

**The Experiment Report of**

***Machine Learning***

**College Software College**

**Subject Software Engineering**

**Members**  Zijie Hong

**Student ID 201530611593**

**E-mail 464538338@qq.com**

**Tutor**   **Qingyao Wu**

**Date submitted** **2017.12 .14**

**1. Topic: Comparison of SVM and logistic regression for solving classification problems**

**2. Time: 2017.12.14**

**3. Reporter: Zijie Hong**

**4. Purposes: To compare SVM and logistic regression for the problem of classification using BSGD**

**5. Data sets and data analysis: a9a**(**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 and Stochastic Gradient Descent*

1. Load the training set and validation set.
2. Initalize logistic regression model parameters, you can consider initalizing zeros, random numbers or normal distribution.
3. Select the loss function and calculate its derivation, find more detail in PPT.
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 losses.
7. Repeate step 4 to 6 for several times, and **drawing graph of losses**

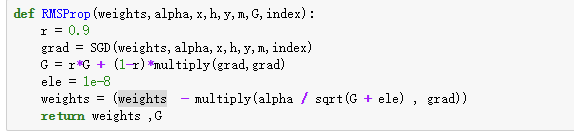
*Linear Classification and Stochastic Gradient Descent*

1. Load the training set and validation set.
2. Initalize SVM model parameters, you can consider initalizing zeros, random numbers or normal distribution.
3. Select the loss function and calculate its derivation, find more detail in PPT.
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 ，， and .
7. Repeate step 4 to 6 for several times, and **drawing graph of those loss**

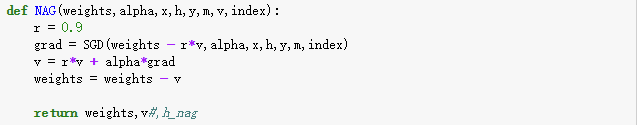
**7. Code:**

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

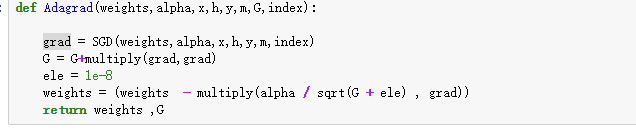
RMSProp:

****

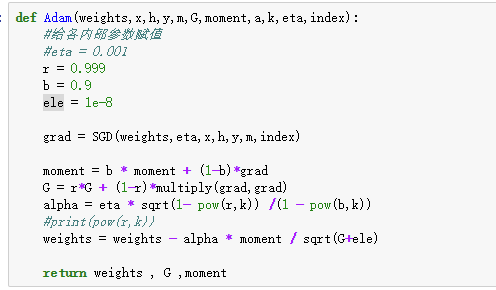
**NAG**

****

**Adagrad**

****

**Adam**

****

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

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

**10. Experimental results and curve:**(Fill in this content for various methods of gradient descent respectively)

## Hyper-parameter selection:

## Predicted Results (Best Results):

## Loss curve:

**11. Results analysis:**

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

**13. Summary:**