(a)

1. Synthetic 1 :
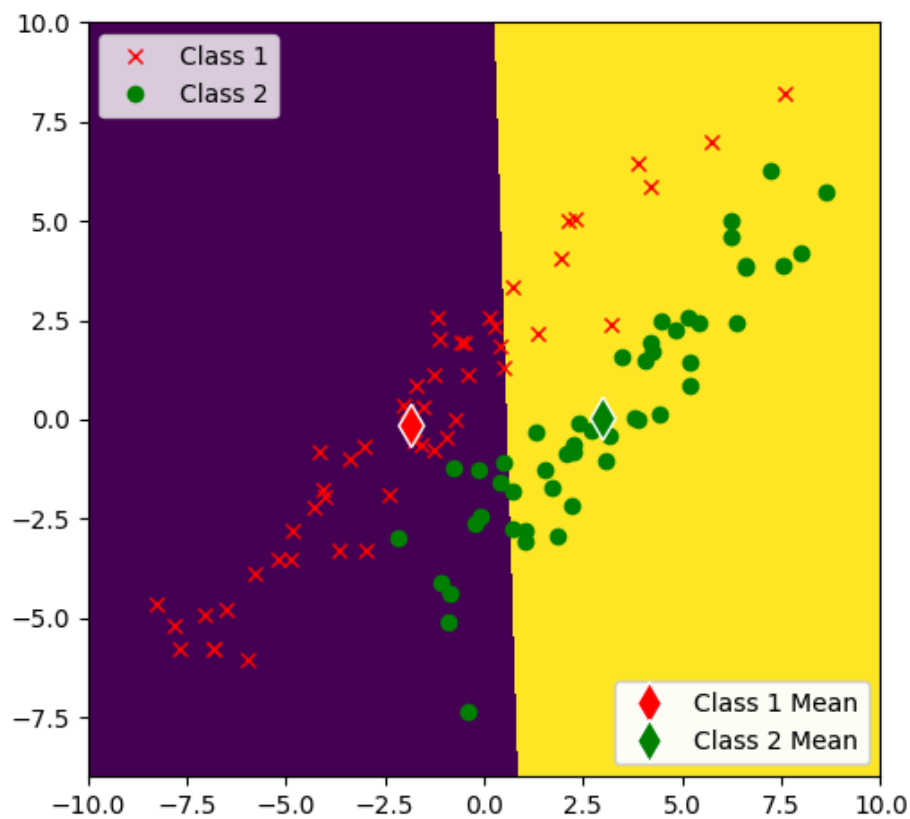
(1)Training data:

In class1, sample mean is (-1.8731151999999998, -0.11664180000000002)

In class2, sample mean is (2.98095798, 0.03548128800000003)

The error rate is 21%.
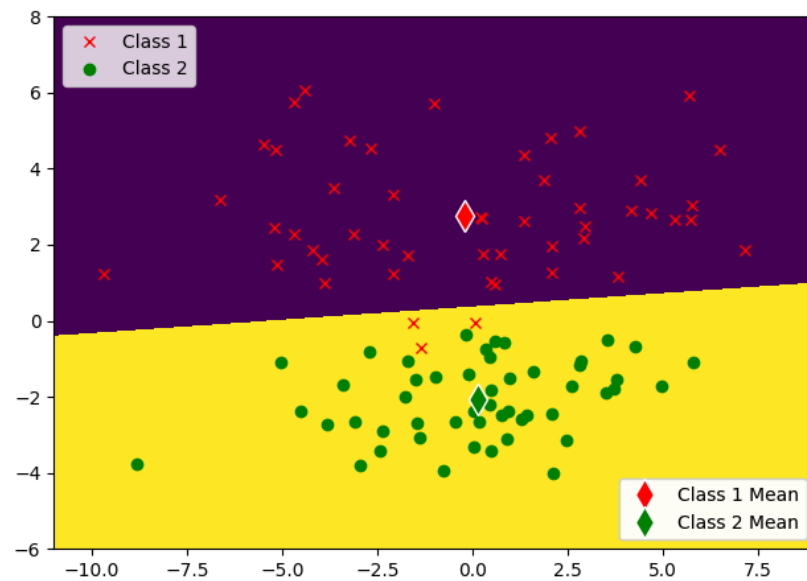
(2)In testing data, the error rate is 24%.

2.Synthetic 2

(1)Training data:

In class1, sample mean is (-0.20326850000000005, 2.7552259199999996)

In class2, sample mean is (0.13275594, -2.0526066)

The error rate is 3%.

(2) In testing data, the error rate is 4%.

(b) Yes, there is much difference in the error rates of the two synthetic data sets. As we can see from the previous two plots, the classification line is almost parallel to the distribution of class 1 and class 2 in plot 2. However, the decision boundary does not follow the direction of the distributions of class 1 and class 2 in plot 1. Therefore, we can see that the error rate of the first synthetic dataset is higher than the second one.

(C)In the wine dataset, the error rate of training data is 20.224719101123595%. And the error rate of the testing data is 22.47191011235955%.
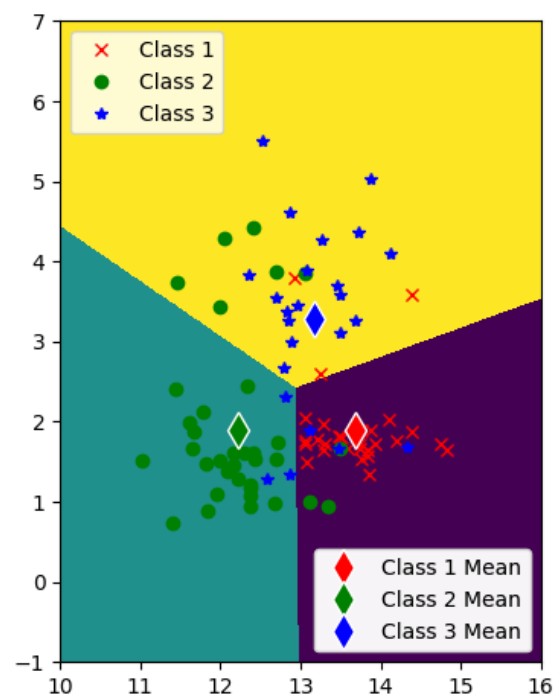
(d) Because there are 13 features in the training dataset, there are $\binom{13}{2} = 78$ pairs to verify which one is the best. First, we need to choose 2 different feature and compute the sample mean of the two features in each class. Then, verifying the distance between the point in a specific class and its mean is the shortest one than to the mean of other classes. If it is not the shortest distance, we would recognize the point as misclassified. After that, we would count the number of misclassified points and divide it by the total number of points to get the error rate as you can see in the following table.

feature 1	feature 2	training err%	testing err%
1	2	0.20224719101123595	0.2247191011235955
1	3	0.3146067415730337	0.2808988764044944
1	4	0.449438202247191	0.4044943820224719
1	5	0.5617977528089888	0.449438202247191
1	6	0.14606741573033707	0.15730337078651685
1	7	0.0898876404494382	0.11235955056179775
1	8	0.33707865168539325	0.2808988764044944
1	9	0.16853932584269662	0.24719101123595505
1	10	0.25842696629213485	0.2247191011235955
1	11	0.25842696629213485	0.2696629213483146
1	12	0.07865168539325842	0.12359550561797752
1	13	0.24719101123595505	0.30337078651685395
2	3	0.39325842696629215	0.38202247191011235
2	4	0.39325842696629215	0.42696629213483145
2	5	0.5730337078651685	0.43820224719101125
2	6	0.29213483146067415	0.29213483146067415
2	7	0.20224719101123595	0.23595505617977527
2	8	0.3258426966292135	0.39325842696629215
2	9	0.4044943820224719	0.3707865168539326
2	10	0.24719101123595505	0.2247191011235955
2	11	0.38202247191011235	0.449438202247191
2	12	0.42696629213483145	0.3707865168539326
2	13	0.24719101123595505	0.30337078651685395
3	4	0.47191011235955055	0.5056179775280899
3	5	0.5730337078651685	0.4606741573033708
3	6	0.3258426966292135	0.2808988764044944
3	7	0.14606741573033707	0.2247191011235955
3	8	0.516853932584269	0.3258426966292135
3	9	0.38202247191011235	0.4044943820224719
3	10	0.30337078651685395	0.23595505617977527
3	11	0.30337078651685395	0.2696629213483146
3	12	0.29213483146067415	0.2808988764044944
3	13	0.24719101123595505	0.30337078651685395
4	5	0.42696629213483145	0.38202247191011235
4	6	0.47191011235955055	0.47191011235955055
4	7	0.42696629213483145	0.4157303370786517
4	8	0.47191011235955055	0.5056179775280899
4	9	0.449438202247191	0.48314606741573035
4	10	0.21348314606741572	0.29213483146067415
4	11	0.47191011235955055	0.5056179775280899
4	12	0.43820224719101125	0.449438202247191
4	13	0.24719101123595505	0.30337078651685395

5	6	0.5617977528089888	0.449438202247191
5	7	0.5617977528089888	0.4157303370786517
5	8	0.5730337078651685	0.4606741573033708
5	9	0.5617977528089888	0.449438202247191
5	10	0.4943820224719101	0.43820224719101125
5	11	0.5730337078651685	0.4606741573033708
5	12	0.5617977528089888	0.43820224719101125
5	13	0.24719101123595505	0.30337078651685395
6	7	0.2247191011235955	0.24719101123595505
6	8	0.34831460674157305	0.3595505617977528
6	9	0.4044943820224719	0.3258426966292135
6	10	0.2808988764044944	0.2247191011235955
6	11	0.3258426966292135	0.2808988764044944
6	12	0.24719101123595505	0.25842696629213485
6	13	0.24719101123595505	0.30337078651685395
7	8	0.16853932584269662	0.24719101123595505
7	9	0.16853932584269662	0.2696629213483146
7	10	0.21348314606741572	0.15730337078651685
7	11	0.15730337078651685	0.24719101123595505
7	12	0.1348314606741573	0.1797752808988764
7	13	0.24719101123595505	0.30337078651685395
8	9	0.42696629213483145	0.47191011235955055
8	10	0.30337078651685395	0.23595505617977527
8	11	0.33707865168539325	0.34831460674157305
8	12	0.4044943820224719	0.3258426966292135
8	13	0.24719101123595505	0.30337078651685395
9	10	0.30337078651685395	0.24719101123595505
9	11	0.3707865168539326	0.39325842696629215
9	12	0.34831460674157305	0.3146067415730337
9	13	0.24719101123595505	0.30337078651685395
10	11	0.30337078651685395	0.23595505617977527
10	12	0.2696629213483146	0.2247191011235955
10	13	0.24719101123595505	0.30337078651685395
11	12	0.4044943820224719	0.3258426966292135
11	13	0.24719101123595505	0.30337078651685395
12	13	0.24719101123595505	0.30337078651685395

In the end, we compared the error rate of each pair and get the best classification which causes the lowest error rate.

The minimum error rate in training data is 0.07865168539325842

feature 1	feature 2	training err%	testing err%
1	12	0.07865168539325842	0.12359550561797752

After computation, we found that the pair of feature x_1 and x_{12} can get the minimum error rate. There are 7 misclassified points and the error rate of training dataset is 7.865%. Besides, the error rate in testing dataset is 12.36%.
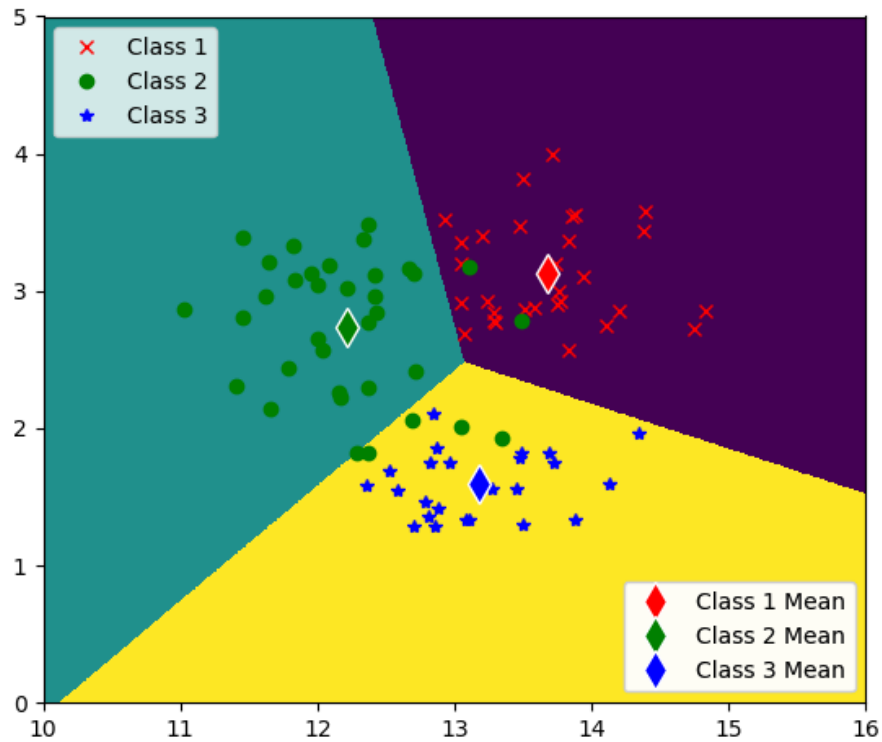
(e)

(1)For the training dataset, the deviation of the error rate for different pairs of features is 0.1284 and we can also see from the previous table in question (d) to know that the maximum value of the error rate is 0.573 and the minimum value is 0.0787. Therefore, it is easy to see that there is much difference between the error rate of different pairs of features.

(2)For the testing dataset, the deviation of the error rate for different pairs of features is 0.0957 and we can also see from the previous table in question (d) to know that the maximum value of the error rate is 0.5056 and the minimum value is 0.1124. Therefore, it is obvious that there is much difference between the error rate of different pairs of features.