INTELLIGENTE SYSTEME, 2. SEMINAR – BAYESSCHE ENTSCHEIDUNGEN, LERNEN

Aufgabe 1. Gegeben sei ein Wahrscheinlichkeitsmodell mit 3 verborgenen Klassen, d.h. $k \in \{1,2,3\}$. Die a-posteriori Wahrscheinlichkeitsverteilung der Klassen für eine Beobachtung x sei $p(k|x) = (0.1, 0.6, 0.3)$.

a) Für welche Klasse entscheidet man sich bei der Maximum A-posteriori Entscheidung?

b) Die Menge der Entscheidungen sei mit der Entscheidung „Rückweisung“ ergänzt (siehe Vorlesung, Folie 9). Die Strafe dafür ist ε. Bei welchen Werten von ε wird zurückgewiesen?

c) Man betrachte eine beliebige diskrete a-posteriori Wahrscheinlichkeitsverteilung der Klassen $k \in \{1,2,\dots,K\}$. Die Entscheidungsstrategie ist die Maximum A-posteriori Entscheidung mit Rückweisung. Bei welchen Werten von ε wird nie zurückgewiesen?

Aufgabe 2. Ein Objekt kann sich mit den bekannten a-priori Wahrscheinlichkeiten $p(k)$ in den zwei Zuständen $k = 1,2$ befinden. Die bedingten Wahrscheinlichkeiten für die Merkmale $x \in \mathbb{R}^n$ sind Gauß-verteilt:

$$p(x|k) = \frac{1}{(\sqrt{2\pi}\sigma_k)^n} \exp\left[-\frac{||x-\mu_k||^2}{2\sigma_k^2}\right].$$

(siehe Vorlesung, Folien 7-8).

a) Beide Verteilungen haben dieselbe Streuung, d.h. $\sigma^1 = \sigma^2 = \sigma$, sowie die gleichen a-priori Wahrscheinlichkeiten, d.h. $p(1) = p(2) = 0.5$. Die Kosten für Fehlklassifikationen $C(k,k')$ sind jetzt aber unsymmetrisch:

$$C(k,k') = \begin{cases} 0 & \text{falls } k = k' \\ a & \text{falls } k = 1, k' = 2 \\ b & \text{falls } k = 2, k' = 1 \end{cases}$$

Leiten Sie die zugehörige Bayessche Strategie ab und geben Sie eine geometrische Interpretation.

b) Beide Verteilungen haben *dasselbe* Zentrum, d.h. $\mu^1 = \mu^2 = \mu$ und *unterschiedliche* Streuungen σ^k. Für dieses Wahrscheinlichkeitsmodell soll der Bayessche Klassifikator konstruiert werden. Die Kostenfunktion für Fehlklassifikationen ist die Deltafunktion $\delta(k \neq k')$. Welche geometrische Form hat die Entscheidungsregel?

c) Betrachten Sie nur einen Gaussian als Wahrscheinlichkeitsverteilung einer Zufallsgröße $x \in \mathbb{R}^n$ (siehe Vorlesung, Folien 18-19). Zum Lernen unbekannter Parameter steht eine Lernstichprobe $L = (x_1, x_2, \ldots, x_{|L|})$ zur Verfügung. Wie ergibt sich daraus die unbekannte Streuung σ nach dem Maximum Likelihood Prinzip?

Hinweis: Gehen Sie zunächst so wie bei der Ermittlung des Zentrums μ vor (siehe Vorlesung). Die Formel auf der Folie 19 (oben) muss jetzt aber nach σ abgeleitet werden.

Aufgabe 3. Ein Objekt kann sich mit den a-priori Wahrscheinlichkeiten $p(k)$ in den Zuständen $k = 1, 2$ befinden. Die bedingten Wahrscheinlichkeiten für das skalare Merkmal $x \in \mathbb{R}$ sind

$$p(x|k) = C \cdot \exp\big[-\tau \cdot |x - \mu_k|\big]$$

(τ und μ_k, $k = 1, 2$ sind reellwertige Parameter).

a) Wie ergibt sich die Bayessche Entscheidung für den Objektzustand k bei bekannten τ, μ^k und $p(k)$ (bei der Maximum A-posteriori Entscheidung)?

b) Geben Sie die Parameter an, bei welchen für eine der Klassen nie entschieden wird. Kann man eine solche Situation auch bei Gausschen bedingten Wahrscheinlichkeitsverteilungen konstruieren?

c) Bestimmen Sie den Normierungskoeffizient C.

d) Die unbekannten Parameter μ_k und und τ sollen nach dem Maximum-Likelihood Prinzip anhand einer Lernstichprobe $L = \big((x^l, k^l), \ldots\big)$ angelernt werden. Wie ergeben sich daraus die gesuchten Größen?