Sensitivity, Specificity, and ROC Curves

Prof. Sharon McNicholas

STATS 780/CSE 780

1

Introduction

- We have discussed stratification of the training/test (or training/validation/test) split.

- Recall that stratified sampling, here at least, means that the proportion of each class in the training set and test set (and validation set, where applicable) is the same.

- For example, we saw that a stratified 50/50 training/test split of the iris data that left 25 flowers of each species in the training set and 25 flowers of each species in the test set.

- I mentioned that stratification can be particularly important in certain situations, e.g., when the "event" is rare in a binary classification problem.

2

Sensitivity and Specificity

- Terminology comes from World War II.

- Today, people tend to think in terms of a (binary) situation where disease is either present or absent.

- Then sensitivity is the proportion of diseased cases correctly predicted as such (true positives)...

- ...and specificity is the proportion of non-diseased cases correctly predicted as such (true negatives).

Rare Events

- I mentioned that stratification can be particularly important in certain situations, e.g., when the "event" is rare (in a binary classification problem).

- Sensitivity and specificity have an important role here too.

- Suppose the event only happens 2% of the time.

- A model that predicts that the event never happens will be 98% correct!

- But it will have 0% sensitivity.

ROC Curve (Alon data)
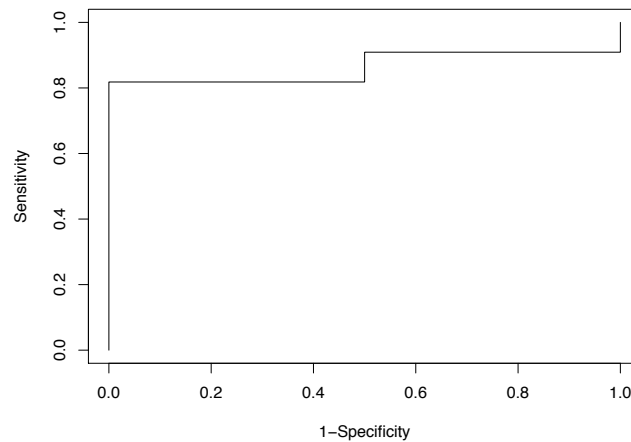
ROC Curve

- The area between the curve and the line $y = x$ can be taken as a measure of the quality of the classifier.

- This is known as area under the curve (AUC).

- Some people use Specificity for the x-axis, rather than $1 -$ Specificity.

- The ROC curve can be used to compare classification methods; let's look at an example from Hastie et al. (2009).
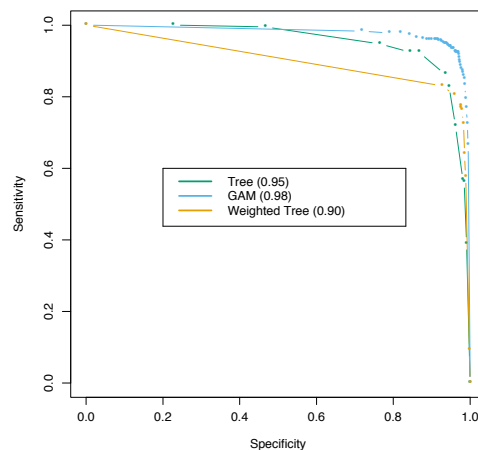
FIGURE 9.6. *ROC curves for the classification rules fit to the* spam *data. Curves that are closer to the north-east corner represent better classifiers. In this case the GAM classifier dominates the trees. The weighted tree achieves better sensitivity for higher specificity than the unweighted tree. The numbers in the legend represent the area under the curve.*

7

Comments

- In a minute, we will look at sensitivity, specificity, and some ROC curves in R.

- Now, I want to mention that this is the end of what I consider the core material for the course.

- I hope you have enjoyed this material; putting it together in a (somewhat) coherent way has been a challenge but I have enjoyed it.

- However, this is not the end of the course material; next, we will look at a smorgasbord of topics.

- The idea here will be to say a little about a lot and, hopefully, to leave you with a good starting place (in each case) for further investigations.

8