Why

We can summarize the (label, prediction) pairs for a particular classifier on a particular dataset in a matrix.

Boolean case

Let A be a nonempty set and $B = \{-1, 1\}$. For a dataset $(a^1, b^i), \ldots, (a^n, b^n)$ in $A \times B$, and classifier $G : A \to B$, the *confusion matrix* C is defined

$$C = \begin{bmatrix} \# \text{ true negatives} & \# \text{ false negatives} \\ \# \text{ false positives} & \# \text{ true positives} \end{bmatrix} = \begin{bmatrix} C_{\text{tn}} & C_{\text{fn}} \\ C_{\text{fp}} & C_{\text{tp}} \end{bmatrix}.$$

Using this notation, $C_{\text{tn}} + C_{\text{fn}} + C_{\text{fp}} + C_{\text{tp}} = n$. $N_{\text{n}} := C_{\text{tn}} + C_{\text{fp}}$ is the number of negative examples. $N_{\text{p}} := C_{\text{fn}} + C_{\text{tp}}$ is the number of positive examples.

The diagonal elements of the confusion matrix give the numbers of correct predictions. The off-diagonal entries give the numbers of incorrect predictions for the two types of errors (see Classifier Errors).

In this notation, the false positive rate is C_{fp}/n, the false negative rate is C_{fn}/n and the error rate is the sum of these, $(C_{\text{fn}} + C_{\text{fp}})/n$.

The true positive rate is $C_{\text{tp}}/(C_{\text{fn}} + C_{\text{tp}})$. The true negative rate is $C_{\text{tn}}/(C_{\text{tn}} + C_{\text{fp}})$. The false alarm rate is $C_{\text{fp}}/(C_{\text{tn}} + C_{\text{fp}})$. The precision is $C_{\text{tp}}/C_{\text{tp}} + C_{\text{fp}}$

Confusion Matrices

Classifier Errors
Real Matrices

Linear Equations

Linear Functions
Real Vectors

N-Dimensional Space
Plane Vector Sums and Differences
Space Vectors

Distance
Plane Vectors

Arrays

Space Distance

Plane Distance

Interval Length

Intervals
Real Space

Real Line
Real Plane

Classifiers
Real Summation
Real Order
Integral Line

Factorials
Real Numbers
Rational Order
Geometry

Predictors
Sequences
Rational Numbers

Lists
Integer Arithmetic

Set Numbers
Natural Multiplicative Identity
Natural Additive Identity
Integer Products

Direct Products
Finite Sets
Arithmetic
Identity Elements

Equivalent Sets
Natural Summation
Natural Powers

Family Unions and Intersections
Function Inverses
Natural Equations
Family Operations
Integer Sums
Natural Products
Integer Order
Comparisons

Function Image
Function Composites
Families
Equations
Natural Sums
Operations
Integer Numbers
Orders

Natural Order
Recursion Theorem
Equivalence Relations
Functions
Converse Relations

Peano Axioms
Relations

Natural Induction
Set Products

Generalized Set Dualities
Natural Numbers
Set Powers

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

Set Intersections
Pair Unions
Set Complements
Ordered Pairs

Pair Intersections
Set Unions
Set Differences

Empty Set
Unordered Pairs

Set Specification

Set Equality

Set Inclusion

Standardized Accounts

Accounts

Deductions

Quantified Statements

Logical Statements

Statements
Identities
Sets

Names

Letters
Objects

4