

**The Experiment Report of**

***Machine Learning***

**College Software College**

**Subject Software Engineering**

**Members Bo Wang**

**Student ID 20130381731**

**E-mail 765246150@qq.com**

**Tutor MingKui Tan**

**Date submitted 2017.12 .15**

**1. Topic:**

Logistic Regression, Linear Classification and Stochastic Gradient Descent

**2. Time:**

2017/12/8

**3. Purposes:**

1. Compare and understand the difference between gradient descent and stochastic gradient descent.

2. Compare and understand the differences and relationships between Logistic regression and linear classification.

3. Further understand the principles of SVM and practice on larger data.

**4. 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.

**5. Experimental 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. Repeat step 4 to 6 for several times, and drawing graph of ，，and . with the number of iterations.

**6. Code:**

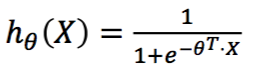
**Logisic Regression**

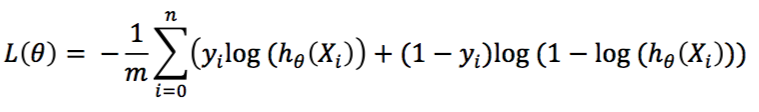
****

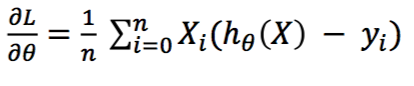
(Fill in the contents of 8-11 respectively for logistic regression and linear classification)

**Logisic Regression**

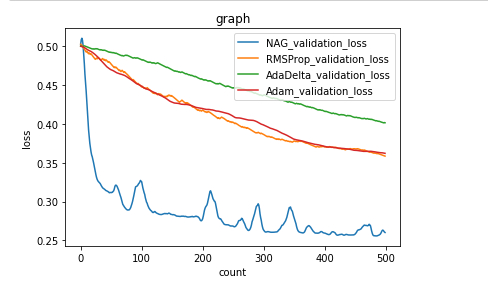
1. **The selected loss function and its derivatives:**







**9. Experimental results and curve:**

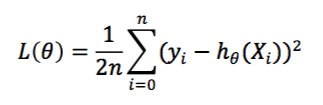
We get the following loss graphs as results after running the program.

1. **Results analysis:**

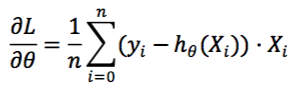
From the graph we find that NAG\_loss reaches the local optimal solution fastest, but it also has vibration. RMSProp and Adam are closely overlapped.

**Linear classification**

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







**9. Experimental results and curve:**

1. **Results analysis:**

.

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

**12. Summary:**