The University of Connecticut **Spring 2021**
Dept. of ECE **KRP**

Problem Set # 2
(Due February 23, 2021)

1. Consider two one-dimensional Gaussian-distributed classes C_1 and C_2 that have a common variance equal to 1. Their mean values are as follows: $\mu_1 = 10$ $\mu_2 = -10$. Find the discriminant. What is the probability of error?
2. Problem 3.16 of Theoderidis, pp. 117.
3. Consider a binary classification problem with two-dimensional features. The two classes are modeled by Gaussian distributions with means and covariance matrices given by:

$$\underline{\mu}_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}; \underline{\mu}_2 = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

$$\Sigma_1 = \Sigma_2 = \begin{bmatrix} 1 & 0.25 \\ 0.25 & 1 \end{bmatrix}$$

   (a) Form and plot a dataset consisting of 500 data points from each class.
   (b) Draw the Bayesian decision rule on this graph.
   (c) Analytically evaluate the probability of error. Also, evaluate the empirical error rate from the data set in (a) using the decision rule in (b).
   (d) Let

$$\Lambda_1 = \begin{bmatrix} 0 & 1000 \\ 1 & 0 \end{bmatrix}; \Lambda_2 = \begin{bmatrix} 0 & 1 \\ 10 & 0 \end{bmatrix}$$

   be two loss matrices. Assign each one of the points of the data set in (a) to either class 1 and class 2 according to the Bayesian risk minimization decision rule and plot the points with different colors, depending on the class they are assigned to.
   (e) Based on (d), estimate the average risk for the above loss matrices.
   (f) Comment on the two decision rules in (b) and (d).
   (g) Suppose you have a reject option that costs you λ_r. Plot the average risk as a function of λ_r for the two loss matrices.
4. Exercises 7.4 & 7.5, Theoderidis, Page 345.
5. Exercise 3.27, Theoderidis, Page 119.
6. Exercise 7.20, Theoderidid, Page 347.
7. Exercise on slide 15 of Lecture 3. For $P(Z=1) = 2/3$; $p(x|Z=1) = N(2,0.5)$, $p(x|Z=2) = N(1.5,0.2)$, compute the decision rule in special case 1. Indicate the regions along the real axis for the two decisions.
8. (This can be done with you and two partners) In preparation for extensive testing of machine learning data, become familiarized with the data sets described in the class and University of California-Irvine (UCI) data sets (ftp://ftp.ics.uci.edu/pub/machine-learning-databases/). Experiment with the quadratic and linear discriminant rules with at least four data sets. Use MATLAB (or any other tool) where necessary. You can use estimated means and covariances in the decision rules.