

**The Experiment Report of**

***Machine Learning***

**College Software College**

**Subject Software Engineering**

**Members**   **zhancongcong**

**Student ID 201530613573**

**E-mail 643033786@qq.com**

**Tutor**  tanmingkui

**Date submitted** **2017. .**

**1. Topic:**

Linear Regression, Linear Classification and Gradient Descent

**2. Time:**

2017,12,02

**3. Reporter:**

zhancongcong

**4. Purposes:**

1).Further understand of linear regression and gradient descent.

2).Conduct some experiments under small scale dataset.

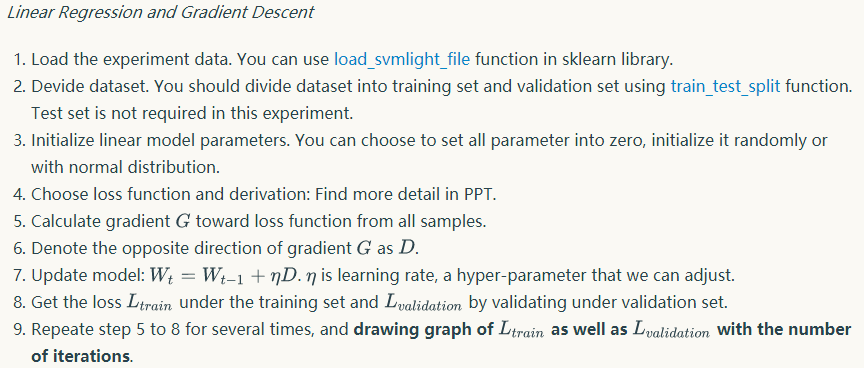
3).Realize the process of optimization and adjusting parameters.

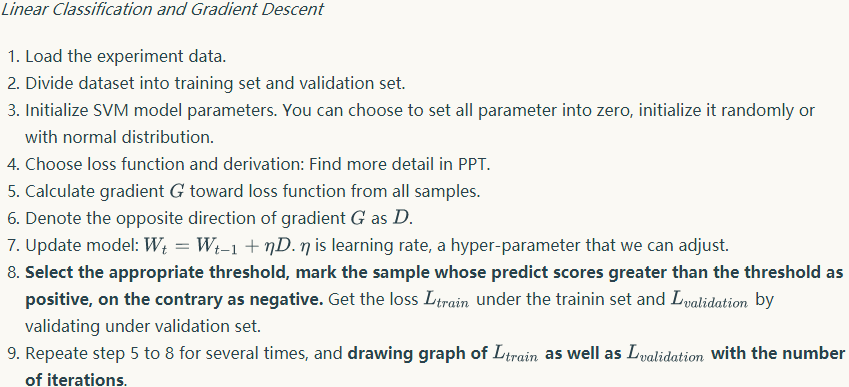
**5. Data sets and data analysis:**

1)Linear Regression uses Housing in LIBSVM Data, including 506 samples and each sample has 13 features. Download scaled edition. After downloading, divide it into training set, validation set.

2)Linear classification uses australian in LIBSVM Data, including 690 samples and each sample has 14 features. Download scaled edition. After downloading, divide it into training set, validation set.

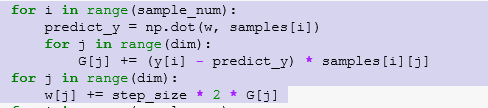
**6. Experimental steps:**



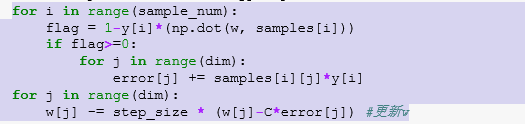


**7. Code:**

The codes to update parament w of Linear Regression ：



The codes to update parament w of Linear Classification :



**Linear Regression(8-12):**

**8. Selection of validation (hold-out, cross-validation, k-folds cross-validation, etc.):**

cross-validation

**9. The initialization method of model parameters:**

full zero

**10. The selected loss function and its derivatives:**





**11. Experimental results and curve:**

## Hyper-parameter selection (η, epoch, etc.):

η=0.01 epoch=500(constant）

## Assessment Results (based on selected validation):

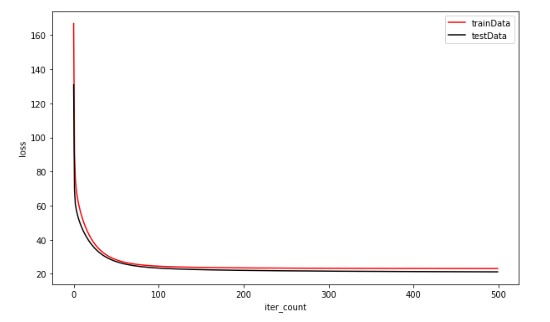
According to the loss curve situation to update the parameters

η=0.001 η=0.0001 η=0.0002 η=0.0003

## Predicted Results (Best Results):

η=0.0002,epoch=500

## Loss curve:



**12. Results analysis:**

The losses of the training set and the loss of the validation set both tend to be smoothed with the number of iterations reduced to the optimum, and the loss value tends to be constant to a minimum.

Linear Classification(8-12):

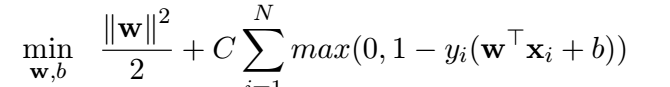
**8. Selection of validation (hold-out, cross-validation, k-folds cross-validation, etc.):**

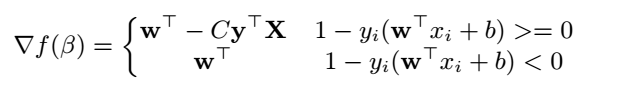
cross-validation

**9. The initialization method of model parameters:**

full zero

**10. The selected loss function and its derivatives:**





**11. Experimental results and curve:**

## Hyper-parameter selection (η, epoch, etc.):

η=0.01 epoch=100(constant），C=1.0

## Assessment Results (based on selected validation):

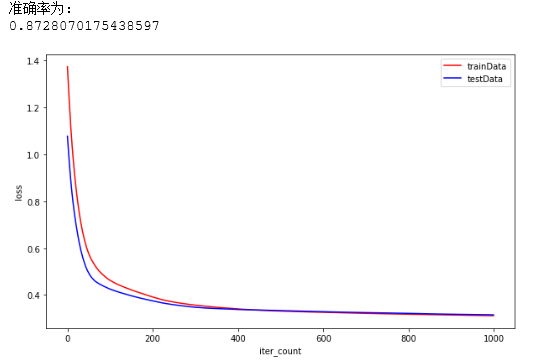
According to the loss curve situation to update the parameters

η=0.001,C=1.0 η=0.001,C=0.9 η=0.0001,C=0.9 η=0.0002,C=0.9

## Predicted Results (Best Results):

η=0.0001,C=0.9,epoch=100

## Loss curve:



**12. Results analysis:**

The losses of the training set and the loss of the validation set both tend to be smoothed with the number of iterations reduced to the optimum, and the loss value tends to be constant to a minimum.

**13. Similarities and differences between linear regression and linear classification:**

Similarities: All are approximations using the gradient descent method to find the best w.

Different points: Linear regression is mainly used to deal with continuous data problems, predict the continuous data function value; and linear classification problems are mainly used to deal with discrete numerical problems for the separation of discrete points, that is, classification problems.

**14. Summary:**

Through this experiment, we can get a deeper understanding of the basic principle of linear regression and linear classification (support vector machine), and how to determine loss function and the derivation of loss function, and finally realize the programming.