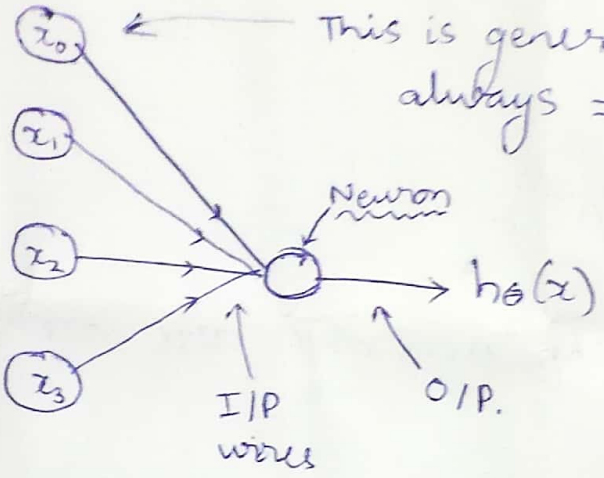
Neural Networks (Week 4)

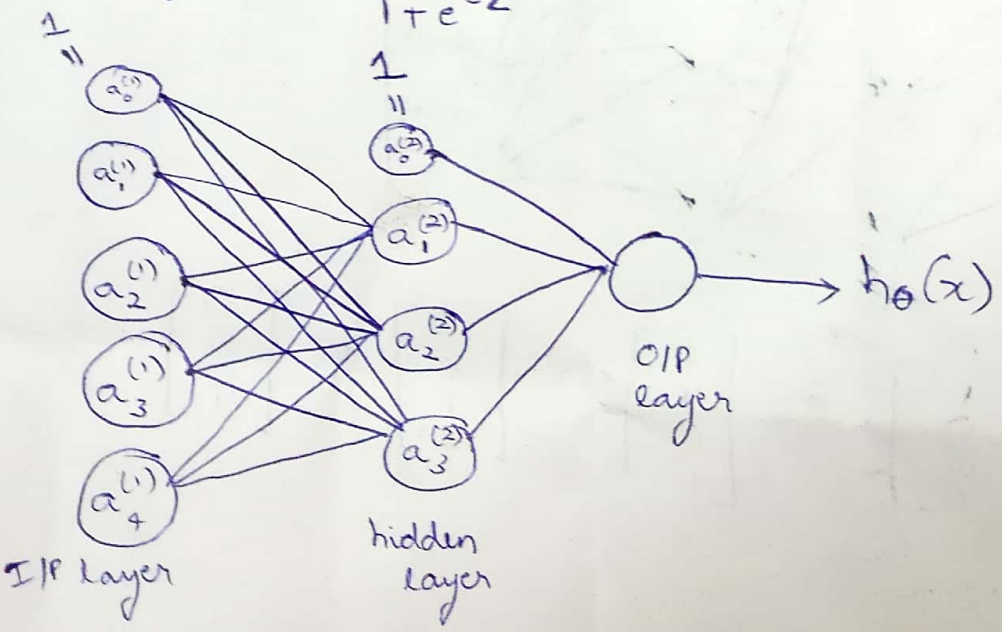
- We consider each neuron as a logistic unit.



$$x = \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix} \qquad \theta = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \\ \theta_3 \end{bmatrix} \qquad h_\theta(x) = \frac{1}{1+e^{-\theta^T x}}$$

- We use sigmoid activn fn. for neural nets.

$$g(z) = \frac{1}{1+e^{-z}}$$



θ_j : matrix of wts controlling fn. mapping from j to j+1.

$\theta_j : (n_{j+1} \times (n_j + 1))$

~~$A_{j+1} = g(\theta_j A_j)$~~

$$A^{(j+1)} = g(\theta^{(j)} A^{(j)})$$

* But this isn't how we write in code. (That'll come later)

∴ For our case,

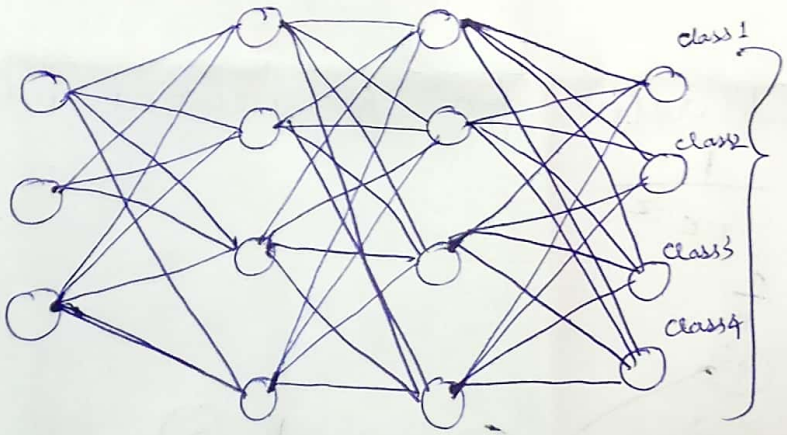
$$A^{(2)} = g(\theta^{(1)} A^{(1)})$$

$$A^{(3)} = h_\theta(x) = g(\theta^{(2)} A^{(2)})$$

* Architecture: How diff neurons are connected together.

Multi-class classifn using Neural nets

- Suppose there are 4 classes.



Output $h_\theta(x) \in \mathbb{R}^4$.

We want:

$$h_\theta(x) = \begin{bmatrix}1\\0\\0\\0\end{bmatrix} \text{ or } \begin{bmatrix}0\\1\\0\\0\end{bmatrix} \text{ or } \begin{bmatrix}0\\0\\1\\0\end{bmatrix} \text{ or } \begin{bmatrix}0\\0\\0\\1\end{bmatrix}$$

Writing in code:

```
X = [ones(m,1) X];
z2 = X * Theta1';
a2 = sigmoid(z2)
```

$a2 = \bullet X \times Theta1'$ | $a3 = a2 \times Theta2'$