

BINARY CLASSIFICATION CONFUSION MATRIX

| | | PREDICTED | |
|-----------------|----------|--|--|
| | | NEGATIVE | POSITIVE |
| ACTUAL EXAMPLES | NEGATIVE | <div><i>a</i></div> <div>TN - True Negative correct rejections</div> | <div><i>b</i></div> <div>FP - False Positive false alarms type I error</div> |
| | POSITIVE | <div><i>c</i></div> <div>FN - False Negative misses, type II error overlooked danger</div> | <div><i>d</i></div> <div>TP - True Positive hits</div> |



Binary Classification Confusion Matrix

A confusion matrix of binary classification is a **two by two table** formed by counting of the number of the four outcomes of a binary classifier

We usually denote them as **TP, FP, TN, and FN** instead of “the number of true positives”, and so on

Various measures can be derived from a confusion matrix
(refer next slide)



First two basic measures from the confusion matrix

Error rate (ERR) is calculated as the number of all incorrect predictions divided by the total number of the dataset. The best error rate is 0.0, whereas the worst is 1.0.

$$\text{ERR} = \frac{\text{FP} + \text{FN}}{\text{TP} + \text{TN} + \text{FN} + \text{FP}} = \frac{\text{FP} + \text{FN}}{\text{P} + \text{N}}$$

Accuracy (ACC) is calculated as the number of all correct predictions divided by the total number of the dataset. The best accuracy is 1.0, whereas the worst is 0.0. It can also be calculated by $1 - \text{ERR}$.

$$\text{ACC} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FN} + \text{FP}} = \frac{\text{TP} + \text{TN}}{\text{P} + \text{N}}$$



Other basic measures from the confusion matrix

Sensitivity (SN) is calculated as the number of correct positive predictions divided by the total number of positives. The best sensitivity is 1.0, whereas the worst is 0.0.

$$SN = \frac{TP}{TP + FN} = \frac{TP}{P}$$

Specificity (SP) is calculated as the number of correct negative predictions divided by the total number of negatives. The best specificity is 1.0, whereas the worst is 0.0.

$$SP = \frac{TN}{TN + FP} = \frac{TN}{N}$$



Other basic measures from the confusion matrix

Precision (PREC) is calculated as the number of correct positive predictions divided by the total number of positive predictions. The best precision is 1.0, whereas the worst is 0.0.

$$\text{PREC} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

False positive rate (FPR) is calculated as the number of incorrect positive predictions divided by the total number of negatives. The best false positive rate is 0.0 whereas the worst is 1.0.

$$\text{FPR} = \frac{\text{FP}}{\text{TN} + \text{FP}} = 1 - \text{SP}$$



Get daily updates on
Data Science concepts

FOLLOW NOW



DATA-DRIVEN
SCIENCE