Rješenje zadatka 3.2 predmeta Strojno učenje

Siniša Biđin

5. siječnja 2013.

(a) (i) Ne možemo je izravno minimizirati jer je "većinom konstantna", pa gradijentni spust nije primjenjiv.

(ii) Kriterij perceptrona konveksna je funkcija jer je i sama korištena funkcija gubitka konveksna funkcija.

$$E_p(\tilde{\mathbf{w}}|\mathcal{D}) = \sum_{i=1}^{N} \max\left(0, -\tilde{\mathbf{w}}^{\mathrm{T}} \boldsymbol{\phi}(\mathbf{x}^{(i)})\mathbf{y}^{(i)}\right)$$

Bitno nam je da je funkcija konveksna, jer želimo imati garanciju da ćemo joj gradijentnim spustom pronaći globalni minimum.

(iii) U slučaju nijedne greške vrijedi $E_p = E_m = 0$, dok za samo jednu grešku vrijedi $E_p = 1, E_m = 1/N, E_p \geq E_m$. Pritom je E_m ograničena na raspon $[0, 1]$, dok E_p nije. Povećanjem broja grešaka razlika između E_p i E_m stoga samo nastavlja rasti, pa je E_p gornja ograda za E_m.

(b) Ispisujemo detaljno samo korake u kojima se vektor težina mijenja, za skup podataka $D = \{((-3, 1), 1), ((-3, 3), 1), ((1, 2), -1), ((2, 1), -1)\}$ i početni vektor težina $\tilde{\mathbf{w}} = [0, 0, 0]$:

$$\begin{aligned}
\mathbf{x}^{(1)}, \mathbf{x}^{(2)} &: \text{Uspješno klasificirani.} \\
\mathbf{x}^{(3)} &: \tilde{\mathbf{w}} = \tilde{\mathbf{w}} + [-0.005, -0.005, -0.01] = [-0.005, -0.005, -0.01] \\
\mathbf{x}^{(4)} &: \text{Uspješno klasificiran.} \\
\mathbf{x}^{(1)} &: \tilde{\mathbf{w}} = \tilde{\mathbf{w}} + [+0.005, -0.015, +0.005] = [0, -0.02, -0.005] \\
\mathbf{x}^{(2)}, \mathbf{x}^{(3)}, \mathbf{x}^{(4)} &: \text{Uspješno klasificirani.} \\
\mathbf{x}^{(i)} &: \text{Uspješno klasificirani; } \textit{konvergencija.}
\end{aligned}$$

Postupak konvergira; konačan $\tilde{\mathbf{w}} = [0, -0.02, -0.005]$.

(c) Perceptron nastavlja s radom sve dok postoji barem jedan neuspješno klasificiran primjer. Pošto primjeri nisu linearno odvojivi, nužno uvijek postoji barem jedan primjer koji će biti neuspješno klasificiran.