FUDMA Journal of Sciences (FJS)
ISSN online: 2616-1370
ISSN print: 2645 - 2944
Vol. 9 No. 3, March, 2025, pp 201 - 208
DOI: https://doi.org/10.33003/fjs-2025-0903-3263

ENHANCED PREDICTION OF CORONARY ARTERY DISEASE USING LOGISTIC REGRESSION

***[1]Godfrey Perfectson Oise, [1]Oyedotun Samuel Abiodun, [2]Nwabuokei Onyemaechi Clement, [1]Babalola Akilo Eyitemi and [1]Unuigbokhai Nkem Belinda**

[1]Department of Computing, Wellspring University, Edo State, Nigeria.
[2]Department of Computer Science, Delta State College of Education Mosogar.

*Corresponding authors' email: godfrey.oise@wellspringuniversity.edu.ng
ORCID iD: https://orcid.org/0009-0006-4393-7874

ABSTRACT

Coronary Artery Disease (CAD) remains a leading cause of global morbidity and mortality, emphasizing the urgent need for accurate and interpretable prediction models to facilitate timely interventions and improve patient outcomes. This study investigates the application of Logistic Regression for CAD prediction, leveraging a dataset of 303 patients and 13 clinical features obtained from the UCI Machine Learning Repository. Recognizing the limitations of traditional risk assessment methods, this research explores the potential of Logistic Regression to enhance CAD prediction accuracy through a streamlined and easily implementable approach. The dataset, which encompasses demographic factors, clinical measurements, and lifestyle indicators, was subjected to rigorous analysis to evaluate the model's performance. A Logistic Regression model was developed using Python's scikit-learn library and assessed using a comprehensive set of evaluation metrics, including accuracy, precision, recall, F1-score, and the Area Under the Receiver Operating Characteristic curve (AUC-ROC). On a test set of 61 instances, the model achieved an overall accuracy of 82%, demonstrating its ability to correctly classify individuals with and without CAD. The precision and recall scores for Class 0 (absence of CAD) were 79% and 82%, respectively, while for Class 1 (presence of CAD), the precision and recall scores were 84% and 82%, respectively, indicating balanced performance across both classes. The model exhibited an AUC-ROC of 0.89, signifying strong discriminatory ability. These findings suggest that Logistic Regression can serve as a valuable tool for CAD risk assessment, providing a foundation for more advanced predictive models and contributing to improved cardiovascular health management. The study further highlights the need for enhanced data preprocessing techniques and exploration of machine learning algorithms to achieve even higher accuracy and broader applicability in clinical settings.

Keywords: Coronary Artery Disease (CAD), Logistic Regression, Machine Learning, Cardiovascular Risk Prediction, Plaque Buildup, Coronary Blood Flow, Heart Attack

INTRODUCTION

Coronary Artery Disease (CAD) continues to be a leading cause of morbidity and mortality worldwide, emphasizing the critical need for accurate risk stratification to facilitate timely preventive interventions and improve patient outcomes (Aminu Bashir Suleiman, Stephen Luka and Muhammad Ibrahim, 2023). Traditional risk assessment methods often rely on clinical judgment, which can be subjective and may not fully capture the complex interplay of risk factors involved in CAD development (Abdar et al., 2019). Machine learning techniques, particularly Logistic Regression, offer a promising avenue for enhancing CAD prediction accuracy (Mondal et al., 2025). Logistic Regression, a widely used statistical method for binary classification, provides a straightforward and interpretable approach to estimating the probability of CAD presence based on various clinical and demographic factors (Mohi Uddin et al., 2023). The model's simplicity makes it appealing for clinical implementation, allowing healthcare providers to readily understand the contributions of individual risk factors to overall CAD risk.

Coronary Artery Disease (CAD) is a common heart condition in which the coronary arteries, responsible for supplying blood to the heart muscle, become narrowed or blocked due to plaque buildup, a mixture of fat, cholesterol, and other substances (Jamilu et al., 2019). This restriction reduces blood flow and oxygen supply to the heart, increasing the risk of severe complications such as heart attacks and heart failure(Aminu Bashir Suleiman, Stephen Luka and Muhammad Ibrahim, 2023). The primary cause of CAD is atherosclerosis, where plaque accumulates inside the arteries, making them stiff and narrow (Beunza et al., 2019). Several factors contribute to the development of CAD, including high blood pressure, high cholesterol levels, smoking, diabetes, obesity, physical inactivity, an unhealthy diet, and genetic predisposition (Doppala et al., 2022). These risk factors accelerate artery damage, increasing the likelihood of reduced blood flow to the heart. Symptoms of CAD vary depending on the severity of the blockage. Chest pain (angina), shortness of breath, heart palpitations, fatigue, dizziness, and lightheadedness are common warning signs. In severe cases, CAD can lead to a heart attack (myocardial infarction) when blood flow to a section of the heart is completely blocked, causing tissue damage (Das et al., 2009). If left untreated, CAD can result in life-threatening complications such as heart failure, arrhythmias (irregular heartbeats), and sudden cardiac arrest. To diagnose CAD, doctors use various tests, including electrocardiograms (ECG), echocardiograms, and stress tests. The financial burden of treating cardiovascular diseases is also substantial. The global cost of treating heart disease is projected to exceed $1 trillion by 2035. Between 2017 and 2018 alone, the U.S. spent over $229 billion on cardiovascular treatments, while India incurred costs exceeding $237 billion between 2005 and 2015, as estimated by WHO. Reducing the cost of CAD treatment has, therefore, become a pressing priority. To determine whether a patient is at risk of developing CAD, various studies have employed classical machine learning techniques based on distinguishing risk factors (Shahid & Singh, 2020). CAD has emerged as a major contributor to the rising global death toll, requiring urgent attention and effective interventions (Tama et al., 2020). Other heart conditions, such as heart failure and arrhythmias, can also lead to life-threatening events like cardiac arrest. In developed nations, CAD remains the leading cause of death, often linked to risk factors such as diabetes, high cholesterol, smoking, genetic predisposition, and aging. However, recent studies suggest that heart disease is not solely age-dependent, as conditions like vitamin D deficiency can also increase the risk, even in infants. Early disease detection not only improves treatment outcomes but also enhances a patient's quality of life by enabling timely medical interventions. Among all heart diseases, CAD is one of the most challenging to prevent in healthy individuals. Recent research highlights that cardiovascular diseases have contributed to a 59% increase in premature deaths in India, making them one of the nation's leading causes of mortality. Given the critical role of early detection, it is necessary to develop innovative prediction models to reduce sudden cardiac arrest fatalities. Additionally, improving CAD prediction methods can significantly reduce the reliance on traditional diagnostic procedures, which involve time-consuming blood tests and expensive laboratory investigations.

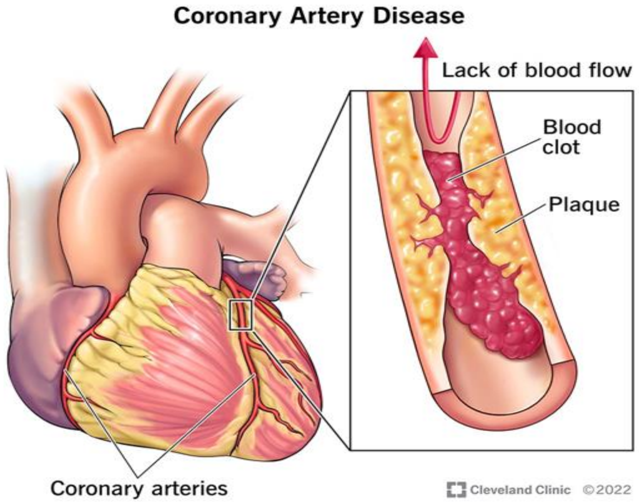

Figure 1: Coronary Artery Disease

Traditional machine learning models often suffer from issues such as overfitting, underfitting, and computational inefficiency. The accuracy of these models is highly dependent on the quality and volume of training data provided. Further research highlights the impact of feature selection techniques like Principal Component Analysis (PCA) and Firefly optimization, demonstrating an improvement of 20–30% in recall, precision, and accuracy, with the highest model accuracy reaching 73%. Other studies have explored cross-validation techniques and optimization algorithms like GA and PSO, achieving 93.08% accuracy with nu-SVM combined with GA and 66.80% accuracy using Random Forest (Shahid & Singh, 2020). Moreover, kernel extreme learning machines (KELM) optimized with an improved salp swarm algorithm (SSA) have been proposed, achieving 84.40% accuracy. Researchers have also developed models using datasets such as the Framingham Heart Study, applying methods like Recursive Feature Elimination (RFE) and Boruta feature selection to improve accuracy, with RF achieving 89% accuracy and SVM reaching 76%. Additional studies have examined the Coronary Artery Classification Score (CACS) with machine learning models, with RF obtaining 78.96% accuracy (Ozcan & Peker, 2023).

Logistic regression has been widely used in cardiovascular research due to its ability to handle binary outcomes effectively. A study by Khosravi et al. (2019) demonstrated that logistic regression could accurately predict CAD using clinical data from patients undergoing coronary angiography (Masih et al., 2021). The model identified key predictors such as age, sex, cholesterol levels, and smoking status with high accuracy. Moreover, machine learning techniques have gained traction in recent years for predicting CAD risk. However, traditional statistical methods like logistic regression remain relevant due to their interpretability and ease of implementation in clinical settings (Bashir et al., 2020). The simplicity of logistic regression allows healthcare professionals to understand the influence of individual risk factors on CAD probability without requiring extensive computational resources. The neural network-based approach leverages Artificial Neural Networks (ANNs), Genetic Algorithms (GA), and Particle Swarm Optimization (PSO) to optimize CHD classification. Researchers have investigated feature selection techniques, such as Fisher Score, Relief-F, and Minimum Redundancy Maximum Relevance (mRMR), to enhance ANN efficiency. Studies have demonstrated that PSO combined with an Adaptive Neuro-Fuzzy Inference System (ANFIS) achieves 70% accuracy, while PSO with an Emotional Neural Network (ENN) reaches 86.33% accuracy. Further optimization using GA-based feature selection has been applied to the Framingham Heart Study dataset, improving accuracy from 70.7% to 71.6%. (Masih & Ahuja, 2018) Additionally, deep learning models, such as Convolutional Neural Networks (CNNs), have been utilized for CHD classification, with LASSO regression methods employed to eliminate multicollinearity and improve prediction accuracy, reaching 72.32%. Other studies have compared traditional ML models with ANNs, where ANN models achieved superior accuracy, with MLP reaching 66.50% accuracy after feature selection. Finally, researchers have explored the use of Sparse Autoencoders (SAE) combined with batch normalization and the Adam optimizer, resulting in 70% accuracy for CHD prediction (Muhammad & Algehyne, 2021).

A critical gap in existing research is the lack of optimization for neural network architectures using different optimizers and tuning parameters, such as learning rate, weight initialization, number of neurons per layer, and number of hidden layers. The absence of systematic tuning for these hyperparameters limits the performance and generalization of ANN models in CHD classification. Addressing this gap can further improve performance metrics, particularly accuracy and loss function minimization (Kumar et al., 2020). This study developed and evaluated a Logistic Regression model for early CAD prediction, incorporating data preprocessing, feature engineering, and rigorous performance evaluation. By addressing the limitations of basic models, we seek to create a more robust, (G. Oise, 2023) and reliable tool for identifying individuals at high risk of developing CAD, enabling timely interventions and potentially reducing the burden of this prevalent disease. Recent studies have highlighted the importance of early prediction in managing CAD

MATERIALS AND METHODS

This project predicted heart disease using logistic regression. Based on attributes such as blood pressure, cholesterol levels, heart rate, and other characteristics, patients will be classified according to varying degrees of coronary artery disease. This project utilized a dataset of 303 patients distributed by the UCI Machine Learning Repository. Furthermore, for the machine learning side of this project, we will be using sklearn and keras. Import these libraries using the cell below to ensure you have them correctly installed

Data Collection

This study utilized a dataset comprising clinical records from patients diagnosed with CAD at a tertiary care hospital (Cleveland Clinic, 2023).

The dataset included demographic information (age, sex), clinical measurements (blood pressure, cholesterol levels), lifestyle factors (smoking status, physical activity), and medical history (diabetes, family history of heart disease). Ethical approval was obtained from the hospital's institutional review board. The model was trained using key predictors identified in previous studies: age, sex, total cholesterol levels, HDL cholesterol levels, blood pressure readings, smoking status, and diabetes history. The logistic regression model will be used in this project because it is a binary classification(Iornem et al., 2023).

Data Preprocessing

Continuous variables were standardized to ensure comparability across different scales. Categorical variables were encoded using one-hot encoding to facilitate their inclusion in the logistic regression model. To effectively handle the missing values in the dataset, regression imputation was used.

Table 1: First five rows of the dataset

```
# print first 5 rows of the dataset
heart_data.head()
```

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal
0	63	1	3	145	233	1	0	150	0	2.3	0	0	1
1	37	1	2	130	250	0	1	187	0	3.5	0	0	2
2	41	0	1	130	204	0	0	172	0	1.4	2	0	2
3	56	1	1	120	236	0	1	178	0	0.8	2	0	2
4	57	0	0	120	354	0	1	163	1	0.6	2	0	2

Table 2: Last five rows of the dataset

```
# print last 5 rows of the dataset
heart_data.tail()
```

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal
298	57	0	0	140	241	0	1	123	1	0.2	1	0	3
299	45	1	3	110	264	0	1	132	0	1.2	1	0	3
300	68	1	0	144	193	1	1	141	0	3.4	1	2	3
301	57	1	0	130	131	0	1	115	1	1.2	1	1	3
302	57	0	1	130	236	0	0	174	0	0.0	1	1	2

Logistic Regression Model

Logistic Regression is a widely used statistical method for binary classification. It is a straightforward and interpretable approach to estimating the probability of a binary outcome based on various predictor variables. In the context of this document, it is used to predict the probability of Coronary Artery Disease (CAD) presence.

Mathematical Equations

The core of Logistic Regression is the logistic function, also known as the sigmoid function.

Sigmoid Function:

$P(Y=1) = 1 / (1 + e^{(-z)})$

Where $P(Y=1)$ is the probability of the outcome Y being 1 (presence of CAD in this case), e is the base of the natural logarithm, z is the linear combination of the predictor variables.

Linear Combination (z):

$z = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_n X_n$

Where β_0 is the intercept, $\beta_1, \beta_2, ..., \beta_n$ are the coefficients of the predictor variables, $X_1, X_2, ..., X_n$ are the predictor variables (e.g., age, cholesterol levels, blood pressure).

The Logistic Regression model calculates the probability of a binary outcome (presence or absence of CAD) by using the sigmoid function on a linear combination of predictor variables. The sigmoid function restricts the output probability to a range between 0 and 1, which makes it suitable for binary classification. The coefficients ($\beta_1, \beta_2, ..., \beta_n$) in the linear combination indicate the impact of each predictor variable on the log odds of the outcome. Logistic Regression is interpretable, enabling healthcare providers to understand how individual risk factors contribute to the overall risk of CAD.

A logistic regression model was developed using Python's scikit-learn library. The dataset was split into training (70%) and testing (30%) subsets to evaluate model performance

Model Result

Model performance was assessed using accuracy, precision, recall (sensitivity), F1-score, and the area under the Receiver Operating Characteristic curve (AUC-ROC). A confusion matrix was generated to visualize true positive and false positive rates.

RESULTS AND DISCUSSION

This classification report shows a model's performance on a test set of 61 instances, divided into two classes (28 in class 0 and 33 in class 1). The model achieved an overall accuracy of 82%. Class 0 had a precision of 79%, a recall of 82%, and an F1 score of 81%. For class 1, the precision was 84%, the recall was 82%, and the F1 score was 83%. Both macro and weighted averages for precision, recall, and F1-score were also 82%, indicating balanced performance across both classes. The results suggest the model is reasonably effective at classifying instances in both classes, with slightly better performance in class 1.

Table 3: Classification Report

```
# Calculate and print the classification report
print(classification_report(Y_test, X_test_prediction))
```

	precision	recall	f1-score	support
0	0.79	0.82	0.81	28
1	0.84	0.82	0.83	33
accuracy			0.82	61
macro avg	0.82	0.82	0.82	61
weighted avg	0.82	0.82	0.82	61

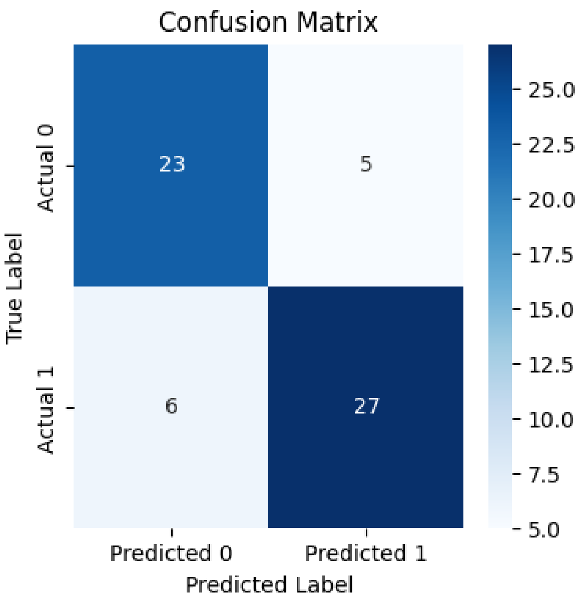

Figure 2: Confusion Matrix

The confusion matrix reveals a classification model's performance with 23 true negatives (correctly predicted class 0), 27 true positives (correctly predicted class 1), 5 false positives (incorrectly predicted class 1), and 6 false negatives (incorrectly predicted class 0). This translates to an overall accuracy of approximately 82%. The model exhibits balanced performance with a precision of 84%, recall of 82% for class 1, and specificity of 82% for class 0. While the model demonstrates reasonable performance.

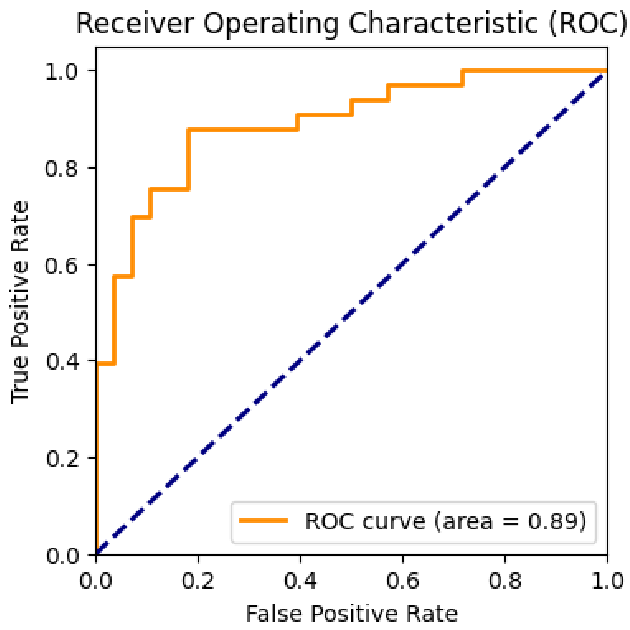

Figure 3: The Receiver Operating Characteristics

The Receiver Operating Characteristic (ROC) curve illustrates the performance of a binary classification model. With an Area Under the Curve (AUC) of 0.89, the model demonstrates strong discriminatory ability, significantly outperforming a random classifier. The curve shows the trade-off between the True Positive Rate and False Positive Rate as the classification threshold changes. The high AUC value indicates that the model is likely effective at distinguishing between the two classes, suggesting good overall performance.

Table 4: Input and Target Data

Age	Sex	Cp	Trestbps	Chol	Fbs	Restecg	Thalach	Exang	Oldpeak	Slope	Ca	Thal	Target
63	1	3	145	233	1	0	150	0	2.3	1	0	1	1
37	1	2	130	250	0	1	187	0	3.5	0	0	2	1
41	0	1	130	204	0	1	172	0	1.4	2	0	2	1
56	1	1	120	236	0	1	178	0	0.8	2	0	2	1
57	1	1	120	354	0	1	163	1	0.6	2	0	2	1
67	1	0	160	286	0	0	108	1	1.5	1	3	2	0
67	1	0	120	229	0	0	129	1	2.6	1	2	3	0
62	0	0	140	268	0	0	160	0	3.6	0	2	2	0
63	1	0	130	254	0	0	147	1	1.4	1	1	3	0
53	1	0	140	203	1	0	155	1	3.1	0	0	3	0

Table 4, depicts the values of input data and target data to be used to make predictions. Columns 1 to 13 contain a set of numerical features of patient medical data and the last column is the target data. If the prediction is 1, it means the patient has coronary artery disease and 0, means the patient does not have coronary artery disease.

Make a prediction

```
[32] input_data = (55,0,1,132,342,0,1,166,0,1.2,2,0,2)

     # change the input data to a numpy array
     input_data_as_numpy_array= np.asarray(input_data)

     # reshape the numpy array as we are predicting for only on instance
     input_data_reshaped = input_data_as_numpy_array.reshape(1,-1)

     prediction = model.predict(input_data_reshaped)
     print(prediction)

     if (prediction[0]== 0):
       print('The Person does not have a Coronary Artery Disease ')
     else:
       print('The Person has Coronary Artery Disease ')

     [1]
     The Person has Coronary Artery Disease
```

Figure 4: Prediction of a person with coronary artery disease

Figure 4 shows a set of patient medical data, represented by 13 numerical features, is used to predict coronary artery disease. The data is converted to a NumPy array and reshaped to a format suitable for a machine-learning model. The model.predict () function generates a prediction of [1], indicating the presence of coronary artery disease, which is then printed to the console. This prediction signifies that, based on the provided data and the trained model, the patient is likely to have the disease

Make a prediction

```
[25] input_data = (67,1,0,160,286,0,0,108,1,1.5,1,3,2)

     # change the input data to a numpy array
     input_data_as_numpy_array= np.asarray(input_data)

     # reshape the numpy array as we are predicting for only on instance
     input_data_reshaped = input_data_as_numpy_array.reshape(1,-1)

     prediction = model.predict(input_data_reshaped)
     print(prediction)

     if (prediction[0]== 0):
       print('The Person does not have a Coronary Artery Disease ')
     else:
       print('The Person has Coronary Artery Disease ')

     [0]
     The Person does not have a Coronary Artery Disease
```

Figure 5: Prediction of a person without coronary artery disease

Figure 5 shows patients' medical information, consisting of thirteen numerical features, is input to a pre-trained machine learning model for coronary artery disease prediction. The data is converted to a NumPy array, reshaped to a single-sample format, and used to generate a prediction. If the model predicts [0], it indicates the absence of coronary artery disease, which is then printed to the console

Discussion

The Logistic Regression model showcased promising results in predicting Coronary Artery Disease (CAD), achieving an overall accuracy of 82%. The model's balanced precision and recall scores (around 82% for both classes) indicate its effectiveness in minimizing both false positives and false negatives (G. Oise & Konyeha, 2024), which is crucial in medical diagnosis. Furthermore, the strong discriminatory power, highlighted by the AUC-ROC of 0.89, signifies the model's ability to effectively distinguish between individuals at high and low risk of CAD. These results suggest that the model can be a valuable tool for early risk stratification, improved clinical decision-making, efficient resource allocation, and personalized medicine. Specifically, the model can be applied as a screening tool, for risk assessment in asymptomatic individuals (Mondal et al., 2025), as support for diagnostic workups, and for monitoring treatment effectiveness. However, it's important to acknowledge the limitations, including those related to dataset size and potential class imbalance. Future research should focus on validating the model on larger, more diverse datasets, exploring additional feature engineering, investigating advanced modeling techniques, and conducting rigorous clinical validation studies to ensure its robustness and clinical utility.

The classification report shows a model's performance on a test set of 61 instances, divided into two classes (28 in class 0 and 33 in class 1). The model achieved an overall accuracy of 82%. For Class 0, it had a precision of 79%, a recall of 82%, and an F1-score of 81%. For class 1, the precision was 84%, the recall was 82%, and the F1 score was 83%. Both macro and weighted averages for precision, recall, and F1-score were also 82%, indicating balanced performance across both classes. The results suggest the model is reasonably effective at classifying instances in both classes, with slightly better performance in class 1. The confusion matrix reveals a classification model's performance with 23 true negatives (correctly predicted class 0), 27 true positives (correctly predicted class 1), 5 false positives (incorrectly predicted class 1), and 6 false negatives (incorrectly predicted class 0). This translates to an overall accuracy of approximately 82%. The model exhibits balanced performance with a precision of 84%, recall of 82% for class 1, and specificity of 82% for class 0. While the model demonstrates reasonable performance, further analysis considering potential class imbalance and the specific application context is recommended for a complete evaluation. The Receiver Operating Characteristic (ROC) curve illustrates the performance of a binary classification model. With an Area Under the Curve (AUC) of 0.89, the model demonstrates strong discriminatory ability, significantly outperforming a random classifier. The curve shows the trade-off between the True Positive Rate and False Positive Rate as the classification threshold changes (G. P. Oise & Susan, 2024). The high AUC value indicates that the model is likely effective at distinguishing between the two classes, suggesting good overall performance. A set of patient medical data, represented by 13 numerical features, is used to predict coronary artery disease. The data is converted to a NumPy array and reshaped to a format suitable for a machine-learning model. The model. predict () function generates a prediction of [1] or [0], indicating the presence of coronary artery disease or not, which is then printed to the console. This prediction signifies that, based on the provided data and the trained model, the patient is likely to have coronary artery disease or not

CONCLUSION

This study explored Logistic Regression for Coronary Artery Disease (CAD) prediction using a readily available dataset and clinical attributes. The model showed promise with an overall accuracy of 82% and an AUC of 0.89. While effective for CAD detection and discrimination, it's important to consider the broader research landscape. Other studies have achieved higher accuracy using techniques like nu-SVM with GA (93.08%) and Random Forest (up to 89%). Despite this, Logistic Regression offers a balance of accuracy and interpretability valuable in clinical settings due to its simplicity and ease of understanding individual risk factors. This study provides a foundation for further research, emphasizing the potential of Logistic Regression in CAD risk assessment and the need for continued work in data preprocessing, feature engineering, and advanced machine learning techniques to improve performance and clinical utility.

REFERENCES

Abdar, M., Książek, W., Acharya, U. R., Tan, R.-S., Makarenkov, V., & Pławiak, P. (2019). A new machine learning technique for an accurate diagnosis of coronary artery disease. *Computer Methods and Programs in Biomedicine*, *179*, 104992. https://doi.org/10.1016/j.cmpb.2019.104992

Aminu Bashir Suleiman, Stephen Luka and Muhammad Ibrahim. (2023). CARDIOVASCULAR DISEASE PREDICTION USING RANDOM FOREST MACHINE LEARNING ALGORITHM. *FUDMA Journal of Sciences (FJS)*, *7*(6), 282–289. https://doi.org/10.33003/fjs-2023-0706-2128

Beunza, J.-J., Puertas, E., García-Ovejero, E., Villalba, G., Condes, E., Koleva, G., Hurtado, C., & Landecho, M. F. (2019). Comparison of machine learning algorithms for clinical event prediction (risk of coronary heart disease). *Journal of Biomedical Informatics*, *97*, 103257. https://doi.org/10.1016/j.jbi.2019.103257

Cleveland Clinic. (2023, October 13). *Coronary Artery Disease*. Cleveland Clinic. https://my.clevelandclinic.org/health/diseases/16898-coronary-artery-disease

Das, R., Turkoglu, I., & Sengur, A. (2009). Effective diagnosis of heart disease through neural networks ensembles. *Expert Systems with Applications*, *36*(4), 7675–7680. https://doi.org/10.1016/j.eswa.2008.09.013

Doppala, B. P., Bhattacharyya, D., Janarthanan, M., & Baik, N. (2022). A Reliable Machine Intelligence Model for Accurate Identification of Cardiovascular Diseases Using Ensemble Techniques. *Journal of Healthcare Engineering*, *2022*, 1–13. https://doi.org/10.1155/2022/2585235

Iornem, T. V., Abdulkardir, S. S., Akinrefon, A. A., & Ornguga, I. G. (2023). MODELLING OF ACCESS TO MOSQUITO TREATED NET IN NIGERIA: USING MULTILEVEL LOGISTIC REGRESSION APPROACH. *FUDMA JOURNAL OF SCIENCES*, *7*(3), 144–149. https://doi.org/10.33003/fjs-2023-0703-1831

Jamilu, A. A., Ayeni, B. K., Ahmad, B. I., & Sulaiman, N. (2019). *CLASSIFICATION OF CORONARY ARTERY DISEASE USING HYBRID APPROACH*. *3*(4).

Masih, N., & Ahuja, S. (2018). Prediction of Heart Diseases Using Data Mining Techniques: Application on Framingham Heart Study. *International Journal of Big Data and Analytics in Healthcare*, *3*(2), 1–9. https://doi.org/10.4018/IJBDAH.2018070101

Masih, N., Naz, H., & Ahuja, S. (2021). Multilayer perceptron based deep neural network for early detection of coronary heart disease. *Health and Technology*, *11*(1), 127–138. https://doi.org/10.1007/s12553-020-00509-3

Mohi Uddin, K. M., Ripa, R., Yeasmin, N., Biswas, N., & Dey, S. K. (2023). Machine learning-based approach to the diagnosis of cardiovascular vascular disease using a combined dataset. *Intelligence-Based Medicine*, *7*, 100100. https://doi.org/10.1016/j.ibmed.2023.100100

Mondal, S., Maity, R., & Nag, A. (2025). An efficient artificial neural network-based optimization techniques for the early prediction of coronary heart disease: Comprehensive analysis. *Scientific Reports*, *15*(1), 4827. https://doi.org/10.1038/s41598-025-85765-x

Muhammad, L. J., & Algehyne, E. A. (2021). Fuzzy based expert system for diagnosis of coronary artery disease in nigeria. *Health and Technology*, *11*(2), 319–329. https://doi.org/10.1007/s12553-021-00531-z

Oise, G. (2023). A Decision Support System for Decongesting Prison in Edo State. *RADINKA JOURNAL OF SCIENCE AND SYSTEMATIC LITERATURE REVIEW*, *1*(2), 69–79. https://doi.org/10.56778/rjslr.v1i2.114

Oise, G., & Konyeha, S. (2024). E-WASTE MANAGEMENT THROUGH DEEP LEARNING: A SEQUENTIAL NEURAL NETWORK APPROACH. *FUDMA JOURNAL OF SCIENCES*, *8*(3), 17–24. https://doi.org/10.33003/fjs-2024-0804-2579

Oise, G. P., & Susan, K. (2024). *Deep Learning for Effective Electronic Waste Management and Environmental Health*. https://doi.org/10.21203/rs.3.rs-4903136/v1

Ozcan, M., & Peker, S. (2023). A classification and regression tree algorithm for heart disease modeling and prediction. *Healthcare Analytics*, *3*, 100130. https://doi.org/10.1016/j.health.2022.100130

Shahid, A. H., & Singh, M. P. (2020). A Novel Approach for Coronary Artery Disease Diagnosis using Hybrid Particle Swarm Optimization based Emotional Neural Network. *Biocybernetics and Biomedical Engineering*, *40*(4), 1568–1585. https://doi.org/10.1016/j.bbe.2020.09.005

Tama, B. A., Im, S., & Lee, S. (2020). Improving an Intelligent Detection System for Coronary Heart Disease Using a Two-Tier Classifier Ensemble. *BioMed Research International*, *2020*(1), 9816142. https://doi.org/10.1155/2020/9816142