STATISTICAL PRINCIPALS AND COMPUTATIONAL METHODS.
3. SEMINAR – INFERENCE

Aufgabe 1. 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 Gaussch verteilt:

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

Leiten Sie die zugehörige Bayessche Strategie ab und geben Sie eine geometrische Interpretation. Welche geometrische Form hat die Entscheidungsregel? Wie ergeben sich die Parameter der Entscheidungsregel aus den bekannten Parametern des Wahrscheinlichkeitsmodells $p(k{=}1)$, $p(k{=}2)$, σ_1, σ_2, μ_1, μ_2? Betrachten Sie dabei die folgenden Spezialfälle:

a) Beide Verteilungen haben *dieselbe* Streuung, d.h. $\sigma_1 = \sigma_2 = \sigma$. Die Kosten für die Fehlklassifikationen ist die Deltafunktion $C(k,k') = \delta(k \neq k')$.

b) Beide Verteilungen haben *dieselbe* Streuung, die Kosten für die Fehlklassifikationen sind als eine (allgemeine) 2×2 Matrix $C_{kk'}$ angegeben.

c) Beide Verteilungen haben *dasselbe* Zentrum, d.h. $\mu_1 = \mu_2 = \mu$ und *unterschiedliche* Streuungen σ_k. Die Kostenfunktion für Fehlklassifikationen ist die Deltafunktion $C(k,k') = \delta(k \neq k')$.

Aufgabe 2. 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 die 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 $p(k)$?

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?

Aufgabe 3. Sei x und y zwei diskrete Zufallsvariablen. Konstruieren Sie eine Wahrscheinlichkeitsverteilung $p(x,y)$ derart, dass die Maximum Marginal Entscheidung

$$x^* = \arg\max_x \sum_y p(x,y)$$
$$y^* = \arg\max_y \sum_x p(x,y)$$

die Wahrscheinlichkeit $p(x^*,y^*) = 0$ hat.

Hinweis: Es genügt solche Variablen x und y zu betrachten, die jeweils nur drei Werte annehmen können. Somit ist die Wahrscheinlichkeitsverteilung $p(x,y)$ eine 3×3 Matrix.

Aufgabe 4. Man betrachte eine Markovsche (a-posteriori) Wahrscheinlichkeitsverteilung

$$p(y) \propto \exp\left[\sum_i \psi_i(y_i) + \sum_{ij} \psi_{ij}(y_i,y_j)\right]$$

Die Erkennungsaufgabe wird als Aufgabe der Bayesschen Entscheidung formuliert:

$$R(y^*) = \sum_y p(y)C(y,y^*) \to \min_{y^*}.$$

Die Kostenfunktion ist dabei die additive Deltakostenfunktion für *Kanten*:

$$C(y,y^*) = \sum_{ij} \delta\big((y_i,y_j) \neq (y_i^*,y_j^*)\big),$$

d.h. sie bewertet Übereinstimmung der *Labelpaare* auf allen Kanten (vergleiche mit der additiven Kostenfunktion $\sum_i \delta(y_i \neq y_i^*)$, die die Übereinstimmung der Label in den Knoten bewertet).

Leiten Sie die daraus folgende Entscheidungsstrategie ab.

Hinweis: Berücksichtigen Sie, dass es in diesem Fall nicht möglich ist, die Entscheidungen auf den Kanten anhand der marginalen Wahrscheinlichkeitsverteilungen der Labelpaare von einander unabhängig zu treffen. Zeigen Sie, dass die gesuchte Entscheidungsstrategie der Lösung eines MaxSum Problems entspricht.