



Why

How should we compare two classifiers?

Definitions

Let a^1, \dots, a^n in A be a dataset of inputs and b^1, \dots, b^n in B be a dataset of labels. Let $G : A \rightarrow B$ be a classifier.

For each $i = 1, \dots, n$, the classifier associates a prediction $G(a^i)$ with a_i . The prediction $G(a^i)$ is *correct* if $G(a^i) = b^i$ and *incorrect* (*wrong*, *error*) if $G(a^i) \neq b^i$. The *error rate* is the proportion of the dataset for which the classifier's prediction is an error. In other words, $1/n |\{i \mid G(a^i) \neq b^i\}|$.

Given two classifiers, we may be interested in the one with a lower error rate. Or given that these two are the results of different inductors applied on a shared dataset, we can use a different separate dataset to *validate* these.

Boolean case

In the case that $B = \{-1, 1\}$, we call the class -1 *negative* and the class $+1$ *positive*. Similarly, we call an example (a^i, b^i) a *negative example* if $b^i = -1$ and a *positive example* if $b^i = 1$.

We call the result of G on a^i a *true positive* if $b^i = 1$ and $G(a^i) = 1$, a *true negative* if $b^i = -1$ and $G(a^i) = -1$, a *false negative* (or *type II error*, read “type two error”) if $b^i = 1$ and $G(a^i) = -1$ and a *false positive* (or *type I error*, read “type one error”) if $b^i = -1$ and $G(a^i) = 1$.

The *false positive rate* (*false negative rate*) of a classifier on a dataset is the proportion of dataset elements for which it predicts a *false positive* (*false negative*).

The *true positive rate* (or *sensitivity*, *recall*) of the classifier G on the dataset is the proportion of positive examples which G labels positive. The *true negative rate* (or *specificity*, *selectivity*) is the proportion of negative examples which G labels negative.

On the other hand, the *false alarm rate* is the proportion of negative examples which G (incorrectly) labels positive. The *precision* (or *positive predictive value*) of G is the proportion of all examples which the classifier labels *positive* whose label is positive.

