

init\_scale = 0.05, learning\_rate = 1.0, max\_grad\_norm = 5.0, num\_layers = 2,  
num\_steps = 35, max\_epoch = 6,  
max\_max\_epoch = 39, keep\_prob = 0.5, lr\_decay = 0.8, batch\_size = 20,  
vocab\_size = 10000, optimizer = GradDesc, loss\_function = sequence\_loss\_by\_example

