Schreiroval

RPZ test 1, varianta A

15. 10. 2021

Úloha 1

Mějme $p(k=1)=0.4$ a sdruženou pravděpodobnost $p(x,k)$ danou tabulkou.

$p(x,k)$	$x=1$	$x=2$	$x=3$	$x=4$
$k=1$	0.01	A	0.11	0.15
$k=2$	0.25	0.24	B	0.01

$W(d,k)$	$k=1$	$k=2$
$d=1$	0	2
$d=2$	1	0

a) Doplňte chybějící hodnoty A, B. *(0.5 bodu)*

b) Najděte optimální Bayesovskou strategii pro pozorování x, rozdělení $p(x,k)$ a ztrátu $W(d,k)$. *(1 bod)*

$q^* = \{2, 2, 2, 1\}$

c) Vypočítejte Bayesovský risk nalezené optimální strategie. *(1 bod)*

$R(q^*) = 0,112$

Úloha 2

Předpokládejme že máme rozhodovací problém o dvou třídách $k \in \{1,2\}$ se spojitými pozorováními $x \in \langle -1;1\rangle$ a známé jsou jen podmíněné pravděpodobnosti $p(x|k=1)=\frac{1}{2}$ a $p(x|k=2)=\max(\frac{1}{2}x+\frac{1}{4}; -\frac{1}{2}x+\frac{1}{4})$.

a) Napište *fomální* definici Neyman-Pearson úlohy. *(1 bod)*

b) Najděte optimální Neyman-Pearson strategii pokud $k=2$ je nebezpečný stav a pravděpodobnost přehlédnutí nebezpečí nesmí být větší než 0.2. (pomůcka: $\sqrt{2.6} \approx 1.6$) *(2.5 bodu)*

Bayesovská formulace

$$p(a,b) = p(a|b)\,p(b) = p(b|a)\,p(a)$$

$P(k=1) = 0.4$

$P(x,k)$

k \ x	1	2	3	4	
1	0,01	0,13	0,11	0,15	0,4
2	0,25	0,24	0,1	0,01	0,6
	0,26	0,37	0,21	0,16	1

W(d,k)	k=1	k=2
d=1	0	2
d=2	1	0

↓ ztráta při daném rozhodnutí

k≠D				
1≠1	0,01 · 1 = 0,01	0,13 · 1 = 0,13	0,11 · 1 = 0,11	0,15 · 1 = 0,15 ✓
2≠2	0,25 · 2 = 0,5 ^	0,24 · 2 = 0,48 ^	0,1 · 2 = 0,2 ^	0,01 · 2 = 0,02
	2	2	2	1

$q^* = \{2,2,2,1\}$

$$R = \sum_k \sum_x p(x,k) \cdot W(q(x),k)$$

$k=1 \quad d=2 \rightarrow 1 \cdot (0,01 + 0,13 + 0,11) = 0,25$

$k=2 \quad d=1 \rightarrow 2 \cdot 0,01 = 0,02$

$R(q^*) = 0,25 + 0,02 = 0,27$

$$q^* = \underset{q\ X \to D}{\operatorname{argmin}}\ R(q)$$

② Neyman-Pearson

$K = \{N, D\}$ $\quad q^* = \arg\min_q \varepsilon_N$; $\quad \varepsilon_N = \sum_{x: q(x) \neq N} p(x|N)$ false alarm

s.t. $\varepsilon_D \leq \varepsilon$ $\quad \varepsilon_D = \sum_{x: q(x) \neq D} p(x|D)$ overlooked alarm

ε ... uživatelem stanovená chybovost

Řešení podle LR

$$LR = \frac{p(x|N)}{p(x|D)}$$

$LR > \mu \rightarrow q(x) = N$

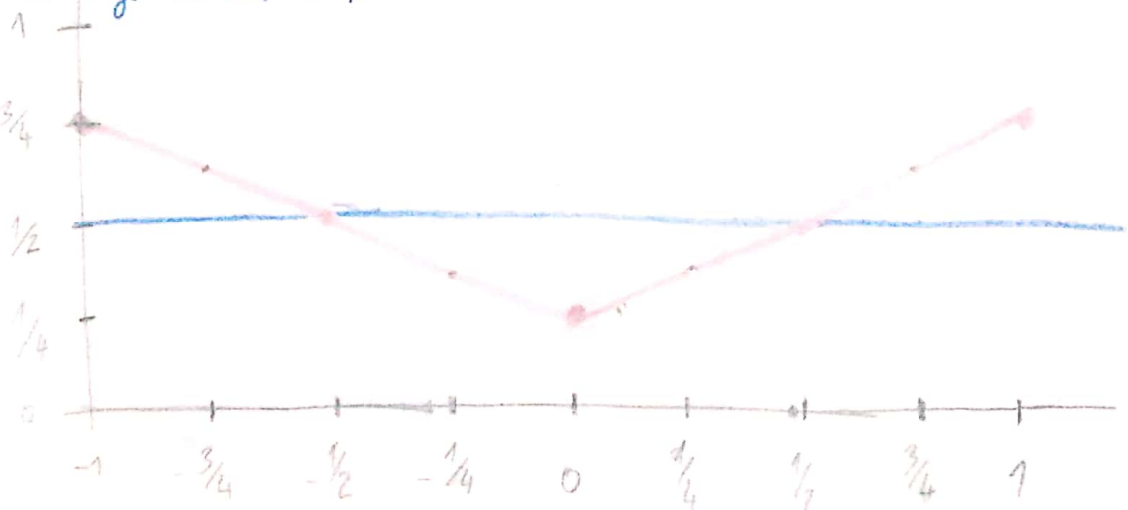
$LR \leq \mu \rightarrow q(x) = D$

Problém: $k \in \{1, 2\}$; $x \in \langle -1, 1 \rangle$

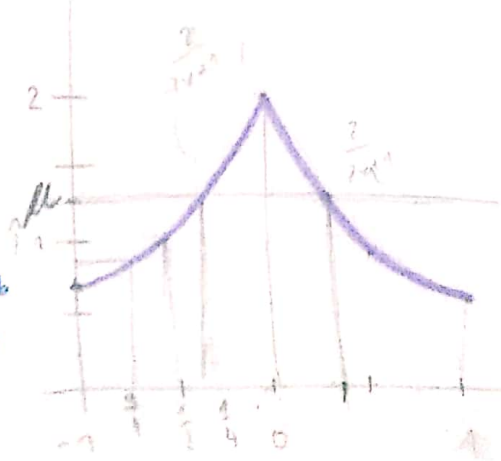
$p(x \mid k=1) = \frac{1}{2}$

$p(x \mid k=2) = \max\left(\frac{1}{2}x + \frac{1}{4}\,;\, -\frac{1}{2}x + \frac{1}{4}\right)$

$k=2$ je nebezpečný stav $\quad \varepsilon = 0{,}2$



$$LR = \frac{p(x|N)}{p(x|D)}$$



$1/2 : 1/4 = 2 \rightarrow 0$

$1/2 \cdot 1/2 = 1 \rightarrow 1/2 \cdot 1/2$

$1/2 \cdot 3/4 = 3/8 \rightarrow -1 \cdot 1$

$$2 \int_0^b \frac{1}{2}x + \frac{1}{4}\,dx =$$

$$= 2\left[\frac{x^2}{4} + \frac{x}{4}\right]_0^b = \frac{1}{2}\left[x^2 + x\right]_0^b$$

$$= \frac{1}{2}(b^2 + b)$$

$\frac{1}{2}(b^2 + b) = 0{,}2$

$b^2 + b - 0{,}4 = 0$ nikoo into

$b_1 = 0{,}3$; $b_2 = -1{,}3$

$$\mu = \frac{2}{2x+1}\Big|_{0,3} = 1{,}25$$

RPZ test 2, varianta A 5. 11. 2021

3.5

Jméno: Schreierova Magdalena

Úloha 1

Stojíte před fotbalovým stadionem, ze kterého vychází agresivní a poklidní fanoušci.

0.5 a) Formálně definujte úlohu odhadu pravděpodobnosti ω, že je vycházející fanoušek agresivní, jako úlohu odhadu pomocí metody maximální věrohodnosti. *(1 bod)*

1.5

1 b) Odvoďte (s postupem) analyticky vzorec pro maximálně věrohodný odhad $\hat{\omega}_{ML}$. *(1 bod)*

Úloha 2

Mějme dána měření $X = \{2, 3, 1, -1, 1, 3\}$.

✓ a) Vykreslete graf neparametrického odhadu hustoty $p(x)$ pomocí metody Parzenových oken. Uvažujte jádrovou funkci (kernel) $K(x, y) = k(x - y)$, kde $k(z)$ je definována jako

2

$$k(z) = 1/h \quad \text{pro } |z| \leq h/2 \,,$$
$$k(z) = 0 \quad \text{pro } |z| > h/2 \,,$$

a $h = 2$. Dbejte na to, aby graf zobrazoval přesné hodnoty distribuce. *(2 body)*

✗ b) Vypočítejte $p(x = 2.5)$ pomocí k-NN odhadu s $k = 3$ *(1 bod)*

Test 2

① MLE

$$\omega^* = \arg\max_{\omega} \prod_i p(x_i \mid \omega)$$

M → počet agresivních

N → počet klidných

ω → pst agresivního

$$L = \omega^M \cdot (1-\omega)^N$$

$$l = M \cdot \ln \omega + N \cdot \ln(1-\omega)$$

$$\frac{\partial l}{\partial \omega} = M \cdot \frac{1}{\omega} + N \cdot \frac{-1}{1-\omega} = 0$$

$$\frac{M}{\omega} = \frac{N}{1-\omega}$$

$$M - M\omega = N\omega$$

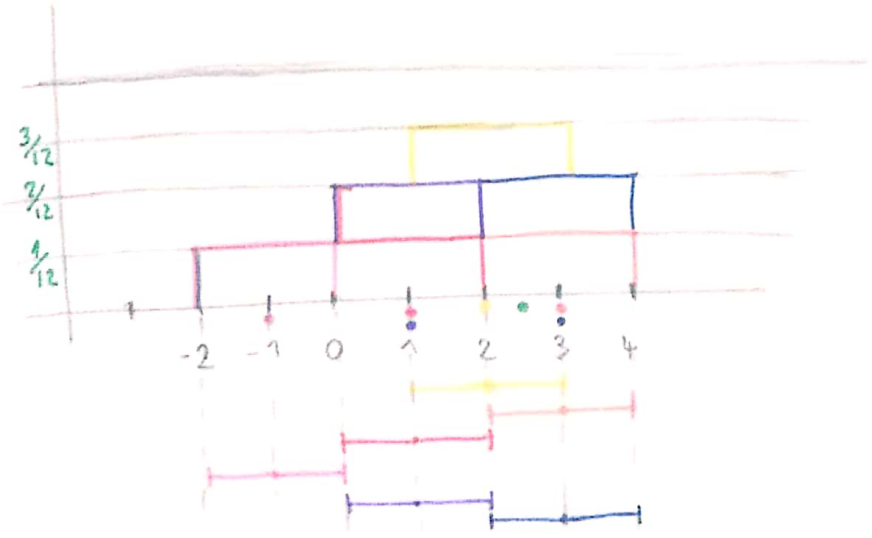
$$\hat{\omega}_{ML} = \frac{M}{M+N}$$

② $X = \{2, 3, 1, 1, -1, 1, 3\}$, $n = 6$

$h = 2$

$k(z) = \frac{1}{h} = \frac{1}{2}$ pro $|z| \leq 1$

výška → $\frac{1}{h} \cdot \frac{1}{n} = \frac{1}{2} \cdot \frac{1}{6} = \frac{1}{12}$



k-NN $k = 3$

$$p(x = 2{,}5) = \frac{3}{6 \cdot 1} = \frac{3}{6} = \frac{1}{2}$$

3 → k

6 → n

1 → velikost boxu, ve kterém jsou všechny body nejblíže a uprostřed je x

RPZ test 3, varianta B

2,5

Jméno: Schreierová

3. 12. 2021

Úloha 1

a) Formálně definujte empirické riziko. Vysvětlete všechny použité symboly. *(1 bod)*

b) Vysvětlete, proč strategie s malým empirický rizikem nemusí v praxi fungovat dobře. *(1 bod)*

Úloha 2

Máme trénovací sadu $\mathcal{T} = \{([-3, 1]; 2), ([-2, 2]; 2), ([-2, 1]; 1), ([-1, 3]; 1)\}$ s měřeními $\mathbf{x}_i \in \mathbb{R}^2$ a labely $y_i \in \{1, 2\}$.

a) Proveďte 3 iterace učení Perceptronu (přehledně vypisujte důležité hodnoty v každé iteraci). *(1 bod)*

b) Vyberte a zdůvodněte (!) jednu možnost. Pro zadaná data Perceptron: *(1 bod)*

0.5

(i) nikdy nenalezne řešení

(ii) nalezne řešení až po 5-ti nebo více krocích

(iii) může nalézt řešení v méně než 5-ti krocích

Úloha 3

Mějme funkci $K(x, y) = 2x - y$, kde $x, y \in \mathbb{R}^2$. Je $K(x, y)$ validní kernel? Pokud ano, najděte odpovídající lifting $\Phi(x)$, pokud ne, ukažte, že žádný lifting neexistuje. *(1 bod)*

Test S B1

① a) $R_{emp} = \frac{1}{N} \sum_{i=1}^{N} W(q(x_i), y_i)$

$\frac{1}{N}$ ↓ počet vzorů; $W(q(x_i), y_i)$ ztrátová funkce

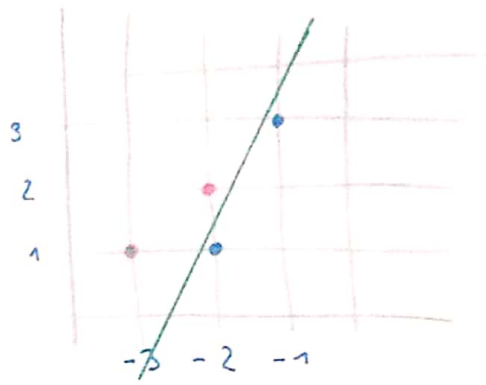
b) p+ž. může dojít k přefitování

② Perceptron třída 2 = (-1)

a)

třída	2	2	1	1
$\tilde{x}$	+3	+2	-2	-1
	-1	-2	1	3
	-1	-1	1	1
i	1	2	3	4

b)



Řešení lze nalézt, ale bude to trvat dlouho, protože přímka je daleko od počátku

$v = (w | b)$

① $v = (0, 0, 0)$

$v\tilde{x}_i \leq 0$ pro všechny $\rightarrow v = v + \tilde{x}_4$

② $v = (-1, 3, 1)$

$v\tilde{x}_1 = -3 - 3 - 1 \leq 0 \rightarrow v = v + \tilde{x}_1 = (2, 2, 0)$

③ $v = (2, 2, 0)$

$v\tilde{x}_3 = -4 + 2 + 0 \leq 0 \rightarrow v = v + \tilde{x}_3 = (0, 3, 1)$

③ $k(x, y) = 2x - y \quad x, y \in \mathbb{R}^2$

$2\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in \mathbb{R}^2$

$k(x, x') = \phi(x) \cdot \phi(x') \in \mathbb{R}$

Lze není kernel

Lifting neexistuje, při zvyšování dimenzí bychom se do $\mathbb{R}$ nedostali.