Cost function : Logloss function

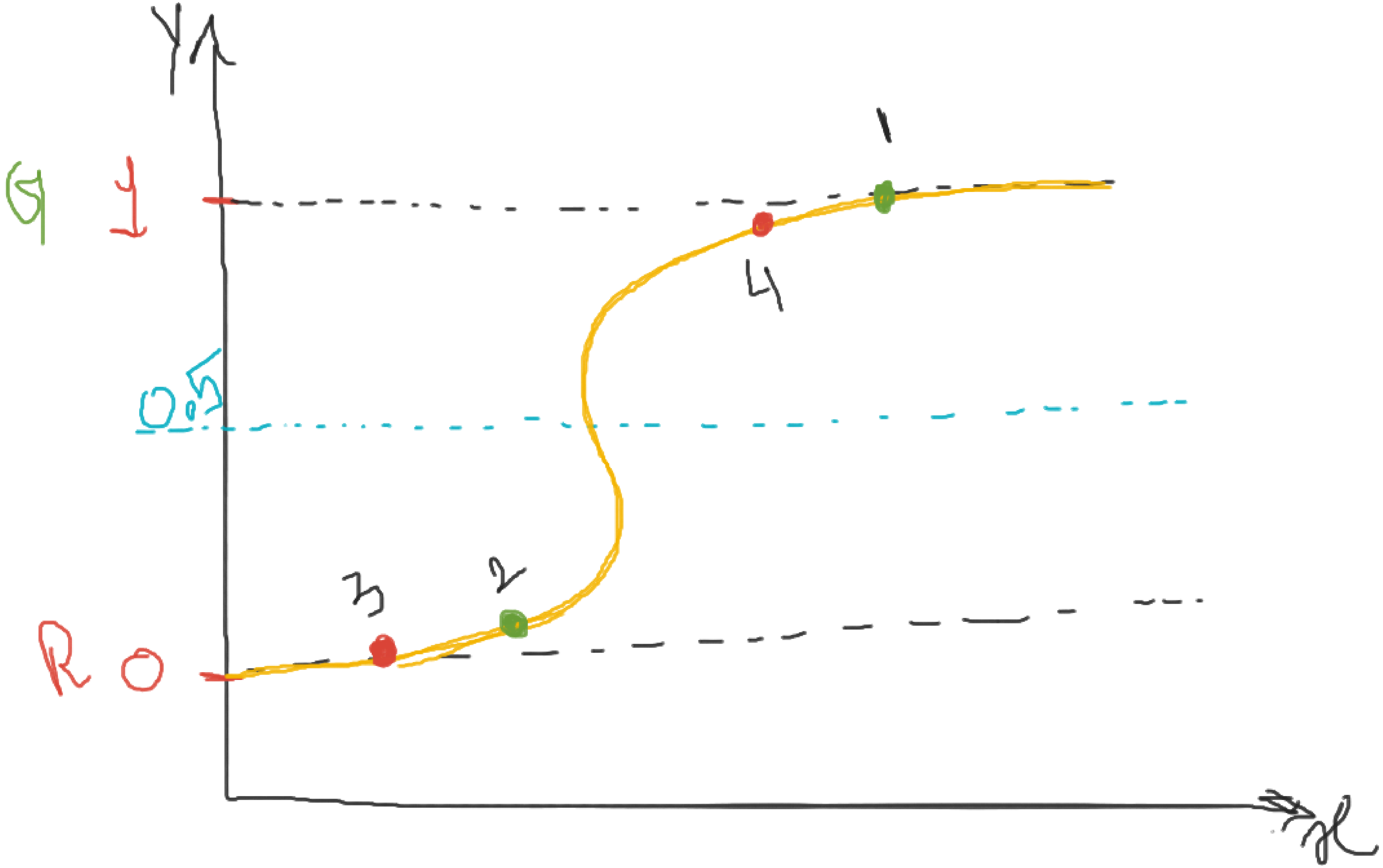
$$LL = -\frac{1}{N}\sum_{i=1}^{n} \underbrace{Y_i \cdot \log P_i}_{\mathrm{I}} + \underbrace{(1-Y_i)\log(1-P_i)}_{\mathrm{II}}$$

0 to 1

$$LL = -\frac{1}{N}\sum_{i=1}^{n} y_i \log p_i + (1-y_i)\log(1-p_i)$$

0/1

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



Y
G 1
0.5
R 0
1
4
3
2
x

i] Logloss $\rightarrow$ Correctly Classified data points

= Low

ii] Log loss $\rightarrow$ Misclassified data points

= High

$$LL = -\frac{1}{N}\sum_{i=1}^{n} y_i \log P_i + (1-y_i)\cdot\log(1-P_i)$$

i] Case I :- $Y_i = 1$, $P_i = 0.9$ → Correctly classified → Low

ii] Case II :- $Y_i = 1$ $P_i = 0.2$ → Miss. → High

iii] Case III :- $Y_i = 0$ $P_i = 0.1$ → correctly classifi.. → Low

iv] Case IV :- $Y_i = 0$ $P_i = 0.8$ → Miss. → High case :- $y_a = 1$ & $p_i = 0.9$

$$= -\frac{1}{N}\sum_{i=1}^{n} y_i \log p_i + (1-y_i)\log(1-p_i)$$

$$= -\frac{1}{N}\left(1 \times \log(0.9)\right) + 0$$

$$= -\log(0.9)$$

$$=$$

Case II : $Y_i = 1$ & $p_i = 0.2$

$$= -\frac{1}{N}\sum_{i=1}^{n} y_i \cdot \log p_i + (1-y_i) \cdot \log(1-p_i)$$

$$= -\left(1 \times \log(0.2)\right) + 0$$

$$= -\log(0.2)$$

$=$

Case iii :- $y_i = 0$ & $p_i = 0.1$

$$= -\frac{1}{N}\sum_{i=1}^{n} y_i \cdot \log p_i + (1-y_i) \cdot \log(1-p_i)$$

$$= -\left[(0) + (1) \cdot \log(1-0.1)\right]$$

$$= -\left[\log(0.9)\right]$$

$$= -\log(0.9)$$

$$=$$

Case IV :- $Y_i = 0$ & $P_i = 0.8$

$$= \frac{-1}{N} \sum_{i=1}^{n} Y_i \cdot \log P_i + (1 - Y_i) \cdot \log(1 - P_i)$$

$$= -\left[(0) + (1) \cdot \log(1 - 0.8) \right]$$

$$= -\log(0.2)$$

$$=$$

Confusion Matrix :-

		Actual	
		1 P	0 N
Predicted	1 P	Tp	FP
	0 N	FN	TN