

Replication

To receive the csv results for each file and generate bar chart comparisons, run the following scripts to receive the corresponding outputs:

```
python baseline.py      (Provides caffe_baseline.csv)
python improvedv2.py    (Provides caffe_results_v2.csv)
python improvedv3.py    (Provides caffe_results_v3.csv)
python improvedv4.py    (Provides caffe_results_v4.csv)
```

Once all the csv files have been generated, use `plot.py` to select any two csv files and compare them as bar charts.

For example, to compare the baseline results against the Random Forest script, enter the following csv file names into lines 9 and 10. (Lines 22, 23 and 39 contain the labels for the graph and should be changed to reflect the files being compared.)

```
8  # read results
9  results = pd.read_csv("caffe_results_baseline.csv")
10 results_rf = pd.read_csv("caffe_results_v2.csv")
11
```

baseline.py = GaussianNB

improvedv2.py = Random Forest

improvedv3.py = Logistic Regression + Class Weighting

improvedv4.py = Logistic Regression + Multiple Training Datasets + Class Weighting