecerdasan Buatan)*, *7*(1), 31–40. https://doi.org/10.47970/siskom-kb.v7i1.464

Kemenkes. (2024). *Cardio Update 2024*. https://lms.kemkes.go.id/courses/35bff824-437e-4557-b37a-94b128c43333

Kurniadi, F. I., & Larasati, P. D. (2022). Light Gradient Boosting Machine untuk Deteksi Penyakit Stroke. *Jurnal SISKOM-KB (Sistem Komputer Dan Kecerdasan Buatan)*, *6*(1), 67–72. https://doi.org/10.47970/siskom-kb.v6i1.328

Mathivanan, N. M. N., Xian, E. F. Z., Xi, D. F. Y., & Kiat, C. H. (2025). Impact Of Feature Standardization On Heart Disease Prediction: A Comparative Analysis Of Logistic Regression And Support Vector Machine Models. *Malaysian Journal of Computing*, *10*(2), 2159–2175. https://doi.org/10.24191/mjoc.v10i1.6835

Nafisah, S., Novianti Nuril Inayah, & Baharuddin Yusuf. (2024). Literatur Review: Penyebab Dan Perkembangan Penyakit Jantung Koroner. *Jurnal Forum Kesehatan : Media Publikasi Kesehatan Ilmiah*, *14*(1), 27–36. https://doi.org/10.52263/jfk.v14i1.254

Naomi, W. S., Picauly, I., & Toy, S. M. (2021). Faktor Risiko Kejadian Penyakit Jantung Koroner. *Media Kesehatan Masyarakat*, *3*(1), 99–107. https://doi.org/10.35508/mkm.v3i1.3622

Nurmasani, A., & Pristyanto, Y. (2021). Algoritme Stacking Untuk Klasifikasi Penyakit Jantung Pada Dataset Imbalanced Class. *Jurnal Pseudocode*, *8*(1), 21–26. https://doi.org/10.33369/pseudocode.8.1.21-26

Prabowo, A. S., & Kurniadi, F. I. (2023). Analisis Perbandingan Kinerja Algoritma Klasifikasi dalam Mendeteksi Penyakit Jantung. *Jurnal SISKOM-KB (Sistem Komputer Dan Kecerdasan Buatan)*, *7*(1), 56–61. https://doi.org/10.47970/siskom-kb.v7i1.468

Pradana, D. G., Alghifari, M. L., Juna, M. F., & Palaguna, S. D. (2022). Klasifikasi Penyakit Jantung Menggunakan Metode Artificial Neural Network. *Indonesian Journal of Data and Science*, *3*(2), 55–60. https://doi.org/10.56705/ijodas.v3i2.35

R, A. R., W, A. W., A, E. P. A., R, S. F., & Rizki, A. M. (2025). Perbandingan Algoritma LightGBM dan ANN untuk

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

E-ISSN : 2807-9035
Volume 5, **Number** 2, November 2025
https://doi.org/10.47709/brilliance.v5i2.7505

Menentukan Kualitas Anggur Merah. *Jurnal Mahasiswa Teknik Informatika*, *9*(1), 1572–1579.

Ramadhanti, A., Putri, A., & Wagyana, A. (2024). Pengembangan Alat Deteksi Gejala Anemia Non-Invasive Berbasis Narrowband Internet of Things Dengan Pendekatan XGBoost. *Jurnal Informatika Dan Teknik Elektro Terapan*, *12*(3), 4483–4492.

Sankar, S., Potti, A., Naga Chandrika, G., & Ramasubbareddy, S. (2022). Thyroid Disease Prediction Using XGBoost Algorithms. *Journal of Mobile Multimedia*, *18*(3), 917–934. https://doi.org/10.13052/jmm1550-4646.18322

Sausan, Pratiwi, D. M., & Mufidah, L. (2024). Perbandingan Metode Decision Tree Classifier dan XGBoost Classifier Dalam Memprediksi Penyakit Jantung. *Conference on Electrical Engineering, Informatics Technology and Creative Media 2024*, *4*, 991–1000.

Wang, Y., & Wang, T. (2020). Application of improved LightGBM model in blood glucose prediction. *Applied Sciences (Switzerland)*, *10*(9). https://doi.org/10.3390/app10093227

Yogianto, A., Homaidi, A., & Fatah, Z. (2024). Implementasi Metode K-Nearest Neighbors (KNN) untuk Klasifikasi Penyakit Jantung. *G-Tech: Jurnal Teknologi Terapan*, *8*(3), 1720–1728. https://doi.org/10.33379/gtech.v8i3.4495

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.