ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 6, June 2019

Analysis of Different Types of Arrhythmia Using Ensemble Model

Harshitha. R [1], Dr. Suresh .L [2]

M.Tech. Student, Department of CSE, Cambridge Institute of Technology, Bangalore, India[1]

Principal, Cambridge Institute of Technology, Bangalore, India[2]

ABSTRACT: Arrhythmia is a condition in which the human heart beats with an irregular or abnormal rhythm which may be immediately fatal or cause irreparable damage to the heart. This type of a fatal abnormality would require a model which gives high accuracy and takes short periods of time for classification.
In this project, a high dimensional clinical dataset consisting of Electrocardio (ECG) results of patients is used by individual classifiers like Multi-Layer-Perceptron(MLP), Support Vector Machine (SVM) and Naïve bayes to classify them into different types of Arrhythmia.

KEYWORDS: NA values, Arrhythmia, Ensemble Model, Accuracy.

I. INTRODUCTION

An Arrhythmia describes an irregular heartbeat - the heart may beat too fast, too slowly, too early, or irregularly. Arrhythmias occur when the electrical signals to the heart that coordinate heartbeats are not working properly. Many heart Arrhythmias are harmless however, if they are particularly abnormal or result from a weak or damaged heart, Arrhythmias can cause serious and even potentially fatal symptoms.

In this paper by using Multilayer perceptron (MLP), Support vector machine (SVM) and Naïve bayes classifiers gives accuracy which are comparatively low. Ensemble methods are techniques that create multiple models and then combine them to produce improved results. They produce more accurate solution than a single model would. Thus an Ensemble model is built which gives higher accuracy but takes more time for classification compared to individual models. The same Ensemble model is run on an environment called Anaconda where it reduce the execution time. Classifying Arrhythmia is a great significance to reduce the uncertainty for doctors on deciding the optimum treatment for patients

II. LITERATURE SURVEY

Liangyuan Li, Mei Chen, Hanhu Wang presents an analysis of the importance of cost sensitive methodology in medical disease prediction .And C5.0 based cost sensitive ensemble approach is presented to support improving the effectiveness and decreasing the misdiagnosis cost in the same time. The verification is conducted on the real case from changhai hoapital renal registered (CHRR) data which consists of 483 records and 25 fields are used through variables selection. The real case expiriments show that their approach works effectively and could be used as an assistant approach for peritonitis analysis in some circumstances.

Hamiderza Ashrafi Esfahani, Morteza Ghazanfari studying the data on cardiovascular patients collected from UCI laboratory is utilized for applying discovery patterns algorithms including decision tree, neural networks, rough set, SVM, naïve bayes, and compare their accuracy in prediction. Finally they proposed a hybrid algorithm to increase

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 6, June 2019

the accuracy of these algorithms. Based on the results the proposed hybrid methods achieved an F-measure of 86.8% which outperforms other competing methods.

III. METODOLOGY

The flow of our design is been shown in the below diagram (Figure 1). The goal for all the data collection is to capture quality evidence that allows analysis to lead to the formulation of convincing and credible answers to the questions that have been proposed. Health survey, administrative enrollment and billing records and medical records used by various entities including hospitals, CHCs, and health plan. We are using this data as it has multiple attribute about 279 with over four hundred records which help us to provide quality.

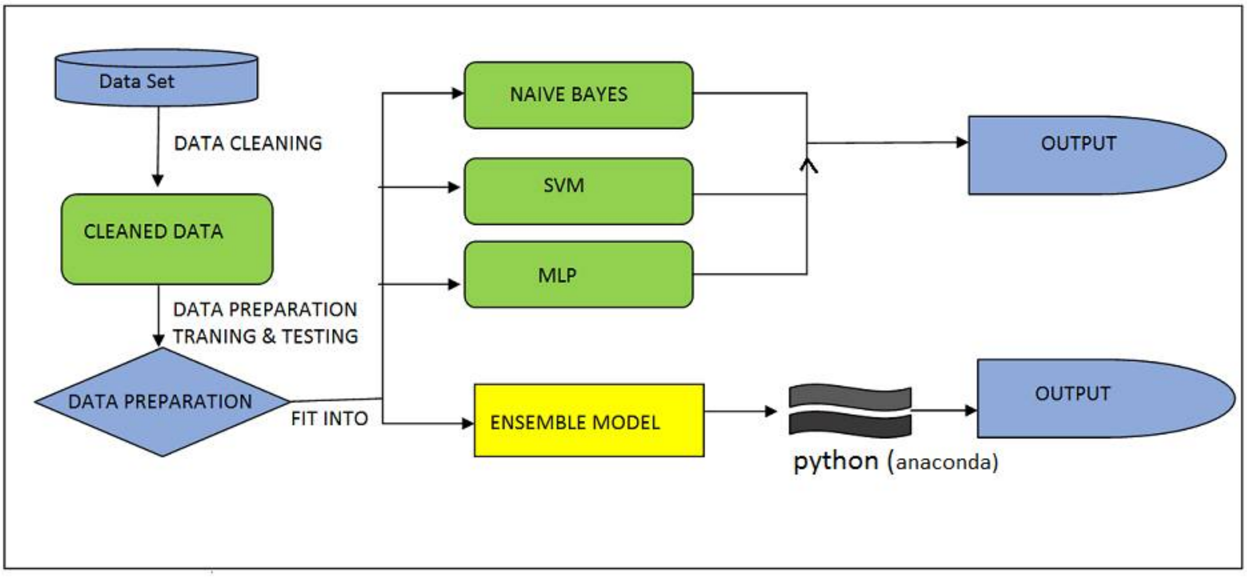


Figure 1: Block diagram of proposed system

The original data being used is UC Irvine machine learning repository which has a data set containing Arrhythmia information for 452 patients (row) .Each of these patients has 279 features (columns) and was classified into one of 16 categories which is 15 abnormal and 1 normal heart conditions. Our feature space comprised of 279 dimensions including patients information such as age, sex and channel information.

A. NAÏVE BAYES

Naive bayes is a simple technique for constructing classifiers: models that assign class labels to problems instances, represented as vectors of features values, where the class lables are drawn from some finite set. All naïve bayes classifiers assume that value of a feature is independent of the value of any other feature giving class variable. Naïve Bayes classifiers are highly scalable, requiring several parameter linear in the number of variables (features/predictors) in a learning problem.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

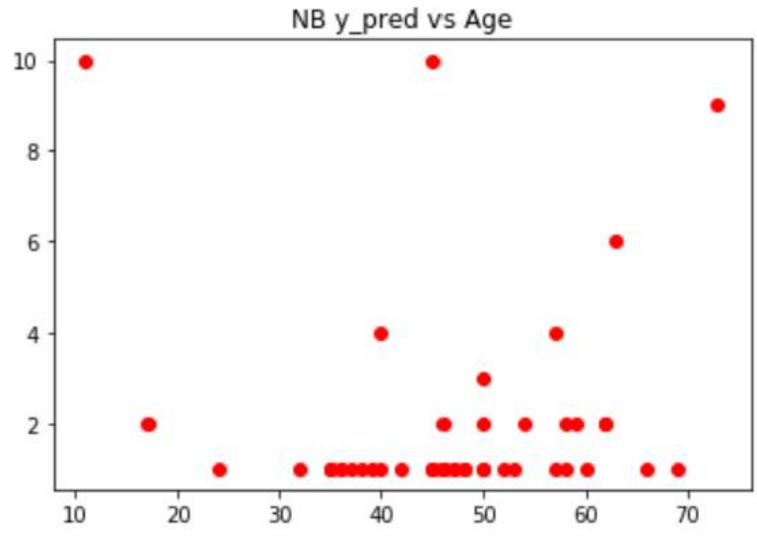
Website: www.ijircce.com

Vol. 7, Issue 6, June 2019

classification report naive bayes

	precision	recall	f1-score	support
1.0	1.00	1.00	1.00	31
2.0	0.71	1.00	0.83	5
3.0	1.00	0.33	0.50	3
4.0	1.00	1.00	1.00	2
6.0	1.00	1.00	1.00	2
9.0	1.00	1.00	1.00	1
10.0	1.00	1.00	1.00	2
micro avg	0.96	0.96	0.96	46
macro avg	0.96	0.90	0.90	46
weighted avg	0.97	0.96	0.95	46

Figure 2. Classification report naive bayes




```
Accuracy 95.65217391304348
execution time of naive bayes model 0.34110093116760254 seconds ---
```

Figure 3: Accuracy of naive bayes

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 6, June 2019

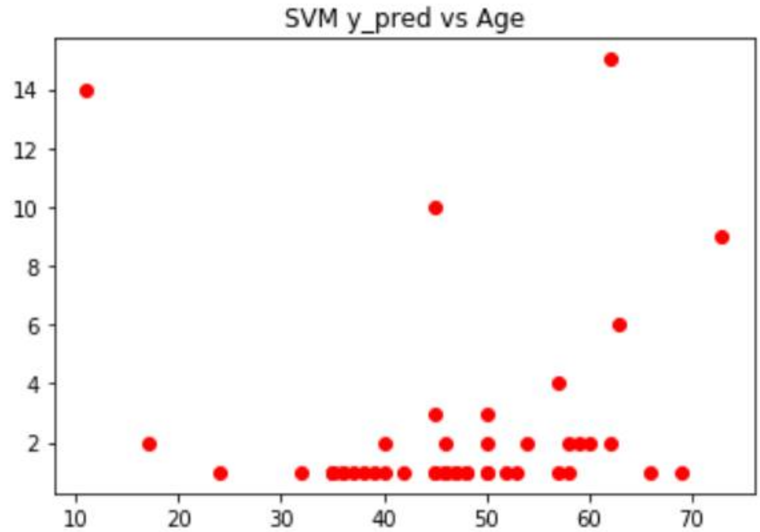
B. SUPPORT VECTOR MACHINE (SVM)

A Support Vector Machine (SVM) is a discriminative classifier formally defined by a separating hyperplane. In other words, given labeled training data (*supervised learning*), the algorithm outputs an optimal hyperplane which categorizes new examples. In two dimentional space this hyperplane is a line dividing a plane in two parts where in each class lay in either side.

classification report SVM

	precision	recall	f1-score	support
1.0	1.00	0.94	0.97	31
2.0	0.44	0.80	0.57	5
3.0	1.00	0.67	0.80	3
4.0	1.00	0.50	0.67	2
6.0	1.00	0.50	0.67	2
9.0	1.00	1.00	1.00	1
10.0	1.00	0.50	0.67	2
14.0	0.00	0.00	0.00	0
15.0	0.00	0.00	0.00	0
micro avg	0.85	0.85	0.85	46
macro avg	0.72	0.54	0.59	46
weighted avg	0.94	0.85	0.87	46

Figure 4: Classification report of SVM



SVM accuracy 84.78260869565217
execution time of svm model 0.5079300403594971 seconds ----

Figure 5: Accuracy of SVM

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 6, June 2019

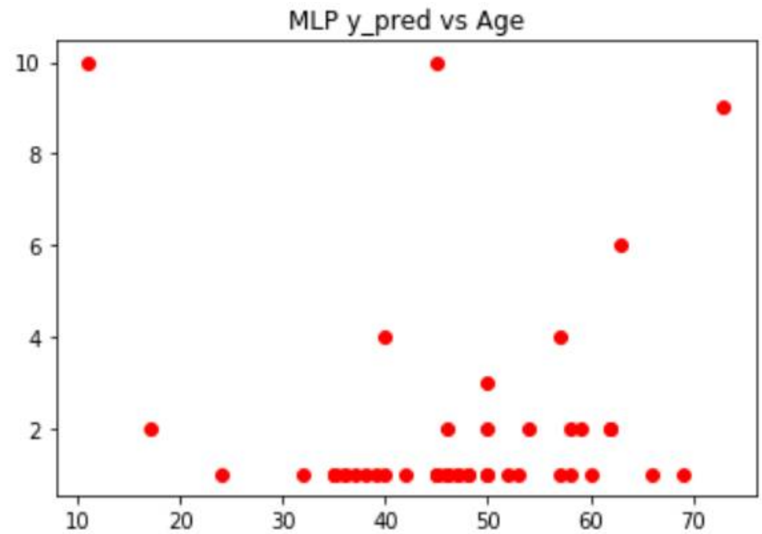
C. MULTILAYER PERCEPRTON (MLP)

An MLP can be viewed as a logistic regression classifier where the input is first transformed using a learnt non-linear transformation. This transformation projects the input data into a space where it becomes linearly separable. This intermediate layer is referred to as a hidden layer. A single hidden layer is sufficient to make MLPs a universal approximator.

Classification report MLP

	precision	recall	f1-score	support
1.0	0.97	0.97	0.97	31
2.0	0.62	1.00	0.77	5
3.0	1.00	0.33	0.50	3
4.0	1.00	1.00	1.00	2
6.0	1.00	0.50	0.67	2
9.0	1.00	1.00	1.00	1
10.0	1.00	1.00	1.00	2
micro avg	0.91	0.91	0.91	46
macro avg	0.94	0.83	0.84	46
weighted avg	0.94	0.91	0.91	46

Figure 6: Classification on MLP



Accuracy 0.9130434782608695
execution time of MLP model 0.3730125427246094 seconds ---

Figure 7: Accuracy of MLP

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

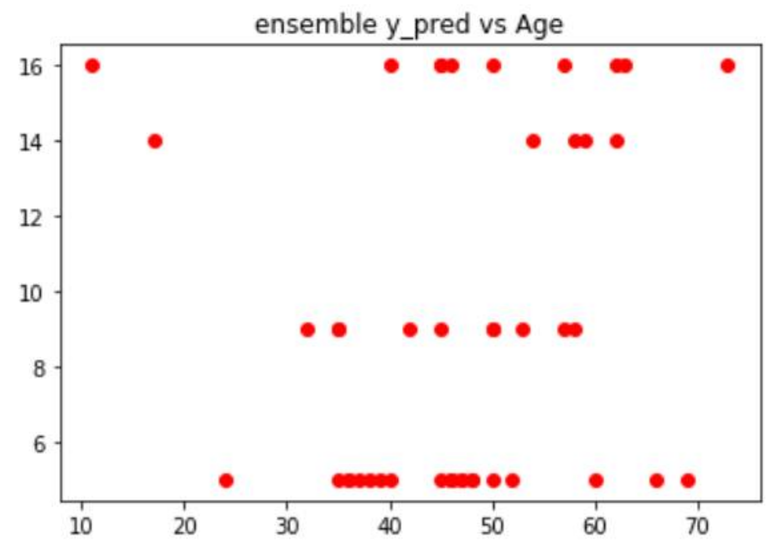
(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 6, June 2019

D. ENSEMBEL MODEL

The field of machine learning keeps getting better and better with time. Predictive models form the core of machine learning. Better the accuracy better the model is and so is the solution to a particular problem. In this post, you are going to learn about something called Ensemble learning which is a potent technique to improve the performance of your machine learning model.




```
ensemble model accuracy 98.00609756097559 %
execution time of ensemble model 10.49934720993042 seconds ---
```

Figure 8: Accuracy of Ensemble model

IV. RESULTS

As machine learning technologies are evolving and becoming more commom an amount of their applications have been found in many sectors one being health care. In this project, we develop a heart prediction system that can assist medical professionals in evealuating a patients heart disease based on clinical data of the patient. By using Naïve bayes, Support Vector Machine (SVM), Multilayer Perceptron (MLP) and combining them in the Ensemble model we get the following accuracy.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 6, June 2019

Algorithm Name	Accuracy	Execution Time
Naive Bayes	95.652%	0.3411(sec)
Support Vector Machine(SVM)	84.786%	0.5075(sec)
Multilayer Perceptron(MLP)	0.91304%	0.37301(sec)
Ensemble Model	**98.006%**	**10.499(sec)**

V. CONCLUSION

In this paper we selected a clinical dataset with 279 clinical features and on record 452 patients. We proposed the data and pruned by removing features having more than 60% NA values and the ones with lesser were modified with mean of the values. Later we developed MLP, SVM and Naïve Bayes model for predicting the accuracy of different kinds of arrhythmia based on these pruned 139 clinical features after data processing. An ensemble model for the same dataset gave accuracy of 98% higher than that of individual models.

REFERENCES

1. Liang Li, Mei Chen, Hunhu Wang, “cost sensitive ensemble method for medical prediction”, IEEE 2016.
2. Hamidreza Ashrafi Esfahani, Moreza Ghazanifari, “Cardiovascular disease detection using ensemble classifier”, 2017 IEEE international conference on knowledge based engineering and innovation(KBEI)
3. Tharaha S, Rashika K, “hybrid artificial neural network and decision tree algorithm foe disease recognition and prediction in human blood cells”, IEEE, 2017.
4. Jianfeng Xu, Dan Li, YuanjianZhang “ cluster analysis of multi-granular time series data for mortality rate rediction”, computing in cardiology, 2016 IEEE.
5. AH Chen, SY Huang, CH cheng, EJ lin, “HDPS: Heart disease prediction system”, computing in cardiology, 2017 IEEE
6. Tahira Mahaboob, Rida irfan, bazelah ghaffar “evaluating ensemble prediction of coronary heart disease using receiver operating characteristics”, IEEE 2017.
7. Tharaha S, Rashika K, “hybrid artificial neural network and decision tree algorithm foe disease recognition and prediction in human blood cells”, IEEE, 2017.
8. R. kavitha, E.Kannan, “an efficient framework for heart disease classification using features extraction and feature slection techniques in data mining”, 2017 IEEE.
9. Tanoy debnath, Md. Mehandi hasan, tanwi biswas, “analysis of ECG signal and classification of heart abnormalities using artificial neural network”,conference on electrical and computer engineering , 2016 IEEE.
10. Paurakh rajbhandary , benjsmin zhou gaspar Garcia jr. “detection heart abnormality using ECG with CART”, 2017 IEEE.
11. Vjay Pappu and panos m pardalos “high dimensional data classification”.
12. Swain m dash s. k and mohapatra “An approach for iris plant classification using neural network”, IBMRD’s computing vol. 3 NO1. Pp.79, 2012.
13. Chaki, Jyotismita and Ranjan Parekh, “plant leaf recognition using shape based features and neural network classifier”, IJACSA 2011.