Assignment 3

Name : Divyanshu Talwar
Roll No : 2015028 (B.Tech. CSE)

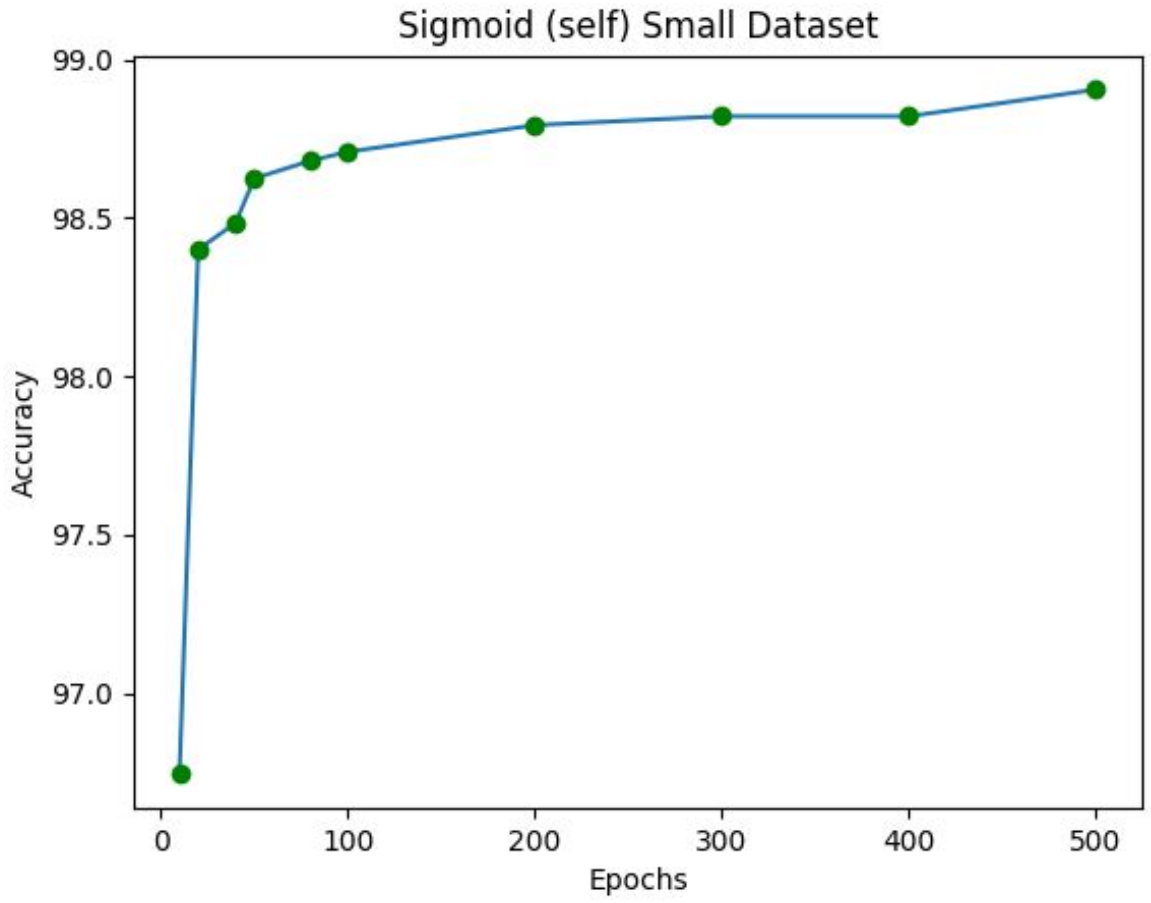
PROGRAMMING ASSIGNMENT:

Part 1: Self Implementing Forward and Back propagation:

- Sigmoid Activation Function (MNIST Subset):

Accuracy (on pickled weights) - 99.6070%

Epochs	10	20	40	50	80	100	200	300	400	500
Accuracy	96.7443	98.4002	98.4844	98.6247	98.6808	98.7089	98.7931	98.8212	98.8212	98.9054



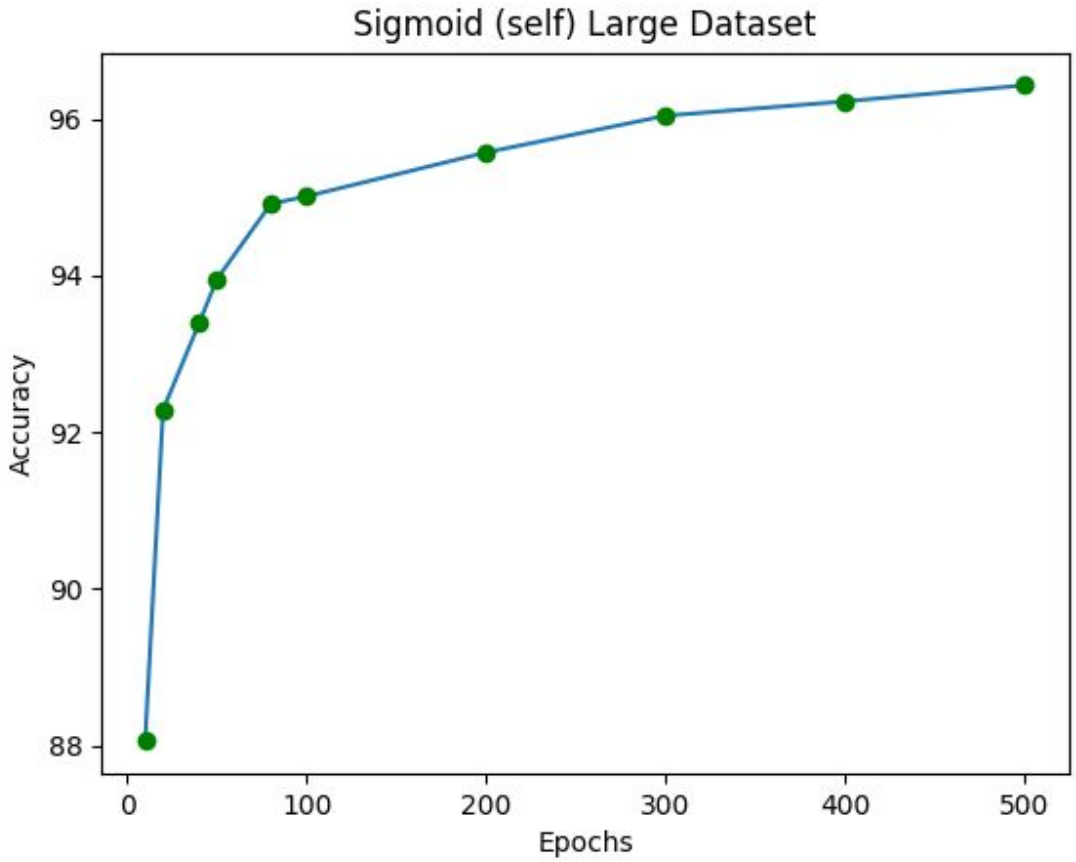
Hyper-parameters : learning rate = 0.01

Sigmoid performs pretty good on the MNIST subset.

❖ Sigmoid + Softmax (at last) Activation Function (MNIST):

Accuracy (on pickled weights) - 96.3771%

Epochs	10	20	40	50	80	100	200	300	400	500
Accuracy	88.0571	92.2857	93.3885	93.9542	94.9142	95.0114	95.5714	96.04	96.2228	96.4285



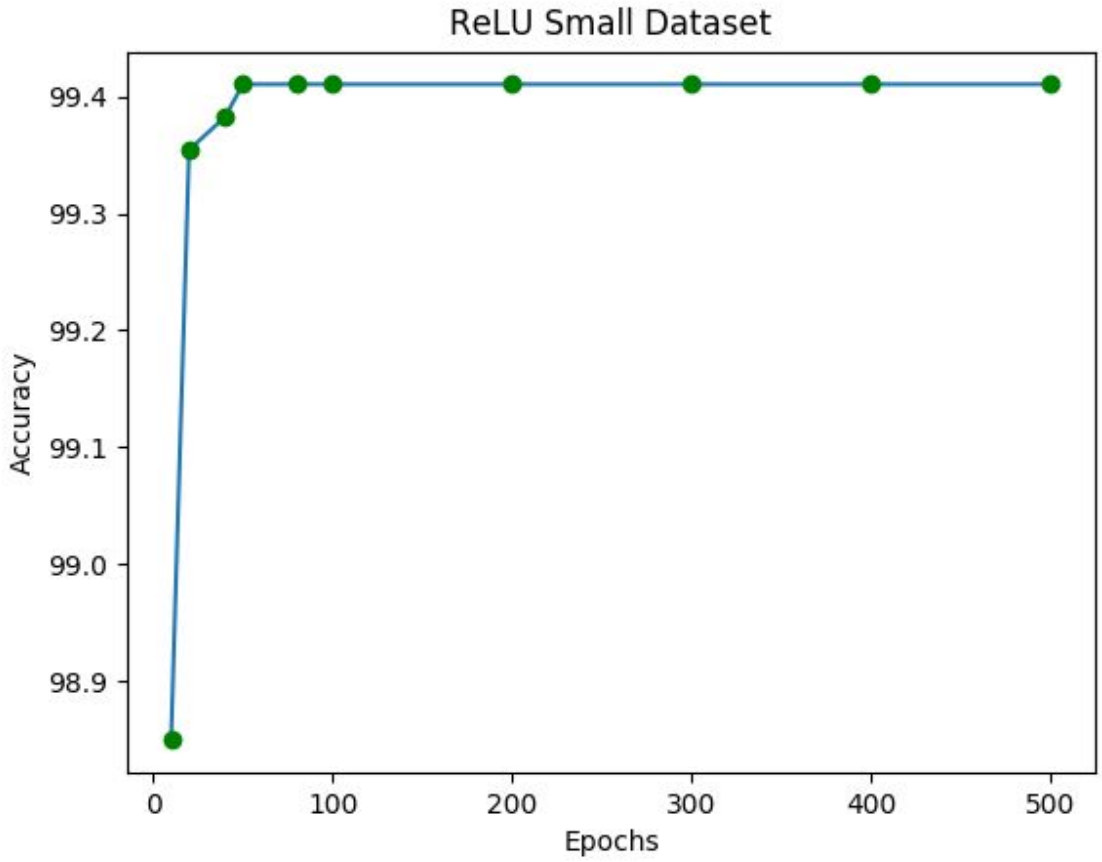
Hyper-parameters : learning rate = 0.01

Sigmoid performs decently well on the MNIST dataset.

❖ ReLU Activation Function (MNIST Subset):

Accuracy (on pickled weights) - 99.8421%

Epochs	10	20	40	50	80	100	200	300	400	500
Accuracy	98.8492	99.3544	99.3825	99.4106	99.4106	99.4106	99.4106	99.4106	99.4106	99.4106



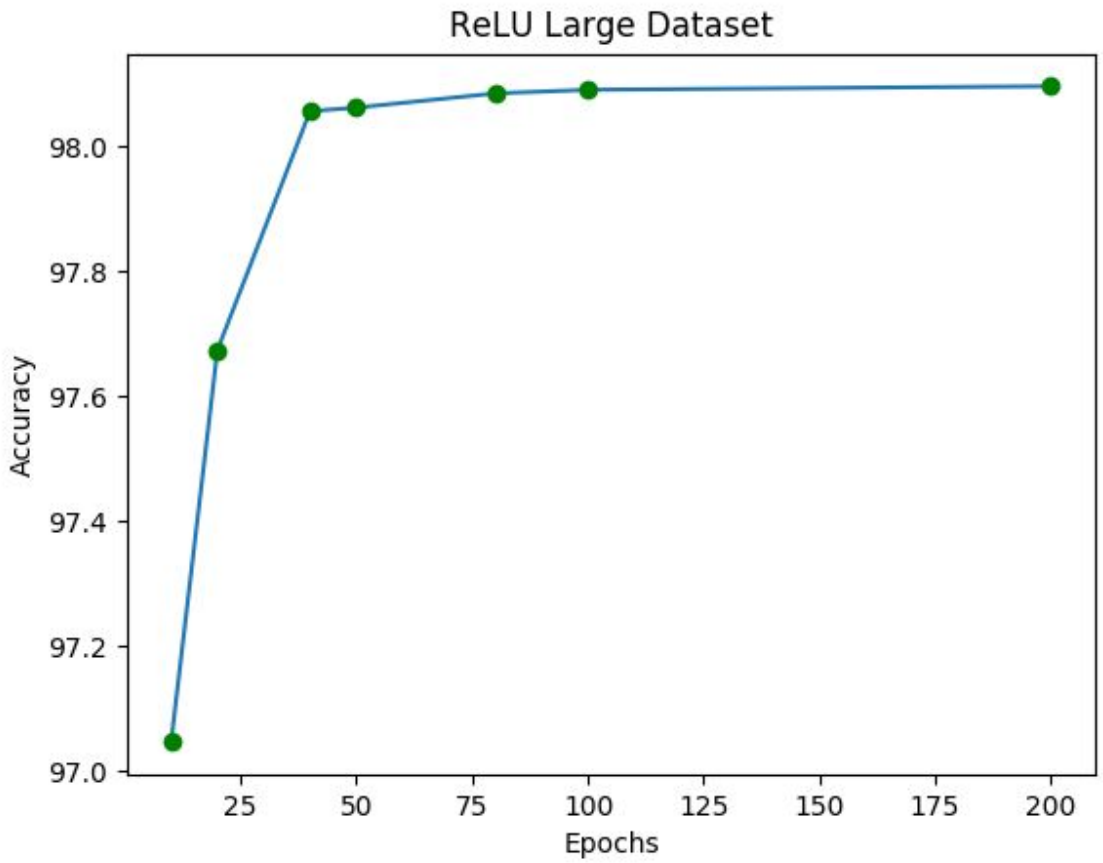
Hyper-parameters : learning rate = 0.01

ReLU performs decently well on the MNIST subset.

❖ ReLU + Softmax (at last) Activation Function (MNIST)

Accuracy (on pickled weights) - 99.5214%

Epochs	10	20	40	50	80	100	200
Accuracy	97.0457	97.6742	98.0571	98.0628	98.0857	98.0914	99.4106



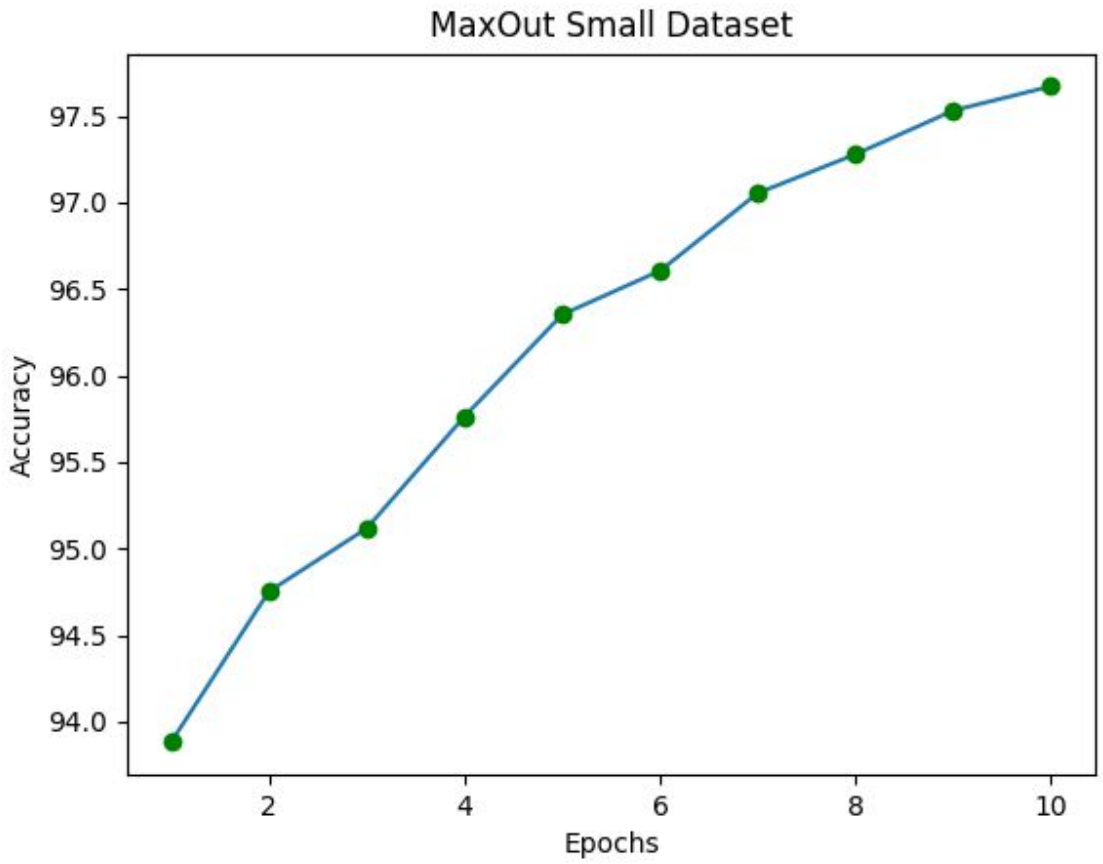
Hyper-parameters : learning rate = 0.01

ReLU performs decently well on the MNIST dataset.

❖ MaxOut Activation Function (MNIST Subset):

Accuracy (on pickled weights) - 97.5442%

Epochs	1	2	3	4	5	6	7	8	9	10
Accuracy	93.8815	94.7516	95.1164	95.7619	96.3513	96.6039	97.053	97.2775	97.5301	97.6705



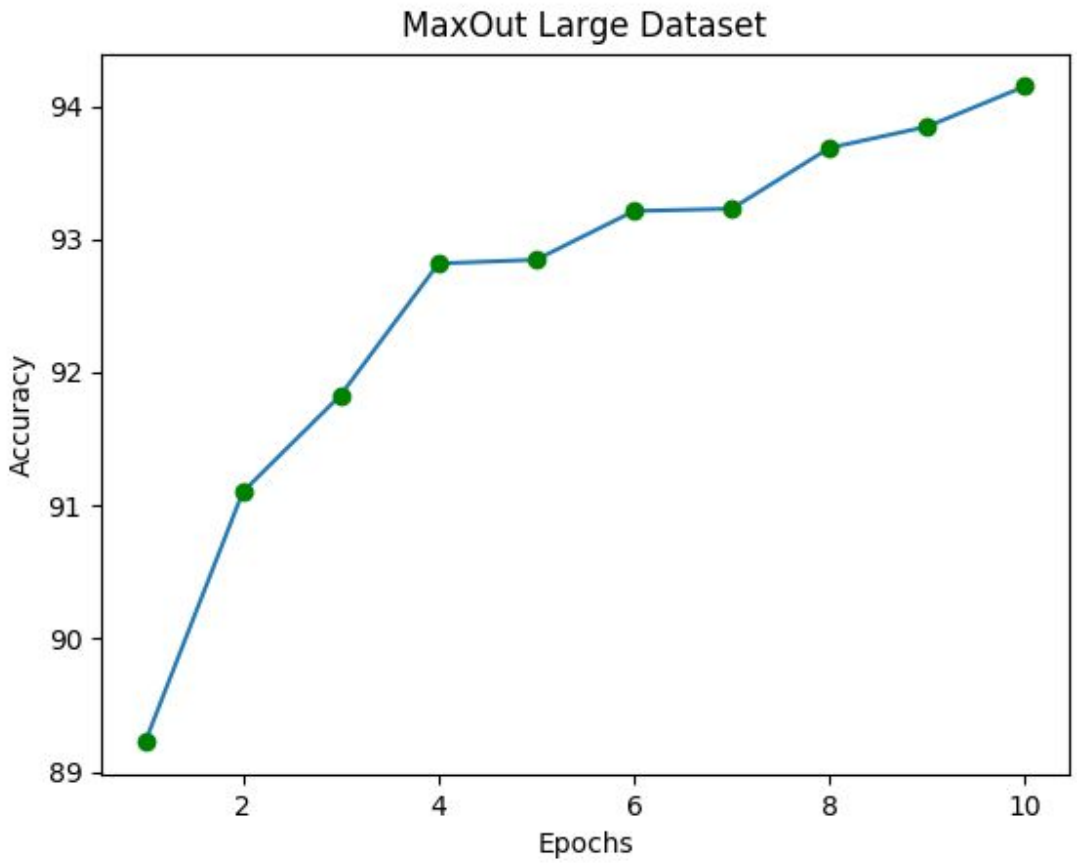
Hyper-parameters : learning rate = 0.01, no. of feature extractions = 3

MaxOut performs decently well on the MNIST subset.

❖ MaxOut + Softmax (at last) Activation Function (MNIST):

Accuracy (on pickled weights) - 94.9428%

Epochs	1	2	3	4	5	6	7	8	9	10
Accuracy	89.2228	91.1028	91.8285	92.8171	92.8457	93.2114	93.2285	93.6857	93.8457	94.1485



Hyper-parameters : learning rate = 0.01, no. of feature extractions = 3

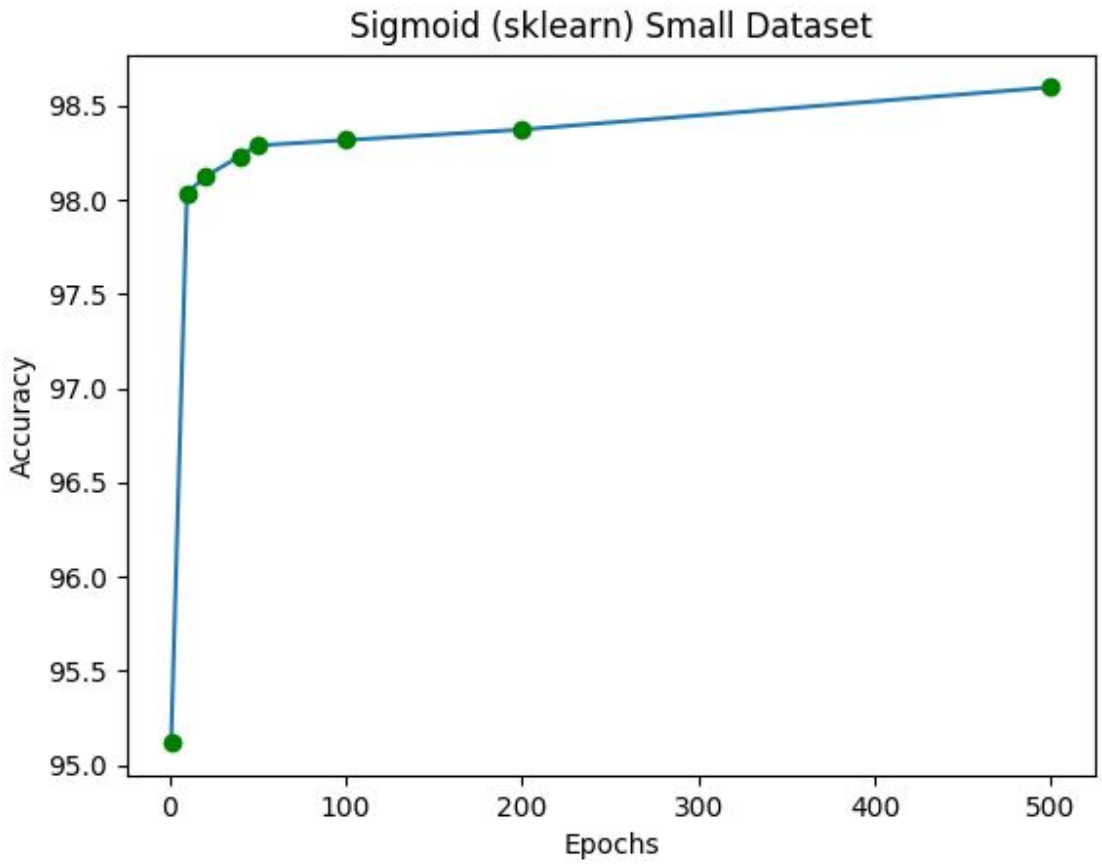
MaxOut performs decently well on the MNIST dataset.

Part 2: SKLearn's MLPClassifier :

- Sigmoid Activation Function (MNIST Subset):

Accuracy - 98.5966%

Epochs	1	10	20	40	50	100	200	500
Accuracy	95.1164	98.0353	98.1195	98.2318	98.2879	98.316	98.3721	98.5966

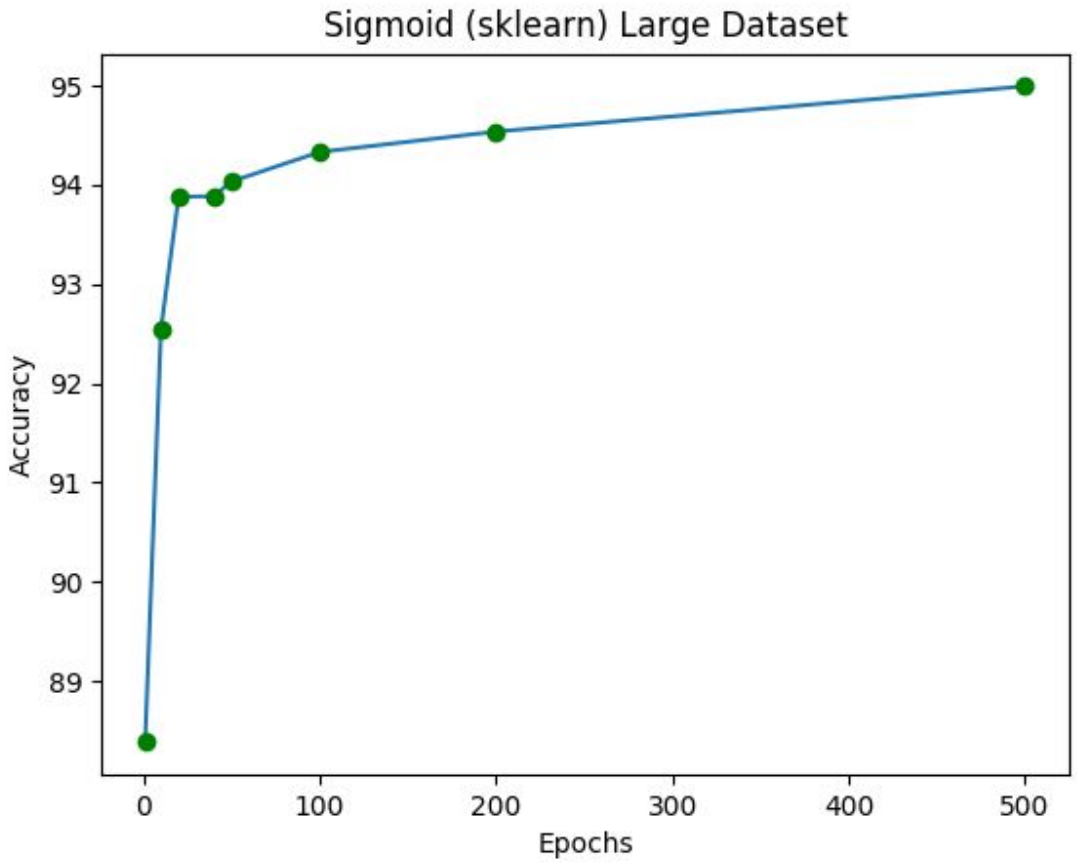


Hyper-parameters : learning rate = 0.001

❖ Sigmoid + Softmax (at last) Activation Function (MNIST):

Accuracy - 94.9942%

Epochs	1	10	20	40	50	100	200	500
Accuracy	88.3828	92.5485	93.8799	93.8857	94.0342	94.3314	94.5371	94.9942



Hyper-parameters : learning rate = 0.01, tol = 1e-04

Observation:

The Self-implemented neural networks with sigmoid activation perform better than or similar to SKlearn's MLP-Classifier. The reason behind the same might be:

- Tolerance value in MLP-Classifier as a convergence check
- Batch size in MLP-Classifier = 200 and in Self-implemented was taken to be 50 samples.