Name - Rahul Sailesh Wadhwa

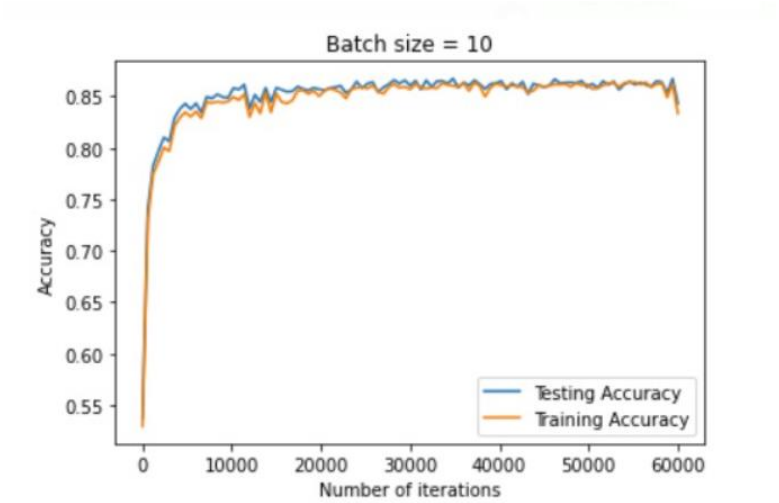
Student ID - 862309846

Problem Set 2 Report

1. **Apply the normalization on the training and test data.**

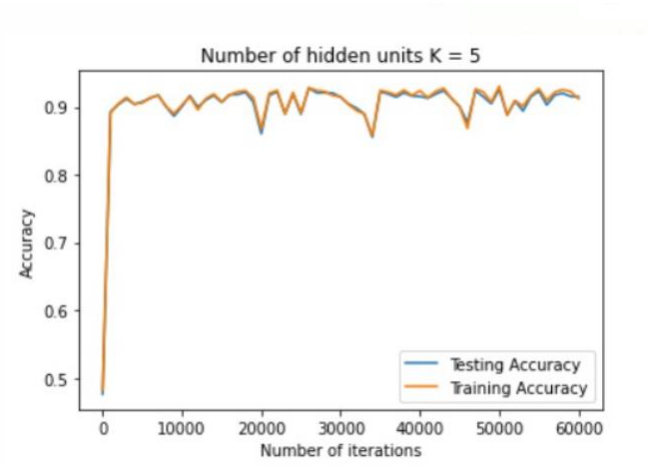
   Implemented in the notebook.

2. **As a baseline, train a linear classifier $y = v^T x$ and quadratic loss. Report its test accuracy.**

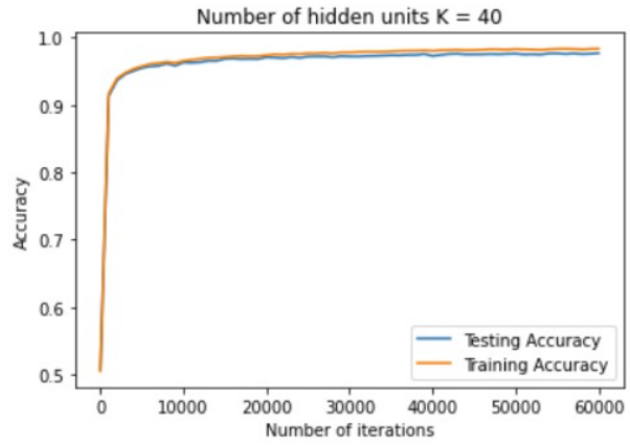


   On making use of a linear classifier with quadratic loss an accuracy of around 84% was obtained.

3. **Implement a neural network classifier with σ(x) = x i.e. f(x) = v T ReLU(W x). Use quadratic loss. Plot the progress of the test and training accuracy (y-axis) as a function of the iteration counter t (x-axis). Report the final test accuracy for the following choices (5,40,200). Comment on the role of hidden units k on the ease of optimization and accuracy.**
   - k=5

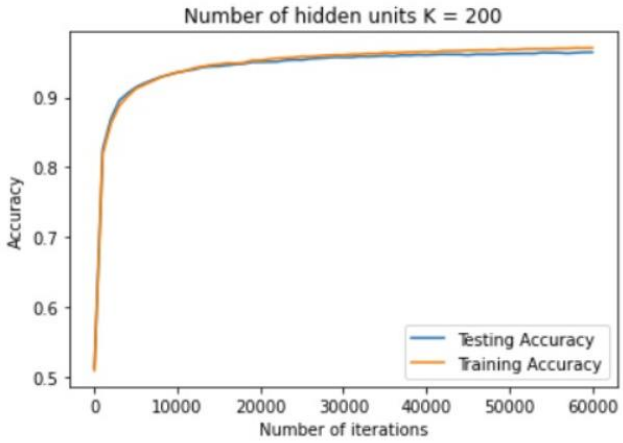


     An accuracy of around 92.3% was achieved when taking k=5 with a learning rate of 0.1
   - k=40



An accuracy of around 97.6% was achieved when taking k=40 with a learning rate of 0.01

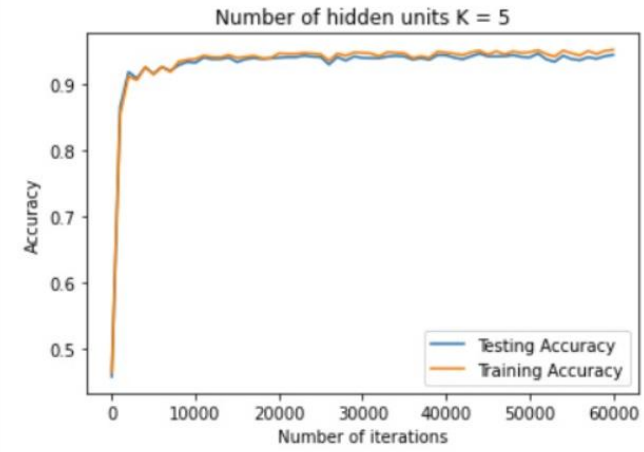
- k=200



An accuracy of around 96.5% was achieved when taking k=200 with a learning rate of 0.001

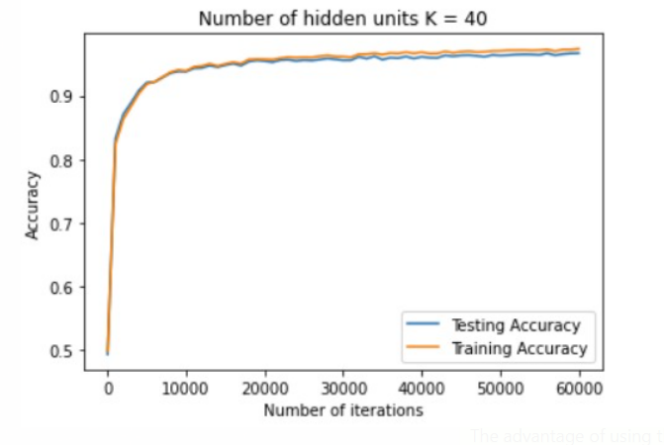
4. Implement a neural network classifier with $\sigma(x) = 1/ 1+e^{-x}$ and logistic loss. Repeat step 3.

- k=5



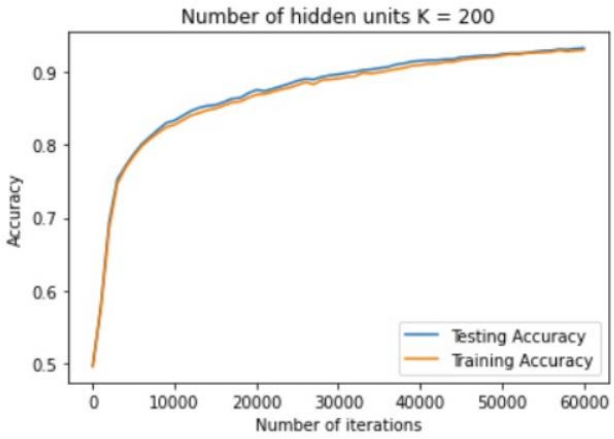
An accuracy of around 94.2% was achieved when taking k=5 with a learning rate of 0.1

- k=40



An accuracy of around 96.6% was achieved when taking k=40 with a learning rate of 0.01

- k=200



An accuracy of around 93.2% was achieved when taking k=200 with a learning rate of 0.001

5. **Comment on the difference between linear model and neural net. Comment on the differences between logistic and quadratic loss in terms of optimization and test/train accuracy.**

   The linear model achieved a relatively lower testing accuracy of around 85% as compared to that of the neural net which had an accuracy of above 90% even in cases where the number of hidden neurons was relatively small (5). The advantage of using the Neural net over the Linear classifier is that neural networks allow us to learn complex dependencies within the data and therefore they will tend to achieve a higher accuracy as compared to simple linear models.

   Using logistic loss gave us a similar but slightly lower training and testing accuracy as compared to making use of quadratic loss and the model took around 20% longer to train when making use of logistic loss.