

Answer:

1. a higher constant learning rate would give faster convergence for dataset B, but it's not really helping assuage the problem that much
2. asdf