MULTISTAGE CLASSIFICATION FOR ARRHYTHMIA RISK PREDICTION

A Dissertation

Submitted in fulfilment of the

requirements for the award of the degree

Of

MASTER OF TECHNOLOGY

in

COMPUTER SCIENCE AND ENGINEERING

Submitted By

BHARAT GOEL

(14535010)

Under the guidance of

DR. DURGA TOSHNIWAL

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

ROORKEE – 247667 (INDIA)

MAY, 2016

CANDIDATE'S DECLARATION

I hereby declare that the work, which is presented in this dissertation report entitled " Multistage Classification For Arrhythmia Risk Prediction" towards the fulfilment of the requirements for the award of the degree of Master of Technology with specialisation in Computer Science And Engineering submitted in the department of Computer Science and Engineering, Indian Institute of Technology, Roorkee (India) , is an authentic record of my own work carried out during the period of July 2015 to May 2016 under the guidance of Dr. Durga Toshniwal, Associate Professor, Department of Computer Science and Engineering, Indian Institute of Technology Roorkee.

I have not submitted the matter embodied in this dissertation for the award of any other degree or Diploma.

Date :

Place : BHARAT GOEL

CERTIFICATE

This is to certify that the above statements made by the candidate is correct to the best of my knowledge and belief.

Date :

Place :

Dr. Durga Toshniwal,
Associate Professor,
Department of Computer Science and Engineering,
IIT Roorkee

ACKNOWLEDGEMENT

It gives me immense pleasure to thank all those people who have, at various stages and in various ways have played a key role in successful completion of this work. I would take this opportunity to extend my heartfelt gratitude to my guide and mentor Dr. Durga Toshniwal, Associate Professor, Indian Institute of Technology, Roorkee for her invaluable advice, guidance, encouragement and for sharing her knowledge . Her wisdom and commitment to the highest standards motivated me throughout. She has been very generous in providing the necessary resources to carry out the research. She is an inspiring professor, a great advisor and most importantly a person.

I am also highly indebted to my colleague Maj. Sumit Prakash Gupta, Ankur Tagra, and my family who gave me the moral support and valuable suggestions. On a personal note, I owe everything to the almighty.

Bharat Goel

ABSTRACT

Arrhythmia is the most common cardiovascular disease (CVD) taking a toll of approximately 10 million cases per year in India. ECG analysis is the most prominent way of detecting the Arrhythmia abnormalities. ECG being complex with many crest and trough, it would be of great help if automated analysis of ECG can done. In this work we propose a framework for prediction of arrhythmia risk by statistically analyzing the ECG of the user. We have proposed novelistic way for training and testing for multi-stage classification to improve the accuracies and sensitivities of the model. Further we also consider the time aspect of the model since the model should be able to put in use in real-time applications. We focuses on to build a model that can be combined with personal holter machines so that it can act as a precursor to the consultation with the doctor.

We have used MIT-BIH Arrhythmia Benchmark dataset following the AAMI recommendations and the work corresponds to inter-patient paradigm.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

Introduction

1.1 Background & Motivation

CVD's are the most prominent cause of deaths globally, as claimed by WHO [1] and Arrhythmia is the most common from it. Reportedly there are approximately 10 million Arrhythmia cases per year in India. The analysis of electrocardiogram (ECG or EKG) can provide clinicians with valuable information about the health condition of the patient. ECG is a set of interpretative signals or vectors which indicates the electrical movements in the heart with respect to time with respect to different reference planes. ECG is recorded by placing electrodes of the holter machine in direct contact with the patient's body[4]. If a person has arrhythmia or other CVD's then his ECG gets deviated from the normal signal. The features of ECG gets distorted from the normal features and similar CVD's have similar change in the ECG features [2,3]. If an unknown ECG has features similar to that of an ECG signal with arrhythmia, it could be deduced that this unknown signal also has the arrhythmia. ECG has proved itself to be non-invasive clinical tool because using the ECG , it is very convenient to predict the abnormality of the heart.

The main cause of mortality in any area is CVD's and many surveys shows the importance of diminishing the time delay involved in the medication of these diseases for improving the results of clinical diagnosis and ECG being the most promising technique for the same . ECG interpretation being complex due to many precise details it is possible to make error by any human. It is also tough for experienced cardiologists to observe the abnormalities in the ECG due to large number of patients and at a time 4-5 ECGs are done for one patients. Hence an automated tool is needed for observing the abnormalities and predicting the risk. The negligence of people towards their health is well known , so a personal health system can be of great help for observing the ECG of a person and recommending accordingly. Moreover remote areas does not have medical assistance required for heart abnormalities and hence such a system can be of great use in both cases.

1.2 ECG Signal

The heart is a muscle that pumps the blood throughout the body by contracting in a rhythmic manner. Augestus Desire Waller recorded the first ECG in 1887 [5]. Electrocardiography is a methodology of recording the electrical activity of the heart using holter monitor and various electrodes. ECG can be done in any of the off-the-person, in-the person, on-the person mode. On-the-person method is the most popular since it is simple, cheap and reliable. Generally 5 electrode configuration is used to obtain the electrical potential between the electrodes placed on the body[6]: one of the electrodes is positioned on the left leg (LL),one on the right leg (RL), one on the left arm (LA), one on the right arm (RA), and one on the chest, to the right of the external (V or V1). From these electrodes several different leads can be constructed to visualize the ECG signal.

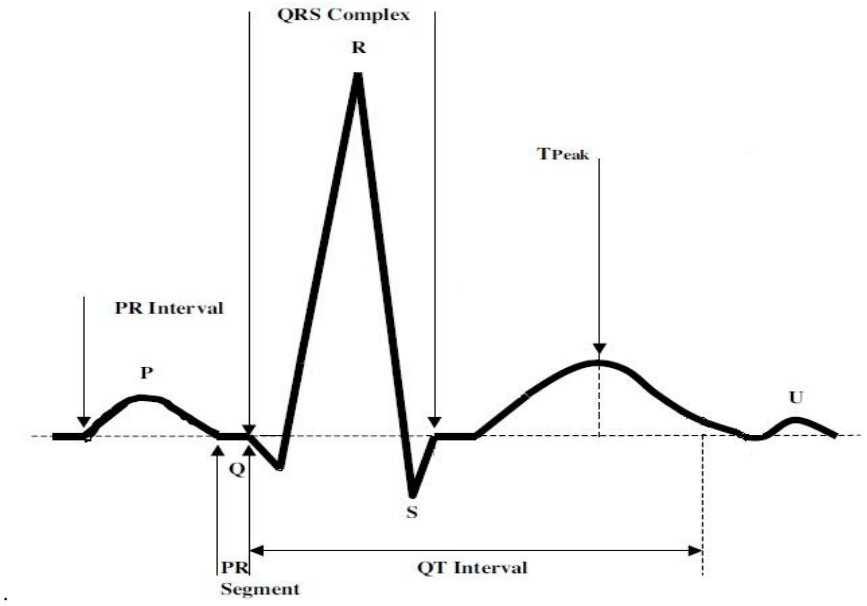


Fig1: Sample ECG [6]

Figure 1 shows a sample ECG. Various ECG features are used for detecting their irregularities in the ECG such as RR-interval, morphological features, Heartbeat interval, QRS width, QRS peak , ST segment etc. From these RR-interval and QRS are the most important ones. These features are capable of giving sufficient information to discriminate heartbeat types and are the key factor for the success of any model.

1.3 Problem Statement

The main objective of this dissertation work is to propose a framework for predicting the arrhythmia risk by analysing the ECG of the users. Since our work is related to the medical domain , the primary concern is to improve the sensitivity of the system. The work further focuses on improving the testing time , accuracy and sensitivity of the system.

1.4 Research Contribution

The work done in this research work is novel in its multistage classification and multistage hybrid classification. Here we do not leave the misclassified instances as misclassified rather we again test them in other stages. The proposed model opens new ways for traditional single stage classification ways. The accuracies , sensitivities and time achieved are better than the previous work.

1.5 Organisation of Thesis

In chapter 2, we discuss the work done so far related to this research work. It also discusses the research gaps , various terminologies used in literature and performance metrics. In chapter 3 , we give an insight to the proposed work of this research. Chapter 4 explains the experiments conducted and their respective results. Chapter 5 concludes the research work and discusses the future work that can be done in this area.

Chapter 2

Literature Review

In this age of automation, a lot of work has been done for arrhythmia prediction using different feature sets, datasets and classification models. Association for the Advancement of Medical Instrumentation (ANSI/AAMI) has developed standards for the evaluation of automatic arrhythmia classification methods ANSI/AAMI EC57:1998/(R)2008 [7].

AAMI Standards:

AAMI recommends five databases, namely MIT-BIH(The Massachusetts Institute of Technology – Beth Israel Hospital Arrhythmia Database), EDB(The European Society of Cardiology ST-T Database . AHA(The American Heart Association Database for Evaluation of Ventricular Arrhythmia Detectors). CU(The Creighton University Sustained Ventricular Arrhythmia Database). NST (The Noise Stress Test Database). Of these databases MIT-BIH , an annotated database is used widely.

AAMI also recommends that records of patients using pacemaker should not be used and records of same patient should not be used for training and testing. The model of evaluation can be divided into two types based on the dataset creation i.e. inter-patient and intra-patient. If the records of same person is used for training as well as testing then it favours the model and called as intra-patient evaluation model. Whereas when training and testing sets have different patient records then it is termed as inter-patient model of evaluation.

The work done so far can be broadly divided on the basis of AAMI standards into Inter-patient paradigm and Intra-patient paradigm[38].

2.1 Inter-patient paradigm

De chazal et al. [8] have proposed a model for extracting various morphological and ECG-interval features and used weighted linear discriminant model for classification obtaining a accuracy 0f 83% and sensitivity of 76%. Soria et al.[9] have also used Vectorcardiogram feature

from both the leads and then applied weighted linear discriminant model to get 90% accuracy. Llamedo et al.[10] have combinedly used Vectorcardiogram and wavelets with the weighted linear discriminant classifier to obtain higher accuracy of 93% sensitivity of 77 %. Mar et al.[11] used the temporal , morphological and statistical features and gained an accuracy of 89% by using weighted linear discriminant combined with multi-layer perceptron .

Luz et al.[12] have used more complex feature set and used optimum path forest classifier and achieved an accuracy of 90%.Ye et al.[13] used morphological features, wavelet transform, ECG-intervals, Independent component analysis (ICA), Principle component analysis (PCA) for gaining features and SVM as classifier obtaining an accuracy of 86.4% and sensitivity of 60 %. De lannoy et al.[14] have used ECG-intervals , morphological features, Higher order statistics and Hermite basis function to obtain then features and applied weighted SVM to obtain an accuracy of 83% and sensitivity of 78%. Park et al.[15] also used ECG-intervals, morphological features, Higher order statistics and Hermite basis function to obtain then features but instead applied hierarchical SVM to obtain an accuracy of 85% and sensitivity of 80%. Zhang et al.[16] used ECG-intervals, morphological features and ECG segments and used a complex combination of SVM to obtain an accuracy of 86%. Zhang et al.[16] again used the combination of SVM but this time they increases the feature set and also used wavelets coefficient for feature set selection and obtained accuracy of 87% and sensitivity of 74%..

2.2 Intra-patient paradigm

Chen et al. [17] used RR-interval as feature set and for classification used some Set of rules to acheive an accuracy of 95%. Dokur et al. [18] used Wavelet and Discrete Fourier Transform to generate feature set and Artificial Neural Networks as a classifier to generate accuracy of 96%. Osowski et al. [19] Higher Order Statistics and wavelets for feature extraction and Fuzzy Neural Networks for classification to obtain accuracy of 96%. Tsipouras et al. [20] used only RR-interval as feature and Deterministic automata for classification to acheive an accuracy of 96%. Mehmet et al. [21] used Higher Order Statistics and Wavelet as feature set and k-Nearest Neighbour and Bayes classifier to achieve accuracy of 98% . Ozbay et al. [22] used Raw-wave and MLP and Fuzzy Cluster to attain average accuracy of 99%. Yu et al. [23] used RR-interval

and Independent Component Analysis for feature set generation and Neural Networks for classification to attain an accuracy of 98%. Ceylan et al. [24] used Principle Component Analysis and Wavelet Transforms to generate feature set and Artificial Neural Network to attain an accuracy of 99%. Wen et al. [25] used RR-interval and raw-wave as feature set and SOCMAC-based Cluster to achieve 98% accuracy. Yu et al. [26] used Independent Component Analysis for feature extraction and SVM classifier to gain 98% accuracy. Ye et al. [27] used Wavelet, ICA,PCA, RR-interval for feature and SVM as classifier to obtain 99% accuracy. Mishra et al. [28] used Local Fractal Dimension for feature extraction and Nearest Neighbor for classification obtaining 89% accuracy. Korurek et al. [29] used RR-interval, QRS-width, Wavelet, PCA for feature set and k-NN for classification achieving 90% accuracy. Yeh et al. [30] used Morphological features, RR-interval as feature set and clustering to achieve an accuracy of 94%. Khazaee [31] used Heartbeat intervals and morphological amplitudes as feature set and Particle Swarm Optimizer ,SVM for classification and achieved 97% accuracy.

Wang et al. [32] used PCA, LDA for feature extraction and Probabilistic Neural Network for achieving 99% accuracy as classifier. Kumar et al. [33] used RR-intervals as feature and ensemble methods for classification to gain 92% accuracy. Chen et al. [34] also used only RR-intervals as feature and SVN, NN for classification to achieve 100% accuracy. Ahmed et al. [35] used Heartbeat intervals ,morphology amplitude, HOS for feature generation and MLP, SVM, TreeBoost for classification attaining 98% accuracy. Sarfraz et al. [36] used RR-intervals, QRS segment for feature generation and Back Propagation Neural Network to attain 99% accuracy. Tran et al. [37] used RR-intervals, Hermite Basis Function for feature set and Ensemble of classifiers attaining 98% accuracy. Hina sharma [39] used a mixture of feature set to attain an accuracy of 99.71%.

2.3 Biased Dataset

In medical sciences individual differences of persons are of very importance. Hence the division of dataset into training and testing should not have records of same person. Random division assumes that the samples represents each cluster but there are chances that random sampling

does not take care of individual differences. If the division has common persons then it favours accuracies , sensitivity, specificity hence becomes biased. Figure 2 explains Biasing .

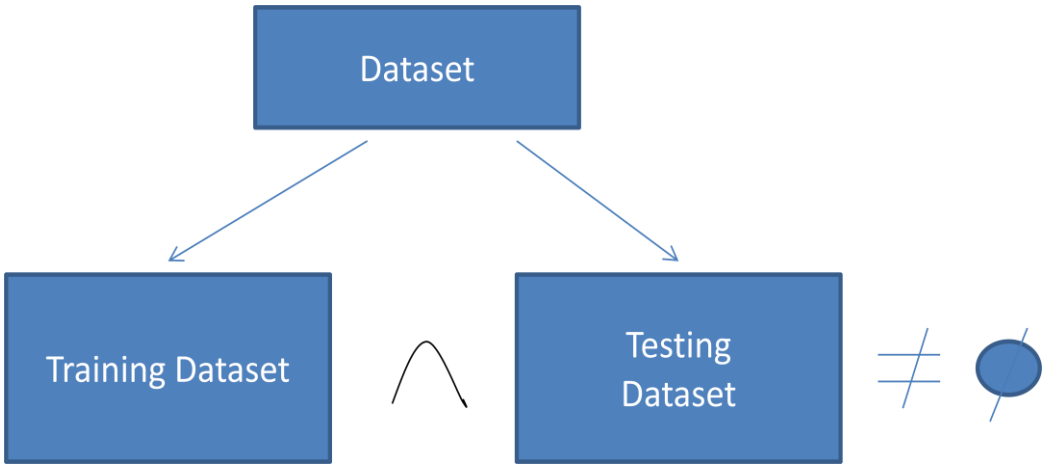


Fig. 2 : Biased Dataset

2.4 Un-Biased Dataset

Since the real life applications will work on unseen records , the training and testing should be done on different patient records. Division of dataset into training and testing set is **Mutually Disjoint** in respect to the person. Figure 3 explains un-biasing.

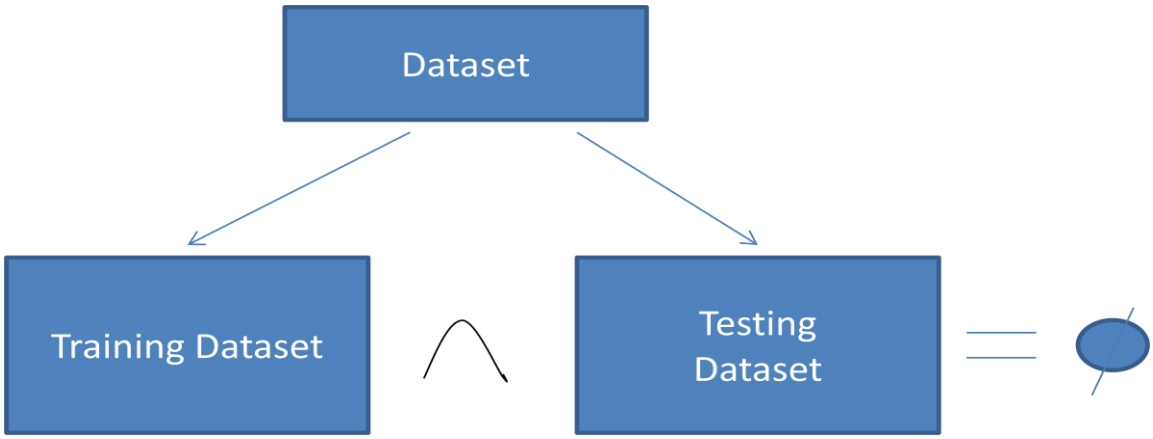


Fig. 3: Unbiased Dataset

2.5 Performance Evaluation Measurements

Predicting performance of a machine learning method based on inadequate data is difficult. Figure 4 shows the sample confusion matrix. In our problem domain it is important to reduce the false negatives since none of the patient with abnormal heartbeat be regarded as normal person. Though it is good to reduce false positives also but if we have a trade-off b/w the two then we need the optimized value of both.

	Predicted_True	**Predicted_False**
Actual_True	True_Positive (TP)	False_Negative (FN)
Actual_False	False _Positive (FP)	True_Negative (TN)

Fig. 4: Sample Confusion Matrix

Here TP means that the patient had the disease and is correctly declared to be diseased. FN means that the patient had the disease but is incorrectly declared to be non-diseased. FP means that the patient was normal but was declared to be diseased. TN means that the non-diseased patient was declared non-diseased correctly.

The number of TN, FN , TP, FP are used to compute the efficiency of the classifier. The sensitivity is statistical measurements of checkout tests.

Sensitivity is the rate of true positive test result,

$$Sensitivity = \left(\frac{TP}{TP+FN}\right) * 100\ \% \qquad \text{eq. 1}$$

Accuracy shows overall measure i.e. the correct prediction ratio, which is:

$$Accuracy = \left(\frac{TP+TN}{TP+FP+TN+FN}\right) * 100\ \% \qquad \text{eq. 2}$$

2.6 Research Gaps

The literature present have used single stage i.e. once misclassified is considered as misclassified and it is not further evaluated. The dataset used in most of the research work has considered biased dataset i.e. used Intra-patient paradigm and hence these works cannot be used for comparing results and there is a huge scope of using the un-biased dataset. Various works using inter-patient paradigm have given reasonable accuracies but their works have focussed on accuracy aspect and neglected the sensitivity aspect , moreover to build an automated system time constraint should also be kept in mind. For designing a system for real world practical use the feature set should be simple and time should also be minimized , the previous works fails to do so.

Chapter 3

Proposed Work

The general framework of our work has been shown in Figure 5. The Multistage-classification phase is the most important and we have used it for our purpose as explained in section 3.3.

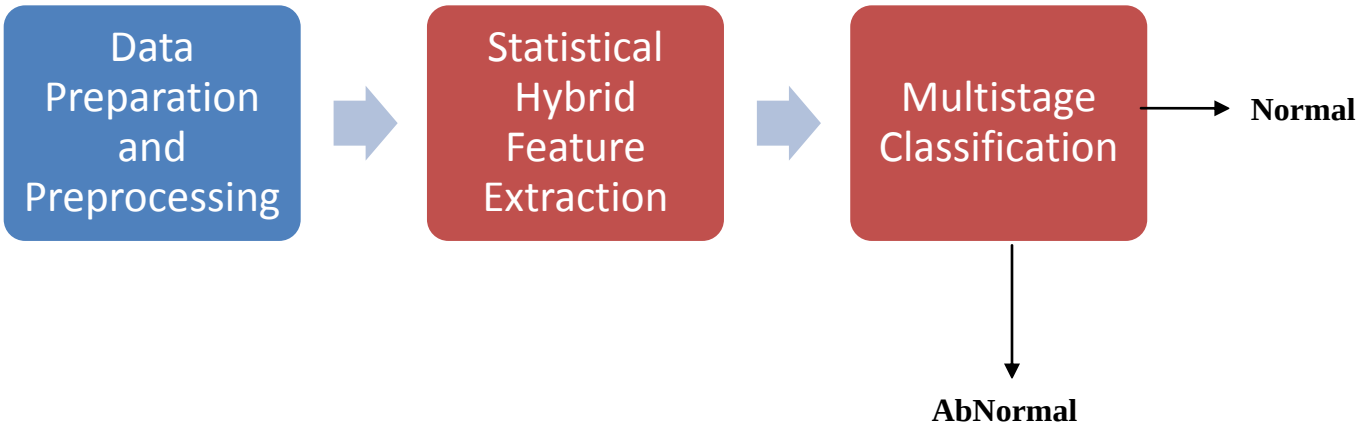


Fig. 5 : Proposed Framework

3.1 Data Preparation and Pre-processing

This is a crucial step in data mining since it directly affects the performance of the model and covers approximately 60% of the efforts required for the complete model. We integrated the annotations and the signal and generated a dataset.

We have made un-biased dataset by randomly dividing patients into 4 groups to make 3 dataset for training and 1 dataset for testing. The pre-processing involves dealing with noise, missing values and the transformation and integration aspects of the data.

3.2 Statistical Hybrid Feature Extraction

Feature set is the non-redundant and informative property of the data that can represent most crucial aspect of the data. These features facilitates the learning model to generalize and instead of complete initial data these features are used only. The more expressive and content the features are the better are the results of the task. We extracted statistical features of an ECG.

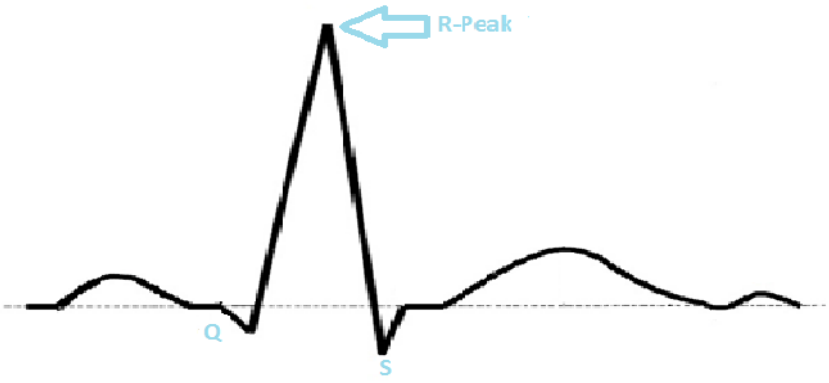


Fig. 6 : R-peak

Firstly we extracted the R-peak value for both the mlii, v1 leads . Secondly we calculated the interval between the previous R-peak found with the current R-peak and termed it as pre-rr interval. Thirdly we calculated the interval between the current R-peak and the next R-peak and termed it as post-RR-interval. Figure 6 shows the R-peak and Figure 7 shows the pre-RR interval and post-RR interval.

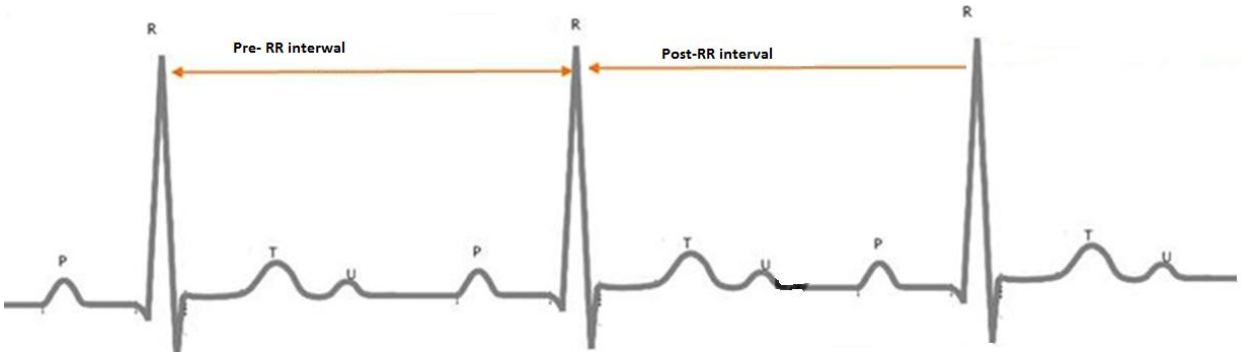


Fig 7 : Pre-RR interval and Post-RR interval

3.3 Multistage Classification:

Classification is a supervised way of grouping data into different categories such that similar data points comes in same category and dissimilar data point does not. The number of category depends on the problem domain and for our purpose number of categories are two i.e. Normal representing the patients whose ECG is normal and Abnormal representing patients with irregularities in ECG. The focus of our work is this classification model. We have modified the general classification model and used modified Boosting for training and Multistage classification for testing.

Figure 6 shows the our modified boosting approach for training. Figure 8 shows the training model for various classifiers.

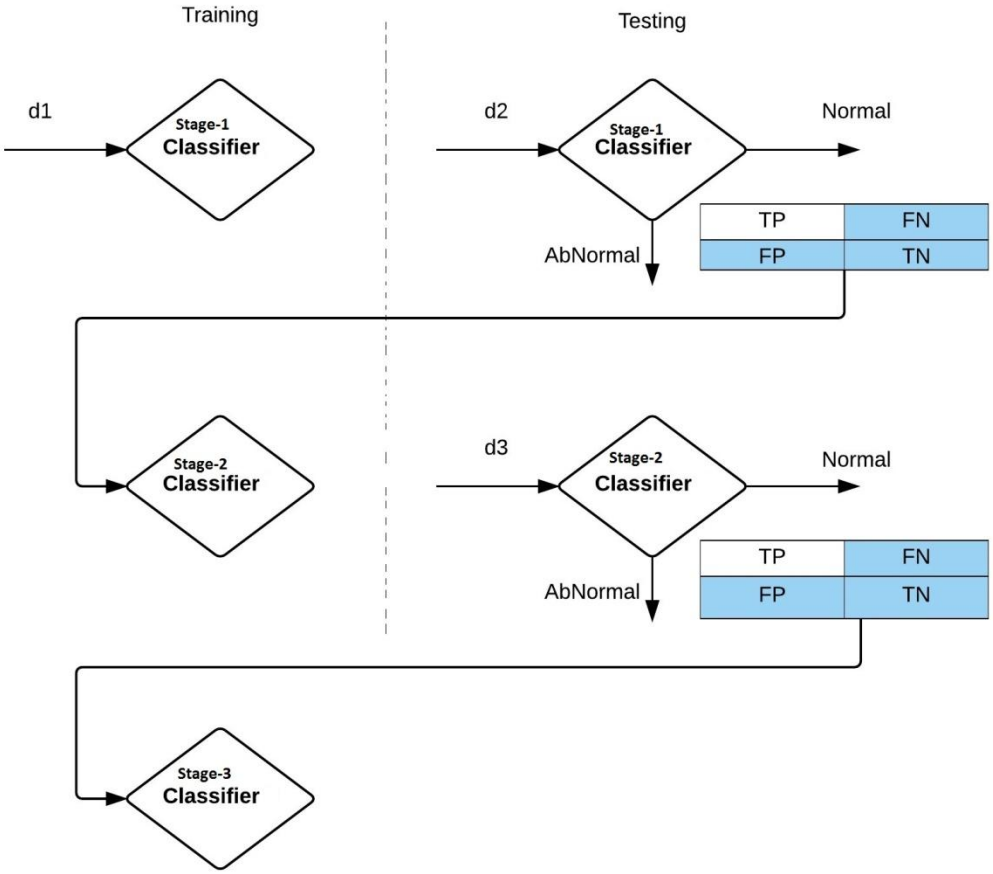


Fig. 8: Training of Multistage classification

We used Multistage for testing as shown in Figure 9 .

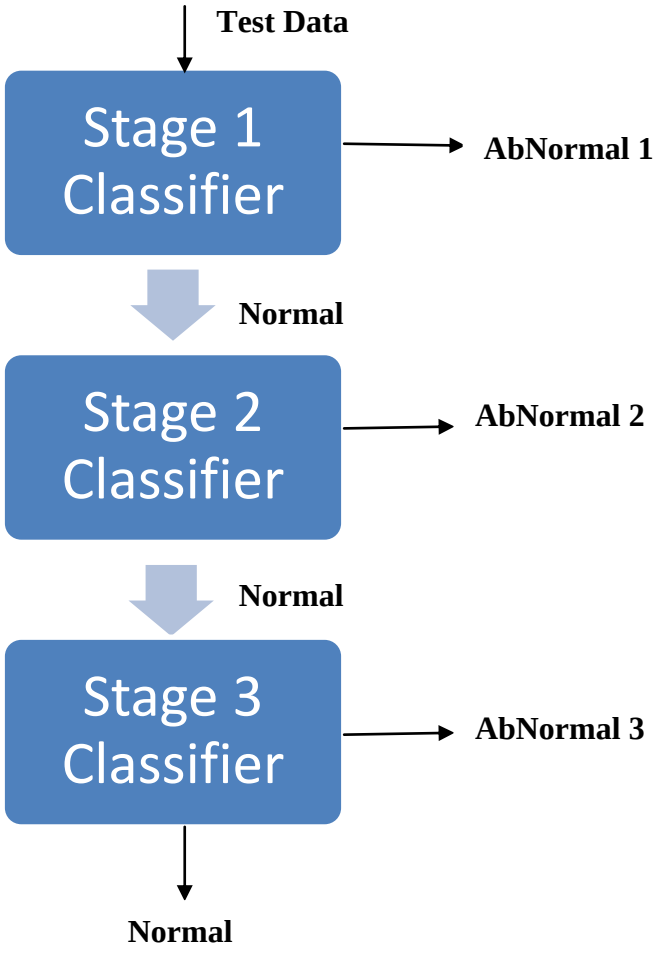


Fig. 9 :Testing for Multistage Classification

Chapter 4

Experiments and Results

4.1 Data Set Used

MIT_BIH: The Massachusetts Institute of Technology – Beth Israel Hospital Arrhythmia Database [39].

We have used MIT-BIH Arrhythmia Database [18] in this work. It contains 48 half hour two lead ambulatory ECG recordings from 47 real patients. The digitization of recordings has been done at 360 samples per second per channel (2 channels). Each record is annotated by a panel of cardiologists and any ambiguity has been resolved with proper consultation or give a n unclassified class. Since some records contain paced beats(102,104,107, and 217) AAMI recommends that these records should be excluded, hence only 44 recordings from the MIT-BIH Arrhythmia Dataset are used for predicting cardiac arrhythmia. The dataset contains 15 classes table [39] of arrhythmia but principal types of heartbeats present are grouped into five broad classes as shown in the figure 10.

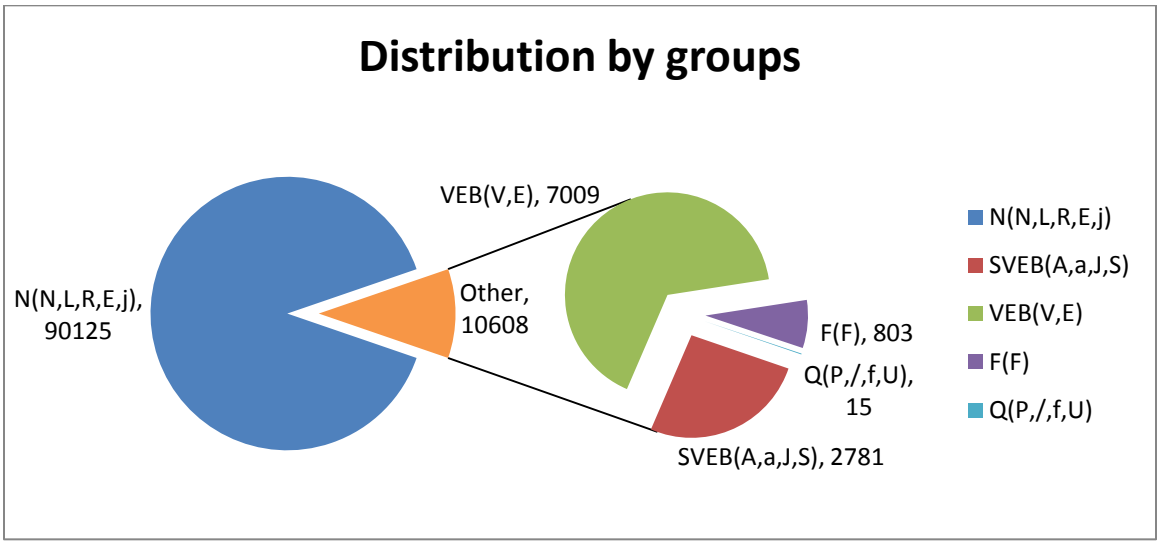


Fig 10: Distribution of instances into various categories

Class	Group	Symbol	Description
Normal	**N** Any heartbeat not categorized as SVEB, VEB, F or Q	N j E L R	Normal beat Nodal(junctional) escape beat Atrial escape beat Left bundle branch block beat Right bundle branch block beat
Abnormal	**SVEB** Supraventricular_ectopic_beat	J S A a	Nodal(junctional) premature beat Supraventricular_premature beat Atrial premature beat Aberrated atrial premature beat
	VEB Ventricular ectopic beat	E V	Ventricular escape beat Premature_ventricular_contraction
	F Fusion beat	F	Fusion of ventricular and normal beat
Unused	**Q** Unknown beat	P, / F U	Paced beat Fusion of paced and normal beat Unclassifiable beat

Table 1: Class and their Labels

4.2 Dataset Preparation and Pre-processing

For preparing the dataset we integrated the digitized signals with the annotations so as to obtain the annotated dataset. We took 40 patient data (having mlii and v1 leads)and divided these 40 patients randomly into 4 datasets creating four sets each having ECG data for 10 patients and named them d1,d2,d3 and test. The first three datasets are used as training data and the fourth one is used for testing. Since there is no overlap of patients in training set and training set this makes it unbiased dataset.

Pre-processing is an important part in data mining for attaining good results. Pre-processing consists of noise removal, missing value imputation etc. The database we used consisted many instances that were having noise for the label or the value was missing. We instead of imputing

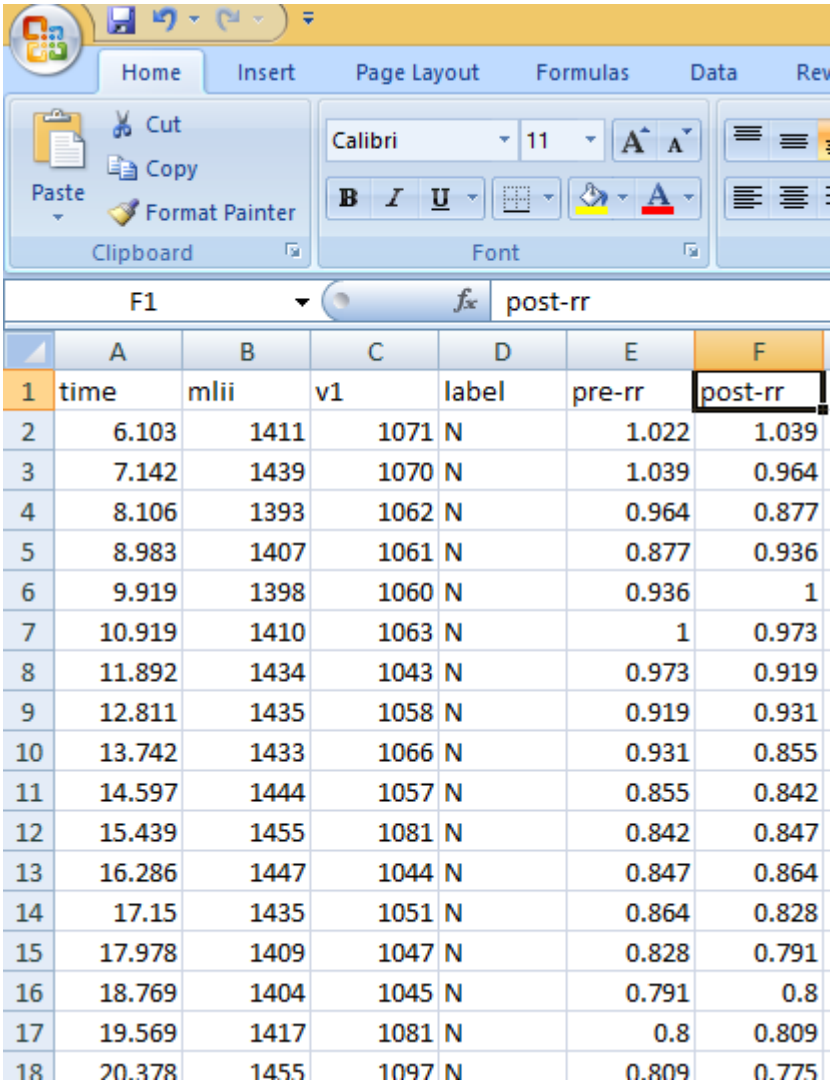
the missing value, removed that instance from our dataset since the number of instances were big and removing some instances does not have any impact on our dataset size. We also removed the unused class instances since they represent paced beat which is not recommended by AAMI, fusion of paced and normal and those beats that does not belong to any class. The number of instances of these unknown beats were very less and their removal did not have any impact on our obtained dataset. Python 3.5 was used for dataset preparation and Pre-processing.

4.3 Statistical Hybrid Feature Extraction

The aim of our research was to get the least possible run time and choose the simplest possible feature that is easy to extract and also has classification capabilities. We have used statistical features of the ECG i.e. the value of R-peak from the QRS Complex found from the ECG wave, Pre-RR interval and Post-RR interval. The database we used was already annotated with the R-peak values using the Pan Tompkins Algorithm so we extracted those values from the database. Python 3.5 were used for feature extraction with the Eclipse IDE.

Total Number of entries in the Database	**30.5 Million**
Total Number of Features Extracted	**0.11 Million**

Figure 11 shows the snapshot of features extracted with labels. We performed experiments on these features , first with R-peak , second with Pre-RR and Post-RR, third with both R-peak and Pre-RR and Post-RR intervals.



	A	B	C	D	E	F
1	time	mlii	v1	label	pre-rr	post-rr
2	6.103	1411	1071	N	1.022	1.039
3	7.142	1439	1070	N	1.039	0.964
4	8.106	1393	1062	N	0.964	0.877
5	8.983	1407	1061	N	0.877	0.936
6	9.919	1398	1060	N	0.936	1
7	10.919	1410	1063	N	1	0.973
8	11.892	1434	1043	N	0.973	0.919
9	12.811	1435	1058	N	0.919	0.931
10	13.742	1433	1066	N	0.931	0.855
11	14.597	1444	1057	N	0.855	0.842
12	15.439	1455	1081	N	0.842	0.847
13	16.286	1447	1044	N	0.847	0.864
14	17.15	1435	1051	N	0.864	0.828
15	17.978	1409	1047	N	0.828	0.791
16	18.769	1404	1045	N	0.791	0.8
17	19.569	1417	1081	N	0.8	0.809
18	20.378	1455	1097	N	0.809	0.775

Fig. 11: Snapshot of features extracted with labels

4.4 Multistage Classification & Multistage Hybrid Classification

For the classification we used three classifiers and their training , testing is done as explained in previous section 3.3 and accuracies and sensitivities are calculated according to the eq(1) and eq(2). We implemented various classifiers i.e. Decision Trees, Nearest neighbour, Adaboost , Random Forest in Python 3.5 using Eclipse IDE Environment [40]. We permuted these classifiers so as to make a more stronger classifier and compared their results. Confusion matrix after each stage have been made along with accuracy and sensitivity scores.

4.5 Results

4.5.1 Results for Multistage Classification

Experiment 1: MultiStage Classification with single feature R-peak

Results 1:

Decision Trees(Max_depth=8, Max_depth=9, Max_depth=8) .We used exponential followed by linear scaling to get the most suitable value of max_depth in the decision tree.

Over all confusion matrix

	Predicted_True	**Predicted_False**
Actual_True	460+1583+75=2118	684
Actual_False	2292+3044+1182=6518	9445

Over all Confusion Matrix

Table 2 summarizes the results of various stages in respect of accuracies and sensitivities achieved and the testing time.

Stage	Accuracy after each stage %	Sensitivity after each stage %
Stage 1 Classifier Decision tree	75.30	16.41
Stage 2 Classifier Decision tree	67.51	72.91
Stage 3 Classifier Decision tree	61	75.56
Total Testing Time	0.2112 s	

Table 2: Stage wise Comparison Decision Trees.

Results 2:

k-Nearest Neighbour (k=1,k=1,k=1)

Over all confusion_matrix

	Predicted_True	**Predicted_False**
Actual_True	736+1271+237=2244	558

Actual_False	3249+3478+2908=9635	6328

Over all Confusion Matrix

Table 3 summarizes the results of various stages in respect of accuracies and sensitivities achieved and the testing time.

Stages	Accuracy after each stage %	Sensitivity after each stage %
Stage 1 Classifier KNN	71.67	26.26
Stage 2 Classifier KNN	59.91	71.62
Stage 3 Classifier KNN	45	80
Total Testing Time	0.3613s	

Table 3: Stage wise Comparison KNN

Results 3:

Using RandomForestClassifier (max_depth=30, n_estimators=30, max_features=1) for all the stages.

Over all confusion matrix Decision tree

	Predicted_True	**Predicted_False**
Actual_True	608+1400+140 =2148	654
Actual_False	2562+3007+2698=8267	7696

Over all Confusion Matrix

Table 4 summarizes the results of various stages in respect of accuracies and sensitivities achieved and the testing time.

Stages	Accuracy after each stage %	Sensitivity after each stage %
Stage 1 Classifier Random Forest	74.65	21.16
Stage 2 Classifier Random Forest	66.09	71.66
Stage 3 Classifier Random Forest	52.45	76.65
Total Testing Time	0.4934 s	

Table 4: Stage wise Comparison Random Forest

Results 4:

Using Adaboost classifier at all the stages.

Over all confusion matrix Adaboost

	Predicted_True	**Predicted_False**
Actual_True	115+1788+28=1931	871
Actual_False	68+ 3094+4=3166	12797

Over all Confusion Matrix

Table 5 summarizes the results of various stages in respect of accuracies and sensitivities achieved and the testing time.

Stages	Accuracy after each stage %	Sensitivity after each stage %
Stage 1 Classifier Adaboost	85.31	4.1
Stage 2 Classifier Adaboost	78.35	67.91
Stage 3 Classifier Adaboost	78.48	68.91
Total Testing Time	0.5935 s	

Table 5: Stage wise Comparison Adaboost

Figure 12 shows comparison of these three classifiers based on accuracy , sensitivity and testing time .

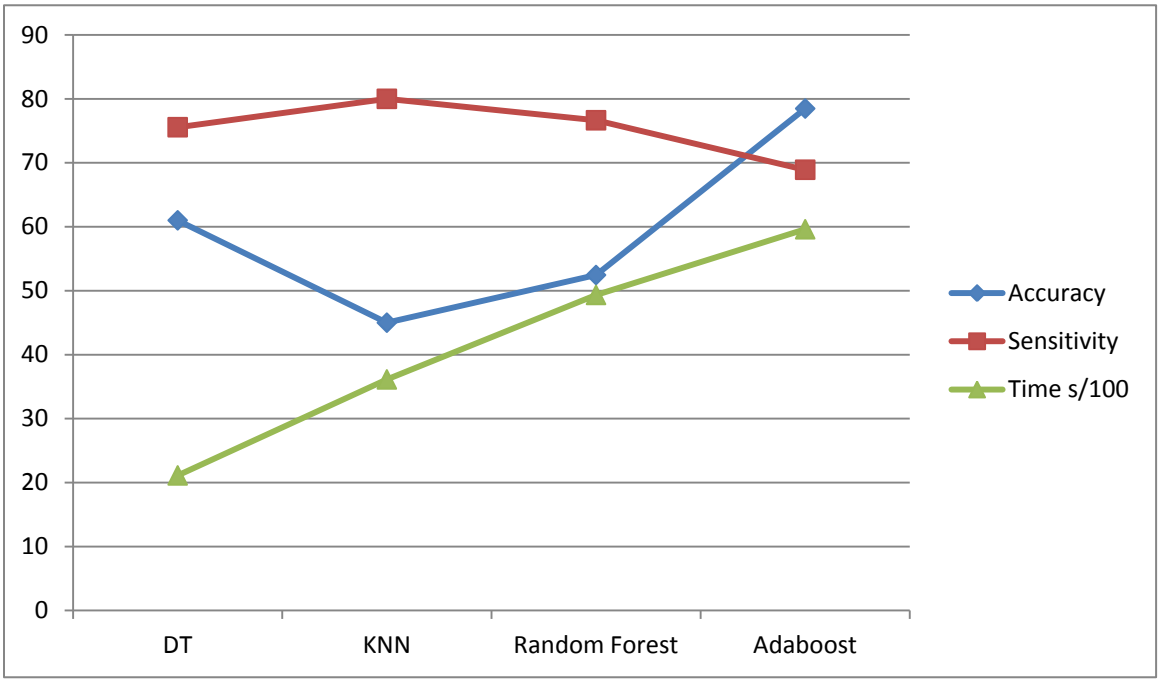


Fig. 12: Comparison of MultiStage Classification with single feature R-peak

In this experiment we conclude that the multistage classification having Decision Trees pose to give optimized accuracy, sensitivity and time.

4.5.2 Results for Multistage Hybrid Classification

Experiment 1: Multistage Hybrid Classification with single feature R-peak

Results 1: Ensemble Techniques: R-peak

Table 6 shows comparison of 24 ensemble classifiers using only R-peak value as a feature set.

S. No.	Classifier_1	Classifier_2	Classifier_3	Accuracy	Sensitivity
1	Adaboost	RandomForest	KNN	58.54	71.27
2	Adaboost	RandomForest	DT	72.07	64.91
3	Adaboost	KNN	RandomForest	59.72	69.55
4	Adaboost	KNN	DT	66.44	62.20
5	Adaboost	DT	RandomForest	63.46	70.34
6	Adaboost	DT	KNN	59.08	73.12
7	RandomForest	Adaboost	KNN	49.91	80.94
8	RandomForest	Adaboost	DT	60.05	77.65
9	RandomForest	KNN	Adaboost	62.12	70.27
10	RandomForest	KNN	DT	56.26	72.55
11	RandomForest	DT	Adaboost	65.19	74.94
12	RandomForest	DT	KNN	50.39	80.33
13	KNN	Adaboost	RandomForest	53.58	80.22
14	KNN	Adaboost	DT	5637	78.37
15	KNN	RandomForest	Adaboost	63.55	73.94
16	KNN	RandomForest	DT	64.83	75.08
17	KNN	DT	Adaboost	64.71	74.41
18	KNN	DT	RandomForest	52.09	78.33
19	DT	Adaboost	RandomForest	55.44	81.33
20	DT	Adaboost	KNN	50.63	83.01

21	**DT**	**RandomForest**	**Adaboost**	**66.17**	**73.76**
22	**DT**	**RandomForest**	**KNN**	**49.34**	**79.22**
23	**DT**	**KNN**	**Adaboost**	**63.20**	**73.69**
24	**DT**	**KNN**	**RandomForest**	**51.15**	**78.40**

Table 6: Multistage Hybrid Classifiers comparison (R-peak as feature set)

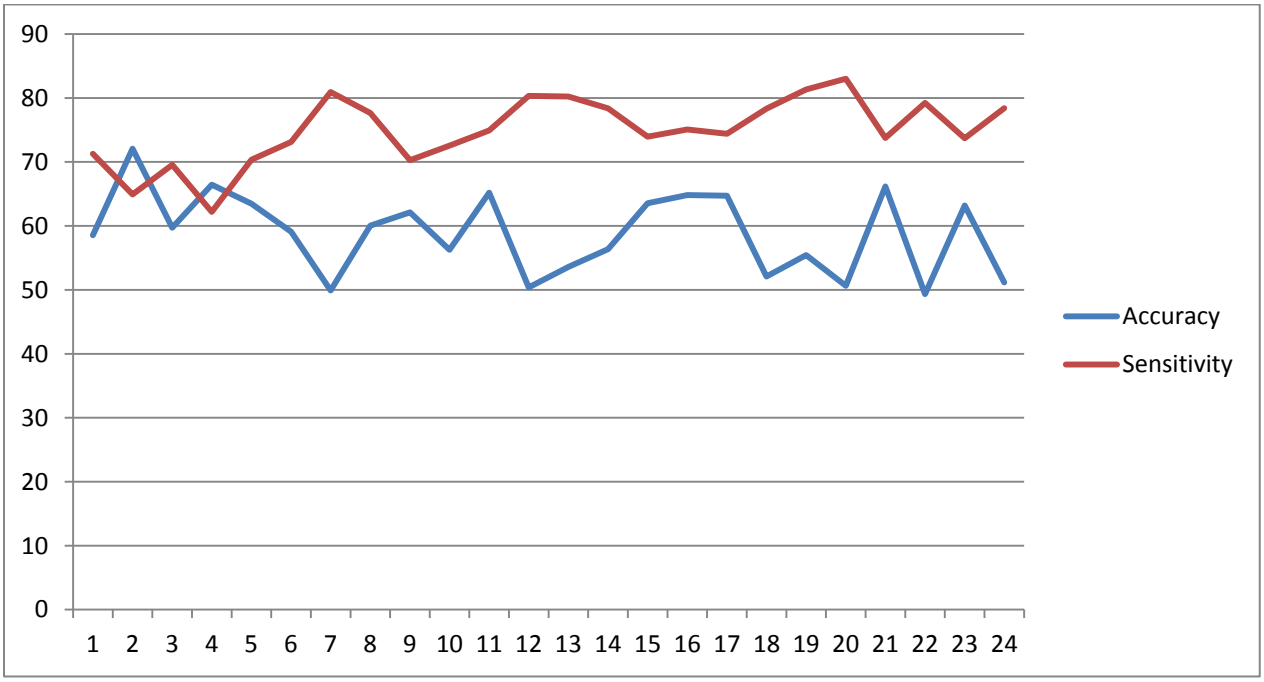


Fig. 13: Multistage Hybrid Comparison of classifiers with R-peak

In this experiment we conclude that the multistage hybrid classification having Adaboost at its first stage , Random Forest at its second stage and Decision Tree at its third stage pose to give optimized accuracy, sensitivity and time.

Experiment 2: Multistage Hybrid Classification with Pre - RR interval and Post - RR interval features

Results 2: Ensemble Technique: With Pre - RR interval and Post - RR interval

Table 7 shows comparison of 24 ensemble classifiers using pre-RR interval and post-RR interval as feature set.

S. No.	Classifier_1	Classifier_2	Classifier_3	Accuracy	Sensitivity
1	Adaboost	RandomForest	KNN	77.96141	63.05299
2	Adaboost	RandomForest	DT	82.45943	61.76471
3	Adaboost	KNN	RandomForest	80.03152	62.78561
4	Adaboost	KNN	DT	82.19108	64.14682
5	Adaboost	DT	RandomForest	80.62785	62.20224
6	Adaboost	DT	KNN	77.49712	63.27175
7	RandomForest	Adaboost	KNN	77.95715	62.0807
8	RandomForest	Adaboost	DT	82.57018	62.17793
9	RandomForest	KNN	Adaboost	81.26251	57.3894
10	RandomForest	KNN	DT	81.55642	61.47302
11	RandomForest	DT	Adaboost	81.53086	56.22265
12	RandomForest	DT	KNN	76.70486	61.76471
13	KNN	Adaboost	RandomForest	79.90374	62.25085
14	KNN	Adaboost	DT	81.53086	60.13612
15	KNN	RandomForest	Adaboost	81.26251	58.11862
16	KNN	RandomForest	DT	81.36474	61.98347
17	KNN	DT	Adaboost	81.33066	58.36169
18	KNN	DT	RandomForest	79.31593	62.71269
19	DT	Adaboost	RandomForest	79.85262	67.40399
20	DT	Adaboost	KNN	77.09673	68.81381

21	**DT**	**RandomForest**	**Adaboost**	**81.76939**	**65.79971**
22	**DT**	**RandomForest**	**KNN**	**76.64523**	**67.42829**
23	**DT**	**KNN**	**Adaboost**	**81.04954**	**66.57754**
24	**DT**	**KNN**	**RandomForest**	**78.90276**	**67.35537**

Table 7: Multistage Hybrid Classifiers comparison (pre - RR interval and post - RR interval as feature set)

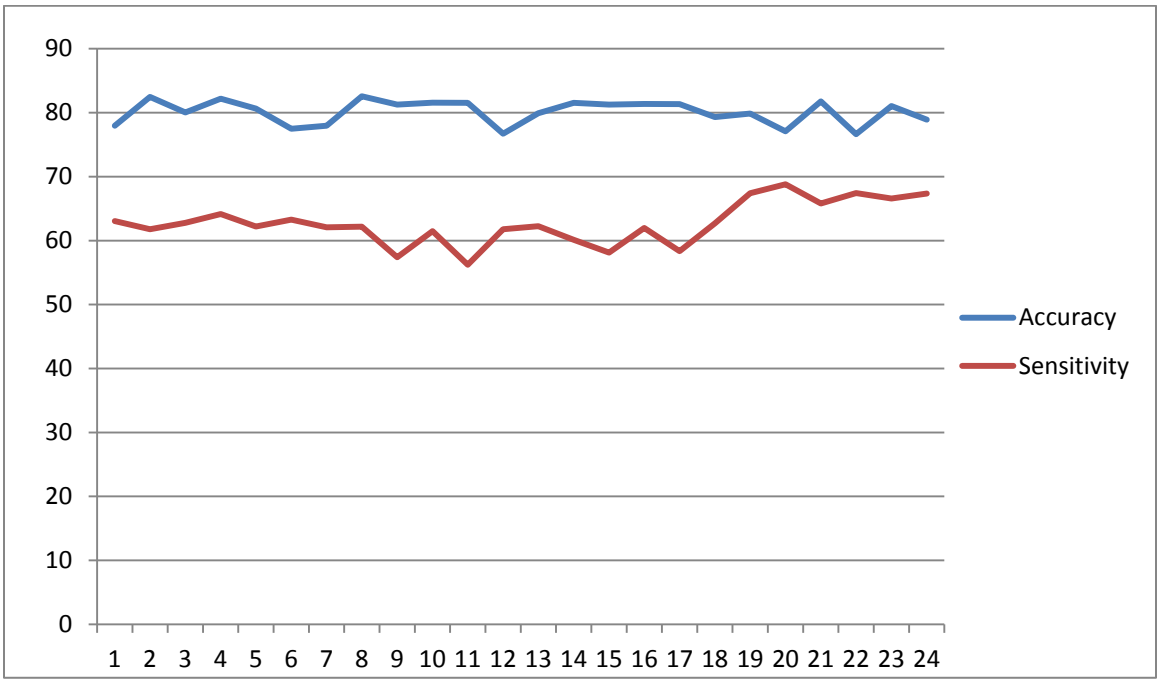


Fig. 14: Multistage Hybrid Comparison of classifiers with pre - RR interval and post - RR interval

In this experiment we conclude that the multistage hybrid classification having Decision Tree at its first stage, Adaboost at its second stage and Random Forest at its third stage pose to give optimized accuracy, sensitivity and time.

Experiment 3: Multistage Hybrid Classification with R-peak, Pre - RR interval and Post - RR interval features

Results 3: Ensemble Technique: With R-peak and Pre-RR interval and Post-RR interval

Table 8 shows comparison of 24 ensemble classifiers using R-peak, pre-RR interval and post-RR interval as feature set.

S. No.	Classifier_1	Classifier_2	Classifier_3	Accuracy	Sensitivity
1	Adaboost	RandomForest	KNN	67.91	64.12
2	Adaboost	RandomForest	DT	74.88	79.26
3	Adaboost	KNN	RandomForest	70.69	69.15
4	Adaboost	KNN	DT	68.19	80.40
5	Adaboost	DT	RandomForest	78.01	59.38
6	Adaboost	DT	KNN	68.60	64.41
7	RandomForest	Adaboost	KNN	69.90	65.84
8	RandomForest	Adaboost	DT	73.26	81.81
9	RandomForest	KNN	Adaboost	67.42	57.36
10	RandomForest	KNN	DT	66.36	80.43
11	RandomForest	DT	Adaboost	79.63	54.39
12	RandomForest	DT	KNN	66.60	65.50
13	KNN	Adaboost	RandomForest	70.11	64.56
14	KNN	Adaboost	DT	63.98	67.93
15	KNN	RandomForest	Adaboost	67.96	56.90
16	KNN	RandomForest	DT	67.54	74.89
17	KNN	DT	Adaboost	69.63	49.80
18	KNN	DT	RandomForest	72.20	69.51
19	DT	Adaboost	RandomForest	76.18	61.27
20	DT	Adaboost	KNN	66.91	61.47

21	**DT**	**RandomForest**	**Adaboost**	**76.61**	**50.19**
22	**DT**	**RandomForest**	**KNN**	**67.12**	**63.00**
23	**DT**	**KNN**	**Adaboost**	**68.26**	**51.38**
24	**DT**	**KNN**	**RandomForest**	**67.54**	**63.00**

Table 8: Multistage Hybrid Classifiers comparison (R-peak, pre-RR interval and post-RR interval as feature set)

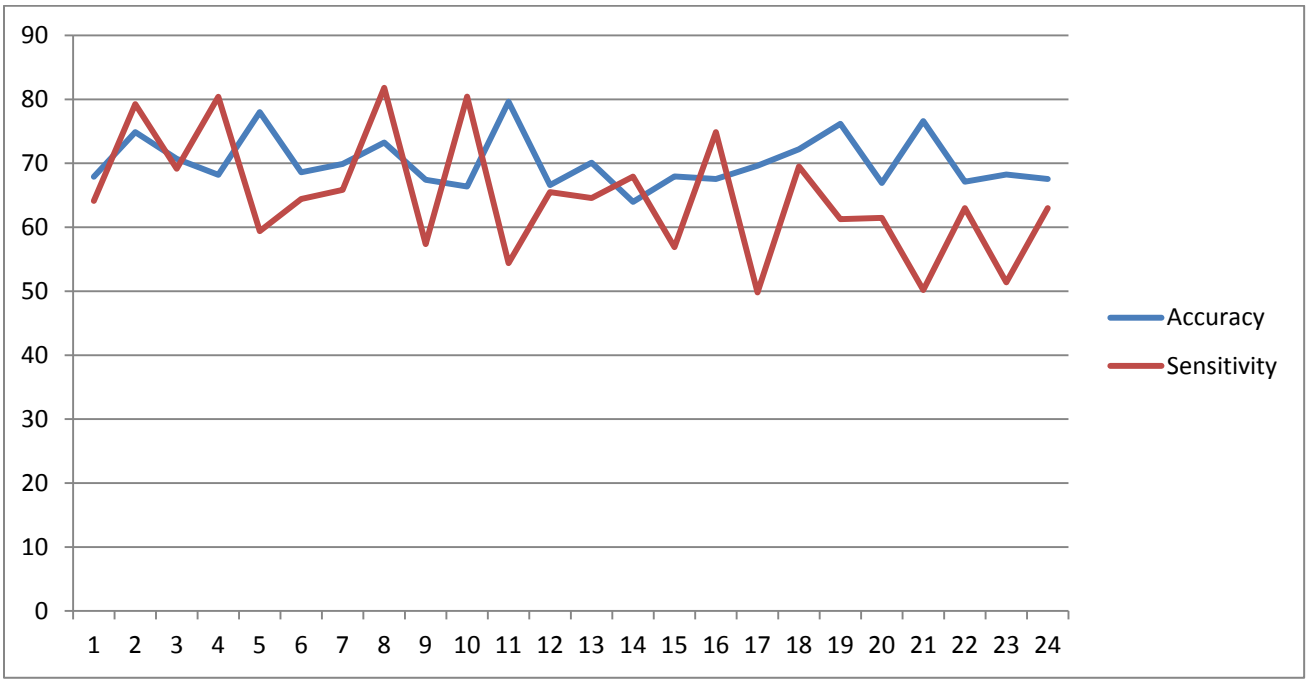


Fig. 15: Multistage Hybrid classifiers comparison (R-peak, pre-RR interval and post-RR interval)

Multistage hybrid classification having Adaboost at its first stage, Random Forest at its second stage and Decision Tree at its third stage pose to give optimized accuracy, sensitivity and time. Multistage hybrid classification having Random Forest at its first stage, Adaboost at its second stage and Decision Tree at its third stage pose to give equally optimized accuracy, sensitivity and time. Hence in this experiment we conclude that the combination with S.No. 2 and S.No. 8 are equally good. Both of these are explained as follows

S. No. 2 Results Description : Adaboost, Random Forest and Decision Trees

Over all confusion_matrix Decision tree

	Predicted_True	Predicted_False
Actual_True	3028	1086
Actual_False	4394	14969

Over all Confusion Matrix

Table 9 summarizes the results of various stages in respect of accuracies and sensitivities achieved and the testing time.

Stages	Accuracy after each stage %	Sensitivity after each stage %
Stage 1 Classifier Adaboost	85.19	46.54
Stage 2 Classifier RandomForest	84.74	48.32
Stage 3 Classifier Decision tree	74.88	79.26
Time	Training time=2.28 s, Testing Time= 0.7965 s	

Table 9: Stagewise comparison multistage hybrid classification

S. No. 8 Results Description: Random Forest, Adaboost, Decision Trees.

Over all confusion_matrix Decision tree

	Predicted_True	Predicted_False
Actual_True	3377	737
Actual_False	5723	13640

Over all Confusion Matrix

Table 10 summarizes the results of various stages in respect of accuracies and sensitivities achieved and the testing time.

Stages	Accuracy after each stage %	Sensitivity after each stage %

Stage 1 Classifier RandomForest	85.11	47.76
Stage 2 Classifier Adaboost	83.08	52.60
Stage 3 Classifier Decision tree	73.26	81.81
Time	Training time=2.43 s, Testing Time= 0.76990 s	

Table 10: Stagewise comparison multistage hybrid classification

Comparison of Time:

Table 11 shows the Training and testing time (in seconds) for MIT-BIH Arrhythmia Dataset for various works as calculated by Luz et. al. [12]. Our above explained approaches i.e. combination of Adaboost, Random Forest and Decision Trees and Random Forest, Adaboost , Decision Trees have very less testing time as compared to the approaches given in literature[12]. The total time required for our work is 3.19 seconds which is a huge improvement from the previous works as explained in Luz et.al.

	SVM			OPF		
Features	Train	Test	Total(sec)	Train	Test	Total(sec)
Chazal et al. (2004)	**190.36**	**173.56**	**363.92**	**609.98**	**891.14**	**1501. 12**
Güler and Übeyli (2005)	**077.90**	**058.93**	**136.83**	**205.76**	**216.58**	**0422.34**
Song et al. (2005)	**101.63**	**057.27**	**158.90**	**216.65**	**228.14**	**0444.78**
Yu and Chen (2007)	**115.23**	**068.12**	**183.35**	**249.83**	**224.41**	**0474.24**
Yu and Chou (2008)	**068.37**	**049.81**	**118.18**	**188.00**	**177.98**	**0365.98**
Ye et al. (2010)	**158.11**	**129.90**	**288.02**	**443.63**	**581.69**	**1025. 32**

Table 11: Training and Testing Time[12]

	Bayesian			**MLP**		
Features	Train	Test	Total(sec)	Train	Test	Total(sec)

Chazal et al. (2004)	**90.17**	**1393.75**	**1483.92**	**3682.04**	**0.21**	**3682.25**
Güler and Übeyli (2005)	**18.75**	**0209.50**	**0228.25**	**1790.11**	**0.13**	**1790.24**
Song et al. (2005)	**18. 16**	**0226.17**	**0244.33**	**1794.33**	**0.13**	**1794.46**
Yu and Chen (2007)	**22. 60**	**0302.08**	**0324.69**	**1947.38**	**0.14**	**1947.51**
Yu and Chou (2008)	**14. 41**	**0168.84**	**0183.25**	**1700.25**	**0.12**	**1700.37**
Ye et al. (2010)	**62.24**	**0944.93**	**1007.17**	**2951.61**	**0.18**	**2951.78**

Table 12: Training and Testing Time[12]

Comparison with Single Stage:

Single stages have slightly better accuracies but they are far worse on the sensitivity aspect. We obtained good sensitivity scores without much loss of accuracy.

Chapter 5

Conclusion and Future Work

Arrhythmia is the most common cardiovascular disease and its early detection can help in saving lives. ECG is the simplest indicator of heart activity and the ECG analysis being complex due to many precise details, it is possible to misinterpret the ECG hence there is a need of automated analysis of the ECG. ECG has various features that act as the deciding factor for the abnormality of the patient. The framework proposed uses multistage classification and gives better accuracies, sensitivity and time . The work proposed can act as a precursor to the consultation with the doctor. The work does not eliminate the need of doctors rather it may also be used by the doctors as a helping tool for ECG analysis. The experiments conducted infers that the proposed framework correctly addresses the problem domain of improving the sensitivity and can be used in real time. Further there is a scope of improving the accuracies and sensitivity scores. The framework can be done with the personal holter machines so as to develop small ECG analyzing tools. Further the model can also be extended on other morphological features, other datasets and their hybridization in future.

REFERENCES

[1] World Health Organization, http://www.who.int/mediacentre/ factsheets/ fs317/en/

[2] B. Qian, X. Wang, N. Cao, H.Li, Y. Jian (2014) A relative similarity based method for interactive patient risk prediction. Data Mining and Knowledge Discovery

[3] Vafaie, M. H., Ataei, M., & Koofigar, H. R. (2014). Heart diseases prediction based on ECG signals' classification using a genetic-fuzzy system and dynamical model of ECG signals. Biomedical Signal Processing and Control

[4] Indian Institute of Technology Roorkee . www.iitr.vlab.co.in,. (2012). ECG signal acquisition and R-Peak detection. Retrieved 28 October 2015, from iitr.vlab.co.in/?sub=52&brch=234&sim =1181&cnt=1.

[5] E.Besterman,R.Creese,Waller-pioneer of electrocardiography,Br. Heart J.42 (1)(1979)61–64.

[6] T. Barill, The Six Second ECG: A Practical Guidebook to Basic ECG Interpretation, nursecom, 2003.

[7] ANSI/AAMI, Testing and reporting performance results of cardiac rhythm and ST segment measurement algorithms, American National Standards Institute, Inc. (ANSI),Association for the Advancement of Medical Instrumentation (AAMI), ANSI/AAMI/ISO EC57,1998-(R)2008, 2008

[8] P. de Chazal, M. O'Dwyer, R.B. Reilly, Automatic classification of heartbeats using ECG morphology and heartbeat interval features, IEEE Trans. Biomed. Eng. 51 (7)(2004) 1196–1206.

[9] M.L. Soria, J.P. Martinez, Analysis of multidomain features for ECG classification, in: Comput. Cardiol., 2009, pp.561–564.

[10] M. Llamedo, J.P. Martí nez, Heartbeat classification using feature selection driven by database generalization criteria, IEEE Trans. Biomed. Eng. 58 (3) (2011) 616–625.

[11] T. Mar, S. Zaunseder, J.P. Martínez, M. Llamedo, R. Poll, Optimization of ECG classification by means of feature selection, IEEE Trans. Biomed. Eng. 58 (8) (2011)2168–2177.

[12] E.J.d.S. Luz, T.M. Nunes, V.H.C. De Albuquerque, J.P. Papa, D.Menotti, ECG arrhythmia classification based on optimum-path forest, Expert Syst. Appl. 40 (9) (2012)3561–3573.

[13] C. Ye, B.V.K. Kumar, M.T. Coimbra, Combining general multi-class and specific two-class classifiers for improved customized ECG heartbeat classification, in: International Conference on Pattern Recognition (ICPR), 2012, pp.2428–2431.

[14] G. de Lannoy, D. Franc¸ois, J. Delbeke, M. Verleysen, Weighted SVMs and feature relevance assessment in supervised heart beat classification, in: Biomedical Engineering Systems and Technologies (BIOSTEC), 2010, pp.212–223.

[15] K.S. Park, B.H. Cho, D.H. Lee, S.H. Song, J.S. Lee, Y.J. Chee, I.Y.Kim, S.I. Kim, Hierarchical support vector machine based heartbeat classification using higher order statistics and hermite basis function, in: Comput. Cardiol., 2008, pp.229–232.

[16] Z. Zhang, J. Dong, X. Luo, K.-S. Choi, X. Wu, Heartbeat classification using disease-specific feature selection, Comput. Biol. Med. 46 (2014) 79–89.

[17] S.-W. Chen, P.M. Clarkson, Q. Fan, A robust sequential detection algorithm for cardiac arrhythmia classification, IEEE Trans. Biomed. Eng. 43 (11) (1996) 1120–1124.

[18] Z. Dokur, T. Ölmez, ECG beat classification by a novel hybrid neural network, Comput. Method Program Biomed.66 (2-3) (2001) 167–181.

[19] S. Osowski, T.H. Linh, ECG beat recognition using fuzzy hybrid neural network, IEEE Trans. Biomed. Eng. 48 (11)(2001) 1265–1271.

[20] M.G. Tsipouras, D.I. Fotiadis, D. Sideris, Arrhythmia classification using the RR-interval duration signal, in:Comput. Cardiol., 2002, pp. 485–488.

[21] E. Mehmet, ECG beat classification using neuro-fuzzy network, Pattern Recogn. Lett. 25 (15) (2004) 1715–1722.

[22] Y. Özbay, R. Ceylan, B. Karlik, A fuzzy clustering neural network architecture for classification of ECG arrhythmias, Comput. Biol. Med. 36 (4) (2006) 376–388.

[23] S.-N. Yu, K.-T. Chou, Integration of independent component analysis and neural networks for ECG beat classification, Expert Syst. Appl. 34 (4) (2008) 2841–2846.

[24] R. Ceylan, Y. Özbay, B. Karlik, A novel approach for classification of ECG arrhythmias: type-2 fuzzy clustering neural network, Expert Syst. Appl. 36 (3) (2009) 6721–6726.

[25] C. Wen, T.-C. Lin, K.-C. Chang, C.-H. Huang, Classification of ECG complexes using self-organizing CMAC, Measurement 42 (3) (2009) 399–407.

[26] S.-N. Yu, K.-T. Chou, Selection of significant independent components for ECG beat classification, Expert Syst. Appl.36 (2) (2009) 2088–2096.

[27] C. Ye, M.T. Coimbra, B.V.K.V. Kumar, Arrhythmia detection and classification using morphological and dynamic features of ECG signals, in: Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2010, pp. 1918–1921.

[28] A.K. Mishra, S. Raghav, Local fractal dimension based ECG arrhythmia classification, Biomed. Signal Process. Control 5(2) (2010) 114–123.

[29] M. Korürek, A. Nizam, A new arrhythmia clustering technique based on ant colony optimization, J. Biomed. Inform. 41 (6) (2008) 874–881.

[30] Y.-C. Yeh, C.W. Chiou, H.-J. Lin, Analyzing ECG for cardiac arrhythmia using cluster analysis, Expert Syst. Appl. 39 (1)(2012) 1000–1010.

[31] A. Khazaee, Heart beat classification using particle swarm optimization, Int. J. Intell. Syst. Appl. 5 (6) (2013) 25.

[32] J.-S. Wang, W.-C. Chiang, Y.-L. Hsu, Y.-T.C. Yang, ECG arrhythmia classification using a probabilistic neural network with a feature reduction method, Neuro computing 116 (20) (2013) 38–45.

[33] R.G. Kumar, Y.S. Kumaraswamy, Investigation and classification of ECG beat using input output additional weighted feed forward neural network, in: International Conference on Signal Processing, Image Processing &Pattern Recognition (ICSIPR), 2013, pp. 200–205.

[34] H. Chen, B.-C. Cheng, G.-T. Liao, T.-C. Kuo, Hybrid classification engine for cardiac arrhythmia cloud servicein elderly healthcare management, J. Vis. Lang. Comput. 25(6) (2014) 745–753.

[35] R. Ahmed, S. Arafat, Cardiac arrhythmia classification using hierarchical classification model, in: International Conference on Computer Science and Information Technology (CSIT), 2014, pp. 203–207.

[36] M. Sarfraz, A.A. Khan, F.F. Li, Using independent component analysis to obtain feature space for reliable ECG arrhythmia classification, in: IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2014, pp. 62–67.

[37] H.L. Tran, V.N. Pham, H.N. Vuong, Multiple neural network integration using a binary decision tree to improve the ECG signal recognition accuracy, Int. J. Appl. Math. Comput. Sci.24 (3) (2014) 647–655.

[38] Luz, Eduardo José da S., et al. "ECG-based heartbeat classification for arrhythmia detection: A survey." Computer Methods and Programs in Biomedicine (2015).

[39] Hina Sharma , "Recommendation system for cardiovascular disease risk prediction": Thesis IITR-2015.

[39] Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng C-K, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: Components of

a New Research Resource for Complex Physiologic Signals. Circulation 101(23):e215-e220 [Circulation Electronic Pages; http://circ.ahajournals.org/cgi/content/full/101/23/e215]; 2000 (June 13).

[40] Scikit-learn: Machine Learning in Python ,Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, Édouard Duchesnay; 12(Oct):2825−2830, 2011.

List of Publications

[1] Bharat Goel, Durga Toshniwal, Hina Sharma ,"Multistage classification for cardiovascular disease risk prediction", Published in Big Data Analytics '2015.