Evaluating the Predictive Performance for Two Classes

- Important R functions are:
  - Display a cross-tabulation of observed (rows) against predicted (column) class memberships: **`gmodels::CrossTable( )`**
  - Provide detailed statistics: **`caret::confusionMatrix( )`**. See online help **`help(confusionMatrix)`**.
  - Receiver operating characteristic (ROC) curve: **`pROC::roc( )`**
- For the sake of terminology let us call one outcome as ***"positive"***, which is usually an outcome of interest leading to action.
  Think as a medical doctor for whom ***"positive"***, means a test indicates the presence of a disease.
  Usually ***rare classes*** (loan default, insurance fraud, disease outcome in a screening test, spam text messages etc.) are labelled "positive"
  Positives are usually coded as factor level 1 whereas the negatives are set to 0.
- This convention makes sense under specific test scenarios, but can be arbitrary if both classes are "value-free".

- For rare positives an intuitive negative prediction will lead to a small error rate equal to the frequency of the rare positives. This provides the motivation for ***conditional error rates***.
- For just two classes we get "***confusion matrix***", which has the observed <u>true classes in its rows</u> and the <u>predicted class in its columns</u>:

	Predicted Negative	**Predicted Positive**
True Negative	*True Negative (TN)*	*False Positive (FP)*
True Positive	*False Negative (FN)*	*True Positive (TP)*

- This allows to calculate several ***key statistics***:

Name	Definition	Synonyms
False pos. rate	$FP/(TN+FP)$	Type I error, $1-$ specificity (row perspective) x-axis of ROC
True pos. rate	$TP/(FN+TP)$	$\boldsymbol{sensitivity}$, $1-$ Type II error, power, ***recall*** (row perspective) y-axis of ROC
True neg. rate	$TN/(TN+FP)$	***specificity*** (row perspective)
Pos. pred. value	$TP/(FP+TP)$	***Precision***, $1-$ false discovery proportion (column perspective)

Total Accuracy	$\dfrac{TP+TN}{TN+FP+FN+TP}$	
Total Error Rate	$\dfrac{FP+FN}{TN+FP+FN+TP}$	$1-\text{accuracy}$

The Radio Operating Characteristics Curve Roc

- Error rates are affected by the threshold probability $\delta \in [0,1]$ of assigning an object y_i either to the positive $\hat{y}_i = 1$ of negative $\hat{y}_i = 0$ class in relation to the observed features $\mathbf{x}_i$:

$$\hat{y}_i = 0 \text{ if } \Pr(y_i = 1|\boldsymbol{x}_i) \leq \delta$$
$$\hat{y}_i = 1 \text{ if } \Pr(y_i = 1|\boldsymbol{x}_i) > \delta$$

- Depending on the threshold δ the total error rate, specificity and sensitivity change:
  - For $\delta = 0$ the ***false positive rate*** $(1-\text{specificity})$ is zero percent and ***the true positive rate*** (sensitivity) is zero percent.
  - In contrast, for $\delta = 1$ the ***false positive rate*** is 100 % and the ***true positive rate*** is 100%.

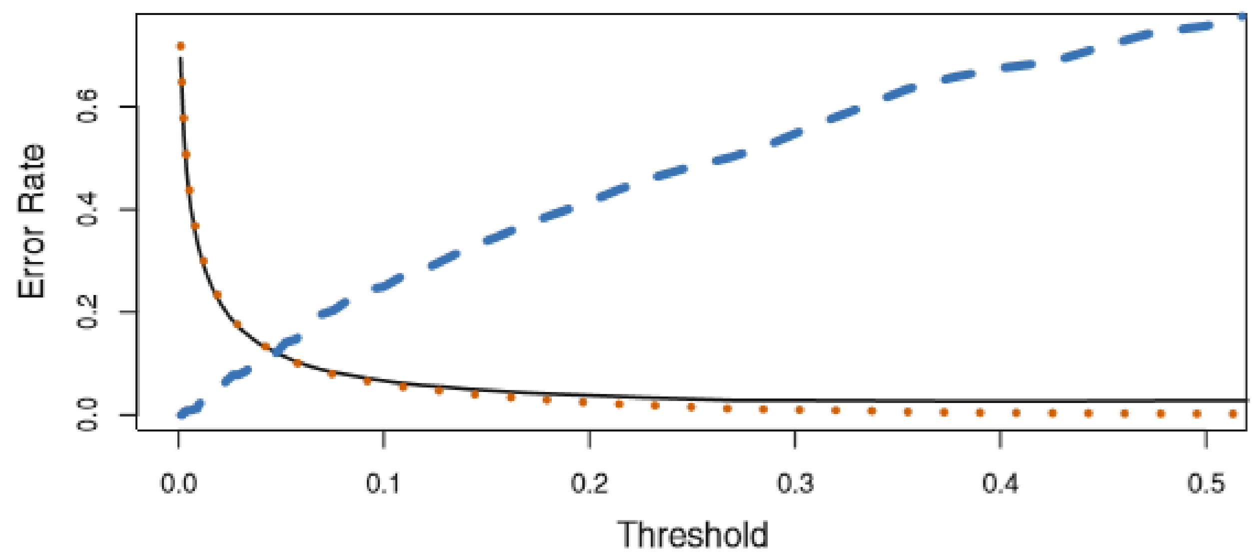

FIGURE 4.7. *For the* `Default` *data set, error rates are shown as a function of the threshold value for the posterior probability that is used to perform the assignment. The black solid line displays the overall error rate. The blue dashed line represents the fraction of defaulting customers that are incorrectly classified, and the orange dotted line indicates the fraction of errors among the non-defaulting customers.*

- Error rates of example:
  - Fraction of defaulting incorrectly classified ⇨ $1 - sensitivity$: $FN/(FN + TP)$
  - Fraction of non-defaulting incorrectly classified ⇨ $1 - specificity$: $FP/(TN + FP)$

- The ROC (radio operating characteristic) plots the false positive rate (1 − specificity) against the sensitivity.

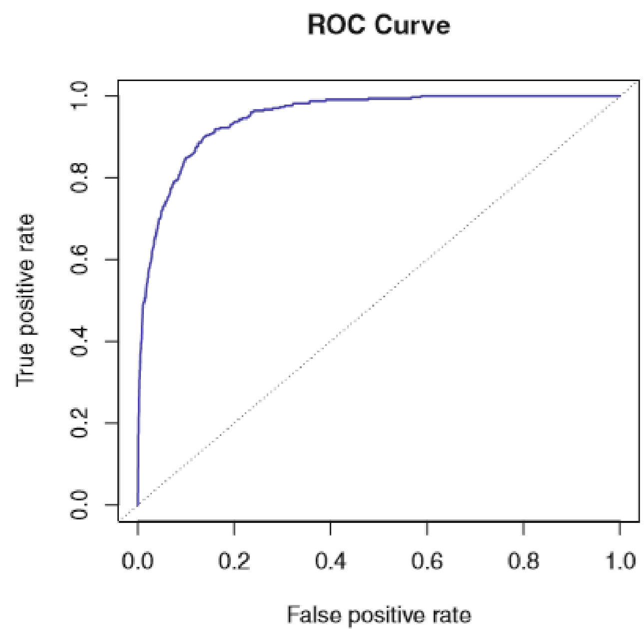

FIGURE 4.8. *A ROC curve for the LDA classifier on the* `Default` *data. It traces out two types of error as we vary the threshold value for the posterior probability of default. The actual thresholds are not shown. The true positive rate is the sensitivity: the fraction of defaulters that are correctly identified, using a given threshold value. The false positive rate is 1-specificity: the fraction of non-defaulters that we classify incorrectly as defaulters, using that same threshold value. The ideal ROC curve hugs the top left corner, indicating a high true positive rate and a low false positive rate. The dotted line represents the "no information" classifier; this is what we would expect if student status and credit card balance are not associated with probability of default.*

- For an uninformative predictor for which $sensitivity = 1 - specificity$ for any threshold $\delta \in [0,1]$.
- For a well-discriminating classifier, the ROC curve takes a rectangular shape:

  The upper left corner denotes that predictor for which all positives are properly classified, and no negative is falsely classified.

- The area underneath the ROC curve (AUC) measures the discriminating power of a classifier:

  A. *0.9 to 1.0*: ***outstanding***
  B. *0.8 to 0.9*: ***good***
  C. *0.7 to 0.8*: ***fair***
  D. *0.6 to 0.7*: ***poor***
  E. *0.5 to 0.6*: ***no discrimination***

- Two competing classifiers may have an identical AUC but different shapes of the ROC curve.

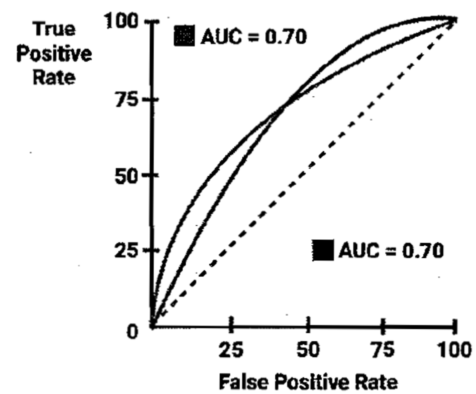

Figure 10.5: ROC curves may have different performance despite having the same AUC