CS 161: Homework 9

Jonathan Woong

804205763

Fall 2017

Discussion 1A

Sunday 3rd December, 2017

Problem 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 accuracy?

By changing mnist.test to mnist.train, the accuracy of the learned model on the training set is 91.3982%. It differs from the test set accuracy because the training set contains different images than the test set.

Problem 2. In Line 62 in mnidst_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.

After changing the number of loops to 10, the accuracy is 81.37%. This decrease in accuracy is expected, since the number of batches (subsets of the test set) used for training is reduced. After changing the number of loops to 10000, the accuracy is 92.04%. This increase in accuracy is expected. The increase of less than 1% shows that 1000 loops is sufficient to ensure reasonable accuracy, while 10000 loops may be costly for a similar level of accuracy.

Problem 3. Lines 40 and 41 in mnist_softmax.py are initializing the model: W and b with zeroes. 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 zeroes (i.e. replace tf.zeroes by tf.ones). What is the test set accuracy in this case? Is it significantly different? Explain the reason behind your observation.

The test set accuracy becomes 91.94%, which is not significantly different. Since softmax exponentiates its inputs and then normalizes them, the initial value of W and b in this case have little effect on the end result. Due to the multiplicative nature of the algorithm, setting W or b to 1 at a later step in the regression would have a larger impact.