Why

How should we compare two classifiers?

Definition

Let $u_1, \ldots, u_n$ in $\mathcal{U}$ be a dataset of inputs and $v_1, \ldots, u_n$ in $\mathcal{V}$ be a dataset of labels. Let $G : \mathcal{U} \to \mathcal{V}$ be a classifier.

For each $i = 1, \ldots, n$, the classifier associates a prediction $G(u_i)$ with input u_i. The prediction $G(u_i)$ is *correct* if $G(u_i) = b_i$ and *incorrect* (*wrong*, an *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 incorrect. In other words, the error rate is $(1/n) |\{i \mid G(a_i) \neq b_i\}|$.

Given two classifiers, it is natural to prefer the one with the lower error rate. If these two classifiers are the results of different inductors applied to the same dataset, we can use a 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 (u_i, v_i) a *negative example* if $b_i = -1$ and a *positive example* if $b_i = 1$.

We call the result of u_i under classifier G a *true positive* if $v_i = 1$ and $G(u_i) = 1$, a *true negative* if $v_i = -1$ and $G(v_i) = -1$. In these two cases, the classifier has predicted the label correctly.

We call the result of u_i under classifier G a *false positive* (or *type I error*, read "type one error") if $v_i = -1$ and $G(u_i) = 1$ and a *false negative* (or *type II error*, read "type two error") if $v_i = 1$ and $G(u_i) = -1$. In these two cases the classifier has predicted the label incorrectly. Notice that the words positive and negative reference the classifier's prediction, not the true label.

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 G on a dataset is the proportion of positive examples which G 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. The *true negative rate* (or *specificity, selectivity*) is the proportion of negative examples which G labels negative. The *false alarm rate* is the proportion of negative examples which G *incorrectly* labels positive.

Classifier Errors

Classifiers

Factorials   Predictors

Size of Direct Product   Sequences

Number of Set Products   Finite Set Examples   Lists

Natural Products   Number of Disjoint Unions   Direct Products

Natural Sums   Set Numbers

Finite Sets

Recursion Theorem   Equivalent Sets

Natural Order   Function Inverses   Family Unions and Intersections

Peano Axioms   Function Images   Function Composites   Families

Equivalence Relations   Functions

Natural Induction

Relations

Ordered Pair Projections

Set Products

Natural Numbers   Set Powers   Generalized Set Dualities

Set Partitions   Intersection of Empty Set   Successor Sets   Unordered Triples   Set Unions and Intersections   Set Dualities

Set Intersections   Pair Unions   Ordered Pairs   Set Complements

Set Unions   Set Differences   Pair Intersections

Empty Set   Unordered Pairs

Set Specification

Set Equality

Set Inclusion

Deductions

Quantified Statements

Logical Statements

Statements

Identities   Sets

Names

Letters   Objects

5