Jeremy Swartwood Machine Learning 02/25/2018 Project 1 (OCR Letter recognition)

**Overall observations:**

1. The higher the learning rate, the worse the training.
2. The lower the learning rate, the longer it takes to get a good learning.
3. The lower the batch size, the better the final training outcome.
4. In general, the higher the number of nodes in the hidden layer, the better out outcome.

|  |  |
| --- | --- |
| **Best performing combination in 500 epochs: 93.8% Accuracy on Validation set**   * Learn = 0.1 * Nodes in Hidden = 256 * Batch Size = 20 | |
| C:\Users\thisisme1\Documents\CSCE415 - ML\Assignments\Project1\runs\0.1_learn_500_epochs_20_batch__confusion_matrix.png | C:\Users\thisisme1\Documents\CSCE415 - ML\Assignments\Project1\runs\Figure1.png |

|  |  |
| --- | --- |
| However, the same test above on 100 nodes in the hidden layer instead of 256, produced very close results at a lower computational and time cost. | |
| C:\Users\thisisme1\Documents\CSCE415 - ML\Assignments\Project1\runs\100 nodes in hidden\0.1_learn_500_epochs_20_batch__confusion_matrix.png | C:\Users\thisisme1\Documents\CSCE415 - ML\Assignments\Project1\Figure2.png |

**To show the run combinations, below is an image of 5 tests per learning rate, on 100 node hidden at 500 epochs.**

* Tests were run on 0.001 (blue [first 5]), 0.01 (purple), 0.1 (yellow), and 0.5(Red [last 5]).
* Each learning rate was run at batch sizes 20, 50, 100, 500, and 1000.\
* Each run produces 3 graphs:
  + Accuracy Rate vs epoch
  + Confusion matrix (Actual Vs Predicted)
  + Error Rate vs epoch

