IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Heart Failure Prediction through Machine Learning

Ms.Ganga D Benal[1], Anjan Gowda S R [2], Rohith S V [3], Sushma V [4] and Varsha S[5]

[1]*Assistant Professor, Department of Computer Science and Engineering, Cambridge Institute of Technology (CITech), Bengaluru, India*

[2,3,4,5]*Student, Department of Computer Science and Engineering, CITech, Bengaluru, India*

Abstract: Machine learning has application in many different fields worldwide. It is crucial for healthcare professionals to utilize machine learning algorithms and data analysis tools in order to improve patient outcomes and provide more accurate diagnoses. Such knowledge, if anticipated well in advance, can provide physicians with vital intuitions, allowing them to adjust their diagnosis and method specific to each patient. Using machine learning techniques, we are attempting to predict potential cardiac disorders in humans. This study compares the effectiveness of variety of classifiers, comprising of the Random Forest, SVM, KNN, and logistic regression. We also provide an ensemble classifier that combines the best features of both strong and weak identified Modes to conduct hybrid classification, as it may need an abundance of features.

Keywords: Machine learning, Heart failure.

INTRODUCTION

A chronic illness called heart failure occurs when the heart muscle weakens or is injured, making it unable to pump blood as effectively as it should. This may result in symptoms like exhaustion, shortness of breath, and an accumulation of fluid in the lungs and other areas of the body. Heart failure is a dangerous condition featuring a high fatality rate, although it is frequently treatable with medicine and lifestyle modifications. The deployment of algorithms for learning enables computers to learn without explicit programming. Large amounts of evaluating data can be accomplished through ML algorithms to find trends and generate predictions. Recently, models that can forecast the risk of cardiac failure produced through machine learning.

By examining a patient's background in medicine combined with additional variables, anyone can use machine learning for this purpose. This covers characteristics like nicotine, hypertension, cholesterol levels, chronological age, sexuality and heart disease in the family history. To find the most significant components of the information that are indicative of heart failure, feature ranking algorithms are employed. This is significant because it permits creation of models for machine learning that are more precise and effective. The most crucial features can be chosen to train a model for machine learning after the features have been ranked. Heart failure can be predicted using a wide range of methods for learning through machine learning, such as KNN, logistic regression, random forests and support vector machines. Learning through machine learning enables computers to learn without explicit programming. Algorithms are employed in the field of ML to create models from data. To find the most significant information that are indicative of heart failure, feature ranking algorithms are employed. To do this, a score showing each feature's significance for predicting the target variable is calculated. The objective variable in cardiac failure prediction often the patient’s condition.

In the early days, heart failure prediction was primarily based on clinical assessments, medical history, and physical examinations. These methods lacked accuracy and often resulted in late stage diagnoses. Heart failure prognosis underwent a dramatic change with the invention of computational intelligence and machine learning in the medical field. Algorithm utilization to evaluate patient data for early detection was investigated by researchers. Electronic Health Records (EHRs) have emerged as a beneficial source of knowledge for heart failure prediction as a result of the digitization of healthcare records. EHRs could be employed by models of computational learning to find pertinent features for prediction.

The theory behind the deployment of intelligent machines in cardiac failure prediction is that specific patterns in patient data are connected to a higher chance of developing heart failure. Algorithms for learning by machine may offer that forecast a patient's likelihood of developing heart failure by finding these patterns.

LITERATURE REVIEW

Purushottam et.al proposed[1] use of decision tree structures and hill climbing algorithms, which are employed in the system for Effective Heart Disease Prediction. Depending on the dependent variables, split conditions which can be either vertical or horizontal are the foundation for the outcomes of techniques like SVM and KNN. Nonetheless, an arrangement is a decision structure that is erected upon the choices made in each tree and has the appearance of a tree with leaves, branches, and a root node. The choice tree also provides an explanation for the contents of the dataset's attributes. They exploited the Cleveland facts collection as well. The set of data has been broken into 70% training and 30% testing using various strategies. The accuracy of this technique is 91%. Bayes, the subsequent algorithm, is employed in classification. Although it is capable of managing intricate, not linear, contingent data, cardiovascular condition dataset—it is seen to be an appropriate fit because it is equally complicated, reliant and unpredictable in nature. This study gives 87% accuracy. **Sonam Nikhar et.al** proposed paper[2] "Prediction of Heart Disease Using Machine Learning Algorithms," a study[3], provides into substantial details regarding the Bayesian models and trees of decisions that are specifically utilized to forecast the probability of coronary heart disease. Studies comparing decision trees with Bayesian classifiers on the same dataset that examined the application of predictive decision-tree models found to function well using data analysis approaches. The multi-layer perceptron neural network approach is used for dataset training and testing underneath the investigation "Prediction of Heart Disease Using Machine Learning" that **Aditi Gavhane et.al** proposed. **Avinash Golande et.al**, proposed[4] "Heart Disease Prediction Using Efficient Machine Learning Approaches," which aids medical professionals in differentiating between various forms of cardiac disease incorporates several mining of data approaches. KNN, reasoning trees and naïve Bayes's are couple of prevalent strategies.

Lakshmana Rao et.al, proposed[5] "Machine Learning Methods for Heart Disease Prediction," which supplies extra details into the reasons behind heart disease. Therefore, it is difficult to differentiate cardiac disease. The degree of coronary artery problems in patients is assessed via range of neural networks techniques. **Maruf Ahmed Tamal et.al** this paper, [6] This study examines a machine learning-based procedure for predicting cardiac disease with 95% precision. Decision trees, SVM, naive bayes, random forests, logistic regression, and QDA are the approaches that are employed. SVM, which had a 95% peak accuracy, while LR and QDA (94%) obtained the second-highest classification accuracy in their investigation. The outcomes show that SVM performed 95% more precisely than any other approach. One of the system's shortcomings is that it leaves out vital variables impacting women specifically, like menopause and additional situation that might have a major impact on heart disease. **Sentilkumar et.al** In this paper, [7] Sentilkumar et al. create a feature to enhance CVD precision prediction by utilizing machine learning techniques. The hybridized linear random forest design model(HRFLM) is the name given to this method. Using the previously indicated method, the categorization of high accuracy errors, precision, F1 measure, sensitivity aredetermined. Ultimately, the author explained that the HRFLM method predicts cardiac diseases more accurately. The main drawback is the lack of real-world data sets—only knowledge and concepts are available.

Pushpavathi T P In this paper, [8] In this research, predictive methods were used to process the unprocessed information. The Neural networks with convolutions have been used to develop another deep learning model known as bug severity and priority classification, demonstrating the effective use of the system in classification. developed a multi-model system to assess deep learning techniques along with methods for ML such as KNN Naive Bayes. Conversely, though enhancing the single-model method for forecasting the bug reports' priority and severity. Lastly, an empirical study of their performance is conducted using criteria for accuracy, precision, and recall. When it came to predicting heart disease, CNN turned out to be fairly accurate. To enhance prediction methods, this research can be carried out eventually using various combinations with intelligent approaches. Moreover, novel feature selection techniques can be created to raise the standard of performance of heart failure prediction and obtain a wider understanding of the features. **Aqat ali** In this paper,[9] the system produces accurate heart failure prediction, optimized SVM stacking enhances performance. Complexity may limit real-time application, data requirements may pose challenges. However, the complexity introduced by the optimized SVM stacking raises considerations, particularly regarding real-time application. Striking a balance between intricate model architecture and the demand for swift predictions is pivotal. To guarantee that a smooth integration into clinical workflows and enable prompt interventions based on precise predictions, it becomes essential to simplify the system's complexity. In conclusion, future investigations ought to concentrate on streamlining the system's complexity for real-time application and resolving potential data-related issues, even though the suggested system shows proficiency in accurate heart failure prediction and performance enhancement through optimal SVM stacking. To fully realize the system's promise of changing proactive patient care and heart failure prediction, an iterative approach will be essential. **Abhay Kishore et.al** In this paper, [10] proposed "Heart Attack Prediction Using Deep Learning" combines a coronary artery bypass forecasting system that comprising a recurrent neural system to predict the patient's risk of heart-related infections. This approach is employed with mining of data and deep learning to produce the best accurate model with the fewest errors. This study acts as a trustworthy reference point for other heart attack prediction algorithms.

METHODOLOGY

The initial stage in running the framework is to choose the most important attributes. The relevant data is subsequently preprocessed into the appropriate format. The data is then divided into both training and testing sets. The procedures are implemented, and the training data is used to train the model. Testing the system with test data determines its accuracy. This system is enforced using the following modules.

1. Collection of Dataset
2. Selection of attributes
3. Data Pre-Processing
4. Balancing of Data
5. Failure Prediction

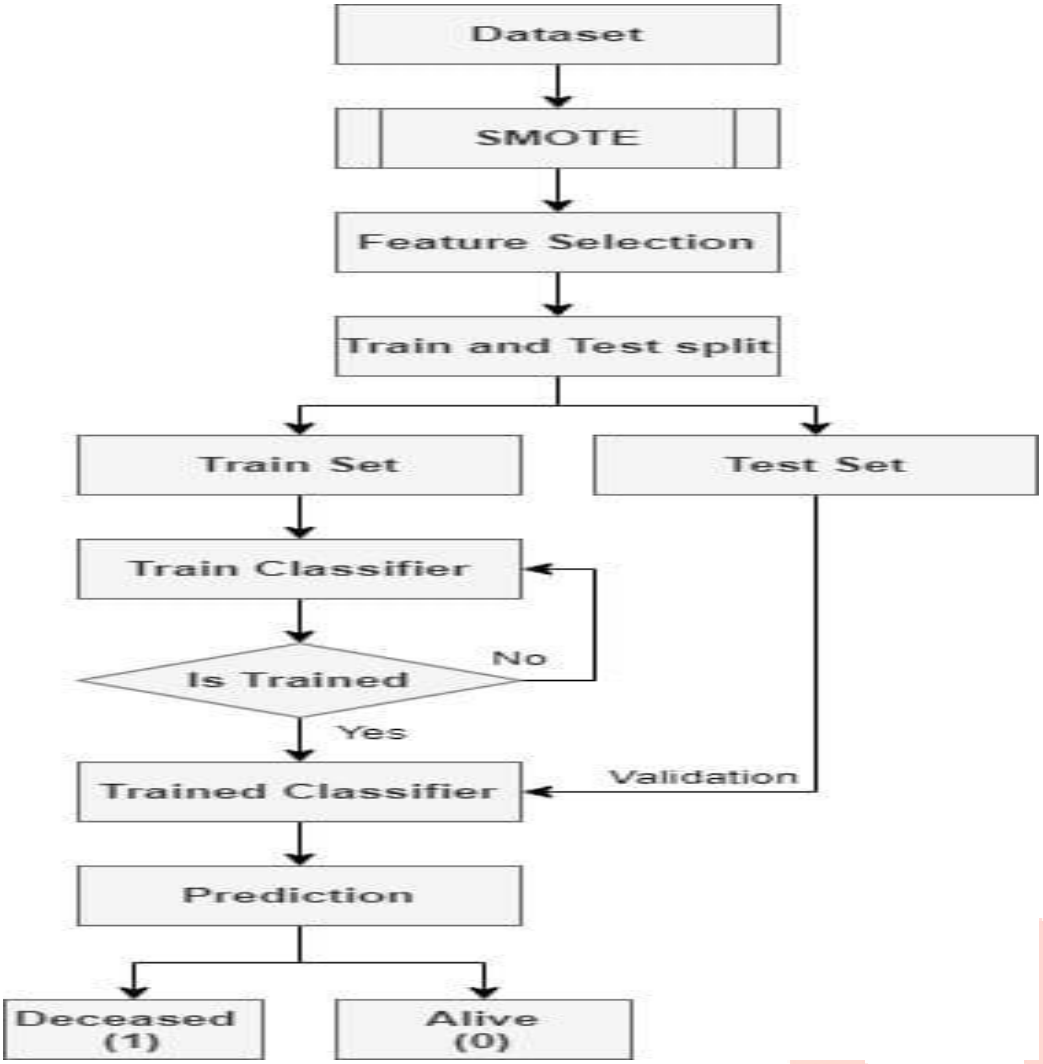

Fig 1. System Architecture

Support Vector Machine (SVM):

The famous supervised learning method referred as SVM is serves as the basis for solving classification and regression issues. However, it is primarily used in Robotic Learning Classification difficulties. The SVM algorithm's objective consists of finding the optimal line or decision boundary for dividing n-dimensional space into classes, allowing us to quickly categorize new data points in the future. The ideal decision boundary is refers as a hyperplane. The support vector machinery, supervised methods of learning serve a purpose in both regression and classification.

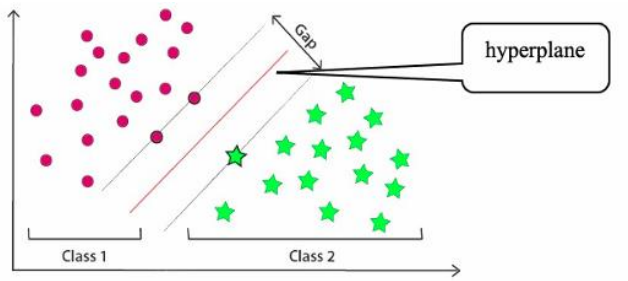

Fig 2. Support Vector Machine

Logistic Regression:

Logistic regression (LR) is a widely used supervised machine-learning technique is capable of being utilized in both classification and regression issues. The LR method uses probability to predict how categorical data will be labeled. LR's learning and prediction mechanism relies on binary classification probability measurements. The class variables in logistic regression models must be binary classified. Similar to the target column in our study dataset, which contains two separate

binary values. In the dataset, the zero represents patients who have no chance of developing heart failure, whereas the one represents anticipated heart patients.

$$z = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \ldots.$$
$$h(\theta) = g(z)$$
$$g(z) = \frac{1}{1+e^{-z}}$$

Random Forest:

Random forests, additionally referred to as random decision forests, are methods of collaborative learning that learning that generate the class-the training mode of the classes(classification) or mean prediction(regression) of the individual trees-by building a large number of decision training times and doing other tasks. The inclination of trees of choice to overfit their training set is compensated for by random decision forests.

KNN:

The controlled learning technique K-Nearest Neighbors(KNN) is utilized primarily for problems with regression and classification. KNN algorithm is non-parametric, which means it doesn't make any assumptions about the underlying data. The KNN algorithm places the new instance into a category comparable to the available classes, assuming that the new and available cases are similar. Most sample information is retrieved using the Euclidean distance metric in KNN.

RESULTS

The following section addresses our proposed exploration results and scientific validity. The creation of machine algorithms is done with the python programming language-based skit-learn library module. Our study's performance metrics include runtime computation, recall, delicacy, perfection and F1-Score. Our research models performance metrics are examined to validate scientific results. We specifically anatomized the effectiveness of the arbitrary timber (RF), Support Vector Machine (SVM), kernel neural network (KNN), and logistic regression.

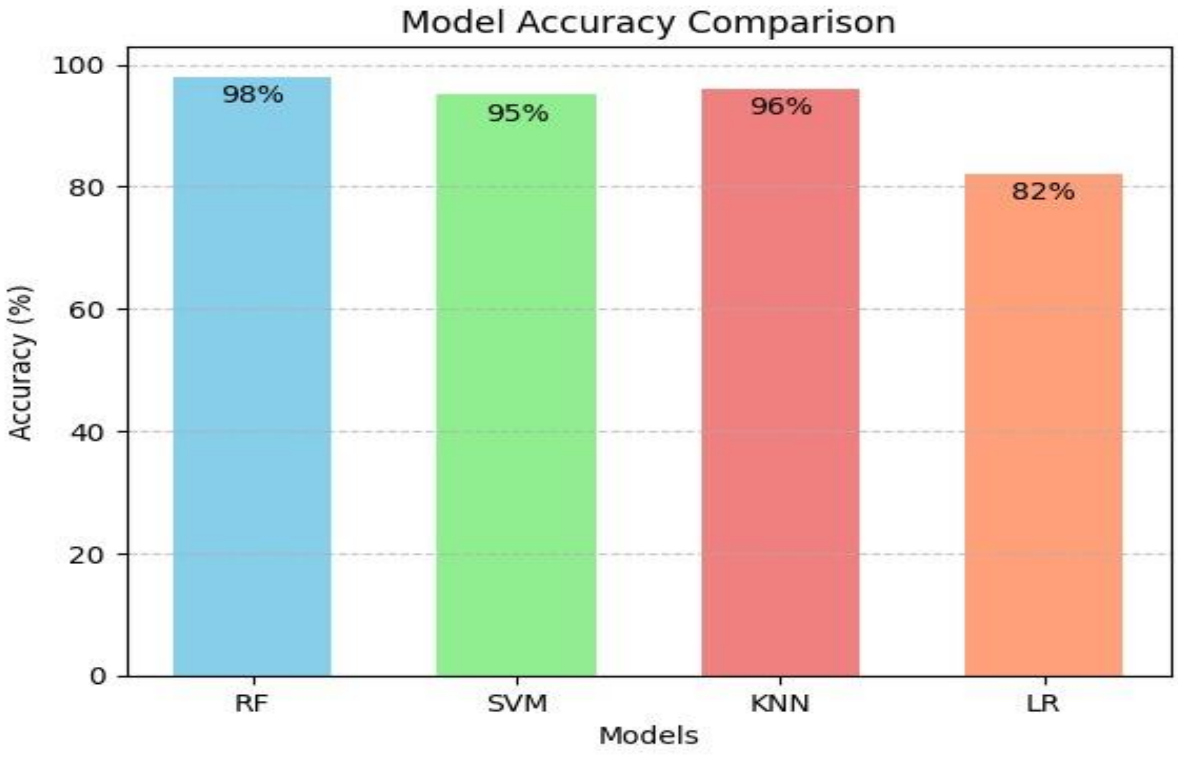

Fig 3. Comparison between various machine learning model

Fig 4. Interface that takes basic medical data

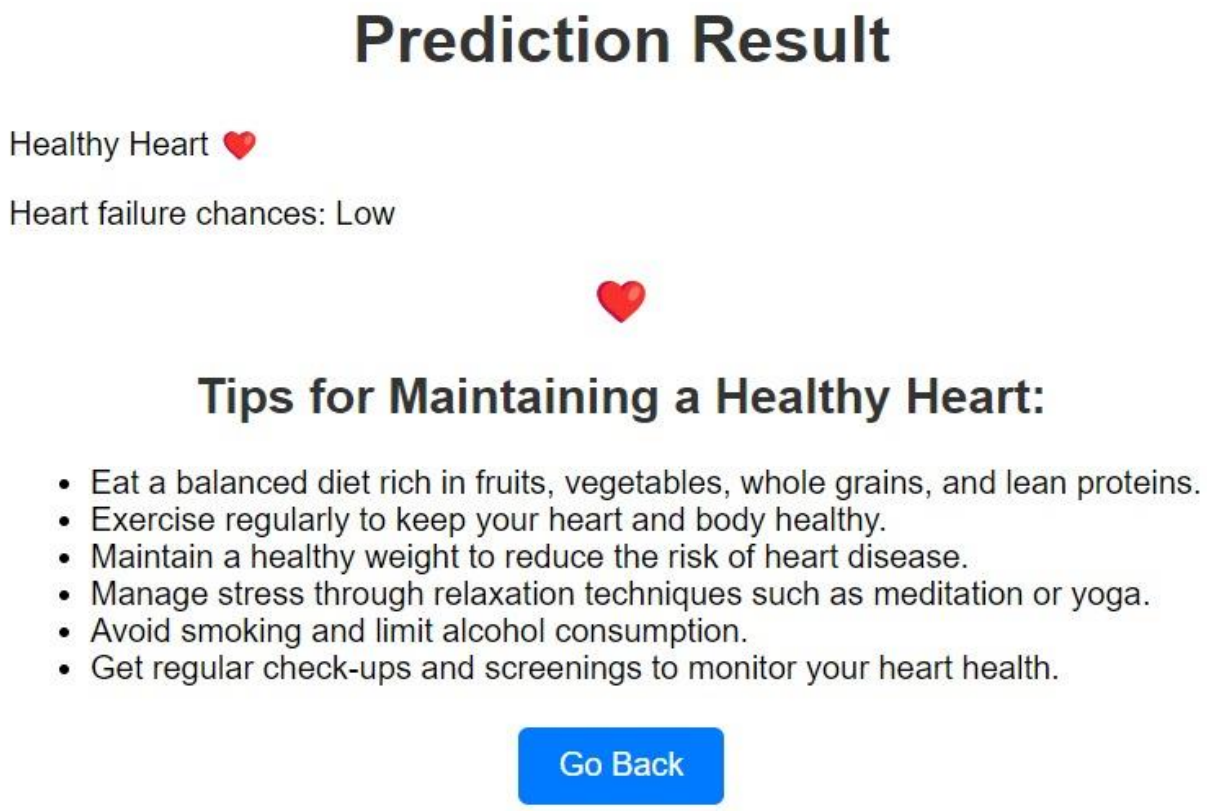

Fig 5. Output and Suggestions when chances of heart failure is low

Methods of machine learning make it easier and faster for medical authorities to make early forecasts for healthcare management objectives. As the number of cardiac fatalities rises, the machine learning paradigm can assist forecast heart failure more reliably and efficiently. This study suggests that using techniques for machine learning to make early forecasts of heart failure could improve the healthcare management system. In this experiment, A The logistic regression, random forest, support vector machine and k-nearest neighbors, are the applied leveraging a ML approaches in comparison. The proposed SVM method

The interface provides users with their cardiac dysfunction prediction results and offers tips for maintaining a healthy heart. Tips for Maintaining a Healthy Heart provides a list of recommendations for maintaining a healthy heart. A balanced diet provides the nutrients your heart needs to function properly.CONCLUSION

REFERENCE

[1] M. Gjoreski, M. Simjanoska, A. Gradišek, A. Peterlin, M. Gams, and G. Poglajen, ‘‘Chronic heart failure detection from heart sounds using a stack of Machine-learning classifiers,’’ in Proc. Int.

[2] G. Savarese and L. H. Lund, ‘‘Global public health burden failure,’’Cardiac Failure Rev., vol. 3, no. 1, p. 7, 2017.

[3] G. Savarese and L. H. Lund, ‘‘Global public health burden failure,’’Cardiac Failure Rev., vol. 3, no. 1, p. 7, 2017.

[4] E. J. Benjamin et al., ‘‘Heart disease and stroke statistics—2019 update:Areport from the American heart association,’’ Circulation, vol. 139, no. 10,pp. e56–e528, 2019.

[5] Soni J, Ansari U, Sharma D & Soni S (2011). Predictive data mining for medical diagnosis: an overview of heart disease prediction. International Journal of Computer Applications, 17(8), 43-8 .

[6] Dangare C S & Apte S S (2012). Improved study of heart disease prediction system using data mining classification techniques. International Journal of Computer Applications, 47(10), 44-8.

[7] Ordonez C (2006). Association rule discovery with the train and test approach for heart disease prediction. IEEE Transactions on Information Technology in Biomedicine, 10(2), 334-43.

[8] Shinde R, Arjun S, Patil P & Waghmare J (2015). An intelligent heart disease prediction system using k-means clustering and Naive Bayes algorithm. International Journal of Computer Science and Information Technologies, 6(1), 637-9.

[9] Bashir S, Qamar U & Javed M Y (2014, November). An ensemble-based decision support framework for intelligent heart disease diagnosis. In International Conference on Information Society (i-Society 2014) (pp. 259-64). IEEE. ICCRDA 2020 IOP Conf. Series: Materials Science and Engineering 1022 (2021) 012072 IOP Publishing doi:10.1088/1757-899X/1022/1/012072 9 .

[10] Takci H (2018). Improvement of heart attack prediction by the feature selection methods. Turkish Journal of Electrical Engineering & Computer Sciences, 26(1), 1-

[11] L. Ali et al.: Optimized Stacked SVMs Based Expert System for the Effective Prediction of HF

[12] Proceedings of the Sixth International Conference on Inventive Computation Technologies [ICICT 2023]

[13] IEEE Xplore Part Number: CFP21F70-ART; ISBN: 978-1 7281-8501-9

[14] Patel S & Chauhan Y (2014). Heart attack detection and medical attention using motion sensing device -kinect. International Journal of Scientific and Research Publications, 4(1), 1-4.