Prediction Analysis on Heart Disease using HNB and NB Techniques

Dr.I.Lakshmi

Department of Computer Science, Stella Maris College, Chennai-600086

Abstract

Data mining is the altogether limitless territory for Scrutinize field. Human services may be the greater part vital association to our planet. The Different information mining systems would used to anticipate the coronary illness. Coronary illness will be exceptionally risky malady for our human body. Heart is paramount some piece to our physique. Information mining prediction device is assuming around key part previously, social insurance. This paper Investigation those Different strategy will anticipate the coronary illness.

Key Words: *Data mining, heart disease, classification algorithm*

I. INTRODUCTION

Those heart is the vast majority noteworthy bulky organ done humans, which pumps blood through those blood vessels of the circulatory disappointment. An aggregation may be fully relying upon the heart. Though the heart influences the disease, it will likewise influence alternate organs of the human body. Information mining may be the methodology about extracting the PC based majority of the data from expansive set for database. Information mining strategies Also devices would utilized for Different association. Information mining instruments need aid used to foresee that ailment done social insurance field. As stated by those WHO accounted similarly as 12 million people groups demise happen in the coronary illness. Over restorative association have manual records for heart tolerant subtle elements. Medicinal specialist needs best electronic records. Information mining strategies will be effectively change over the manual records wills electronically records. There would amount for elements on increment the coronary illness.

- ➢ Hyper tension
- ➢ Poor diet
- ➢ Obesity
- ➢ High blood pressure
- ➢ High cholesterol
- ➢ Family history

II. TYPES OF HEART DISEASE

The disease of heart and blood vessels within it. The various types of heart disease given below. They are

- ➢ Cardiovascular disease
- ➢ Arrythametic disease
- ➢ Coronary artery disease
- ➢ Disease of Heart valves

III. DATA MINING ALGORITHM

A. Classification

Arrangement may be dependent upon machine taking in. Order is used to arrange each thing over a situated about information under a standout amongst predefined situated of classes alternately gatherings. Arrangement method makes utilization of scientific strategies for example, choice trees, straight programming, and neural system.

B. Prediction

Prediction is an information mining method that uncovers those relationships the middle of free variables What's more relationship Around indigent Also autonomous variables.

IV. RESEARCH OBJECTIVES

The majority healing facilities support an clinic data framework. This framework holds an extensive amount for tolerant information. This data will be generally not accessed. Information mining strategies need aid used to change over An information under suitable data. The fundamental goal from claiming this Examine may be produced a one information mining displaying strategy to be specific similarly as moved forward concealed naïve bayes. It cam wood extricate the concealed learning starting with coronary illness chronicled dataset. V.Hidden Naive Bayes

Stowed away Naïve bayes is dependent upon machine Taking in What's more information mining strategies. It the entire more complex orders system. It may be used to make those prediction models. It indicates the predictable state for likelihood of every quality. This model produces a a greater amount effective yield think about with different yield. The primary point for naïve bayes obliges a little measure for preparing information on evaluate those parameter for order. It is skilled of figuring the A large portion possible yield rely on upon those enter. It will be not difficult should include new information during runtime to a exceptional classifier.

A. Steps

- ➢ Convert the dataset into frequency table.

➢ Create likelihood table by finding the probability.
➢ Hidden Naïve bayes to calculate the posterior probability of each class.

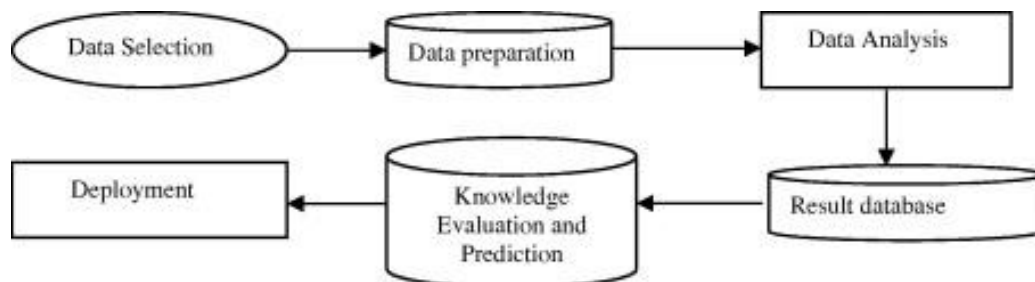➢ The class with highest priority probability is the outcome of prediction

Fig1. Naïve Bayes implementation diagram

B. Input Attributes

The dataset are available on UCI machine repository. This dataset contains 303 records.

1. Age
2. Gender
3. CP (chest pain)
4. treetops : resting blood pressure
5. cholesterol
6. fbs: fasting blood sugar>120 ? yes=1,no = 0
7. restecg: resting electrocardiographic results 0,1,2
8. thalach : maximum heart rate achieved
9. exang : exercise induced angina (1= yes ; 0=no)
10. old peak = ST depression induced by exercise relative to rest
11. slope : the slope of the peak exercise ST segment
12. ca: no. of major vessels (0 to 3) colored fluoroscopy
13. thal :3 =normal ,6=fixed defect ,7= reversible defect
14. diagnosis of heart disease

C. Proposed Method

This proposed method is used to improve the naïve bayes performance. This method takes only two values for prediction. the following steps are implemented for this method.

➢ Enter the patient dataset
➢ we have to classify the data as two classes.
➢ 0-Ahsent   1-Present
➢ Find the probability of each attribute of both classes.
➢ Calculate the probability of attribute.
➢ Proposed algorithm is used for identify where values ranges from 0 to 1.
➢ This value is used for better performance.
➢ Calculate the maximum probability for both classes.
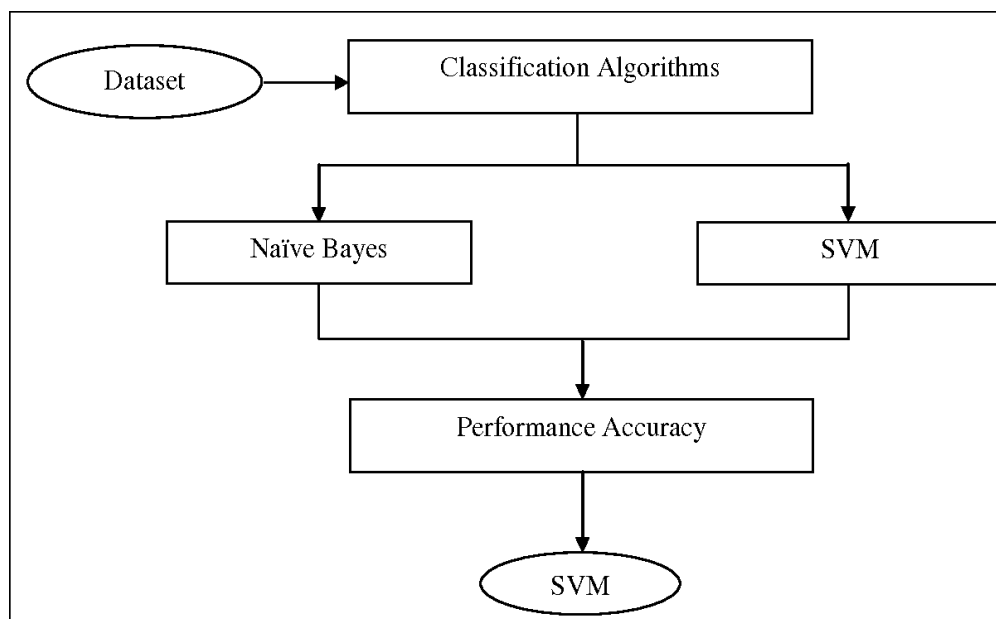➢ Decide the class for patient record.

Fig 2. Proposed method diagram

VI. RESULT ANALYSIS

The following table provides the proposed algorithm result

Sl. no	Data set	Instances	Number of Attributes
1	Heart disease	270	14

Table 1: Heart disease data set information

Data set	Instances cross validation (10 fold)	
Heart disease data set	Test	Train
	27	243

Table 2: Heart disease stalog data statistics

		Predicted	
		Yes	No
Actual	Yes	TP	FP
	No	FN	TN

Table 3: Confusion matrix

MEASURE	HNB		NB	
	Cross validation	Full training	Cross validation	Full training
Sensitivity	90.56	92.5	79.16	80
Specificity	86	96	85.6	87.85
Accuracy	83.3	94.4	83.70	85.18
PPV	82.0	94.8	83	85
NPV	81	94.1	88	84
TPR	80	92.5	79	80.83
FPR	14	4	14	12.1

Table 4: Heart disease for data set analysis of various algorithms

Sl.no	Author	Method	Accuracy
1	Peter john [12]	CFS+NB	85.18
2	Maishouman[13]	NB	84.5
3	Rupali [14]	NB	78
4	Elma [15]	KNN+NB	82.96
5	**Our approach**	**HNB+IQR**	**100**

From the table that the proposed new algorithm gave an highest accuracy value of 97%.Naive bayes algorithm accuracy value as 89%.In this above two algorithm result ,the newly one achieve a highest accuracy.
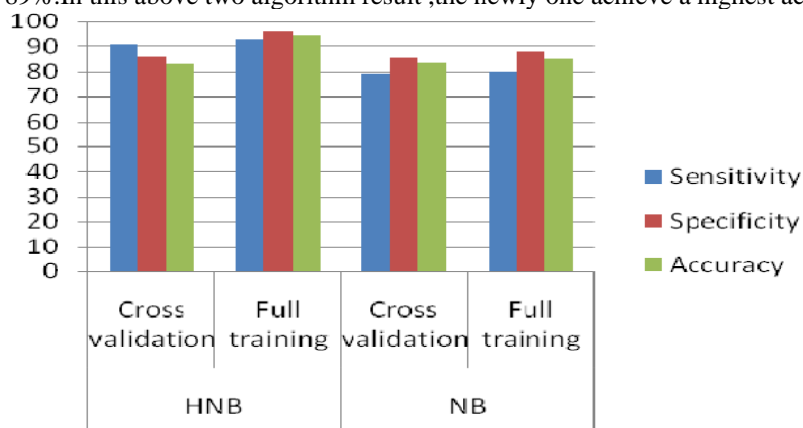

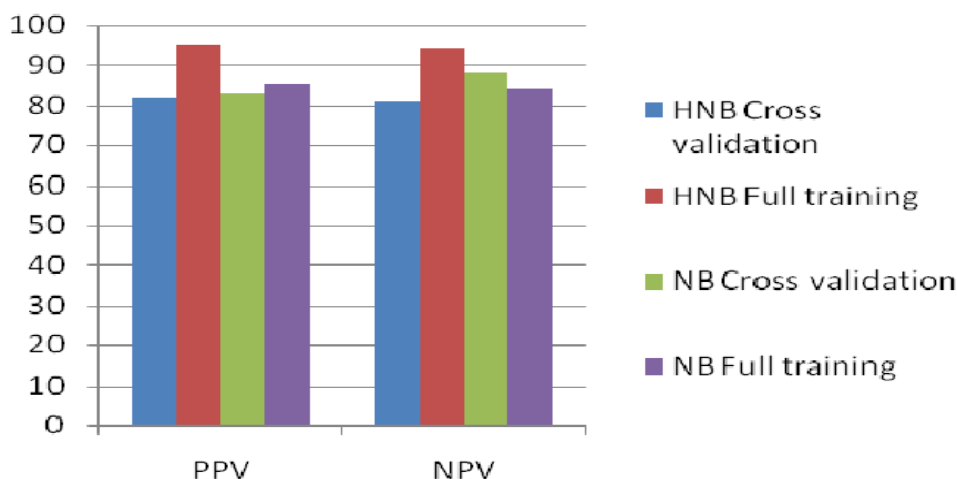**Fig 3. Sensitivity, Specificity and Accuracy comparison of various algorithms**

Fig 4: PPV and NPV calculation for HNB and NB

VII. CONCLUSION

Heart disease is one of the leading causes of deaths worldwide and the early prediction of heart disease is very important. In this study prove that the proposed new algorithm achieve a highest accuracy compare with other algorithm.

REFERENCES

[1] J.Vijayashree and N.Ch.SrimanNarayanaIyengar Heart Disease Prediction System Using Data Mining and Hybrid Intelligent Techniques: A Review International Journal of Bio-Science and Bio- Technology Vol.8, No.4 (2016),

[2] S.Kiruthika Devi* , S. Krishnapriya and Dristipona Kalita Prediction of Heart Disease using Data Mining Techniques Indian Journal of Science and Technology, Vol 9(39), DOI: 10.17485/ijst/2016/v9i39/102078, October 2016

[3] Mr.K.AravinthanA Novel Cluster and Rank Based Method for Prediction of Heart Diseases International Journal of Advanced Information Science and Technology (IJAIST)

[4] V.Krishnaiah Heart Disease Prediction System using Data Mining Techniques and Intelligent Fuzzy Approach: A Review International Journal of Computer Applications (0975 – 8887) Volume 136 –No.2, February 2016

[5] Syed Immamul Ansarullah Heart Disease Prediction System using DaMining Techniques: A study international Research Journal of Engineering and Technology (IRJET) Volume: 03 Issue: 08 | Aug-2016

[6] Sonam Nikhar "prediction of heart disease using data mining techniques" - A Review International Research Journal of Engineering and Technology (IRJET) Volume: 03 Issue: 08 | Aug-2016.

[7] S.B.Bhalerao1, Dr. B.L.Gunjal2 "Survey of Heart Disease Prediction Based on Data Mining Algorithms", ijariievol-2 Issue-2, 2016.

[8] Umair safique "Data mining in healthcare for heart disease" International Journal of Innovation and Applied Studies,10 March 2014.

[9] Ms.Shide Swati B "Decision support system on heart disease using data mining techniques "International journal of engineering research and general science volume 3March-April 205.

[10] Jyoti Soni" Predictive Data Mining for Medical Diagnosis: An Overview of Heart Disease Prediction "International Journal Computer Applications, 17, 8, March 2011.