A gépi tanulás elmélete - Házi feladat v2.0

Danner Gábor

December 16, 2016

Feladat: Legfeljebb hány iterációt hajt végre a perceptron algoritmus az n-változós többségi függvény tanulásakor?

Feltettem, hogy n páratlan. Szükségünk lesz bias-re is, ezért legyen $X = \{0,1\}^n \times \{1\}$. Továbbá legyen $Y = \{-1, 1\}$ a címkék halmaza, $Z = \{(x, y) : x \in X, y \in Y, \sum_{i=1}^{n} x_i > \frac{n}{2} \iff y = 1\}$ a lehetséges címkézett példák halmaza.

Legyen $B = \min\{\|w\| : \forall(x, y) \in Z, y\langle w, x\rangle \geq 1\}$. Teljesül, hogy $B \leq \|(2, \ldots, 2, -n)\| = \sqrt{n2^2 + (-n)^2} = \sqrt{n^2 + 4n}$, mivel a $w = (2, \ldots, 2, -n)$ vektorra teljesül, hogy $\forall(x, y) \in Z, y\langle w, x\rangle \geq 1$. (Belátható[1], hogy $B = \|(2, \ldots, 2, -n)\|$ is teljesül.)

Legyen $R = \max\{\|x\| : x \in X\}$. Teljesül, hogy $R = \|(1, \ldots, 1, 1)\| = \sqrt{n+1}$.

Alkalmazva a tankönyv 9.1-es tételét megkapjuk, hogy a Perceptron algoritmus legfeljebb $R^2B^2 \leq (n + 1)(n^2 + 4n) = n^3 + 5n^2 + 4n$ lépésben megáll.

[1] Legyen $(x, y) \in Z, y\langle w, x\rangle \geq 1$. Tegyük fel, hogy az x vektorban (a bias-t nem számítva) eggyel több 1-es van, mint 0-s. Ekkor $y = 1$, és $\langle w, x\rangle \geq 1$. Továbbá legyen $(x', y') \in Z, y'\langle w, x'\rangle \geq 1$ úgy, hogy x'-t megkaphassuk x-ből egy 1-es 0-ra változtatásával, a változatás pozícióját jelöljük k-val ($1 \leq k \leq n$). Ekkor $y' = -1$, és $\langle w, x'\rangle \leq -1$. Ebből azt kapjuk, hogy $\langle w, x\rangle - \langle w, x'\rangle \geq 2$, azaz $\langle w, x - x'\rangle \geq 2$. Az $x - x'$ vektor a k pozíción 1, a többin 0, így $w_k \geq 2$. Ennek tetszőleges k-ra teljesülnie kell, mivel minden k-hoz létezik megfelelő $(x, y), (x', y') \in Z$.

Legyen $(x, y) \in Z, y\langle w, x\rangle \geq 1$. Tegyük fel, hogy az x vektorban (a bias-t nem számítva) pontosan $\frac{n-1}{2}$ 1-es van. Ekkor $y = -1$, és $-1 \geq \langle w, x\rangle \geq 2\frac{n-1}{2} + w_{n+1}$, mivel w első n komponense legalább 2. Átrendezve azt kapjuk, hogy $w_{n+1} \leq -n$.

Belátható, hogy $w = (2, \ldots, 2, -n)$ eleget tesz a feltételeknek, továbbá minimális hosszúságú is, mivel a feltételeket figyelembe véve komponensei egyenként minimálisak abszolútértékben, és a vektorhossz monoton nő a vektor komponenseinek abszolútértékének függvényében.