CSE574 Introduction to Machine Learning
Programming Assignment 2 Report

Neural-Networks

Group_18: Tianyu Bao, Yikai Lu, Xianzhe Meng

Part 1

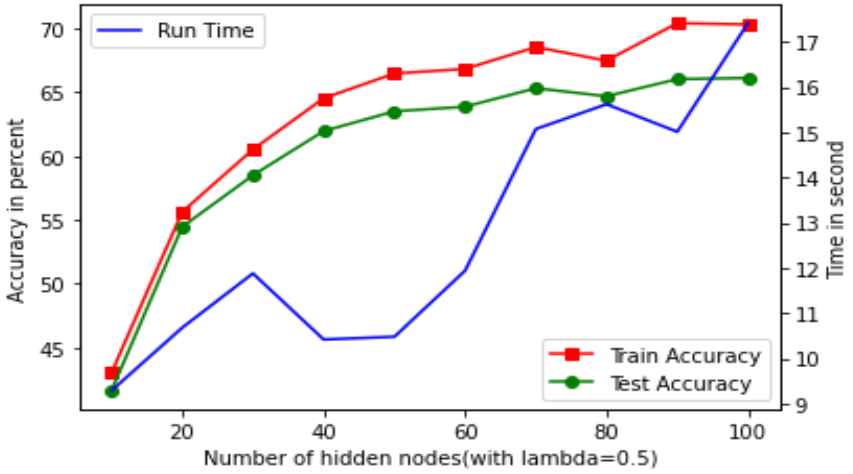
Passed all 8 unit tests in 0.031s.

Part 2

1. **Run the evaluation of the implemented neural network in the notebook - PA2-Part2.ipynb and report the training and test accuracy and the run time.**

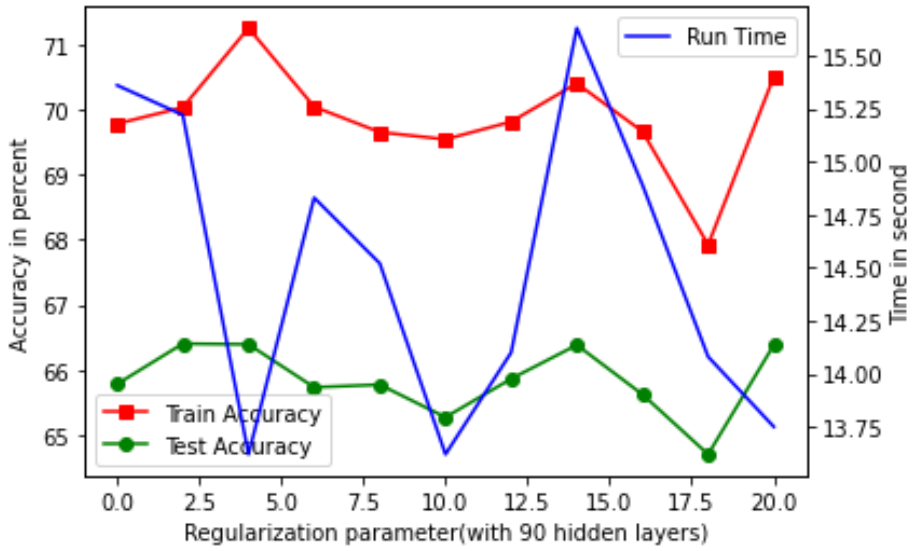
   Run time: 10.71 seconds
   Train Set Accuracy: 66.89%
   Test Set Accuracy: 63.76%

2. **Compare the performance when the number of hidden layer units (M) is increased from 10 to 100, in increments of 10. Plot the training and test accuracies and training time, as a function of M. Make your observations and state the optimal value of M that you would finally choose, along with the reason.**



The line chart of training and testing accuracies, running time with the different numbers of hidden nodes shown as above. It is clear that training and testing accuracies, running time are positively related to the number of hidden nodes from the figure. Therefore, the optimal value of M(number of hidden nodes) is 90, which has a good performance in training and test accuracy and an acceptable running time.

3. **For the optimal setting of M found above, rerun your analysis by modifying λ from 0 to 20, in steps of 2. Again, plot the training and test accuracies and the training time as a function of λ and make your observations. Which value of λ is optimal and why?**



From the figure above, there is not enough evidence to show that training and testing accuracies, running time have a relationship with the value of lambda. Therefore, only considering the performance in this chart, I will pick 4 as the value for optimal lambda because it provides a relatively higher accuracy rate and shorter running time.

4. **For the optimal settings for M and λ, study the performance of your model on the test data. What kind of objects does it make more mistakes on? Briefly discuss how the performance of your model can be improved further.**

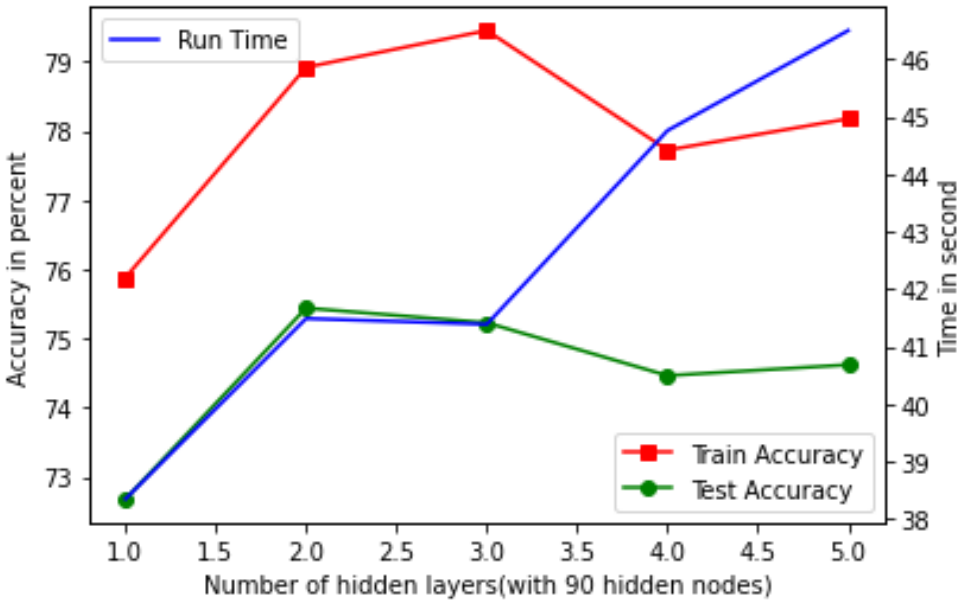
Training completed in 16.21 seconds.
Training set Accuracy: 70.57%
The objects in train set most mistakes making on: arm, mistakes rate is 24.0%
Test set Accuracy : 66.08%
The objects in test set most mistakes making on: arm, mistakes rate is 22.0%

The outcome shown above is with the optimal M = 90 and λ = 4. Training and testing set accuracy rates are improved obviously. The object making the most mistakes is the “arm”, which mistakes rates are 24.0% and 22.0% in train and test sets respectively.

For further improving the performance, we can create new models to process the objects that have high mistake rates in the original model, then we can combine these models as a new model, which may have better performance than before.

Part 3

5. **Fixing the number of units in each hidden layer (M) to the optimal value found in Part II, run the evaluation of the implemented neural network in the notebook - PA2-Part3.ipynb for different numbers of hidden layers (L), from 1 to 5. Plot the training and test accuracies and training time, as a function of L. Make your observations and state the optimal value of L that you would finally choose, along with the reason.**



The figure shows the performance of a neural network with different hidden layers. The run time is positively related to the number of hidden layers: the model takes longer with more hidden layers. However, we cannot tell that the more hidden layers could bring better performance on the accuracy of training and testing sets. Too many layers could let the model over-fitted and decrease

its accuracy on test sets. In summary, I will choose 2 hidden layers because since which the testing accuracy rate starts to going down.

6. **Using the optimal M and L from the previous part, compare the performance of the model (in terms of training and testing accuracies and the training time) for different choices of the activation function (try sigmoid, tanh, and relu). Report the best choice**

Sigmoid:

Training completed in 42.76 seconds
Training set Accuracy: 78.80%
Testing set Accuracy: 75.25%

Tanh:

Training completed in 44.82 seconds.
Training set Accuracy: 79.52%
Testing set Accuracy: 75.70%

Relu:

Training completed in 41.43 seconds.
Training set Accuracy: 86.87%
Testing set Accuracy: 71.76%

Among these 3 different activation functions, relu has the shortest training time and the highest training accuracy rate. But relu did not perform well in testing sets(lowest in these three), which indicates the model could be overfitted. Therefore I prefer to choose tanh, which has the highest testing accuracy rate.