#B241/RPZ

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)*

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

N ... počet všech

$W(q(x_i), y_i)$ ... ztrátová funkce

(b) model může jednoduše přetrénovat

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)*

(i) nikdy nenalezne řešení

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

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

Nápověda: ~~Novikoff~~, w?, b!

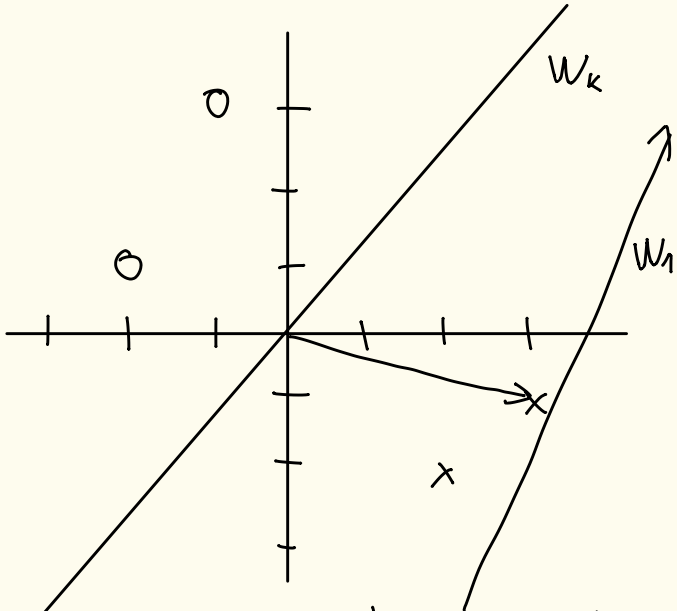
(a) $\mathcal{T} = \{([-3,1];2), ([-2,2];2), ([-2,1];1), ([-1,3];1)\}$

tříd 2 změním na -1

$\mathcal{T}' = \{([3,-1];-1), ([2,-2],-1), ([-2,1],1), ([-1,3];1)\}$

$\mathcal{T}'' = \left\{\begin{bmatrix}3\\-1\\-1\end{bmatrix}, \begin{bmatrix}2\\-2\\-1\end{bmatrix}, \begin{bmatrix}-2\\1\\1\end{bmatrix}, \begin{bmatrix}-1\\3\\1\end{bmatrix}\right\}$

0. it	$v = (0;0;0)$	$(3,-1,-1)$
1. it	$v = (3;-1;-1)$	$(-2,1,1)$
2. it	$v = (1,0,0)$	$(-2,1,1)$
3. it	$v = (-1,1,1)$	$(3,-1,-1)$
4. it	$v = (2,0,0)$	$\vdots$



(b) jelikož máme lineárně separabilní data, algoritmus perceptronu nalezne řešení, jenom to může chvíli trvat, tedy (ii)

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.

Kernel funkce musí být symetrická a pozitivně semidef. Symetrie:

$$k(x,y) = k(y,x)$$
$$2x - y \neq 2y - x \quad \lightning$$

Funkce není kernel.

Soft margin SVM klasifikátor

SVM: (primár)

$$w^*, b^*, \xi_i^* = \underset{w,b,\xi_i}{\arg\min} \left(\frac{1}{2}\|w\|^2 + C\sum_{i=1}^{m} \xi_i \right)$$

$$\text{s.t. } \langle w_i, x_i \rangle + b \geq +1 - \xi_i \qquad y_i = +1$$
$$\langle w_i, x_i \rangle + b \leq -1 + \xi_i \qquad y_i = -1$$
$$\xi_i \geq 0, \forall i$$

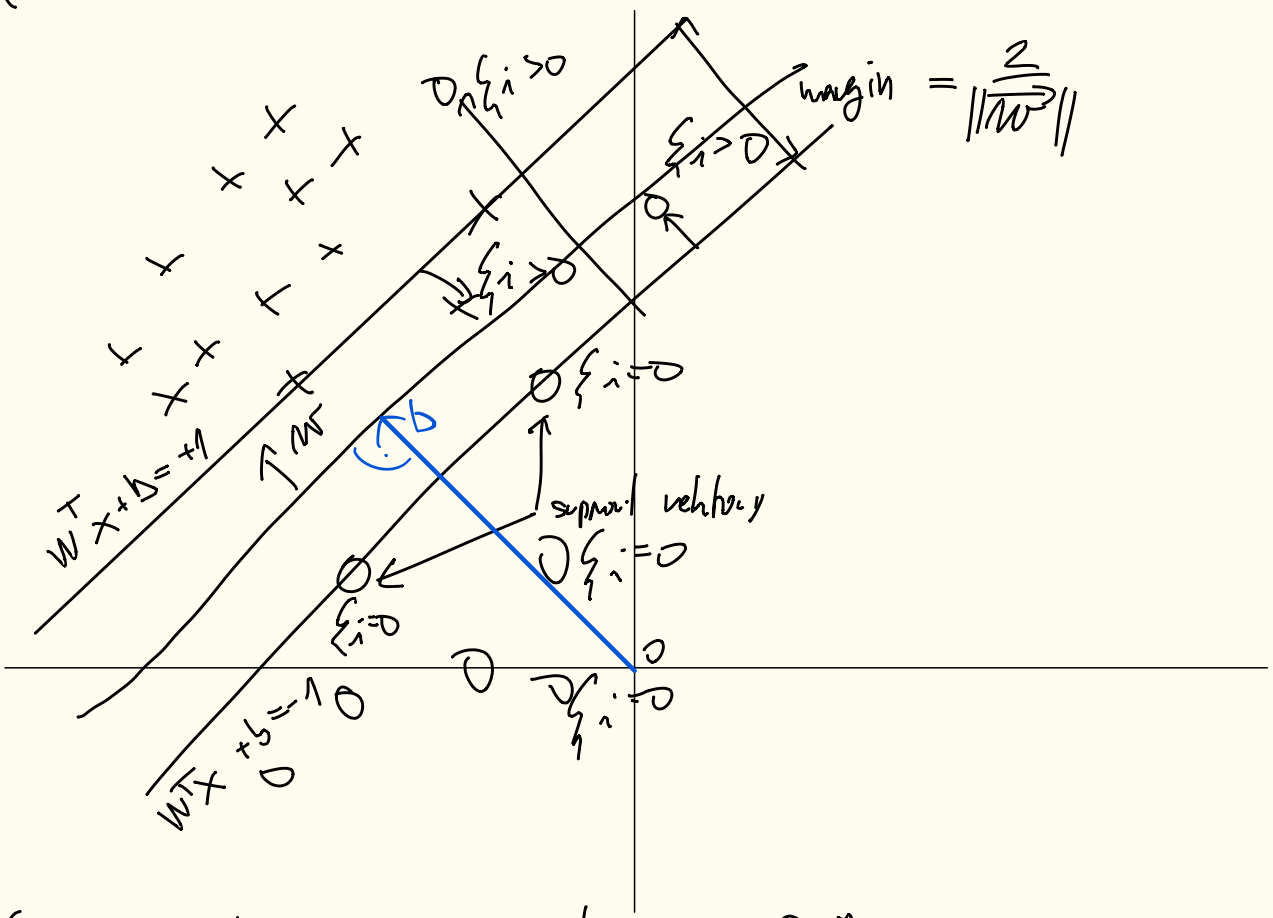
(duál)

$$\alpha^* = \underset{\alpha}{\arg\max} \left(\sum_{i=1}^{m} \alpha_i - \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m} \alpha_i y_i k(x_i, x_j) y_j \alpha_j \right)$$
$$\text{s.t. } C \geq \alpha_i \geq 0 \quad i = 1 \dots m \qquad \sum_{i=1}^{m} \alpha_i y_i = 0$$

klasifikace SVM

$$h(x', \alpha^*, b^*) = \text{sign}\left(\sum_{i=1}^{m} \alpha_i^* y_i k(x_i', x_i) + b^*\right)$$

(b) chceme se obecně vrátit mezi všech možných parametrů

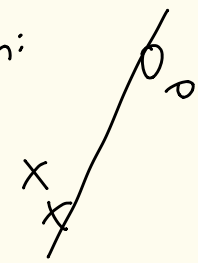


(c) rozdíly mezi perceptronem a SVM

Perceptron nevrátí lineárně neseparabilní data (soft margin SVM ano)

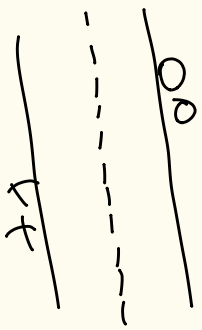
SVM maximalizuje strukturu jako optimální hranici rozdělení, perceptron se zastaví jakmile najde řešení, třeba

perceptron:



je validní rozdělení, ale není optimální

SVM:

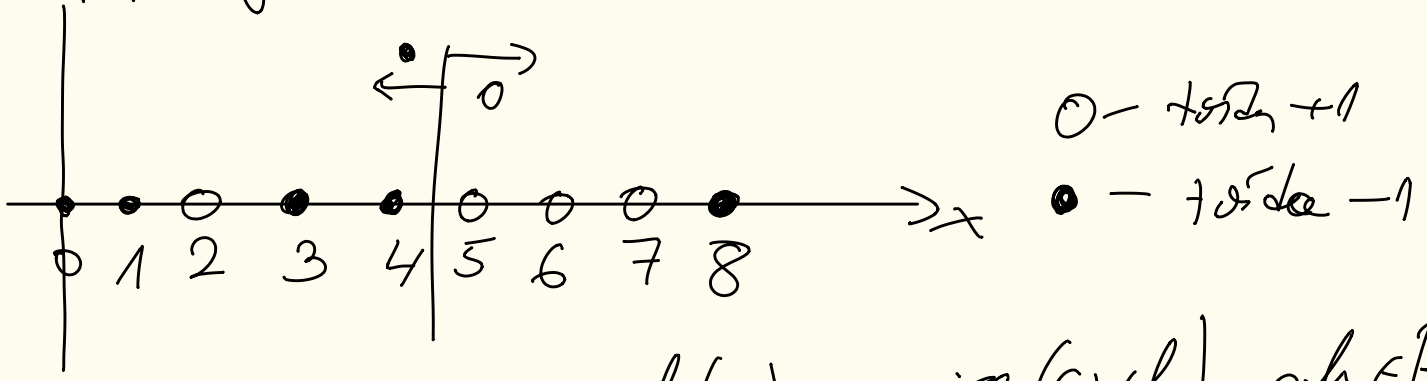


SVM využívá kernel k (perceptron ale jednoduchý lze)

perceptron má jednodušší implementaci

AdaBoost

pro tato jednorozměrná data



a sadu klasifikátorů $h(x) = \text{sign}(ax+b)$, $a,b \in R$

najděte první slabý klasifikátor, jeho váhu a spočítejte váhu dat pro následující kolo

$D_1(i) = 1/N$ pro $i = 1 \dots N$

$D_1(i) = 1/9 \quad \forall i$

první slabý klasifikátor bude například $\text{sign}(ax+b)$

$a = 1$

$b = -4.5$

$\varepsilon_t = \sum_i D_t(i)[\![h_t(x_i) \neq y_i]\!]$

$\varepsilon_1 = \frac{2}{9}$ → dvě misklasifikace

$h_1 = \text{sign}(x - 4.5)$

$\alpha_1 = \frac{1}{2}\ln\left(\frac{1-\varepsilon_1}{\varepsilon_1}\right)$

$D_2(i) = D_1(i) \cdot e^{-\alpha_1 y_i h_1(x_i)}$

-1 dá ... správná klasifikace

$+1$ dá ... špatná klasifikace

$+1$

$y_i \; \text{sign}(x_i - 4.5)$

$D_2(i) = D_1(i) \cdot e^{-1/2 \ln(7/2)}$

D_2:

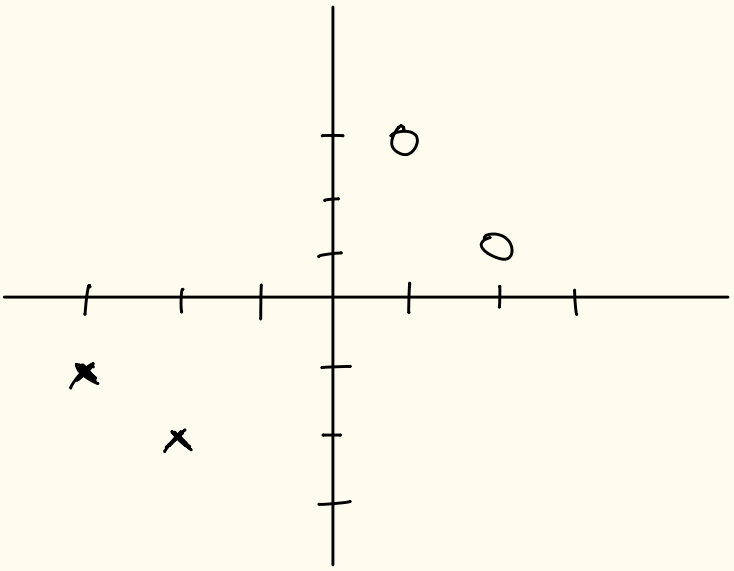
0: $1/9 \cdot 7/2^{-1/2}$
1: $1/9 \cdot 7/2^{-1/2}$
2: $1/9 \cdot 7/2^{1/2}$
3: $1/9 \cdot 7/2^{-1/2}$
4: $1/9 \cdot 7/2^{-1/2}$
5: $1/9 \cdot 7/2^{-1/2}$
6: $1/9 \cdot 7/2^{-1/2}$
7: $1/9 \cdot 7/2^{-1/2}$
8: $1/9 \cdot 7/2^{1/2}$

$D_2(i) /= \sum_i D_2(i)$

$\gamma =$
$\{([3,1],2), ([2,2],2), ([2,1],1), ([1,3],1)\}$

$\gamma' = \left\{\begin{pmatrix}-3\\-1\\-1\end{pmatrix}, \begin{pmatrix}-2\\-2\\-1\end{pmatrix}, \begin{pmatrix}2\\1\\1\end{pmatrix}, \begin{pmatrix}1\\3\\1\end{pmatrix}\right\}$

0. it.	$v_0 = (0,0,0)$	$(-3,-1,-1)$
1. it.	$v_1 = (-3,-1,-1)$	$(2,1,1)$
2. it.	$v_2 = (-1,0,0)$	$(1,3,1)$
3. it.	$v_3 = (0,3,1)$	$(-3,-1,-1)$
4. it	$v_4 = (-3,2,0)$	$(2,1,1)$
5. it	$v_5 = (-1,3,1)$	$(-3,-1,-1)$
6. it	$v_6 = (-4,2,0)$	$(2,1,1)$
7. it	$v_7 = (-2,3,1)$	$(-2,-2,-1)$
8. it	$v_8 = (-4,1,0)$	$(1,3,1)$
9. it	$v_9 = (-3,4,1)$	$(-2,-2,-1)$
10. it	$v_{10} = (-5,2,0)$	$(2,1,1)$
11. it	$v_{11} = (-3,3,1)$	$(-2,-2,-1)$
12. it	$v_{12} = (-5,1,0)$	$(1,3,1)$
13. it	$v_{13} = (-4,4,1)$	$(2,1,1)$
14. it	$v_{14} = (-2, \ldots$	



Cíchy to dokonvergujú pretože to je lin separabilní ...

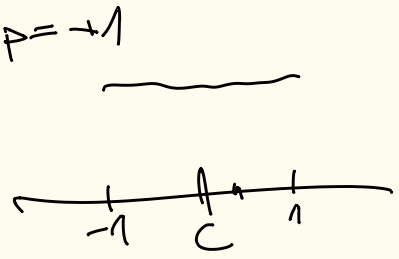
$$T = \{(1,+1), (3,+1), (0,+1), (5,+1), (3,+1), (6,-1), (-3,-1), (-4,-1), (8,-1), (-3,-1)\}$$

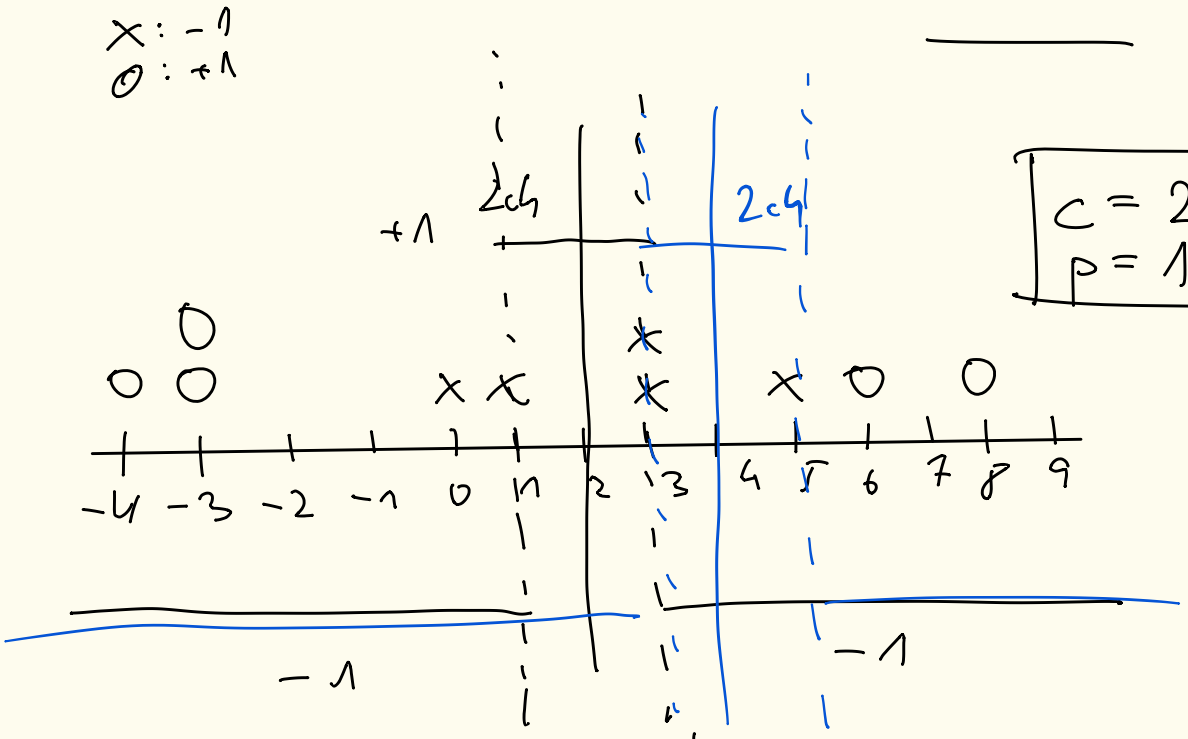
$$y = \{-1, +1\}$$

$\mathcal{H} = \{h(x; c, p)\}$ parametry označují jejich možné hodnoty

$c \in \mathbb{R}$ a pravidlo $p \in \{-1, +1\}$

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





$c = 2$
$p = 1$

$$\varepsilon_1 = \sum_i D_1(i) \cdot [\![y_i \neq h(x_i)]\!] = 0.2$$

$$\alpha_1 = 1/2 \ln\left(\frac{1-\varepsilon_1}{\varepsilon_1}\right) = 1/2 \ln 4$$

$$D_2(i) = D_1(i) \cdot e^{-1/2 \ln 4 \cdot y_i \cdot h(x_i, c, p)}$$ pro $c=2$, $p=1$

pro správně klasifik.

$D_2(i) = 1/10 \cdot 4^{-1/2} = 1/20$

$h(x_i, c, p) = y_i$

pro špatně klasifik.

$D_2(i) = 1/10 \cdot 4^{1/2} = 1/5$

$h(x_i, c, p) \neq y_i$

$\sum_i D_2(i) = 8 \cdot 1/20 + 2 \cdot 1/5 = 4/5$

$D_2(i) = 1/16$

$h(x_i, c, p) = y_i$

$D_2(i) = 1/4$

$h(x_i, c, p) \neq y_i$