Loss function - LMS function.

Least Mean Square

$h(x) = \theta_0 + \theta_1 x_1$

↗ Predicted.

$Y \rightarrow$ Actual

$\text{Loss} = |h(x) - y|$ ×

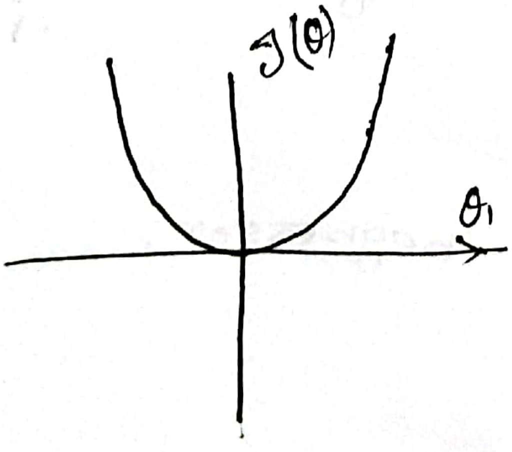
$\text{Loss} = (h(x) - y)^2$

$= (\theta_0 + \theta_1 x_1 - y)^2$

$Y = a + bx$

$$a = \frac{[(\Sigma y)(\Sigma x^2) - (\Sigma x)(\Sigma xy)]}{[n(\Sigma x^2) - (\Sigma x)^2]}$$

$$b = \frac{[n(\Sigma xy) - (\Sigma x)(\Sigma y)]}{[n(\Sigma x^2) - (\Sigma x^2)]}$$



Learning Rate.

$$\theta_1 = \theta_1 - \alpha \frac{d}{d\theta_1} J(\theta_1)$$

$$= \theta_1 - \alpha(+ve)$$

$$= \theta_1 + (\ldots\ldots)$$

$$\theta := \theta - \alpha \frac{d}{d\theta} J(\theta_0, \theta_1)$$ Partial derivative

For $\theta_0 \to \theta_0 := \theta_0 - \alpha \frac{1}{2} \frac{d}{d\theta_0} (h(x) - y)^2$

$$= \theta_0 - \alpha \frac{1}{2} \frac{d}{d\theta_0} (\theta_0 + \theta_1 x - y)^2$$

$$= \theta_0 - \alpha (\theta_0 + \theta_1 x - y) . 1$$

$$= \theta_0 - \alpha (h(x) - y) . x$$

$$h(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \cdots \theta_j x_j .$$

Repeat until Convergence {

$$\theta_j = \theta_j - \alpha (h(x) - y) x_j .$$

}

ML. Linear Regression plot diagram. .ipynb

why change and update.

classification and logistic regression.

ML

25-09-2023

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \dots \\ x_4 \end{bmatrix} \qquad \theta = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \end{bmatrix}$$

$x_1, x_2, x_3 \dots\dots$ $\qquad h(x) = \theta^T x.$

$$h(x) = \theta_0 + \theta_1 x + \theta_2 x_2 + \dots$$

$$J(\theta) = \frac{1}{2m} \Sigma \left(h(x) + y\right)^2$$

$$\approx \frac{1}{2m} (X\theta - y)^T (X\theta - y)$$

$$\frac{d}{d\theta} \left\{ (X\theta)^T X\theta - (X\theta)^T y - y^T (X\theta) + y^T y \right\}$$

$$2X^T X\theta - 2X^T y = 0$$

$$\theta = \theta - \alpha \frac{J}{J\theta}(h(x) - y)$$

$$\theta = (X^T X)^{-1} X^T y$$

$$h(x) = \theta^T x$$

Optimox

Classification

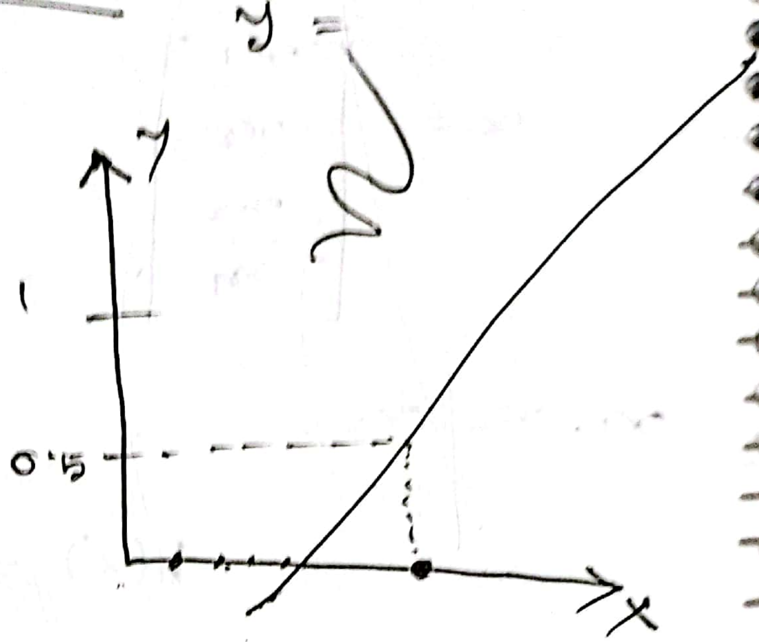
$y = (-\alpha, \alpha)$

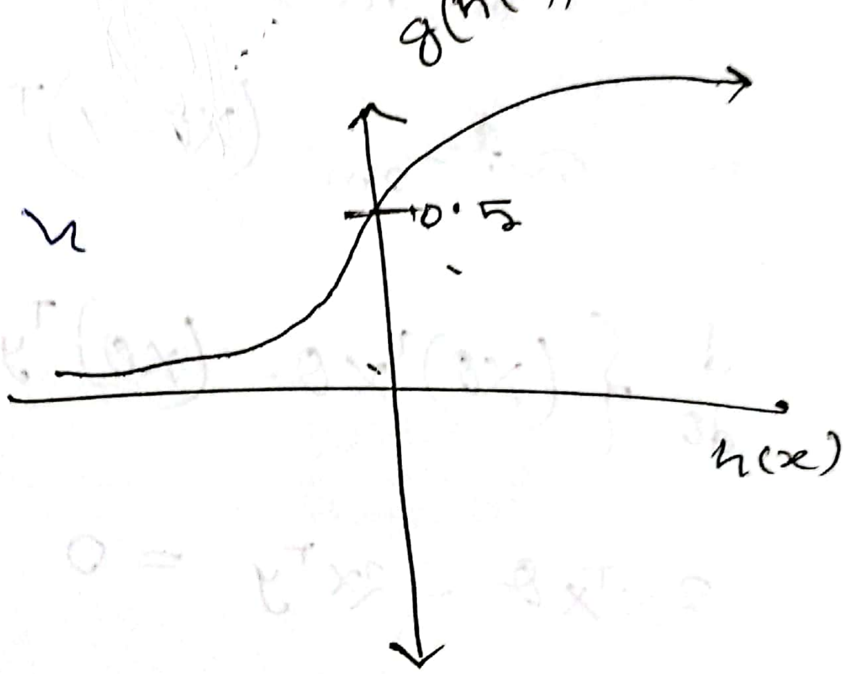
$h(x) = \theta_0 + \theta_1 x$

$$g(h(x)) = \frac{1}{1 + e^{-h(x)}}$$

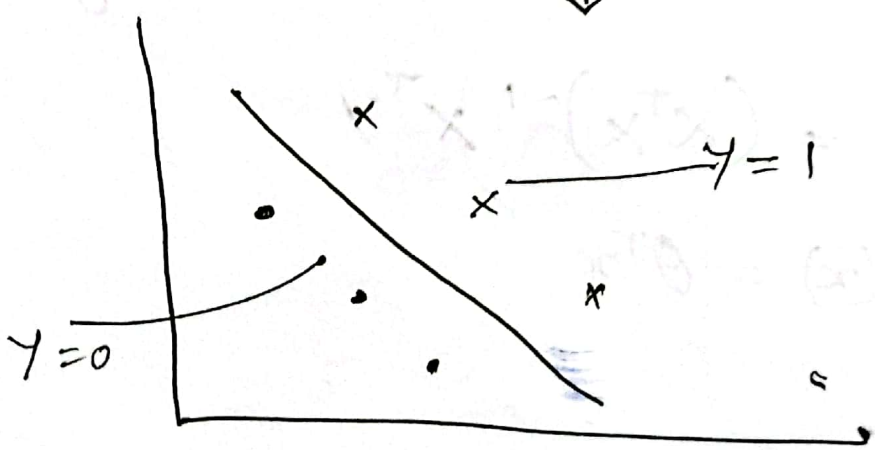
$$g(z) = \frac{1}{1 + e^{-z}}$$

for $y = \{0, 1\}$





x_1	x_2
0.5	1.5
1	1
1.5	0.5
3	0.5
2	2
1	2.5



* Logistic Regression

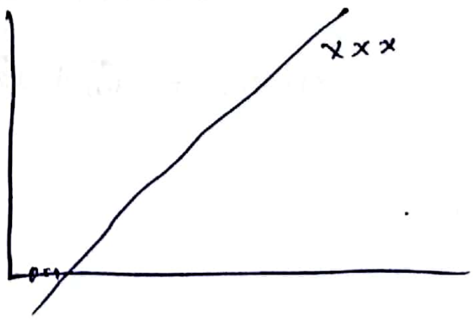
$$y = h(x) = \theta^T X = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \ldots$$

$$y - \{0, 1\}$$

Logistic function.

* Sigmoid function.

$$g(z) = \frac{1}{1 + e^{-z}}$$

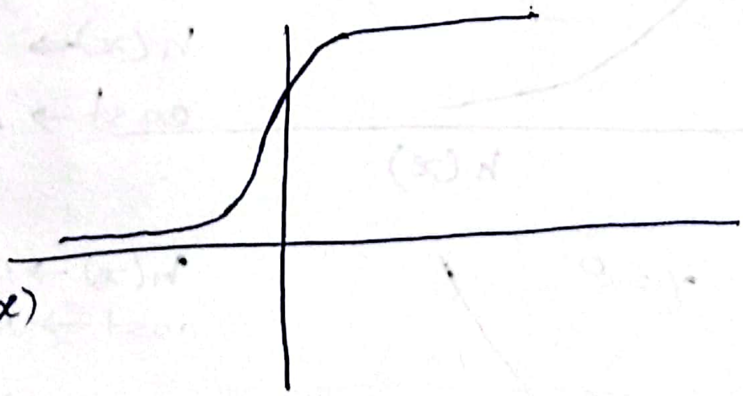


$$h(x) = \theta_0 + \theta_1 x$$

→ Predicted hypothesis

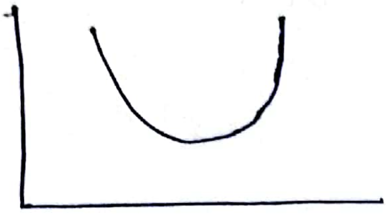
$$y = \{0, 1\}$$

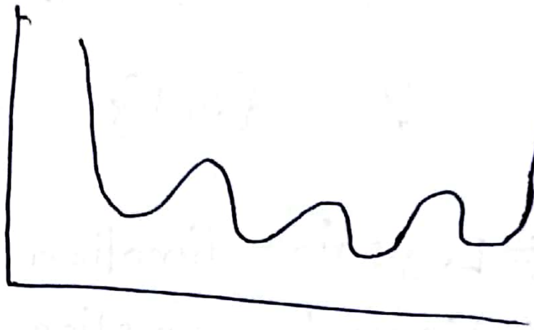
$$0 \leq h(x) \leq 1 \quad , \quad h(x) = \frac{1}{1 + e^{-\theta^T X}}$$



$\frac{1}{1+e^{-\theta^T x}}$

$$J(\theta) = \frac{1}{2m} \sum \left(h(x) - y\right)^2$$





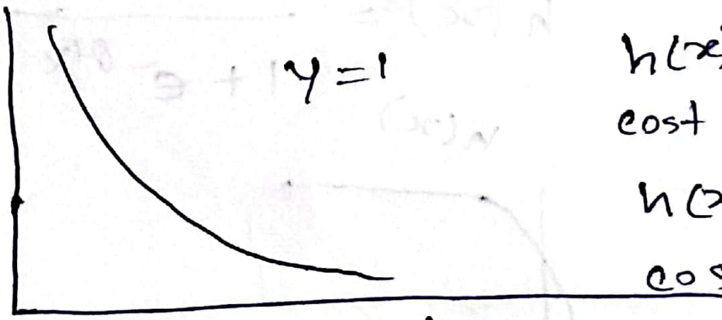
$h(x) = \theta_0 + \theta_1 x_1$

$\log_e e^x$

$$J(\theta) = \begin{cases} -\log(h(x)) ; & y = 1 \\ -\log(1 - h(x)) ; & y = 0 \end{cases}$$

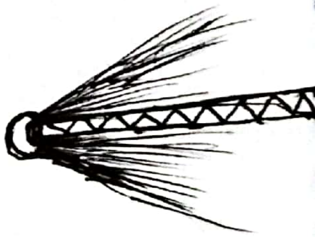
cost function.

$-\log(h(x))$
$J(\theta)$.

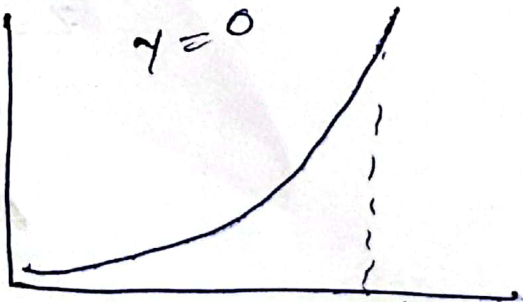


$h(x) \Rightarrow 0$
cost $\rightarrow \infty$

$h(x) \Rightarrow 1$
cost $\rightarrow 0$



$J(\theta)$
$-\log(1 - h(x))$



$h(x) \rightarrow 0$.
cost $\rightarrow 0$

$h(x) \rightarrow 1$
cost $\rightarrow \infty$

Merge two equation:

$$J(\theta) = -y \log(h(x)) - (1-y)\log(1-h(x))$$

$y = 1$; $J(\theta) = -\log(h(x))$.

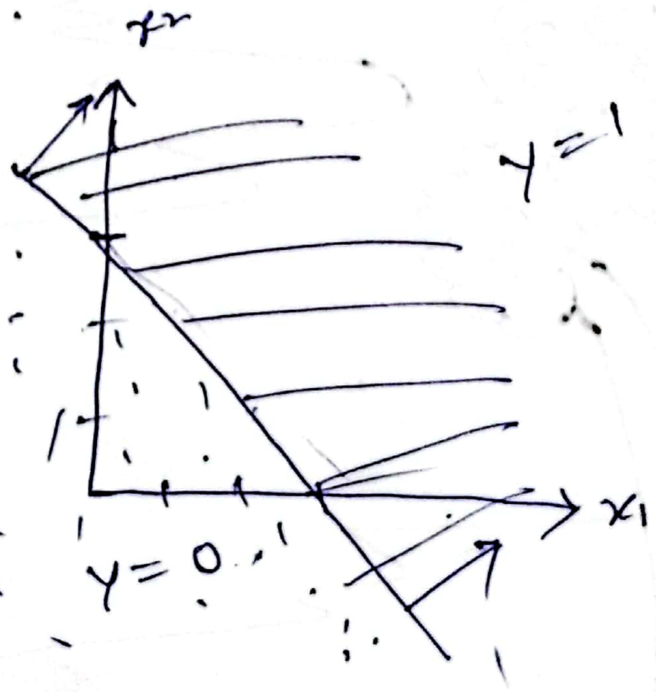
$y = 0$; $J(\theta) = -\log(1-h(x))$.

⇒ skewed graph;

⇒ dataset observe:

⇒ threshold may be not always 0'5

⇒ May 0'95.



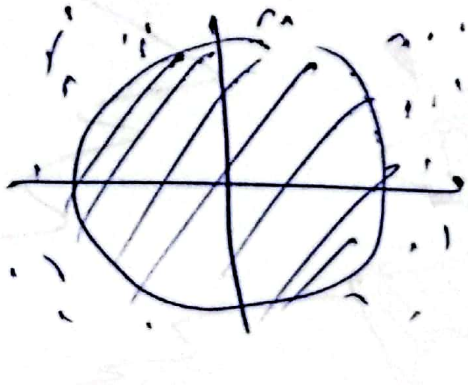
$0 \leq h(x) \leq 1$

$h(x) \geq 0.5$

$\theta^T x \geq 0$

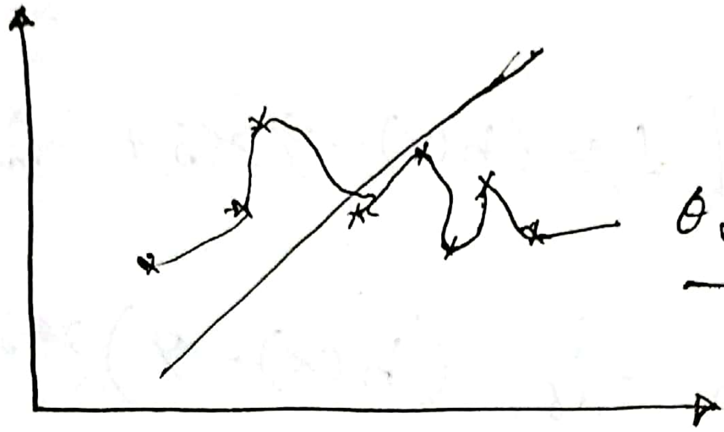
$-3 + x_1 + x_2 \geq 0$.

$x_1 + x_2 \geq 3$



not only linear.

ML 09-10-2023



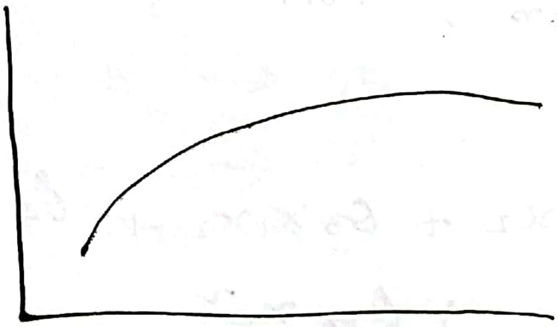
$\theta_0 + \theta_1 x + \theta_2 x^2 + \theta_3 x^3$

Overfit

High varience...

Underfit

High bias



$\theta_0 + \theta_1 x + \theta_2 x^2$

Just Perfect

Low varience

Less bias

⇒ Regularization

⇒ Boosting

⇒ Bagging

$\Sigma (h(x) - y)^2 + \boxed{1000}\,\theta_3^2$

Penalty add

Regularization term.

$\Sigma (h(x) - y)^2 + \lambda \Sigma \theta_j^2$

$\theta_j := \theta_j - \alpha \frac{1}{2m} \Sigma (h(x) - y)^2 x_j$

Updated cost function

Optimex

Regularization $\theta_0 + \theta_1 x + \theta_2 x + \theta_3 x$

$$\theta_j = \theta_j - \alpha \left[\frac{1}{m} \Sigma (h(x) - y) x_j + \frac{\alpha}{m} \theta_j \right]$$

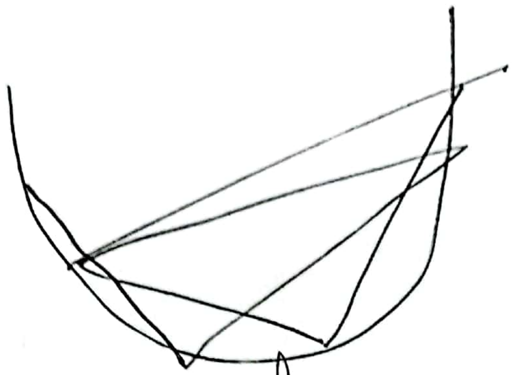
$$\theta_j = \theta_j - \alpha \frac{1}{2m} \frac{d}{d\theta_j} (h(x) - y) x_j$$

$$\theta_j = \theta_j \left(1 - \alpha \frac{\lambda}{m}\right) - \alpha \frac{1}{m} \Sigma (h(x) - y) x_j$$

Non-linear regression, Polynomial regression

$$Y = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_1 x_2 + \theta_4 x_1^2 + \theta_5 x_2^2.$$

$x = 1.0$ and $y = 3.0$ | ML | 11-10-23



$l \Rightarrow$ learning rate.

$L = 0.5$.

1st training data : $x = 3.0$ and $y = 1.0$.

Initial A ;

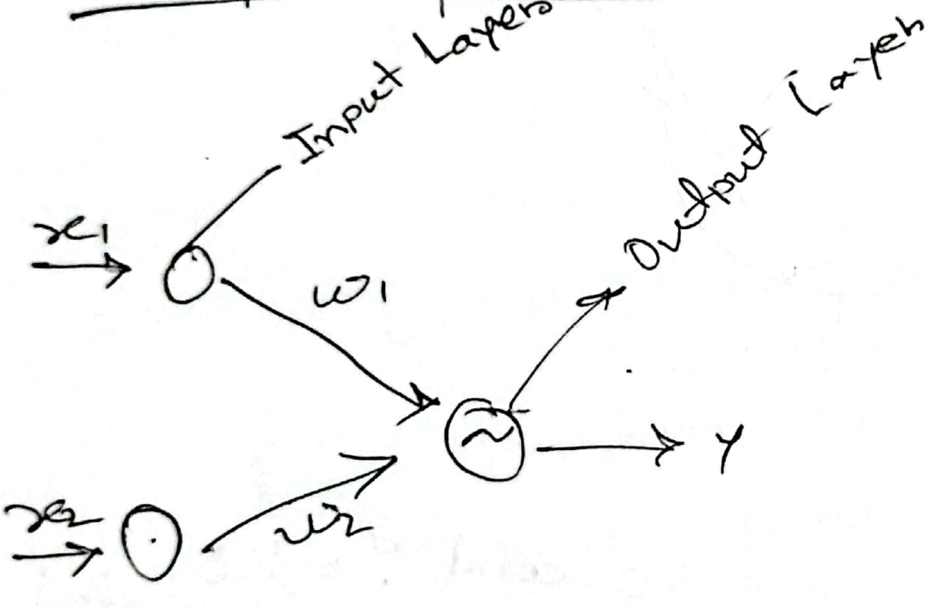
Optimox

Neural network

Complex polynomial function.

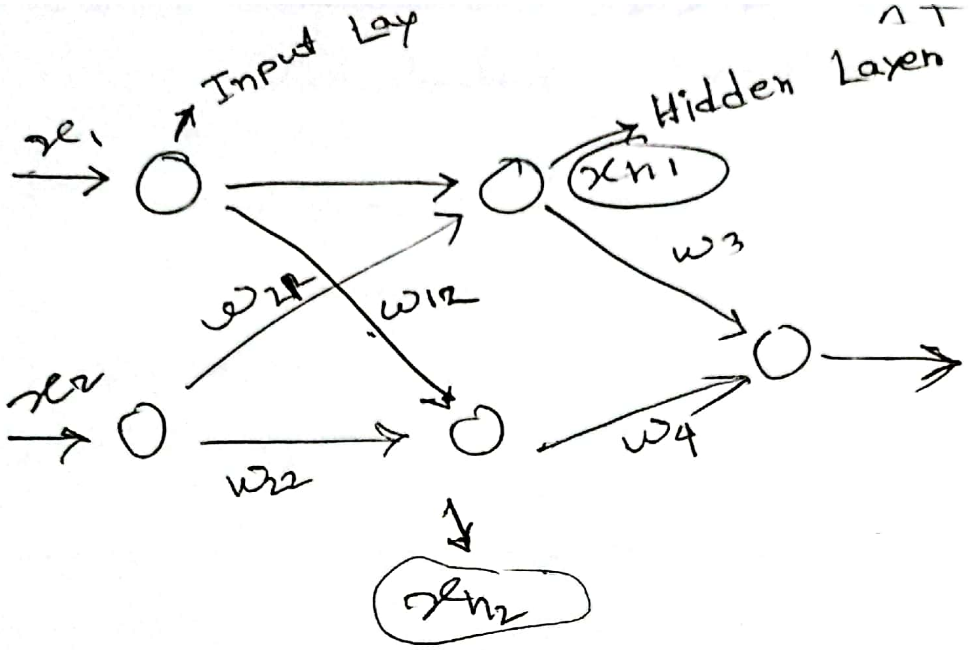
$x_1 x_2 x_3^2 x_4^4 x_n^2$

x_1	x_2	y
10	5	0
12	0	1
15	10	0



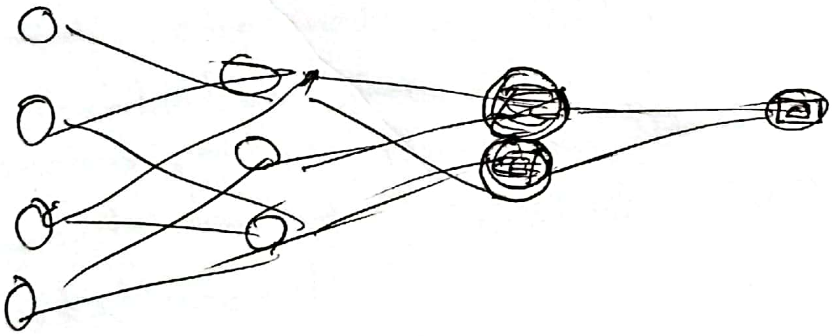
$$y = g(x_1 y_1 + x_2 w_2)$$

$$g(z) = \frac{1}{1+e^{-z}}$$



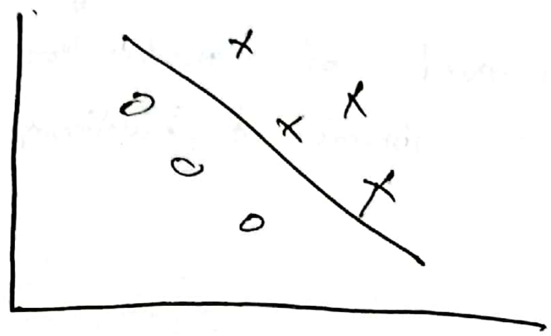
3.5

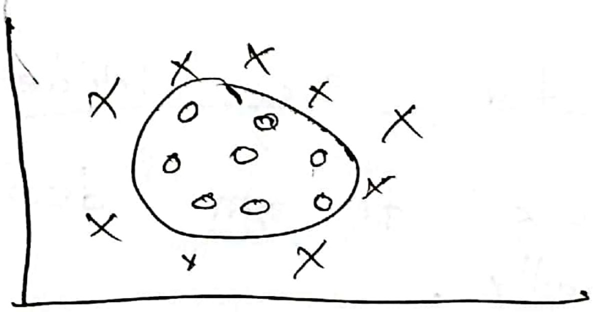
$$Y = g(w_3 x_{h1} + w_4 x_{h2})$$



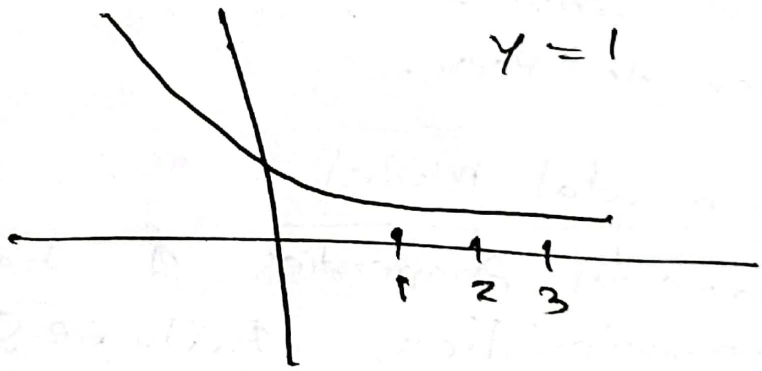
The ~~...~~

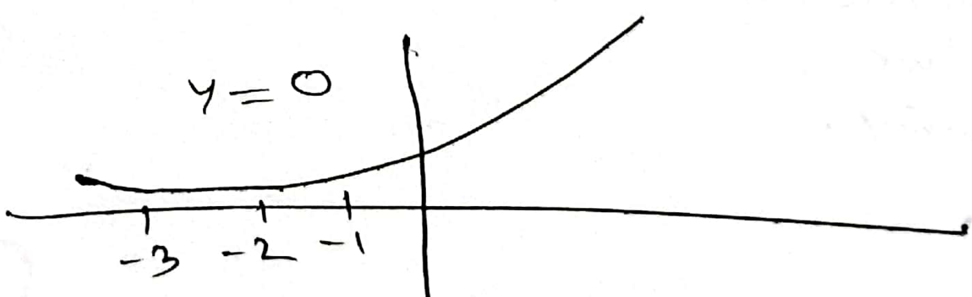
Support vectora Machine (SVM)

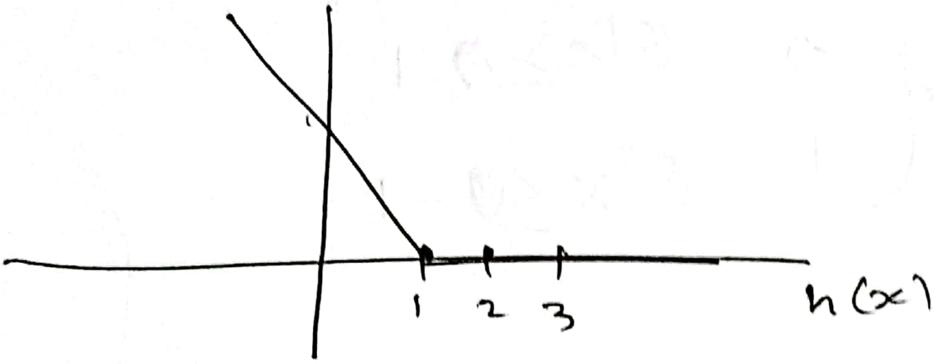


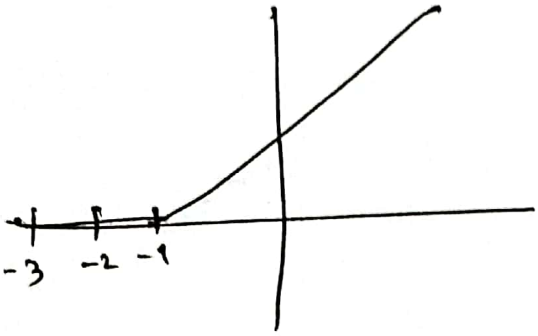


$$cost = -y \log h(x) - (1-y) \log (1 - h(x)).$$









$$-\min \frac{1}{m}\left[\sum y \log h(x) + (1-y)\log\left(1-h(x)\right)\right]$$

$$+ \frac{\lambda}{2m}\theta_y^2 \,]$$

$\min(\cancel{4}-5)^2+1$ $\qquad u = 5.$

$\min \{(u-5)^2$ $\qquad$ too large

$A + \lambda B$. $\qquad c = \frac{1}{\lambda}$

$CA + B$

too small.

Optimox

$$h(x) = \begin{cases} 0 & \theta^T x \geq 0 \quad 1 \\ 1 & \theta^T x \leq 0 \quad -1 \end{cases}$$

$$\frac{1}{1 + e^{\theta^T x}}.$$