

**The Experiment Report of**

***Deep Learning***

**College Software College**

**Subject Software Engineering**

**Student ID 201720145037\_**

**E-mail 790567025@qq.com**

**Tutor**   **Mingkui Tan**

**Date submitted** **2017.12 .7**

**1. Topic:** Linear Regression, Linear Classification and Gradient Descent

**2. Time:** 2017-12-02

**3. Reporter:** Yihan Zheng

**4. Purposes:**

1. Further understand of linear regression and gradient descent.
2. Conduct some experiments under small scale dataset.
3. Realize the process of optimization and adjusting parameters.

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

1. Linear Regression uses 'Housing' in LIBSVM Data, including 506 samples and each sample has 13 features. You are expected to download scaled edition. After downloading, you are supposed to divide it into training set, validation set.
2. Linear classification uses 'australian' in LIBSVM Data, including 690 samples and each sample has 14 features. You are expected to download scaled edition. After downloading, you are supposed to divide it into training set, validation set.

**6. Experimental steps:**

1. Load the experiment data and divide the dataset into training set and validation set.
2. Initialize linear model parameters. Set all parameter into zero, initialize it randomly or with normal distribution.
3. Define the loss function of the linear regression to be Least squared loss, and the loss function of the linear classification to be Hingle loss.
4. Compute the gradient of the loss function with respect to the weight W and bias b.
5. Update the parameters W and b.
6. Repeat above steps for several times until convergence.
7. **Code:**

linear regression:

****

****

****

****

linear classification**:**

****

****

****

1. **Selection of validation (hold-out, cross-validation, k-folds cross-validation, etc.):**

hold-out

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

np.random.normal()

1. **The selected loss function and its derivatives:**
2. Linear regression
3. Loss function
4. Gradient with the respect of the W
5. Linear classification
6. Loss function
7. Gradient with the respect of the W

**11. Experimental results and curve:**

## Hyper-parameter selection (η, epoch, etc.):

linear regression as example : learning-rate and regularization parameter both try the value in [10-4, 10-3, 10-2, 10-1, 1, 10], and the result shows that lr:0.0001,reg:0.001 get the best result

## Assessment Results (based on selected validation):

linear regression:

avg\_train\_loss:73.49406089058121,

avg\_val\_loss:73.55045788432349

linear classification:

avg\_train\_loss:0.522114676852641,

avg\_val\_loss:0.5618017028780596

## Predicted Results (Best Results):

linear regression:

min train loss:55.50353308622947,

min val loss:55.50353308622947

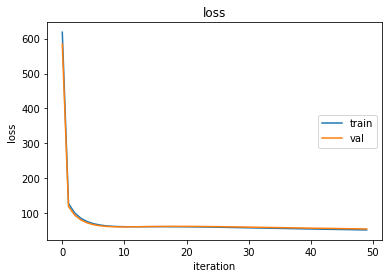
linear classification:

min train loss:0.33386760749395455,

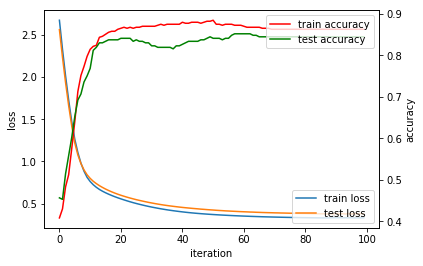
min val loss:0.37745552469952004

## Loss curve:

linear regression:



linear classification:



1. **Results analysis:**
   * If the learning rate is too small, the update of the model will be very slow.
   * If the regularization parameter is too small, the proportion of the regularization term will be small and the model may be easy to fall into over-fitting.
   * If the learning rate is too large, the step of the update will be too fast that may miss the minimal.
   * If the regularization parameter is too small, the proportion of the error term will be small and the model may be under-fitting.

**13. Similarities and differences between linear regression and linear classification:**

Linear regression and linear classification belong to supervised learning.

The difference between classification and regression is the type of output variable, Quantitative output is called regression and Qualitative output is called classification. The classification model can discretize the output of the regression model, regression models also allow the output of the classification model to be continuous.