CS 531 / ECE 517 Final

Name:_____

May 7, 2007

Be sure to clearly indicate both answers and work (equations etc). Do not work with others, violators will receive a zero on this final!!! Due Friday the 11^{th}.

1 Bayes Rule

1.1 Write the expression for Bayes Rule

1.2 What do each of the terms represent?

1.3 Plot the following likelyhoods on the same graph.

$$p(x|\omega_1) = N(-1,1), \tag{1}$$
$$p(x|\omega_2) = N(1,2), \tag{2}$$

where $N(\mu,\sigma)$ is the normal distribution.

1.3.1 Plot the posterior distribution for Eqns 1 and 2

Use the priors

$$P(\omega_1) = .2 \tag{3}$$
$$P(\omega_2) = .8 \tag{4}$$

1.3.2 Plot the Risk Functions

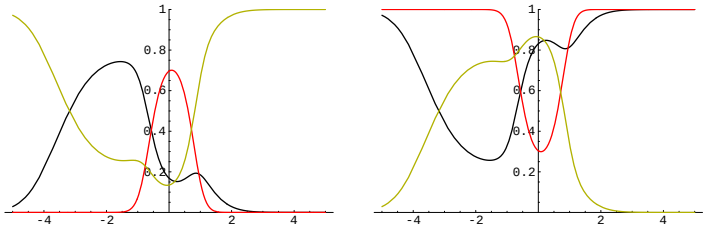
Use the previously defined likelyhoods and priors with the following risk functions

$$\lambda(\omega_1, \omega_1) = 0 \tag{5}$$
$$\lambda(\omega_1, \omega_2) = 100 \tag{6}$$
$$\lambda(\omega_2, \omega_2) = 0 \tag{7}$$
$$\lambda(\omega_2, \omega_1) = 20 \tag{8}$$

2 Decision Boundaries

2.1 Please identify the decision boundaries on each of the previous plots.

2.2 Please identify the decision boundaries on the following RISK plots.



3 Risk

3.1 Please describe the meaning of each of these terms:

1. $R(\alpha_i|\vec{x})$
2. $\lambda(\alpha_i, \omega_j)$
3. $\lambda(\alpha_i, \omega_j) = \begin{cases} 0 & i = j \\ 1 & i \neq j \end{cases}$

3.2 What is the general expression for (conditional) risk?

Given c classes (ω_j) and n actions α_i.

$R(\alpha_i|x) =$

3.3 Simplify this expression using $\lambda(\alpha_i|\omega_j)$ from term 3 in part 3.1 above.

3.4 Describe briefly what is the purpose of ”minimax” risk.

4 Discriminants

4.1 How is the set of discriminant functions $\{g_i(\vec{x})\}$ used to classify $\vec{x}$?

4.2 Write a discriminant using only posterior distributions.

4.3 Write a discriminant for Normal Distributions that does not involve an exponential.

4.4 Write a discriminant using risk.

4.5 What is the valid range of values for a discriminant function?

4.6 Write the equation for a linear discriminant. What does it mean geometrically?

4.7 What role does $\Phi(x)$ play in discriminants and classification?

4.8 How are linear regression and linear discriminant training related?

5 Classification

Given:

$$p(\vec{x}|\omega_i) = \beta_i \exp\left[-\frac{1}{2}(\vec{x}-\vec{\mu}_i)^T \Sigma_i^{-1}(\vec{x}-\vec{\mu}_i)\right] \tag{9}$$

$$\mu_1 = [5,2]^T, \quad \Sigma_1^{-1} = \begin{bmatrix} .5 & 0 \\ 0 & .5 \end{bmatrix}, \quad \beta_1 = \frac{1}{4\pi} = .0795775, \quad P(\omega_1) = 1/3 \tag{10}$$

$$\mu_2 = [2,6]^T, \quad \Sigma_2^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & .25 \end{bmatrix}, \quad \beta_2 = \frac{1}{4\pi} = .0795775, \quad P(\omega_2) = 1/3 \tag{11}$$

$$\mu_3 = [8,6]^T, \quad \Sigma_3^{-1} = \begin{bmatrix} .25 & 0 \\ 0 & 1 \end{bmatrix}, \quad \beta_3 = \frac{1}{4\pi} = .0795775, \quad P(\omega_3) = 1/3 \tag{12}$$

5.1 Classify the following points.

1. $\vec{x} = [7, 3.5]^T$
2. $\vec{x} = [4,6]^T$
3. $\vec{x} = [2,4]^T$

5.2 Repeat the classification with the following:

$$P(\omega_1) = \frac{1}{7} = .142857, \quad P(\omega_2) = \frac{1}{7} = .142857, \quad P(\omega_3) = \frac{5}{7} = .714286 \tag{13}$$

1. $\vec{x} = [7, 3.5]^T$
2. $\vec{x} = [4, 6]^T$
3. $\vec{x} = [2, 4]^T$

5.3 What is the difference between Parzen Window and K-Nearest Neighbor?

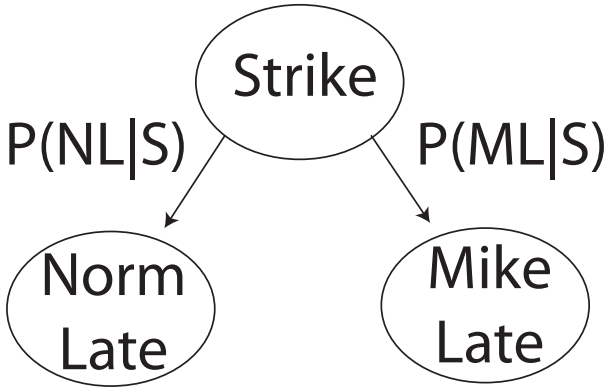
5.4 How do Parzen Window and K-Nearest differ in computational complexity?

5.5 Describe the relationship between Parzen Window and Support Vector Machines.

6 Decision Trees

6.1 Compute probability of being late.

Given the following decision network, What is the probablity of Norman being late, with the following probabilites:



1. Probability that Norman is late if there is a strike: .8
2. Probability that Norman is late if there is no strike: .1
3. Probability that Michael is late if there is a strike: .6
4. Probability that michael is late if there is no strike: .5
5. Probability that there is a strike: .2

6.2 What is the probability that Norman is Late?

6.3 What is the probability that there is a strike if Norman is Late?

6.4 What is the probability that Michael is Late because of a Strike, if Norman is Late?

7 Extra Credit

7.1 Name and discuss two advantages of SVMs over Parzen Window

7.2 Discuss the compuational complexity of SVMs

7.3 Discuss the advantages and disadvantages of Decision Trees for classification

7.3.1 What does a purity measure tell us?

7.3.2 If our tree branches on yes, no, and sometimes, what is the depth of a balanced tree?

7.4 Discuss a method for making linear discriminants provide probabilistic information.