Logistic Regression.

→ Classification Algorithm.

→ Target variable → $\boxed{0, 1}$

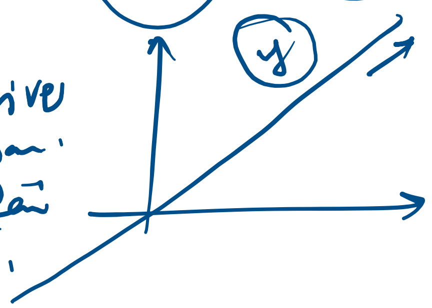
Give loan, don't give loan.

Rain, Not Rain

True, False.

$-\infty$ to ∞

y



Binary

dichotomous classification

Multiclass

Cu.. S..	Ag	M	Sal	Cu..	loan

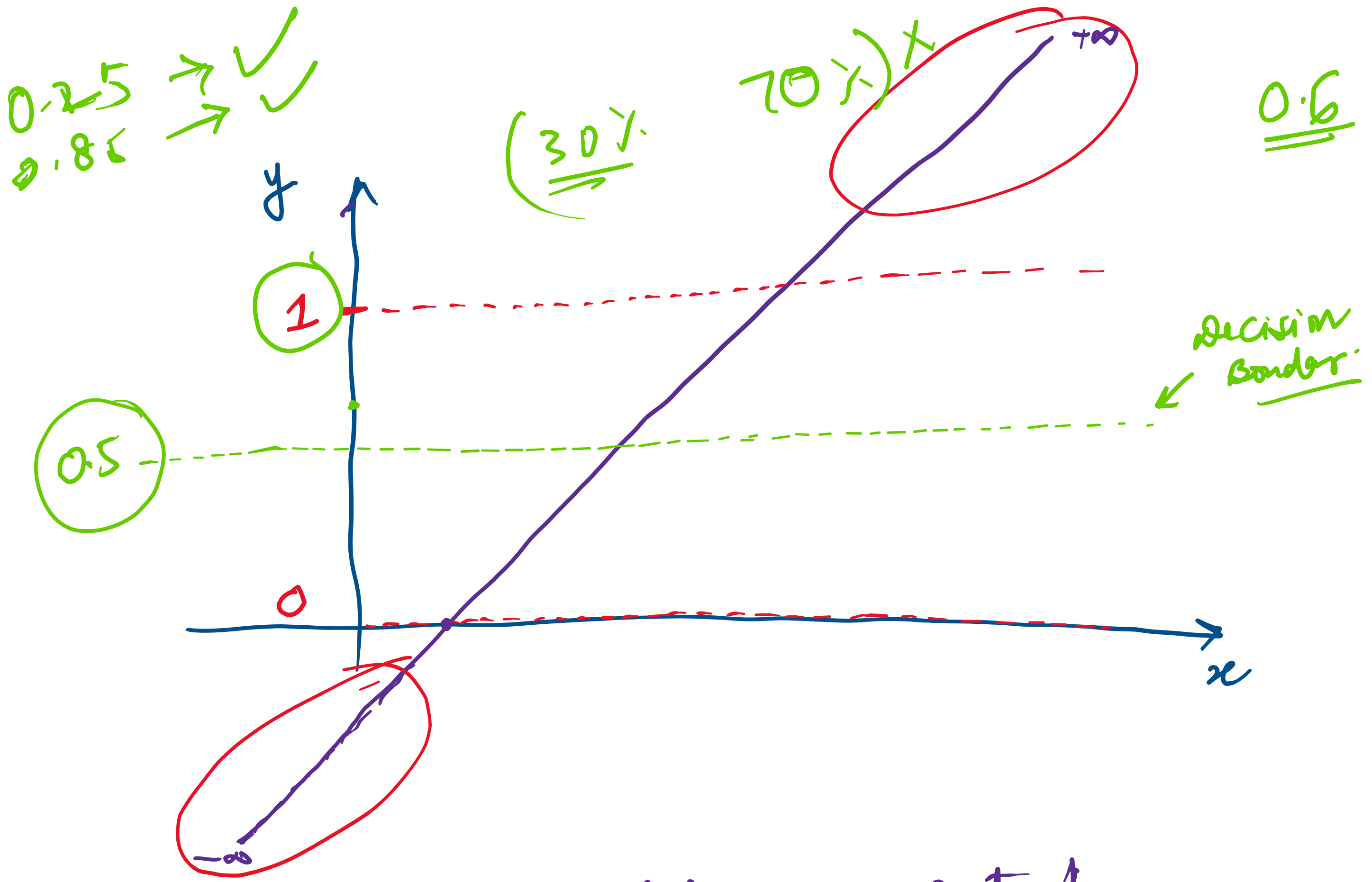
1

0

Decision:

↳ 0 : not happen / not survived

↳ 1 : happen / survived



Range of y for L.R. ⇒ 0 to 1

filtering ⇒ $y(-\infty, \infty) \rightarrow \boxed{f(x)} \rightarrow y(0,1)$

Logit

Sigmoid function

$$\boxed{P = \frac{1}{1+e^{-y}}}$$

$$p(1+e^{-y}) = 1$$

$$p + pe^{-y} = 1$$

$$pe^{-y} = 1-p$$

$$\boxed{e^{-y} = \frac{1-p}{p}}$$

$$\log_e e^{-y} = \log_e\left(\frac{1-p}{p}\right)$$

$$-y = \log_e\left(\frac{1-p}{p}\right)$$

$$-y = \ln\left(\frac{1-p}{p}\right)$$

$$\boxed{y = \underbrace{\ln\left(\frac{p}{1-p}\right)}_{\text{log odd}}}$$

$b_0 + b_1 x$ ← y

$\left(\frac{p}{1-p}\right)$ → Odd's Ratio

⇓

$$\frac{\text{Prob. of Success}}{\text{〃 of Failure}}$$

log odd

||

Logistic f^n

Ref:–

$\log_{\square} 10$

$\log_{10}$

$\log_e$

$\log_2$

$\log_{10} 10^2$

⇓

2

$\log_e = \ln$

sigmoid $f^n = \dfrac{1}{1+e^{-y}}$
p

$y = -\infty$

$$\frac{1}{1+e^{-(-\infty)}}$$

$$= \frac{1}{1+e^{\infty}}$$

$$= \frac{1}{1+(2.71)^{\infty}}$$

$$= \frac{1}{1+\infty}$$

$$= \frac{1}{\infty}$$

$$= 0.$$

$y = +\infty$

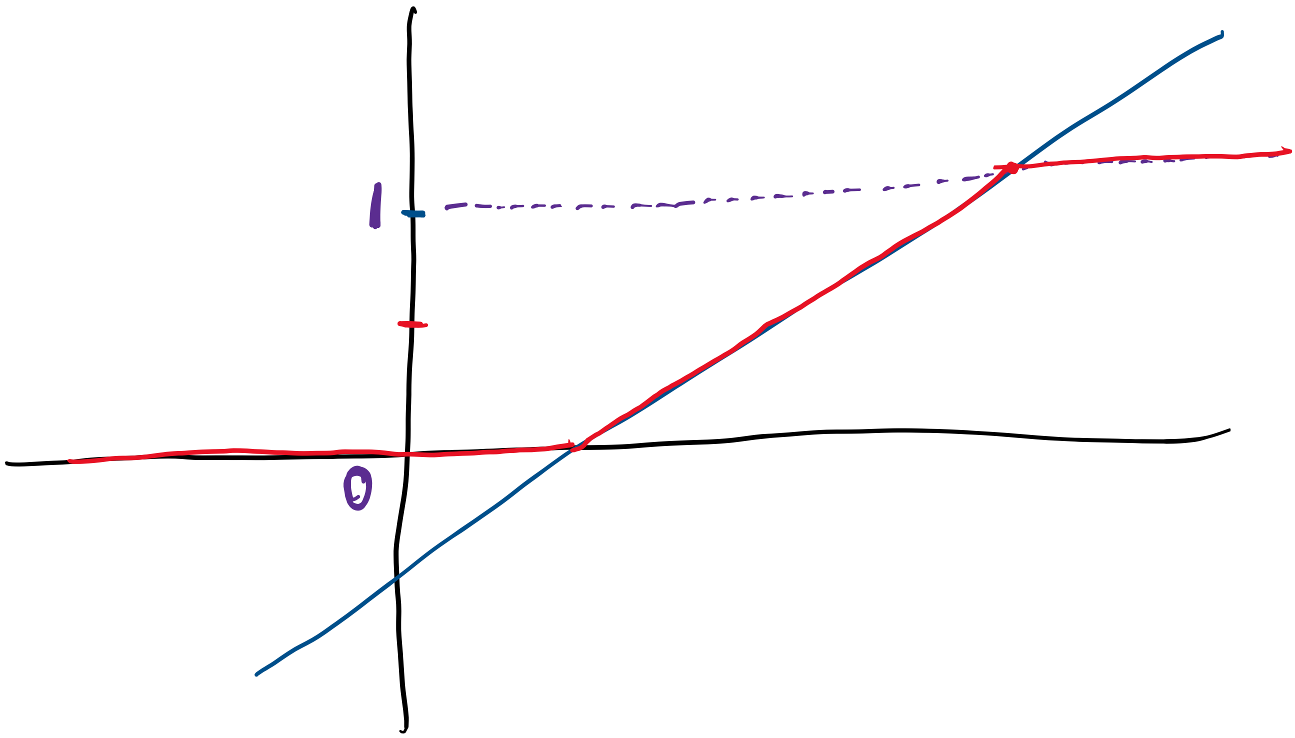
$$\frac{1}{1+e^{-\infty}}$$

$$= \frac{1}{1+\frac{1}{e^{\infty}}}$$

$$= \frac{1}{1+0}$$

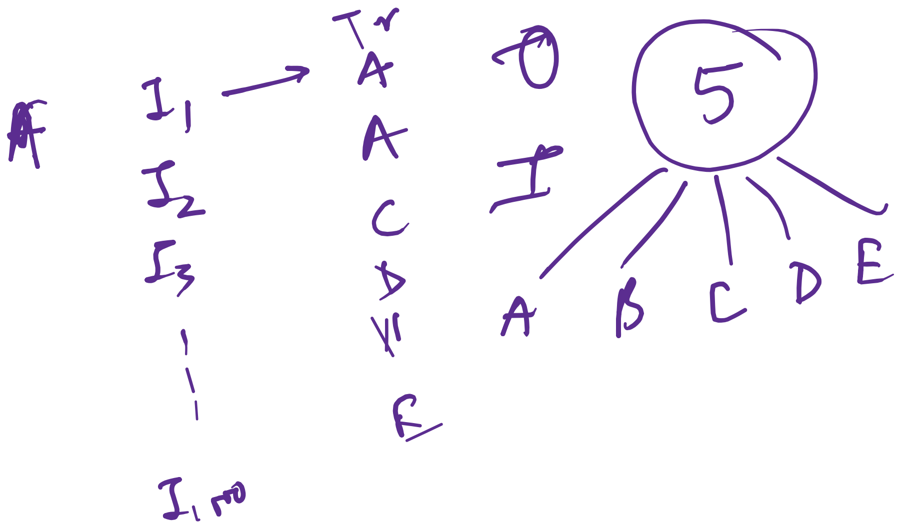
$$= \frac{1}{1}$$

$$= 1$$



① sigmoid f^n → log odd → logit f^n

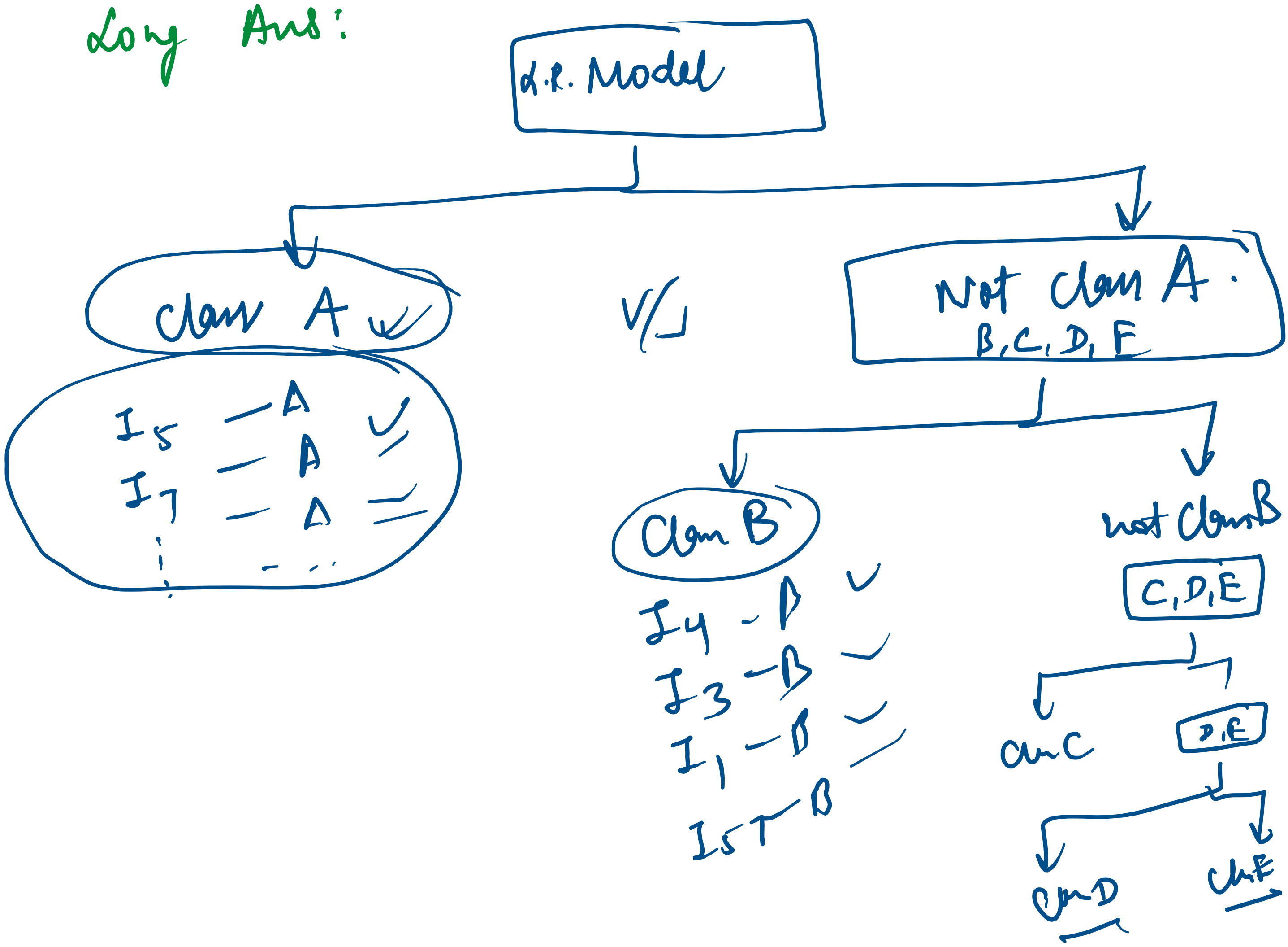
② Proved how sigmoid f^n changes the range of $y(-\infty, \infty)$ to $y(0, 1)$.



Q. Log. Reg. is doing dichotomous/binary classification. Discuss how can it do the multiclass classification.

— Short Ans: "one-v/s-All" approach.

Long Ans:

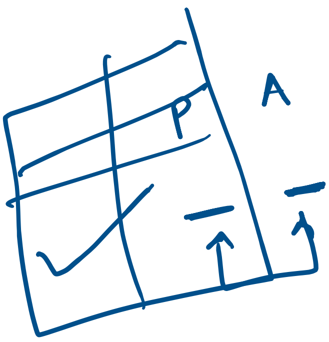


Evaluation Matrix for classification

⊛ Confusion Matrix

1. TP : (18) Predicting +ve Cases Correctly
2. TN : (12) Pred. -ve cases correctly Pred.
3. FP : (3) Predicted to have covid when actually they didn't have.
4. FN : (2) predicted to not having covid when actually they had.

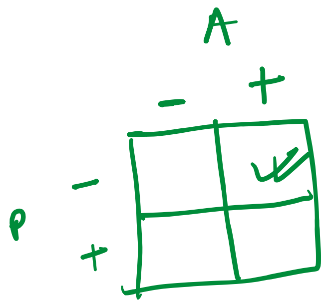
	Actual +	Actual −
+	18 ✓	3 ✗
−	2 ✗	12 ✓



TN

T → about my prediction result

N → about what I predicted.



FN

F → Predi~ False

N → P = N
A = P

Practice

① FN

F → My Pred. is false (incorrect)

N → −

② FP

		P: COVID −	COVID +
A	COVID +	FN (A)	X (B) ✓
	COVID −	(C)	FP (D) ✓

TN

TP ? where Actual = Predicted

		P: −	+
A	−	10 TN ✓	FP 2
	+	5 FN	8 TP ✓

what if

10 → Person is Tested −ve, when had been no Corona.

8 → " " " +ve " had corona.

not Corona

Died

−ve

+ve

find medi

Side eff

2

C

I

	FP or	FN
C-1	✓	X
C-II	X	✓

TP	TN

30

★ Accuracy $= \dfrac{\text{Total Correct Pred.}}{\text{Total no. of Pred.}} = \dfrac{TN + TP}{\text{Total Pred.}}$

$= \dfrac{12+18}{35}$

$= \dfrac{30}{35} =$

★ **Misclassification:**

(1 − Accuracy)

$= \dfrac{\text{Incorrect Pred}}{\text{Total Pred}} = \dfrac{FP + FN}{35} = \dfrac{2+3}{35}$

$= \dfrac{5}{35}$

* **True Positive Rate (TPR)**

How often its predicts +ve

$$\frac{TP}{\text{Total Actual +ve}} = \frac{18}{20} \quad \checkmark$$

* **False Positive rate (FPR)**

$$\frac{FP}{\text{Total Actual -ve}} = \frac{3}{15}$$

* **True Negative rate (TNR)**

$$\frac{TN}{\text{Total Actual -ve}} = \frac{12}{15}$$

Precision : when it Predicts yes, how often is it correct.

$$\frac{TP}{(\text{Total Prediction as +ve})} = \frac{18}{21}$$

$$T\boxed{P} + F\boxed{P}$$

Pred. Accuracy

Recall : It tells us abt how many actual +ve cases we were able to predict.

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

(precision ⇓ Total Pred. + , Recall ⇓ Total Actual +)

Precision

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

* use Precision when wrong results could lead to loss of business.

* Prec. is more useful when FP is a higher concern than FN.

* email : spam :-

FP :
Actual -ve → Predicted spam(+)
⇓
FP

Recall

$$\uparrow \frac{TP}{TP + FN}$$

* It doesn't matter much if we detect false, But Actual +ve cases should not go undetected

* Recall is more useful when FN is of a higher concern than FP.

* Fraud Transaction

AP —— Pred -ve
⇓
FN

F1-Score:

~ is a harmonic mean of Precision & Recall.

$$f1\text{-}score = \frac{2}{\frac{1}{(preci.)} \uparrow + \frac{1}{(Recall)} \uparrow}$$

→ It gives combined approximation of both Prec. & Recall.

Drawback:

f1 score's Interpretability,

---

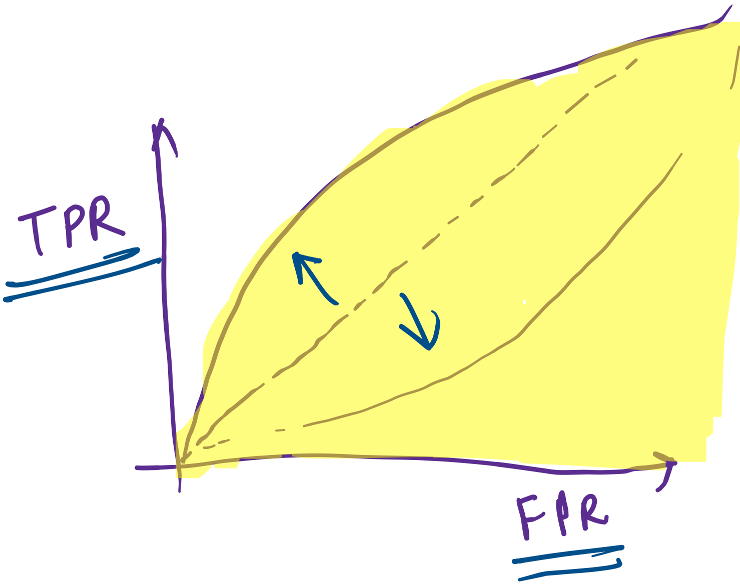
1. Confusion Matrix (TP, TN, FP, FN)
2. Accuracy
3. Mis-classification.
4. Precision
5. Recall
6. F1-Score ✓
7. ROC curve.

⇓

(Receiver operating characteristics Curves).

ROC- Curve

* TPR ↑ Area ↑



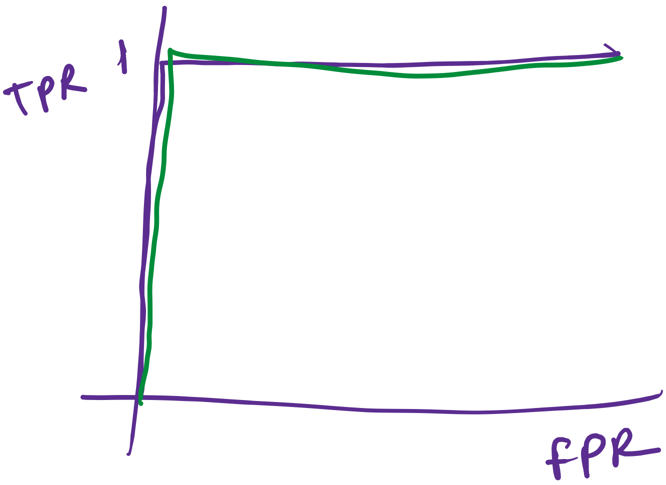
* ROC: TPR v/s FPR
  y x
* AUC: Area

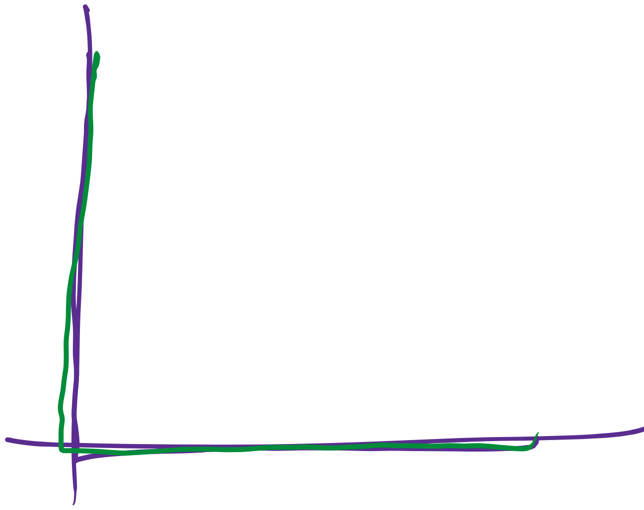
→ ROC is a Prob. Curve.
→ what it tells: model classification capability.

AUC ↑ Model is better at 0 as 0, 1 as 1

Q. what do you mean by AUC = 0.7 or 70%.

– If AUC is 0.7 ⇒ 70% chance that the model would be able to distinguish betⁿ +ve and -ve cla..



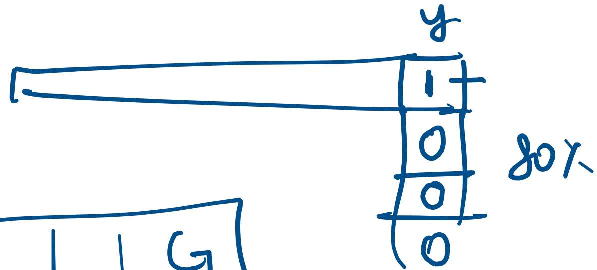


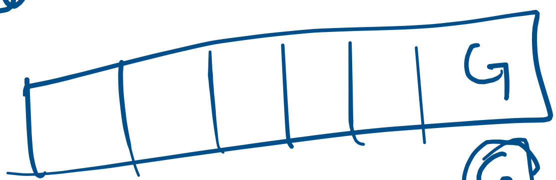
Majority - more in number :

Minority - less in number.

2 → 1
8 → 0

100%
↓
0.1% → Fraudulent = F
99.9% → genuine = G





G X

EOF>>