**Machine Learning: Assignment 6**

# 

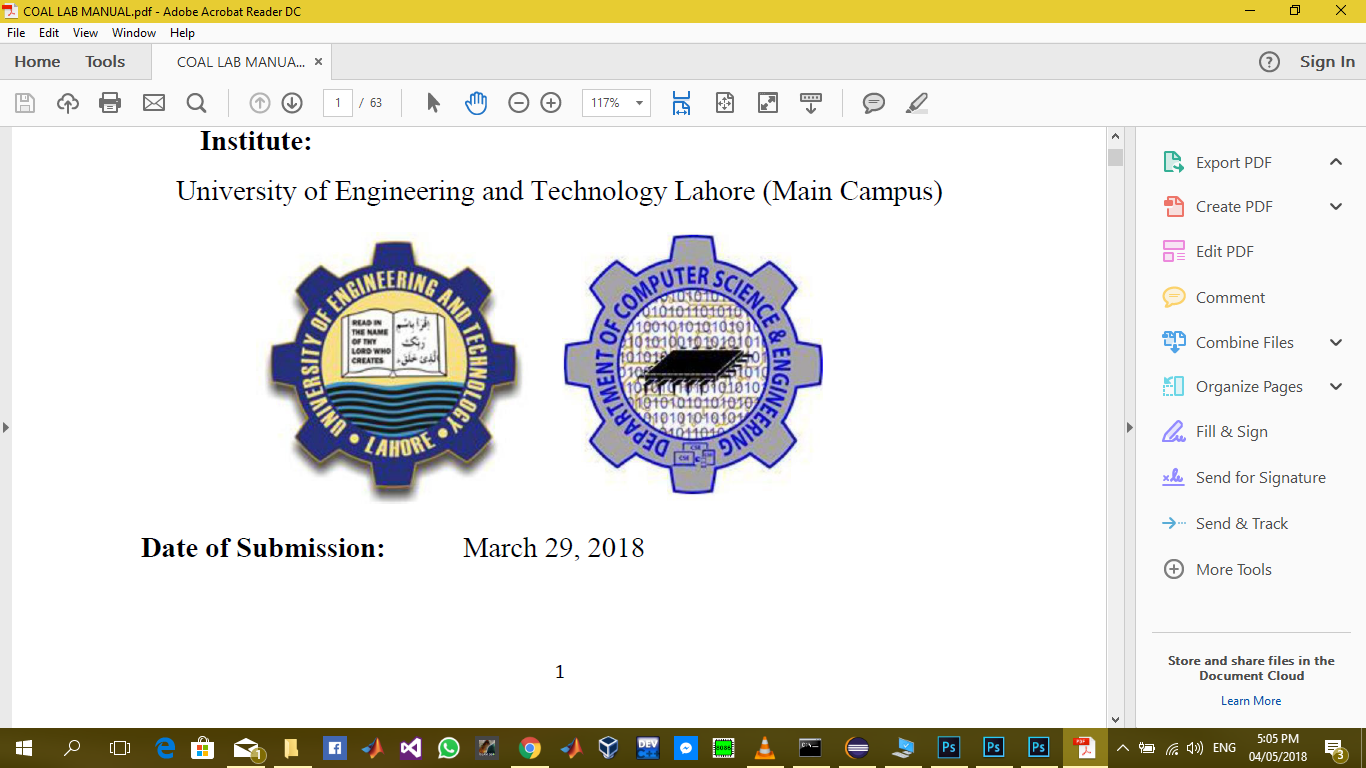
# Submitted to:

Dr. Awais Hassan

# 

# Submitted by:

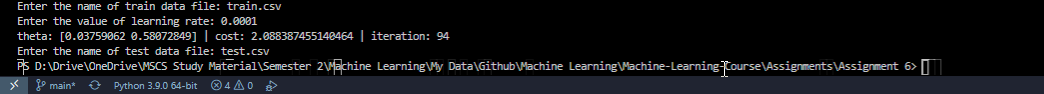
Komal Shehzadi (2020-MSCS-590)

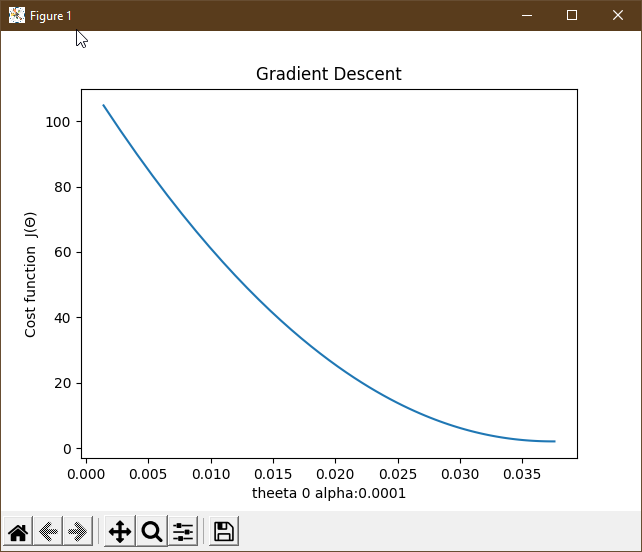


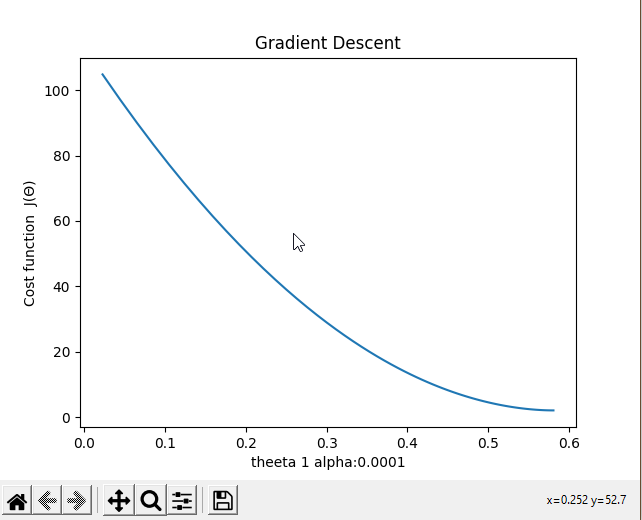
**Sample Output:**

Note\* for the sample values of alpha in problem statement, I am getting overflow error so I am using those values which are safe and giving results.

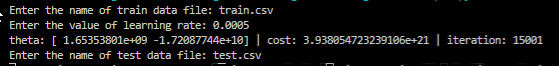
**Learning rate: 0.0001**

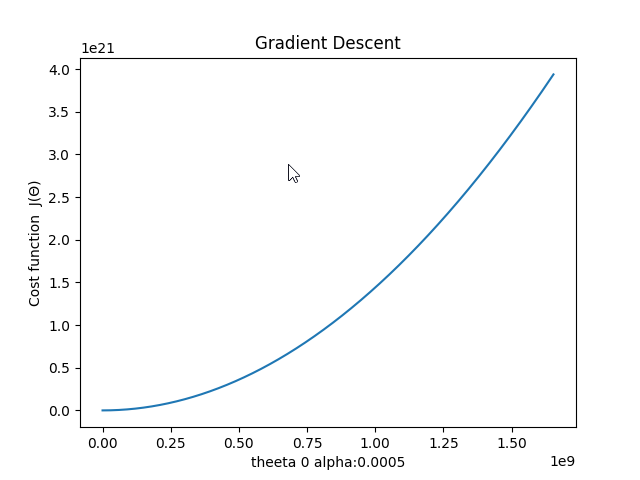


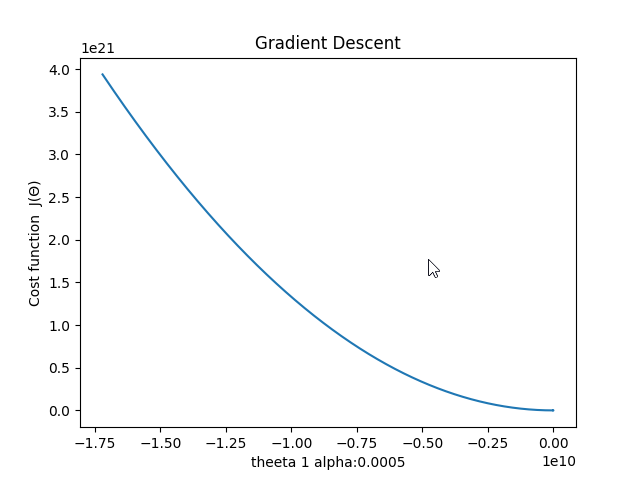




**Learning rate: 0.0005**







**Discussion:**

Learning rate is a hyper-parameter that controls how much we are adjusting the weights of our network with respect the loss gradient. The lower the value, the slower we travel along the downward slope. While this might be a good idea (using a low learning rate) in terms of making sure that we do not miss any local minima, it could also mean that we’ll be taking a long time to converge — especially if we get stuck on a plateau region. Increasing the learning rate is not safe in some cases as you can see the cost is much greater for slightly greater learning rate. When it takes bigger steps to find the local minima it is possible it will cross the region and instead fails to find the minima of function that’s why keeping low learning rate is quite useful it can be slower but chances of finding minima is much higher.