

Paskaita 5

Gilieji neuronu tīkli

Aprēķināt drosmeis, kurus
neuroņi apstrādā NT.

Taip pat šādu formātu drosmeis
apstrādā NT, kuris pateiks pro-
gnozē, kas ~~jā~~ pateiks nācējo
drosmeis. Tās sauc par metoju.

$$A = (A[m][j]) \quad \begin{matrix} \text{rindas skaits} \\ \uparrow \\ 0 \leq m < M, \quad 0 \leq j < J. \end{matrix}$$

kai m -fiksētas, šķēršļi J drosmeis
rindas (vektoris - element)

$$[A[m][0], A[m][1], \dots, A[m][J-1]]$$

Tevu fiksojums j (konkrēti parametri)
tas ir kopā vektoris - stulpelis

$$\begin{bmatrix} A[0][j], \\ A[1][j], \\ \dots \\ A[M-1][j] \end{bmatrix}$$

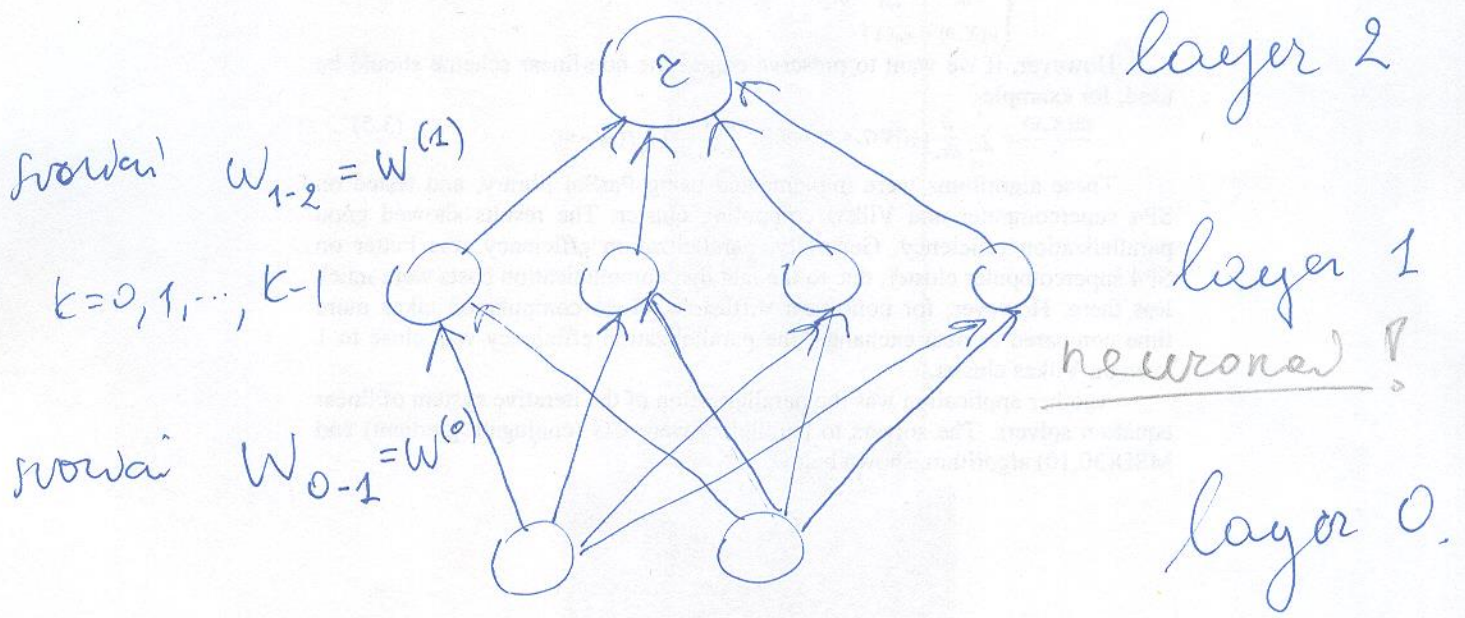
Jis apibūdina j -tojo parametro reikšmes skirtingais duomenų rinkiniais (bendryniais, natūraliais, ...).

Tada parostume eksperto / mokytojo nuomonę apie šviesos rezultatus

$$\begin{bmatrix} r[0], \\ r[1], \\ \dots \\ r[M-1] \end{bmatrix}$$

Vėl apšviename atvejį, kur rezultatai surašomi kaip vektorius - stulpelis.

Sudarome papildomus (pasleptą) sliekšnių, keičiame gelmė geriau atspindėti ryšius (koreliacijas) tarp įėjimų duomenų A ir rezultato (prognozės) z.



Turime tarpinių sliekšnių.

Jų gelmė būtų ir keli.

! Parametrus arba padidėjo ir jie išdėstyti specialiu būdu. Neuronai smegenų modelyje

Taip parinkti parametrai, kad galėtume minimizuoti prognozes paklaidas

$$e = \sum_{m=0}^{M-1} (\text{pred}[m] - r[m])^2$$

(kaip ~~naudojome~~ ^{naudojome} paprastai DNT atveju)

$$\text{pred}[m] = \sum_{k=0}^{K-1} p[m][k] \underbrace{W^{(1)}[k]}_{\text{svoriai}} + b^{(1)}$$

$p[m][k]$ - paslėpto sluoksnio neurono k -tojo reikšmė m -tajam bandymui

Jas skaičiuojami pamašodami įėjimo informaciją ir aktyvumo funkciją $f(x)$:
(pvz. $f(x) = \tanh(x)$) → neuronų įėjimą arba lėlinės aktyvumo

$$p[m][k] = f\left(\sum_{j=0}^{J-1} \underbrace{A[m][j]}_{\text{įėjimo duon.}} \underbrace{W^{(0)}[j][k]}_{\text{svoriai}} + b^{(0)}[k]\right)$$

Kai kada apšviekiama aktyvumo funkcija

$$f(x) = x, \text{ jei } x > 0, \text{ ir } b^{(0)}[k] = 0.$$

Kota darinei naudojama alityvumo funkcija yra ReLU (relu).

rectified Linear Unit

(elektroninės lygtinės)

Tai dalinis tiesinė funkcija:

- jei argumentas teigiamas, tai rezultate gaunama reikšmė proporcinga argumentui
- jei argumentas neigiamas - $f(x) = 0$

$$f(x) = \begin{cases} cx, & x \geq 0 \\ 0, & x < 0, \end{cases} \quad (\text{kur } c=1).$$

Uzrēķināme pakļautos izteikto matricāle formā. ($b_k^{(0)}$ vai $b^{(1)} = 0$ iekšāve).
 (Note: The original text has a typo "matricāle" which has been corrected to "matricāle" in the transcription.)

$$e = \sum_{m=0}^{M-1} \left(\sum_{k=0}^{K-1} f \left(\sum_{j=0}^{J-1} a_{mj} W_{jk}^{(0)} \right) W_k^{(1)} - z_m \right)^2$$

A: skaicēsim gradientus δ .

$$\frac{\partial e}{\partial W_{k_0}^{(1)}} = 2 \sum_{m=0}^{M-1} \left[\sum_{k=0}^{K-1} f \left(\sum_{j=0}^{J-1} a_{mj} W_{jk}^{(0)} \right) W_k^{(1)} - z_m \right] \times$$

δ

$$\times f \left(\sum_{j=0}^{J-1} a_{mj} W_{jk_0}^{(0)} \right)$$

$$\frac{\partial e}{\partial W_{j_0 k_0}^{(0)}} = 2 \sum_{m=0}^{M-1} \left[\sum_{k=0}^{K-1} f \left(\sum_{j=0}^{J-1} a_{mj} W_{jk}^{(0)} \right) W_k^{(1)} - z_m \right]$$

$$\times \frac{\partial f}{\partial x} \left(\sum_{j=0}^{J-1} a_{mj} W_{jk_0}^{(0)} \right) a_{mj_0} W_{k_0}^{(1)}$$

$$W_{k_0}^{(2), n+1} = W_{k_0}^{(2), n} - \alpha \frac{\partial e}{\partial W_{k_0}^{(1)}} (W^{(0), n}, W^{(1), n})$$

$$W_{j_0 k_0}^{(0), n+1} = W_{j_0 k_0}^{(0), n} - \alpha \frac{\partial e}{\partial W_{j_0 k_0}^{(0)}} (W^{(0), n}, W^{(1), n})$$