

Deep Learning → 1940

↳ Classification  
↳ Regression

1980-1990

Neural Network

2010 + Very popular → DL



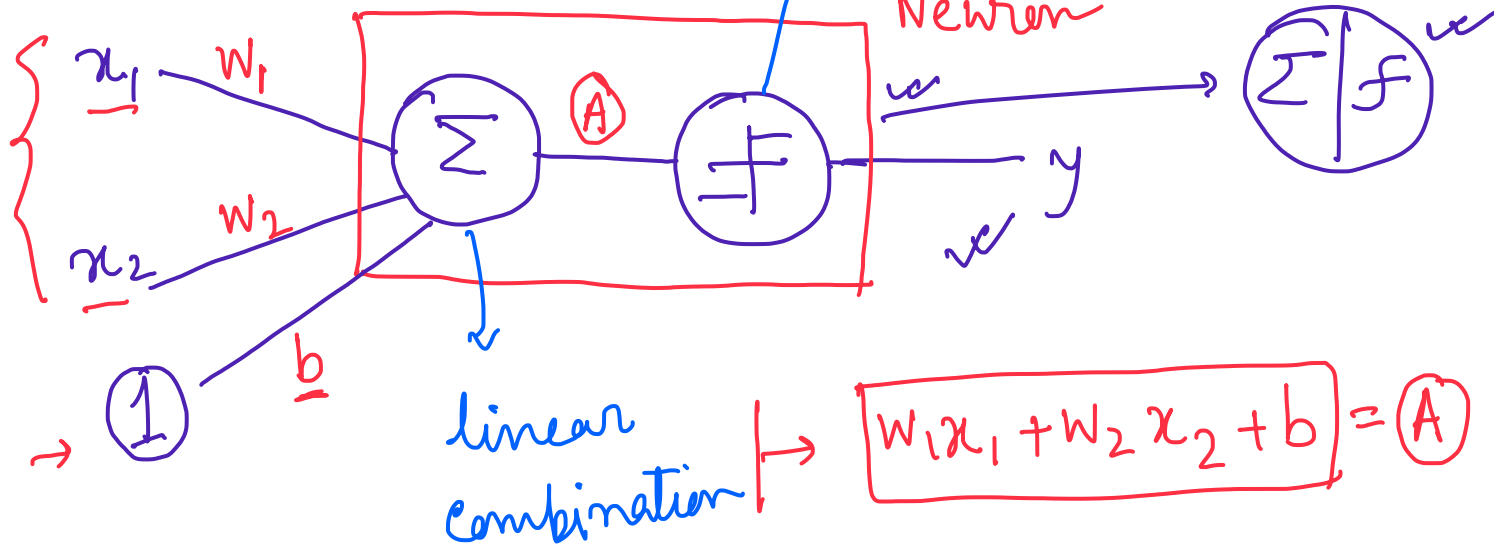
Activation function

Thresholding

$A > 0 \rightarrow y = 1$   
 $A < 0 \rightarrow y = 0$

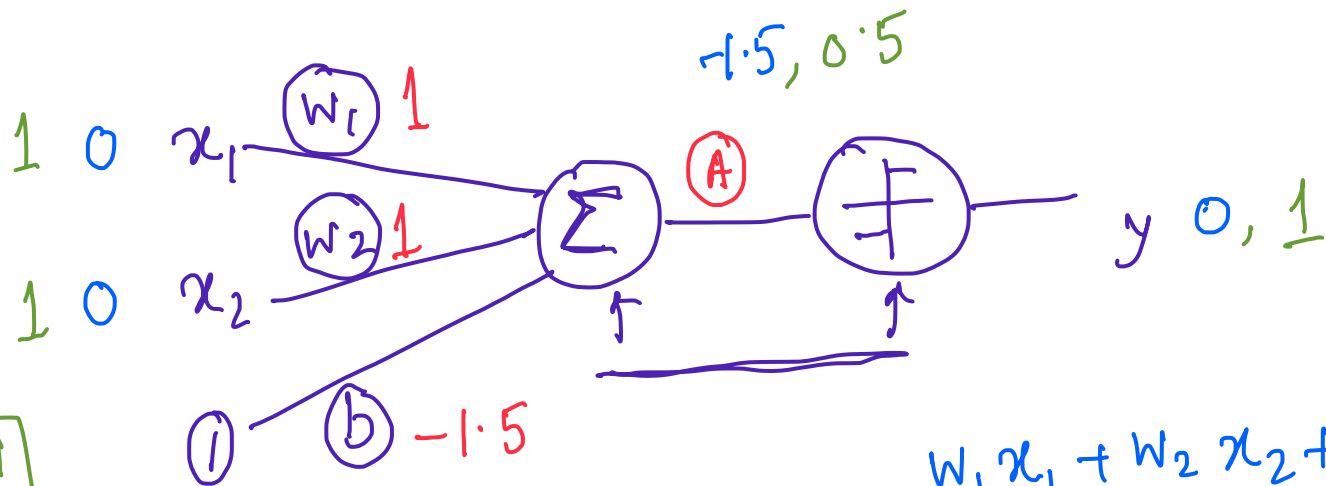


non-linear transformation!

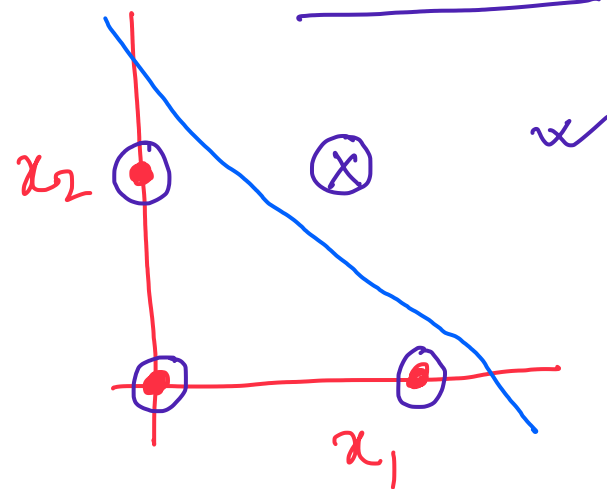


linear combination  $\rightarrow W_1 x_1 + W_2 x_2 + b = A$

AND  $x_1, x_2$



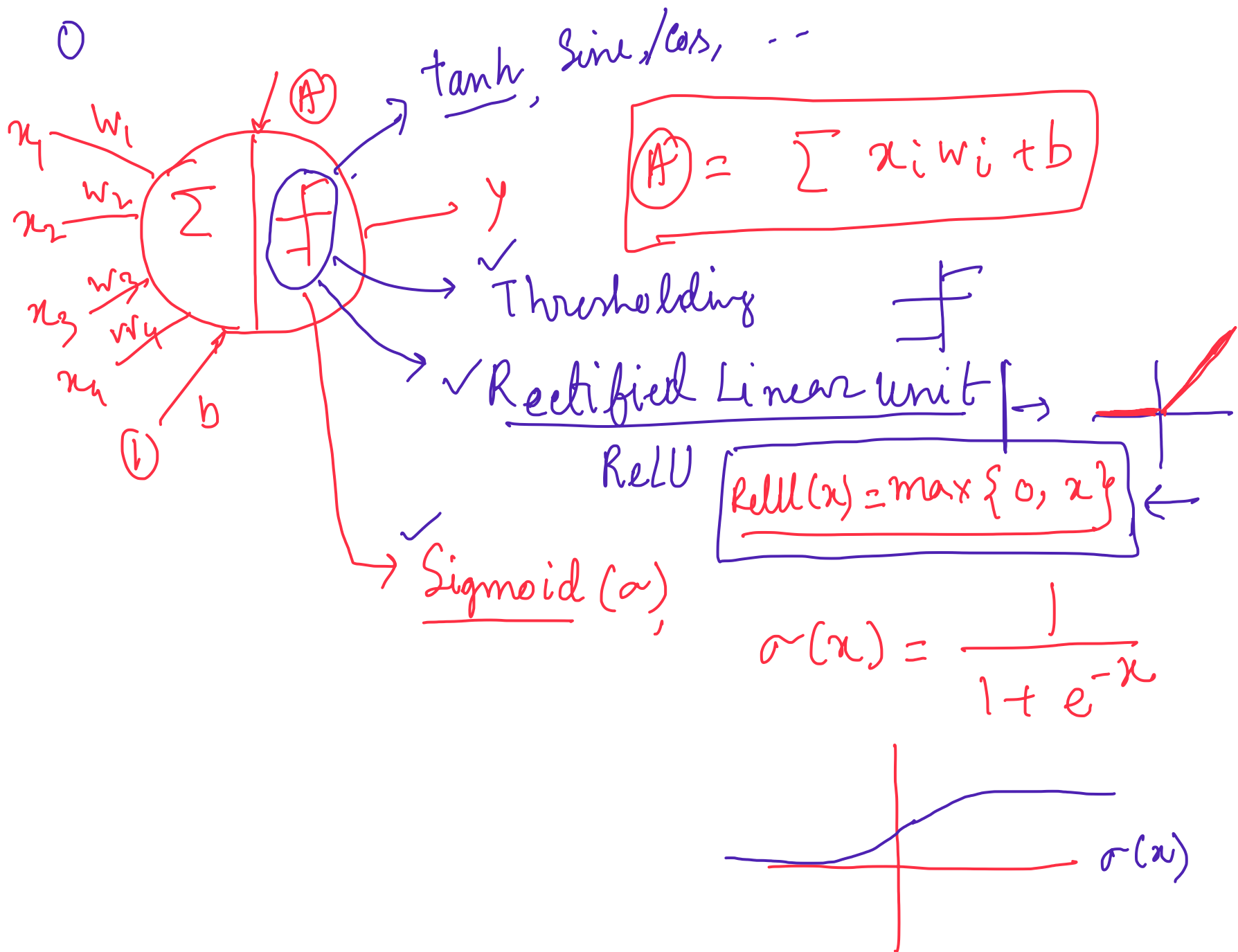
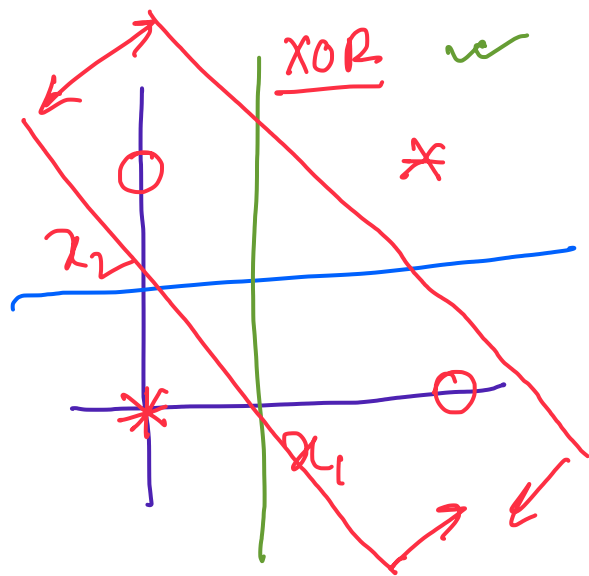
$W_1 x_1 + W_2 x_2 + b > 0$   
 $\rightarrow x_1 + x_2 - 1.5 > 0$

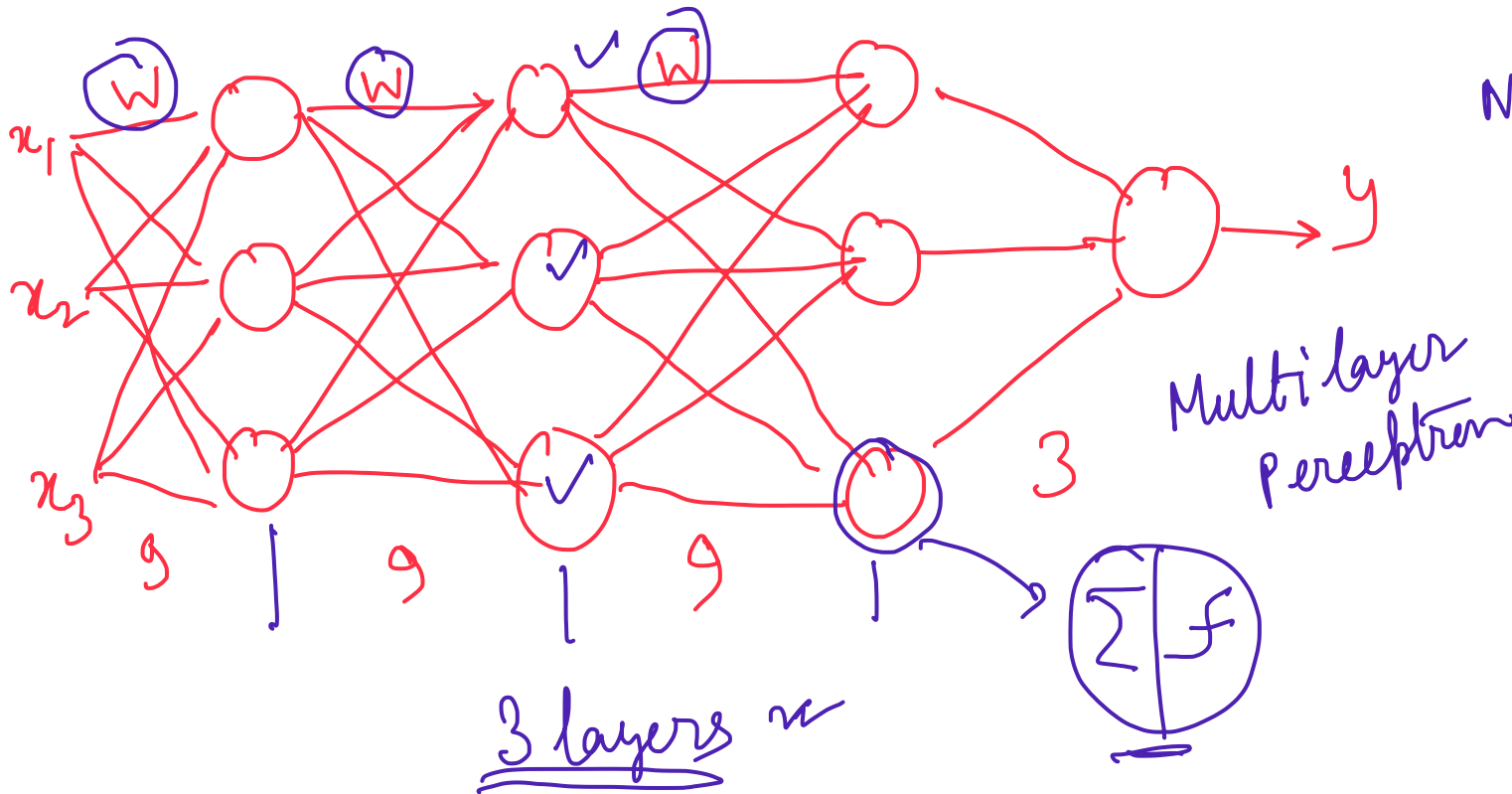


01  
10

# XOR

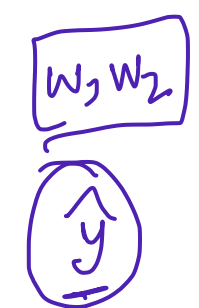
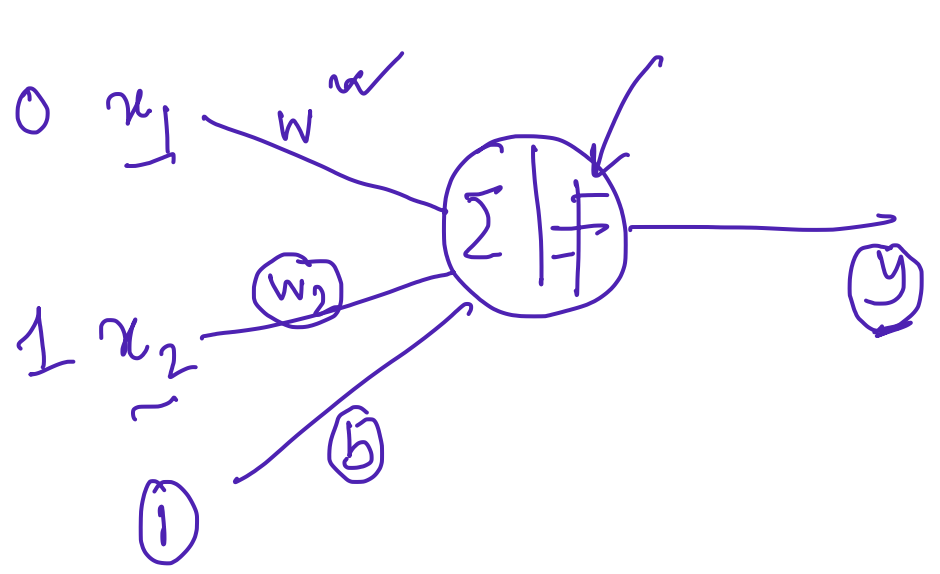
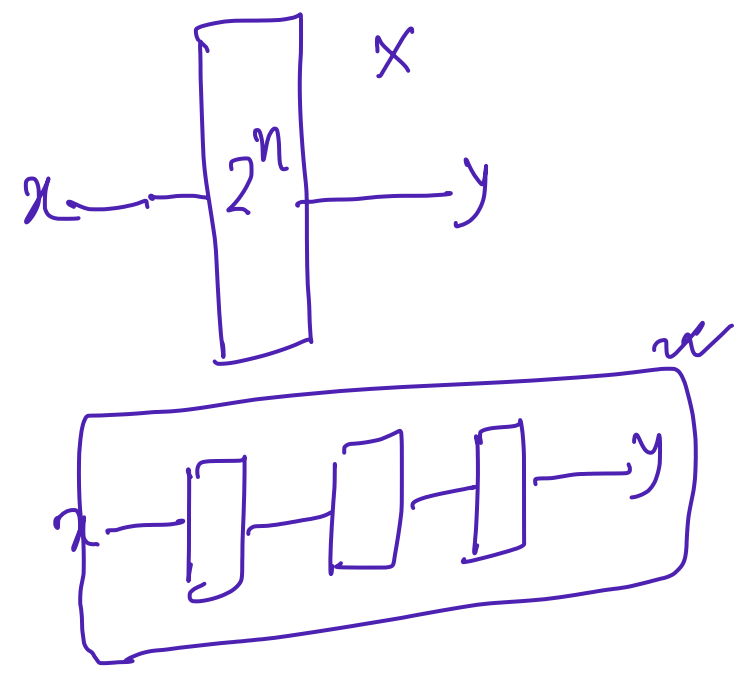
$x_1$	$x_2$	o/p
0	0	0
0	1	1
1	0	1
1	1	0





✓ # neurons in a layer

Neuron  $\Rightarrow$  Perceptron



$$E = \frac{1}{2} (y - \hat{y})^2$$

sub. predicted

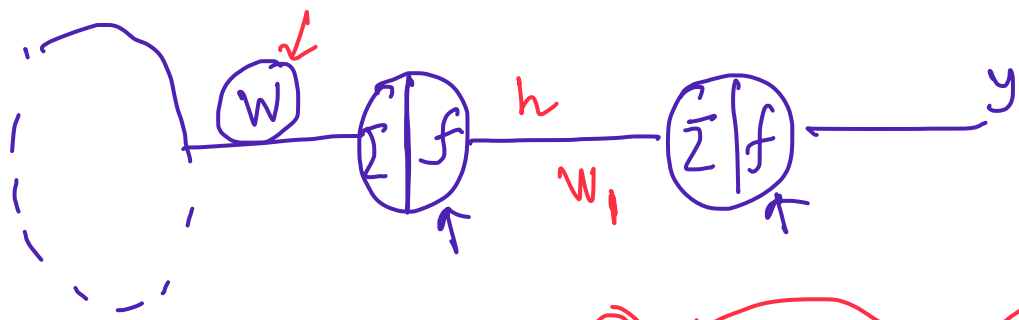
$$E = \frac{1}{2} (y - f(w, w_2))^2$$

MS E

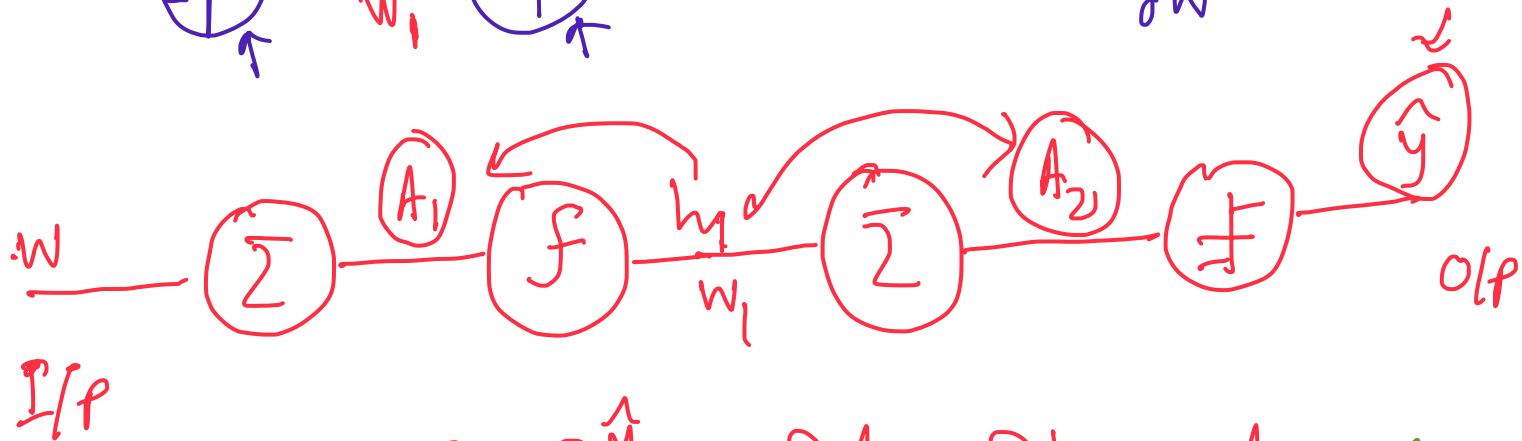
$$\frac{\partial E}{\partial w}$$

gradient descent

$$\frac{\partial E}{\partial w_2}$$



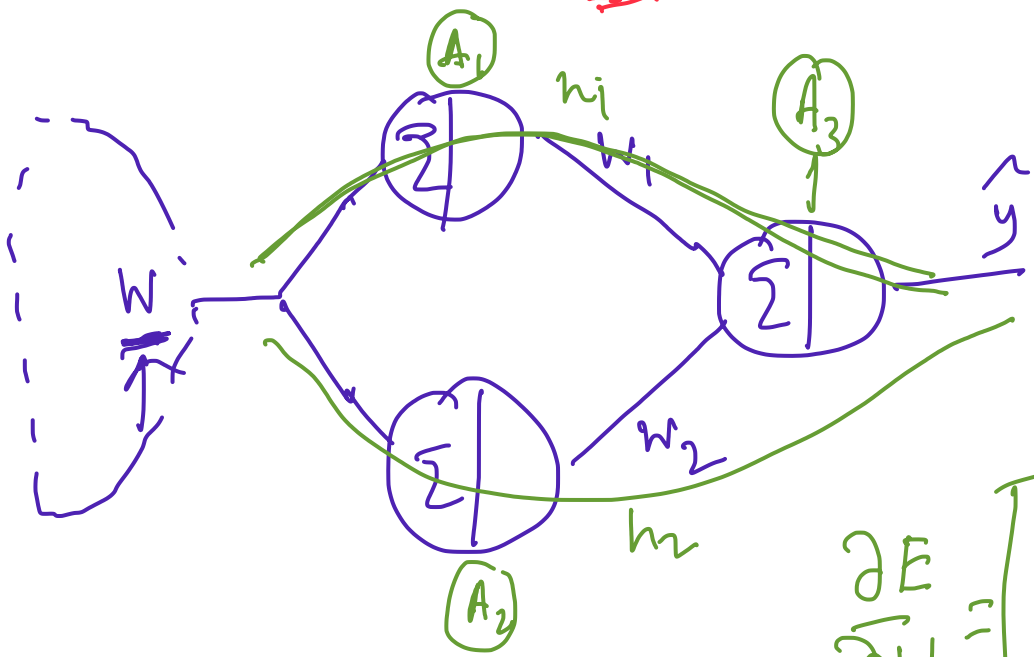
$$E, \frac{\partial E}{\partial w}$$



$$E \Rightarrow (y - \hat{y})^2$$

$$\frac{\partial E}{\partial w}$$

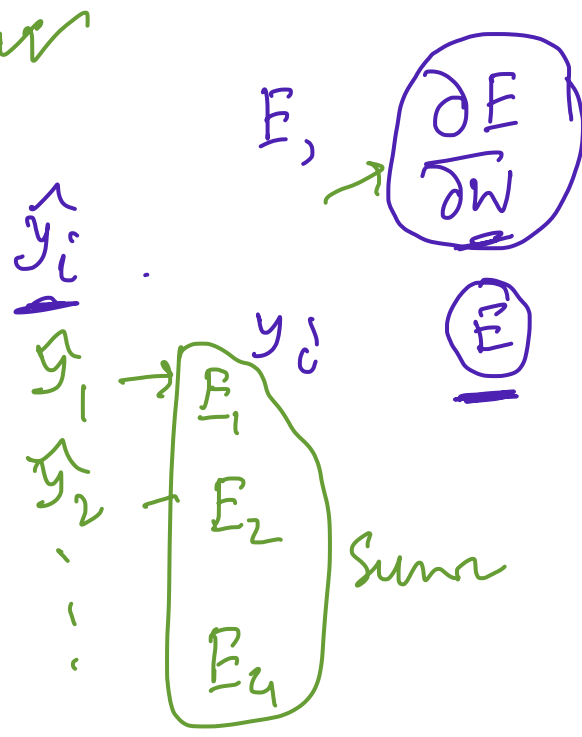
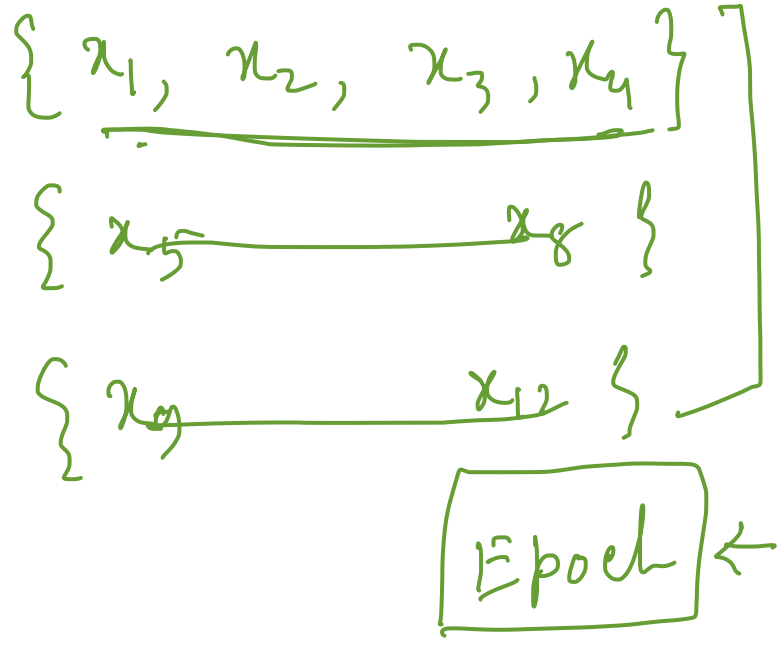
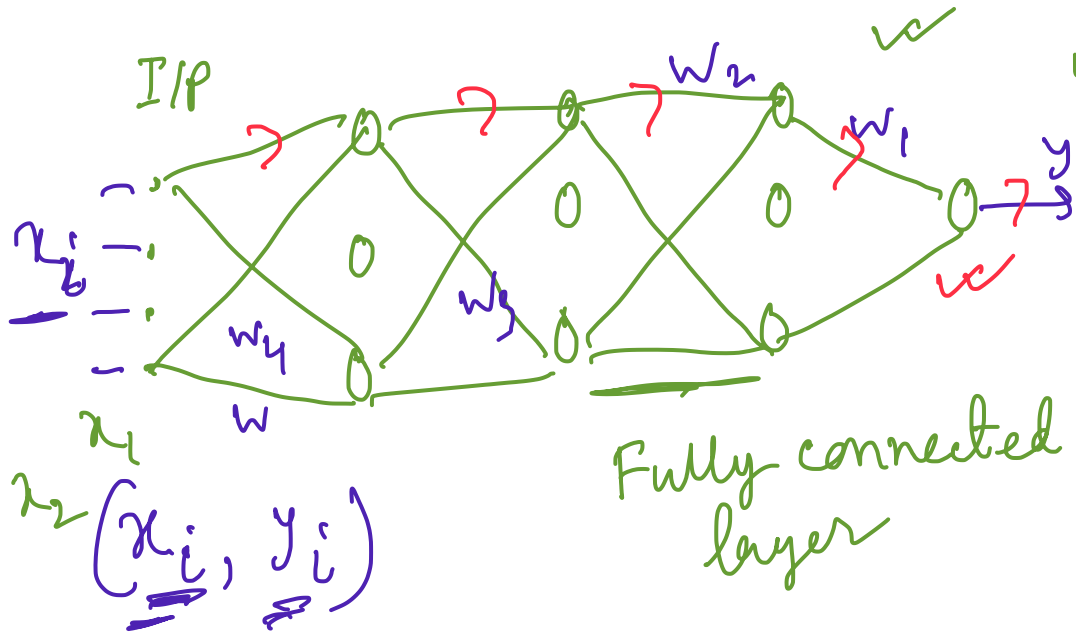
$$\frac{\partial E}{\partial w} = \frac{\partial E}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial A_2} \cdot \frac{\partial A_2}{\partial h_1} \cdot \frac{\partial h_1}{\partial A_1} \cdot \frac{\partial A_1}{\partial w}$$



$$E = (y - \hat{y})^2$$

$$\frac{\partial E}{\partial w}$$

$$\frac{\partial E}{\partial w} = \frac{\partial E}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial A_3} \cdot \frac{\partial A_3}{\partial h_1} \cdot \frac{\partial h_1}{\partial A_1} \cdot \frac{\partial A_1}{\partial w} + \frac{\partial E}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial A_3} \cdot \frac{\partial A_3}{\partial h_2} \cdot \frac{\partial h_2}{\partial A_2} \cdot \frac{\partial A_2}{\partial w}$$

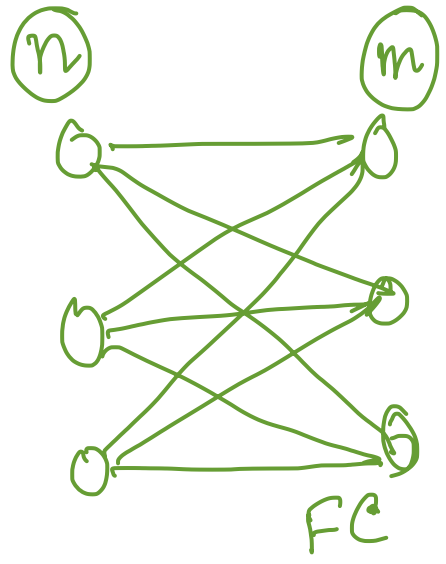


$$\frac{\partial E}{\partial w_1}, \frac{\partial E}{\partial w_2}, \dots$$

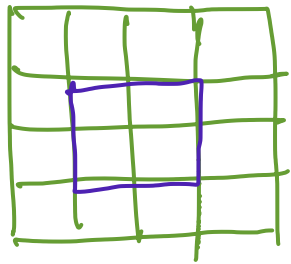
SGD

minibatch GD

All Examples  
 $\Downarrow$   
 GD

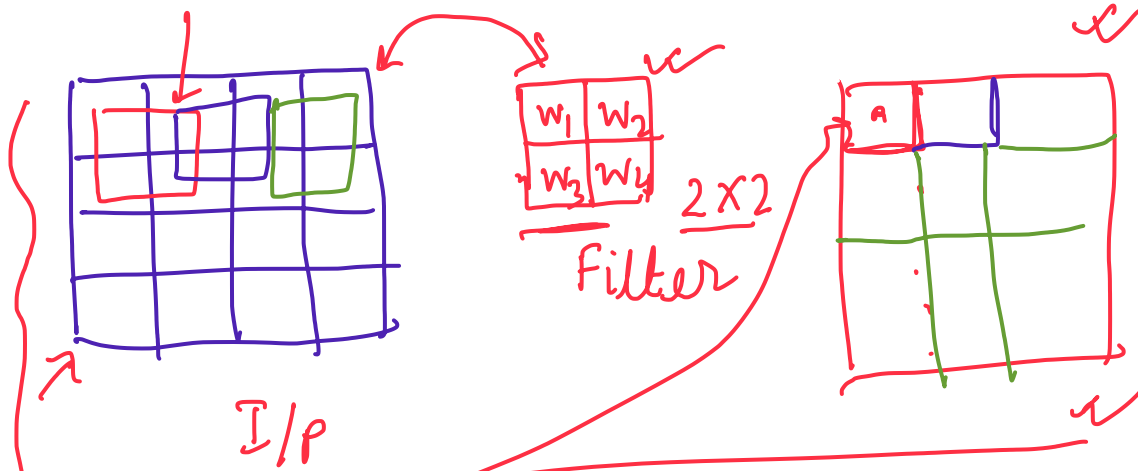


$nm$



16 x 4

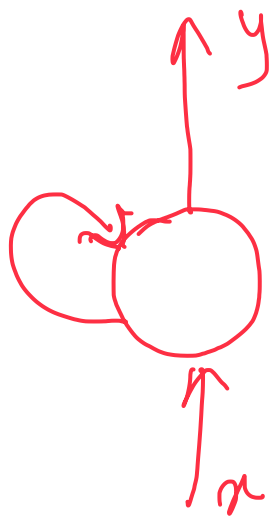
$x_{ij}$



Convolution

$$\begin{cases} x_{11}w_1 + x_{12}w_2 + x_{21}w_3 + x_{22}w_4 \\ x_{12}w_1 + x_{13}w_2 + x_{22}w_3 + x_{23}w_4 \end{cases}$$

16 I/P      (4)



Recurrent NN

Sequence

Natural language processing

Computer Vision

Convolutional NN

