Oppgave 1

Innledning

a) Forklar hva som er formålet med faget *mønstergjenkjenning* (Pattern Recognition), og nevn noen eksempler på praktisk bruk av faget.

b) Beskriv hovedkomponentene i et typisk mønstergjenkjenningssystem. Ta gjerne utgangspunkt i et system for klassifisering av objekter på et samlebånd.

c) Gi eksempler på inngangsdata (rådata) til et slikt system, og forklar hva som er formålet med *egenskapsuttrekking*.

d) Forklar hva som menes med *ledet læring* (Supervised learning), og redegjør for fordeler og ulemper ved *parametriske* og *ikke-parametriske* metoder.

Oppgave 2

Beslutningsteori

a) Forklar hva som menes med *klassebetinget sannsynlighetstetthet*, *a priori sannsynlighet* og *a posteriori sannsynlighet*, og sett opp *Bayes regel* (Bayes formel) som knytter disse størrelsene sammen.

b) Gjør rede for *minimum feilrateprinsippet*, og sett opp en optimal (minimum feilrate) beslutningsregel for et problem med vilkårlig antall klasser, uttrykt ved klassenes a posteriori sannsynligheter (det kreves ingen utledning).

c) I et klassifiseringsproblem med to klasser ω_1 og ω_2, er de klassebetingede sannsynlighetstetthetene gitt ved univariate (endimensjonale) normalfordelinger

$$p(x|\omega_i) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left[-\frac{(x-\mu_i)^2}{2\sigma_i^2}\right], \quad i = 1,2$$

Redegjør for størrelsene som inngår i disse fordelingene, og forklar hva som er parametrene.

d) Utled en minimum feilrate beslutningsregel for dette problemet uttrykt ved parametrene i fordelingene og klassenes a priori sannsynligheter $P(\omega_1)$ og $P(\omega_2)$. Hint: her kan det være en fordel å ta utgangspunkt i diskriminantfunksjoner.

e) La forventningsverdiene være $\mu_1 = 1$ og $\mu_2 = 3$ og standardavvikene være $\sigma_1 = \sigma_2 = 1$. Finn desisjonsgrensen (terskelen) mellom klassene i dette tilfellet, dersom $P(\omega_1) = 1/3$ og $P(\omega_2) = 2/3$.

f) Et ukjent objekt har egenskapsverdien $x = 2$. Hvilken klasse blir dette objektet klassifisert til, med terskelverdien fra punktet ovenfor?

Oppgave 3

Parametriske metoder

a) Gjør rede for parametriske metoder i ledet læring, og forklar hva som er forskjellen på *maksimum-likelihoodmetoden* for parameterestimering og *Bayesisk* estimering. I hvilke tilfeller kan det være gunstig å bruke Bayesisk estimering fremfor maksimum-likelihoodmetoden?

b) Sett opp *likelihoodfunksjonen* $p(\mathscr{X}|\boldsymbol{\theta})$ og utled et generelt likningssystem for maksimum-likelihoodestimatet av parametervektoren $\boldsymbol{\theta}$ i fordelingsfunksjonen $p(\boldsymbol{x}|\boldsymbol{\theta})$, basert på treningssettet $\mathscr{X} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ trukket fra fordelingen. Hva må det forutsettes om disse samplene?

c) Finn maksimum-likelihoodestimatet av parameteren θ i den univariate fordelingen gitt ved

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

der $x \geq 0$ og $\theta > 0$. La treningssettet være gitt ved $\mathscr{X} = \{x_{1,\ldots,}x_n\}$.

Oppgave 4

Diskriminantfunksjoner

a) Forklar hva som menes med *diskriminantfunksjoner*, og sett opp en generell beslutningsregel basert på diskriminantfunksjoner for et klassifiseringsproblem med c klasser. Hvor mange diskriminantfunksjoner har man for et slikt problem?

b) La $g(\boldsymbol{x}) = \boldsymbol{w}^t\boldsymbol{x} + w_0$ være en lineær diskriminantfunksjon for et toklasseproblem. Forklar størrelsene som inngår, og vis hvordan $g(\boldsymbol{x})$ kan omskrives til *utvidet* form, som indreproduktet av en utvidet vektvektor $\boldsymbol{a}$ med en utvidet egenskapsvektor $\boldsymbol{y}$. Anta at det opprinnelige egenskapsrommet er av dimensjon d. Hvilken dimensjon har det utvidede egenskapsrommet?

c) Gjør rede for *minste kvadraters metode* for trening av vektvektoren $\boldsymbol{a}$ i et toklasseproblem, ved hjelp av treningssettet $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n$, ved å sette opp et likningssystem der $\boldsymbol{a}$ inngår. Begrunn hvorfor det er ønskelig å finne en best mulig løsning av dette likningssystemet.

d) Sett opp en kriteriefunksjon basert på likningssystemet, og vis hvordan man kan komme frem til en minste kvadraters løsning for $\boldsymbol{a}$ ved å minimalisere denne kriteriefunksjonen med hensyn på vektvektoren (Pseudoinvers løsningsmetode).

Oppgave 5

Ikke-parametriske metoder

a) Gjør rede for fordeler og ulemper ved *ikke-parametriske* metoder fremfor parametriske metoder, og sett opp et uttrykk for tetthetsestimatet i et punkt $\boldsymbol{x}$ basert på et treningssett av egenskapsvektorer $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_n$.

b) Forklar hvordan slike estimater kan brukes til klassifisering av ukjente sampler, og nevn to hovedtyper av metoder.

c) For et todimensjonalt klassifiseringsproblem med tre klasser ω_1, ω_2 og ω_3 er treningssettet $\mathscr{X} = \mathscr{X}_1 + \mathscr{X}_2 + \mathscr{X}_3$ gitt ved følgende delmengder av sampler fra klassene:

$$\mathscr{X}_1 = \left\{ \begin{bmatrix} 1 \\ 9 \end{bmatrix}, \begin{bmatrix} 2 \\ 7 \end{bmatrix}, \begin{bmatrix} 3 \\ 8 \end{bmatrix}, \begin{bmatrix} 4 \\ 8 \end{bmatrix}, \begin{bmatrix} 5 \\ 6 \end{bmatrix} \right\} \quad \text{(egenskapsvektorer fra } \omega_1\text{)}$$

$$\mathscr{X}_2 = \left\{ \begin{bmatrix} 3 \\ 5 \end{bmatrix}, \begin{bmatrix} 5 \\ 7 \end{bmatrix}, \begin{bmatrix} 6 \\ 4 \end{bmatrix}, \begin{bmatrix} 7 \\ 6 \end{bmatrix} \right\} \quad \text{(egenskapsvektorer fra } \omega_2\text{)}$$

$$\mathscr{X}_3 = \left\{ \begin{bmatrix} 4 \\ 3 \end{bmatrix}, \begin{bmatrix} 5 \\ 1 \end{bmatrix}, \begin{bmatrix} 5 \\ 4 \end{bmatrix}, \begin{bmatrix} 7 \\ 2 \end{bmatrix}, \begin{bmatrix} 8 \\ 4 \end{bmatrix} \right\} \quad \text{(egenskapsvektorer fra } \omega_3\text{)}$$

Estimer a priorisannsynlighetene $P(\omega_1)$, $P(\omega_2)$ og $P(\omega_3)$ ut fra antall treningssampler i hver klasse, og bruk *vindumetoden* (Parzen-metoden) med hyperkubisk vindufunksjon (i dette tilfellet et kvadrat) med side $h = 3$ til å klassifisere et ukjent objekt i punktet $[4, 6]^t$ i egenskapsrommet.