

The Experiment Report of *Machine Learning*

SCHOOL: SCHOOL OF SOFTWARE ENGINEERING

SUBJECT: SOFTWARE ENGINEERING

*Author: Supervisor:*

ZhiXiang Xue Qingyao Wu

*Student ID: Grade:*

201530613313 Undergraduate

December 14, 2017

1

# Linear Regression, Linear Classification and Gradient Descent

*Abstract*—This experiment focuses on Linear Regression, Linear Classification and Gradient Descent.

## I. INTRODUCTION

w

e made this experiment in order to Compare and understand the difference between gradient descent and stochastic gradient descent, Compare and understand the differences and relationships between Logistic regression and linear classification and Further understand the principles of SVM and practice on larger data.

## II. METHODS AND THEORY

Experiment steps:

Logistic Regression and Stochastic Gradient Descent

1. Load the training set and validation set.
2. Initialize logistic regression model parameters, you can consider initializing zeros, random numbers or normal distribution.
3. Select the loss function and calculate its derivation, find more detail in PPT.
4. Calculate gradient G toward loss function from partial samples.
5. Update model parameters using different optimized methods(NAG, RMSProp, AdaDelta and Adam).
6. Select the appropriate threshold, mark the sample whose predict scores greater than the threshold as positive, on the contrary as negative. Predict under validation set and get the different optimized method loss , , and .
7. Repeat step 4 to 6 for several times, and drawing graph of , , and with the number of iterations.

Linear Classification and Stochastic Gradient Descent

1. Load the training set and validation set.
2. Initialize SVM model parameters, you can consider initializing zeros, random numbers or normal distribution.
3. Select the loss function and calculate its derivation, find more detail in PPT.
4. Calculate gradient G toward loss function from partial samples.
5. Update model parameters using different optimized methods(NAG, RMSProp, AdaDelta and Adam).
6. Select the appropriate threshold, mark the sample whose predict scores greater than the threshold as positive, on the contrary as negative. Predict under validation set and get the different optimized method loss , , and .
7. Repeate step 4 to 6 for several times, and drawing graph of , , and with the number of iterations.

## III. EXPERIMENTS

1. *Dataset*

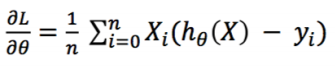
Experiment uses a9a of LIBSVM Data, including 32561/16281(testing) samples and each sample has 123/123 (testing) features.

1. *Implementation*

Logistic Regression and Stochastic Gradient Descent

Loss：屏幕剪辑

Gradient：

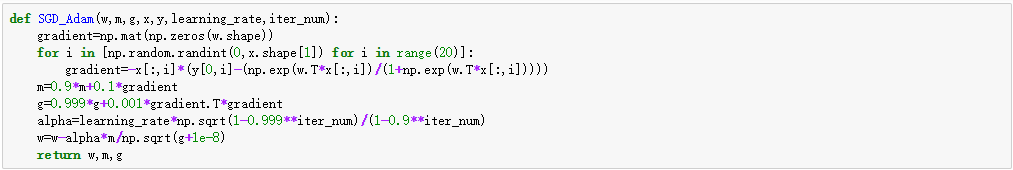


The parameters of four methods we select are as follows.

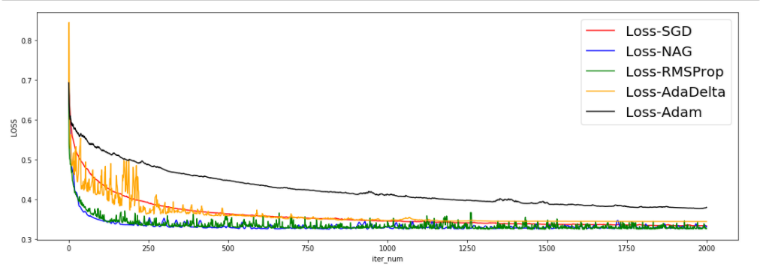
|  |  |
| --- | --- |
| **method** | **parameters** |
| NAG | Iter\_num = 2000  Learning\_rate = 0.002  gamma = 0.9 |
| RMSProp | gamma = 0.9  iter\_num = 2000  learning\_rate = 0.1 |
| AdaDelta | gamma = 0.95  delta = 0.003  iter\_num = 2000 |
| Adam | gamma = 0.9  learning\_rate = 0.1  iter\_num = 2000 |

Main code:





Result:



Linear Classification and Stochastic Gradient Descent

Loss:

屏幕剪辑

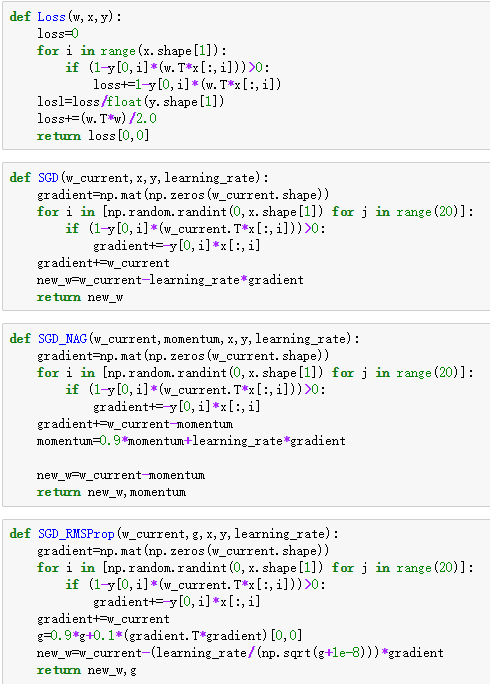
Gradient:

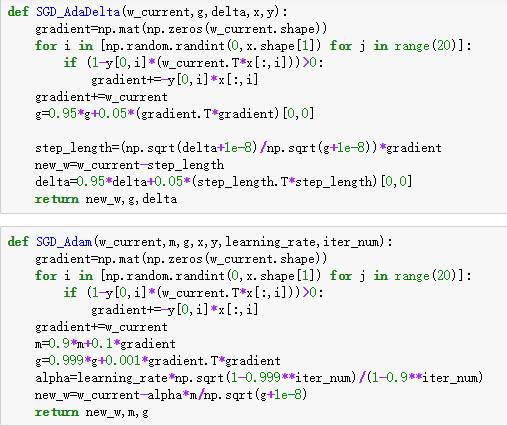
屏幕剪辑

The parameters of four methods we select are as follows.

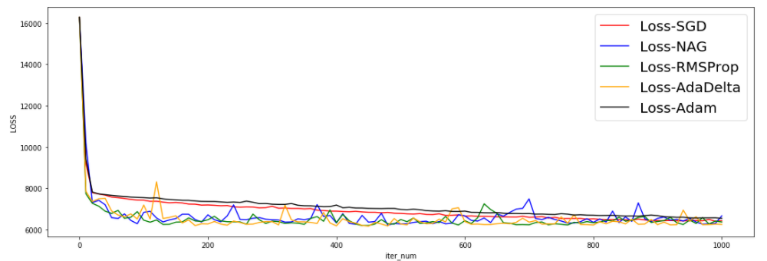
|  |  |
| --- | --- |
| **method** | **parameters** |
| NAG | Iter\_num = 2000  Learning\_rate = 0.002  gamma = 0.9 |
| RMSProp | gamma = 0.9  iter\_num = 2000  learning\_rate = 0.1 |
| AdaDelta | gamma = 0.95  delta = 0.003  iter\_num = 2000 |
| Adam | gamma = 0.9  learning\_rate = 0.1  iter\_num = 2000 |

Main code:





Result:



## IV. CONCLUSION

In this experiment, I implemented logistic regression and linear classification on larger data, using SGD and four different optimization methods. During this process, I found that differences in result of using different optimization methods or just using SGD may be not large. It can’t be said that four different optimization methods must be better than SGD, we should Analyze specific issues to choose the best method. But the speeds of convergence are certainly different.

Through the experiment, I compared and understand the differences and relationships between Logistic regression and linear classification and further understood the principles of SVM and practice on larger data. I feel that I have a deeper understanding of machine learning.