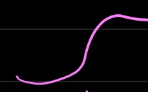
Till now

→ Logistic regression
→ L_1 and L_2 regularisation.
→ indepth explanation of logistic Regression.

→* Why Logistic Regress? ??

Sigmoid
⇓

⇓
$$\frac{1}{1+e^{-z}}$$

Agenda

→ Logistic Regression implementation
→ Regularisation in logistic Regression
→ Multiclass classification
→ " " implementation

Probability → 0 to 1

$$h_\theta(x) = \frac{1}{1+e^{-z}}$$

⇓

(0.5)

x_1	x_2	x_3	y_{data}	y_{prob}	y_{pred}
—	—	—	1	{0.8, 0.2}	1
—	—	—	0	{0.6, 0.4}	0
—	—	—	1	{0.3, 0.7}	1
—	—	—	0		0
—		—	1		:

$(x_1 x_2 x_3) \approx 1$
≈ 0

$$y_p = \theta_0 + \theta_1 x$$

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

$$\Rightarrow \frac{1}{1+e^{-(\theta_0+\theta_1 x_1)}} = P$$

$$\Rightarrow p \times \left(1 + e^{-(\theta_0 + \theta_1 x_1)}\right) = 1$$

$$\boxed{5^{-1} = \frac{1}{5}}$$

$$\Rightarrow p \times \left(1 + \frac{1}{e^{(\theta_0 + \theta_1 x_1)}}\right) = 1$$

$$\Rightarrow p \times \left(\frac{e^{\theta_0 + \theta_1 x_1} + 1}{e^{(\theta_0 + \theta_1 x_1)}}\right) = 1$$

$$\Rightarrow p \cdot e^{\theta_0 + \theta_1 x_1} + p = e^{\theta_0 + \theta_1 x_1}$$

$$\Rightarrow p \cdot e^{\theta_0 + \theta_1 x_1} - e^{\theta_0 + \theta_1 x_1} = -p$$

$$\Rightarrow e^{\theta_0 + \theta_1 x_1}(p - 1) = -p$$

$$(3-5) \;\Big|\; -(5-3)$$
$$-2 \;\Big|\; -2$$

$$\Rightarrow \cancel{-}\, e^{\theta_0 + \theta_1 x_1}(1 - p) = +p$$

$$e^{\theta_0 + \theta_1 x} = \frac{p}{1-p}$$

taking log on both sides

$$\boxed{\theta_0 + \theta_1 x = \log\left(\frac{p}{1-p}\right)}$$

$$\Downarrow$$

$$y = \theta_0 + \theta_1 x$$

Sigmoid fn + Linear reg. = Logistic reg.

$$\frac{p}{1-p} \Rightarrow \text{odds} \Rightarrow p(\text{success}) / P(\text{failure})$$

$$\log\left(\frac{p}{1-p}\right) \Rightarrow \log(\text{odds})$$

L.R

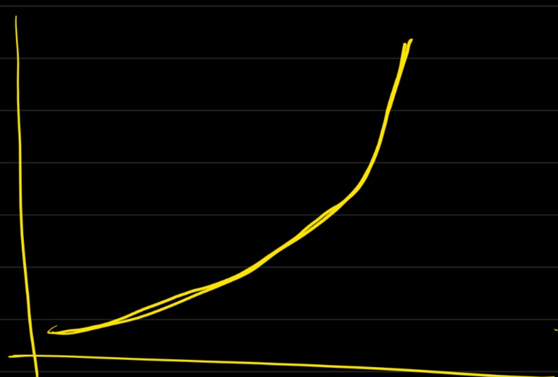
$$y = \theta_0 + \theta_1 x_1$$

Logistic Reg Coeffic $\boxed{\log \frac{p}{1-p} = \theta_0 + \theta_1 x}$

$\log \text{odds} = \theta_0 + \theta_1 x$

$$\boxed{\log \frac{p}{1-p} = \theta_0 + \theta_1 x}$$

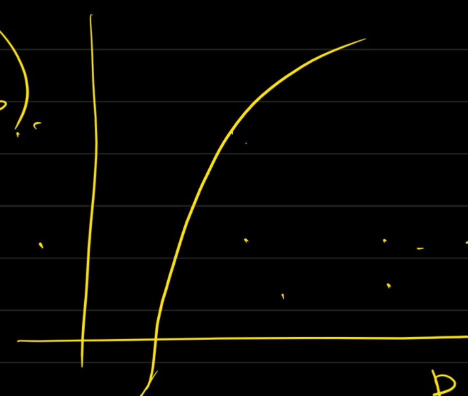
$\frac{p}{1-p}$



$p \rightarrow$

$p \uparrow \quad \frac{p}{1-p} \uparrow$

$\log\left(\frac{p}{1-p}\right)$



p

$p \uparrow \quad \log \text{odds} \uparrow$

$$\boxed{p \uparrow \quad \text{odds} \uparrow \quad \log(\text{odds}) \uparrow \quad y \uparrow}$$

montonoic relatioship.

One direction

Evaluation metrics for classification Problem

1. Confusion matrix
2. Accuracy | mis classification rate
3. Precision
4. Recall
5. F-Beta score.
6. True Positive rate (Sensitivity)
7. False Positive rate
8. True Negative rate (Specificity)
9. ROC - AUC
10. Precision-Recall / Sensitivity - specificity trade off.

X_1	X_2	y_{act}	y_{pred}	
—	—	0	1	→ wrong
		1	1	→ Correct
		0	0	→ correct
		1	0	— wrong
		0	1	— wrg
		1	0	— wrong.
		1	1	→ Correct.

		Actual value 1	Actual value 0
Predicted value	1	2	2
	0	2	1

1 → positive   True Positive.   Actual is 1 and Predicted is 1

0 → negative.   → or " "Positive" " is positive

if and condition is true ⇒ True

Pred \ Act	1	0
1	TP	FP
0	FN	TN.

TP - True Positive.

→ Actual is Neg and Predicted in Neg   True.

FP, FN ⇒ Think in terms of Actual what you have predicted wrong.

F.

② Accuracy ⇒ How many are correctly predicted from all the datapoint

		1	0	Actual
Predicted	1	TP	FP	
	0	FN	TN	

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

$$\frac{2+1}{2+2+2+1} = \frac{3}{7} = 0.4 \Rightarrow 40\%$$

* Misclassification rate = Opposite of Accuracy

$$\frac{FP+FN}{\text{total}} = 1 - \text{Accuracy} = 1 - \frac{3}{7} = \frac{4}{7}$$

0.6

③ Precision

⇓

{ Out of all predicted ones, how many are actual ones. }

imbalance class

1000 dP's → 900 class 1, 100 class 0 } imbalanced data

$$\text{Accuracy} \Rightarrow \frac{900+0}{1000} = 90\%$$

⇓

Accuracy doesn't give more priority to minority class.

		1	0	Act
Pred	1	900	100	
	0	0	0	

		1	0	Act
Pred	1	TP	FP	
	0	FN	TN	

→ Out of all actual value, how many are correctly predicted.

$$P = TP / TP+FP$$

Precision of +ve cases.

④ Recall = Out of all Actual one how many are correctly predicted one.

		1	0	Act
Pred	1	T.P	FP	
	0	FN	TN	

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

Precision ——— Positive class ——— Recall

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

Use case-1 Spam classification.

text → model → spam | Not spam

	1	0 Act.
Pred 1	TP	FP
0	FN	TN

Mail (goodmail Job) → Spam, inbox

TP ⇒ Mail – Spam(1), Model – Spam(1) } Correct case.

TN ⇒ Mail – Not spam(0), model – Not spam(0) } Correct case.

FP ⇒ Mail → Not a spam, model → Spam } → Blunder.

FN ⇒ Mail – Spam, Mode – Not a spam

wrong Predict

Precision

probability ≥0.5 → 1, <0.5 → 0

→ We want a model with high precision.

→ threshold | cutoff

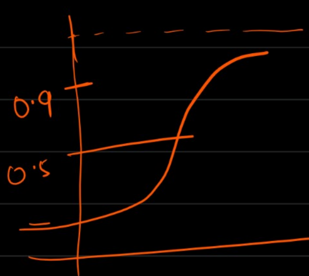
0.5

increase the cutoff

Logistic → 1

≥ 0.5 → 1

< 0.5 → 0



50% → spam

After increasing the cutoff

30% → spam

↓

minimise the case ⇒ FP

* Use case-2.

	1	0 Actual
Pred 1	TP	FP
0	FN	TN

TP = Actual diabetic, model – diabetic

TN = Act Not diabetic, mode " "

} Correct cases.

FP – Act – No diabetic
Model – Predicted diabetic

FN – Act diabetic
Model – No diabetic → Blunder

FN ←

why ??

Recall

US → AI model → Guilty / Not guilty

H/W

→ Conviction of a person in the court trial.

→ Stock market will crash or not.

⑤ F-beta Score

$$(1+\beta^2)\frac{P \times R}{P+R}$$

P – Precision
R – Recall

① if FP and FN both are important

$\beta = 1$, F_1 Score $= \boxed{2 \cdot \frac{P \times R}{P+R}}$

② if FP is more important than FN

$\beta = 0.5$

$$F_{0.5} \text{ Score} = (1+0.25)\frac{P \times R}{P+R}$$

③ if FN is more important than FP,

$\beta = 2$

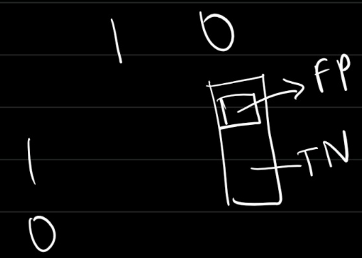
$$F_2 \text{ Score} = (1+4)\frac{P \times R}{P+R}$$

6. True Positive Rate ⇒ out of all actual 1, it is actually predicted one

↳ Sensitivity, Recall

$$\boxed{TPR = \frac{TP}{TP+FN}}$$

7. False Positive rate = $\frac{FP}{FP+TN}$



8. True Negative Rate $TNR = \frac{TN}{FP+TN}$

→ Specificity

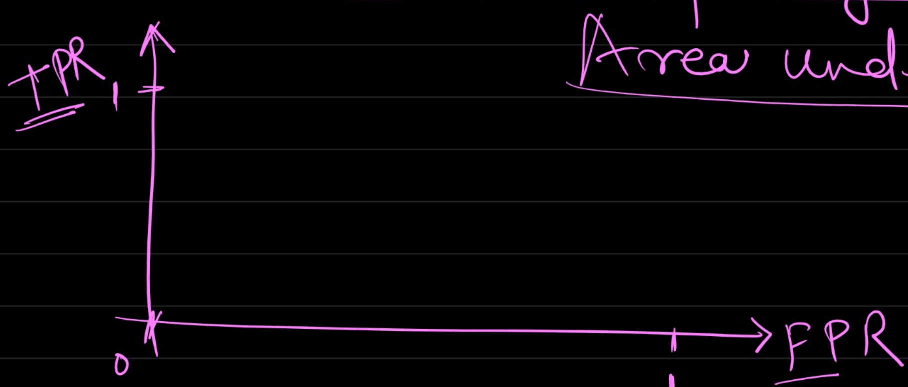
$$TNR = 1 - FPR$$

9. ROC AUC

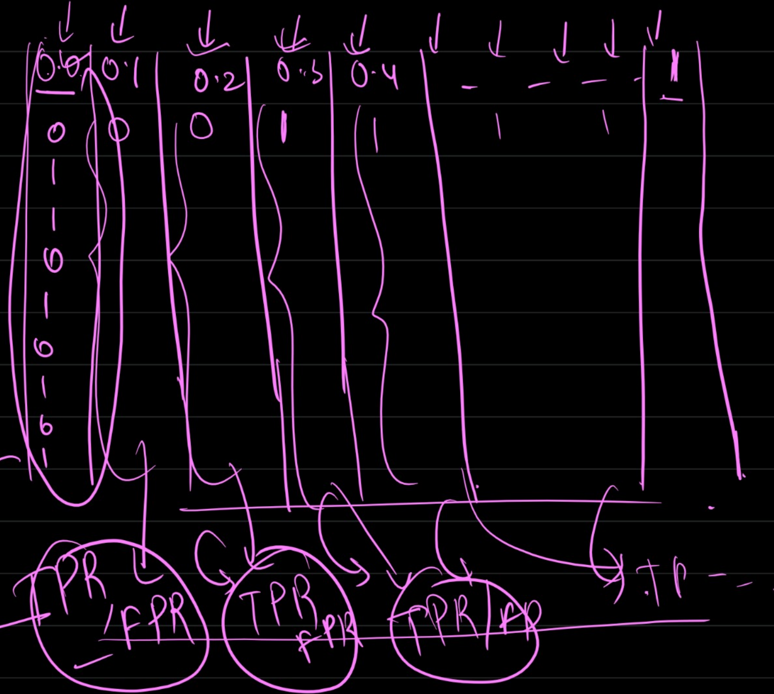
TPR Vs FPR

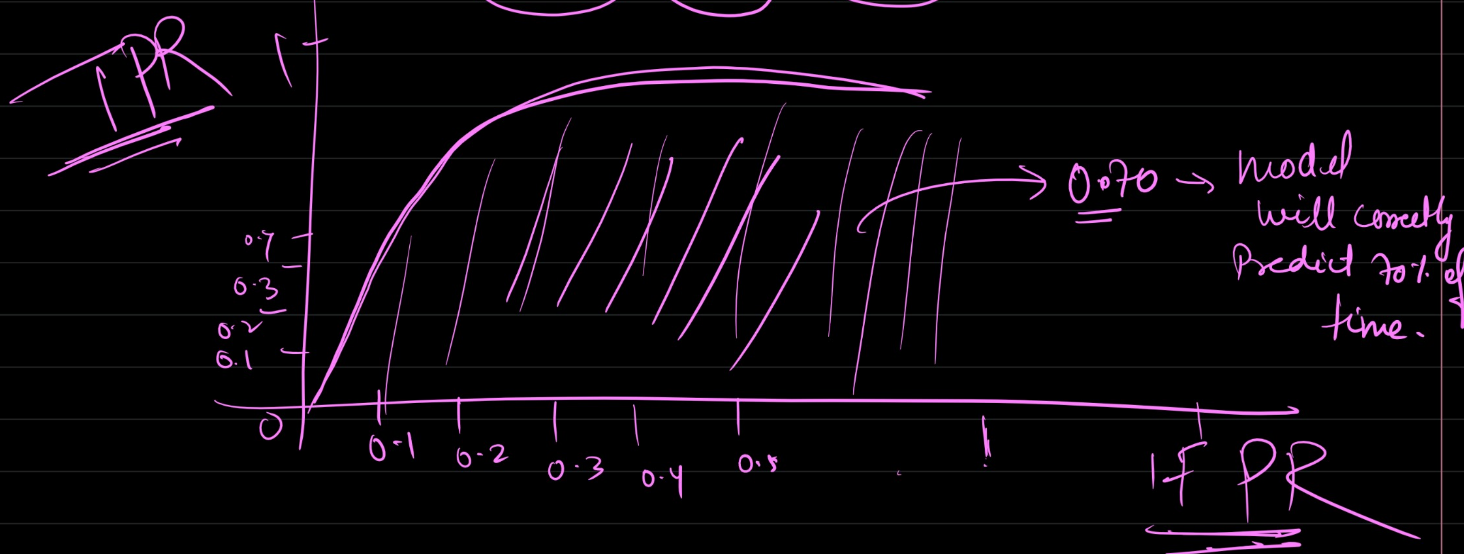
↳ Reciever operating Characteristic

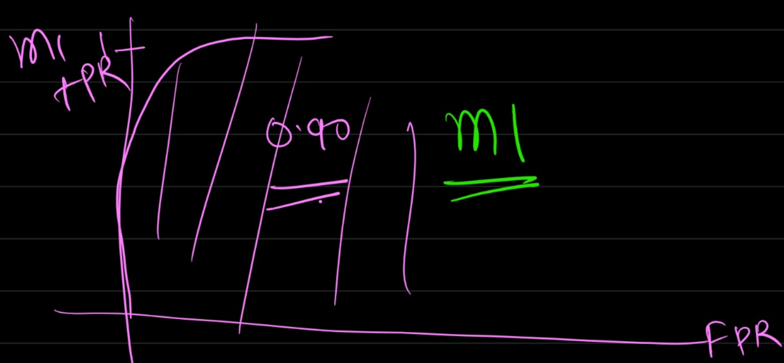
Area under Curve.



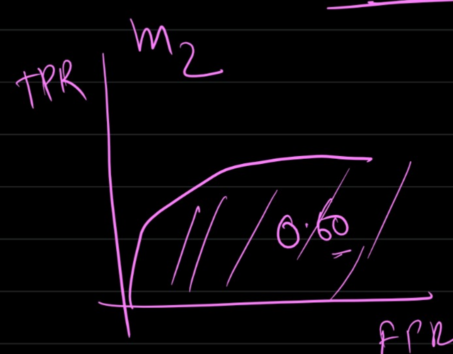
X	y_{act}	Prob	0.5 y_{pred}
5	1	0.9	1
6	0	0.8	1
7	1	0.6	1
8	0		0
4	1	1	1
1	0	1	0
1	1	1	1



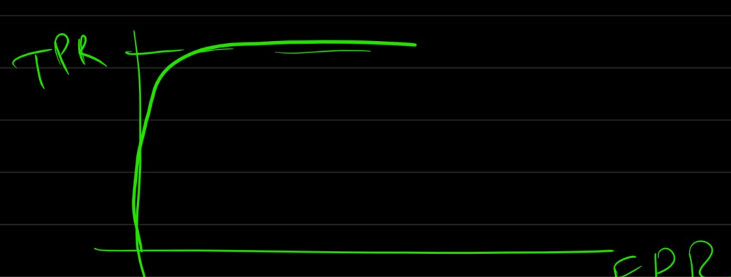




m1

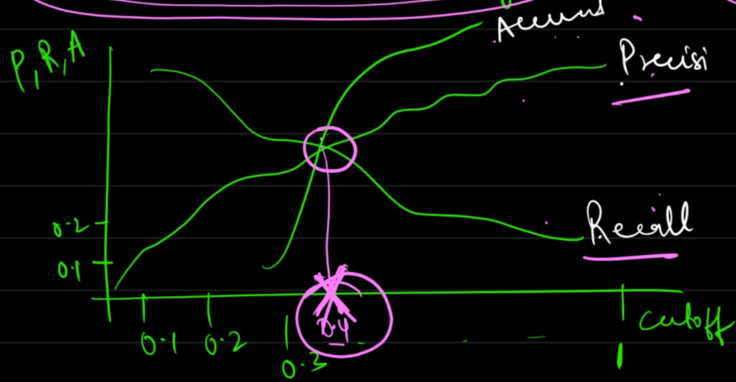


Higher the AUC better the model will will be.

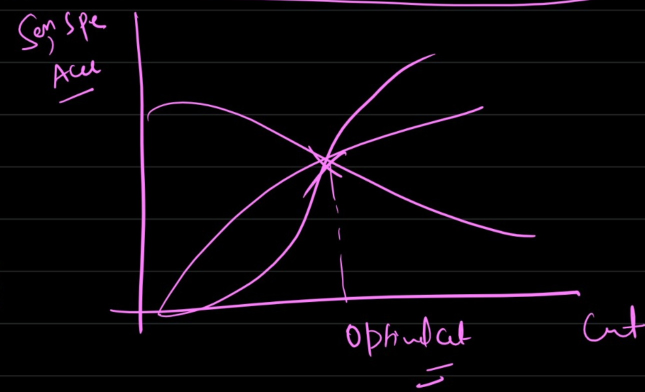


10. Cutoff → business team (by default 0.5)

Precision - Recall - Accuracy



Sensitivity - Specificity - Accuracy.



* ROC AUC implementation