Donald E. Brown
Bill Basener   DS 6014

Exercise 3.2
Posterior Probabilities with LDA and QDA

Instructions: You may discuss this assignment with other students in the class, but you must submit your own answers to the questions below. Include an honor pledge with your submission. Submit on-line and in pdf. This exercise is ungraded.

1. Download the required files:

    - Linear_Discriminant_Analysis.py,

    - Quadratic_Discriminant_Analysis.py, and

    - iris_data.csv

    into a single folder, and watch the video explaining the Python code.

2. The code for the LDA and QDA classes are the same except for two lines (one line in the __init__ method and one in the compute_likelihoods method). What are these two lines, and what is the functional/mathematical differences in what they do.

3. In each of the Python files, complete the code for the method compute_probabilities, and use your resulting code to answer the following two questions. (Submit your code with your solution to this exercise.)

4. Your friend Carl calls you and says he has measured two iris flowers and wants your help determining the species. The first flower has dimensions [SepalLength, SepalWidth, PetalLength, PetalWidth] = [5.5, 2.4, 3.8, 1.1] and the second has dimensions [5.5, 3.1, 5, 1.5]. Assuming noninformative priors, what are the most probable species according to LDA and QDA, and what are the probabilities for each species.

5. You realize Carl is working in the country of Bagend, and in this country 70% of the irises are Iris-virginica, 20% are Iris-versicolor, and 10% are Iris-setosa. Use this information to create informative priors, and use your Python code to provide an updated answer to the previous question.