Yang Liu
3323643

Lab Exercise 4

COMP6248
yl58n21@soton.ac.uk

1 Wider MLPs

I think the number of hidden units of MLP should not be more than the pixels of a single input image, otherwise it will start to overfit.

By adjusting the number of hidden units parameters, **Figure 1** shows the Loss of 5 different model on training and test set while training. **Figure 2** shows the Accuracy of 5 different model on training and test set while training. The divergence between the results on training set and test increases as the MLP gets wider(since hidden size is more than 784). This in some ways proves our initial thought.

The reason of this phenomenon might be that ideal models should only learn everything they can get from the input. When they try to learn something that an input is not be able to tell, they will always learn unrelated noise thus decreases the generalisation ability.

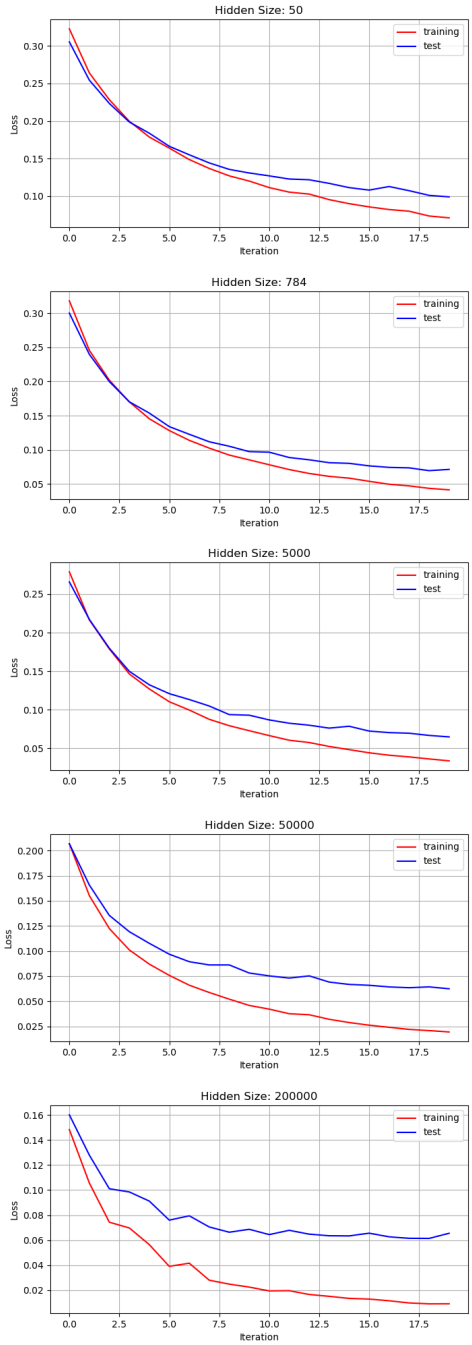

Figure 1: Loss

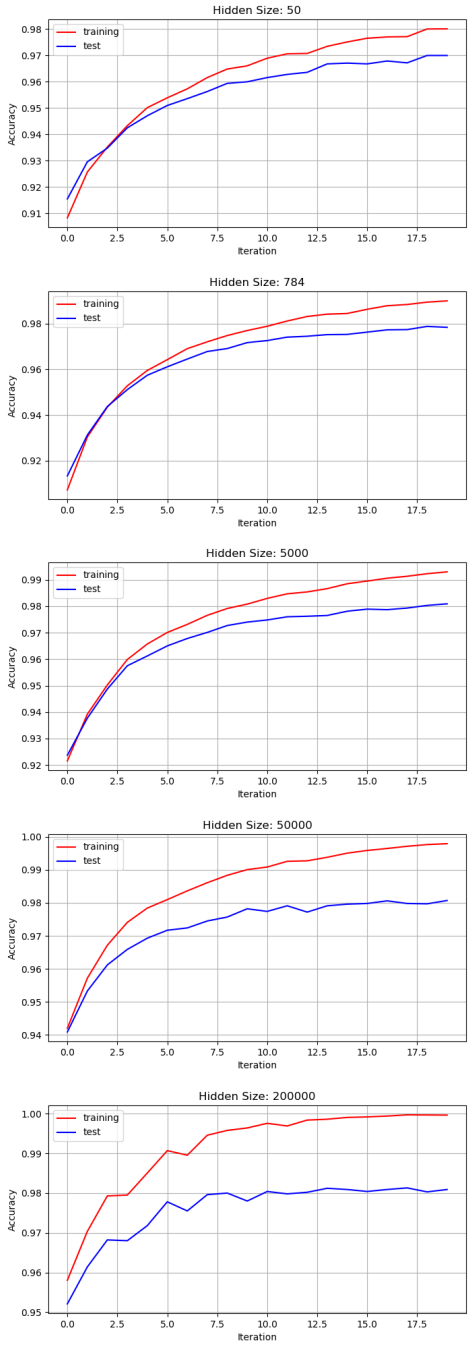

Figure 2: Accuracy