Nanotechnology Perceptions
ISSN 1660-6795
www.nano-ntp.com

Implementation Of A Heart Disease Risk Prediction Model Using Machine Learning

Shailendra Chaurasia[1], Dr.Megha Kamble[2*]

[1,2]Department of Computer Science Engineering,,LNCT University,Bhopal, Madhya Pradesh,India
[1]chaurasia.shailendra@gmail.com,[2] meghak@lnct.ac.in

Still, heart disease claims a disproportionate number of lives in modern times. When it comes to clinical data analysis, cardiovascular disease prediction is one of the toughest topics. In such a situation, ML may play a significant role in making judgments and predictions about cardiovascular disease based on individual critical indicators, such as BP, cholesterol, smoking, diabetes, obesity, stroke, and alcohol use. This study makes use of ML models that draw on the Cleveland Heart Disease dataset and is housed in the UCI ML repository. A given ML models—SVM, RF, DT, and LR—have been thoroughly compared. The Cleveland Heart Disease dataset is analyzed using principle component analysis (PCA) to extract the useful information about the condition. The model's performance was assessed using metrics generated from the confusion matrix, including recall, accuracy, precision, and F1 score. Among the models tested, the SVM model was the most successful in predicting the likelihood of cardiovascular illness, its 96.66%accuracy, 96.4%precision, 96.4%recall, and 96.4%F1 score attest to this. Future work should involve validating the model on larger and more diverse datasets, integrating additional relevant features, and exploring more sophisticated algorithms to improve accuracy and reliability in heart disease risk prediction, thereby contributing to better healthcare outcomes.

Keywords: Heart Disease, Risk Prediction, Cardiovascular Diseases, Diagnosis, Machine Learning, Cleveland Dataset.

INTRODUCTION

The prognosis of cardiac disease is among the most intricate and difficult illness situations in the world of medical research. An essential organ for every human body is the heart[1]. Heart disease is a term that appropriately describes circumstances in which the faulty function of the heart may be attributed to several causes, such as blood clots or heart arteries. The main causes of the rise in the incidence of heart disease include modifications to lifestyle, work-related stress, and eating habits. Worldwide, 17.7 million individuals have heart disease each year, according to the WHO [2]. The Global Burden of Disease report estimates that almost 1.7 million individuals in India were affected by CVD in 2016 [3]. The most common kind of cardiovascular disease is called (CVD), which happens when the coronary arteries are obstructed in some way. Because of this, heart attacks and vascular ruptures are the results of inadequate blood flow, which harms the brain, legs, and heart. Approximately 2% of the global population is dealing with CVD, and 2% of the yearly expenditure has gone into CVD treatment

[4]. Health care costs rise and productivity falls as a result of cardiac conditions such CVD. Therefore, in order to lower the CVD mortality rate, accurate diagnostic and preventative methods are required.

A doctor would usually make a diagnosis of heart disease after reviewing the patient's symptoms, medicalhistory, and physicalexam findings, as well as other risk factors such age, sex, family history, and lifestyle. Unfortunately, it is computationally costly to analyze such massive components, and manual study leads to erroneous prediction since certain parameters to be analyzed may stay concealed [5]. Although angiography is the gold standard for CVD diagnosis, it comes with a hefty price tag and a host of serious side effects[6]. The healthcare industry is facing significant challenges in providing affordable, dependable diagnoses. Also, screening big populations using image-based detections isn't practical because of how expensive they are. Many academics have thus worked to create an automated diagnostic method for early identification of CVD that is cost-effective and quick, using ML Algorithms.

Several ML algorithms have shown encouraging outcomes in the early identification of cardiac illness by uncovering hidden patterns, leading to a higher accuracy rate and faster performance[7][8] . Despite these improvements, ML algorithms still can't handle computationally intensive jobs because of a computational bottleneck that occurs when working with more and more complicated data on conventional computers[9][10]. The notion of quantum computing has now extended a helping hand in increasing processing capacity. As a result, evaluating and treating complex health conditions is made easier by healthcare firms using quantum enhanced ML algorithms. The primary purpose behind the classification algorithms is to identify the level to which the new data will fall. Classifications are known as a supervised learning algorithm with a computer program that could be learned using data[25].

CONTRIBUTION OF STUDY

This study makes significant contributions to a field of heart disease risk prediction through its comprehensive methodology and findings. The key contribution of study is:

- Employs the Cleveland Heart Disease dataset, a well-recognized resource in the field, facilitating comparisons with existing research.
- Conduct data preprocessing techniques, including outlier detection, one-hot encoding, and normalization, ensuring data quality.
- Used ML models such as SVM, RF, DT and LR to compare and contrast.
- Determine model performance with F1-score, recall, accuracy, and precision.
- Predictive model may assist healthcare professionals in risk assessment and early detection of heart disease, contributing to improved patient outcomes.

STRUCTURE OF PAPER

The following is an outline of a paper: Section 1 introduces a concept of cardiac disease detection; Section 2 reviews the literature and analyses other studies that have addressed this topic; and Section 3 emphasizes any gaps in this knowledge. Section 4 contains the methodology as well as the results and discussion sections. Section 5 discusses more research.

LITERATURE REVIEW

A creation of the risk prediction system for heart disease has been aided by several researchers. Based on data mining methods, a risk of heart disease is forecasted. There are several articles and study materials available to researchers on this cardiac condition these days. The most recent works by various writers and academics are shown in this area.

The study Dibaji and Sulaimany, (2024) employs ML, graph conversion, and community finding approaches to evaluate a sizable dataset that has been aggregated from numerous sources. The suggested method outperforms conventional ML models, according to the results, which show an accuracy94%.

This study demonstrates the possibility forimproving the classification of heart illness using graph-based approaches. It offers perceptions into prospects, constraints, and directions for further study in this area. The results have effects on therapy and diagnosis[11].

LITERATURE WORK COMPARISON FOR HEART DISEASE RISK PREDICTION USING MACHINE LEARNING

Ref	Data	Methods	Performance	Limitation/Future Work
Dibaji and Sulaimany, 2024	A comprehensive dataset amalgamated from multiple sources	Graph conversion, community detection, and ML algorithms	Accuracy: 94%	Future study should focus on enhancing real-world applicability, addressing limitations related to sample size and difficulties in deployment
Singh, Guleria, and Sharma, 2023	Dataset of 1025 tuples, 11 attributes (e.g., age, sex, thalach)	NB, MLP, DT, LR	DT: 95.04%, MLP: 95.51%, NB: 83.12%	Focus on improving model scalability; expand dataset to include more diverse patient characteristics
Das et al., 2023	Survey data from 400k US residents	XGBoost, Bagging, RF, DT, KNN, NB	XGBoost: 91.30%	Larger datasets may enhance model generalizability; further tuning of models to achieve higher sensitivity and accuracy
Elghalid et al., 2022	Kaggle dataset (12 attributes/features)	KNN, NB, LR, SVM, RF, DT, NN	LR, SVM, NN: 94.57%	Future work could explore ensemble learning and deep learning to further

				improve classification performance
Iqbal et al., 2020	UCI Cleveland Heart Disease dataset (14 attributes)	Ordinary Learning Method (OLM), compared with C5.0, SVM, KNN, Neural Network	OLM: 95.46%	Investigate more sophisticated feature engineering methods, testing on broader datasets for better generalizability

This workSingh, Guleria and Sharma, (2023) offers a range of ML and DL models for cardiac disease prediction using a large dataset of patient attributes and health indicators that point to the presence of cardiac disease. This study develops an efficient cardiac disease prediction model using a variety of classifiers, including NB, MLP, DT, and LR. All ML and DL models performed poorly compared to the Decision Tree model, which had an accuracy rate95.04%. When it comes to binary heart disease categorization, nevertheless, MLP has shown a performance of 95.51%, making it the second best model. On top of that, logistic regression has shown an accuracy of 84.48% and Naive Bayes of 83.12% [12].

This study Das et al., (2023)created and assessed six ML models for aforecasting of CVD using survey data from 400,000 US people. Xgboost, Bagging, RF, DT, KNN, and NB are the six ML models that were compared in this research. Additionally, six ML methods are given and assessed based on their accuracy, sensitivity, F1-score, and AUC. With a 91.30% accuracy rate, the Xgboost model demonstrated optimal performance[13].

This paper aims Elghalid et al., (2022) use several ML techniques to forecast whether a patient is likely to get a cardiac diagnosis. A dataset of twelve properties (Features) was utilized in this article and was obtained from the Kaggle repository. KNN, NB, LR, SVM, RF, DT, and NN techniques are used to execute this task. The results showed that the three top methods, namely LR, SVM, and NN, achieved an accuracy of up to 94.57% [14].

This research paper Iqbal et al., (2020) uses clinical data in conjunction with the Ordinary Learning Method to accurately diagnose heart disease. Using 14 characteristics, the suggested technique is evaluated on the Standard UCI Cleveland Heart Disease dataset. At 95.4615% accuracy, the suggested method surpasses state-of-the-art algorithms such as DT, SVM, KNN, and NN. The results demonstrate that the suggested OLM method outperforms other data mining strategies that have been suggested in the literature for the diagnosis of cardiac diseases [15].

Table 1 shows the results of a comparison of studies that have used ML to predict the likelihood of heart disease.

RESEARCH GAP

Despite significant advancements in heart disease prediction using machine learning, several gaps remain that limit the full potential of these methods. The following research gap are:

- **Model Scalability**: Limited exploration of how well models scale with larger or more complex datasets.
- **Data Diversity:** Many models are trained on homogenous datasets, limiting their applicability across diverse populations.
- **Feature Engineering**: Lack of focus on sophisticated feature engineering methods to improve model performance.
- **Real-World Deployment**: Few studies address challenges related to integrating machine learning models into clinical settings, such as interpretability and usability.
- **Generalizability**: Many studies focus on specific datasets, which limits their findings' generalizability to broader populations or different regions.

METHODOLOGY

Several crucial elements are included in the technique for deploying a ML-based heart disease risk prediction model to guarantee good model development and assessment. The whole steps and phases of methodology are present in figure 1.

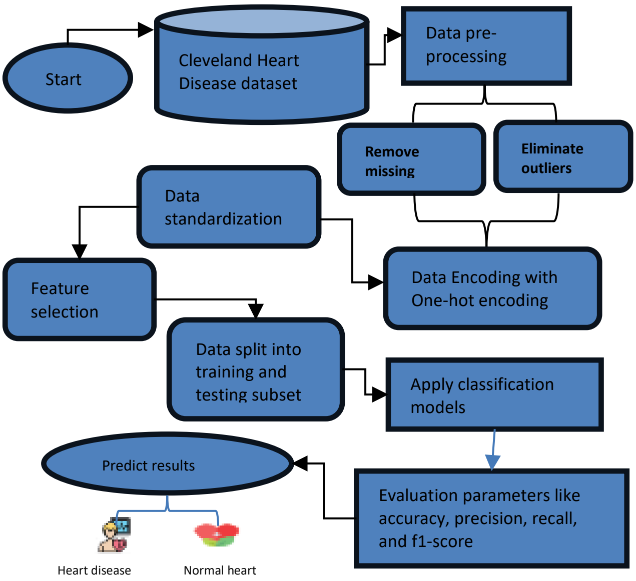


Fig. 1. Flowchart for Heart disease risk prediction

With 303 records and 75 attributes—13 of which are relevant to heart disease prediction—found in the UCI ML repository's Cleveland Heart Disease dataset, we begin by using it. Data pre-processing begins with the removal of records with missing values, reducing the dataset to 297 entries, followed by the identification and elimination of outliers to enhance model performance. Data standardization is implemented by Z-score normalization to guarantee all characteristics are on a similar scale, and one-hot encoding is used to translate categorical data into numerical representation. Feature selection is performed to retain attributes and decrease dimensionality for disease classification by using Principle Component Analysis (PCA)[24] The dataset is then split into training(80%) and testing(20%) sets to facilitate model training and evaluation. Various classification algorithms, including DT, SVM, RF, and LR, are employed to build the predictive model. Measures like accuracy, precision, recall and F1 score are produced by using a confusion matrix to assess the models performance, these metrics offer a comprehensive view of how effectively each categorization model forecasts the risk of cardiovascular disease.

DATASET DESCRIPTION

This study makes use of the UCI ML repository's Cleveland Heart Disease dataset as its heart disease dataset. Four distinct databases gathered from four different medical facilities make up the UCI heart disease dataset. There are 75 attributes and 303 records in the collection. The existence of heart disease in a patient can be predicted using one of thirteen features in the heart disease prediction model.

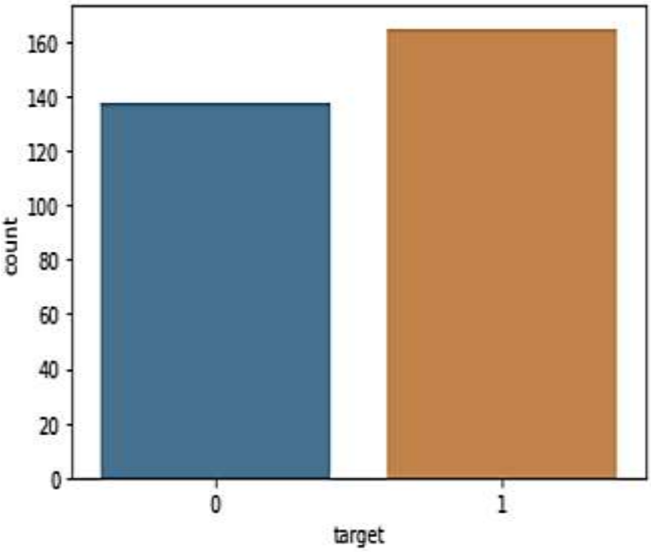


Fig. 2. Count of target Variables

Figure 2 shows a bar plot representing the count of target variables, likely from a classification dataset. It indicates two classes, 0 and 1. The count for each class seems slightly imbalanced, with the count for class "1" being slightly higher than class "0." Both bars represent counts in the range of approximately 140-160. This plot might help visualize class distribution, often useful when working with classification problems, especially when addressing class imbalance.

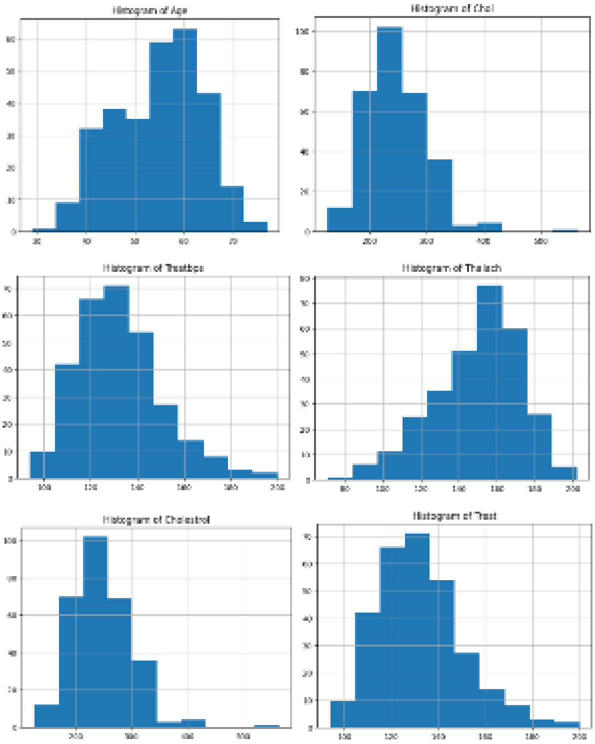


Fig. 3. Histogram of continuous-valued attributes

Figure 3 presents histograms for 6 continuous attributes. Each histogram displays the frequency distribution of these variables, helping to identify patterns, trends, or potential outliers in the dataset, which is valuable for exploratory data analysis.

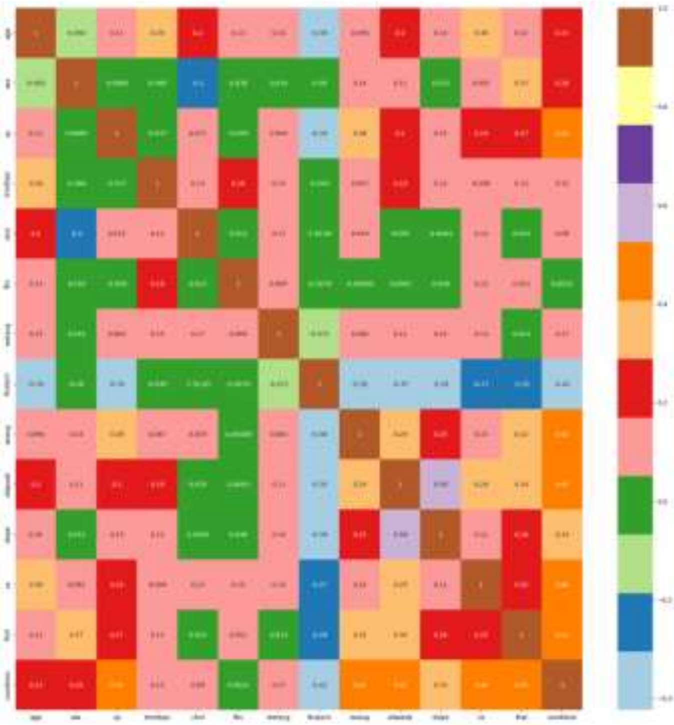


Fig. 4. heat map for correlation features data

A heat map of the dataset's correlations is shown in Figure 4, using a color scale going to show the relationships between characteristics. Each square shows the correlation between two features, while the diagonal represents perfect positive correlations, as each feature is compared with itself. This visualization helps identify relationships and dependencies between variables.

DATA PRE-PROCESSING

Data preparation involves analyzing, filtering, modifying, and encoding data in order to make it understandable and usable by a ML system. Removing errors like missing values, improving data quality, and preparing data for ML are the main objectives of data preprocessing. Further pre-processing steps are given below:

- **Remove missing values:** After removing all records with blank or incomplete information, the dataset shrinks from 303 to 297 records.
- **Eliminate outliers:** Outliers can distort a performance of ML models, so it's common to identify and eliminate them.

DATA ENCODING WITH ONE-HOT ENCODING

Data encoding refers to converting categorical data into a numerical format that can be fed into machine learning models. Nominal categorical data is the primary focus of one-hot encoding, a feature engineering technique. Data conversion to numerical form is necessary for ML to be applied to categorical data in the absence of a tree-based method.

DATA STANDARDIZATION

Data standardization is a typical preprocessing step in ML. It makes sure that features are all on the same scale, which improves the model's performance and speeds up its convergence. The Z-score normalization is a popular tool for standardizing data and is defined as (1):

$$z = \frac{x - \mu}{\sigma} \tag{1}$$

where the original value is represented by x, the standardized value by z, the feature's mean by μ, and its standard deviation by σ.

FEATURE SELECTION

There are a total of fourteen characteristics included in the dataset; thirteen of these are used for the purpose of forecastingan existence of heartdisease in patients, and one of these attributes acts as either a class label or the predicted attribute. In this context, "age" and "sex" are patient-specific identifiers. There are eleven more clinical characteristics. Characteristics are drawn from a separate variable that is more than 30% related to the dependent variable. This study will make use of the eight characteristics that were chosen.

DATA SPLITTING

Proceed to split a dataset into training and testing data after a completion of a chosen features from a dataset. In this research, 20% of the data are used for testing and the remaining 80% are used to train our model.

CLASSIFICATION WITH SVM

A support vector machine is a type of model used to analyze data and discover patters in classification and regression analysis. Support vector machine (SVM) is used when your data has exactly two classes[27]. An SVM classifies data by finding the best hyper plane that separates all data points of one class from those of the other class. The larger margin between the two classes, the better the model is. A margin must have no points in its interior region. The support vectors are the data points that on the boundary of the margin[30]. SVM is based on mathematical functions and used to model complex, and real world problems[28]. SVM performs well on data sets that have many attributes. Support Vector Machines map the training data into kernel space. In addition, there are multiple methods of implementing SVM, such as quadratic programming, sequential minimal optimization, and least squares. The challenging aspect of SVM is kernel selection and method selection such that your model is not over optimistic or pessimistic. Considering that the Cleveland Heart Disease Database (CHDD)[24] has a large number of instances as well as features, it is arguable whether the kernel chosen is RBF or linear. Although the relation between the attributes and class labels are nonlinear, due to the large number of features, RBF kernel may not improve performance. It is recommended that both kernels be tested and the more efficient one be finally selected.

Model Evaluation

To assess how well a classification model is doing, a confusion matrix can be a useful tool in ML. A brief synopsis of the forecasted and actual class labels in the test data is made available in this matrix. Quantities of true positives, false positives, true negatives, and false negatives are shown for every class in a confusion matrix. By definition, TP is the number of positively predicted members of the class, FP is the number of negatively predicted members of the class, TN is the number of positively predicted members of the negative class, and FN is the number of negatively predicted members of the negative class. A confusion matrix can be utilized to compute evaluation measures including F1 score, recall, accuracy, and precision.

Accuracy: The accuracy score, which is short for "classification accuracy rating," is calculated by splitting the total numberof predictions by the numberof correct predictions. Accuracy is defined by Equation (3).

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{2}$$

Precision: A measure of precision can be found by splittinga proportion of correct diagnoses by a total numberof positive results (including those that were incorrectly identified). In order to determine P, we must use Equation (4):

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{3}$$

Recall: A measure of recall is the ratio of true positive samples to true positive outcomes. To determine the recall, one uses Equation (5):

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{4}$$

F1-Score: The model's accuracy in each class is determined by the F1 score. It is common practice to use the F1-score measure when dealing with an imbalanced dataset. Here, it shows how the proposed method works by evaluating it using the F1 score [19]. The F1-score is obtained using Equation (6).

$$\text{F1} - \text{Score} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \tag{5}$$

The ML models and their corresponding scores were thoroughly assessed using the aforementioned criteria.

RESULTS ANALYSIS AND DISCUSSION

The efficiency of the ML algorithms used to identify an existence of heartdisease is shown by a several performance metrics that are revealed by the examination of the HeartDisease Risk Prediction Model's predictions. We assess the ML model's overall performance and compare it to other advance approaches in this section. Thefollowing table 2 provides the SVM model performance across heart disease dataset.

TABLE I. SVM MODEL PERFORMANCE FOR HEART DISEASE RISK PREDICTION

Matrix	**Support Vector Machine**

Accuracy	96.66
Precision	96.4
Recall	96.4
F1-score	96.4

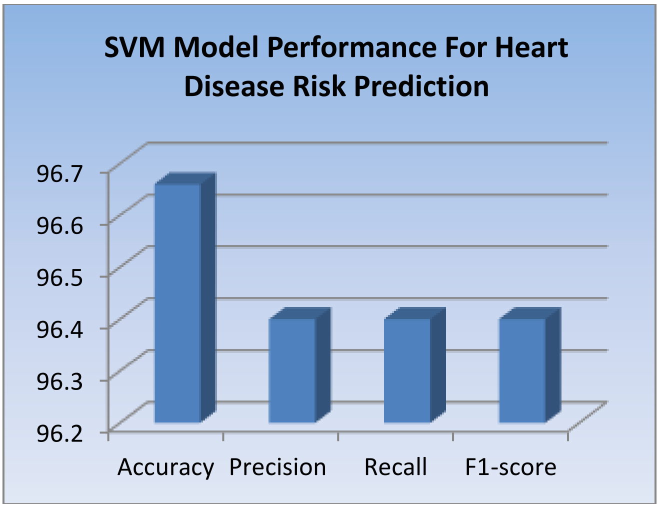
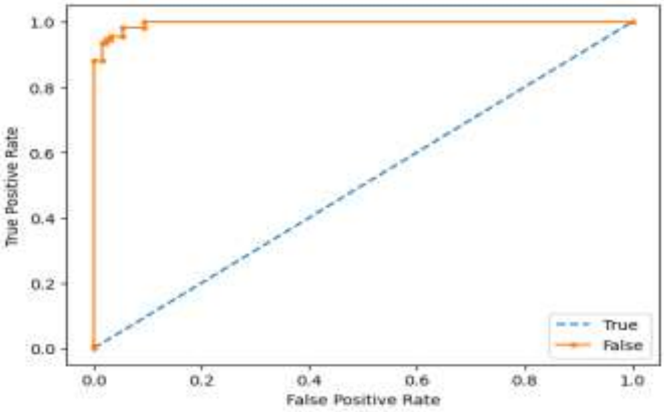


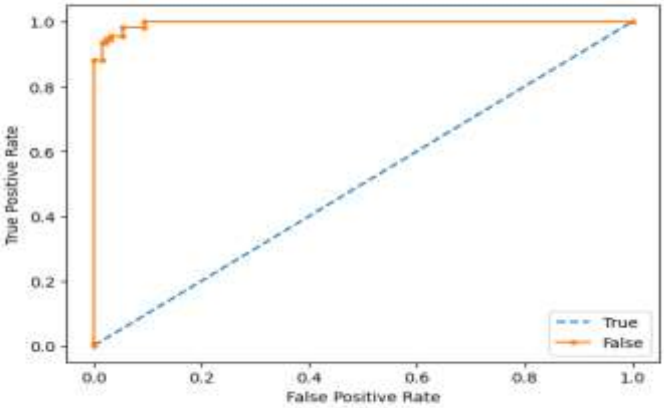
Fig. 5. Bar graph for model performance on data

The above table 1 and figure 5 show the SVM model performance. The model achieves a high accuracy of 96.66%, indicating it correctly classified the vast majority of instances. Its precision is 96.4%, showing an excellent ability to correctly identify negative cases. The model's efficacy in accurately detecting affirmative instances is shown by its 96.4% recall (or sensitivity). With a strong F1-score of 96.4%, the SVM model balances precision and recall, making it highly reliable for heart disease prediction.

ROC Curve of svm for testing
AUC: 0.967
True: ROC AUC=0.500
False: ROC AUC=0.995



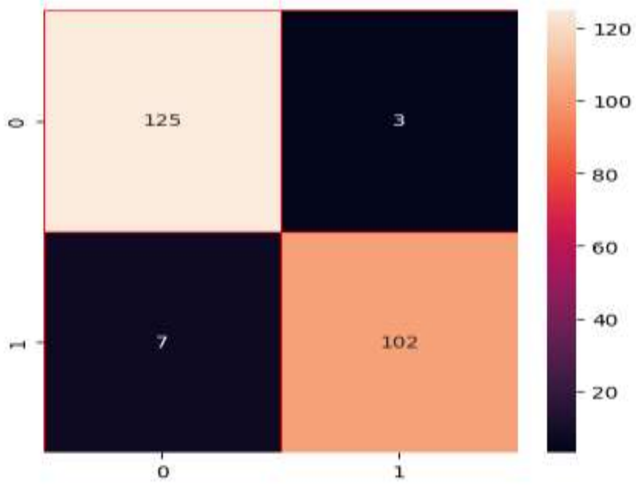
ROC Curve of svm for training
AUC: 0.956
True: ROC AUC=0.500
False: ROC AUC=0.995



AUC and ROC curve for SVM

Figure 6 shows the ROC curve, which is used to evaluate the classification performance of SVM models. A y-axis displays the TPR and a x-axis a FPR) Two lines are plotted: the Train AUC (0.956) and the Test AUC (0.967) with a dashed diagonal line indicating the performance of a random classifier for comparison. The model's class discrimination capabilities on the training and test datasets are shown by this graphic.

Confusion matrix of SVM for training



Confusion matrix of svm for testing

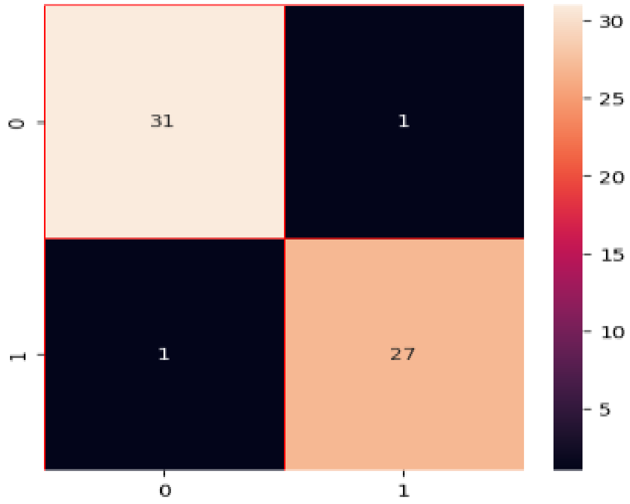


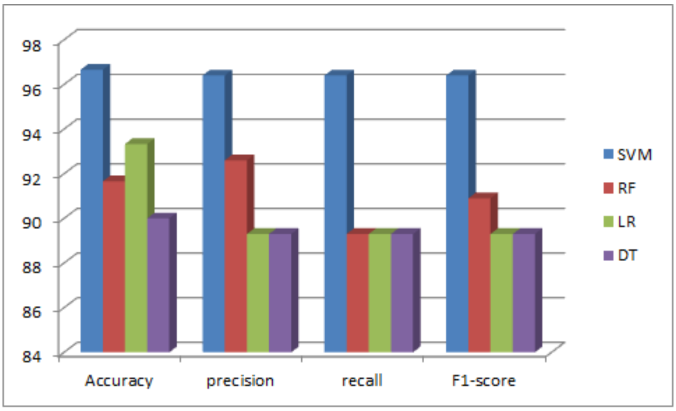
Fig. 6. Confusion matrix for SVM

Figure 7 displays a confusion matrix for a SVM model, where TN are 31, FP are 1, FN is 1, and TP are 27 for testing and TN are 125, FP are 3, FN is 7, and TP are 102 for training. This matrix aids in assessing the efficacy of the predictions by highlighting the model's performance by displaying the right and wrong classifications.

TABLE II. COMPARATIVE ANALYSIS OF MODELS' PERFORMANCE ON CLEVELAND HEART DISEASE DATASET

Model	**Accuracy**	**precision**	**recall**	**F1-score**

SVM	**96.66**	**96.4**	**96.4**	**96.4**
RF	91.66	92.6	89.3	90.9
LR	93.33	89.3	89.3	89.3
DT	90	89.3	89.3	89.3



The comparative analysis for heart disease risk prediction on the dataset is shown in Table 3. In this comparison, the SVMmodel outperforms others with an exceptional accuracy96.66%, precision96.4%, recall96.4%, and F1-score96.4. LR also performs well with 93.33% accuracy, 89.3% precision, 89.3% recall, and an 89.3% F1score. DT shows strong precision at 89.3% but lower recall at 89.3%, yielding an90% accuracy and an F1-score of 89.3%.RF follows with 91.66% accuracy and balanced precision 92.6% recall 89.3%, and F1-score 90.9. SVM stands out for its superior performance in predicting heart disease risk.

CONCLUSION AND FUTURE SCOPE

Heartdisease is a significant cause of death each year and one of the public's main health concerns in this globe. The WHO reports that in 2019, cardiovascular disease (CVD) claimed the lives of more than 17 million individuals. Early detection is key to effectively preventing and treating heart disease, and ML algorithms demonstrate promise in predicting the probability of heartdisease due to several harmful factors. Modern methods based on ML has greatly simplified the process of diagnosing heart diseases. ML methods such as DT, RF, SVM, and LR are employed for a purpose of predicting heartdisease. It was necessary to finish data preparation and feature selection before constructing the models. Included in the model's performance calculation are metrics likeaccuracy, recall, f1-score, and precision. In conclusion, this study successfully developed a HeartDisease Risk Prediction Model, highlighting the effectiveness of the SVM algorithm, which achieved remarkable metrics: accuracy96.66%, precision96.4%, recall96.4%, and an F1 score96.4%. These results indicate that the model can effectively identify individuals at risk of heart disease, potentially aiding in timely medical interventions. Future work may also involve the development of a tool to predict the risk of disease of a prospective patient. The framework can also be extended by using the hybrid machine learning classifiers on same or hybrid dataset.

REFERENCES

[1] C. S.Dangare and S. S. Apte, "Improved Study of Heart Disease Prediction System using Data Mining Classification Techniques," Int. J. Comput. Appl., 2012, doi: 10.5120/7228-0076.

[2] V. V. Ramalingam, A. Dandapath, and M. Karthik Raja, "Heart disease prediction using machine learning techniques : a survey," Int. J. Eng. Technol., vol. 7, no. 2.8, p. 684, Mar. 2018, doi: 10.14419/ijet.v7i2.8.10557.

[3] S. Mendis, P. Puska, and B. Norrving, "Global atlas on cardiovascular disease prevention and control," World Heal. Organ., 2011.

[4] R. Katarya and P. Srinivas, "Predicting Heart Disease at Early Stages using Machine Learning: A Survey," in Proceedings of the International Conference on Electronics and Sustainable Communication Systems, ICESC 2020, 2020. doi: 10.1109/ICESC48915.2020.9155586.

[5] J. P. Li, A. U. Haq, S. U. Din, J. Khan, A. Khan, and A. Saboor, "Heart Disease Identification Method Using Machine Learning Classification in E-Healthcare," IEEE Access, 2020, doi: 10.1109/ACCESS.2020.3001149.

[6] Z. Arabasadi, R. Alizadehsani, M. Roshanzamir, H. Moosaei, and A. A. Yarifard, "Computer aided decision making for heart disease detection using hybrid neural network-Genetic algorithm," Comput. Methods Programs Biomed., 2017, doi: 10.1016/j.cmpb.2017.01.004.

[7] S. K. Mehdiev, I. I. Mustafaev, and M. N. Mamedov, "Features of glycemic status and risk factors in patients with coronary artery disease and type 2 diabetes in Azerbaijan population," Russ. J. Cardiol., 2019, doi: 10.15829/1560-4071-2019-6-85-91.

[8] A. El-Ibrahimi, O. Terrada, O. El Gannour, B. Cherradi, A. El Abbassi, and O. Bouattane, "Optimizing Machine Learning Algorithms for Heart Disease Classification and Prediction," Int. J. online Biomed. Eng., 2023, doi: 10.3991/ijoe.v19i15.42653.

[9] S. S. Kavitha and N. Kaulgud, "Quantum K-means clustering method for detecting heart disease using quantum circuit approach," Soft Comput., 2023, doi: 10.1007/s00500-022-07200-x.

[10] R. Tandon, "The Machine Learning Based Regression Models Analysis For House Price Prediction," Int. J. Res. Anal. Rev., vol. 11, no. 3, 2024.

[11] A. Dibaji and S. Sulaimany, "Community Detection to Improve Machine Learning Based Heart Disease Prediction," in 2024 20th CSI International Symposium on Artificial Intelligence and Signal Processing (AISP), IEEE, Feb. 2024, pp. 1–6. doi: 10.1109/AISP61396.2024.10475216.

[12] G. Singh, K. Guleria, and S. Sharma, "Machine Learning and Deep Learning Models for Early Detection of Heart Disease," in 2023 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), IEEE, Nov. 2023, pp. 419–424. doi: 10.1109/ICCCIS60361.2023.10425392.

[13] R. C. Das, M. C. Das, M. A. Hossain, M. A. Rahman, M. H. Hossen, and R. Hasan, "Heart Disease Detection Using ML," in 2023 IEEE 13th Annual Computing and Communication Workshop and Conference, CCWC 2023, 2023. doi: 10.1109/CCWC57344.2023.10099294.

[14] R. A. M. Elghalid, A. Alwirshiffani, A. A. I. Mohamed, F. H. A. Aldeeb, and A. Andiasha, "Comparison of Some Machine Learning Algorithms for Predicting Heart Failure," in Proceedings - 2022 International Conference on Engineering and MIS, ICEMIS 2022, 2022. doi: 10.1109/ICEMIS56295.2022.9914325.

[15] J. Iqbal, M. M. Iqbal, U. Khadam, and A. Nawaz, "Ordinary Learning Method for Heart Disease Detection using Clinical Data," in 2020 3rd International Conference on Computing, Mathematics and Engineering Technologies: Idea to Innovation for Building the Knowledge Economy, iCoMET 2020, 2020. doi: 10.1109/iCoMET48670.2020.9074056.

[16] Y. Y. Song and Y. Lu, "Decision tree methods: applications for classification and prediction," Shanghai Arch. Psychiatry, 2015, doi: 10.11919/j.issn.1002-0829.215044.

[17] J. Thomas, "Enhancing Supply Chain Resilience Through Cloud-Based SCM and Advanced Machine Learning : A Case Study of Logistics," J. Emerg. Technol. Innov. Res., no. September 2021, 2021.

[18] H. Sinha, "Predicting Bitcoin Prices Using Machine Learning Techniques With Historical Data," Int. J. Creat. Res. Thoughts, vol. 12, no. 8, 2024, doi: 10.3390/e25050777.

[19] Y. Qian, G. Zeng, Y. Pan, Y. Liu, L. Zhang, and K. Li, "A Prediction Model for High Risk of Positive RT-PCR Test Results in COVID-19 Patients Discharged From Wuhan Leishenshan Hospital, China," Front. Public Heal., 2021, doi: 10.3389/fpubh.2021.778539.

[20] P. Nandakumar and R. Subhashini, "Heart Disease Prediction Using Convolutional Neural Network with Elephant Herding Optimization," Comput. Syst. Sci. Eng., vol. 48, no. 1, pp. 57–75, 2024, doi: 10.32604/csse.2023.042294.

[21] A. Anderies, J. A. R. W. Tchin, P. H. Putro, Y. P. Darmawan, and A. A. S. Gunawan, "Prediction of Heart Disease UCI Dataset Using Machine Learning Algorithms," Eng. Math. Comput. Sci. J., 2022, doi: 10.21512/emacsjournal.v4i3.8683.

[22] L. Gamage, "Heart Disease Prediction Using Machine Learning Techniques: A Comparative Analysis," Int. Conf. Adv. Technol. Comput., no. December, pp. 999–1005, 2021.

[23] J. Premsmith and H. Ketmaneechairat, "A predictive model for heart disease detection using data mining techniques," J. Adv. Inf. Technol., 2021, doi: 10.12720/jait.12.1.14-20.

[24] J. Vijayashree and H. P. Sultana, "A Machine Learning Framework for Feature Selection in Heart Disease Classification Using Improved Particle Swarm Optimization with Support Vector Machine Classifier," Program Comput Soft, vol. 44, no. 6, pp. 388–397, Nov. 2018, doi: 10.1134/S0361768818060129.

[25] D. Shah, S. Patel, and S. K. Bharti, "Heart Disease Prediction using Machine Learning Techniques," SN COMPUT. SCI., vol. 1, no. 6, p. 345, Nov. 2020, doi: 10.1007/s42979-020-00365-y.

[26] A. Saboor, M. Usman, S. Ali, A. Samad, M. F. Abrar, and N. Ullah, "A Method for Improving Prediction of Human Heart Disease Using Machine Learning Algorithms," Mobile Information Systems, vol. 2022, pp. 1–9, Mar. 2022, doi: 10.1155/2022/1410169.

[27] V. Sharma, S. Yadav, and M. Gupta, "Heart disease prediction using machine learning techniques," in 2020 2nd international conference on advances in computing, communication control and networking (ICACCCN), IEEE, 2020, pp. 177–181. Accessed: Apr. 07, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9362842/

[28] B. Duraisamy, R. Sunku, K. Selvaraj, V. V. R. Pilla, and M. Sanikala, "Heart disease prediction using support vector machine," Multidiscip. Sci. J., vol. 6, p. 2024ss0104, Dec. 2023, doi: 10.31893/multiscience.2024ss0104.

[29] J. A. Quesada et al., "Machine learning to predict cardiovascular risk," Int J Clin Pract, vol. 73, no. 10, Oct. 2019, doi: 10.1111/ijcp.13389.

[30] J. L. Herrera Salazar, M. Talla Linderman, and O. Iparraguirre-Villanueva, “Predicting Heart Failure: A Comparative Approach between Artificial Neural Networks and Support Vector Machines,” Int. J. Onl. Eng., vol. 21, no. 01, pp. 41–55, Jan. 2025, doi: 10.3991/ijoe.v21i01.52869.

[31] A. H. Gonsalves, F. Thabtah, R. M. A. Mohammad, and G. Singh, “Prediction of Coronary Heart Disease using Machine Learning: An Experimental Analysis,” in Proceedings of the 2019 3rd International Conference on Deep Learning Technologies, Xiamen China: ACM, Jul. 2019, pp. 51–56. doi: 10.1145/3342999.3343015.

[32] M. Pal, S. Parija, G. Panda, K. Dhama, and R. K. Mohapatra, “Risk prediction of cardiovascular disease using machine learning classifiers,” Open Medicine, vol. 17, no. 1, pp. 1100–1113, Jun. 2022, doi: 10.1515/med-2022-0508.

[33] D. E. Salhi, A. Tari, and M.-T. Kechadi, “Using Machine Learning for Heart Disease Prediction,” in Advances in Computing Systems and Applications, vol. 199, M. R. Senouci, M. E. Y. Boudaren, F. Sebbak, and M. Mataoui, Eds., in Lecture Notes in Networks and Systems, vol. 199. , Cham: Springer International Publishing, 2021, pp. 70–81. doi: 10.1007/978-3-030-69418-0_7.

[34] C. M. Bhatt, P. Patel, T. Ghetia, and P. L. Mazzeo, “Effective heart disease prediction using machine learning techniques,” Algorithms, vol. 16, no. 2, p. 88, 2023.

[35] R. Hoque, M. Billah, A. Debnath, S. S. Hossain, and N. B. Sharif, “Heart disease prediction using SVM,” International Journal of Science and Research Archive, vol. 11, no. 2, pp. 412–420, 2024.