Bevezetés a lágy számítás módszereibe

Neurális hálózati modellek

Werner Ágnes

Villamosmérnöki és Információs Rendszerek Tanszék

e-mail: werner.agnes@virt.uni-pannon.hu

Egy neuronos perceptron hálózat

Bináris perceptron

- az x_i bemeneti jeleket összeszorozzák a hozzájuk tartozó w_i súlyokkal
- ezeket összegezve nyerik az aktivizációs függvény net bemenetét
- a **kemény limtáló aktivizációs függvény** $o \in \{-1, 1\}$ bipoláris bináris **aktuális kimenet**et szolgáltat
- ezt összevetjük a $d \in \{-1, 1\}$ **kívánt bináris** jellel
- a kettő különbsége az a **hiba**, amelyet a súlyok változtatására használnak
- Rosenblatt-féle tanulási szabály valamely k időpontban valamelyik w_i **súly módosítására**:
  $w_{ij}^{k+1} = w_{ij}^k + \eta e^k x_j^k$ ahol $e^k = d^k - o^k$ és
  $o^k = sgn(net^k) = sgn(\sum w_{ij} x_j)$
  e^k a hiba, o^k az aktuális kimenet és η a tanulási ráta
  (szabályozza az algoritmus konvergenciájának sebességét)

Példa 1

Feladat egy olyan diszkrét perceptron tervezése, amelyik a kétdimenziós bementi tér négy mintáját az alábbi módon sorolja 2 osztályba:

$$\mathbf{x}^{(1)} = \begin{bmatrix} -1 \\ 0 \end{bmatrix} \qquad \mathbf{x}^{(2)} = \begin{bmatrix} -1.5 \\ -1 \end{bmatrix} \qquad 1.\ osztaly \quad d^{(1)} = d^{(2)} = 1$$

$$\mathbf{x}^{(3)} = \begin{bmatrix} 2.5 \\ -1 \end{bmatrix} \qquad \mathbf{x}^{(4)} = \begin{bmatrix} 1 \\ -2 \end{bmatrix} \qquad 2.\ osztaly \quad d^{(3)} = d^{(4)} = -1$$

Példa megoldása

- a két osztály mintái lieárisan szétválaszthatók
- az osztályokat szétválasztó **döntési egyenes** pl.

$$e(\mathbf{x}) = -x_1 + 2x_2 + 2 = 0$$

- két féltér: az $e(\mathbf{x}) > 0$ féltér az 1. osztály, az $e(\mathbf{x}) < 0$ féltér a 2. osztály mintáit tartalmazza
- a döntési egyenes nem halad át az origón: T küszöb érték szükséges

$$o^k = sgn[w_1 x_1^k + w_2 x_2^k - T]$$

- $w_1 = -1$, $w_2 = 2$ és $T = -2$

Példa megoldás - diszkrét perceptron tanulása 1

A diszkrét perceptron tanulási szabály szerint:

$$\mathbf{w}^{k+1} = \mathbf{w}^k + \eta(d^k - o^k)\mathbf{x}$$

ahol

$$o^k = sgn[w_1 x_1^k + w_2 x_2^k - w_3]$$

Válasszuk az indulási súlyvektort: $\mathbf{w}^t = [-1\ 2\ \ -2]$
Ekkor pl. az 1. bemeneti mintavektor kimenete:

$$o^1 = sgn[\mathbf{w}^{1t}\mathbf{x}^1] = sgn[-1\ 2\ \ -2]\begin{bmatrix} -1 \\ 0 \\ -1 \end{bmatrix} = sgn[1+0+2] = +1$$

A hiba: $d^1 - o^1 = 1 - 1 = 0$

Példa megoldás - diszkrét perceptron tanulása 2

Hasonlóan helyes osztályozást kapunk a többi bemeneti mintára.
Válasszuk az indulási súlyvektort: $\mathbf{w}^t = [-1\ 1\ 2]$

$$o^1 = sgn[-1\ 1\ 2]\begin{bmatrix} -1 \\ 0 \\ -1 \end{bmatrix} = sgn[1\ 0\ -2] = -1$$

pl. az $\eta = 0.25$ értékkel a

$$\mathbf{w}^{2t} = \mathbf{w}^{1t} + \eta(d^1 - o^1)\mathbf{x}^1 = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} + 0.25[1-(-1)]\begin{bmatrix} -1 \\ 0 \\ -1 \end{bmatrix} =$$

$$\begin{bmatrix} -1.5 \\ 1 \\ 1.5 \end{bmatrix}$$ súlymódosítás szükséges

Ezt a többi mintára is mindaddig el kell végezni, amíg a helyes osztályozást el nem érjük.

Példa megoldás - diszkrét perceptron tanulása 3

Helyes osztályozás esetén a 3D mintatérben döntési sík lesz, amely a 2 osztály mintáit különválasztja.
Ennek a síknak át kell mennie az
$e(\mathbf{x}) = o = sgn[w_1 x_1 + w_2 x_2 - w_3] = 0$ egyenesen
és a $\mathbf{w}^t\mathbf{x} = 0$ szerint az $x = 0$ kezdőponton.