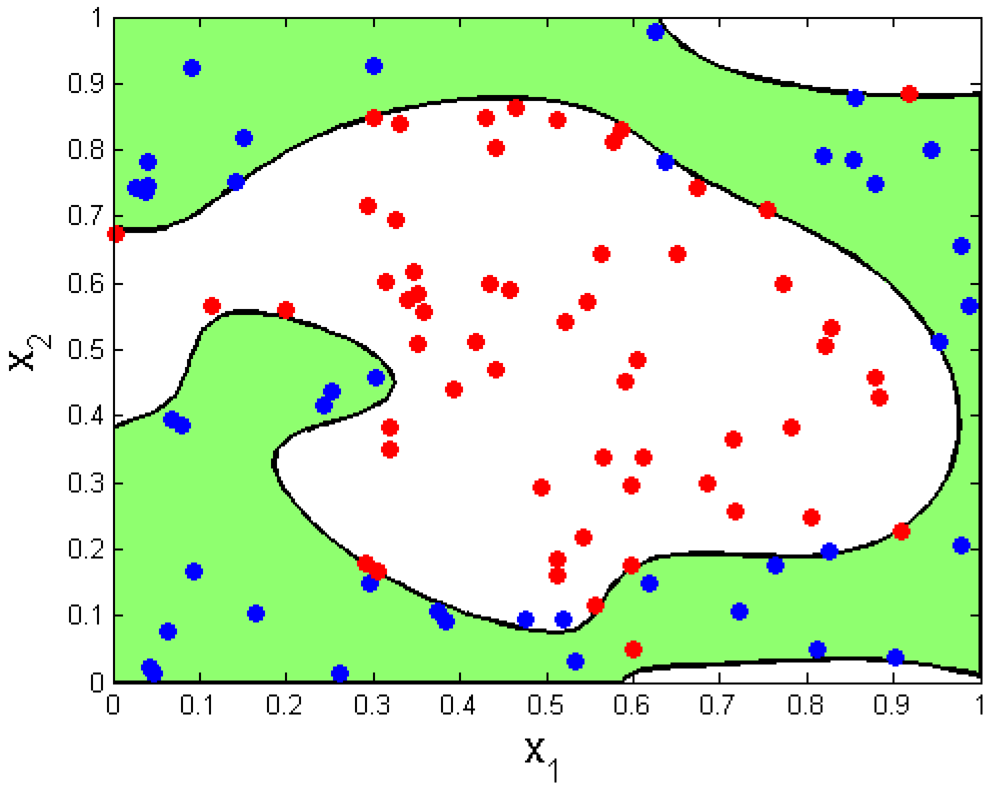
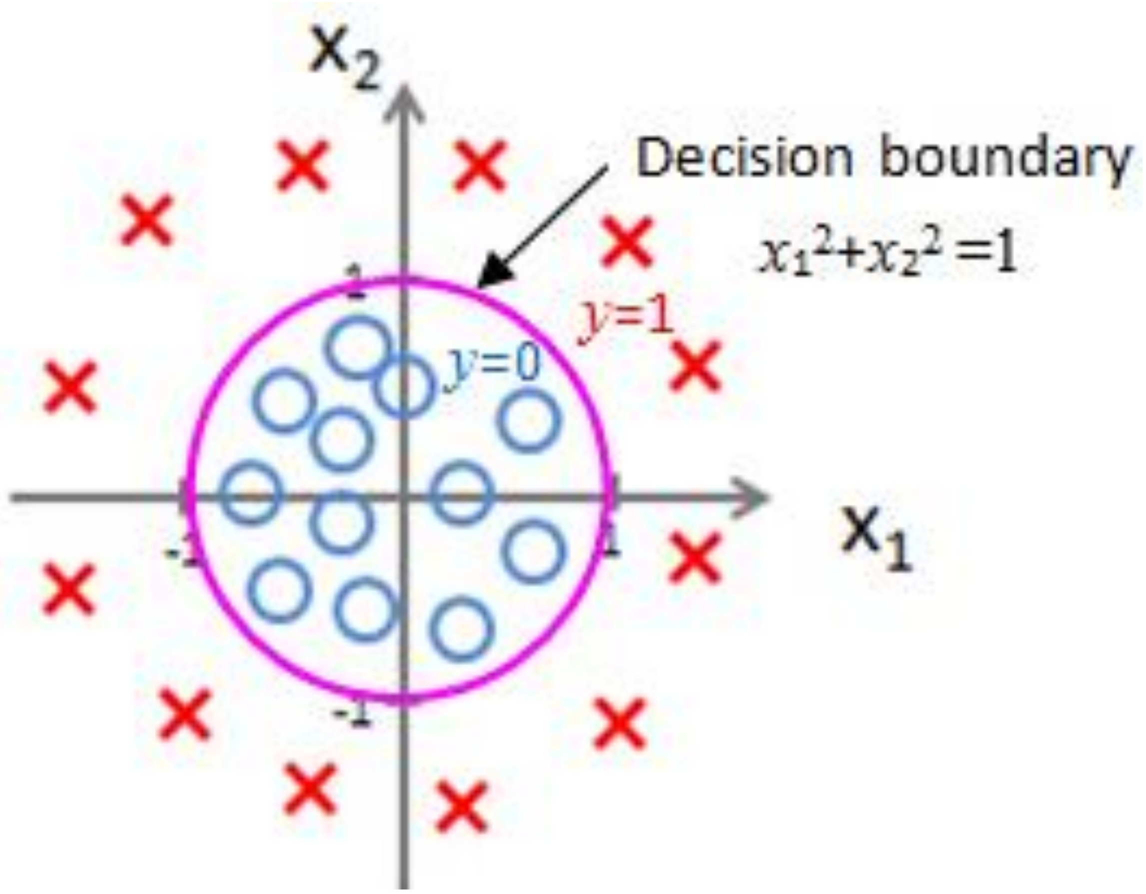


Nelinearna granica odluke

Nelinearna granica odluke

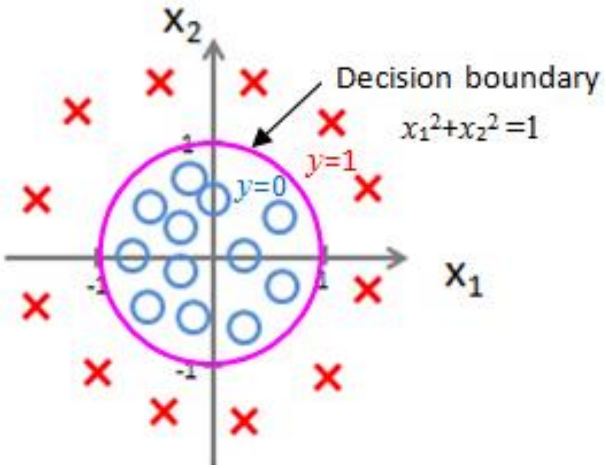


Nelinearna granica odluke

U našu hipotezu možemo dodati obeležja koja odgovaraju višem stepenu polinoma:

$$h_\theta(x) = \mathrm{g}(\theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_1^2 + \theta_4 x_2^2)$$

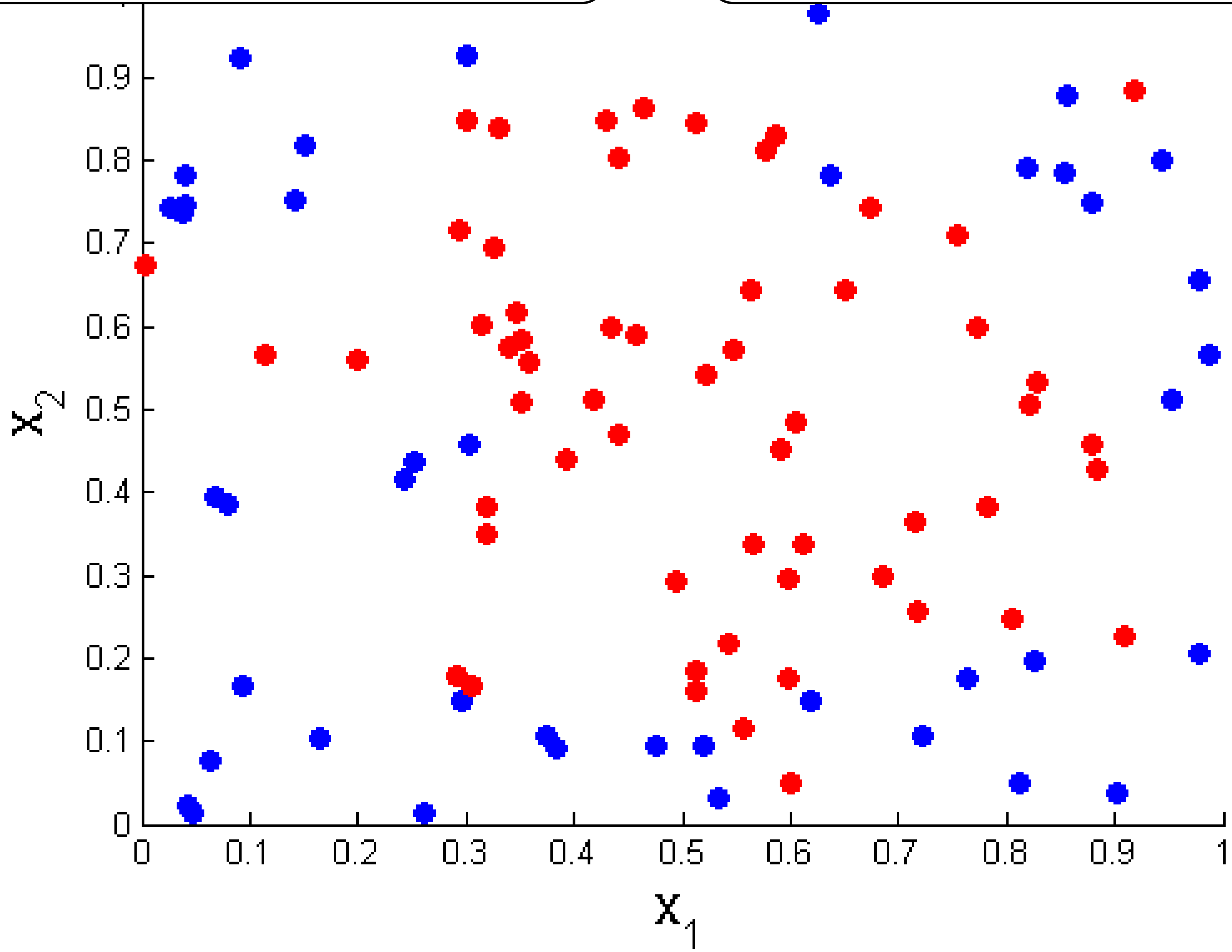
$$\theta = \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \Rightarrow y = 1 \text{ ako je } -1 + x_1^2 + x_2^2 \geq 0$$



Primer

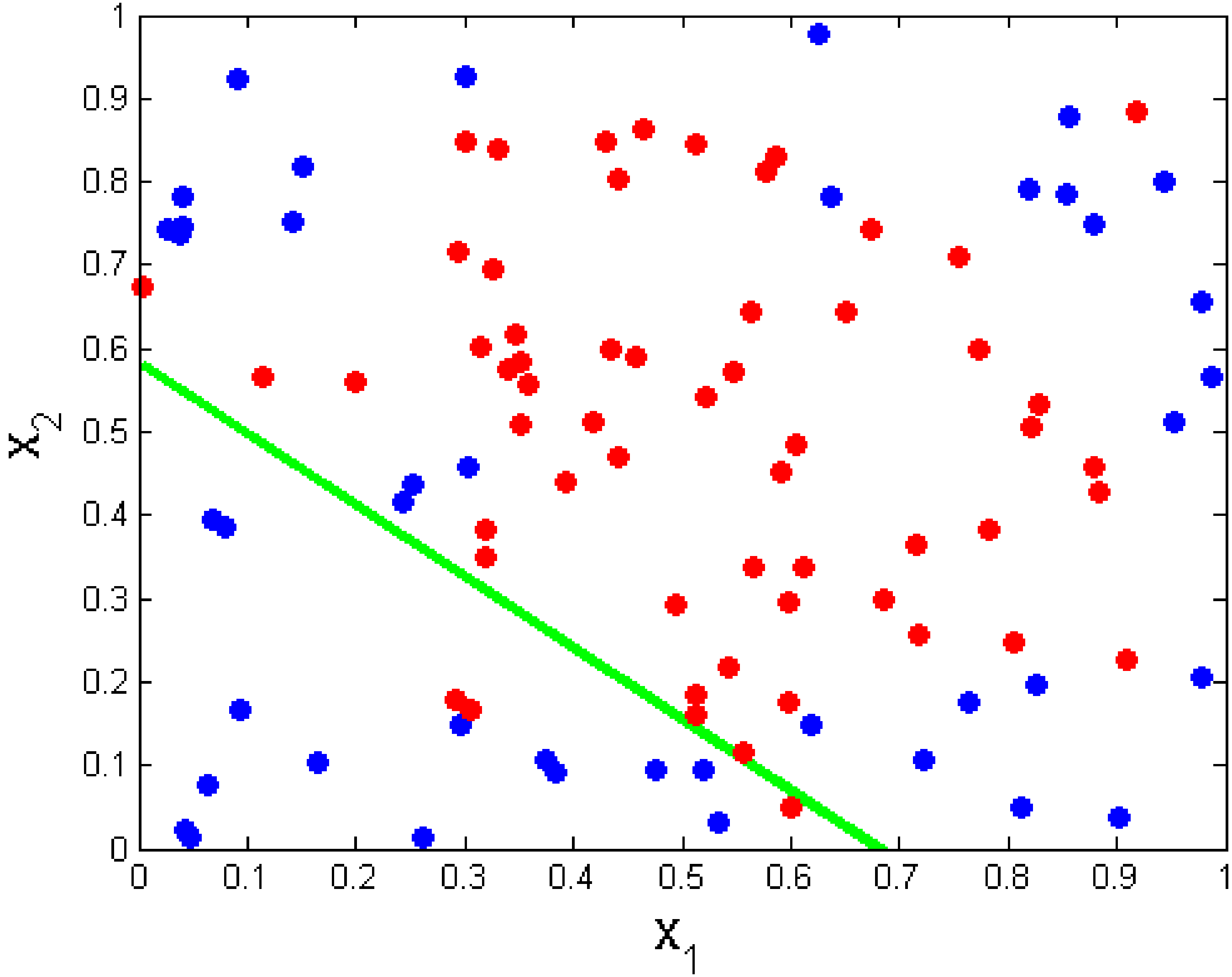
Stvarna granica odluke je kružnica sa centrom u (0.5, 0.5) i poluprečnikom 0.4

Dodali smo i šum: od 100 tačaka odabrali smo 10 i zamenili 0->1 i 1->0



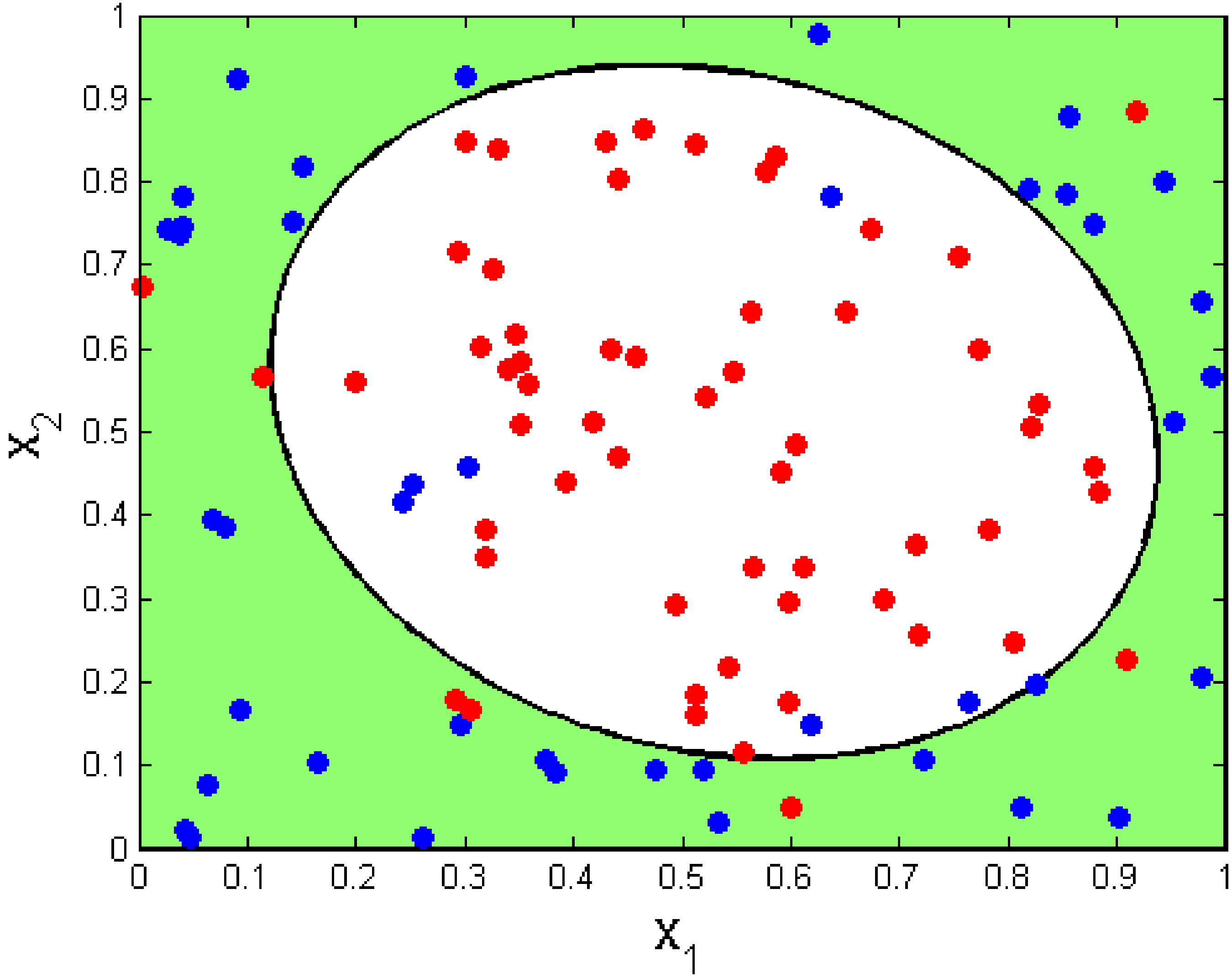
Primer

(x_1, x_2) accuracy=69%



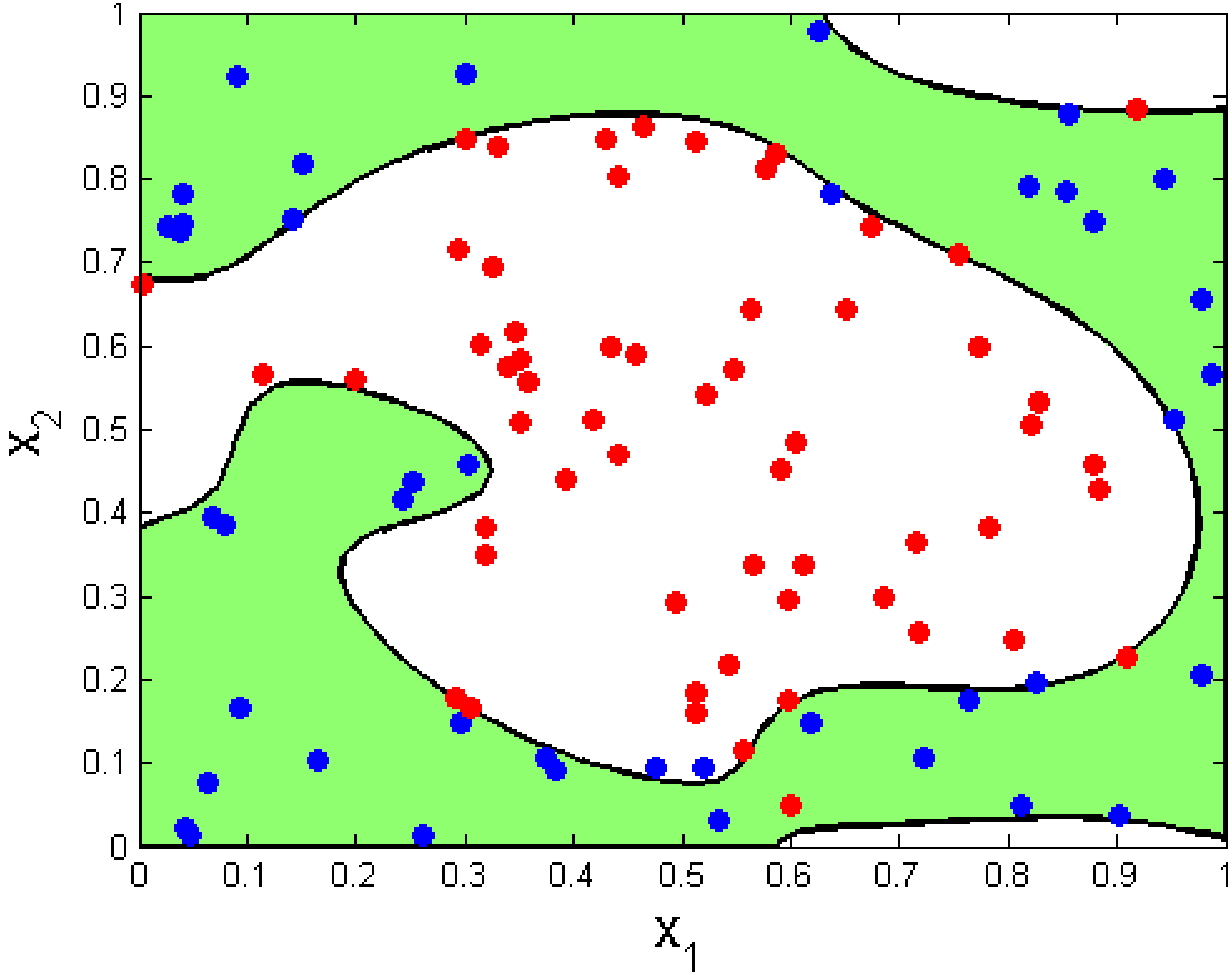
Primer

$(x_1, x_2, x_1 \cdot x_2, x_1^2, x_2^2)$ accuracy=86%

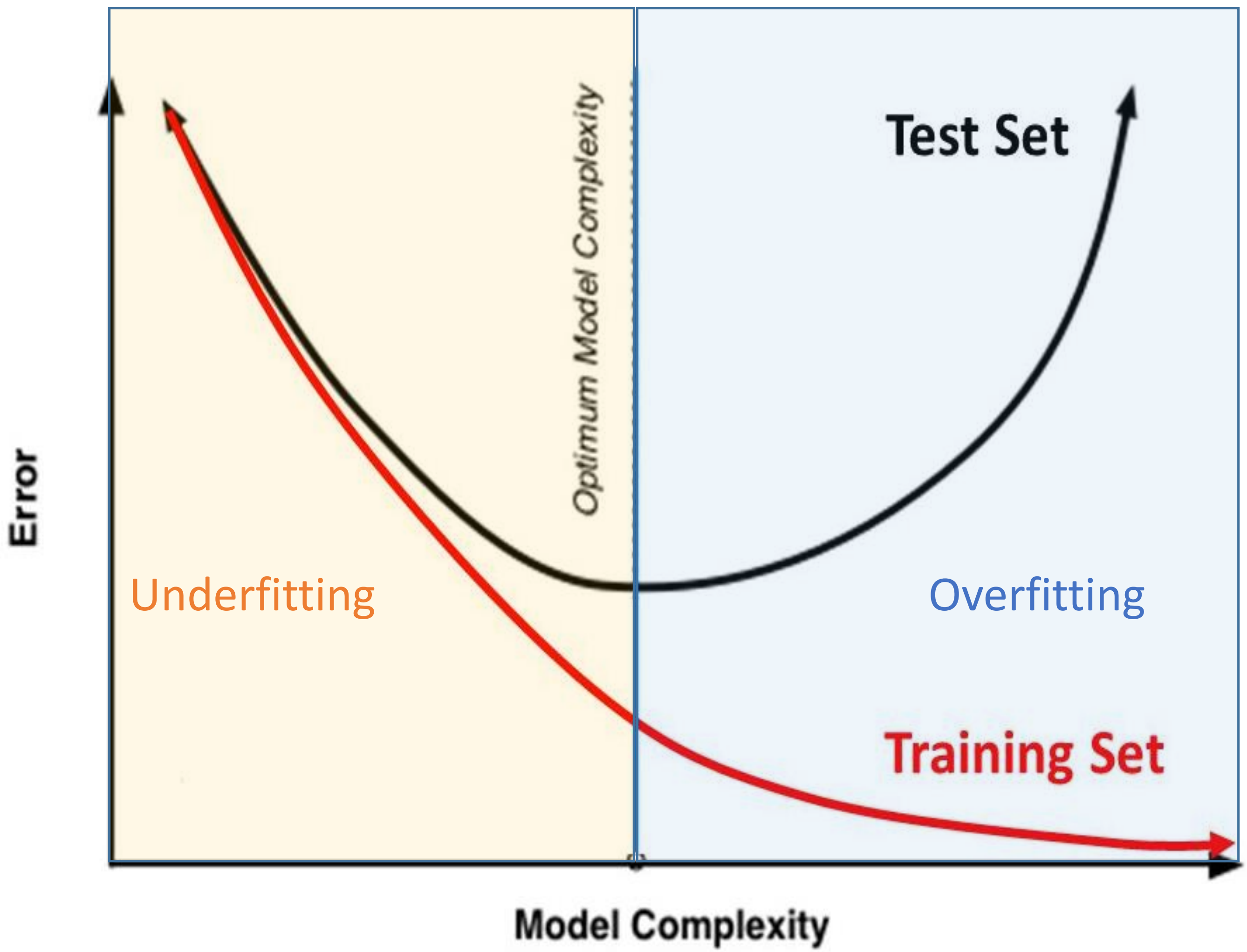


Primer

$\left(x_1, x_2, x_1 \cdot x_2, x_1^2 \cdot x_2, x_2^2, \dots, x_1^5, x_2^5, \dots\right)$ accuracy=94%



Overfitting (podsetnik)



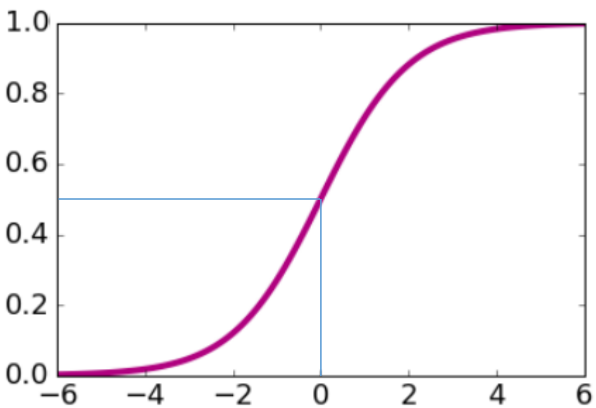
Overfitting kod klasifikatora

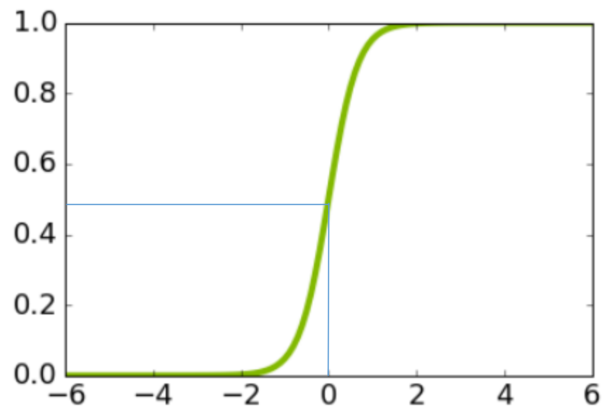
- Kakav efekat ima overfitting na procenjene verovatnoće?
  - Koeficijenti θ će biti velikih magnituda
  - Ovo čini da θx bude velika pozitivna ili velika negativna vrednost
  - U tom slučaju će

$$h_\theta(x) = \sigma(\theta x) = 1/(1 + \exp\{-\theta x\})$$

  Biti veoma blizu 0 ili 1

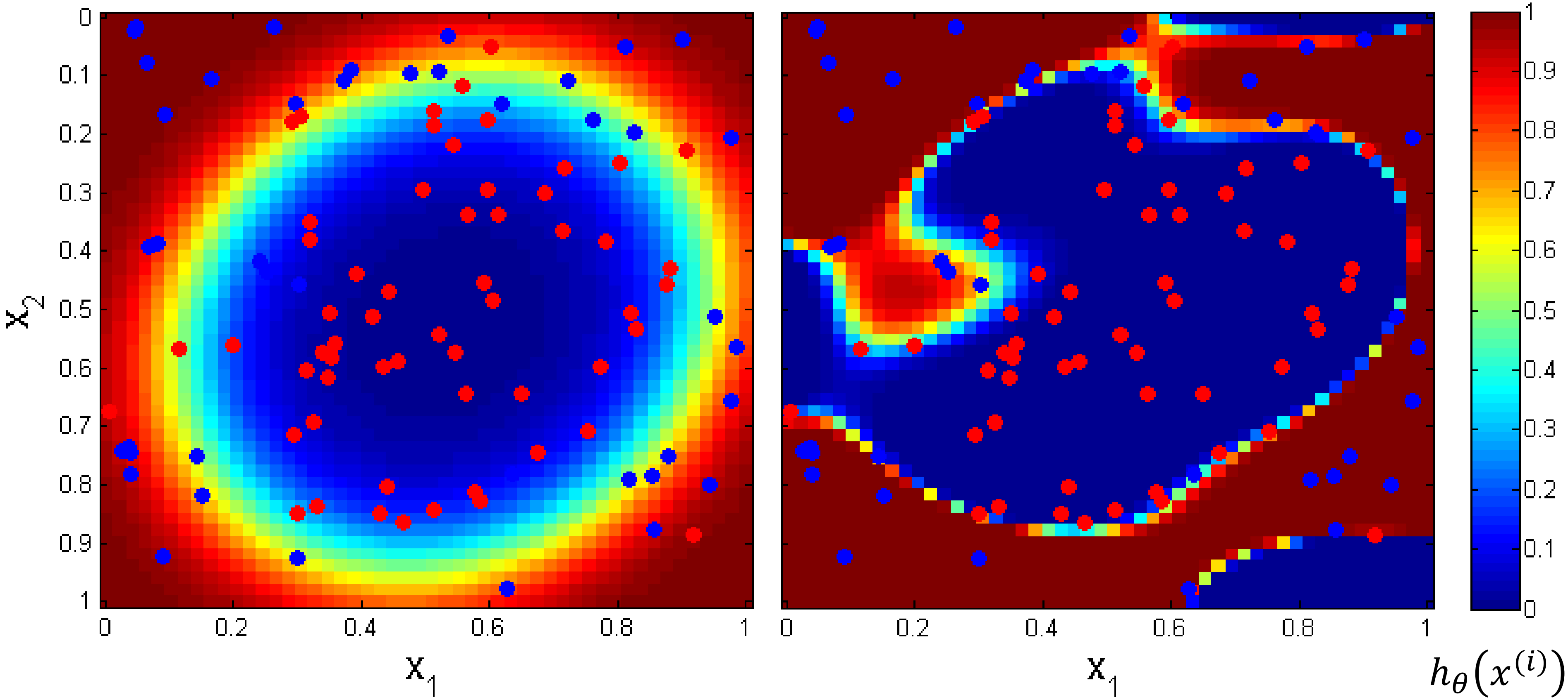
  - Naš model postaje previše pouzdan u svoje predikcije





Što je veća magnitude θ, kriva je strmija – brže postajemo sigurni u predikciju

Overfitting kod klasifikatora



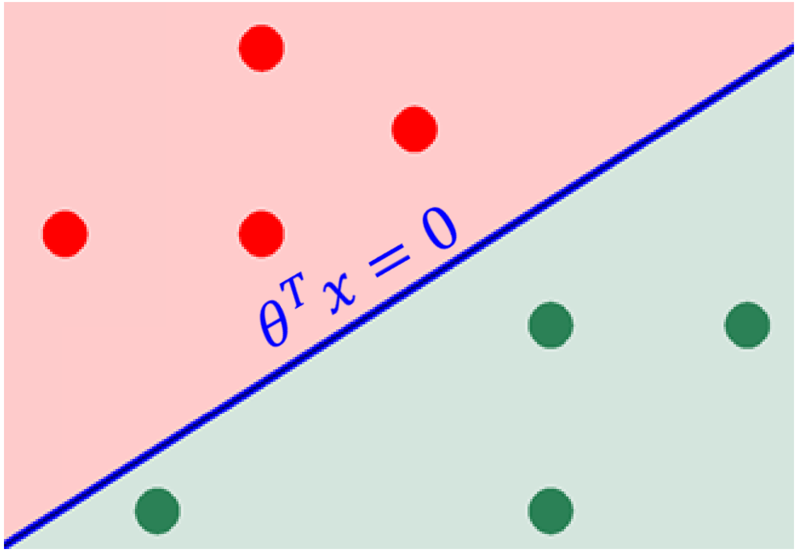
drugi stepeni polinoma

peti stepeni polinoma

nesigurnost oko granice odluke je veća

Uske regije nesigurnosti – prevelika pouzdanost u predikcije

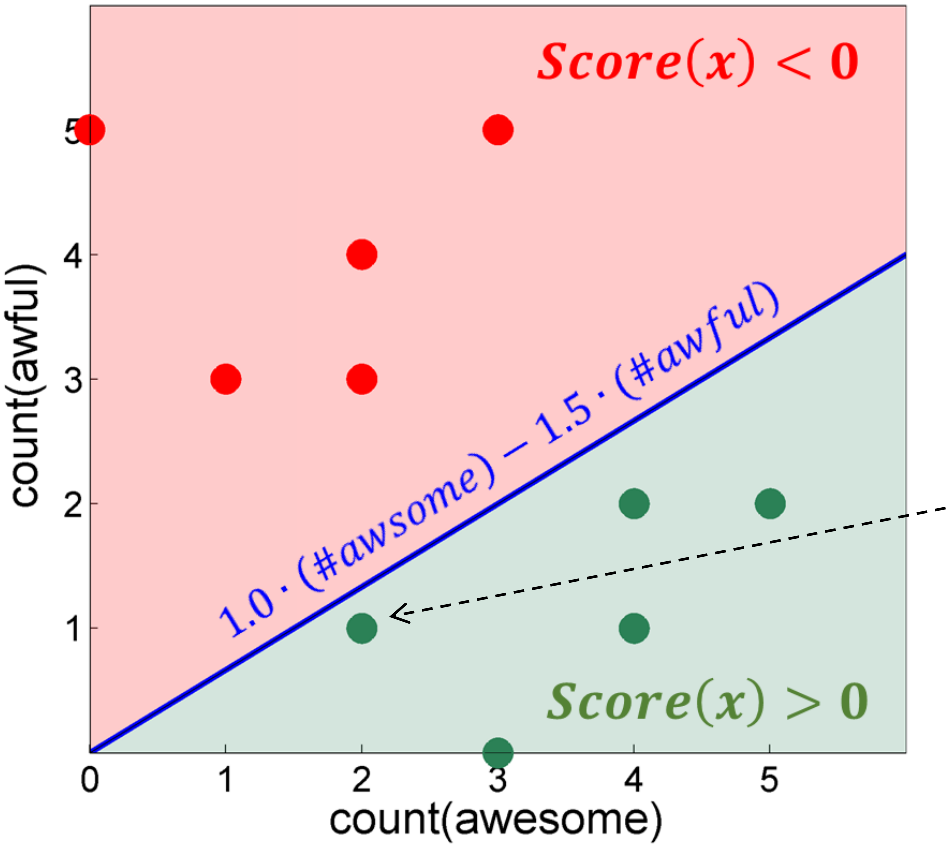
Linearno separabilni podaci



- Podaci su linearno separabilni ako postoje koeficijenti θ takvi da:
  - Za sve pozitivne primere važi $\theta X > 0$
  - Za sve negativne primere važi $\theta X < 0$
- Ovo znači da će greška na trening skupu biti 0

- Ako imamo *D* obeležja, linearna separabilnost se dešava u *D*-dimenzionom prostoru
- Sa dovoljno obeležja, podaci će gotovo uvek biti linearno separabilni!

Efekat linearne separabilnosti na θ



Podaci su linearno separabilni za koeficijente:

a) $\theta_1 = 1, \theta_2 = -1.5$

b) $\theta_1 = 10, \theta_2 = -15$

c) $\theta_1 = 10^9, \theta_2 = -1.5 \cdot 10^9$

a) $P(y = 1|x, \theta) = \frac{1}{1+e^{-0.5}} = 0.62$

b) $P(y = 1|x, \theta) = \frac{1}{1+e^{-5}} = 0.99$

c) $P(y = 1|x, \theta) = \frac{1}{1+e^{-0.5 \cdot 10^9}} \approx 1$

Efekat linearne separabilnosti na θ

- Maximum Likelihood preferira da verovatnoća bude što veća
- Zbog toga će težiti da koeficijenti budu što veći
- Kod linearno separabilnih podataka koeficijenti mogu otići u beskonačnost!