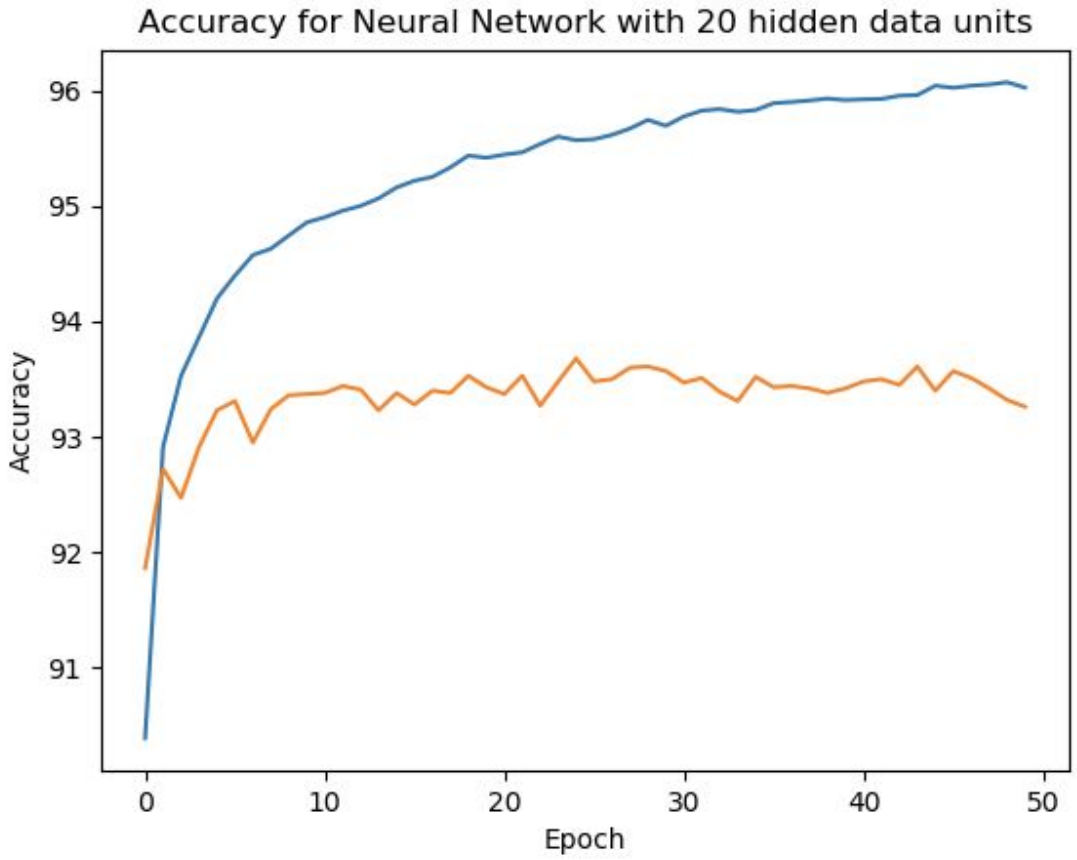
CS 445/545: Machine Learning, Spring 2018
Programming Assignment #1

Reference:
Blue - Train dataset
Orange - Test dataset

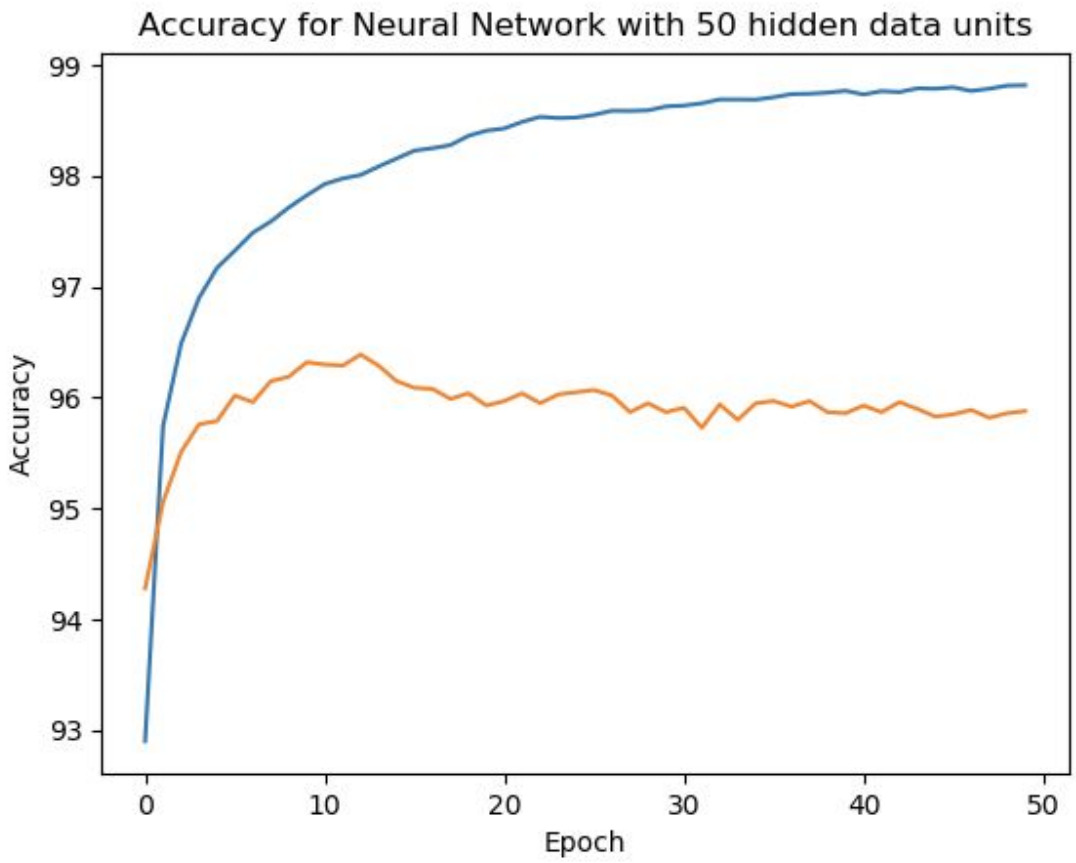
Experiment 1: Vary number of hidden units with momentum = 0.9 and epoch = 50

1. 20 hidden units



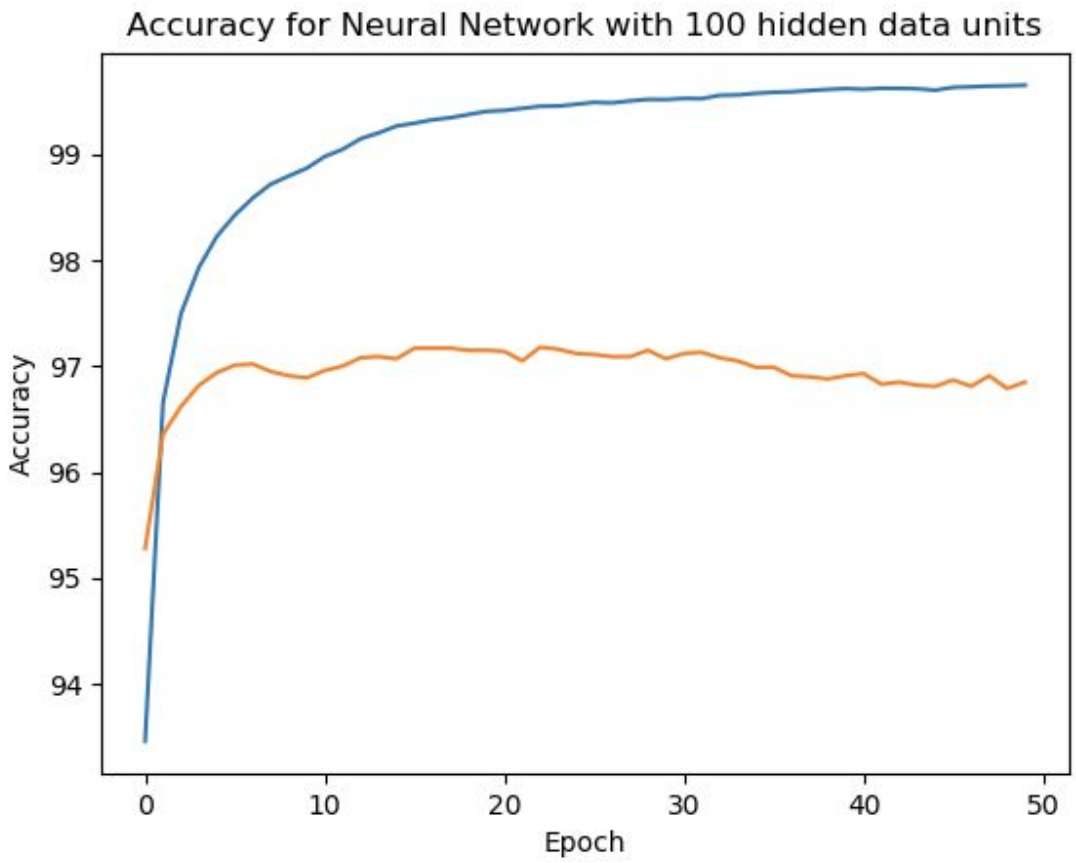
```
Srilakshmis-MacBook-Pro:hw1 srilakshmishivakumar$ python ques2.py

Confusion matrix for train data with 20 hidden data points

[[5739    0   62    2    9   18   11    0   76    6]
 [   2 6548   45   39    7   13    4   12   58   14]
 [  22   19 5698   19   39   15   28   44   63   11]
 [   4    3   92 5815    3   54   10   34   87   29]
 [   3   15   29    2 5614    3   33    7   21  115]
 [  17    3   29  135    9 5117   43    5   39   24]
 [  22   10   42    1   17   60 5713    1   50    2]
 [   8   12   35    7   40   20    3 6034   31   75]
 [  10   40   28    5   10   37   15    6 5669   31]
 [  19    6   17   45   64   35    2   38   52 5671]]

Train data accuracy

[90.385      92.91333333 93.52833333 93.86333333 94.2        94.39833333
 94.57666667 94.63166667 94.74666667 94.86       94.905      94.96166667
 95.00333333 95.06833333 95.16333333 95.22166667 95.255      95.33833333
 95.44       95.42166667 95.44833333 95.46833333 95.54       95.605
 95.57333333 95.58166667 95.62       95.675      95.75       95.69833333
 95.78       95.83       95.84333333 95.82       95.835      95.895
 95.905      95.91833333 95.935      95.92166667 95.92833333 95.93166667
 95.96       95.965      96.04666667 96.02833333 96.04666667 96.05666667
 96.07666667 96.03      ]

Confusion matrix for test data with 20 hidden data points

[[ 963    0    1    0    0    5    3    3    5    0]
 [   0 1110    5    5    1    1    1    2   10    0]
 [  11    8  941    4    7    2    5    9   38    7]
 [   4    0   18  927    1   23    0    5   22   10]
 [   1    2    5    0  898    1    9    1    5   60]
 [   9    1    5   38    4  788   14    6   19    8]
 [  15    3    7    0    2   12  902    0   17    0]
 [   2    7   23    4    2    0    0  942   12   36]
 [  10   14    4    1    7   10    9    3  908    8]
 [   7    6    2    8   11    5    0    8   15  947]]

Test data accuracy

[91.86 92.72 92.47 92.91 93.23 93.31 92.95 93.24 93.36 93.37 93.38 93.44
 93.41 93.23 93.38 93.28 93.4  93.38 93.53 93.43 93.37 93.53 93.27 93.48
 93.68 93.48 93.5  93.6  93.61 93.57 93.47 93.51 93.39 93.31 93.52 93.43
 93.44 93.42 93.38 93.42 93.48 93.5  93.45 93.61 93.4  93.57 93.51 93.42
 93.32 93.26]
Srilakshmis-MacBook-Pro:hw1 srilakshmishivakumar$
```

2. 50 hidden units



```
Srilakshmis-MacBook-Pro:hw1 srilakshmishivakumar$ python ques2.py

Confusion matrix for train data with 50 hidden data points

[[5841    0    8    4    6    0    8    2   48    6]
 [   2 6687   22    4    6    0    1    4   15    1]
 [   9    1 5895   12    7    4    4   17    3    6]
 [   2    0   13 6050    3    9    3   10   28   13]
 [   4    7   10    2 5778    1    7    1    7   25]
 [  13    4    5   27    2 5316   15    3   25   11]
 [  12    3    3    2    3   11 5868    1   15    0]
 [   0    4   20    1    5    0    1 6211    4   19]
 [   6    6   12    3    2    5    4    1 5806    6]
 [  10    0    2   19   13    9    3   17   36 5840]]

Train data accuracy

[92.90166667 95.745      96.48666667 96.9        97.17166667 97.32833333
 97.49       97.59       97.71666667 97.82666667 97.92833333 97.97833333
 98.01       98.08666667 98.15833333 98.23       98.25166667 98.28
 98.36333333 98.41       98.43       98.49       98.535      98.52333333
 98.52833333 98.55333333 98.59       98.58833333 98.59333333 98.62833333
 98.63666667 98.65666667 98.69166667 98.69166667 98.69       98.71166667
 98.74       98.74333333 98.75333333 98.77       98.73666667 98.76666667
 98.75833333 98.79166667 98.78833333 98.80166667 98.76833333 98.78833333
 98.81666667 98.82      ]

Confusion matrix for test data with 50 hidden data points

[[ 965    0    3    1    0    0    1    2    7    1]
 [   0 1120    4    3    1    1    1    3    2    0]
 [   5    2  984   11    4    1    1   13    9    2]
 [   3    0    5  960    0   11    0    8   15    8]
 [   1    0    4    0  940    0    3    3    5   26]
 [   5    1    1    7    1  846    8    3   14    6]
 [  10    4    2    1    4   16  912    0    9    0]
 [   1    2   12    4    5    0    0  982    5   17]
 [   4    1    3    5    4    6    3    5  927   16]
 [   6    6    1    5   15    2    1    6   15  952]]

Test data accuracy

[94.28 95.06 95.51 95.76 95.79 96.02 95.96 96.15 96.19 96.32 96.3  96.29
 96.39 96.29 96.15 96.09 96.08 95.99 96.04 95.93 95.97 96.04 95.95 96.03
 96.05 96.07 96.02 95.87 95.95 95.87 95.91 95.73 95.94 95.8  95.95 95.97
 95.92 95.97 95.87 95.86 95.93 95.87 95.96 95.9  95.83 95.85 95.89 95.82
 95.86 95.88]
Srilakshmis-MacBook-Pro:hw1 srilakshmishivakumar$
```

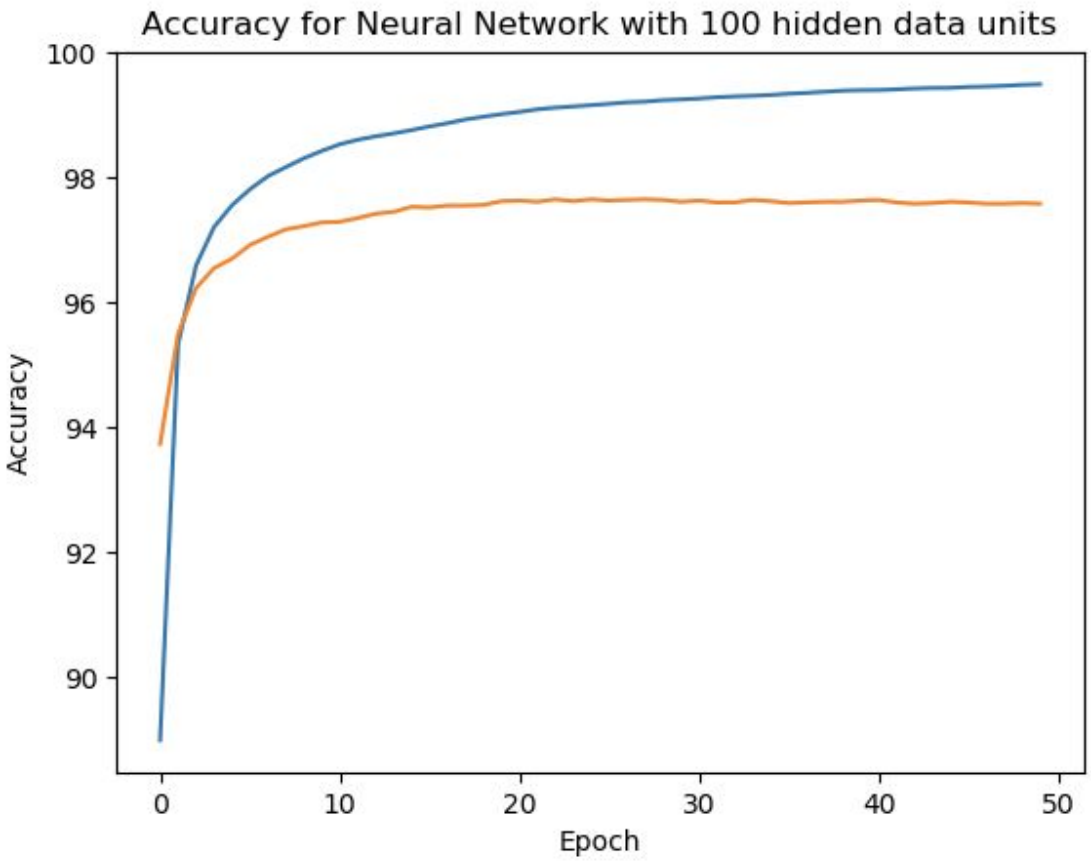
3. 100 hidden units



```
Srilakshmis-MacBook-Pro:hw1 srilakshmishivakumar$ python ques2.py

Confusion matrix for train data with 100 hidden data points

[[5895    1    4    0    0    1    3    0   15    4]
 [   0 6727    4    1    2    3    0    3    2    0]
 [   4    1 5939    0    0    0    1    1   10    2]
 [   0    0    3 6107    0    0    0    4   11    6]
 [   2    1    2    0 5825    0    1    1    4    6]
 [   2    1    3    6    0 5398    2    1    5    3]
 [   5    0    1    0    1    7 5895    0    9    0]
 [   0    2    4    3    0    1    0 6244    7    4]
 [   2    1    0    1    0    0    0    0 5843    4]
 [   5    1    0   10    1    1    1    0   12 5918]]

Train data accuracy

[93.46166667 96.66166667 97.49833333 97.93333333 98.22666667 98.42666667
 98.58833333 98.71833333 98.795      98.86833333 98.97833333 99.04666667
 99.14666667 99.20166667 99.26666667 99.29333333 99.325      99.345
 99.375      99.405      99.41333333 99.43333333 99.45333333 99.455
 99.47166667 99.49       99.485      99.50333333 99.51833333 99.51666667
 99.52666667 99.525      99.55666667 99.56       99.575      99.585
 99.59       99.60166667 99.61166667 99.62       99.61333333 99.62333333
 99.62166667 99.61833333 99.605      99.63166667 99.63666667 99.64333333
 99.64666667 99.65166667]

Confusion matrix for test data with 100 hidden data points

[[ 968    1    2    0    0    3    1    2    3    0]
 [   1 1121    3    2    0    1    1    1    5    0]
 [   6    1  990    6    2    1    1    9   14    2]
 [   1    0    5  969    0   18    0    5   10    2]
 [   1    0    2    0  951    0    6    1    1   20]
 [   4    1    1   10    0  860    4    3    5    4]
 [   5    3    1    1    3   13  922    0   10    0]
 [   0    5   11    2    3    1    0  987    6   13]
 [   4    1    3    3    3    2    4    3  946    5]
 [   3    5    1    4   12    0    1    2   10  971]]

Test data accuracy

[95.28 96.36 96.62 96.82 96.94 97.01 97.02 96.95 96.91 96.89 96.96 97.
 97.08 97.09 97.07 97.17 97.17 97.17 97.15 97.15 97.14 97.05 97.18 97.16
 97.12 97.11 97.09 97.09 97.15 97.07 97.12 97.13 97.08 97.05 96.99 96.99
 96.91 96.9  96.88 96.91 96.93 96.83 96.85 96.82 96.81 96.87 96.81 96.91
 96.79 96.85]
Srilakshmis-MacBook-Pro:hw1 srilakshmishivakumar$
```

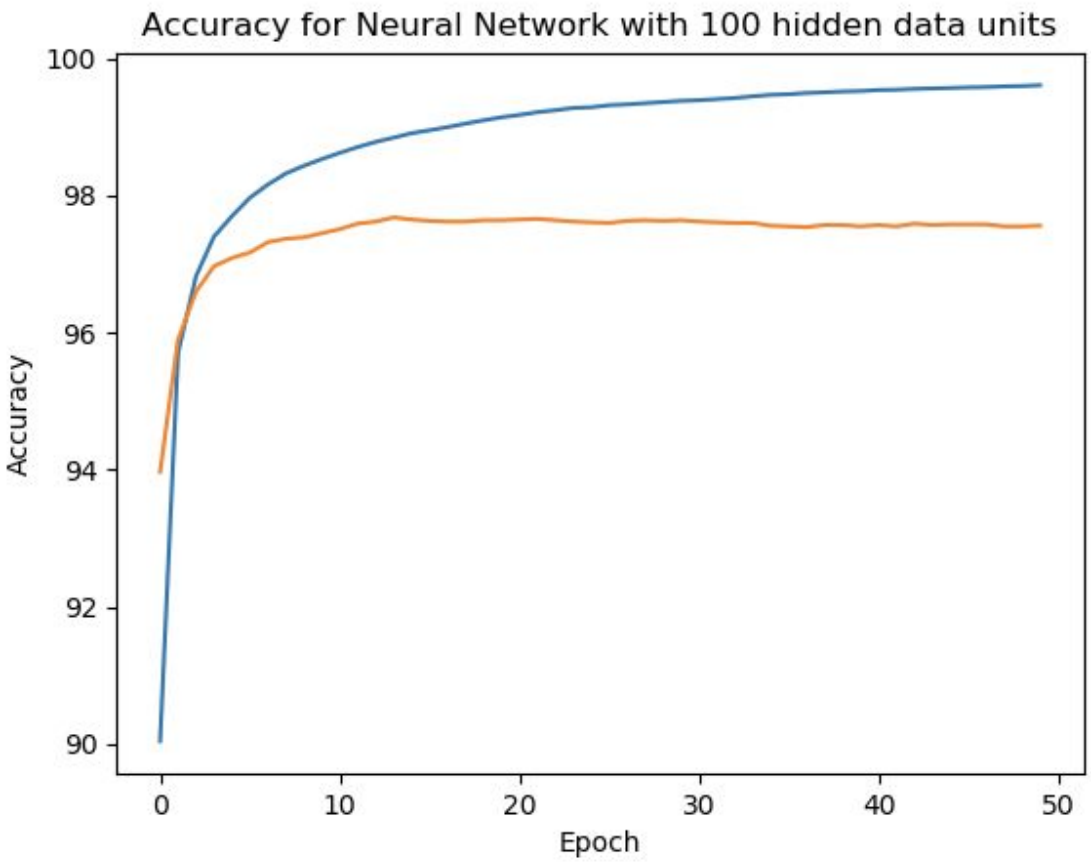
Experiment 2: Vary the momentum value with 100 hidden units and epoch = 50

1. Momentum = 0



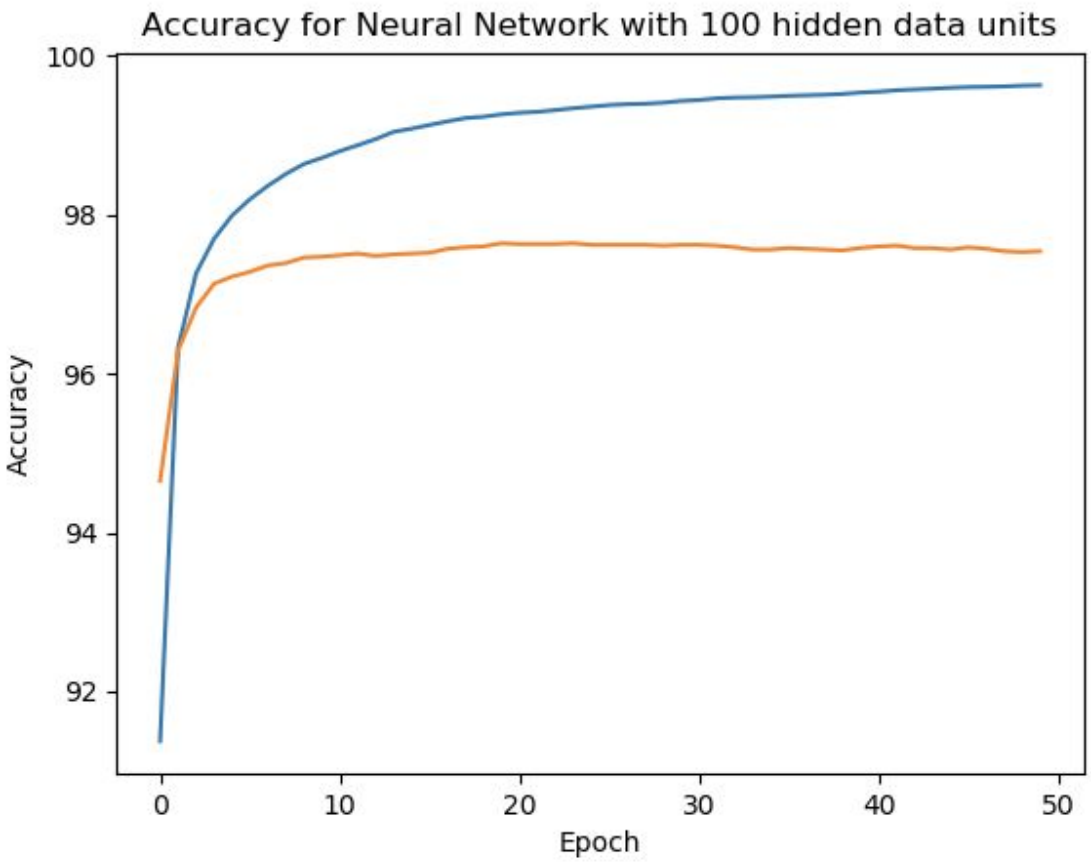
```
Srilakshmis-MacBook-Pro:hw1 srilakshmishivakumar$ python ques2.py

 momentum - 0

Confusion matrix for train data with 100 hidden data points

[[5912    0    1    0    1    0    4    0    3    2]
 [   1 6692   12    8    5    1    2   12    6    3]
 [   5    3 5938    0    3    0    1    5    1    2]
 [   1    0    5 6088    0    6    0   13    8   10]
 [   1    4    1    0 5821    0    3    1    1   10]
 [   4    1    1    2    0 5401    5    0    3    4]
 [   8    3    1    0    2    6 5891    0    7    0]
 [   5    7   15    3    0    0    0 6224    3    8]
 [   3    5    1    4    1    4    2    1 5828    2]
 [   6    2    1    3    7    8    0   14    9 5899]]

Train data accuracy

[89.00333333 95.33166667 96.59333333 97.205      97.555      97.81
 98.02       98.165      98.305      98.42333333 98.53       98.60166667
 98.65666667 98.70333333 98.75666667 98.815      98.86666667 98.92666667
 98.97       99.01166667 99.04666667 99.08833333 99.115      99.135
 99.155      99.17666667 99.2        99.21166667 99.235      99.24666667
 99.26166667 99.28166667 99.295      99.305      99.31833333 99.33833333
 99.35166667 99.37       99.38666667 99.395      99.39666667 99.40833333
 99.42166667 99.43       99.43166667 99.44666667 99.45333333 99.465
 99.48       99.49      ]

Confusion matrix for test data with 100 hidden data points

[[ 973    1    0    1    0    1    0    1    2    1]
 [   0 1122    3    4    0    1    1    1    3    0]
 [   5    2 1010    0    0    0    3    3    7    2]
 [   1    0    5  991    0    4    0    2    3    4]
 [   1    0    1    0  957    0    6    0    1   16]
 [   5    0    0    9    2  861    1    1    5    8]
 [   9    3    0    1    2    6  931    0    6    0]
 [   2    4   12    4    3    0    0  988    3   12]
 [   4    1    2    1    5    6    1    3  947    4]
 [   3    6    1    5    6    2    0    4    4  978]]

Test data accuracy

[93.74 95.5  96.23 96.55 96.7  96.92 97.05 97.17 97.22 97.28 97.29 97.35
 97.42 97.45 97.53 97.52 97.55 97.55 97.56 97.62 97.63 97.61 97.65 97.62
 97.65 97.63 97.64 97.65 97.64 97.61 97.63 97.6  97.6  97.64 97.62 97.59
 97.6  97.61 97.61 97.63 97.64 97.6  97.58 97.59 97.61 97.6  97.58 97.58
 97.59 97.58]
Srilakshmis-MacBook-Pro:hw1 srilakshmishivakumar$
```

2. Momentum = 0.25



```
Srilakshmis-MacBook-Pro:hw1 srilakshmishivakumar$ python ques2.py

 momentum - 0.25

Confusion matrix for train data with 100 hidden data points

[[5912    0    1    0    0    0    3    0    5    2]
 [   1 6707    8    3    3    0    3    5    9    3]
 [   3    1 5945    0    1    0    0    2    4    2]
 [   1    1    5 6101    0    2    0    6    7    8]
 [   1    7    1    0 5822    0    1    1    1    8]
 [   2    1    3    1    2 5392    6    0    7    7]
 [   7    1    0    0    1    2 5901    0    6    0]
 [   2    4   10    1    2    1    0 6230    0   15]
 [   3    4    0    1    1    0    1    0 5837    4]
 [   6    1    1    3    5    1    1    9    4 5918]]

Train data accuracy

[90.045      95.70833333 96.825      97.40166667 97.7        97.96833333
 98.16       98.32166667 98.43333333 98.53166667 98.62166667 98.70666667
 98.78166667 98.84333333 98.90833333 98.95166667 98.99666667 99.04833333
 99.09666667 99.14       99.175      99.215      99.24333333 99.27666667
 99.28666667 99.31666667 99.32833333 99.34666667 99.36333333 99.38
 99.39       99.405      99.42166667 99.44666667 99.46833333 99.47666667
 99.495      99.505      99.51666667 99.52333333 99.53833333 99.54333333
 99.555      99.56166667 99.56833333 99.57666667 99.58166667 99.59166667
 99.59666667 99.60833333]

Confusion matrix for test data with 100 hidden data points

[[ 973    1    0    0    0    0    1    2    3    0]
 [   0 1120    2    2    0    3    4    1    3    0]
 [   3    1 1007    3    0    0    5    5    7    1]
 [   0    0    4  990    0    4    0    3    5    4]
 [   2    0    0    0  955    0    5    0    2   18]
 [   3    0    1    9    0  862    4    0    7    6]
 [   6    3    2    0    2    7  936    0    2    0]
 [   0    4   11    6    3    0    0  988    1   15]
 [   5    1    2    2    4    4    3    3  948    2]
 [   2    6    0    5    8    0    1    5    5  977]]

Test data accuracy

[93.97 95.9  96.6  96.97 97.09 97.17 97.32 97.37 97.39 97.45 97.51 97.59
 97.62 97.68 97.65 97.63 97.62 97.62 97.64 97.64 97.65 97.66 97.64 97.62
 97.61 97.6  97.63 97.64 97.63 97.64 97.62 97.61 97.6  97.6  97.56 97.55
 97.54 97.57 97.57 97.55 97.57 97.55 97.59 97.57 97.58 97.58 97.58 97.55
 97.55 97.56]
Srilakshmis-MacBook-Pro:hw1 srilakshmishivakumar$
```

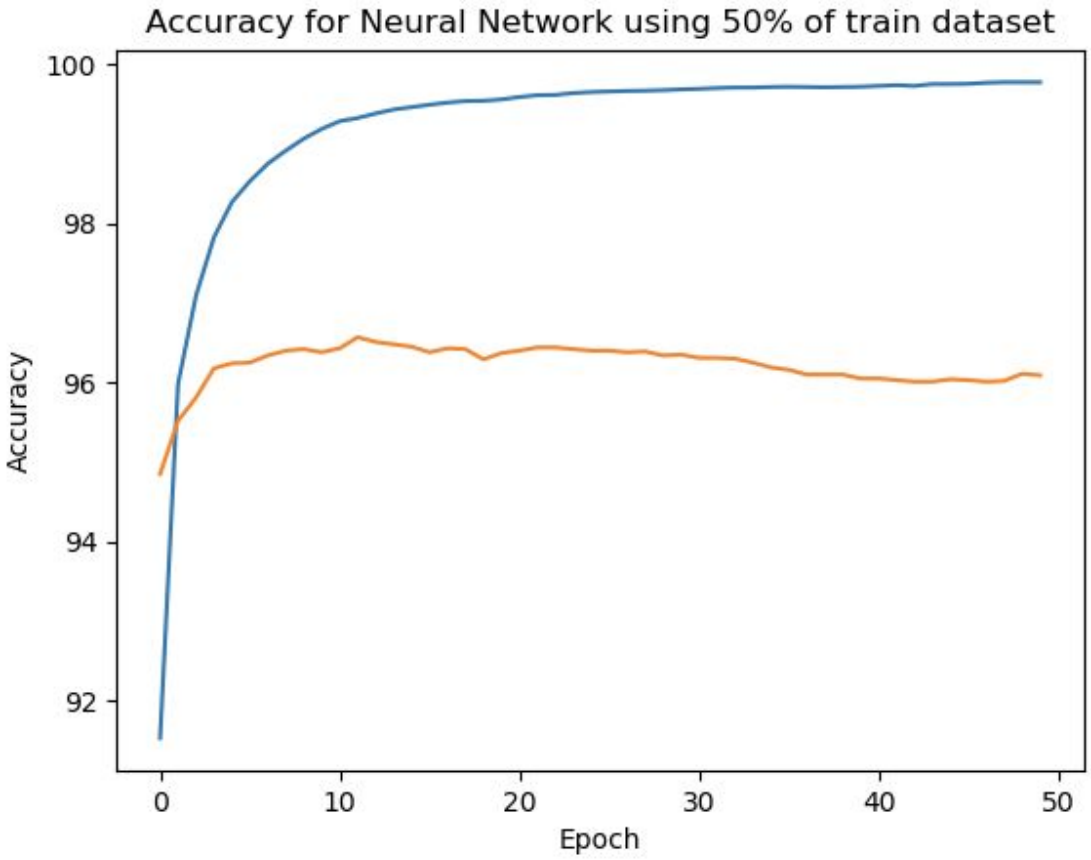
3. Momentum = 0.5



```
Srilakshmis-MacBook-Pro:hw1 srilakshmishivakumar$ python ques2.py

 momentum - 0.5

Confusion matrix for train data with 100 hidden data points

[[5906    1    0    0    0    1    5    0    7    3]
 [   1 6689   20    5    4    2    5    5    9    2]
 [   2    2 5948    2    0    0    0    1    3    0]
 [   0    0    3 6110    1    3    0    5    5    4]
 [   2    6    0    0 5821    0    4    1    4    4]
 [   5    1    3    2    0 5401    4    0    3    2]
 [   9    2    0    0    2    2 5898    0    5    0]
 [   1    4    9    1    1    0    0 6239    1    9]
 [   1    2    1    0    0    0    1    1 5844    1]
 [   8    1    0    3    3    4    0    4    5 5921]]

Train data accuracy

[91.375      96.32833333 97.26       97.695      97.98833333 98.19333333
 98.36333333 98.51166667 98.63666667 98.71333333 98.79833333 98.87333333
 98.95       99.04166667 99.08166667 99.12833333 99.17333333 99.215
 99.23166667 99.26166667 99.28166667 99.29333333 99.315      99.33833333
 99.35833333 99.37833333 99.39       99.395      99.40833333 99.43
 99.44166667 99.46333333 99.47333333 99.47666667 99.485      99.495
 99.50166667 99.50833333 99.51833333 99.53666667 99.54666667 99.565
 99.575      99.585      99.59666667 99.605      99.60833333 99.61333333
 99.62333333 99.62833333]

Confusion matrix for test data with 100 hidden data points

[[ 970    1    1    1    0    2    1    1    2    1]
 [   0 1122    2    2    0    2    4    1    2    0]
 [   4    3 1001    6    0    0    2   10    5    1]
 [   0    0    1  990    0    8    0    4    3    4]
 [   2    0    1    1  963    0    4    0    1   10]
 [   4    0    0   10    1  864    4    1    4    4]
 [   6    3    2    0    0   10  933    0    3    1]
 [   1    4   15    4    2    1    0  992    2    7]
 [   5    2    2    1    3    5    3    3  947    3]
 [   4    6    1    4   10    4    0    4    4  972]]

Test data accuracy

[94.65 96.29 96.83 97.13 97.22 97.28 97.36 97.39 97.46 97.47 97.49 97.51
 97.48 97.5  97.51 97.52 97.57 97.59 97.6  97.64 97.63 97.63 97.63 97.64
 97.62 97.62 97.62 97.62 97.61 97.62 97.62 97.61 97.59 97.56 97.56 97.58
 97.57 97.56 97.55 97.58 97.6  97.61 97.58 97.58 97.56 97.59 97.57 97.54
 97.53 97.54]
Srilakshmis-MacBook-Pro:hw1 srilakshmishivakumar$
```

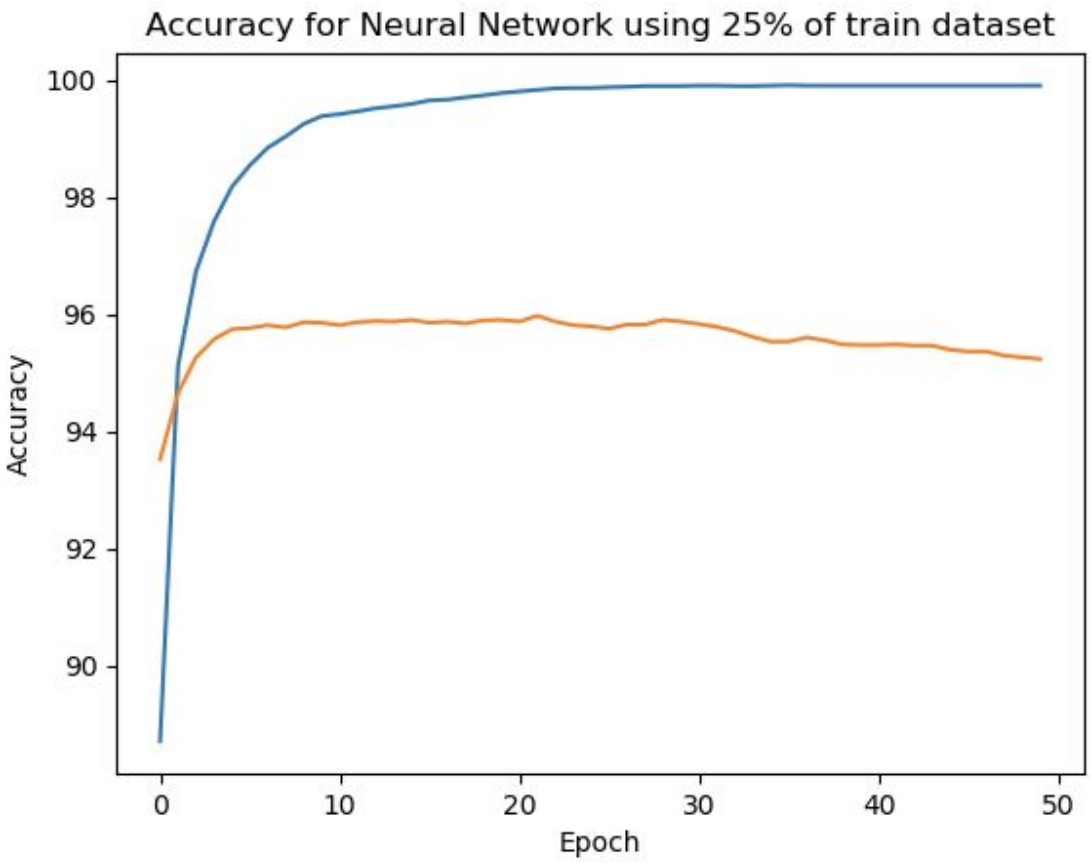
3. Vary the number of training examples with 100 hidden units and momentum = 0.9

1. Neural network using one half of training dataset



```
Srilakshmis-MacBook-Pro:hw1 srilakshmishivakumar$ python ques2.py

 momentum - 0.9

Confusion matrix for train data with 100 hidden data points

[[2956    0    0    0    0    0    1    0    3    1]
 [   0 3414    2    2    0    0    0    1    3    1]
 [   0    0 2945    0    0    0    0    0    3    0]
 [   0    1    2 3058    0    0    0    1    9    2]
 [   1    0    2    0 2917    0    0    0    1    5]
 [   0    0    2    2    0 2697    1    0    4    3]
 [   1    0    0    0    0    0 2973    0    1    0]
 [   0    0    3    1    0    0    0 3099    2    2]
 [   0    0    0    0    0    0    0    0 2874    1]
 [   1    0    0    0    0    0    0    0    2 3000]]

Train data accuracy

[91.52666667 95.98333333 97.09333333 97.82666667 98.27       98.53333333
 98.75333333 98.91666667 99.06666667 99.19       99.28666667 99.32666667
 99.38333333 99.43333333 99.46333333 99.49333333 99.52       99.54
 99.54333333 99.56       99.59       99.61333333 99.61666667 99.64
 99.65333333 99.66       99.66666667 99.67       99.67666667 99.68666667
 99.69333333 99.70333333 99.71       99.71       99.71666667 99.72
 99.71666667 99.71333333 99.71666667 99.72       99.73       99.74
 99.73       99.75333333 99.75333333 99.75666667 99.77       99.77666667
 99.77666667 99.77666667]

Confusion matrix for test data with 100 hidden data points

[[ 970    2    1    0    0    1    3    0    2    1]
 [   1 1112    2    4    0    1    2    2   10    1]
 [   8    2  976    5    5    0    1   19   10    6]
 [   3    0    3  965    0    6    0    3   22    8]
 [   1    0    1    0  962    0    2    1    4   11]
 [   6    0    4   24    0  812   12    6   18   10]
 [   9    4    0    0    8    7  915    0   14    1]
 [   0    6    8    6    8    0    0  986    1   13]
 [   7    0    1    0    3    5    4    5  946    3]
 [   6    2    0    6   11    1    0    6   12  965]]

Test data accuracy

[94.85 95.52 95.81 96.18 96.24 96.25 96.34 96.4  96.42 96.38 96.43 96.57
 96.51 96.48 96.45 96.38 96.43 96.42 96.29 96.37 96.4  96.44 96.44 96.42
 96.4  96.4  96.38 96.39 96.34 96.35 96.31 96.31 96.3  96.25 96.19 96.16
 96.1  96.1  96.1  96.05 96.05 96.03 96.01 96.01 96.04 96.03 96.01 96.02
 96.11 96.09]
^Z
[3]+  Stopped                 python ques2.py
Srilakshmis-MacBook-Pro:hw1 srilakshmishivakumar$
```

2. Neural network using one quarter of training dataset



```
Srilakshmis-MacBook-Pro:hw1 srilakshmishivakumar$ python ques2.py

 momentum - 0.9

Confusion matrix for train data with 100 hidden data points

[[1492    0    0    0    0    0    2    0    1    1]
 [   0 1688    0    0    1    0    0    0    1    0]
 [   0    0 1461    0    0    0    0    0    1    0]
 [   0    0    0 1546    0    0    0    0    1    1]
 [   0    0    0    0 1468    0    0    0    0    0]
 [   0    0    1    1    0 1316    0    0    0    0]
 [   0    0    0    0    0    0 1490    0    0    0]
 [   0    1    1    1    0    0    0 1590    0    0]
 [   0    0    0    0    0    0    0    0 1432    0]
 [   0    0    0    0    0    0    0    0    1 1502]]

Train data accuracy

[88.70666667 95.12       96.73333333 97.58       98.17333333 98.54
 98.84       99.03333333 99.24666667 99.38       99.41333333 99.46
 99.51333333 99.54666667 99.58666667 99.64666667 99.66       99.7
 99.73333333 99.77333333 99.8        99.82666667 99.85333333 99.86
 99.86       99.87333333 99.88       99.89333333 99.89333333 99.89333333
 99.9        99.9        99.89333333 99.89333333 99.9        99.90666667
 99.9        99.9        99.9        99.9        99.9        99.9
 99.9        99.9        99.9        99.9        99.9        99.9
 99.9        99.9       ]

Confusion matrix for test data with 100 hidden data points

[[ 958    1    1    1    0    1    8    2    6    2]
 [   0 1115    5    1    1    2    3    0    8    0]
 [   8    1  974    9    2    2    7    7   19    3]
 [   1    0   11  961    0    4    1    5   19    8]
 [   1    2    4    0  923    0    9    3    3   37]
 [   6    0    2   33    3  809    7    1   26    5]
 [  11    3    4    0    1    3  922    1   12    1]
 [   2    5   13    3    4    0    0  978    4   19]
 [   7    0    4    4    3    2    5    3  935   11]
 [   4    4    1   12   14    3    1    9   13  948]]

Test data accuracy

[93.52 94.65 95.26 95.57 95.74 95.76 95.81 95.77 95.86 95.85 95.81 95.86
 95.88 95.87 95.9  95.85 95.87 95.84 95.89 95.9  95.87 95.97 95.87 95.81
 95.79 95.75 95.82 95.82 95.9  95.87 95.83 95.78 95.71 95.61 95.53 95.53
 95.6  95.55 95.48 95.47 95.47 95.48 95.46 95.46 95.39 95.36 95.36 95.29
 95.26 95.23]
Srilakshmis-MacBook-Pro:hw1 srilakshmishivakumar$
```