1. What is the accuracy, Precision, and Recall for each class prediction with 10-fold cross validation?

Code:

%Load DNA Data

cellDNAData= xlsread('C:\tmp\CellDNA.xls');

X = cellDNAData(:,1:13);

Y = cellDNAData(:,14);

%Normalizied all predicators

ZX=zscore(X);

% 0 = bacteria not interesting 0, 1 = bacteria interesting

for r = 1:length(Y)

if Y(r,1) > 0

Y(r,1) = 1;

else

Y(r,1) = 0;

end

end

%SVM with RBF kernel

SVMdl = fitcsvm(ZX,Y,'KernelFunction','rbf','ClassNames',[0,1]);

%10 fold validation

CVSVMModel = crossval(SVMdl);

[labels scorePred] = kfoldPredict(CVSVMModel);

%Confusion Matrix

CFM = confusionmat(Y, labels)

accuracy = sum(diag(CFM))/sum(CFM(:));

Recall = CFM([1],[1]) / (CFM([1],[1]) + CFM([1],[2]));

Precision = CFM([1],[1]) / (CFM([1],[1]) + CFM([2],[1]));

[XFP,YTP,T,AUC]=perfcurve(Y,scorePred(:,1),'0');

figure

plot(XFP,YTP)

xlabel('False positive rate');

ylabel('True positive rate');

title('ROC for SVM [Class 0]');

annotation('eclipse','String',strcat('Area under the curve = ',num2str(AUC)));

[XFP,YTP,T,AUC]=perfcurve(Y,scorePred(:,1),'1');

figure

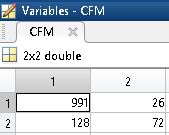
plot(XFP,YTP)

xlabel('False positive rate');

ylabel('True positive rate');

title('ROC for SVM [Class 1]');

annotation('eclipse','String',strcat('Area under the curve = ',num2str(AUC)));



accuracy = 0.8735

Recall = 0.9744

Precision = 0.8856

1. Create an ROC curve plot for each class prediction.

[XFP,YTP,T,AUC]=perfcurve(Y,scorePred(:,1),'0');

figure

plot(XFP,YTP)

xlabel('False positive rate');

ylabel('True positive rate');

title('ROC for SVM [Class 0]');

annotation('textbox','String',strcat('Area under the curve = ',num2str(AUC)));

[XFP,YTP,T,AUC]=perfcurve(Y,scorePred(:,1),'1');

figure

plot(XFP,YTP)

xlabel('False positive rate');

ylabel('True positive rate');

title('ROC for SVM [Class 1]');

annotation('textbox','String',strcat('Area under the curve = ',num2str(AUC)));

