HOW CAN WE DERIVE METRICS FROM CONFUSION MATRIX

The metrics which can be derived from confusion matrix is

* Accuracy
* Precision
* F1 score

Accuracy:

The formula for calculating Accuracy is

(TP+TN)/(TP+FP+FN+TN)

Where

TP – True Positive

TN – True Negative

FP – False Positive

FN – False Negative

These classification models are used in seggregating an email as spam or non-spam.

Precision&Recall:

The model does not care about something irrelevant and not retrieved.

Therefore only TP,FP,FN are used in precision &Recall.

Precision:

Out of all the positive predicted,What percentage is truly positive.

Precision = TP/(TP+FP)

The precision value lies between 0 and 1.

Recall:

Out of the total positive,What percentage are predicted positive.It is the same as TPR

True positive Rate

Recall = TP/(TP +FN)

Recall is more important than precision in credit card Fraud detection and Medical Application.

False Positive--> Missing a critical mail because it is classified as spam.

False Positive should be as low as possible. Here,Precision is more vital as compared to recall.

When comparing different models it is difficult to decide which is better high precision or high recall.There should be a metric that combines both,one such metric is the F1 score.

F1 score:

It is the harmonic mean of Precision & Recall.It takes both false Positive and false negatives into Account.It performs well on an imbalanced dataset.

F1score = 2/(1/Precision)+2/(1/Recall)

= 2\*(Precision\*Recall)/(Precision+Recall)