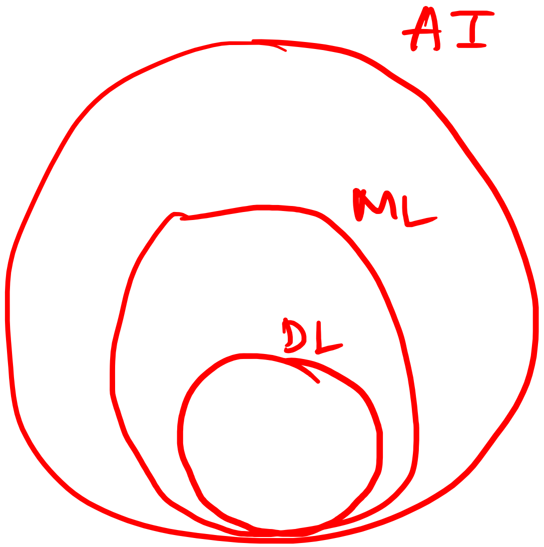
Deep learning:

X → ANN - tabular

→ CNN - image & video

→ RNN - text



x_1	x_2	x_3	y
Screentime	Rest time	Study time	Passing criteria
3	9	6	1
6	7	2	0

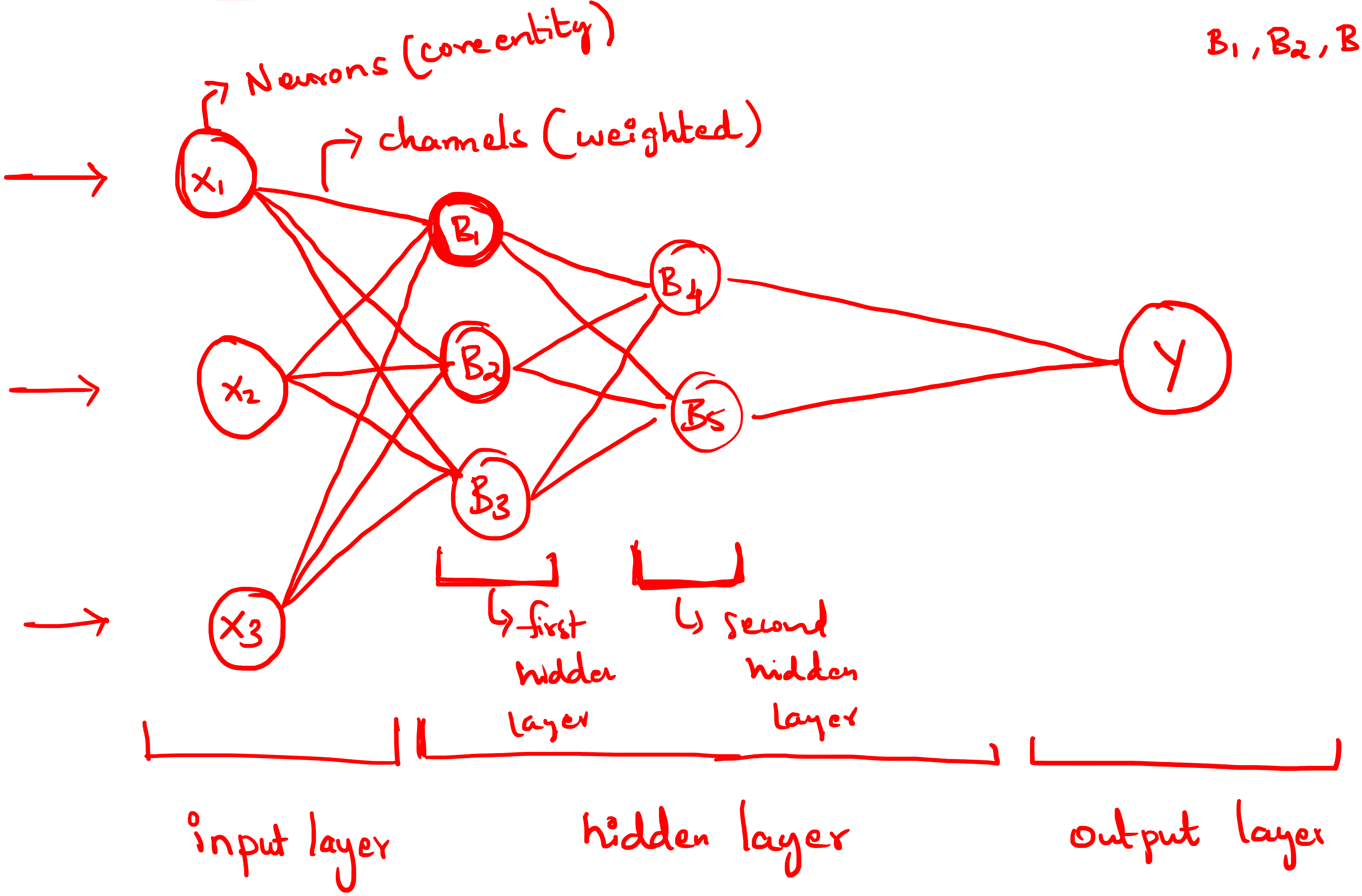
} [binary Classification]

Architecture: → ANN

X_n, B_n, Y – neurons

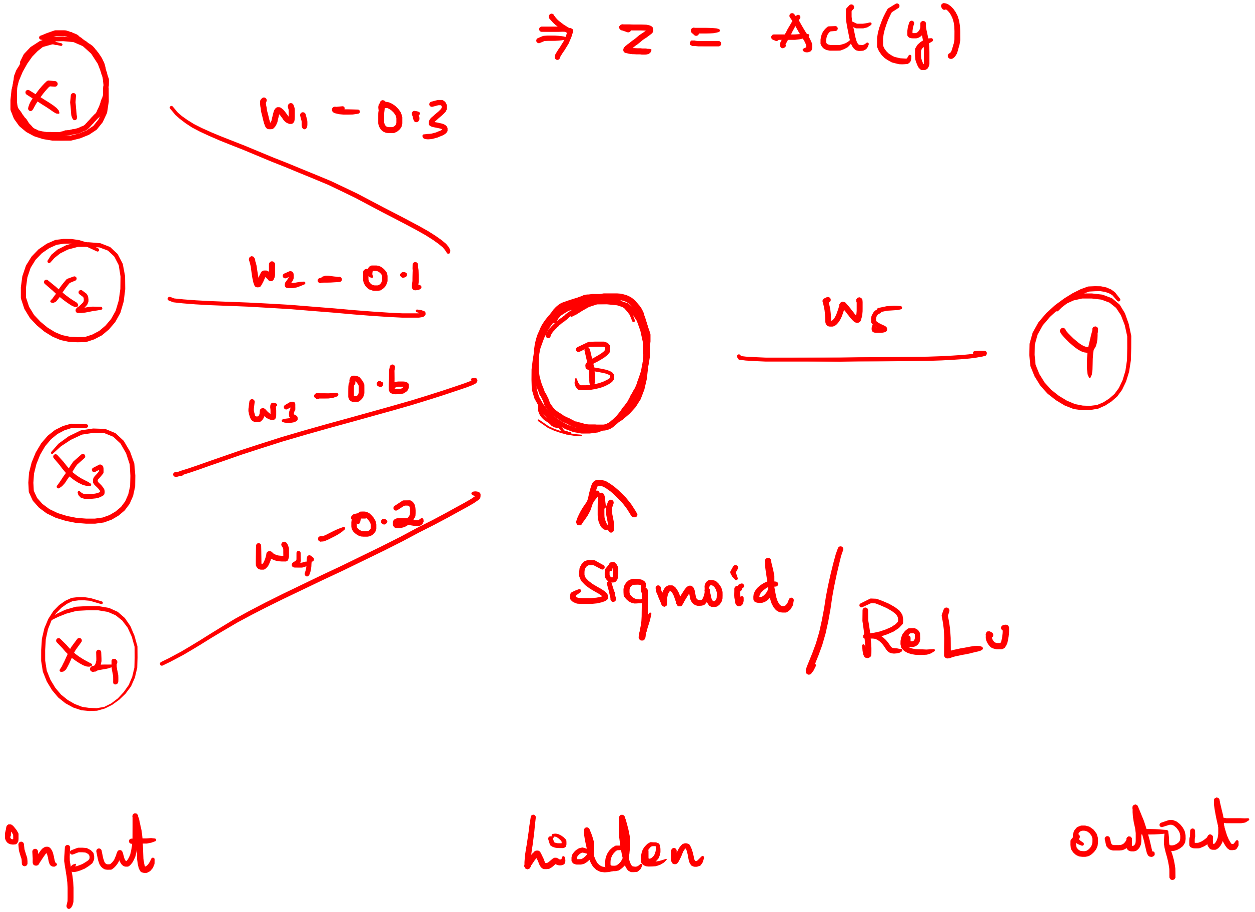
B_1, B_2, B_n – bias

unique number associated with each neuron in the hiddenlayer



$$\Rightarrow \{ y = w_1x_1 + w_2x_2 + w_3x_3 + w_4x_4 + \text{Bias} \}$$

$$\Rightarrow z = \text{Act}(y)$$



$$\overbrace{\{\text{Weighted sum of Input}\} + \{\text{Bias}\}}^{y} \Rightarrow \{\text{applied to known activation function}\}$$

$$[(x_1w_1) + (x_2w_2) + (x_3w_3) + (x_4w_4)]$$

$$+$$

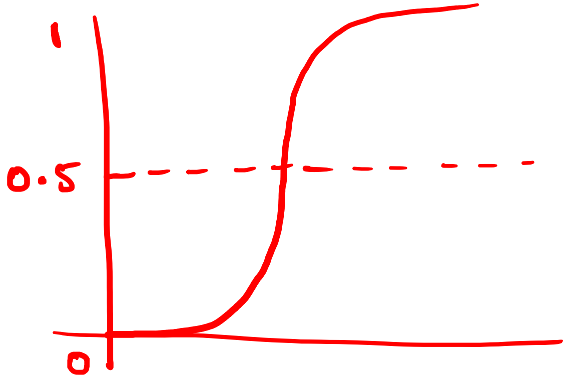
$$[(B)]$$

{ Refer Next slide }

$\{(x_1 \times 0.3) + (x_2 \times 0.1) + (x_3 \times 0.6) + (x_4 \times 0.2)\} + \underline{\{B\}} \rightarrow$ { Applied to the known activation function }

Sigmoid → (0 − 1) → classification

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



$Z < 0.5$ → Not activated

$Z \geq 0.5$ → activated

Sigmoid

ReLu

↓

{ The result of the activation function determines if the neuron gets activated }

ReLu ⇒ $(0-\infty)$ ⇒ Regression

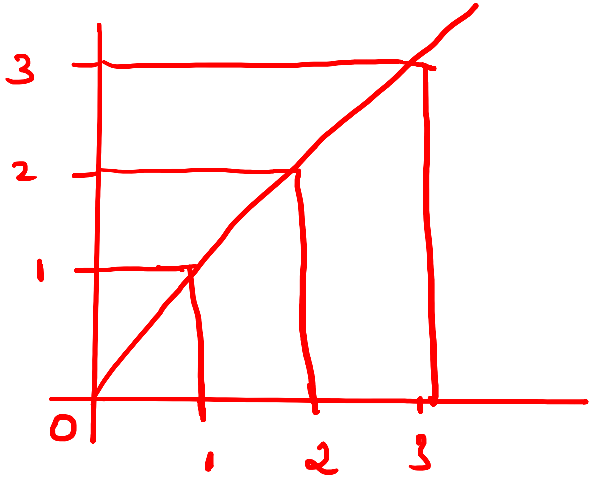
$Z = \max(y, 0)$

$z = \max(285, 0)$

$= 285$

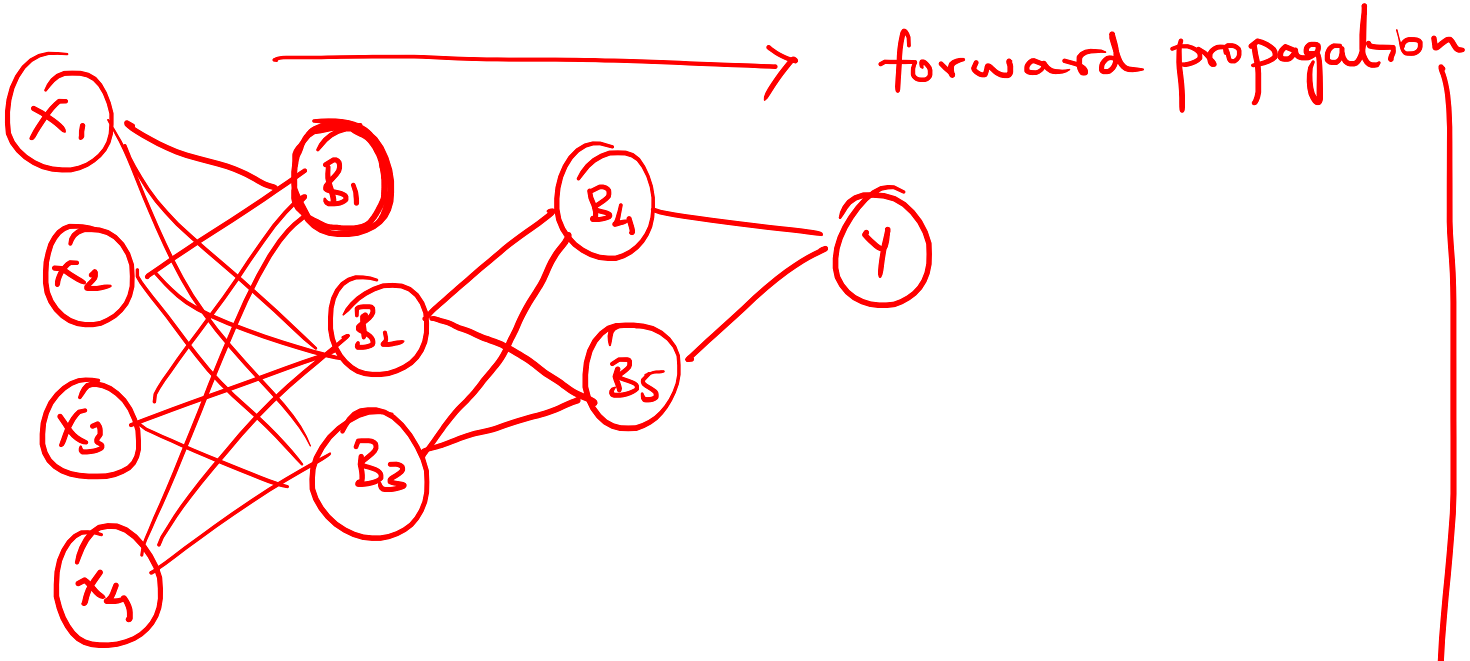
$z = \max(-16, 0)$

$= 0$



$Z \leq 0$ → Not activated

$Z > 0$ → Activated



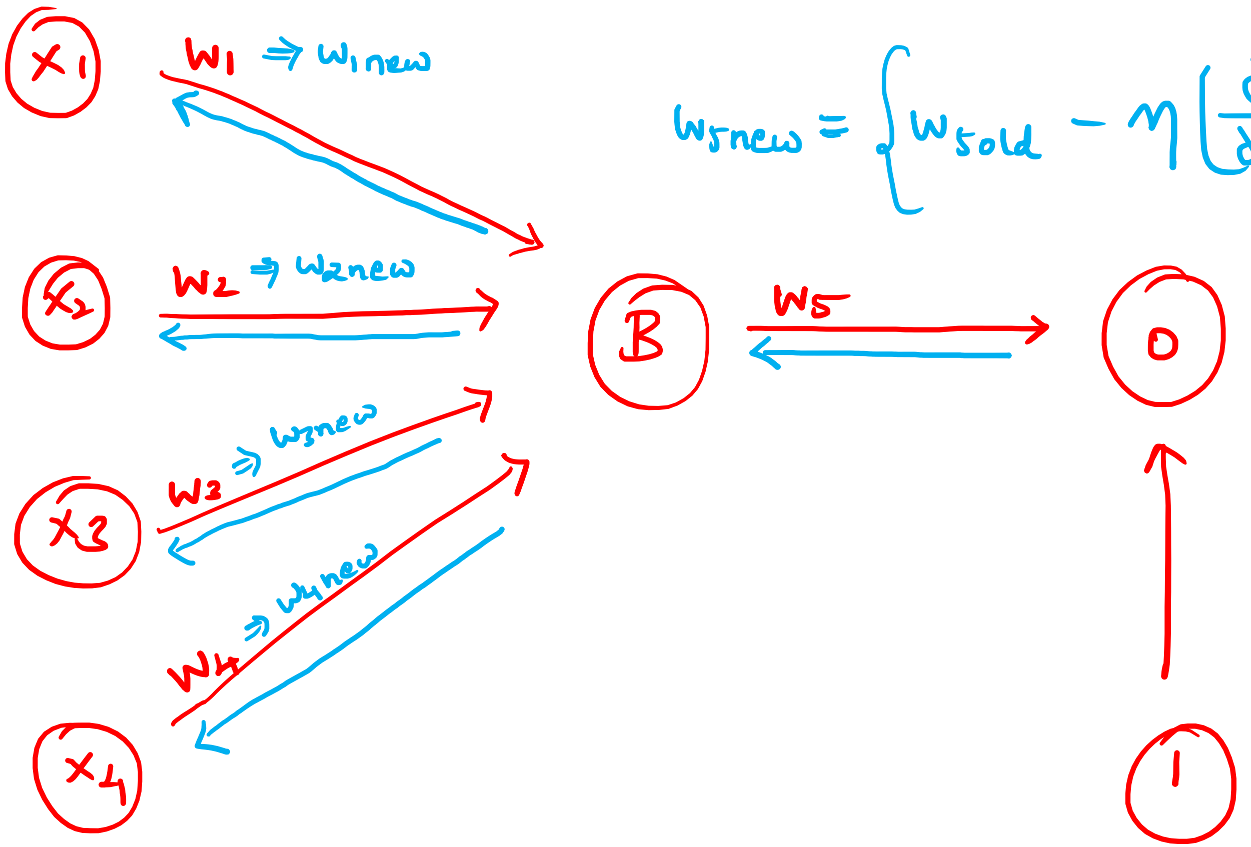
$$(x_n W_n) + (B_1) \Rightarrow (\text{ReLu}) \Rightarrow 0$$

$$(x_n w_n) + (B_2) \Rightarrow (\text{ReLu}) \Rightarrow 1$$

$$(x_n w_n) + (B_3) \Rightarrow (\text{ReLu}) \Rightarrow 1$$

Between the layers different Activation functions can be used. But, within one layer activation function can't be changed.

→ forward propagation



$$w_{5new} = \left\{ w_{5old} - \eta \left(\frac{\partial L}{\partial w_{5old}} \right) \right\} \Rightarrow$$ optimizer(GD, SGD, Adam)

$$\left. \begin{aligned} \text{Loss} &= (y - \hat{y})^2 \\ &= (1-0)^2 \\ &= 1 \end{aligned} \right\}$$

$\hat{y}$ → pred output

y = actual output

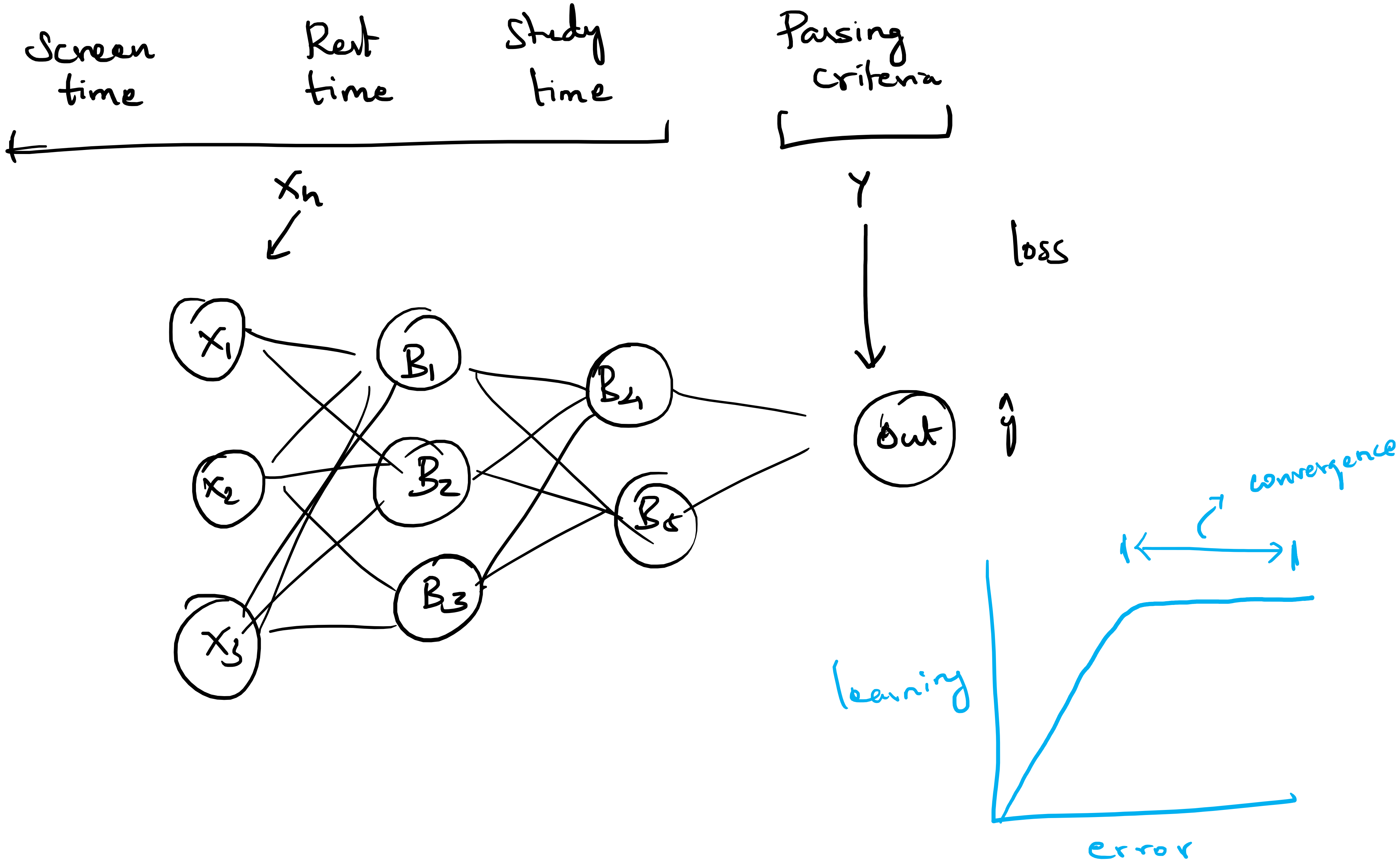
(η - learningrate)

its very small

$$\left\{ \begin{array}{c} 0.001 \\ 0.002 \\ \vdots \end{array} \right\}$$

← Backward propagation

forward & backward propagation will happen till convergence. ⓧ


Screen time
Rest time
Study time
Passing criteria
X_n
Y
loss
X_1
X_2
X_3
B_1
B_2
B_3
B_4
B_5
Out
$\hat{y}$
convergence
learning
error

Square = 1

○ ⇒ 0.6 ⇒ $(0 - 0.6)^2$ ⇒ 0.36

△ ⇒ 0.2 ⇒ $(0 - 0.2)^2$ ⇒ 0.04

□ → 0.2 → $(1 - 0.2)^2$ ⇒ 0.64