International Journal of Innovative Research in Engineering
Volume 3, Issue 3 (May-June 2022), PP: 207-210.
www.theijire.com
ISSN No: 2582-8746

Hybrid Classification Algorithm for Heart Disease Prediction

M.Kartheeswari[1], G.Ramkumar[2]
[1] *Department of Computer Science, Sri Kaliswari College (Autonomous), Sivakasi, Tamilnadu, India.*
[2] *Assistant Professor, Department of Computer Science, Sri Kaliswari College (Autonomous), Sivakasi, Tamilnadu, India*

How to cite this paper:
M.Kartheeswari[1], G.Ramkumar[2] "Hybrid Classification Algorithm for Heart Disease Prediction", IJIRE-V3I03-207-210.

Abstract: *Coronary heart disorder is the major important reason of mortality in the globe today. The gauge of coronary contamination is a most fundamental test inside the clinical records investigation sector. a few strategies are proposed to find out the effects of disorder at earlier degree that is as but getting looked at. The records mining is normally used to separate the crucial huge and desired statistics from the affected person's datasets. The few characterization strategies are utilized inside the standard techniques for the coronary contamination forecast wherein the facts mining ascribes are looked after it. on this paper, to acquire a best end result and additionally for the prediction of heart sickness in the sooner degree, the radical hybrid model proposed. The proposed, hybrid mixture of device learning algorithms Logistic regression, KNN, Random wooded area, SVM and Naive bayes Algorithms which has been analyzed that produces the excellent end result for all kinds of health related datasets. consequently, the outcome of Hybrid version affords the most effective outcomes in phrases of precision, accuracy and bear in mind with the evaluation of traditional approach.*

Key Word: *Machine Learning, Classification Algorithms, Heart disease prediction, hybrid model.*

I. INTRODUCTION

One of the significant illnesses that influence sever a individuals all through mature age is coronary illness and probably it in the end activates deadly confusions. heart problems are extra commonplace in men that in women. As consistent with measurements from who, it has been evaluated that 24% of passing's predicted to non-communicable infections in India are completed via coronary heart hardships. along those traces 33% of all worldwide passing's are because of heart problems. half of the humans inside the global's are suffered because of coronary heart disorder. round 17 million people died due to cardiovascular illness (CVD) consistently around the world, and the disease is profoundly accepted in Asia. The cardiovascular pay attention ailment Database (CHD) is viewed as the proper records base for coronary infection research.

The heart is the primary little bit of human's body. lifestyles is itself situation to extraordinary operating inside the heart. within the occasion that project of the heart isn't true; it will have an effect on the other frame segments of human, as an example, cerebrum, kidney, and so on. Coronary soreness is a problem that ramifications for the improvement of the coronary heart. there's unique segments which makes threat of coronary heart problem. it's far difficult to recognize coronary ailment due to multiple contributory risk factors, for instance, diabetes, hypertension, raised cholesterol, sporadic heartbeat price and various factors.

More than one methodologies in facts mining and neural institutions have been used to find the truth of coronary sickness among individuals. The seriousness of the sickness is grouped dependent on specific strategies like selection Tree (DT), ok-Nearest Neighbor set of rules (KNN), Naive Bayes (NB), assist Vector device (SVM), Logistic Regression (LG), Random forest (RF), mild GBM, Gaussian Naive Bayes and Genetic calculation (GA). The died of coronary illness is unpredictable and hence, the disorder must be looked after cautiously. no longer doing so can also have an effect on the coronary heart or reason surprising passing. The attitude of clinical technology and records digging are used for finding diverse kinds of metabolic situations, records mining with collecting expects a exquisite part within the gauge of coronary ailment and facts evaluation.

In this paper, the not unusual method of statistics mining is executed with the hybrid set of rules of system mastering and optimization strategies (i.e)., We considered the overall performance measures of accuracy, precision and don't forget for comparative evaluation. the very best average accuracy (88.15%), accompanied by the ensemble approach Hybrid models (88%) category algorithm. for used to development of coronary heart sickness identification inside the early level. The concept of hybrid approach i.e., aggregate of system studying techniques using weighted average. These procedures are quite concentrated on the consideration, precision and Accuracy of the sickness prediction which are compared with the traditional methods.

II.RELATED WORKS

AUTHOR NAME	PERPAR NAME	PUBLISHED YEAR	RESULT
K. embandasamy, R. Sasipriya AND E. Deepa	Heart Diseases Detection using Naïve Bayes Algorithm	2015	86.4198%
Mirpouya Mirmozaffari, Alireza Alinezhad	Data Mining Classification Algorithms for Heart Disease Prediction	2016	97.6077%
V.Ravikumar, M.Bhavani	Effective Heart Disease Prediction using Hybrid machine learning	2021	88.7%
Rohit Bharti, Adityo Khamparia, Mohammad Shabaz, GauravDhiman, Sagarpande, and Parneetsingh	Prediction of Heart Disease using a combination of Machine Learning and Deep Learning	2021	94.2%
Monther Tarawneh, Ossama Embarak	Hybrid Approach for Heart Disease Prediction for using Data Mining Techniques.	2019	89.2%
Pratiksha Shetgaonkar SREIT-Goa, Dr. Shailendra Aswale SRIEIT-Goa	Heart Disease Prediction using Data Mining Techniques	2021	85.01%
Archana Singh, Rakesh Kumar	Heart Disease Prediction Using Machine Learning Algorithms	2020	81.3%

III.PROPOSED METHODOLOGY

The heart ailment class dataset of the sufferers in kaggle UCI repository. It makes it simple to make use of visible illustrations of the dataset. operating circumstances and building predictive analytics. hybrid mixture of system learning algorithms Logistic regression, KNN, Random woodland, SVM and Naive bayes Algorithms which has been analyzed that produces the exceptional result for all sorts of fitness associated datasets. Therefore, the final results of the Hybrid version offer the optimum outcomes in phrases of precision, accuracy and recall with the evaluation of traditional methods. classification of modeling execution is done by using the proposed hybrid fashions class algorithms with improved precision.

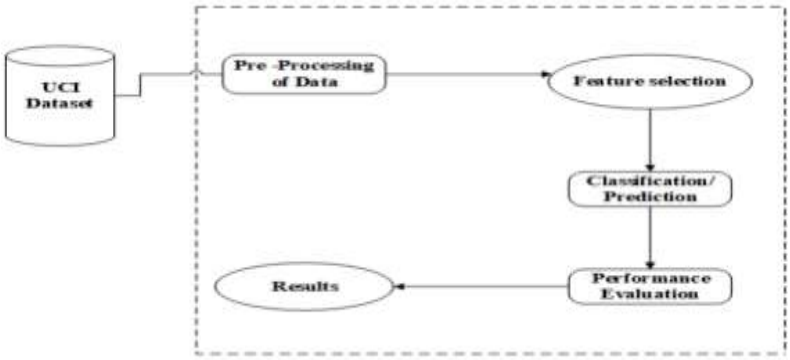

Figure 1: Proposed workflow

Data Pre processing

Datasets (Information) Pre-processing
heart disease Prediction the use of the dataset-Kaggle UCI .
The dataset consists of a total of 305 affected person information.

Table 1: UCI dataset attributes

1	Age	Age in years
2	Sex	Male or female
3	Cp	Chest pain type
4	Thestbps	Resting blood pressure
5	Chol	Cholestral
6	Restecg	Resting electrographic results
7	Fbs	Fasting blood sugar
8	Thalach	Maximum heart rate achieved
9	Exang	Exercise induced gain
10	Oldpeak	ST depression induced by exercise relative to rest

11	Solpe	Peak exercise slope ST segment
12	Ca	Major vessels colored by floursopy
13	Thal	Defect type
14	Num	Diagnosis of heart disease

Feature Selection

Function choice is essential for device learning (ML). The function willpower techniques are the order exactness and reduce the version execution time.

To include dedication on this framework, this k-Nearest Neighbor (KNN) algorithm is used.

Classification

In the traditional technique, the classifiers are utilized for specific prediction. Anyways, it has been analyzed that there may be no single classifier that gives the best final results for each dataset and now not a single records mining approaches which provide precise consequences to cardiovascular infection related statistics. In this manner, a set of rules is needed that can give optimum outcome.

For this class the proposed techniques can have a classifier on the way to hybrid with the optimization strategies. This is inspired from the literature assessment wherein it's far classified as an optimization technique which can improve the additives of the classifiers to enhance their category belongings. As an instance in KNN technique, if the beginning load of the business enterprise are refreshed, the order precision is moreover going to vary. consequently those weight esteemed may be upgraded making use of numerous optimization algorithms.

Those findings motivate the proposed paintings to provide a hybrid version as a fine classifier.

Hybrid Model

In Machine learning, the combining of model is done by using two approaches “Ensemble models” and “hybrid models”. Ensemble models use multiple machine learning algorithms to bring out better predictive results, as compared to using a single algorithm.

In this task, the five different types of machine learning models are used as weak learners to build a hybrid ensemble learning model. These models are – Logistic Regression Model, Decision Tree, Support Vector Machine, K-Nearest Neighbor Model, and the Naive Bayes Model. The term hybrid is used here because, in other ensemble models, a homogeneous collection of weak learners is used but in this task, a heterogeneous collection of weak learners is used.

Algorithm

Step 1: Collection of input data from UCI to train the software.
Step 2: Read the input data.
Step 3: Applying KNN algorithm for feature selection.
Step 4: Applying the Five algorithms (KNN, Logistic Regression, Support Vector Machine, Naïve Bayes, and Decision Tree) for selected classifier factor updation.
Step 5: Retrain the network to improve accuracy.
Step 6: Get final results from proposed hybrid model.

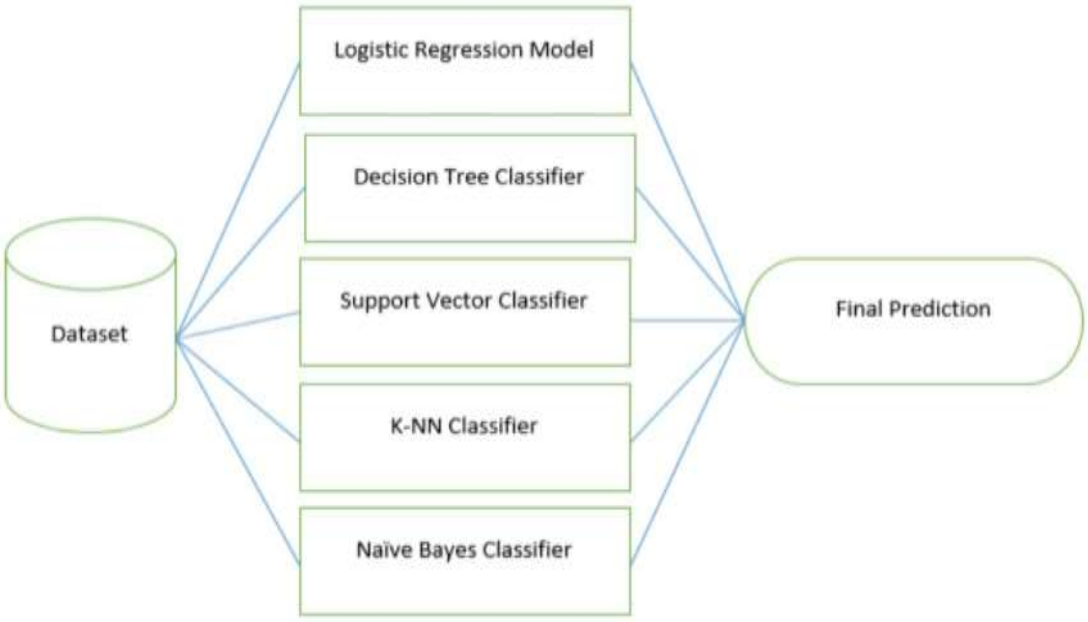

Figure 2 : Using Hybrid Models

The hybrid ensemble learning model, built using these weak learning models, is applied in the task of

classification for the bank's customer churn modelling.

O/P

The Hybrid model score is: 88.1578947368421

	precision	recall	f1-score	support
0	0.90	0.82	0.86	33
1	0.87	0.93	0.90	43
accuracy			0.88	76
macro avg	0.88	0.87	0.88	76
weighted avg	0.88	0.88	0.88	76

Accuracy for training set for Hybrid model=0.920704845814978

Accuracy for testing set for Hybrid model=0.881578947368421

IV.CONCLUSION & FUTURE WORK

Coronary heart disease prediction is the maximum common motive of demise in the international world currently. It is essential to predict coronary heart issues in advance so that the disorder may be prevented. Much preceding research is used for the coronary heart disease prediction namely SVM, KNN, RF, LG AND NAIVE BAYES where the green overall performance is executed via the use of KNN. even though it isn't enough for all kinds of datasets. Consequently, a few novel hybrid version classifiers are provided on this paper for which the assessment for the Algorithms. This hybrid version is used for locating the optimum outcomes that can be accomplished for the coronary heart infection facts. The proposed classifier is in comparison with the traditional classifiers with recognition to three parameters i.e., accuracy, precision and recall. From the obtained results of this paper, the proposed hybrid model classifier is efficient in terms of accuracy, precision and don't forget which offers the choicest solutions. Those type methods can be additionally improved via growing the quantity of attributes for the more specific prediction to be carried out. There are many likely improvements to decorate the scalability and precision of this prediction method that can be implemented in future.

References

1. Khan, M., Ding, Q., & Perrizo, W. (2002). K-nearest neighbor classification on spatial data streams using p- trees. In Advances in Knowledge Discovery and Data Mining (pp. 517-528). Springer Berlin Heidelberg.

2. Lavanya, D., & Rani, K. U. (2012). Ensemble decision tree classifier for breast cancer data. International Journal of Information Technology Convergence and Services (IJITCS), 2(1), 17-24.

3. Picek, S., Golub, M.: On the Efficiency of Crossover Operators in Genetic Algorithms with Binary Representation. In: Proceedings of the 11th WSEAS International Conference on Neural Networks (2010)

4. Rish, I. (2001). An empirical study of the naive Bayes classifier. In IJCAI 2001 workshop on empirical methods in artificial intelligence (Vol. 3, No. 22, pp. 41-46). IBM New York.

5. Nidhi Bhatla Kiran Jyoti, An Analysis of Heart Disease Prediction using Different Data Mining Techniques, International Journal of Engineering Research & Technology (IJERT), 2012

6. Mohammad Taha Khan, Dr. Shamimul Qamar and Laurent F. Massin, A Prototype of Cancer/Heart Disease Prediction Model Using Data Mining, International Journal of Applied Engineering Research, 2012.

7. K. embandasamy, R. Sasipriya AND E. Deepa Heart Diseases Detection using Naïve Bayes Algorithm
IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 2 Issue 9, September 2015.

8. MirpouyaMirmozaffari, Alireza Alinezhad Data Mining Classification Algorithms for Heart Disease Prediction Conference Paper · November 2016.

9. V.Ravikumar, M.Bhavani Effective Heart Disease Prediction using Hybrid machine learning ISSN NO:0377-92 Vol 12, Issue 12, Dec/2021.

10. Rohit Bharti, AdityoKhamparia, Mohammad Shabaz, GauravDhiman, Sagarpande, and Parneetsingh Prediction of Heart Disease using a combination of Machine Learning and Deep Learning Received 16 May 2021; Revised 15 June 2021; Accepted 21 June 2021; Published 1 July 2021.

11. Ishtake S.H ,Prof. Sanap S.A., "Intelligent Heart Disease Prediction System Using Data Mining Techniques", International J. of Healthcare & Biomedical Research,2013.

12. Dr. K. Usha Rani,analysis of heart diseases dataset using neural network approach,International Journal of Data Mining & Knowledge Management Process, 2011.

13. Y.S.Kumaraswamy and Shantakumar B.Patil "Intelligent and Effective Heart Attack Prediction System Using Data Mining and Artificial Neural Network‖"; European Journal of Scientific Research, ISSN 1450-216X Vol.31 No.4, 2009.

14. Dilip Roy Chowdhury, Mridula Chatterjee & R. K. Samanta, An Artificial Neural Network Model for Neonatal Disease Diagnosis, International Journal of Artificial Intelligence and Expert Systems (IJAE), Volume (2): Issue (3), 2011.

15. Milan Kumari, Sunila Godara, Comparative Study of Data Mining Classification Methods in Cardiovascular Disease Prediction, IJCST Vol. 2, Iss ue 2, June 2011