华南理工大学

South China University of Technology

The Experiment Report of *Machine Learning*

SCHOOL: SCHOOL OF SOFTWARE ENGINEERING

SUBJECT: SOFTWARE ENGINEERING

Author:
Peizhuo Xie

Supervisor:
Mingkui Tan

Student ID:
201530613207

Grade:
Undergraduate

December 15, 2017

Linear Regression, Linear Classification and Gradient Descent

***Abstract*—We do a experiment and here is our result.**

I. Introduction

Classification problems play a important role in machine learning and mean a lot in our daily life. Therefore, we conduct a experiment on classification problems. In our experiment, we address a classification problems through logistic regression and support-vector-machine(svm). We use the a1a dataset from LIBSVM. Our methods have a good performance and achieve 86% accuracy. Although the accuracy is not outstanding compared with state-of-the-art method, our experiment is still meaningful ——– it teaches us a lot. Let us make a brighter future together.

II. Methods and Theory

We try both logistic regression and svm to address this problem.
In logistic regression, we use

$$P_{is\_positive} = \frac{1}{1 + exp{-}(\mathbf{w^T x} + b)} \tag{1}$$

to calculate the probability of a positive sample. Therefore, whole Probability can be written to

$$\begin{aligned} P &= P(Y|\mathbf{w}, b)P(\mathbf{w}) \\ &= \prod_{i=1}^{N} \{\Omega(y=1)P_{y_i=1} + \Omega(y=-1)P_{y_i=-1}\}P(\mathbf{w}) \end{aligned} \tag{2}$$

After taking the negative logarithm of Eqn.(2), we get our final optimization problem

$$\min_{\mathbf{w},b} \sum_{i=1}^{N} log(1 + exp - (y_i(\mathbf{w^T x_i} + b))) + \lambda \|w\|_2^2 \tag{3}$$

In svm, we simply have our loss function

$$\min_{\mathbf{w},b} \|w\|_2^2 + C \sum_{i=1}^{N} max(0, y_i(\mathbf{w^T x_i} + b))) \tag{4}$$

III. Experiments

A. Dataset

In this experiment, we use a1a dataset from LIBSVM. It has 123 features. Training set has 32561 samples and testing set has 16281 samples.

B. Implementation

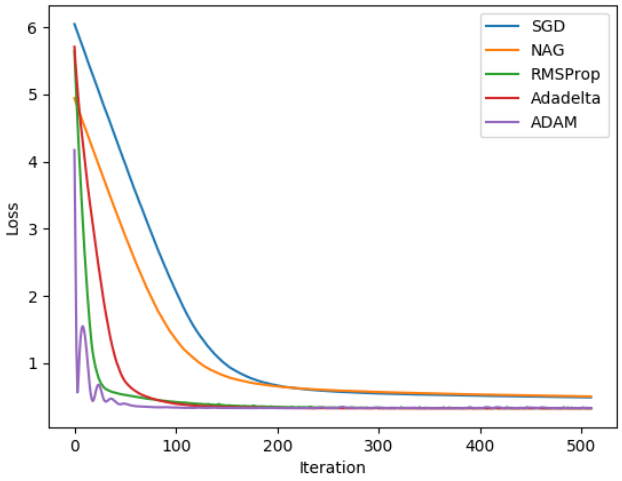
We do the experiment by these steps:

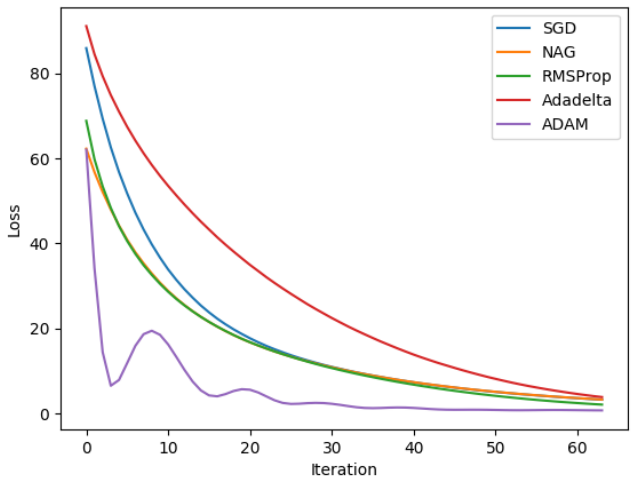
- Using the function load_svmlignt_file from sklearn read the data set.
- Initializing the parameters randomly.
- Using Cross-entropy loss / Hinge loss and compute all the samples loss.
- Update the loss function by methods = ['SGD', 'NAG', 'RMSProp', 'Adadelta', 'ADAM']
- Repeat the above steps for some times and draw the figure of loss-iteration.

Especially, the parameters are seted as follow:

- sgd_para_list = {'learning_rate': 0.01}
- nag_para_list = {'miu': 0.9, 'learning_rate': 0.01}
- rmsprop_para_list = {'delta': 10e-7, 'rho': 0.95, 'learning_rate': 0.1}
- adam_para_list = {'delta': 10e-8, 'rho1': 0.9, 'rho2': 0.999, 'learning_rate': 0.1}
- adadelta_para_list = {'delta': 10e-7, 'miu': 0.9, 'learning_rate': 10}

All the implementation details can be seen in https://github.com/unbreading/ML-02. The results are as follow:





IV. Conclusion

After this experiment, we strengthen our coding ability and we have a better understanding about logistic regression and svm. We appreciate all the teachers and teaching assistants' works a lot.Thank you very much.