

**The Experiment Report of**

***Machine Learning***

**College Software College**

**Subject Software Engineering**

**Student ID 201722800100**

**E-mail Az.yamen@gmail.com**

**Tutor**   **Mingkui Tan**

**Date submitted** **2017.12 .08**

**Linear Regression, Linear Classification and Gradient Descent**

**Time:** 2017-12-08

**Reported By:**

**Ezzaddin Ahmed Othman Saeed**

**Purposes:**

* Use data set to identify relationships among variables and use these relationships to make predictions(Regression).
* To estimate likely performance of a model on out-of-sample data.
* To realize the process of optimization and adjusting parameters and further understand of liner regression and gradient descent.

**Data sets and data analysis:**

In this work, I used two LIBSVM datasets which are pre-processed data originally from UCI data repository.

* *Data set 1:*  Housing (Boston) :
  + (data: 506 , features: 13):

*Linear Regression*uses (Housing) data set in LIBSVM Data, including 506 samples and each sample has 13 features.

This data set  contains information about the housing values in suburbs of Boston.

* *Data set 2* : Australian
  + ( classes: 2 , data: 690 of features: 14 )

*Linear classification* uses (Australian) data set  in LIBSVM Data, including 690 samples and each sample has 14 features.

**Experimental steps:**

Experiment 1 : (Linear Regression and Gradient Descent )

***step 1:*** Load the experiment data. You can use [load\_svmlight\_file](http://scikit-learn.org/stable/modules/generated/sklearn.datasets.load_svmlight_file.html" \t "_blank) function in sklearn library.

***step 2:*** Divide dataset. Here I divide dataset into training set and validation set using train\_test\_split function. Of course Test set is not required in this experiment.

***step 3:*** Initialize linear model parameters. Here I choose normal distribution to initialize the model .

***step 4:*** Choose loss function and derivation: Find more detail in PPT.

***step 5:*** from all samples Icalculate gradient  toward loss function.

***step 6:*** Denote the opposite direction of gradient  as  G -> D.

***step 7:*** Update model.

***step 8:*** Get the loss  under the training set and  by validating under validation set.

***step 9:*** Repeat step 5 to 8 for several times, and drawing graph of  as well as  with the number of iterations.

Experiment 2 : *( Linear Classification and Gradient Descent* )

*step 1:*Load the experiment data.

*step 2:*Divide dataset into training set and validation set by using train\_test\_split function .

*step 3:*Initialize SVM model parameters with normal distribution.

*step 4:*I identify loss function and derivation.

*step 5:*Calculate gradient  toward loss function from all samples.

*step 6:*Denote the opposite direction of gradient  G as D .

*step 7:*Update model.

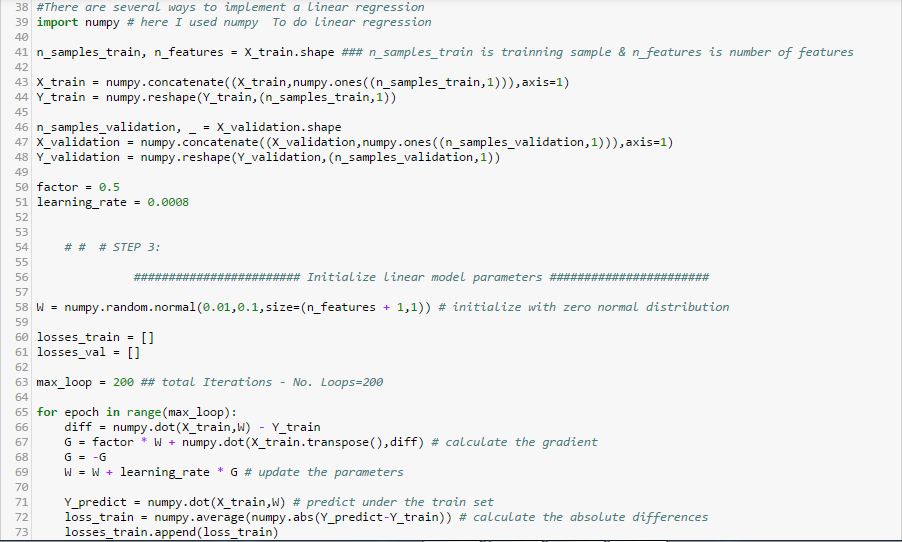
*step 8:*Get the loss  under the train in set and  by validating under validation set.

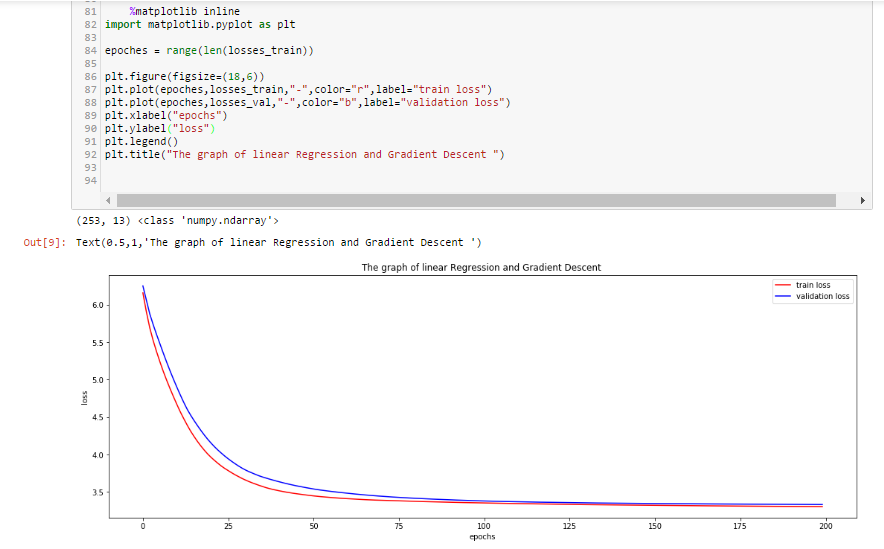
*step 9:*repeat step 5 to 8 for several times, and drawing graph of  as well as  with the number of iterations.

**Code:**

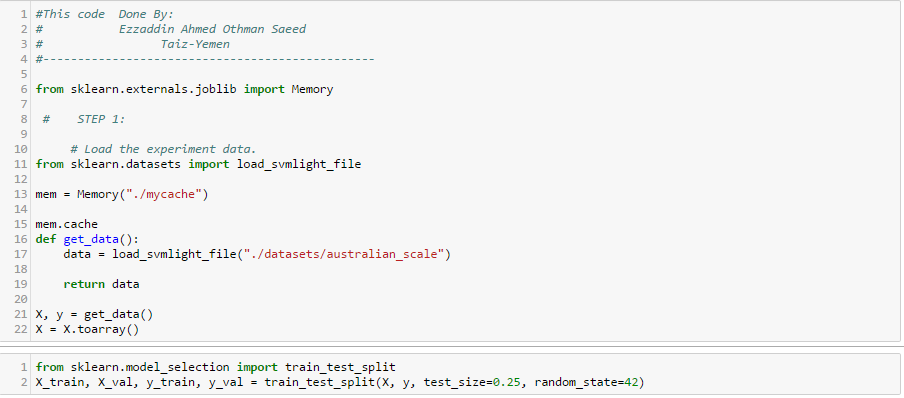
For linear regression :

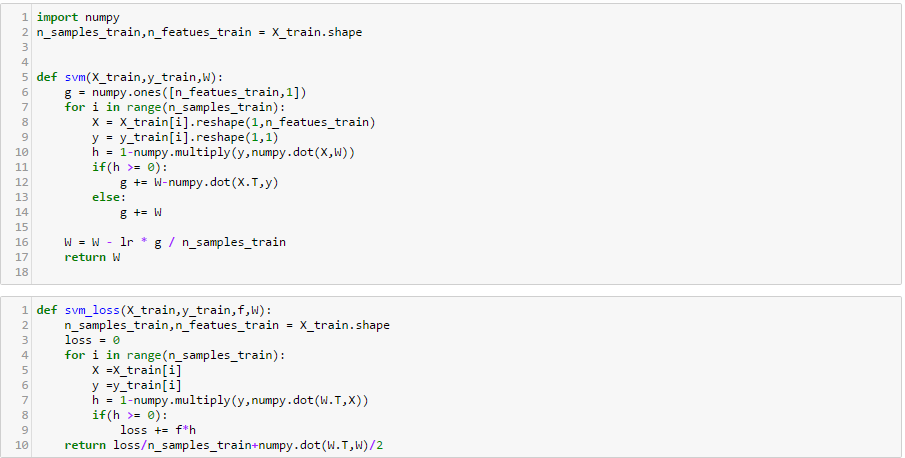






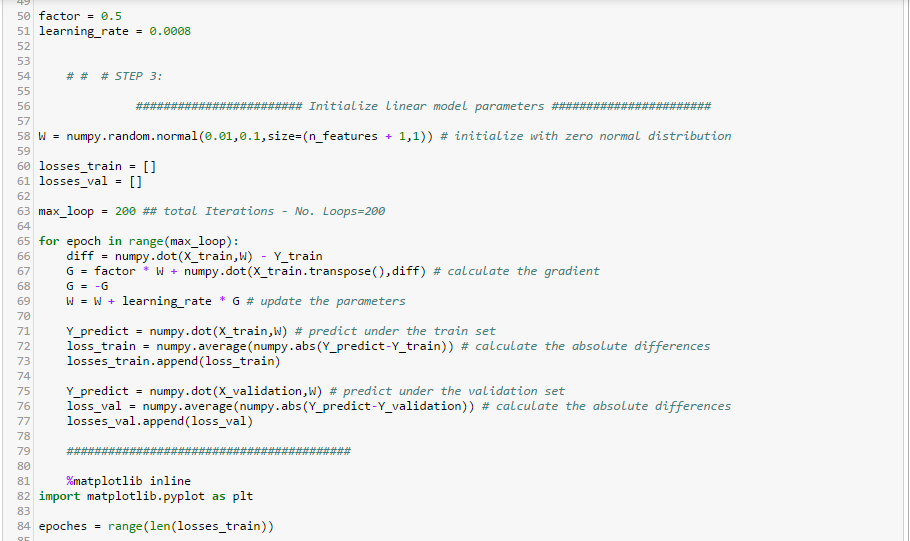
For linear classification







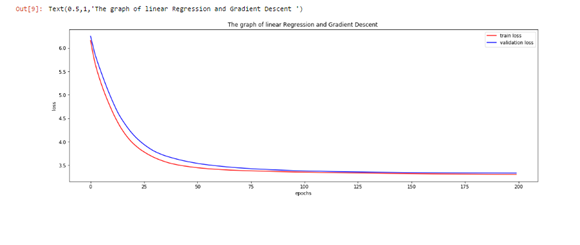
**The initialization method of model parameters:**

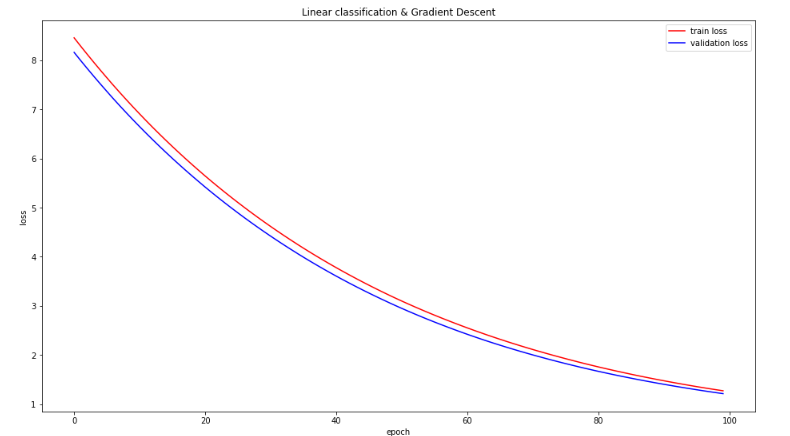


**Experimental results and curve:**

## Hyper-parameter selection (η, epoch, etc.): η= 0.0008 and 0.01 , epoch=200 and 100

## Assessment Results (based on selected validation):





**Results analysis:**

From the graph its easily understandable that the train curve and the test curve are almost same.

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

|  |  |  |
| --- | --- | --- |
| **Type** | **linear regression** | **linear classification** |
| type of output | Continuous | discrete |
| Category | supervised | supervised |

Regression is used to predict **continuous values**. Classification is used to predict which class a data point is part of (**discrete value**).

Both regression and classification problems belong to the supervised category of machine learning. In Supervised machine learning, a model or a function is learnt from the data to predict the future data. In simple terms, y=f(x) is a predictive model learnt from the data set D = {(X1,y1),…(Xn,y2} where X is the input vector and y is the output.

Based on the type of output y, the learning problems are classified into regression and classification. In case of classification, the output variable is discrete and in regression, the output variable is continuous.

**Summary:**

In this Report, after apply Experiment, we understand similarity and differences between linear regression and gradient descent through conduct some experiments under small scale dataset. In Experiment 1 we use regression for predicting housing prices in the boston dataset present in the sklearn datasets .