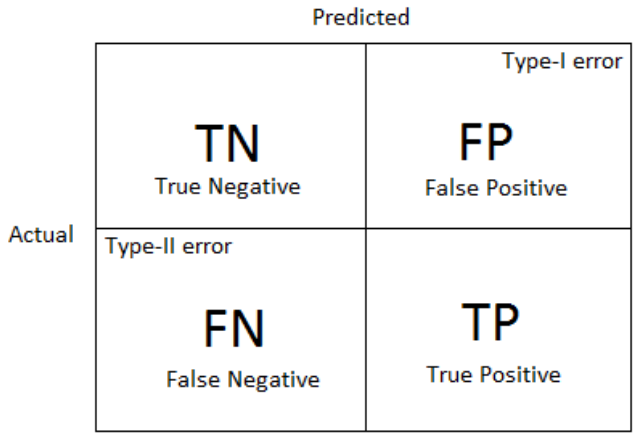
Confusion Matrix:



This is the fundamental structure of the statistical matrix. If actual does not match with the predicted we call it as Type-I and Type-II kind errors. If actual is true but our prediction goes wrong we suffer Type-II error and if actual is not true but our prediction declares it true we go into Type-I error.

Let's understand the confusion matrix with an example:

We have a report of the patients. Let's see what the actual values and predicted values we have collected.

Total: 165

		Predicted		
Actual	TN	50	FP 10	60
	FN	5	TP 100	105
		55	110	

$$\textbf{Accuracy} = \frac{TP + TN}{Total} = \frac{100 + 50}{165} = \textbf{0.91}$$

Error rate = 1 - Accuracy **Type-I,II errors**

OR

$$\frac{FN + FP}{Total} = \textbf{0.09}$$

$$\textbf{Precision} = \frac{TP}{Predicted\ Positive} = \frac{100}{110} = \textbf{0.64}$$

$$\textbf{Recall} = \frac{TP}{Actual\ Yes} = \frac{100}{105} = \textbf{0.95}$$

What we have shown above is just the simplest explanation of the confusion matrix in Machine Learning, with appropriate formulas and calculations. Now we discuss the above phenomenon concerning the Machine Learning aspect.

The head thing is that we apply an algorithm to predict our aim. Let's say we have a dataset of people and we have to predict whether the person has heart disease or not. We fitted the algorithm of **Random Forest** to the testing data, let say. The algorithm predicted our assumption 150 times true and 15 times wrong.

NOTE: We are doing this on a Categorical assumption, hence it false under the Classification ML Problem.

Let say if we apply another algorithm like **K- Nearest Neighbors** and we got a different accuracy. Now the target is to look out for the best model for the classification problem we will approach the values that, which one is giving a good accuracy for our assumption. If suppose **K-Nearest Neighbors** gives good accuracy, we'll go with it, otherwise, we approach **Random Forest**. We have multiple algorithms for our data predictions; here I defined only two for example purpose.

One most important thing I must elaborate the dimension of the confusion matrix is totally dependent upon the categories we are about to deal with for our predictions. In the above example, we have only two categories, **Yes/ No**. Hence we got a 2x2 confusion matrix form.

Let's say, if we have an example of multiple categories, for example: If we have been given the data about market stuff, and we have to predict which stuff will be sold by the customer, then our confusion matrix will be based on that of all kinds of stuff we are taken in our data. *Here data is nothing but our shop*.

Just a simple structure I would love to say that, all the diagonals of the matrix will be the data, machine predicted true with no error, remaining all non-diagonals are Type-I/ II errors.

A confusion matrix is more precisely defined as accuracy elaboration of machine algorithms for specific classification problem.