

Assignment 1

Please refer to the code for implementation of `mlParams()`. Figure 1 shows the result.

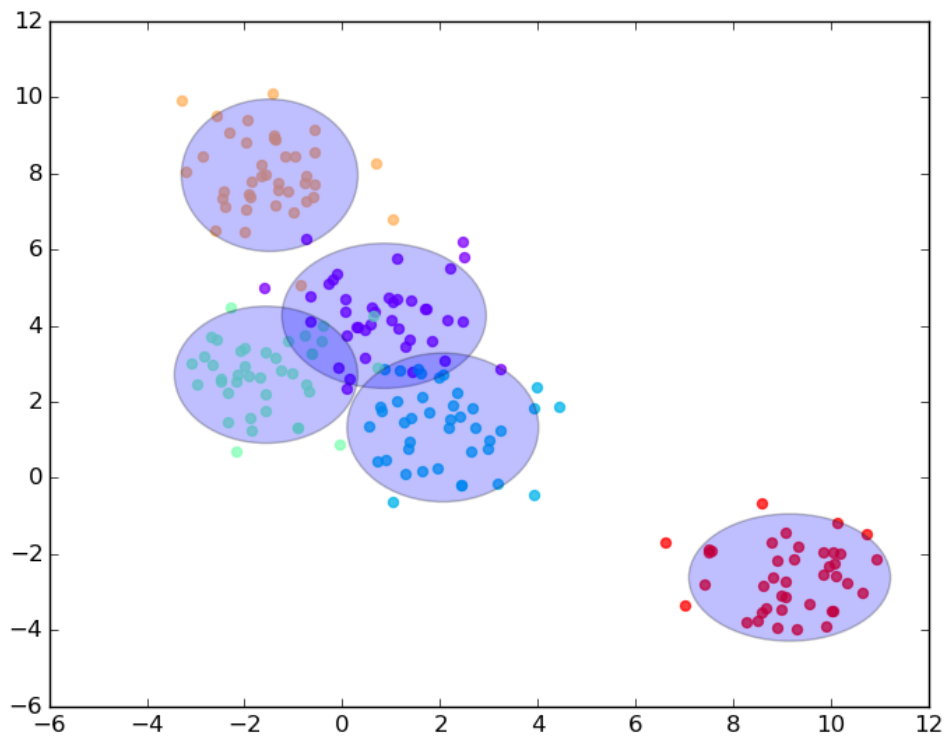


Figure 1: 95 % confidence intervals for data generated by `genBlobs(centers = 5)`

Assignment 2

Refer to code for implementations of `computePrior()` and `classifyBayes()`.

Assignment 3

The result of running `testClassifier(BayesClassifier(), dataset='iris', split=0.7)` and `testClassifier(BayesClassifier(), dataset='vowel', split=0.7)` is given:

```

Trial: 0 Accuracy 84.4
Trial: 10 Accuracy 95.6
Trial: 20 Accuracy 93.3
Trial: 30 Accuracy 86.7
Trial: 40 Accuracy 88.9
Trial: 50 Accuracy 91.1
Trial: 60 Accuracy 86.7
Trial: 70 Accuracy 91.1
Trial: 80 Accuracy 86.7
Trial: 90 Accuracy 91.1
Final mean classification accuracy 89 with standard deviation 4.16
Trial: 0 Accuracy 61
Trial: 10 Accuracy 66.2
Trial: 20 Accuracy 74
Trial: 30 Accuracy 66.9
Trial: 40 Accuracy 59.7
Trial: 50 Accuracy 64.3
Trial: 60 Accuracy 66.9
Trial: 70 Accuracy 63.6
Trial: 80 Accuracy 62.3
Trial: 90 Accuracy 70.8
Final mean classification accuracy 64.7 with standard deviation 4.03

```

The decision boundary for the 2D-iris data is depicted in Figure 2.

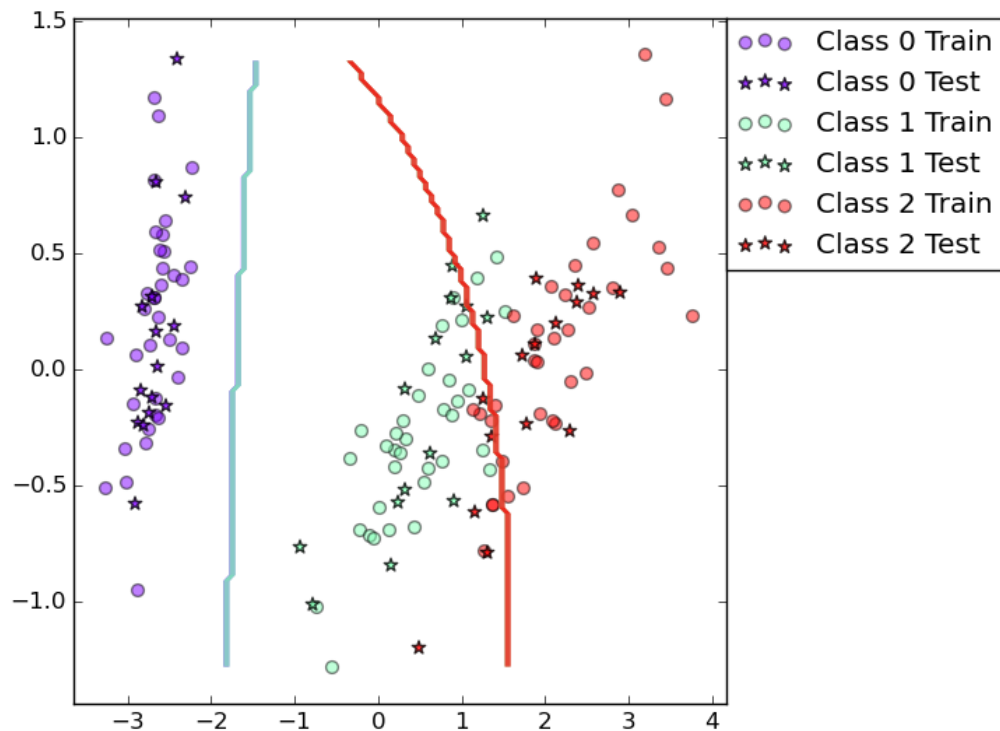


Figure 2: Boundary for the iris data.

Questions:

- 1) When can feature independence assumption be reasonable?
- 2) How does the decision boundary look for the iris dataset? How could one improve the results for this scenario by changing the classifier, or by manipulating the data?

Assignment 4

Refer to code for the augmented functions `mlParams()`

Assignment 4