AGENDA :-

→ (i) More about Perceptron. { Mathematically }

→ (ii) Python implimentation ✓

{ Jupyter notebook }

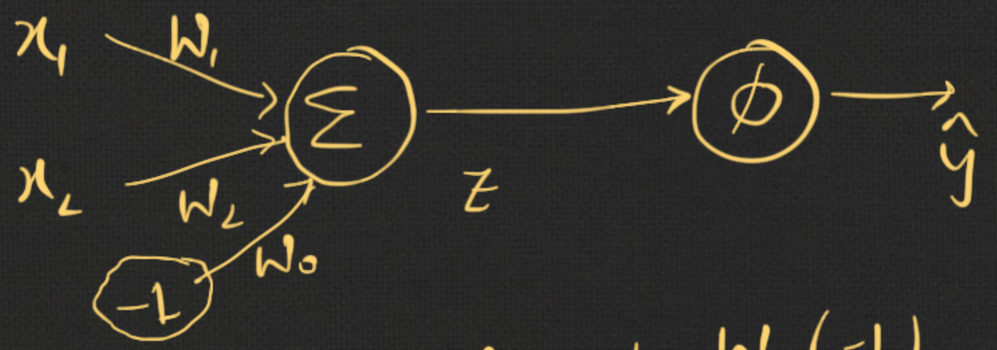
VS code By simple

Dataset –

AND, OR, XOR

↓

| $x_1$ | $x_2$ | $y$ |
|---|---|---|
| 0 | 0 | 0 |
| 0 | 1 | 0 |
| 1 | 0 | 0 |
| 1 | 1 | 1 |


$$z = w_1 x_1 + w_2 x_2 + w_0(-1)$$

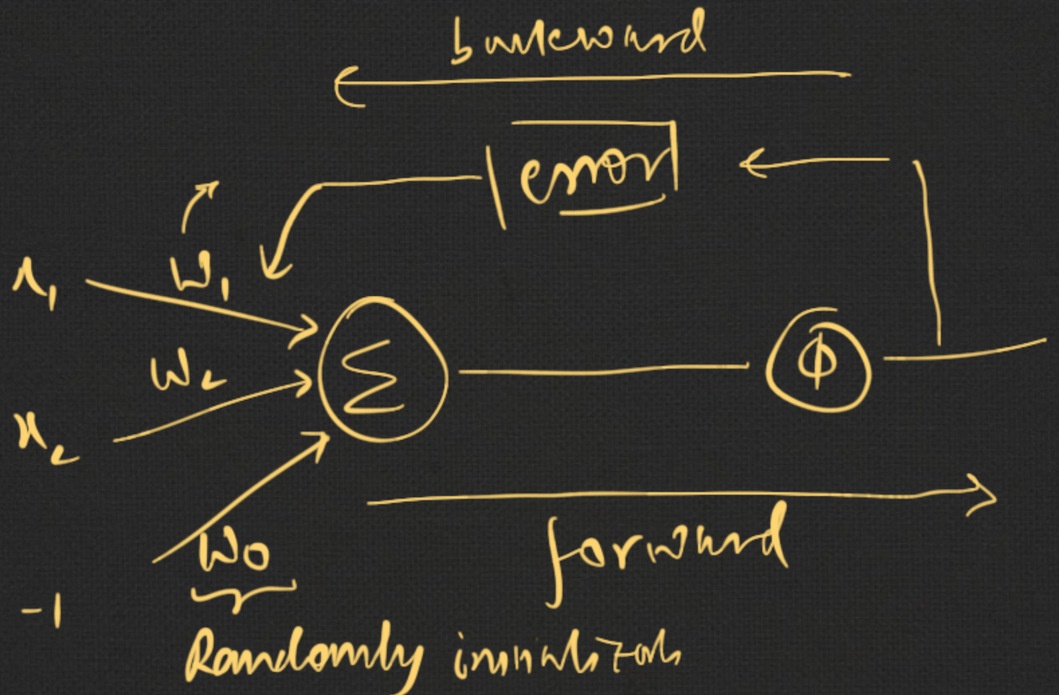
$$\hat{y} = \Phi(z) = \begin{cases} 0 & z < 0 \\ 1 & z \geq 0 \end{cases}$$

$$\text{error} = y - \hat{y}$$

$$\overline{w} = w + \Delta w \quad , \quad \Delta w = \eta (y - \hat{y}) x$$

AND

| $x_1$ | $x_2$ | bias | $y$ |
|---|---|---|---|
| ↑ | ↑ | ↑ | ↑ |




threshold = 0

1 epoch = 1 iteration = forward + backward

| | | |
|---|---|---|
| 0 | 0 | → 0 |
| 0 | 1 | → 0 |
| 1 | 1 | → 1 |

AND

| $x_1$ | $x_2$ | $\hat{y}$ |
|---|---|---|
| 0 | 0 | 0 |
| 0 | 1 | 0 |
| 1 | 0 | 0 |
| 1 | 1 | 1 |

$X$ $\quad$ $y$

| $x_1$ | $x_2$ | bias |
|---|---|---|
| 0 | 0 | -1 |
| 0 | 1 | -1 |
| 1 | 0 | -1 |
| 1 | 1 | -1 |

$X$ with bias

$\Rightarrow X_{4\times3}$ $\qquad W = \begin{bmatrix} w_1 \\ w_2 \\ w_0 \end{bmatrix}_{3\times1}$

Transpose operation in matrix

$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}_{2\times3}$ $\xrightarrow{A^T}$ $\begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}_{3\times2}$

Multiplication in matrix

$A_{3\times2}$ $B_{4\times2}$ $= X$

$A_{3\times2}$ $B_{2\times4}$ $= C_{3\times4}$

$$W = W + \Delta W \Rightarrow W = W + \eta \underset{4\times3}{X^T} \cdot \underset{4\times1}{e}$$

$$\begin{pmatrix} w_1 \\ w_2 \\ w_0 \end{pmatrix} = \begin{pmatrix} w_1 \\ w_2 \\ w_0 \end{pmatrix} + \eta \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ -1 & -1 & -1 & -1 \end{bmatrix}_{3\times4} \begin{pmatrix} y_1 - \hat{y}_1 \\ y_2 - \hat{y}_2 \\ y_3 - \hat{y}_3 \\ y_4 - \hat{y}_4 \end{pmatrix}_{4\times1}$$

$$\underbrace{\begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix}}_{3\times1} = \underbrace{\begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix}}_{3\times1} + \eta \underbrace{\begin{pmatrix} \Delta w_1 \\ \Delta w_2 \\ \Delta w_3 \end{pmatrix}}_{\Delta W}{}_{3\times1}$$

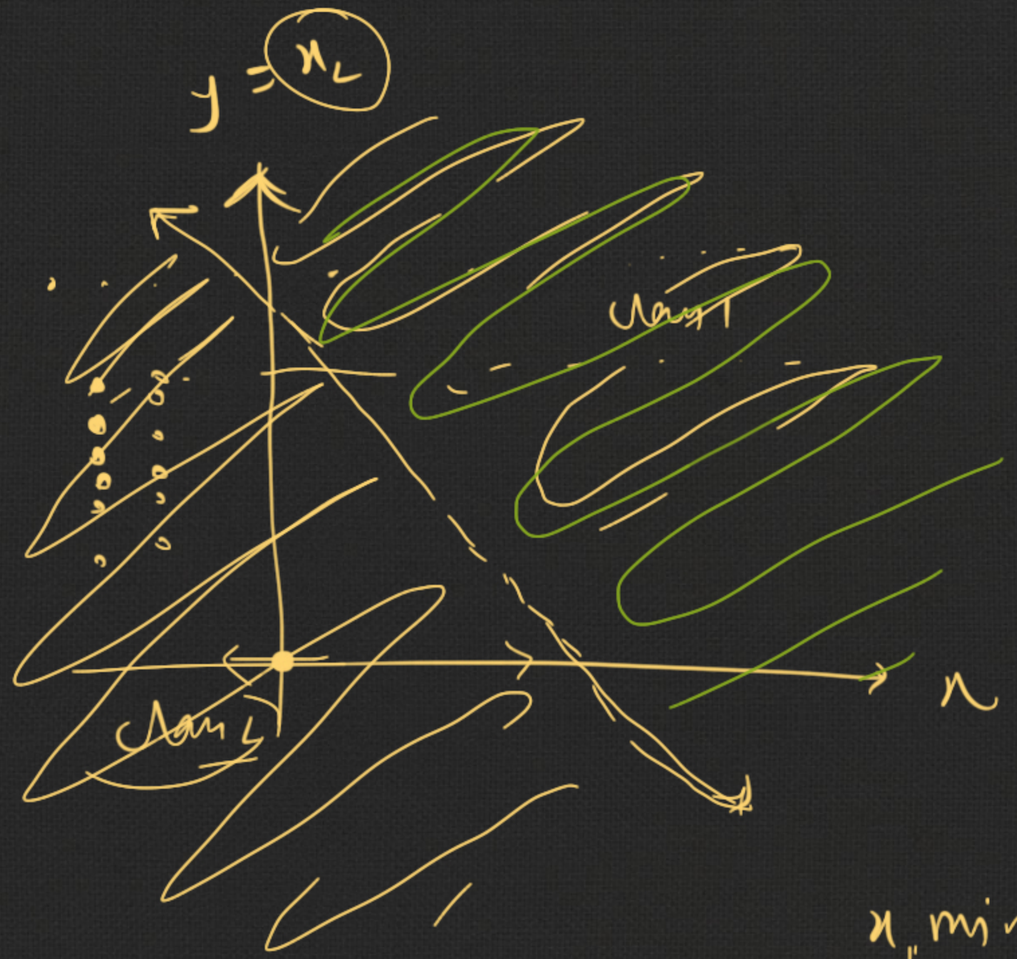
$$W_{new} = W_{old}$$

Stochastic Gradient Descent

Random
↓
choose datapoint → calculate error → update weights

Random batches
↓
calculate avg. error → update weights.

$$y = \phi(z)$$
$$= \phi(W \cdot x)$$




$x_1 \to 0,1$

$x_2 \to 0,1$

$-1, 2$

modul.

$x = 0$

$y =$

$y > 0$

$x_1 min$ $y min$

$x max$ $y max$