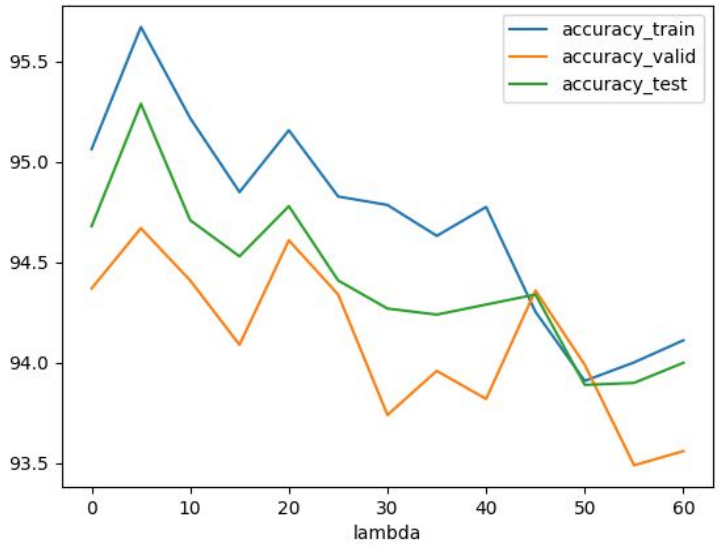
CSE 474/574 Assignment 2

Calvin Kyi
Kyle Shuttleworth
Yuanyi Yang

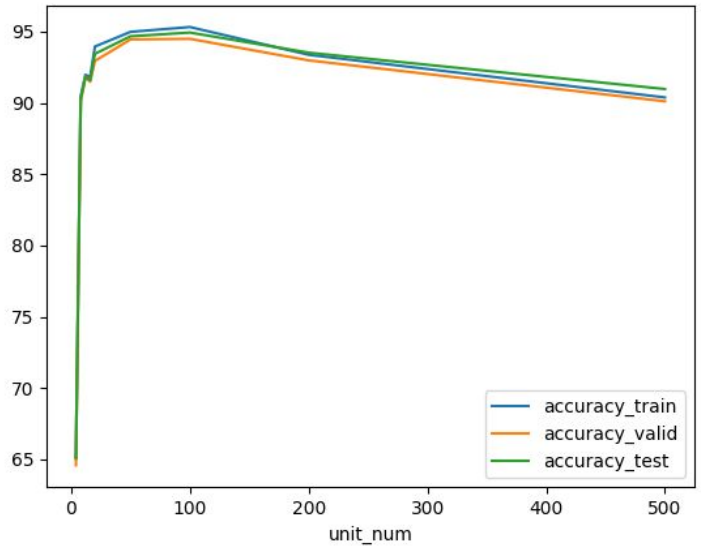
Supporting figures of how to choose the hyper-parameter for Neural Network

Impact to the accuracy from different hyper-parameter (lambda) values:



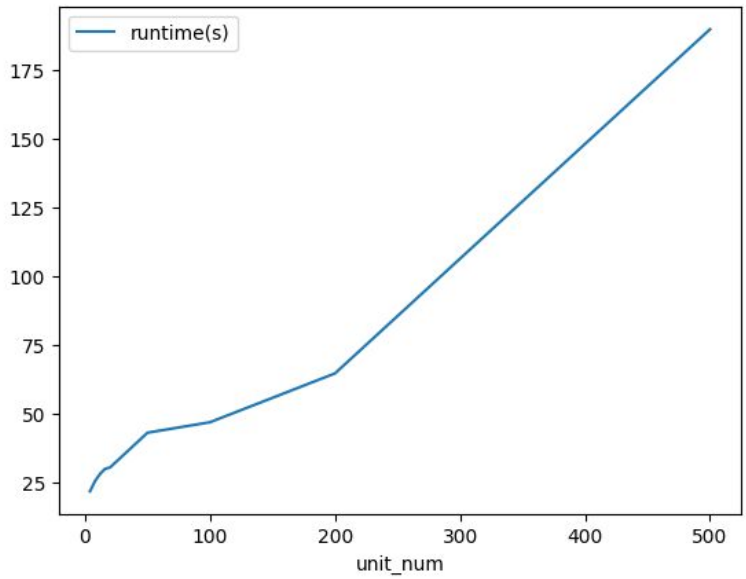
We found that the accuracy between train, valid, test data were very similar. Therefore we believe that we aren't overfitting or underfitting the dataset. The best regularization hyper-parameter (lambda) value we found for nnScript.py is 5 because it yields the highest accuracy for train, valid and test data.

Impact of accuracy from different hidden nodes in hidden layer :



As we increased the number of nodes in the hidden unit in our model the accuracy of train, valid and test data increased until we reached 100. As we increased the number of nodes in the hidden unit to over 100 the accuracy of train, valid and test data decreased. Therefore the optimal amount of nodes in the hidden unit we found for our model was 100.

Impact of training time in seconds from different number of hidden nodes in hidden layer:



The graph above shows that as the number of nodes in the hidden unit increases the training time increases. This is because if there is a higher amount of nodes in the hidden unit more calculations must be performed. We found the optimal amount of hidden nodes was 100 because training time was still fast 47 seconds and it had the highest accuracy for train, validation and test data.

Based on our results in the graphs above we found 100 nodes in the hidden unit and a lamda of 5 were optimal for our model.

Accuracy of classification method on the handwritten digits test data:

nnScript results:
Test set Accuracy with lambda = 5 and number of nodes in hidden unit = 100: 95.62%

Accuracy of classification method on the celebA dataset:

deepnnScript (deep neural network) results on the celebA dataset**:**

Three Hidden Layers
Accuracy: 79.44%
Training Time: 6 minutes 3 seconds

Five Hidden Layers
Accuracy: 81.07%
Training Time: 5 minutes 44 seconds

Seven Hidden Layers
Accuracy: 80.55%
Training Time: 5 minutes 57 seconds

facennScript (single hidden layer neural network) results on the celebA dataset:

Accuracy: 50%
Training Time: 3 minutes 6 seconds

Comparison of our neural network with a deep neural network (using tensorflow) in terms of accuracy and training time

Deep neural networks are able to create deeper representations than a single hidden layer neural network since they have multiple hidden layers which allow them to create deep representations at every layer. Because of this the deep neural networks are able to learn more and make better predictions on the input data compared to the single hidden layer neural network. However there is a tradeoff: the more hidden layers a network has the more computations that need to be performed which leads to longer training times. There can also be diminishing or even negative returns when adding hidden layers. More hidden layers doesn't necessarily mean higher accuracy after adding a certain amount of hidden layers there will be no benefit to the accuracy of the deep neural networks predictions and accuracy can actually go down if there are too many. In the Accuracy of Classification method on the celebA dataset section above we can see our conclusion on the comparison of deep neural networks and single

hidden layer neural networks take place. The deep neural network has almost twice the training time as the single hidden layer neural network but is much more accurate 81.07% for a deep neural network with 5 hidden layers compared to the single hidden layer neural network which had an accuracy of 50%. We can also see that 5 hidden layers is the optimal amount of hidden layers for the deep neural network since it has a higher accuracy than 3 and 7 hidden layers.

Results from convolutional neural network in terms of accuracy and training time
cnnScript (convolutional neural network) results:

Accuracy on Test-Set: 98.7% (9869 / 10000)
Training Time: 7 min 52 seconds