

یادگیری ماشین با پایتون

فرزاد حسینیان

دانشگاه محقق اردبیلی

دانشکده علوم کامپیوتر

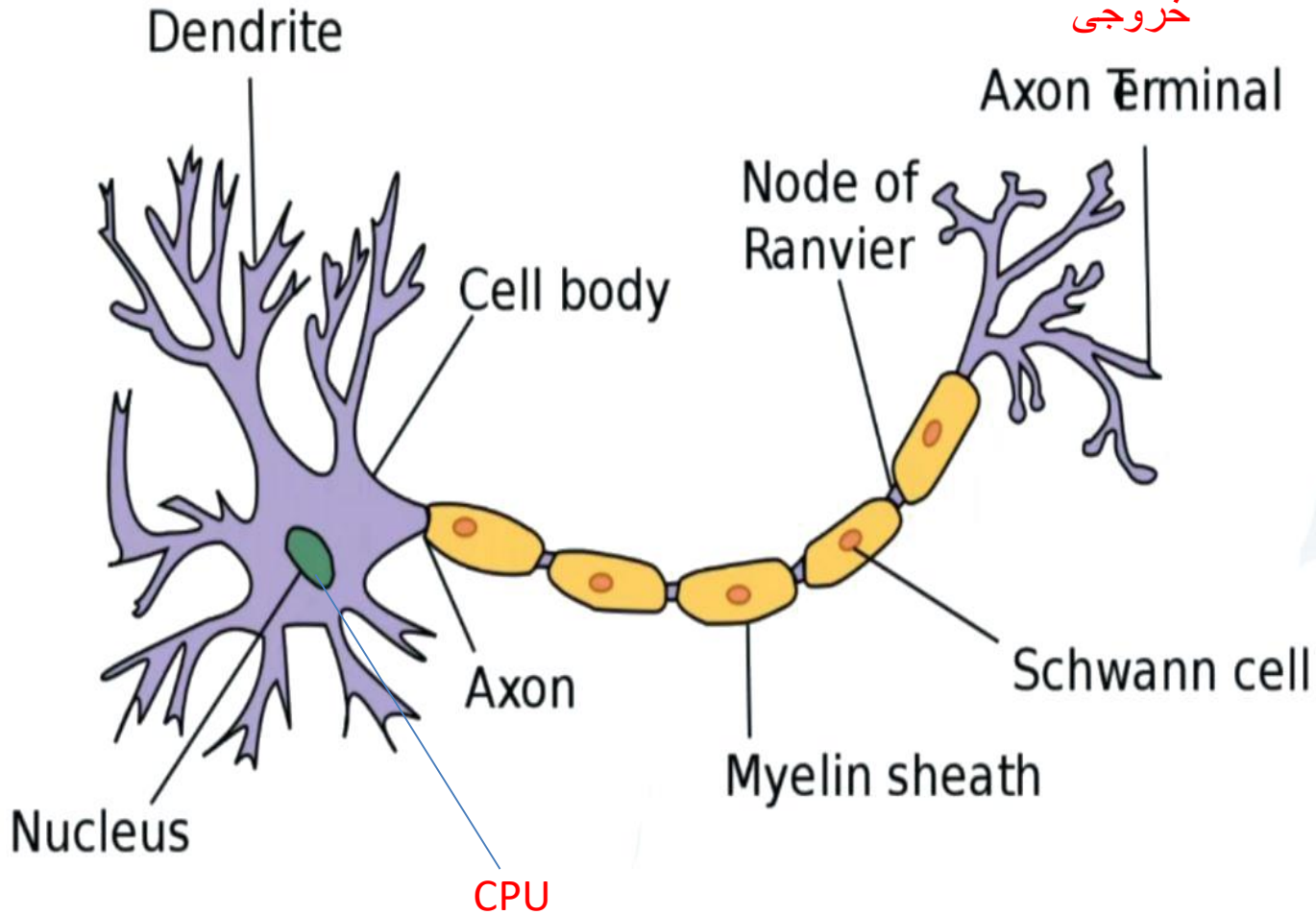
آبان 98

جلسه پنجم

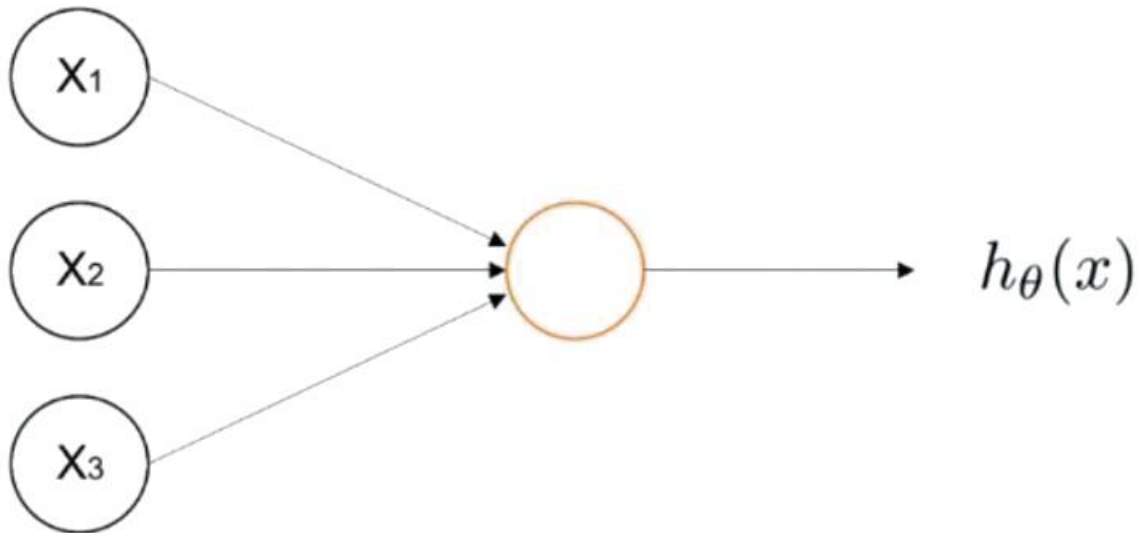
شبکه عصبی مصنوعی

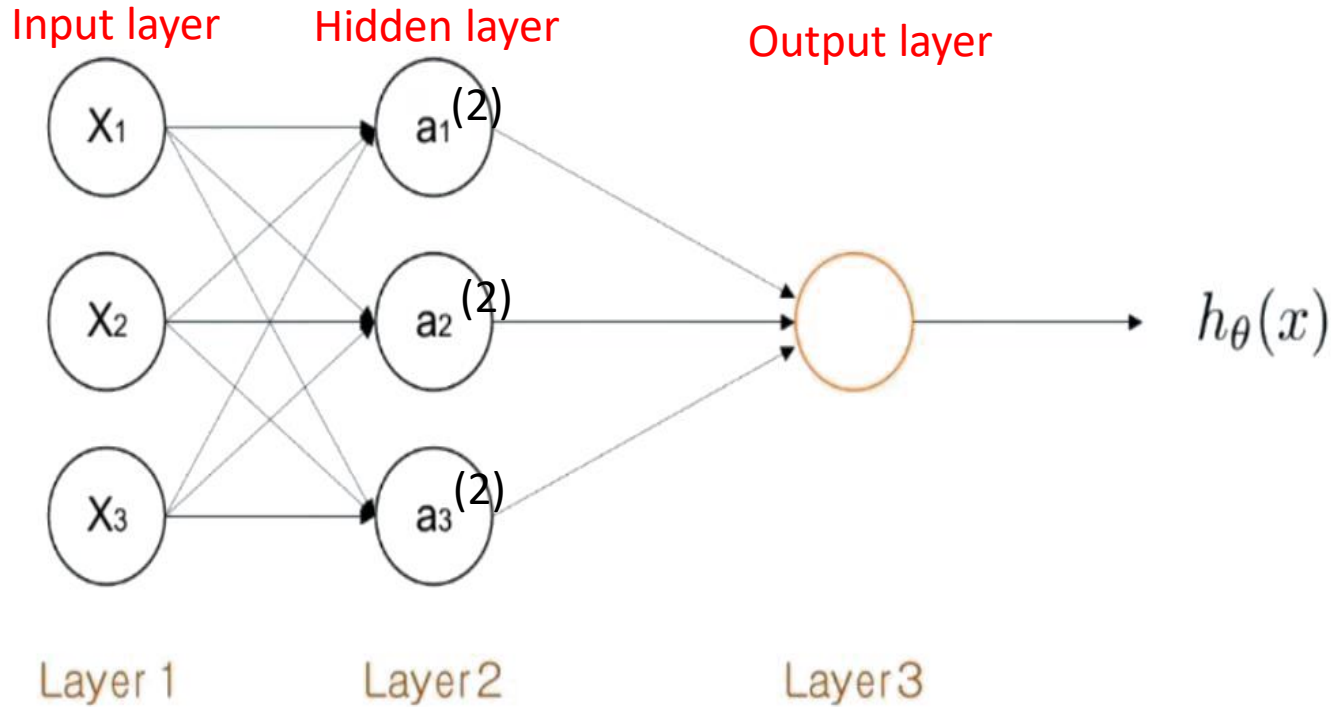
ورودی

خروجی



Neural Model: Logistic unit





$a_i^{(j)}$: تابع فعالسازی — activation function

$\theta^{(j)}$: weight matrix

وزن بین لایه j و $j+1$

$$a_1^{(2)} = g(\Theta_{10}^{(1)} x_0 + \Theta_{11}^{(1)} x_1 + \Theta_{12}^{(1)} x_2 + \Theta_{13}^{(1)} x_3)$$

$$a_2^{(2)} = g(\Theta_{20}^{(1)} x_0 + \Theta_{21}^{(1)} x_1 + \Theta_{22}^{(1)} x_2 + \Theta_{23}^{(1)} x_3)$$

$$a_3^{(2)} = g(\Theta_{30}^{(1)} x_0 + \Theta_{31}^{(1)} x_1 + \Theta_{32}^{(1)} x_2 + \Theta_{33}^{(1)} x_3)$$

$$h_{\Theta}(x) = a_1^{(3)} = g(\Theta_{10}^{(2)} a_0^{(2)} + \Theta_{11}^{(2)} a_1^{(2)} + \Theta_{12}^{(2)} a_2^{(2)} + \Theta_{13}^{(2)} a_3^{(2)})$$

$$z_1^{(2)}$$

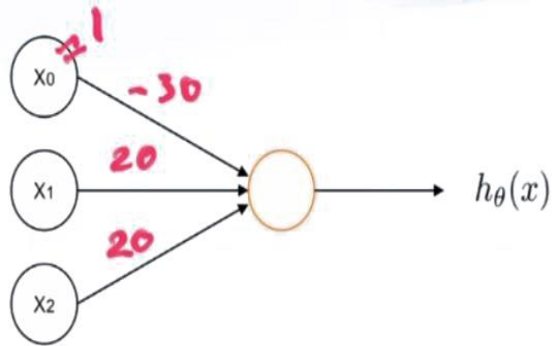
$$a_1^{(2)} = g(\Theta_{10}^{(1)} x_0 + \Theta_{11}^{(1)} x_1 + \Theta_{12}^{(1)} x_2 + \Theta_{13}^{(1)} x_3)$$

$$a_2^{(2)} = g(\Theta_{20}^{(1)} x_0 + \Theta_{21}^{(1)} x_1 + \Theta_{22}^{(1)} x_2 + \Theta_{23}^{(1)} x_3)$$

$$a_3^{(2)} = g(\Theta_{30}^{(1)} x_0 + \Theta_{31}^{(1)} x_1 + \Theta_{32}^{(1)} x_2 + \Theta_{33}^{(1)} x_3)$$

$$h_{\Theta}(x) = a_1^{(3)} = g(\Theta_{10}^{(2)} a_0^{(2)} + \Theta_{11}^{(2)} a_1^{(2)} + \Theta_{12}^{(2)} a_2^{(2)} + \Theta_{13}^{(2)} a_3^{(2)})$$

مثال: گیت AND



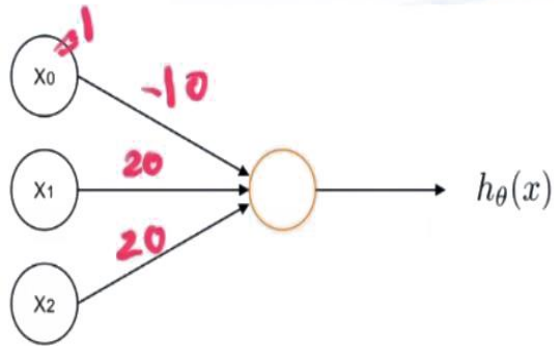
X_1	X_2	y	h
0	0	0	$g(-30) \approx 0$
0	1	0	$g(-10) \approx 0$
1	0	0	$g(-10) \approx 0$
1	1	1	$g(10) \approx 1$

$$h_{\theta}(x) = g(-30 + 20(0) + 20(0)) = g(-30) \approx 0$$

$$h_{\theta}(x) = g(-30 + 20(0) + 20(1)) = g(-10) \approx 0$$

$$h_{\theta}(x) = g(-30 + 20(1) + 20(1)) = g(+10) \approx 1$$

مثال: گیت OR



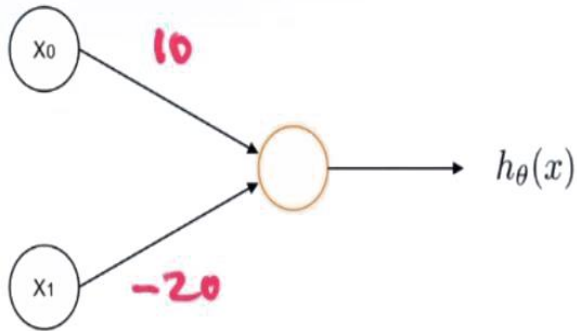
X_1	X_2	y	h
0	0	0	$g(-10) \approx 0$
0	1	1	$g(+10) \approx 1$
1	0	1	$g(+10) \approx 1$
1	1	1	$g(30) \approx 1$

$$h_{\theta}(x) = g(-10 + 20(0) + 20(0)) = g(-10) \approx 0$$

$$h_{\theta}(x) = g(-10 + 20(0) + 20(1)) = g(+10) \approx 1$$

$$h_{\theta}(x) = g(-10 + 20 + 20) = g(30) \approx 1$$

مثال: گیت NOT



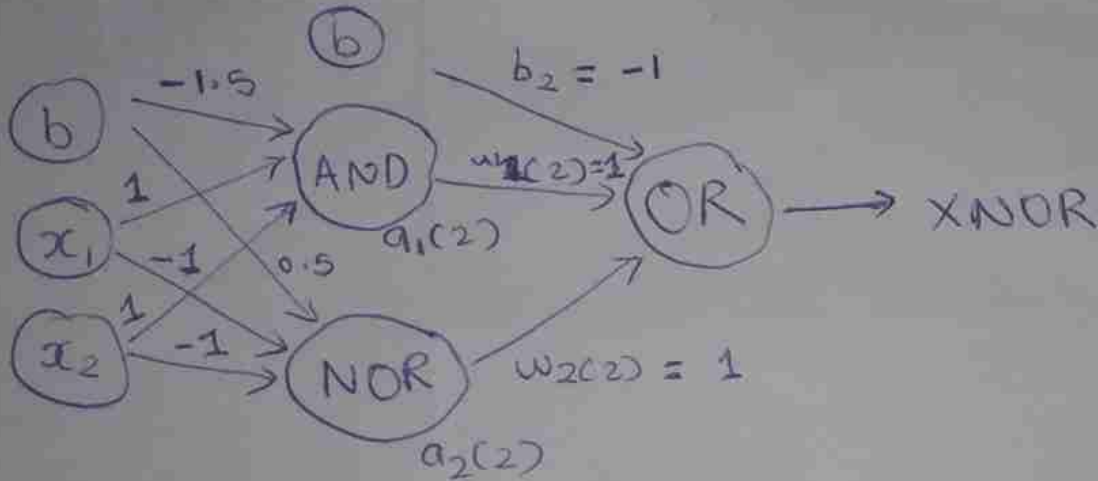
x_1	y	h
0	1	$g(10) \approx 1$
1	0	$g(-10) \approx 0$

$$h_{\theta}(x) = g(10 - 20(0)) = g(10) \approx 1$$

$$h_{\theta}(x) = g(10 - 20(1)) = g(-10) \approx 0$$

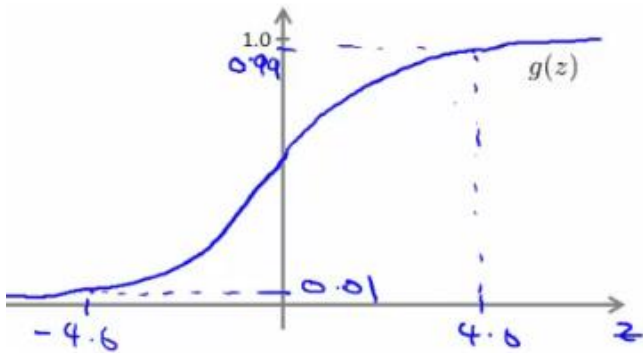
نکته: شبکه عصبی تک لایه معادل رگرسیون لاجستیک (با مرز تصمیم خطی) هستند

مثال: گیت XNOR



x_1	x_2	$a_1(2)$	$a_2(2)$	XNOR
0	0	0	1	1
0	1	0	0	0
1	0	0	0	0
1	1	1	0	1

مثال: گیت XNOR

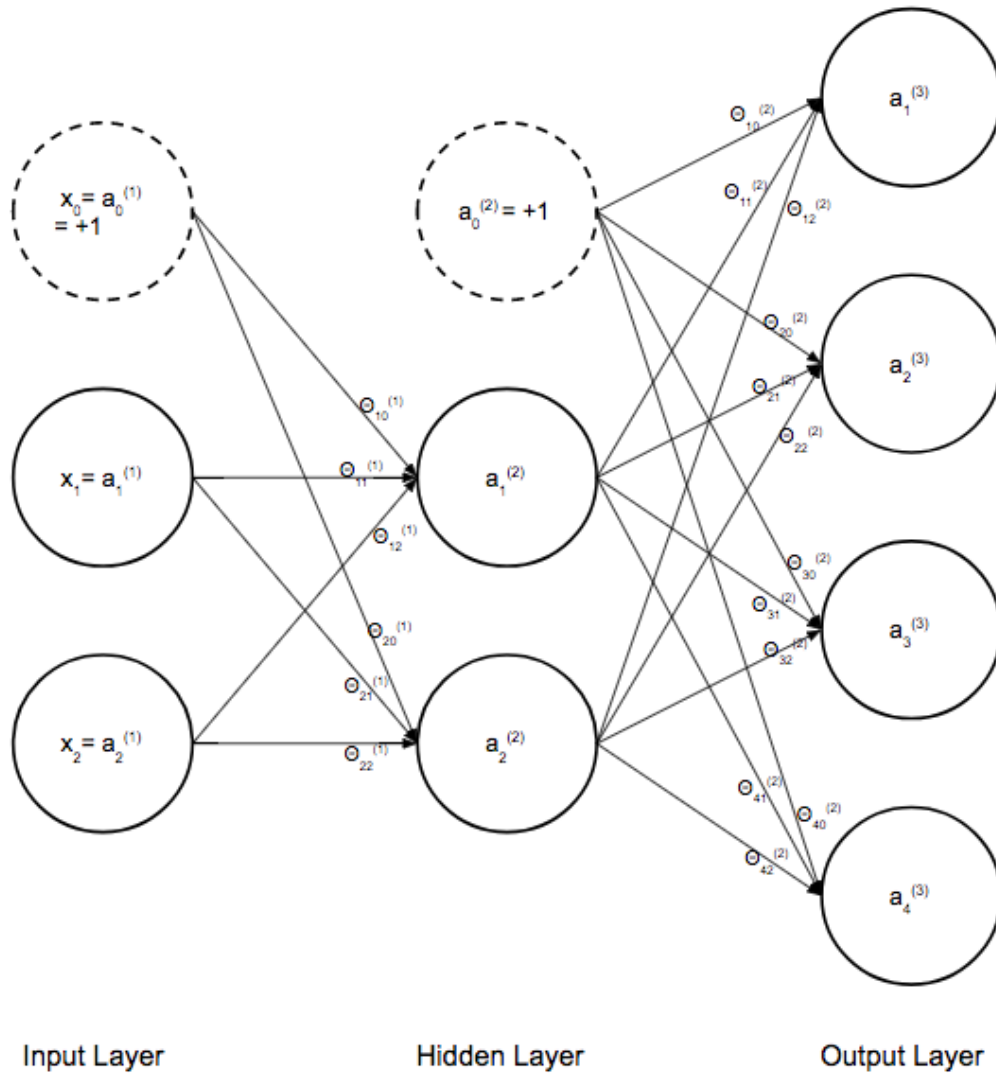


Sigmoid function (reminder)

x_1	x_2	$h_{\Theta}(x)$
0	0	$g(-30) \approx 0$
0	1	$g(-10) \approx 0$
1	0	$g(-10) \approx 0$
1	1	$g(10) \approx 1$

$h_{\Theta}(x) \approx x_1 \text{ AND } x_2$

Forward propagation



هدف شبکه عصبی مصنوعی: پیدا کردن وزن ها

a = output of last layer

$$e = \vec{y} - \vec{a}$$

$$J(\theta) = e^T e$$

هدف شبکه عصبی مصنوعی: پیدا کردن وزن ها

a = output of last layer

$$e = \vec{y} - \vec{a}$$

$$J(\theta) = e^T e$$

$$e = \vec{y} - \vec{a} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} - \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} y_1 - a_1 \\ y_2 - a_2 \\ \vdots \\ y_n - a_n \end{bmatrix}$$

$$\begin{aligned} e^T e &= (y_1 - a_1)^2 + (y_2 - a_2)^2 + \dots + (y_n - a_n)^2 \\ &= \sum_{i=1}^n (y_i - a_i)^2 \end{aligned}$$

Python code