

Tentamen: Introduction to Pattern Recognition for AI (NB054B)

5 juli 2006, 9:00 - 12:00

Schrijf boven elk vel je naam, studentnummer en studierichting. Schrijf duidelijk en motiveer je antwoorden en schrijf ook eventuele tussenstappen op: dit kan punten schelen!

Opgave 1

Een fabriek produceert halffabrikaten X . 25% van de geproduceerde halffabrikaten is van kwaliteit $x = 1$ en de rest is kwaliteit $x = 2$. Er is een test Z om halffabrikaten te testen. De uitslag van Z is een getal z tussen 0 en 1. Stel dat de testuitslag, afhankelijk van de kwaliteit, als volgt verdeeld is:

$$\begin{aligned}p(z|x=1) &= c_1(1-z) \\ p(z|x=2) &= c_2(1+2z)\end{aligned}$$

voor $0 \leq z \leq 1$. c_1 en c_2 zijn constanten. $p(z|x) = 0$ voor z buiten het interval.

Vraag 1.1 Laat zien dat $c_1 = 2$ en $c_2 = 1/2$.

Vraag 1.2 Bereken met de regel van Bayes de kans op kwaliteit $x = 1$ en $x = 2$ voor het geval dat de testuitslag $z = 0$ is. Doe hetzelfde voor het geval dat de testuitslag $z = 0.5$ en voor het geval dat $z = 1$ is.

Afhankelijk van de testuitslag z wordt er een beslissing genomen, namelijk om het halffabrikaat als $x = 1$ of $x = 2$ te classificeren. De classificatie als $x = k$ is (Bayes) optimaal voor alle z in het gebied R_k waarvoor

$$P(z, x = k) > P(z, x = j)$$

waarbij j de andere kwaliteit is.

Vraag 1.3 Laat zien dat $P(z, x = 1) = \frac{1}{2}(1 - z)$ en $P(z, x = 2) = \frac{3}{8}(1 + 2z)$.

Er zijn dus twee gebieden, R_1 en R_2 . Het punt waar zij elkaar raken is de beslissingsgrens z^* . We gaan z^* op twee manieren bepalen: eerst grafisch, dan analytisch.

Vraag 1.4 (1) Reken voor beide k eerst uit wat $p(z = 0, x = k)$ en $p(z = 1, x = k)$ is. Teken vervolgens in één figuur voor beide k de grafieken van $p(z, x = k)$ door tussen deze punten te interpoleren.

(2) Geef in dezelfde figuur aan waar de gebieden R_k liggen, en bepaal grafisch waar de beslissingsgrens z^* ligt.

Vraag 1.5 Bereken z^* analytisch door de vergelijking

$$p(z, x = 1) = p(z, x = 2)$$

op te lossen. Vergelijk het resultaat met wat je in de vorige vraag hebt gevonden.

Opgave 2

Beschouw de kansverdeling

$$p(x, k) = p(x|k)p(k) = \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp\left(-\frac{(x - \mu_k)^2}{2\sigma_k^2}\right) P(k) \quad (1)$$

met x een 1-d continue variabele en k een discrete variabele die waarden $1, \dots, K$ kan aannemen.

We willen de parameters $\{\mu_k, \sigma_k, P(k)\}$ van dit model schatten met behulp van een gegeven set van N datapunten (x^n, k^n) die onderling onafhankelijk zijn getrokken uit de verdeling (1).

Vraag 2.1 Laat zien dat de negative log-likelihood bij deze data wordt gegeven door

$$E = \sum_{n=1}^N \frac{(x^n - \mu_{k^n})^2}{2\sigma_{k^n}^2} + \sum_{n=1}^N \log \sigma_{k^n} + \frac{N}{2} \log \pi - \sum_{n=1}^N \log P(k^n)$$

en laat zien dat deze vervolgens omgeschreven kan worden als

$$E = \sum_{k=1}^K \sum_{n=1}^N \delta_{kk^n} \frac{(x^n - \mu_k)^2}{2\sigma_k^2} + \sum_{k=1}^K N_k \log \sigma_k + \frac{N}{2} \log \pi - \sum_{k=1}^K N_k \log P(k)$$

met $N_k = \sum_{n=1}^N \delta_{kk^n}$, d.w.z. het aantal datapunten met $k^n = k$.

Vraag 2.2 Laat zien door minimalisatie van E dat de maximum likelihood schattingen van μ_k , σ_k en $P(k)$ worden gegeven door

$$\begin{aligned} \mu_k &= \frac{1}{N_k} \sum_{n=1}^N \delta_{kk^n} x^n \\ \sigma_k^2 &= \frac{1}{N_k} \sum_{n=1}^N \delta_{kk^n} (x^n - \mu_k)^2 \\ P(k) &= \frac{N_k}{N} \end{aligned}$$

Opgave 3

Vraag 3.1 Bespreek de voor en nadelen van de methode van maximum likelihood (hfdstk 2.2) ten opzichte van Bayesiaanse inferentie (hdstk 2.3).

Opgave 4

Beschouw een geregularizeerde error functie van de vorm

$$\tilde{E}(w) = E(w) + \nu \Omega(w)$$

waarbij w een 1-dimensionale parameter is (d.w.z., $w \in \mathbb{R}$). Neem aan dat de ongeregularizeerde error E geminimaliseerd wordt door de gewichtenvector w^* .

Vraag 4.1 Laat m.b.v. een Taylor ontwikkeling rond w^* zien dat als de regularizatie coefficient ν klein is, de gewichten vector \tilde{w} die de geregularizeerde error minimalizeert geschreven kan worden als

$$\tilde{w} = w^* - \nu \frac{\Omega'(w^*)}{E''(w^*)}$$