

**The Experiment Report of**

***Machine Learning***

**College Software College**

**Subject Software Engineering**

**Members**   **Yu Guo**

**Student ID 201530611500**

**E-mail G624786883@gmail.com**

**Tutor**   **Mingkui Tan**

**Date submitted 2017. 12 . 2**

**1. Topic:**

Logistic Regression, Linear Classification and Stochastic Gradient Descent.

**2. Time:**

2017.12.2

**3. Reporter:**

Yu Guo

**4. Purposes:**

Achieve logistic regression and linear classification with different methods of gradient descent in python.

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

Experiment uses a9a of LIBSVM Data, including 32561/16281(testing) samples and each sample has 123/123 (testing) features. Please download the training set and validation set.

**6. Experimental steps:**

**Logistic Regression**

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.
4. Calculate gradient 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.
7. Repeat step 4 to 6 for several times, and drawing graph of loss with the number of iterations.

**Linear Classification**

1. Load the training set and validation set.
2. Initialize SVM model parameters, you can consider initalizing zeros, random numbers or normal distribution.
3. Select the loss function and calculate its derivation.
4. Calculate gradient 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.
7. Repeat step 4 to 6 for several times, and drawing graph of loss with the number of iterations.

**7. Code:**

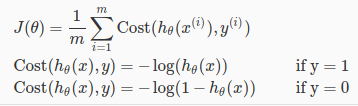
**8. The initialization method of model parameters:**

Logistic regression: Initial with all zeros.

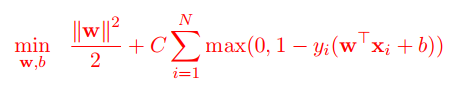
Linear classification: Initial with all zeros.

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

Logistic regression:

Linear classification:

loss function:

 Gradient = ||w|| + 0 yi(Wtxi+b)>=1

||w|| - Cxiyi else

**10. Experimental results and curve:**

## Hyper-parameter selection:

## Parameters are the same for logistic regression and classification.

## NAG: no hyper-parameter

## RMSprop: gamma = 0.99, epslion = 10e-7

## AdaDelta: gamma = 0.9, epslion = 10e-7

## Adam: gamma1 = 0.9, gamma2 = 0.99, epslion = 10e-7

## Predicted Results (Best Results):

## Logistic regression:

## Error rate:

## nag: 0.1595

## rmsProp: 0.1592

## adaDelta: 0.1482

## Adam: 0.1517

## Linear classification:

## Error rate:

## nag: 0.1970

## rmsProp: 0.1973

## adaDelta: 0.1968

## adam: 0.2035

## Loss curve:

## Logistic regression:

## Linear classification:

## 

**11. Results analysis:**

The curves show that Adadelta and NAG are more stable gradient descent ways since the curves have less fluctuate. RMSprop and Adam are fast in convergence but with more fluctuate. In this experiment Adadelta shows the best result.

**12. Similarities and differences between logistic regression and linear classification：**

similarities:

1.handle classification problem and the result is discrete

2.supervised learning

3.both arelinear classifier

Differences:

1.different loss function

**13. Summary:**