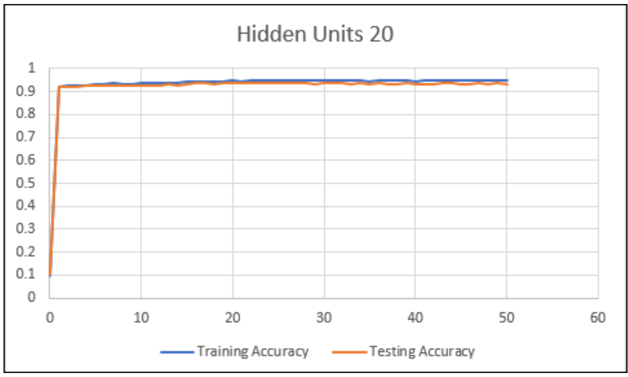
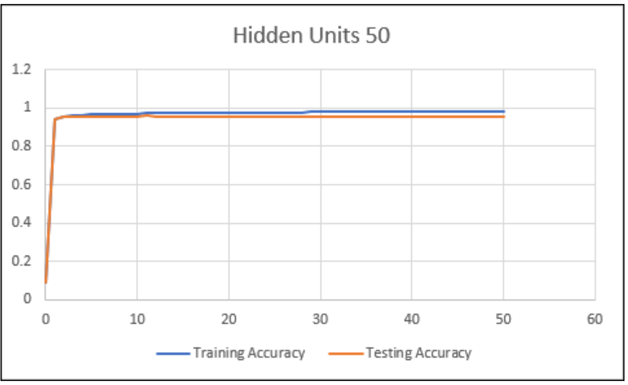
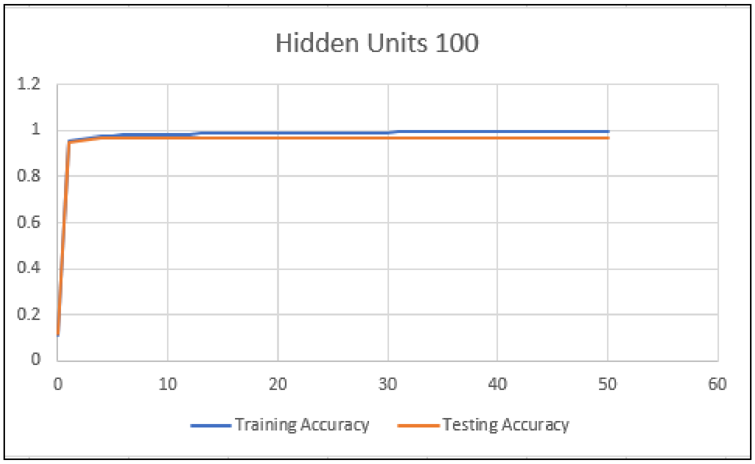
CS 545 Machine Learning
Homework 2 Neural Networks
Haomin He

Experiment 1: Vary number of hidden units.
Do experiments with n = 20, 50, and 100.
Discuss your results in a paragraph in your report. Include answers to the following questions:
(1) How does the number of hidden units affect the final accuracy on the test data?
(2) How does it affect the number of epochs needed for training to converge?
(3) Is there evidence that any of your networks has overfit to the training data? If so, what is that evidence?
(4) How do your results compare to the results obtained by your perceptron in HW 1?
Answer: The higher number of hidden units increases the final accuracy on the test data. As you can see from the graph below. Accuracy is very close to 1 with hidden units 100. The higher number of hidden units requires less number of epochs to train. In hidden units 100, the training converges after 30 epochs. Networks have overfit to the training data as there is no much fluctuations between training accuracy and testing accuracy. Results of perceptron has more variance compares to network results.

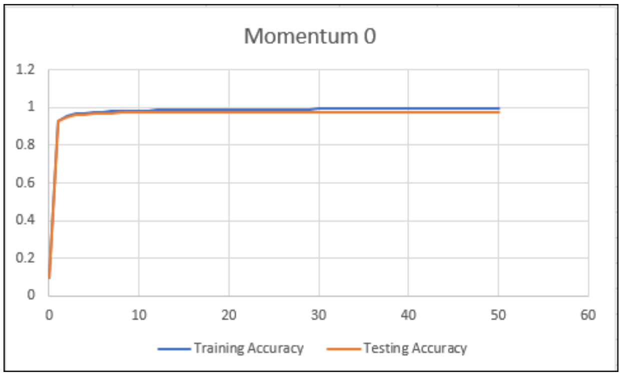


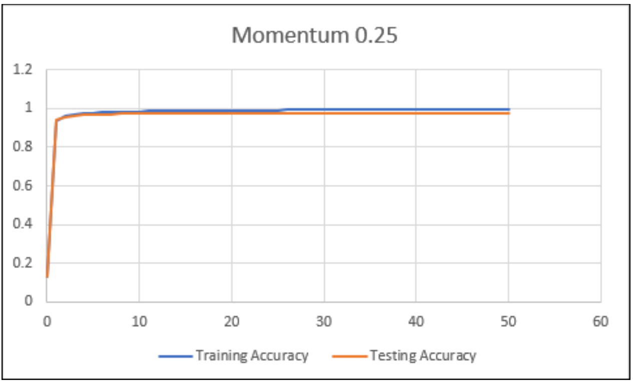


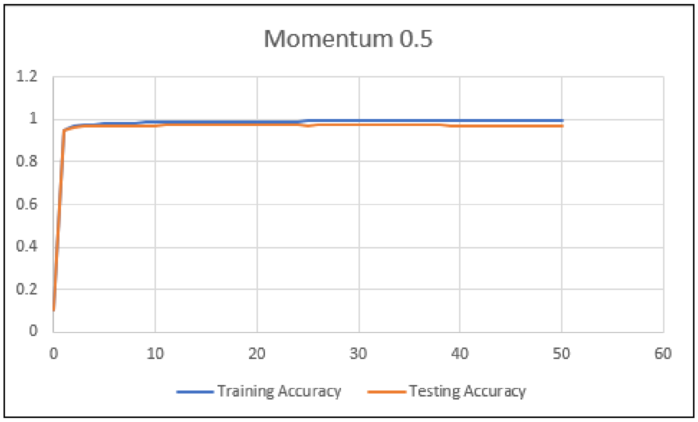


Experiment 2: Vary the momentum value. Here, fix the number of hidden units to 100, and vary the momentum value during training. Use momentum values of 0, 0.25, and 0.5.
Discuss your results in a paragraph in your report. Include answers to the following questions:
(1) How does the momentum value affect the final accuracy on the test data?
(2) How does it affect the number of epochs needed for training to converge?
(3) Again, is there evidence that any of your networks has overfit to the training data? If so, what is that evidence?

Answer: The smaller momentum value makes the final accuracy on the test data closer to 1. The smaller momentum value requires less number of epochs to train. In momentum 0, the training converges after 20 epochs. Networks have overfit to the training data as there is no much fluctuations between training accuracy and testing accuracy.

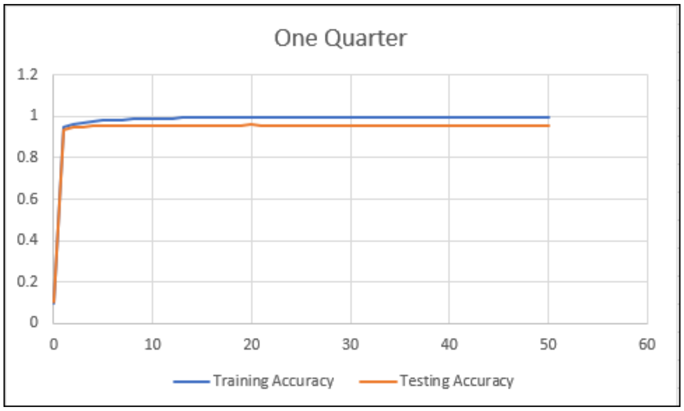


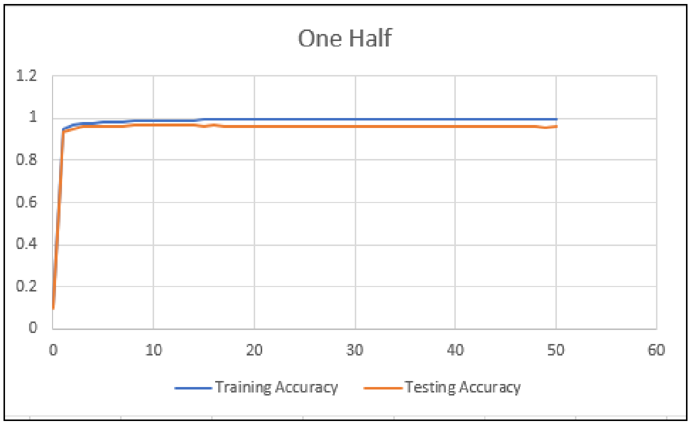




Experiment 3: Vary the number of training examples. In this experiment, fix the number of hidden units to 100 and momentum 0.9. Instead of using all of the training examples, train two networks, using respectively one quarter and one half of the training examples for training.
Discuss your results in a paragraph in your report. Include answers to the following questions:
(1) How does the size of the training data affect the final accuracy on the test data?
(2) How does it affect the number of epochs needed for training to converge?
(3) Again, is there evidence that any of your networks has overfit to the training data? If so, what is that evidence?

Answer: The bigger size of the training data makes accuracies closer to 1. In other word, prediction is more accurate with one half of the training examples. The bigger size of the training data requires less number of epochs to train. In one half of the training examples, the training converges after 18 epochs. Networks have overfit to the training data as there is no much fluctuations between training accuracy and testing accuracy.


One Quarter
Training Accuracy
Testing Accuracy


One Half
Training Accuracy
Testing Accuracy