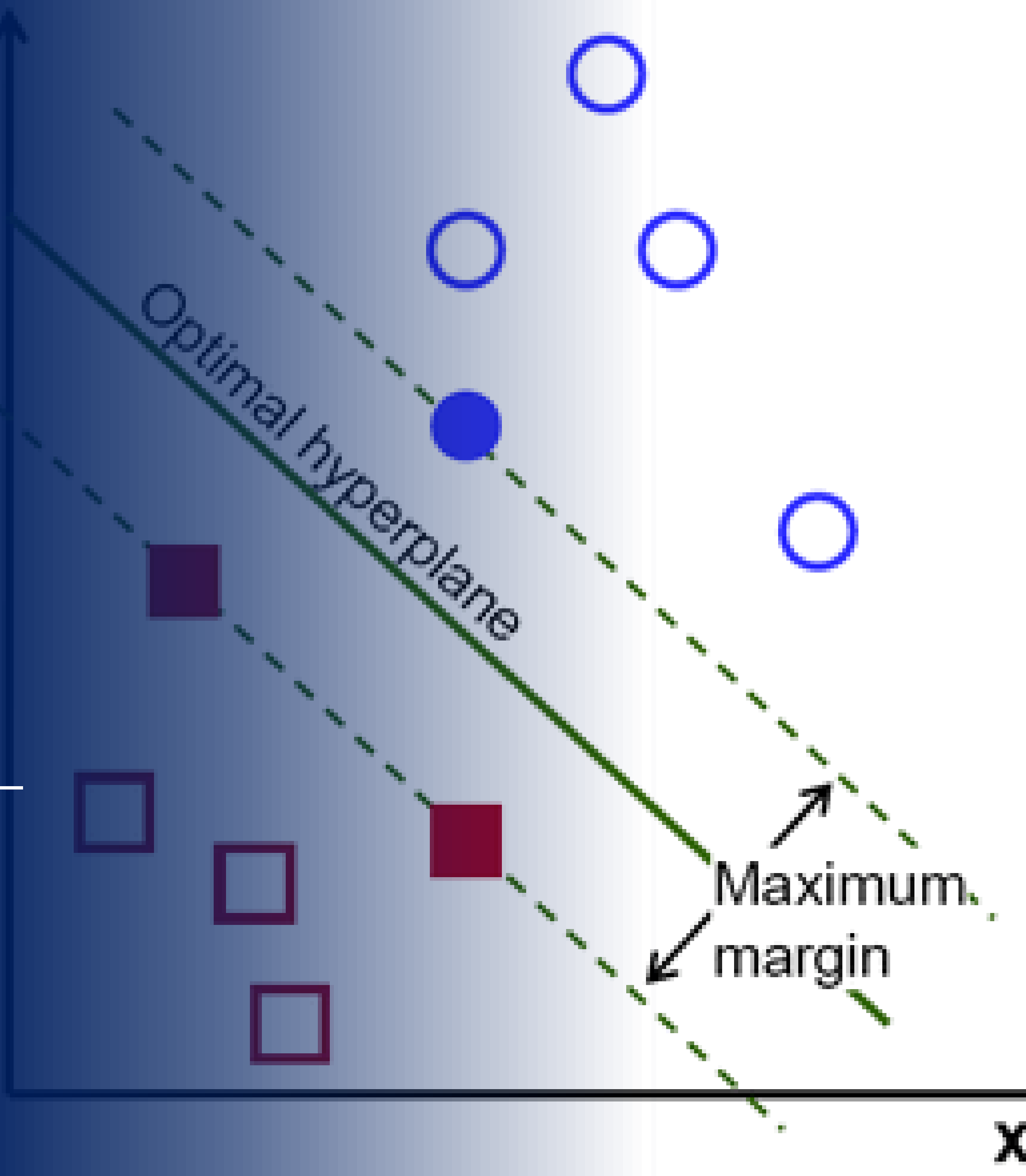
Metoda potpornih vektora

Support Vector Machines (SVM)



Metoda potpornih vektora

- Direktan pristup klasifikacionom problemu (kako to bi uradio inženjer – bez uvođenja probabilističkog modela)

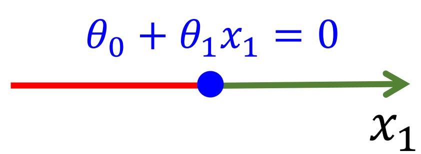
- Razmatraćemo problem binarne klasifikacije
  - Imamo dve klase – „pozitivnu“ i „negativnu“
  - Svejedno je kako ćemo obeležiti klase (bitno nam je da ih razlikujemo)
  - Mi ćemo odabrati da „pozitivnu“ klasu obeležimo sa $y = +1$, a „negativnu“ klasu obeležimo sa $y = -1$

- Granica odluke (koja razdvaja primere „pozitivne“ klase od primera „negativne“ klase) je hiperravan

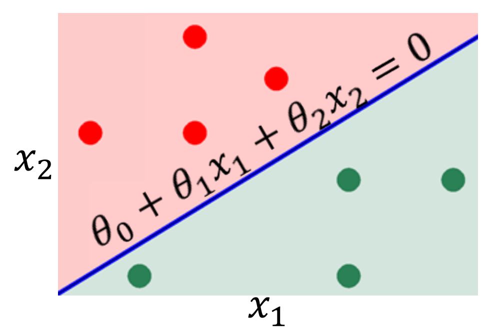
Šta je hiperravan?

- Koncept u geometriji kojim se generalizuje koncept
- Deli D-dimenzionalni prostor na dva dela
- Hiperravan je podprostor dimenzije $D - 1$
- Može se opisati linearnom jednačinom oblika:

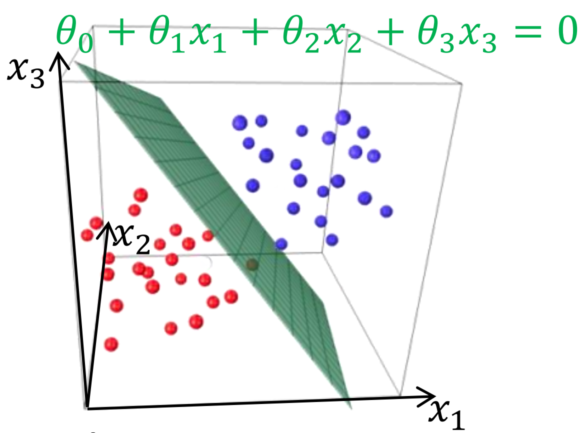
$$f(x) = \theta^T x = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_D x_D = 0$$



1-dimenzioni prostor:
- Tačka
- Deli pravu na 2 dela



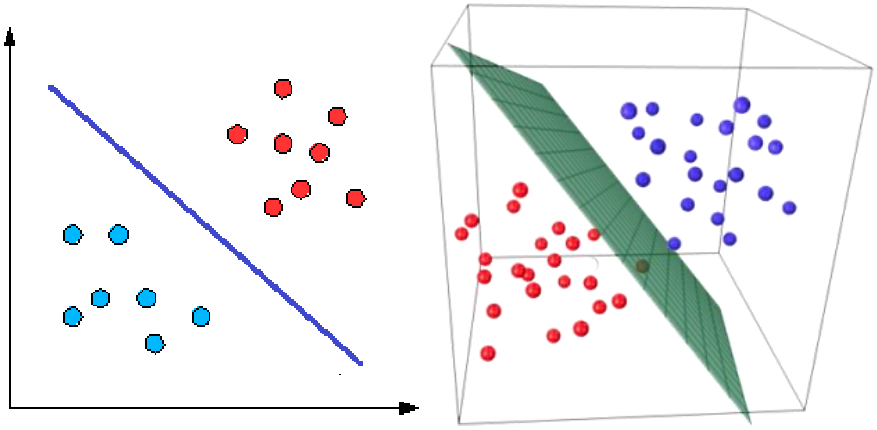
2-dimenzioni prostor:
- Prava linija
- Deli ravan na dve poluravni



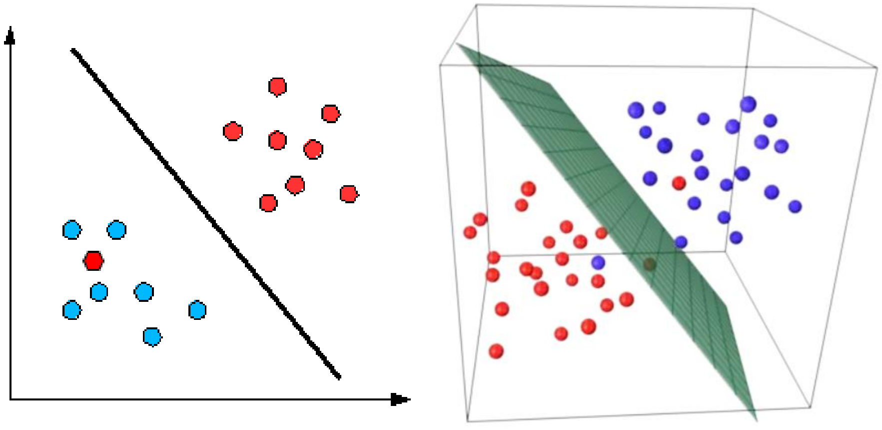
3-dimenzoni prostor:
- Ravan
- Deli prostor u dva poluprostora

Linearna separabilnost

- Razmatraćemo slučaj linearno separabilnih podataka
- (Klase možemo savršeno razdvojiti pomoću hiperravni)

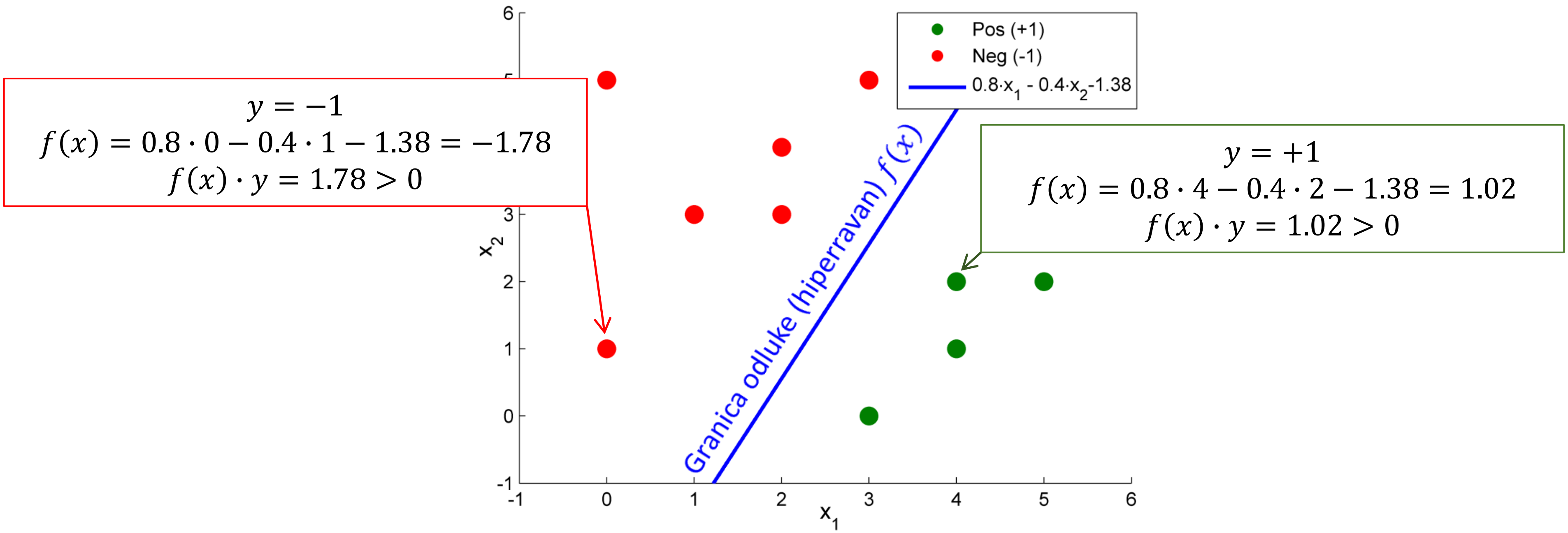


Primeri linearno separabilnih podataka



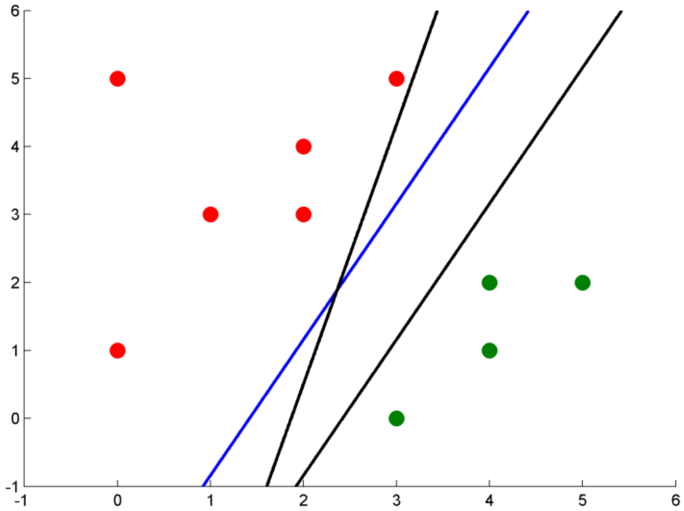
Primeri podataka koji nisu linearno separabilni

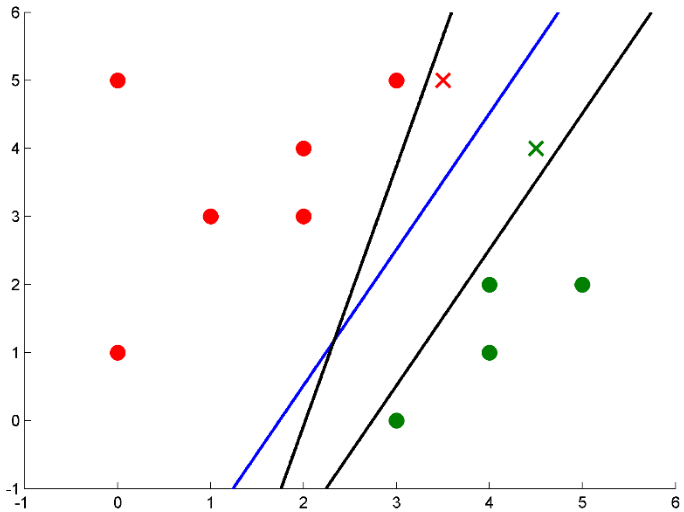
Linearna separabilnost – formalnije



- Neka je data hiperravan opisana jednačinom $f(x) = \theta^T x = 0$
  - Za zelene tačke (koje smo obeležili sa $y^{(i)} = +1$) važi $f\left(x^{(i)}\right) > 0$
  - Za crvene tačke (koje smo obeležili sa $y^{(i)} = -1$) važi $f\left(x^{(i)}\right) < 0$
- Ako postoji takva hiperravan da za **sve** primere važi $y^{(i)} \cdot f\left(x^{(i)}\right) > 0$, podaci su linearno separabilni

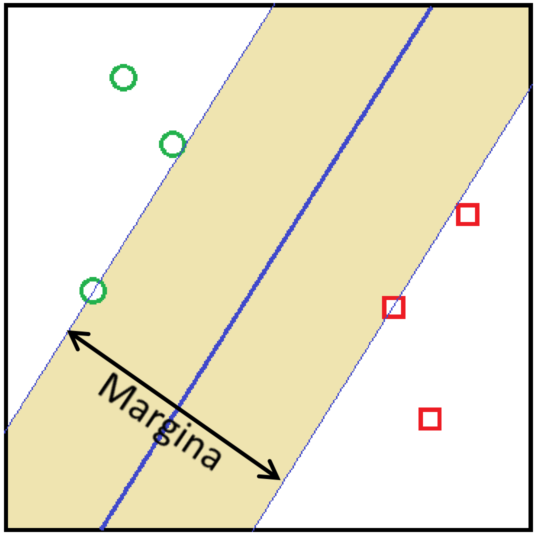
Kako pronaći razdvajajuću hiperravan?

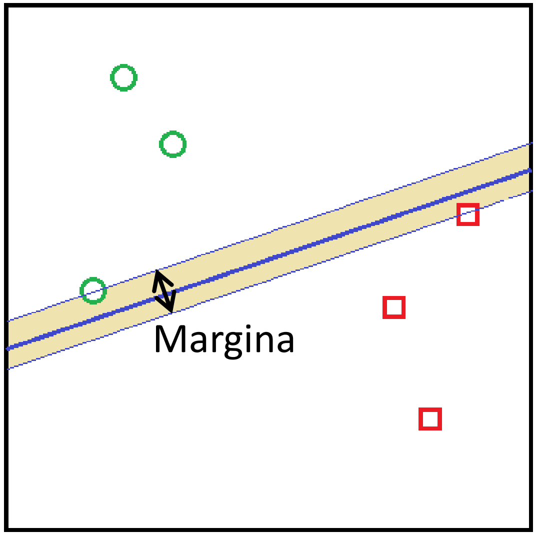




- Postoji beskonačno mnogo linija koje savršeno mogu da razdvoje dve klase
- Koju da odaberemo?
- Intuitivno, plava je najbolja jer je najudaljenija od primera obe klase
- Uzorak na kome treniramo je ograničen

Margina razdvajajuće hiperravni

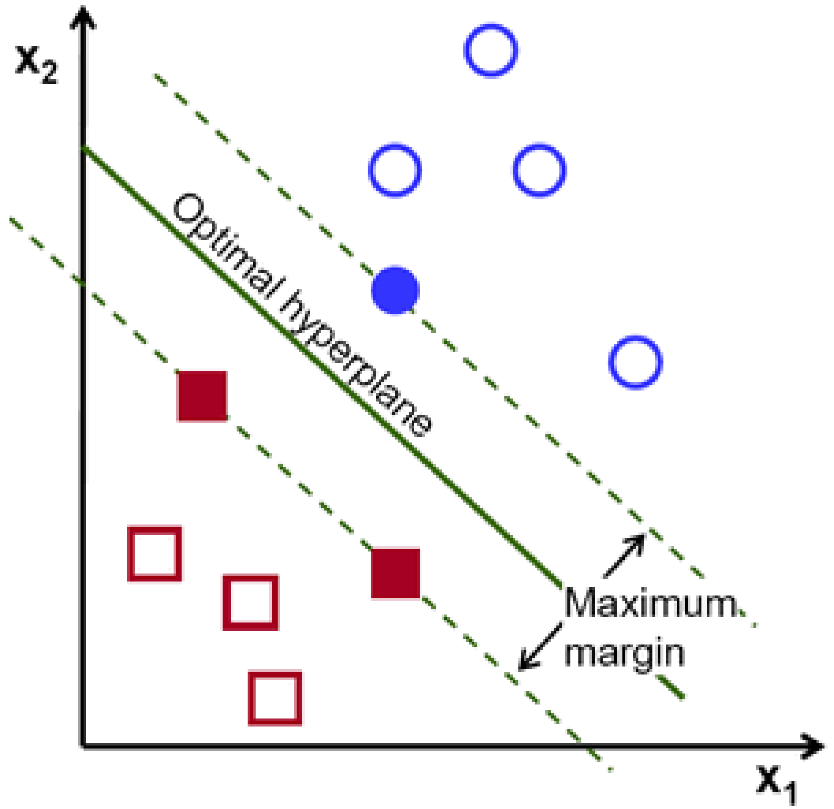




- Obe razdvajajuće hiperravni savršeno klasifikuju (isti) skup podataka
- Ali hiperravan na slici levo ima veću marginu

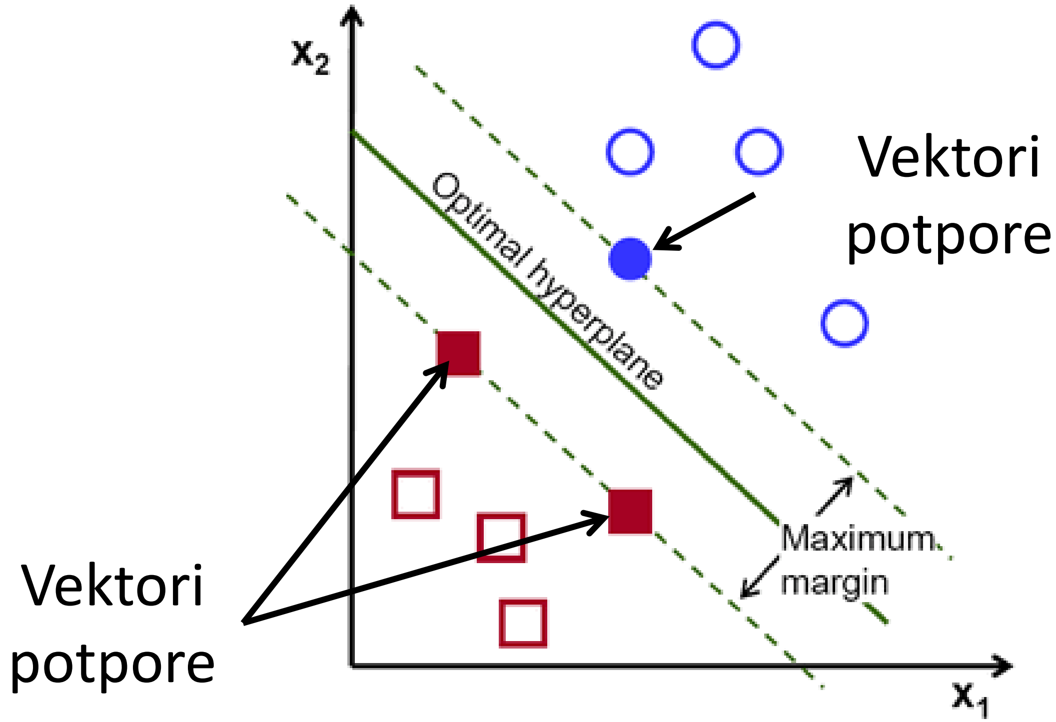
- Margina razdvajajuće hiperravni je minimum rastojanja od te hiperravni do neke od tačaka skupa podataka
- Za datu hiperravan (npr. plava linija)
  - Pomeramo liniju paralelno na jednu stranu dok ne udari o prvi zeleni krug
  - Pomeramo liniju paralelno na drugu stranu dok ne udari o prvi crveni kvadrat
  - Rastojanje ovako dobijene dve linije (paralelne sa datom hiperravni) predstavlja marginu te hiperravni

SVM optimizacioni algoritam



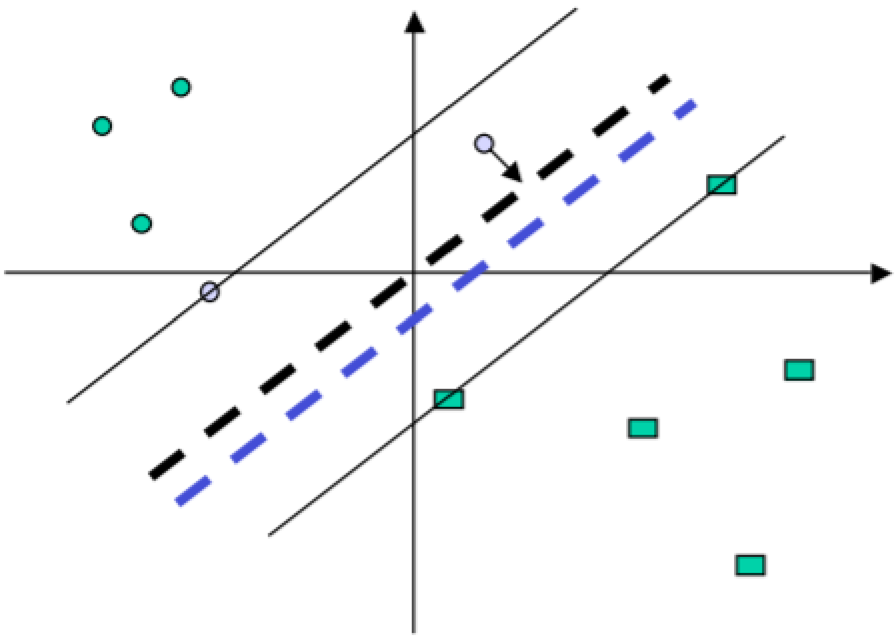
- Među svim hiperravnima koje savršeno razdvajaju podatke na dve klase, pronaći onu sa najvećom marginom

Vektori potpore

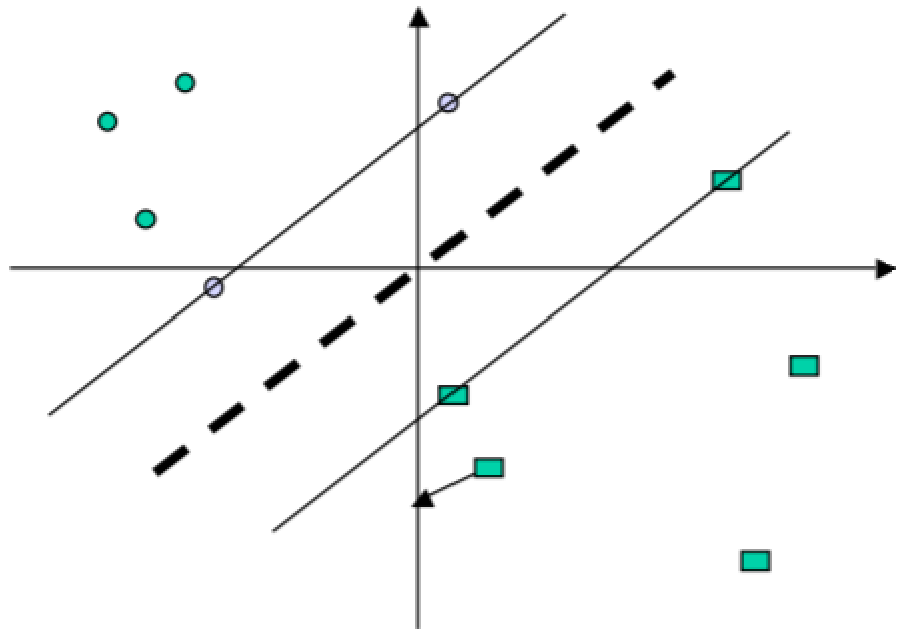


- Vektori potpore su tačke $x^{(n)}$ najbliže hiperravni – kritične tačke koje određuju marginu

Vektori potpore



Pomeranje vektora potpore pomera granicu odluke



Pomeranje ostalih vektora ne utiče na granicu odluke

- Optimizacioni algoritam generiše parametre θ na taj način da samo vektori potpore utiču na njih