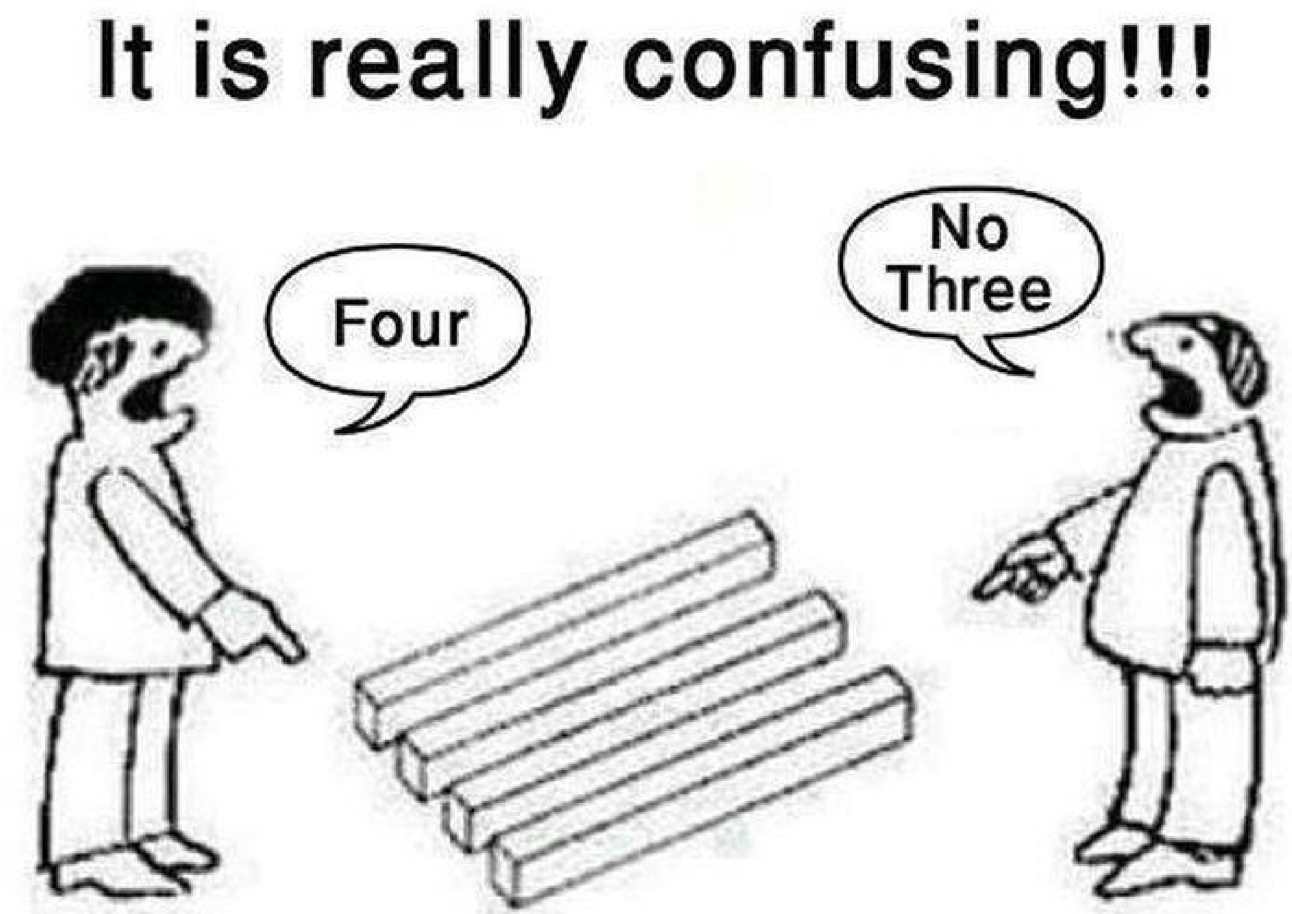
Confusion Matrix no longer Confusing!

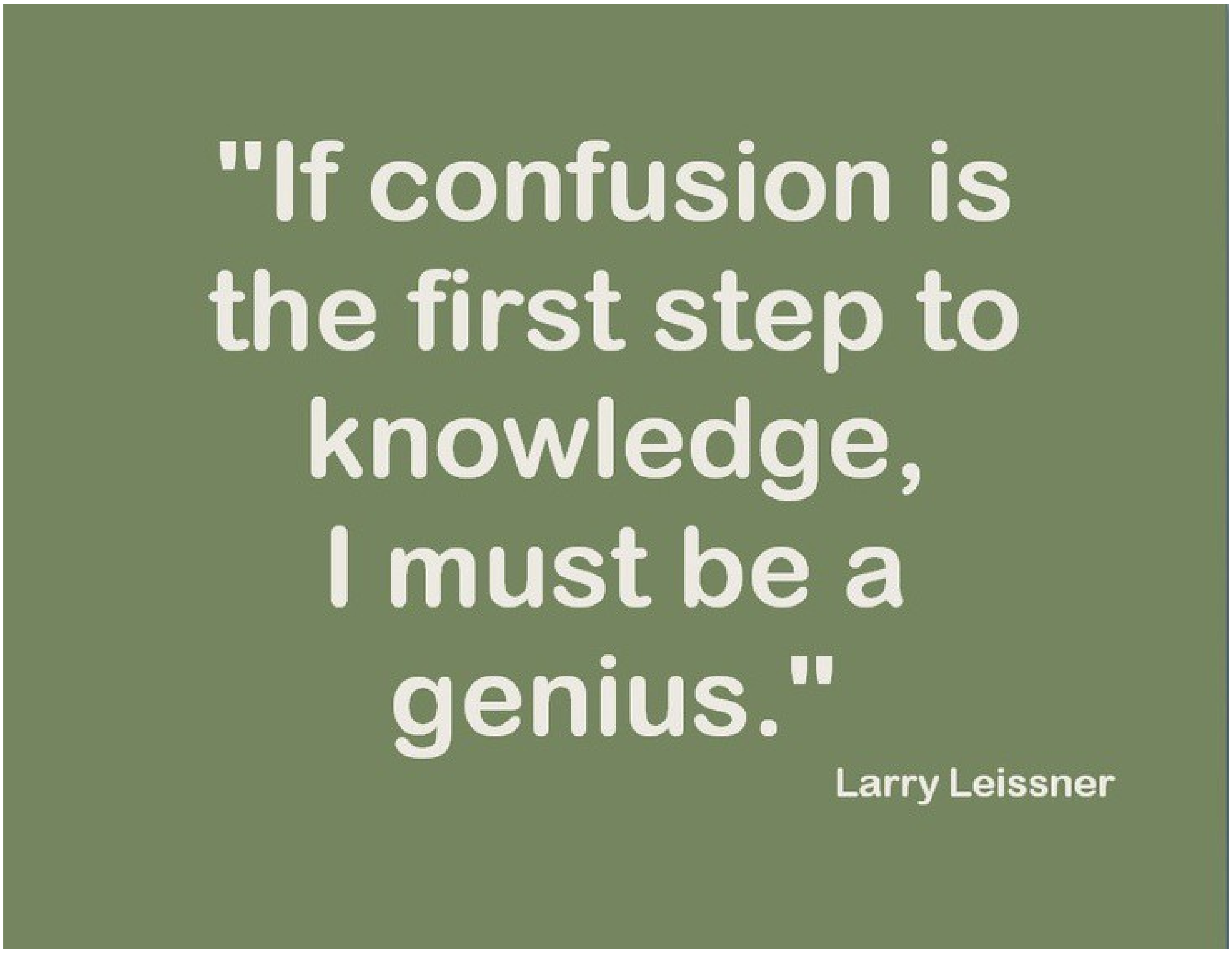
medium.com/@bhushanshewale45/confusion-matrix-no-longer-confusing-dde1615b3586

April 10, 2018



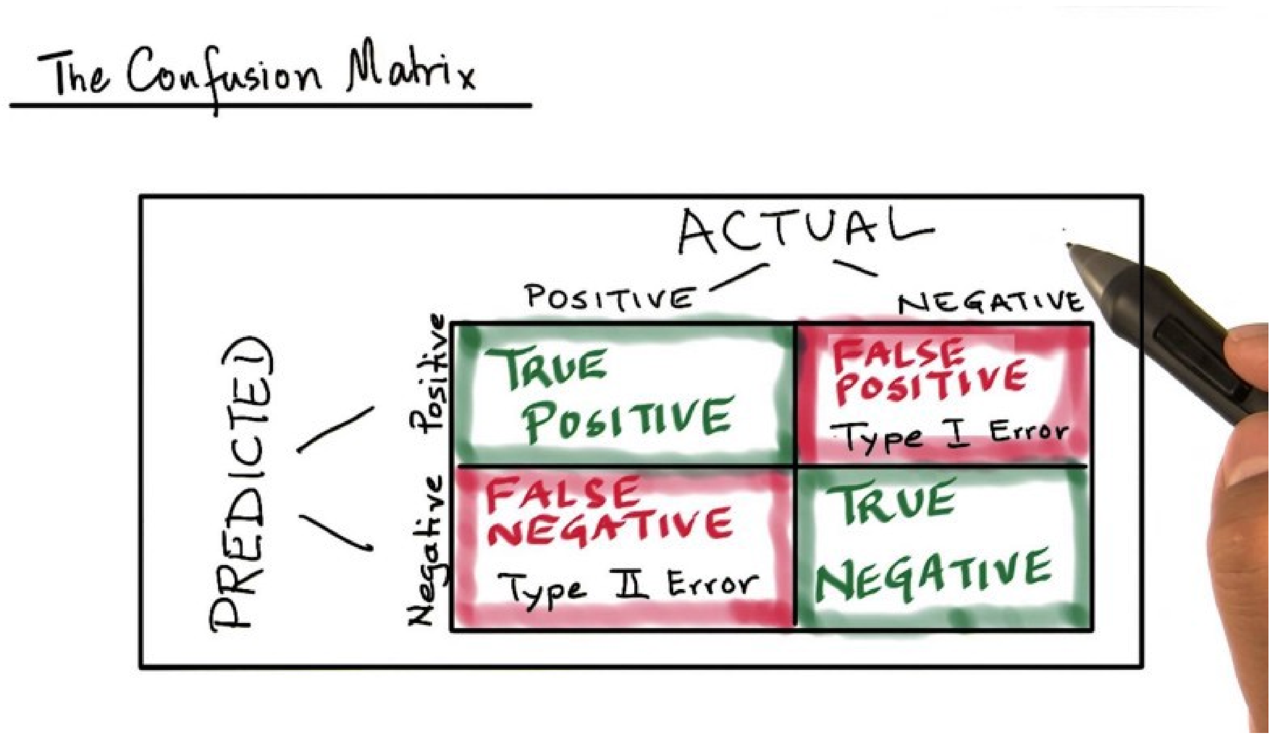
Why we use confusion matrix table?

We use confusion matrix table to describe the performance of a classification model on a set of test data for which the true values are known. It allows the visualization of the performance of an algorithm.The confusion matrix is very easy to understand, but the terminology can be confusing.



After reading this blog your reaction must be like

Without wasting too much time, Let’s start confusion matrix.



Source: Udacity

This is how confusion matrix looks like. Let’s understand each box .

> **True Positive**—we predicted “+” and the true class is “+”.

Example: we predicted yes (they have the disease), and they do have the disease.

> **True Negative**—we predicted “-” and the true class is “-”.

Example: We predicted no, and they don’t have the disease.

> **False Positive**—we predicted “+” and the true class is “-” (Type I error).

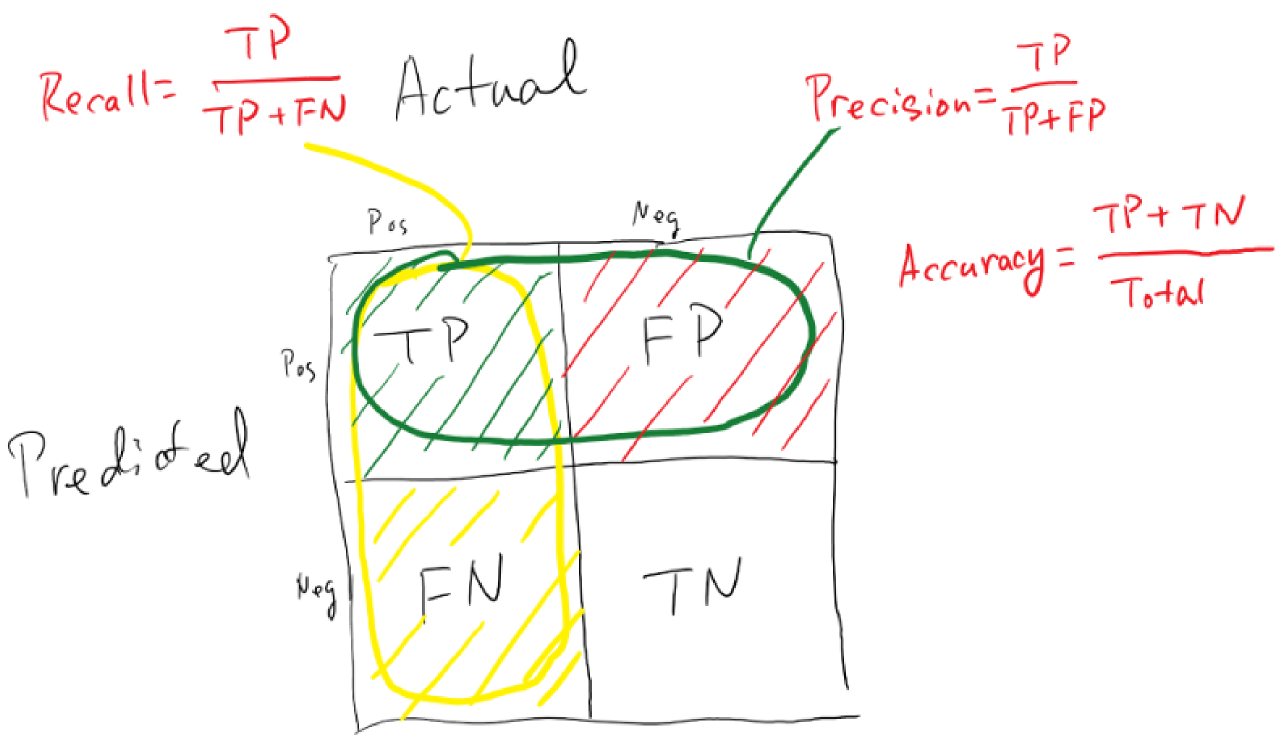
Example: We predicted yes, but they don’t actually have the disease. (Also known as a “Type I error.”)

> **False Negative**—we predicted “-” and the true class is “+” (Type II error).

Example: We predicted no, but they actually do have the disease. (Also known as a “Type II error.”)

This is one of the most popular example to understand the type 1 and type 2 error.

Measures:



Recall = TP / (TP + FN)

Interpretation

- Out of all the people that do actually have disease, how much we identified?
- The higher the better.
- We don't fail to spot many people that actually have disease.

Precision = TP / (TP + FP)

Interpretation

- Out of all the people we thought have disease, how many actually had it?
- High precision is good.
- we don't tell many people that they have disease when they actually don't.

F1 Score = 2(Precision * Recall) / (Precision + Recall) [1-Best, 0-Worst]

> f1-score is the harmonic mean of the precision and recall. As when we create a classifier we always make a compromise between the recall and precision, it is kind of hard to compare a model with high recall and low precision versus a model with high precision but low recall. *f1-score is measure that we can use to compare two models.*