

# Mašinsko učenje – Domaći 3

## Support vector machine

Nemanja Saveski 2023/3163

29. novembar 2023.

# 1 Primalni problem

Za dati skup podataka (*svmData.csv*,  $n$  prediktora,  $m$  primera) vidimo da nije potrebno standardizovati podatke. U slučaju rešavanja primalnog problema potrebno je minimizovati sledeću funkciju:

$$\min_{\omega, b, \xi} \left( \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^m \xi_i \right)$$

$$\text{p.u. } y^{(i)}(\omega^T x^{(i)} + b) \geq 1 - \xi_i, \quad \xi_i \geq 0; \quad i = 1..m.$$

Ovaj problem može da se reši kvadratnim programiranjem:

$$\min \frac{1}{2} x^T P x + q^T x$$

$$\text{p.u. } Gx \leq h, \quad \text{ i } Ax = b$$

gde ćemo naći matrice  $P, q, G, h$  iz datog optimizacionog problema iznad. U ovom delu zadatka nemamo uslov  $Ax = b$ , dok za dualnu formu imamo.

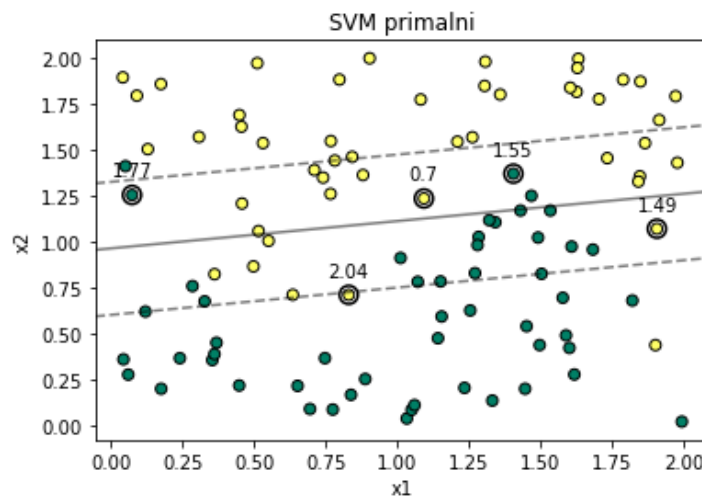
$$P = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}, \quad q = \begin{bmatrix} 0 & C & 0 \end{bmatrix}^T,$$

$$G = \begin{bmatrix} -\text{diag}(y)X & -I & -y \\ 0 & -I & 0 \end{bmatrix}, \quad h = \begin{bmatrix} -1 & 0 \end{bmatrix}^T$$

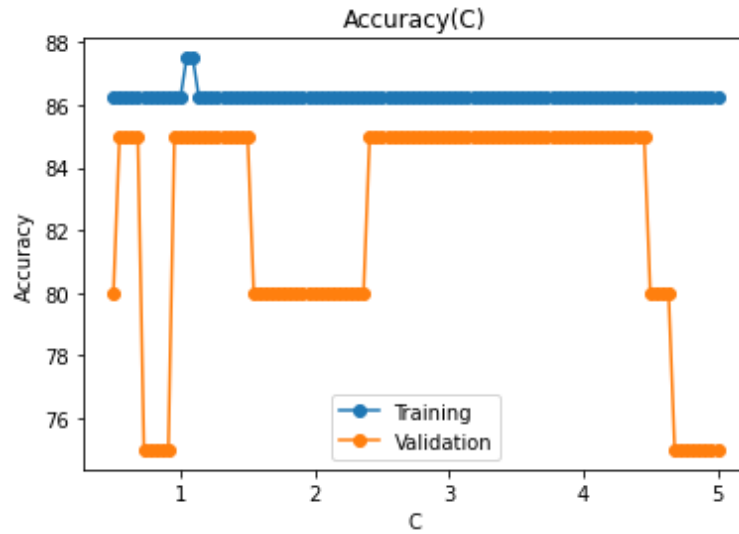
Vektor  $x$  (vektor promenljivih koje minimizujemo) izgleda ovako:

$$x = [\omega_1, \omega_2, \xi_1, \dots, \xi_m, b]^T$$

Podaci su podeljeni u trening i validacioni skup u odnosu 4:1. Dobijamo sledeći klasifikator:



Za ovaj klasifikator dobijeno je optimalno  $C = 1.045$ , prateći tačnost na trening i validacionom skupu:



## 2 Dualni problem sa Gausovskim kernelom

I u ovom delu zadatka će se koristiti kvadratno programiranje kao metod numeričke optimizacije. Izabran je Gausovski kernel:

$$K(x, z) = \exp\left(-\frac{\|x - z\|^2}{2\sigma^2}\right)$$

U ovom slučaju moramo maksimizovati sledeću funkciju:

$$\max_{\alpha} \left( \sum_i \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y^{(i)} y^{(j)} K(x^{(i)}, x^{(j)}) \right)$$

$$\text{p.u.} \sum_{i=1}^m \alpha_i y^{(i)} = 0, \quad 0 \leq \alpha_i \leq C, \quad i = 1..m$$

Matrice potrebne za kvadratno programiranje:

$$P = [y y^T K(X, X)], \quad q = [-1]^T,$$

$$G = [-1 \mid 0]^T, \quad h = [-1 \mid 0]^T$$

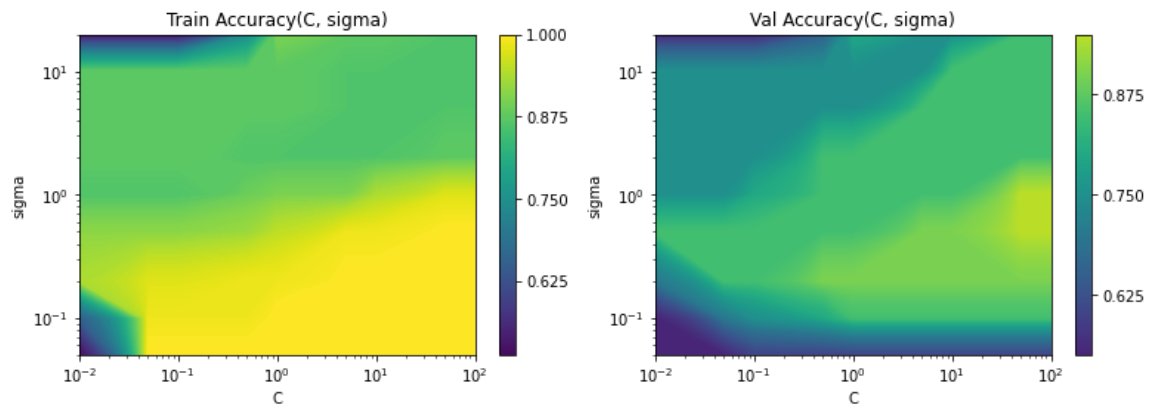
$$A = y^T, \quad b = 0$$

Ovde imamo da je nepoznati vektor  $x$  oblika:

$$x = [\alpha_1, \dots, \alpha_m]^T$$


---

Za izbor parametara  $\sigma$  i  $C$  ćemo ponovo pratiti tačnost na trening i test skupu:



Vidimo sa grafika da su tražene vrednosti  $\sigma = 1$  i  $C = 100$ . Klasifikator sada izgleda ovako:

