Day 8

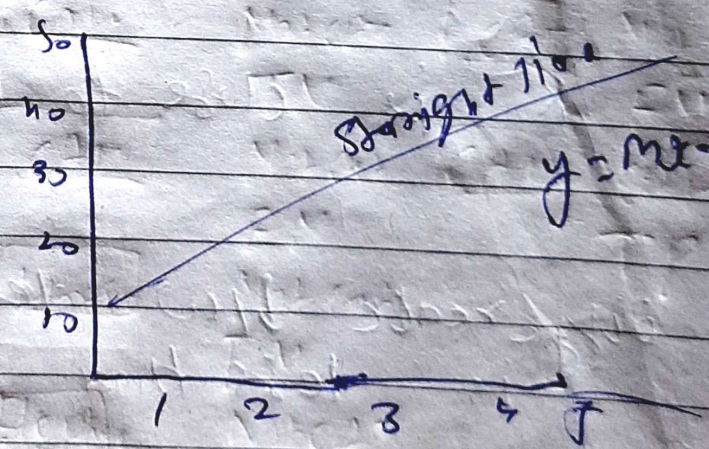
ML with Python

Logistic Regression

Linear	Logistic
→ y is continuous (numerical)	y is categorical
→ y (-∞ to ∞)	y (0 to 1)
→ Line Straight	Sigmoid curve

No. hours	marks
1	20
2	30
3	40
4	50
5	60

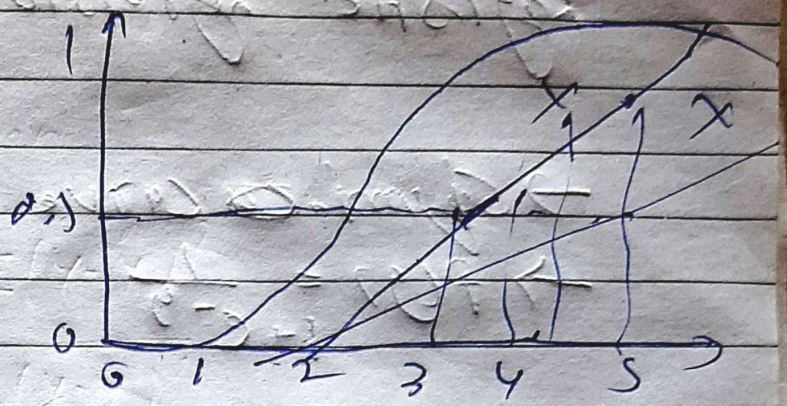
Threshold



किफायती
Date
Pages
Rituraj

No.	Hours	Result
1	20	F
2	30	F
3	50	P
4	70	P
5	80	P

misclassification.



eg. of Sigmoid curve.

$$f(y) = \frac{1}{1+e^{-y}}$$

$$y = mx + c$$

$-\infty$ to ∞

when $y = -\infty \Rightarrow \frac{1}{1+\infty}$

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

$y = \infty$

$\frac{1}{1+e^{-\infty}} \Rightarrow$ 0 to 1

$\frac{1}{1+e^{-2}} \Rightarrow -\log$ Neg. $\Rightarrow$ 0 to 1

Date
Pages
Rituraj

Logistic Regression ⇒ Classification

→ Sigmoid curve

→ $P(y) = \frac{1}{1+e^{-y}}$ $(y \to -\infty \text{ to } \infty)$

$P(y) \Rightarrow 0$ to 1

④ Titanic
- Survied → 0 → not survied
- 1 → survied

⑤ Movie ⇒ 0 (not cancelled), 1 (cancelled)

~~X~~ Metrics classification

Actual	Predicted					
CAT	CAT TP	COW	Not CAT TN	CAT	Not CAT FN	
CAT	Not CAT FN	CAT	Not CAT FN	CAT	CAT TP	
CAT	CAT TP	Horse	Not CAT TN			
COW	CAT FP	Elephant	Not CAT TN			

Date
Pages
Rituraj

TP = 3
TN = 3
FP = 1
FN = 3
} 10

Confusion matrix

Predict \ Actual	0	1
1	FP	TP
0	TN	FN

Predict \ Actual	0	1
1	FP	TP
0	TN	FN

1 - Positive
0 - Negative

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

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

$$\text{precision} = \frac{TP}{PP} = \frac{TP}{TP + FP}$$