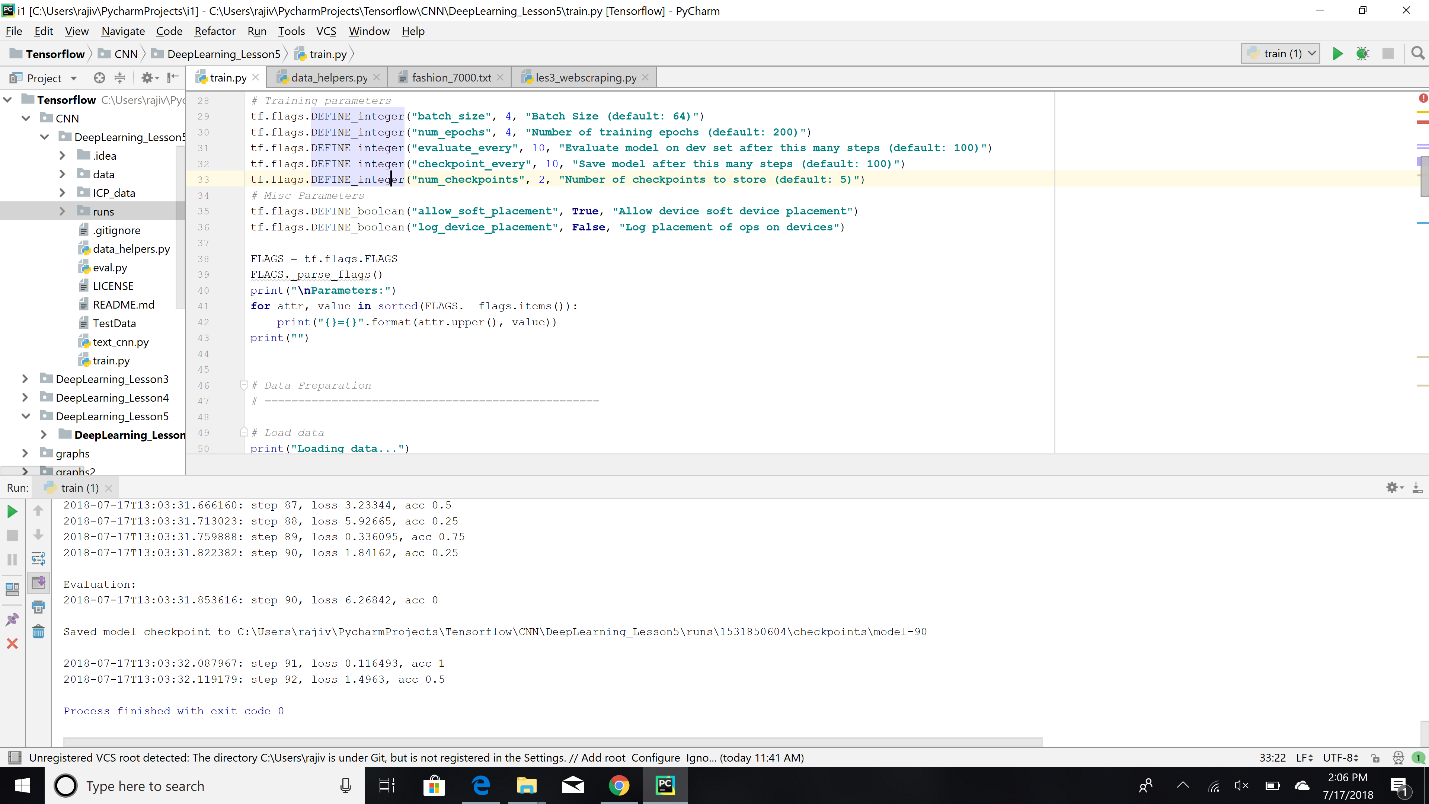
**DEEP LEARNING ICP5**

Here the vocabulary size is 1146.

1. **Using the Adam Optimizer:**

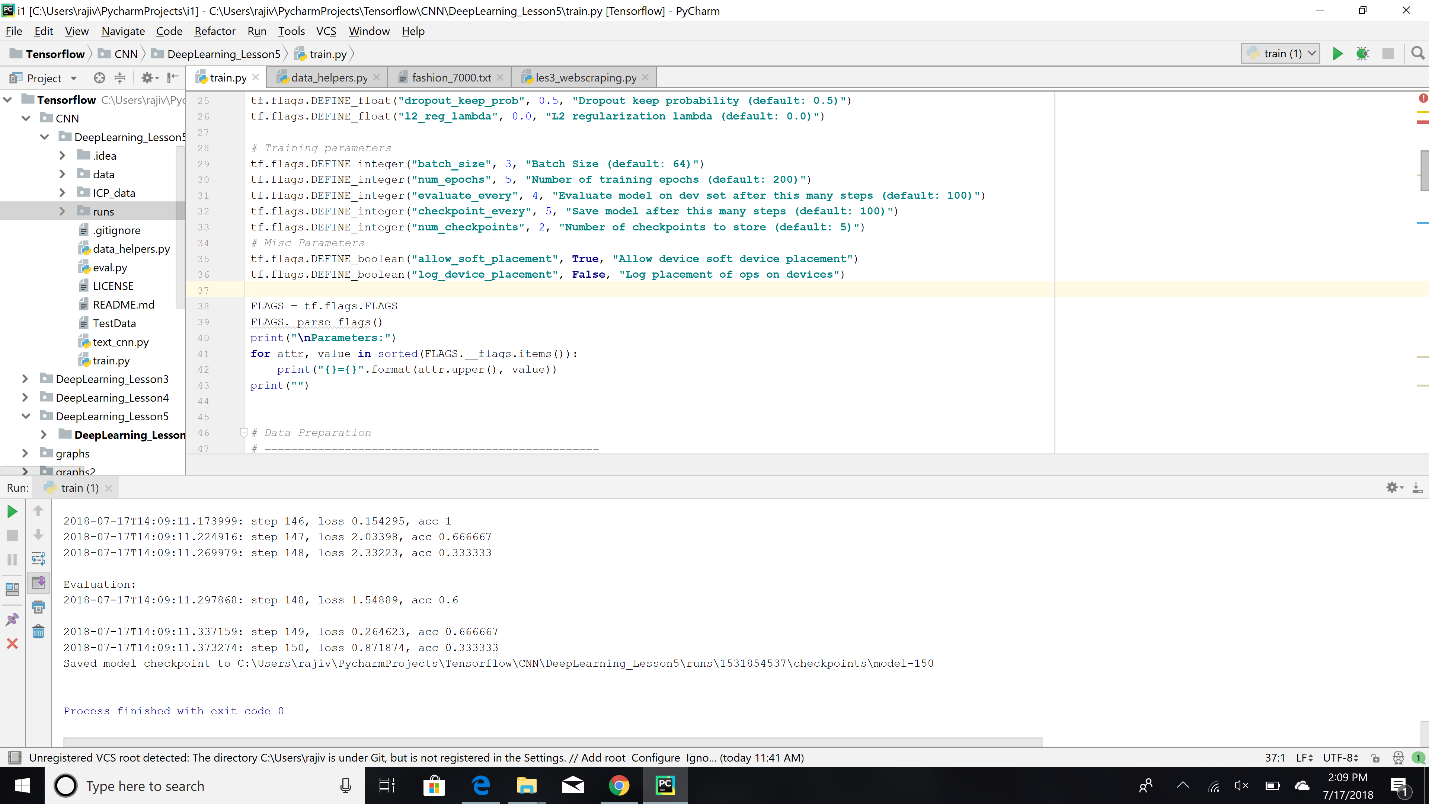
After significant change in the hyper parameters from the original stage.

We see low accuracy each time.



**2)Gradient Descent Optimizer :**

Loss is greater than Adam Optimizer



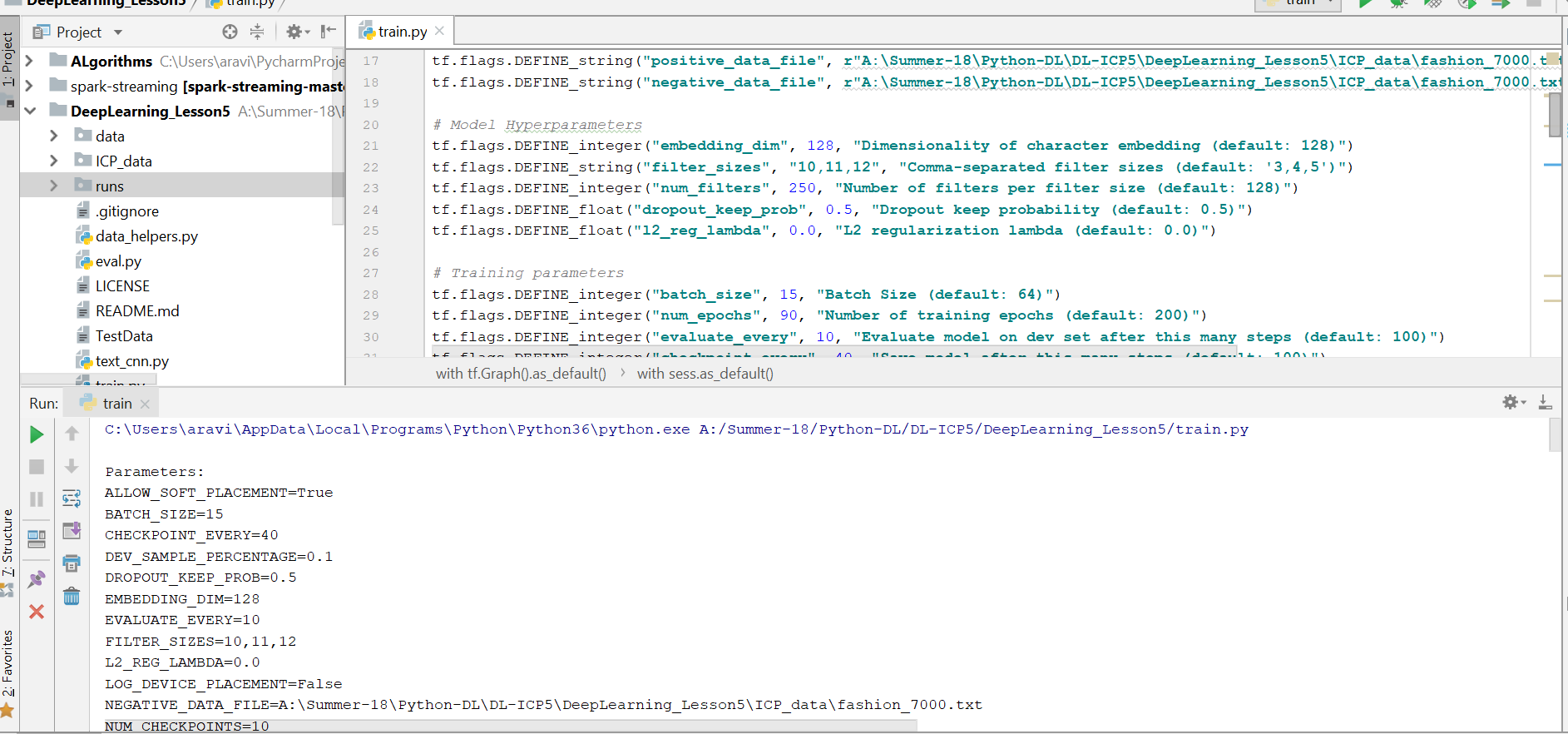
**3)Proximal Gradient Descent Optimizer:**

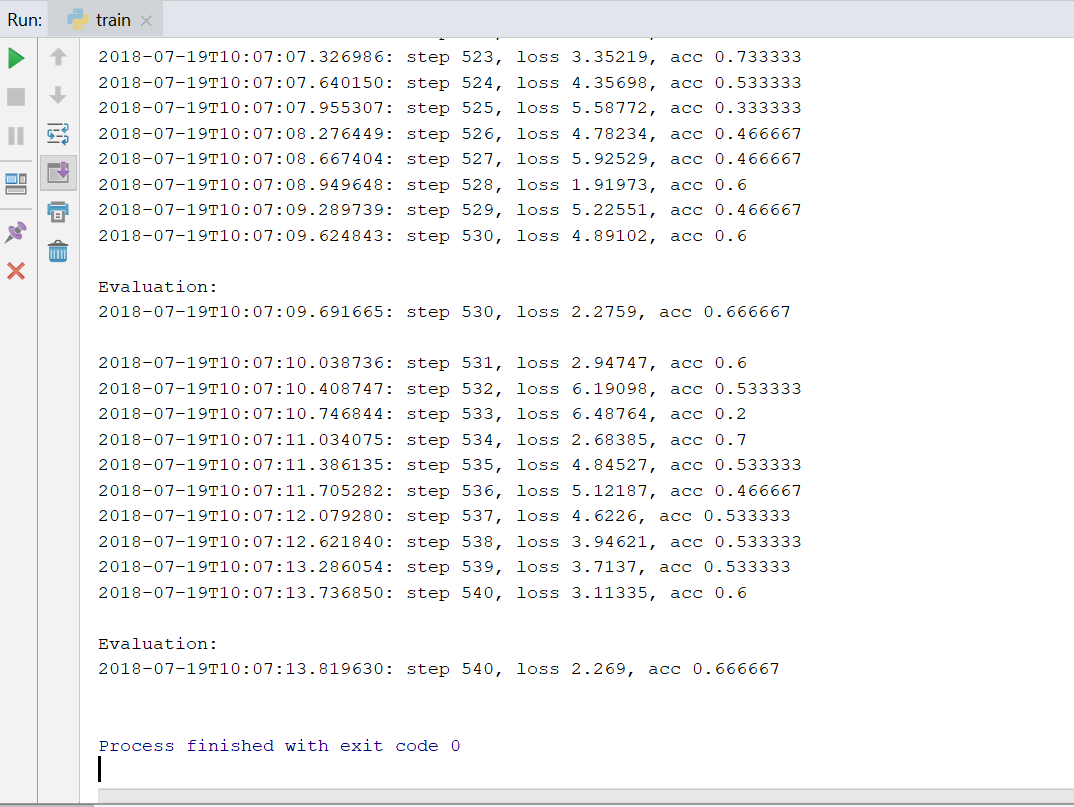
Loss is higher when the Proximal Gradient Descent Optimizer is used.



**4) Adadelta Optimizer:**

When Adadelta Optimizer is used the loss is very high when compared to the above optimizers.





**Graph :**

