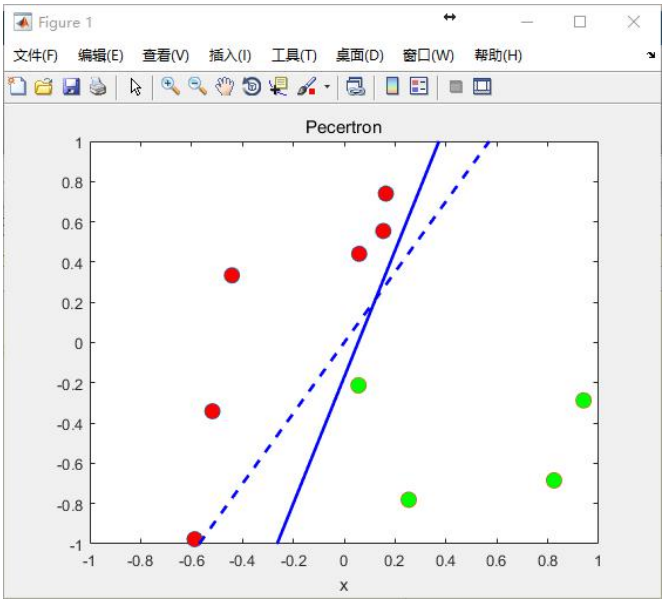
Machine Learning: Assignment #2

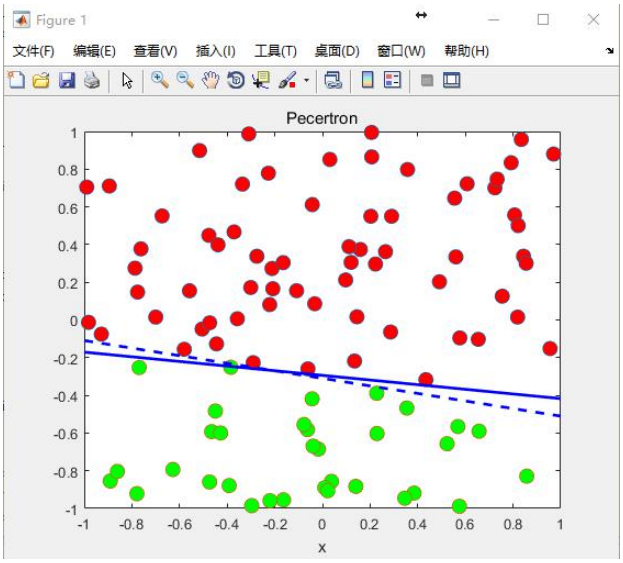
(a)

(i)

When the size of training set is 10 and the size of testing set is 100, the training error is zero, and the testing error is 10.78%.



When the size of training set is 100 and the size of testing set is 1000, the training error is zero, and the testing error is 1.4217%.



(ii)

The average number of iterations is 5.5240 when the size of training set is 10.

The average number of iterations is 51.5230 when the size of training set is 100.

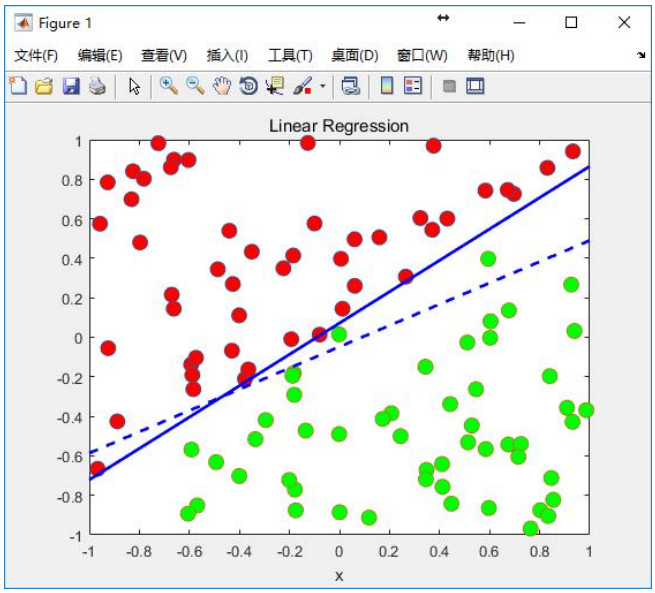
(iii)

If the training data is not linearly separable, the perceptron algorithm run without stopping. The perceptron algorithm cannot converge.

(b)

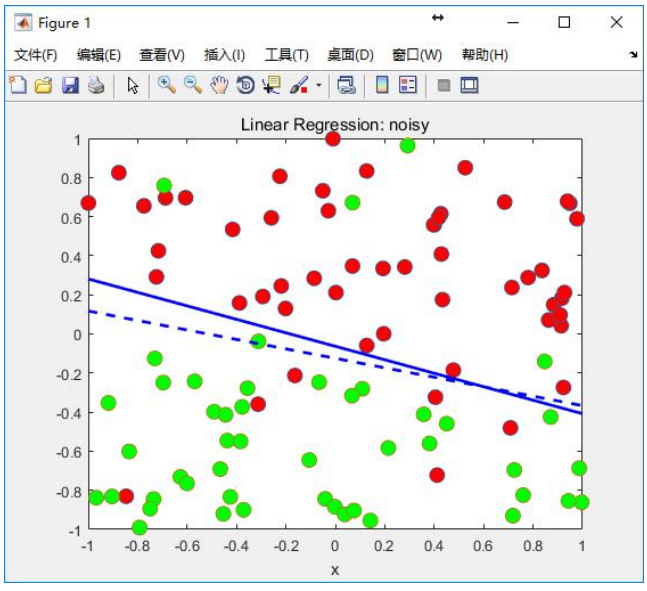
(i)

When the size of training set is 100, and the testing set is 1000, the training error rate is 3.977%, and the testing error rate is 4.89%.



(ii)

When the size of training set is 100, and the testing set is 1000, the training error rate is 13.885%, and the testing error rate is 15.016%.



(iii)

My w is [-0.0241, 0.2356, -0.1151], and the training error rate is 49.00%, the testing error rate is 54.96%.

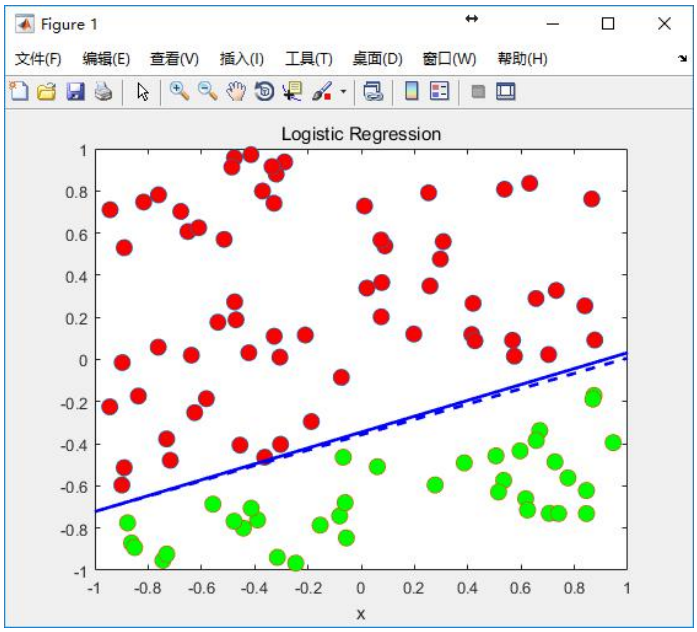
(iv)

My w is [-0.0687, -0.0269, 0.2246, 0.1316, -1.8483, 2.0042], and the training error rate is 5.00%, the testing error rate is 6.60%.

(c)

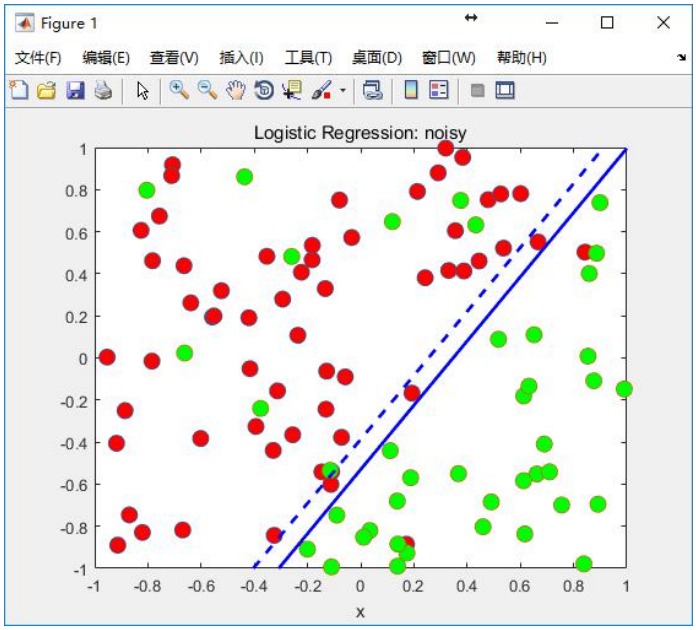
(i)

When the size of training set is 100, and the testing set is 1000, the training error rate is 1.140%, and the testing error rate is 1.870%.



(ii)

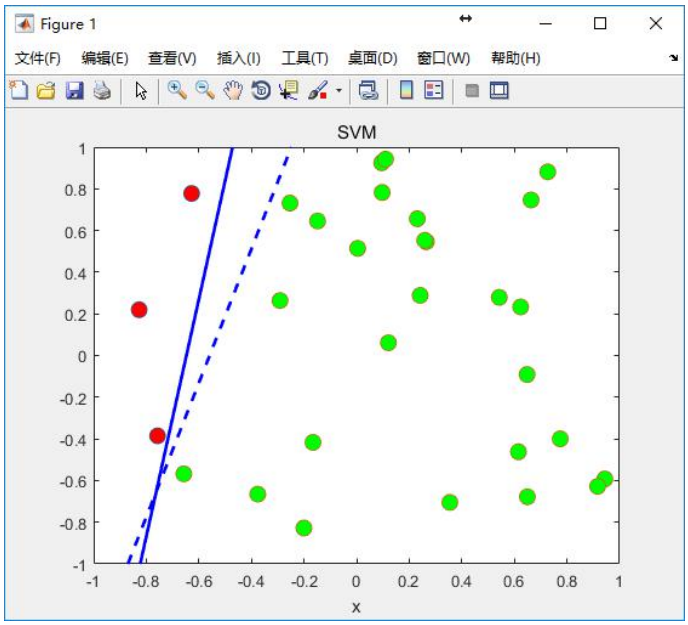
When the size of training set is 100, and the testing set is 1000, the training error rate is 13.6300%, and the testing error rate is 14.4210%.



(d)

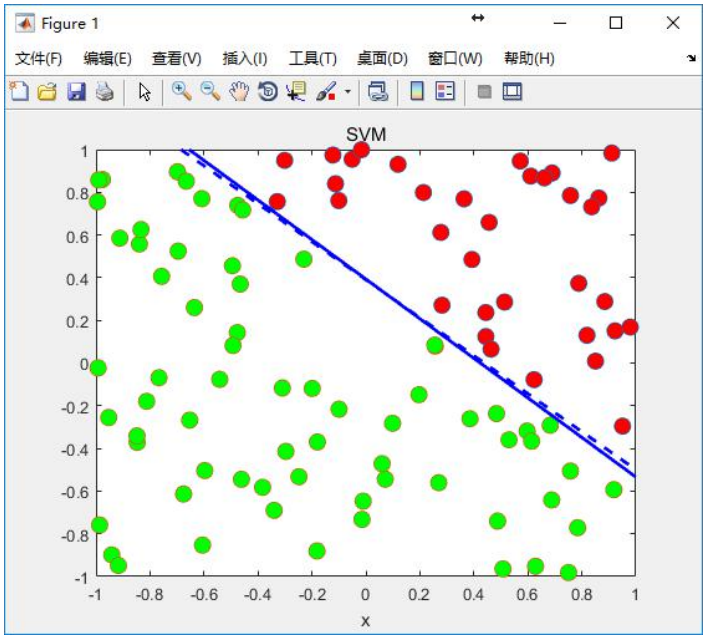
(i)

When the size of training set is 30, and the testing set is 300, the training error rate is 0.0000%, and the testing error rate is 3.4503%.



(ii)

When the size of training set is 100, and the testing set is 1000, the training error rate is 0.0000%, and the testing error rate is 1.0446%.



(iii)

When the size of training set is 100,, the average number of support vectors in my trained SVM models is 2.998.

2.

(a)

(i)

Lambda = 100 is chosen by LOOCV.

(ii)

The sigma(w^2) is 0.89286 without regularization.

The sigma(w^2) is 0.11351 with regularization when the lambda = 100

(iii)

The training error rate is 0% without regularization.

The training error rate is 0% without regularization.

The testing error rate is 9.04% with regularization when the lambda = 100.

The testing error rate is 6.48% with regularization when the lambda = 100.

(b)

The training error rate is 4.00% without regularization.

The training error rate is 3.00% without regularization.

The testing error rate is 8.59% with regularization when the lambda = 10.

The testing error rate is 8.04% with regularization when the lambda = 10.

3.

(a)

(i)

False. The test error is not only decided by bias, add more training examples will reduce the variance. So the test error may be reduced.

(ii)

False. We prefer models with low variance and low bias, because it will be able to better fit the training set and testing set.

(iii)

True.

(iv)

False. Introducing regularization to the model may be results in equal or better performance on the testing set. But on the training set, it may be worse.

(v)

False. Using a very large value of regularization will reduce the importance of our optimization function, and hurt the performance of our hypothesis.