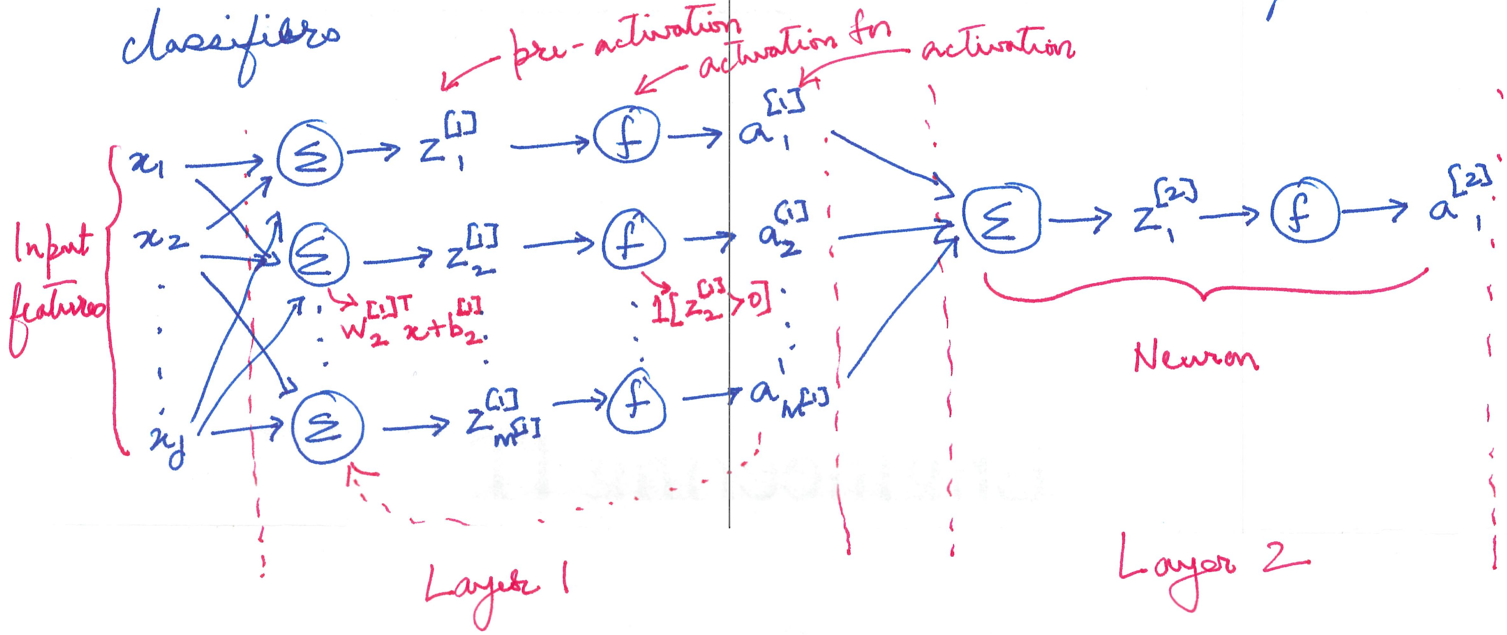
CS 307 Lecture 21

Neural Nets: Add features using step functions / linear classifiers

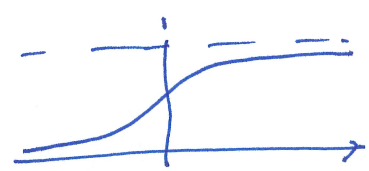


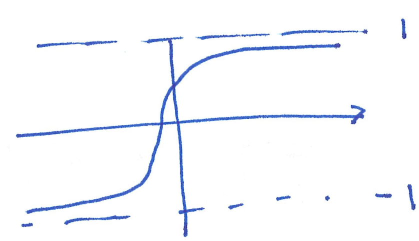
Full connected network.

Feed forward network / Recurrent Nework (when edges go back)

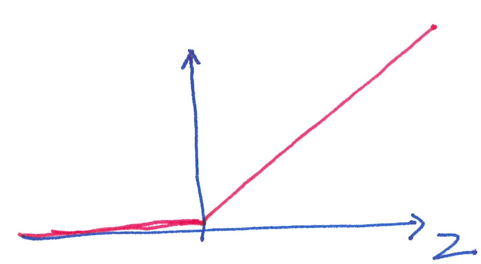
Activation functions:

- Step function : $1[z > 0]$
- Sigmoid : $\sigma(z) = \dfrac{1}{1+e^{-z}}$
- tanh : $\tanh(z) = \dfrac{e^{z} - e^{-z}}{e^{z} + e^{-z}}$





- RELU : $RELU(z) = \max(0, z)$



Input to layer i:
$$a^{[i-1]} = \left.\begin{bmatrix} a_1^{[i-1]} \\ \vdots \\ a_{m^{[i-1]}}^{[i-1]} \end{bmatrix}\right\} m^{[i-1]}$$

Input to layer 1:
$$a^{[0]} = x = \begin{bmatrix} x_1 \\ \vdots \\ x_d \end{bmatrix} \qquad m^{[0]} = d.$$

$$\left.\begin{aligned} z_j^{[i]} &= w_j^{[i]T} a^{[i-1]} + b_j^{[i]} \\ a_j^{[i]} &= f_j^{[i]}(z_j^{[i]}) \end{aligned}\right\} j \in \{1, 2 \ldots m^{[i]}\}$$

$$z^{[i]} = W^{[i]} a^{[i-1]} + b^{[i]}$$
$$a^{[i]} = f^{[i]}(z^{[i]})$$

$$W^{[i]} = \begin{bmatrix} \text{---} & w_1^{[i]T} & \text{---} \\ \text{---} & w_2^{[i]T} & \text{---} \\ & \vdots & \\ \text{---} & w_{m^{[i]}}^{[i]T} & \text{---} \end{bmatrix} \qquad b^{[i]} = \begin{bmatrix} b_1^{[i]} \\ \vdots \\ b_{m^{[i]}}^{[i]} \end{bmatrix}$$

$$m^{[i]} \times m^{[i-1]} \qquad\qquad m^{[i]} \times 1$$

Output:

$$\begin{aligned} z^{[1]} &= W^{[1]} x + b^{[1]} \\ a^{[1]} &= f^{[1]}(z^{[1]}) \\ z^{[2]} &= W^{[2]} a^{[1]} + b^{[2]} \\ a^{[2]} &= f^{[2]}(z^{[2]}) \\ &\vdots \\ a^{[k]} &= f^{[k]}(z^{[k]}) \end{aligned}$$

Back Propagation: $f^{[1]} = \text{RELU}$

$f^{[2]} = \text{Identity} \; : \; a^{[2]} = z^{[2]}.$

$$J[W, b] = \frac{1}{n} \sum_{l=1}^{n} \Big(\underbrace{h(x^{(l)})}_{} - y^{(l)} \Big)^2 \longrightarrow \text{highly non-convex.}$$

Linear Reg $\quad \theta^T x^{(l)}$

$$J(W;b) = (h(x) - y)^2 = (a_1^{[2]} - y)^2$$

$$W_1^{[2]} = \begin{bmatrix} (W_1^{[2]})_1 \\ \vdots \\ (W_1^{[2]})_{m^{[1]}} \end{bmatrix}$$

$$\frac{\partial J}{\partial (W_1^{[2]})_j} = \frac{\partial J}{\partial z_1^{[2]}} \cdot \frac{\partial z_1^{[2]}}{\partial (W_1^{[2]})_j}$$

$$= \frac{\partial J}{\partial a_1^{[2]}} \cdot \frac{\partial a_1^{[2]}}{\partial z_1^{[2]}} \cdot \frac{\partial z_1^{[2]}}{\partial (W_1^{[2]})_j}$$

$$= 2(a_1^{[2]} - y) \cdot 1 \cdot a_j^{[1]}$$

$$\rightarrow z_1^{[2]} = (W_1^{[2]})_1 a_1^{[1]} + (W_1^{[2]})_2 a_2^{[1]} + \cdots (W_1^{[2]})_{m^{[1]}} a_{m^{[1]}}^{[1]} + b^{[2]}$$

$$a_1^{[2]} = f^{[2]}(z_1^{[2]})$$

$$\frac{\partial a_1^{[2]}}{\partial z_1^{[2]}} = (f^{[2]})'(z_1^{[2]}) = 1$$

$$\frac{\partial J}{\partial (W_k^{[1]})_j} = \frac{\partial J}{\partial z^{[2]}} \cdot \frac{\partial z^{[2]}}{\partial (W_k^{[1]})_j}$$

$$= \frac{\partial z^{[2]}}{\partial a_k^{[1]}} \cdot \frac{\partial a_k^{[1]}}{\partial (W_k^{[1]})_j}$$

$$= \frac{\partial a_k^{[1]}}{\partial z_k^{[1]}} \cdot \frac{\partial z_k^{[1]}}{\partial (W_k^{[1]})_j}$$