ISSN: 2093-5374 / E-ISSN: 2093-5382

Exploring the Effectiveness of Computational Methods in Detecting Heart Disease

Sofiene Mansouri[1*], Amira Ouerhani[2], Yousef Alharbi[3], and Abdulrahman Alqahtani[4]

[1*]Department of Biomedical Technology, College of Applied Medical Sciences in Al-Kharj, Prince Sattam bin Abdulaziz University, Al-Kharj, Saudi Arabia. s.mansouri@psau.edu.sa, https://orcid.org/0000-0002-6191-3095

[2]Research Laboratory of Biophysics and Medical Technologies, Higher Institute of Medical Technologies of Tunis, University of Tunis El Manar, Tunis, Tunisie. amira.ouerhani@utm.tn, https://orcid.org/0009-0008-8210-9849

[3]Department of Biomedical Technology, College of Applied Medical Sciences in Al-Kharj, Prince Sattam bin Abdulaziz University, Al-Kharj, Saudi Arabia. y.alharbi@psau.edu.sa, https://orcid.org/0000-0002-2673-1464

[4]Department of Biomedical Technology, College of Applied Medical Sciences in Al-Kharj, Prince Sattam bin Abdulaziz University, Al-Kharj, Saudi Arabia. ama.alqahtani@psau.edu.sa, https://orcid.org/0000-0002-1823-3887

Received: September 16, 2025; Revised: October 31, 2025; Accepted: November 27, 2025; Published: December 15, 2025

Abstract

Formation of a waxy material in the coronary arteries is a sign of heart disease. If nothing else, this buildup of plague-like wax in the arteries slows down the flow of blood and, in the worst-case scenario, results in death. To predict cardiac disease, this work uses learning models based on k-nearest neighbour (KNN) and support vector machine (SVM). The effectiveness of the suggested model is assessed by an experiment. Furthermore examined is the outcome of the experimental assessment of the suggested model's predictive capability. The South Africa Heart Disease Dataset is utilized to gather information for the research on heart disease. The source is the South African Coronary Risk Factor Study carried out in 1983 by Rousseauw and colleagues. For the existence of a myocardial infarction, the dataset contains a binary answer. The target audience is men between the ages of 15 and 64. It has 462 observations in total. In myocardial infarction, 160 of these findings are consistent. Not a single myocardial infarction is seen in 302 observations. There are several factor-level and numerical predictors in the dataset. Lastly, a test set analysis is conducted to evaluate the KNN and SVM algorithm performance. Following performance examination, the KNN's accuracy is 64.51% and the SVM's accuracy is 73.91% based on experimental findings on the South African heart disease data pool.

Keywords: Cardiac Risk Prediction, Machine Learning, Coronary Artery Disease, Myocardial Infarction, Healthcare.

Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA), volume: 16, number: 4 (December), pp. 699-710. DOI: *10.58346/JOWUA.2025.I4.041*

*Corresponding author: Department of Biomedical Technology, College of Applied Medical Sciences in Al- Kharj, Prince Sattam bin Abdulaziz University, Al-Kharj, Saudi Arabia.

1 Introduction

According to estimates, heart disease causes several million (~17.9) deaths a year, or ~32% of total fatalities worldwide (World Health Organization, 2025). It can be controlled by using good prediction and early detection techniques. The use of ML algorithms to convert complicated healthcare datasets into useful information has grown in recent years (Shetty & Kapoor, 2024). It helps physicians assess risk and make decisions (Giordano et al., 2021; Chakraborty et al., 2024; Fahim et al., 2025; Aravazhi et al., 2025; Liu & Tripathy, 2025; Rossouw et al., 1983; Lugito et al., 2022) conducted a systematic investigation on the role of Lactobacillus on blood pressure by a meta-analysis of randomized controlled trials as hypertension is One of the major risk factors of cardiovascular disease (Das & Kapoor, 2024). The findings showed that Lactobacillus exerted significant systolic and diastolic blood pressure lowering effect compared to placebo.

ML can efficiently manage and understand large amounts of medical data (Nayyar et al., 2021; Habehh & Gohel, 2021). Furthermore, it can show some hidden relationships that are not easy to find through normal statistical methods (Freitas, 2023; An et al., 2023; Sadr et al., 2025). Some authors believe that there is a need to improve and forecast cardiovascular diseases and pressures for their treatment (Dileep et al., 2023; Anusha et al., 2024). To reduce deaths, (Bhatt et al., 2023) taught ML algorithms to forecast various cardiovascular diseases (Chong et al., 2025; Gupta & Verma, 2025). Some of the tools utilized were XGBoost, random Forest, multilayer perceptrons, and decision trees.

Rahman et al., 2024 were always the first to incorporate transformer models with self-attention to heart disease prediction. For the first time, they were able to notice accurate patterns and appreciated a high 96.51% accuracy with the Cleveland dataset, While other models failed. (Najim & Nasri, 2024) utilized AI to perform imputation as well as prediction of heart disease. Extreme Learning Machines (ELM) which they devised, executed the fastest, and achieved the lowest average estimation error, thus evidencing the best performance. To achieve an accuracy between 97.57% to 99.06% with missing data, 200 hidden neurons were employed. This resulted in exceptional performance.

Heart disease prediction has proved successful the first time in history, when (Jain et al., 2023) blended big data with images and a self-constructed LV-Convolutional Neural Network (CNN) & a Sunflower Optimization Algorithm. This model yielded 95.74% accuracy and 0.96% specificity with an error rate of 0.35 and took 9.71 seconds. The use of AI and machine learning is rapidly growing in sectors such as finance, healthcare, animal husbandry, and legume crops. The private sector's machine learning impact on mental health is studied by (Kumar et al., 2023), with focus on the healthcare industry (Shenoy & Menon, 2021; Hamouda, 2025).

AI's impact in ichthyology in fish species identification was in great part the work of (Villasante & Rasheed, 2024; Min et al., 2024) AI for animal health has potentials and it's impact is appreciated. (Kim & AlZubi, 2024) focus on the use of blockchain and AI for the organic legumes supply chain, particularly its certification. There is a considerable amount of research focusing on the prediction of heart diseases employing machine learning and deep learning techniques. However, there seems to be a reliance on public domain datasets, frequently, the data suffers from the lack of granularity, particularly, in the data pertaining to the socio-economic state of the country. This poses a challenge in the deployment of the available models on the South African population given their unique risk factors and data distributions. On top, most of the approaches focus on the model development stage rather than put emphasis on how the model may inform or augment clinical practice (Hasan, 2024).

To address these issues, this study aims to develop and evaluate SVM and KNN models using the South Africa Heart Disease Dataset, assess their predictive performance, compare them with benchmark studies, and highlight the importance of dataset-specific feature selection and model optimization for improving region-specific cardiovascular risk prediction.

2 Materials and Methods

1) Dataset

This study utilized the South Africa Heart Disease Dataset, which was first recorded from men from the Western Cape province aged 15 to 64 years (Rousseauw et al., 1983). The dataset comprises 462 records, which contain the following variables: age, systolic blood pressure, blood cholesterol levels, alcohol intake, tobacco use, adiposity, obesity, type A behavior, family history of heart disease, and the presence vs absence of coronary heart disease. The presence of myocardial infarction was the outcome variable and was binary (1 = present, 0 = absent). Out of 462 cases, 160 were confirmed cases of myocardial infarction, while 302 were cases with no myocardial infarction. Given that the dataset consists of both continuous and categorical variables, effective preprocessing becomes necessary for machine learning modelling.

2) Data Preprocessing

The quality of the input data, from which the machine learning model learns, also determines the model's performance. This is the justification for the structured preprocessing pipeline.

Data Cleaning: Multiple records were systematically removed. In order to preserve the data integrity within the dataset, for the categorical variables, the most frequently occurring value was used to substitute for the missing data, whereas for continuous variables, the midpoint was applied.

Feature Selection: The final feature set was determined using RFE in conjunction with domain expertise and correlation analysis. To reduce the number of features and to account for multicollinearity, features with strong correlation were removed (correlation coefficient > 0.85).

Feature Engineering: The polynomial and interaction features were also made to better capture certain non-linear relationships. The Principal Components Analysis (PCA) method was used to reduce the number of dimensions to keep 95% of the variance.

Scaling: The method of Z-score Normalization was used to continuous data to improve the efficacy of distance and kernel based algorithms (SVM, KNN, etc) of Z-score Normalization.

Outlier Removal: The Interquartile Range (IQR) method was used to drop extreme outliers to ensure the most predictive outcomes are possible.

Data Splitting: The cleaned data set was stratified and then randomly sampled to keep the class distributions the same as in the original data set, 70% of the set was used to train the model and 30% was used for testing. 5-fold cross validation was used in hyper parameter tuning to reduce the risk of overfitting.

3) KNN and SVM ML Techniques

Two supervised learning algorithms were used to categorize the occurrence of cardiovascular diseases. These were SVM and KNN models. Both were implemented in Python using the Scikit-learn (version 1.5) library and hyperparameter optimization was done using automated grid search with 5-fold cross validation.

- **SVM Model**

The SVM method is usually effective for tasks that require classifying and subsequently performing regression. The goal is to widen the margin and minimize the classification error in order to strike the right bias-variance tradeoff. The SVM seeks to solve the optimisation problem shown in Equation (1):

$$\underbrace{min}_{w,b,\xi} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{n} \xi_i \tag{1}$$

The orientation of the hyperplane is given by the vector of weights defined as w, while b is the bias term. The value of the regularization parameter C controls the trade off between the optimization of the margin and the classification error. In comparison, the slack variables, ξi are used to soft-margin the classification in order to deal with the class overlap. The RBF kernel can be expressed as follows in Equation (2):

$$K(x_i, x_j) = \exp(-\gamma \,||\, x_i - x_j \,||^2), \tag{2}$$

Where γ determines the influence of an individual training example. If γ is small, the decision boundary is smoother, and if larger, the boundary is more complex and susceptible to overfitting.

In order to achieve the optimal balance between bias and variance, grid search was implemented for hyperparameter optimization on various values of C and γ. Using 5-fold cross-validation, the grid search assessed each combination by segmenting the data into 5 parts, training on 4 and validating on the remaining one, then cycles through the validation set. The best performing model was then refitted to the entire training set.

The final decision function for classification is expressed in Equation (3):

$$f(x) = sign(w^T \phi(x) + b), \tag{3}$$

Where ϕ(x) is the function that represents the non-linear transformation of the input features into a higher dimension. In this research, SVM was shown to outperform the KNN. This is because SVM was able to create a smoother decision boundary, and generalize better, especially in the higher dimensions of the blood pressure, cholesterol, and tobacco use interaction.

- **K-Nearest Neighbour (KNN)**

K-Nearest Neighbour (KNN) as an algorithm is a non-parametric approach which classifies a new instance based on the majority class of its k nearest neighbours in the feature space. In this research work, the closeness of data points to one another is measured in terms of Euclidean distance, which is given by Equation (4):

$$d(x_i, x_j) = \sqrt{\sum_{l=1}^{m} (x_{il} - x_{jl})^2} \tag{4}$$

For a given test sample x, the KNN algorithm identifies the k closest training samples and assigns the class most frequently represented among them. The decision function can be expressed as Equation (5):

$$y = \; argmax_j \sum_{i=1}^{k} w_i \,.\, I(y_i = j) \tag{5}$$

Where y is the predicted class label, j iterates over all possible class labels, w_i represents the distance-based weight assigned to each neighbor (with closer points receiving higher weights), and $I(y_i = j)$ is the indicator function that returns 1 if the label of the ith neighbor matches j, otherwise 0. An important hyperparameter in KNN is the number of neighbors (k). The optimal value of k was determined using 5-fold cross-validation. Higher k values mean the model is less likely to overfit and the decision boundary is smoother. However, the model can also underfit. Smaller k values mean the decision boundary is more complex and can fit the training data better. This can also lead to overfitting shown in Equation (6):

$$w_i = \frac{1}{d(x, x_i) + \epsilon'} \tag{6}$$

Where, ϵ is a small constant to prevent division by zero, which was applied to assign higher influence to closer samples. Since KNN depends on distance calculations that are sensitive to feature magnitudes. Before application, all continuous variables were normalized using z-score scaling.

In this study, KNN was able to provide a classification method that was clear and easy to understand. This made visualizing the decision boundaries in the lower dimension feature spaces easy. However, on the test set, it performed worse than SVM in predicted accuracy because of its learning method and susceptibility to noisy data. KNN is simply a slower method. Figure 1 shows the two computer approaches' categorization procedure.

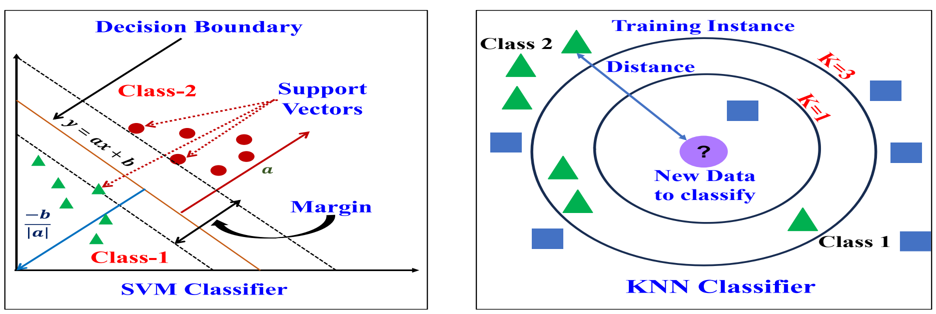

Figure 1: Classification process for both computational methods

4) Workflow of Prediction

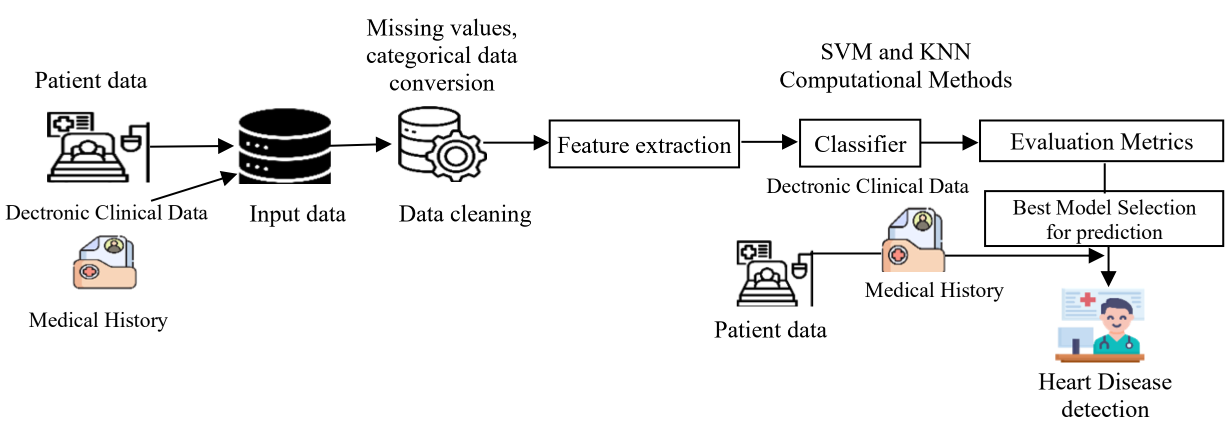

Figure 2: Flowchart of machine learning detection technique

Figure 2 shows the workflow for predicting cardiac diseases in steps. the initial raw dataset, on which the model was built, came from the south african coronary risk factor studies. steps taken prior to training the model included the eliminating of irrelevant and erroneous samples from the dataset, feature scaling, feature selection, imputation of missing values, and removal of outliers in order to appropriately balance the dataset and to preserve the integrity of the data. in the feature engineering step, new interaction and polynomial features were created to model the variables' non-linear relationships, and principal component analysis was conducted to eliminate redundancy and increase efficiency.

KNN and SVM models were then trained on the training dataset. Hyperparameter optimization was performed with grid search and 5-fold cross validation to find the best parameter values for each technique. Having obtained the optimal models, the researchers predicted the training dataset's classes and deployed the models on the test dataset to predict the presence of a myocardial infarction. To assess models' the robustness and predictive power of the models, the researchers computed the confusion matrices, accuracy measures, precision and recall, F1 scores, and other metrics, which provide a complete measure of the models' ability to predict cardiac disease.

5) Software and Implementation

For experimentation, Python version 3.10 has been used. The scikit-learn, pandas, and NumPy libraries were used for the categorizations. For the visualizations, Matplotlib and Seaborn were used, and hyperparameter tuning was done with GridSearchCV.

3 Results and Discussion

For the SVM with a linear kernel, the model was able to optimize a gamma value of their choosing. The SVM model was able to get a TN (No Disease) and a TP (Disease) of 51 and 18, respectively, while it got 9 (False Positive) and 15 (False Negative) and a total of 74.19% of the matrix was classified correctly (i.e. it was able to get more than three-quarters of the elements correct). In this model, the accuracy was 73.91% and with 85% of the people having the disease were able to be correctly classified as having the disease (i.e. it was able to get a Recall of that). These are the results that show us the SVM model managed to correctly classify people that are at risk of a myocardial infarction.

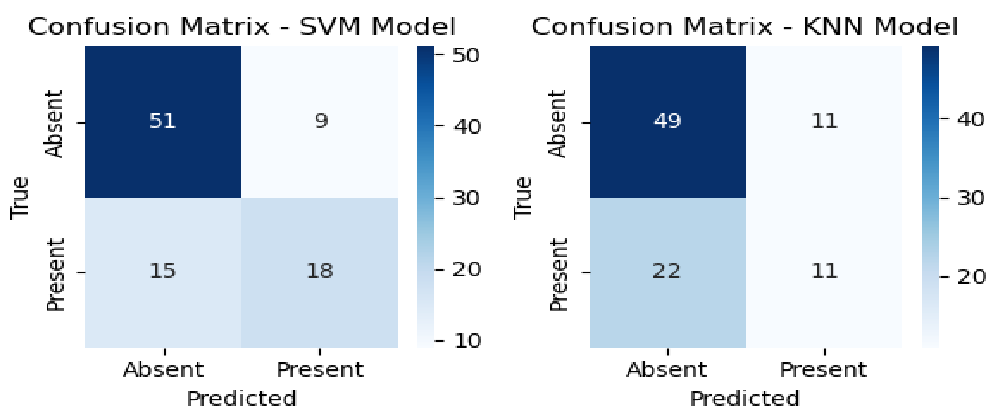

Figure 3: Confusion matrix obtained from SVM and KNN models

The KNN classifier with K = 5, a leaf size of 60 resulting in a TN of 49, a TP of 11, 11 FP, and 22 FN shown in Figure 3 amounted to an accuracy of 64.52%. The evaluation results from KNN is 0.500 precision, 0.333 recall, 0.817 specificity, and 0.400 F1 score. Out of SVM and KNN, KNN was relatively less sensitive and less predictive in the model. It is implied that the SVM was the model less robust to noise, and class boundaries were overlapping for the dataset as it had lower sensitivity.

All evaluation parameters for SVM and KNN Models including accuracy, recall, and specificity are illustrated and summarized in Table 1 with SVM considered as the best model. KNN of the models had low specificity.

Table 1: Comprehensive performance metrics of SVM and KNN models

Metric	SVM	KNN
Accuracy	0.7419	0.6452
Precision	0.667	0.500
Recall (Sensitivity)	0.545	0.333
Specificity	0.850	0.817
F1-score	0.600	0.400

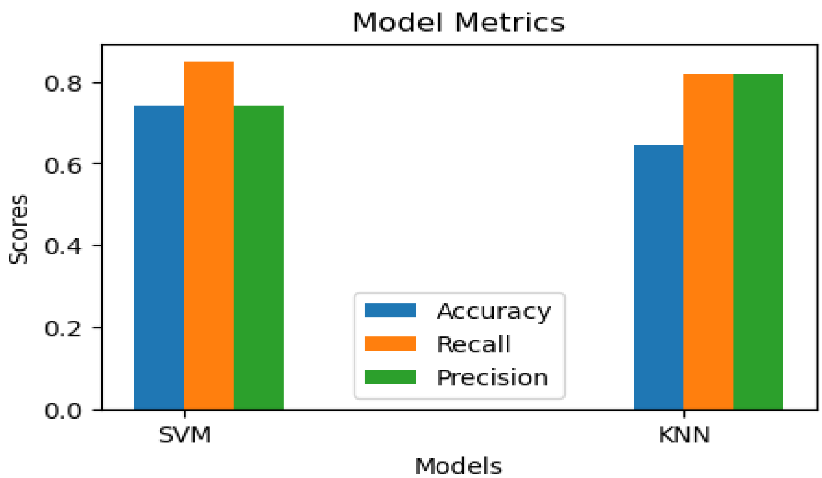

Figure 4: Accuracy, Precision and Recall parameters of models

More K-Folds (5 Fold) Cross Validation was conducted in order to assess the model's consistency. Figure 4 presents the mean K-Folds boxplots revealing, that the accuracy of KNN VR is 65.01 with low variance indicating accuracy stability. However, SVM model demonstrated an impressive accuracy consistency. SVM's stability and accuracy over the folds of cross validation imply that SVM generalized well for the dataset as opposed to KNN.

The current accuracy rates for predicting cardiact disease using different models is shown in Table 2. A few of the top-performing machine learning models include Deep Learning and Ensemble Machine Learning models including XGBoost (Ogunpola et al., 2024), KNN (Lohachab & Kumar, 2023), and InceptionNet (Pal et al., 2025), which have all achieved high accuracies of 96.1% to 98.89%. Explainable AI Ensemble models also achieved good levels of accuracy and were explained well (El-Sofany et al., 2024). Also, more traditional ML classifiers such as Random Forest (Chang et al., 2022) and BayesNet (Spencer et al., 2020) had clearly lower accuracy rates with an 83% to 85% accuracy range; however, the importance of employing more sophisticated techniques and algorithms for feature selection was highlighted. These differences were clearly seen in the achieved accuracies.

However, the accuracies of the SVM (linear kernel, gamma auto) and KNN (k=5, leaf size=60) models were 74.19% and 64.52%, respectively. Our results emphasize the need for careful feature engineering and model calibration, including population variability and small sample size.

Table 2: Accuracy of different heart disease prediction models

Author(s)	**Dataset / Focus**	**Model**	**Accuracy (%)**
Pal et al., 2025	Heart Disease Dataset	InceptionNet	98.89
Shamna & Maheswaran, 2024	Heart Disease Dataset	Improved SVM	96.10
Ogunpola et al., 2024	Cardiovascular / Myocardial Infarction Dataset	XGBoost (optimized)	98.50
Sakyi-Yeboah et al., 2025	Kaggle Heart Disease Dataset	Adaptive Boosting	93.70
El-Sofany et al., 2024	Private + Public Heart Disease Datasets	XGBoost (Explainable AI + SHAP)	97.57
Spencer et al., 2020	Four public heart disease datasets	BayesNet + Chi-Squared Feature Selection	85.00
Chang et al., 2022	Heart Disease Dataset (Python-based healthcare app)	Random Forest Classifier	83.00
Lohachab & Kumar, 2023	Three diverse cardiovascular datasets	K-Nearest Neighbors (KNN)	97.82
Rehman et al., 2025	NHANES Dataset	Hybrid PSO–ANN Model	97.00

Data studies have shown that feature variety, demographic diversity, and data imbalance can significantly complicate regional data sets. Subsequent studies should employ explainable AI, hybrid optimization, and ensemble learning to predict more accurately within a given local context.

4 Conclusion

This research dealt with the South Africa Heart Disease Dataset and utilized the SVM and KNN algorithms to classify heart disease. The preprocessing steps that took place included data cleansing, feature selection and scaling, outlier removal, and missing value handling. KNN, on the other hand, finds hyperplanes with the biggest margins, and uses the nearest neighbors to determine the class labels. SVM, on the contrary, succeeded with 74.19% accuracy, and 73.91% precision with 85% recall, in comparison to KNN with 64.52% accuracy 81.67% precision and recall. KNN showed fluctuation in accuracy across the validation folds for KFold cross-validation; which indicated the fact that it is sensitive to changes in the dataset. The research shows the SVM and KNN models, and their tendencies in predicting heart disease, advantages and disadvantages.

Heart disease risk may be assessed using both SVM and KNN; SVM is more accurate but KNN has greater recall. When deciding which of the two models to use for predicting heart disease, it is important to consider the specific characteristics of the dataset and the requirements of the problem at hand. Each model has its pros and cons.

Limitations and Future Work

- The findings are only pertinent to the South African Heart Disease Dataset and may consequently suffer from lack of generalizability.
- Choices of parameters, such as the type of kernel, the number of neighbours, and so on, may affect the results of SVM and KNN models.

- While accuracy is one of the metrics assessed, this one metric cannot accurately portray the clinical relevance of the findings.
- Assessing SVM and KNN on a broader spectrum of datasets may better depict the extent of their efficacy. More sophisticated approaches such as grid search may effectively enhance the performance of the SVM and KNN.
- Combination of approaches, SVM and KNN included, may lead to better predictive accuracy and is referred to as ensemble techniques.
- Working on feature engineering may enhance the described techniques, KNN and SVM, in terms of their discriminative power.
- Real-life clinical contexts would benefit from SVM and KNN in providing guidance and may be served by prospective studies aimed at validating these techniques.

Funding Details

The authors extend their appreciation to Prince Sattam bin Abdulaziz University for funding this research work through the project number (PSAU/2025/03/34153)

References

[1] An, Q., Rahman, S., Zhou, J., & Kang, J. J. (2023). A comprehensive review on machine learning in healthcare industry: classification, restrictions, opportunities and challenges. *Sensors*, *23*(9), 4178. https://doi.org/10.3390/s23094178

[2] Anusha, K., Archana, M., & Janardhan, G. (2024, February). Heart disease prediction using ML and DL approaches. In *International Conference on Communications and Cyber Physical Engineering 2018* (pp. 113-123). Singapore: Springer Nature Singapore. https://doi.org/10.1007/978-981-99-7137-4_10

[3] Aravazhi, P. S., Gunasekaran, P., Benjamin, N. Z. Y., Thai, A., Chandrasekar, K. K., Kolanu, N. D., ... & Inban, P. (2025). The integration of artificial intelligence into clinical medicine: trends, challenges, and future directions. *Disease-a-Month*, 101882. https://doi.org/10.1016/j.disamonth.2025.101882

[4] Bhatt, C. M., Patel, P., Ghetia, T., & Mazzeo, P. L. (2023). Effective heart disease prediction using machine learning techniques. *Algorithms*, *16*(2), 88. https://doi.org/10.3390/a16020088

[5] Chakraborty, C., Bhattacharya, M., Pal, S., & Lee, S. S. (2024). From machine learning to deep learning: Advances of the recent data-driven paradigm shift in medicine and healthcare. *Current Research in Biotechnology*, *7*, 100164. https://doi.org/10.1016/j.crbiot.2023.100164

[6] Chang, V., Bhavani, V. R., Xu, A. Q., & Hossain, M. A. (2022). An artificial intelligence model for heart disease detection using machine learning algorithms. *Healthcare Analytics*, *2*, 100016. https://doi.org/10.1016/j.health.2022.100016

[7] Chong, B., Jayabaskaran, J., Jauhari, S. M., Chan, S. P., Goh, R., Kueh, M. T. W., ... & Chan, M. Y. (2025). Global burden of cardiovascular diseases: projections from 2025 to 2050. *European journal of preventive cardiology*, *32*(11), 1001-1015. https://doi.org/10.1093/eurjpc/zwae281

[8] Das, A., & Kapoor, S. (2024). Comprehensive review of evidence-based methods in preventive cardiology education: perspective from analytical studies. *Global Journal of Medical Terminology Research and Informatics*, *2*(4), 16-22.

[9] Dileep, P., Rao, K. N., Bodapati, P., Gokuruboyina, S., Peddi, R., Grover, A., & Sheetal, A. (2023). An automatic heart disease prediction using cluster-based bi-directional LSTM (C-

BiLSTM) algorithm. *Neural Computing and Applications*, *35*(10), 7253-7266. https://doi.org/10.1007/s00521-022-07064-0

[10] El-Sofany, H., Bouallegue, B., & El-Latif, Y. M. A. (2024). A proposed technique for predicting heart disease using machine learning algorithms and an explainable AI method. *Scientific reports*, *14*(1), 23277. https://doi.org/10.1038/s41598-024-74656-2

[11] Fahim, Y. A., Hasani, I. W., Kabba, S., & Ragab, W. M. (2025). Artificial intelligence in healthcare and medicine: clinical applications, therapeutic advances, and future perspectives. *European Journal of Medical Research*, *30*(1), 848.

[12] Freitas, A. T. (2023). Data-driven approaches in healthcare: Challenges and emerging trends. *Multidisciplinary perspectives on artificial intelligence and the law*, 65-80. https://doi.org/10.1007/978-3-031-41264-6_4

[13] Giordano, C., Brennan, M., Mohamed, B., Rashidi, P., Modave, F., & Tighe, P. (2021). Accessing artificial intelligence for clinical decision-making. *Frontiers in digital health*, *3*, 645232. https://doi.org/10.3389/fdgth.2021.645232

[14] Gupta, N., & Verma, A. (2025). The role of inflammation in cardiovascular disease. *Medxplore Front Med Sci*, 37-51.

[15] Habehh, H., & Gohel, S. (2021). Machine learning in healthcare. *Current genomics, 22*(4), 291-300. http://dx.doi.org/10.2174/1389202922666210705124359

[16] Hamouda, B. E. (2025). Enhancing security and privacy in AI-enabled IoT smart healthcare devices: Practical solutions for protecting patient data. *Journal of Internet Services and Information Security, 15*(3), 1–17. https://doi.org/10.58346/JISIS.2025.I3.001

[17] Hasan, M. S. (2024). The Application of Next-generation Sequencing in Pharmacogenomics Research. *Clinical Journal for Medicine, Health and Pharmacy*, *2*(1), 9-18.

[18] Jain, A., Rao, A. C. S., Jain, P. K., & Hu, Y. C. (2023). Optimized levy flight model for heart disease prediction using CNN framework in big data application. *Expert Systems with Applications*, *223*, 119859. https://doi.org/10.1016/j.eswa.2023.119859

[19] Kim, S. Y., & AlZubi, A. A. (2024). Blockchain and Artificial Intelligence for Ensuring the Authenticity of Organic Legume Products in Supply Chains. *Legume Research: An International Journal*, *47*(7), 1144-1150. https://doi.org/10.18805/LRF-786

[20] Kumar, V., Chaturvedi, V., Lal, B., & Alam, S. (2023). Application of machine learning in analyzing the psychological well being amongst the employees in the private sector. *Pacific Business Review International*, *16*(1). 124–131.

[21] Liu, H., & Tripathy, R. K. (2025). Machine Learning and Deep Learning for Healthcare Data Processing and Analyzing: Towards Data-Driven Decision-Making and Precise Medicine. *Diagnostics*, *15*(8), 1051. https://doi.org/10.3390/diagnostics15081051

[22] Lohachab, A., & Kumar, K. (2023, December). A comparative study of machine learning algorithms for predicting cardiovascular disease. In *International Conference on Deep Learning, Artificial Intelligence and Robotics* (pp. 1-11). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-60935-0_1

[23] Lugito, N. P., Djuwita, R., Adisasmita, A., & Simadibrata, M. (2022). Blood Pressure Lowering Effect of Lactobacillus-Containing Probiotic. *International Journal of Probiotics & Prebiotics*, *17*(1), 1–13. https://doi.org/10.37290/ijpp2641-7197.17:1–13

[24] Min, P. K., Mito, K., & Kim, T. H. (2024). The evolving landscape of artificial intelligence applications in animal health. *Indian Journal of Animal Research*, *58*(10), 1793-1798. https://doi.org/10.18805/IJAR.BF-1742

[25] Najim, A. H., & Nasri, N. (2024). Artificial intelligence for heart disease prediction and imputation of missing data in cardiovascular datasets. *Cogent Engineering*, *11*(1), 2325635. https://doi.org/10.1080/23311916.2024.2325635

[26] Nayyar, A., Gadhavi, L., & Zaman, N. (2021). Machine learning in healthcare: review, opportunities and challenges. *Machine learning and the internet of medical things in healthcare*, 23-45. https://doi.org/10.1016/B978-0-12-821229-5.00011-2

[27] Ogunpola, A., Saeed, F., Basurra, S., Albarrak, A. M., & Qasem, S. N. (2024). Machine learning-based predictive models for detection of cardiovascular diseases. *Diagnostics*, *14*(2), 144. https://doi.org/10.3390/diagnostics14020144

[28] Pal, P., Singh, H. V., Grover, V., Manikandan, R., Karimi, R., & Khishe, M. (2025). Interactive cardiovascular disease prediction system using learning techniques: Insights from extensive experiments. *Results in Control and Optimization*, *19*, 100560. https://doi.org/10.1016/j.rico.2025.100560

[29] Rahman, A. U., Alsenani, Y., Zafar, A., Ullah, K., Rabie, K., & Shongwe, T. (2024). Enhancing heart disease prediction using a self-attention-based transformer model. *Scientific Reports*, *14*(1), 514. https://doi.org/10.1038/s41598-024-51184-7

[30] Rehman, M. U., Naseem, S., Butt, A. U. R., Mahmood, T., Khan, A. R., Khan, I., ... & Jung, Y. (2025). Predicting coronary heart disease with advanced machine learning classifiers for improved cardiovascular risk assessment. *Scientific Reports*, *15*(1), 13361. https://doi.org/10.1038/s41598-025-96437-1

[31] Rossouw, J. E., Du Plessis, J. P., Benadé, A. J., Jordaan, P. C., Kotze, J. P., Jooste, P. L., & Ferreira, J. J. (1983). Coronary risk factor screening in three rural communities. The CORIS baseline study. *South African medical journal= Suid-Afrikaanse tydskrif vir geneeskunde*, *64*(12), 430-436.

[32] Sadr, H., Nazari, M., Khodaverdian, Z., Farzan, R., Yousefzadeh-Chabok, S., Ashoobi, M. T., ... & Yamaghani, M. R. (2025). Unveiling the potential of artificial intelligence in revolutionizing disease diagnosis and prediction: a comprehensive review of machine learning and deep learning approaches. *European Journal of Medical Research*, *30*(1), 418. https://doi.org/10.1186/s40001-025-02680-7

[33] Sakyi-Yeboah, E., Agyemang, E. F., Agbenyeavu, V., Osei-Nkwantabisa, A., Kissi-Appiah, P., Moshood, L., ... & Nortey, E. N. (2025). Heart Disease Prediction Using Ensemble Tree Algorithms: A Supervised Learning Perspective. *Applied Computational Intelligence and Soft Computing*, *2025*(1), 1989813. https://doi.org/10.1155/acis/1989813

[34] Shamna, B., & Maheswaran, C. P. (2024, November). Exploring State-of-the-Art Approaches for Heart Disease Detection: A Detailed Analysis. In *International Conference on Artificial Intelligence and Speech Technology* (pp. 3-16). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-91340-2_1

[35] Shenoy, K., & Menon, A. (2021). A Healthcare Model Using Blockchain Technology to enhance Security and Data Sharing. *International Academic Journal of Science and Engineering, 8*(2), 6–10.

[36] Shetty, V., & Kapoor, B. (2024). The Role of Participatory Governance in Strengthening Community Health Systems. *International Journal of SDG's Prospects and Breakthroughs*, 10-12.

[37] Spencer, R., Thabtah, F., Abdelhamid, N., & Thompson, M. (2020). Exploring feature selection and classification methods for predicting heart disease. *Digital health*, *6*, 2055207620914777. https://doi.org/10.1177/2055207620914777

[38] Villasante, A., & zaib Rasheed, A. (2024). Unlocking the secrets of ichthyology with artificial intelligence: A taxonomic revolution. *FishTaxa-Journal of Fish Taxonomy*, *31*, 1-10.

Authors Biography

Sofiene Mansouri is an Associate Professor at Department of Biomedical Technology, Prince Sattam bin Abdelaziz University, Saudi Arabia. He obtained his PhD from AL-Manar University, Tunis (Tunisia) in Medical Electronics in 2011. His current research interests include Biomedical Engineering and Devices, Instrumentation; Signal Processing, Bioimpedance and AI in medical diagnosis. He is the author of more than 60 scientific papers.

Amira Ouerhani received the M.Sc. degree in biomedical engineering and a Ph.D. in biophysics, medical physics, and medical imaging from the University of Tunis ElManar. Her research work focuses on radiation biology and medical image processing with a focus on lung disease, applying deep learning and machine learning techniques. She is a member of Radiation Protection and Medical Imaging research group in Biophysics and Medical Technologies Research Laboratory (LR13ES07) in Tunis, Tunisia. She is a member of the Tunisian Association of Biomedical Engineering.

Yousef Alharbi received his B.S. degree in Biomedical Technology from the University of King Saud, Riyadh, Saudi Arabia in 2009. He went on to complete his M.S. and Ph.D. degrees in Biomedical Engineering from the University of New South Wales, Sydney, Australia, in 2013 and 2019, respectively. Currently, he works as an Assistant Professor at the Department of Biomedical Technology, College Applied of Medical Sciences, Prince Sattambin Abdulaziz University. Yousef Alharbi's research interests are broad and diverse, with a focus on several areas, including bioimpedance, artificial intelligence, image processing, computational modeling, and simulation. Through his work, he aims to contribute to the development of new technologies and techniques that can be used to improve healthcare outcomes for patients.

Abdulrahman Alqahtani is an Assistant Professor at the Department of Medical Equipment Technology, College of Applied, Medical Science, Majmaah University and the Department of Biomedical Technology, College Applied of Medical Sciences, Prince Sattam bin Abdulaziz University. He received his B.S. degree in Biomedical Technology from the University of King Saud, Riyadh, Saudi Arabia in 2006. He received his M.S. and Ph.D. degrees in Biomedical Engineering from the University of New South Wales, Sydney, Australia, in 2013 and 2019, respectively. His research interests include computational neuroscience, neuroprosthetics, neural engineering, medical microwave imaging, bioimpedance and artificial intelligence.