Keith Mitchell
Discussion 1D-Shi, F.
CS 161 HW 9
704281781

3.1 Question 1
--------------------------------------------------------------------------------
Modify the script to print the accuracy of the learned model on the training set.
What is the accuracy you got? Why is it different from the test set accuracy?

->Well given that the learned model optimizes itself using the training set then the
  accuracy of this would be higher. Then the trained model would attempt a test set
  to evaluate itself on a new set of parameters. Especially when the number of iterations
  of the training is lower the difference between the two accuracies would be higher.

3.2 Question 2
--------------------------------------------------------------------------------
In Line 62 in mnist softmax.py, we loop 1000 times, taking small steps towards our
final optimized model. Try changing 1000 to 10, and reprint the accuracy on the test set.
What is the accuracy you got? Now try increasing the number of steps to 10000, and report
the accuracy in this case. Comment briefly on the results.

->Changing the loop from 1000 to 10 times:
 Accuracy changes from 0.9181 to 0.772.

->Changing the loop from 1000 times to 10000 times:
   Running time was definitely longer.
   Accuracy changes from 0.9181 to 0.9247.

->This overall is due to the ability to further train the model and increase the
  accuracy of the network. More iterations, or more use of computing power(like gpus)
  the more accuracy that can be achieved. This is done by creating an optimizer and running
  the optimization operation to update network weights.

3.3 Question 3
--------------------------------------------------------------------------------
Lines 40 and 41 in mnist softmax.py are initializing the model: W and b with zeros.
These are the values that the optimization algorithm (e.g. gradient descent) starts
with and then takes small steps towards new values for W and b that can perform better
in recognizing digits. Try initializing W and b with ones rather than zeros
(i.e. replace tf.zeros by tf.ones). What is the test set accuracy in this case?
Is it significantly different? Explain the reason behind your observation.

->Changing W and b to ones from zeroes does not change the accuracy by much.

Since the optimization algorithm takes small steps towards the proper value of W and b then they will eventually end up in a similar place with similar values and therefore similar accuracy.