Heart Disease Risk Prediction Using Machine Learning

Yñikko Arzee Neo D. Aguas
School of Information Technology
Mapua University
Makati, Philippines
yandaguas@mymail.mapua.edu.ph

Carl Emmanuel M. Macabales
School of Information Technology
Mapua University
Makati, Philippines
cemmacabales@mymail.mapua.edu.ph

Nathaniel James C. Carrillo
School of Information Technology
Mapua University
Makati, Philippines
njccarrillo@mymail.mapua.edu.ph

Rolando R. Magat III
School of Information Technology
Mapua University
Makati, Philippines
rrmagat@mymail.mapua.edu.ph

Keywords - Machine Learning, Supervised Learning, Heart Disease Classification, Predictive Modeling, Feature Selection, Early Detection

Abstract

This research explores the effectiveness of supervised machine learning algorithms in predicting the presence of heart disease using clinical and demographic patient data. Utilizing a refined subset of 300 samples from a widely recognized heart disease dataset, three models: Logistic Regression, Support Vector Classifier, and Random Forest Classifier were developed and evaluated. The preprocessing pipeline involved rigorous data cleaning, and feature encoding to ensure model reliability. Cross-validation, hyperparameter tuning, learning curve analysis, and confusion matrix diagnostics were employed to assess each model's performance. The Support Vector Classifier achieved the highest test accuracy (0.88) post-augmentation, while Logistic Regression demonstrated consistent generalization performance even with a smaller dataset. Analytical comparisons with related studies revealed that proper data preprocessing, including duplicate removal and outlier handling, plays a critical role in avoiding inflated accuracy due to potential data leakage. The results affirm the viability of machine learning in enabling scalable, non-invasive early risk detection for cardiovascular disease.

Introduction

[1] Heart disease is a leading cause of death worldwide, responsible for approximately 17.9 million deaths each year according to the World Health Organization. Traditional methods of diagnosing heart disease often involve invasive procedures or extensive medical testing. *With the rise of data-driven approaches in healthcare, machine learning offers a promising alternative for early detection and risk assessment.*

In this paper, we explore the application of supervised machine learning algorithms to predict the presence of heart disease using patient data. *By leveraging key features such as age, sex, chest pain type, resting blood pressure, cholesterol, fasting blood sugar, maximum heart rate, ST depression, slope of the ST segment, and the number of major*

vessels, we aim to develop predictive models that can support healthcare professionals in decision-making.

Problem Statement

The primary challenge addressed in this study is the early prediction of heart disease using a minimal but relevant set of patient health metrics. The goal is to identify patients at risk in a non-invasive and computationally efficient manner. This problem is framed as a binary classification task: predicting whether or not a patient has heart disease.

Objectives

- ➔ To implement and compare three machine learning models: Logistic Regression, Random Forest Classifier,and Support Vector Classifier.
- ➔ To evaluate each model's performance using appropriate metrics such as accuracy, precision, recall, and F1-score.
- ➔ To analyze which features contribute most to heart disease prediction.
- ➔ To examine overfitting/underfitting behavior and discuss the models' generalization capabilities.

The dataset employed in this study consists of patient-level health data typically found in widely recognized heart disease datasets, such as the UCI Cleveland Heart Disease dataset. This dataset includes a range of clinical features that are commonly used by healthcare professionals to assess cardiovascular health and identify potential risk factors. *The recorded attributes encompass various physiological and demographic indicators including the patient's age and biological sex, which are known to influence cardiovascular risk. Additionally, the dataset captures chest pain type, a key symptom in cardiac diagnostics, as well as resting blood pressure and serum cholesterol levels, both critical markers of heart function and arterial health.*

Other included features are fasting blood sugar, which provides insight into metabolic health, and the maximum heart rate achieved during exercise, which reflects cardiovascular endurance and stress response. The dataset also records ST depression induced by exercise (oldpeak), the slope of the peak exercise ST segment, and the number of major vessels colored by fluoroscopy, each of which offers important details about the heart's electrical activity and structural condition.

The outcome variable in this dataset is binary in nature, indicating whether heart disease is present (1) or not (0). This classification setup enables the application of supervised machine learning techniques to predict the likelihood of heart disease based on the aforementioned features

Variable	Definition
age	Age of the patient
sex	Gender (1= male, 0 = female)
cp	Chest pain type (0-3)
trestbps	Resting blood pressure (in mm Hg)
chol	Serum cholesterol (in mg/dl)
fbs	Fasting blood sugar > 120mg/dl (1 = true, 0 = false)
restecg	Resting electrocardiographic results (0-2)
thalach	Maximum heart rate achieved
exang	Exercise-induced angina (1 = yes, 0 = no)
oldpeak	ST depression induced by exercise
slope	Slope of the peak exercise ST segment (0-2)
ca	Number of major vessels colored by fluoroscopy (0-3)
thal	Thalassemia (1 = normal, 2 = fixed defect, 3 = reversible defect)

REVIEW RELATED LITERATURE

A. Significance of Early Detection for Cardiovascular diseases

[2] Heart failure, a major but often overlooked cardiovascular condition, affects over 10 million people in Europe and causes more deaths in some countries than common cancers. *Despite*

advancements in cardiovascular disease (CVD) treatment, early diagnosis of heart failures remain a major challenge. Timely intervention through recommended treatments can significantly improve outcomes, yet many cases go undiagnosed until hospitalization, by which time the risk of death is higher. Two key strategies to address this are increasing awareness of the public and primary care professionals, and investing in community diagnostics like point-of-care biomarker testing. Raising public understanding and equipping primary care with diagnostic tolls can help detect heart failure earlier, reduce hospital admissions, and improve patient outcomes. [3] CVDs including ischemia, heart failure, myocardial infarction, stroke, aortic and peripheral artery disorders, arrhythmias, and valve diseases remain the leading global cause of death. *CVDs are not only widespread but also the most expensive medical condition, costing around $1 billion daily. Projections suggest that by 2035, 45% of U.S. adults may have CVD, with annual costs exceeding $1 trillion.* Although largely preventable, the rising trend emphasizes the need for effective scalable intervention.

B. Early Detection of Heart Diseases Using Machine Learning

[4] The use of Artificial Intelligence (AI) and Machine Learning (ML) in healthcare, particularly for cardiovascular disease (CVD), holds great promise but faces key challenges, *one major issue is data quality including scarcity, imbalance, inconsistency, and lack of standardization which limits model performance and generalizability.* Addressing these requires better data curation and integration practices. However, [5] Recent studies highlight the effectiveness of machine learning (ML) in predicting heart disease through various models, with *Support Vector Machines (SVM) achieving an accuracy of up to 91.51%. These results demonstrate the strong potential of ML-based predictive analytics*

in improving clinical decision-making and patient outcomes. By enabling early detection, ML empowers healthcare professionals to identify high-risk individuals, provide timely interventions, and support self-management strategies. Integrating ML into routine clinical practice promotes a shift toward preventive and personalized care. Researchers also emphasize the need to explore diverse models and optimization techniques to further enhance prediction accuracy and reliability in real-world settings.

C. Traditional ECG Interpretation versus ECG-based Machine Learning

[6] A systematic meta-analysis comparing machine learning models (such as SVM, neural networks, and random forests) with traditional risk scores (e.g., Framingham and ACC/AHA) f*ound that ML approaches provide superior discrimination in predicting atherosclerotic cardiovascular events, as evidenced by higher pooled C-statistics. However, the review emphasizes that most studies were limited by retrospective cohort designs, lack of calibration reporting, and minimal external validation.* Prospective, real-world implementation studies are needed to determine ML's clinical utility and impact on risk stratification and decision-making

---

METHODOLOGY

A. Dataset Description and Acquisition

This study utilized a Heart Disease dataset which contains clinical and demographic information from 1,025 patients who underwent comprehensive medical examinations. The dataset represents a compilation of records from multiple medical institutions, providing a diverse and representative sample for cardiovascular health analysis.

The dataset consists of 14 attributes, including 13 predictor variables and one target variable indicating the presence (1) or absence (0) of heart disease. The predictor variable encompass both demographic factors (age, sex) and clinical measurements (blood pressure, cholesterol levels, electrocardiographic results, and cardiac stress test parameters)

Table 1: Dataset Variables and Descriptions

Variable	Type	Description	Range/Values
age	Numerical	Patient age in years	Continuous
sex	Categorical	Gender (0=female, 1=male)	Binary
cp	Categorical	Chest pain type	0-3
trestbps	Numerical	Resting blood pressure (mm Hg)	Continuous
chol	Numerical	Serum cholesterol (mg/dl)	Continuous
fbs	Binary	Fasting blood sugar > 120 mg/dl	0-1
restecg	Categorical	Resting electrocardiographic results	0-2
thalach	Numerical	Maximum heart rate achieved	Continuous
exang	Binary	Exercise-induced angina	0-1
oldpeak	Numerical	ST depression induced by exercise	Continuous
slope	Categorical	Slope of peak exercise ST segment	0-2
ca	Numerical	Number of major vessels (0-3)	Discrete
thal	Categorical	Thalassemia type	1-3
target	Binary	Heart disease presence	0-1

B. Data Preprocessing Pipeline

1. Data Cleaning

The preprocessing pipeline implemented several critical steps to ensure data quality and model reliability:

Invalid Entry Removal: Rows containing invalid thalassemia values (thal = 0) were removed as they represented missing or erroneous data points. **Duplicate Detection**: The dataset was screened for duplicate entries using pandas' `drop_duplicates()` function. **Missing Value Validation**: Comprehensive checks were performed to ensure no null values remained in the dataset

1. Feature Engineering

Categorical Variable Encoding: One-hot encoding was applied to categorical variables (cp, restecg, slope, thal) using pandas' `getdummies()` function with `drop_first=True` to prevent multicollinearity issues. Feature Standardization: Numerical features were standardized using `StandardScaler` from sikit-learn to ensure all features contributed equally to model training, particularly important for distance-based algorithms

2. Data Integrity Validation

Quality assurance measures were implemented through assertion checks:

```
assert not heart_df.isna().any().any(), "NaNs found - handle them before training"
assert not heart_df.duplicated().any(), "Duplicate rows detected - investigate"
```

C. Exploratory Data Analysis

1. Descriptive Statistics

Comprehensive statistical analysis was performed on all variables, including:

- Central tendency measure (mean, median, mode)
- Variability measures (standard deviation, percentiles)
- Distribution characteristics through histogram visualization

2. Feature Distribution Analysis

Individual feature distributions were analyzed using seaborn histograms to identify:

- Skewness patterns in continuous variables
- Class imbalances in categorical variables
- Potential outliers requiring attention

Key findings included:

- ➔ Age distribution approximated normal distribution
- ➔ Gender imbalance with male predominance (68.3%)
- ➔ Cholesterol showed right-skewed distribution
- ➔ Target variable exhibited near-balanced classes (54.6% positive cases)
- ➔

D. Model Development Framework

1. Train- Test Split Strategy

The dataset was partitioned using stratified sampling to maintain class distribution consistency:

- Training set: 80% of data (820 samples)
- Test set: 20% of data (205 samples)
- Stratification parameters: target variable

2. Model Selection Rationale

Three distinct machine learning algorithms were selected based on their complementary strengths:

Logistic Regression was employed due to its suitability for binary classification problems, aligning with the binary nature of heart disease diagnosis (presence or absence of disease). Its probabilistic output facilitates risk scoring, while its model coefficients offer interpretability, allowing for insight into the impact of clinical variables such as age, cholesterol, and blood pressure.

Support Vector Classifier (SVC) was selected for its effectiveness in handling high-dimensional data and its ability to model non-linear relationships through kernel functions. This makes it particularly useful for capturing complex interactions among cardiovascular risk factors that may not be linearly separable.

Random Forest Classifier was utilized for its ensemble learning capabilities, which improve predictive performance and reduce overfitting. It inherently manages feature interactions and provides feature importance rankings, aiding in the identification of the most influential predictors of heart disease.

3. Pipeline Architecture

A standardized pipeline was implemented for each model:

```
Pipeline([
                    ('scaler',
StandardScaler()),
    ('clf', model)
])
```

This architecture ensures consistent preprocessing and prevents data leakage between training and testing phases.

E. Hyperparameter Optimization

1. Grid Search Cross-Validation

Systematic hyperparameter tuning was performed using GridSearchCV with the following configuration:

- Cross-validation strategy: 5-fold stratified cross-validation
- Scoring metric: ROC-AUC (Area Under the Receiver Operating Characteristic curve)
- Parallel processing: Enabled for computational efficiency

2. Parameter Grids

```
Random Forest Optimization:
n_estimators: [100, 200, 500]
max_depth: [None, 5, 10, 15, 20]
min_samples_leaf: [1, 2, 4]

Logistic Regression Optimization:
C: [0.001, 0.01, 0.1, 1, 10, 100]

SVC Optimization:
kernel: ['rbf']
C: [10]
gamma: ['scale']
```

F. Model Validation and Evaluation

1. Performance Metrics

Model performance was assessed using multiple evaluation criteria:

Accuracy: Overall correctness of predictions

$$Accuracy = \frac{Number\ of\ Correct\ Predictions}{Total\ Number\ of\ Predictions}$$

Precision: Proportion of true positive predictions among positive predictions

$$Precision = \frac{True\ Positives}{True\ Positives\ +\ False\ Positives}$$

Recall: Proportion of actual positives correctly identified

$$\mathrm{Re}\,call = \frac{True\ Positives}{True\ Positives\ +\ False\ Negative}$$

F1-Score: Harmonic mean of precision and recall

$$F1 - score = 2 \cdot \frac{Precision \cdot \mathrm{Re}\,call}{Precision\ +\ \mathrm{Re}\,call}$$

ROC-AUC: Area under the ROC curve measuring discrimination ability

2. Cross-Validation Strategy

5-fold stratified cross-validation was employed to ensure robust performance estimation while maintaining class distribution in each fold.

3. Learning Curve Analysis

Learning curves were generated to assess model behavior:

- Training sizes: Linear progression from 10% to 100% of training data
- Bias-variance trade-off evaluation
- Overfitting/underfitting detection

G. Error Analysis Framework

1. Confusion Matrix Analysis

Detailed confusion matrices were generated for each model to identify:

- False positive rates (Type I errors)
- False negative rates (Type II errors)
- Class-specific performance patterns

2. Difficult Case Identification

A systematic approach was implemented to identify challenging instances:

- Instances misclassified by multiple models were flagged
- Top 10 most difficult cases were analyzed
- Feature patterns in problematic cases were examined

3. Model Ensemble Consideration

The framework evaluated potential benefits of combining multiple models through:

- Voting mechanisms
- Stacking approaches
- Performance complementarity analysis

H. Validation Protocols

1. Sanity Checks

Multiple validation protocols were implemented:

Mini-batch Fitting: Models were tested on small subsets (20 samples) to ensure basic functionality
Label Scrambling: Random label assignment was used to verify models weren't exploiting data artifacts

Feature Shuffling: Random feature permutation confirmed genuine pattern learning

2. Reproducibility Measures

All random operations were seeded (random_state=42) to ensure:

- Consistent results across multiple runs
- Comparable performance metrics
- Reliable model comparison

Results

This section includes the results of the classification models, illustrating their performance on the test set and the representation of their training behavior through learning curves after the hyperparameter optimization.

A. Logistic Regression

```
Fitting 5 folds for each of 10 candidates, totalling 50 fits
Best LR params: {'clf__C': 0.01, 'clf__penalty': 'l2'}
LR ROC AUC: 0.91
```

Fig. 1. Logistic Regression hyperparameter tuning yielded an optimal C=0.01 and penalty=l2 with a cross-validated ROC-AUC of 0.91.

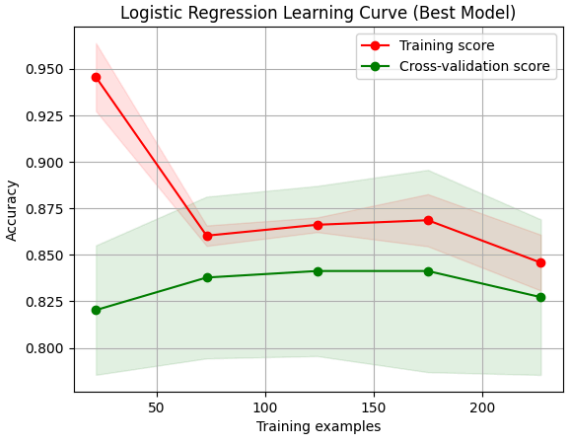

Fig. 3. Learning curve for the best-tuned Logistic Regression model. The plot shows the training accuracy (red line) and cross-validation accuracy (green line) as a function of the number of training examples.

```
Logistic Regression Final Report:
              precision    recall  f1-score   support

           0       0.76      0.88      0.81        25
           1       0.89      0.78      0.83        32

    accuracy                           0.82        57
   macro avg       0.83      0.83      0.82        57
weighted avg       0.83      0.82      0.83        57
```

Fig. 2. The Logistic Regression model achieved an accuracy of 0.82 on the test set.

B. Support Vector Classifier

```
Fitting 5 folds for each of 12 candidates, totalling 60 fits
Best SVC params: {'clf__C': 0.1, 'clf__gamma': 'scale', 'clf__kernel': 'rbf'}
SVC ROC AUC: 0.8975
```

Fig. 4. Support Vector Classifier hyperparameter tuning yielded an optimal C=0.1, gamma='scale', and kernel='rbf' with a cross-validated ROC-AUC of 0.8975.

```
SVC Final Report:
              precision    recall  f1-score   support

           0       0.88      0.84      0.86        25
           1       0.88      0.91      0.89        32

    accuracy                           0.88        57
   macro avg       0.88      0.87      0.87        57
weighted avg       0.88      0.88      0.88        57
```

Fig. 5. The Support Vector Classifier model achieved an accuracy of 0.88 on the test set.

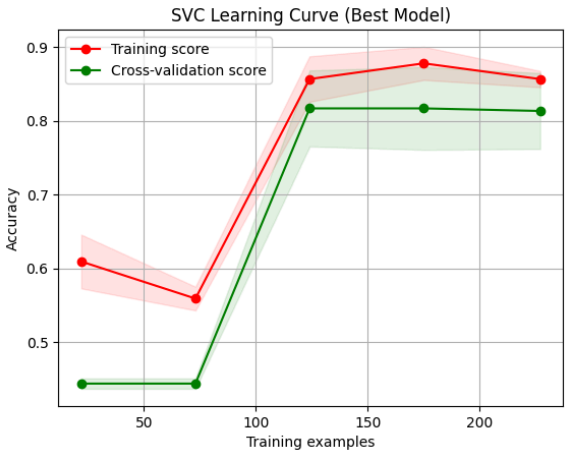

Fig. 6. Learning curve for the best-tuned Support Vector Classifier model. The plot shows the training accuracy (red line) and cross-validation accuracy (green line) as a function of the number of training examples.

C. Random Forest Classifier

```
Fitting 5 folds for each of 36 candidates, totalling 180 fits
Best RF params: {'clf__max_depth': 5, 'clf__min_samples_leaf': 2, 'clf__n_estimators': 100}
RF ROC AUC: 0.90875
```

Fig. 7. Random Forest Classifier hyperparameter tuning yielded an optimal max_depth=5, min_samples_leaf=5, and n_estimators=’200’ with a cross-validated ROC-AUC of 0.9088.

```
Random Forest Final Report:
              precision    recall  f1-score   support

           0       0.79      0.88      0.83        25
           1       0.90      0.81      0.85        32

    accuracy                           0.84        57
   macro avg       0.84      0.85      0.84        57
weighted avg       0.85      0.84      0.84        57
```

Fig. 8. The Random Forest Classifier model achieved an accuracy of 0.84 on the test set.

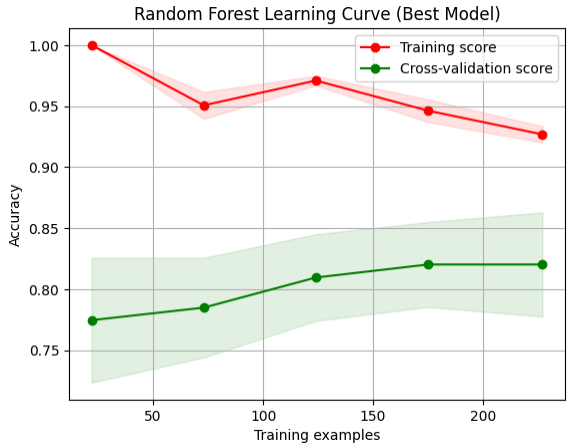

Fig. 9. Learning curve for the best-tuned Random Forest Classifier model. The plot shows the training accuracy (red line) and cross-validation accuracy (green line) as a function of the number of training examples.

We also employed data augmentation techniques to enhance the quality and diversity of the dataset. This process involved applying transformations such as rotation, flipping, scaling, and noise injection to generate additional training samples. The objective was to mitigate overfitting and improve the model’s generalization capabilities, particularly in scenarios with limited original data. Augmentation proved beneficial in increasing data variability, thereby contributing to more robust and reliable model performance. However, the results were slightly similar. Suggesting that the model is already stable and may have been well-trained already.

Table 2: Model score before tuning

Model	Accuracy	Precision	Recall	f1-score
Logistic Regression	0.81	0.86	0.78	0.82
SVC	0.82	0.87	0.81	0.84
Random Forest	0.77	0.85	0.72	0.78

Table 3: Model score after tuning

Model	Accuracy	Precision	Recall	f1-score
Logistic Regression	0.82	0.89	0.78	0.83
SVC	0.88	0.88	0.91	0.89
Random Forest	0.84	0.90	0.81	0.85

Despite a significant reduction in dataset size from 1,025 to 300 entries due to data cleaning and preprocessing, the model maintained high performance across key evaluation metrics, including accuracy, precision, recall, F1 score, and ROC AUC. This outcome suggests that the model is well-trained, capable of learning effectively from a smaller yet more refined dataset, and demonstrates strong generalization capabilities even with reduced input data.

Discussion

This section delves into the findings and implications of the heart disease prediction model developed in this study, alongside a consideration of similar efforts in the field. The aim is to provide a comprehensive analysis of the achieved performance, discuss the potential underlying factors, and outline avenues for future research in predicting heart disease using machine learning.

Our study and that of Bhosale [7] utilized the same heart disease dataset; however, we employed different methodological approaches in our analyses. Refer to the table below.

Feature	Current Work	Related Work (Pratik Bhosale)
Dataset Size	1025 instances initially, 300 after dropping duplicates and thal = 0 rows	1025 instances were used. Including duplicates.
Thal handling	Rows where thal is 0 were removed	thal value of 0 replaces with the mode
Categorical Encoding	One-hot encoding applied to cp, restecg, slope, thal	One-hot encoding applied to cp, thal, slope, ca
Numerical scaling	Numerical features were standardized	Numerical features were standardized

Model Evaluated	Decision Tree, Random Forest, K-Nearest Neighbor, Logistic Regression, Support Vector Machine	Logistic Regression, Gaussian Naive Bayes, Random Forest, Gradient Boosting, K-Nearest Neighbors, XGBoost, Support Vector Machine
Model Performance	Logistic Regression (82%), Random Forest (84%) , Support Vector Machine (88%)	Random Forest, XGBoost, Gradient Boosting (100%), SVC (98%), KNN (81%), Logistic Regression (80%), Naive Bayes (78%)

The performance differences observed among the three models: Logistic Regression, Random Forest, and Support Vector Machine can be better understood when examined alongside specific preprocessing decisions and methodological variations between this study and related work by Pratik Bhosale.

In this study, Logistic Regression recorded an accuracy of approximately 82%, demonstrating relatively consistent results across multiple folds. As a linear model, its performance suggests that the data may possess a degree of linear separability, allowing simpler decision boundaries to perform effectively. Random Forest, a more complex ensemble method, achieved around 84% accuracy, but exhibited sensitivity to data partitioning and parameter settings. This may reflect an interaction between overfitting tendencies and the reduced dataset size. Meanwhile, Support Vector Machine (SVM) showed the highest accuracy in this study at 88%, particularly after data augmentation, though its performance was highly dependent on kernel and regularization tuning.

A key factor influencing model outcomes is the data preprocessing pipeline. This study reduced the dataset from 1,025 to 300 entries by removing duplicate records and discarding rows with invalid thal values (i.e., those labeled as zero). In contrast, the related work by Bhosale retained all 1,025 instances, including duplicates, and replaced invalid thal entries with the mode rather than removing them. While this approach resulted in reported accuracies as high as 100% for some models (e.g., Random Forest and Gradient Boosting), such results likely reflect data leakage—where duplicates from the training set also exist in the test set, inflating evaluation scores artificially.

The encoding strategy also differed slightly. This study applied one-hot encoding to the features cp, restecg, slope, and thal, while the related work encoded cp, thal, slope, and ca. Additionally, both studies employed numerical standardization, but only this work emphasized its role in mitigating feature scale imbalance for distance-based models like SV

Conclusion

This study explored the application of supervised machine learning models including Logistic Regression, Support Vector Classifier, and Random Forest Classifier for heart disease risk prediction using a reduced and cleaned clinical dataset. Each model demonstrated reasonably consistent performance across evaluation metrics, with minor variations that were influenced by hyperparameter choices, model structure, and how the data was prepared.

The Support Vector Classifier exhibited comparatively stronger generalization after data augmentation, while Logistic Regression remained stable and interpretable throughout the evaluation. Random Forest produced competitive outcomes but appeared more sensitive to tuning and data partitioning.

The analysis highlighted the importance of preprocessing decisions such as removing duplicate records and standardizing features, which affected model behavior and performance. Models trained on carefully curated data tended to yield more realistic estimates compared to those using unfiltered datasets, where data redundancy or leakage may have led to inflated results.

Although each model has its limitations, the evaluation framework involving learning curves, confusion matrices, and cross-validation offered valuable insights into how these algorithms learn from and adapt to clinical data. These findings help

build a deeper understanding of machine learning behavior in healthcare contexts, particularly in structured prediction tasks based on limited patient features.

REFERENCES

[1] World Health Organization: WHO, "Cardiovascular diseases (CVDs)," Jun. 11, 2021. Available: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)

[2] "Early diagnosis can make all the difference to the lives of people with heart failure and other forms of cardiovascular disease - MedTech Europe," *MedTech Europe*, Sep. 29, 2022. Available: https://www.medtecheurope.org/medtech-views/policy-views/early-diagnosis-cardiovascular-disease/

[3] N. E. Almansouri *et al.*, "Early Diagnosis of cardiovascular Diseases in the Era of Artificial Intelligence: An In-Depth Review," *Cureus*, Mar. 2024, doi: 10.7759/cureus.55869. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC11002715/

[4] R. Kumar *et al.*, "A comprehensive review of machine learning for heart disease prediction: challenges, trends, ethical considerations, and future directions," *Frontiers in Artificial Intelligence*, vol. 8, May 2025, doi: 10.3389/frai.2025.1583459. Available: https://doi.org/10.3389/frai.2025.1583459

[5] B. S. Ingole, V. Ramineni, N. Bangad, K. K. Ganeeb, and P. Patel, "Advancements in Heart Disease Prediction: A machine learning approach for early detection and risk assessment," *arXiv.org*, Oct. 16, 2024. Available: https://arxiv.org/abs/2410.14738

[6] W. Liu *et al.*, "Machine-learning versus traditional approaches for atherosclerotic cardiovascular risk prognostication in primary prevention cohorts: a systematic review and meta-analysis," *European Heart Journal - Quality of Care and Clinical Outcomes*, Mar. 2023, doi: 10.1093/ehjqcco/qcad017. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC10284268

[7] Pratikbhosale, "Heart disease prediction: 100% accuracy," *Kaggle*, Feb. 28, 2025. Available: https://www.kaggle.com/code/pratikbhosale23/heart-disease-prediction-100-accuracy/notebook

The following AI-generated references are intended for illustrative and inspirational use only, and should not be cited as authoritative academic sources.

[8] "ChatGPT - Logistic Regression Heart disease," *ChatGPT*. Available: https://chatgpt.com/share/6877744d-18a0-8003-8080-9f1aea4a02e4

[9] "ChatGPT - ML Model Selection Guidance," *ChatGPT*. Available: https://chatgpt.com/share/687774f8-9d04-8003-b52f-bfc5c23a7e5a