Machine Learning

Assignment 2

Question 6c & 6d

IIT2019001
Karan Kunwar

Solution Link : https://github.com/southpole01/ML_Assignments

Problem Statement 6c: Design Predictor using Batch Gradient Descent Algorithm, Stochastic Gradient Algorithm and mini batch Gradient Descent algorithms (determining mini-batch size is your choice- here it could be 10, 20, 30 etc.) with and without regularisation and compare their performances in terms of % error in prediction.

Observations :

1) Using scaled batch gradient without regularisation
Final coefficients are:
[7730.872053867435, 8069.264303687423, 11077.015405893277, 18485.569122447192]
Mean absolute percentage error is : **20.34046542497449 %**

2) Using scaled batch gradient with regularisation
Final coefficients are:
[5037.585668619078, 11147.667574879839, 10378.580439168689, 22647.298983883848]
Mean absolute percentage error is : **19.92701396456417 %**

3) Using Stochastic gradient without regularisation
Final coefficients are:
[18648.663069990776, 15073.501985961251, 15766.862790309351, 22357.23427068568]
Mean absolute percentage error is : **32.87903970192347 %**

4) Using Stochastic gradient with regularisation
Final coefficients are:
[70204.71424890438, 187.30113153910685, 890.065454652648, 300.99208452729107]
Mean absolute percentage error is : **22.61033001770984 %**

5) Using Mini-batch gradient without regularisation for batch size = 20
Final coefficients are:
[6340.552295015755, 2827.87646125345, 15916.90159915714, 10968.772912396124]
Mean absolute percentage error is : **20.275262089497147 %**

6) Using Mini-batch gradient with regularisation for batch size = 20
Final coefficients are:
[2723.6897250938937, 4233.213711870577, 17165.79483996975, 13584.30048250552]
Mean absolute percentage error is : **19.666044244420977 %**

Result: **In all the cases the Mean absolute percentage error is minimised after regularisation.**

Problem Statement 6d: Implement the LWR algorithm on the Housing Price data set with different tau values. Find out the tau value which will provide the best fit predictor and hence compare its results with a) , b) and c) above.

Observations :

Using Locally Weighted Linear Regression for Tau = **5e-05**
Mean absolute percentage error is : **5.40731848 %**

Using scaled batch gradient with regularisation
Final coefficients are:
[5037.585668619078, 11147.667574879839, 10378.580439168689, 22647.298983883848]
Mean absolute percentage error is : **19.92701396456417 %**

Using Stochastic gradient with regularisation
Final coefficients are:
[68851.74253913165, 154.03672870633724, 825.4095648111318, 275.1862520790242]
Mean absolute percentage error is : **22.391487213486133 %**

Using Mini-batch gradient with regularisation for batch size = 20
Final coefficients are:
[2723.6897250938937, 4233.213711870577, 17165.79483996975, 13584.30048250552]
Mean absolute percentage error is : **19.666044244420977 %**

Result : We can easily observe that Using Locally Weighted Linear Regression the mean absolute percentage error is considerably minimised as compared to any other methods mentioned for this particular use case.