Løsningsforslag, TEK5020, eksamen 2018H

Oppgave 1

a) Kort beskrivelse av prosessen sensor → egenskapsuttrekking → klassifisering → klasse; litt mer om hva klassifikatoren gjør.

b) Kort om ledet læring, nevne parametriske og ikke-parametriske metoder, eventuelt trening av vekter i diskriminantfunksjoner.

c) Empirisk metode; teste på uavhengige data (testsett), estimert feilrate = antall feil / antall objekter testet.

d) Forklare overtrening pga. for mange egenskaper/parametre i forhold til antall treningsobjekter.

Oppgave 2

a) Kort forklaring; sette opp $P(\omega_i|\bar{x}) = \frac{p(\bar{x}|\omega_i)P(\omega_i)}{p(\bar{x})}$

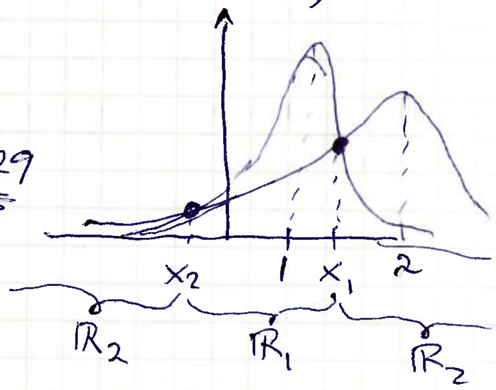
b) Kort forklaring av prinsippet; velge klassen med størst a posteriori sannsynlighet:

$$\boxed{\text{Velg } \omega_i \text{ hvis } P(\omega_i|\bar{x}) \geq P(\omega_j|\bar{x}) \ \forall\, j \neq i}$$

c) Sette inn i likningen $p(x|\omega_1)P(\omega_1) = p(x|\omega_2)P(\omega_2)$ og løse at forskjellene er gitt ved

$$3x^2 - 4x - 2\ln 2 = 0$$

som gir løsningene $x_1 \approx 1,62$ og $x_2 \approx -0,29$



d) Forklar handlinger α_i, $i = 1, \dots, a$ og kostnader $\lambda(\alpha_i|\omega_j)$ til handlinger gitt klasse.
Sett opp betinget risiko:

$$R(\alpha_i|\bar{x}) = \sum_{j=1}^{c} \lambda(\alpha_i|\omega_j)P(\omega_j|\bar{x})$$ mhp. målt egenskapsvektor $\bar{x}$.

Her velger man handlingen med minst betinget risiko; dermed minimaliseres den totale risikoen.

Oppgave 3

a) Kort beskrivelse av ML-metoden:

Ønsker å minimalisere $p(\mathcal{X}|\bar{\theta}) = \prod_{k=1}^{n} p(\bar{x}_k|\bar{\theta})$ med hensyn til parametervektor $\bar{\theta}$. Her er $\bar{x}_1, \ldots, \bar{x}_n$ treningssamplene. Kort forklaring på hvorfor man har ~~t~~ heller kan minimalisere log-likelihoodfunksjonen:

$$\mathcal{L}(\bar{\theta}) = \ln p(\mathcal{X}|\bar{\theta}) = \sum_{k=1}^{n} \ln p(\bar{x}_k|\bar{\theta}).$$

Nødvendig betingelse for minimum:

$$\nabla_{\bar{\theta}} \mathcal{L}(\bar{\theta}) = \sum_{k=1}^{n} \nabla_{\bar{\theta}} \ln p(\bar{x}_k|\bar{\theta}) = 0$$

b) Sette Rayleigh-fordelingen inn i likningen ovenfor, og vis at parameterestimatet blir:

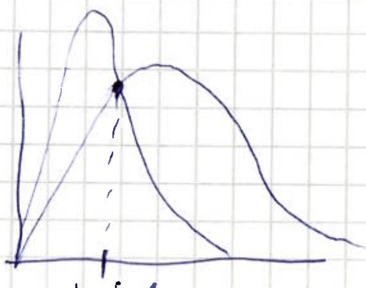
$$\hat{\theta} = \frac{n}{\sum_{k=1}^{n} x_k^2}$$

c) Minimum feilrateprinsippet ~~sier~~ leder her til at man skal velge klassen med størst estimert tetthet, dvs.

Velg ω_1 hvis $\ln p(x|\hat{\theta}_1) - \ln p(x|\hat{\theta}_2) > 0$ (f.eks.) og ω_2 ellers

Innsetting av fordelingen gir da:

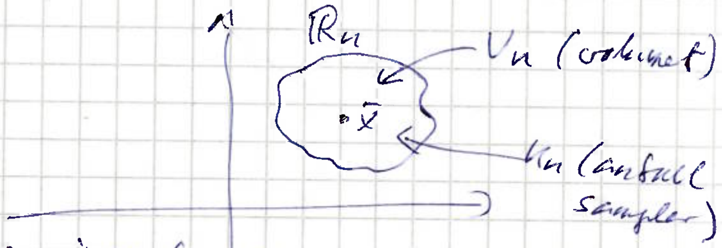
Velg ω_1 hvis $\ln(\hat{\theta}_1/\hat{\theta}_2) + (\hat{\theta}_2 - \hat{\theta}_1)x^2 > 0$, ω_2 ellers.



Oppgave 4

a) Kort beskrivelse av prinsippet; telle opp antall treningsobjekter (fra aktuell klasse) innenfor region som omslutter punktet $\bar{x}$. Sette opp estimatet:

$$P_n(\bar{x}) = \frac{k_n/n}{V_n}$$



k_n = antall sample fra klassen innenfor regionen $\mathbb{R}_n$

V_n = volumet av $\mathbb{R}_n$

n = antall treningssampler fra klassen.

b) Utled estimatet

$$P_n(\omega_i|\bar{x}) = \frac{k_i}{k}$$ da

k_i = antall sampler fra ω_i innenfor regionen $\mathbb{R}_n$

k = totalt antall sampler innenfor $\mathbb{R}_n$

ved å sette inn tetthetsestimatet fra a) i Bayes formel.

c) nærmeste-nabo-regelen: Velg samme klasse som det nærmeste treningssamplet i egenskapsrommet.

k-nærmeste-naboregelen: Velg klassen med flest representanter blant de k nærmeste naboene.

d) NNR krever ingen antakelse om fordelingene, samt har en øvre grense ($< 2 \times$ optimal feilrate) for den asymptotiske feilraten.

e) Bedre egnet for praktisk bruk; metodene gir normalt en raskere og mer kompakt klassifikator.

Oppgave 5

a) Funksjoner $g_i(\bar{x})$ av egenskapsvektoren $\bar{x}$, én funksjon pr. klasse, åpner for generell formulering av beslutningsregler:

$$\boxed{\text{Velg } \omega_m \text{ hvis } g_m(\bar{x}) \geq g_j(\bar{x}) \ \forall \ j \neq m}$$

b)

$$g(\bar{x}) = \bar{w}^t \bar{x} + w_0 \quad \begin{cases} \bar{w} = \text{vektvektor} \\ w_0 = \text{skalarvekt (terskelingsvekt)} \end{cases}$$

$$\boxed{\text{Velg } \omega_1 \text{ hvis } g(\bar{x}) > 0, \ \omega_2 \text{ ellers}}$$

c)

$$g(\bar{x}) = w_0 + \bar{w}^t \bar{x} = [w_0, \bar{w}] \begin{bmatrix} 1 \\ \bar{x} \end{bmatrix} = \bar{a}^t \bar{y}$$

der $\bar{a} = \begin{bmatrix} w_0 \\ \bar{w} \end{bmatrix}$ og $\bar{y} = \begin{bmatrix} 1 \\ \bar{x} \end{bmatrix}$ er utvidede vektorer i et $d+1$ dimensjonalt egenskapsrom.

d) En minste kvadraters løsning av $Y\bar{a} = \bar{b}$

gir $\bar{a} = (Y^t Y)^{-1} Y^t \bar{b} = Y^+ \bar{b}$.

Her er Y = dahematrisen $= \begin{bmatrix} \bar{y}_1^t \\ \vdots \\ \bar{y}_n^t \end{bmatrix}$ der hver linje er en utvidet egenskapsvektor, men med snudd fortegn for sample fra ω_2!

og $\bar{b} = [1, 1, \ldots, 1]^t$

e) Her er den Pseudoinverse:

$$Y^+ = \begin{bmatrix} 1/4 & 1/4 & -1/4 & -1/4 \\ -1/10 & -2/10 & -1/10 & -2/10 \end{bmatrix}$$

og $\bar{b} = [1, 1, 1, 1]$

Slik at vektvektoren $\bar{a}$ blir:

$$\bar{a} = Y^+ \bar{b} = \begin{bmatrix} 0,0 \\ -0,6 \end{bmatrix}$$

Desisjonsgrensen (hyperplanet) går da langs y_1-aksen i det utvidede rommet.