



Strojové učenie a neurónové siete - **zadanie 2** - Stroj s podpornými vektormi

Zadanie 2

- Zadanie a bodovanie je k dispozícii na dokumentovom serveri v AIS
- Máte sa naučiť:
 1. Používať validačnú množinu
 2. Vedieť identifikovať podtrénovanie/ pretrénovanie
 3. Sledovať a regulovať proces trénovania siete
 4. Trénovať Stroj s podpornými vektormi
- Dataset:
 - dostupný v AIS
 - testovacia množina je oddelená, aby sa dali systémy navzájom porovnávať
 - popisy stĺpcov podľa Spotify API



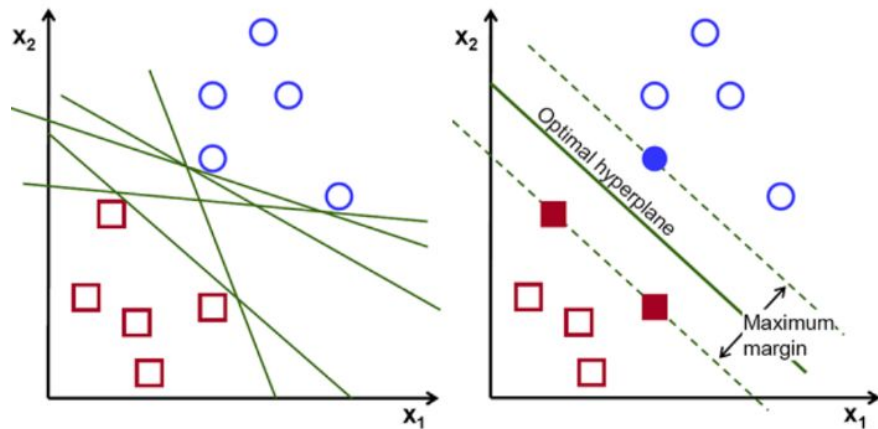
Obsah tejto prezentácie

- Stroj s podpornými vektormi:
 - Trénovanie
 - Kernely
 - Parametre
- Vzorová neurónka (opäť), dokumentácie
- O týždeň - Zadanie 3



Metóda podporných vektorov - intuícia

- Učenie s učiteľom
- Hľadá hyperroviny oddeľujúce triedy (klasifikácia) / približnú funkciu (regresia) - podobné neurónkam
- Hľadá váhy w pre rovnice, ktorých kombinácia so vstupmi predikuje výstup - podobné neurónkam
- Predstavuje koncept podporných vektorov, ktoré pomáhajú hľadať dobré hyperroviny - optimalizujem pre "ťažké" vstupy, nie pre všetky
- Vyberáme hyperparametre:
 - Druh kernelu
 - C
 - Gamma
 - Stupeň (pri polynomiálnom kerneli)



Kritériálna funkcia/trénovanie

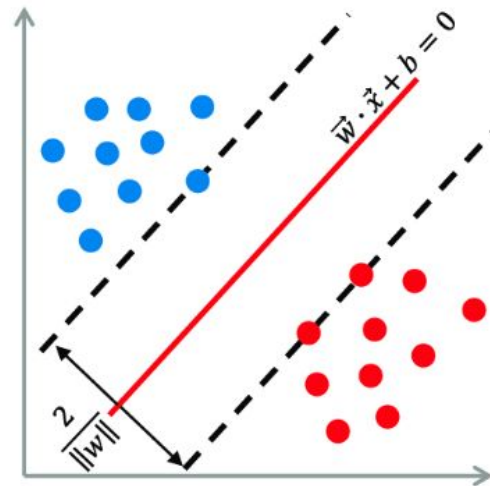
- Hinge loss:

$$L(\mathbf{w}, b) = \sum_i^n \max(0, 1 - y_i(\mathbf{w}^T \mathbf{x}_i + b)) + \lambda \|\mathbf{w}\|^2;$$

$$y = \pm 1$$

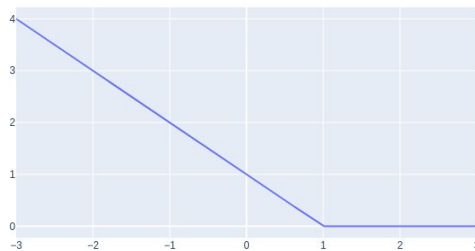
- Resp.: $\min_{w,b} \|\mathbf{w}\|^2$

$$\text{ak platí } y_i(\mathbf{w}^T \mathbf{x}_i + b) \geq 1$$

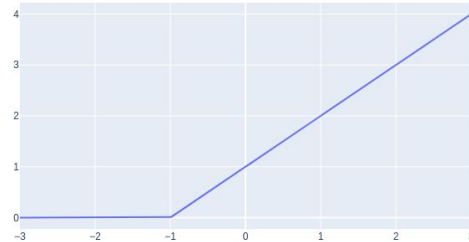


- Parabola má jedno minimum :)

y=1



y=-1



Hard a soft margin

Zmeníme:

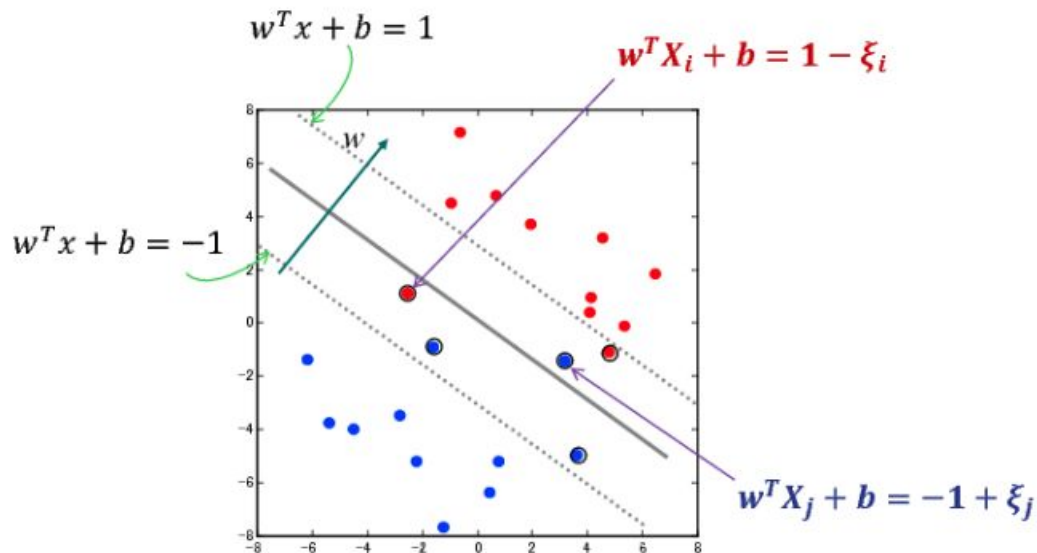
$$\min_{w,b} \|w\|^2$$

$$y_i(w^T x_i + b) \geq 1$$

na:

$$\min_{w,b} \|w\|^2 + C \sum_i^n \xi_i$$

$$y_i(w^T x_i + b) \geq 1 - \xi_i; \xi_i \geq 0$$

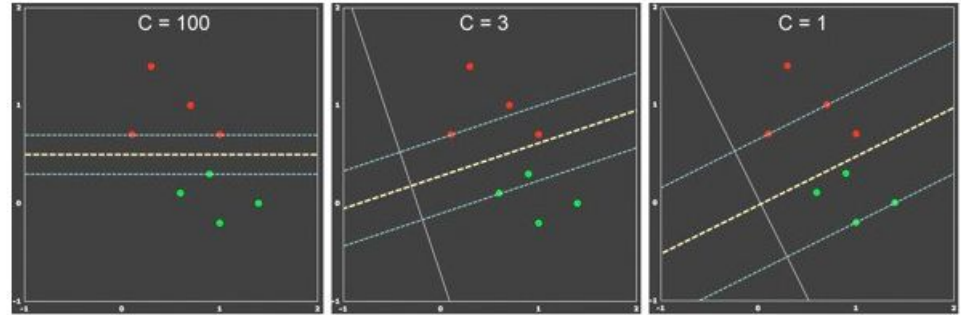
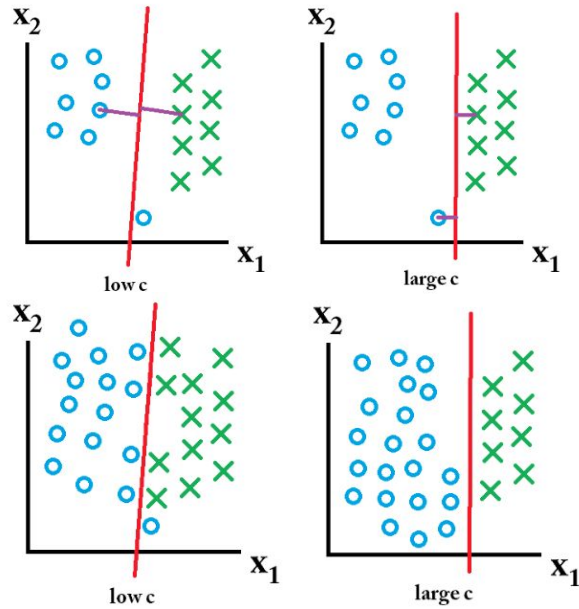


a do hyperparametrov pribudne *cost* C .

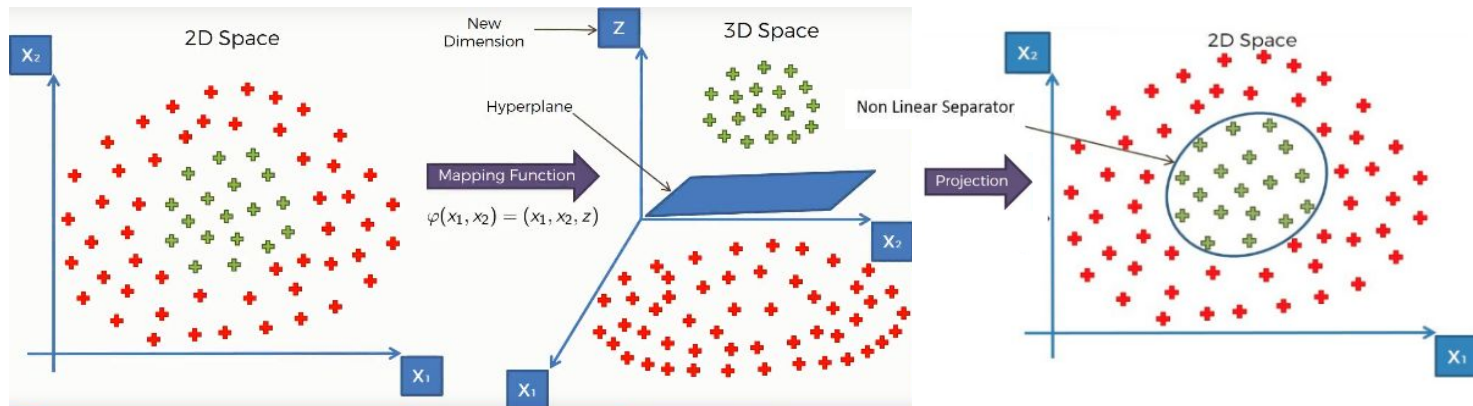


Hyperparametre - C

- Určuje počet správne klasifikovaných vzoriek vs maximalizovanie rozdeľovacej hranice
- Vysoká hodnota C = väčšia hranica
- Nízka hodnota C = nízka hranica



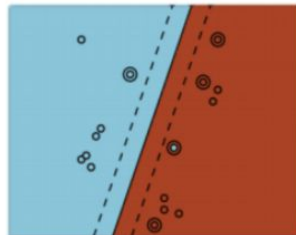
Kernelový trik



Hyperparametre - druh kernelu

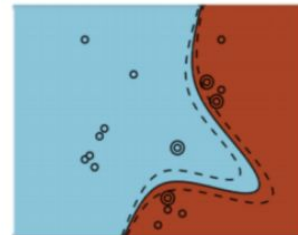
- Lineárny kernel
- Polynomiálny kernel
- RBF kernel

Linear Kernel



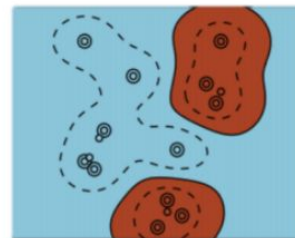
C hyperparameter

Polynomial Kernel



C plus gamma, degree and coefficient hyperparameters

RBF Kernel

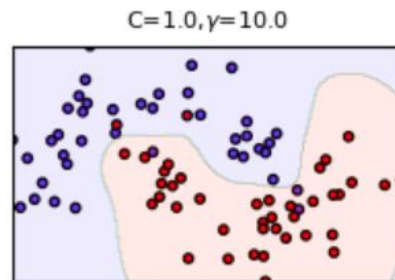
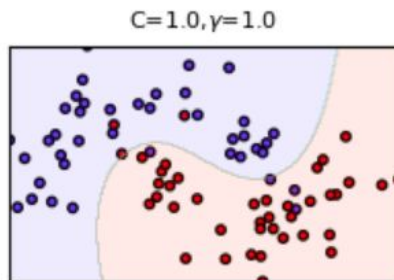
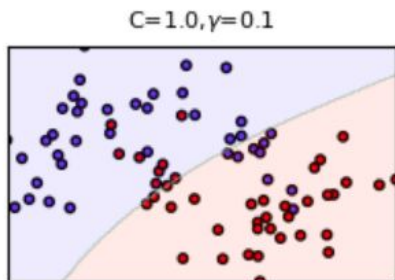
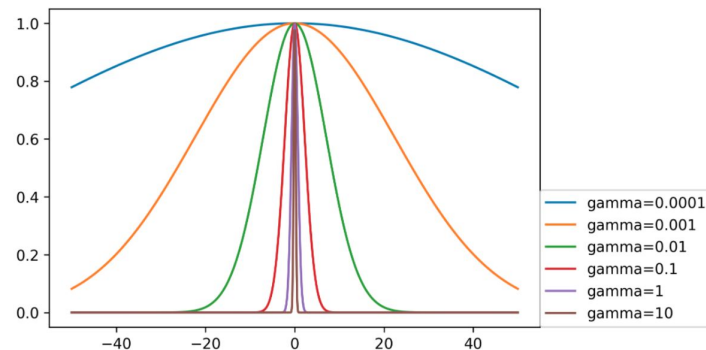


C plus gamma hyperparameter

Hyperparametre - γ

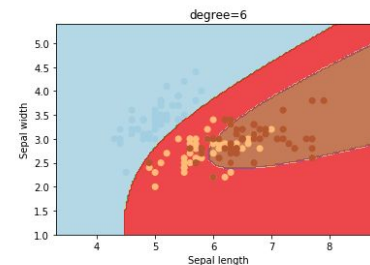
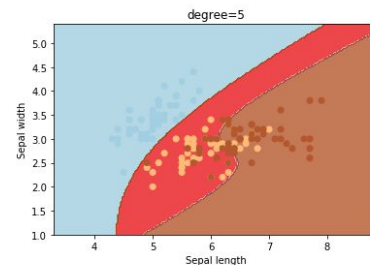
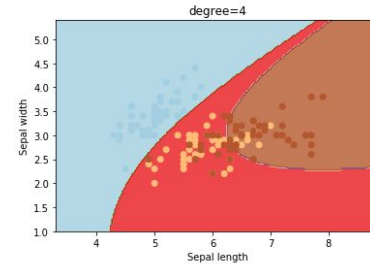
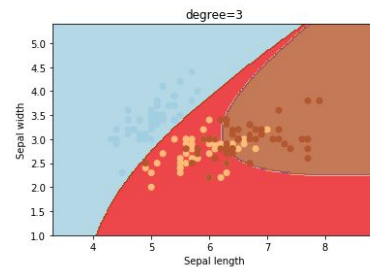
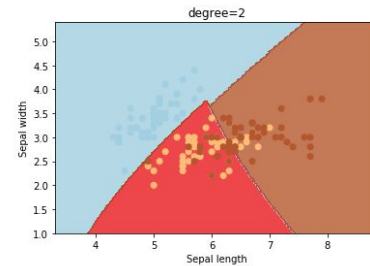
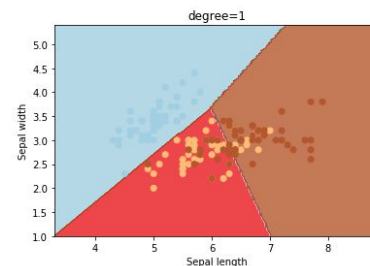
- Súvisí so šírkou RBF funkcií
- Určuje, aký “ďaleký dosah” má jedna vzorka
- Čím nižšia hodnota γ , tým väčší dosah majú vzorky

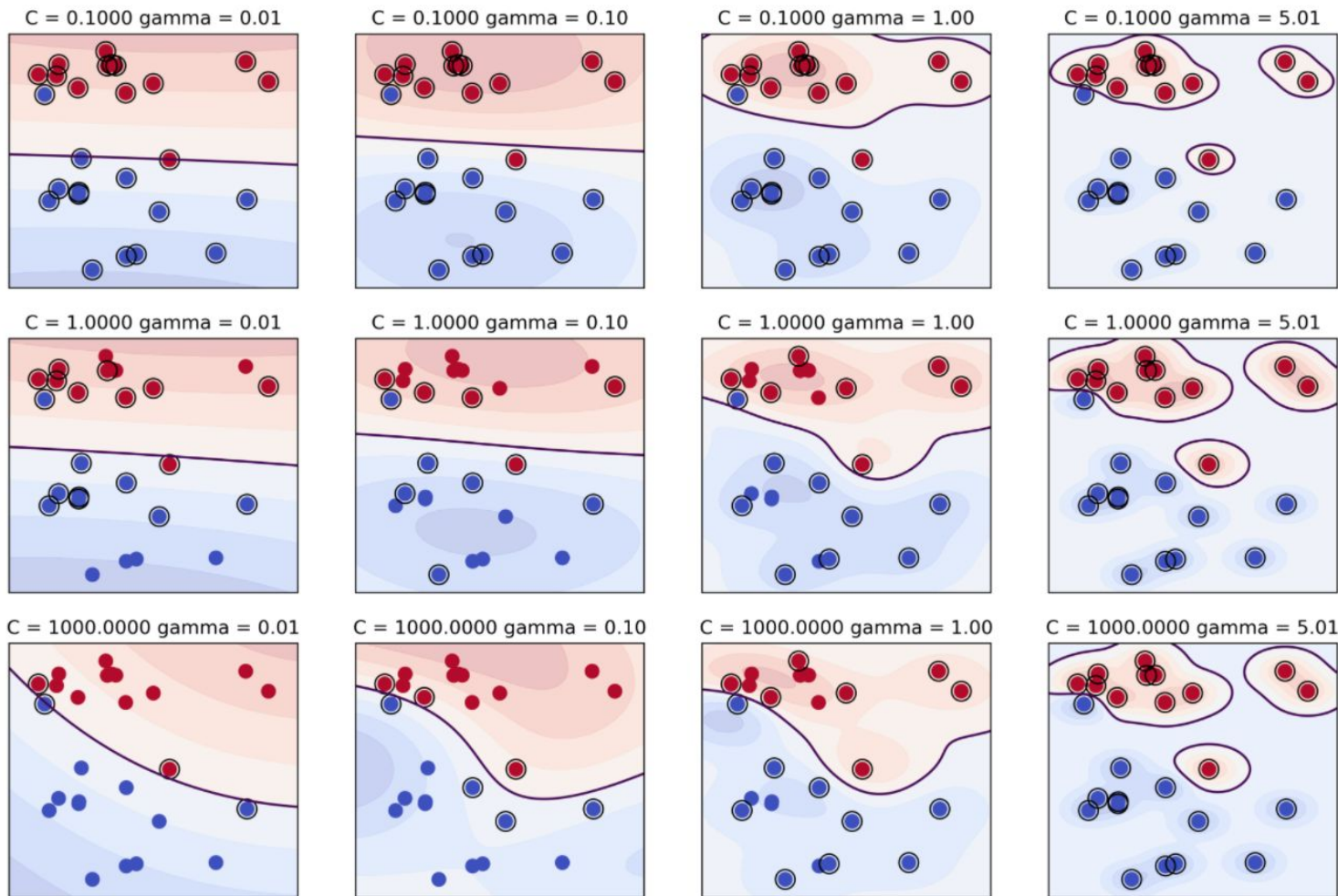
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$



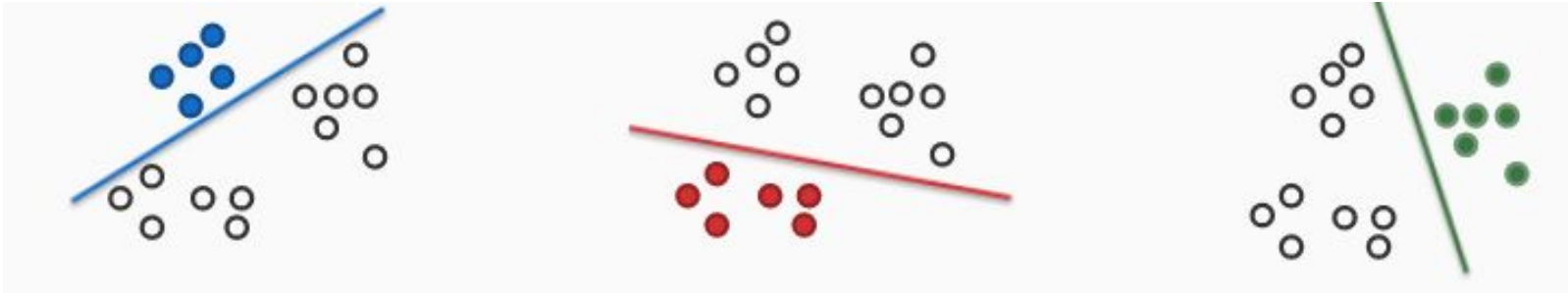
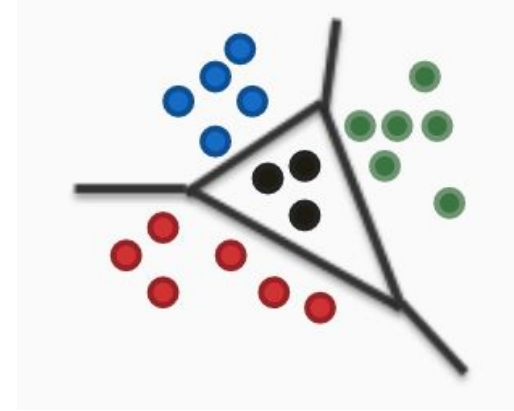
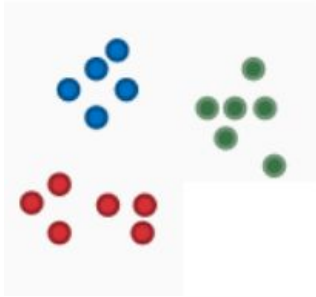
Hyperparameter - stupeň

- Pri polynomiálnom kerneli
- Vyšší stupeň - dlhší čas tréovania, väčšia komplexnosť

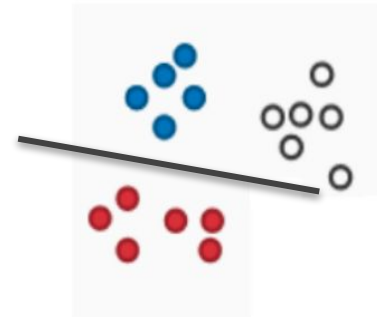
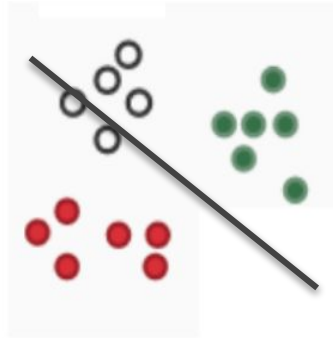
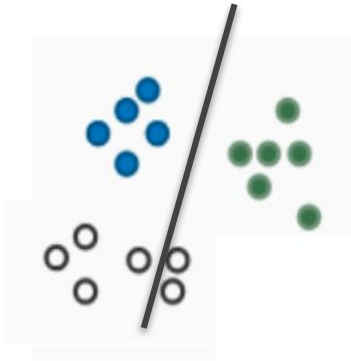
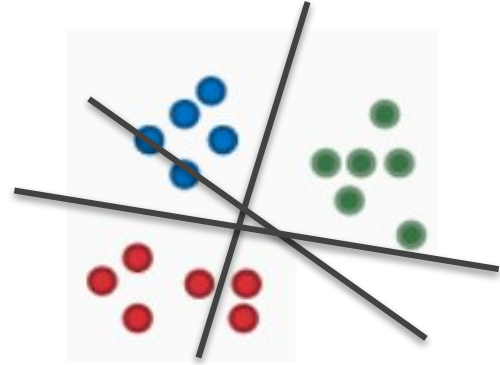
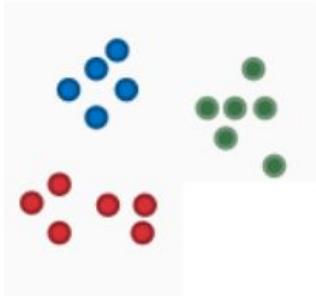




Viac-triedne rozpoznávanie - one vs all



Viac-triedne rozpoznávanie - one vs one



Poznámky

- Čím komplexnejší model, tým dlhší čas tréovania a tým väčšia možnosť pretrénovania
- SVM sa dá lepšie trénovať na malej množine ako ANN (pretrénovanie) - ale pozor na outliers a nereprezentatívnu množinu
- SVM obvykle rýchlejšie ako ANN
- Takisto vie riešiť aj regresné problémy

