NPTEL ONLINE CERTIFICATION COURSES

# Course Name: Deep Learning

Faculty Name: Prof. P. K. Biswas
Department : E & ECE, IIT Kharagpur

## Topic

Lecture 25: Back propagation Learning – Examples

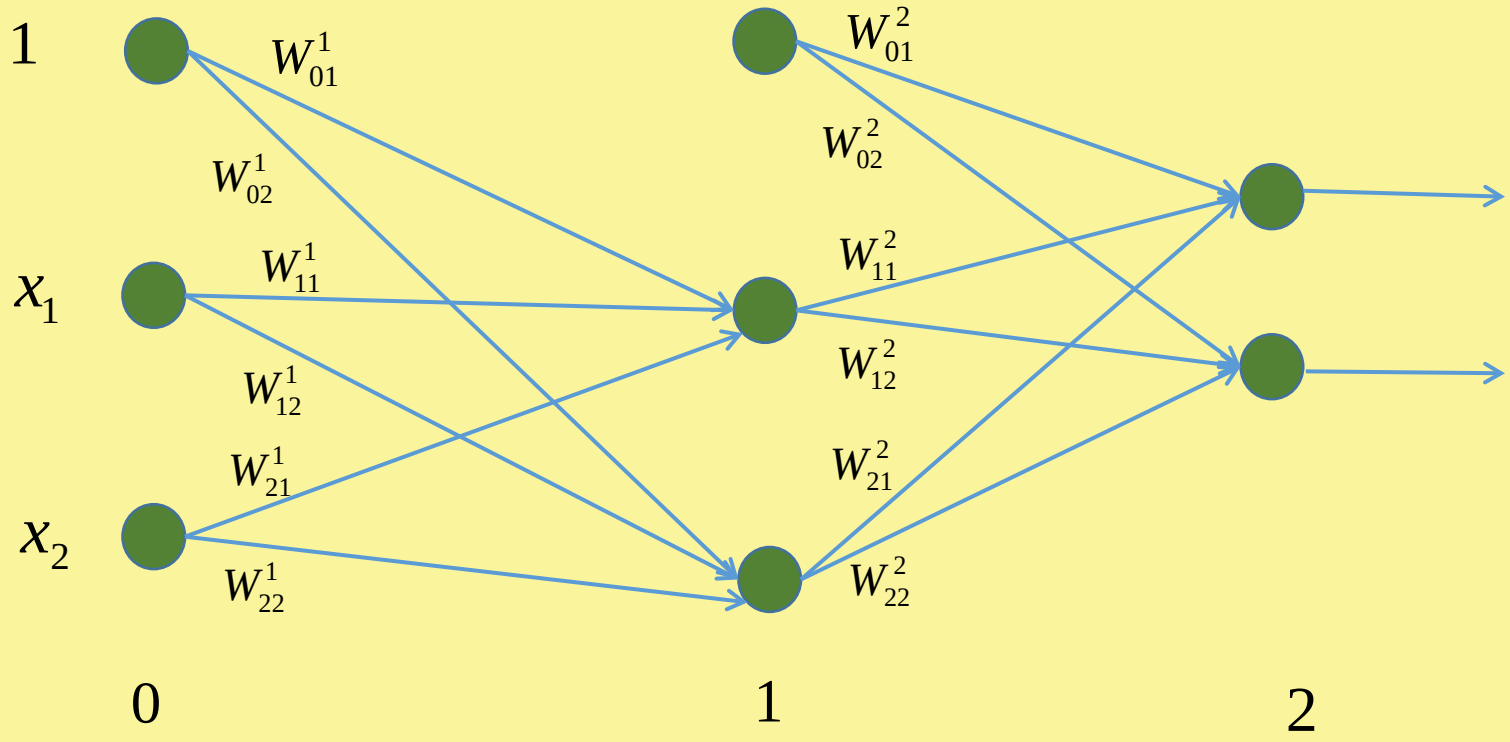
# CONCEPTS COVERED

Concepts Covered:

- Back Propagation Learning in MLP
- Different Loss Functions
- Back Propagation Learning - Example
- Back Propagation – Node Level

# Back Propagation Learning an Example

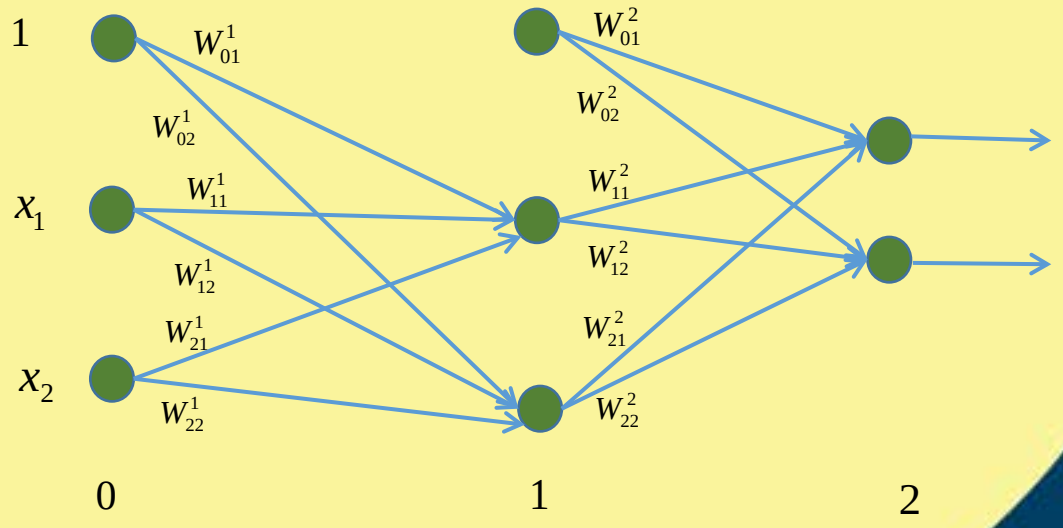
# Multilayer Perceptron



# Multilayer Perceptron

| $W_{01}^1$ | $W_{11}^1$ | $W_{21}^1$ |
|---|---|---|
| 0.5 | 1.5 | 0.8 |
| $W_{02}^1$ | $W_{12}^1$ | $W_{22}^1$ |
| 0.8 | 0.2 | -1.6 |

| $W_{01}^2$ | $W_{11}^2$ | $W_{21}^2$ |
|---|---|---|
| 0.9 | -1.7 | 1.6 |
| $W_{02}^2$ | $W_{12}^2$ | $W_{22}^2$ |
| 1.2 | 2.1 | -0.2 |

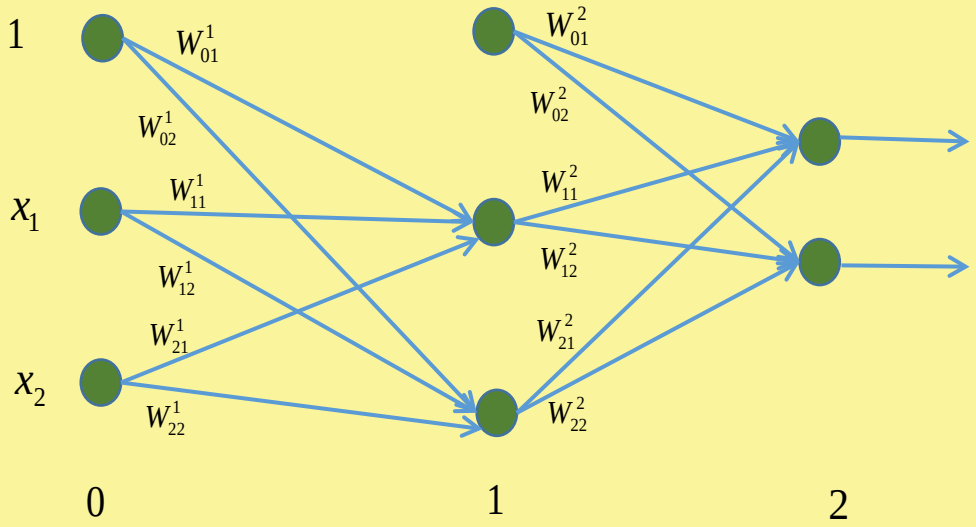



$$X = \begin{bmatrix} 0.7 \\ 1.2 \end{bmatrix} \text{ from category } 1 \Rightarrow t = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

# Feed Forward Pass

$W^1$ $\quad x_i^0 \quad \theta_j^1 = \sum W_{ij}^1 x_i^0 \quad x_j^1 = \dfrac{1}{1+e^{-\theta_j^1}}$

$$\begin{bmatrix} 0.5 & 1.5 & 0.8 \\ 0.8 & 0.2 & -1.6 \end{bmatrix} \begin{bmatrix} 1 \\ 0.7 \\ 1.2 \end{bmatrix} = \begin{bmatrix} 2.51 \\ -9.8 \end{bmatrix} \Rightarrow \begin{bmatrix} 0.92 \\ 0.27 \end{bmatrix}$$

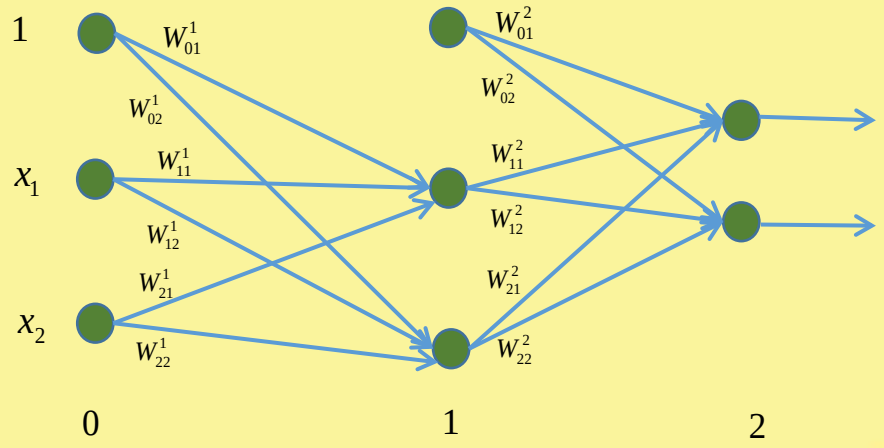



$W^2$ $\quad x_i^1 \quad \theta_j^2 = \sum_i W_{ij}^2 x_i^1 \quad x_j^2 = \dfrac{1}{1+e^{-\theta_j^2}}$

$$\begin{bmatrix} 0.9 & -1.7 & 1.6 \\ 1.2 & 2.1 & -1.-0.26 \end{bmatrix} \begin{bmatrix} 1 \\ 0.92 \\ 0.27 \end{bmatrix} = \begin{bmatrix} -0.232 \\ 3.057 \end{bmatrix} \Rightarrow \begin{bmatrix} 0.44 \\ 0.95 \end{bmatrix}$$

# Back Propagation Learning:- Output Layer

$$E = \frac{1}{2}\sum_{j=1}^{2}\left(x_j^2 - t_j\right)^2 \qquad x_j^2 = \frac{1}{1+e^{-\theta_j^2}} \qquad \theta_j^2 = \sum_{i=0}^{2} W_{ij}^2 x_i^1$$




$$\frac{\partial E}{\partial W_{ij}^2} = \frac{\partial E}{\partial x_j^2}\cdot\frac{\partial x_j^2}{\partial \theta_j^2}\cdot\frac{\partial \theta_j^2}{\partial W_{ij}^2} = (x_j^2 - t_j)x_j^2(1-x_j^2)x_i^1$$

We set $\delta_j^2 = x_j^2(1-x_j^2)(x_j^2 - t_j) \Rightarrow \frac{\partial E}{\partial W_{ij}^2} = \delta_j^2 x_i^1$

$$W_{ij}^2 \leftarrow W_{ij}^2 - \eta\frac{\partial E}{\partial W_{ij}^2}$$

NPTEL ONLINE CERTIFICATION COURSES

*Thank you*