# Semestrální test 4 - oprava

pondělí 16. ledna 2023 22:32

$T = \{(1;+1), (3;+1), (0;+1), (5;-1), (3;+1), (6;-1), (-3;-1), (-4;-1), (8;-1), (-3;-1)\}$

dvojic $(x;y)$, kde $x \in \mathbf{R}$ je měření a $y \in \{-1,+1\}$ je třída. Dále máme k dispozici nekonečnou množinu slabých klasifikátorů $\mathcal{H} = \{h(x;c,p)\}$ parametrizovaných jejich centrální pozicí $c \in \mathbf{R}$ a paritou $p \in \{-1,+1\}$:

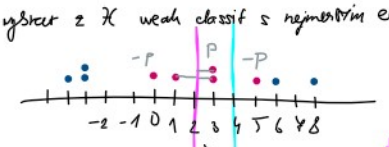
$$h(x;c,p) = \begin{cases} p, & \text{if } |x-c| \le 1 \\ -p, & \text{if } |x-c| > 1 \end{cases}$$

a) Napište formálně AdaBoost algoritmus *(1 bod)*

1, $D_1(i) = \frac{1}{10}$

$N \leftarrow 10$

2, chci vybrat z $\mathcal{H}$ weak classif s nejmenším errorem




$\boxed{c=2}$

$c=-1$

$\varepsilon_t = \sum D_t(i) \cdot [\![ y_i + h(x_i) ]\!] = \frac{1}{10} \cdot 2 = \frac{2}{10} = 0.2$

$\alpha_t = \frac{1}{2} \ln\left(\frac{1-\varepsilon_t}{\varepsilon_t}\right) = \frac{1}{2} \cdot \ln\left(\frac{1-0.2}{\underbrace{0.2}_{4}}\right) = \boxed{0.693}$

$D_{t+1}(i) = \frac{1}{Z_t} \cdot \underbrace{D_t(i)}_{\frac{1}{10}} \cdot e^{-\alpha_t y_i h_t(x_i)}$

$\underbrace{2 \cdot 0.1994 + 8 \cdot 0.05}_{= 0.7988}$

pro dobře klasif. $\underbrace{e^{-0.69}}_{\frac{1}{10} \cdot e^{-0.69} = 0.05}$

špatně $\underbrace{e^{0.69}}_{\dots = 0.1994}$

$D_{t+1}(i) = 0.0625$, $0.2496$

2. iterace

$\varepsilon_t = 0.2496 \cdot 2$

$f(x) = \sum_{i=1}^{5} \alpha_t h_t(x) = \alpha_1 \cdot h_1(x,2,p) + \alpha_2 \cdot (x, \ ,p)$

$\alpha_1 = 0.693$

**Úloha 2.** Mějme následující dvourozměrná data:

$\{(8,5), (1,4), (3,0), (10,4), (7,6), (0,1), (4,3), (9,7), (2,2), (9,4)\}$.

a) Popište formálně algoritmus k-medians. *(1 bod)*

b) Proveďte dvě iterace algoritmu, kde počáteční pozice center jsou $(5,3), (2,5)$. *(1 bod)*

a) k-medians

↳ inicializace clusterů $\{c_i\}_{i=1}^{k}$

dist $= \|x_1 - c_i\|_1$

argmin dist

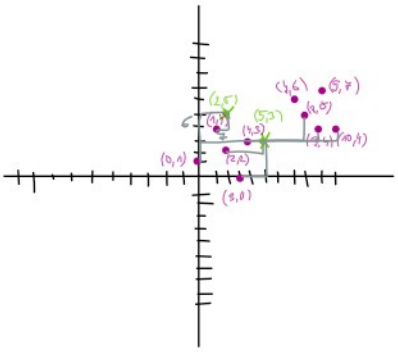
} minimalizuju vzdálenost ke clusterům $c_i$

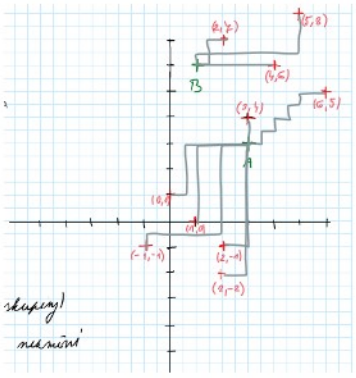
b) 2 iterace k-medians

$c_1 = (5,3)$

$c_2 = (2,5)$

$\{(8,5), (1,4), (3,0), (10,4), (7,6), (0,1), (4,3), (9,7), (2,2), (9,4)\}$




→ B

x: 2, 4, 5

y: 6, 7, 8

} nové B=(4,7)

medián pro A

osa x : −1, 0, 1, 2, 2, 3, 6

y : −2, −1, −1, 0, 1, 4, 5

→ nové A=(2,0)