Reading for problem set 1: DHS (Duda, Hart, & Stork) chapter 2

Written Problems
DHS problems 2.2, 2.6, 2.13, 2.22, 2.23, 2.30, 2.31, 2.39, 2.43.

Library Problem
Choose an application of pattern recognition that is interesting to you. Read at least one long peer-reviewed papers describing an approach that other researchers have taken to solve your problem or a related problem ("long" means 7-8 pages or more – e.g., journal articles, or long conference papers at conferences such as NIPS). Write a paragraph: tell me why the application is important or interesting, what makes the application difficult, and what was the approach taken in the paper you read. Cite the paper.

Computer Problems
DHS computer exercises 2.1, 2.2, and 2.7.

Find a source for labeled data in an application area that interests you (presumably a sub-problem of the area in which you’ve been reading papers). Choose a two-class subset of your data, with feature vectors that are at least three-dimensional. Repeat computer exercise 2.2 using these data. Design your dichotomizer under whichever of the following assumptions is least inappropriate: (a) the data are Gaussian distributed, or (b) the data are discrete. You may find that performance on real-world data is not as good as performance on the canned data in the textbook. In order to convince me that your code works, plot (a) histograms \( p(x_1 | \omega_1) \) and \( p(x_1 | \omega_2) \) of the first feature given the two classes, (b) the Gaussian PDFs or discrete PMFs that you used to approximate \( p(x_1 | \omega_1) \) and \( p(x_1 | \omega_2) \) for the purpose of designing a classifier, and (c) the discriminant function. On the last of these plots, draw the decision threshold.