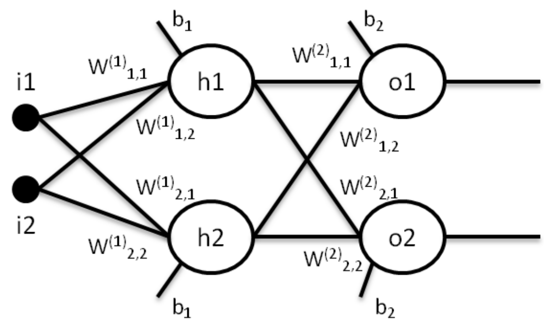
1. Neurális hálók gyakorlat

(Mesterséges intelligenca VIMIAC10 - 2018.nov.8.)

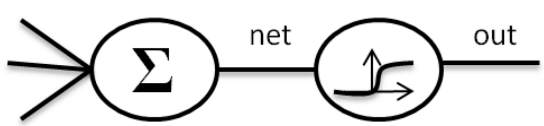


1. ábra.

1.1. Paraméterek

- i_1, i_2: bemenet
- h_1, h_2: rejtett réteg neuronjai
- o_1, o_2: kimeneti réteg neuronjai
- $w^k_{j,l}$: k-dik réteg j-dik neuronjának l-edik bemenete,
  pl.: $w^1_{1,2}$ 1-es réteg 1. neuronjának 2. bemenete ($i_2 \rightarrow h_1$)
- b_1, b_2: bias (jelen esetben egy adott rétegben ugyanaz a bias minden neuronra)
- t_1, t_2: elvárt kimenet

1.2. Egy node felépítése



2. ábra.

$$net_j = \sum_l (w^{(k)}_{j,l} \cdot in^{(k)}_l) + b_k \tag{1}$$

Aktivációs függvény: szigmoid

$$out_j = \frac{1}{1 + \exp^{-net_j}} \tag{2}$$

Szigmoid függvény deriváltja:

$$(out_j)^{'} = out_j \cdot (1 - out_j) \tag{3}$$

1.3. Rejtett réteg neuronjainak kimenete

h1 neuron:

$$net_{h_1} = w_{1,1}^{(1)} \cdot i_1 + w_{1,2}^{(1)} \cdot i_2 + b_1 \tag{4}$$

$$out_{h_1} = \frac{1}{1 + \exp^{-net_{h1}}} \tag{5}$$

h2 neuron:

$$net_{h_2} = w_{2,1}^{(1)} \cdot i_1 + w_{2,2}^{(1)} \cdot i_2 + b_1 \tag{6}$$

$$out_{h_2} = \frac{1}{1 + \exp^{-net_{h2}}} \tag{7}$$

1.4. Kimeneti réteg neuronjainak kimenete

o1 neuron:

$$net_{o_1} = w_{1,1}^{(2)} \cdot out_{h1} + w_{1,2}^{(2)} \cdot out_{h2} + b_2 \tag{8}$$

$$out_{o_1} = \frac{1}{1 + \exp^{-net_{o1}}} \tag{9}$$

o2 neuron:

$$net_{o_2} = w_{2,1}^{(2)} \cdot out_{h1} + w_{2,2}^{(2)} \cdot out_{h2} + b_2 \tag{10}$$

$$out_{o_2} = \frac{1}{1 + \exp^{-net_{o2}}} \tag{11}$$

1.5. Hiba számítása

$$E_{total} = \sum_j (t_j - out_{oj})^2 \tag{12}$$

$$E_{total} = E_{o_1} + E_{o_2} = (t_1 - out_{o_1})^2 + (t_2 - out_{o_2})^2 \tag{13}$$

1.6. Gradiens számítása egy kimeneti rétegbeli neuron súlyára

o_1 neuron $w_{1,1}^{(2)}$ súlyára, a láncszabály alkalmazásával:

$$\frac{\partial E_{total}}{\partial w_{1,1}^{(2)}} = \frac{\partial E_{total}}{\partial out_{o_1}} \cdot \frac{\partial out_{o_1}}{\partial net_{o_1}} \cdot \frac{\partial net_{o_1}}{\partial w_{1,1}^{(2)}} \tag{14}$$

$$\tag{15}$$

A hibára vonatkozó parciális derivált:

$$\frac{\partial E_{total}}{\partial out_{o_1}} = \frac{\partial (E_{o_1} + E_{o_2})}{\partial out_{o_1}} = \frac{\partial ((t_1 - out_{o_1})^2 + (t_2 - out_{o_2})^2)}{\partial out_{o_1}} \tag{16}$$

$$= (2 \cdot (t_1 - out_{o_1}) \cdot (-1)) + 0 \tag{17}$$

Láthatóan o_2 kimenete nem játszik szerepet, az ehhez tartozó parciális derivált értéke: 0
Az aktivációs függvényhez kapcsolódó parciális derivált:

$$\frac{\partial out_{o_1}}{\partial net_{o_1}} = \left(\frac{1}{1 + \exp^{-net_{o_1}}} \right)' = out_{o_1} \cdot (1 - out_{o_1}) \tag{18}$$

A súlyozott bemenethez kapcsolódó parciális derivált:

$$\frac{\partial net_{o_1}}{\partial w_{1,1}^{(2)}} = \frac{\partial (w_{1,1}^{(2)} \cdot out_{h1} + w_{1,2}^{(2)} \cdot out_{h2} + b_2)}{\partial w_{1,1}^{(2)}} \tag{19}$$

$$= out_{h1} + 0 + 0 \tag{20}$$

Tehát összességében:

$$\frac{\partial E_{total}}{\partial w_{1,1}^{(2)}} = -2 \cdot (t_1 - out_{o_1}) \cdot out_{o_1} \cdot (1 - out_{o_1}) \cdot out_{h1} \tag{21}$$

A kifejezés első két tagjához o_1 kimenete szükséges, összevontan δ_{o_1} -ként jelölhetjük. A harmadik tag pedig o_1 bemenete, ami az előző réteg kimenete out_{h1}. A hibára vonatkoztatott parciális derivált tehát az alábbi alternatív (delta szabály) formában is megadható:

$$\frac{\partial E_{total}}{\partial w_{1,1}^{(2)}} = \frac{\partial E_{total}}{\partial out_{o_1}} \cdot \frac{\partial out_{o_1}}{\partial net_{o_1}} \cdot \frac{\partial net_{o_1}}{\partial w_{1,1}^{(2)}} = \tag{22}$$

$$\frac{\partial E_{total}}{\partial w_{1,1}^{(2)}} = \frac{\partial E_{total}}{\partial net_{o_1}} \cdot \frac{\partial net_{o_1}}{\partial w_{1,1}^{(2)}} = \delta_{o_1} \cdot out_{h_1} \tag{23}$$

Adott α bátorsági tényezőt feltételezve a $w_{1,1}^{(2)}$ súly módosítását az alábbi képlet szerint számíthatjuk:

$$w_{1,1}^{(2)*} = w_{1,1}^{(2)} - \alpha \cdot \frac{\partial E_{total}}{\partial w_{1,1}^{(2)}} \tag{24}$$

1.7. Gradiens számítása egy rejtett rétegbeli neuron súlyára

h_1 neuron $w_{1,1}^{(1)}$ súlyára, a láncszabály alkalmazásával:

$$\frac{\partial E_{total}}{\partial w_{1,1}^{(1)}} = \frac{\partial E_{total}}{\partial out_{h_1}} \cdot \frac{\partial out_{h_1}}{\partial net_{h_1}} \cdot \frac{\partial net_{h_1}}{\partial w_{1,1}^{(1)}} \tag{25}$$

Ebben az esetben viszont mindkét kimenetre számított hibának szerepe lesz a hibára vonatkozó parciális deriváltban, mivel mindkettőt befolyásolja h_1 kimenete:

$$\frac{\partial E_{total}}{\partial out_{h_1}} = \frac{\partial E_{o_1}}{\partial out_{h_1}} + \frac{\partial E_{o_2}}{\partial out_{h_1}} \tag{26}$$

Az első tag:

$$\frac{\partial E_{o_1}}{\partial out_{h_1}} = \frac{\partial E_{o_1}}{\partial net_{o_1}} \cdot \frac{\partial net_{o_1}}{\partial out_{h_1}} \tag{27}$$

, ahol a szorzat első tényezője a korábbi számításokból rendelkezésre áll, hiszen:

$$\frac{\partial E_{o_1}}{\partial net_{o_1}} = \frac{\partial E_{o_1}}{\partial out_{o_1}} \cdot \frac{\partial out_{o_1}}{\partial net_{o_1}} \tag{28}$$

aminek két tényezője a 16. egyenlet és a 18. egyenlet alapján számítható. A 27. egyenlet második tényezője pedig az alábbi alakra egyszerűsödik:

$$\frac{\partial net_{o_1}}{\partial out_{h_1}} = \frac{\partial (w_{1,1}^{(2)} \cdot out_{h1} + w_{1,2}^{(2)} \cdot out_{h2} + b_2)}{\partial out_{h_1}} = w_{1,1}^{(2)} \tag{29}$$

Mindezek alapján a 27. egyenlet átalakítható a következő alakra:

$$\frac{\partial E_{o_1}}{\partial out_{h_1}} = \frac{\partial E_{o_1}}{\partial out_{o_1}} \cdot \frac{\partial out_{o_1}}{\partial net_{o_1}} \cdot \frac{\partial net_{o_1}}{\partial out_{h_1}} = \tag{30}$$

$$= \frac{\partial E_{o_1}}{\partial out_{o_1}} \cdot \frac{\partial out_{o_1}}{\partial net_{o_1}} \cdot w_{1,1}^{(2)} = \delta_{o_1} \cdot w_{1,1}^{(2)} \tag{31}$$

Hasonlóan a 26. egyenlet második tagja:

$$\frac{\partial E_{o_2}}{\partial out_{h_1}} = \frac{\partial E_{o_2}}{\partial out_{o_2}} \cdot \frac{\partial out_{o_2}}{\partial net_{o_2}} \cdot \frac{\partial net_{o_2}}{\partial out_{h_1}} = \tag{32}$$

$$= \frac{\partial E_{o_2}}{\partial out_{o_2}} \cdot \frac{\partial out_{o_2}}{\partial net_{o_2}} \cdot w_{2,1}^{(2)} = \delta_{o_2} \cdot w_{2,1}^{(2)} \tag{33}$$

mivel:

$$\frac{\partial net_{o_2}}{\partial out_{h_1}} = \frac{\partial (w_{2,1}^{(2)} \cdot out_{h1} + w_{2,2}^{(2)} \cdot out_{h2} + b_2)}{\partial out_{h_1}} = w_{2,1}^{(2)} \tag{34}$$

Így tehát a teljes hibára vonatkozó parciális deriváltat h_1 vonatkozásában:

$$\frac{\partial E_{total}}{\partial out_{h_1}} = \frac{\partial E_{o_1}}{\partial out_{h_1}} + \frac{\partial E_{o_2}}{\partial out_{h_1}} = \delta_{o_1} \cdot w_{1,1}^{(2)} + \delta_{o_2} \cdot w_{2,1}^{(2)} = \sum_{j \in O} \delta_j \cdot w_j \tag{35}$$

, ahol O a kimeneti réteg neuronjainak halmaza, továbbá δ_j és w_j egy adott j neuronhoz tartozó parciális deriváltak illetve a rejtett réteg beli h1-hez kapcsolódó súly.

A 25. egyenlet további tényezői a 18. és a 19. egyenletek alapján:

$$\frac{\partial out_{h_1}}{\partial net_{h_1}} = out_{h_1} \cdot (1 - out_{h_1}) \tag{36}$$

h_1 súlyozott bemenetéhez kapcsolódó parciális derivált:

$$\frac{\partial net_{h_1}}{\partial w_{1,1}^{(1)}} = \frac{\partial (w_{1,1}^{(1)} \cdot i_1 + w_{1,2}^{(1)} \cdot i_2 + b_1)}{\partial w_{1,1}^{(1)}} \tag{37}$$

$$= i_1 + 0 + 0 \tag{38}$$

Összességében tehát a $w_{1,1}^{(1)}$ súlyhoz tartozó gradiens:

$$\frac{\partial E_{total}}{\partial w_{1,1}^{(1)}} = \frac{\partial E_{total}}{\partial out_{h_1}} \cdot \frac{\partial out_{h_1}}{\partial net_{h_1}} \cdot \frac{\partial net_{h_1}}{\partial w_{1,1}^{(1)}} = \tag{39}$$

$$\left(\sum_{j \in O} \delta_j \cdot w_j \right) \cdot \frac{\partial out_{h_1}}{\partial net_{h_1}} \cdot \frac{\partial net_{h_1}}{\partial w_{1,1}^{(1)}} \tag{40}$$

, amely átírható a 22. egyenlethez hasonlóan (δ-szabály)

$$\frac{\partial E_{total}}{\partial out_{h_1}} \cdot \frac{\partial out_{h_1}}{\partial net_{h_1}} = \delta_{h_1} \tag{41}$$

$$\left(\sum_{j \in O} \delta_j \cdot w_j \right) \cdot \frac{\partial out_{h_1}}{\partial net_{h_1}} = \delta_{h_1} \tag{42}$$

$$\tag{43}$$

Így a 39. egyenlet alternatív formája:

$$\frac{\partial E_{total}}{\partial w_{1,1}^{(1)}} = \delta_{h_1} \cdot \frac{\partial net_{h_1}}{\partial w_{1,1}^{(1)}} \tag{44}$$

$$\frac{\partial E_{total}}{\partial w_{1,1}^{(1)}} = \delta_{h_1} \cdot i_1 \tag{45}$$

Ennek megfelelően a súlymódosítás a következő:

$$w_{1,1}^{(1)*} = w_{1,1}^{(1)} - \alpha \cdot \frac{\partial E_{total}}{\partial w_{1,1}^{(1)}} \tag{46}$$

$$= w_{1,1}^{(1)} - \alpha \cdot \delta_{h_1} \cdot i_1 \tag{47}$$