Deep Learning

Twitter: @streamingnology

What is Deep Learning?

Artificial Intelligence

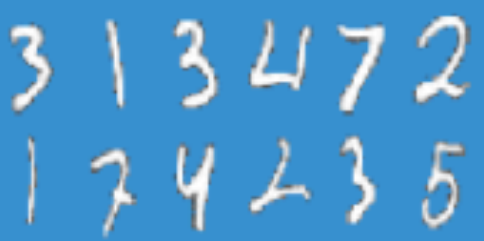
Any technique that enables computers to mimic human behavior

Machine Learning

Ability to learn without explicitly being programmed

Deep Learning

Extract patterns from data using neural networks

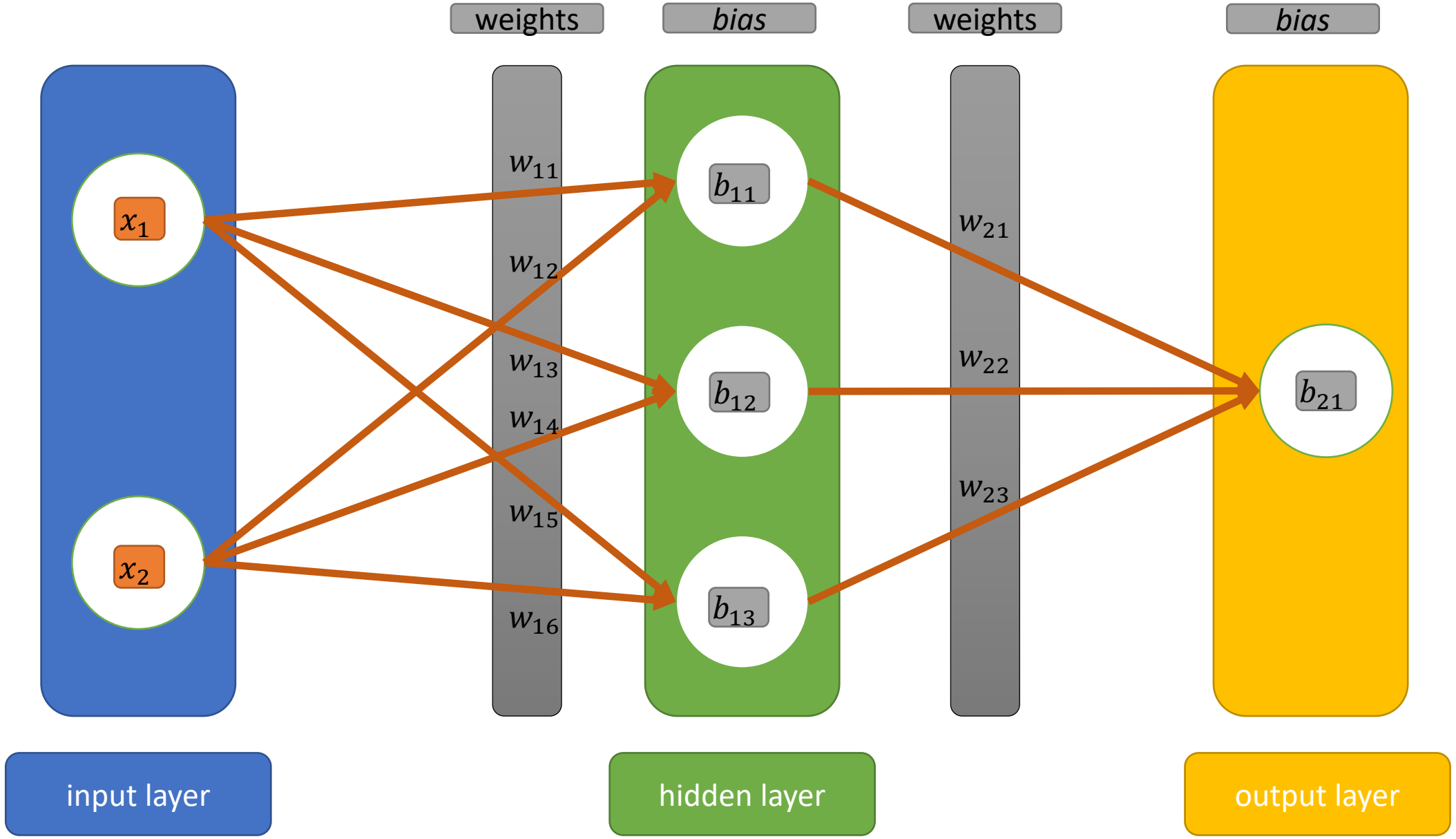


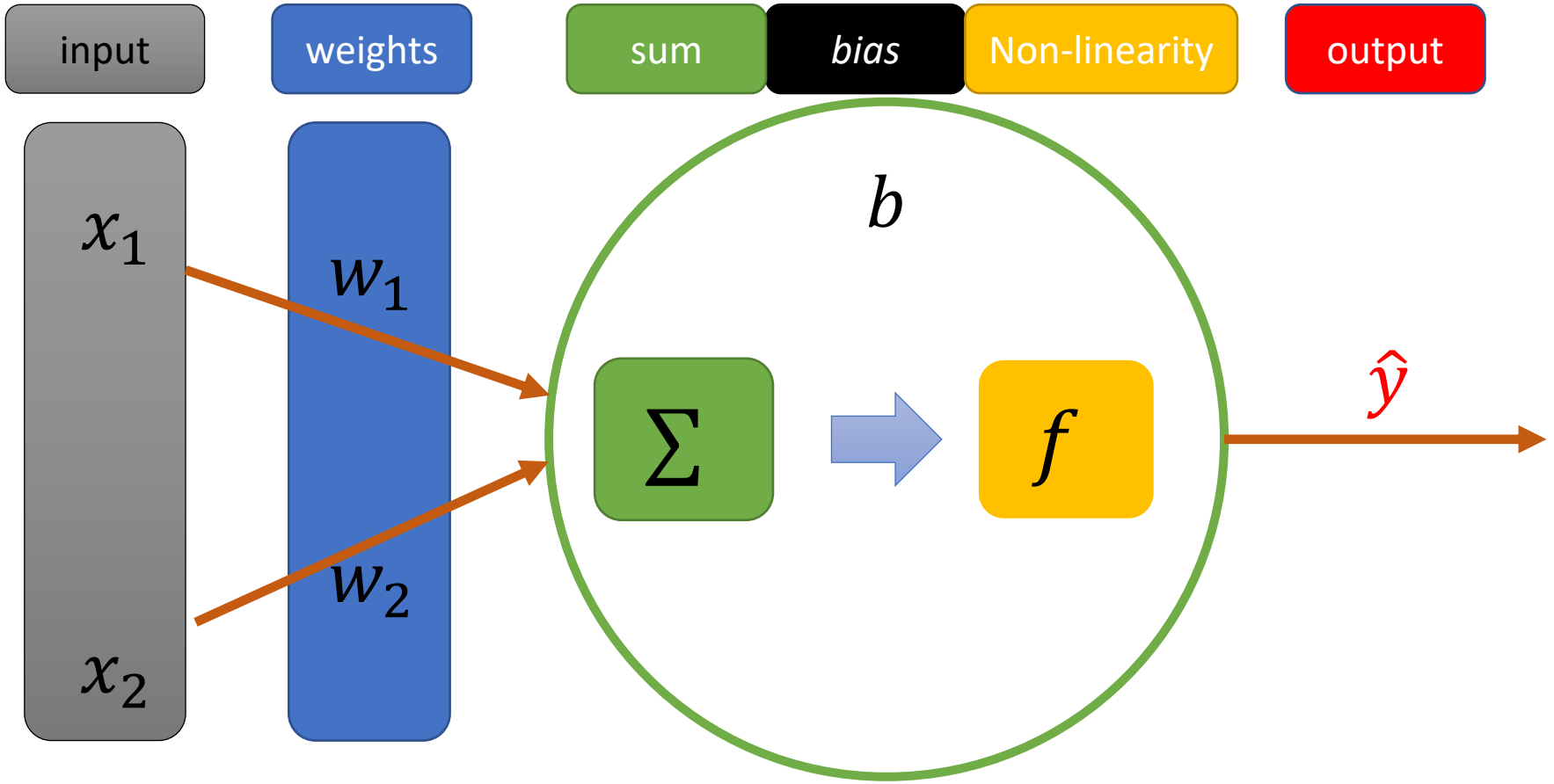
[0 1 2 3 4 5 6 7]

[1 3 5 7 9 11 13 15 17]

$f(x) = 2x + 1$


[0 1 2 3 4 5 6 7]
数据
结果
[1 3 5 7 9 11 13 15 17]
机器学习
$f(x) = 2x + 1$
输出规则


weights
bias
weights
bias
x_1
x_2
w_{11}
w_{12}
w_{13}
w_{14}
w_{15}
w_{16}
b_{11}
b_{12}
b_{13}
w_{21}
w_{22}
w_{23}
b_{21}
input layer
hidden layer
output layer

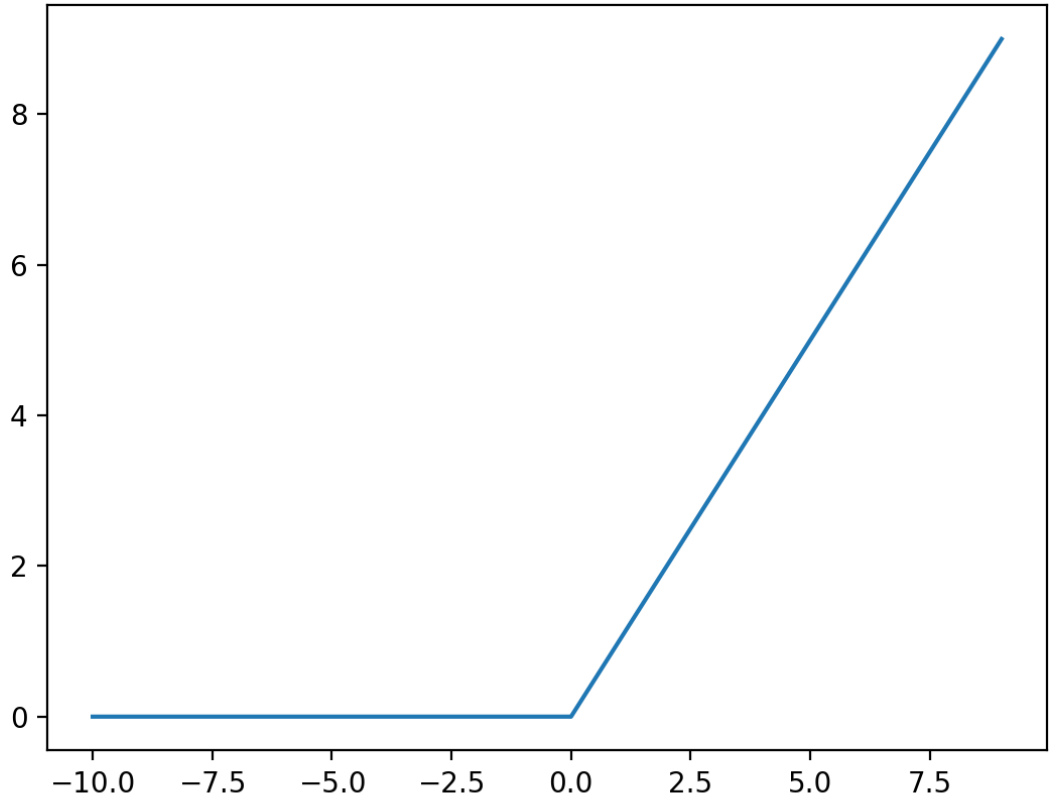


$$\hat{y} = f(x_1 * w_1 + x_2 * w_2 + b) = f\left(X^{\mathrm{T}}W + b\right)$$

$$X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \text{ and } \mathrm{W} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$$

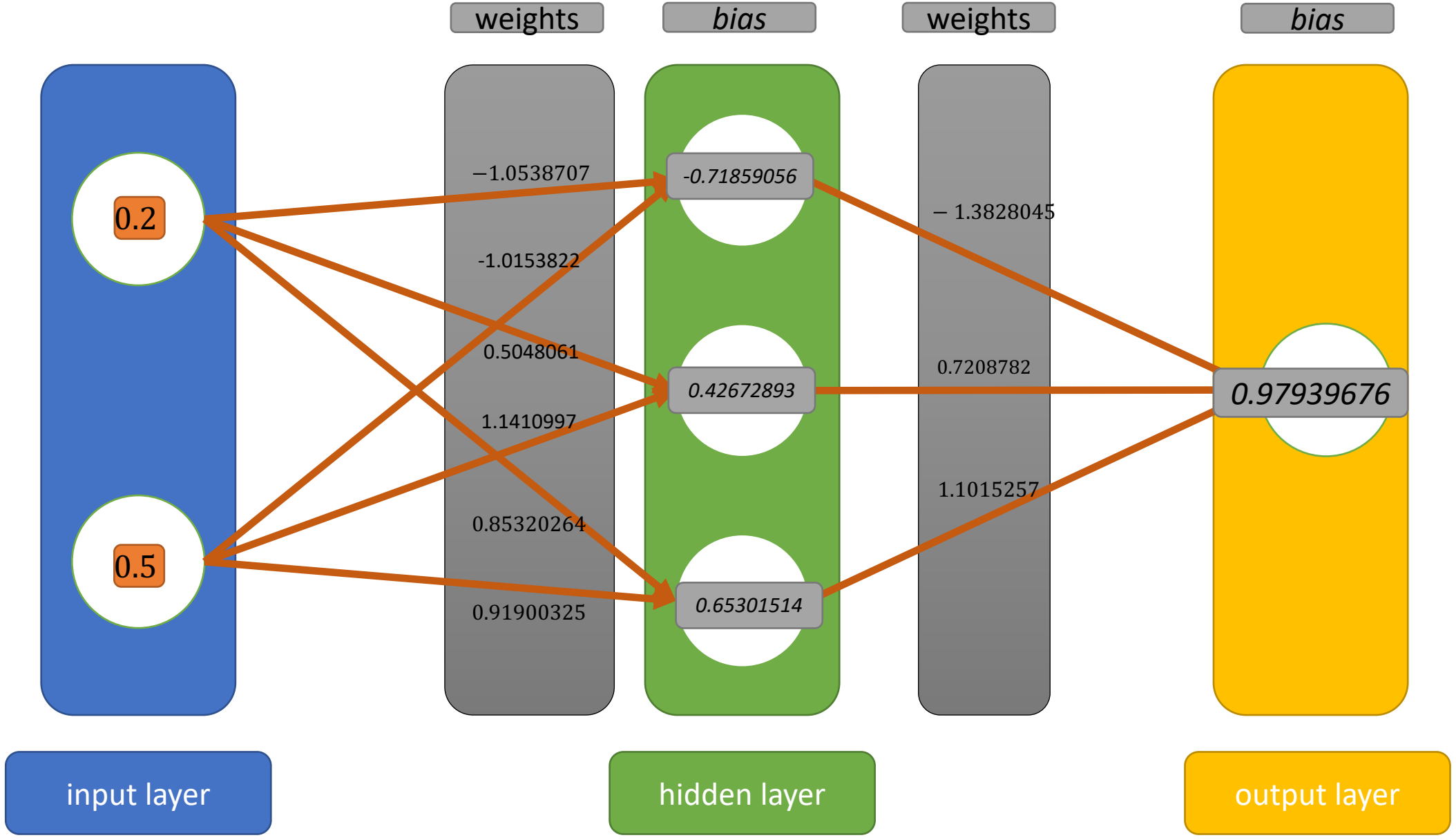
f 非线性函数 ReLU and sigmoid

$$ReLU(x) = \max(0, x)$$



$$X = \begin{bmatrix} 0.1 \\ 0.2 \end{bmatrix} \quad \mathrm{W} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad b = 0.3 \quad f(x) = ReLU$$

$$\hat{y} = f(X^T W + b) = f(0.1 * 1 + 0.2 * 2 + 0.3) = f(0.8) = 0.8$$



$$f(x_1, x_2) = 2x_1 + 4x_2 + 3$$

$$X = \begin{bmatrix} 0.2 \\ 0.5 \end{bmatrix}$$

$$W1 = \begin{bmatrix} -1.0538707 & 0.5048061 & 0.85320264 \\ -1.0153822 & 1.1410997 & 0.91900325 \end{bmatrix}$$

hidden layer

$$b1 = [-0.71859056 \quad 0.42672893 \quad 0.65301514]$$

$$o1 = X^T W1 + b$$
$$= \begin{bmatrix} 0.2 \\ 0.5 \end{bmatrix}^T \begin{bmatrix} -1.0538707 & 0.5048061 & 0.85320264 \\ -1.0153822 & 1.1410997 & 0.91900325 \end{bmatrix} + [-0.71859056 \quad 0.42672893 \quad 0.65301514]$$
$$= [-0.71846522 \quad 0.67151107 \quad 0.63014215] + [-0.71859056 \quad 0.42672893 \quad 0.65301514]$$
$$= [-1.43705578 \quad 1.09824 \quad 1.28315729]$$

$$o1 = \begin{bmatrix} -1.43705578 \\ 1.09824 \\ 1.28315729 \end{bmatrix}$$

$$W2 = \begin{bmatrix} -1.3828045 \\ 0.7208782 \\ 1.1015257 \end{bmatrix}$$

$$b2 = [0.97939676]$$

output layer

$$o2 = o1^{T} W2 + b2$$

$$= \begin{bmatrix} -1.43705578 \\ 1.09824 \\ 1.28315729 \end{bmatrix}^{T} \begin{bmatrix} -1.3828045 \\ 0.7208782 \\ 1.1015257 \end{bmatrix} + [0.97939676]$$

$$= [4.19229516] + [0.97939676]$$

$$= [5.17169192]$$

END