**Report**

**Introduction**

I partition the original dataset to TRdata and TEdata as training data and testing data ,and then keep the columns hr ,atemp , hum ,windspeed as independent variables , cnt as dependent variable to build the model.

Then set the original beta0,1,2,3,4=1, alpha=0.5, conv\_threshold=100, max\_iterations=1000, then we got the cost function value (MSE)=22328.59

Then I run the loop, get the initial parameter values is as follows:

Optimal slope and optimal intercept for my linear regression model:

"Optimal intercept: 2.03021435807894 Optimal slope: 13.5922336772969 1.63430647791061 1.54267718481769 1.21075621884279"

Then, we obtain that when alpha=0.5, conv\_threshold=100, the optimal linear regression model is

yhat=2.0302+13.592\*hr+1.634 \*temp+1.542\*hum+1.2107\*winspeed

MSEt =28849.3526

**Experimentation**

**1**.

I keep the conv\_threshold =100, then change four different alpha values to gradient descent function

Then we get the different cost function values (MSE\_alph).

When Alpha=0.005 , the MSE0=29110.39

When Alpha=0.001,the MSE1=29141.34

When Alpha=0.015,the MSE2=29110.04

When Alpha=0.012,the MSE3=132045.57

Then we can get the graph as follows. Since the MSE get to the minimum value at it shows that when the best alpha value is 0.01



**2.**

I keep alpha=0.01, then change four different conv\_threshold values to the gradient descent function.

Then we obtain the different cost function values (MSE\_thre).

When conv\_threshold =0.001, the MSE\_a=29110.39

When conv\_threshold =0.01, the MSE\_b=29141.34

When conv\_threshold =0.05, the MSE\_c=29110.04

When conv\_threshold =0.1, the MSE\_d=132045.57

Then we can get the graph as follows, it shows that when the best threshold value is 0.001 and the corresponding minimum cost function result is 29110.39



**3.**

I select hr, atemp, hum as our three features to build the linear regression model and still obtain that alpha=0.5, conv\_threshold=100, then the optimal linear regression model is

• yhat=2.0284+13.49\*hr+1.6327 \*atemp+1.5427\*hum

• The training and test error result of the three features model is 28858.9462.

When there are four features hr,atemp,hum,winspeed, the linear regression model is

• yhat=2.0302+13.592\*hr+1.634 \*atemp+1.542\*hum+1.2107\*winspeed

• The training error result and the testing error result is 28849.3526.

As we presenting above, we can get to know that as the number of features reduces, the training and testing error result will be growingly higher.

**4.**

I pick up hr, hum, windspeed as the three features which are best suited to predict the output. And then

Then we get the linear regression model as follows:

• yhat =2.0367+ 13.6423\*hr+ 1.6373\*hum+ 1.2118\*windspeed

• The training and testing error result is 28867.4135.

Compared with all feature and random three features, the variables that I picked up are not the best suitable model. Maybe because the combinations of windspeed and huminity are not that good enough compared to the combinations of temperature and humidity factor to predict the number of renting bikes.

**Discussion**

For my perspective, the selection of feature, learning rate (alpha) and threshold matters during the process of model building. The more related features we selected like temperature feature, the more precise model we can obtain. What’s more, we can set the learning rate to 0.01 and set the threshold value to 0.001 to obtain the minimal cost function value. Therefore, our linear regression model can be much closer to the perfect and with minimal bias to the fact.