AHSANULLAH UNIVERSITY OF SCIENCE AND TECHNOLOGY

Assignment-2 (Deep Neural Networks)

Course No: CSE 4238

Course Name: Soft Computing Lab

Section: C

Semester: Fall 2020

Submitted to:

Sanjana Karim Lora

Lecturer,

Department of CSE, AUST.

Submitted by:

Name: Kanak Barua

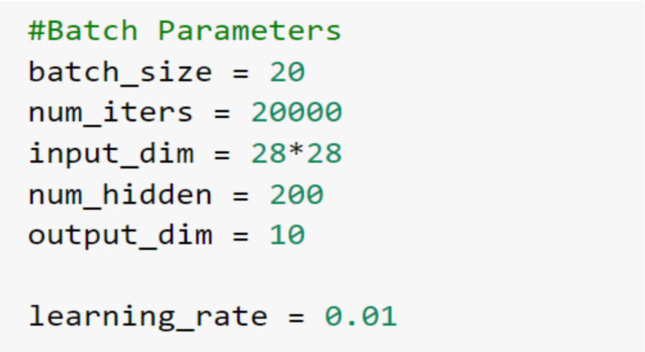
ID: 170104109

Group: C1

Email: 170104109@aust.edu

## Make necessary tables to show the comparison between two experiments for the first dataset and discuss them.

Ans: In experiment-1, we applied the base parameters as followed:

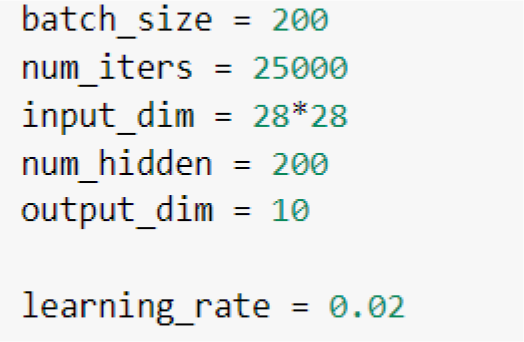


```
#Batch Parameters
batch_size = 20
num_iters = 20000
input_dim = 28*28
num_hidden = 200
output_dim = 10

learning_rate = 0.01
```

As a result, we got around 9 to 12% of accuracy during the very first attempt.

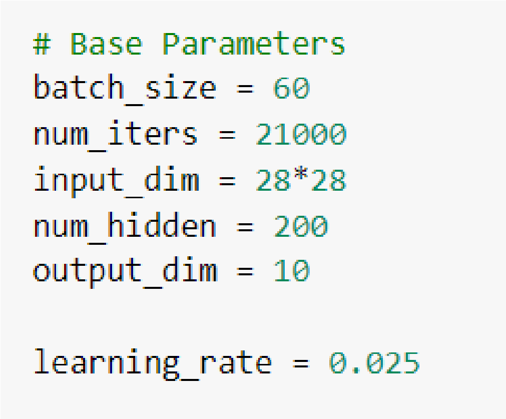
In Experiment 2, we did some fine tuning in parameters as followed:



```
batch_size = 200
num_iters = 25000
input_dim = 28*28
num_hidden = 200
output_dim = 10

learning_rate = 0.02
```

We got at max 62.54% accuracy but eventually it gave 21.85% only.

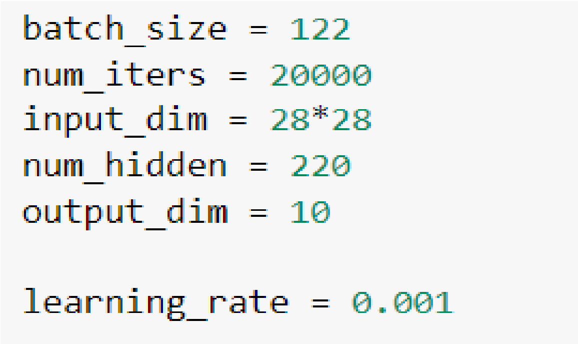


```
# Base Parameters
batch_size = 60
num_iters = 21000
input_dim = 28*28
num_hidden = 200
output_dim = 10

learning_rate = 0.025
```

Using this parameter, we got only 16.7% of accuracy.

After some fine tuning and changes in parameters, we got 83.71%. The parameters were given as followed:



```
batch_size = 122
num_iters = 20000
input_dim = 28*28
num_hidden = 220
output_dim = 10

learning_rate = 0.001
```

Hidden Layers:

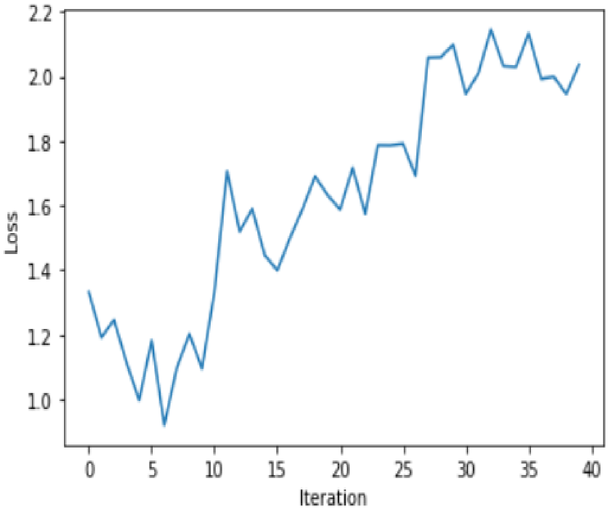
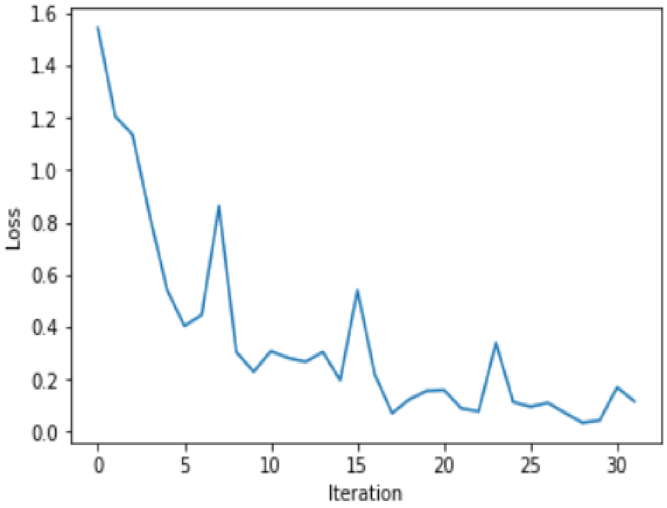
We used six hidden layers in Experiment 1 and five hidden layers in Experiment 2.

In tabular data, we showed the comparisons between experiment-1 and experiment-2.

Constrains	Exp-1	Exp-2
Batch Size	20	122
Number of hidden layers	6	4
Number of nodes in hidden layers	200	220
iteration	20000	20000
Learning rate	0.01	0.001
Optimizer	Relu(all layers)	Softmax(in 2nd layer), rest are Relu.
Epoch	205	100
Accuracy	62.31%	83.206%

## Show the Loss vs Iteration graph.

Here we showed the Loss VS Iteration Graph comparison between Experiment 1 and Experiment 2.

Experiment-1	Experiment-2
	

## Make a comparison of the results between the first dataset and second dataset by tables/graph/charts. Then discuss the results of those if they vary. Discuss the reasons behind them.

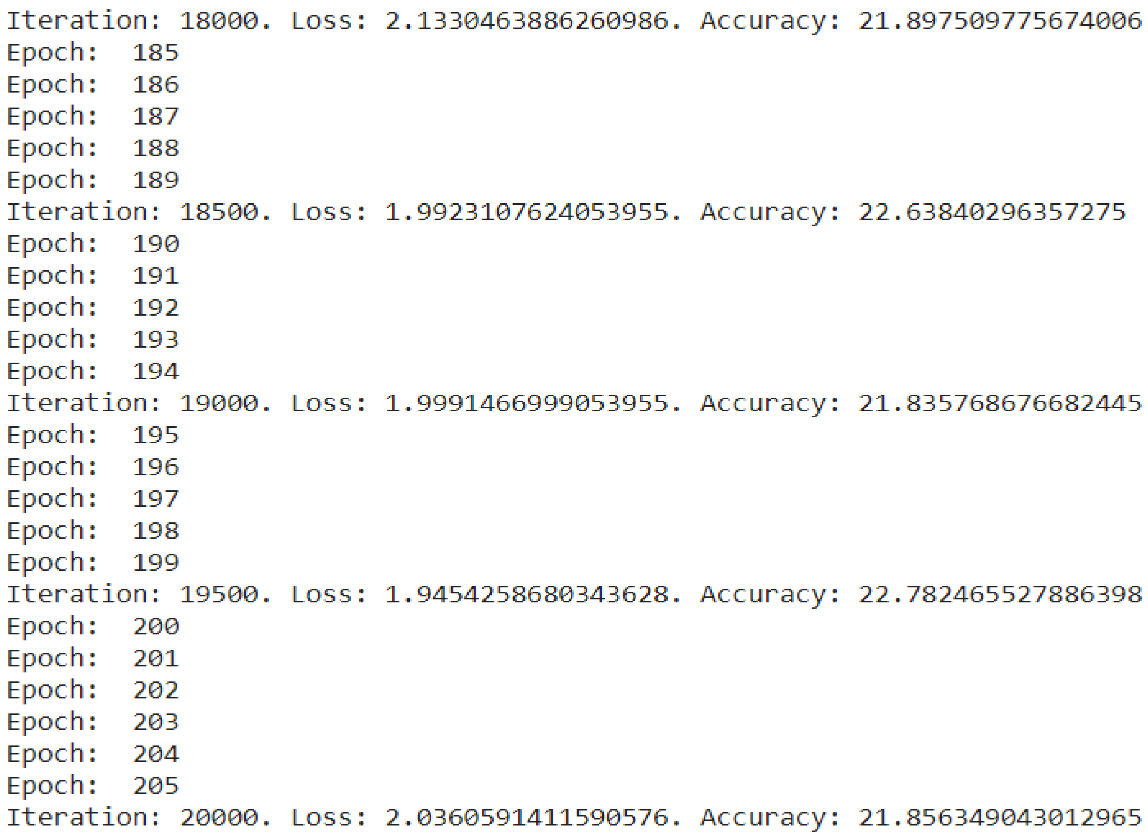
Ans:

Dataset-1:

Base parameters:

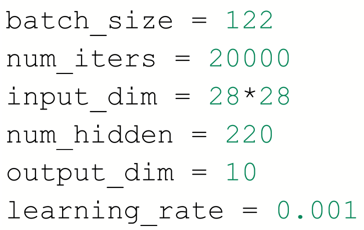
```
batch_size = 200
num_iters = 25000
input_dim = 28*28
num_hidden = 200
output_dim = 10
learning_rate = 0.02
```

Iterations and Epochs:



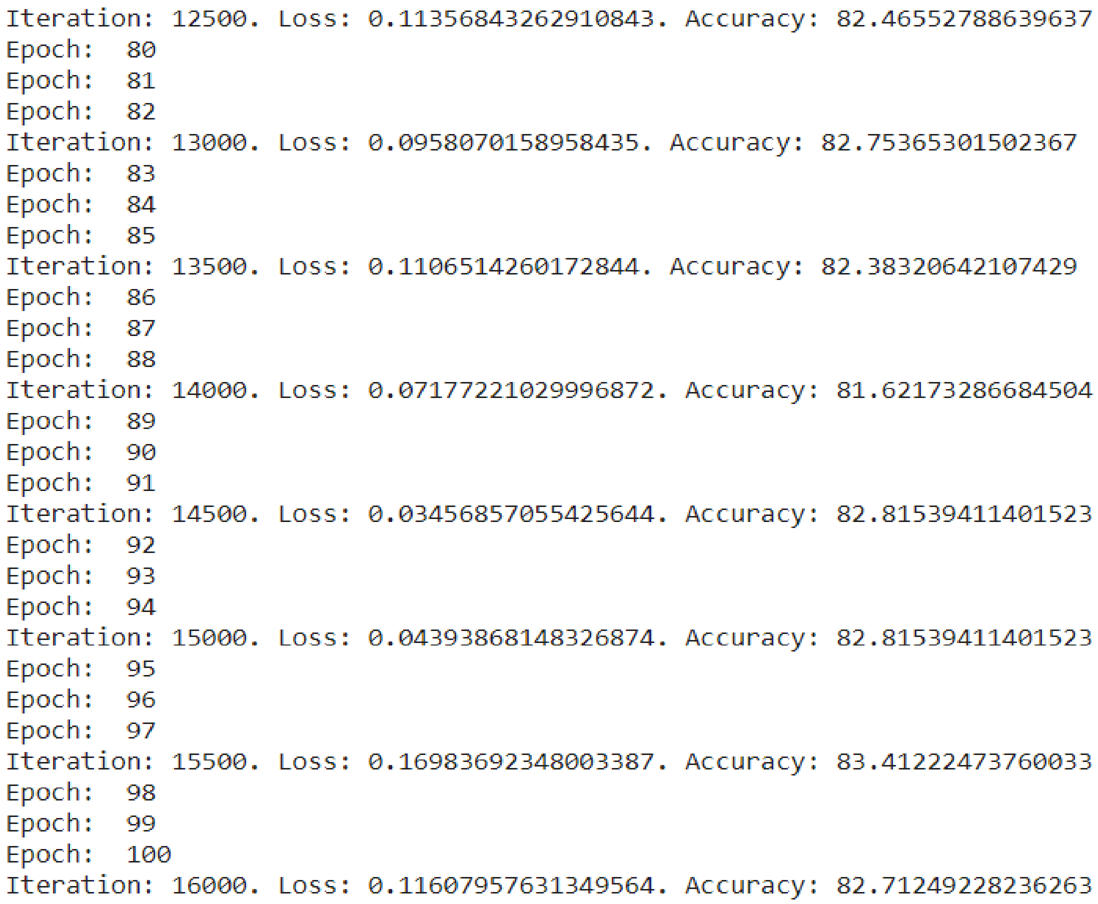
```
Iteration: 18000. Loss: 2.1330463886260986. Accuracy: 21.897509775674006
Epoch:  185
Epoch:  186
Epoch:  187
Epoch:  188
Epoch:  189
Iteration: 18500. Loss: 1.9923107624053955. Accuracy: 22.63840296357275
Epoch:  190
Epoch:  191
Epoch:  192
Epoch:  193
Epoch:  194
Iteration: 19000. Loss: 1.9991466999053955. Accuracy: 21.835768676682445
Epoch:  195
Epoch:  196
Epoch:  197
Epoch:  198
Epoch:  199
Iteration: 19500. Loss: 1.9454258680343628. Accuracy: 22.782465527886398
Epoch:  200
Epoch:  201
Epoch:  202
Epoch:  203
Epoch:  204
Epoch:  205
Iteration: 20000. Loss: 2.0360591411590576. Accuracy: 21.856349043012965
```

Modified parameters:



```
batch_size = 122
num_iters = 20000
input_dim = 28*28
num_hidden = 220
output_dim = 10
learning_rate = 0.001
```

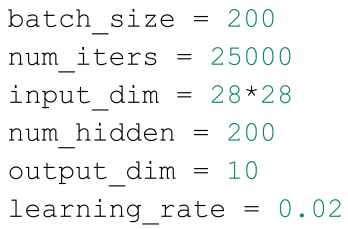
Iterations and Epochs:



```
Iteration: 12500. Loss: 0.11356843262910843. Accuracy: 82.46552788639637
Epoch:  80
Epoch:  81
Epoch:  82
Iteration: 13000. Loss: 0.0958070158958435. Accuracy: 82.75365301502367
Epoch:  83
Epoch:  84
Epoch:  85
Iteration: 13500. Loss: 0.1106514260172844. Accuracy: 82.38320642107429
Epoch:  86
Epoch:  87
Epoch:  88
Iteration: 14000. Loss: 0.07177221029996872. Accuracy: 81.62173286684504
Epoch:  89
Epoch:  90
Epoch:  91
Iteration: 14500. Loss: 0.03456857055425644. Accuracy: 82.81539411401523
Epoch:  92
Epoch:  93
Epoch:  94
Iteration: 15000. Loss: 0.04393868148326874. Accuracy: 82.81539411401523
Epoch:  95
Epoch:  96
Epoch:  97
Iteration: 15500. Loss: 0.16983692348003387. Accuracy: 83.41222473760033
Epoch:  98
Epoch:  99
Epoch:  100
Iteration: 16000. Loss: 0.11607957631349564. Accuracy: 82.71249228236263
```

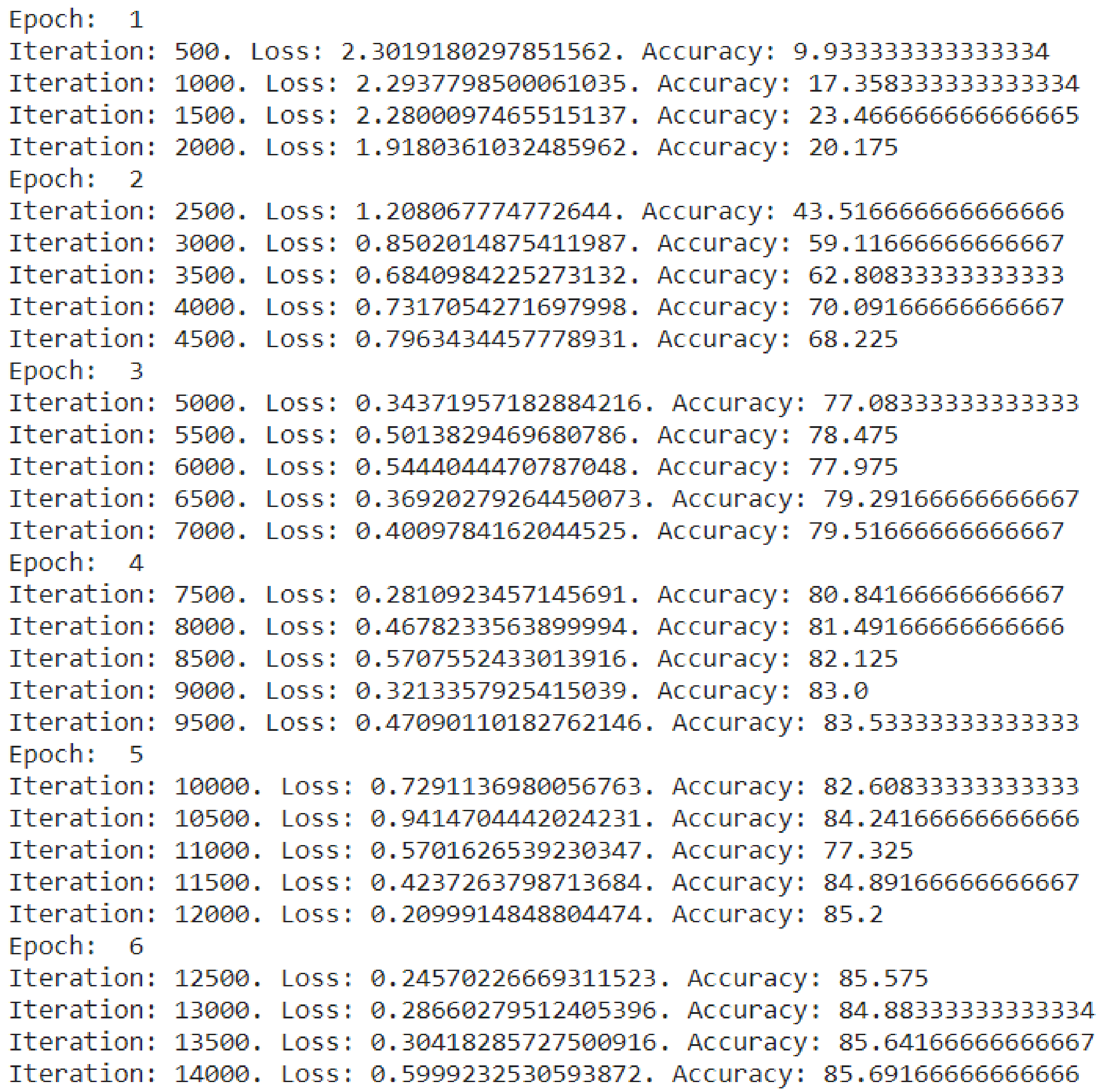
Dataset-2:

Base parameters:



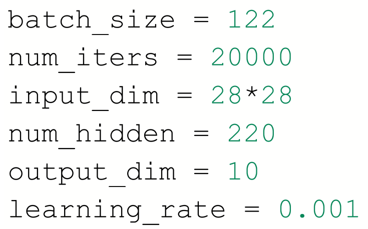
```
batch_size = 200
num_iters = 25000
input_dim = 28*28
num_hidden = 200
output_dim = 10
learning_rate = 0.02
```

Iterations and Epochs:



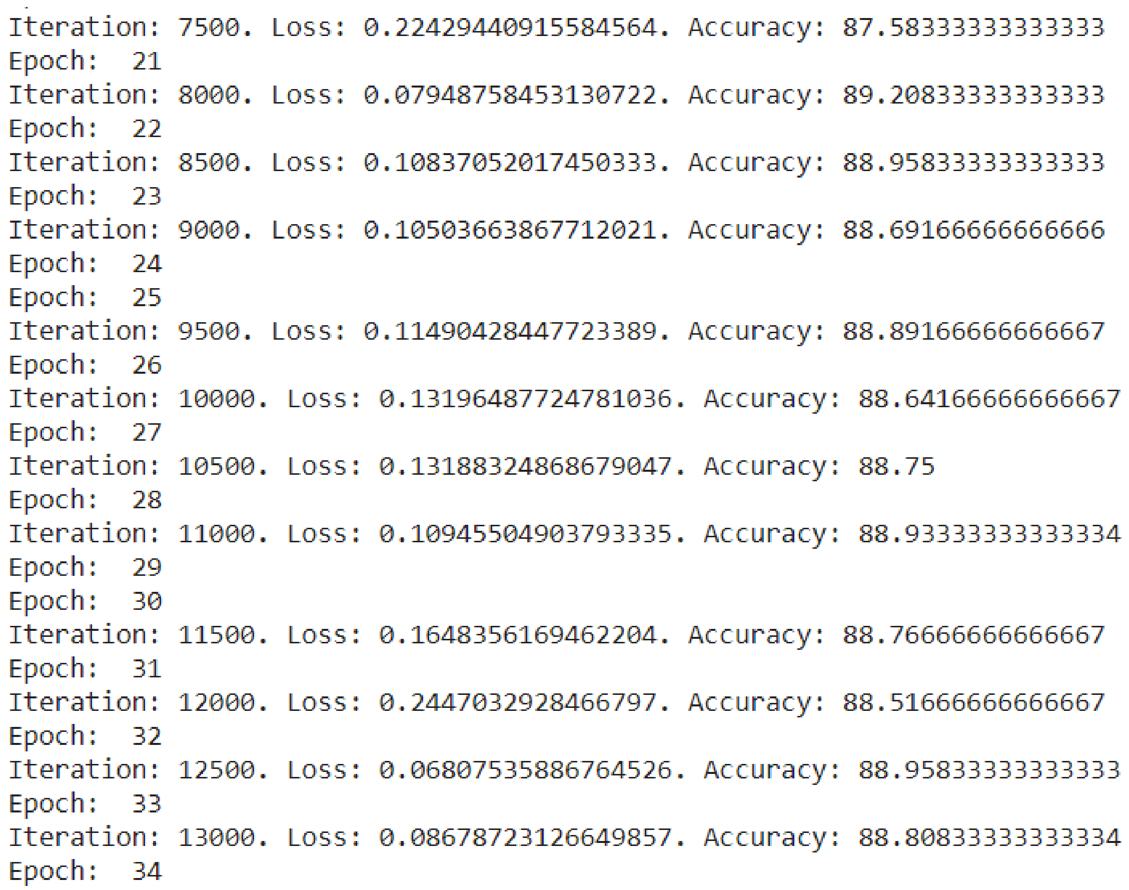
```
Epoch:  1
Iteration: 500. Loss: 2.3019180297851562. Accuracy: 9.933333333333334
Iteration: 1000. Loss: 2.2937798500061035. Accuracy: 17.358333333333334
Iteration: 1500. Loss: 2.2800097465515137. Accuracy: 23.466666666666665
Iteration: 2000. Loss: 1.9180361032485962. Accuracy: 20.175
Epoch:  2
Iteration: 2500. Loss: 1.208067774772644. Accuracy: 43.516666666666666
Iteration: 3000. Loss: 0.8502014875411987. Accuracy: 59.11666666666667
Iteration: 3500. Loss: 0.6840984225273132. Accuracy: 62.80833333333333
Iteration: 4000. Loss: 0.7317054271697998. Accuracy: 70.09166666666667
Iteration: 4500. Loss: 0.7963434457778931. Accuracy: 68.225
Epoch:  3
Iteration: 5000. Loss: 0.34371957182884216. Accuracy: 77.08333333333333
Iteration: 5500. Loss: 0.5013829469680786. Accuracy: 78.475
Iteration: 6000. Loss: 0.5444044470787048. Accuracy: 77.975
Iteration: 6500. Loss: 0.36920279264450073. Accuracy: 79.29166666666667
Iteration: 7000. Loss: 0.4009784162044525. Accuracy: 79.51666666666667
Epoch:  4
Iteration: 7500. Loss: 0.2810923457145691. Accuracy: 80.84166666666667
Iteration: 8000. Loss: 0.4678233563899994. Accuracy: 81.49166666666666
Iteration: 8500. Loss: 0.5707552433013916. Accuracy: 82.125
Iteration: 9000. Loss: 0.3213357925415039. Accuracy: 83.0
Iteration: 9500. Loss: 0.47090110182762146. Accuracy: 83.53333333333333
Epoch:  5
Iteration: 10000. Loss: 0.7291136980056763. Accuracy: 82.60833333333333
Iteration: 10500. Loss: 0.9414704442024231. Accuracy: 84.24166666666666
Iteration: 11000. Loss: 0.5701626539230347. Accuracy: 77.325
Iteration: 11500. Loss: 0.4237263798713684. Accuracy: 84.89166666666667
Iteration: 12000. Loss: 0.2099914848804474. Accuracy: 85.2
Epoch:  6
Iteration: 12500. Loss: 0.24570226669311523. Accuracy: 85.575
Iteration: 13000. Loss: 0.28660279512405396. Accuracy: 84.88333333333334
Iteration: 13500. Loss: 0.30418285727500916. Accuracy: 85.64166666666667
Iteration: 14000. Loss: 0.5999232530593872. Accuracy: 85.69166666666666
```

Modified Parameters:



```
batch_size = 122
num_iters = 20000
input_dim = 28*28
num_hidden = 220
output_dim = 10
learning_rate = 0.001
```

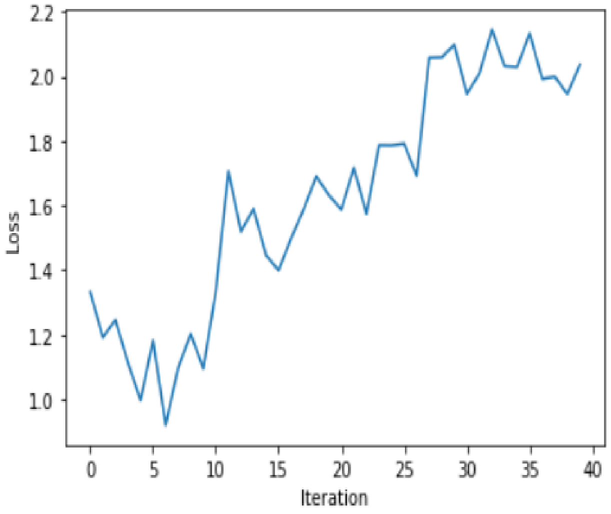
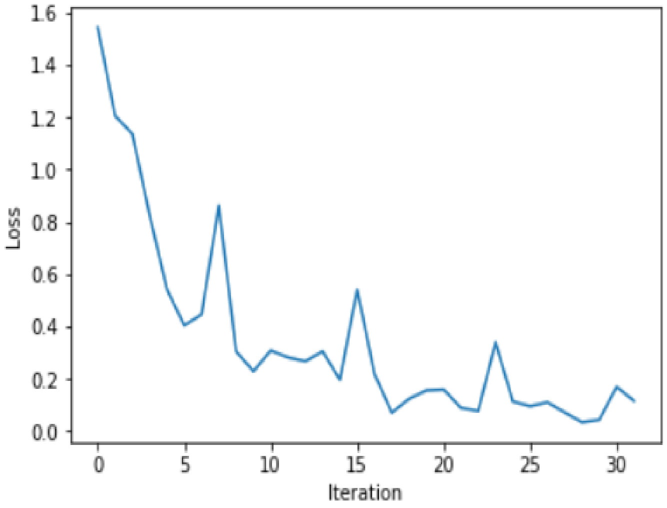
Iterations and Epochs:



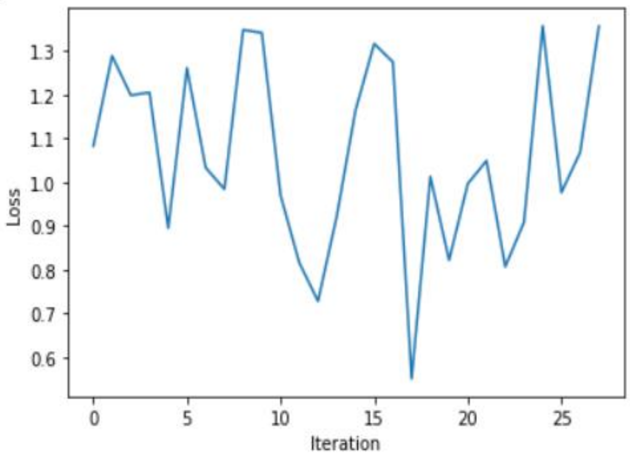
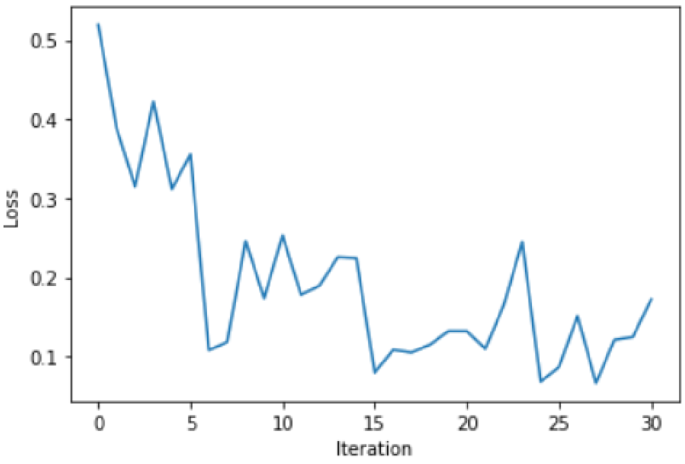
```
Iteration: 7500. Loss: 0.22429440915584564. Accuracy: 87.58333333333333
Epoch:  21
Iteration: 8000. Loss: 0.07948758453130722. Accuracy: 89.20833333333333
Epoch:  22
Iteration: 8500. Loss: 0.10837052017450333. Accuracy: 88.95833333333333
Epoch:  23
Iteration: 9000. Loss: 0.10503663867712021. Accuracy: 88.69166666666666
Epoch:  24
Epoch:  25
Iteration: 9500. Loss: 0.11490428447723389. Accuracy: 88.89166666666667
Epoch:  26
Iteration: 10000. Loss: 0.13196487724781036. Accuracy: 88.64166666666667
Epoch:  27
Iteration: 10500. Loss: 0.13188324868679047. Accuracy: 88.75
Epoch:  28
Iteration: 11000. Loss: 0.10945504903793335. Accuracy: 88.93333333333334
Epoch:  29
Epoch:  30
Iteration: 11500. Loss: 0.1648356169462204. Accuracy: 88.76666666666667
Epoch:  31
Iteration: 12000. Loss: 0.2447032928466797. Accuracy: 88.51666666666667
Epoch:  32
Iteration: 12500. Loss: 0.06807535886764526. Accuracy: 88.95833333333333
Epoch:  33
Iteration: 13000. Loss: 0.08678723126649857. Accuracy: 88.80833333333334
Epoch:  34
```

Loss vs Iteration Curves:

Dataset-1:

Experiment-1	Experiment-1
	

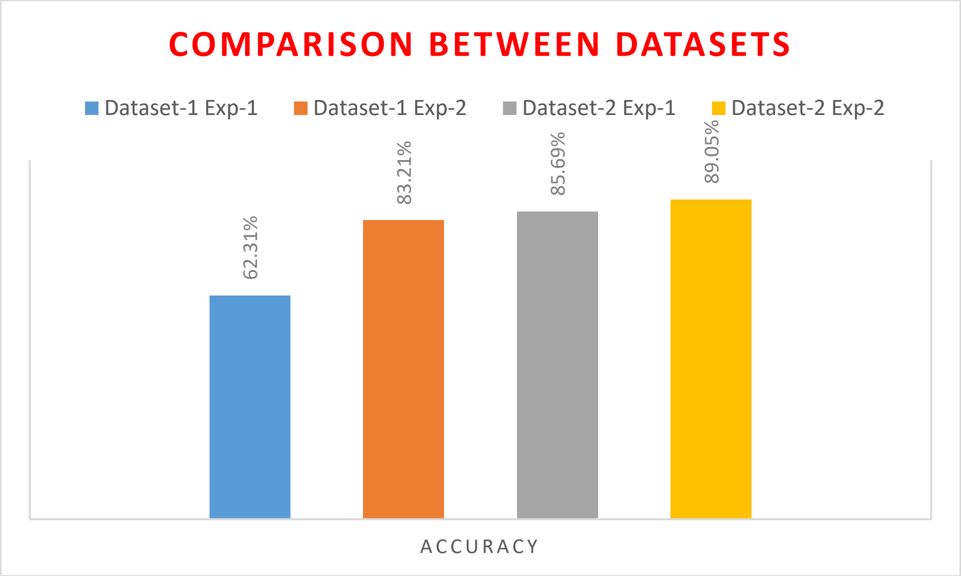
Dataset-2:

Experiment-1	Experiment-2
	

Comparison Table:

	Experiments	Batch Size	No. hidden layers	Number of nodes	iteration	Learning rate	Optimizer	Epoch	Accuracy
Dataset -1	Exp-1	20	6	200	20000	0.01	Relu(all layers)	205	62.31%
	Exp-2	122	4	220	20000	0.001	Softmax(in 2^{nd} layer), rest are Relu.	100	83.206%
Dataset -2	Exp-1	20	6	200	20000	0.01	Relu(all layers)	6	85.69%
	Exp-2	122	4	220	20000	0.001	Softmax(in 2^{nd} layer), rest are Relu.	40	89.05%

Comparison Chart:



Here we can see a significant difference in Experiment 1 and Experiment 2. When we applied six hidden layers and above shown hyperparameters, we did not get expected accuracy in Experiment 1. We got 62.31% highest accuracy. But after changing the hyperparamters, reducing the hidden layers by 4, using Adam optimizer and also using Softmax activation functions in one of the hidden layers, we got highest accuracy as 83.206% in Experiment 2. Because in our first experiment, the base parameters has 205 epochs and it eventually overfits the model. But in experiment-2, we reduced the epoch, gave used 4 hidden layers. Also, we used Softmax optimizer in 2nd layer and batch_size 122. By this approach, we got 83.2% accuracy.

Github code links:

https://github.com/kanak109/softcom-assignments.git