

Hybrid Machine Learning Based Heart Disease Prediction System

¹ J.Tejaswini , ²A Ramesh Babu

¹M-Tech, Dept. of CSE, J.B.Institute of Engineering and Technology, Email id

juluritejaswini.2000@gmail.com

²Associate professor, Dept. of CSE, J.B.Institute of Engineering and Technology, Email id

rameshbabu.cse@jbiet.edu.in

Abstract

Predicting heart disease is a crucial task in today's healthcare landscape due to the alarming rate of fatalities caused by heart attacks every minute. Physicians face challenges in predicting heart disease accurately given the vast amount of health records. To address this complexity, an automated heart disease prediction system is essential to alert patients and facilitate their recovery. Leveraging machine learning techniques is key to efficiently predicting heart disease using extensive data. Machine learning methods encompass both unsupervised and supervised learning classifiers. While unsupervised learning is adept at predicting with unstructured data, supervised learning, which deals with structured data, is recommended for implementing this classifier. Hence, this system utilizes supervised machine learning techniques such as KNN, RF, NN, DT, NB, and SVM classifiers. To facilitate heart disease prediction, the system employs a training dataset sourced from the UCI machine learning repository. Additionally, the system evaluates the accuracy performance of various machine learning algorithms and presents the results graphically for comparison.

Keywords:- ML , Dataset, HDP,

1. INTRODUCTION

1.1 Overview

Step by step realities the volume of wellbeing realities inside the clinical region is expanding. Consequently, it's miles advised to oversee monstrous realities units and transform them into material realities for reasonable navigation. The medical care locale is making an endeavor to clear up this issue through introducing area a mechanized contraption approach as a method for creating helpful choices from a major dataset. In this way, these styles of difficulties can be effectively settled via device learning strategies. Since it can offer down to earth ways to deal with separate basic records without the need to check a tremendous data set out. In the clinical region, clinical audits and the indications of various victims might yield significant realities for clinicians to

research. A huge amount of people as of late experience cardiovascular

breakdown signs in some vague time later on in their lives. Nonetheless, differentiating going before

Perhaps of the main requesting circumstance inside the medical care venture is properly foreseeing the standing of a visualization and offering striking contributions. Many individuals with coronary heart ailment kicked the bucket, in accordance with the polling form , despite the situation being successfully checked and controlled with the guide of a computerized machine. Here, each disorder might be overseen on the off chance that it's far situated brilliantly. Consequently, the recommended framework in this contraption can possibly figure the standing of warmth illnesses in advance of time, alarming victims and helping of their rebuilding. Clinical experts make sizeable amounts of logical realities, which they then, at that point, investigate and remove relevant records from. Most of the data inside the wellness care data set is detached, making it hard to conjecture coronary heart problem. Subsequently, inside the area of medical care

2. LITERATURE SURVEY

A machine dominating essentially based cardiovascular infirmity expectation device transformed into proposed with the guide of B. Gomathy et al. The guide diminishing strategy was utilized through the author in this method to anticipate or recognize heart jumble. This overview got a gigantic scope of clinical case models, and it utilized a guide rebate way to deal with extricate material impacted individual measurements from the enormous wellbeing data. Contrasting the exploratory discoveries with fluffy brain organizations, the past well-known shows the best precision at more than 80%. As needs be, the creator guided the use of the Guide Decrease Calculation to properly estimate heart sickness issues in clinical settings.

The forecast of cardiovascular breakdown affliction framework transformed into cutting edge through the creator

Fahd Saleh Alotaibi et al. The essayist of this strategy examines how extreme it's far to become mindful of cardiovascular breakdown in its initial reaches from the point of a heart subject matter expert. In any case, the logical area gives first asset with the choice to apply expectation designs. The main objective of this machine is to upgrade the precision of coronary heart infection disappointment expectation the utilization of coronary illness datasets. Furthermore, the essayist utilized a few calculations, like Irregular Woodland and Strategic Relapse, to figure the opportunity of coronary cardiovascular breakdown the use of schooling insights. To exhibit that the cutting-edge procedure can producing right discoveries inside the forecast of heart problem, he additionally presents examination records with different calculations.

The expectation of cardiovascular breakdown illness contraction changed into advanced through the creator Fahd Saleh Alotaibi et al. The maker of this approach examines how troublesome it's miles to see cardiovascular breakdown in its initial levels from the disposition of a cardiologist. Nonetheless, the clinical portion gives first asset with the choice to apply expectation styles. The essential objective of this gadget is to upgrade the exactness of coronary heart infection disappointment forecast the utilization of heart infirmity datasets. Furthermore, the author utilized a few

calculations, which incorporates Irregular Woodland and Strategic

Relapse, to figure the likelihood of coronary cardiovascular breakdown the utilization of schooling insights. To outline that the cutting-edge approach can creating precise discoveries in the expectation of coronary heart affliction, he moreover bears the cost of evaluation data with various calculations. Eventually, the utilization of a contraction getting to know classifier inside the

The creators, Ramya G. Franklin et al., included AES cryptography for secure realities trade at some stage in HD expectation, and they additionally utilized guileless bayes framework getting to know procedure to anticipate HD for the kind of clinical records. As per the creator, to foresee coronary illness, she amassed various capabilities from impacted individual logical measurements, which incorporate age, orientation, ECG, and various measurements. Utilizing these capacities, a gullible bayes classifier rendition changed into made, and the clinical data have been cut up into two sections: 70% had been utilized for the tutoring dataset and the last 30% had been utilized for evaluating the arrangement of rules' general presentation. In abstract, the result results of this strategy are important in determining the peril factors related with coronary heart problem.

3. PROPOSED METHOD

3.1 System Model

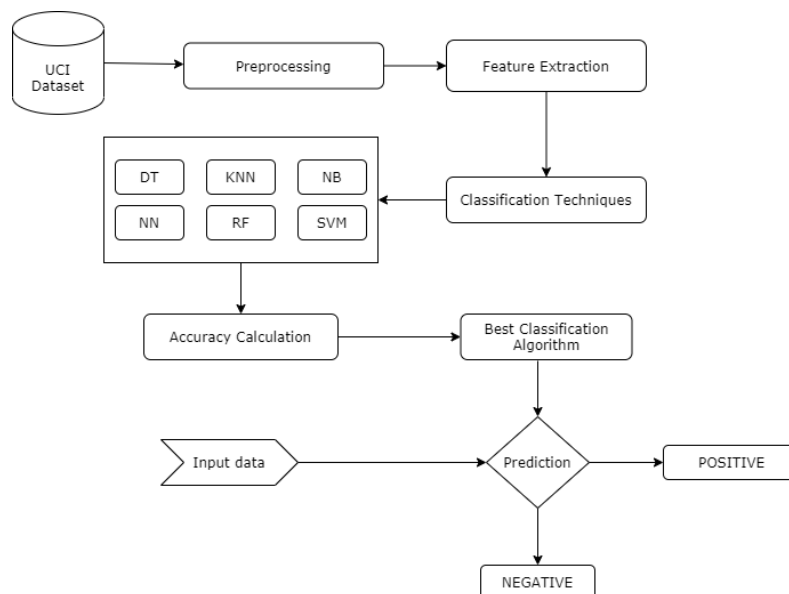


Figure.1 System Architecture

A portrayal of our forewarned machine variant is displayed in decide 1. The coronary heart infirmity schooling dataset

applied on this gadget model was gotten from the UCI library. Hence, by means of preprocessing, it can unravel the

tutoring dataset, separate the fair-minded and laid out credits through trademark extraction, after what collect a tutoring variant with a coronary heart infection characterization set of rules through offering enter records. At last, it processes the precision between six machine concentrating on classifiers.

3.2 Dataset Collection

We are utilize the UCI coronary heart problem dataset in our machine, which is gotten to through the Kaggle web storehouse and is portrayed in figure 2(<https://www.kaggle.com/ronitf/heart-disease-uci?select=heart.csv>). This schooling dataset has 303 records, of which 164 records compare to NEGATIVE illustrations or goals and 139 realities to POSITIVE preparation or targets. It likewise comprises of 14 ascribes or works, which can be referenced in Table.1.

SYSTEM TESTING

Unit checking out and validation checking out are two subcategories of system testing. The next levels offer concise causes of the two testing classes.

Unit testing

One basic or starter looking at device for programming looking at is unit looking at. The brief capability or modules may be evaluated consistently in this looking at. In this evaluating way, each the designer and the analyzer might be concerned in assessing the usefulness, investigating any issues which are found, and endeavoring to fix any bugs right on time inside the SDLC.

Validation Testing

Approval giving a shot inspects the abilities that meet up to decide if or presently not expected outcomes are gotten. These types of tests should be possible at the buyer's surrender. Instances of approving administrative work that might be used incorporate enrollment structures, login organization, and some other structure filling circumstance. Just while the approval giving a shot is finished might the individual at any point visit the accompanying step of the utility functionalities. We are utilizing this framework to approve dataset transfers, comprehensive of requiring the executive to pick the dataset before it's miles saved inside the data set.

Integration testing

One of the most critical and accommodating parts of the product testing technique is trying. This is where it can test how various functionalities' points of interaction cooperate. Furthermore, it might test different programming frameworks, working frameworks, and equipment to perceive how the framework associates.

Transparent testing

Straightforward looking at, regularly called white compartment testing, is a subset of unit testing. The product's conduct added substances will be tried, and the programming designers will manage the testing to ensure that the down to earth working float is showed.

Black-box Testing

The analyzer in this testing isn't educated about the internal functions regarding the abilities. For it to totally permit giving a shot of the capacity' primary components, including the shopper interface.

4. EXPERIMENTSL RESULTS

Table.1 Classification Metrics of Classifiers

Algorithm	Accuracy	Precision	Recall	F1_score	ROC
KNN	69	70.37	63.34	66.67	68.76
SVM	74	88.89	53.34	66.67	73.44
NB	77	80.76	70.0	75.0	76.93
NN	85	80.0	93.34	86.15	85.37
RF	98.0	98.0	97.0	98.0	98.0
DT	90.0	89.0	88.0	90.0	89.0

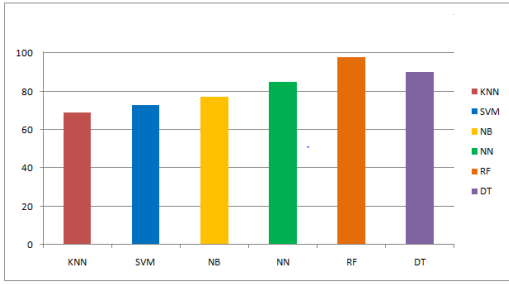


Figure.2

Accuracy Comparison between six Classifier

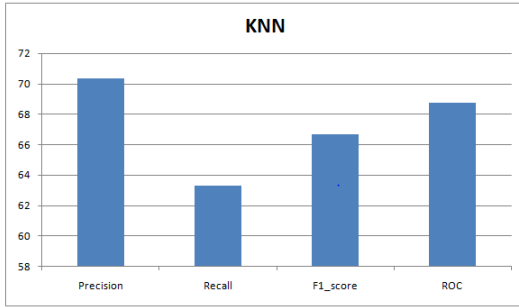


Figure.3

KNN Algorithm Metrics

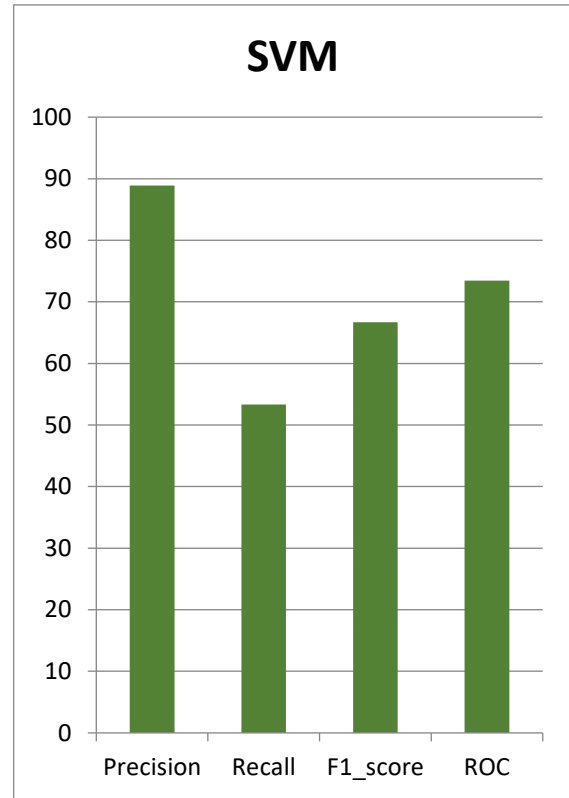


Figure.4

SVM Algorithm Metrics

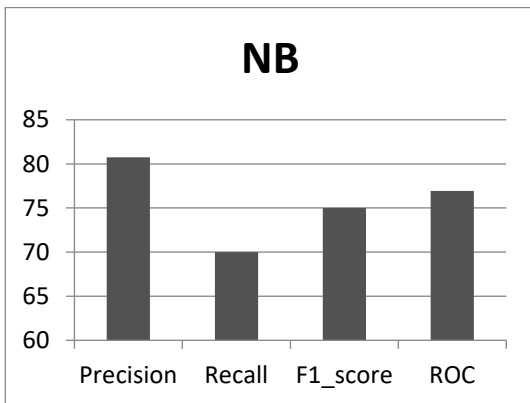


Figure.5 Naïve Bayes Algorithm Metrics

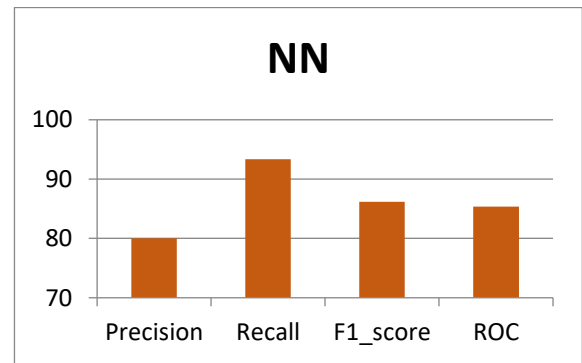


Figure.6 Neural Networks Algorithm Metrics

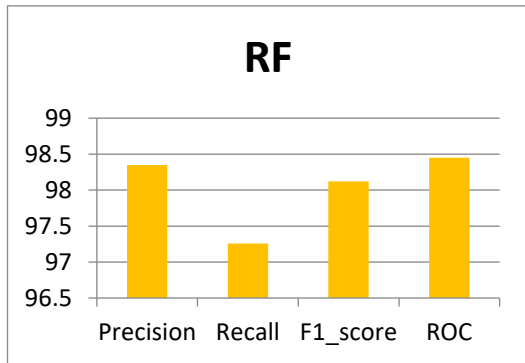


Figure.7

Random Forest Algorithm Metrics

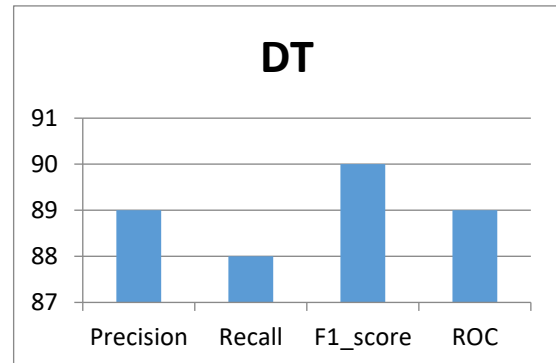


Figure.8

Decision Tree Algorithm

CONCLUSION

Identifying the medical history data, extracting and analyzing this information has become imperative for

disease prediction. Particularly in the case of heart diseases, the escalating fatality rates from heart attacks underscore the need for proactive disease prediction through the analysis of patients' medical histories. This project aims to conduct a comparative analysis of heart disease prediction using widely recognized classification algorithms. The study evaluates and compares the accuracy of these algorithms. Specifically, KNN, SVM, NB, NN, DT, and Random Forest algorithms are employed to classify heart attack medical data and calculate accuracy scores. Among these algorithms, Random Forest stands out with an impressive accuracy rate of 98%. Consequently, the Random Forest algorithm is selected for deploying user heart disease predictions.

REFERENCES

- [1] T.Nagamani, S.Logeswari, B.Gomathy, "Heart Disease Prediction using Data Mining with Mapreduce Algorithm", International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8 Issue-3, January 2019.
- [2] Fahd Saleh Alotaibi, "Implementation of Machine Learning Model to Predict Heart Failure Disease", (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 10, No. 6, 2019.
- [3] Avinash Golande, Pavan Kumar T, "Heart Disease Prediction Using Effective Machine Learning Techniques", International Journal of Recent Technology and Engineering, Vol 8, pp.944-950,2019.
- [4] Anjan Nikhil Repaka, Sai Deepak Ravikanti, Ramya G Franklin, "Design And Implementation Heart Disease Prediction Using Naives Bayesian", International Conference on Trends in Electronics and Information(ICOEI 2019).
- [5] Nagaraj M Lutimath, Chethan C, Basavaraj S Pol., "Prediction Of Heart Disease using Machine Learning", International journal Of Recent Technology and Engineering, 8,(2S10), pp 474-477, 2019.
- [6] Theresa Princy R, J. Thomas, "Human heart Disease Prediction System using Data Mining Techniques", International Conference on Circuit Power and Computing Technologies, Bangalore, 2016.
- [7] C. B. Rjeily, G. Badr, E. Hassani, A. H., and E. Andres, "Medical Data Mining for Heart Diseases and the Future of Sequential Mining in Medical Field", in Machine Learning Paradigms, 2019, pp. 71–99.
- [8] Puneet Bansal and Ridhi Saini et al. "Classification of heart diseases from ECG signals using wavelet transform and kNN classifier", International Conference on Computing, Communication and Automation (ICCCA2015).
- [9] V. Krishnaiah, G. Narsimha, and N. Subhash, "Heart disease prediction system using data mining techniques and intelligent fuzzy approach: A review," Int. J. Comput. Appl., vol. 136, no. 2, pp. 43–51, 2016.
- [10] S. Radhimeenakshi, "Classification and prediction of heart disease risk using data mining techniques of support vector machine and artificial neural network," in Proc. 3rd Int. Conf. Comput. Sustain. Global Develop. (INDIACom), New Delhi, India, Mar. 2016, pp. 3107–3111.
- [11] T. Vivekanandan and N. C. S. N. Iyengar, "Optimal feature selection using a modified differential evolution algorithm and its effectiveness for prediction of heart disease," Comput. Biol. Med., vol. 90, pp. 125–136, Nov. 2017.
- [12] J. Nahar, T. Imam, K. S. Tickle, and Y.-P. P. Chen, "Association rule mining to detect factors which contribute to heart disease in males and females," Expert Syst. Appl., vol. 40, no. 4, pp. 1086–1093, 2013.
- [13] A. Gavhane, G. Kokkula, I. Pandya, and K. Devadkar, "Prediction of heart disease using machine learning," in Proc. 2nd Int. Conf. Electron., Commun. Aerosp. Technol. (ICECA), Mar. 2018, pp. 1275–1278.
- [14] R. Das, I. Turkoglu, and A. Sengur, "Effective diagnosis of heart disease through neural networks ensembles," Expert Syst. Appl., vol. 36, no. 4, pp. 7675–7680, May 2009.