

Using Machine Learning to Predict Heart Failure: A Comparative Analysis of Various Classification Algorithms

Isaac Osei¹, Acheampong Baafi-Adomako²

¹Amity University

²University of Ghana

DOI: <https://doi.org/10.51244/IJRSI.2024.1101026>

Received: 29 December 2023; Revised: 10 January 2024; Accepted: 13 January 2024; Published: 14 February 2024

ABSTRACT

This study presents a comprehensive analysis of machine learning algorithms for predicting heart failure, a significant cause of morbidity and mortality worldwide. Employing a robust dataset of patient records with diverse clinical features, the performance of several widely-used classification algorithms, including K-Nearest Neighbours (KNN), Decision Trees, Support Vector Machines (SVM), Random Forests (RF), and Gaussian Naive Bayes etc. are systematically evaluated and compared. The methodology encompasses loading the dataset, data pre-processing, feature selection, model training, and validation. The performance of each algorithm is rigorously assessed using metrics such as accuracy, precision, recall, F1-score, and the area under the ROC curve with Random Forest outperforming the other classification algorithms. The study further investigates the influence of hyper-parameter tuning and cross-validation on model efficacy. Additionally, an interpretative analysis of the models, offering insights into feature importance and the clinical relevance of the prediction outcomes were provided. The findings aim to contribute to the field of medical informatics by identifying the most effective machine learning strategies for heart failure prediction, thereby facilitating early intervention and personalized patient care. This research not only underscores the potential of machine learning in healthcare but also highlights the challenges and considerations in developing reliable predictive models for complex medical conditions.

Keywords: Machine Learning, Classification, Heart Failure, Classification Algorithm, Random Forest, Support Vector Machine.

INTRODUCTION

Globally, cardiovascular diseases (CVDs) stand as the leading cause of death, claiming approximately 17.9 million lives annually and accounting for 31% of all global deaths. Heart attacks and strokes are responsible for 80% of these CVD fatalities, with a third occurring prematurely in individuals below 70 years of age. Heart failure, a frequent consequence of CVDs, is potentially predictable using a dataset with 11 distinct features. Early detection and management are crucial for individuals with cardiovascular disease or those at elevated risk due to factors like hypertension, diabetes, or existing conditions. In such scenarios, machine learning models can play a pivotal role in facilitating timely and effective interventions.

Problem Statement

Cardiovascular diseases, including heart failure, remain a leading cause of mortality globally, presenting a significant challenge for early diagnosis and effective treatment. Traditional diagnostic methods, while

effective, often rely on a combination of clinical assessments and patient history, which may not always detect heart failure at an early stage. The emergence of machine learning (ML) presents a novel opportunity to enhance the predictive accuracy and timeliness of heart failure diagnosis. The primary problem addressed in this research is the identification and evaluation of the most effective ML classification algorithms for predicting heart failure. This involves analyzing various algorithms' ability to accurately classify and predict heart failure based on a set of clinical and patient features.

Objectives

The global objective of this research, is to create an effective and best machine learning model to predict the heart failure using clinical dataset. This is achieved by creating several models and comparing their accuracies. The specific objectives are as follow:

- To predict whether a patient will suffer heart failure or not
- To determine at what age can one be affected by heart failure
- To determine which gender or sex are being affected most by heart failure
- To determine whether some factors like chest pain, cholesterol, blood pressure, blood sugar, heart rate, etc. can contribute to heart failure
- To discover other hidden knowledge or insight in the clinical dataset use

LITERATURE REVIEW

Review of Related Works

The table below depicts some of the related works.

Table I. A Summary Review of Related Works

#	Year	Author	Title	Classifier	Accuracy
1	2019	Mohan, Thirumalai & Srivastava	Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques	HRFLM	88.7%
2	2021	Pandita et al.	Prediction of Heart Disease using Machine Learning Algorithms	LR KNN SVM Naïve Bayes RF	84.4% 89.1% 87.5% 85.9% 87.5%
3	2020	Nikhar & Karandikar	Prediction of heart disease using machine learning algorithms	DT Naïve Bayes	N/A N/A

4	2019	Awan et al.	Machine learning-based prediction of heart failure readmission or death: implications of choosing the right model and the right metrics	LR RF SVM MLP	62.3% 76.4% 71.8% 64.9%
5	2019	Lutimath, Chethan & Pol	Prediction of heart disease using machine learning	Naïve Bayes SVM	N/A N/A
6	2019	Alotaibi	Implementation of machine learning model to predict heart failure disease	DT LR RF Naïve Bayes SVM	93.2% 87.4% 89.1% 87.3% 92.3%
7	2020	Adler et al.	Improving risk prediction in heart failure using machine learning	MARKER-HF	N/A
8	2021	Padmaja et al.	Early and Accurate Prediction of Heart Disease Using Machine Learning Model	Random Forest LR KNN SVM Decision Tree Gradient Boosting Naïve Bayes	93.4% 88.5% 91.8% 83.6% 86.9% 85.3% 85.3%
9	2020	Princy et al.	Prediction of Cardiac Disease using Supervised Machine Learning Algorithms	Naïve Bayes Decision Tree LR Random Forest SVM KNN	60% 73% 72% 71% 72% 66%

10	2020	Choudhary & Singh	Prediction of Heart Disease using Machine Learning Algorithms	Decision Tree AdaBoost	97.1% 89.9%
11	2017	Sharma & Rizvi	Prediction of Heart Disease using Machine Learning Algorithms: A Survey	N/A	N/A
12	2018	Haq et al.	A hybrid intelligent system framework for the prediction of heart disease using machine learning algorithms	LR KNN ANN SVM Naïve Bayes Decision Tree Random Forest	84% 76% 74% 86% 83% 74% 83%
13	2020	Olsen et al.	Clinical applications of machine learning in the diagnosis, classification, and prediction of heart failure: Machine learning in heart failure	N/A	N/A
14	2021	Bharti et al.	Prediction of Heart Disease Using a Combination of Machine Learning and Deep Learning	LR K Neighbors SVM Random Forest Decision Tree Deep Learning	83.3% 84.8% 83.2% 80.3% 82.3% 94.2%
15	2019	Bhardwaj et al.	Prediction of Heart Attack Using Machine Learning	LR	N/A

Conclusion

From the review of the related works, most of the accuracies are around 80% to 90%. Some used few classifiers. Others also used good number of classifiers. Only two of them included Artificial Neural Networks (ANN) or Multi-Layer Perceptron (MLP) but with very low accuracies (Haq et al., 74% and Awan et al., 64.93%). None of them used parametric or non-parametric statistical test on the features in order to test for causality. This is where a research gap was discovered.

RESEARCH METHODOLOGY

The task of classification is employed to predict future instances using historical data. In past research,

experts have applied numerous data mining methods, including clustering and classification techniques to accurately diagnose heart diseases and stroke. The researcher utilized various machine learning algorithms (classifiers), including K-Nearest Neighbors (KNN), Decision Trees, Gaussian Naïve Bayes, Support Vector Machine (SVM), Logistic Regression, Multi-Layer Perceptron (ANN), Gradient Boosting, and Random Forest (RF). To assist physicians and their patients in the medical field, this research aims to predict the probability of getting stroke. The researcher has used statistical analysis especially parametric and non-parametric testing on the various features to confirm whether they likely to cause heart failure. The researcher finally talked about using several machine learning algorithms on clinical dataset.

Data Source

The dataset used was downloaded from Kaggle (uploaded by fedesoriano, Kaggle Datasets Grandmaster). This involves an authentic dataset of size 36 KB comprising 918 data instances or observations, encompassing 12 distinct features (11 predictive features and 1 class). Age, Sex, Chest Pain Type, Resting BP, Cholesterol etc. (Fig. 1).

	Age	Sex	ChestPainType	RestingBP	Cholesterol	FastingBS	RestingECG	MaxHR	ExerciseAngina	Oldpeak	ST_Slope	HeartDisease
0	40	M	ATA	140	289	0	Normal	172	N	0.0	Up	0
1	49	F	NAP	160	180	0	Normal	156	N	1.0	Flat	1
2	37	M	ATA	130	283	0	ST	98	N	0.0	Up	0
3	48	F	ASY	138	214	0	Normal	108	Y	1.5	Flat	1
4	54	M	NAP	150	195	0	Normal	122	N	0.0	Up	0

Fig. 1. Attributes and details of the dataset

Feature Description

Some features (Age, RestingBP, Cholesterol, MaxHR, and Oldpeak) have numerical data values while the remaining (Sex, ChestPainType, FastingBS, ExerciseAngina, ST_Slope, and HeartDisease, which is the class) features have categorical values. The table below (TABLE 2) depicts a short description of the various features in the dataset.

Table II. A Short Description Of The Features In The Dataset

#	Feature	Description
1	Age	Age of the person in years [28 – 77]
2	Sex	Sex of the person [M: Male, F: Female]
3	Chest Pain Type	Chest pain type [TA: Typical Angina, ATA: Atypical Angina, NAP: Non-Anginal Pain, ASY: Asymptomatic]
4	RestingBP	Resting blood pressure [mm Hg]
5	Cholesterol	Serum cholesterol [mm/dl]
6	Fasting BS	Fasting blood sugar [1: If FastingBS > 120 mg/dl, 0: Otherwise]
7	Resting ECG	Resting electrocardiogram results [Normal: Normal, ST: having ST-T wave abnormality (T wave inversions and/or ST elevation or depression of > 0.05 mV), LVH: showing probable or definite left ventricular hypertrophy by Estes’ criteria]
8	Max HR	Maximum heart rate achieved [60 – 202]
9	Exercise Angina	Exercise-induced angina [Y: Yes, N: No]
10	Oldpeak	Oldpeak (ST measured in depression) [-2.6 – 6.2]

11	ST_Slope	The slope of the peak exercise ST segment [Up: up sloping, Flat: flat, Down: down sloping]
12	Heart Disease	Class or target attribute [1: Heart disease, 0: No heart disease]

Process Model (Working Process)

The dataset is loaded and pre-processed. It was then analyzed to discover hidden patterns and insights. The dataset was then split into two namely; training and testing sets with the ratio of 4:1. Eighty percent (80%) for training and twenty percent (20%) for testing. The training set was used for the training of the various models (classifiers) whiles the testing set was used to evaluate the various models created so the best one could be chosen. The diagram below (Fig. 2) depicts the process flow of the proposed model.

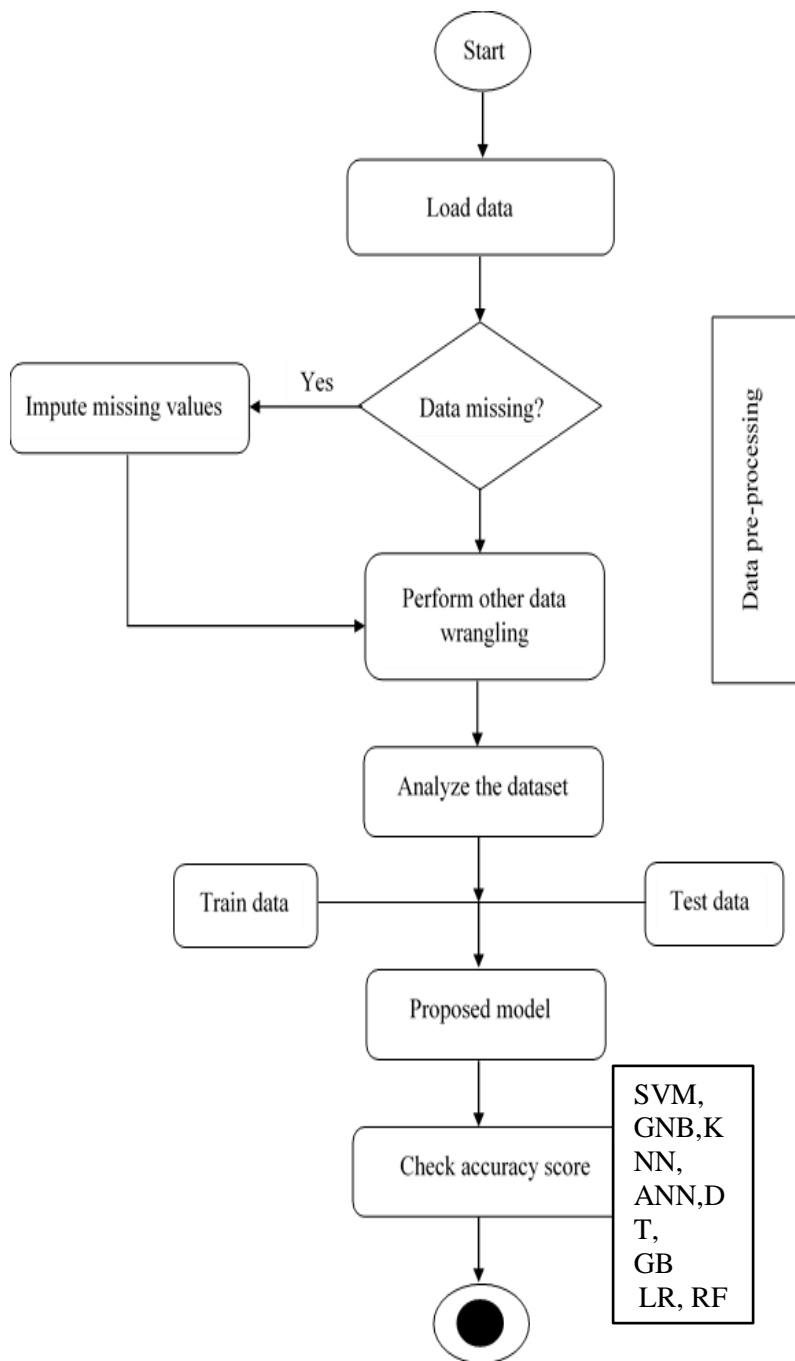


Fig. 2. Flow chart (Process flow) of the proposed model

Data Pre-processing

Data from the real-world do not come in the format suitable for machine learning. This may include noise and missing data. To get around these problems and create strong predictions, these data have to be processed. For this reason, the dataset was processed very well. The activities in the processing included; data cleaning, transformation, normalization, handling imbalance data, etc.

Data cleaning typically includes looking for noise, fake data, duplicated values, and missing values. These data need to be cleansed of noise and the missing values filled in, in order to produce an accurate and useful result. Fortunately, the dataset had no missing value or noise.

Transformation is the process of converting data from one format to another so that it is easier to understand. It entails duties for aggregation, data type casting, encoding, and smoothing. All numeric and categorical features or variables were converted to their appropriate data types and forms. Again, all categorical variables were encoded.

Normalization involves scaling of numerical features to bring them to specific range or distribution. Min-max scaling (normalization) was used on all the numerical features.

Dimensionality Reduction (Data Reduction) involves dropping or taking out of unwanted or less related features. Here no feature was dropped.

Handling Imbalance Data entails balancing the data distribution to prevent biases during analysis and modelling. The data was fairly balanced with 501 observations being patients with heart failure and 410 without heart failure. Therefore, there was no need to balance the dataset.

Metrics for Evaluating Machine Learning Models

1. *Confusion Matrix*: A Confusion Matrix is a straightforward method to assess the effectiveness of a classification model. It does this by comparing the number of positive and negative instances that were accurately or inaccurately classified (Azam, Habibullah, & Rana, 2020).

Table III. Confusion Matrix

	Predicted Positive	Predicted Negative
Actual Positive	TP	FN
Actual Negative	FP	TN

True Positives (TP):

True positives are the instances where both the predicted class and actual class are true.

True Negatives (TN):

True negatives are the instances where both the predicted class and actual class are false.

False Negatives (FN):

False negatives are the instances where the predicted class is false (0) but actual class is true (1).

False Positives (FP):

False positives are the instances where the predicted class is true (1) while actual class is false (0).

From the confusion matrix, metrics like accuracy, precision, recall, and f1-score can be calculated with the following formulae.

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN} \tag{3.1}$$

Where TP = True Positive, TN = True Negative,

FP = False Positive, and FN = False Negative

$$\text{Precision} = \frac{TP}{TP+FP} \tag{3.2}$$

$$\text{Recall} = \frac{TP}{TP+FN} \tag{3.3}$$

$$F1 = 2 * (\text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall}) \tag{3.4}$$

2. *Area under Curve*: The Area under Curve (AUC) is an effective metric with values ranging from 0 to 1. The closer it is to 1, the more capable the machine learning model is at differentiating between heart failure and non-heart failure cases. A model that perfectly discriminates between the two classes has an AUC of 1. Conversely, if all non-heart failure instances are incorrectly classified as heart failure and vice versa, the AUC is 0 (Dritsas & Trigka, 2022).

Deployment of the Proposed Model

With the help of Flask framework, HTML, and CSS, the model was deployed in a web based. The figures below show the respective interfaces.

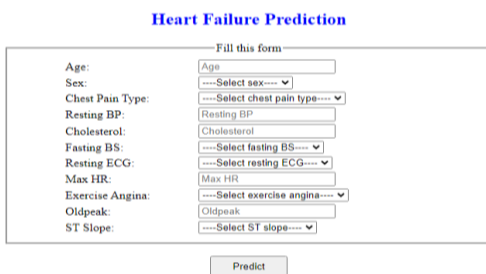


Fig. 2. Homepage of the web application

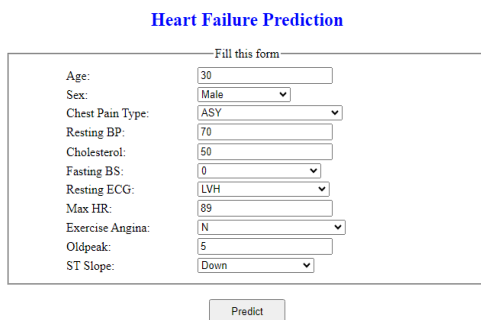


Fig. 3. Prediction phase

Heart Failure Prediction

Heart Failure Predicted

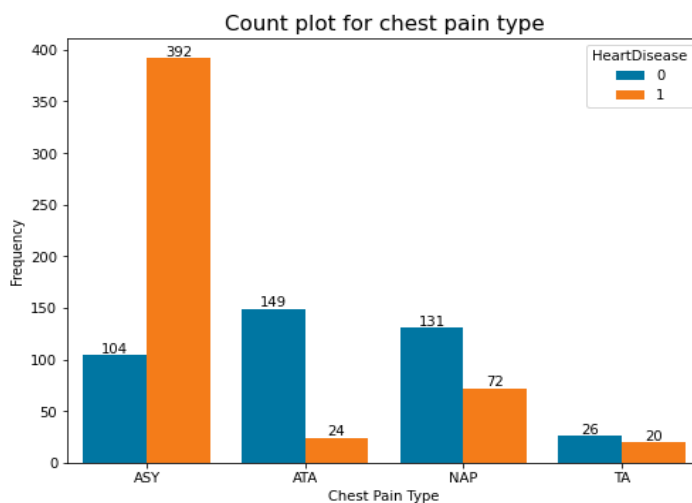
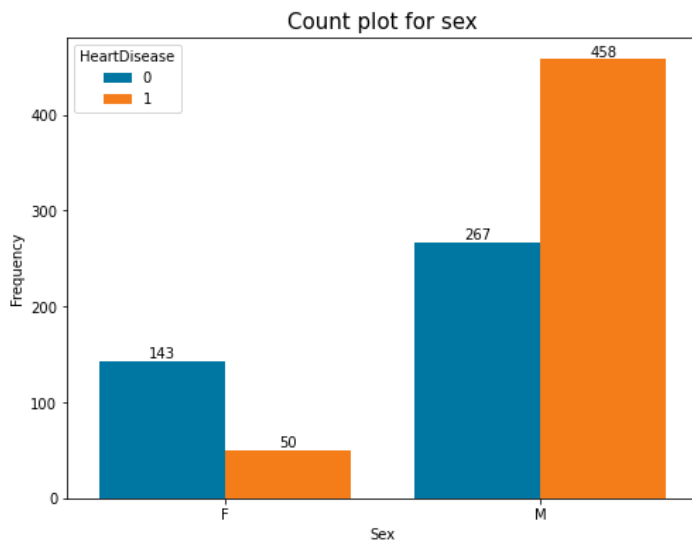
Fill this form

Age:	<input type="text" value="Age"/>
Sex:	<input type="text" value="---Select sex---"/>
Chest Pain Type:	<input type="text" value="---Select chest pain type---"/>
Resting BP:	<input type="text" value="Resting BP"/>
Cholesterol:	<input type="text" value="Cholesterol"/>
Fasting BS:	<input type="text" value="---Select fasting BS---"/>
Resting ECG:	<input type="text" value="---Select resting ECG---"/>
Max HR:	<input type="text" value="Max HR"/>
Exercise Angina:	<input type="text" value="---Select exercise angina---"/>
Oldpeak:	<input type="text" value="Oldpeak"/>
ST Slope:	<input type="text" value="---Select ST slope---"/>

Fig. 4. Results or output phase

DATA ANALYSIS AND INTERPRETATION

• Exploratory Data Analysis (EDA)



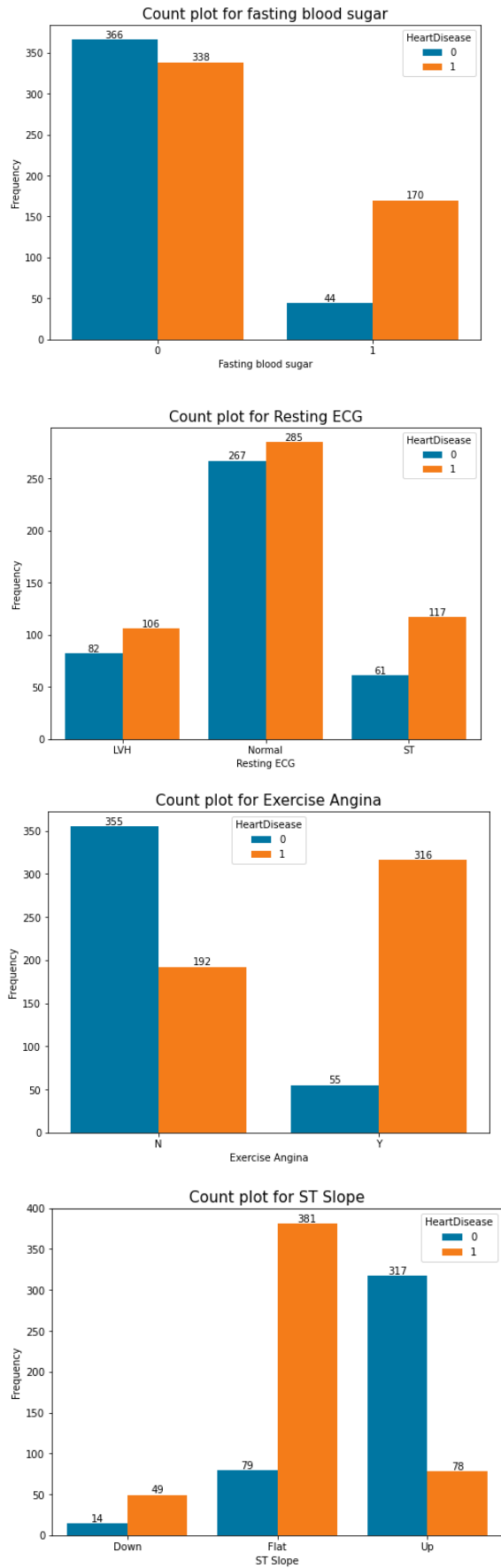
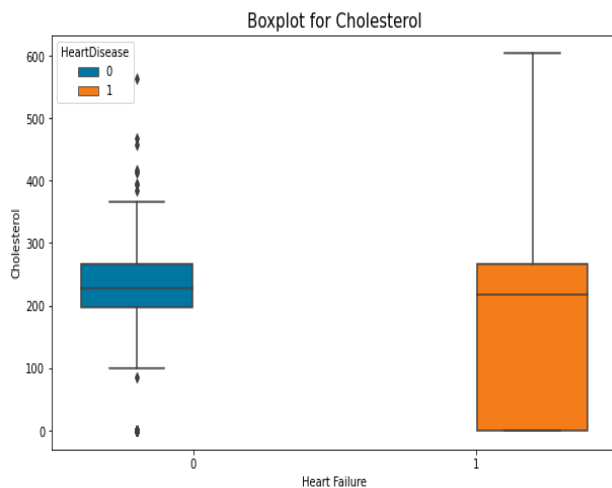
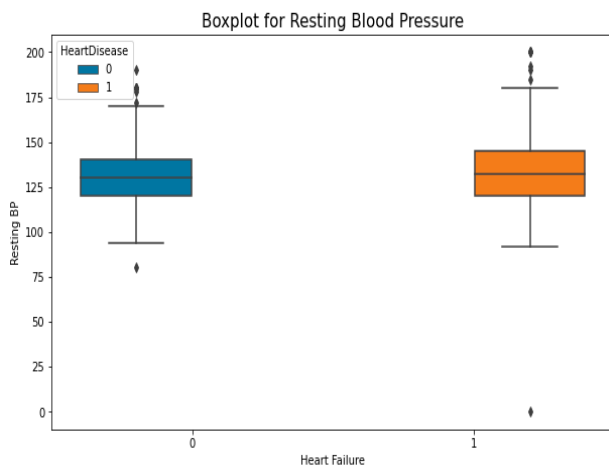
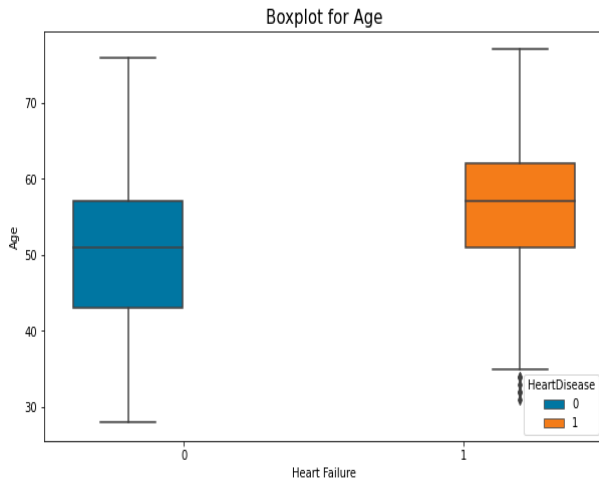


Fig. 5. Count plots for categorical variables in the dataset

Based on the above count plots, the following inferences were made;

1. Male had more heart failure than female.
2. Surprisingly those with asymptomatic (ASY) chest pain had more heart failure.
3. Again, those with fasting blood sugar not greater than 120 mm/dl tend to have more heart failure.
4. Those with normal resting ECG tend to have more heart failure. This is interesting.
5. Those with exercise-induced angina tend to have more heart failure.
6. Those with flat slope of the peak exercise ST segment tend to have more heart failure



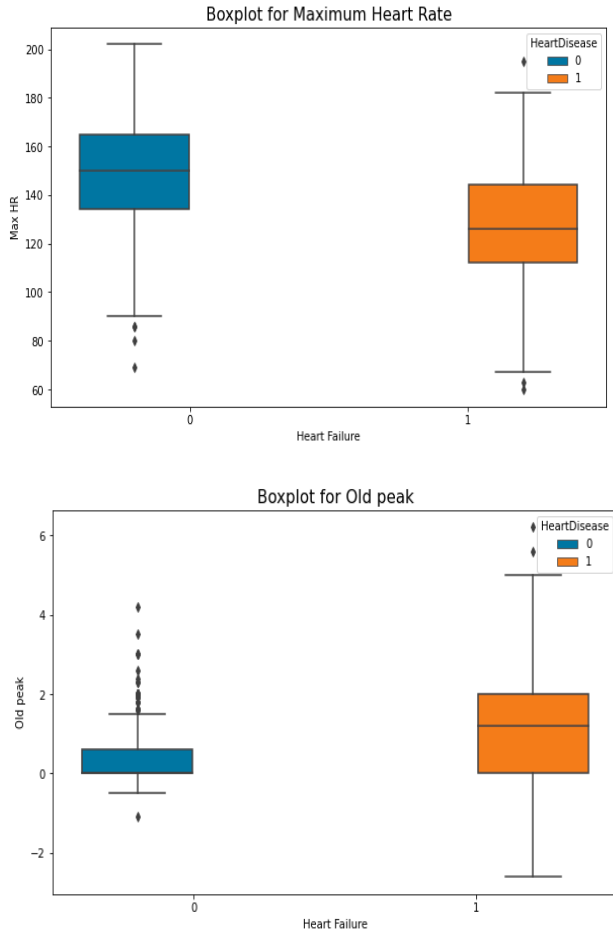


Fig. 6. Boxplots for continuous variables in the dataset

From the above box plots, the following inferences were made;

1. Older people are at high risk of getting heart failure. This is because heart failure affected older people even though there were few outliers.
2. Resting blood pressure is not a major cause of heart failure. Both high and low levels are affected.
3. Cholesterol is not a major cause of heart failure. Both high and low levels are affected.
4. Maximum heart rate is not a major factor to heart failure. Both high and low levels are affected.
5. Old peak is not a major factor to heart failure. But on the average, people with heart failure had higher value than those who did not have

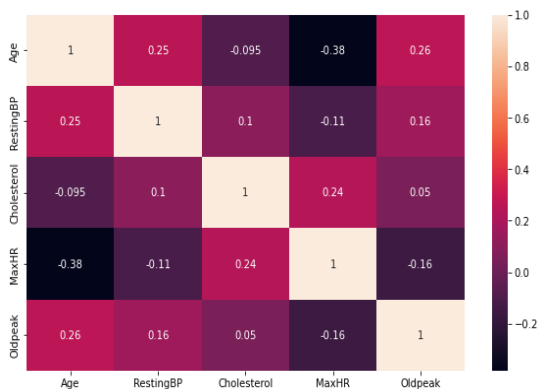


Fig. 7. Correlation coefficients between the continuous variables

From Fig. 7, there is no strong positive or negative correlation coefficient between the continuous variables.

• **Confirmatory Data Analysis (CDA)**

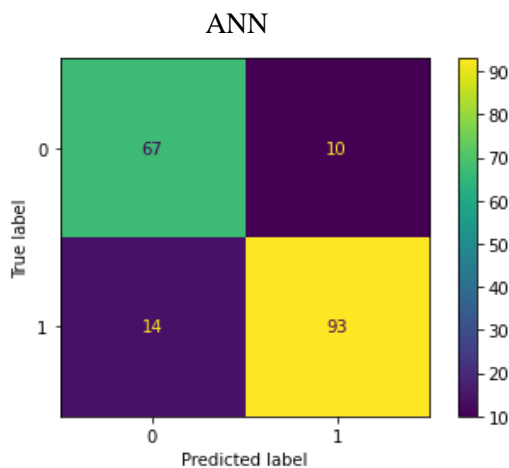
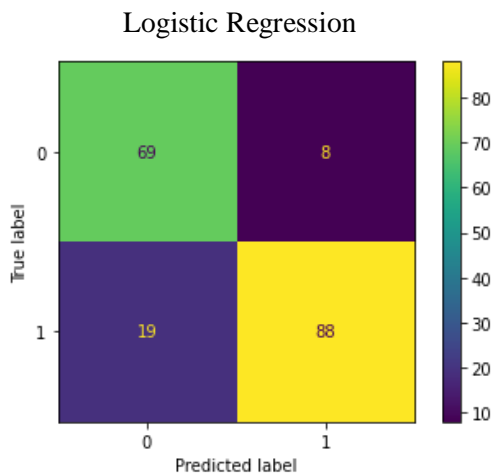
A non-parametric statistical test (Chi-square) was used on most of the categorical variables to test for causality. Again, a parametric statistical test (One-way ANOVA) was also conducted on the continuous variables and the following deductions were made;

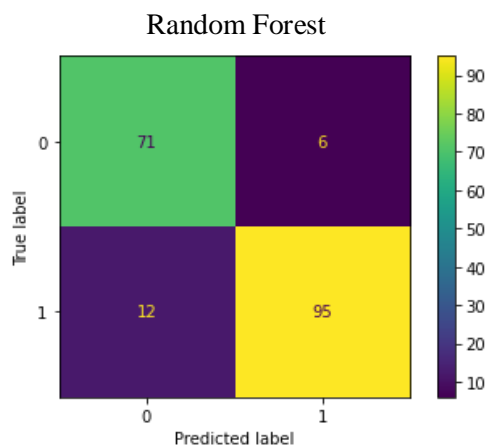
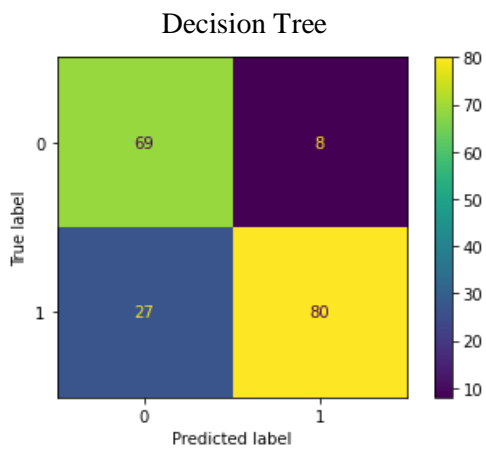
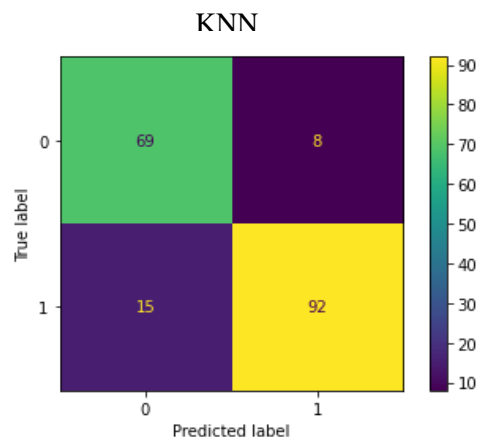
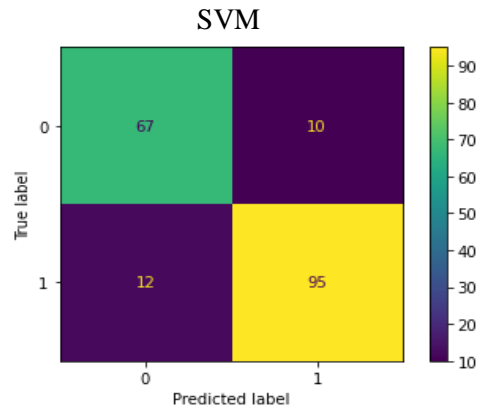
1. Sex is among the causes of heart failure.
2. Chest pain type is among the causes of heart failure.
3. Fasting BS is among the causes of heart failure.
4. Resting ECG does not contribute to heart failure.
5. Exercise angina is among the causes of heart failure.
6. ST slope is among the causes of heart failure.
7. Those with and without heart failure do not share the same mean or variance.

RESULTS ANALYSIS AND DISCUSSIONS

• **Confusion Matrix**

The figure below (Fig. 8) depicts the confusion matrices for the various classifiers.





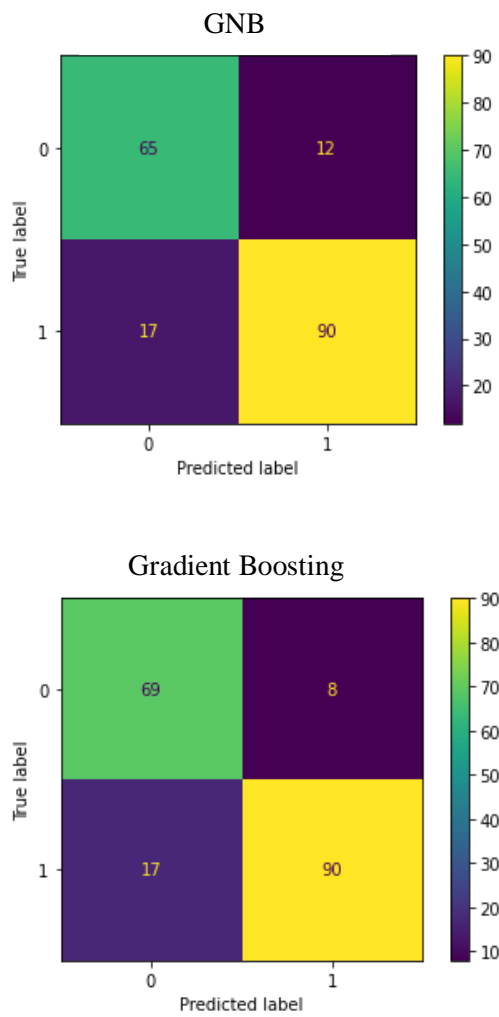


Fig. 8. Confusion Matrices for the classifiers

• **Performance Comparison of Various Classifiers**

The table below shows the comparison of the various metrics for evaluating the classifiers

Table IV. Comparison Of Classifiers Using Various Evaluation Metrics

#	Classifier	ACC	PRE	REC	F1-Score	AUC
1	Logistic Regression (LR)	85%	86%	85%	85%	91%
2	ANN	87%	87%	87%	87%	93%
3	SVM	88%	88%	88%	88%	94%
4	KNN	88%	88%	88%	88%	92%
5	Decision Tree (DT)	81%	83%	81%	81%	82%
6	Random Forest (RF)	90%	90%	90%	90%	95%
7	Gaussian Naïve Bayes (GNB)	84%	84%	84%	84%	91%
8	Gradient Boosting (GB)	86%	87%	86%	86%	93%

ACC = Accuracy, PRE = Precision, REC = Recall, AUC = Area under Curve

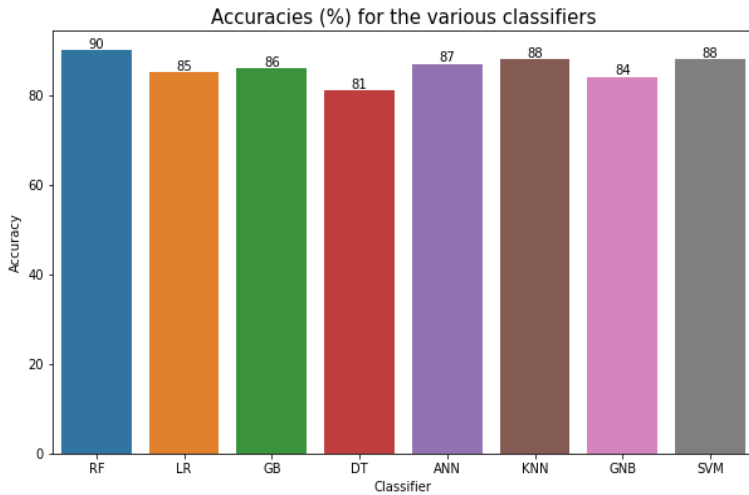


Fig. 9. Accuracies for the classifiers

Results and Performance Analysis

The confusion matrices in Fig. 8 revealed the rates of false positives (FP) and false negatives (FN), which are critical considerations for any model. A false positive might lead to unnecessary treatment, while a false negative, particularly in the case of undetected heart failure, could result in a severe misdiagnosis. The Random Forest classifier demonstrated a low incidence of FP and FN enhancing its reliability. The false positives suggest that these non-heart failure patient records may exhibit heart failure-like characteristics. While the false negatives also suggest that these heart failure patients may exhibit non-heart failure-like characteristics. TABLE IV provides an evaluation of accuracy, precision, recall, F-1 score, and AUC for various classification methods, as detailed in equations (3.1) to (3.4). The Random Forest (RF) model achieved a 90% accuracy rate, outperforming the other classifiers. Precision, which is the ratio of correctly predicted positive observations to the total predicted positives, was the highest in RF (0.90), indicating a lower false-positive rate for RF.

Recall, the measure of correctly predicted positive cases relative to all cases in the class. RF showed a superior recall. The F-1 score, which is the harmonic mean of precision and recall, takes into account both false positives and negatives. Although not as straightforward as accuracy, the F-1 score is often more informative, especially in imbalanced class distributions. RF scored highest too in this metric. The final metric, Area under Curve (AUC), assesses the total area under the ROC curve, stretching from (0, 0) to (1, 1). A higher score, closer to 1, indicates better performance. Here again, RF excelled with a score of 0.95.

From all the metrics used, RF outperformed all the other classifiers which suggests that RF work well on the dataset used for this research work.

Benchmarking

The table below (TABLE V) shows the accuracy of some related work as compared to this work.

Table V. Results Comparison Of The Related Works

#	Reference	Classifier Used	Accuracy
1	Pandita et al.	Random Forest	87.50%
		Support Vector Machine	87.50%

2	Awan et al.	Random Forest	76.39%
		Support Vector Machine	71.80%
3	Alotaibi	Random Forest	89.14%
		Support Vector Machine	92.30%
4	Haq et al.	Random Forest	83%
		Support Vector Machine	86%
5	Princy et al.	Random Forest	71%
		Support Vector Machine	72%
6	This work (Proposed)	Random Forest	90%
		Support Vector Machine	88%

CONCLUSION AND RECOMMENDATIONS

This section summarizes the entire research work, the findings, challenges encountered throughout the research process. It also talks about recommendations and some appropriate and useful future works that can complement this work.

Summary

Heart failure, also known as congestive heart failure, is a long-term, gradually worsening condition in which the heart muscle doesn't pump blood as effectively as required. This decreased efficiency can stem from a range of underlying factors and may impact the left side, right side, or both sides of the heart. It's important to note that this condition doesn't imply the heart has ceased functioning; instead, it indicates that the heart isn't operating efficiently enough to fulfil the body's demands. It is the leading cause of death globally. Heart failure is a critical health issue that requires continuous medical care and changes in lifestyle habits. Detecting and treating it early can greatly enhance the life quality of individuals affected by this condition. Medical practitioners and physicians are able to diagnose and prescribe some medications for people with heart failure. People can get to know whether they have heart failure only if they have undergone through proper medical check-ups and tests. Financial constraints pose a lot challenges on people to be able to undergo these proper medical check-ups. This research focuses on the use of Machine Learning in predicting heart failure. Both individuals and medical practitioners can benefit from this system. A lot of works has been done already in this field but in this research work, the focus was on the comparative analysis on several classification algorithms. Random Forest outperformed all the other algorithms with an accuracy of **90%**.

Challenges and problems encountered

The following were the challenges and problems encountered throughout the research work;

- Inaccessibility of local (Ghanaian) medical dataset
- Choice of appropriate dataset for this research

- Some external knowledge was required from some health workers before some of the features in the dataset were fully understood
- It took longer period for some of the models to complete the training during the fine-tuning process hence the processes were halted.
- The appropriate hyper parameter to fine-tune in order to produce a higher accuracy.

Recommendations

- Ghana government should be able to fund or sponsor individuals and corporate bodies who are willing to embark on such research.
- Local clinical dataset should be made available for researchers who want to use secondary data in their work.
- An online portal should be setup to store local clinical datasets for researchers to utilize.

Future Works

The field of Artificial Intelligence keeps on evolving day in day out so it is also in Machine Learning. There is always an emerging technology in these fields. For these reasons, the researcher proposes the following as future works;

- Multiple datasets would be used for same or similar research
- Deep Learning, Deep Neural Networks, and Transfer Learning would also be utilized for future works
- There would be more research into the new emerging technologies in the field of Machine Learning like Generative AI, Transformers, Large Learning Models (LLMs), etc.

Conclusion

On a whole, this research has been a successful one in tackling or solving the stated problem. The best model has been chosen to solve the problem. Both the aim and objectives of this research have been achieved.

REFERENCES

1. fedesoriano. (September 2021). Heart Failure Prediction Dataset. Retrieved [September 12, 2023] from <https://www.kaggle.com/fedesoriano/heart-failure-prediction>.
2. Adler, E. D., Voors, A. A., Klein, L., Macheret, F., Braun, O. O., Urey, M. A., ... Yagil, A. (2020). Improving risk prediction in heart failure using machine learning. *European Journal of Heart Failure*, 22(1), 139–147. <https://doi.org/10.1002/ejhf.1628>
3. Alotaibi, F. S. (2019). Implementation of machine learning model to predict heart failure disease. *International Journal of Advanced Computer Science and Applications*, 10(6), 261–268. <https://doi.org/10.14569/ijacsa.2019.0100637>
4. Awan, S. E., Bennamoun, M., Sohel, F., Sanfilippo, F. M., & Dwivedi, G. (2019). Machine learning-based prediction of heart failure readmission or death: implications of choosing the right model and the right metrics. *ESC Heart Failure*, 6(2), 428–435. <https://doi.org/10.1002/ehf2.12419>
5. Azam, M. S., Habibullah, M., & Rana, H. K. (2020). Performance Analysis of Machine Learning Approaches in Stroke Prediction. *Proceedings of the 4th International Conference on Electronics, Communication and Aerospace Technology, ICECA 2020*, 175(21), 1464–1469. <https://doi.org/10.1109/ICECA49313.2020.9297525>
6. Bhardwaj, A., Kundra, A., Gandhi, B., Kumar, S., Rehalia, A., & Gupta, M. (2019). Prediction of Heart Attack Using Machine Learning. *IITM Journal of Management and IT*, 10, 20–24.
7. Bharti, R., Khamparia, A., Shabaz, M., Dhiman, G., Pande, S., & Singh, P. (2021). Prediction of Heart Disease Using a Combination of Machine Learning and Deep Learning. *Computational*

- Intelligence and Neuroscience, 2021. <https://doi.org/10.1155/2021/8387680>
8. Choudhary, G., & Singh, D. S. N. (2020). Prediction of Heart Disease using Machine Learning Algorithms, 197–202. <https://doi.org/10.1109/ICSTCEE49637.2020.9276802>
 9. Dritsas, E., & Trigka, M. (2022). Stroke Risk Prediction with Machine Learning Techniques. *Sensors*, 22(13). <https://doi.org/10.3390/s22134670>
 10. Haq, A. U., Li, J. P., Memon, M. H., Nazir, S., Sun, R., & Garcíá-Magarinõ, I. (2018). A hybrid intelligent system framework for the prediction of heart disease using machine learning algorithms. *Mobile Information Systems*, 2018. <https://doi.org/10.1155/2018/3860146>
 11. Lutimath, N. M., Chethan, C., & Pol, B. S. (2019). Prediction of heart disease using machine learning. *International Journal of Recent Technology and Engineering*, 8(2 Special Issue 10), 474–477. <https://doi.org/10.35940/ijrte.B1081.0982S1019>
 12. Mohan, S., Thirumalai, C., & Srivastava, G. (2019). Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques. *IEEE Access*, 7, 81542–81554. <https://doi.org/10.1109/ACCESS.2019.2923707>
 13. Nikhar, S., & Karandikar, A. M. (2020). Prediction of heart disease using machine learning algorithms. *Proceedings of the International Conference on Smart Technologies in Computing, Electrical and Electronics, ICSTCEE 2020*, (6), 197–202. <https://doi.org/10.1109/ICSTCEE49637.2020.9276802>
 14. Olsen, C. R., Mentz, R. J., Anstrom, K. J., Page, D., & Patel, P. A. (2020). Clinical applications of machine learning in the diagnosis, classification, and prediction of heart failure: Machine learning in heart failure. *American Heart Journal*, 229, 1–17. <https://doi.org/10.1016/j.ahj.2020.07.009>
 15. Padmaja, B., Srinidhi, C., Sindhu, K., Vanaja, K., Deepika, N. M., & Rao Patro, E. K. (2021). Early and Accurate Prediction of Heart Disease Using Machine Learning Model, 12(6), 4516–4528.
 16. Pandita, A., Vashisht, S., Tyagi, A., & Yadav, P. S. (2021). Prediction of Heart Disease using Machine Learning Algorithms, (May).
 17. Princy, R. J. P., Jose, P. S. H., Parthasarathy, S., Lakshminarayanan, A. R., & Jeganathan, S. (2020). Prediction of Cardiac Disease using Supervised Machine Learning Algorithms, (*Iciccs*), 570–575.
 18. Sharma, H., & Rizvi, M. A. (2017). Prediction of Heart Disease using Machine Learning Algorithms : A Survey.