Technische Hochschule Ostwestfalen-Lippe
Fachbereich Elektrotechnik und Technische Informatik
Prof. Dr. Helene Dörksen

Exercises and Labs 5 for Lecture “Authentication „ (M.Sc.)

Exercise 5.1 Suppose that the class-conditional probability functions for ω_1 and ω_2 are Gaussians with (μ_i, σ_i) of $(4, 2)$ and $(10, 1)$, and that they have equal prior probabilities ($P_1 = P_2 = 0.5$). What is the optimal decision threshold? (Try by calculation, and then check using CondprobEx5.xls).

Exercise 5.2 What are the decision thresholds for two class-conditional probabilities which are Gaussian in shape, with means $\mu_1 = 4$ and $\mu_2 = 10$, variances $\sigma_1^2 = 4$ and $\sigma_2^2 = 1$, and prior probabilities $P(\omega_1) = 2/3$ and $P(\omega_2) = 1/3$? (Try by calculation, and then check using CondprobEx5.xls).

Exercise 5.3 Select the optimal decision where the class-conditional probabilities are Gaussians (i.e., $\mathcal{N}(\mu, \sigma^2)$), given by $\mathcal{N}(2, 0.5)$ and $\mathcal{N}(1.5, 0.2)$ and the corresponding priors are 2/3 and 1/3. (Try by calculation, and then check using CondprobEx5.xls).

Exercise 5.4 Let us consider the problem of classifying two types of fish: sea bass (class ω_1) and salmon (class ω_2) as they travel down a conveyor belt in a canning factory. Let us say that the lightness of the fish are measured for samples of the two types of fish, and the resulting class-conditional probability density functions, $p(x|\omega_1)$ and $p(x|\omega_2)$ are calculated as presented in lecture (s. slide 13). The season and the locale (and many other variables) determine directly the probability of the two different types of fish being caught, but suppose that for this particular catch, twice as many sea bass as salmon were caught, so that $P(\omega_1) = 2/3$ and $P(\omega_2) = 1/3$. We can determine the posterior probabilities as presented on the slide 13 in the lecture. The lightness of each new fish coming along the conveyor belt will be measured, and we will be able to classify it as sea bass or salmon according to the posterior probabilities.

For this so-named *fish canning problem* customers may not mind some of the more expensive salmon turning up in their cans of sea bass, but may have more objections to misclassified sea bass turning up in their salmon cans. We can factor this into the analysis with the loss function: which term should we set to be the larger λ_{21} [the loss associated with misclassifying a sea bass (ω_1) as salmon (ω_2)] or λ_{12} [the loss associated with misclassifying a salmon (ω_2) as a sea bass (ω_1)]?

Exercise 5.5 In a two-class problem with a single feature, x, the PDFs are Gaussians, (i.e., $\mathcal{N}(\mu, \sigma^2)$), given by $\mathcal{N}_1(0, 0.5)$ and $\mathcal{N}_2(1, 0.5)$. If $P(\omega_1) = P(\omega_2) = 0.5$, find the decision thresholds for:
(i) minimum error probability and
(ii) minimum risk if the loss matrix is $\lambda = \begin{vmatrix} 0 & 0.5 \\ 1.0 & 0 \end{vmatrix}$.
(Use CondprobEx5.xls).

Exercise 5.6 In a two-class, two-feature classification task, the feature vectors described by Gaussian distributions with the same covariance matrix $\Sigma = \begin{vmatrix} 1.1 & 0.3 \\ 0.3 & 1.9 \end{vmatrix}$. The means are $\mu_1 = [0, 0]$ and $\mu_2 = [3, 3]$.
Classify the vector $[1.0, 2.2]$ according to the Bayesian classifier. (Hint: use the Mahalanobis distances to the two means).

Exercise 5.7 advanced self-study

In Fig. 1 the labeled samples are given.

Class ω_1	Class ω_2	Class ω_3
(2.491, 2.176)	(4.218, −2.075)	(−2.520, 0.483)
(1.053, 0.677)	(−1.156, −2.992)	(−1.163, 3.161)
(5.792, 3.425)	(−4.435, 1.408)	(−13.438, 2.414)
(2.054, −1.467)	(−1.794, −2.838)	(−4.467, 2.298)
(0.550, 4.020)	(−2.137, −2.473)	(−3.711, 4.364)

Figure 1: Table for Exercise 5.7

a) Familiarize yourself with the theoretical basics of *k-NN classification*.

b) Classify each of the following vectors:
(2.543, 0.046)
(4.812, 2.316)
(-2.799, 0.746)
(-3.787, -1.400)
(-7.429, 2.329)
by using
(i) the 1-NN rule
(ii) the 3-NN rule