

Paskaita 4

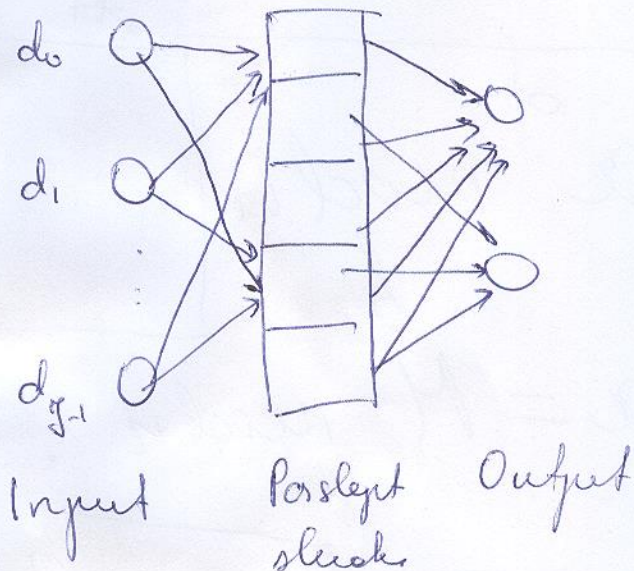
Dirbtiniai Neuroniniai

Tinklai su keliais svorais
- parametrais.

Klasiniai (kartais be gelutinis
atšakymų):

- 1) Apribotas procesas, kai turime
kelis (daug) duomenų rinkinius.
Duomenų vektoriai (d_1, d_2, \dots, d_m)
- 2) Gradientinis nes. leidimo metodas,
kai turime kelis rezultatus
(multiple outputs).
- 3) Aibų, variantų kombinacija.

DNT paprasči. schema



Aistra, kad galime sudaryti
įvairias architektūras

- keli paslepti sluoksniai
- įvairūs neuronų jungimo
funkcijos (aktyvavimo f-jos)
- sąrūkos
- skirtingų neuronų sluoksniai
skirtinguose sluoksniuose.

Prz. Apmo lygmu dromenys

$$d_0 = [8.5, 9.5, 9.9, 9.0]$$

$$d_1 = [0.65, 0.8, 0.8, 0.9]$$

$$d_2 = [1.2, 1.3, 0.5, 1.0]$$

$$r = [1, 1, 0, 1] \quad (\text{laimejo, pralinieto})$$

Reikiamas dromenys yra vienas rezultatas

$$w = [0.1, 0.2, -0.1] \quad (\text{juos reikia optimizuoti param})$$

Galime skaičiuoti paklaidą (pr.

0-io reikiamo dromenys).

$$\text{pred} = d_0[0] \cdot w[0] + d_1[0] \cdot w[1] + d_2[0] \cdot w[2]$$

$$\text{error} = (\text{pred} - r[0])^2$$
$$\text{delta} = \text{pred} - r[0] \rightarrow \begin{pmatrix} 0.020 \\ -0.14 \end{pmatrix}$$

Skaičiuojame svorius (koparametrus)
pataisais

Prismunskime gradientinio nusileidimo
metodą

$$e(w_0, w_1, w_2) = \left(\sum_{j=0}^2 w_j d_j - r \right)^2$$

Sprendžiamie uždaviniai

$$\arg \min e(w_0, w_1, w_2) = (\bar{w}_0, \bar{w}_1, \bar{w}_2)$$

$$e(\bar{w}_0, \bar{w}_1, \bar{w}_2) \leq e(w_0, w_1, w_2)$$

Iteracijos: (anti gradiento kryptimi)

$$w_j^{n+1} = w_j^n - \gamma \frac{\partial e}{\partial w_j} (w_0^n, w_1^n, w_2^n) \quad j=0,1,2$$

$$\frac{\partial e}{\partial w_{j_0}} = 2 \left(\sum_{j=0}^2 w_j d_j - r \right) d_{j_0}$$

$$= 2 \text{ delta } d_{j_0}$$

Imlūme $y = \frac{1}{2}$ ir suskaičiuojame

~~w_j~~ w_j-deltas, o tada naujus norus

~~$w_0^1 = 8.5 + 0.14 \cdot 8.5 = 9.69$~~

~~$w_1^1 = 0.65 + 0.14$~~

f

$w_0^1 = 0.1 + 0.14 \cdot 8.5 = 1.29$

$w_1^1 = 0.2 + 0.14 \cdot 0.65 = 0.291$

$w_2^1 = -0.1 + 0.14 \cdot 1.2 = 0.068$

ir t.t.

Tęgnu algoritmas neišsivergvoja, tai
~~ta~~ (t.y. paklaida error nesumašėja,
tai sumušume y ir patikūčiuoja
me norus (parametrus w_j^1)).

a) DNT apmūlyme turime daugiam
duomenų (jūs reideris išuodst)

b) kai kurines svoris parametrus galdue
"užaldyti".

Āia pamīnēsime vīep sverbīp
pastabā.

P1. Jēgu "uzsālolome" - fīkruojame
vīep īs parametru, pr. \tilde{w}_0 ir vīlīan
pavīksta līkumīn parametru (pr.
 w_1, w_2) parūklīti šolūn, kēd palīlīdā
parūklītiem špūslīguš dīroumēnīn
grū līgu 0, šai papīldomēnī dījīn-
gībant w_0 parūklīnīp, mēs jān ve-
gūmēme papīldomēnī vīslīguš plīnīp.

Apibendrinimas (pašaulinimas
formules), kai duomenų rinkinys
skaitinis yra M .

• Apibūdiname palydovę

$$\text{error} = \sum_{m=0}^{M-1} \left(\sum_{j=0}^{J-1} w_j d_j [m] - z [m] \right)^2$$

gali
būti
svertini
vidurkis

$$\text{pred} [m] = \sum_{j=0}^{J-1} w_j d_j [m]$$

$$\text{delta} [m] = \text{pred} [m] - z [m]$$

$$\begin{aligned} \frac{\partial e}{\partial w_{j_0}} &= 2 \sum_{m=0}^{M-1} \left(\sum_{j=0}^{J-1} w_j d_j [m] - z [m] \right) d_{j_0} [m] \\ &= 2 \sum_{m=0}^{M-1} \text{delta} [m] d_{j_0} [m] \end{aligned}$$

$$w_j^{n+1} = w_j^n - \gamma \frac{\partial e}{\partial w_j} (w_0^n, \dots, w_{j-1}^n)$$

Nagrinėjame uzduotį, kai rezultatas (output) yra rinkinys (vektorius)

$$z[m] = [z_0[m], z_1[m], \dots, z_{K-1}[m]]$$

Proz. imleme biojūtklę, kecis turi atpažinti medžiagų rinkinio koncentracijas.

Sudarome paklaidos formulę

$$\text{error} = \sum_{k=0}^{K-1} \sum_{m=0}^{M-1} \left(\sum_{j=0}^{J-1} w_j^k d_j[m] - z_k[m] \right)^2$$

Turime parametrų aibę

$$w_j^k, j=0, \dots, J-1, k=0, \dots, K-1.$$

Skaičiuojame gradientus:

$$\frac{\partial e}{\partial w_{j_0}^{k_0}} = 2 \sum_{m=0}^{M-1} \left(\sum_{j=0}^{J-1} w_j^{k_0} d_j[m] - z_{k_0}[m] \right) d_{j_0}[m]$$

Gauname K DNT, atskirai kiekvienai komponentei. †