

**The Experiment Report of**

***Machine Learning***

**College Software College**

**Subject Software Engineering**

**Members**  黄品超

**Student ID 201530611760**

**E-mail 201530611760@mail.scut.edu.cn**

**Tutor**   **Mingkui Tan**

**Date submitted** **2017. 12 . 12**

1. **Topic:**

Logistic Regression, Linear Classification and Stochastic Gradient Descent

1. **Time:**

2017-12-02 2:00-5:00 PM B7-138/238

**Reporter:**

黄品超

1. **Purposes:**

In this experiment, the first target is to compare and understand the difference between gradient descent and stochastic gradient descent.

Secondly, after the whole process, a realization of

differences and connection between Logistic regression and linear classification should have cultivated.

Finally, implement practicing over larger data, for further understanding principles of SVM.

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

The following content is completely came from the experiment guide.

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.

1. **Experimental steps:**

The following content is completely came from the experiment guide.

The experimental code and drawing are completed on jupyter.

*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 loss ，， and .
7. Repeate step 4 to 6 for several times, and drawing graph of ，， and with the number of iterations.

*Linear Classification and Stochastic Gradient Descent:*

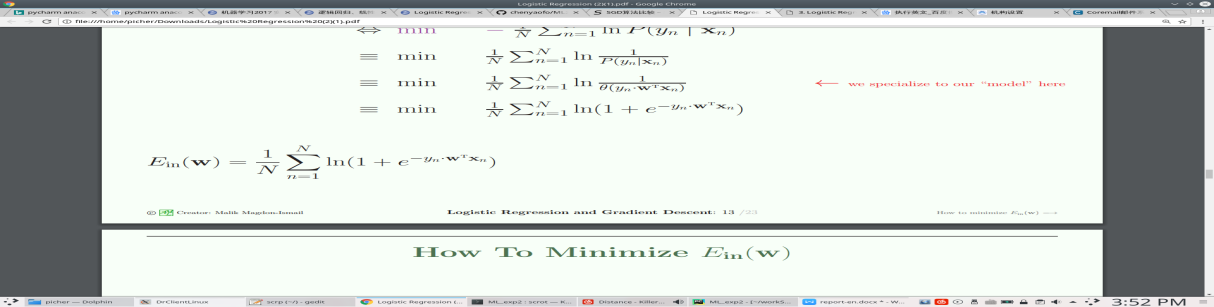
1. Load the training set and validation set.
2. Initalize SVM model parameters, you can consider initalizing zeros, random numbers or normal distribution.
3. Select the loss function and calculate its derivation, find more detail in PPT.
4. Calculate gradient toward loss function from partial samples.
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 ，， and with the number of iterations.

**7. Code:**

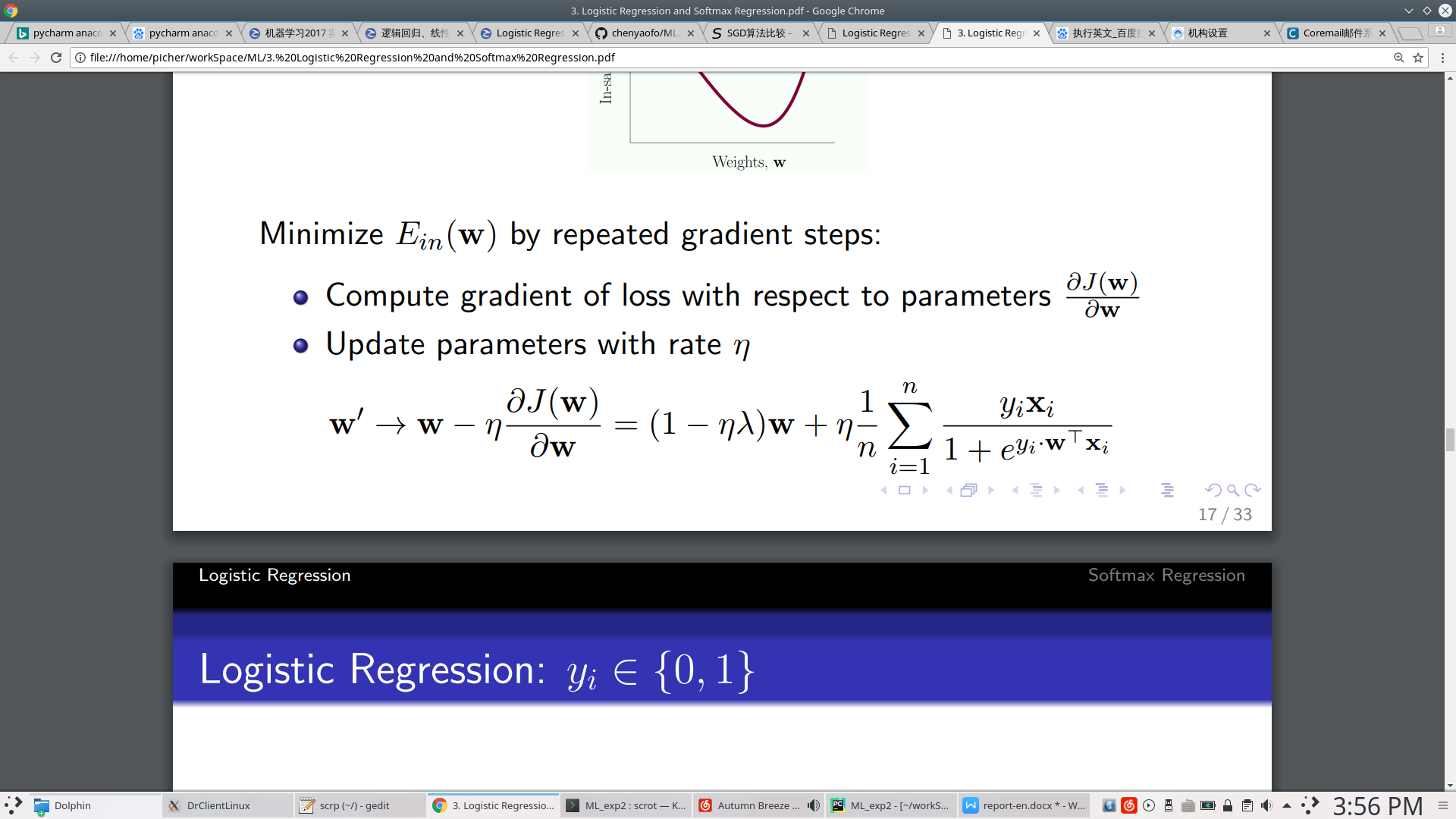
(Fill in the contents of 8-11 respectively for logistic regression and linear classification)

1. **The initialization method of model parameters:**
2. **The selected loss function and its derivatives:**

The loss function I have chosen is Cross Entropy Error, as follow shows:

****

**Derivatives:**

****

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

NAG

## Hyper-parameter selection:

## Predicted Results (Best Results):

## Loss curve:

## RMSProp

## Hyper-parameter selection:

## Predicted Results (Best Results):

## Loss curve:

## AdaDelta

## Hyper-parameter selection:

## Predicted Results (Best Results):

## Loss curve:

## Adam

## Hyper-parameter selection:

## Predicted Results (Best Results):

## Loss curve:

1. **Results analysis:**
2. **Similarities and differences between logistic regression and linear classification：**

Similarities:

1. These two method share the same target of classification.
2. Data of the two method is still binary.
3. The model functions they based on are all nonlinear.

Differences:

1. Logistic regression chooses a really different model function at first of the mission, which it introduced the sigmoid function to make the function continuous.
2. Logistic regression chooses a different loss function, as a result of which, the derivative of them are different as well.

**13. Summary:**