

W04

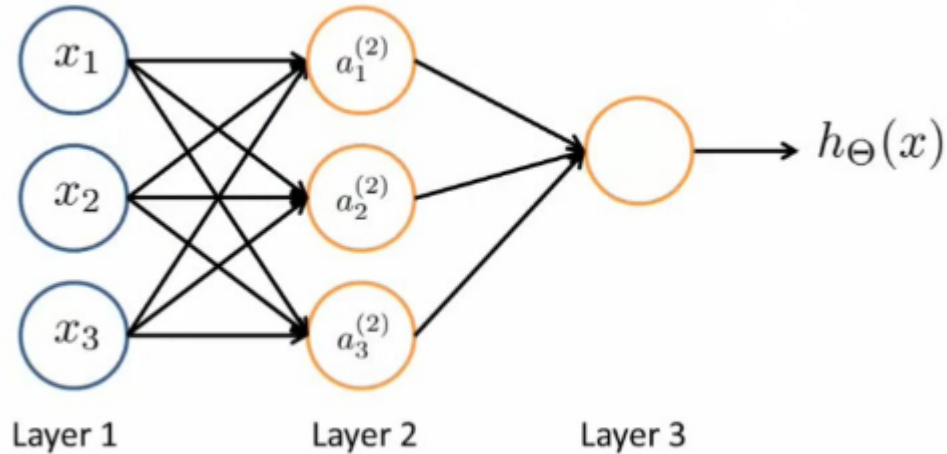
Fekete Máté

2020 Április

Tartalomjegyzék

1	Neurális hálózatok	2
2	Neurális hálózatok - Osztályozás több osztállyal	3

1 Neurális hálózatok



A neurális hálózatok rétegre bomlanak, amik ezen belül neuronokra. Az első réteg mindig a bemenet, az utolsó a kimenet, illetve közte tetszőleges mennyiségű rejtett réteg. Két réteg közötti kapcsolatot a neuronjaik (csúcsaik) közötti θ súlyok adnak meg.

Jelölések:

$a_i^{(j)}$ - a j -edik réteg i -edik neuronjának "aktivációja"

A $a_0^{(j)}$ alakú neutronokat eltolásnak nevezzük, ezeknek valamilyen konstans az értéke

$\theta^{(j)}$ - a súlyok mátrixa, amik a j -edik és $j+1$ -edik réteget kötik össze

$$a_1^{(2)} = g(\theta_{1,0}^{(1)} \cdot x_0 + \theta_{1,1}^{(1)} \cdot x_1 + \theta_{1,2}^{(1)} \cdot x_2 + \theta_{1,3}^{(1)} \cdot x_3)$$

Tehát a második réteg első neuronjának értéke az előző (jelen esetben a bemeneti) réteg páronként vett érték-súly szorzatainak összegére hívott, már ismert szigmoid (aktivációs) függvény.

Pl.: $\theta_{1,1}^{(1)}$ jelen esetben az x_1 -ből $a_1^{(2)}$ -be való átment súlya, ezt megszorozzuk x_1 értékével és ha épp az $a_1^{(2)}$ értéket számoljuk, akkor ugyan ezt megteesszük minden 1. rétegbeli értékkel, majd a kapott értékeket összeadjuk.

Ugyan ilyen mintára:

$$a_2^{(2)} = g(\theta_{2,0}^{(1)} \cdot x_0 + \theta_{2,1}^{(1)} \cdot x_1 + \theta_{2,2}^{(1)} \cdot x_2 + \theta_{2,3}^{(1)} \cdot x_3)$$

$$a_3^{(2)} = g(\theta_{3,0}^{(1)} \cdot x_0 + \theta_{3,1}^{(1)} \cdot x_1 + \theta_{3,2}^{(1)} \cdot x_2 + \theta_{3,3}^{(1)} \cdot x_3)$$

$$h_\theta(x) = a_1^{(3)} = g(\theta_{1,0}^{(2)} \cdot a_0^{(2)} + \theta_{1,1}^{(2)} \cdot a_1^{(2)} + \theta_{1,2}^{(2)} \cdot a_2^{(2)} + \theta_{1,3}^{(2)} \cdot a_3^{(2)})$$

2 Neurális hálózatok - Osztályozás több osztállyal



Ha egy olyan osztályozási problémánk van, amiben egy példa több mint két kategória egyikébe tartozik, minden osztályunkhoz lesz egy neuron a kimeneti rétegünkben (jelen esetben ez 4 osztályt jelent).

Tfh. meg szeretnénk állapítani egy képről, hogy mi szerepel rajta.

A négy opciónk a következő:

Gyalogos, autó, motorkerékpár, kamion

Azt szeretnénk hogy a kimenetünk valami ilyesmi módon nézzen ki:

$$h_\theta(x) \approx \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad h_\theta(x) \approx \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad h_\theta(x) \approx \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad h_\theta(x) \approx \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

ha gyalogos

autó

motor

kamion