

Home Work 2: CAP 6610 Spring '07

Due Date: Mar 1st 2007

Show all steps. Be as concise as possible.

Here is the first programming project that is meant to cement your understanding of the differences between identifying and imitating the supervisor. This time we focus on imitation. Save your work for use in a later assignment.

1. Code for a generative model in \mathbb{R}^3 , i.e., 3-dimensions, that has two classes each modeled by normal distributions. Class-1 is centered at $\langle 1, 1, 1 \rangle$ and Class-2 is centered at $\langle 2, 0, 5 \rangle$. Both normal distributions have the same covariance matrix

$$\begin{bmatrix} 62 & 8 & -26 \\ 8 & 29 & -5 \\ -26 & -5 & 35 \end{bmatrix}$$

Explain how you can generate data from either distribution if you had access to a random number generator for a 1-dimensional normal distribution (as in matlab). Finally write code to generate i.i.d samples from both distributions.

2. What is the optimal discriminant for the above generative model? (You will have to study chapter 2 of the text) What is the expected classification error of this discriminant?
3. Write an incremental perceptron learning algorithm where you train on 50 data points generated i.i.d (with equal probability for each class, i.e., $p = 0.5$), freeze the hyperplane and test on 50 data points generated i.i.d, and cycle thru this 100 times. Report your classification error on the test set in a graph. Comment on how the perceptron testing error compares to the expected classification error for the optimal discriminant.