

**The Experiment Report of**

***Machine Learning***

**College Software College**

**Subject Software Engineering**

**Members**   **林杨**

**Student ID 201530612279**

**E-mail 291192462@qq.com**

**Tutor**  谭明奎

**Date submitted** **2017.Dec.8th**

**1. Topic:** Linear Regression, Linear Classification and Gradient Descent

**2. Time:** Dec.2nd.2017

**3. Reporter:**林杨

**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:**

Linear Regression uses [Housing](https://www.csie.ntu.edu.tw/~cjlin/libsvmtools/datasets/regression.html" \l "housing" \t "https://www.zybuluo.com/chenyaofo/note/_blank) in [LIBSVM Data](https://www.csie.ntu.edu.tw/~cjlin/libsvmtools/datasets/" \t "https://www.zybuluo.com/chenyaofo/note/_blank), including 506 samples and each sample has 13 features.

Linear classification uses [australian](https://www.csie.ntu.edu.tw/~cjlin/libsvmtools/datasets/binary.html" \l "australian" \t "https://www.zybuluo.com/chenyaofo/note/_blank) in [LIBSVM Data](https://www.csie.ntu.edu.tw/~cjlin/libsvmtools/datasets/" \t "https://www.zybuluo.com/chenyaofo/note/_blank), including 690 samples and each sample has 14 features.

1. **Experimental steps:**

**Linear Regression and Gradient Descent**

(1).Load the experiment data. You can use [load\_svmlight\_file](http://scikit-learn.org/stable/modules/generated/sklearn.datasets.load_svmlight_file.html" \t "https://www.zybuluo.com/chenyaofo/note/_blank) function in sklearn library.

(2).Devide dataset. You should divide dataset into training set and validation set using [train\_test\_split](http://scikit-learn.org/stable/modules/generated/sklearn.model_selection.train_test_split.html" \t "https://www.zybuluo.com/chenyaofo/note/_blank) function. Test set is not required in this experiment.

(3).Initialize linear model parameters. You can choose to set all parameter into zero, initialize it randomly or with normal distribution.

(4).Choose loss function and derivation: Find more detail in PPT.

(5).Calculate gradient G toward loss function from all samples.

(6).Denote the opposite direction of gradient G as D.

(7).Update model: Wt=Wt-1+ηD. η is learning rate, a hyper-parameter that we can adjust.

(8).Get the loss Ltrain under the training set and  Lvalidation by validating under validation set.

(9).Repeate step 5 to 8 for several times, and drawing graph of  Ltrain as well as  Lvalidation with the number of iterations.

**Linear Classification and Gradient Descent**

(1).Load the experiment data.

(2).Divide dataset into training set and validation set.

(3).Initialize SVM model parameters. You can choose to set all parameter into zero, initialize it randomly or with normal distribution.

(4).Choose loss function and derivation: Find more detail in PPT.

(5).Calculate gradient G toward loss function from all samples.

(6).Denote the opposite direction of gradient G as D.

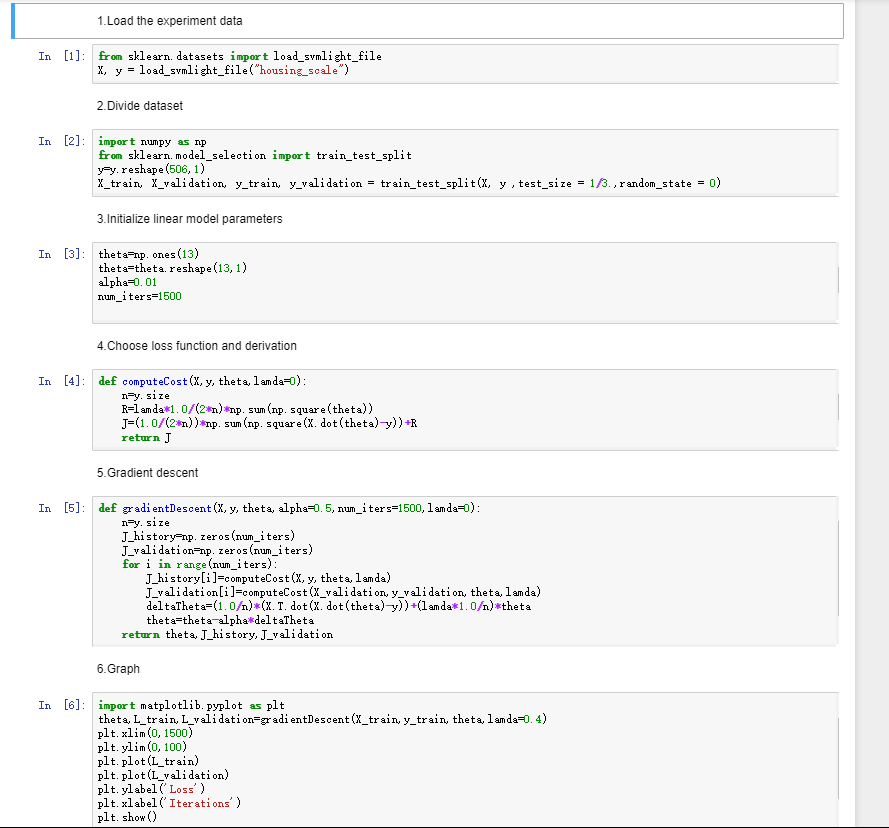
(7).Update model:Wt=Wt-1+ηD. ηis learning rate, a hyper-parameter that we can adjust.

(8).Select the appropriate threshold, mark the sample whose predict scores greater than the threshold as positive, on the contrary as negative. Get the loss Ltrain under the trainin set and Lvalidation by validating under validation set.

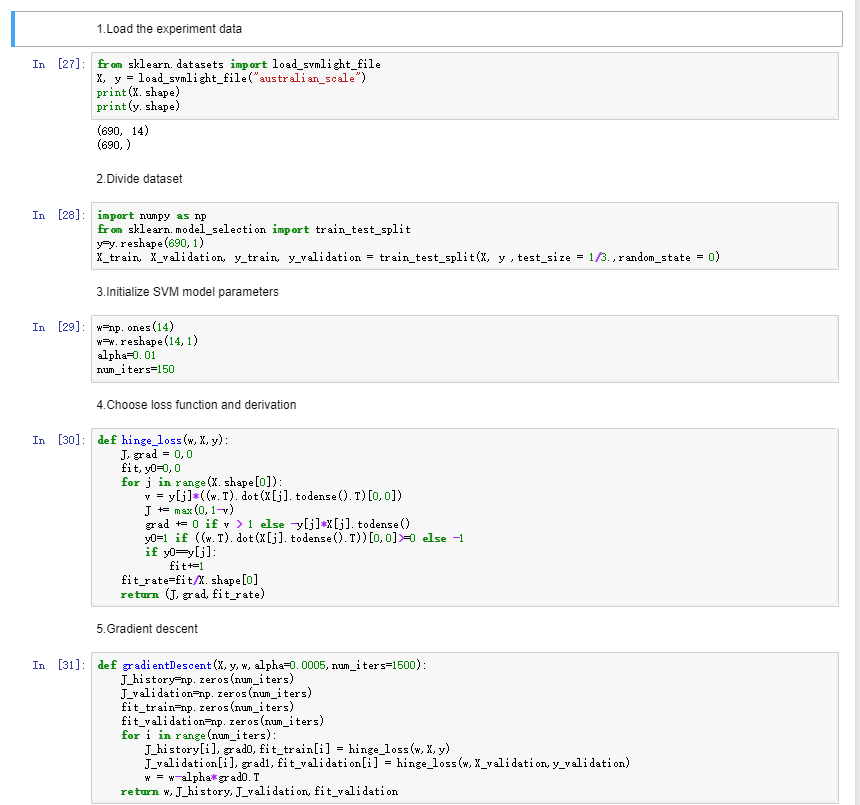
(9).Repeate step 5 to 8 for several times, and drawing graph of Ltrain as well as Lvalidation with the number of iterations.

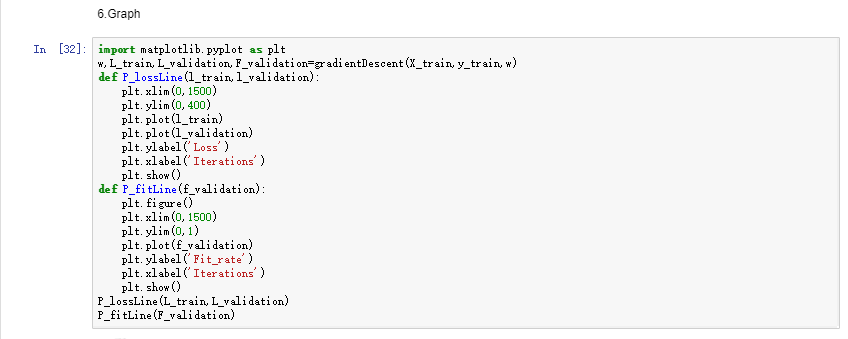
**7. Code:**

**Linear Regression and Gradient Descent**

****

**Linear Classification and Gradient Descent**

****

****

**8. Selection of validation :**

**Linear Regression and Gradient Descent:**hold-out

**Linear Classification and Gradient Descent:**hold-out

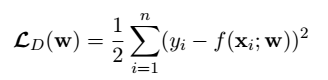
1. **The initialization method of model parameters:**

**Linear Regression and Gradient Descent:**random

**Linear Classification and Gradient Descent:**random

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

**Linear Regression and Gradient Descent:**



**Linear Classification and Gradient Descent:**

Hingeloss=max(0,1-yi(wT\*xi))

Derivatives=={−y xi if y x⋅w<1

0 if y x⋅w>1

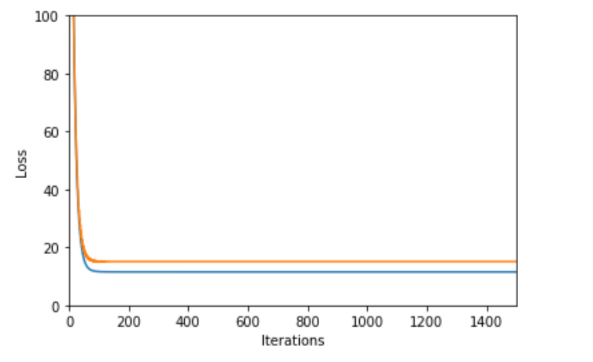
1. **Experimental results and curve:**

**Linear Regression and Gradient Descent**

## Hyper-parameter selection:η=0.5 epoch=1500

## Assessment Results :Ltrain=17

## Predicted Results :Lvalidation=14

Loss curve:

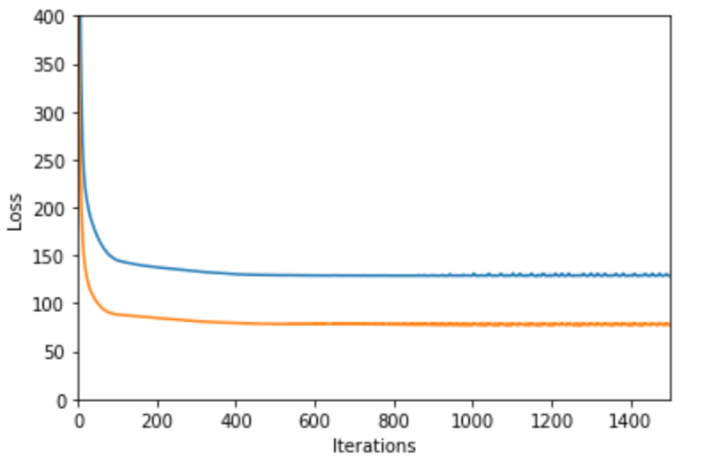
**Linear Classification and Gradient Descent**

## Hyper-parameter selection :η=0.0005 epoch=1500

## Assessment Results :Ltrain=140

## Predicted Results :fit rate=85% Lvalidation=90

## Loss curve:

****

1. **Results analysis:**

Linear Regression and Gradient Descent:The loss of validation data is lower than the loss of training data in the end, and both have reached a stable level at about 15. Thus we have managed to minimize the loss.

Linear Classification and Gradient Descent:The fit rate has reached almost 85%,so it’s good for the classification of the dataset ‘australian’

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

Similarities: both use linear loss function with the way of gradient descent.

Differences:In linear regression, one data is connected with one lable.

In linear classification, lots of data in some area is connected with one lable.

1. **Summary:**

In this section, we conduct some experiments under small scale dataset.By adjusting parameters we realize the process of optimization, and have further understand of linear regression and gradient descent.