SGN-2506 Introduction to Pattern Recognition, Fall 2010, Exam April 2011 / Jari Niemi

Perform all the Problems 1-5. Each of them is worth of 6 points. No literature. Needed formulas are given. Use of your own functional calculator allowed. <u>All computational steps must be shown and explained briefly.</u> You can have the exam question paper also after the exam.

<u>NOTE: There is an extra examination possibility of SGN-2506 in May 2011. Sign-up via POP.</u>

Problems:

1. Let E and F be arbitrary events in the same probability space with the probability function P. "Prove" by using a suitable Venn diagram that $P(E \cup F) = P(F) + P(E) - P(E \cap F)$. Give a special example of a random experiment and events such that $P(E \cup F) = P(F) + P(E)$. (6 p.)

2. Consider the 1-dimensional, 2-class classification problem with equal priors and the continuous class-conditional density functions (a_1 and a_2 are fixed real numbers)

$$p(x|\omega_i) = \frac{1}{\pi(1 + (x - a_i)^2)}, \quad i = 1,2.$$

Find the Bayes minimum error rate decision boundary. (6 p.)

Useful formula: $P(\omega|x) = \frac{P(\omega)p(x|\omega)}{p(x)}$.

3. Classify the point $\mathbf{x} = [1,2,7]^T$ with the 7-nearest-neighbors classifier (based on the Bayes minimum risk classification), as: (6 p.)

1. the actions are the classification decisions: $\alpha_i = \omega_i, \ i = 1,2,3,$
2. the loss function is: $\lambda = \begin{bmatrix} \lambda(\alpha_1|\omega_1) & \lambda(\alpha_1|\omega_2) & \lambda(\alpha_1|\omega_3) \\ \lambda(\alpha_2|\omega_1) & \lambda(\alpha_2|\omega_2) & \lambda(\alpha_2|\omega_3) \\ \lambda(\alpha_3|\omega_1) & \lambda(\alpha_3|\omega_2) & \lambda(\alpha_3|\omega_3) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 2 \\ 3 & 1 & 0 \end{bmatrix},$
3. the training data is:

1 2 3 4 5 6 7 8 9 10

ω_1: $[1,0,1]^T, [0,2,1]^T, [0,0,1]^T, [0,0,0]^T, [2,2,1]^T, [7,5,1]^T, [1,2,1]^T, [1,0,6]^T, [0,0,7]^T, [0,2,2]^T$

1 2 3 4 5 6 7 8

ω_2: $[3,2,1]^T, [4,2,1]^T, [3,4,4]^T, [2,2,5]^T, [4,2,3]^T, [3,3,1]^T, [1,2,8]^T, [5,0,0]^T$

1 2 3 4 5 6 7 8 9 10

ω_3: $[3,2,9]^T, [4,3,7]^T, [5,2,6]^T, [5,2,8]^T, [4,2,9]^T, [3,2,4]^T, [4,2,7]^T, [4,2,8]^T, [3,2,5]^T, [5,3,6]^T$

Useful formulas: $\hat{P}(\omega_i|\mathbf{x}) = \frac{k_i}{k}, \ R(\alpha_i|\mathbf{x}) = \sum_{j=1}^{c} \lambda(\alpha_i|\omega_j)P(\omega_j|\mathbf{x}), \ ||\mathbf{x}|| = \sqrt{\mathbf{x}^T\mathbf{x}}.$

Please, turn over →

4. Let $\mathbf{x}_1 = [4,5]^T$, $\mathbf{x}_2 = [1,4]^T$, $\mathbf{x}_3 = [0,1]^T$. Find the two-class partition of the data $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ that the K-means criterion favors. (6 p.)

Useful formulas: $J(D_1, D_2, \ldots, D_c) = \sum_{i=1}^{c} \sum_{\mathbf{x} \in D_i \neq \emptyset} ||\mathbf{x} - \boldsymbol{\mu}_i||^2$, $\boldsymbol{\mu}_i = \frac{1}{|D_i|} \sum_{\mathbf{x} \in D_i \neq \emptyset} \mathbf{x}$.

5. Write an essay (200-600 words) on the topic *Linear classifiers.* (6 p.)

Instruction: Discuss (at least) the following things (as related to the linear classifiers): definition, discriminant functions, decision boundaries, training error, test error, perceptron algorithm for two-class classification problem. Give an example of a classification problem, where the ideal Bayes minimum error rate classifier is linear.