

Øvingsoppgaver

TEK5020/9020 – Mønstergjenkjenning

Del 3 – Ikke-parametriske metoder

Høsten 2023

Oppgave 1

La den univariate fordelingen $p(x)$ være uniform fra 0 til a , og la vindufunksjonen være gitt ved

$$\varphi(u) = \begin{cases} e^{-u} & u > 0 \\ 0 & u \leq 0. \end{cases}$$

a) Vis at forventningen $\tilde{p}_n(x)$ til vinduestimatet er gitt ved

$$\tilde{p}_n(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{a}(1 - e^{-x/h_n}) & 0 \leq x \leq a \\ \frac{1}{a}(e^{a/h_n} - 1)e^{-x/h_n} & a < x. \end{cases}$$

b) Skisser $\tilde{p}_n(x)$ som funksjon av x for $h_n = a$, $a/4$ og $a/16$.

c) Hvor liten må h_n være for at feilen (differansen mellom estimatet og den faktiske fordelingen) skal være mindre enn én prosent over 99 prosent av intervallet $0 < x < a$?

Oppgave 2

For et todimensjonalt klassifiseringsproblem med to klasser ω_1 og ω_2 er treningssettet gitt ved

$$\mathcal{X}_1 = \left\{ \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 3 \end{bmatrix}, \begin{bmatrix} 4 \\ 4 \end{bmatrix}, \begin{bmatrix} 5 \\ 1 \end{bmatrix}, \begin{bmatrix} 5 \\ 3 \end{bmatrix}, \begin{bmatrix} 6 \\ 2 \end{bmatrix} \right\} \quad (\text{sampler fra } \omega_1)$$

og

$$\mathcal{X}_2 = \left\{ \begin{bmatrix} 3 \\ 6 \end{bmatrix}, \begin{bmatrix} 5 \\ 7 \end{bmatrix}, \begin{bmatrix} 6 \\ 5 \end{bmatrix}, \begin{bmatrix} 7 \\ 1 \end{bmatrix}, \begin{bmatrix} 7 \\ 3 \end{bmatrix}, \begin{bmatrix} 8 \\ 6 \end{bmatrix}, \begin{bmatrix} 9 \\ 2 \end{bmatrix} \right\} \quad (\text{sampler fra } \omega_2).$$

- a) Beregn Parzen-estimetet i punktet $[7.5, 2.0]^t$ for hver av klassene. Anta en hyperkubisk vindu-funksjon (i dette tilfellet et kvadrat) med side $h = 4$.
- b) Klassifiser det ukjente samplet gitt ved egenskapsvektoren $[7.5, 2.0]^t$. Anta at klassenes a priori sannsynligheter er $P(\omega_1) = 0.8$ og $P(\omega_2) = 0.2$.

Oppgave 3

La treningssettet for et univariat toklasseproblem består av 5 sampler fra hver klasse, der

$$\mathcal{X}_1 = \{1.0, 2.9, 4.2, 5.1, 7.3\} \text{ er samplesettet fra klasse } \omega_1 \text{ og}$$

$$\mathcal{X}_2 = \{5.4, 7.9, 8.8, 9.8, 11.2\} \text{ er samplesettet fra klasse } \omega_2.$$

- a) Bruk k-nærmeste-nabo regelen (k-NNR) med $k = 3$ til å klassifisere et objekt med egenskapen $x_0 = 6.1$.
- b) Hva blir klassifiseringsresultatet for det samme objektet med nærmeste-nabo-regelen (NNR)?