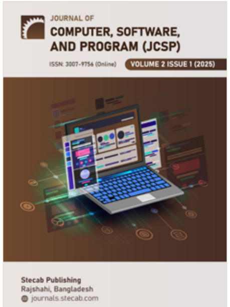

Journal of Computer, Software, and Program (JCSP)
ISSN: 3007-9756 (Online)
Volume 2 Issue 1, (2025)
https://doi.org/10.69739/jcsp.v2i1.744
https://journals.stecab.com/jcsp

Published by
Stecab Publishing

Research Article

Performance Evaluation of Machine Learning Models for Cardiovascular Disease Prediction

*[1,3]Nuraini Usman, [2]Babawuro Usman, [3]Zahriah Binti Sahri

About Article

Article History
Submission: June 28, 2025
Acceptance : August 02, 2025
Publication : August 12, 2025

Keywords
Cardiovascular Disease, CRISP-DM Framework, Machine Learning Models, Performance Evaluation, Prediction Models

About Author

[1] Department of International Programmes, Jigawa State Polytechnic for Information and Communication Technology (JSPICT), Kazaure, Nigeria

[2] Department of Computer Science, Faculty of Computing and Mathematical Sciences (FACMS), Aliko Dangote University of Science and Technology (ADUSTECH), Wudil, Kano State, Nigeria

[3] Department of Intelligence Computing, Universiti Teknikal Malaysia Melaka (UTeM), Durian Tunggal, Melaka, Malayisa

Contact @ Nuraini Usman
message2nuraini@gmail.com

ABSTRACT

Cardiovascular disease (CVD) is a deadly health issue that requires urgent attention due to the increasing global population. Using technological solutions, Machine Learning (ML) could be used to tackle health challenges in medical institutions. However, CVD diagnosis often involves multiple diagnostic procedures, leading to high medical errors. Using high-frequency and gamma rays in diagnostic devices exposes patients to high risks of other diseases. The acquisition of these devices is challenging in Low and Middle-Income Countries (LMICs), where patients are less privileged, and families lose their lives due to the inability to afford them. Healthcare institutions are utilizing the adoption of Artificial Intelligence (AI) to provide innovative solutions to these issues. In this research, the dataset was obtained from Kaggle. The Cross Industry Standard Process for Data Mining (CRISP-DM) framework was used. We selected five phases excluding the Deployment phase from the framework. Then, experiments conducted starting with exploratory data analysis, data cleansing, data visualization, transforming the dataset then splitting the CVD dataset, allocating 80% of the dataset to training and the remaining 20% as a testing dataset. We applied the different ML algorithms, where we achieved the best accuracy from Random Forest, Extra Tree Classifier, and Decision Tree all with 98.18% on the training dataset, for further experimental selection we tested the dataset where we obtained the following percentages: Random Forest 79.22%, Extra Tree Classifier 78.68% and Decision Tree 73.48%. Finally, the experiment featured the Random Forest algorithm as the best classifier compared to the rest as mentioned above, as it is more robust with the ability to handle non-linear and complex relationships making it more effective for CVD models.

Citation Style:

Usman, N., Usman, B., & Sahri, Z. B. (2025). Performance Evaluation of Machine Learning Models for Cardiovascular Disease Prediction. *Journal of Computer, Software, and Program, 2*(1), 39-48. https://doi.org/10.69739/jcsp.v2i1.744

1. INTRODUCTION

CVD is a group of diseases that involve the heart or blood vessels and other internal body organs such as the Liver, Kidney, and Brain. CVD is a prolonged disease that stands as the greatest threat to humanity, and now it has become one of the major causes of death around the globe with an exponential trend (Junwei *et al.*, 2019). Initially, CVD and other related diseases are identified by an inflammation in the body part at the beginning of the pathology. The ways to detect CVD based on laboratory investigations are long procedures (Behera *et al.*, 2015). Based on the research conducted by (Srivenkatesh, 2020), an estimate from WHO stated that about 17 million people die every year as a result of CVD, especially strokes and respiratory failures. Another research showed that (Tripathi *et al.*, 2020), bout 31% of people die from CVDs, and roughly 85% of deaths are because of heart attacks and strokes. Based on (Ruwanpathirana *et al.*, 2015), the report projected that CVDs would be leading the death of about 17.5 million patients each year, also expected to increase up to 24.2 million by 2030. CVD is a cause of death both in underdeveloped, and developing countries (Chunko *et al.*, 1996), more than 30 percent of women die annually in the USA. CVD has affected global populations, necessitating AI and ML prediction models to reduce CVD threats. AI is now established as a transformative technology used in various institutions for scientific research (Dasgupta & Wendler, 2019), and adopting this in clinical trials enhances productivity and decision-making. AI refers to machines that behave like humans, or machines that are capable of actions that need intelligence without any intervention from any human being (Samoili *et al.*, 2020), utilizing these technologies effectively boosts clinical efficiencies. AI (Rong *et al.*, 2020), is well-defined as the ability of machines to execute tasks intelligently, in contradiction to human intelligence. In situations where machines mimic the human brain in learning and terms of analysis when initiating solutions to a given problem, sympathetic intelligence is considered ML. The Centre for Technology and Global Affairs (Dasgupta & Wendler, 2019), stated that; AI enables organizations to replace relative static models with dynamic decision and interaction models to enhance decision-making processes, the detailed report enables organizations to take the right actions precisely by leveraging the report for the data supplied to the system. ML is the subset of AI and consists of methods enabling computer systems to recognize data. ML is among the most proficient and dependable techniques to treat any problem in the medical domain to cut down on medical errors. Multiple ML applications have labeled data popularly known as supervised ML (Konieczny & Idczak, 2016). The advancement of the field of ML makes it the leading field (Nasteski, 2017). ML typically refers to the dissimilarities in systems that do specific tasks that are associated with AI (Nilsson, 2005). As healthcare institutions have accumulated huge medical data which requires data mining techniques to discover the hidden patterns of the data and make effective decisions in the prediction of diseases (Kanikar, 2016). The CVD have difficult diagnosis methods as it required multiple laboratory investigations, also have high chance of clinical errors due to manual classification techniques. These procedues require a highly skilled medical practitioner to interpret the results that are generated during the medical investigations while using the Standard Operation Procedures (SOPs) (Cardiovascular Diseases - Repatriation Medical Authority, n.d.), it is also time consuming based on the inclusive investigations (Khan Mamun & Elfouly, 2023), in Low and Middle-Income Countries (LMICs) we have scarcity of resources across health institutions (Khan Mamun & Elfouly, 2023). The diagnosis procedures are use harmful radiation signals. Theses methods such as X-ray machines (Spiegel, 1995), Nuclear Imagers, CT- Scans (Radiation Risk from Medical Imaging - Harvard Health, n.d.) which trigger high risk of premature death. The machines used in the detection are also expensive making them unaffordable in the LMICs, in an advanced country like the USA, the yearly cumulative cost of CVD was approximately 351.2 billion dollars from 2014 to 2015, with approximately, 213.8 billion dollars in indirect costs, including inpatient care of 46% (Ej *et al.*, 2019), CVD is costly disease with calculated indirect costs of 237 billion dollars per year and a projected increased to 368 billion dollars by 2035 (Lopez *et al.*, 2023), this makes it significantly expensive. As healthcare institutions generate a lot of voluminous data, the need for data analytics arises with the help of AI and ML techniques to develop a sophisticated prediction model to effectively detect CVD patients early, having the systems will reduce the risk of medical errors and will boost efficiency in the decision-making process with minimal time consumption (Ekwonwune *et al.*, 2023). Also minimize the procedural exposure to harmful radiation signals, the model will use the dataset collected from various people for effective predictions and classifications of CVD patients. The main aim of this research is to enhance the detection procedure for CVD by investigating the best machine learning algorithms for the prediction and classification of the disease. This will enhance the decision-making process for medical doctors and reduce the impact of medical errors that occur in the process of diagnostic procedures, and hence significantly reducing premature death of CVD patients across the LMICs.

2. LITERATURE REVIEW

The rapid changes in AI and its adoptions in healthcare institutions have led to significant improvement in CVD early predictions, as some similar experimental research has identified a lot of contributions. The author (Balakrishnan *et al.*, 2021), used four different datasets (Cleveland, Hungaria, Switzerland with Long Beach VA heart Disease) obtained from UCI and also used algorithms (Logistics Regression, Naïve Bayes, Random Forest, Gradient Boosting, and SVM) to perform the experiments, finally the; classifier logistics regression accuracy obtained 87% score and the paper used accuracy as a performance measure. The researchers (Ware *et al.*, 2020), used one dataset (Cleveland) that was obtained from Kaggle, and several algorithms (SVM, Random Forest, KNN, Decision Tree, Naïve Bayes, and Logistics Regression) were used in the experiments, and the author made a conclusion that SVM has a score with a performance measure accuracy 89%, Precision 96%, Recall 81%, and F1 Score 88%). The authors Louridi *et al.* (2019), used a dataset from UCI (Heart Disease), in the experiments, several classifiers (SVM-linear kernel, SVM-RBF kernel, KNN, Naïve Bayes) were used and the authors concluded SVM-linear kernel with the performance measures (Accuracy 87%, Precision

87%, Sensitivity 87%, Specificity 86%, F1-Score 87%). The researchers Ravindhar *et al.* (2019), used Cardiac Disease from UCI to perform the experiments using the classifiers (Logistics Regression, Naïve Bayes, Fuzzy KNN, K-Means Clustering, BP-Neural Network), and finally, the result showed that the BP-Neural Network has a good score with the performance measures (Accuracy 98%, Precision 90%, and Recall 88%). The authors Yang *et al.* (2020), used a questionnaire to survey the residents of Zhejiang Province, the survey documents were distributed over a hundred thousand (X100,000), and the experiments used several algorithms (Multivariate Regression, CART, Naïve Bayes, Bagging Trees, AdaBoost, Random Forest, and Framingham Score), in the conclusion Random Forest obtained the highest score with the performance measure (AUC 0.787). The researchers (Srivenkatesh, 2020), conducted experiments with Kidney Disease for CVD obtained from Kaggle, and the algorithms used were KNN, SVM, Logistic Regression, Naïve Bayes, and Random Forest, finally after the conclusion of the experiments Logistic Regression was featured, with the performance measures of Precision 81%, Recall 74%, and F1-Score 71%. The authors (Sabab *et al.*, 2017), used the Goldsmiths University of London dataset and conducted experiments using three classifiers (SMO, C4.5 Decision Tree, and Naïve Bayes), the researchers featured Naïve Bayes with a performance measure of Accuracy of 87% and AUC of 0.909. The author Kanikar (2016), used the Heart Disease dataset from UCI, and two classifiers were used to conduct experiments (SVM and Naïve Bayes), in the conclusion SVM was featured with the performance measures (Accuracy 57%, Sensitivity 35%, and Specificity 87%). The researchers (Tsipouras *et al.*, 2008), the dataset was obtained from the University Hospital of Ioannina, and the experiments were conducted using four classifiers (Crisp Rule-Based Classifier, Optimized Fuzzy Model, Adaptive Neuro-fuzzy Inference System, and ANN), finally the featured classifier was ANN with performance measures (Accuracy 74%, Sensitivity 80%, Specificity 60%. The authors (Kolukisa *et al.*, 2020), the datasets used in the experiments are Cleveland and Z-Alizadeh Sani both carried out independently, for the Cleveland dataset the algorithms used are (KNN, Linear Regression, Linear Discriminant Analysis (LDA), Naïve Bayes, SVM, and Ensemble) and the featured algorithm is Ensemble with the performance measures (Accuracy 83% and AUC 0.824), for the Z-Alizadeh Sani dataset the algorithms used are (KNN, Linear Regression, Linear Discriminant Analysis (LDA), Naïve Bayes, SVM, and Ensemble) and the featured algorithm is also Ensemble with the performance measures (Accuracy 88% and AUC 0.824). The researchers (Ali, 2017), two types of datasets (Diabetes and Cardiovascular Disease) obtained from UCI, to conduct experiments, the authors used two different algorithms (ANNs and BNs) and both of which the scored algorithms are ANNs Diabetes dataset (Accuracy of 87%) and Cardiovascular Disease dataset (Accuracy 96%) are obtained. The authors (Ahmad *et al.*, 2021), conducted experiments through primary data collection from five Hospitals in Saudi with a certain collection from 2016 to 2018 which cumulatively gives the total sum of 3,000 samples, the experiments were conducted using interval 9) features of the dataset with a labeled (HbA1c) and other nine (9) features of the dataset with a labeled (FPG), during the experiments several classifiers are used (Logistic Regression, SVM, Decision Tree, Random Forest, and Ensemble) on both datasets, finally the scored algorithms are SVM and Random Forest with a performance means (Accuracy 82%, Precision 82%, Recall 82%, and F1-Score 82% (HbA1c)), and Radom Forest (Accuracy 88%, Precision 88%, Recall 88%, and F1-Score 88% (FPG)). The researchers (Gürfidan & Ersoy, 2021), used a dataset from the University of California Irvine Machine Learning Repository to conduct experiments using several algorithms (SVM, Logistics Regression, Decision Tree, KNN, LDA, GNB), finally, the authors concluded SVM as scored with a performance measure (Accuracy 83%). The researcher (Khan & Saboor, 2020), used a dataset Cleveland Heart Disease obtained from UCI to conduct experiments using the following classifiers (Logistic Regression, KNN, ANN, SVM (rbf), SVM (linear), Naïve Bayes, and Decision Tree), at the end of the experiments SVM (linear) is the featured algorithm with a performance measure (Accuracy 85%, Sensitivity 75%, and Specificity 95%). The authors (Hussain *et al.*, 2020), Congestive Heart Failure RR Interval from Physionet, an experiment carried out where several algorithms are used (Naïve Bayes, Decision Tree, SVM, Gaussian, SVM RBF, and SVM poly), finally the featured classifier is Naïve Bayes with performance measures (Accuracy 89%, Sensitivity 89% Specificity 89% and AUC 0.929. The researcher (Terrada *et al.*, 2020), used three datasets (Cleveland (the University of California Irvine), Hungarian Institute of Cardiology (Hungarian Institute of Cardiology, Budapest), Z- Alizadeh Sani (Tehran's Shaheed Rajaei Cardiovascular, Medical and Research Centre) to conduct the experiments using the following algorithms: ANN, Decision Tree, and AdaBoost, all are applied to the three datasets, the authors featured classifier ANN with best performance measures, for Cleveland dataset (Accuracy 94%, Precision 80%, and Recall 80%, for Hungarian Institute of Cardiology (Accuracy 90%, Precision 85%, Recall 78%), Z- Alizadeh Sani (Accuracy 94%, Precision 93%, Recall 98%). The authors (Yadav *et al.*, 2020), used Cleveland (University of California, Irvine) to experiment using the following algorithms: K-Means, KNN, Naïve Bayes, NB & K-Mean (Hybrid), Logistic Regression, Fuzzy KNN, and NN, conclusively the authors featured classifier with best performance measure NN 98%. The researchers (Zhang *et al.*, 2020), used Heart Disease obtained from UCI, to conduct experiments using the classifiers (KNN, Logistic Regression, SVC, Decision Tree, MLP, Random Forest, LGB, and Gradient Boosting), finally, the authors featured Logistic Regression with the highest performance measures (Accuracy 88%, Precision 91%, Recall 89%, and F1-Score 90%.

Based on the literatures the key limitations are: inconsistenet evaluation metrics and reporing as shown many studies reported accuracy only neglecting the critical metrics; overreliance on a single-source, small datasets obtained from UCI Cleveland (n < 300>, this lacks demographic and clinical diversity making the model not to perform well on the real-world heterogeneous populations; lack of rigorous model comparism under uniform conditions; and limitation in the use of AUC-ROC for model validation. However, this study addresses these limitations directly through evaluating multiple algorithms on the CVD dataset, developing a comprehensive performance evaluation metrics providing a holistic view of model critical performance. Hence, the study provide a transparent, regorous, and clincal informed comparision, Random Forest emerged as highest performing and interpretable model that is well-suited for accurate and actionable model on CVD risk prediction.

3. METHODOLOGY

This research adopted the CRISP-DM Framework which is a popular methodology with the highest adoption rate by data mining experts. It is a process model that describes the data science project lifecycle. CRISP-DM helps to provide a uniform framework for experimental documentation and guidelines. It is cost-effective by taking out simple data mining step-by-step tasks with good processes established across the industry and cooperatively it encourages excellent project replication practice. CRISP-DM provides planning and managing any project uniform procedures by following the phases. CRISP-DM can be implemented in any project in any domain across the world. CRISP-DM is to formulate the reasons for the goals of Knowledge in the Data Discovery process. The researchers diversified areas in both the private and public sectors in which data mining is used as a simplified (Bošnjak *et al.*, 2009). Based on my understanding CRISP-DM framework is adaptable by many different companies across the world which is why it has the word "CROSS". The phases are flexible and easier when doing reverse engineering, much information that is featured, and getting insights becomes simplified.

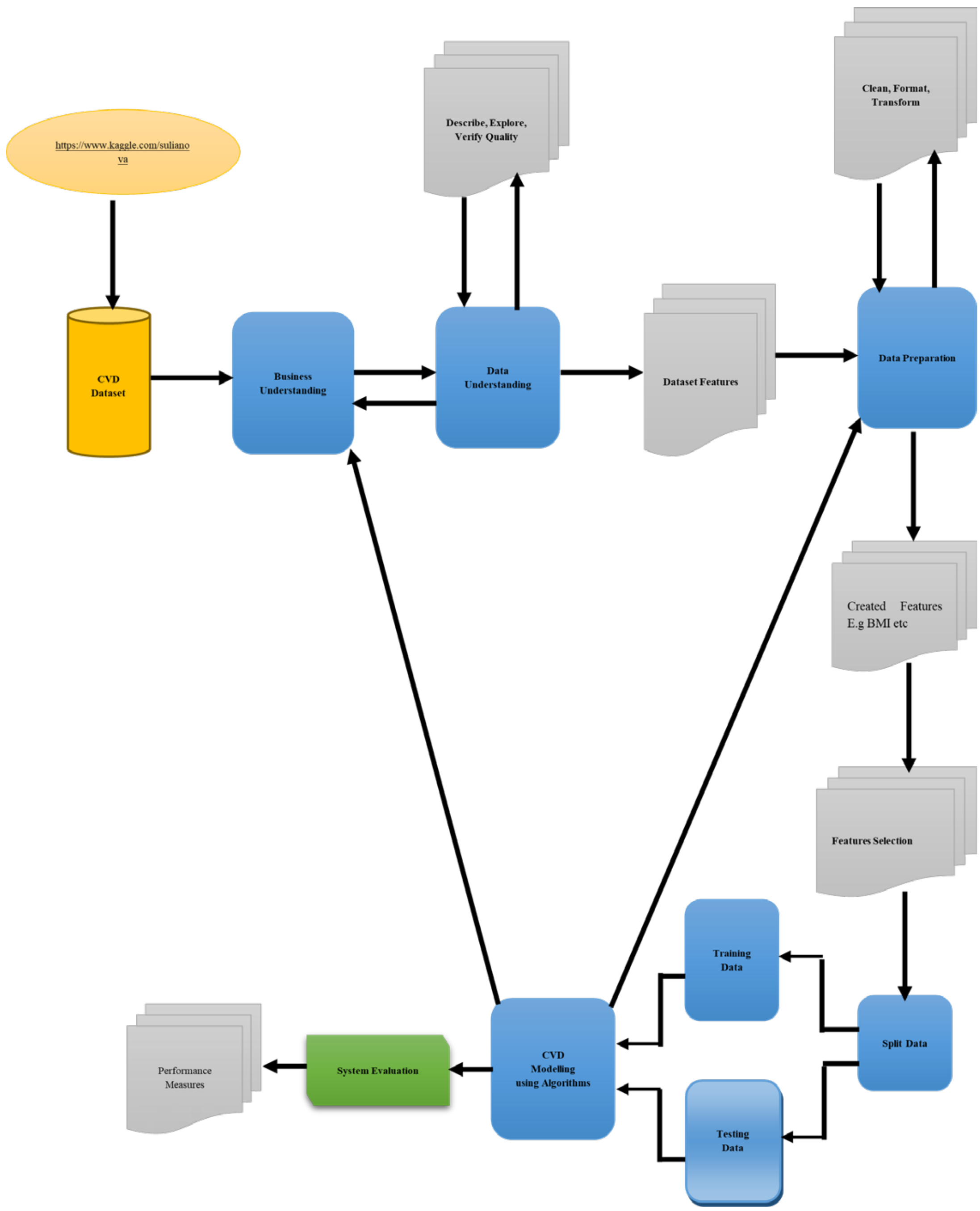

Figure 1. The Proposed Architecture of CVD Prediction Models

The benefits of using CRISP-DM are essential in any domain of ML or Data Science projects.

i. Open-Source Framework: It is available at research disposals.

ii. Reusability Framework: It is easier to be used on the next project in the future.

iii. Adaptable Framework: It is easier to implement with a detailed report.

iv. Cross-Platform Framework: It can be easily adopted in any kind of research.

3.1. The CRISP-DM Phases in the Experiment

3.1.1. Business Understanding

The business understanding phase is critical in any project, as it involves identifying the business requirements and transforming them into analytics for technical transformation. This stage aims to achieve project objectives by determining the business objectives, assessing the situation, and determining the goals of data analysis. The CVD model's primary objective is to investigate the best ML algorithms for prediction, make recommendations based on the developed model, and assess its performance measures. The situation assessed by identifying available resources, targeted objectives, and risks, and implementing contingency plans.

3.1.2. Data Understanding

The data analysis process involves extracting meaningful rules, patterns, and models from the CVD dataset. Data mining is a trending area that provides a deeper understanding of the CVD dataset, enabling the development of a CVD prediction system that supports informed decision-making in healthcare institutions. This phase of the project involves a deeper understanding and the data exploration of the CVD dataset, which includes four tasks. The dataset was published by Svetlana Ulianova at Ryerson University, Toronto, Ontario, Canada in 2019 on Kaggle. The dataset contains 70,000 data points with 12 features, which are age, height, weight, gender, systolic and diastolic blood pressure, cholesterol, glucose, smoking, alcohol, physical activities, and cardiology disease. The tasks include identifying the presence or absence of CVD, analyzing patient social information, and predicting the presence of CVD.

3.1.3. Data Preparation

This stage involves achieving research objectives from theoretical to implementation. This involves 'Data Munging' or 'Data Wrangling', which involves mining deeper into the CVD dataset to understand patterns and extract insights for modeling. Data cleaning is a major step, confirming noise and problems are removed to improve the model. Data transformation encompasses creating dummy attributes to transform the initial dataset, while data formatting ensures accuracy without missing values. Data quality verification ensures data is accurate and meets high accuracy standards. The preliminary dataset has 70,000 samples, after going through data preparation, the noise was removed. Then, we reduced the sample size to 57,138. The next is to supply the CVD dataset for training allocating 45,710 (80%) of the data and testing allocating 11,428 (20%) of the dataset to develop the CVD prediction models.

3.1.4. System Modelling

This is among the major steps of the ML project which allows the development of a system using different types of supervised ML algorithms to effectively give the classification of the patients with or without CVD. We applied experimental algorithms to the models.

3.1.5. System Evaluation

In this stage, the researchers have assessed the performance of the developed model to effectively identify the insight and how the model is most likely to perform. We assess all the algorithms used in the system model to see how it performs on the CVD dataset.

3.1.5.1. Performance Measures

We evaluated our CVD prediction models with the following:

- Specificity = TP / (TN + FP) or TN / Overall Negatives.
- Sensitivity = TP / (TP + FN) or TP / Overall Positives.
- Precision = TP / (TP + FP).
- Classification Error = (FP + FN) / float (TP + TN + FP + FN).
- Accuracy = TP + TN / TP + FP + FN + TN (Total Samples).
- F1- Score = 2 x (TP / TP + FP) x TP / TP + TN) / (TP / TP + FP) + (TP/TP +TN).
- F1_Score = 2 x Precision x Recall / Precision + Recall

Table 1. The Confusion Matrix Default

		YES (Positive)	NO (Negative)
Actual Class (Predicted Values)	YES (Positive)	True Positives (TP)	False Negatives (FN)
	NO (Negative)	False Positives (FP)	True Negative (TN)

TP = Patients with CVD correctly predicted as sick people by the model.

FN = Patients with CVD incorrectly predicted as not sick by CVD by the model.

FN = Healthy Patients incorrectly predicted as CVD patients by the model.

TN = Healthy Patients correctly predicted as patients without CVD by the model.

Note:

False Positive – Type 1 Error

False Negative – Type 2 Error

Type 1 Error: The actual value was negative, but the model predicted a positive value.

Type 2 Error: The actual value was positive, but the model predicted a negative value.

4. RESULTS AND DISCUSSION

This section of the study presents the key findings of the CVD data analysis and interprets based on the correlation and feature importance evaluation done to understand the key factors associated with the CVD occurrence.

4.1. Experiments and Results

The CRISP-DM framework on CVD experimental analysis.

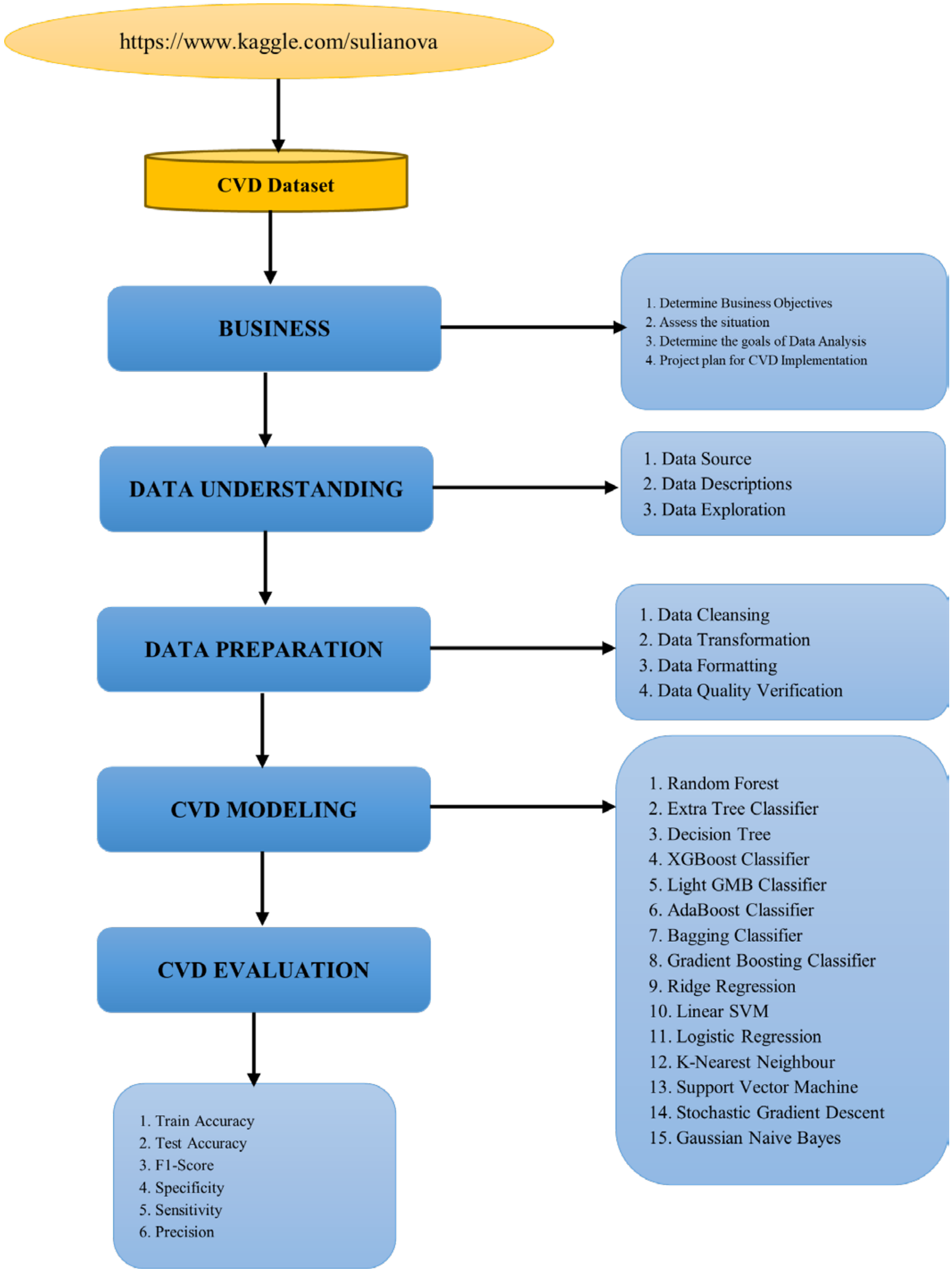

Figure 2. The Experimental Architecture for CVD Prediction Mode

4.2. Preparation of Tables and Figures

Table 2. System Performance Evaluation

Models	Train Accuracy (%)	Test Accuracy (%)	F1-Score	Specificity	Sensitivity	Precision	Classification Error	True Positives (TP)	True Negatives (TN)	False Positives (FP)	False Negatives (FN)
Random Forest	98.18	79.22	0.8	0.726	0.749	0.688	0.265	1070	4353	1305	4700
Extra Tree Classifier	98.18	78.68	0.79	0.781	0.806	0.769	0.208	1294	4516	1142	4476
Decision Tree	98.18	73.48	0.75	0.75	0.682	0.782	0.289	1290	3917	1741	4480
XGBoost Classifier	79.34	74.29	0.73	0.763	0.655	0.816	0.304	1743	4463	1195	4027
Light GMB Classifier	75.38	73.42	0.72	0.765	0.698	0.793	0.273	1859	4480	1178	3911
AdaBoost Classifier	72.22	72.27	0.71	0.739	0.68	0.767	0.294	1883	4372	1286	3887
Bagging Classifier	72.95	72.73	0.71	0.765	0.698	0.793	0.273	1943	4485	1173	3827
Gradient Boosting Classifier	72.95	72.71	0.71	0.758	0.563	0.897	0.396	1945	4484	1174	3825
Ridge Regression	71.05	71.6	0.7	0.759	0.659	0.831	0.305	1911	4323	1335	3859
Linear SVM	71.37	71.8	0.7	0.771	0.749	0.789	0.257	1929	4364	1294	3841
Logistic Regression	70.71	70.62	0.69	0.769	0.719	0.792	0.266	2042	4342	1316	3728
K-Nearest Neighbour	71.82	71.15	0.69	0.796	0.707	0.792	0.213	2062	4423	1235	3708
Support Vector Machine	69.99	69.64	0.66	0.751	0.775	0.773	0.277	2430	4619	1039	3340
Stochastic Gradient Descent	69.93	69.51	0.65	0.743	0.698	0.764	0.284	2542	4716	942	3228
Gaussian Naive Bayes	61.04	60.44	0.45	0.748	0.693	0.771	0.282	3936	5073	585	1834

4.3. Discussions (Data Analysis Key findings)

We used both heatmap of Pearson correlations and random forest classifier, we were able to identify and ranked features based on their statistical relationships in context with the variable (disease) and its predictive contribution to the developed model. The correlation matrix has shown that age, Body Mass Index (BMI), and weight have the strongest association with CVD (disease), the correlation were all under 0.20. Similarly, the feature importance derived from the Random Forest Classifier identified BMI, age, height, weight, and systolic blood pressure as the topset predictors. Whereas BMI consistently ranked as highest across both techniques with a significant role in the disease prediction. These conclusions provide a solid foundation for interpreting how person health indicators contribute to the disease risk factor and inform targeted interventions.

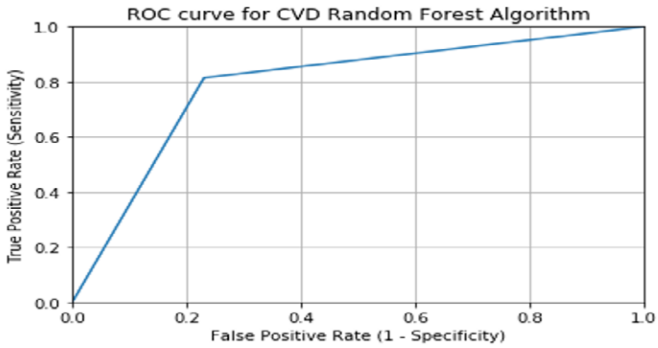

Figure 3. ROC Curve for CVD Random Forest Algorithm

The ROC/AUC curve score for the CVD Extra Tree Classifier model is 78% which is also the second-best score.

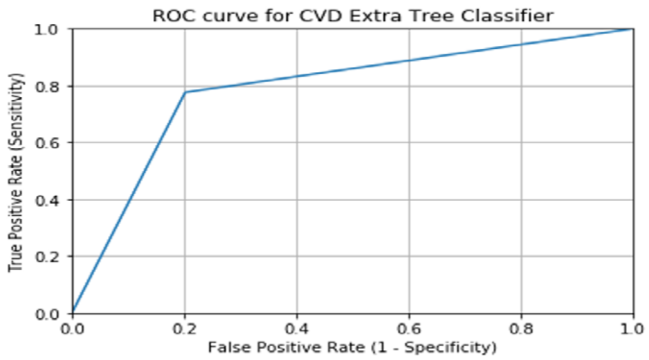

Figure 4. ROC Curve for CVD Extra Tree Classifier

The numerical features such as systolic blood pressure, diastolic blood pressure, and pressure possessed a wide range of values and potential outliers. The categorical features like smoke, gender, alcohol, and active (cardio exercises) show fluctuating frequencies across their groups, with some exhibiting imbalances notation. The disease variable has a near 50/50 class balance distributions (approximately 49.72% for disease=0 and 50.28% for disease=1). Age, BMI, height, weight, and systolic blood pressure were identified as the most influenced features for predicting disease in relation to the correlation between and Random Forest Classifier feature importance. BMI and age features have shown the strongest importance in respect of the Random Forest Classifier, also BMI and age had the highest absolute correlation within CVD disease among the numerical analyzed features.

4.4. ROC curve for CVD Prediction Model

In this stage we focused on the top two ROC curves among others, this allows us to identify the most relevant best ROC curve that we have in our CVD prediction model.

The ROC/AUC curve score for the CVD Random Forest model is 79% which is relatively good for the model and it has the highest score among all other algorithms used in the experiments.

5. CONCLUSION

The report has shown that the prediction model may be used as a decision-support system in healthcare institutions. The CVD prediction model can provide more insight into the data starting with data cleaning, data transformation, and data visualization. During the experiment, the researchers adopted the CRISP-DM framework on top of the ML algorithms and then performed the best modeling on the CVD dataset, after the experiments we achieved the best based on the performance measures from the Random Forest with 98.18% accuracy from training, 79.22% accuracy from testing. Then, the Extra Tree Classifier with 98.18% accuracy from training, and 78.68% accuracy from testing. Finally, the Decision Tree with 98.18% accuracy from training, and 73.48% accuracy from testing. So, the research finally featured the Random Forest Algorithm as the best based on the performance measures used in the experiments. The researchers suggested that future work for the CVD prediction model based on the research investigations will add more values and functionalities to the model by making it more intelligent. The recommendations are as follows:

i. The CVD prediction model should have a well-designed interface to enable friendly interaction with the model, by allowing the user to insert new data into the model.

ii. The CVD prediction model should store more data so that the model would be smart and provide a detailed report related to the patient's supplied information.

iii. For the model to become more sophisticated in the future, we recommended the model have a big data repository for storing the data for all patients with related CVD cases across the LMICs.

iv. The model should have login credentials to avoid unwanted users across the LMICs and for the data protection of the patients.

v. The model should record all user log information, and it should have a dashboard that will provide adequate information related to the CVD patients across the healthcare intuitions across the LMICs.

vi. The model to deployed as a web application, mobile application, and or desktop application over the internet with limited access to LMICs.

vii. The LMICs may have the challenge of having internet access. We recommend the CVD prediction model work offline and later synchronize the data to the online server.

REFERENCES

Ahmad, H. F., Mukhtar, H., Alaqail, H., Seliaman, M., & Alhumam, A. (2021). Investigating health-related features and their impact on the prediction of diabetes using machine learning. *Applied Sciences, 11*(3), 1173.

Alić, B., Gurbeta, L., & Badnjević, A. (2017, June). Machine learning techniques for classification of diabetes and cardiovascular diseases. In *2017 6th mediterranean conference on embedded computing (MECO)* (pp. 1-4). IEEE.

Balakrishnan, M., Arockia Christopher, A. B., Ramprakash, P., & Logeswari, A. (2021). Prediction of Cardiovascular Disease using Machine Learning. *Journal of Physics: Conference Series, 1767*(1), 1–7. https://doi.org/10.1088/1742-6596/1767/1/012013

Behera, S., Pramanik, K., & K Nayak, M. (2015). Recent advancement in the treatment of cardiovascular diseases: Conventional therapy to nanotechnology. *Current pharmaceutical design, 21*(30), 4479-4497. https://doi.org/10.2174/1381612821666150817104635

Bosnjak, Z., Grljevic, O., & Bosnjak, S. (2009, May). CRISP-DM as a framework for discovering knowledge in small and medium sized enterprises' data. In *2009 5th International Symposium on Applied Computational Intelligence and Informatics* (pp. 509-514). IEEE. https://doi.org/10.1109/SACI.2009.5136302

Cardiovascular Diseases - Repatriation Medical Authority. (n.d.). Retrieved July 2, 2025, from https://www.rma.gov.au/sops/category/cardiovascular-diseases

Chunko, M. T., Hayunga, E., Rothenberg, K., Rudick, J., & Pinn, V. W. (1996). Office of Research on Women's Health National Institutes of Health. *Journal of Women's Health, 5*(4), 317–323. https://doi.org/10.1089/jwh.1996.5.317

Dasgupta, A., & Wendler, S. (2019). *AI adoption strategies.* Retrieved September, 18, 2022.

Ej, B., Muntner, P., & Alonso, A. (2019). AHA 2019 Heart Disease and Stroke Statistics. *American College of Cardiology*, 2–5.

Ekwonwune, E. N., Ubochi, C. I., Duruoha, A. E., Duru, N. A., Ekwonwune, E. N., Ubochi, C. I., Duruoha, A. E., & Duru, N. A. (2023). Development of a Data Mining Model to Detect Cardiovascular Disease. *International Journal of Communications, Network and System Sciences, 16*(5), 77–96. https://doi.org/10.4236/IJCNS.2023.165006

Gürfidan, R., & Ersoy, M. (2021). Classification of death related to heart failure by machine learning algorithms. *Advances in Artificial Intelligence Research, 1*(1), 13-18.

Hussain, L., Aziz, W., Khan, I. R., Alkinani, M. H., & Alowibdi, J. S. (2021). Machine learning based congestive heart failure detection using feature importance ranking of multimodal features. *Mathematical Biosciences and Engineering, 18*(1), 69-91. https://doi.org/10.3934/mbe.2021004

Junwei, K., Yang, H., Junjiang, L., & Zhijun, Y. (2019). Dynamic prediction of cardiovascular disease using improved LSTM. *International Journal of Crowd Science, 3*(1), 14-25. https://doi.org/10.1108/IJCS-01-2019-0002

Kanikar, P., & Shah, D. R. (2016). Prediction of cardiovascular diseases using support vector machine and Bayesian classification. *International Journal of Computer Applications, 156*(2), 9-13.

Jevin, M. J., Jayant, H., Sanjay, R., Hemasai, V., & Venkatasrinivas, P. V. (2023). Heart disease identification method using machine learning classification in e-healthcare. *Heart Disease, 10*(3). https://doi.org/10.1109/ACCESS.2020.3001149

Khan Mamun, M. M. R., & Elfouly, T. (2023). Detection of Cardiovascular Disease from Clinical Parameters Using a One-Dimensional Convolutional Neural Network. *Bioengineering, 10*(7). https://doi.org/10.3390/BIOENGINEERING10070796,

Kolukisa, B., Yavuz, L., Soran, A., Bakir-Gungor, B., Tuncer, D., Onen, A., & Gungor, V. C. (2020). Coronary artery disease diagnosis using optimized adaptive ensemble machine learning algorithm. *International Journal of Bioscience, Biochemistry and Bioinformatics, 10*(1), 58-65. https://doi.org/10.17706/ijbbb.2020.10.1.58-65

Konieczny, R., & Idczak, R. (2016). Mössbauer study of Fe-Re alloys prepared by mechanical alloying. *Hyperfine Interactions, 237*(1), 1–8. https://doi.org/10.1007/s10751-016-1232-6

Lopez, E. O., Ballard, B. D., & Jan, A. (2023). *Cardiovascular Disease.* StatPearls Publishing. https://www.ncbi.nlm.nih.gov/books/NBK535419/

Louridi, N., Amar, M., & Ouahidi, B. El. (2019). Identification of Cardiovascular Diseases Using Machine Learning. *7th Mediterranean Congress of Telecommunications 2019, CMT 2019* (1–6). https://doi.org/10.1109/CMT.2019.8931411

Nasteski, V. (2017). An overview of the supervised machine learning methods. *Horizons. b, 4*(51-62), 56. https://doi.org/10.20544/horizons.b.04.1.17.p05

Nilsson, N. J. (2005). Introduction to machine learning an early draft of a proposed textbook Department of Computer Science. *Machine Learning, 56*(2), 387–399.

Radiation risk from medical imaging - Harvard Health. (n.d.). Retrieved July 2, 2025, from https://www.health.harvard.edu/cancer/radiation-risk-from-medical-imaging

Ravindhar, N. V., Anand, K., Hariharan, S., Ragavendran, & Winster, S. G. (2019). Intelligent diagnosis of cardiac disease prediction using machine learning. *International Journal of Innovative Technology and Exploring Engineering, 8*(11), 1417–1421. https://doi.org/10.35940/ijitee.J9765.0981119

Rong, G., Mendez, A., Bou, E., Zhao, B., & Sawan, M. (2020). Artificial Intelligence in Healthcare : Review and Prediction Case Studies. *Engineering, 6*(3), 291–301. https://doi.org/10.1016/j.eng.2019.08.015

Ruwanpathirana, T., Owen, A., & Reid, C. M. (2015). Review on Cardiovascular Risk Prediction. *Cardiovascular Therapeutics, 33*(2), 62–70. https://doi.org/10.1111/1755-5922.12110

Sabab, S. A., Munshi, M. A. R., Pritom, A. I., & Shihabuzzaman, S. (2017). Cardiovascular disease prognosis using effective classification and feature selection technique. *1st International Conference on Medical Engineering, Health Informatics and Technology, MediTec 2016* (pp. 2–7). https://doi.org/10.1109/MEDITEC.2016.7835374

Samoili, S., López Cobo, M., Gómez, E., De Prato, G., Martínez-Plumed, F., & Delipetrev, B. (2020). AI Watch - Defining Artificial Intelligence. Towards an operational definition and taxonomy of artificial intelligence. In *Joint Research Centre* (European Commission). https://doi.org/10.2760/382730

Spiegel, P. K. (1995). The first clinical X-ray made in America--100 years. *American Journal of Roentgenology, 164*(1), 241–243. https://doi.org/10.2214/AJR.164.1.7998549

Srivenkatesh, M. (2020). Prediction of cardiovascular disease using machine learning algorithms. *International Journal of Engineering and Advanced Technology, 9*(3), 2404-2414. https://doi.org/10.35940/ijeat.B3986.029320

Terrada, O., Hamida, S., Cherradi, B., Raihani, A., & Bouattane, O. (2020). Supervised machine learning based medical diagnosis support system for prediction of patients with heart disease. *Advances in Science, Technology and Engineering Systems, 5*(5), 269–277. https://doi.org/10.25046/AJ050533

Tripathi, K., Garg, H., & Sharma, H. (2020). A comprehensive survey on cardiovascular disease. *International Journal of Grid and Distributed Computing, 13*(1), 1772-1780.

Tsipouras, M. G., Exarchos, T. P., Fotiadis, D. I., Kotsia, A. P., Vakalis, K. V., Naka, K. K., & Michalis, L. K. (2008). Automated diagnosis of coronary artery disease based on data mining and fuzzy modeling. *IEEE Transactions on information technology in biomedicine, 12*(4), 447-458.

Ware, S., Rakesh, S., & Choudhary, B. (2020). Heart Attack Prediction by using Machine Learning Techniques. *International Journal of Recent Technology and Engineering, 8*(5), 1577–1580. https://doi.org/10.35940/ijrte.d9439.018520

Yadav, S. S., Jadhav, S. M., Nagrale, S., & Patil, N. (2020). Application of Machine Learning for the Detection of Heart Disease. *2nd International Conference on Innovative Mechanisms for Industry Applications, ICIMIA 2020 - Conference Proceedings, Icimia* (pp. 165–172). https://doi.org/10.1109/ICIMIA48430.2020.9074954

Yang, L., Wu, H., Jin, X., Zheng, P., Hu, S., Xu, X., Yu, W., & Yan, J. (2020). Study of cardiovascular disease prediction model based on random forest in eastern China. *Scientific Reports, 10*(1), 1–8. https://doi.org/10.1038/s41598-020-62133-5

Zhang, C., Qu, J., Li, W., & Zheng, L. (2020, December). Predicting cardiovascular events by machine learning. In *Journal of Physics: Conference Series* (Vol. 1693, No. 1, p. 012093). IOP Publishing. https://doi.org/10.1088/1742-6596/1693/1/012093