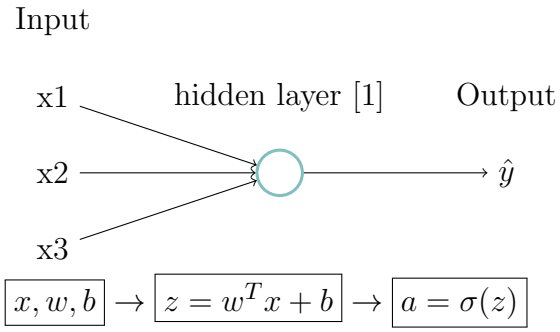
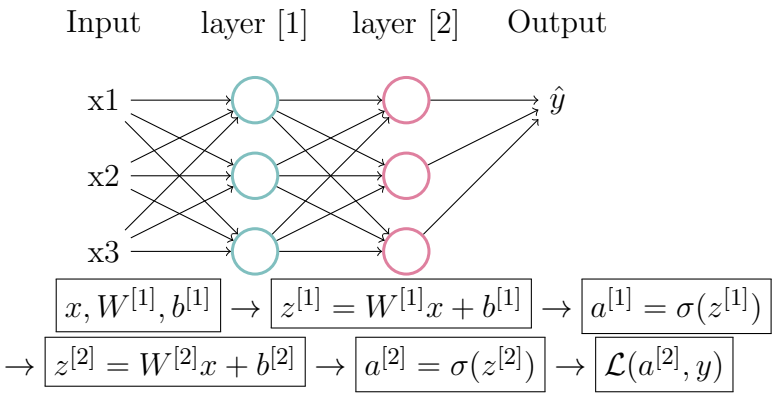
1 Week3

Shallow Neural Network.

1.1 Neural Network Overview

What is a Neural Network?

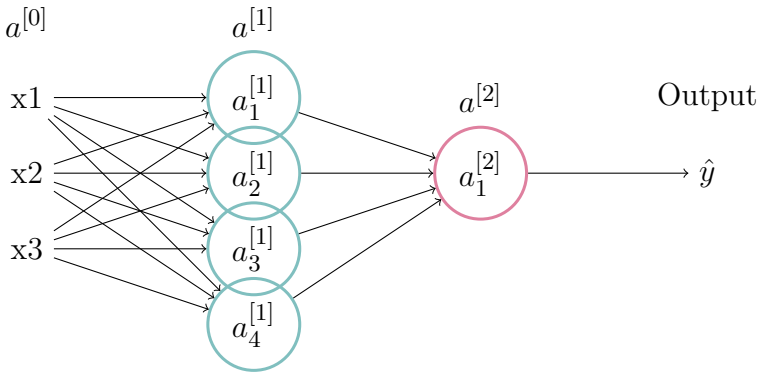




1.2 Neural Network Representations

Values of the input features (activation): $X = a^{[0]}$

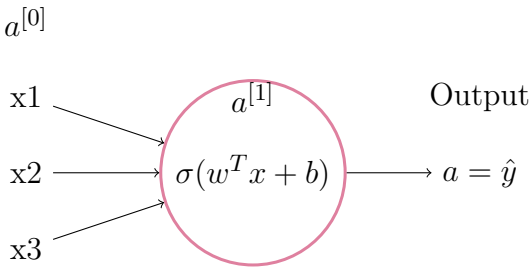
The following is the 2-Layer Neural Network:



$w^{[1]}$ in $(4, 3)$, $b^{[1]}$ in $(4, 1)$

$w^{[2]}$ in $(1, 4)$, $b^{[2]}$ in $(1, 1)$

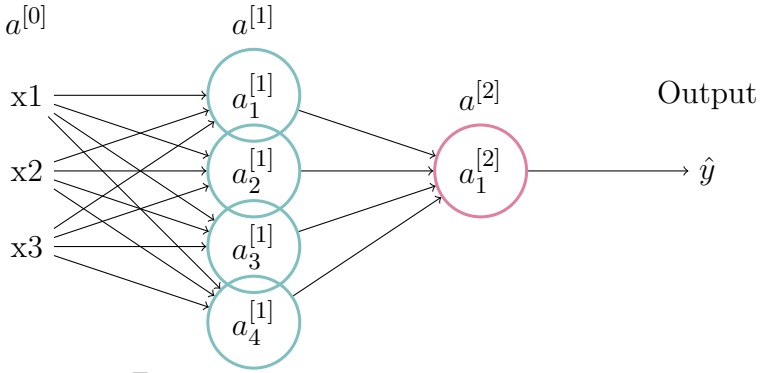
1.3 Computing Neural Network Output



Each circle(node) represents 2 steps of calculation:

$$\boxed{z = w^T x + b}$$

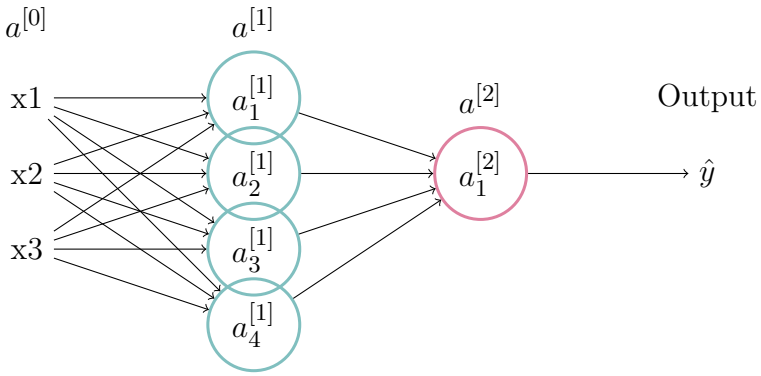
$$\boxed{a = \sigma(z)}$$



$$z_1^{[1]} = {w_1^{[1]}}^T x + b_1^{[1]} \rightarrow a_1^{[1]} = \sigma(z_1^{[1]})$$

$$z_2^{[1]} = {w_2^{[1]}}^T x + b_2^{[1]} \rightarrow a_2^{[1]} = \sigma(z_2^{[1]})$$

$$a^{[1]} = \begin{bmatrix} a_1^{[1]} \\ a_2^{[1]} \\ a_3^{[1]} \\ a_4^{[1]} \end{bmatrix} = \begin{bmatrix} \sigma(z_1^{[1]}) \\ \sigma(z_2^{[1]}) \\ \sigma(z_3^{[1]}) \\ \sigma(z_4^{[1]}) \end{bmatrix} = \begin{bmatrix} \sigma({w_1^{[1]}}^T x + b_1^{[1]}) \\ \sigma({w_2^{[1]}}^T x + b_2^{[1]}) \\ \sigma({w_3^{[1]}}^T x + b_3^{[1]}) \\ \sigma({w_4^{[1]}}^T x + b_4^{[1]}) \end{bmatrix} = \sigma[\begin{bmatrix} \dots & {w_1^{[1]}}^T & \dots \\ \dots & {w_2^{[1]}}^T & \dots \\ \dots & {w_3^{[1]}}^T & \dots \\ \dots & {w_4^{[1]}}^T & \dots \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} b_1^{[1]} \\ b_2^{[1]} \\ b_3^{[1]} \\ b_4^{[1]} \end{bmatrix}]$$



Given input x:

$z^{[1]} = W^{[1]}a^{[0]} + b^{[1]}$, where $_{(4,1)=(4,3)(3,1)+(4,1)}$

$a^{[1]} = \sigma(z^{[1]})$

$z^{[2]} = W^{[2]}a^{[1]} + b^{[2]}$, where $_{(1,1)=(1,4)(4,1)+(1,1)}$

$a^{[2]} = \sigma(z^{[2]})$

1.4 Vectorizing Across Multiple Examples

for $i = 1$ to m:

$$z^{[1](i)} = W^{[1]}x^{(i)} + b^{[1]}$$

$$a^{[1](i)} = \sigma(z^{[1](i)})$$

$$z^{[2](i)} = W^{[2]}a^{[1](i)} + b^{[2]}$$

$$a^{[2](i)} = \sigma(z^{[2](i)})$$

1.5