MTCS-205 Assignment 2: Chapter 5

Welcome to the second assignment in the MTCS-205 lab!

Prerequisites: chapter 5.

Skills to be tested: Python coding ability

Knowledge of prerequisites

Processing data using libraries

Assignments:

1. Generate data from two multivariate normal distributions for form datasets. Using the training and test versions of these, tabulate the training and test accuracies under the conditions enumerated in section 5.5 of the course book (Etham Alpaydin).

Details about the generation of the data:

The following are the means and the variances of the distributions:

a. μ1 = [0.1,0.8], Σ1 = I

b. μ2 = [0.5,0.2], Σ1 = [1, 1.5]

[0, -1 ]

Generate, once for a condition, a ‘training’ set of size 50000 and a ‘test’ set of size 10000 using the following conditions:

a. Each distribution is equally likely (This is not the same as equal number of samples)

b. Use the following priors: P(C\_1) = 0.2, P(C\_2) = 0.8

2. Perform classification of flower species using the Iris dataset using the Sepal length, Sepal width, Petal length and the Petal width. Assume a gaussian model for the classes and using the conditions on the covariance and the priors. Tabulate the train and the test accuracies.