↓
Multicollinearity

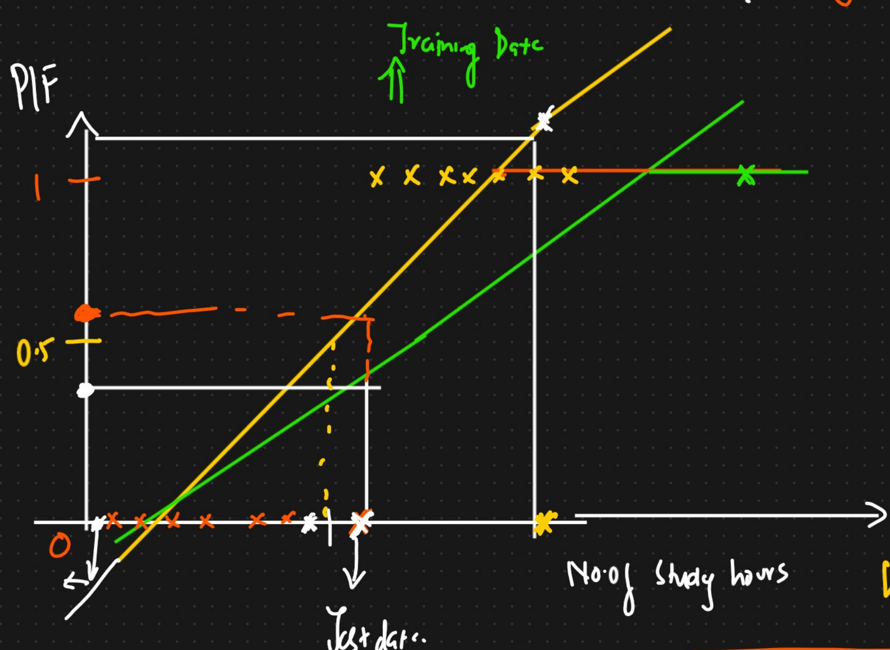
0.95%

-0.95

~~f_1~~ ✓ $\longleftrightarrow$ f_2 ✓ f_3 O/P

600 features $\Rightarrow$ Lasso

② Logistic Regression (Classification) $\rightarrow$ {Binary classification}



No of study	No of play	P/F
—	—	P
—	—	F

$h_\theta(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 + \dots + \theta_n x_n$

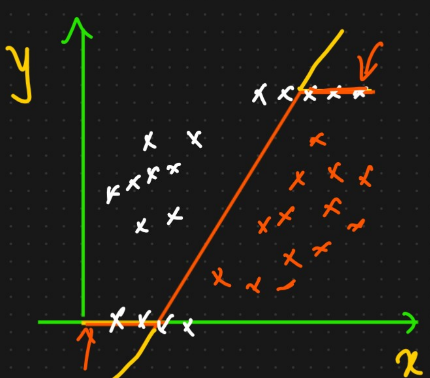
$$\boxed{h_\theta(x) = \theta_0 + \theta_1 x_1}$$

$\Downarrow$

$h_\theta(x) = g(\theta_0 + \theta_1 x_1)$

$\Downarrow$

Let $z = \theta_0 + \theta_1 x_1$



Sigmoid Activation fn
⇑

$h_\theta(x) = g(z)$ $\rightarrow$ function on z

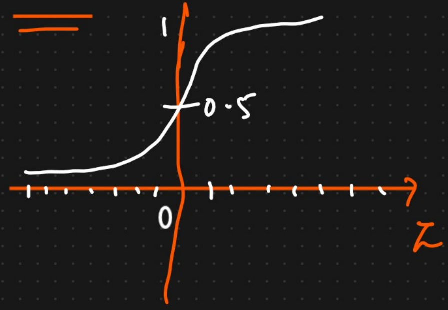
$\Downarrow$

{Activation function}

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

O{0 to 1}

exponential



$$\boxed{h_\theta(x) = \frac{1}{1+e^{-(\theta_0+\theta_1 x_1)}}}$$

$\Downarrow$

Best fit line + Squashing

$$\boxed{h_\theta(x) = \theta_0 + \theta_1 x_1}$$

Training Set

$$\{(x^1,y^1),(x^2,y^2),(x^3,y^3) \ldots (x^n,y^n)\}$$

$y = \{0,1\}$ → 2 Output → Binary classification

$$h_\theta(z) = \frac{1}{1+e^{-z}} \qquad z = \theta_0 + \theta_1 x_1 \qquad \theta_0 = 0 \qquad \text{intercept} = 0$$

Aim : Change θ_1 → It classifies points.

Cost function

$\hat{y}$ MSE

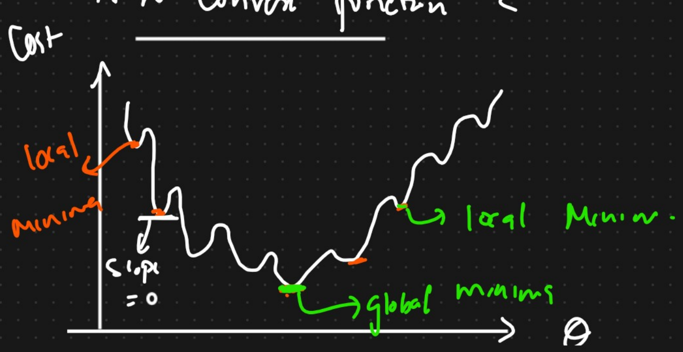
Logistic Regression $$J(\theta_1) = \frac{1}{2m}\sum_{i=1}^{m}\left(h_\theta(x)^i - y^{(i)}\right)^2 \qquad h_\theta(x) = \theta_1 x_1$$

$\Downarrow$

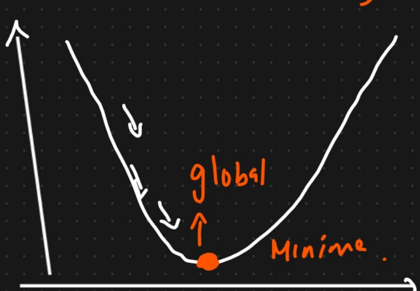
Gradient Descent

$$h_\theta(x) = \frac{1}{1+e^{-(\theta_1 x)}}$$

Non Convex function ←

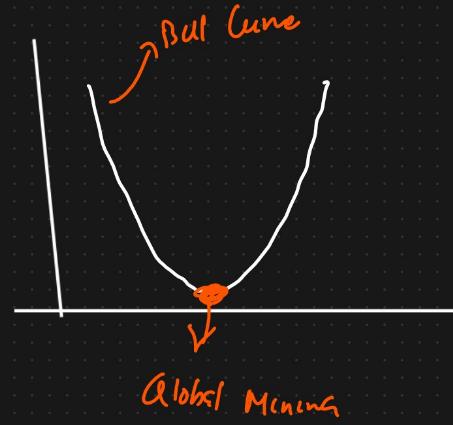


{ Convex function }



Logistic Regression Cost function (Log Loss)

$$J(\theta_1) = \begin{cases} \boxed{-\log(\underbrace{h_\theta(x)}_{\hat{y}})} & y=1 \\ \boxed{-\log(1-\underbrace{h_\theta(x)}_{\hat{y}})} & y=0 \end{cases}$$



$y=1$

Cost function

$$J(\theta_0,\theta_1) = -y\log(h_\theta(x^i)) - (1-y)\log(1-h_\theta(x))$$ {log loss cost function}

Cost function for Logistic Regression

$$J(\theta_1) = -\frac{1}{2m}\sum_{i=1}^{m}\left[\left(y^i\log(h_\theta(x^i)\right) + (1-y^{(i)})\log(1-h_\theta(x^i))\right]$$

$$\boxed{h_\theta(x^i) = \frac{1}{1+e^{-(\theta_1 x_1)}}} \Rightarrow \text{hypothesis.}$$

θ_1

repeat until Convergence

$$\left\{ \theta_j := \theta_j - \alpha\frac{\partial}{\partial\theta_j}(J(\theta_1)) \right\}$$

Performance Metrics (Binary Classification).

x_1	x_2	y (Act)	$\hat{y}$ (predi)
–	–	0	1
–	–	1	1
–	–	0	0
–	–	1	1
–	–	1	1
		0	1
		1	0

{Actual}

{Predicted}	1	0
1	3	2
0	1	1

⇒ Confusion matrix

Predicted \ Actual	1	0
1	↑TP	FP↓
0	FN↓	TN↑

TP & TN Are correct predictions.

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+FN+TN} = \frac{3+1}{3+2+1+1} = \frac{4}{7}$$

$= 57\%$

Training

1000 datapoint points

0 → 900 datapoint
1 → 100 datapoint } Imbalanced Dataset

0 → 600 datapoints
1 → 400 datapoints } Balanced

Not Spam

Predicted \ Actual	1	0
1	TP	FP
0	FN	TN

→ FP↓↓
→ FN↓

① Precision

$$= \boxed{\frac{TP}{TP+FP\downarrow\downarrow}}$$

② Recall (TPR).

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

③ F-Score

Spam classification {FP↓↓↓} → Precision

Has cancer or not {FN↓↓↓} → Recall

Company ⇒ FP

People ⇒ FN

Tomorrow Stock market is going to crash $\left\{ \begin{array}{l} \text{Both FP} \downarrow\downarrow \\ \text{\& FN} \downarrow\downarrow \end{array} \right\}$.

F-Beta Score: $(1+\beta^2) \dfrac{\text{Precision} * \text{Recall}}{\beta^2 * [\text{Precision} + \text{Recall}]}$

$\boxed{\text{FP \& FN}}$ ✓

$\boxed{\beta = 1}$ $= (1+1) \dfrac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$ } Harmonic Mean

$\beta \downarrow\downarrow$

$\boxed{\text{FP} >> \text{FN}}$

$\beta = 0.5$ $= (1+0.25) \dfrac{P \times R}{(0.25)[P+R]}$

$\boxed{\text{FN} >> \text{FP}}$

$\boxed{\beta = 2}$ $= \boxed{(1+(2)^2) \dfrac{PR}{(4)[P+R]}}$.