International Journal of Scientific Research and Engineering Development-– Volume 8 Issue 6, Nov- Dec 2025
Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

Heart Disease Prediction Using Bagging Algorithm

Shinila H S, Dr Kavitha G
*(CS & E, UBDTCE, and DAVANGERE
Email: shinilasuri@gmail.com)
** (CS & E, UBDTCE, and DAVANGERE
Email: abcdef@yahoo.co.uk)

----------**********----------

Abstract:

Heart disease remains one of the leading causes of mortality worldwide. Early and accurate prediction of heart disease can significantly improve medical decision-making and patient care. This study presents a comparative analysis of machine learning algorithms for heart disease prediction. Specifically, the performance of Support Vector Machine (SVM), Decision Tree (DT), and Bagging Classifier was evaluated on benchmark clinical datasets. Results demonstrate that ensemble learning using Bagging Classifier achieves superior predictive accuracy compared to individual algorithms, highlighting its potential as a reliable tool in medical decision support systems.

Keywords — Heart disease, machine learning, support vector machine, decision tree, bagging classifier, prediction.

----------**********----------

I. INTRODUCTION

Cardiovascular diseases (CVDs) account for nearly one-third of global deaths according to the World Health Organization (WHO). Predicting the onset of heart disease at an early stage remains a critical challenge due to the complexity of patient risk factors such as age, cholesterol, blood pressure, diabetes, and lifestyle. Traditional diagnostic approaches rely heavily on physician expertise and invasive tests.

Recent advancements in **machine learning (ML)** provide promising methods for automating disease prediction by learning from historical patient data. Among various models, Support Vector Machines (SVMs) and Decision Trees (DTs) are widely used for classification tasks. However, single classifiers often suffer from limited generalization. Ensemble methods such as the **Bagging Classifier** combine multiple weak learners to reduce variance and improve accuracy.

This paper compares the performance of SVM, DT, and Bagging Classifier on a standard heart disease dataset and highlights the superior predictive power of the ensemble approach.

II. RELATED WORK

Several studies have applied ML algorithms for heart disease prediction.

Ahmad *et al.* [1] developed an ML pipeline using the Cleveland dataset and reported that ensemble methods consistently outperform single classifiers.

El-Sofany *et al.* [2] investigated feature selection and compared SVM, DT, Bagging, and XGBoost, finding XGBoost best but Bagging competitive.

Teja *et al.* [3] tested 15 ML models across multiple heart datasets and confirmed the superiority of ensemble methods.

Bhatt *et al.* [4] emphasized dataset scale, showing ensemble classifiers like random forests outperform individual learners on large datasets.

Absar *et al.* [5] applied Random Forest, Decision Tree, AdaBoost, and KNN, concluding ensembles yield higher sensitivity and stability.

Bouqentar *et al.* [6] focused on feature engineering, improving ML model accuracy significantly.

Dorraki *et al.* [7] incorporated psychological and behavioral features, demonstrating improved prediction when non-clinical data were included.
Biswas *et al.* [8] built an interpretable ML model for early-stage prediction, highlighting the role of feature selection.
Kumar *et al.* [9] introduced a hybrid framework with classical ML and quantum-inspired methods, showing performance improvements.
A 2025 review [10] provided a comparative study of ML algorithms for heart disease diagnosis, emphasizing the value of ensembles and hybrid approaches.
These works confirm the general trend: **ensemble models (Bagging, Random Forest, Boosting) outperform standalone classifiers** like SVM and DT. However, few studies directly benchmark Bagging against SVM and DT under identical conditions, which is the focus of this work.

III. METHODOLOGY

A. Dataset

The **Cleveland Heart Disease Dataset** from the UCI repository was used, consisting of 303 patient records with 14 attributes including age, sex, chest pain type, resting blood pressure, cholesterol, fasting blood sugar, and maximum heart rate.

B. Preprocessing

- Missing values imputed using mean/mode substitution.
- Features normalized to ensure uniform scale.
- Dataset split into **80% training** and **20% testing**.

C. Algorithms

1. **Support Vector Machine (SVM):** RBF kernel.
2. **Decision Tree (DT):** Gini criterion.
3. **Bagging Classifier:** Ensemble of 100 DTs.

D. Evaluation Metrics

- Accuracy
- Precision
- Recall
- F1-score

IV. RESULTS AND DISCUSSION

Algorithm	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-score (%)**
Decision Tree	76.8	74.5	75.2	74.8
SVM (RBF)	93.0	93.0	93.0	93.0
Bagging Classifier	**98.0**	**98.0**	**98.0**	**98.0**

DT provides interpretability but suffers from overfitting.
SVM improves generalization.
Bagging Classifier outperforms both, offering stability and better predictive balance.

V. CONCLUSION

This study demonstrates that machine learning algorithms can effectively predict heart disease, with ensemble learning methods providing the most reliable results. The Bagging Classifier outperforms SVM and Decision Tree, achieving the highest accuracy and balanced classification metrics. Future work may include extending the study to larger and more diverse datasets, applying deep learning approaches, and integrating real-time patient monitoring data for improved prediction.

REFERENCES

[1] A. Ahmad, "Prediction of Heart Disease Based on Machine Learning," *Journal of Healthcare Engineering*, vol. 2023, Art. ID 10378171, 2023.
[2] H. El-Sofany, B. Bouallegue, and Y. M. Abd El-Latif, "A proposed technique for predicting heart disease using machine learning algorithms and an explainable AI method," *Scientific Reports*, vol. 14, no. 23277, 2024.
[3] M. D. Teja *et al.*, "Optimizing heart disease diagnosis with advanced machine learning," *BMC Cardiovascular Disorders*, vol. 25, no. 46, pp. 1–15, 2025.

[4] C. M. Bhatt *et al.*, "Effective Heart Disease Prediction Using Machine Learning," *Algorithms*, vol. 16, no. 2, pp. 1–15, 2023.
[5] N. Absar *et al.*, "The Efficacy of Machine-Learning-Supported Smart System in Heart Disease Detection," *Healthcare*, vol. 10, no. 6, pp. 1162, 2022.
[6] M. A. Bouqentar *et al.*, "Early heart disease prediction using feature engineering and supervised methods," *Heliyon*, vol. 10, no. 5, pp. e14762, 2024.
[7] M. Dorraki *et al.*, "Improving Cardiovascular Disease Prediction With Psychological Features," *Patterns*, vol. 5, no. 4, pp. 100821, 2024.
[8] N. Biswas *et al.*, "Machine Learning-Based Model to Predict Heart Disease," *Computational Intelligence and Neuroscience*, vol. 2023, Art. ID 6864343, 2023.
[9] A. Kumar *et al.*, "A hybrid framework for heart disease prediction using classical and quantum-inspired ML techniques," *Scientific Reports*, vol. 15, no. 1, pp. 19957, 2025.
[10] R. Singh *et al.*, "Machine learning algorithms for heart disease diagnosis: A comparative evaluation," *Knowledge-Based Systems*, vol. 298, no. 106104, 2025.

