In The Name of God, The Merciful, The Compassionate

Midterm Exam (120 Minutes – 115 Points)

Statistical Pattern Recognition: CE 40-725 - Spring 2008
Department of Computer Engineering
Sharif University of Technology

---

1. Fundamental Concepts (20 points)

(a) Draw a block diagram for a typical pattern recognition system and briefly explain functionality of each subsystem. What are the general characteristics for suitable features? Specify, what are some good features in a typical image segmentation problem.

b) Class w_1 is normally distributed with $\mu = 2$ and $\sigma^2 = 2$, while class w_2 is uniformly distributed in the range $0 \le x \le 5$. The priors are equal. What decision regions are optimal for this problem? Do you prefer linear or nonlinear classifiers? Why? Sketch the densities and decision regions.

2. Linear Classification (15 points)

You are given the following prototypes:

S1: {(-1, 1), (0, 1), (-1, 0)} , S2: {(1, 0), (1, 1)}, S3{(0, -1), (1, -1)}

(a) Is it possible to linearly separate the data without an augmented weight vector? Why?

(b) If the answer to (a) is yes, determine an unaugmented weight vector that linearly separates the data . Draw the decision boundaries in feature space. Show data points and class domains.

3. Statistical Classification (20 points)

(a) Determine the Bayes minimum error decision boundary if

$$p(x|S_i) = \frac{1}{\pi b}\left(\frac{1}{1+(\frac{x-a_i}{b})^2}\right) \quad i = 1,2$$

For P(S1) = P(S2).

(b) Sketch P(S1|x) for a1 = 3, a2=5, b=1.

(c) How does P(S1|x) behave as x approaches infinity.

4. Statistical Classification (30points)

Two one-dimensional distributions are uniform in [0, 2] for w_1 and [1, 4] for w_2 with equal priors.

a) Plot the ROC curve.
b) Find the Neyman-Pearson boundary with E2=0.25.
c) Find the minimax boundary.

5. Data Reduction Techniques (15 points)

(a) Given the following prototypes; S1: {(1,1), (0,2), (1,3)}, S2: {(1,0), (2,0), (3,0)}, find the Fisher's linear discriminant, and give the resulting prototype in 1-D space.

(b) In general, under what condition it is accurate enough to choose $\mu_2 - \mu_1$ as the direction of Fisher's discriminant for projection of data down to one dimension? Explain your reasons.

6. Extra Points (15 points)

Show that if $g_i(x), = 1, \ldots, M$, are the discriminant functions of an M-Class problem, we can construct from them M-1 new functions that are, in principle, sufficient for the classification (Derive new discriminant functions).

Good Luck