

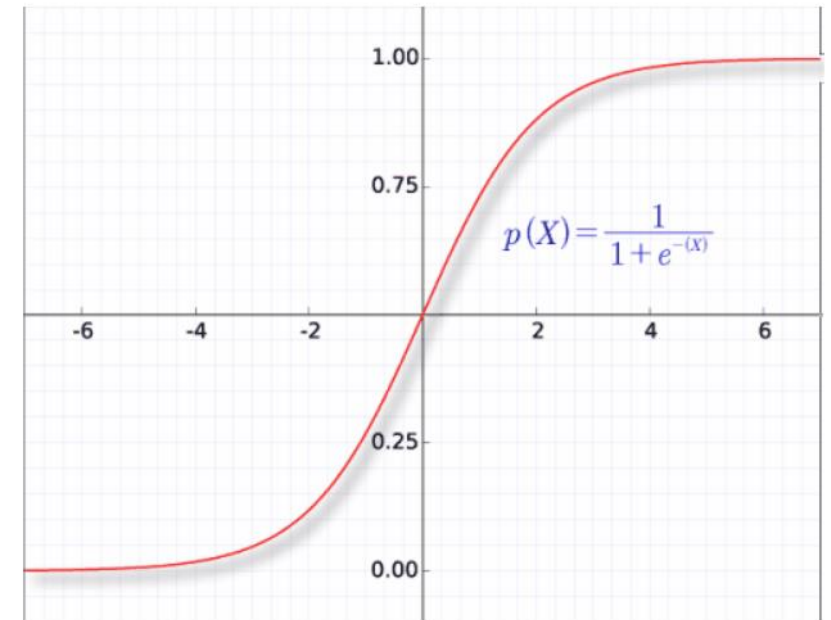
Suppose we have a dataset(classification) with 4 features **and only one label**

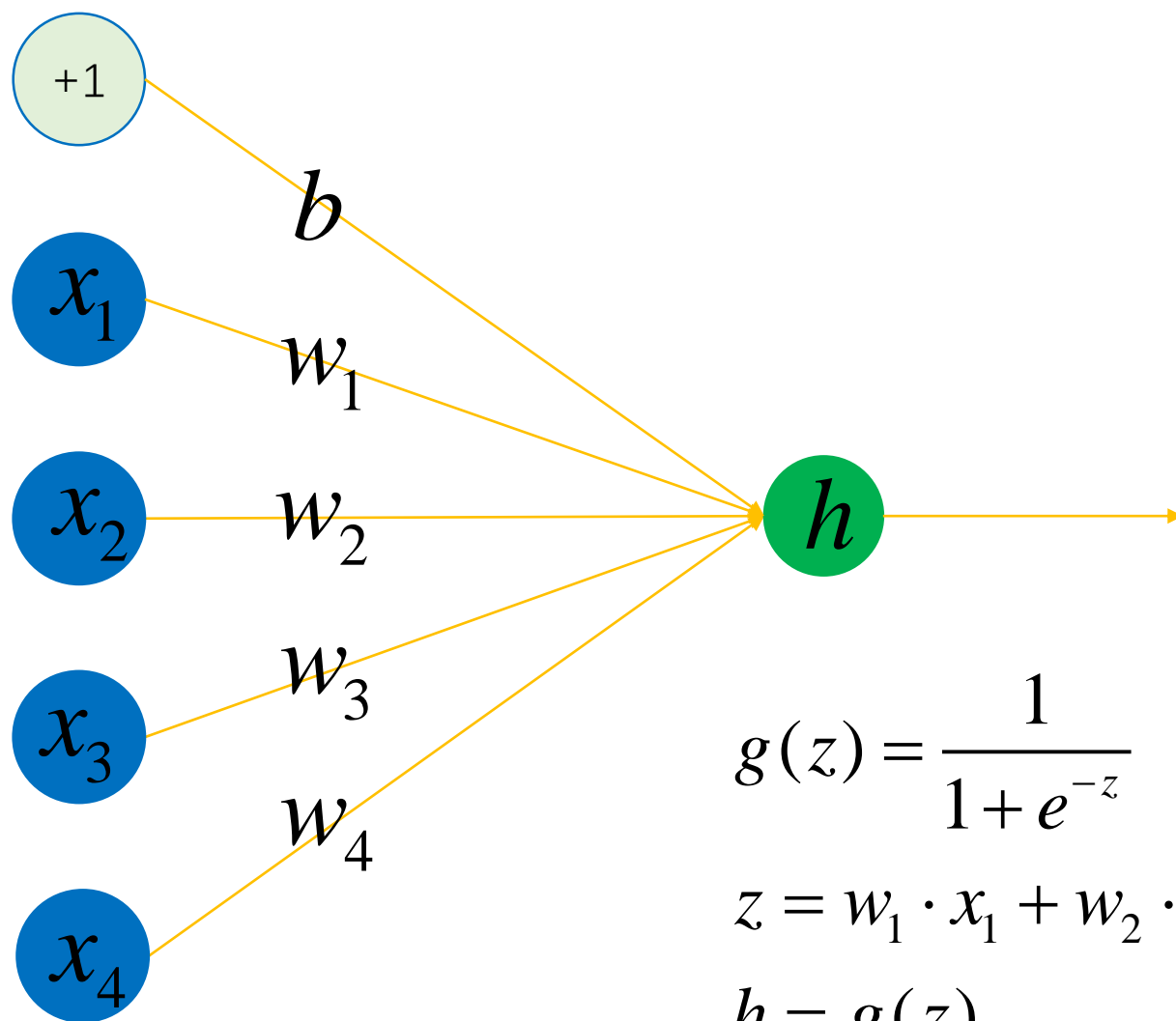
Logistic Regression Model

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

$$z = w_1 \cdot x_1 + w_2 \cdot x_2 + w_3 \cdot x_3 + w_4 \cdot x_4 + 1 \cdot b$$

$$h = g(z)$$



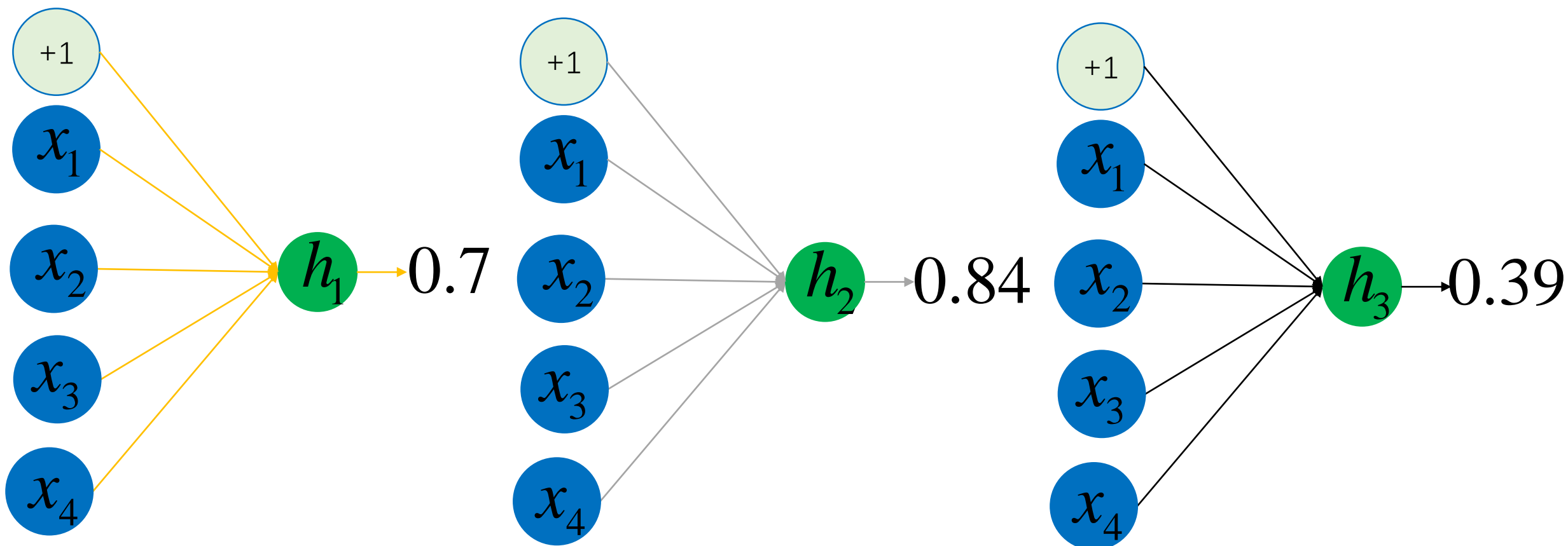


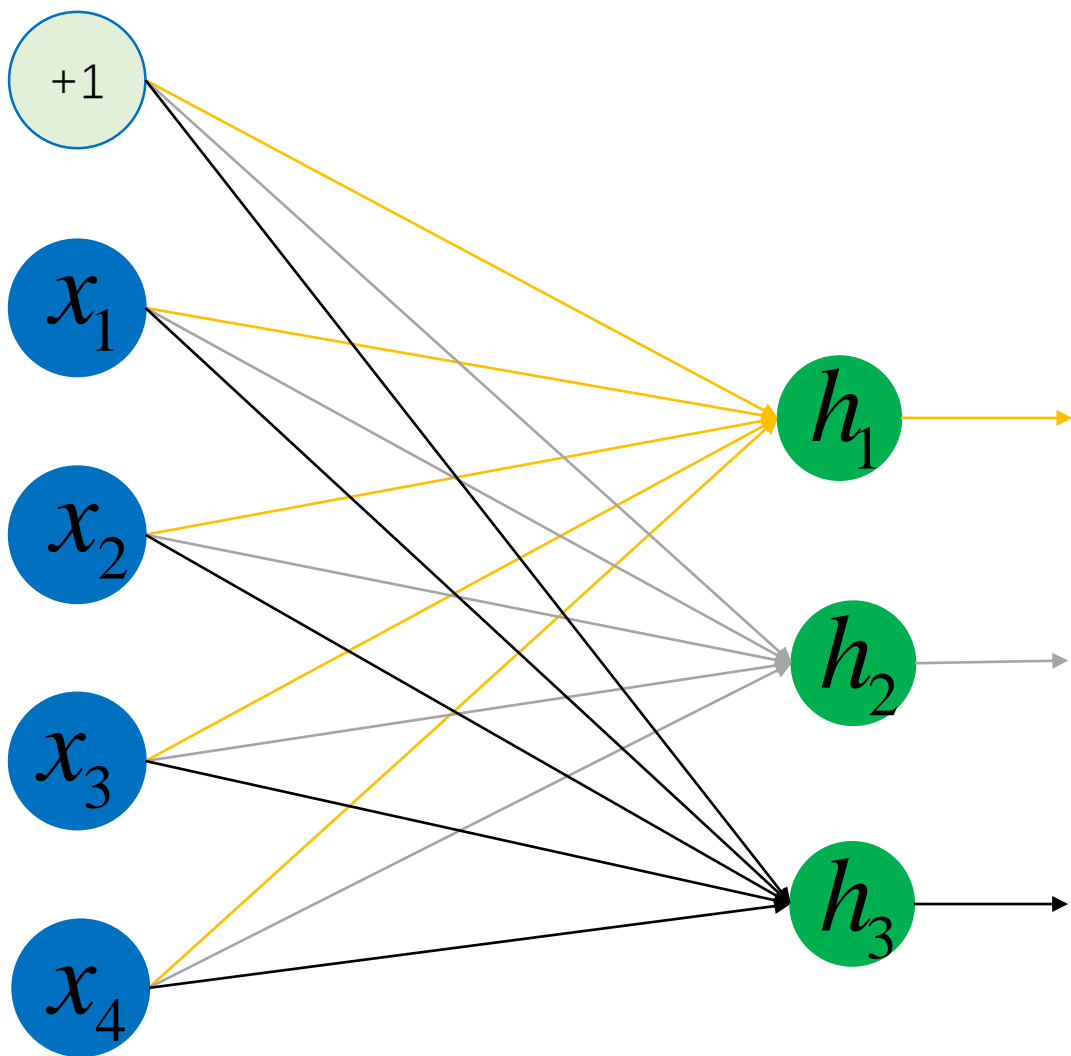
If labels more than one?

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

$$z = w_1 \cdot x_1 + w_2 \cdot x_2 + w_3 \cdot x_3 + w_4 \cdot x_4 + 1 \cdot b$$

$$h = g(z)$$





$X : [m, 4]$

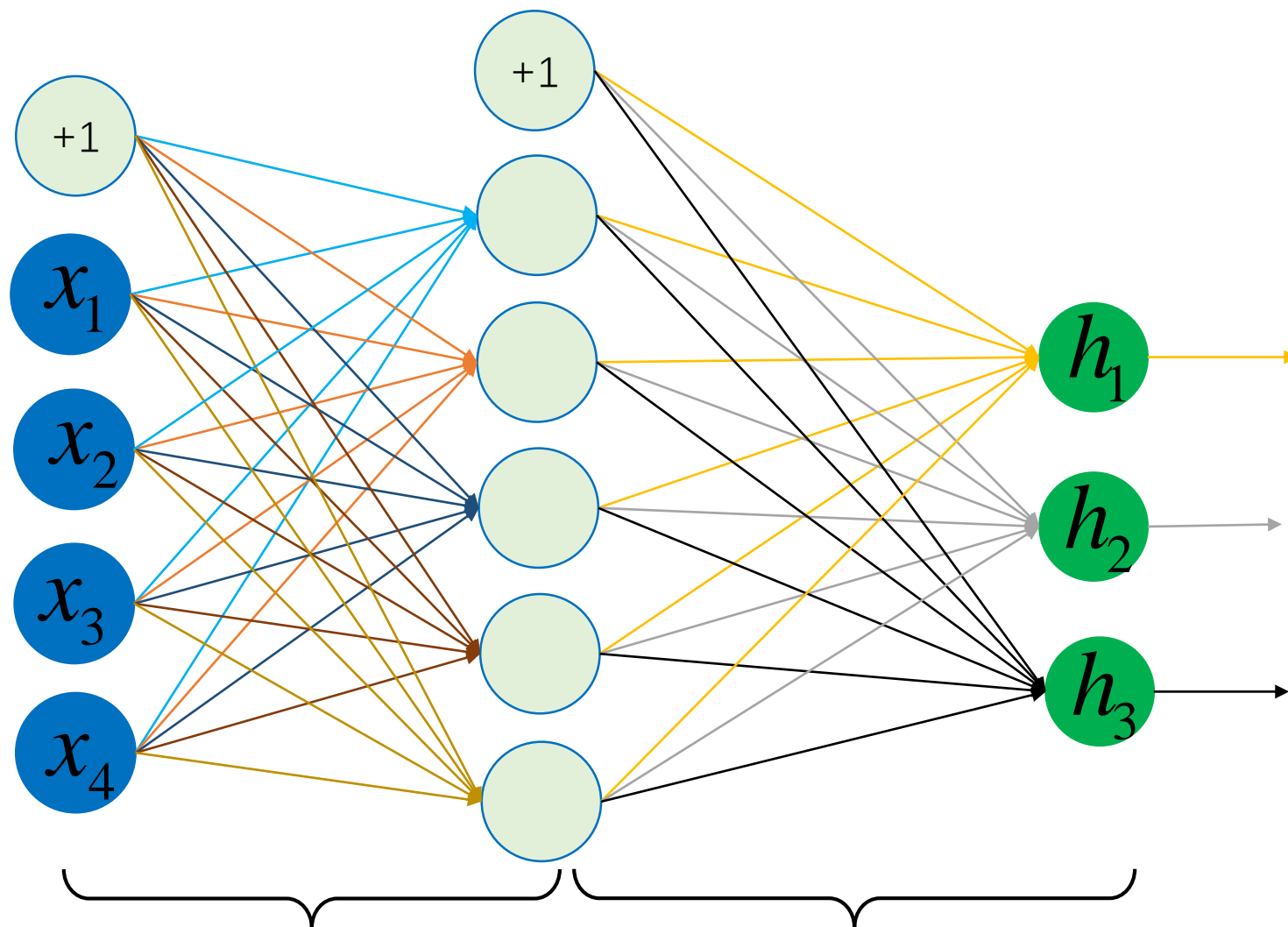
$W : [4, 3]$

$h : [m, 3]$

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

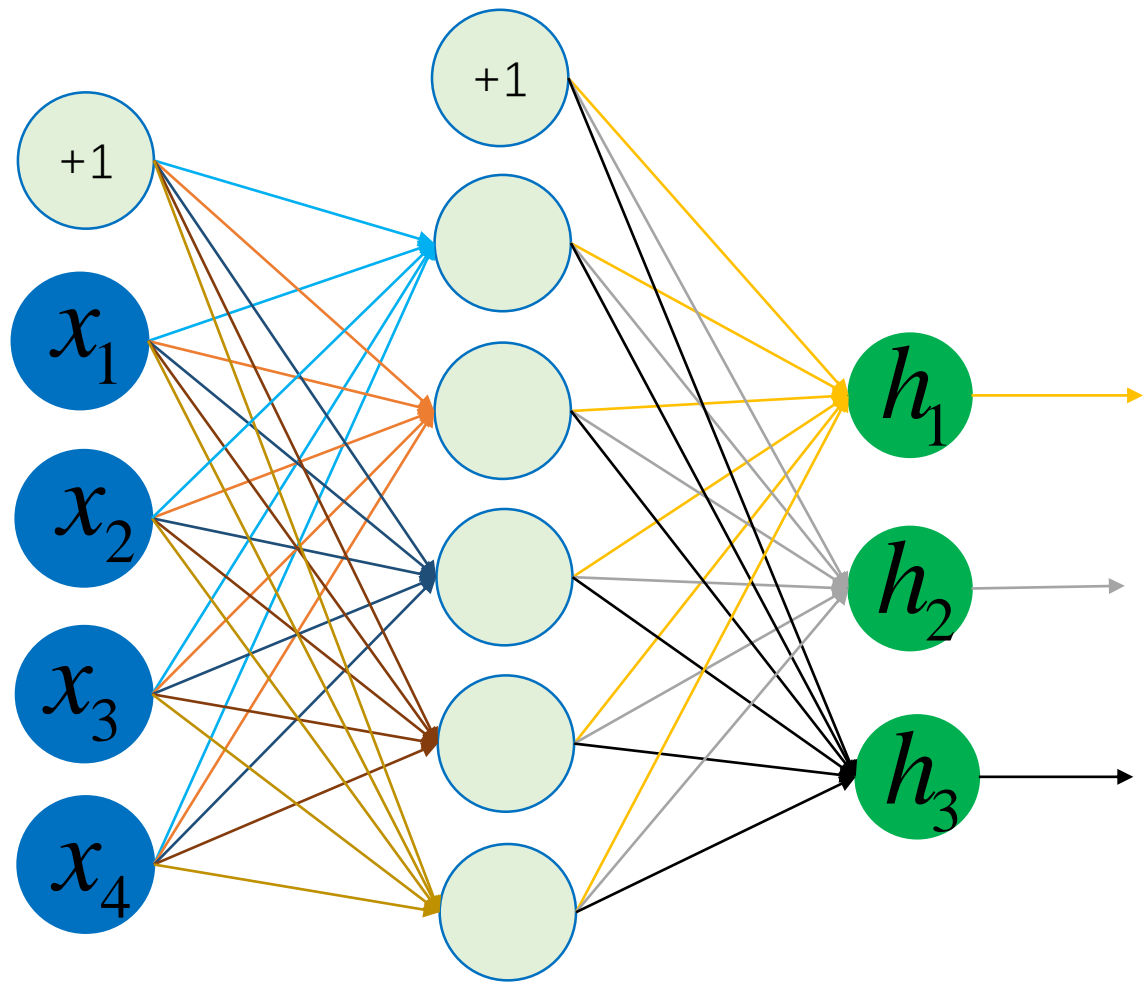
$$Z = X \cdot W + b$$

$$h = g(Z)$$

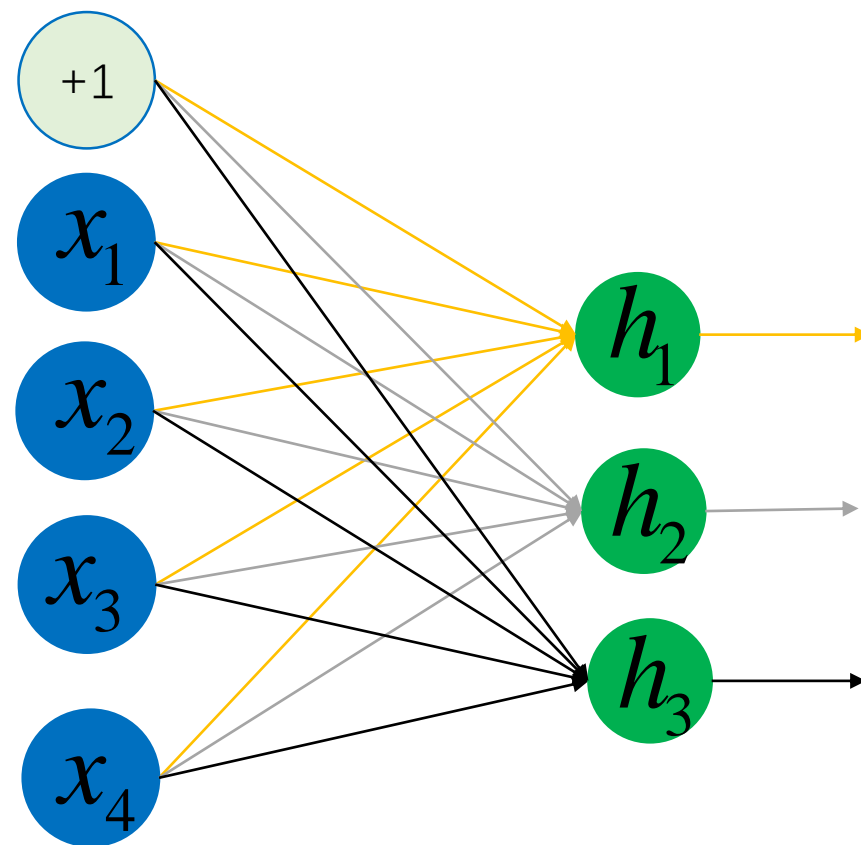


feature extraction

classification

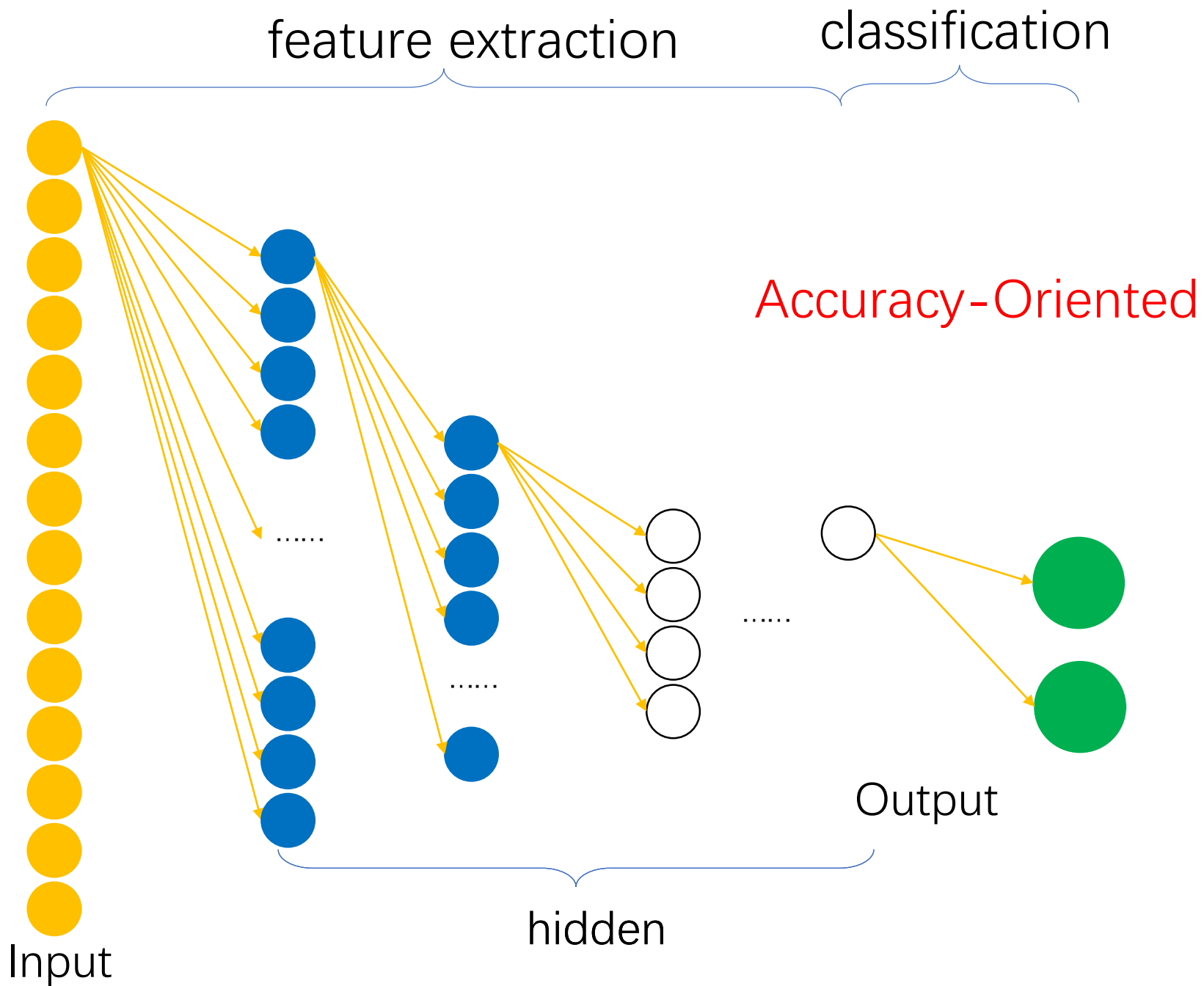
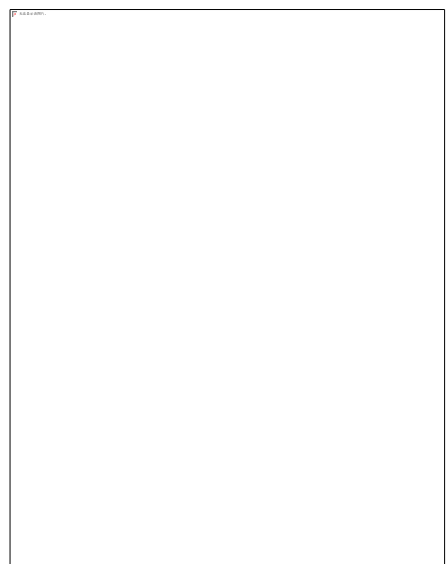


Deep Neural Network(DNN)



Perceptron

How many layers could be added ?





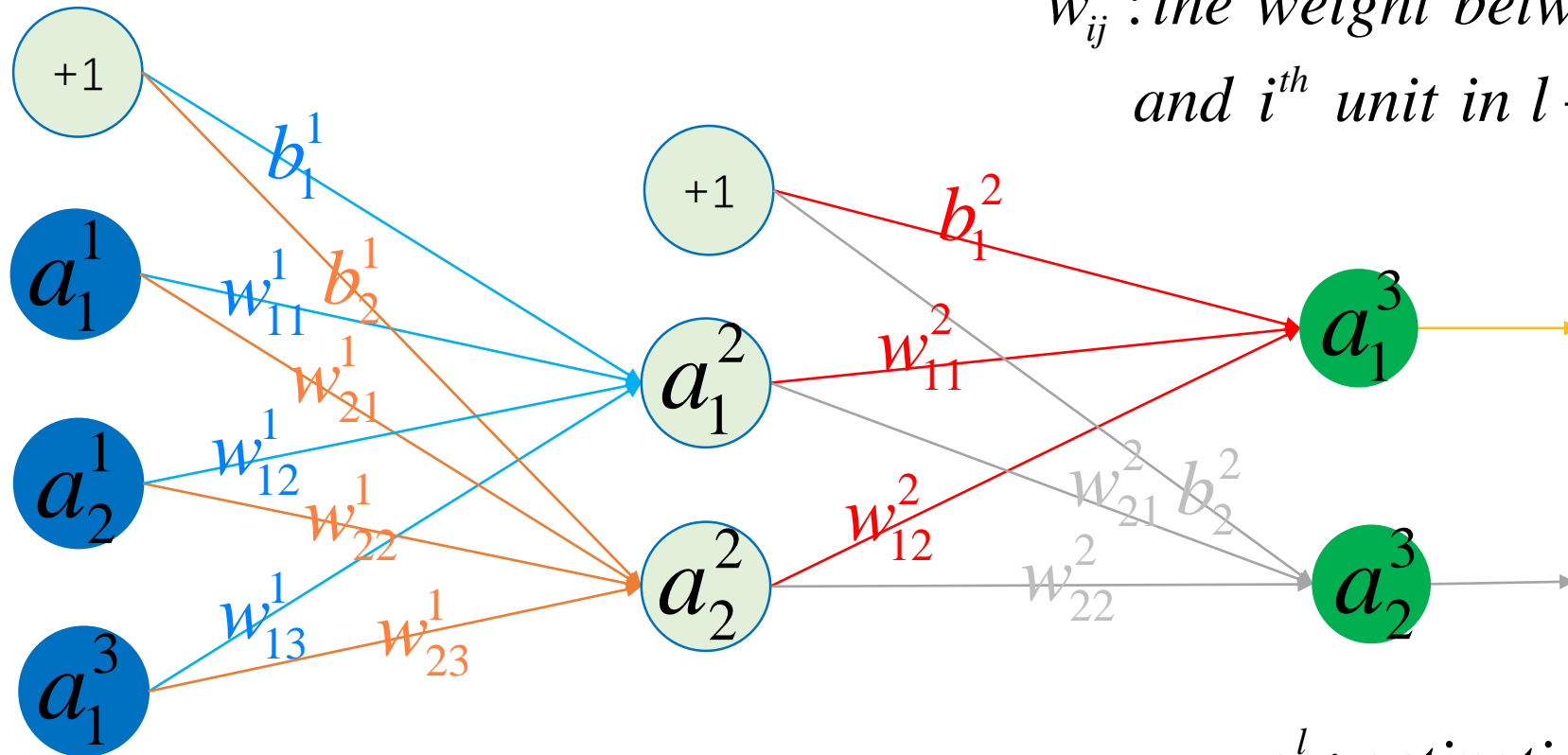
# A simplified model

$L$ : layers of full nn

$S_l$ : units in  $l^{\text{th}}$  layer

$K$ : units in output layer

$w_{ij}^l$ : the weight between  $j^{\text{th}}$  unit in  $l$  layer  
and  $i^{\text{th}}$  unit in  $l+1$  layer

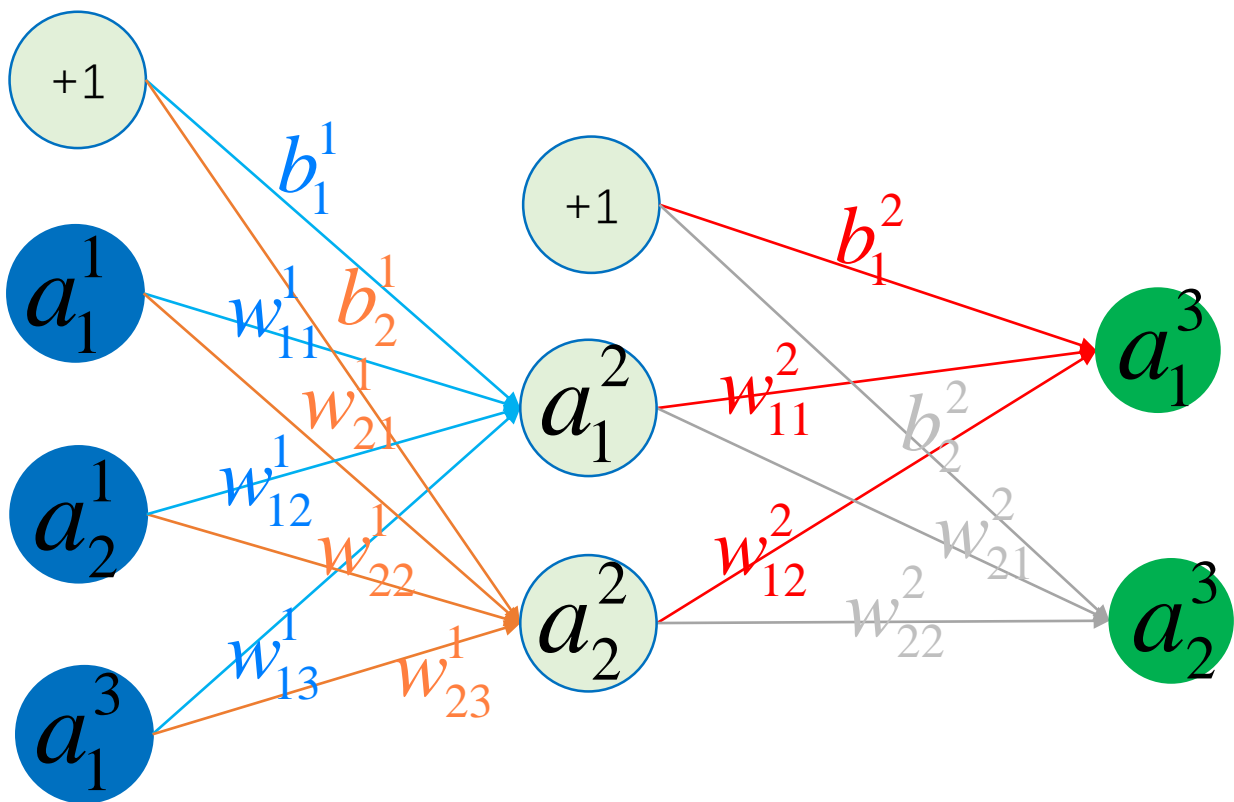


$$L = 3$$

$$S_1 = 3, S_2 = 2, S_3 = K = 2$$

$a_i^l$ : activation value of  $i^{\text{th}}$  unit in  $l$  layer

$b^l$ : bias in  $l$  layer



$$z_1^2 = a_1^1 w_{11}^1 + a_2^1 w_{12}^1 + a_3^1 w_{13}^1 + b_1^1$$

$$z_2^2 = a_1^1 w_{21}^1 + a_2^1 w_{22}^1 + a_3^1 w_{23}^1 + b_1^1$$

$$\Rightarrow \begin{bmatrix} z_1^2 \\ z_2^2 \end{bmatrix} = \begin{bmatrix} w_{11}^1 & w_{12}^1 & w_{13}^1 \\ w_{21}^1 & w_{22}^1 & w_{23}^1 \end{bmatrix}_{2 \times 3} \times \begin{bmatrix} a_1^1 \\ a_2^1 \\ a_3^1 \end{bmatrix}_{3 \times 1} + \begin{bmatrix} b_1^1 \\ b_2^1 \end{bmatrix}$$

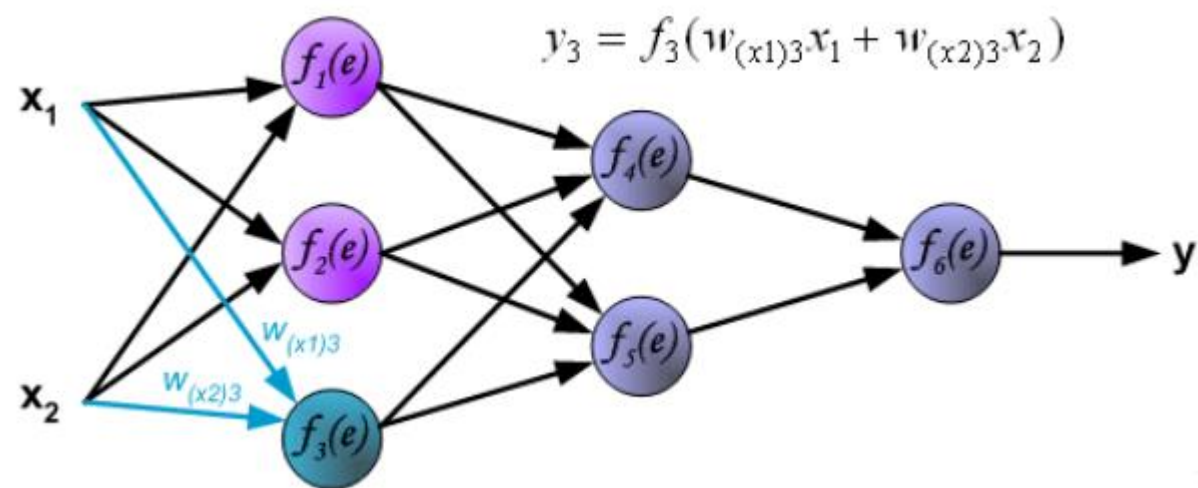
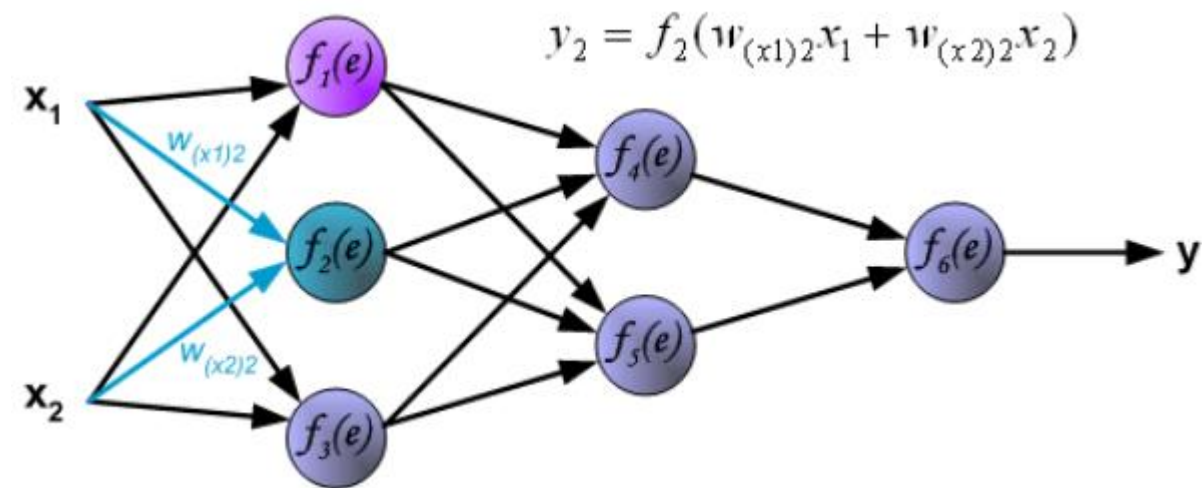
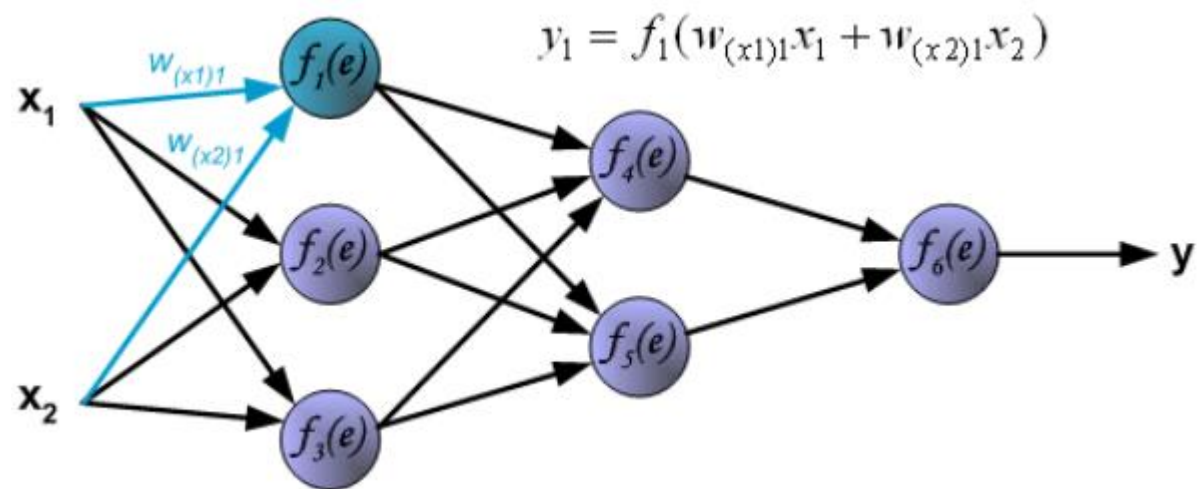
$$\Rightarrow z^2 = a^1 w^1 + b^1 \Rightarrow a^2 = f(z^2)$$

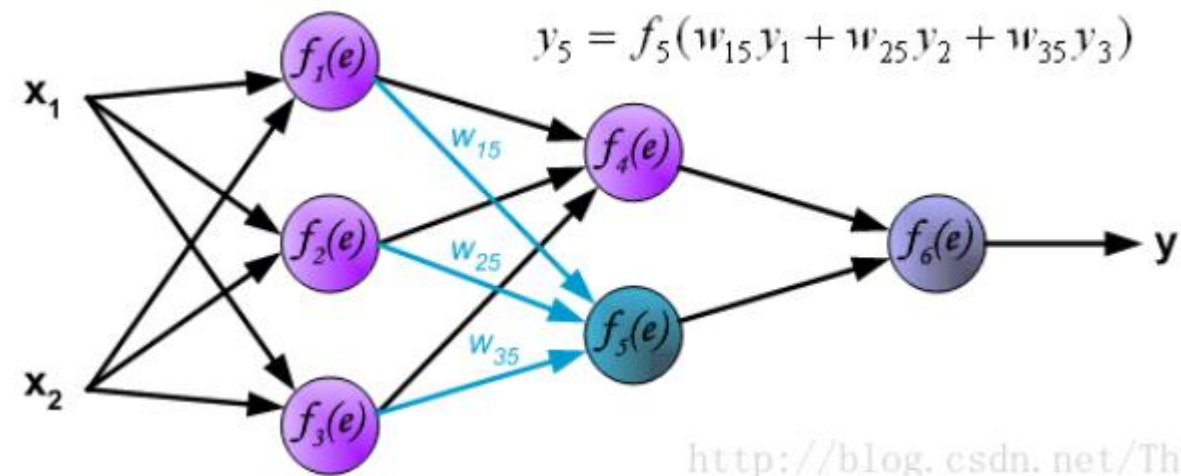
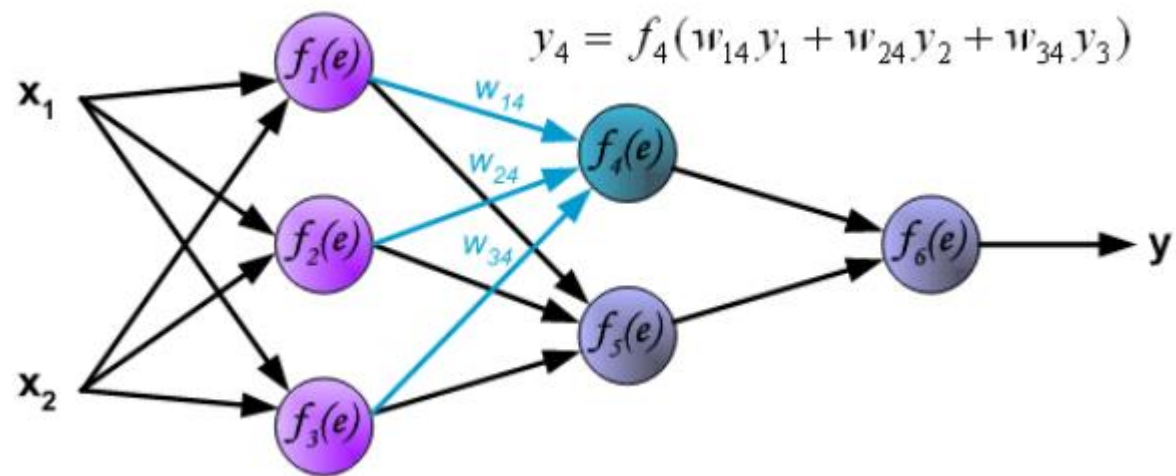
$$\Rightarrow z^3 = a^2 w^2 + b^2 \Rightarrow a^3 = f(z^3)$$

$$z_i^{l+1} = a_1^l w_{i1}^l + a_2^l w_{i2}^l + \cdots + a_{S_l}^l w_{iS_l}^l + b^l$$

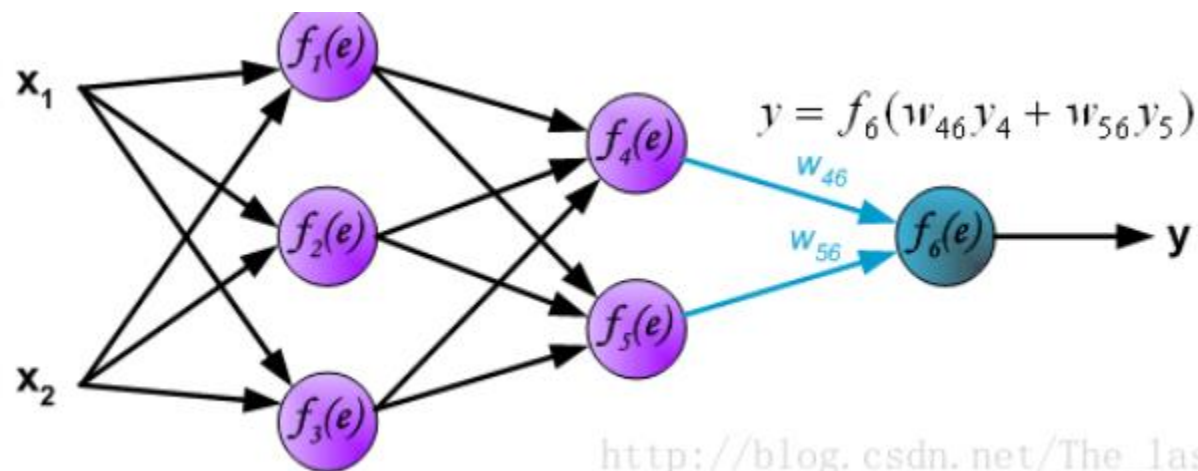
$$z^{l+1} = a^l w^l + b^l$$

$$a^l = f(z^l)$$

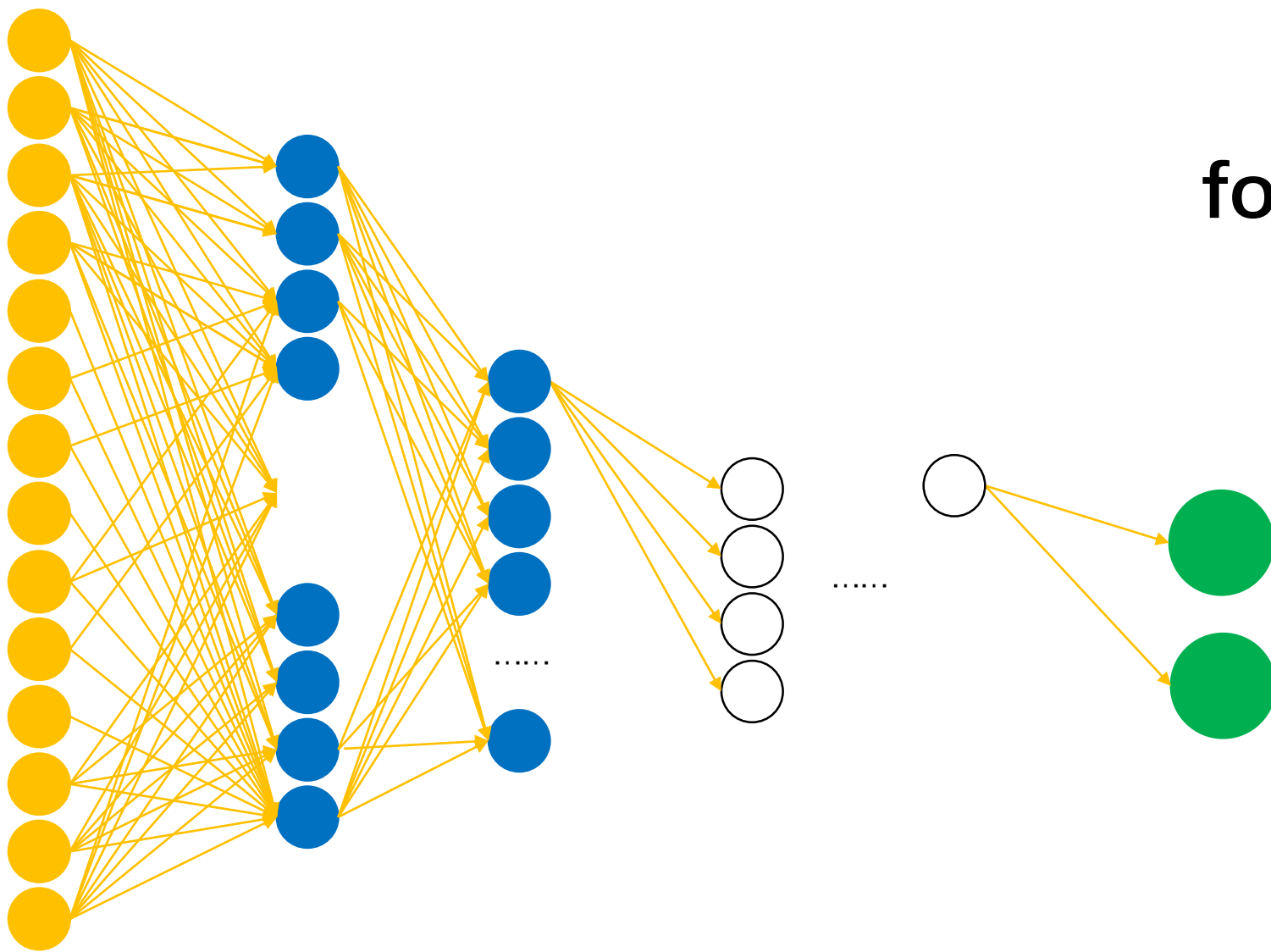




<http://blog.csdn.net/Th>



[http://blog.csdn.net/The\\_las](http://blog.csdn.net/The_las)



**forward propagation**

$$z^{l+1} = a^l w^l + b^l$$

$$a^l = f(z^l)$$