CLASSIFICATION METRICS - 1

Spam / Non Spam filter

↳ Classification model

100 → 70 Non spam
→ 30 Spam

⇒ Logistic Regression Accuracy.

90% Accurate

Quiz time!

Quiz Ended!

if data1: 20 Cancer Patients and 100 non-Cancer Patients and data2: 80 Cancer Patients 100 non-Cancer Patients, then:

35 users have participated

A	Data1 = Imbalance, Data2 = balance	60%
B	Data1 = balance, Data2 = imbalance	29%
C	Data1 = balance, Data2 = balance	3%
D	Data1 = Imbalance, Data2 = Imbalance	9%

Fraud detection

100 → 99 Genuine
1 Not Genuine / Fraud.

Actual → 100 → 99 NF
1 F

dumbest Model

100

100
NF

Reality → 99 ✓
1 ✗

Accuracy $\frac{99}{99+1} = 99\%$

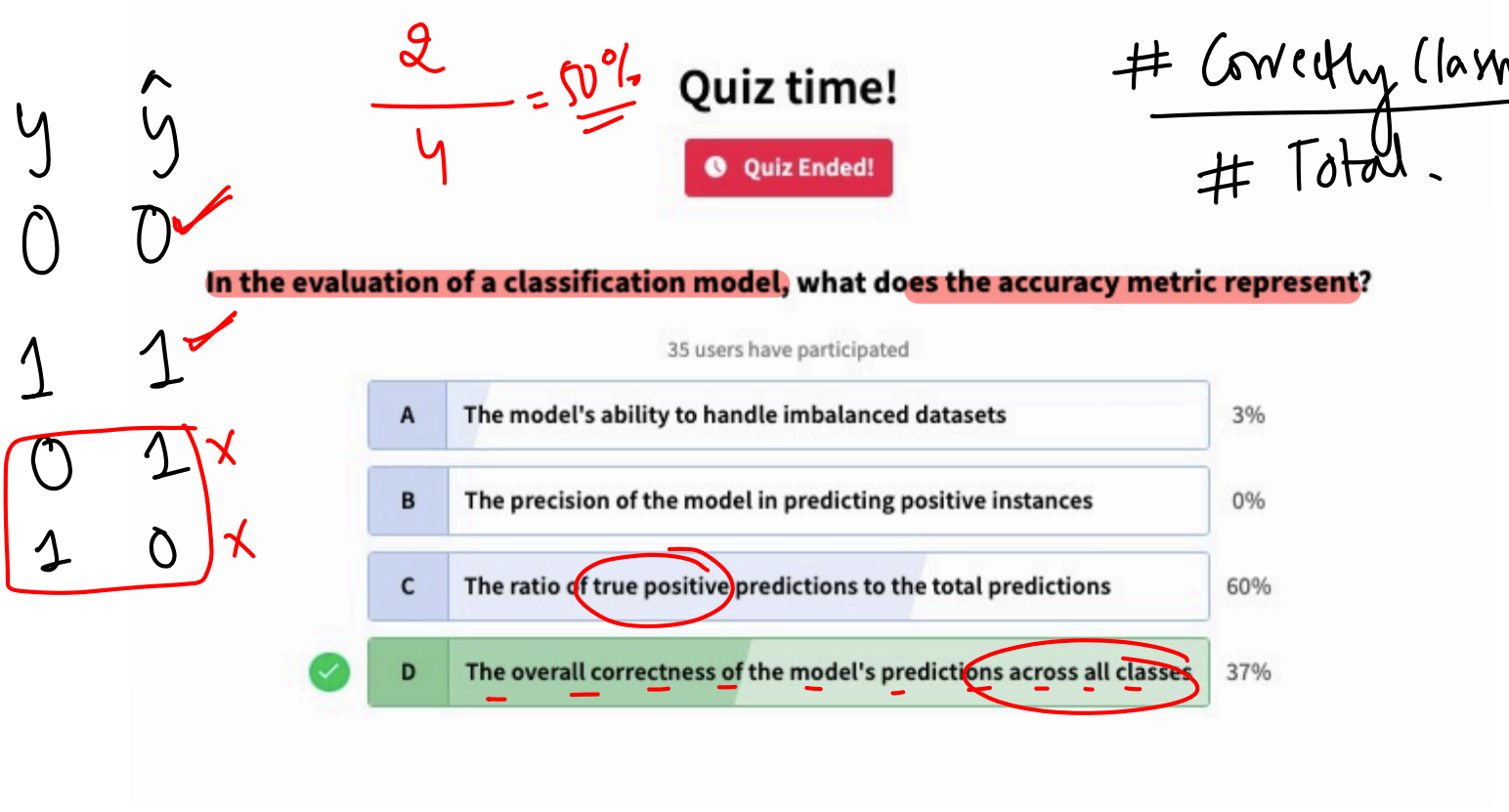
misleading in this car

Problems

$$\text{Accuracy} = \frac{\#\ \text{Correctly Classified}}{\text{Total}\ \#\ \text{Classified}}$$

(1) Doesn't work with imbalance dataset.

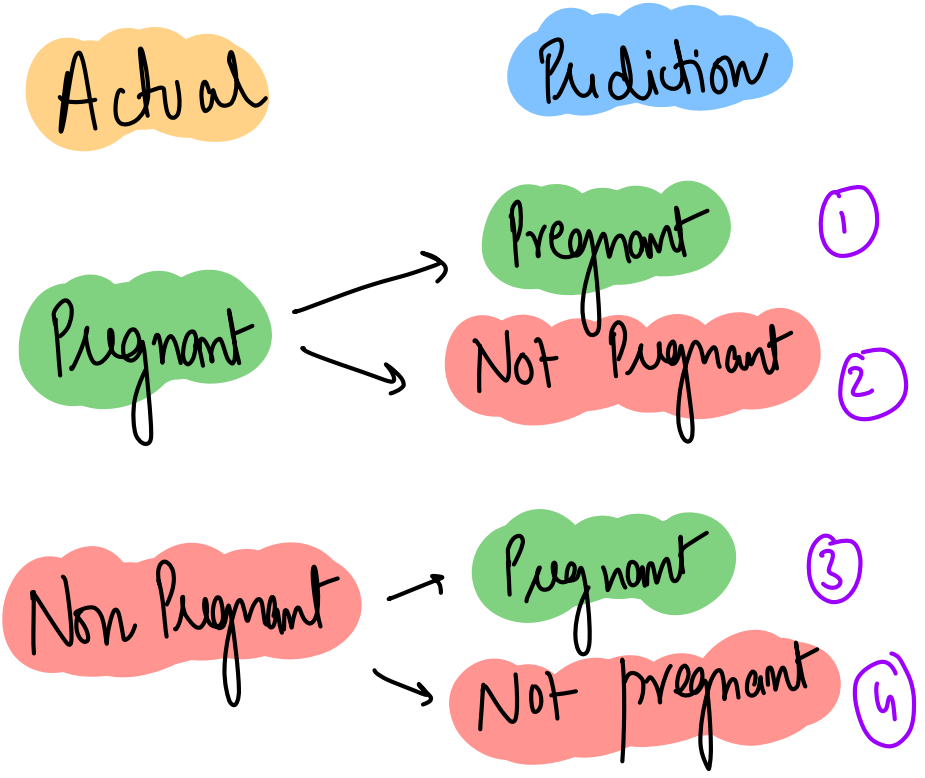
(2) Fails to capture class wise performance.


y ŷ
0 0
1 1
0 1 x
1 0 x
2/4 = 50%
Correctly Classified / # Total.
Quiz time!
Quiz Ended!
In the evaluation of a classification model, what does the accuracy metric represent?
35 users have participated
A The model's ability to handle imbalanced datasets 3%
B The precision of the model in predicting positive instances 0%
C The ratio of true positive predictions to the total predictions 60%
D The overall correctness of the model's predictions across all classes 37%

Classification ~~Matrix~~ Matrices.

Pregnant — positive

Not pregnant — negative



TRUE POSITIVE

FALSE NEGATIVE
Type II error.

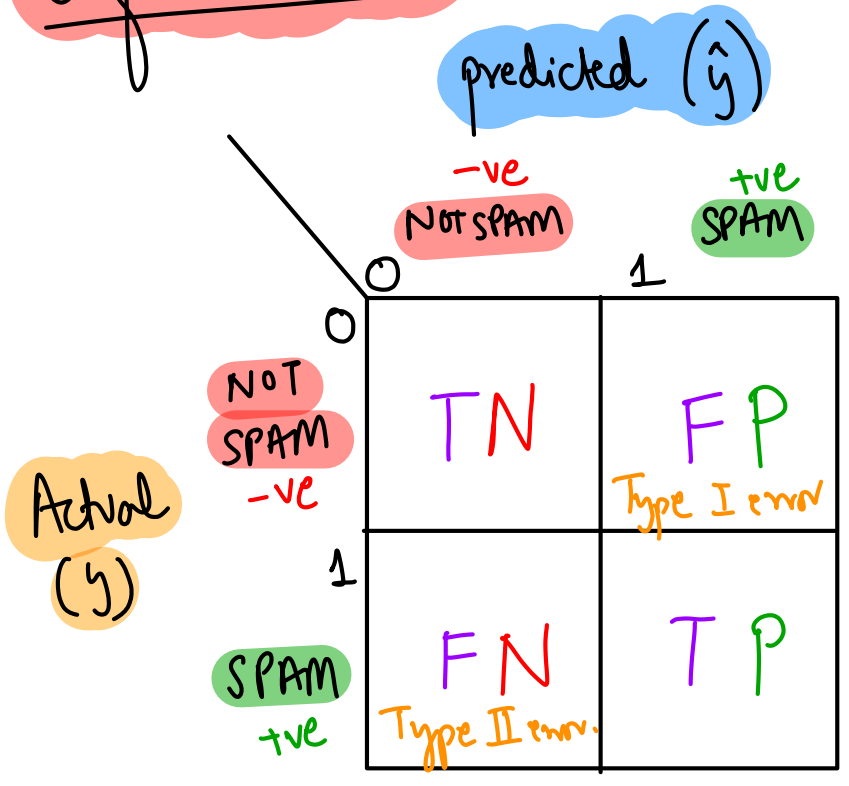
FALSE POSITIVE
Type I error

TRUE NEGATIVE

Confusion Matrix

Spam (1) +ve

Nonspam (0) -ve

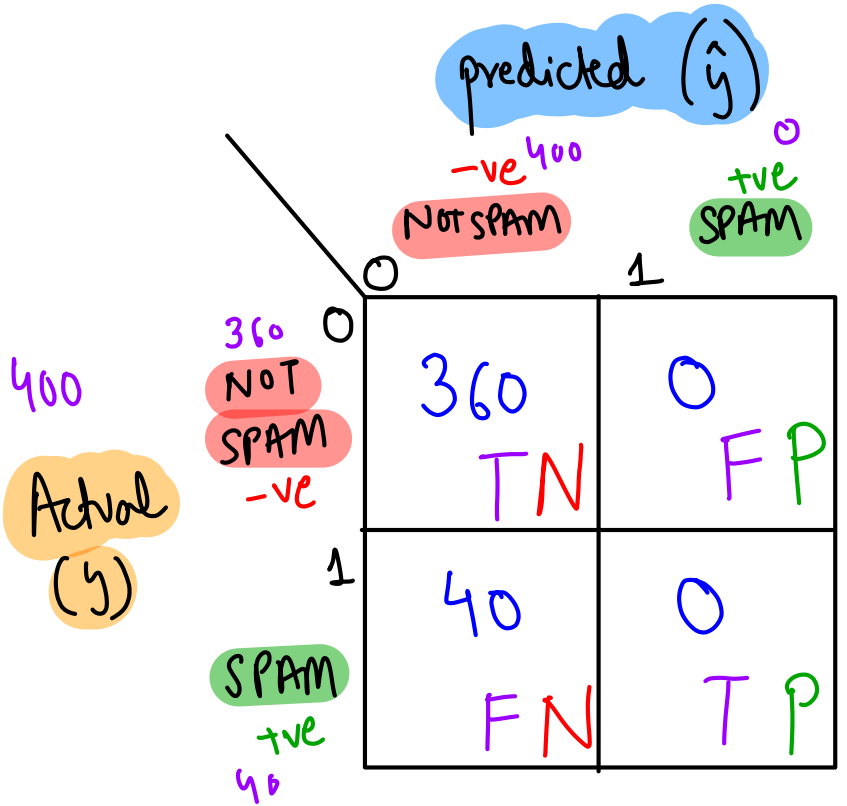


TP $y = 1$ $\hat{y} = 1$

TN $y = 0$ $\hat{y} = 0$

FP $y = 0$ $\hat{y} = 1$

FN $y = 1$ $\hat{y} = 0$



Actual
400 → 360 NS, 40 S

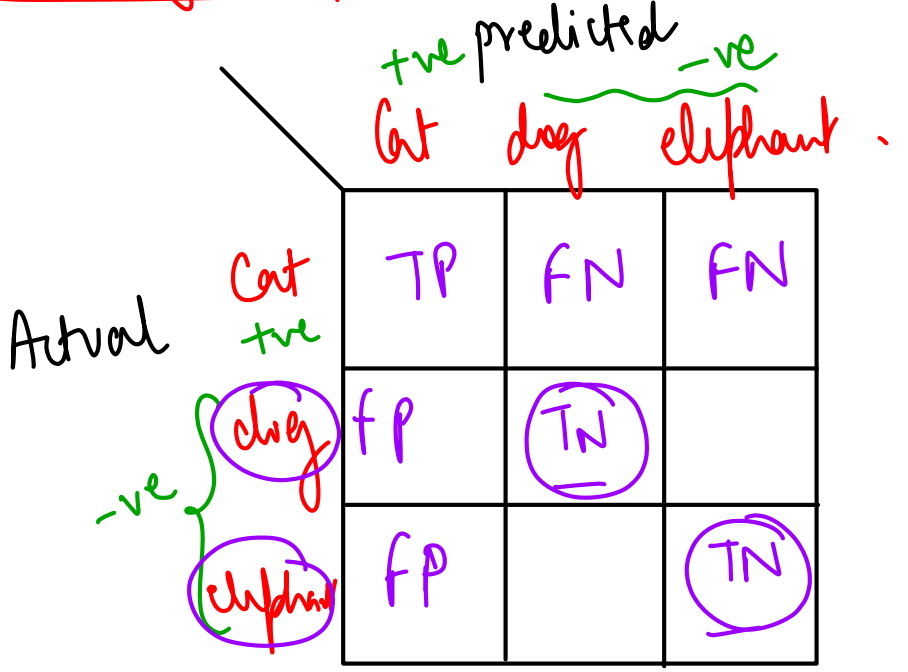
dumbest model

Predicted
400 → 400 NS, 0 S

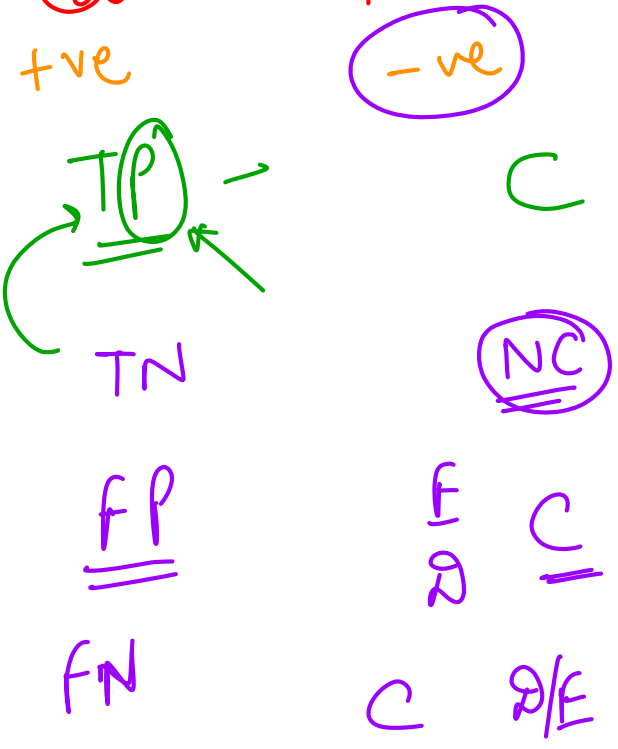
$$\text{Accuracy} = \frac{TN + TP}{TN + TP + FN + FP}$$

$$= \frac{360 + 0}{360 + 0 + 40 + 0} = \frac{360}{400} = 90\%$$

Cat dog elephant



Cat or NonCat

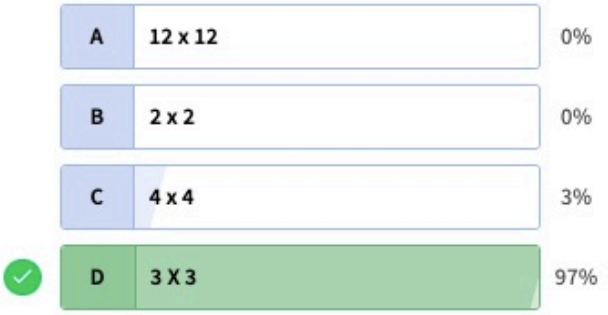


Quiz time!

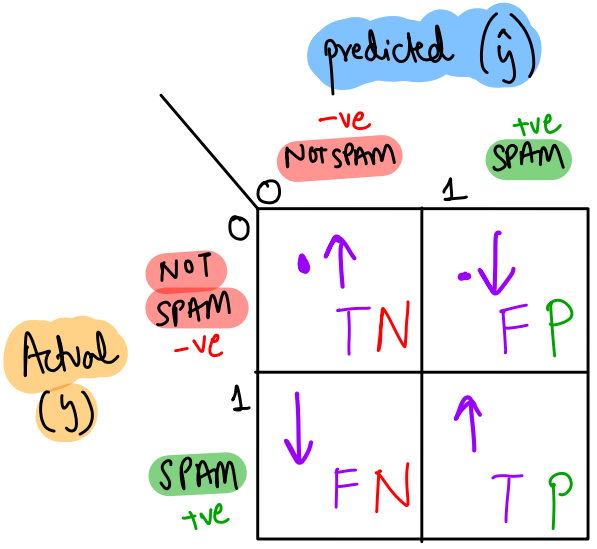
Quiz Ended!

if model classifies students into classes A,B and C , then Confusion matrix looks like ?

34 users have participated



400 → 360 NS → 360 N
→ 40 S → 40



Quiz time!

Quiz Ended!

For Ideal Model, which of the following is true?

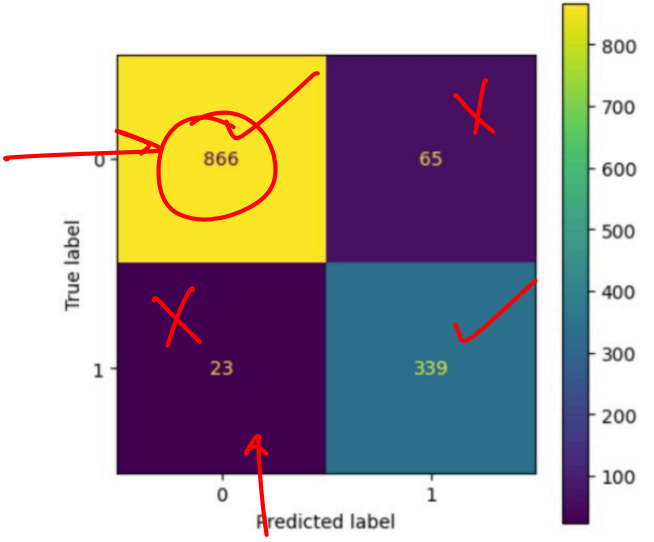
35 users have participated

A	FP and FN ⇓, while TP and TN ⇑	60%
B	TP and TN ⇓, while FP and FN ⇑	20%
C	TP and FN ⇓, while FP and TN ⇑	11%
D	FP and TN ⇓, while TP and FN ⇑	9%

Quiz time!

Quiz Ended!

Based on the Confusion matrix we saw, what is the total number of erroneous points?



65 + 23

= 88

24 users have participated

	A	31	0%
✓	B	88	79%
	C	106	8%
	D	1187	13%

=> 2 Scenario

Spam +ve Non Spam -ve

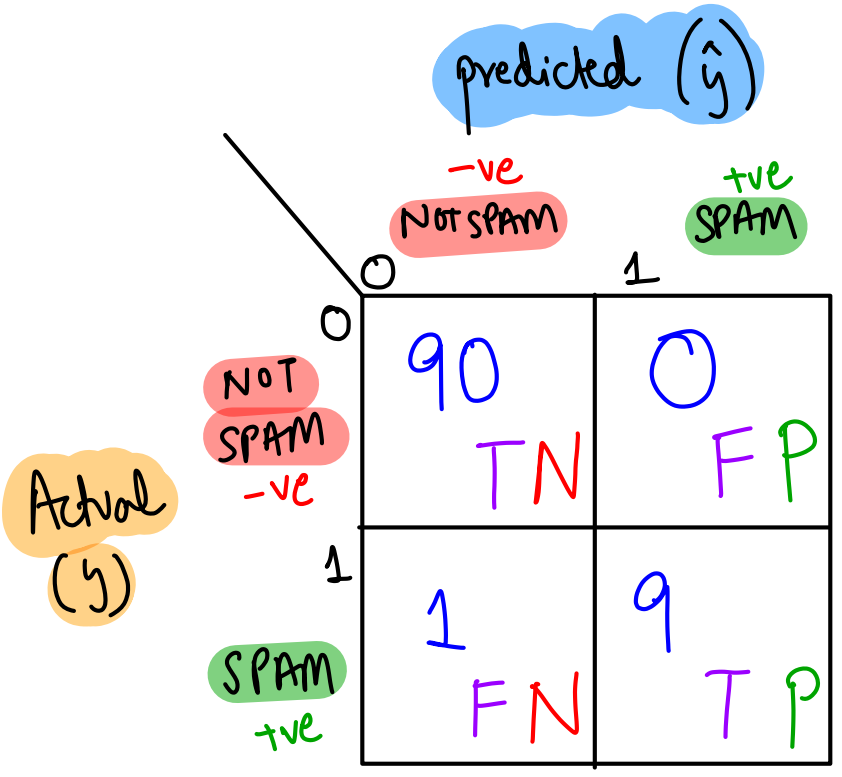
(1) Recieving spam email in your inbox → FN

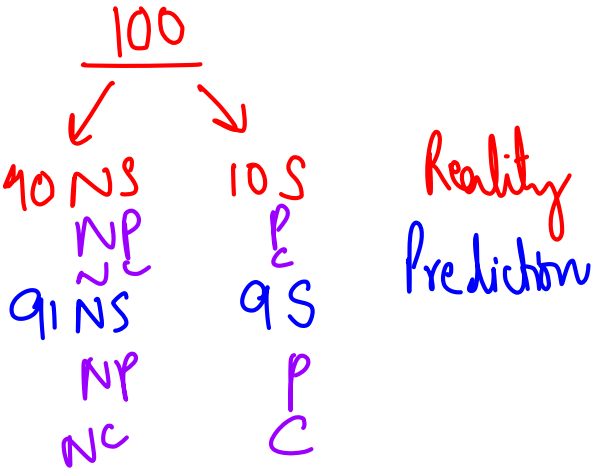
(2) Recieving appraisal letter in spam folder. → FP

Conclusion FP is more dangerous

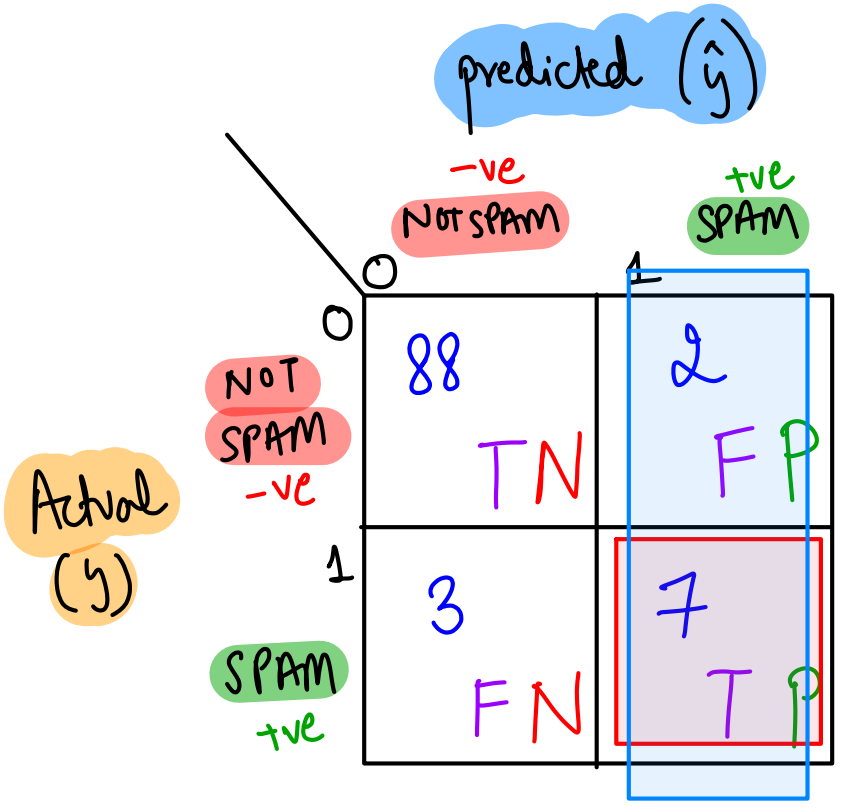
→ Minimise FP

Precision : $\frac{\#\text{ of actual spam emails}}{\text{Out of all predictions as spam}}$





$$\text{Precision} = \frac{TP}{TP + FP} = \frac{9}{9} = 100\%$$



FORTIS → 1000 →

$$\text{precision} = \frac{7}{9}$$

Accuracy → All classes

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

2nd Scenario

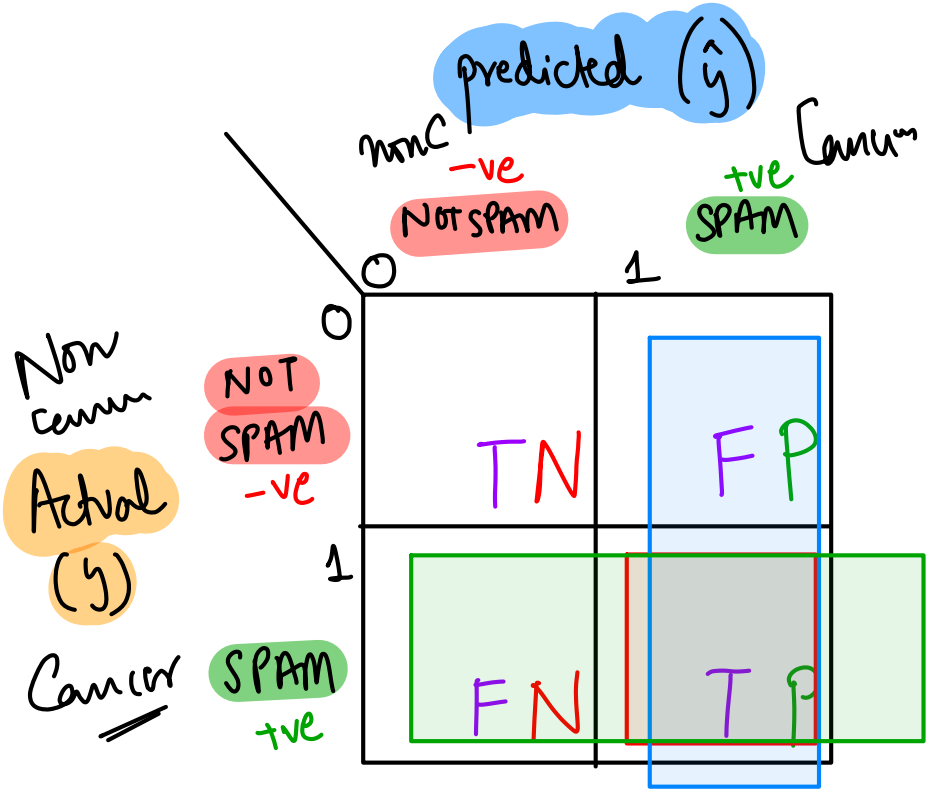
Cancer +ve NC -ve

(1) healthy person is identified as Cancerous FP

(2) Cancerous patient identified as healthy. FN

Minimise FN

$$\Rightarrow \frac{\text{How many of them were identified by my hospital.}}{\text{Out of all Cancerous patients that arrive at hospital}} = \text{Recall}$$



Recall: How many were predicted as cancer / All Cancerous Patient Reality

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

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

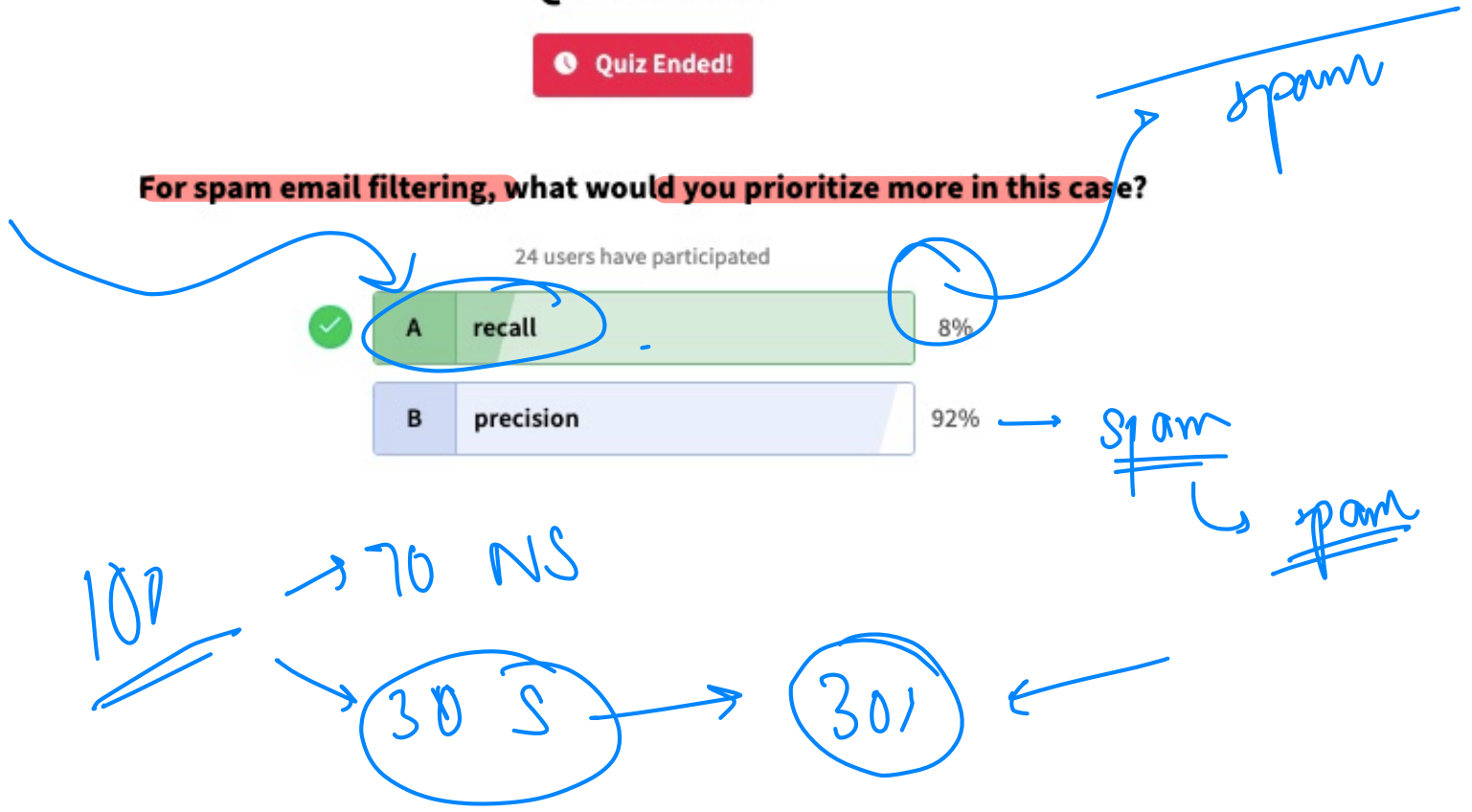
Quiz time!

Quiz Ended!

For spam email filtering, what would you prioritize more in this case?

24 users have participated

A	recall	8%
B	precision	92%



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

# of spam I predicted

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

I predicted spam

All actual spam emails

3 models

	Precision	Recall	Acc	
M1	0.3	0.8	0.55	0.24
M2	0.2	0.9	0.51	0.18
M3	0.7	0.4	0.55	0.28

Harmonic Mean

$$HM(P,R) = \frac{1}{\frac{1}{2}\left(\frac{1}{P} + \frac{1}{R}\right)} \Rightarrow \frac{2PR}{P+R}$$

F1 Score

0.2×0.9 0.18

0.4×0.6 0.24

0.5×0.5 0.25

0.9×0.9 0.81

Quiz time!

Quiz Ended!

Why does the F-1 score use Harmonic Mean (HM) instead of Arithmetic Mean (AM) ?

22 users have participated

	A	AM penalizes models the most when even Precision and Recall are low.	14%
✓	B	HM penalizes models the most when even Precision and Recall are low.	64%
	C	HM penalizes models the most when even Precision and Recall are high.	23%
	D	AM penalizes models the most when even Precision and Recall are high.	0%

Worst model = ρ & $R \Rightarrow 0$

$$FI = \frac{2\rho R}{\rho + R} = \frac{2 \times 0 \times 0}{0 + 0} = \frac{0}{0}$$

$$= \frac{2 \times 0 \times 0}{0 + 0 + 10^{-6}} =$$

BEST P & R = 1

$$F1 = \frac{2 \times 1 \times 1}{1+1} = 1$$

Range of F1 = [0, 1]