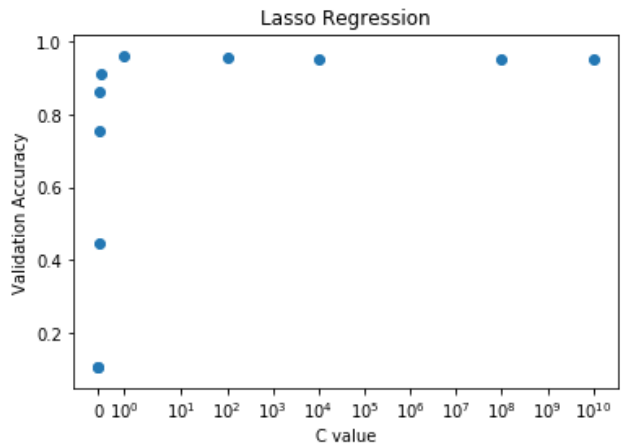
Tyler Larsen
5 April 2018
IEOR 135 (Data-X)
HW03 Write Up

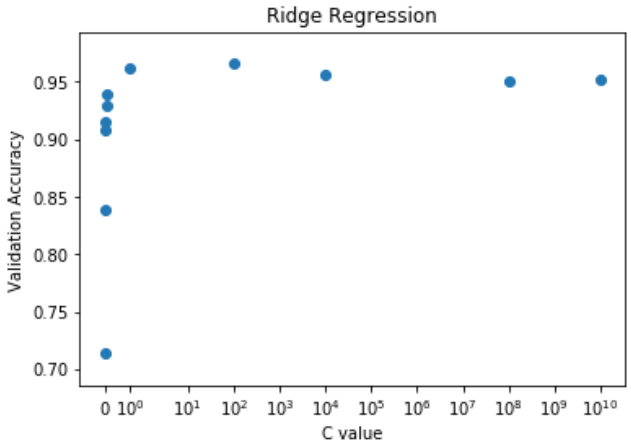
Regularization and Neural Networks

1: Lasson and Ridge Regression Models, Hyperparameter Tuning

In order to get a an idea how the validation accuracy fluctuates with a changing C value, I trained models on a wide range of C values (ranging from 1/1e-10 up to 1e5), and then plotted these C value – Validation accuracies pairs. The plots are displayed below.




```
Optimal c value:  1.0
Validation Accuracy:  0.9637921449761316
```




```
Optimal c value:  100.0
Validation Accuracy:  0.9659331917923104
```

I then took the C value that corresponding to the highest accuracy for both the Lasso and Ridge Regression models. While this wasn't a truly analytical method of optimizing C, in both cases it still produced high validation accuracies and well performing models.

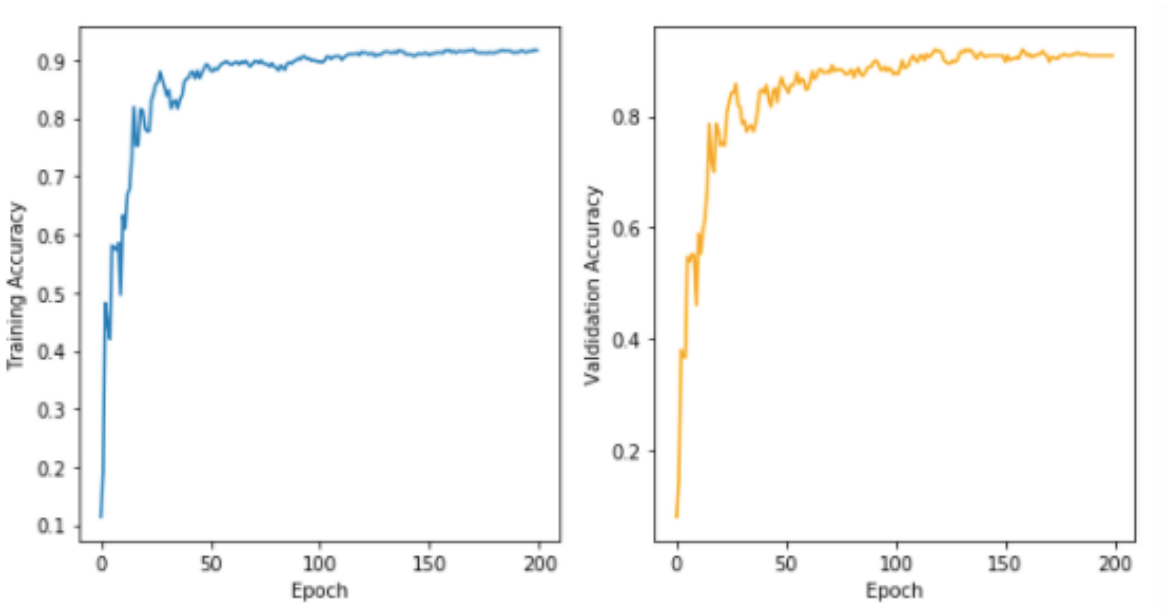
In both cases, the optimized model outperformed the non-optimized Logistic Regression model.

```
# compared to non-normalized, non-regularized, non-cross validated model:
model2 = LogisticRegression()
model2.fit(x_train*16, y_train)
y_pred = model2.predict(x_test*16)
score = accuracy_score(y_test, y_pred)
print ("Non-optimized LRG Accuracy: ", score)
```

```
Non-optimized LRG Accuracy:  0.95
```

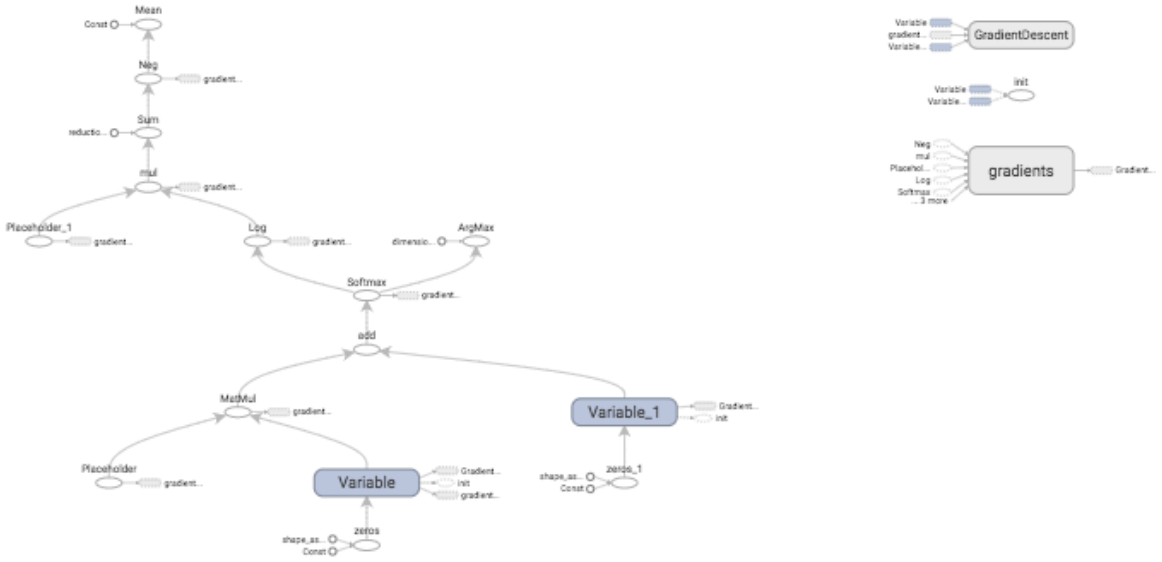
2.1: Simple Multiclass Logistic Regression Model

When training a multiclass logistic regression model using the DIGITS dataset in tensorflow, we can see that the training and validation accuracies rapidly improved over the first 100 iterations, after which time the gains were marginal, but still significant.



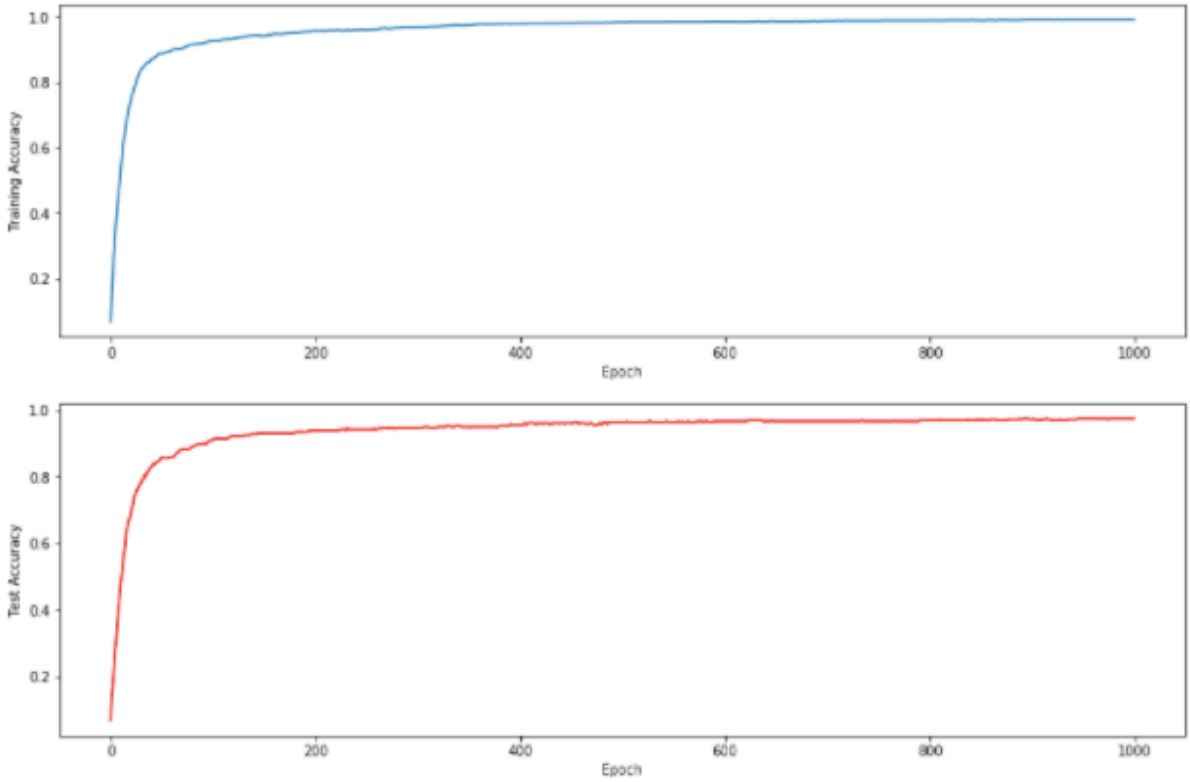

```
Training accuracy: 0.9171885873347251
Validation accuracy: 0.9083333333333333
```

Here is the final Tensorboard graph:

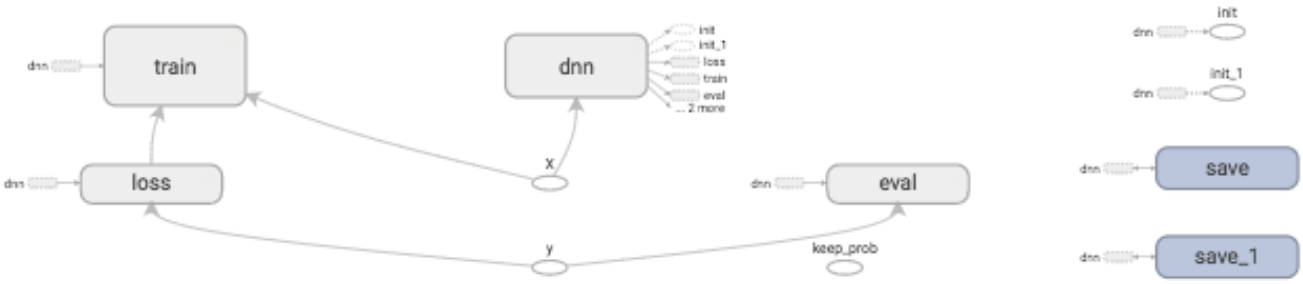


2.2: Vanilla Dense Neural Network

Similarly, the Training (top) and Test (bottom) accuracies rapidly impved in the first 20% of iteration, and gains were maginal, but still significant, after that. The final Test Accuracy of the Vanilla DNN was .975.



Here is the final Tensorboard graph for the Vanilla DNN:



The DNN Model was far more complex than the multiclass logistic training model. The DNN had several more trainable model parameters, and it was passed through three deep layers, as opposed to the only one layer in the multiclass logistic training model. As a result, the DNN out performed, leading to a .975 Test accuracy, compared to the .91 achieved by the single layer network.