Bayes Threshold

Bookmark this page

For the two models (Logistic Regression and kNN Classification) trained previously, we can now compute the values for each metric discussed.

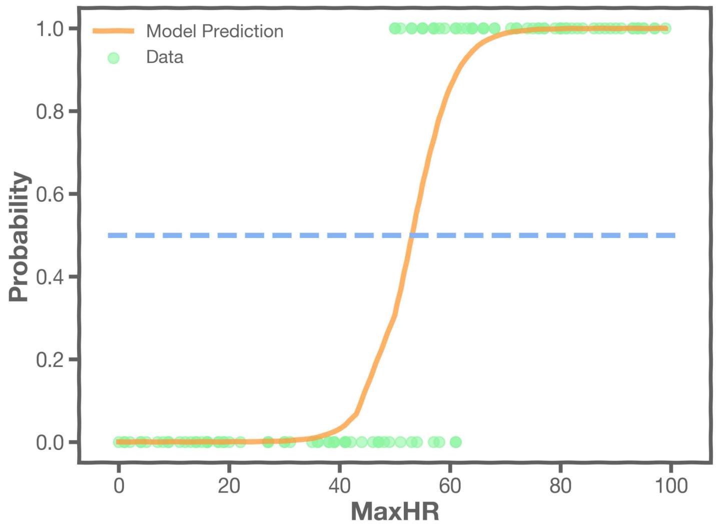
Classification Metric	Formula	Logistic Regression	kNN Classification
Accuracy	$\frac{TN+TP}{TP+TN+FP+FN}$	62%	62%
Sensitivity/Recall/TPR	$\frac{TP}{TP+FN}$	72%	75%
Specificity/TNR	$\frac{TN}{TN+FP}$	50%	47%
Precision	$\frac{TP}{TP+FP}$	62%	62%
F1 Score	$\frac{2*Precision*Recall}{Precision+Recall}$	67%	68%

Bayes threshold

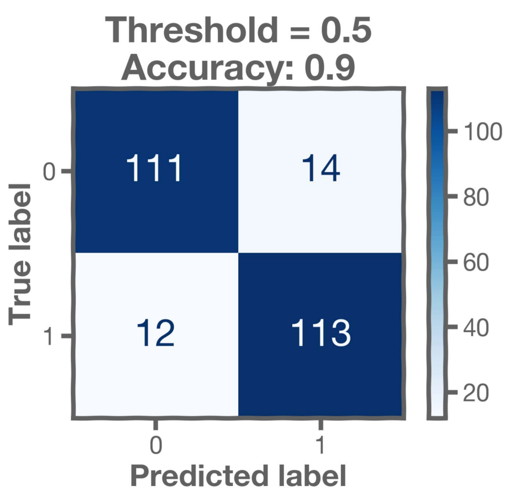
All of these metrics are based on classification. The models return probabilities which we turn into classifications by introducing a threshold, the Bayes threshold.

EXAMPLE:

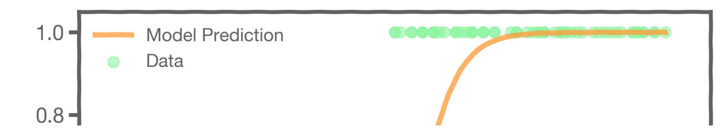
We decide that a probability threshold of $P \geq 0.5$ implies a decision of $y = 1$ (that is, a "yes" value).

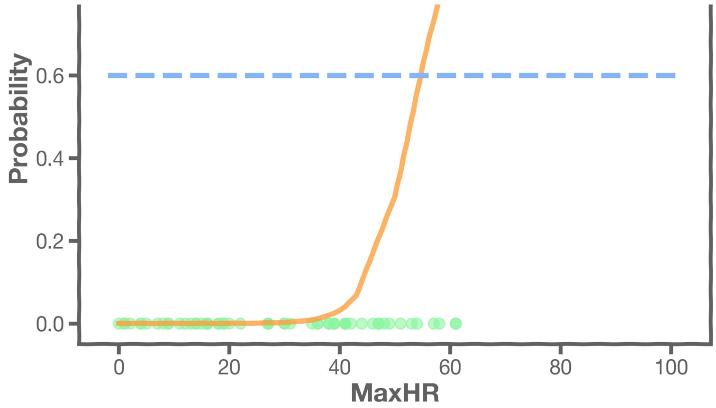


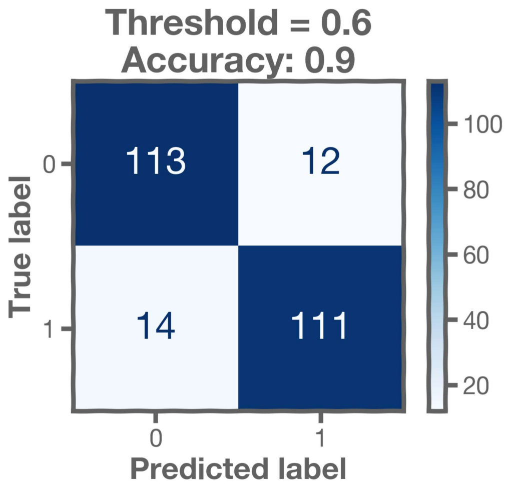
This threshold results in a confusion matrix as follows:



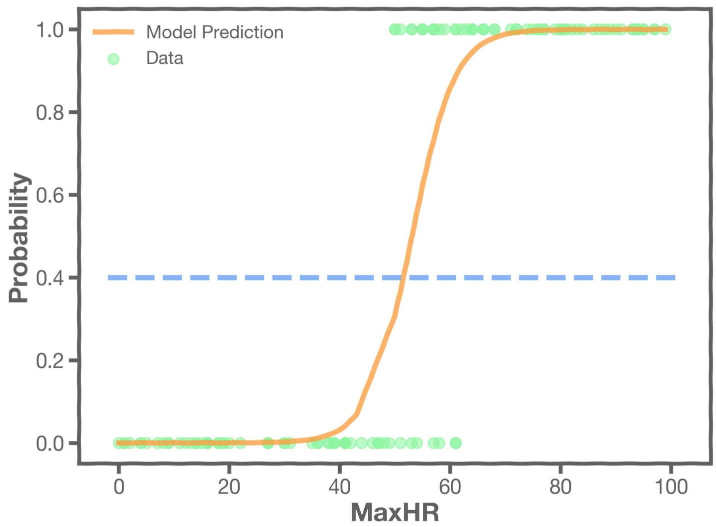
If we change that threshold, the elements on the confusion matrix change. If we increase it, we predict fewer positives and more negatives. Here is the case for $P \geq 0.6$.

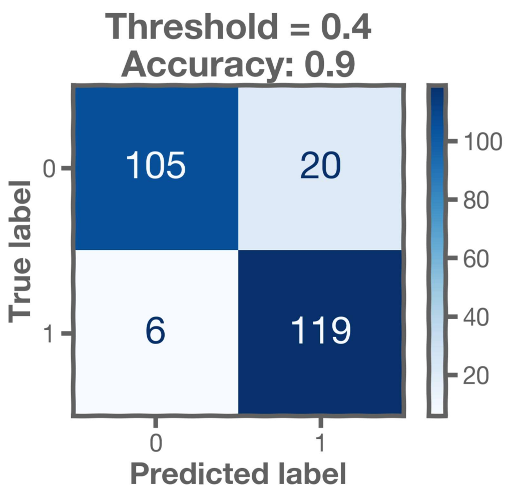






If the threshold is decreased, then we predict more positives and less negatives. Here is the case for $P \geq 0.4$.

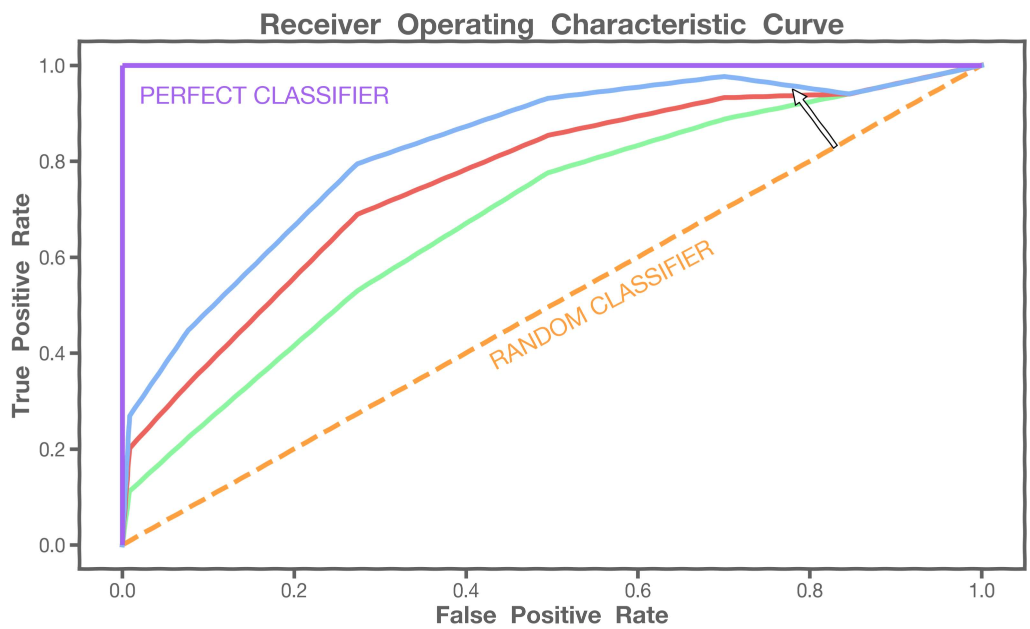




The Receiver Operating Characteristics Curve

We can visualize how things change as we vary the threshold with the Receiver Operating Characteristics curve (ROC). You saw this earlier in the course, in Max's video!

ROC curves are a very powerful tool as a statistical performance measure in detection/classification theory. The ROC curve is created by plotting the true positive rate (TPR) against the false positive rate (FPR) at various threshold settings. They were first developed by radar engineers during World War II for detecting enemy objects on battlefields.



As the image indicates, a perfect classifier has an ROC curve that goes straight up and then across to the right at a true positive rate of 100%. A completely random classifier has a curve that is a diagonal line from the bottom left (0%, 0%) to the top right (100%, 100%).

Discussion Board (External resource)

Click OK to have your username and e-mail address sent to a 3rd party application.

OK

< Previous Next >

Connect

Calculator