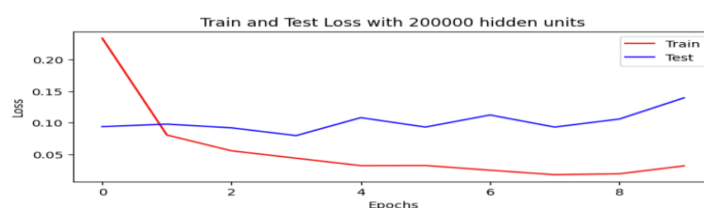
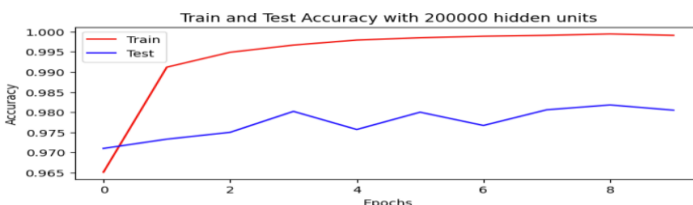
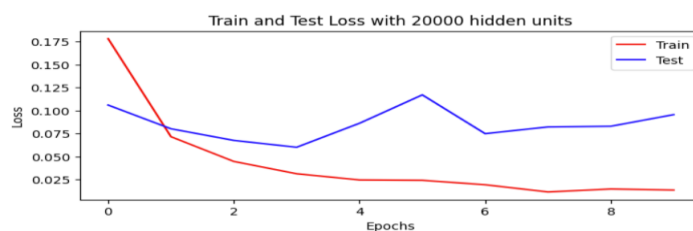
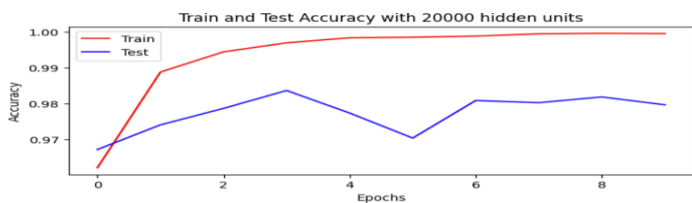
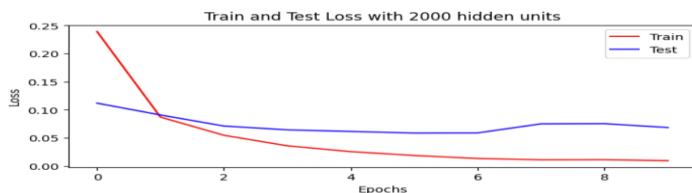
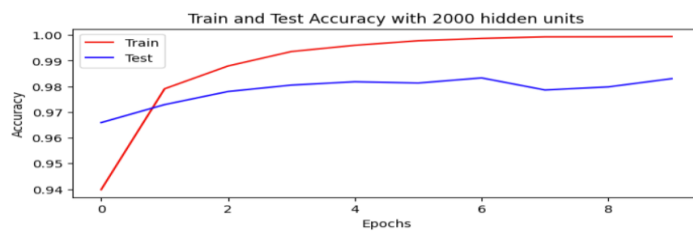


LAB 4

Student ID: 32770995



Based on the experimental evidence above, the difference between the train and test accuracy on the dataset increases when the number of hidden units are above 20000 for each epoch. The difference between train and test errors for each epoch on the dataset also increases when the number of hidden units are 20000 and above. The training accuracy reaches maximum at around 1.0 when the number of hidden units is 2000 neurons. In this case, the difference between train and test accuracy is not very large based on the plot above. Hence, when the number of hidden units was increased from 2000 to 20000 and 200000, the train and test accuracy as well as the train and test loss start to diverge higher. As the number of hidden units increases, the model can capture the underlying structure of the data better as the model complexity also increases. However, at some point, the model starts to overfit the train dataset and is not able to generalize well to unseen dataset. So while the train accuracies remains high and the train loss remains low for models with hidden units 20000 and 200000, the test accuracies and the test errors starts to diverge greatly as model overfit the training set data which may have slightly different structure than the test set data. The test data was treated as validation data in this experiment.