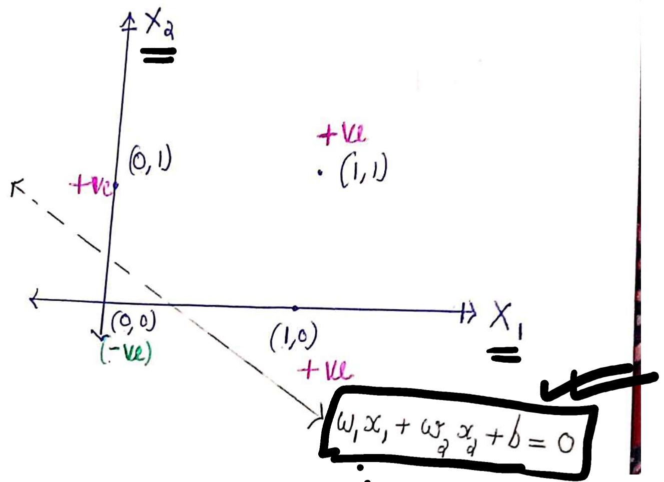
Perceptron

Perceptron is a linear supervised learning algorithm which can efficiently perform a binary classification task
The predictions made by a perceptron are based on linear predictor functions

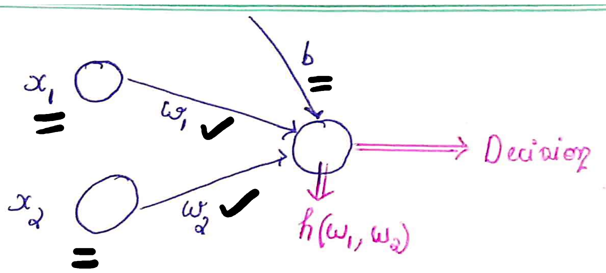
X_1	X_2	Y
1	0	1
0	1	1
1	1	1
0	0	$\boxed{-1}$



$$\boxed{w_1 x_1 + w_2 x_2 + b = 0}$$

$$Xw = 0$$

$$X = \begin{array}{c} X_1 \quad X_2 \\ \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}_{4\times 3} \end{array} \qquad w = \begin{pmatrix} w_1 \\ w_2 \\ b \end{pmatrix} \qquad Y = \begin{pmatrix} 1 \\ 1 \\ 1 \\ -1 \end{pmatrix}$$



Decision rule :-

$$h(x_1, x_2) = \text{sign}(w_1 x_1 + w_2 x_2 + b)$$

$$= \begin{cases} +1 & \text{if } w_1 x_1 + w_2 x_2 + b > 0 \\ -1 & \text{if } w_1 x_1 + w_2 x_2 + b < 0 \\ 0 & \text{if } w_1 x_1 + w_2 x_2 + b = 0 \end{cases}$$

$(Y = 1)$

Correct if

$z = w_1 x_1 + w_2 x_2 + b > 0$ } $y(w_1 x_1 + w_2 x_2 + b) > 0$

$\Rightarrow h(z) = 1$

Incorrect if

$z = w_1 x_1 + w_2 x_2 + b < 0$ } $y(w_1 x_1 + w_2 x_2 + b) < 0$

$\Rightarrow h(z) = -1$

$(Y = -1)$

Correct if

$z = w_1 x_1 + w_2 x_2 + b < 0$ } $y(w_1 x_1 + w_2 x_2 + b) > 0$

$\Rightarrow h(z) = -1$

Incorrect if X

$z = w_1 x_1 + w_2 x_2 + b > 0$ } $y(w_1 x_1 + w_2 x_2 + b) < 0$

$\Rightarrow h(z) = +1$

$Y^T X w$

$$f(\underline{w}) = Y^T X w$$

$$\max_{\underline{w}} f(w) = \max_{\underline{w}} \sum_{i=1}^{4} y_i (x_{i1} w_1 + x_{i2} w_2 + b)$$

$$\nabla_w f(\underline{w}) = \begin{pmatrix} \frac{\partial}{\partial w_1} f(w) \\ \frac{\partial f(w)}{\partial w_2} \\ \frac{\partial f(w)}{\partial b} \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{4} y_i x_{i1} \\ \sum_{i=1}^{4} y_i x_{i2} \\ \sum_{i=1}^{4} y_i \end{pmatrix}$$

$$\nabla_\omega f(\omega) = \begin{pmatrix} y_1 x_{11} + y_2 x_{21} + y_3 x_{31} + y_4 x_{41} \\ y_1 x_{12} + y_2 x_{22} + y_3 x_{32} + y_4 x_{42} \\ y_1 + y_2 + y_3 + y_4 \end{pmatrix}$$

$$\nabla_\omega f(\omega) = \begin{pmatrix} x_{11} & x_{21} & x_{31} & x_{41} \\ x_{12} & x_{22} & x_{32} & x_{42} \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}$$

$$\nabla_\omega f(\omega) = X^T Y.$$

$$\omega_{n+1} = \omega_n + \alpha \nabla_\omega f(\omega)$$

$$\omega_{n+1} = \omega_n + \alpha X^T Y$$

ω_n := weight vector at nth iteration.

α := learning rate

ω_0 ✓
α ✓
n ✓

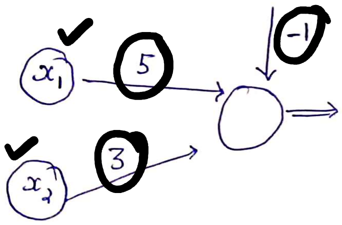
We get eventually $w_1 = 5$, $w_2 = 3$, $b = -1$

equation of decision boundary $= 5x_1 + 3x_2 - 1$

The decision rule will be

$$h(x_1, x_2) = \begin{cases} +1 & \text{if } 5x_1 + 3x_2 - 1 > 0 \Rightarrow 5x_1 + 3x_2 > 1 \\ -1 & \text{if } 5x_1 + 3x_2 - 1 < 0 \Rightarrow 5x_1 + 3x_2 < 1 \\ 0 & \text{if } 5x_1 + 3x_2 - 1 = 0 \Rightarrow 5x_1 + 3x_2 = 1 \end{cases}$$

X_1	X_2	$5x_1 + 3x_2$	$h(x_1, x_2)$		
1	1	8	1	$8 > 1$	1
1	0	5	1	$5 > 1$	1
0	1	3	1	$3 > 1$	1
0	0	0	−1	$0 < 1$	−1



$$\begin{cases} +1 & \text{if } 5x_1 + 3x_2 > 1 \\ -1 & \text{else} \end{cases}$$