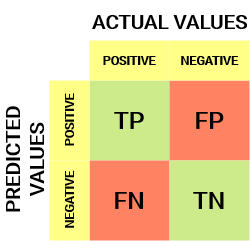
## Confusion Matrix

A confusion matrix is a performance evaluation tool in machine learning, representing the accuracy of a classification model. It displays the number of true positives, true negatives, false positives, and false negatives. This matrix aids in analyzing model performance, identifying mis-classifications, and improving predictive accuracy.

A Confusion matrix is an N x N matrix used for evaluating the performance of a classification model, where N is the total number of target classes. The matrix compares the actual target values with those predicted by the machine learning model. This gives us a holistic view of how well our classification model is performing and what kinds of errors it is making.

For a binary classification problem, we would have a **2 x 2 matrix,** as shown below, with 4 values:

[](https://cdn.analyticsvidhya.com/wp-content/uploads/2020/04/Basic-Confusion-matrix.png)

Let’s decipher the matrix:

* The target variable has two values: **Positive**or **Negative**
* The **columns**represent the **actual values** of the target variable
* The **rows**represent the **predicted values**of the target variable

But wait – what’s TP, FP, FN, and TN here? That’s the crucial part of a confusion matrix. Let’s understand each term below.

## Important Terms in a Confusion Matrix

**True Positive (TP)**

* The predicted value matches the actual value, or the predicted class matches the actual class.
* The actual value was positive, and the model predicted a positive value.

**True Negative (TN)**

* The predicted value matches the actual value, or the predicted class matches the actual class.
* The actual value was negative, and the model predicted a negative value.

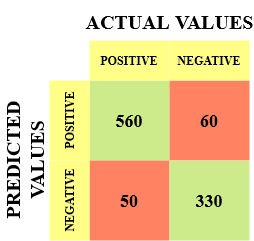
**False Positive (FP) – Type I Error**

* The predicted value was falsely predicted.
* The actual value was negative, but the model predicted a positive value.
* Also known as the type I error.

**False Negative (FN) – Type II Error**

* The predicted value was falsely predicted.
* The actual value was positive, but the model predicted a negative value.
* Also known as the type II error.

Let me give you an example to better understand this. Suppose we had a classification dataset with 1000 data points. We fit a classifier (say logistic regression or decision tree) on it and get the below confusion matrix:



The different values of the Confusion matrix would be as follows:

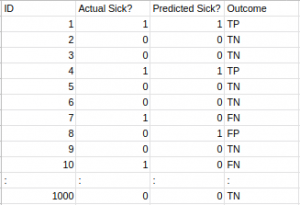
* True Positive (TP) = 560, meaning the model correctly classified 560 positive class data points.
* True Negative (TN) = 330, meaning the model correctly classified 330 negative class data points.
* False Positive (FP) = 60, meaning the model incorrectly classified 60 negative class data points as belonging to the positive class.
* False Negative (FN) = 50, meaning the model incorrectly classified 50 positive class data points as belonging to the negative class.

This turned out to be a pretty decent classifier for our dataset, considering the relatively larger number of true positive and true negative values.

## Significance of Confusion Matrix

[Equation_Accuracy](https://cdn.analyticsvidhya.com/wp-content/uploads/2020/04/Equation_Accuracy.png)

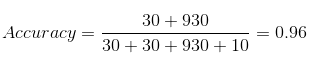
Let’s see how our model performed:



The total outcome values are:

TP = 30, TN = 930, FP = 30, FN = 10

So, the accuracy of our model turns out to be:

[](https://cdn.analyticsvidhya.com/wp-content/uploads/2020/04/Confusion-matrix_Accuracy.png)

96%! Not bad!

## Calculate Confusion Matrix for a 2-class classification problem

To calculate the confusion matrix for a 2-class classification problem, you will need to know the following:

* **True positives (TP)**: The number of samples that were correctly predicted as positive.
* **True negatives (TN)**: The number of samples that were correctly predicted as negative.
* **False positives (FP)**: The number of samples that were incorrectly predicted as positive.
* **False negatives (FN)**: The number of samples that were incorrectly predicted as negative.

Once you have these values, you can calculate the confusion matrix using the following table:

|  |  |  |
| --- | --- | --- |
| Predicted | TRUE | FALSE |
| Positive | True positives (TP) | False positives (FP) |
| Negative | False negatives (FN) | True negatives (TN) |

The confusion matrix can be used to calculate a variety of metrics, such as accuracy, precision, recall, and F1 score.