Test Exam: Pattern Recognition

Prof. Dr. Hauke Schramm

Exercise 1. (Questions on the lecture)

a. Which of the presented decision rules is equivalent (i.e. same classification result) to the shown MAP-decision rule? It is possible to mark more than one entry.

$$\hat{k} = \arg\max_k \; p(k \mid x)$$

Answer	**Mark your choice here**
$\hat{k} = \arg\max_k p(k)$	
$\hat{k} = \arg\max_k \; p(x \mid k) \cdot p(k)$	
$\hat{k} = \arg\max_k \; p(x \mid k)$	
$\hat{k} = \arg\max_k \; p(x, k)$	
$\hat{k} = \arg\max_k p(x \mid k) \cdot p(k) / p(x)$	
$\hat{k} = \arg\max_k \{p(x \mid k) \cdot p(k) + C\}$ C: constant	

(1.5 points for each correct cross, minus 1.5 points for each mistake)

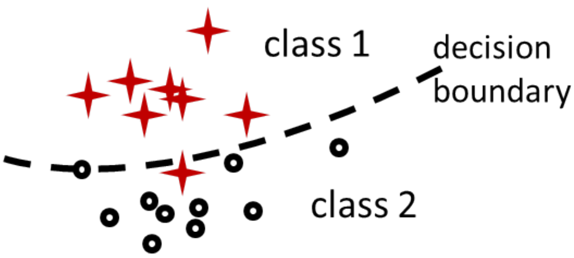
b. Explain, how Bayes rule (Attention: ***not*** Bayes decision rule!) is used in classification and why it is important.

Answer:

c. Give one example for a parametric and for a non-parametric classification technique and discuss the differences between them. What are the advantages / disadvantages of non-parametric techniques?

Answer:

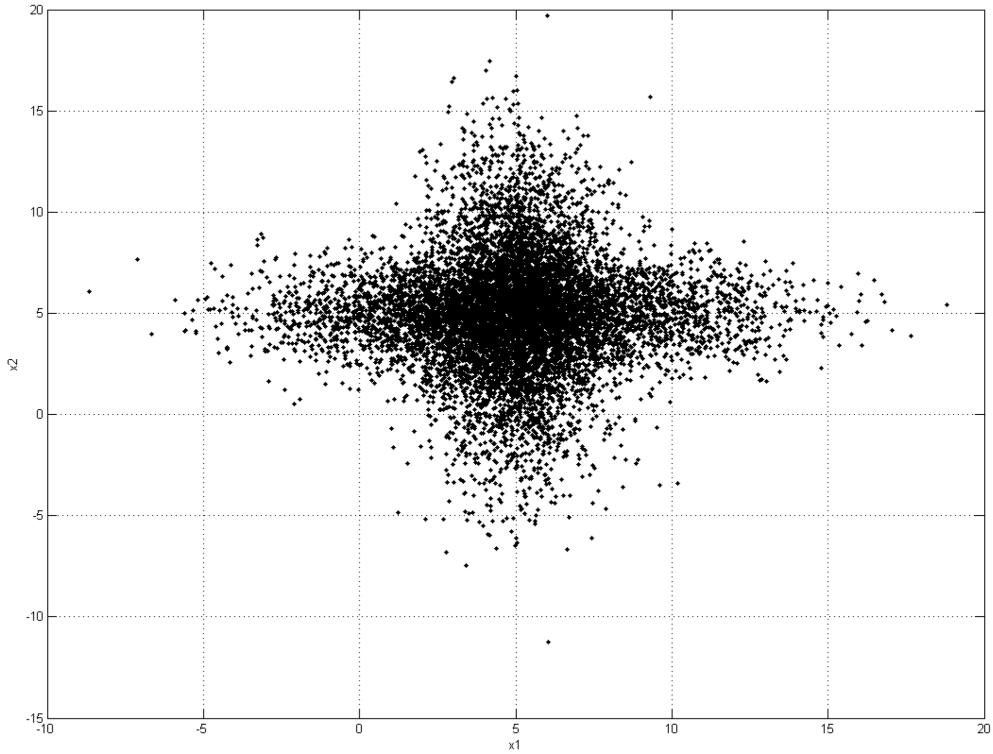
d. Could it happen, using the ***Parzen Window*** approach, that single ***training data*** samples of a class lie on the wrong side of the decision boundary (see illustration)? Explain in detail! (No points without correct and exact justification.)



Answer:

Exercise 2. (Joint distribution)

A measurement of 10000 data samples of two sensor signals x1 and x2 shows the distribution of the figure below:



a. Are x1 and x2 statistically independent? Explain!

b. Estimate the form of the marginal distributions from the given figure and illustrate them. Hint: The maximum probability of the marginals amounts to about 0.05.

c. How would you ***calculate*** the marginals from the given data points? Write a short Matlab script that performs this task. Take into consideration that the marginals are probability distributions! Could you reconstruct the joint probability from the marginals? If yes: How? If no: Why not?

d. Assume, you want to model the point cloud with a

   i. multivariate Gaussian distribution with ***diagonal*** covariance matrix
   ii. multivariate Gaussian distribution with ***full*** covariance matrix

   Describe how you would do that and estimate and illustrate the resulting joint probability for both cases (i. and ii.). Use contour lines for illustration. Is one of the two models better suited to represent the data? Explain!

e. Do you achieve a good representation of the data using the models applied in d.? If not: How could you improve the modeling? Describe your approach in detail and illustrate the expected probability contour image of the improved model.

Answer a.:

Answer b.:

Answer c.:

Generation of marginals:

-
-

Matlab script:

Can the joint probability be constructed from the marginals?

Answer d.:

i.

ii.

Illustration of the distributions:

Are the distributions suited well to model the data?

Answer e.:

Exercise 3. (Maximum-likelihood estimation)

a. Explain the maximum-likelihood parameter estimation. What is the purpose of this technique and how is it applied? Also write down the central equation(s) and explain the symbols.

Answer a.:

For the following two exercise parts you are given

- a simple ***log-likelihood function*** with unknown parameter a:

  $\log p(x \mid a) = -(a-x)^2 + 1$

  (You may ignore the fact that this function is not normalized.)

- ***training samples***:

  $x_1 = 2,\ x_2 = 1,\ x_3 = 4,\ x_4 = 2$

b. Write a ***Matlab script*** that computes the optimization of parameter `a` by brute-force trial and error strategy. You may assume that `a` lies in the range between -10 and 10.

c. ***Mathematically*** perform the maximum-likelihood optimization of parameter `a` for the given likelihood function and training samples. What is the optimal value of a?

Answer b.:

Answer c.: