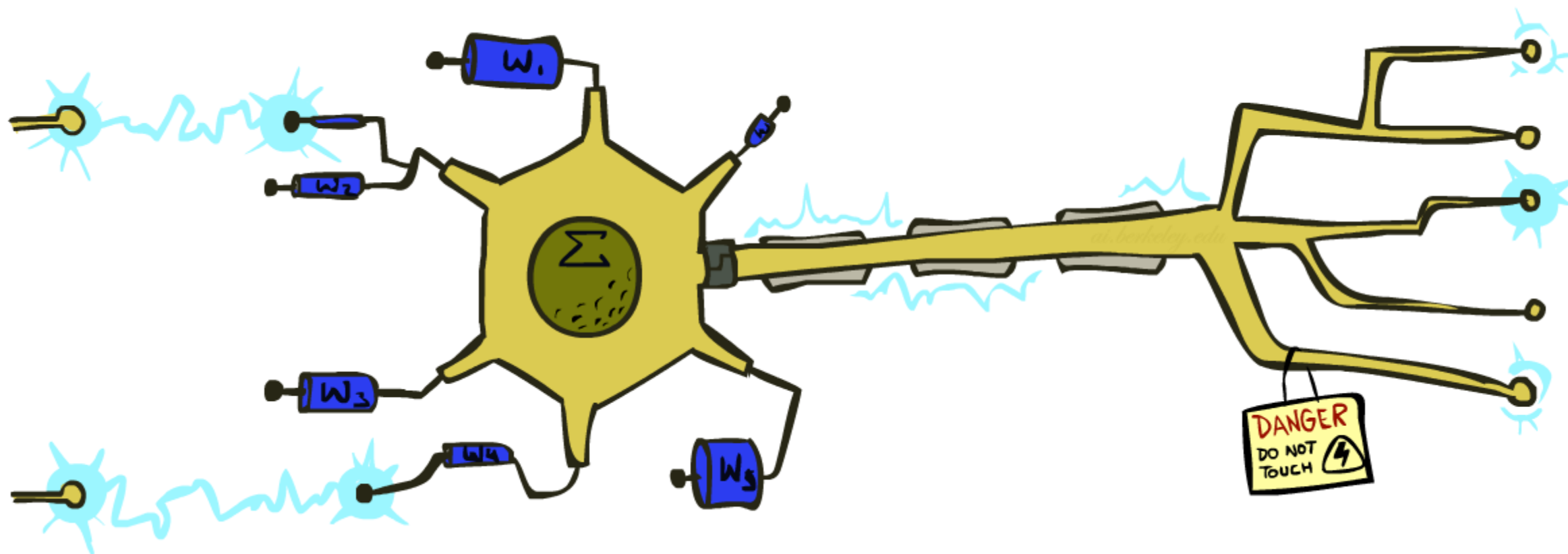


Osnovi Računarske Inteligencije

Perceptron

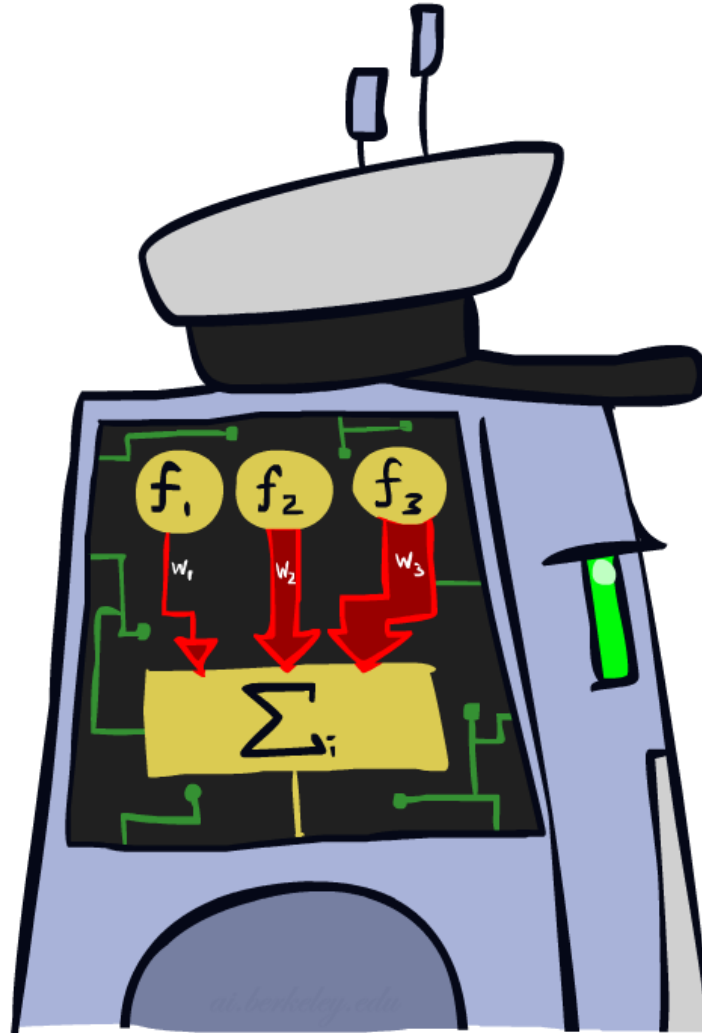


Predavač: Aleksandar Kovačević

Slajdovi preuzeti sa kursa CS188, University of California, Berkeley

<http://ai.berkeley.edu/>

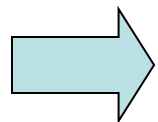
Linearni Klasifikatori



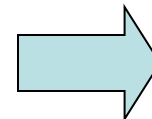
Vektori Osobina (*Feature Vectors*)

 x $f(x)$ y

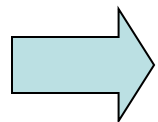
```
Hello,  
  
Do you want free printr  
cartridges? Why pay more  
when you can get them  
ABSOLUTELY FREE! Just
```



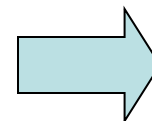
```
# free      : 2  
YOUR_NAME   : 0  
MISPELLED   : 2  
FROM_FRIEND : 0  
...
```



SPAM
or
HAM



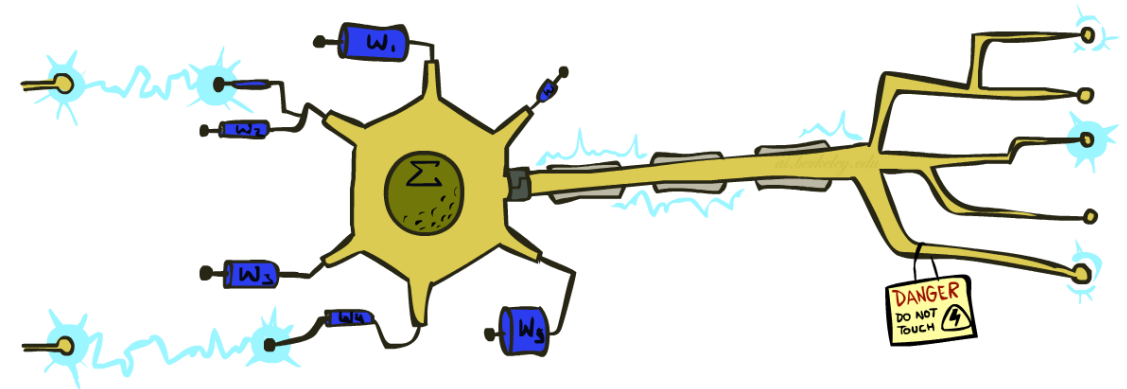
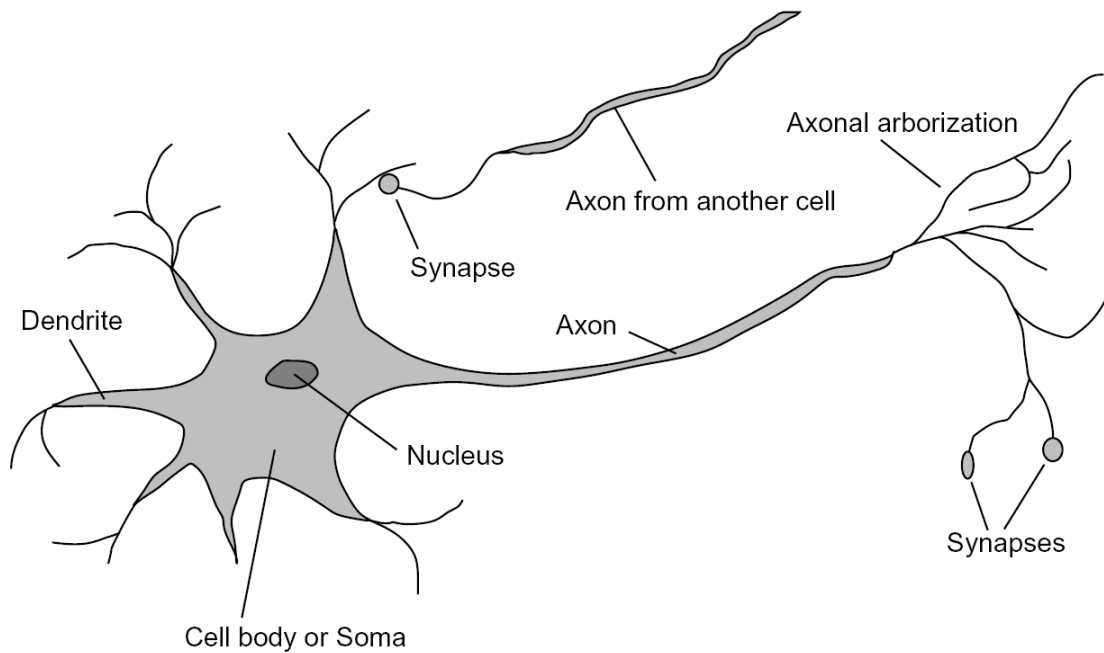
```
PIXEL-7,12 : 1  
PIXEL-7,13 : 0  
...  
NUM_LOOPS  : 1  
...
```



"2"

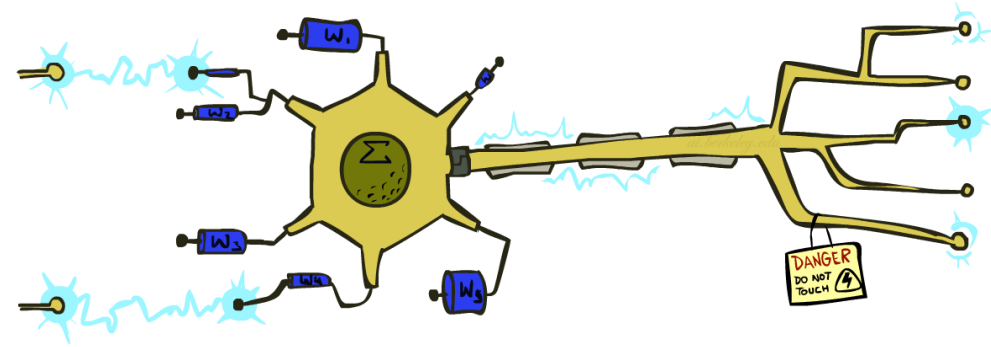
Inspiracija iz Biologije

- Perceptron se može posmatrati kao vrlo uprošćen ljudski neuron



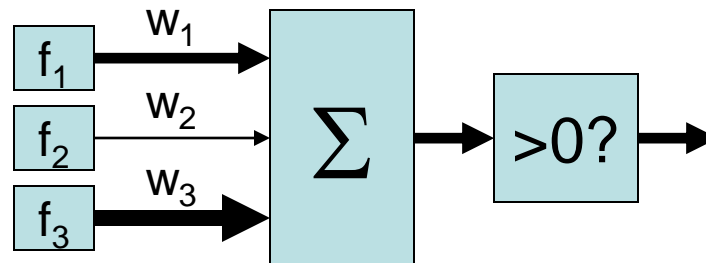
Linearni Klasifikatori

- Ulazi su **vrednosti osobina**
- Svaka osobini dodljena je **težina**
- Suma težina i vrednosti je **aktivacija**



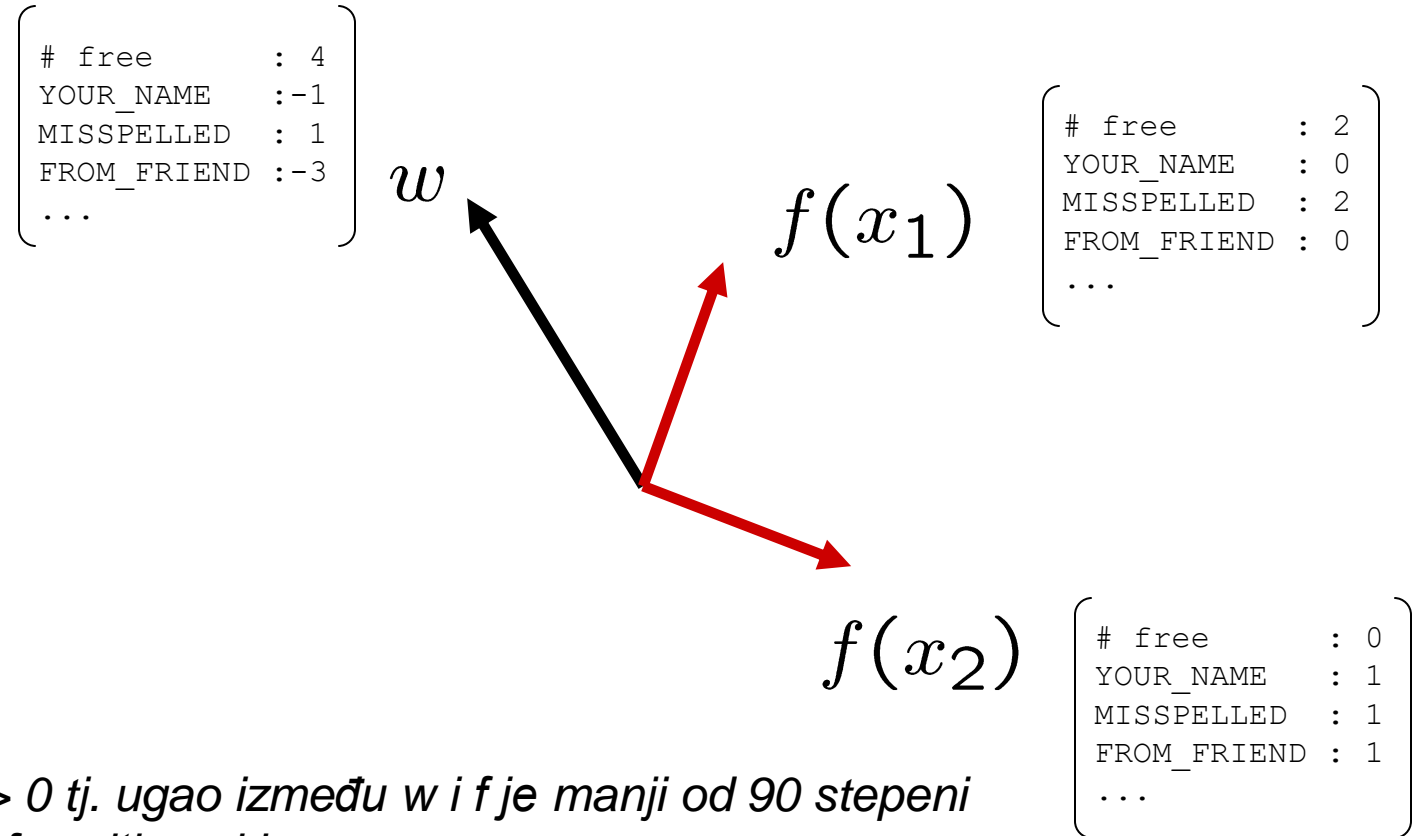
$$\text{activation}_w(x) = \sum_i w_i \cdot f_i(x) = w \cdot f(x)$$

- Kod preceptrona:
- Ako je aktivacija:
 - Pozitivna, izlaz je +1
 - Negativna, izlaz je -1



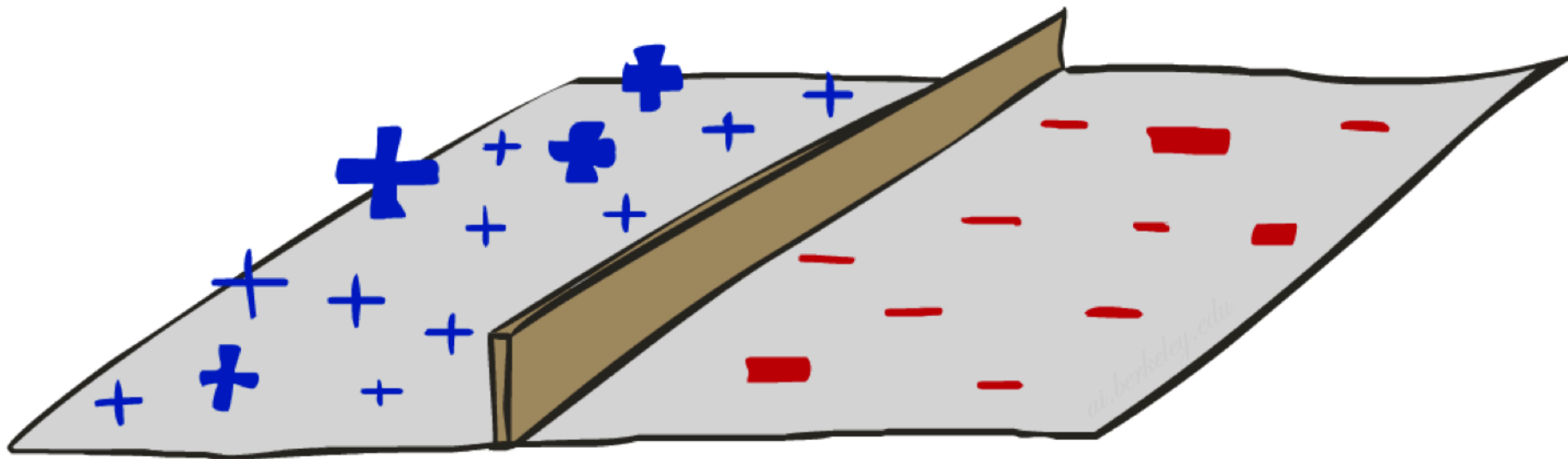
Težine

- Ako imamo dve klase: poredimo vektor osobina sa vektorom težina (gledamo ugao između njih)
- Učenje: Određivanje vektora težina na osnovu obučavajućeg skupa



Skalarni proizvod $w \cdot f > 0$ tj. ugao između w i f je manji od 90 stepeni znači da je predikcija za f pozitivna klasa

Granica Odluke (*Decision Boundary*)



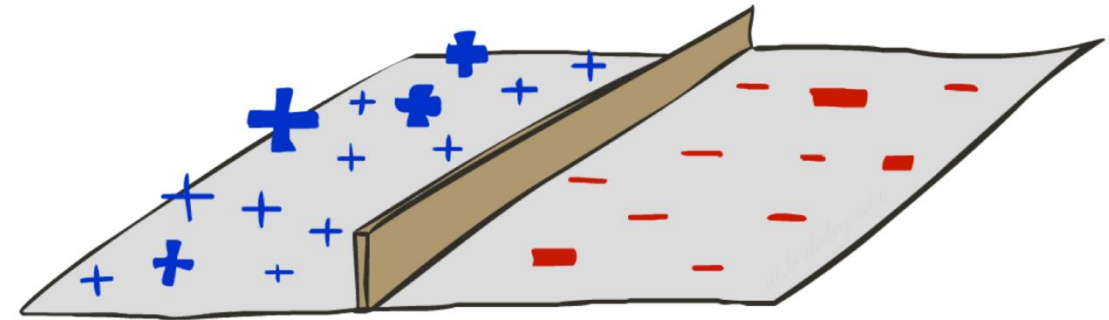
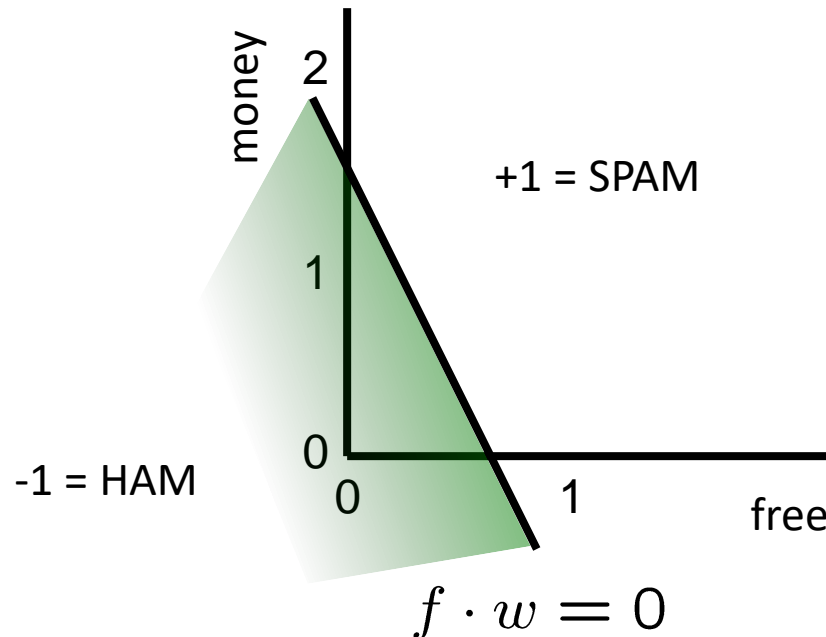
Binarna Granica Odluke

- U prostoru vektora osobina

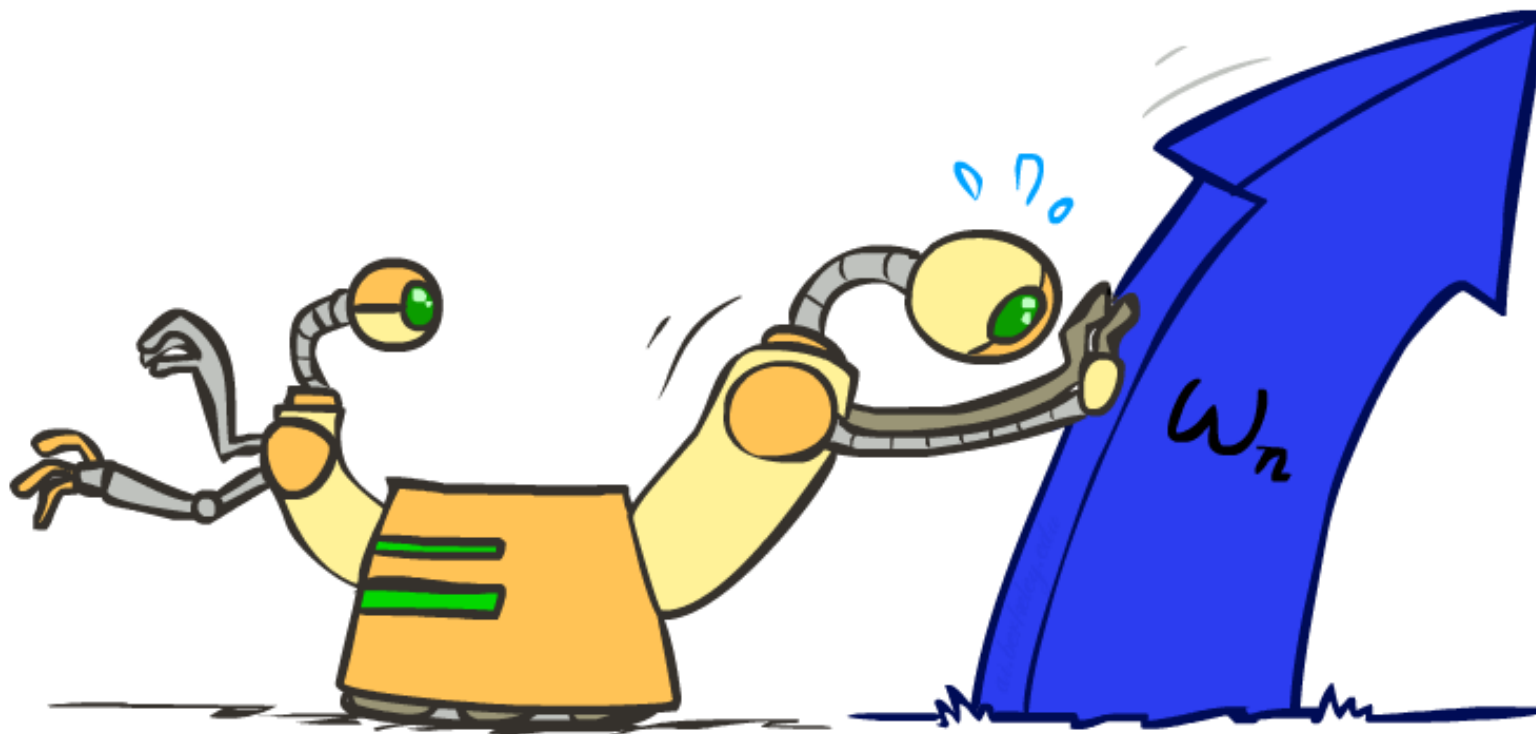
- Primeri iz ob. skupa su tačke
- Svaki vektor težina je hiper-ravan
- Jednoj strani odgovara pozitivna klasa $Y=+1$
- Drugoj negativna $Y=-1$

w

| | | |
|-------|---|----|
| BIAS | : | -3 |
| free | : | 4 |
| money | : | 2 |
| ... | | |

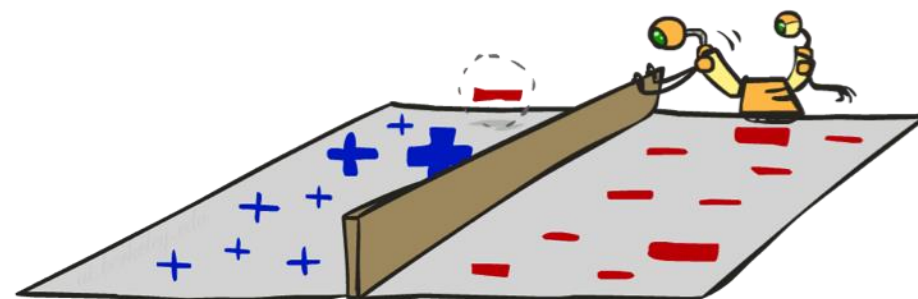
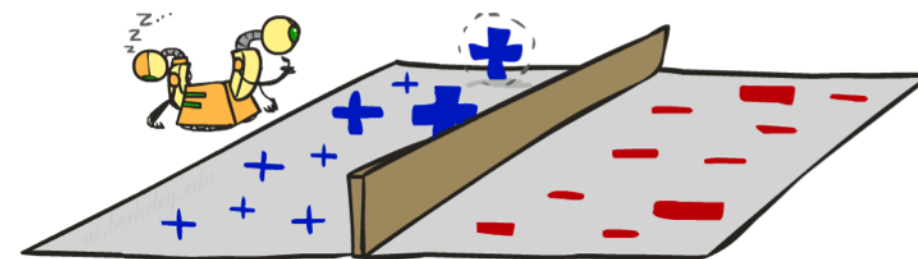
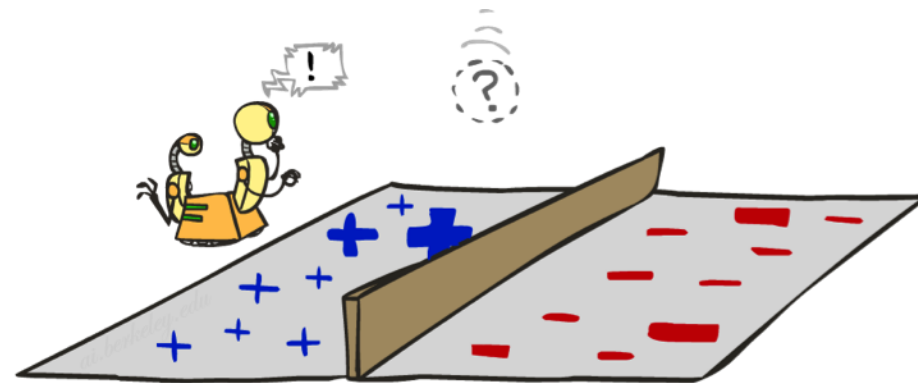


Promene (Korekcije) Težina



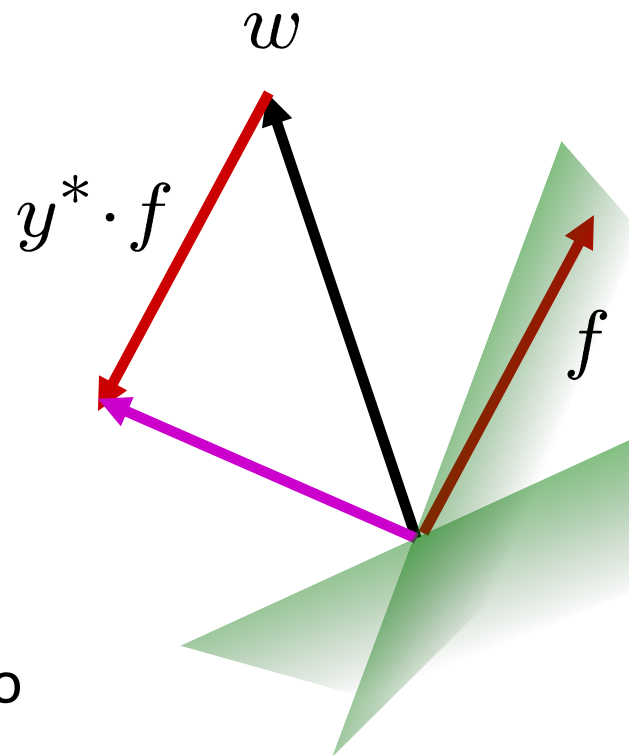
Obučavanje Binarnog Perceptrona

- Na početku sve težine su = 0 (ili slučajno odabrane vrednosti)
- Za svaki primer iz ob skupa radimo:
 - Klasifikujemo ga pomoću trenutnih težina
- Ako je klasifikacija tačna (tj. $y=y^*$) onda nema promene težina!
- y^* je prava klasa primera
- Ako je pogrešna: menjamo vektor težina



Obučavanje Binarnog Perceptrona

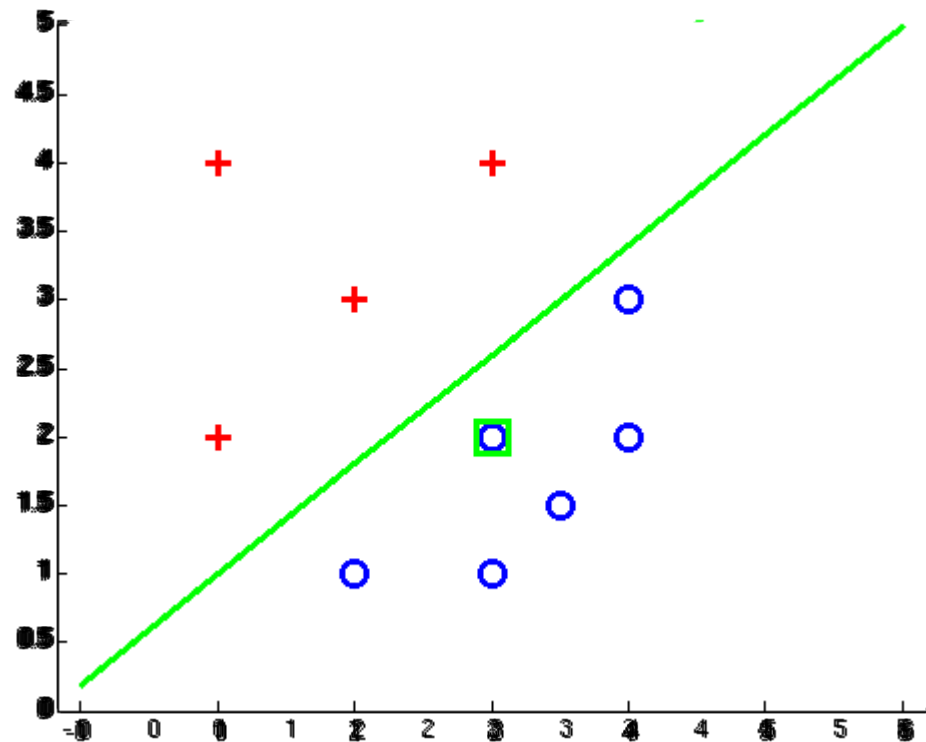
- Na početku sve težine su $= 0$
- Za svaki primer iz ob skupa radimo:
 - Klasifikujemo ga pomoću trenutnih težina
$$y = \begin{cases} +1 & \text{if } w \cdot f(x) \geq 0 \\ -1 & \text{if } w \cdot f(x) < 0 \end{cases}$$
 - Ako je klasifikacija tačna (tj. $y=y^*$) onda nema promene težina!
 - Ako je pogrešna: menjamo vektor težina tako što na njega dodajemo ili od njega oduzimamo vektor osobina tog primera. To je realizovano kroz formulu dole. Npr. ako je $y^* = -1$ biće oduzimanje f .



$$w = w + y^* \cdot f$$

Primer

- Pimeri su linearno razdvojivi po klasi



Više-klasna Granica Odluke

- Ako imamo više klasa:

- Imamo vektor težina za svaku klasu:

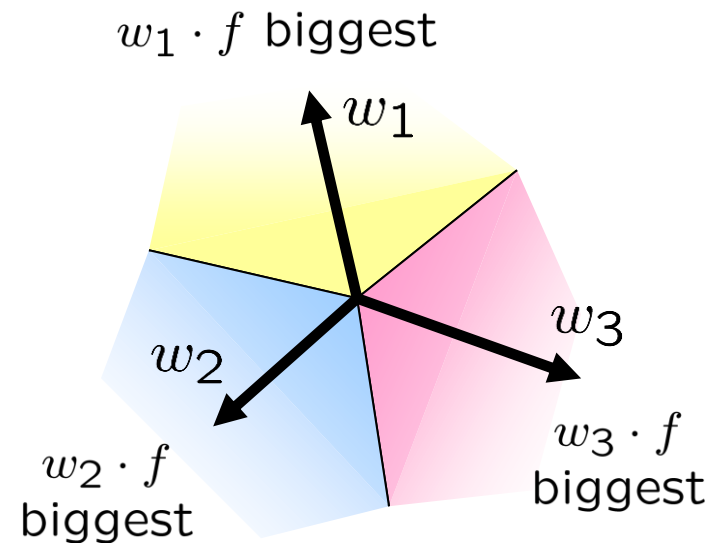
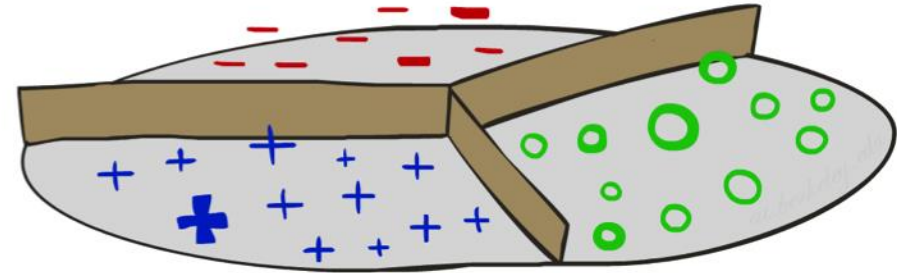
$$w_y$$

- Za svaku klasu računamo aktivaciju:

$$w_y \cdot f(x)$$

- Predikcija sa najvećom aktivacijom je klasa koju vraćamo

$$y = \arg \max_y w_y \cdot f(x)$$



Obučavanje Više-klasnog Perceptrona

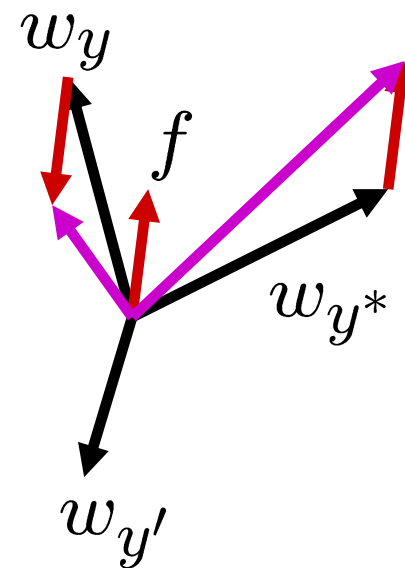
- Na početku sve težine su = 0
- Idemo redom po primerima iz ob. skupa
- Određujemo klasu pomoću formule:

$$y = \arg \max_y w_y \cdot f(x)$$

- Ako je tačna, ništa ne menjamo!
- Ako je pogrešna: smanjujemo aktivaciju pogrešnog odgovora, povećavamo aktivaciju tačnog odgovora

$$w_y = w_y - f(x)$$

$$w_{y^*} = w_{y^*} + f(x)$$

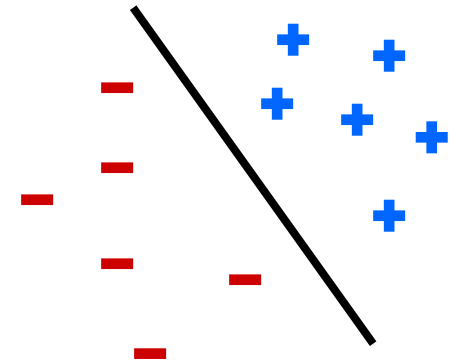


- Pogrešan odgovor je vektor težina zbog koga je primer dobio pogrešnu klasu (u argmax).
- Tačan odgovor je vektor težina koji odgovara pravoj klasi primera.

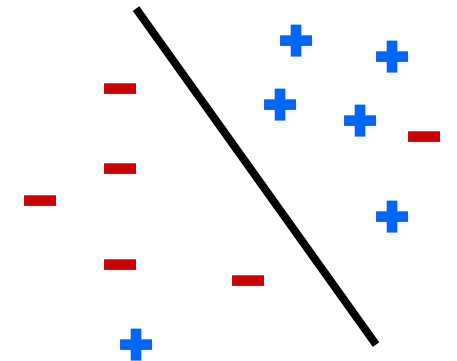
Karakteristike Perceptrona

- Može linearno da razdvoji dve ili više klasa
- Postoji teorijski dokaz da ako su podaci linearno separabili perceptron algoritam će sigurno da pronađe rešenje tj. da konvergira

Separable

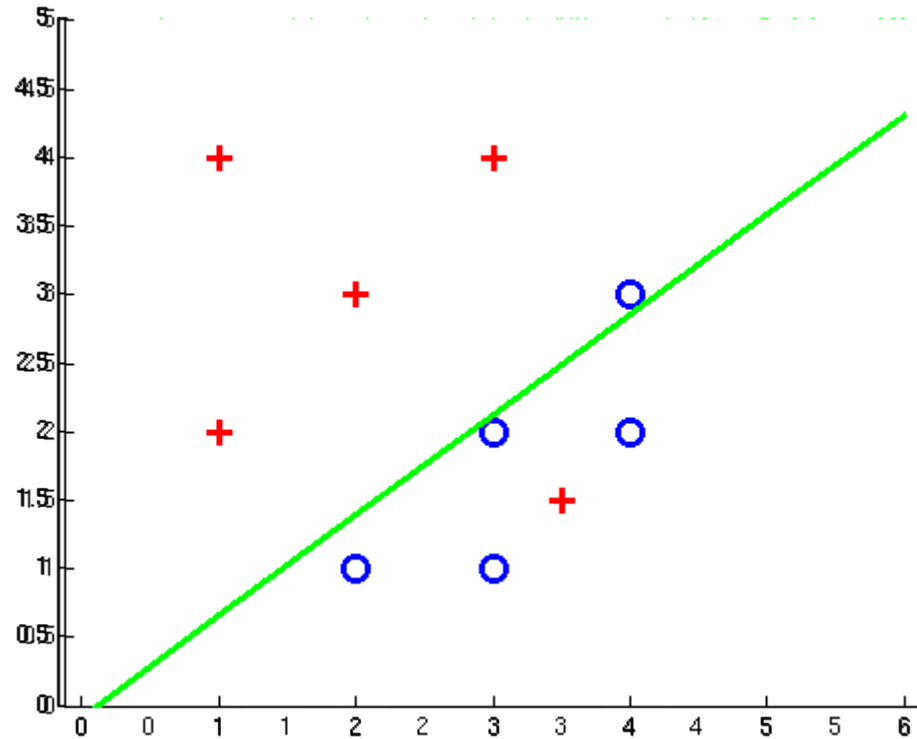


Non-Separable



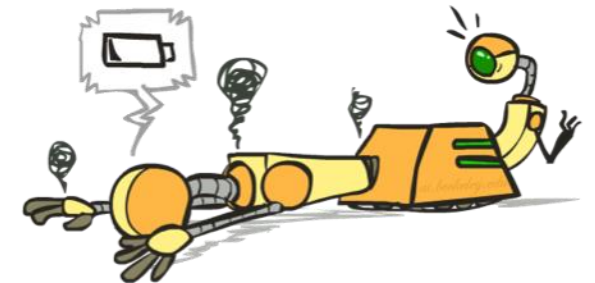
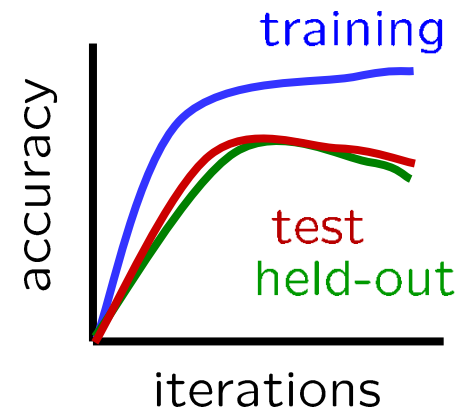
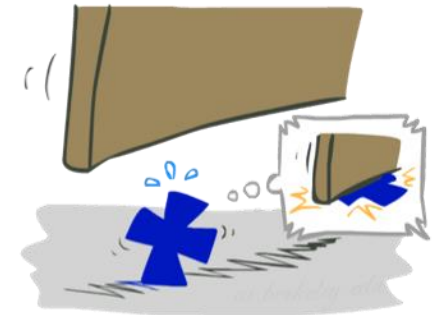
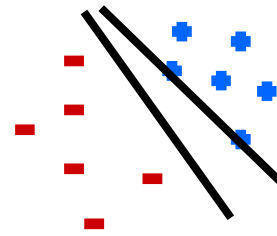
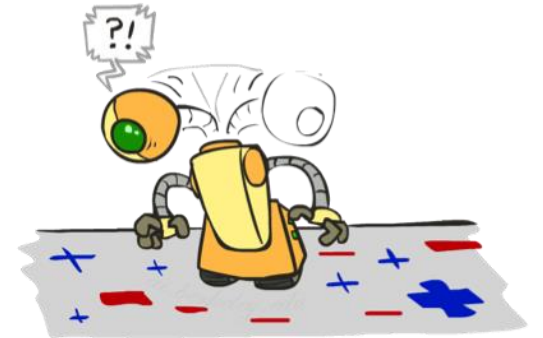
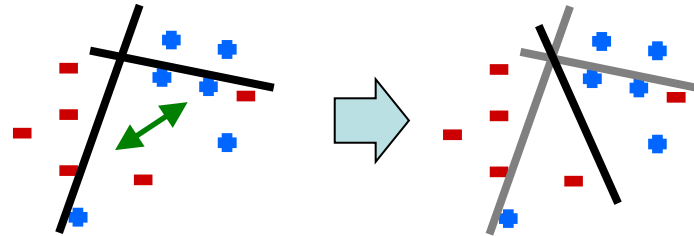
Primer

- Linearno ne-separabilni podaci



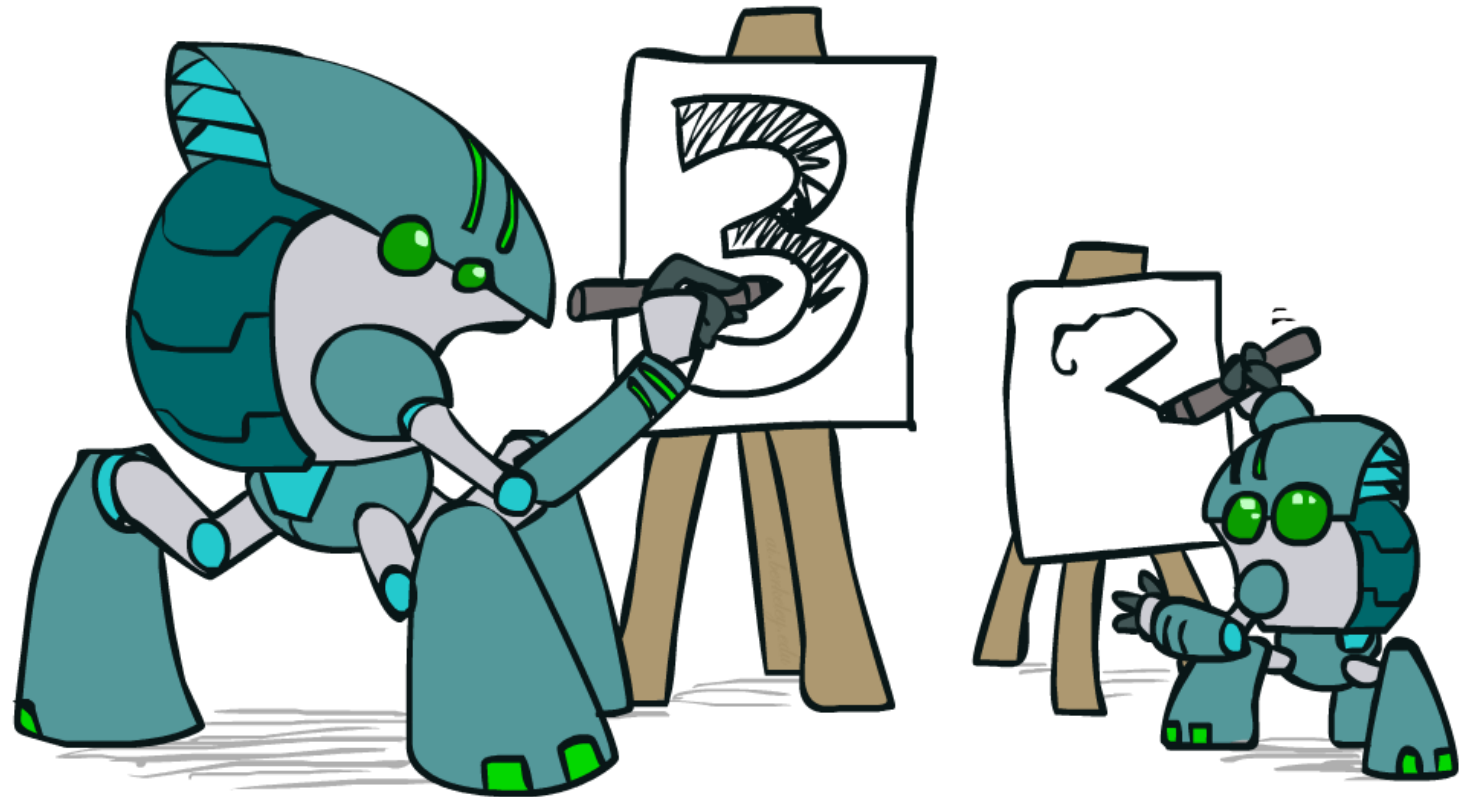
Problemi sa Perceptronom

- Šum: ako su podaci skoro linearno razdvojeni možemo dobiti većito „šetanje“ vektora težina tj. divergenciju
- Generalizacija nije baš dobra: nalazi rešenje, ali ne ostavlja mnogo „lufta“ između klasa (termin iz literature za „luft“ je margina separacije)
- Često se dešava da tokom iteracija tačnost modela raste do nekog momenta, pa onda krene da opada
 - Dobro je pamtiti iteraciju u kojoj je model imao najveću tačnost

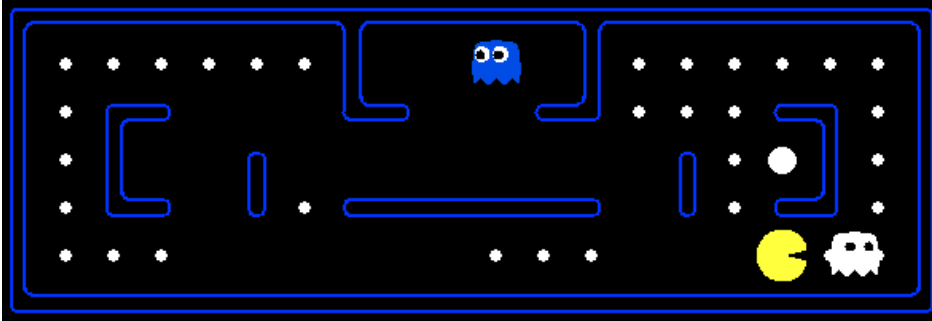


„Šegrtovanje“ - *Apprenticeship*

- Ideja kod ovog načina učenja je da učimo od eksperta tj. da mu budemo šegrt
- Ekspertovo znanje može da se reprezentuje kao klasa,
- dok se stanje sveta tj. opis problema može reprezentovati kao vektor osobina
- Klasa kaže šta bi ekspert uradio u situaciji koja je reprezentovana vektorom osobina

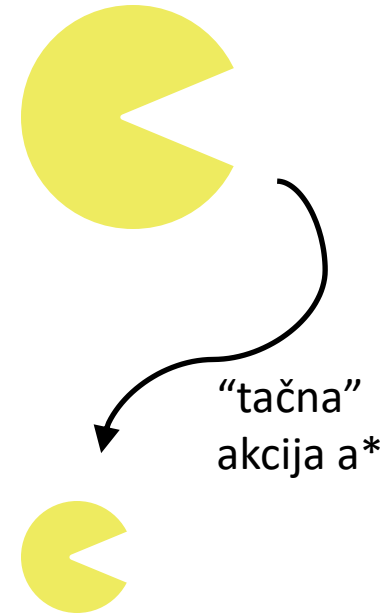


Primer - Pacman!



- Kandidati za klasifikator su parovi stanje-akcija (s,a)
- Ekspert svakom paru stanje-akcija (s,a) dodeljuje +1 ili -1, u zavisnosti da li je to akcija koju bi on uradio u stanju s
- Ceo vektor osobina je reprezentacija i stanja i akcije (npr. udaljenost duhova, broj pojedenih kuglica itd.)
- Aktivacija za svaki par (s,a) data je sa:

$$w \cdot f(s, a)$$



$$\forall a \neq a^*, \\ w \cdot f(a^*) > w \cdot f(a)$$

Demo

