Lab 3

Angeline Chen

The purpose of the lab is to use LDA and QDA to classify iris plants.

Features Selected	LDA Train Error Rate	QDA Train Error Rate	LDA Test Error Rate	QDA Test Error Rate
sepal length, sepal width, petal length, petal width	2.52%	1.68%	0%	0%
sepal width, petal length, petal width	1.68%	2.52%	0%	0%
sepal length, petal length, petal width	2.52%	1.68%	0%	0%
sepal length, sepal width, petal width	5.88%	4.2%	0%	0%
sepal length, sepal width, petal length	4.2%	5.88%	3.33%	3.33%

The train error rate is usually greater than the test error rate. Because the training data is more than the testing data, the model fits the testing data better than the training data. The LDA and QDA error rates are approximately the same. Because LDA learns linear boundaries, and QDA learns quadratic boundaries, the data has a fixed covariance, and the decision boundary is linear.

Sepal length and sepal width are not important in classifying iris type because the test error rate stays the same when the model is not run on sepal length, sepal width, and petal length. Petal width is important in classifying iris type because the test error rate increases when the model is not run on petal width. Petal length is also important in classifying iris type because the train error rate increases when the model is not run on petal length. Petal width is more important than petal length in classifying iris type because differences in the test error rate are more important than differences in the train error rate.