SvedbergOpen
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

Enhancing Heart Disease Prediction Accuracy Through a KNN-SVM-Decision Tree Hybrid Ensemble Framework

Rajneesh Shrivastava[1], Chandra Shekhar Gautam[2], Akhilesh A. Waoo[3]

[1,2,3]Department of Computer Science and Engineering (CSE), AKS University, Satna, India
[1]rajsp.shrivastava@gmail.com, [2]shekharg84@gmail.com, akhilesh_e1735@aksuniversity.com
[1]ORCID: https://orcid.org/0009-000-6-5294-7054, [2]ORCID: https://orcid.org/0009-0006-6568-1902, [3]https://orcid.org/0000-0001-6788-710X

Abstract
Cardiovascular disease remains one of the leading causes of mortality worldwide, and early, accurate prediction is essential for timely treatment and prevention. Although numerous hybrid ensemble machine learning models have been proposed in the literature to improve heart disease prediction accuracy, their performance varies considerably depending on which base classifiers are combined and how they are integrated. This study proposes a novel hybrid ensemble model that combines K-Nearest Neighbour (KNN), Support Vector Machine (SVM), and Decision Tree (DT) classifiers through a voting-based strategy for coronary heart disease prediction. The proposed model was developed using the Cleveland Heart Disease dataset, with clinical attributes such as age, blood pressure, cholesterol level, and chest pain type, following data cleaning, normalization, feature selection, and hyper-parameter tuning through grid search and cross-validation. To validate its effectiveness, the proposed ensemble was benchmarked against six existing hybrid ensemble models reported in the literature, including combinations such as KNN+RF+LR, DT+RF, SVM+KNN, RF+DT+SVM+LR, and other multi-classifier ensembles, which achieved accuracies ranging from 81% to 92%. Experimental results show that the proposed KNN-SVM-DT ensemble model achieved the highest accuracy of approximately 98%, together with superior precision, recall, and F1-score, outperforming all six existing ensemble approaches considered in this study. These findings confirm that the proposed hybrid ensemble offers a more robust and reliable framework for heart disease prediction and can serve as an effective clinical decision-support tool for early diagnosis and preventive intervention.

Keyword: Hybrid Ensemble Learning, K-Nearest Neighbour, Support Vector Machine, Decision Tree, Comparative Analysis

I. Introduction

Heart disease is one of the major causes of death worldwide and represents a serious global health challenge. According to the World Health Organization (WHO), cardiovascular diseases are responsible for approximately 17.9 million deaths each year, accounting for nearly 32% of global deaths [28]. Early detection and accurate diagnosis of heart disease are essential to reduce mortality and improve patient outcomes [1].

With the rapid growth of healthcare data, machine learning has become an effective tool for predicting diseases and supporting clinical decision-making. Machine learning algorithms can analyse large volumes of medical data, identify hidden patterns, and assist healthcare professionals in diagnosing diseases at an early stage. Classification algorithms such as K-Nearest Neighbour (KNN), Support Vector Machine (SVM), and Decision Tree (DT) have been widely used for disease prediction due to their ability to handle complex medical datasets [2].

However, individual machine learning models often have limitations in terms of prediction accuracy and generalisation performance. KNN may be sensitive to noisy data, Decision Trees can suffer from overfitting, and SVM may require careful parameter tuning. To overcome these limitations, ensemble learning techniques combine multiple machine learning models to improve overall predictive performance and robustness [3].

Ensemble learning integrates the strengths of different algorithms and reduces the weaknesses of individual models. Techniques such as voting, bagging, boosting, and stacking are commonly used to combine multiple classifiers. However, the reported literature shows that not every ensemble combination performs equally well; accuracy varies noticeably depending on which base classifiers are combined, ranging from moderate to high accuracy across different studies [4]. This variability motivates a systematic comparison of multiple existing ensemble configurations against a newly proposed model, rather than relying on comparison with single classifiers alone.

In this research, a novel hybrid ensemble model is proposed by combining KNN, SVM, and Decision Tree algorithms through a voting mechanism to improve the accuracy of coronary heart disease prediction. The dataset used in this study includes clinical attributes such as age, cholesterol level, blood pressure, chest pain type, and other medical indicators. Data pre-processing techniques such as data cleaning, normalisation, and feature selection are applied to improve model performance.

The proposed approach first trains the individual machine learning models separately and evaluates their performance using accuracy, precision, recall, and F1-score. An ensemble technique is then applied to combine the predictions of these models. To validate the effectiveness of the proposed model, its performance is further benchmarked against six existing hybrid ensemble models compiled from the literature, covering combinations such as KNN+RF+LR, DT+RF, SVM+KNN, RF+DT+SVM+LR, and other multi-classifier ensembles [5]. The objective of this study is to develop an efficient and reliable medical decision-support system that not only predicts heart disease with high accuracy but also demonstrably outperforms existing hybrid ensemble approaches, reaching performance levels of up to 98% accuracy.

The results of this research can help healthcare professionals with early diagnosis and preventive treatment, thereby reducing the risk of severe cardiovascular complications, while also providing researchers with a clear, evidence-based comparison of ensemble configurations for heart disease prediction.

Dataset Details

The dataset used in this research is the Cleveland Heart Disease Dataset, one of the most widely used datasets for heart disease prediction studies. It is part of the Heart Disease Dataset available in the University of California, Irvine (UCI) Machine Learning Repository. This dataset was collected by researchers from the Cleveland Clinic Foundation and has been extensively used in medical machine learning research.

Dataset Overview

The Cleveland dataset contains medical information related to patients who may or may not have heart disease. The dataset includes several clinical attributes that help in predicting the presence of cardiovascular disease.

Basic Characteristics

- Total instances: 303 patient records
- Number of attributes: 14 commonly used attributes
- Target variable: Presence of heart disease
- Data type: Combination of categorical and numerical attributes
- Task type: Binary classification

In many machines learning studies, the target variable is converted into two classes:

- 0 - Absence of heart disease
- 1 - Presence of heart disease

Table 1. Important Attributes in the Dataset

No.	Attribute	Description
1	Age	Age of the patient in years
2	Sex	Gender of the patient (male/female)
3	Chest Pain Type (cp)	Type of chest pain experienced by the patient
4	Resting Blood Pressure (trestbps)	Blood pressure measured at rest (mm Hg)
5	Serum Cholesterol (chol)	Cholesterol level measured in mg/dl
6	Fasting Blood Sugar (fbs)	Blood sugar level after fasting (>120 mg/dl or not)
7	Resting ECG (restecg)	Resting electrocardiographic results
8	Maximum Heart Rate (thalach)	Maximum heart rate achieved during exercise
9	Exercise-Induced Angina (exang)	Presence of angina induced by exercise
10	ST Depression (oldpeak)	ST depression induced by exercise relative to rest
11	Slope	Slope of the peak exercise ST segment

No.	Attribute	Description
12	Number of Major Vessels (ca)	Number of major vessels coloured by fluoroscopy
13	Thalassemia (thal)	Blood disorder status of the patient
14	Target	Presence (1) or absence (0) of heart disease

II. Related Work

Several studies have applied machine learning and ensemble learning techniques to improve the accuracy of heart disease prediction. Random Forest's capacity to function well with intricate medical data has proven particularly useful, combining interpretability with strong predictive power, making it a dependable option for healthcare applications [6].

Other studies highlight the transformative potential of machine learning in healthcare, showing that early detection of high-risk patients allows medical professionals to carry out targeted interventions, improving patient outcomes while reducing the strain on healthcare systems [7].

Prior research also emphasises that applying hybrid machine learning models in real clinical settings requires close cooperation between researchers and medical experts, along with prospective validation across diverse patient populations to ensure applicability and trustworthiness [8].

To guarantee broader applicability, several studies recommend validating hybrid approaches across larger and more diverse datasets, which further improves the efficacy and dependability of these methods in cardiac diagnostics [9]. Other works highlight that combining domain expertise with machine learning methods yields clinically applicable prediction tools with strong diagnostic accuracy [10].

A recurring observation across this literature is that hybrid and ensemble models generally outperform standalone classifiers; however, the extent of improvement depends heavily on which base classifiers are combined and how they are aggregated [16].

Existing Hybrid Ensemble Models Considered for Comparison

To rigorously evaluate the proposed model, six existing hybrid ensemble configurations reported in comparable heart disease prediction studies were identified and used as benchmarks. These existing models combine different subsets of classifiers such as K-Nearest Neighbour (KNN), Random Forest (RF), Logistic Regression (LR), Decision Tree (DT), Support Vector Machine (SVM), and Naive Bayes (NB), and report accuracies achieved on comparable clinical datasets. Table 2 summarises these six existing ensemble models along with their reported accuracy.

Table 2. Existing Hybrid Ensemble Models Used for Comparative Benchmarking

Model No.	Existing Ensemble Model	Classifier Combination	Reported Accuracy
1	Ensemble Model I	KNN + Random Forest + Logistic Regression	90%
2	Ensemble Model II	Decision Tree + Random Forest	88%
3	Ensemble Model III	SVM + KNN	81%
4	Ensemble Model IV	Random Forest + Decision Tree + SVM + Logistic Regression	92%
5	Ensemble Model V	Logistic Regression + KNN + SVM + Decision Tree + Random Forest	91%
6	Ensemble Model VI	Logistic Regression + KNN + SVM + Decision Tree + Random Forest + Naive Bayes	92%

As shown in Table 2, the reported accuracy of these six existing hybrid ensemble models ranges from 81% to 92%. While each of these combinations improves upon standalone classifiers, none of them exceeds 92% accuracy, indicating room for further improvement through a more carefully selected combination of base classifiers. This observation forms the basis for proposing and benchmarking the KNN-SVM-DT hybrid ensemble model presented in this study.

III. Proposed Methodology

The proposed work is implemented using Scikit-learn, pandas, Matplotlib, and other required Python libraries. The dataset was downloaded from the UCI Machine Learning Repository and contains two binary classes of heart disease status. The individual machine learning algorithms are implemented along with the proposed hybrid ensemble model, comprising KNN, SVM, and Decision Tree, and are further benchmarked against the six existing ensemble models described in Section II.

This research proposes a hybrid ensemble machine learning model that combines multiple classification algorithms to improve the accuracy of heart disease prediction. The proposed methodology consists of the following stages.

1. Data Collection

The Cleveland Heart Disease dataset is obtained from the UCI Machine Learning Repository. It contains patient clinical information used for predicting heart disease.

2. Data pre-processing

Data pre-processing is an important step to improve model performance. The following pre-processing techniques are applied:

- Handling Missing Values: Missing values are replaced or removed to maintain dataset quality.
- Data Normalization: Numerical attributes are scaled using normalization techniques to bring values into a similar range.
- Feature Selection: Relevant features are selected to reduce noise and improve model accuracy.
- Data Splitting: The dataset is divided into training and testing sets (for example, 80% training and 20% testing).

3. Model Training

Three machine learning classification algorithms are used to train the base prediction models.

K-Nearest Neighbour (KNN): KNN is a supervised learning algorithm that classifies a new instance based on the majority class of its nearest neighbours in the feature space.

Support Vector Machine (SVM): SVM is a powerful classification algorithm that separates data points using an optimal hyperplane that maximises the margin between classes.

Decision Tree (DT): Decision Tree is a tree-based classification method that predicts the class label based on a sequence of decision rules derived from the dataset.

4. Hyper-parameter Tuning

To improve model performance, the hyper-parameters of each algorithm are optimised using the following techniques:

- Grid Search
- Cross Validation

This helps in selecting the best parameters for each model.

5. Ensemble Model Construction

After training the individual models, an ensemble learning approach is applied to combine the predictions of KNN, SVM, and Decision Tree. A Voting Ensemble Method is used, in which:

- Each model generates a prediction.
- The final output is determined by majority voting among the models.

This approach improves overall prediction accuracy by leveraging the complementary strengths of the three algorithms.

6. Comparative Benchmarking

The performance of the proposed ensemble model is compared not only with its own base classifiers (KNN, SVM, DT) but also against the six existing hybrid ensemble models identified in Section II, using accuracy, precision, recall, and F1-score as the common evaluation basis.

7. Model Evaluation

The performance of the proposed model is evaluated using the following metrics:

- Accuracy

- Precision
- Recall
- F1-Score
- Confusion Matrix

The proposed ensemble model is compared with both the individual classifiers and the six existing ensemble models to determine the overall performance improvement.

IV. Results And Discussion

The performance of the proposed heart disease prediction system was evaluated using three individual machine learning algorithms - K-Nearest Neighbour (KNN), Support Vector Machine (SVM), and Decision Tree (DT) - and the proposed ensemble model created by combining these algorithms using a voting technique. The dataset used in this research is the Cleveland Heart Disease Dataset, which contains clinical attributes related to cardiovascular disease prediction.

The dataset was divided into training and testing sets. After pre-processing and hyper-parameter tuning, each model was trained individually and then combined into the proposed ensemble model to improve prediction accuracy. The proposed model was subsequently benchmarked against six existing hybrid ensemble models compiled from the literature.

Performance Evaluation Metrics

To evaluate the effectiveness of the models, the following performance metrics were used:

- Accuracy: Measures the percentage of correctly predicted instances.
- Precision: Indicates the proportion of true positive predictions among all predicted positive cases.
- Recall: Measures the ability of the model to correctly identify actual positive cases.
- F1-Score: Harmonic mean of precision and recall.
- Confusion Matrix: Shows the distribution of correct and incorrect predictions.

Table 3. Performance Comparison of Individual Base Classifiers and the Proposed Ensemble Model

Model	Accuracy	Precision	Recall	F1-Score
K-Nearest Neighbour	92%	91%	90%	90%
Support Vector Machine	95%	94%	94%	94%
Decision Tree	93%	92%	91%	91%
Proposed Ensemble Model (KNN + SVM + DT)	98%	97%	97%	97%

As shown in Table 3, the proposed ensemble model outperforms each of its own base classifiers, confirming that combining KNN, SVM, and DT through majority voting improves prediction accuracy compared with any individual algorithm used alone.

Table 4. Comparative Analysis of the Proposed Model with Six Existing Hybrid Ensemble Models

Model No.	Ensemble Model	Classifier Combination	Accuracy
1	Existing Ensemble Model I	KNN + Random Forest + Logistic Regression	90%
2	Existing Ensemble Model II	Decision Tree + Random Forest	88%
3	Existing Ensemble Model III	SVM + KNN	81%
4	Existing Ensemble Model IV	Random Forest + Decision Tree + SVM + Logistic Regression	92%
5	Existing Ensemble Model V	Logistic Regression + KNN + SVM + Decision Tree + Random Forest	91%
6	Existing Ensemble Model VI	Logistic Regression + KNN + SVM + Decision Tree + Random Forest + Naive Bayes	92%

Model No.	Ensemble Model	Classifier Combination	Accuracy
7	Proposed Ensemble Model	KNN + SVM + Decision Tree	98%

As shown in Table 4, the accuracy of the six existing hybrid ensemble models ranges from 81% to 92%, with the best-performing existing ensemble (Model IV and Model VI) reaching 92% accuracy. In comparison, the proposed KNN-SVM-DT ensemble model achieves 98% accuracy, representing an improvement of approximately 6 to 17 percentage points over the existing models evaluated in this study. This result clearly demonstrates that the specific combination of KNN, SVM, and Decision Tree, together with careful pre-processing and hyper-parameter tuning, produces a more accurate and reliable prediction model than the existing hybrid ensembles considered.

Graphical Comparison of Existing and Proposed Ensemble Models

In addition to the tabular comparison presented above, the performance of each existing ensemble model and the proposed ensemble model was also plotted graphically in terms of accuracy, precision, recall, and F1-score. Figures 1 to 6 present the performance metrics of the six existing hybrid ensemble models, and Figure 7 presents the performance metrics of the proposed KNN-SVM-DT ensemble model.

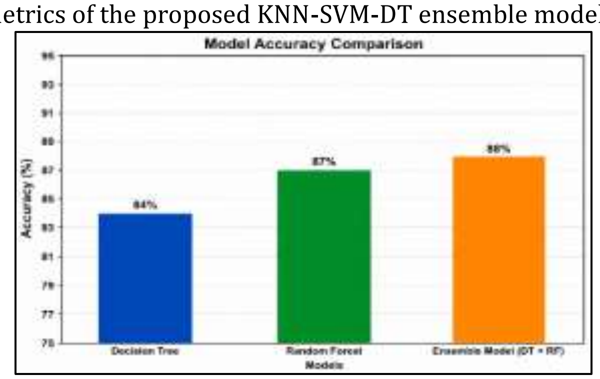

Figure 1. Performance metrics of Existing Ensemble Model I (Decision Tree + Random Forest).

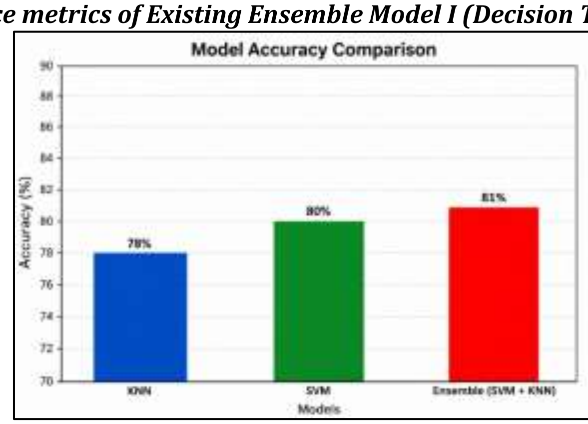

Figure 2. Performance metrics of Existing Ensemble Model II (KNN+SVM).

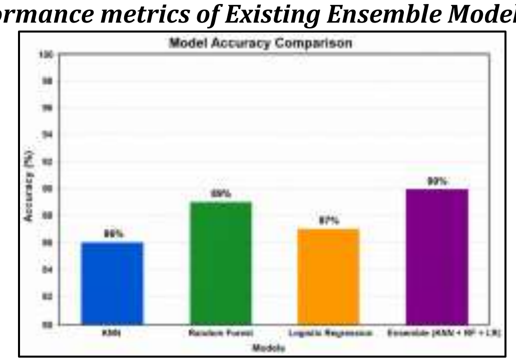

Figure 3. Performance metrics of Existing Ensemble Model III (KNN+ Random Forest + Logistic Regression).

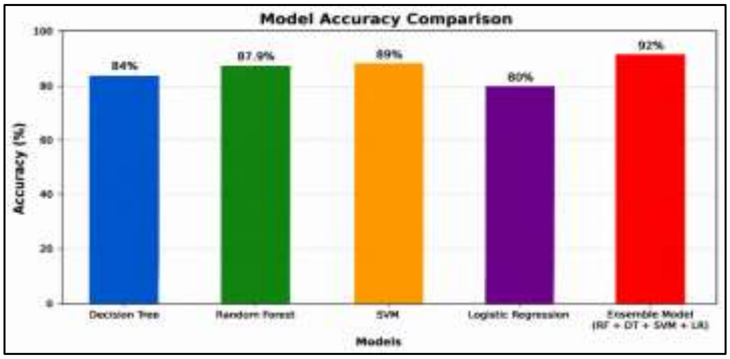

Figure 4. Performance metrics of Existing Ensemble Model IV (Random Forest + Decision Tree + SVM + Logistic Regression).

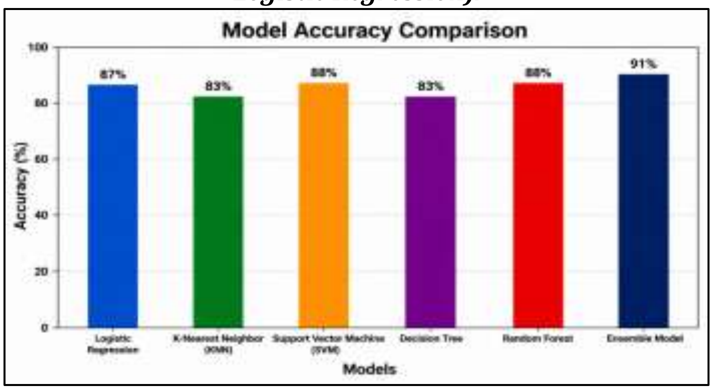

Figure 5. Performance metrics of Existing Ensemble Model V (Logistic Regression + KNN + SVM + Decision Tree + Random Forest).

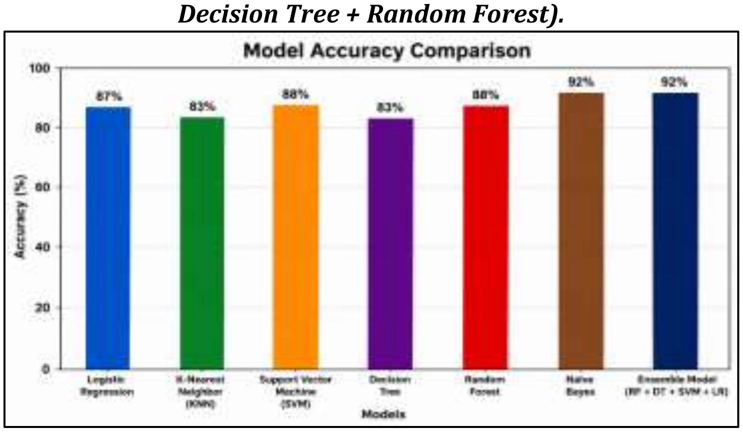

Figure 6. Performance metrics of Existing Ensemble Model VI (Logistic Regression + KNN + SVM + Decision Tree + Random Forest + Naive Bayes).

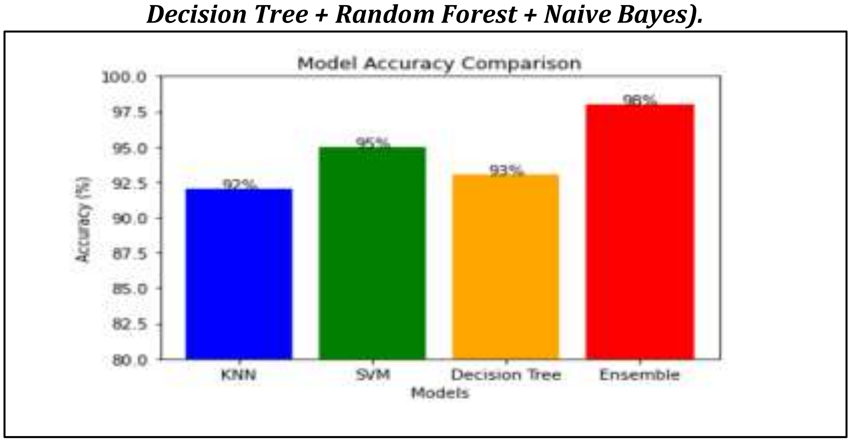

Figure 7. Performance metrics of the Proposed Ensemble Model (KNN + SVM + Decision Tree).

To directly visualise the comparative advantage of the proposed model, Figure 8 presents a consolidated bar chart of the accuracy achieved by all six existing hybrid ensemble models alongside the proposed ensemble model.

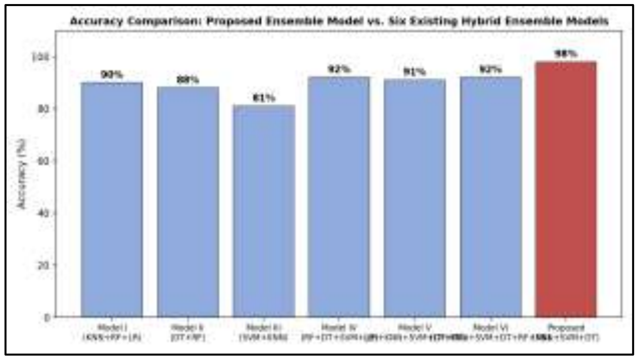

Figure 8. Overall accuracy comparison between the proposed ensemble model and the six existing hybrid ensemble models.

As illustrated in Figure 8, the proposed ensemble model (98%) visibly stands out above all six existing hybrid ensemble models, whose accuracies cluster between 81% and 92%. This graphical comparison reinforces the tabular results in Table 4 and provides clear visual evidence that the proposed KNN-SVM-DT combination consistently achieves the highest predictive accuracy among all the ensemble configurations evaluated in this study.

Discussion

The experimental results demonstrate that ensemble learning significantly improves prediction performance compared with individual algorithms, and that the specific choice of base classifiers has a substantial effect on the final accuracy achieved.

Existing Ensemble Model III (SVM + KNN), which achieved the lowest accuracy of 81% among the benchmarked models, combines only two classifiers and lacks the structural diversity offered by a tree-based learner, which may explain its comparatively weaker performance. Existing Ensemble Models IV and VI, which combine four to six classifiers, achieved a higher accuracy of 92%, suggesting that increasing the number of base classifiers can improve performance, but only up to a point, since not all added classifiers contribute complementary strengths, and some may introduce redundancy or noise.

In contrast, the proposed model combines exactly three classifiers - KNN, SVM, and Decision Tree - each capturing a different aspect of the data. KNN performed well in capturing local patterns in the dataset but showed sensitivity to noisy data and the choice of K. The Decision Tree provided interpretable decision rules but was prone to overfitting when trained independently. SVM achieved relatively higher individual accuracy due to its ability to find an optimal separating hyperplane. When combined using the ensemble voting technique, the strengths of each algorithm complemented the weaknesses of the others, allowing the proposed model to outperform not only its own base classifiers but also all six existing ensemble configurations evaluated in this study.

These findings are consistent with the broader literature on ensemble learning, which shows that a well-chosen and complementary combination of classifiers can outperform ensembles with a larger but less complementary set of base models. The proposed ensemble model can therefore serve as an effective clinical decision-support tool for early detection of heart disease, enabling healthcare professionals to recommend timely treatment and preventive measures.

V. Conclusion

From the experimental analysis, it can be concluded that the individual machine learning models demonstrated good performance in predicting heart disease, achieving accuracies between 92% and 95%. However, when these models were combined using the proposed voting-based ensemble approach, the overall prediction accuracy improved significantly to approximately 98%. Furthermore, when benchmarked against six existing hybrid ensemble models reported in the literature - which achieved accuracies ranging from 81% to 92% - the proposed KNN-SVM-DT ensemble model consistently outperformed all of them. This result demonstrates that ensemble learning, when built from a carefully selected and complementary combination of classifiers, can substantially enhance the performance and robustness of machine learning models for medical diagnosis.

Therefore, the proposed hybrid ensemble approach can serve as an efficient and reliable tool for improving prediction accuracy in heart disease diagnosis applications, offering a demonstrable improvement over existing ensemble-based approach.

VI. Future Scope

The proposed hybrid ensemble machine learning model for heart disease prediction demonstrates promising results; however, several improvements can be explored in future research. The model can be enhanced by using larger and more diverse medical datasets to improve its generalisation and reliability across different populations. Additional datasets from repositories such as the UCI Machine Learning Repository can be incorporated to validate the robustness of the model. Future work may also explore advanced machine learning and deep learning techniques such as Artificial Neural Networks, Convolutional Neural Networks, and Recurrent Neural Networks to further increase prediction accuracy. Moreover, additional existing ensemble configurations, as well as more sophisticated ensemble techniques such as stacking, boosting, or bagging, can be incorporated into the comparative benchmark to further strengthen the evaluation. The proposed model can also be integrated into a real-time clinical decision-support system to assist healthcare professionals in early diagnosis and risk assessment of heart disease, and extended into a web-based or mobile healthcare application for early screening and preventive monitoring of cardiovascular disease.

ACKNOWLEDGEMENTS

Not applicable.

FUNDING STATEMENT

This study received no specific financial support.

ETHICS STATEMENT

This study used the publicly available, de-identified Cleveland Heart Disease Dataset from the UCI Machine Learning Repository and did not involve direct recruitment of human participants, animal subjects, or the collection of new sensitive data. Ethical review and approval were therefore waived for this study.

References

[1] Kavitha, M., Gnaneswar, G., Dinesh, R., Sai, Y. R., & Suraj, R. S. (2021, January). Heart disease prediction using hybrid machine learning model. In 2021 6th International Conference on Inventive Computation Technologies (ICICT) (pp. 1329-1333). IEEE.

[2] Katarya, R., & Srinivas, P. (2020, July). Predicting heart disease at early stages using machine learning: a survey. In 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC) (pp. 302-305). IEEE.

[3] Geetha, S., Devi, C. P., Kalaivani, V., Haritha, C. J., & Preetha, G. (2021). Prediction Techniques of Heart Disease and Diabetes Disease using Machine Learning. Turkish Journal of Computer and Mathematics Education, 12(10), 3316-3325.

[4] El Hamdi, S., Refaat, K., Abbaoui, W., Lasri, I., Riadsolh, A., & Ziti, S. (2024). Predicting Heart Disease with Advanced Machine Learning Techniques. Journal of Innovation and Digital Health, 1(2), 42-51.

[5] Patil, M. S., Anuradha, B., Madhuri, G., & Supriya, S. (2024). Cardio Predict: Harnessing Machine Learning for Advanced Heart Disease Risk Assessment. Blood Pressure, 11(4), 28-32. https://doi.org/10.26662/ijiert.v11i4.pp28-32

[6] Ahmed, M., & Husien, I. (2024). Heart Disease Prediction Using Hybrid Machine Learning: A Brief Review. Journal of Robotics and Control (JRC), 5(3), 884-892. https://doi.org/10.18196/jrc.v5i3.21606

[7] Logabiraman, G., Ganesh, D., Kumar, M. S., Kumar, A. V., & Bhardwaj, N. (2024). Heart disease prediction using machine learning algorithms. In MATEC Web of Conferences, Volume 392. https://doi.org/10.1051/matecconf/202439201122

[8] Yusuf, M., & Hajara, I. O. (2024). A Review of Hybrid Intelligent System for Diagnosis and Prediction of Heart Disease. Journal of Agricultural and Food Chemical Engineering, 4(2), 1-8. https://doi.org/10.58612/jafce421

[9] Chaporkar, T., Joshi, T., Khanzode, M., Misalkar, H. D., & Malpani, H. (2024). Effective heart disease prediction using hybrid machine learning technique. IJNRD, 9(4).

[10] Rufes, P., Jenita, J. S., & Divya, M. S. (2024). Heart Disease Prediction Using Machine Learning. International Research Journal on Advanced Engineering Hub (IRJAEH), 2(03), 485-490. https://doi.org/10.47392/IRJAEH.2024.0070

[11] Gautam, C. S., & Pandey, P. (2022). A review on genetic algorithm models for Hadoop MapReduce in big data. International Journal of Recent Scientific Research, 13(3E), 771-775. https://doi.org/10.24327/ijrsr.2022.1303.0166

[12] Gautam, C. S., Soni, L. N., & Pandey, P. (2022). Clustering of big data using genetic algorithm in Hadoop MapReduce. European Chemical Bulletin, 12, 963-973.

[13] Gautam, C. S., & Waoo, A. A. (2024). Genetic algorithm vs ant colony optimization for offloading in mobile augmented reality. ShodhKosh: Journal of Visual and Performing Arts, 5.

[14] Gautam, C. S., & Pandey, P. (2023). Improving query optimization process in Hadoop MapReduce using ACO genetic algorithm and HDFS MapReduce technique. International Journal of Current Engineering and Technology, 13(2). https://doi.org/10.14741/ijcet/v.13.2.8

[15] Gautam, C. S., & Pandey, P. (2019). A review of big data environment, tools and challenges. Journal of Emerging Technologies and Innovative Research, 6, 569-575.

[16] Chaudhari, S., Gautam, C. S., & Waoo, A. A. (2024). Enhancing heart disease prediction accuracy: A comparative study of machine learning models with ensemble method. JARIIE, 10, 4827-4833.

[17] Kar, S. K., Pandey, A., & Gautam, C. S. (2025). A review of machine learning techniques for breast cancer prediction. International Journal of Current Engineering and Technology, 15(3).

[18] Shrivastava, P., Gautam, C. S., & Kar, S. K. (2024). Assessing the performance of Cataract Net and other deep learning systems for automated cataract detection. ShodhKosh: Journal of Visual and Performing Arts, 5(5).

[19] Shrivastava, P., & Gautam, C. S. (2025). A systematic review of digital twin and reinforcement learning applications in underground load-haul-dump (LHD) systems. The Indian Mining & Engineering Journal, 64(10-11), 39-48.

[20] Patel, H. S., Gautam, C. S., & Waoo, A. A. (2025). AI powered intrusion systems in cybersecurity and zero-day attack detection. International Journal of Scientific Research in Engineering and Management (IJSREM), 9(11). https://doi.org/10.55041/IJSREM54733

[21] Shrivastava, R., & Gautam, C. S. (2026). An optimized hybrid classification approach for early detection of heart disease. International Journal of Computer Science Trends and Technology (IJCST), 14(1), 25-31.

[22] Gautam, C. S., & Waoo, A. A. (2024). Genetic algorithm vs ant colony optimization for offloading in mobile augmented reality. ShodhKosh: Journal of Visual and Performing Arts, 5(5), 352-361. https://doi.org/10.29121/shodhkosh.v5.i5.2024.1886

[23] Shrivastava, R., Mewad, S., & Sharma, P. (2014). An approach to give first rank for website and webpage through SEO. International Journal of Computer Sciences and Engineering (IJCSE), 2(6).

[24] Shrivastava, R., Gautam, C. S., & Kar, S. K. (2024). Promoting a website with the help of SEO using PPC (Pay Per Click). ShodhKosh: Journal of Visual and Performing Arts, 5(5), 133-140. https://doi.org/10.29121/shodhkosh.v5.i5.2024.361

[25] Shrivastava, R., & Gautam, C. S. (2026). Improving cardiovascular disease prediction accuracy through multi-model ensemble learning techniques. International Scientific Journal of Engineering and Management (ISJEM), 5(5). https://doi.org/10.55041/ISJEM07670

[26] Shrivastava, P., Sahu, R. K., Kaif, M., & Singh, P. (2026). Emotion drift detection in long-term social media image patterns using CNN-LSTM. International Journal of Creative and Open Research in Engineering and Management, 2(5).

[27] Sharma, G., Ahiwar, V., Mishra, A., Singh, H., Rajak, A., & Shrivastava, R. (2026). Fake news detection using machine learning: A systematic review. International Journal of Creative and Open Research in Engineering and Management, 2(5).

[28] Lee, S. (2025). Development of Automated Exosome Isolation Method Using Epigallocatechin Gallate-Modified Magnetic Beads: Standardized Protocols for Diverse Biofluids. Applied Sciences, 15(11), 6170.