

**The Experiment Report of**

***Machine Learning***

**College Software College**

**Subject Software Engineering**

**Members 1**

**Student ID 201530612101**

**E-mail 2435489989@qq.com**

**Tutor Mingkui.Tan**

**Date submitted2017. 12 . 8**

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

**2. Time: 2017-12-02 9:00-12:00 AM**

**3. Reporter: 梁功达**

**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:**

**Dataset:** **Linear Regression uses**[**Housing**](https://www.csie.ntu.edu.tw/~cjlin/libsvmtools/datasets/regression.html#housing)**in**[**LIBSVM Data**](https://www.csie.ntu.edu.tw/~cjlin/libsvmtools/datasets/)**;** Linear classification uses [australian](https://www.csie.ntu.edu.tw/~cjlin/libsvmtools/datasets/binary.html#australian) in [LIBSVM Data](https://www.csie.ntu.edu.tw/~cjlin/libsvmtools/datasets/)

**Data analysis: “Housing” dataset including 506 samples and each sample has 13 features. “Australian”dataset including 690 samples and each sample has 14 features.**

**6. 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) 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) 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  toward loss function from all samples.
6. Denote the opposite direction of gradient  as .
7. Update model: .  is learning rate, a hyper-parameter that we can adjust.
8. Get the loss  under the training set and  by validating under validation set.
9. Repeate step 5 to 8 for several times, and **drawing graph of  as well as  with the number of iterations**.

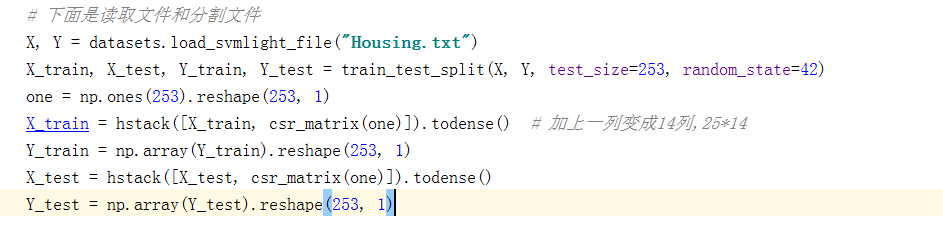
**Linear Classification and Gradient Descent:**

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  toward loss function from all samples.
6. Denote the opposite direction of gradient  as .
7. Update model: .  is 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  under the trainin set and  by validating under validation set.
9. Repeate step 5 to 8 for several times, and **drawing graph of  as well as  with the number of iterations**.

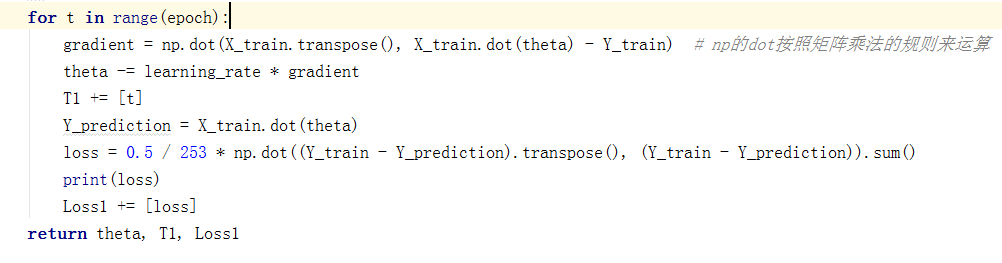
**7. Code:**

**Linear Regression and Gradient Descent :**

**Read file and split it:**

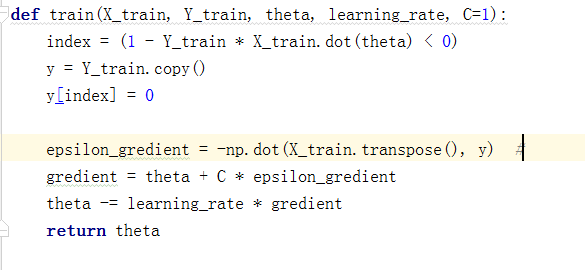
******

**Compute gradient and loss:**

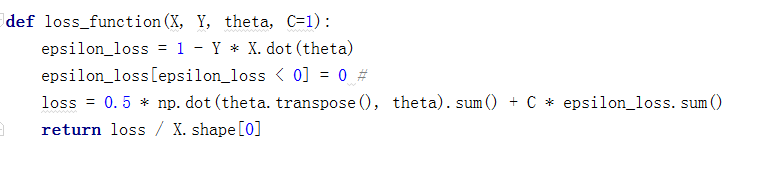
******

**Linear Classification and Gradient Descent:**

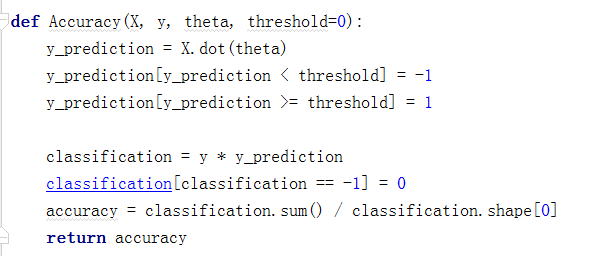
**Train the dataset and get the theta:**

****

**Compute the loss :**

****

**Compute the accuracy of the classification:**

****

(Fill in the contents of 8-12 respectively for linear regression and linear classification)

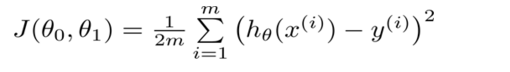
**For linear regression:**

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

**9. The initialization method of model parameters:set all parameter into zero**

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

**Loss function :**

****

**Dericatives:**

****

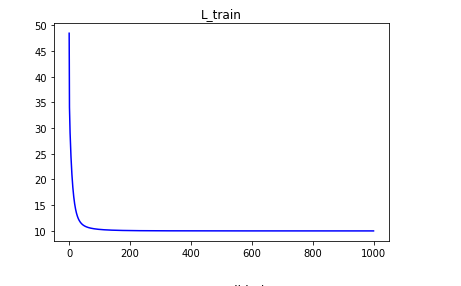
**11.Experimental results and curve:**

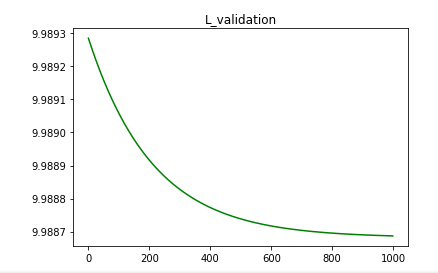
## Hyper-parameter selection (η, epoch, etc.):learning rate=0.001;epoch=1000

## Assessment Results (based on selected validation):Loss of “L\_validation” is smaller than loss of “L\_train”

## Predicted Results (Best Results): Loss of “L\_validation “ is smaller than loss of “L\_train”

## Loss curve:





**12. Results analysis:**

**According the result, we know that the loss has became smaller though the training of train\_dataset. So we know that the model of dataset analysis is useful.**

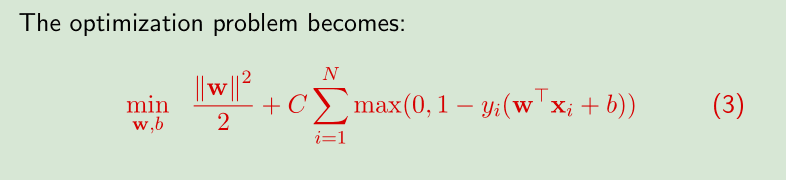
**For linear classification:**

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

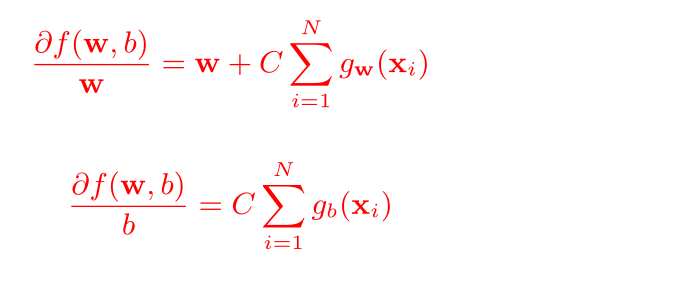
**9. The initialization method of model parameters:** initialize it randomly

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

**Loss function :**



**Derivation:**

****

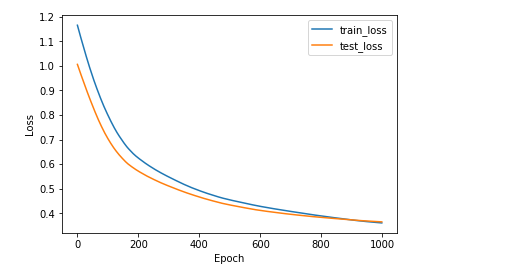
**11.Experimental results and curve:**

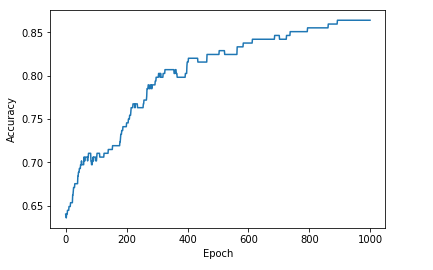
## Hyper-parameter selection (η, epoch, etc.): learning rate = 0.00001;epoch=1000

## Assessment Results (based on selected validation):

## Predicted Results (Best Results): test\_loss is a little smaller than train\_loss , and the accuracy is growing up.

## Loss curve:

****

****

**12. Results analysis:**

**Train-loss maybe a little smaller than Test-loss because theta was trained before used to test dataset. After training, the accuracy of classification is growing up, obviously. As the result says, the training of this dataset is successful.**

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

**Similarities :They do the same thing :model fitting and model matching .**

**Differences: linear regression: the same label corresponds to a X, but the same label of linear classification corresponds to a large number of X**

**14. Summary:**

**Though the experiment, I have known some methods to analyze datasets like splitting dataset, computing gradient descent and calculating loss function. The hardest for me is to code the formulation of loss and gradient .Cause the experiment is not so hard, I have reviewed the contents of class again and I’m felling getting more knowledge though the experiment.**