Classification Probability

[View all machine learning examples](https://www.mathworks.com/solutions/machine-learning/examples.html)

It is often useful to analyze the probability of data classification instead of relying on labels to reveal the class that a particular observation falls into. This example demonstrates how to visualize classification probabilities for the Naive Bayes classification algorithm.

Load Data

load fisheriris

X = meas(:,1:2);

y = categorical(species);

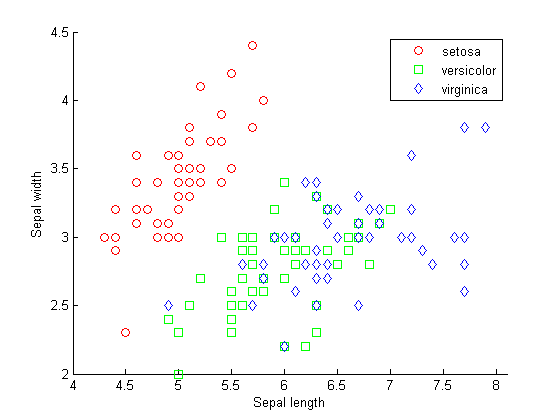
labels = categories(y);

gscatter(X(:,1), X(:,2), species,'rgb','osd');

xlabel('Sepal length');

ylabel('Sepal width');

N = size(meas,1);



Train a Naive Bayes Classifier

mdlNB = NaiveBayes.fit(X,y);

Predict Species Using the Naive Bayes Model

[xx1, xx2] = meshgrid(4:.01:8,2:.01:4.5);

ypred = predict(mdlNB,[xx1(:) xx2(:)]);

postNB = posterior(mdlNB,[xx1(:), xx2(:)]);

Visualize Posterior Distribution for Each Class

sz = size(xx1);

s = max(postNB,[],2);

figure(1),

surf(xx1,xx2,reshape(postNB(:,1),sz),'EdgeColor','none'), hold on

surf(xx1,xx2,reshape(postNB(:,2),sz),'EdgeColor','none')

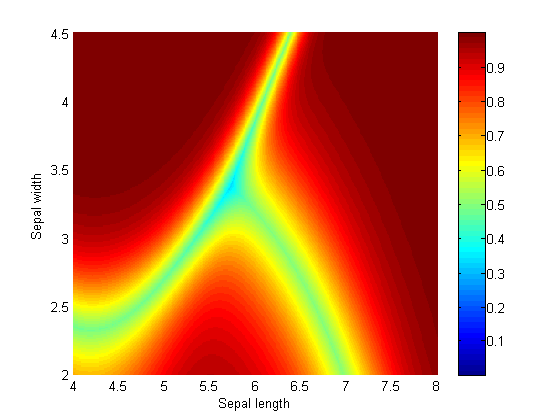
surf(xx1,xx2,reshape(postNB(:,3),sz),'EdgeColor','none')

xlabel('Sepal length');

ylabel('Sepal width'); colorbar

set(gcf,'renderer','painters')

view(2)



The closer the data gets to the decision surface, the less probable it is that the data belongs to a certain class.

figure('Units','Normalized','Position',[0.25,0.55,0.4,0.35])

surf(xx1,xx2,reshape(postNB(:,1),sz),'FaceColor','red','EdgeColor','none'), hold on

surf(xx1,xx2,reshape(postNB(:,2),sz),'FaceColor','blue','EdgeColor','none')

surf(xx1,xx2,reshape(postNB(:,3),sz),'FaceColor','green','EdgeColor','none')

xlabel('Sepal length');

ylabel('Sepal width');

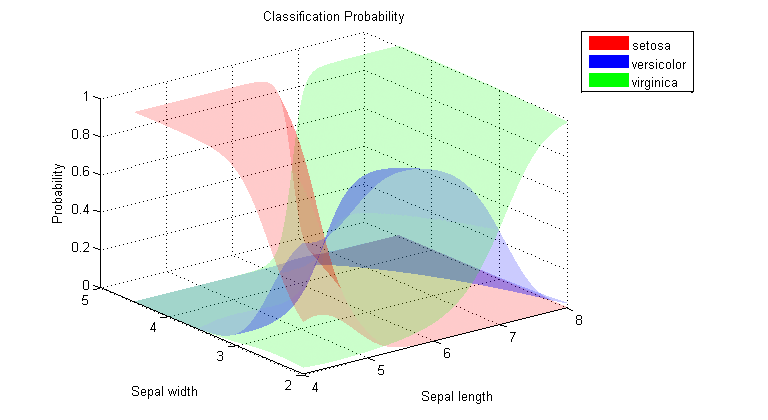
zlabel('Probability');

legend(labels)

title('Classification Probability')

alpha(0.2)

view(3)



The classification probability distributions visualized individually.

Datasets and References

Fisher's iris data consists of measurements on the sepal length, sepal width, petal length, and petal width for 150 iris specimens. There are 50 specimens from each of three species. This dataset is shipped with [Statistics and Machine Learning Toolbox™](https://www.mathworks.com/products/statistics.html) .