Introduction to Data Mining: 21191441
Zhejiang University
Professor Deng Cai

May 9, 2019
Homework 2

Homework 2

Collaborators:
Name: Zhuo Chenn
Student ID: 3170101214

Problem 2-1. A Walk Through Linear Models

(a) Perceptron

Answer:

1. 10: 0%, 10.98%; 100: 0%, 1.34%.
2. 10: 13; 100: 214.
3. The program will not stop because the algorithm cannot converge.

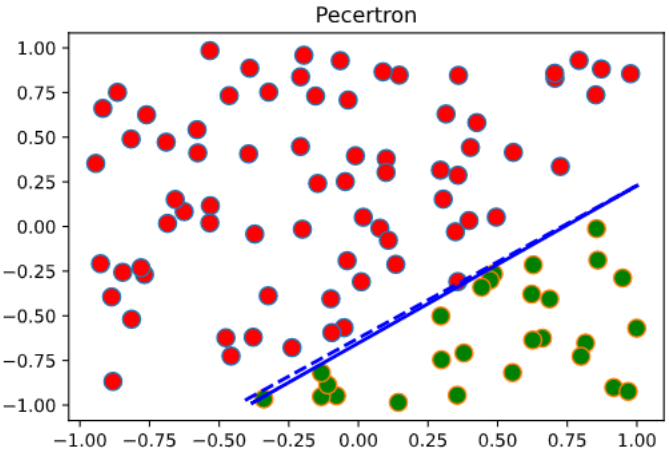


Figure 1: The plotting result for perceptron when nTrain = 100.

(b) Linear Regression

Answer:

1. The training error rate is 3.30%, and test error rate is 3.96%.
2. 13.1% for training set, 5.58% for test set.
3. 49% for training set, 54.79% for test set.
4. 5% for training set, 5.29% for test set.

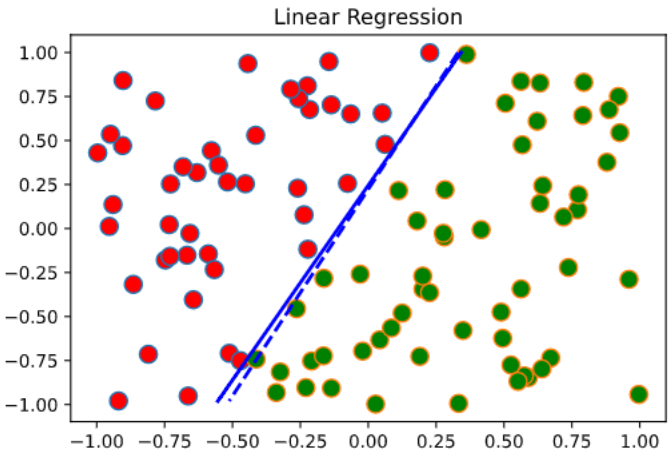


Figure 2: The plotting result for linear regression.

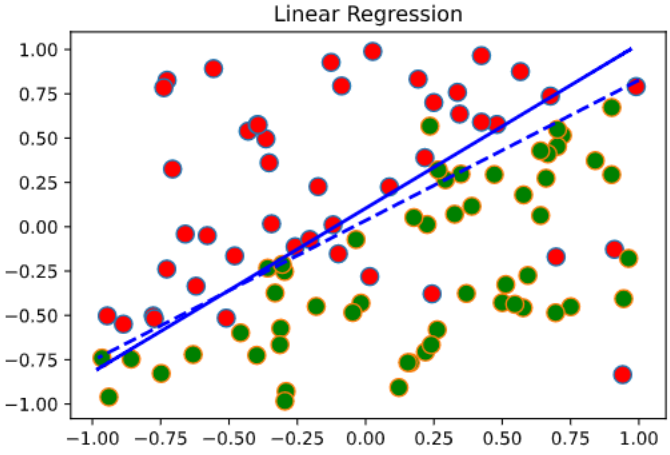


Figure 3: The plotting result for linear regression when training data is not linearly seperable.

(c) Logistic Regression

Answer:

1. 7.52% for training set, 8.55% for test set.
2. 17.1% for training set, 10.8% for test set.

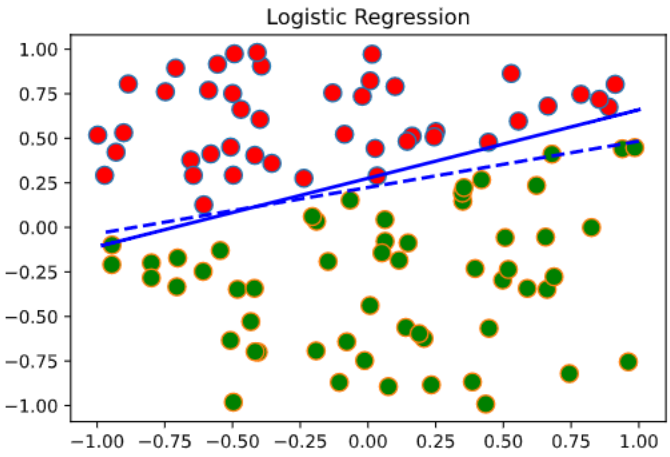


Figure 4: The plotting result for logistic regression.

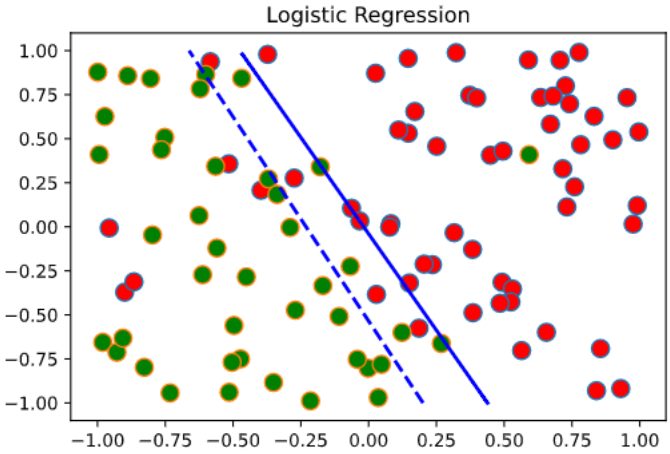


Figure 5: The plotting result for logistic regression when training data is not linearly seperable.

(d) Support Vector Machine

Answer:

1. 0% for training set, 3.22% for test set.
2. 0% for training set, 1.01% for test set.
3. 2.989.
4. (bonus) When the training data is noisy and not linearly separable (nTrain = 100), SVM without slack variables has an avarage error rate of 47.5% for training set and 47.6% for test set, meanwhile SVM with slack variables (C = 4) has an avarage error rate of 13.1% for training set and 5.48% for test set.

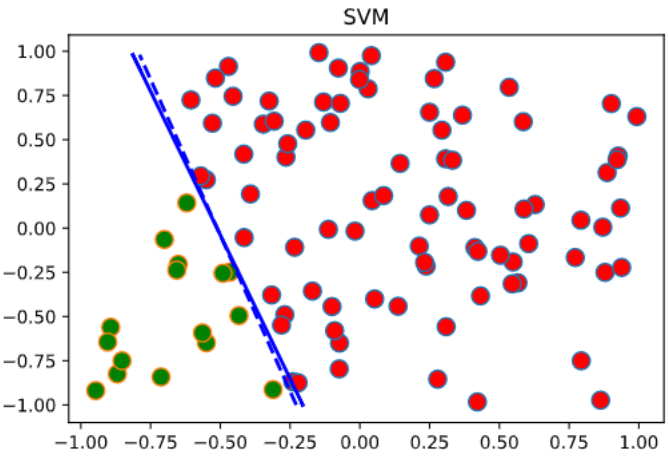
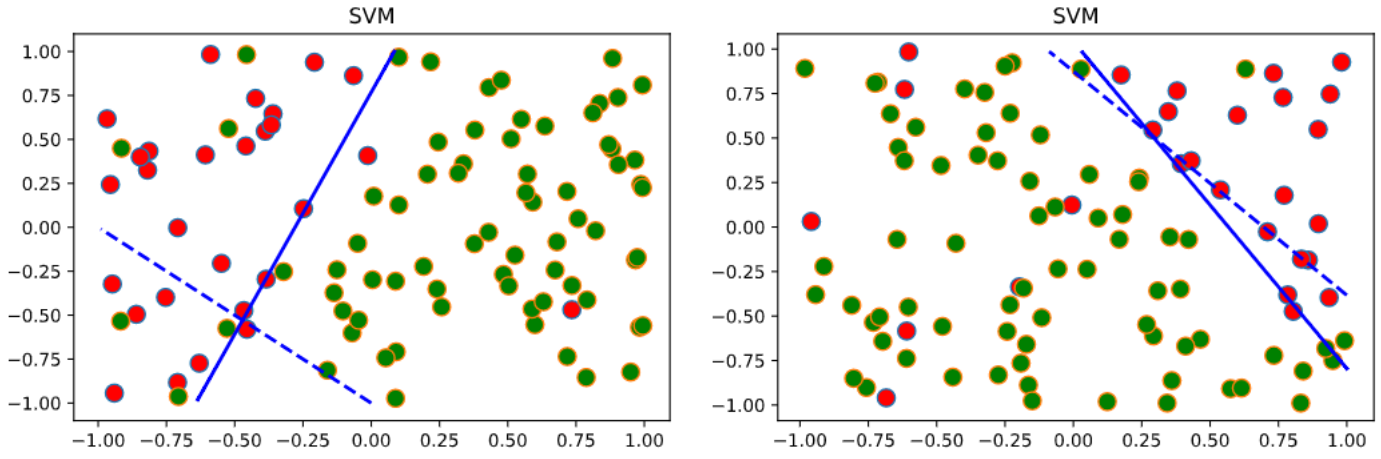


Figure 6: The plotting result for SVM when nTrain is 100.



(a) without slack variables

(b) with slack variables, C = 1

Figure 7: The plotting result for SVM when training data is not linearly seperable.

Problem 2-2. Regularization and Cross-Validation

(a) Implement Ridge Regrssion, and use LOOCV to tune the regularization parameter λ.

Answer:

1. 100
2. 201.78 for $\lambda = 0$, 26.51 for $\lambda = 100$
3. 0%/12.6% for $\lambda = 0$, 0%/5.98% for $\lambda = 100$

(b) Implement Logistic Regrssion, and use LOOCV to tune the regularization parameter λ.

Answer: 0%/6.58% for $\lambda = 0$, 0%/5.02% for $\lambda = 1$.

Problem 2-3. Bias Variance Trade-off

Let's review the bias-variance decomposition first. Now please answer the following questions:

(a) True of False

Answer:

1. F
2. F
3. T
4. F
5. F