

Comparative Evaluation of XGBoost and Support Vector Machine Algorithms for Heart Disease Prediction Using Healthcare Datasets

Oluwasogo Adekunle Okunade

Department of Computer Science, National Open University of Nigeria, Abuja, Nigeria

Corresponding Author's Email: aokunade@noun.edu.ng

Received on: October, 2022

Revised and Accepted on: November, 2022

Published on: December, 2022

ABSTRACT

This study examines the efficacy of advanced machine learning algorithms—Support Vector Machines (SVM) and Extreme Gradient Boosting (XGBoost)—in predicting heart disease using the IEEE DataPort Heart Disease Dataset. The analytical approach included rigorous data preprocessing, such as handling missing values and normalising features, followed by correlation-based feature selection to retain variables strongly associated with heart disease. Both models underwent extensive hyperparameter optimisation using GridSearchCV to enhance predictive performance. Evaluation metrics encompassed accuracy, precision, recall, F1 score, and AUC-ROC. SVM achieved notable results, including an accuracy of 0.8613 and an AUC-ROC of 0.9092, while XGBoost outperformed it across all metrics, with an accuracy of 0.8950 and an AUC-ROC of 0.9526. These findings underscore XGBoost's superior robustness and classification ability. The study highlights the potential of these models to advance diagnostic precision in cardiology, enabling earlier intervention and personalised treatment strategies for improved patient outcomes.

Keywords: Cardiovascular risk, Clinical, Demographic, Lifestyle factors, Symptomatology

1.0 INTRODUCTION

Heart disease persists as a global health crisis, contributing substantially to both morbidity and mortality worldwide. Despite ongoing progress in medical science and technology, the accurate prediction and diagnosis of cardiac conditions remain challenging. Conventional diagnostic approaches—often dependent on invasive procedures, symptomatic evaluation, and patients' medical histories—can fall short of providing timely and reliable assessments of cardiovascular health. This underscores an urgent need for more precise and efficient tools to identify individuals at risk and to facilitate early intervention.

The rise of machine learning (ML) presents a paradigm-shifting opportunity in healthcare, enabling data-driven approaches to clinical decision-making. ML algorithms excel at analysing complex and large-scale patient datasets—including demographic variables, clinical histories, and diagnostic metrics—to uncover subtle patterns and predict disease susceptibility. This capability holds the potential to revolutionise cardiovascular medicine by promoting proactive, individualised treatment strategies.

This research focuses on two prominent ML models: Support Vector Machine (SVM) and Extreme Gradient Boosting (XGBoost). SVM is a powerful classifier particularly suited for binary classification problems, as it determines optimal decision boundaries by maximising the margin between data classes—making it ideal for

differentiating individuals based on cardiovascular risk. XGBoost, known for its ensemble learning approach, builds upon gradient boosting techniques by iteratively correcting prediction errors, thereby enhancing classification performance. Prior studies have demonstrated XGBoost's efficacy in diagnostic systems, highlighting its superior accuracy and reliability in clinical applications.

By integrating the strengths of both SVM and XGBoost, this study aims to improve the robustness and interpretability of heart disease prediction models. The combined approach is expected to better handle complex data distributions and identify critical predictive features, thereby contributing to more resilient diagnostic frameworks. Ultimately, this work seeks to provide actionable insights for clinicians through effective predictive analytics, enabling earlier interventions and tailored patient care. The objective of this study is to develop and compare the performance of SVM and XGBoost models applied to a healthcare dataset for heart disease prediction. Specific goals include conducting a comprehensive literature review, designing a data preprocessing pipeline, implementing both algorithms, and evaluating model performance using metrics such as accuracy, precision, recall, and F1-score.

The methodology of this research encompasses the acquisition and preprocessing of heterogeneous healthcare data, including demographic attributes, clinical indicators (for instance, blood pressure,

cholesterol levels), and lifestyle variables (for instance, smoking habits, physical activity). The study emphasises feature selection and hyperparameter tuning to optimise the performance of SVM and XGBoost models. Evaluative metrics include accuracy and AUC-ROC, as well as comparisons of model interpretability, computational efficiency, and diagnostic precision.

This research also acknowledges limitations, such as the quality and completeness of the available data, which may affect model generalizability and robustness, with missing values or population biases potentially introducing uncertainty. Moreover, given the multifactorial nature of heart disease, variables such as genetic predisposition, environmental exposures, and socio-economic factors may not be fully captured, affecting the comprehensiveness of risk assessments. Finally, the dynamic landscape of clinical data demands ongoing model validation and adaptability to maintain predictive relevance across diverse healthcare contexts.

2.0 LITERATURE REVIEW

A comprehensive synthesis of current research on heart disease prediction using machine learning techniques, with a particular focus on Support Vector Machine (SVM) and Extreme Gradient Boosting (XGBoost), was provided. It examines various studies that leverage demographic, clinical, and lifestyle factors to assess cardiovascular risk, and it identifies key trends in model development, validation, and deployment strategies. Recent advancements in optimisation methods and feature selection are also discussed in light of their impact on predictive accuracy, computational efficiency, and model interpretability. By critically evaluating the strengths and limitations of SVM and XGBoost across diverse research contexts, this section lays the groundwork for the study's methodological approach and underscores the relevance of machine learning in clinical predictive analytics.

2.1 Conceptual Framework

Several studies have demonstrated the efficacy of machine learning in improving the accuracy of heart disease prediction. A machine learning-based web application translates complex models into practical diagnostic tools for clinicians and patients, demonstrating the real-world applicability of adaptive learning systems. Srinivas and Katarya (2022) introduced the OPTUNA-based "hyOPTXg" framework for hyperparameter tuning in XGBoost. Their work reduced overfitting and improved model generalisation, reinforcing the importance of algorithmic fine-tuning. Yang and Guan (2022) integrated SMOTE with XGBoost to address class imbalance and improve accuracy. Their method enhanced minority class

representation, resulting in more equitable and precise predictions. Anbuselvan (2020) analysed models such as SVMs, Neural Networks, and k-NN. The study highlighted the trade-off between accuracy and interpretability, suggesting SVMs as ideal for clinical scenarios requiring transparent decision processes. Patro et al. (2021) presented a novel optimisation algorithm to improve feature selection and model tuning in supervised learning. Their framework increased accuracy by refining predictive features and streamlining model parameters. Wang et al. (2020) applied XGBoost to type 2 diabetes prediction, offering transferable insights into feature importance and algorithm adaptability for heart disease modelling. Nalluri et al. (2020) utilised clustering and association rule mining to extract patterns from chronic heart disease datasets, demonstrating the potential of unsupervised methods in enriching predictive frameworks. Mienye and Sun (2022) employed Principal Component Analysis (PCA) for dimensionality reduction and Genetic Algorithms for feature selection in a hybrid XGBoost-GA model, achieving balanced performance and clinical applicability. Nagavelli et al. (2022) emphasised the impact of preprocessing and algorithm selection, concluding that tailored model optimisation substantially improves predictive accuracy. Jadevi et al. (2021) combined XGBoost with the Black Hole Optimisation Algorithm to improve feature selection and classification, illustrating the power of metaheuristic approaches in medical diagnostics. Budholiya et al., (2020). developed an improved XGBoost-based model incorporating sophisticated preprocessing and tuning strategies, validating the model's robustness against complex medical data.

2.2 Empirical Review

Empirical studies offer valuable insights into the application of machine learning for heart disease prediction, especially regarding feature engineering, algorithm performance, and system implementation. Ngong & Baykan (2021) explored advanced feature extraction methods and concluded that effective preprocessing significantly enhances model accuracy, highlighting the central role of feature engineering. Sang et al. (2020) demonstrated the efficacy of well-tuned Support Vector Machine (SVM) models in predicting survival among patients with heart failure, emphasising the influence of kernel selection and parameter optimisation.

Deployment in Clinical Practice: Many models remain theoretical or lab-tested, with limited adoption in operational healthcare settings.

This paper helps fill these gaps by offering a direct comparison between XGBoost and SVM, evaluating not

only accuracy, precision, recall, and F1-score but also AUC-ROC and feature relevance. By highlighting key predictors and assessing algorithmic trade-offs, the research contributes to the development of interpretable and scalable models for real-world applications. Furthermore, this work advocates for the future inclusion of dynamic patient data and ensemble strategies to foster personalised, proactive cardiac care.

3.0 METHODOLOGY

This study utilised Support Vector Machine (SVM) and XGBoost algorithms to perform predictive analysis on the IEEE DataPort Heart Disease Dataset. The dataset comprises 1,191 samples and includes a wide array of clinical and demographic variables, thereby enabling the identification of complex patterns associated with the manifestation of heart disease.

3.1 Dataset Description

The IEEE DataPort Heart Disease Dataset ([link](#)) serves as a crucial resource for advancing predictive modelling in cardiology. It contains a diverse set of patient-level features spanning clinical and demographic domains.

which are instrumental in training robust machine learning models for heart disease prediction.

Key variables include, Demographic Data, c (Resting blood pressure, Serum cholesterol levels, Fasting blood sugar status), Clinical Measurements (resting blood pressure, Serum cholesterol levels, Fasting blood sugar status), Diagnostic Features (resting electrocardiographic (ECG) results, maximum heart rate during exercise, presence of exercise-induced angina), Specialized Markers (ST depression induced by exercise, slope of the ST segment during peak exercise, number of major blood vessels visualized through fluoroscopy, and thalassemia status). Each patient record in the dataset corresponds to a binary outcome indicating the presence or absence of heart disease. The combination of continuous and categorical variables enables nuanced analyses that go beyond conventional medical assessments. The dataset structure allows machine learning models to capture intricate, non-linear relationships and interactions among features. Algorithms like SVM and XGBoost are particularly well-suited for this domain due to their ability to manage high-dimensional data and complex decision boundaries. Table 1 outlines the dataset’s attributes and structure.

Table 1: Dataset Attributes and Value Encoding

Name	Description	Values
age	Age of the patient in years	Integer (e.g., 29, 40, 60)
sex	Gender of the patient	1 = Male, 0 = Female
chest_pain_type	Type of chest pain experienced	1 = Typical angina, 2 = Atypical angina, 3 = Non-anginal pain, 4 = Asymptomatic
resting_bp_s	Resting blood pressure (in mm Hg)	Integer (e.g., 120, 140)
cholesterol	Serum cholesterol level (in mg/dL)	Integer (e.g., 200, 250)
fasting_blood_sugar	Fasting blood sugar >120 mg/dL	1 = True (>120 mg/dL), 0 = False (≤ 120 mg/dL)
resting_ecg	Resting electrocardiographic results	0 = Normal, 1 = ST-T abnormality, 2 = LV hypertrophy by Estes' criteria
max_heart_rate	Maximum heart rate achieved	Integer (e.g., 150, 170)
exercise_angina	Exercise-induced angina	1 = Yes, 0 = No
oldpeak	ST depression induced by exercise relative to rest	Float (e.g., 0.0, 1.5)
ST_slope	Slope of the ST segment during peak exercise	1 = Upsloping, 2 = Flat, 3 = Downsloping
target	Presence of heart disease	1 = Disease present, 0 = No disease

3.2 Proposed System Design and Workflow

The system architecture integrates two high-performance machine learning algorithms—Support Vector Machine (SVM) and XGBoost—to classify the presence of heart disease using patient-level clinical data. Support Vector Machine (SVM): SVM constructs optimal decision boundaries (hyperplanes) for binary classification. It is effective for both linear and non-linear relationships using kernel functions such as linear, polynomial, and Radial Basis Function (RBF). XGBoost: An ensemble learning algorithm based on gradient boosting, XGBoost iteratively builds decision trees by minimising prediction errors of previous iterations, thus improving accuracy

and resilience. The overall workflow consists of the following steps: Data Preparation involves loading the dataset, handling missing values by dropping rows containing NaN entries, then normalising feature values using StandardScaler to ensure consistent feature impact, and finally separating features (X) from the target variable (y). Feature Selection involved correlation analysis to identify variables with Pearson coefficients > 0.2 or < -0.2 relative to the target variable, and to retain only the most predictive features for model training. Data splitting was performed using `train_test_split`, allocating 80% for training and 20% for testing. Model Training using SVM involved defining a parameter grid for C,

kernel, and gamma, then employing GridSearchCV with cross-validation to identify optimal hyperparameters. The final step involved fitting the best SVM model to the training data. Model Training using XGBoost involved configuring parameters such as `n_estimators`, `max_depth`, and `learning_rate`, then applying GridSearchCV to optimise XGBoost and train the final model on the training set. Model Evaluation and Comparison were performed for both models based on predicted outcomes on the test set, calculating Accuracy, Precision, Recall, F1-Score, and AUC-ROC for both models, and extracting feature importance scores from the XGBoost model

Finally, performance metrics for SVM and XGBoost were visualised side by side, and plots of key clinical predictors were created to interpret the model rationale and diagnostic relevance. This helped to determine comparative strengths based on clinical interpretability and statistical robustness. This modular and replicable workflow ensures systematic evaluation while enabling direct comparison of algorithmic efficiency in heart disease classification.

3.3 Evaluation Metrics

The following metrics are used for the comparative analysis of SVM and XGBoost: Accuracy measures the proportion of correctly classified instances among all samples. While informative, accuracy alone can be misleading in imbalanced datasets. Precision describes the ratio of true positives to the total predicted positives. Precision is vital in medical diagnostics, where false positives can lead to unnecessary treatments. Recall was measured as the sensitivity or true positive rate, which measures the proportion of actual positives correctly identified. High recall reduces the number of undetected cases of heart disease. F1-Score measured the harmonic mean of precision and recall, offering a balanced assessment when class distribution is uneven. It is especially useful when the costs of false negatives and false positives must be jointly minimised, as in high-stakes domains such as medical diagnostics. The Area Under the Receiver Operating Characteristic Curve (AUC-ROC) quantified each model's ability to distinguish between classes at various threshold settings. A higher AUC indicates superior model discrimination. For granular error analysis and model calibration, a Confusion Matrix was used to provide a tabular summary of prediction outcomes, detailing true positives, false positives, false negatives, and true negatives. A 5-fold Cross-Validation Score was used to assess the generalizability of each model across different data

subsets, thereby reducing variance from train-test splits and ensuring a robust evaluation.

3.4 Feature Importance Interpretation

Understanding which clinical variables most influence model predictions is crucial for translating machine learning outcomes into actionable insights for cardiovascular diagnosis and management. Feature importance interpretation also contributes to Explainable AI (XAI) that improves audit of automated decision making and improves trustworthiness in AI models used in healthcare.

3.4.1 XGBoost Feature Importance

XGBoost inherently supports feature importance evaluation using methods like; Weight (frequency with which a feature is used in tree splitting, Gain (contribution of a feature to model improvement through splits and Cover (the number of samples impacted by splits using the feature). In this study, Gain is prioritised as the metric for feature ranking due to its direct relationship with predictive performance. A feature importance plot is generated post-training to visualise the top predictive indicators—such as age, cholesterol levels, resting ECG values, and ST depression (oldpeak).

3.4.2 SVM Feature Ranking

SVM models do not inherently compute feature importance unless combined with linear kernels. For this study, Recursive Feature Elimination (RFE) is employed. This method iteratively removes the least significant features; the model is retrained at each step, and the final rankings indicate the features most valuable for classification. This approach yields a refined understanding of which inputs best differentiate between patients with and without heart disease.

4.0 RESULT AND DISCUSSION

4.1 Correlation Analysis for Dimensionality Reduction

Pearson correlation coefficients were computed for all features relative to the target.

Features with absolute correlation values greater than 0.2 or less than -0.2 were selected, indicating meaningful predictive relationships.

This reduced dimensionality while retaining clinically relevant variables, thereby improving model efficiency and generalisation.

This process is illustrated in Figure 1 of the original source, which shows the impact of data preparation and feature selection.

```
# Feature Selection
correlation = data.corr()['target'].sort_values()
print("Feature Correlation with Target:\n", correlation)

Feature Correlation with Target:
max heart rate      -0.413278
cholesterol         -0.198366
resting ecg         0.073059
resting bp s        0.121415
fasting blood sugar  0.216695
age                 0.262029
sex                 0.311267
oldpeak             0.398385
chest pain type     0.460127
exercise angina     0.481467
ST slope            0.505608
target              1.000000
Name: target, dtype: float64
```

Figure 1. Data Preparation and Feature Selection

4.2 Hyperparameter Tuning using GridSearchCV

For both Support Vector Machine (SVM) and Extreme Gradient Boosting (XGBoost) models, hyperparameters were optimised using GridSearchCV with k-fold cross-

validation (typically set to 5 or 10 folds). This method exhaustively explores combinations of model parameters to identify the optimal configuration Table 2.

Table 2. Parameter Grids

Model	Grids
SVM	i. C: Regularization parameter controlling margin size
	ii. kernel: Type of kernel function (e.g., linear, RBF)
	iii. gamma: Kernel coefficient for non-linear kernels
XGBoost	i. n_estimators: Number of trees in the ensemble
	ii. max_depth: Maximum depth of individual trees
	iii. learning_rate: Step size shrinkage for boosting

4.3 Model Training

Each model was trained on the designated training data using the best parameters identified from GridSearchCV. Performance metrics from cross-validation were recorded for comparative analysis and generalisation insights, as shown in Figure 2.

```
# Grid search for SVM
svm = SVC(probability=True)
grid_search_svm = GridSearchCV(svm, param_grid_svm, cv=5, scoring='accuracy')
grid_search_svm.fit(X_train, y_train)
best_svm = grid_search_svm.best_estimator_

# MODEL OVERVIEW: XGBoost Implementation
# Define the parameter grid for XGBoost
param_grid_xgb = {
    'n_estimators': [100, 200],
    'max_depth': [3, 5, 7],
    'learning_rate': [0.01, 0.1],
}

# Grid search for XGBoost
xgb = XGBClassifier(use_label_encoder=False, eval_metric='logloss')
grid_search_xgb = GridSearchCV(xgb, param_grid_xgb, cv=5, scoring='accuracy')
grid_search_xgb.fit(X_train, y_train)
best_xgb = grid_search_xgb.best_estimator_
```

Figure 2: Data, Splitting and Model, Training

4.4 SVM Model Evaluation Results

The Support Vector Machine (SVM) model demonstrated strong predictive performance for heart disease classification on the test dataset:

- **Accuracy: 0.8613**, correctly classified **86.13%** of instances.
- **Precision: 0.8769**, when heart disease was predicted, the diagnosis was correct **87.69%** of the time, which is an important factor for minimising false alarms in medical settings.
- **Recall (Sensitivity): 0.8702**, detected **87.02%** of actual heart disease cases, ensuring that few true cases go unnoticed.

- **F1 Score: 0.8736**, balances precision and recall, indicating overall reliability.
- **AUC-ROC: 0.9092**, reflects the model's excellent ability to distinguish between the presence and absence of heart disease across different thresholds. A value close to **1.0** indicates high discriminative power.

These results reflect the model's excellent ability to distinguish between the presence and absence of heart disease across various thresholds, with a value close to 1 indicating high discriminative power, as shown in Table 3.

Table 3: Support Vector Machine Classification Report

Metric	Value
Accuracy	0.8613
Precision	0.8769
Recall	0.8702
F1 Score	0.8736
AUC-ROC	0.9092

4.4.1 SVM Confusion Matrix

The SVM confusion matrix for heart disease prediction reveals the model's performance breakdown:

- **True Negatives (TN): 91**, the model correctly identified the absence of heart disease.
- **False Positives (FP): 16**, the model incorrectly predicted heart disease when it wasn't present, which could lead to unnecessary medical interventions.
- **False Negatives (FN): 17**, actual heart disease cases that the model failed to detect, potentially delaying necessary treatment.

- **True Positives (TP): 114**, correct predictions of heart disease presence.

Overall, the matrix demonstrates strong model performance in distinguishing cases, but highlights areas for improvement—particularly in reducing false positives and false negatives to further enhance diagnostic accuracy. This matrix suggests that while the model performs well in identifying both the presence and absence of heart disease, there is room to reduce false positives and false negatives to enhance diagnostic accuracy, as shown in Figure 3.

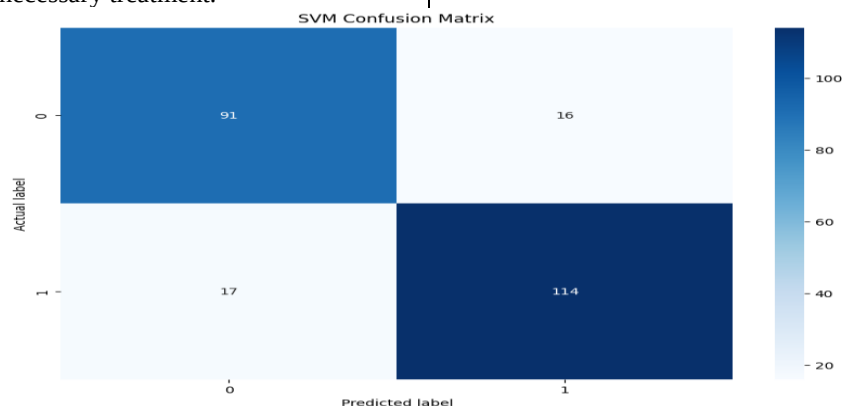


Figure 3: Support Vector Matrix (SVM) Confusion Matrix

4.4.2 XGBoost Model Evaluation

The **XGBoost model** demonstrated robust performance in predicting heart disease from the test dataset.

- **Accuracy: 0.8950**, correctly classified **89.50%** of instances.
- **Precision: 0.9141**, when heart disease was predicted, the diagnosis was correct **91.41%** of the time—crucial for minimising false alarms.
- **Recall (Sensitivity): 0.8931**, identified **89.31%** of actual heart disease cases, ensuring reliable detection.

- **F1 Score: 0.9035**, balanced trade-off between precision and recall, reflecting consistent reliability.
- **AUC-ROC: 0.9526**, **indicating** excellent model ability to distinguish between patients with and without heart disease. A value close to 1 indicates high discriminative power across thresholds in Table 4.

Table 4: XGBoost Classification Report

Metric	Value
Accuracy	0.8950
Precision	0.9141
Recall	0.8931
F1 Score	0.9035
AUC-ROC	0.9526

4.3.3 XGBoost Confusion Matrix

The XGBoost confusion matrix provides insights into prediction breakdown:

- **True Negatives (TN): 96**, correctly identified patients without heart disease.
- **False Positives (FP): 11**, incorrectly predicted heart disease when absent, potentially leading to unnecessary medical concern.

- **False Negatives (FN): 14**, missed actual cases of heart disease, which may delay essential treatment.
- **True Positives (TP): 117**, accurately identified the presence of heart disease.

This matrix indicates that XGBoost has a strong performance in both identifying heart disease and ruling it out, yet there's still potential for enhancing the model to further minimize false positives and false negatives as shown in Figure

4.

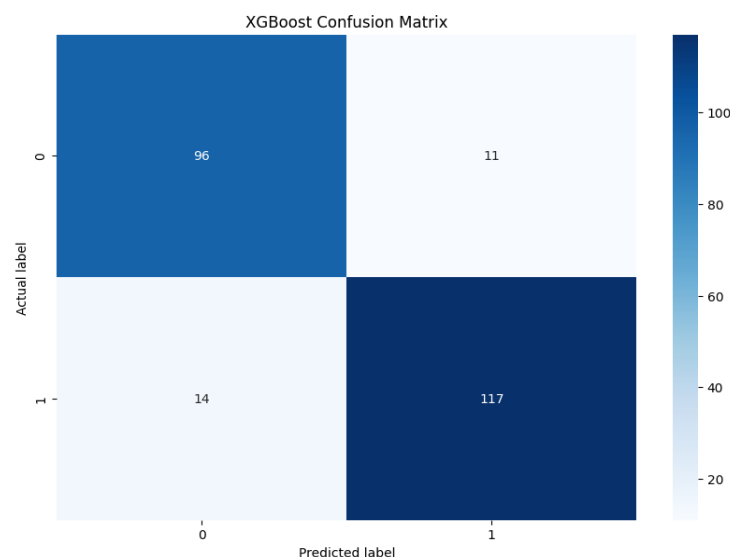


Figure 4: XGBoost Confusion Matrix

4.4.4 Model Comparison

Comparative analysis of the **Support Vector Machine (SVM)** and **XGBoost** models for heart disease prediction reveals that while both demonstrate commendable performance, **XGBoost consistently outperforms SVM** across all evaluation metrics. Most notably, XGBoost achieved an **AUC-ROC score of 0.9526**, significantly surpassing SVM's 0.9092. This demonstrates a greater ability to discriminate between patients with and without heart disease across different threshold settings—a critical attribute in medical diagnostics, where decision thresholds must be carefully calibrated for treatment. Higher precision and recall rates mean that XGBoost not only identifies a greater proportion of true heart disease

cases but also reduces false alarms—minimising unnecessary treatment and alleviating patient anxiety. Additionally, its robust AUC-ROC performance ensures adaptability across various clinical risk thresholds. Further advantages include: **Built-in handling of missing values, Feature importance ranking for clinical interpretability and scalability for real-time or batch prediction tasks.** Together, these attributes position XGBoost as a powerful tool for improving diagnostic precision and enhancing patient outcomes by enabling more precise, reliable diagnostics, as shown in Table 5. Also highlighted below are the model comparison showing various evaluation and performance metrics (Figures 5 to 9)

Table 5 Model Comparison

Metric	SVM	XGBoost
Accuracy	0.8613	0.8950
Precision	0.8769	0.9141
Recall	0.8702	0.8931
F1 Score	0.8736	0.9035
AUC-ROC	0.9092	0.9526

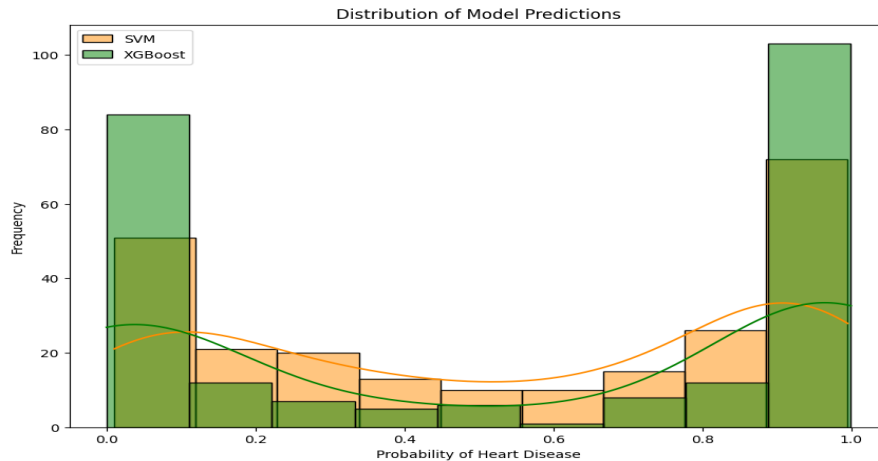


Figure 5: Distribution of Model Predictions

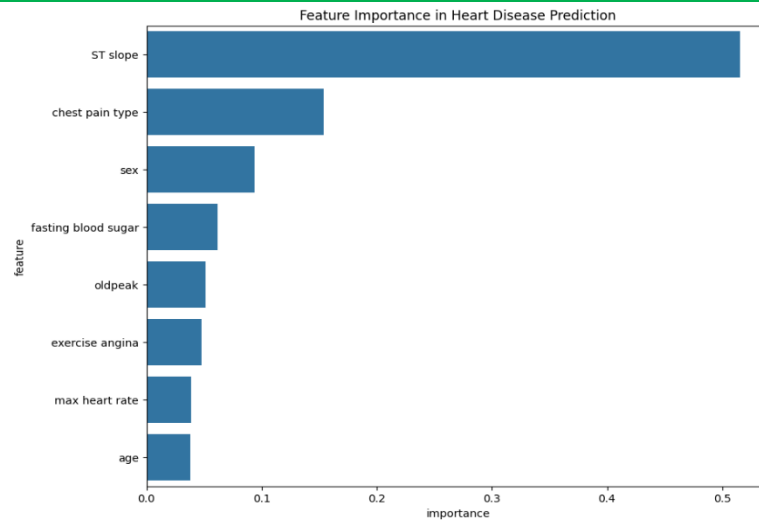


Figure 6: Feature Importance in Heart Disease Prediction

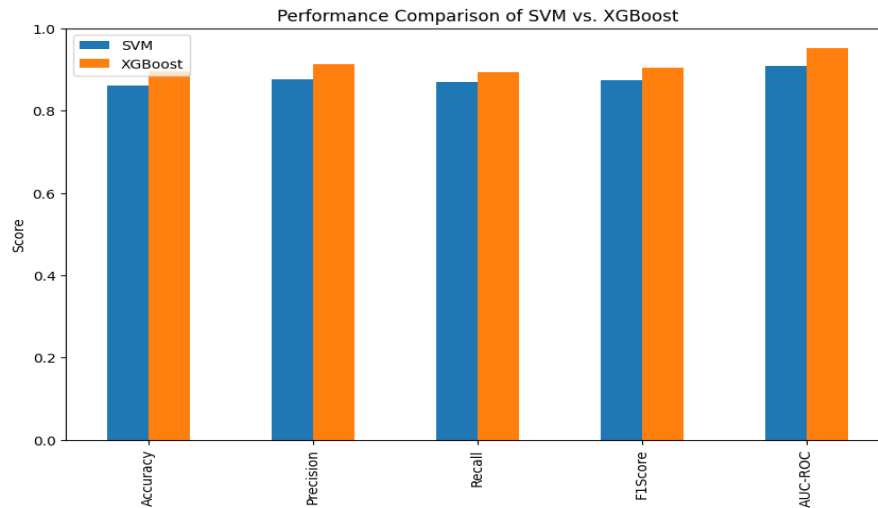


Figure 7: Performance comparison of SVM and XGBoost

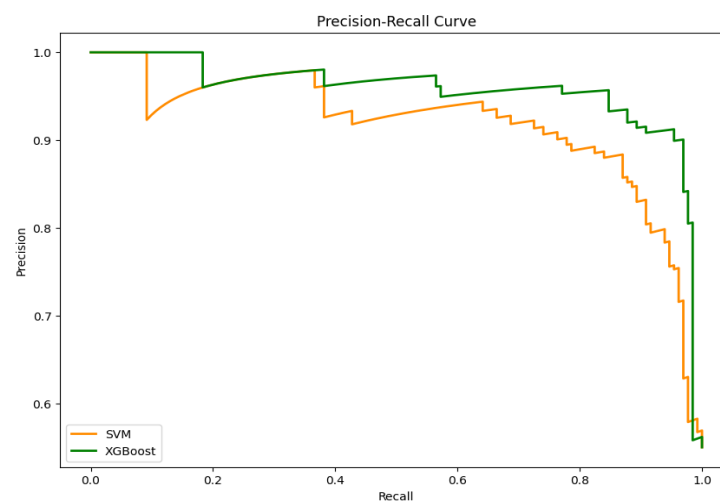


Figure 8: Precision-Recall Curve

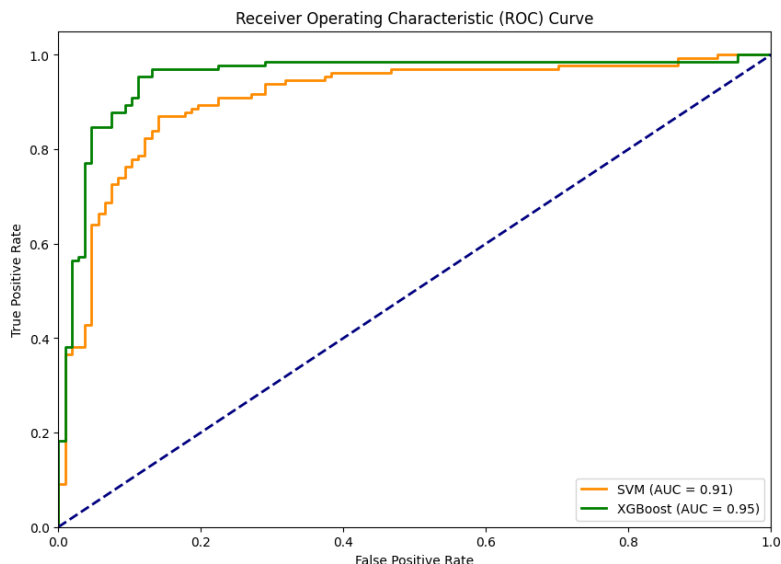


Figure 9: ROC Curve

5.0 CONCLUSION

The heart disease prediction system described in this paper involved a comprehensive pipeline beginning with loading the IEEE DataPort Heart Disease Dataset into a pandas DataFrame. Data preprocessing steps included removing entries with missing values and normalising feature values using *StandardScaler*. Feature selection was performed using correlation analysis, retaining variables with strong coefficients for the target outcome. Two machine learning models—**Support Vector Machine (SVM)** and **XGBoost**—were trained, with hyperparameter tuning performed through *GridSearchCV*. **SVM** achieved an accuracy of 0.8613, precision of 0.8769, recall of 0.8702, F1 score of 0.8736, and an AUC-ROC of 0.9092. **XGBoost** outperformed SVM with an accuracy of 0.8950, precision of 0.9141, recall of 0.8931, F1 score of 0.9035, and an AUC-ROC of 0.9526.

Confusion matrices confirmed strong predictive performance for both models, with XGBoost exhibiting fewer false positives and negatives. The detailed comparison clearly established XGBoost as the superior model, given its higher scores across all key metrics and its ability to capture complex relationships in the data. These characteristics position XGBoost as a viable candidate for practical application in heart disease diagnostics. This study demonstrates the efficacy of machine learning models—particularly SVM and XGBoost—in predicting heart disease. Among them, **XGBoost consistently emerged as the stronger performer**, excelling across all evaluated metrics. Its ability to handle complex, nonlinear medical data and deliver high precision, recall, and AUC-ROC scores

makes it a powerful diagnostic tool for early disease detection. The findings underscore the importance of rigorous data preprocessing, strategic model tuning, and the use of ensemble techniques in clinical machine learning applications. To capitalise on XGBoost's capabilities, future work should focus on integrating it with Electronic Health Records (EHRs) by building intuitive interfaces that enable seamless incorporation into EHR systems and enable real-time risk assessments. Recommendations for further research in this area may include hybrid and ensemble methods, such as combining XGBoost with other models or deep learning architectures to enhance predictive power. Dataset expansion to incorporate broader and more diverse patient demographics to improve model generalizability and minimise bias. Sophisticated Feature engineering, which employs advanced techniques, including time-series analysis, to extract hidden patterns crucial for chronic disease modelling. Explainable AI (XAI) methods can make model decisions transparent to clinicians, improving trust, adoption, and the actionable interpretation of those decisions in clinical workflows. Ethical considerations address fairness, data privacy, and equitable healthcare delivery across populations using AI technologies. Cross-disciplinary collaboration can strengthen partnerships between data scientists and healthcare practitioners, aligning model development with real-world needs and continuously refining outcomes.

References

- Alqahtani, A., Almuqhim, F., Alsubhi, N., Alyami, M., Alshahrani, T., & Alhumam, K. (2022). Ensemble learning for cardiovascular disease detection.
- Anbuselvan, P. (2020). Heart disease prediction using machine learning techniques. *Int. J. Eng. Res. Technol*, 9, 515-518.
- Bharti, R., Khamparia, A., Shabaz, M., Dhiman, G., Pande, S., & Singh, P. (2021). Prediction of heart disease using a combination of machine learning and deep learning. *Computational intelligence and neuroscience*, 2021.
- Budholiya, K., Shrivastava, S. K. and Sharma, V. (2020). An optimised XGBoost-based diagnostic system for effective prediction of heart disease. *Journal of King Saud University – Computer and Information Sciences* 34 (2022) 4514–4523
- Butarbutar, S. Y., Harahap, M., Sitompul, H., Pasaribu, A. L., & Bu'ulolo, M. G. (2022). Classification of electrocardiogram (ECG) waveforms associated with cardiac illness using the XGBoost technique.
- ÇELİK, G. (2022). Heart failure prediction with the proposed DNN-based method.
- Hasan, O. S., & Saleh, I. A. (2021). DEVELOPMENT OF HEART ATTACK PREDICTION MODEL BASED ON ENSEMBLE LEARNING. *Eastern-European Journal of Enterprise Technologies*, 112.
- Jadevi, D. H., Kulkarni, V., & Varthini, A. M. (2021). Feature selection with black hole optimization for warning heart disease.
- Karthiga, A. S., & Mary, D. M. S. (2022). Cardiovascular Disease Detection Using Machine Learning Techniques. *Int. J. Mech. Eng.*, 7.
- Lv, H., Zhang, J., Wu, W., Jiang, R., Ma, C., Liu, X., ... & Ding, D. (2021). Machine learning-Driven models to predict prognostic outcomes in patients hospitalised with heart failure, using electronic health records (EHRs).
- Mienye, I. D and Sun, Y. (2022). Survey of Ensemble Learning: Concepts, Algorithms, Applications and Prospects. *IEEE Access, Open Access Journal*. Vol 10, pp 99129 – 99149. DOI: 10.1109/ACCESS.2022.3207287
- Nagavelli, A. K., Samanta, S., & Chakborty, D. (2022). Machine learning-based models for heart disease detection.
- Nalluri, S., Bheemanathi, K. V., Bhupal, A. S., & Gutta, S. M. (2020). Data mining techniques for predicting chronic heart disease.
- Ngong, I. C., & Baykan, N. A. (2021). Various feature extraction methods for predicting the prevalence of heart disease.
- Patro, D., Mish, S., & Singh, V. K. (2021). Novel optimisation algorithm for the prediction of heart disease in the context of supervised learning.
- Ptama, M. A. W. R. (2021). Influence of up-sampling methods in predicting disease types using a combination of Natural Language Processing (NLP), Naive Bayes, XGBoost, and SVM.
- Sang, X., Li, Y., Zhao, Y., Li, Q., & Wang, Y. (2020). Support Vector Machine (SVM) algorithms for predicting patient survival in heart failure.
- Shorewala, V. (2021). Early detection of coronary heart disease employing ensemble techniques.
- Simegn, G. L., Gebeyehu, T. H., & Degu, B. A. (2022). Computer-aided decision support system for diagnosing heart diseases.
- Singh, V., & Aro, A. (2022). Performance evaluation to predict cardiac disease using machine learning algorithms.
- Srinivas, P., & Katarya, R. (2022). The XGBoost algorithm is designed to predict cardiovascular disease using OPTUNA in Srinivas and Katarya's (2022) hyper-parameter optimisation framework, hyOPTXg.
- Valarmathi, R., & Sheela, T. (2021). Prediction of heart disease by utilising hyperparameters optimisation (HPO) tuning.
- Wang, H., Zhao, X., Xu, X., & Li, Y. (2020). XGBoost model for type 2 diabetes prediction.
- Yang, J., & Guan, J. (2022). SMOTE-XGBoost algorithm in conjunction with feature optimisation for a heart disease prediction model.

