Model Performance Metric

Best Example of Accuracy, Precission.

1. These are basically called your model metrics. To determine how well your model has performed we depend on these metrices.
2. → Accuracy
   → Precision
   → Recall
   → F1 score.

Let's see examples:

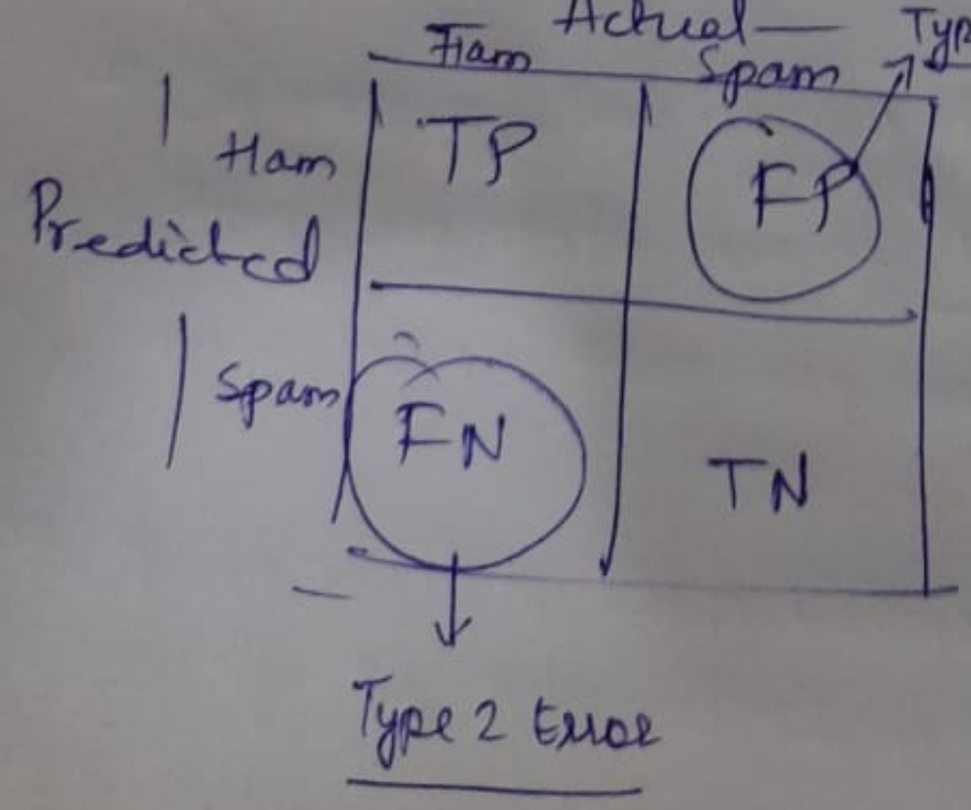
3. True weight of bag of rice is 20kg but weighing machine display weight as 8kg.
4. ↓↓↓↓↓
   → "Poor Accuracy"

Accuracy

Accuracy is defined as closeness of measured value to true value

5.
6. Precision: He weighs same bag 3 more times but still it is 8kg. This says that weighing machine is precise.

* Accuracy & Precision are independent of each other.



True positive: TP — Ham & correctly predicted as Ham

TN: Spam & Coorectly predicted as Spam.

FN: It is actually Ham but predicted as Spam.

FP: A Spam but predicted as Ham.

① Accuracy → Evaluation metrics that determines number of correct predictions made by the model.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

* Accuracy should be looked only if we have balanced dataset.

② Precision → Used when we have Imbalanced dataset

↳ Precision will tell for number of times a model predicted positive, how often it was correct.

$$\text{Precision} = \frac{TP}{TP + FP}$$

(out of number of times a model predicted positive, how often it was correct)

	AP	AN
PP	TP	FP → Pimp is ↑
PN ↓ imp ↓	FN	TN

ex: Spam detection email

③ Recall → (out of total actual positive values, how many of them were correctly predicted as positive)

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

	AP	AN
PP	TP	FP → Pimp ↓
PN	FN → imp high	TN

ex: Cancer detection

④ F1 Score →

* In some case FN & FP both are important & we need to create balance between FP & FN.

In such cases we use F1 score

F1 score is harmonic mean of Precision & Recall

	AP	AN
PP	TP	FP
PN	FN	TN

$$F1\ score = 2 * \frac{Precision * Recall}{Precision * Recall}$$

* F1 score should be used if number of samples in positive class is very less

ex: In fraud detection,

num of fraud ~~detections~~ transactions (positive) is very less as compared to number of genuine transactions (negative)

Key points

* If we have balanced dataset go ahead with accuracy metric

* In case of Imbalanced dataset, If "FP" has more importance then we make use of "Precision" & if "FN" has more importance then we make use of "Recall"

* If we want to keep balance between Precision & recall & Also, if the true values is very less then we go ahead with F1 Score.