

Oppgave 1

Innledning

- Gjør rede for hva som er oppgaven til de ulike komponentene i et typisk mønstergjenkjenningssystem, og lag en skisse som illustrerer et slikt system.
- Nevn et par eksempler på praktisk bruk av mønstergjenkjenning i systemer man kan møte i hverdagslivet.
- Forklar hvordan man kan gå fram for å estimere feilraten til en klassifikator. Hvorfor er det viktig å benytte seg av et uavhengig testsett? Beskriv *leave-one-out* metoden.
- Beskriv problemstillingen i *klyngeanalyse*, og nevnt to hovedtyper av metoder som kan brukes her.

Oppgave 2

Beslutningsteori

- Betinget risk* for en gitt handling α_i kan uttrykkes ved hjelp av kostfunksjoner og a posteriori sannsynligheter ved uttrykket

$$R(\alpha_i|\mathbf{x}) = \sum_{j=1}^c \lambda(\alpha_i|\omega_j)P(\omega_j|\mathbf{x}), \quad i = 1, \dots, a.$$

Gjør rede for størrelsene som inngår i dette uttrykket, og forklar hvilket valg av handling som leder til minimum *total risk* (minimum kostnad). Formulér den tilhørende beslutningsregelen.

- I et endimensjonalt (univariat) problem med to klasser ω_1 og ω_2 , der $a = c$ (antall handlinger lik antall klasser), er sannsynlighetstetthetsfunksjonene på formen

$$p(x|\theta) = \theta^2 x e^{-\theta x},$$

der $\theta > 0$ og $x \geq 0$. Her er θ parameteren som bestemmer fordelingen. La parametrene for de to klassene være henholdsvis θ_1 og θ_2 . Vis at terskelen x_0 som minimaliserer den totale risken er gitt ved

$$x_0 = \frac{1}{\theta_1 - \theta_2} \ln \left[\frac{\lambda_{21} \theta_1^2 P(\omega_1)}{\lambda_{12} \theta_2^2 P(\omega_2)} \right],$$

når $\lambda_{11} = \lambda_{22} = 0$ (null kostnad for feilfri klassifisering).

- Anta $\theta_1 = 1$ og $\theta_2 = 2$. Hva blir terskelen (desisjongsgrensen) x_0 dersom alle feilklassifiseringer koster like mye og apriorisannsynlighetene er like?
- Lag en skisse som viser fordelingsfunksjonene, terskelen og desisjonsregionene i dette tilfellet.

Oppgave 3

Parametriske metoder

a) Den multivariate normalfordelingen er gitt ved

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} \exp \left[-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^t \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right].$$

Forklar størrelsene som inngår. Hva er parametrene i denne fordelingen?

b) Gjør rede for *maksimum-likelihood* metoden for estimering av parametervektoren $\boldsymbol{\theta}$ i en antatt fordelingsfunksjon $p(\mathbf{x}|\boldsymbol{\theta})$ ved ledet læring.

c) Bruk maksimum-likelihood metoden til å finne et estimat for $\boldsymbol{\mu}$ i den multivariate normalfordelingen i deloppgave a, ved hjelp av treningssettet bestående av egenskapsvektorene $\mathbf{x}_1, \dots, \mathbf{x}_n$, når kovariansmatrisen i fordelingen antas kjent. Hva må man forutsette om vektorene i treningssettet?

Oppgave 4

Diskriminantfunksjoner

a) I et todimensjonalt (bivariat) problem med to klasser ω_1 og ω_2 er klassene multivariat normalfordelte med felles kovariansmatrise

$$\Sigma = \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix},$$

og forventningsvektorene

$$\boldsymbol{\mu}_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \text{ og } \boldsymbol{\mu}_2 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}.$$

La apriorisannsynlighetene for klassene være $P(\omega_1) = 1/3$ og $P(\omega_2) = 2/3$. Utled en toklasse diskriminantfunksjon for dette problemet, og finn en likning for desisjongsgrensen.

b) Hvilken form har desisjongsgrensen mellom klassene i dette tilfellet? Forklar hvorfor.

c) Klassifiser egenskapsvektoren

$$\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

ved hjelp av diskriminantfunksjonen fra deloppgave a.

d) Lag en figur som viser forventningsvektorene $\boldsymbol{\mu}_1$ og $\boldsymbol{\mu}_2$ og punktet \mathbf{x}_0 i egenskapsrommet. Skissér desisjongsgrensen i figuren.

e) Hvordan ville desisjongsgrensen endre seg dersom apriorisannsynlighetene for klassene hadde vært like, og hva blir klassifiseringsresultatet av \mathbf{x}_0 i dette tilfellet?

Oppgave 5

Ikke-parametriske metoder

- a) Gjør rede for fordeler og ulemper ved ikke-parametriske metoder sammenliknet med parametriske metoder, og sett opp et uttrykk for tetthetsestimatet i et punkt \mathbf{x} basert på et treningssett av egenskapsvektorer $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$.
- b) Hvilke betingelser må oppfylles for å oppnå konvergens av estimatet til den sanne tettheten $p(\mathbf{x})$ i punktet \mathbf{x} når $n \rightarrow \infty$?
- c) Sett estimatet fra deloppgave a inn i Bayes formel for å utlede et estimat av a posteriorsannsynlighet i et vilkårlig punkt \mathbf{x} for hver av klassene i et problem med c klasser. Forklar hvordan dette estimatet leder til *k-nærmeste-nabo* regelen, og formulér denne beslutningsregelen med ord.
- d) Formulér *nærmeste-nabo* regelen, og angi en øvre grense for den asymptotiske feilraten P som funksjon av den optimale feilraten P^* .