

Replication

To replicate the results reported all you need to do is run the script and just adjust the 'project' variable on line 64 each time to change the dataset. You can also print to the csv either the average or each value from each of the 10 runs for each of the metrics. Currently the code prints each value from each of the 10 runs, to print out the average over 10 runs you just need to change those parameters to say "final_accuracy", "final_precision" etc.

To replicate the mann-whitney u test results, you will have to uncomment out the block of code at the bottom and for each metric of each dataset, you will need to paste in an array of 10 values from the baseline code for that metric of that dataset, and compare this to the solution's array of 10 values for the same metric. So for recall of Caffe, you need to get the array of 10 recall values from the caffe dataset of the baseline solution, paste it into the code over the variable "baseline_scores" and change the solution_scores parameter to recalls.