STAT 562 Lecture 6 Evaluating Classification Model

Beidi Qiang

SIUE

Thresholding

Many classification models returns a probability. You can use the returned probability "as is" or convert the returned probability to a binary value.

- To map a logistic regression value to a binary category, you must define a classification threshold.
- Value above that threshold indicates positive class; a value below indicates negative class 0.
- Thresholds are problem-dependent, and are therefore values that you must tune.
- Choosing a threshold is assessing how much you'll suffer for making a mistake.

Accuracy

Accuracy is one metric for evaluating classification models. Informally, accuracy is the fraction of predictions our model got right.

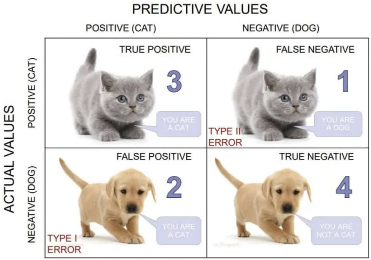
$$Accuracy = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}}$$

In many cases, accuracy is a poor or misleading metric.

- Most often when different kinds of mistakes have different costs
- Class imbalance, when positives or negatives are extremely rare

True Positives and False Positives

For class-imbalanced problems, useful to separate out different kinds of errors using a 2×2 confusion matrix (error-matrix):



- A true positive (TP) is an outcome where the model correctly predicts the positive class.
- A true negative (TN) is an outcome where the model correctly predicts the negative class.
- A false positive (FP) is an outcome where the model incorrectly predicts the positive class.
- A false negative (FN) is an outcome where the model incorrectly predicts the negative class.

Precision and Recall

Using the terms of positives and negatives, we define the following metrics.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

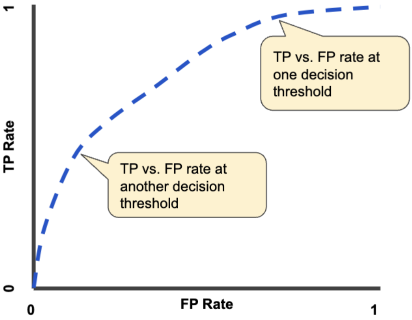
- Precision is the proportion of positive identifications was actually correct.
- Recall is the proportion of actual positives was identified correctly.

ROC curve

An ROC curve (receiver operating characteristic curve) is a graph showing the performance of a classification model at all classification thresholds. This curve plots two parameters:

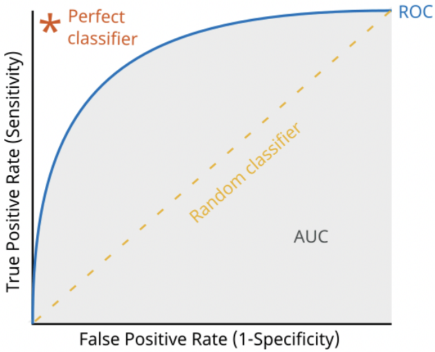
- True Positive Rate (TPR) = TP/(TP+FN). This is the same as recall.
- False Positive Rate (FPR) = FP/(FP+TN).

An ROC curve plots TPR vs. FPR at different classification thresholds.



AUC

Area under the ROC Curve (AUC) measures the entire two-dimensional area underneath the entire ROC curve (0,0) to (1,1).



AUC provides an aggregate measure of performance across all possible classification thresholds. It represents the probability that a random positive example has higher score (predicted probability to be positive) than a random negative example.