Pattern Recognition
Examination on 2010-11-02

NO OPEN BOOK! GEEN OPEN BOEK! - It is not allowed to use the course book(s) or slides or any other (printed, written or electronic) material during the exam.
Give sufficient explanations to demonstrate how you come to a given solution or answer!
The 'weight' of each problem is specified below by a number of points, e.g. (15 p).

1. (10 p) Receiver operating characteristics. What do you understand by a receiver operating characteristics (ROC)? To which class of problems does it apply? What is the common property of points that lie on the same ROC curve?

2. (15 p) Iris pattern recognition. Assume that you are given a set of 100 000 binary feature vectors, each of which is a 2000-dimensional binary code of the iris pattern of the right eye of a person. The set contains 100 iris codes of each of 1000 persons.
a) Describe how you would use this data to design an authentication system based on statistical decision theory.
b) What would you do to achieve that the false acceptance rate is below a given level ε.
c) Assume that half of the bits in a test iris code of a person are missing. How would you use this incomplete iris code for authentication at the same confidence level as with complete data (i.e. with the same false acceptance rate)?

3. (15 p) Vector quantization. Consider the following set of data points: S = {(1, 9.5), (3, 8.5), (4, 8), (7, 6.5), (9, 5.5), (14, 3), (16, 2), (17, 1.5)}. Explain the vector quantization algorithm for clustering using three prototypes (cluster centroids) initialized as follows: (0, 6), (8, 8), (15, -5) and learning rate $\eta = 0.3$. Compute the new positions of the prototypes for one epoch and the assignment of points to clusters after that epoch. *Hint*: Plot the data.

4. (15 p) Describe the basic LVQ1 algorithm and the LVQ algorithm with global relevance learning (RLVQ). Which advantages does the RLVQ algorithm offer compared to the basic LVQ1 algorithm? Describe a method that you would use to estimate the classification error of a RLVQ classifier.

5. (15 p) Bayesian decision theory for normal distributions. Let us consider a two-category classification problem, with categories A and B with prior probabilities $P_A = 1/3$ and $P_B = 2/3$. The class-conditional probability densities $p_{x|A}$ and $p_{x|B}$ are one-dimensional normal distributions:

$$p_{x|A} \sim N(\mu_A, \sigma^2_A), \qquad p_{x|B} \sim N(\mu_B, \sigma^2_B)$$

Let us consider the sets of observations {-1, 0.5, 1, 1.5, 3} for category A and {2, 3.5, 4, 4.5, 6} for category B.

a) Compute *unbiased* maximum likelihood estimations of μ_A, σ_A, μ_B, σ_B. $\rightarrow \frac{1}{n-1}$
b) Plot sketches of the two probability density functions.
c) Classify the following points: -2, 0, 2, 3, 5, 7.

analytical boundary!

6. (15 p) Hierarchical clustering.

Consider the following set of points S = {(1, 9.5), (3, 8.5), (4, 8), (7, 6.5), (9, 5.5), (14, 3), (16, 2), (17, 1.5)}. The dissimilarity between two points is defined as the Euclidean distance between them. The dissimilarity between two clusters of points is defined by the dissimilarity of their least dissimilar elements (single linkage algorithm). → min

a) Build a dendrogram for this set.

b) Using the dendrogram, cluster the points in two clusters.

c) Using the dendrogram, cluster the points in three clusters.

Hint: Using the dissimilarity matrix approach will take you a lot of time. You can build the dendrogram faster if you plot the data and decide visually how to cluster data agglomeratively.

Math reminder:

$$\frac{1}{\sqrt{2\pi}\sigma}\exp(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2})$$

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

x	0.5	1	1.5	2	2.5	3	3.5	4
$\exp(-\frac{1}{2}x^2)$	0.8825	0.6065	0.3247	0.1353	0.0439	0.0111	0.0022	0.0003
$\ln(x)$	-0.6931	0	0.4055	0.6931	0.9163	1.0986	1.2528	1.3863

$\sigma = 1/(2n^{1/2})$