
Assignment 2

V Nidhin Krishnan¹

Abstract

Assignment 2 submission for ML with Large Dataset. Implemented code for local and distributed version of L2 regularized Logistic Regression.

1. Local

For local model a mini-batch of 200 size was used. Training with mini-batch size 1 was taking too much time. Even after two days there was barely any progress for mini-batch size 1. For learning rate 0.001 was used for constant setting.

For increasing the learning rate the initial learning rate was set to $1e-06$. After every 2 epoch the learning rate was multiplied by 10. The final learning rate was 0.1.

For decreasing the learning rate the initial learning rate was set to 1.0. After every 2 epoch the learning rate was divided by 10. The final learning rate was $1e-05$.

1.1. Accuracy

Learning rate	Train(%)	Test(%)
constant	97.65	97.08
increasing	97.65	97.07
decreasing	97.66	97.10

1.2. Run time

Due to gradient computation the training time is large compared to test time.

Learning rate	Train(sec)	Test(sec)
constant	1740.87	8.73
increasing	1725.90	9.19
decreasing	1751.37	9.37

¹IISc, Bangalore. Correspondence to: V Nidhin Krishnan <nidhinkrishnanv@gmail.com>.

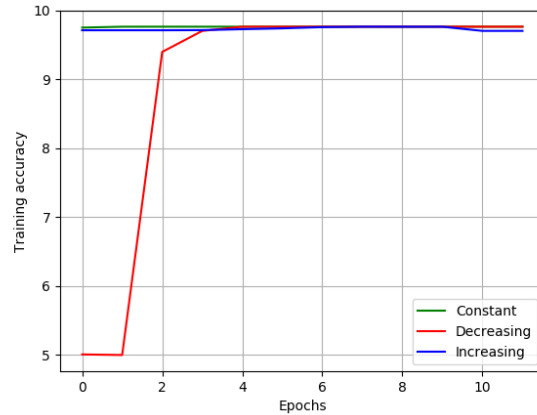


Figure 1. Run time for number of reducers from 2 to 10

1.3. Epoch vs training accuracy

From Figure 1 we can see that the constant and decreasing learning rate given better accuracy. From the experiments the increasing learning rate is unstable for SGD and gives poor accuracy on occasions. Even though constant and decreasing learning rate are giving similar accuracy. For fixing constant learning rate multiple experiments needs to be performed. So its better to choose decreasing learning rate for this experiment.

2. Distributed