ECE 365 **Fall 2020**

September 22, 2020

QUIZ 1: 8:00–8:40 PM (*60 points total*)

Note: This is a **closed-book** exam. Please write your answers on clean sheets of letter paper using a pen or pencil, and upload the scanned pages on Gradescope, indicating the page numbers corresponding to your answers when you upload on Gradescope.

1. [**Bayes Classifier (10 points)**]
   Consider the binary hypothesis testing problem:

$$H_{-1}: \ X \sim p(x|-1)$$
$$H_1: \ X \sim p(x|1)$$

   where

$$p(x|-1) = \begin{cases} 0.1 & x=0 \\ 0.2 & x=1 \\ 0.3 & x=2 \\ 0.4 & x=3 \\ 0 & \text{otherwise} \end{cases} \qquad p(x|1) = \begin{cases} 0.25 & x=0 \\ 0.25 & x=1 \\ 0.2 & x=2 \\ 0.3 & x=3 \\ 0 & \text{otherwise} \end{cases}$$

   with priors $\pi_{-1} = 0.5, \pi_1 = 0.5$.

   (a) Find the Bayes classifier (MAP decision rule).
   (b) Calculate the probability of error for the Bayes classifier.

2. [**Classifier performance (10 points)**]
   (a) Design a classifier f that takes input $\underline{x}$ to produce an estimate of the label y, to have a training error of 0 on a training set $\mathcal{T}$, for the *0-1 loss* function.
   (b) Explain in words what the prediction error of a classifier is. How is it estimated from a validation data set $\mathcal{V}$ for the *0-1 loss* function?

3. [**Linear Discriminant Functions(10 points)**]
   Consider a 3-ary linear classifier, with classes 1, 2, and 3, for which the three linear discriminant functions are:

$$g_{12}(\underline{x}) = x_1 + 2x_2 - 3$$
$$g_{13}(\underline{x}) = 2x_1 + x_2 - 1$$
$$g_{23}(\underline{x}) = x_1 - x_2 + 2$$

   Classify the input $\underline{x}$ with $x_1 = 1, x_2 = 2$.

4. [**Linear Classifiers (10 points)**]
   (a) Identify which among the following classifiers are linear: k-Nearest Neighbors, Logistic Regression, Naive Bayes, Linear Discriminant Analysis, SVM.
   (b) Given a linear classifier explain clearly how you might use it to separate data that are not linearly separable.

5. [**Linear Separability (10 points)**]
   Draw an example of a data set that is linearly separable, and one that is not linearly separable. You can use symbols such as "x" and "o" to denote the data under each class.

6. [**Naive Bayes and Logistic Regression (10 points)**]
   (a) What assumption does Naive Bayes make about the distribution of the components of the feature vector $\underline{x}$?
   (b) Explain why the logistic regression classifier is considered to be a *discriminative* model rather than a *generative* model.