

T-61.231 Hämmontunnistuksen perusteet

Laskuharjoitus 4: 14.10.2002

1. Luokittimen virhettä pyritään usein estimoimaan testijoukolla (virheprosentti lasketaan testivektoreista). Joskus testijoukko on sama kuin millä luokitin rakennettiin. Osoita, että tämä ei ole hyvä tapa erityisesti jos kyseessä on 1-NN-luokitin.
2. Oletetaan että kaksi pistejoukkoa H_1 ja H_2 R^d :ssä ovat lineaarisesti separoituvia. Tämä tarkoittaa että on olemassa vektori $\underline{\omega}' \in R^d$ ja skalaari $\omega_0 \in R$ siten että

$$\underline{\omega}'^T \underline{x} - \omega_0 \begin{cases} > 0 & \forall \underline{x} \in H_1 \\ < 0 & \forall \underline{x} \in H_2 \end{cases}$$

Kasvatetaan vektoreita seuraavasti:

$$\hat{\underline{x}} = \begin{pmatrix} \underline{x} \\ -1 \end{pmatrix}, \quad \underline{\omega} = \begin{pmatrix} \underline{\omega}' \\ \omega_0 \end{pmatrix}.$$

Ja vaihdetaan lopuksi merkit joukon H_2 vektoreista $\hat{\underline{x}}$.

- a) Osoita että $\underline{\omega}^T \hat{\underline{x}} > 0 \quad \forall \hat{\underline{x}}$.
 - b) Näytä tilanne graafisesti, kun $d = 1$ ja $H_1 = \{-3, -2, -1\}$, $H_2 = \{5, 6, 7\}$.
3. Perceptronin oppimissäännössä päivitys $\underline{\omega}^{(n+1)} = \underline{\omega}^{(n)} + \alpha_n \hat{\underline{x}}_i$ tehdään, kun $\hat{\underline{x}}_i$ on luokiteltu väärin. Eli $e_i = \underline{\omega}^{(n)T} \hat{\underline{x}}_i < 0$. Kuinka α_n pitäisi alita jotta $\hat{\underline{x}}_i$ luokiteltaisiin varmasti oikein päivityksen jälkeen?
 4. Olkoon H_1 ja H_2 lineaarisesti separoituvia ja syötteen $x_1 = -2 \in H_1$, $x_2 = -1 \in H_2$ ja $x_3 = 2 \in H_2$ opetusnäytteitä. Olkoon tehtävä 3:n yhtälön painovektorin alkuarvo $\underline{w}^0 = [1, 1]^T$ ja parametri $\alpha = 0.5$. Opetta perceptroni käyttäen näitä näytteitä munneltuina tehtävän 2 mukaisesti.