How to Run the code:

For the first datasets there is a syntax you have to follow in order to run the code:

python dtclassifier.py <c-value> <d-value> <mode>

where:

c-value : any of the c values, {300, 500, 1000, 1500, 1800}

d-value : any of the d values, {100, 1000, 5000}

mode: one of the following four methods, {dt, bagging, rf, gb}

examples:

python dtclassifier.py 300 100 dt

python dtclassifier.py 500 1000 bagging

python dtclassifier.py 1500 100 rf

python dtclassifier.py 1000 5000 gb

Alternatively, you can run 'script.py'

python script.py

This will run through every single combination of c-value, d-value, and method in one go, and write all the outputs to a file 'complete_results.txt'

DISCLAIMER:

I know that this may be misleading because the program is called ‘dtclassifier.py’. However, it is designed to run every method, I just didn’t change the name of the file after adding the other methods.

One thing I must also note is that there is an element to randomness in these experiments because I didn’t define a random state for them. So results may vary.

For the MNIST dataset:

Run the following code

python part2.py <mode>

where mode is one of the same four values listed above.

This one will write its output to a file corresponding to its method, the files follow the following format:

‘mnist_<mode>.txt’

Alternatively, you can run ‘mnist_script.py’ which will do it all at once. This however, will take a very long time.

Have a nice day :)