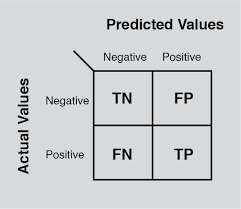
**Confusion matrix**

Define confusion matrix?

Confusion matrix is an matrix used for evaluating the performance of the model built by us, it generally compares the train dataset with our predicted value. This gives the result of how our model is performing & tell us what kind of errors it have.



a.Visualization of confusion matrix

1. **True Positive**

Actual and predicted values are equal, if actual is positive and predicted value is also positive then it is called “True Positive”.

1. **True Negative**

If actual value is negative and the predicted value is also negative then it is called “True Negative”.

1. **False Positive**

If actual value is negative but the prediction value becomes positive then it is called “False Negative”, It is also known as Type -1 error.

1. **False Negative**

If actual value is positive but the predicted value becomes negative then it is called as “False Negative”, It is also known as Type-2 error.

**Key points :**

* True positive and true negative should be high, comparatively false positive and false negative should be as low as possible.
* Type-1 error occurs when model prediction is incorrect, i.e., model predict an positive value for the actual negative value.
* Type-2 error occurs when model prediction is incorrect, i.e., model predict an negative value for the actual positive value.

Our model predict some wrong case and correct cases, but in medical related dataset or the dataset is important for feature consideration and our predicted value doesn’t gives us proper info about the dataset, we need a way to check this matrix for this purpose we will come across this

**Precision**

It tell us how many correctly predicted cases actually turned out positive, formula for this precision is ,

Precision = TP/ TP+FP

**Recall**

It tell us how many actual positive cases were able to predict correctly with our model formula for this is,

Recall = TP / TP+FN

Note : model doesn’t care about irrelevant and not retrieved T.N(true negative)

**F1-Score**

Comparing two model is difficult to decide which one better high precision and low recall or vice-versa , metric used is f1score

It take FP & FN in account and make harmonic mean of precision and recall, it perform well in imbalanced dataset.

F1-score = 2 ( Precision \* Recall) / Recall + Precision

**Code to import confusion matrix**

* **from** **sklearn.metrics** **import** confusion\_matrix