Eric Becerril-Blas

CS 489 HW 3

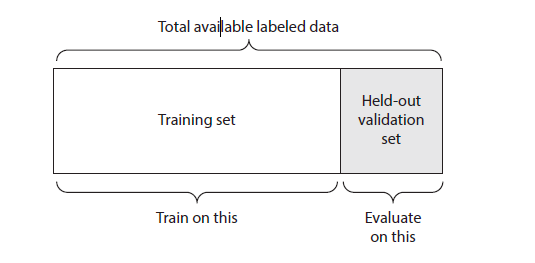
03/04/20

Dr. Mingon Kang

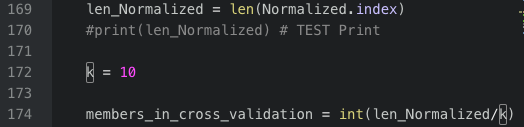
HW 3

**OVERVIEW:**

For this homework assignment I trained a linear model for classification on the 15 by 15 MNIST dataset given to us by Dr. Kang. I used 10 Fold cross validation which means I split the data up into 10 groups and used one as the test dataset while the others were training. It looks something like this:



I only used 330 out of 335 of my data points because I needed to evenly split them up into 10. However you can change this on line 172 of my code. Changing this number in my code will make the script go from a 10 fold cross validation to a K-fold cross validation. Here is line 172 of my code.



This means I had 10 Experiments in which I did this to. The tables will always change as I randomly shuffle my data. But as I write this paper, they are the following:

Experiment 1 Experiment 2

A close up of text on a black background

Description automatically generated A close up of text on a screen

Description automatically generated

Experiment 3 Experiment 4

A close up of text on a black background

Description automatically generated A close up of text on a black background

Description automatically generated

Experiment 5 Experiment 6

A close up of text on a black background

Description automatically generated A screenshot of a cell phone

Description automatically generated

Experiment 7 Experiment 8

A close up of text on a screen

Description automatically generated A picture containing text

Description automatically generated

Experiment 9 Experiment 10

A close up of text on a screen

Description automatically generated A close up of text on a screen

Description automatically generated

And the total averages for each experiment is the following:

|  |  |
| --- | --- |
| **EXPERIMENT** | **AVERAGED ACCURACIES** |
| 1 | 0.952862 |
| 2 | 0.855219 |
| 3 | 0.919192 |
| 4 | 0.966330 |
| 5 | 0.922559 |
| 6 | 0.885522 |
| 7 | 0.905724 |
| 8 | 0.915825 |
| 9 | 0.946128 |
| 10 | 0.919192 |

**HOW TO RUN**

I used Python 2.7 and utilized a virtual environment to install the packages for pandas. To run this script you do the following command on command line:



Notice how (env) is there meaning I am in a virtual environment.

Running this will yield each table shown above in the overview as well as the averaged accuracies at the end of the script.