Today's Agenda

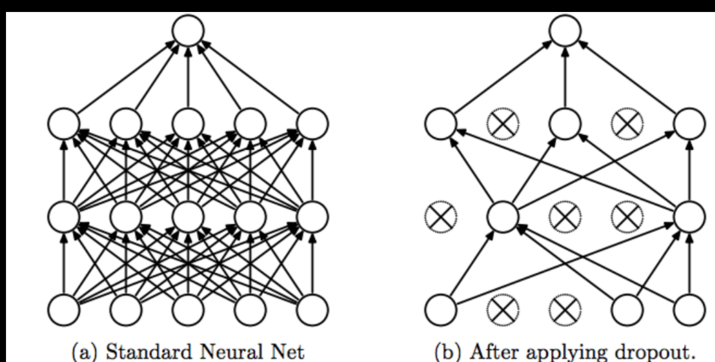
1) Dropout
2) Backprop

What is Dropout ?

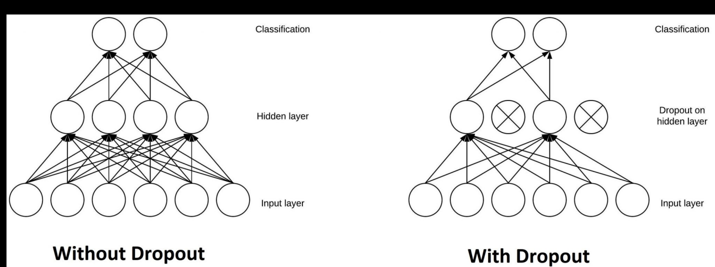
→ Dropout Regularization

↳ Problems of overfitting

Dropout layer



(a) Standard Neural Net (b) After applying dropout.



The Dropout layer randomly sets input units to 0 with a frequency of rate at each step during training time, which helps prevent overfitting.

```
keras.layers.Dropout(rate, noise_shape=None, seed=None, **kwargs)
```

Usage of Dropout

Rate = 0 - 1 in float

The general value are (0.1 - 0.5)

Hidden Layer = 256 nodes

Dropout = 0.5

Deactivated Neurons = 128

Info Loss

Model Generalization

https://keras.io/api/layers/regularization_layers/dropout/

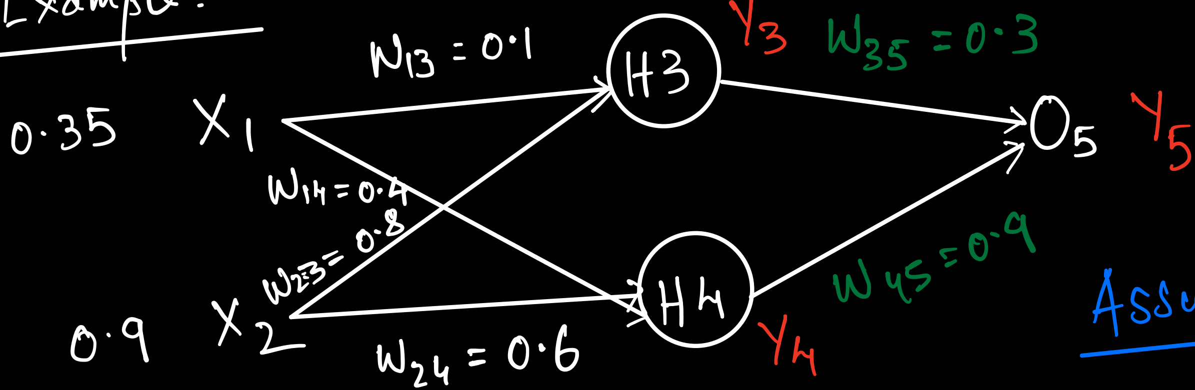
Backpropogation

Why ?

1) Update weights ⟶ Local minima

↳ less Error

Example :-



Assumptions

1) LR = 1
2) Act Output = 0.5
3) Act. Func
   ↳ Sigmoid

$z = wx + b$

$y = f(z)$

$$Y_3 = (W_{13} * X_1) + (W_{23} * X_2)$$
$$= (0.1 \times 0.35) + (0.8 \times 0.9)$$
$$= 0.755$$
$$a_3 = \text{Sigmoid}(0.755)$$
$$Y_3 = 0.68$$

$$Y_4 = (W_{14} \times X_1) + (W_{24} \times X_2)$$
$$Y_4 = (0.4 \times 0.35) + (0.6 \times 0.9)$$
$$Y_4 = 0.68$$
$$Y_4 = \text{Sigmoid}(.68) = 0.66$$

$$Y_5 = (W_{35} \times Y_3) + (W_{45} \times Y_4)$$
$$= (0.3 \times 0.68) + (0.9 \times 0.66)$$
$$= 0.798$$
$$= \text{Sigmoid}(0.798)$$
$$= \boxed{0.689 \approx 0.69} \rightarrow \text{Network Output}$$

$$\left\{ \begin{aligned} \text{ERROR} &= \text{Target} - Y_5 \\ &= 0.5 - 0.69 \end{aligned} \right\} (y - \hat{y})$$
$$= -0.19$$

Weight Updates

Each weight change by :-

$$\Delta w_{ij} = \underbrace{\eta}_{\text{LR}} \delta_j O_i$$

$$\delta_j = O_j(1 - O_j)(t_j - O_j)$$

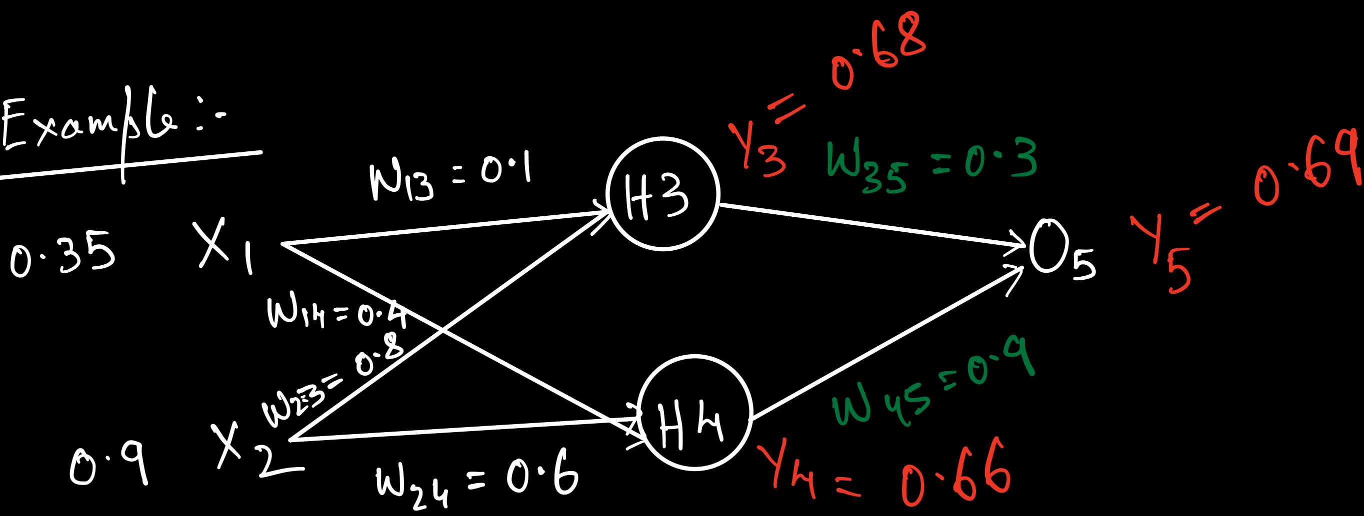
if j is a output unit

$$\delta_j = O_j(1 - O_j) \sum \delta_K W_{Kj}$$

if j is a hidden unit

δ_j = Error Measure

Example :-



$$\delta_5 = O_j(1 - O_j)(t_j - O_j)$$

$$= Y_5(1 - Y_5)(Y_t - Y_5)$$

$$= 0.69(1 - 0.69)(0.5 - 0.69)$$

$$= -0.0406$$

$$\delta_3 = O_j(1 - O_j)\sum \underline{\delta_k} W_{kj}$$

$$= Y_3(1 - Y_3)\, W_{35} * (-0.0406)$$

$$= 0.68(1 - 0.68) \times (0.3 \times -0.0406)$$

$$= -0.00265$$

$$\delta_4 = O_j(1 - O_j) \sum \delta_k W_{jk}$$

$$= 0.66(1 - 0.66)\, W_{45}(-0.0406)$$

$$= 0.66(1 - 0.66)(0.9 \times -.0406)$$

$$= -0.008$$

Weight Updation

$$\Delta w_{ij} = \eta\, \delta_j\, O_i$$

LR = 1

$$\Delta W_{45} = 1 \times -0.0406 \times 0.66$$

$$= -0.0267$$

$$\text{New } W_{45} = -0.0267 + 0.9$$

$$= 0.873$$

$$\Delta w_{14} = 1 \times -0.008 \times 0.35$$

$$= -.0028$$

$$\text{New } w_{14} = 0.4 + (-.0028)$$

$$= 0.3972$$

$\Delta_{35} = 1 \times -0.0406 \times 0.68$

$= -0.0276$

$New_{35} = 0.3 - .0276$

$= 0.272$

$W_{35} = 0.272$

$W_{14} = 0.3972$

$W_{45} = 0.873$

$x \rightarrow wx+b \rightarrow \hat{y} \rightarrow (\hat{y}-y)^2$

$\downarrow$

Loss

$\leftarrow \dfrac{\partial \, Loss}{\partial \hat{y}} \leftarrow 2(\hat{y} - y)$

$$\frac{\partial \, loss}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial w}$$

$$\frac{\partial \, loss}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial b}$$