Classification Matrix

How much accuracy is good?

↳ depends on problem we are solving.

Ex: 90% accuracy is not good if we are predicting cancer through scan.

∵ still 10% is wrongly predicted with cancer kind of life-threatening disease.

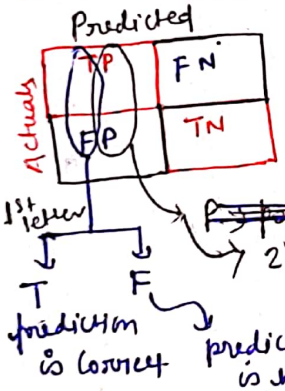
Problem with Accuracy?

↳ It does not says, which kind (FP, FN) wrong prediction has been done in %ge.

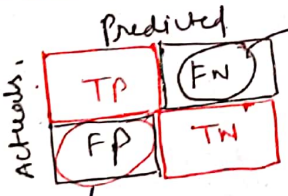
It says over all %age of accuracy.

Confusion metrics

↳ Solved by confusion metrix



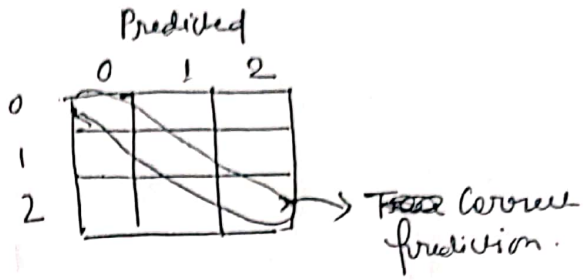
Type 1 Error vs Type 2 Error



Type2 error:- If its +ve but model predicting -ve.

Type1 Error:- i.e. If its Negative but model predicting positive

* Confusion metrix for multiclassification



* When accuracy is misleading?

Imbalanced dataset

* Precision

Model 1	Sent to spam	Not sent to spam
Spam	100	170
Not Spam	30	700

Model 2	Sent to spam	Not sent to spam
Spam	100	190
Not Spam	10	700

FP:- Mail which is Not spam but Model predicted spam. → Here in this problem FP is more important

FN:- Mail which is spam but predicted Not spam.

∴ Calculate Precision

$$\frac{TP}{TP+FP}$$

∴ $\frac{100}{100+30\uparrow} < \frac{100}{100+10}$

Model 2 is better.

* Recall

M1.	Cancer	Not Cancer
Has Cancer	1000	200
No Cancer	800	8000

M2.	Cancer	Not Cancer
Has Cancer	1000	500
No Cancer	500	8000

FP → person have no Cancer but predicted Cancer.

FN → person have Cancer predicted No. Cancer.

Here FN is more important

$$\text{recall} = \frac{TP}{TP+FN}$$

$\therefore \frac{1000}{1000+200} > \frac{1000}{1000+500}$

$\therefore$ (M_1) is better than (M_2).

F_1 score

↳ Such problem where we can not define Type1 (FP) or Type2 (FN) is more dangerous

↳ in this case we will find F_1-score.

F_1 score = Harmonic mean of Precision & recall

$= \frac{2PR}{P+R}$

↳ why HM not arithmatic mean?

∵ HM has characteristic to lie nearer to lower value

Ex:- P=0, R=100

$HM = \frac{2*0+100}{0+100} = 0$

$AM = \frac{0+100}{2} = 50$

So, it always tell F_1 score is towards lower one and we have chance to improve.

```
from sklearn metrics import
    recall_score, precision_score, f1_score.
```

Multiclass classification

Predicted

Actuals	Dog	Cat	rabbit	Total
Dog	25	5	10	40
Cat	0	30	4	34
rabbit	4	10	20	34
Total	29	45	34	

Precision: Precision for multiclass

$P_{Dog} = \frac{TP_{dog}}{(\text{Total Prediction})_{dog}} = \frac{25}{29} = 0.86$

$P_{cat} = \frac{30}{45} = 0.66$, $P_R = \frac{20}{34} = 0.58$

Combined Precision

↳ Macro precision

↳ Avg (P_{cat}, P_{dog}, P_R)

↳ weighted Precision

$\frac{\text{Total Actual dog}}{\text{Total Actuals}} * P_{dog} + \dots$

$\Rightarrow \frac{40}{108} * 0.86 + \frac{34}{108} * 0.66 + \frac{34}{108} * 0.58$

(108 = 40+34+34)

~~Precision for~~ Recall for multiclass

$R_{dog} = \left(\frac{TP_{Dog}}{\text{Total Actual of Dog}}\right) = \frac{25}{40}$

$R_c = \frac{30}{34}$, $R_R = \frac{20}{34}$

F_1-score for multiclass

$F_{1\,Dog} = \frac{2 P_D R_D}{P_D + R_D}$, $F_{1\,c} = \dots$, $F_{1\,R} = \dots$

Combined F_1-score

↳ Macro F_1 score

↳ weighted F_1 score.