Homework 1

CSE 802: Pattern Recognition and Analysis
Instructor: Dr. Arun Ross
Due Date: February 8, 2016

Note: You are permitted to discuss the following questions with others in the class. However, you *must* write up your *own* solutions to these questions. Any indication to the contrary will be considered an act of academic dishonesty. Copying from *any source* constitutes academic dishonesty. A hard-copy of the homework must be submitted before lecture begins on the due date.

---

1. The iris (flower) dataset consists of 150 4-dimensional patterns (i.e., feature vectors) belonging to three classes (setosa=1, versicolor=2, and virginica=3). There are 50 patterns per class. The 4 features correspond to sepal length in cm (x_1), sepal width in cm (x_2), petal length in cm (x_3), and petal width in cm (x_4). Note that the class labels are indicated at the end of every pattern.

   (a) [5 points] For *each feature*, plot the histograms pertaining to the 3 classes. Your output should contain 4 graphs corresponding to the 4 features; each graph should contain 3 histograms corresponding to the 3 classes (choose a bin size of your choice for the histograms). Based on these plots, indicate (a) the *features* that are likely to be useful for distinguishing the 3 classes, and (b) the *classes* that are likely to overlap with each other to a great extent. Provide an *explanation* for your answer.

   (b) [5 points] Assume that each pattern can be represented by features x_1 and x_2. This means, each pattern can be viewed as a point in 2-dimensional space. Draw a scatter plot showing all 150 patterns (use a different label/marker to distinguish between classes). Draw another scatter plot based on features x_1 and x_4. Based on these scatter plots, *explain* which of the two feature *subsets* ((x_1, x_2) or (x_1, x_4)) is likely to be useful for separating the 3 classes.

   (c) [5 points] Assume that each pattern can be represented by features (x_1, x_2, x_4). Draw a 3-dimensional scatter plot showing all 150 patterns. Based on this scatter plot, *explain* which classes overlap with each other to a great extent?

2. [10 points] What type of learning scheme - supervised, unsupervised, or reinforcement - can be used to address each of the following problems. You must *justify* your answer.

   (a) Teaching a computer to play chess.

   (b) Given a set of sea-shells, determining if they can be grouped into multiple categories.

   (c) Determining the make and model of a car based on its side-view image.

   (d) Predicting whether it would rain or not in the next 24 hours based on current weather conditions such as precipitation, humidity, temperature, wind, pressure, etc.

   (e) Dividing a digital image into multiple regions such that each region has a distinct color or texture.

3. [15 points] Describe each of the following terms with an example: (a) generalization, (b) overfitting, (c) decision boundary.

4. [20 points] The paper *Can accelerometry be used to distinguish between flight types in soaring birds?* by Williams et al. discusses a pattern classification system that determines the flight type of a bird based on accelerometry data.

   (a) Briefly describe this system based on the pattern recognition terminology developed in class: (i) sensors used; (ii) features extracted; and (iii) classification scheme. How many features (i.e., d) and classes (i.e., c) are present?

   (b) How was training accomplished? How many data points were available in the training set? How were labels assigned to the data points?

   (c) What metrics were used to evaluate classifier performance?

   (d) In your opinion, did the proposed pattern recognition system perform well? Why or why not?

5. [10 points] Consider the following probability density function which is non-zero in the range $0 \leq x \leq 10$:

   $$p(x) = K.x^3(10-x).$$

   Here, K is a constant. Determine the value of the constant K.

6. [15 points] Consider a 1-dimensional classification problem involving two categories ω_1 and ω_2 such that $P(\omega_1) = 2/3$ and $P(\omega_2) = 1/3$. Assume that the classification process can result in one of three actions:

   α_1 - choose ω_1;
   α_2 - choose ω_2;
   α_3 - do not classify.

   Consider the following loss function, λ:

   $\lambda(\alpha_1|\omega_1) = \lambda(\alpha_2|\omega_2) = 0$;
   $\lambda(\alpha_2|\omega_1) = \lambda(\alpha_1|\omega_2) = 1$;
   $\lambda(\alpha_3|\omega_1) = \lambda(\alpha_3|\omega_2) = 1/4$.

   For a given feature value x, assume that $p(x|\omega_1) = \frac{2-x}{2}$ and $p(x|\omega_2) = 1/2$. Here, $0 \leq x \leq 2$.

   Based on the Bayes minimum risk rule, what action will be undertaken when encountering the value $x = 0.5$?

7. [15 points] Consider two categories, ω_1 and ω_2, of one-dimensional patterns whose class conditional densities are of the form:

   $$p(x|\omega_i) = 2\theta_i x e^{-\theta_i x^2},$$

   θ_i is a constant and denotes the parameter of the class-conditional density, and $i = 1, 2$ is the class label.

   Derive the Bayes decision boundary and the Bayes decision rule for determining the class of a pattern x^* if the two categories are equiprobable and a 0-1 loss function is adopted. You may assume $\theta_1 > \theta_2$ for definiteness.