Løsningsforslag Unik4590/9590/TTK4205 - 2014H

11. desember 2016

Oppgave 1 - Innledning

a) Se forelesningsnotatene avsnitt 1.2.

b) Se forelesningsnotatene avsnitt 1.3 og kapittel 7.

c) Se forelesningsnotatene avsnitt 2.1.

d) Se forelesningsnotatene avsnitt 2.1.

(litt prosa på hvert spørsmål)

Oppgave 2 - Beslutningsteori

a) Se forelesningsnotatene avsnitt 2.1 (nederst på side 17).

b) Se forelesningsnotatene avsnitt 2.1.2.

c) Minimum-feilrate beslutningsregel:

$$\text{Velg } \omega_1 \text{ hvis } p(x|\omega_1)P(\omega_1) > p(x|\omega_2)P(\omega_2), \text{ ellers } \omega_2$$

Innsetting av normalfordelingen i ulikheten gir

$$\frac{1}{\sqrt{2\pi}\sigma_1}\exp\left[-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2\right]P(\omega_1) > \frac{1}{\sqrt{2\pi}\sigma_2}\exp\left[-\frac{1}{2}\left(\frac{x-\mu_2}{\sigma_2}\right)^2\right]P(\omega_2) \quad \text{(for valg av } \omega_1\text{)}.$$

Forkorting og bruk av logaritmen på begge sider av ulikhetstegnet gir da

$$\left[-\ln\sigma_1 - \frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2\right]P(\omega_1) > \left[\ln\sigma_2 + \frac{1}{2}\left(\frac{x-\mu_2}{\sigma_2}\right)^2\right]P(\omega_2) \quad \text{(for valg av } \omega_1\text{)}.$$

Kvadrering og stryking av ápriorisannsynlighetene (siden de er like) leder til

$$2\ln\frac{\sigma_2}{\sigma_1} > \frac{x^2-2\mu_1 x+\mu_1^2}{\sigma_1^2} - \frac{x^2-2\mu_2 x+\mu_2^2}{\sigma_2^2} \quad \text{(for valg av } \omega_1\text{)}.$$

Kvadrering og stryking av ápriorisannsynlighetene (siden de er like) leder til

$$2\ln\frac{\sigma_2}{\sigma_1} > \left(\frac{1}{\sigma_1^2}-\frac{1}{\sigma_2^2}\right)x^2+2\left(\frac{\mu_2}{\sigma_2^2}-\frac{\mu_1}{\sigma_1^2}\right)x+\left(\frac{\mu_1^2}{\sigma_1^2}-\frac{\mu_2^2}{\sigma_2^2}\right) \quad \text{(for valg av } \omega_1\text{)},$$

som leder til at man skal velge ω_1 dersom

$$\left(\frac{1}{\sigma_2^2}-\frac{1}{\sigma_1^2}\right)x^2+2\left(\frac{\mu_1}{\sigma_1^2}-\frac{\mu_2}{\sigma_2^2}\right)x+\left(\frac{\mu_2^2}{\sigma_2^2}-\frac{\mu_1^2}{\sigma_1^2}\right)+2\ln\frac{\sigma_2}{\sigma_1}>0$$

og ω_2 ellers.

d) Her er $\sigma_1=\sigma_2=\sigma$, og beslutningsregelen forenkles til å velge ω_1 hvis

$$\frac{2}{\sigma^2}(\mu_1-\mu_2)x+\frac{\mu_2^2-\mu_1^2}{\sigma^2}>0$$

og ω_2 ellers.

Desisjonsgrensen er da gitt ved

$$2(\mu_1-\mu_2)x=\mu_1^2-\mu_2^2 \quad \Rightarrow \quad x_0=\underline{\underline{\frac{\mu_1+\mu_2}{2}}}.$$

Oppgave 3 - Parametriske metoder

a) Se forelesningsnotatene avsnitt 3.1 (side 34).

b) Eksponensialfordelingen er gitt ved

$$p(x|\theta)=\begin{cases}\theta\,\mathrm{e}^{-\theta x} & x\geq 0\\ 0 & ellers\end{cases}$$

der $\theta>0$. Likelihoodfunksjonen blir da

$$p(\mathscr{X}|\theta)=\prod_{k=1}^{n}\theta e^{-\theta x_k},$$

der treningssettet er gitt ved $\mathscr{X}=\{x_1,...,x_n\}$. Log-likelihoodfunksjonen blir

$$\mathscr{L}(\theta)=\sum_{k=1}^{n}(\ln\theta-\theta x_k)$$

og gradienten til $\mathscr{L}$ (i dette tilfellet den deriverte mht. θ skal settes til null for å finne maksimum, dvs:

$$\sum_{k=1}^{n}\left(\frac{1}{\theta}-x_k\right)=0 \quad \Rightarrow \quad \hat{\theta}=\underline{\underline{\frac{n}{\sum_{k=1}^{n}x_k}}}.$$

c) Velger diskriminantfunksjoner på formen

$$g_i(x)=\ln p(x|\omega_i)+\ln P(\omega_i).$$

To-klasse diskriminantfunksjonen blir da

$$\begin{aligned}
g(x) =& g_1(x) - g_2(x) \\
=& \ln p(x|\omega_1) + \ln P(\omega_1) - [\ln p(x|\omega_2) + \ln P(\omega_2)] \\
=& \ln(\theta_1 e^{-\theta_1 x}) + \ln P(\omega_1) - \ln(\theta_2 e^{-\theta_2 x}) - \ln P(\omega_2) \\
=& \ln \theta_1 - \theta_1 x - \ln \theta_2 + \theta_2 x + \ln [P(\omega_1)/P(\omega_2)] \\
=& \underline{\underline{(\theta_2 - \theta_1)x + \ln[\theta_1 P(\omega_1)/\theta_2 P(\omega_2)]}}.
\end{aligned}$$

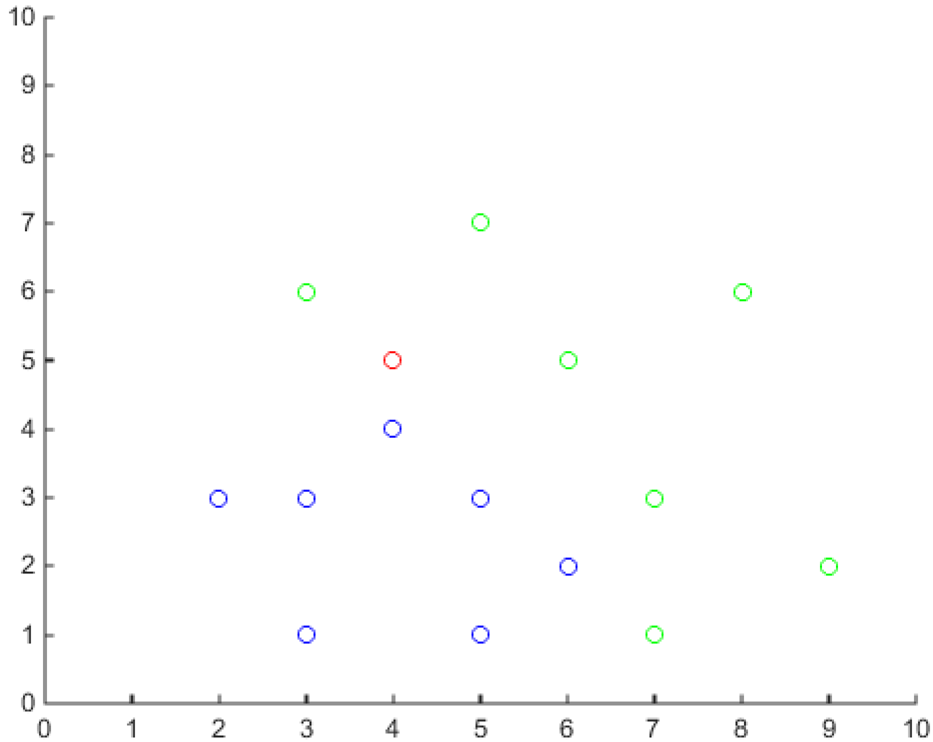
Oppgave 4 - Ikke-parametriske metoder

a) Ikke-parametriske metoder gjør, i motsetning til parametriske metoder, ingen antagelse om formen på de klassebetingede tetthetsfunksjonene. Treningssamplene brukes direkte til å gjøre et punktestimat av tettheten.

b) Se forelesningsnotatene, avsnitt 4.1 (estimatet og figuren på side 50).

c) Se forelesningsnotatene, svsnitt 4.2.1 og 4.2.2.

d) Se figuren. De blå sirklene er samplene fra ω_1, de grønne fra ω_2. Den røde sirkelen er det



ukjente sampelet. Den nærmeste naboen er samplet (4,4) fra ω_1 og de to neste er (3,6) og (6,5) fra ω_2. Det ukjente samplet klassifiseres derfor til ω_1 med NNR og til ω_2 med k-NNR med k=3.

e) Se forelesningsnotatene, avsnitt 4.1 (nederst på side 50).

Oppgave 5 - Trening av lineær diskriminantfunksjon

a) Se forelesningsnotatene, avsnitt 5.3, side 71.

b) Se forelesningsnotatene, avsnitt 5.3.3, nederst på side 72.

c) Se forelesningsnotatene, avsnitt 5.4.3, nederst på side 75.

d) Se løsningsforslag til øvingsoppgave 4.4 under Ø-4.