Übungsblatt 3

Übung 3.1 Betrachten Sie das Lachs/Barsch Unterscheidungsproblem der Vorlesung anhand des Features "Größe". Nehmen Sie an, dass die Größen normalverteilt seien mit *gleicher Streuung* σ, aber verschiedenen Mittelwerten:

$$\begin{aligned} p(x|L) &= \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-(x-\mu_L)^2/2\sigma^2} \\ p(x|B) &= \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-(x-\mu_B)^2/2\sigma^2} \end{aligned}$$

Nehmen Sie ferner an, dass die mittlere Größe von Lachsen größer als die von Barschen ist, d.h. $\mu_L > \mu_B$. Die "a Priori" Wahrscheinlichkeiten seien $P(Lachs) = p_L$ und $P(Barsch) = p_B$.

a) Bestimmen Sie die Entscheidungsgrenze (*decision boundary*) für die Größe, unterhalb der für Barsch, und oberhalb der für Lachs entschieden wird.

b) Gibt es Parameterwerte, bei denen niemals für Barsch, sondern immer für Lachs entschieden wird? Oder umgekehrt? An welcher Ungenauigkeit der Modellierung liegt diese Asymmetrie?

Übung 3.2 Die Verteilungen des Features x zweier Klassen ω_1, ω_2 seien beide Cauchyverteilt zu den Mittelwerten (Median) $\mu_1 < \mu_2$, d.h.

$$p(x|\omega_i) \;=\; \frac{\lambda}{\pi} \cdot \frac{1}{\lambda^2 + (x-\mu_i)^2}$$

Die a priori Wahrscheinlichkeiten beider Klassen seien gleich.

a) Zeigen Sie, dass die Bayessche Entscheidungsregel zu einer Entscheidungsgrenze
$(\mu_1 + \mu_2)/2$ führt, also genau in der Mitte zwischen den Maxima der Verteilungen.

b) Berechnen Sie die Bayessche Fehlerrate, d.h. die Fehlerrate, die sich aus der Entscheidungsregel von a) ergibt. <u>Hinweise:</u>

$$\int \frac{1}{x^2+a}\,dx \;=\; \frac{1}{\sqrt{a}}\,\text{arctan}(x/\sqrt{a})$$

$$\text{arctan}(\pm\infty) = \pm\frac{\pi}{2} \qquad \text{und} \qquad \text{arctan}(-x) = -\text{arctan}(x)$$

c) Was ist die maximale Fehlerrate und unter welchen Bedingungen an μ_1 und μ_2 tritt sie auf? Erklären Sie das Ergebnis anschaulich.