EE559

Mathematical Pattern Recognition—

Homework 1

Jionghao Fang

Due: 02/05/2021

a. For each of the two synthetic datasets

1) Plot the training-set data plots, the resulting class means, decision boundaries, and decision regions.

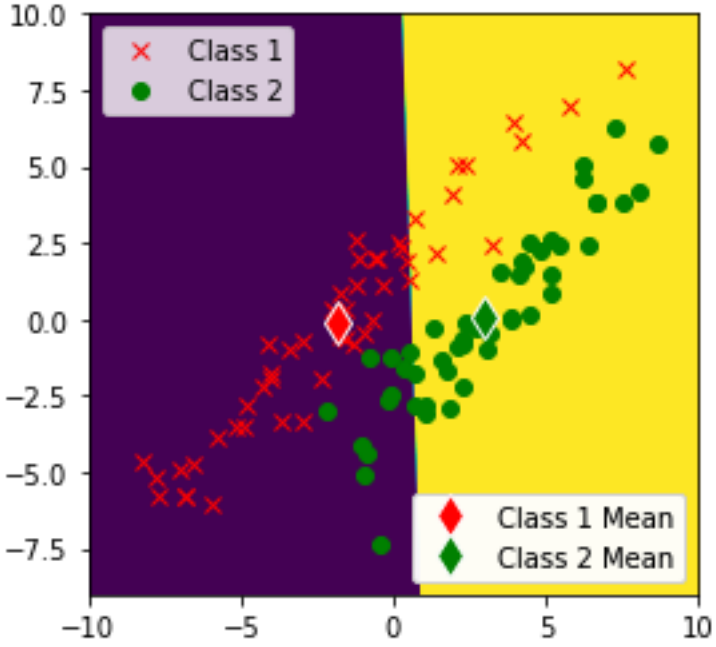

Fig.1 Plot for synthetic 1.

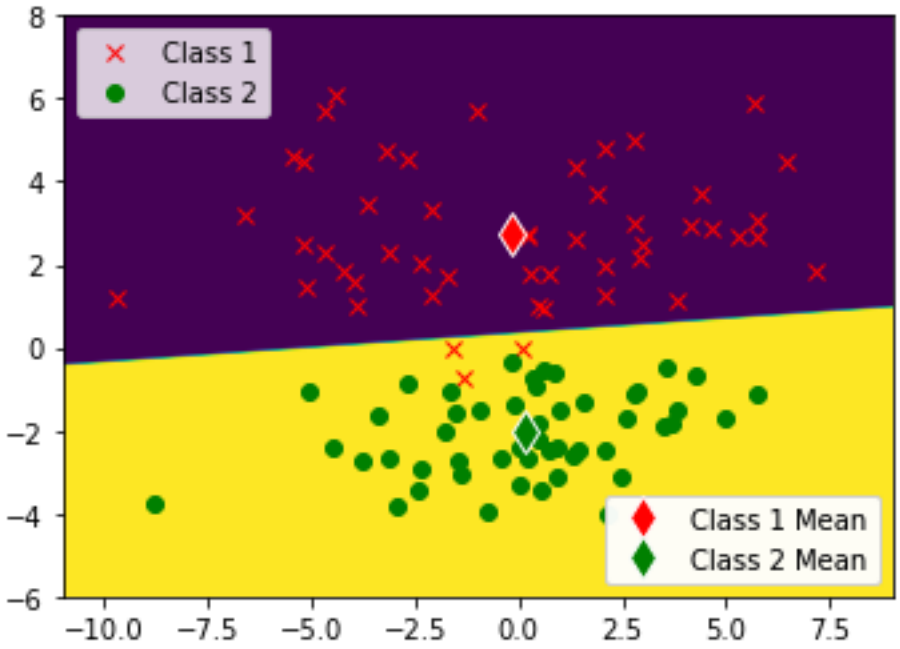

Fig.2 plot for synthetic 2

2) Give the classification error rate on the training set and test set.

```
The error rate for training set in syn1=
0.21
The error rate for test set in syn1=
0.24
```

Fig.3 results for synthetic 1

```
The error rate for training set in syn2=
0.03
The error rate for test set in syn2=
0.04
```

Fig.4 results for synthetic 2

b. Is there much difference in error rate between two synthetic datasets?

Yes, there is much difference.

The plot for synthetic1 shows that there are plenty of points (nearly one third) that are not classified as they are supposed to be.

But from the plot for synthetic2, it indicates that hardly had the points that are misclassified.

Maybe this classification method is quite suit for the dataset in synthetic2, and which is not the suitable method for synthetic1.

c. For the wine dataset, pick the first two features and repeat the procedure of part a.

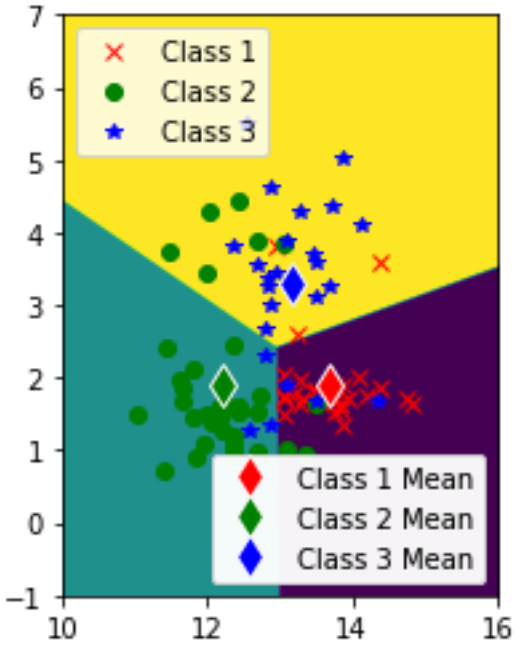

Fig.5 plot for wine

```
The error rate for training set in wine=
0.20224719101123595
The error rate for test set in wine=
0.2247191011235955
```

Fig.6 results for wine

d. Find the 2 features among 13 that achieve the minimum classification error on the training set.

Two FOR loop are used for finding the minimum classification error among 13 features (C^{13}_2 combinations).

```
The minimum training error rate is
0.07865168539325842 ,which appears at
feature 1 and feature 12, meanwhile its
testing error rate is 0.12359550561797752
The minimum testing error rate is
0.11235955056179775 ,which appears at
feature 1 and feature 7, meanwhile its
training error rate is 0.0898876404494382
```

Fig.7 results for minimum error rate

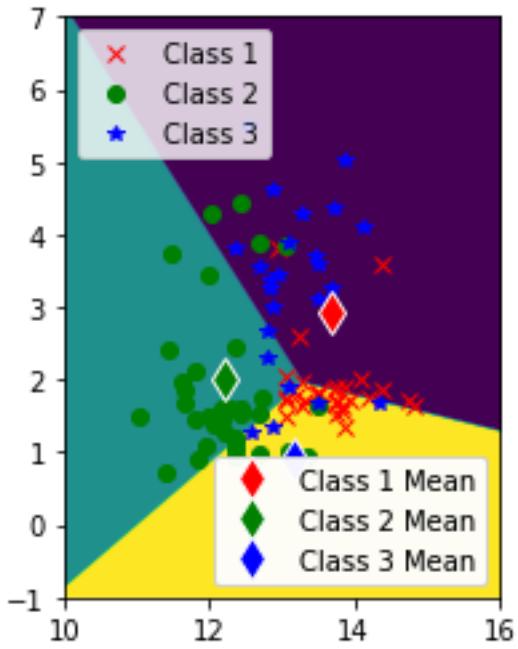

Fig.8 plot for minimum error rate in training set

The two features that could achieve the minimum classification error on the training set is feature 1 and feature 12.

e. Are there much differences in training-set error rate and testing-set error for different pairs of features?

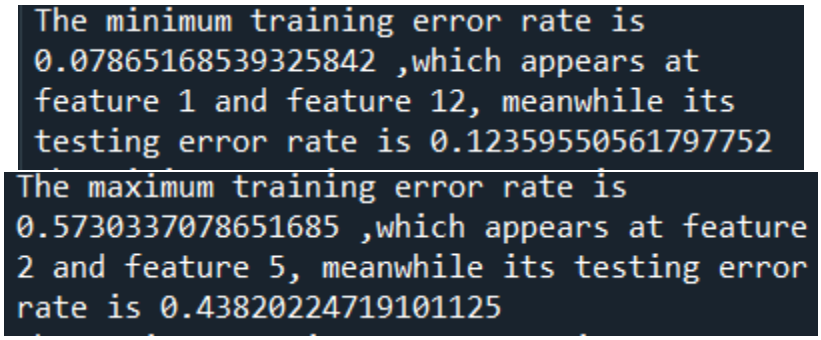

Fig.9 the comparison between the minimum and maximum error rate in training-set

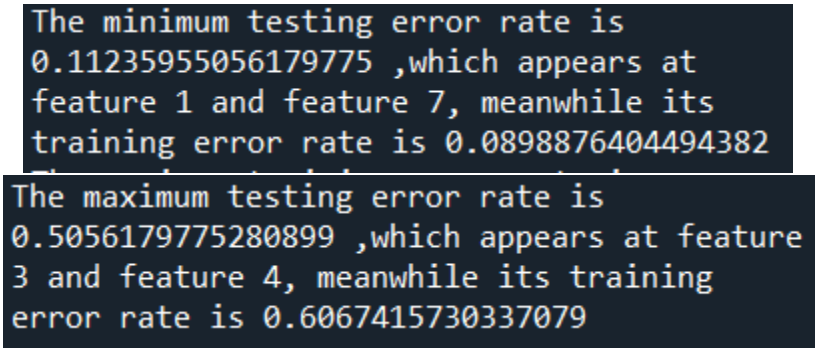

Fig.10 the comparison between the minimum and maximum error rate in testing-set

There are many differences for different feature pairs that can be read from the comparisons above. The training error rate is fluctuated from 0.124 to 0.438 and the testing error rate is fluctuated from 0.090 to 0.607.