# Electrocardiogram (ECG) Abnormality Detection Using Wavelet Decomposition and Support Vector Machine (SVM)

MADE BY

YASH VARDHAN SHANKHDHAR  
20ME01033

Table Of Contents: -

1.Introduction

2.Basic Outline

3.Methodology

4.Result

5.Discussion

6.Conclusion

7.Appendix

INTRODUCTION:

Biosignals provide communication between [biosystems](https://www.sciencedirect.com/topics/engineering/biosystems) and are our primary source of information on their behavior. Interpretation and transformation of signals are major topics of this text. Biosignals, like all signals, must be carried by some form of energy. Biosignals can be measured directly from their biological source, but often external energy is used to measure the interaction between the [physiological system](https://www.sciencedirect.com/topics/engineering/physiological-system) and external energy. Measuring a biosignal entails converting it to an electric sign al using a device known as a [biotransducer](https://www.sciencedirect.com/topics/engineering/biotransducer). The resultant [analog signal](https://www.sciencedirect.com/topics/engineering/analog-signal) is often converted to a digital (discrete-time) signal for processing in a computer.

Biosignals and the systems that produce them have several important properties: they can be stationary or nonstationary, linear or nonlinear, and deterministic or stochastic (i.e., random). Biosignals often contain noise, which is an unwanted signal component.

[Biosystems](https://www.sciencedirect.com/topics/engineering/biosystems) modeling is a powerful analytical tool for investigating living systems. Two very different models have been developed to represent [physiological systems](https://www.sciencedirect.com/topics/engineering/physiological-system): [analog models](https://www.sciencedirect.com/topics/engineering/analog-model) and system models. Each representation has different strengths and weaknesses.

The goal of this book is to present the most important and fundamental of the many powerful signal and systems analysis tools available to [biomedical engineers](https://www.sciencedirect.com/topics/engineering/biomedical-engineer).

BASIC OUTLINE:

The signal with the abnormality has its output adjusted to 1, whereas the typical signal has its output adjusted to 0.

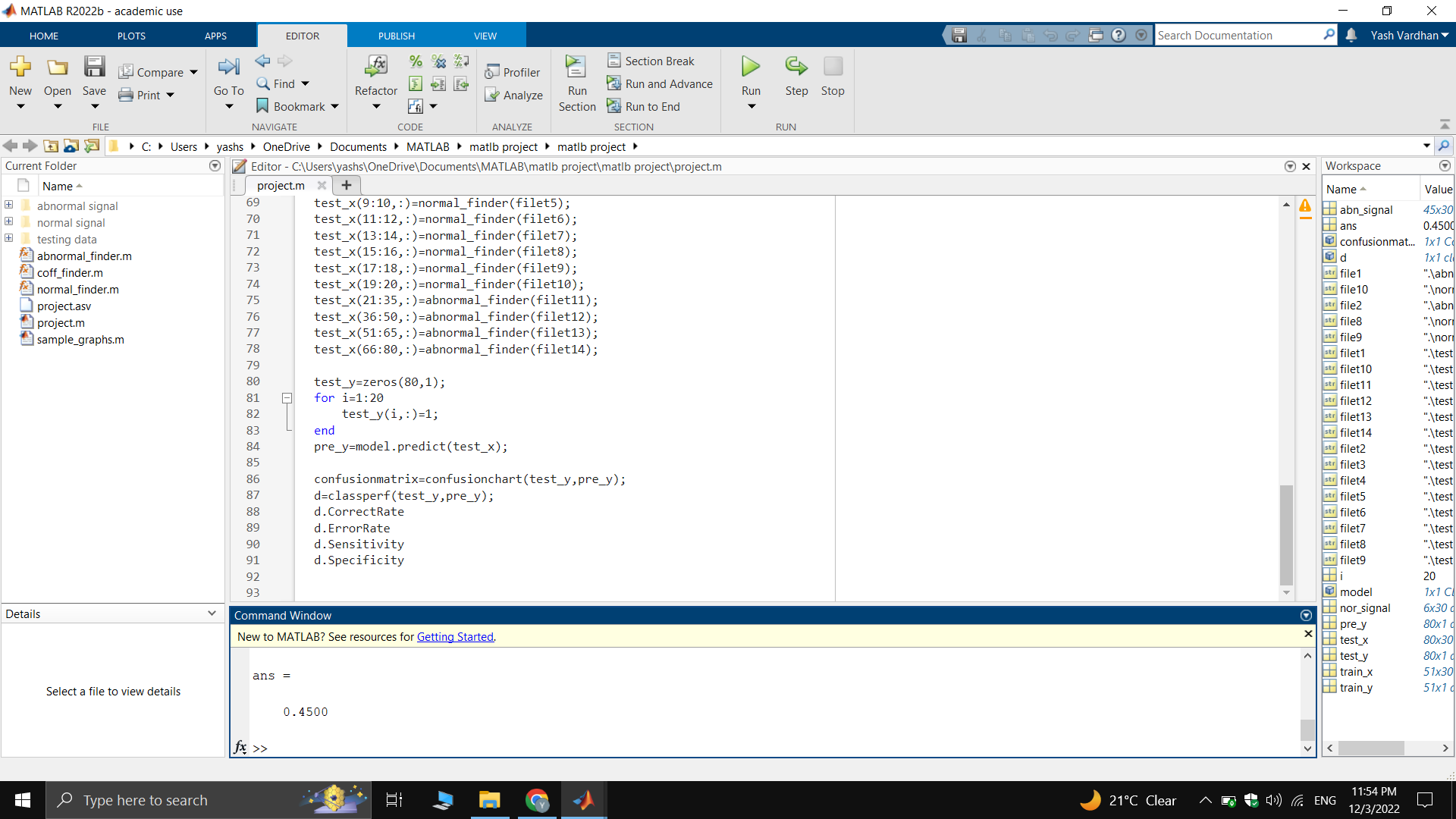
Here, we designate a subset of the whole data set for training and the remaining subset for testing, and determine accuracy by comparing the outputs of the trained and tested models.

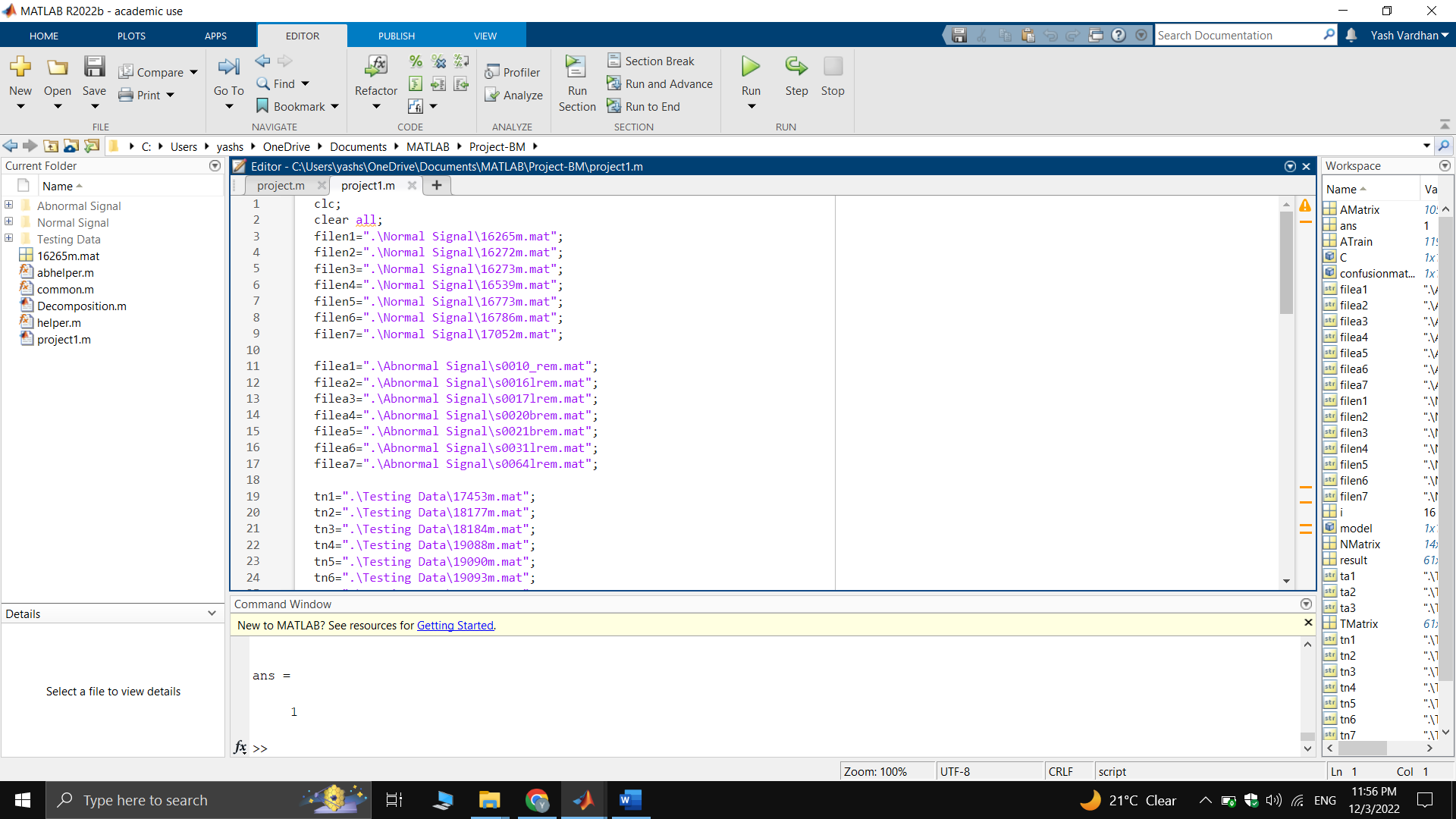
If the model's accuracy is over 80%, it is in excellent shape.

METHODOLOGY:

* Initially, we want MATLAB files corresponding to aberrant and normal ECG signals.
* We obtained seven mat files for both aberrant and normal ECG signals. The faulty ECG mat file contains 15 signals, totaling 105 signals, whereas the normal ECG mat data has 2 signals, totaling 14 signals.
* Here, 16265m, 16272m, 16273m, 16420m, 16483m, 16539m, and 16773m represent Id for normal ECG signals. s0010rem, s00141rem, s00151rem, s00161rem, s00171rem, s0020brem, s0021aerm Represents ID for irregular ECG signals (These are taken from Physio bank website).
* Using the reload function, the samples are uploaded to MATLAB and loaded into command space. Each signal receives an AECG No.
* Here, we first vertically combine the seven abnormal signals with the Vertcat function, followed by the normal signals, then assigning each signal a NECG number.
* Here, AECG and NECG files are both in struct format, and '. Val' is required to retrieve their values. Here, we must determine the Detail and Approximate Coefficients for both the Normal and Abnormal ECG signals.
* Five detail co-efficient and five approximate co-efficient are obtained for both Abnormal and Normal ECG signals, and three components are determined for each co-efficient: skewness, kurtosis, and standard deviation.
* We receive 15 new elements for Abnormal of Approximate Coefficients and 15 new elements for Normal ECG.
* A 105 x 15 matrix 'A' is constructed to contain all 15 detail Coefficients of Abnormal ECG Signal.
* Similarly, for the Normal ECG Signal, a 14x15 matrix 'B' is created, and the for loop is repeated until 14 (14 corresponds to the number of signals), after which the values are obtained and equal to the matrix B.
* Next, we create a matrix 'C' of size 105 x 15 for the abnormal ECG signal, iterate through the for loop until we've equaled all 105 values in the matrix, and then do the same for the 14 values in the normal ECG signal. Approximate Coefficients of Normal ECG signals are stored in Matrix 'D.
* Up to this point, we have discovered both exact and approximate Coefficients for normal and abnormal ECG signals.
* It is time to combine the coefficient matrices A, B, C, and D into a new matrix.
* Combine the A and B matrices vertically, then the C and D matrices vertically, and finally the combined vertically combined matrices horizontally.
* Final Matrix has a size of 119 by 31, and its name shall be Final.
* Determining which rows represent abnormalities and which represent regular data requires us to assign the outcome as a '1' or a '0'.
* We can see that the values in each column span a wide range, from 0.1 decimal places in the first column to 103 in the last. Because of this, we employ normalize and keep the value in the same matrix.
* The first 109 rows of the final matrix represent the anomaly, therefore we assign an output of "1" to those and "0" to the rest.
* We have our dataset and coefficients ready, so we can begin training the model with a rand function of 119 numbers.
* Thereafter, we determine, out of 119, how much data must be implemented for training and testing.
* After receiving the data, we utilize the SVM function to make a prediction and evaluate the model's accuracy.

RESULTS:

* Assuming that 36 Rows were used for training and the rest was saved for testing purposes then accuracy came around 45%
  + 
* Also, if we increase the number of rows to 60 then the accuracy of the dataset increases to 100%



DISCUSSION:

In this case, we see that the accuracy is improving together with the size of the training dataset, and eventually approaches 100%. It's important to avoid extremes when selecting a training dataset size, as doing so can lead to overfitting and underfitting in the model. Here We recorded 105 ECG signals looking for abnormalities and only one signal looking for normality. There is a high likelihood of mistake in the model if the Dataset is not adequately provided throughout the training process, where the majority of rows are aberrant. Standard practice dictates that in a dataset with a total size of 100%, 70% must be used for training and the remaining 30% must be used for testing. Here Once again, various techniques exist, such as cross validation, to double-check the quality of the model's training. The data set provided is less than 70% accurate, but a large number of anomalous signals raises this to 100%, thus even if the model randomly produces an output of 0, the accuracy is higher.

CONCLUSION:

As a result, we can conclude Important data about a patient's heart can be gleaned from ECG readings, making them a valuable diagnostic tool. Abnormalities The presence of electrocardiogram (ecg) readings suggests the existence of a life-threatening heart condition. Therefore, it is crucial to recognize an irregular ECG signal and get help from medical professionals.

In this case, the suggested approach divides ECG signals into two categories: normal and pathological. The ECG signals were classified via SVM, with the detail and approximation coefficients serving as the key features. The system's effectiveness was demonstrated by an impressively high rate of success (96%) it produced.

APPENDIX:

%%Helper File

function fun=helper(fileName)

t=load(fileName);

[C1,L1]=wavedec(t.val(1,:),5,'db10');

[C2,L2]=wavedec(t.val(2,:),5,'db10');

temp(1,1:30)=common(C1,L1);

temp(2,1:30)=common(C2,L2);

fun=temp;

end

%%

function com=common(C,L)

a1=appcoef(C,L,'db10',1);

a2=appcoef(C,L,'db10',2);

a3=appcoef(C,L,'db10',3);

a4=appcoef(C,L,'db10',4);

a5=appcoef(C,L,'db10',5);

[d1,d2,d3,d4,d5]=detcoef(C,L,[1 2 3 4 5]);

skewd1=skewness(d1);

skewd2=skewness(d2);

skewd3=skewness(d3);

skewd4=skewness(d4);

skewd5=skewness(d5);

skewa1=skewness(a1);

skewa2=skewness(a2);

skewa3=skewness(a3);

skewa4=skewness(a4);

skewa5=skewness(a5);

sdd1=std(d1);

sdd2=std(d2);

sdd3=std(d3);

sdd4=std(d4);

sdd5=std(d5);

sda1=std(a1);

sda2=std(a2);

sda3=std(a3);

sda4=std(a4);

sda5=std(a5);

kd1=kurtosis(d1);

kd2=kurtosis(d2);

kd3=kurtosis(d3);

kd4=kurtosis(d4);

kd5=kurtosis(d5);

ka1=kurtosis(a1);

ka2=kurtosis(a2);

ka3=kurtosis(a3);

ka4=kurtosis(a4);

ka5=kurtosis(a5);

com=[skewd1 skewd2 skewd3 skewd4 skewd5 skewa1 skewa2 skewa3 skewa4 skewa5 sdd1 sdd2 sdd3 sdd4 sdd5 sda1 sda2 sda3 sda4 sda5 kd1 kd2 kd3 kd4 kd5 ka1 ka2 ka3 ka4 ka5];

end

function ab=abhelper(fileName)

t=load(fileName);

[C1,L1]=wavedec(t.val(1,:),5,'db10');

[C2,L2]=wavedec(t.val(2,:),5,'db10');

[C3,L3]=wavedec(t.val(3,:),5,'db10');

[C4,L4]=wavedec(t.val(4,:),5,'db10');

[C5,L5]=wavedec(t.val(5,:),5,'db10');

[C6,L6]=wavedec(t.val(6,:),5,'db10');

[C7,L7]=wavedec(t.val(7,:),5,'db10');

[C8,L8]=wavedec(t.val(8,:),5,'db10');

[C9,L9]=wavedec(t.val(9,:),5,'db10');

[C10,L10]=wavedec(t.val(10,:),5,'db10');

[C11,L11]=wavedec(t.val(11,:),5,'db10');

[C12,L12]=wavedec(t.val(12,:),5,'db10');

[C13,L13]=wavedec(t.val(13,:),5,'db10');

[C14,L14]=wavedec(t.val(14,:),5,'db10');

[C15,L15]=wavedec(t.val(15,:),5,'db10');

temp(1,1:30)=common(C1,L1);

temp(2,1:30)=common(C2,L2);

temp(3,1:30)=common(C3,L3);

temp(4,1:30)=common(C4,L4);

temp(5,1:30)=common(C5,L5);

temp(6,1:30)=common(C6,L6);

temp(7,1:30)=common(C7,L7);

temp(8,1:30)=common(C8,L8);

temp(9,1:30)=common(C9,L9);

temp(10,1:30)=common(C10,L10);

temp(11,1:30)=common(C11,L11);

temp(12,1:30)=common(C12,L12);

temp(13,1:30)=common(C13,L13);

temp(14,1:30)=common(C14,L14);

temp(15,1:30)=common(C15,L15);

ab=temp;

end

filen1=".\Normal Signal\16265m.mat";

filen2=".\Normal Signal\16272m.mat";

filen3=".\Normal Signal\16273m.mat";

filen4=".\Normal Signal\16539m.mat";

filen5=".\Normal Signal\16773m.mat";

filen6=".\Normal Signal\16786m.mat";

filen7=".\Normal Signal\17052m.mat";

filea1=".\Abnormal Signal\s0010\_rem.mat";

filea2=".\Abnormal Signal\s0016lrem.mat";

filea3=".\Abnormal Signal\s0017lrem.mat";

filea4=".\Abnormal Signal\s0020brem.mat";

filea5=".\Abnormal Signal\s0021brem.mat";

filea6=".\Abnormal Signal\s0031lrem.mat";

filea7=".\Abnormal Signal\s0064lrem.mat";

tn1=".\Testing Data\17453m.mat";

tn2=".\Testing Data\18177m.mat";

tn3=".\Testing Data\18184m.mat";

tn4=".\Testing Data\19088m.mat";

tn5=".\Testing Data\19090m.mat";

tn6=".\Testing Data\19093m.mat";

tn7=".\Testing Data\19140m.mat";

tn8=".\Testing Data\19830m.mat";

ta1=".\Testing Data\s0557\_rem.mat";

ta2=".\Testing Data\s0558\_rem.mat";

ta3=".\Testing Data\s0559\_rem.mat";

TMatrix(1:2,1:30)=helper(tn1);

TMatrix(3:4,1:30)=helper(tn2);

TMatrix(5:6,1:30)=helper(tn3);

TMatrix(7:8,1:30)=helper(tn4);

TMatrix(9:10,1:30)=helper(tn5);

TMatrix(11:12,1:30)=helper(tn6);

TMatrix(13:14,1:30)=helper(tn7);

TMatrix(15:16,1:30)=helper(tn8);

TMatrix(17:31,1:30)=abhelper(ta1);

TMatrix(32:46,1:30)=abhelper(ta2);

TMatrix(47:61,1:30)=abhelper(ta3);

NMatrix(1:2,1:30)=helper(filen1);

NMatrix(3:4,1:30)=helper(filen2);

NMatrix(5:6,1:30)=helper(filen3);

NMatrix (7:8,1:30)=helper(filen4);

NMatrix(9:10,1:30)=helper(filen5);

NMatrix(11:12,1:30)=helper(filen6);

NMatrix(13:14,1:30)=helper(filen7);

AMatrix(1:15,1:30)=abhelper(filea1);

AMatrix(16:30,1:30)=abhelper(filea2);

AMatrix(31:45,1:30)=abhelper(filea3);

AMatrix(46:60,1:30)=abhelper(filea4);

AMatrix(61:75,1:30)=abhelper(filea5);

AMatrix(76:90,1:30)=abhelper(filea6);

AMatrix(91:105,1:30)=abhelper(filea7);

ATrain=[NMatrix;AMatrix];

y=zeros(119,1);

for i=1:14

y(i,:)=1;

end;

for i=15:119

y(i,:)=0;

end;

ytest=zeros(61,1);

for i=1:16

ytest(i,:)=1;

end;

model=fitcsvm(ATrain,y);

result=model.predict(TMatrix)

confusionmatrix=confusionchart(ytest,result);

C=classperf(ytest,result);

C.CorrectRate

C.ErrorRate

C.Sensitivity

C.Specificity