

Seminar 6

K-NN, K-Means, Estimare ML

DEDP

1. Fie următorul set de 10 vectori, compus din 5 vectori din clasa A și 5 vectori din clasa B:

- Clasa A:

$$\vec{v}_1 = \begin{bmatrix} 2 \\ -4 \end{bmatrix} \quad \vec{v}_2 = \begin{bmatrix} 1 \\ -5 \end{bmatrix} \quad \vec{v}_3 = \begin{bmatrix} -2 \\ 6 \end{bmatrix} \quad \vec{v}_4 = \begin{bmatrix} -3 \\ 4 \end{bmatrix} \quad \vec{v}_5 = \begin{bmatrix} 2 \\ -5 \end{bmatrix}$$

- Clasa B:

$$\vec{v}_6 = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \quad \vec{v}_7 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \quad \vec{v}_8 = \begin{bmatrix} -4 \\ -3 \end{bmatrix} \quad \vec{v}_9 = \begin{bmatrix} -3 \\ 0 \end{bmatrix} \quad \vec{v}_{10} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$$

Calculați clasa vectorului $\vec{x} = \begin{bmatrix} -2 \\ 5 \end{bmatrix}$ folosind algoritmul k-NN, pentru diverse valori ale lui k : $k = 1$, $k = 3$, $k = 5$, $k = 7$ and $k = 9$

2. Fie următoarele zece valori numerice:

$$\vec{v} = \{v_i\} = [1.1, 0.9, 5.5, 0.6, 5, 6, 1.3, 4.8, 6, 0.8]$$

Efectuați cinci iterații ale algoritmului k-Means pentru a găsi doi centroizi \vec{c}_1 și \vec{c}_2 , pornind de la două valori aleatoare $\vec{c}_1 = 0.95$ și $\vec{c}_2 = 0.96$.

3. Se recepționează un semnal constant de amplitudine necunoscută A, afectat de zgomot gaussian, $r(t) = \underbrace{A}_{s_{\Theta}(t)} + zgomot$, unde zgomotul este de tip gaussian $\mathcal{N}(\mu = 0, \sigma^2 = 2)$.

Semnalul este eșantionat la momentele $t_i = [0, 1.5, 3, 4]$ și se observă valorile $r_i = [4.6, 5.2, 5.35, 4.8]$.

- a. Estimați valoarea lui A folosind estimarea Maximum Likelihood

4. Un semnal de forma $r(t) = A \cdot t^2 + 2 + zgomot$ este eşantionat la momentele $t_i = [1, 2, 3, 4, 5]$, şi valorile obţinute sunt $r_i = [1.2, 3.7, 8.5, 18, 25.8]$. Distribuţia zgomotului este $\mathcal{N}(0, \sigma^2 = 1)$.

a. Estimaţi parametrul A folosind estimarea ML

5. Valorile măsurate ale unei funcţii liniare $y = ax$, unde a este necunoscut, sunt următoarele: $(x_i, y_i) = (1, 1.8), (2, 4.1), (2.5, 5.1), (4, 7.9), (4.3, 8.5)$. Presupunând că zgomotul are distribuţia $\mathcal{N}(0, \sigma^2 = 1)$

a. Estimaţi valoarea lui a folosind estimarea ML

6. Un robot se deplasează pe o traiectorie liniară cu o viteză necunoscută V centimetri/secundă, pornind de la poziţia $x = 0$ la momentul iniţial. La intervale de o secundă, robotul măsoară distanţa parcursă folosind un senzor, afectat de zgomot gaussian $\mathcal{N}(0, \sigma^2 = 0.1)$. Valorile măsurate la momentele $t_i = [1, 2, 3, 4, 5]$ sunt $r_i = [4.9, 9.8, 14.3, 21.2, 25.7]$

a. Estimaţi viteza V a robotului folosind estimarea ML

Hint: Dacă viteza e constantă, distanţa parcursă ar trebui să fie $x = V \cdot t$

b. Preziceţi poziţia robotului la momentul 6.

c. Dacă presupunem la momentul iniţial poziţia robotului nu este 0, ci o valoare necunoscută x_0 , estimaţi perechea de parametri $[V, x_0]$ folosind estimarea ML. Preziceţi poziţia robotului la momentul 6.

d. Ştiind că legea de mişcare este $x(t) = a \cdot t^2 + v_0 \cdot t + x_0$, scrieţi sistemul de ecuaţii pentru găsirea necunoscutelor $[a, v_0, x_0]$. (aceleraţia constantă a , viteza iniţială v_0 , poziţia iniţială x_0).

7. Repetaţi exerciţiul 1 în următoarele condiţii:

a. Considerând zgomotul de tip uniform $U[-2, 2]$. Se mai poate găsi o valoare precisă? Care este intervalul valorilor posibile?

b. Considerând ca zgomotul are distribuţie de forma următoare:

$$w(x) = \begin{cases} 1.5 - \frac{x}{3}, & x \in [0, 3] \\ 0, & elsewhere \end{cases}$$

c. Considerând ca zgomotul are distribuţie de tip exponenţial (λ este doar o constantă)

$$w(x) = \begin{cases} \lambda \cdot e^{-\lambda \cdot x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

Notă: Distribuţia impune ca zgomotul să fie mereu pozitiv, adică valoarea originală e întotdeauna mai mică decât valoarea observată, afectată de zgomot