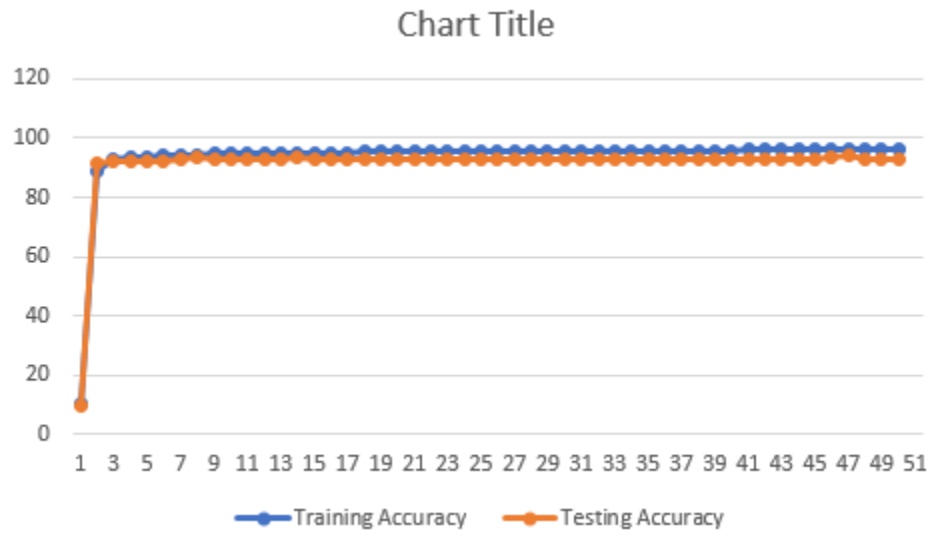
Experiment 1-

Description-

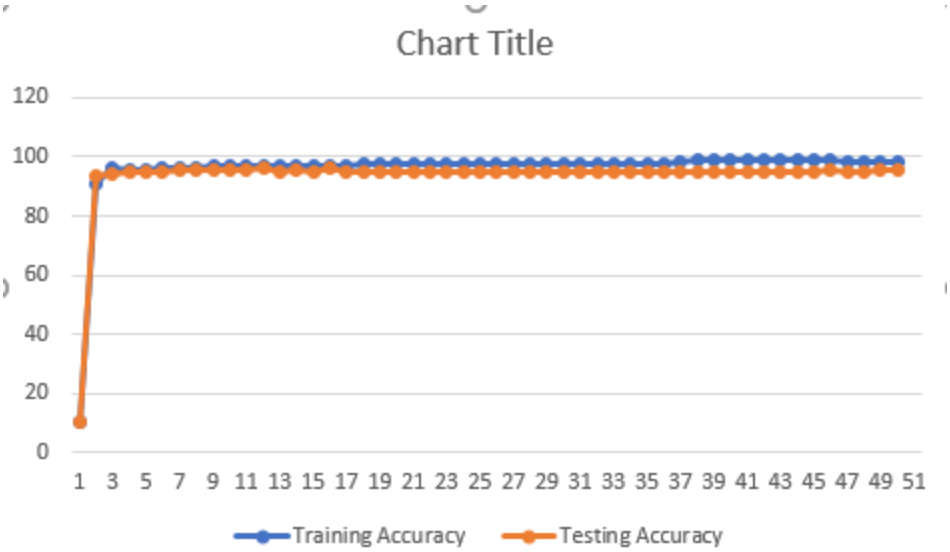
In this experiment we have to check the accuracy of neural network by changing the number of hidden layers, keeping momentum and learning rate same i.e. .9 and .1 resp.. We have tested on three hidden layers that are 20, 50 and 100 and calculated testing and training accuracies after each epoch and plotted them on the graph, also we have created confusion matrix at the end of the training on the test data.Data set used is mnist data for the hand written digit recognition we have scaled the data by dividing by 255 on the scale of 0,1. There were 784 inputs and 10 outputs the image is of size 28* 28.There were 60000 training data and testing data is 10000.

Plots and Matrix with hidden layer 20(93.11 final accuracy)-



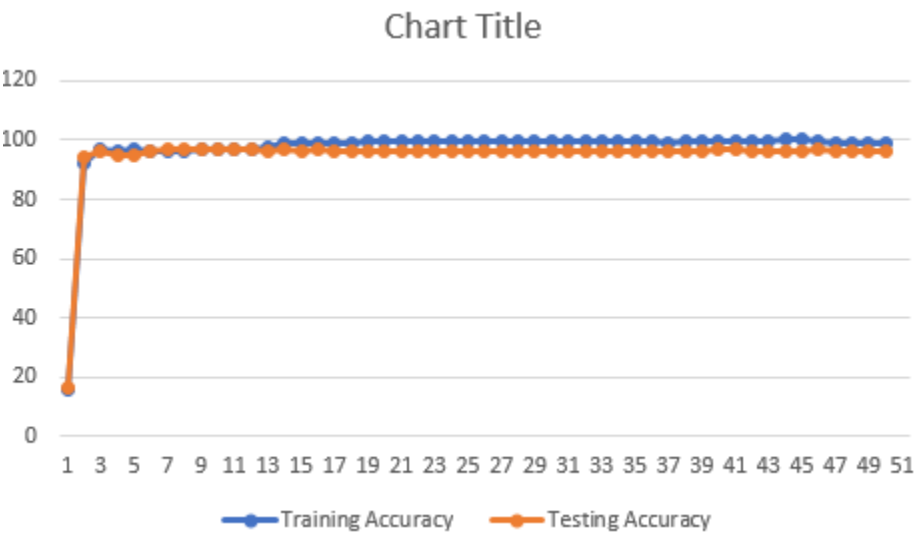
964	0	0	0	4	2	3	1	4	2
0	1103	2	3	0	1	4	0	22	0
15	3	928	17	9	8	14	7	26	5
9	0	8	924	0	19	4	7	32	7
1	3	4	0	895	0	8	2	7	62
9	3	1	23	7	800	15	3	24	7
19	5	2	2	3	13	905	0	7	2
2	5	13	10	3	0	4	955	7	29
12	7	2	11	13	17	9	2	895	6
9	7	1	16	17	2	1	11	3	942

Plots and Graphs for hidden layer 50(95.6 final accuracy)-




```
Confusion Matrix After Training on Testing data is as below:
    966          0          2          0          0          0          3          2          6          1
      0       1115          1          2          0          1          5          1         10          0
      8          5        955         22          1          0          8          7         25          1
      0          1          7        961          0         15          0          6         16          4
      2          0          0          1        939          0          9          1          3         27
      5          1          0         17          1        823         15          4         20          6
      8          2          3          1          3          9        923          1          8          0
      1          3         11          5          1          2          0        968          5         32
      2          3          1         10          3          7          2          3        938          5
      6          6          0          8         15          8          0          5         10        951
```

Plots and matrix after 100 hidden layers(96.8-Final accuracy)-




```
    967          0          2          0          0          1          3          1          4          2
      0       1117          4          2          0          1          3          2          6          0
      4          1        989         11          4          0          6          7          8          2
      0          0          5        983          0          5          1          5          8          3
      1          0          1          0        946          0          7          1          1         25
      5          1          2         28          1        817         15          3         14          6
      7          4          2          2          1          3        933          0          6          0
      3          3          7          6          3          0          0        986          2         18
      5          4          1         10          4          6          1          3        935          5
      3          5          1         11          9          2          0          2          4        972
```

(1) How does the number of hidden units affect the final accuracy on the test data?
The accuracy on test data increases with the number of hidden layers, which can be seen from the obtained results in the above experiment.

(2) How does it affect the number of epochs needed for training to converge?
The number of epochs needed for convergence decreased as the number of layers are increased. for 20-37,50-31,100-24

(3) Is there evidence that any of your networks has overfit to the training data? If so, what is that evidence?
Yes it overfits some of the times as the training accuracy is higher than the testing accuracy for some runs.

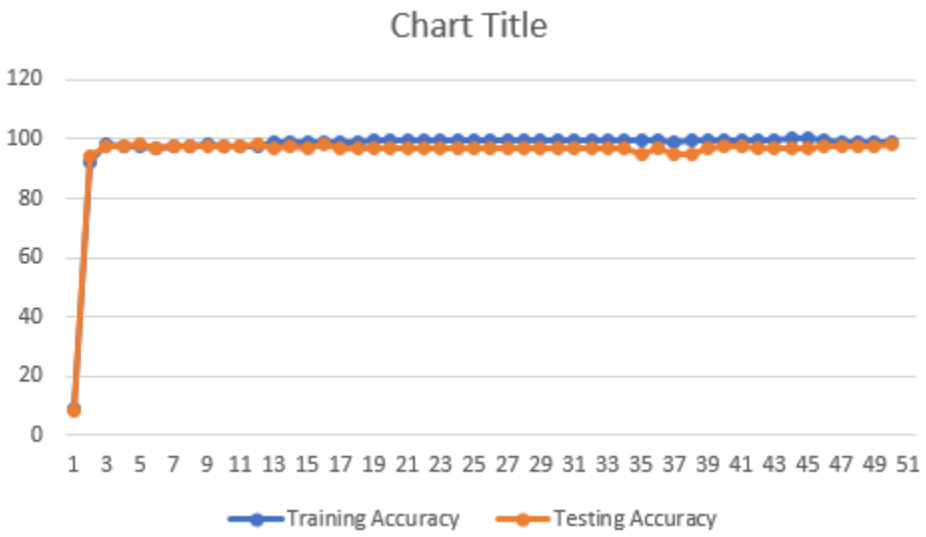
(4) How do your results compare to the results obtained by your perceptron in HW 1?
The Accuracy range is higher as compared to the perceptron. We were getting accuracy of around 80-90% in perceptron and here we are getting 93-97 %

Experiment 2-

Description-

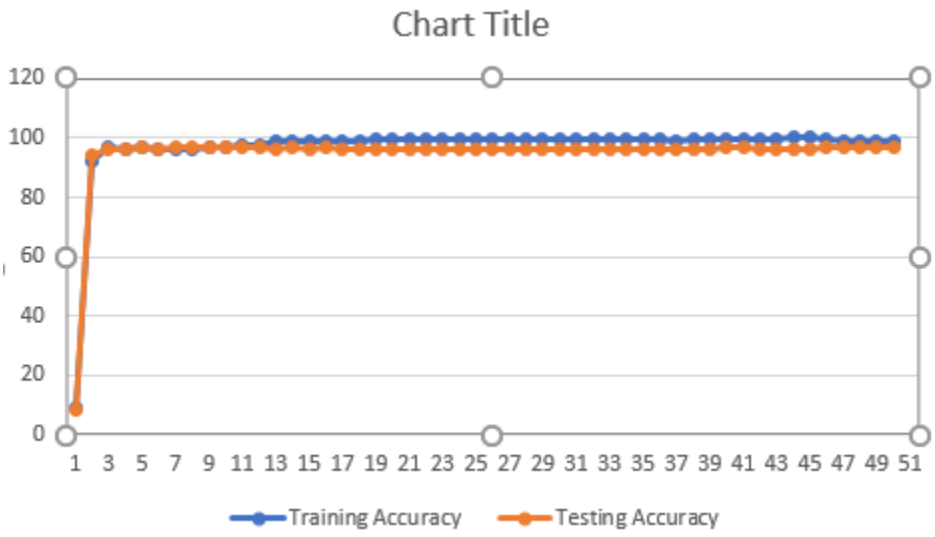
In this experiment we have to check the accuracy of neural network by changing momentum, keeping the number of hidden layers and learning rate same i.e. 100 and .1 resp.. We have tested on three momentum that are 0, .25 and .9 and calculated testing and training accuracies after each epoch and plotted them on the graph, also we have created confusion matrix at the end of the training on the test data. Data set used is mnist data for the hand written digit recognition we have scaled the data by dividing by 255 on the scale of 0,1. There were 784 inputs and 10 outputs the image is of size 28* 28.There were 60000 training data and testing data is 10000.

Plots and matrix for momentum 0-



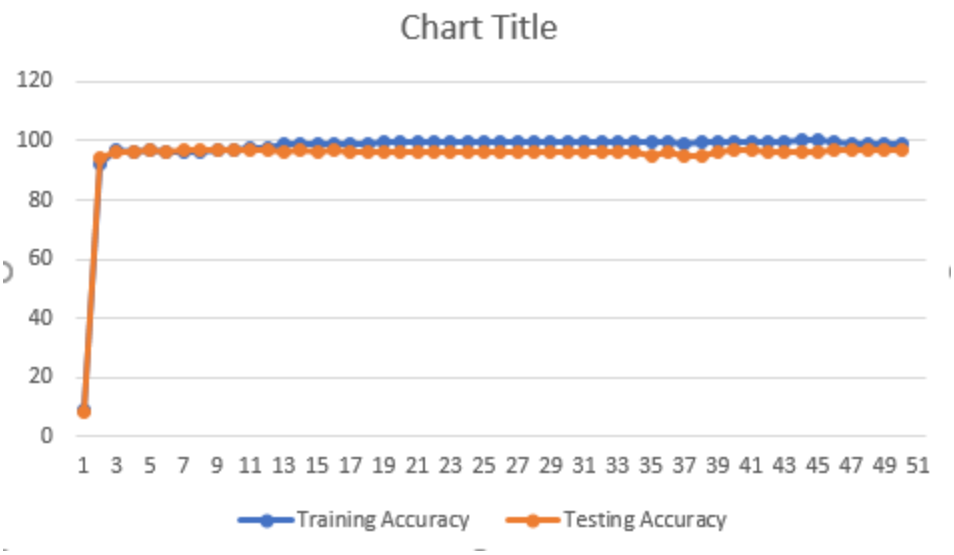
970	0	0	0	1	0	3	1	4	1
0	1113	2	2	0	0	4	1	12	1
5	2	995	9	1	0	4	6	9	1
0	0	5	976	0	10	1	4	10	4
2	0	2	0	943	0	6	1	3	25
3	0	0	12	1	858	5	2	6	5
7	2	2	1	0	3	938	0	5	0
1	4	12	3	4	0	0	985	6	13
3	2	0	13	3	5	1	2	940	5
4	6	0	11	10	2	0	4	3	969

Plots and matrix for momentum .25



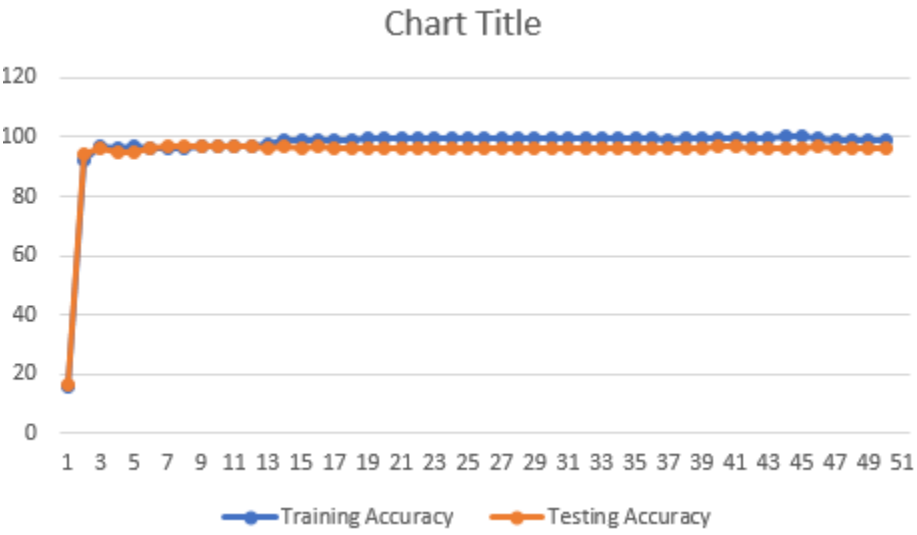
970	0	1	0	1	0	1	1	4	2
0	1118	3	2	1	1	3	1	6	0
5	2	1000	4	2	0	2	7	8	2
0	0	3	981	0	5	0	6	9	6
2	1	0	0	939	0	10	0	3	27
5	2	0	22	2	842	5	1	7	6
8	2	3	1	3	10	927	0	4	0
2	3	8	5	4	0	0	986	4	16
4	2	2	11	2	7	3	2	936	5
4	6	0	8	6	4	1	4	3	973

Plots and matrix for momentum .5-



964	0	0	0	4	2	3	1	4	2
0	1103	2	3	0	1	4	0	22	0
15	3	928	17	9	8	14	7	26	5
9	0	8	924	0	19	4	7	32	7
1	3	4	0	895	0	8	2	7	62
9	3	1	23	7	800	15	3	24	7
19	5	2	2	3	13	905	0	7	2
2	5	13	10	3	0	4	955	7	29
12	7	2	11	13	17	9	2	895	6
9	7	1	16	17	2	1	11	3	942

Plots and matrix after 100 hidden layers(96.8-Final accuracy)-



967	0	2	0	0	1	3	1	4	2
0	1117	4	2	0	1	3	2	6	0
4	1	989	11	4	0	6	7	8	2
0	0	5	983	0	5	1	5	8	3
1	0	1	0	946	0	7	1	1	25
5	1	2	28	1	817	15	3	14	6
7	4	2	2	1	3	933	0	6	0
3	3	7	6	3	0	0	986	2	18
5	4	1	10	4	6	1	3	935	5
3	5	1	11	9	2	0	2	4	972

(1) How does the momentum value affect the final accuracy on the test data?
The accuracy value was higher for network with no momentum (0) and there was little difference between the other momentum accuracy .25-96.7, .5-96.8 and for .9-96.89

(2) How does it affect the number of epochs needed for training to converge?
The one with higher momentum converged earlier..9 around 20 epochs,.5-24 .25-31 and 0-35

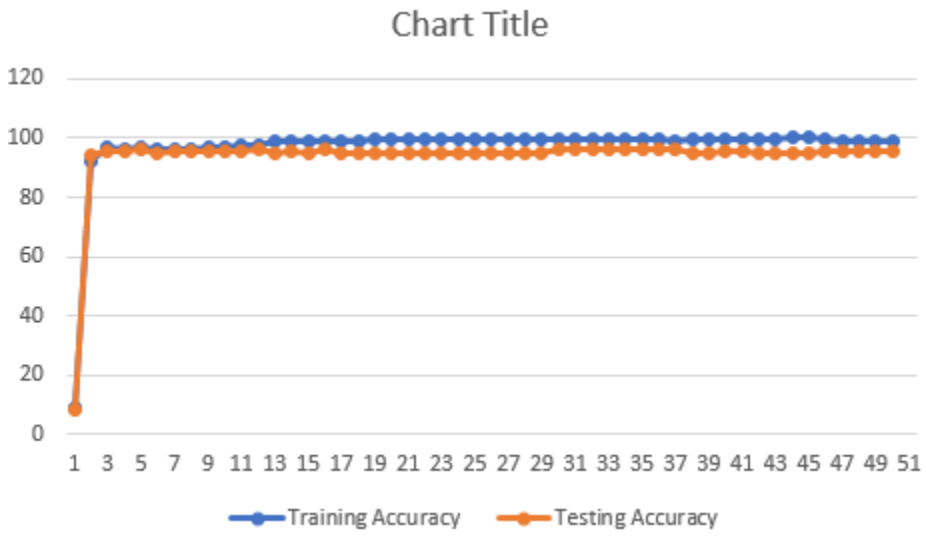
(3) Again, is there evidence that any of your networks has overfit to the training data? If so, what is that evidence?
Yes there is overfitting as the training accuracy is more than testing by around 1-2%

Experiment 3-

Description-

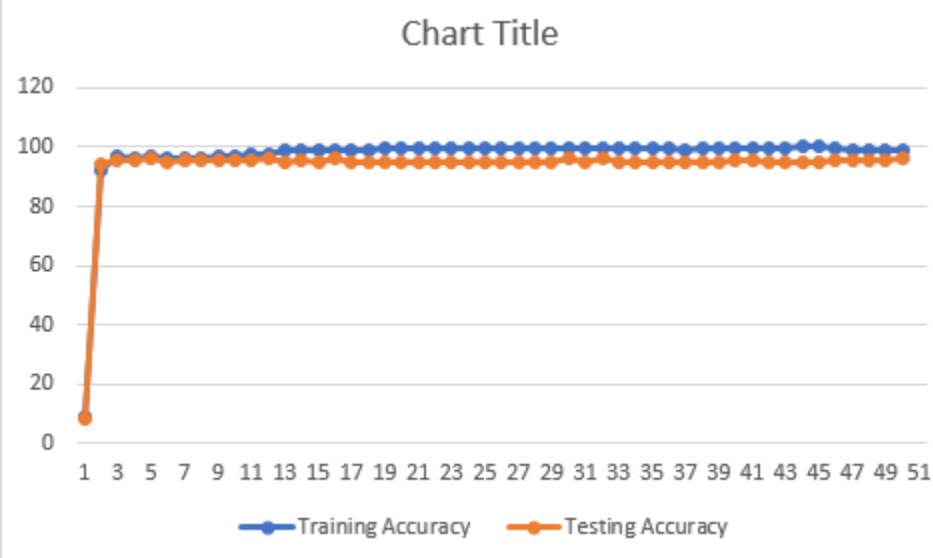
In this experiment we have to check the accuracy of neural network by changing size of the training data, keeping the number of hidden layers and learning rate same and momentum same i.e. 100 and .1 and .9 resp.. We have tested on two sets one half and one quarter and calculated testing and training accuracies after each epoch and plotted them on the graph, also we have created confusion matrix at the end of the training on the test data. Data set used is mnist data for the hand written digit recognition we have scaled the data by dividing by 255 on the scale of 0,1. There were 784 inputs and 10 outputs the image is of size 28* 28.There were 15000 training data and testing data is 10000 for first training and 30000 training data and 10000 test data for second one.

Plots and matrix for one quarter-



```
960      0      1      0      1      1      4      3      8      2
  0   1116      5      3      0      1      2      1      7      0
  4      0    997      4      0      0      2      8     15      2
  2      1     18    947      1     13      1      3     15      9
  1      2      2      0    921      1      8      0      3     44
  4      2      1     17      0    839      7      3     14      5
  9      4      1      1      4      7    925      0      7      0
  1      2     12      5      6      0      0    981      2     19
  6      1      6     10      4      5      5      4    926      7
  4      7      1      9      8      3      0      4     12    961
```

Plots and matrix for One half-




```
970      0      0      0      1      0      3      1      5      0
  0   1106      4      2      2      1      5      1     14      0
  7      3    987     10      1      0      4      8     11      1
  1      1     10    961      0      7      0      6     15      9
  2      1      2      0    958      0      5      1      2     11
  4      0      2     23      1    829      7      4     18      4
 11      2      2      1      4      3    928      0      7      0
  3      2     10      7      6      0      0    978      4     18
  7      1      3      9      3      8      1      3    931      8
  8      7      0      5     17      2      0      7     12    951
```

(1) How does the size of the training data affect the final accuracy on the test data?
There is little difference on the test data accuracy at the end of the 50 epochs. The accuracy is around 95.7 in case of one quarter and 95.99 in case of one half.

(2) How does it affect the number of epochs needed for training to converge?
The network with the one quarter(23 epochs) converge earlier as compared to the one half(26 epochs)

(3) Again, is there evidence that any of your networks has overfit to the training data? If so, what is that evidence?
Yes there was overfitting as the training accuracy is higher than the testing accuracy (1-2%)