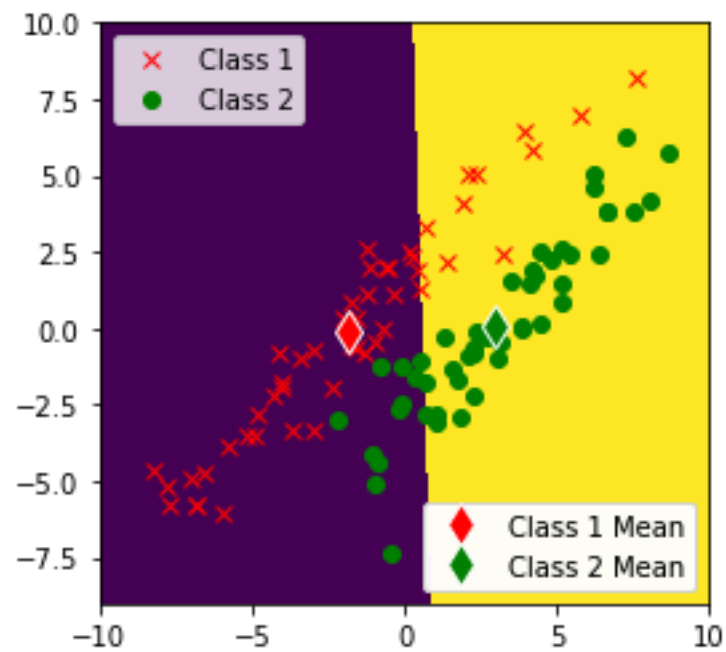Name: Zheng Wen

USC ID: 7112807212

(a)
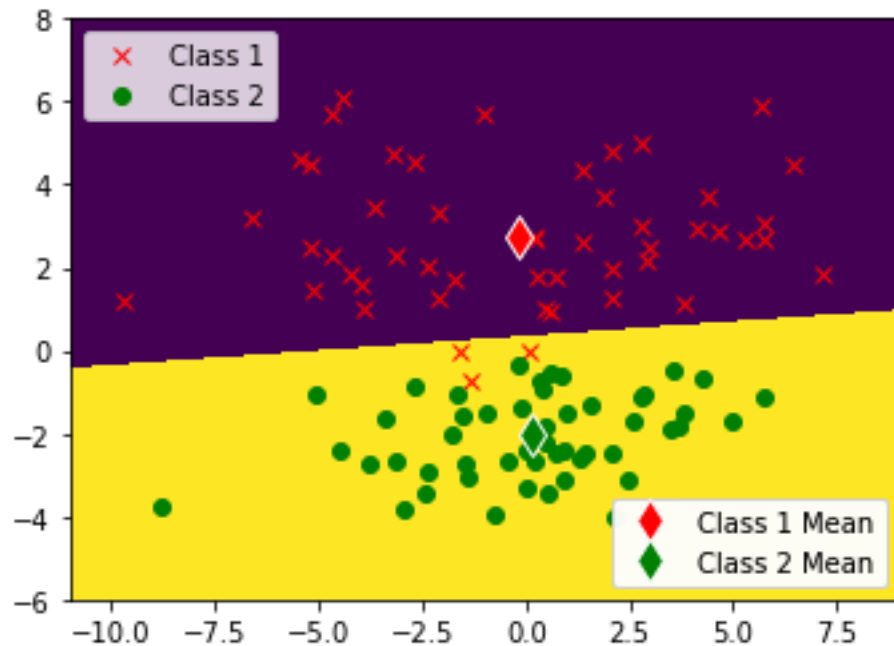
For synthetic1:

Error Rate of synthetic1 training set = 21.0%

Error Rate of synthetic1 test set = 24.0%

For synthetic2:

Error Rate of synthetic2 training set = 3.0%

Error Rate of synthetic2 test set = 4.0%

(b)

The error rate of sythentic2 is much less than that of sythentic1, because the decision boundary of sythentic2 is almost parallel with the true class separation. However, the decision boundary of sythentic1 is not parallel with the true class separation. The decision boundary of sythentic1 should be rotated a certain angle to fit the dataset.
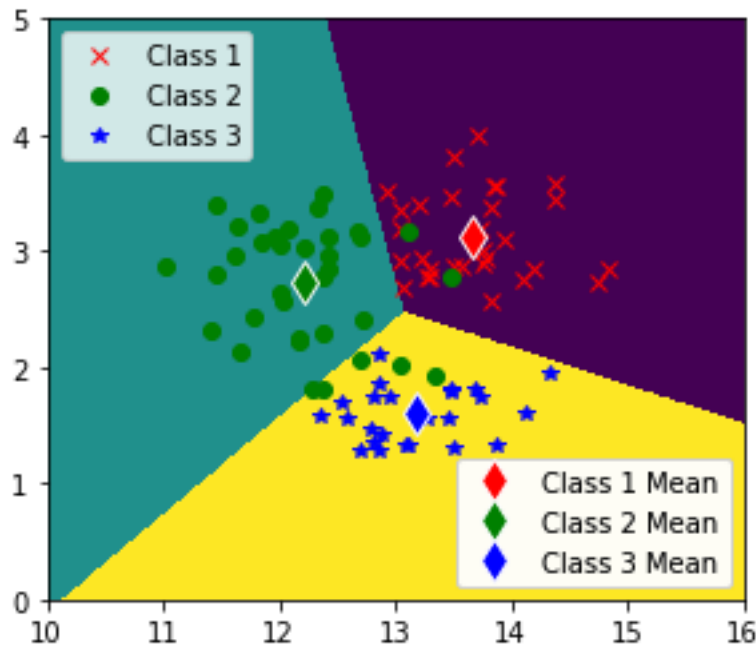
(c)
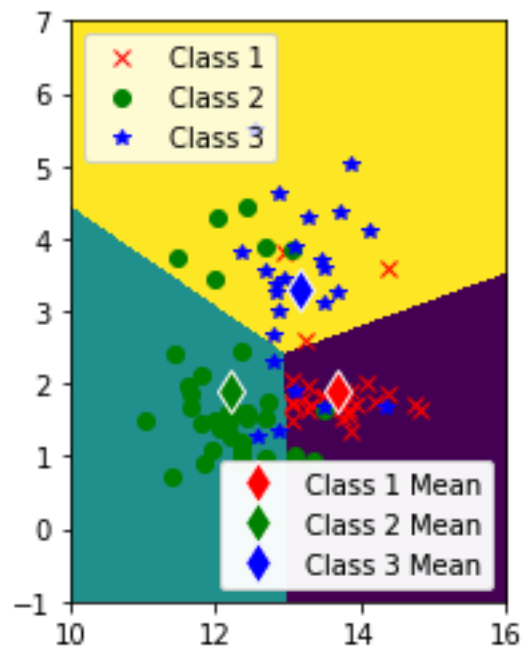
Error Rate of wine training set = 20.224719101123593%

Error Rate of wine test set = 22.47191011235955%

(d)

The feature chosen are [0 11]

Minimum error rate of wine training set = 7.865168539325842%

Minimum error rate of wine test set = 12.359550561797752%

To get the best features, because there are only 13 features in total, choose 2 features between them without repetitions, so there are only $\binom{13}{2}$ = 78 ways to combine different features. Just enumerate all of them and construct 78 different training sets and test sets with the corresponding features, calculate the error rate on test set for each combination, record the combination rendering the lowest error rate on test set, the result is obtained.

(e)

For test set:

      The standard deviation is 9.571906553548871%

      The maximum error rate is 50.56179775280899%

      The minimum error rate is 11.235955056179774%

      The average error rate is 32.72832036876979%

For training set:

      The standard deviation is 12.836178879103917%

      The maximum error rate is 57.30337078651685%

      The minimum error rate is 7.865168539325842%

      The average error rate is 33.57821953327571%

The data above suggest there is much difference in training-set error rate and test-set error rate for different pairs of features.