Elements of Machine Learning

Assigment 2 - Problem 6

Sangeet Sagar(7009050), Philipp Schuhmacher(7010127)
{sasa00001,phsc00003}@stud.uni-saarland.de

December 1, 2021

Problem 6 (P, 20 Points)

4

(5P) Repeat steps (2) and (3) using QDA and report your findings. Would you prefer LDA or QDA in this example? Why?
Let's see the results before analysing it.

1. Error statistics when model trained using all features (i.e. all phonemes).

```
## LDA
Train error:  0.05598802
Test error:  0.08041061

## QDA
Train error:  0
Test error:  0.1582549
```

2. Error statistics when model trained on only two features (i.e. only two phonemes aa, ao).

```
## LDA
Train error:  0.1064163
Test error:  0.214123

## QDA
Train error:  0
Test error:  0.3394077
```

In the given setting and dataset, LDA should be preferred over QDA since the test error is lower when the former is used. This is true when the model is trained on all features, i.e. all phonemes, as well as when the model is trained on only two features.

5

Generate four confusion matrices: for the LDA and QDA model for aa and ao on test and training data. Which differences can you observe between the models?

```
############ Confusion Matrix: LDA
       Train_predicted
Actual  aa  ao
    aa 439  80
    ao  56 703
      Test_predicted
Actual  aa  ao
    aa 121  55
    ao  39 224
```

```
############## Confusion Matrix: QDA
       Train_predicted
Actual   aa   ao
     aa 519    0
     ao   0  759
       Test_predicted
Actual   aa   ao
     aa   29  147
     ao    2  261
```

Test accuracy for LDA = (224+121)/439 = 78.5%
Test accuracy for QDA = (29+261)/439 = 66%
It can be pointed out easily that QDA suffers from overfitting, i.e. train error is exactly 0 while the test error is quite significant to ignore.

Another interesting observation to note is that QDA misclassified `aa` as `ao` only 2 times and `ao` as `aa` 147 times, which is quite significant. While LDA performed better here i.e. it correctly identified `aa` 121 times. Moreover, QDA correctly predicted `aa` only 29 times. This suggests the inefficiency of QDA over LDA.

6

Compare your estimates obtained from cross-validation to the error obtained from the test set and argue about your findings. Which of the methods is (theoretically) fastest?

	Runtime (s)	Estimated test error	Actual test error
LOOCV	0.171	0.657	2.265
5−fold CV	0.020	0.661	2.012
10−fold CV	0.034	0.656	2.494

5-fold CV is the fastest because it only performs 5 folds of the training data, while LOOCV is the slowest because it considers each observation as a validation set and put rest into training. This is repeated until all observations are at most validated once.
It is interesting to see that estimates of test error in all three cases are almost similar. However, their actual test error varies pretty differently. 5-fold CV has the least test error, and 10-fold CV has the highest test error. It is known that LOOCV has the least bias because it generates similar models for each split as the training splits are highly intersecting but suffers from high variance. 10-fold CV, on the other hand, has more bias which is reflected from the fact that it has the highest test error.