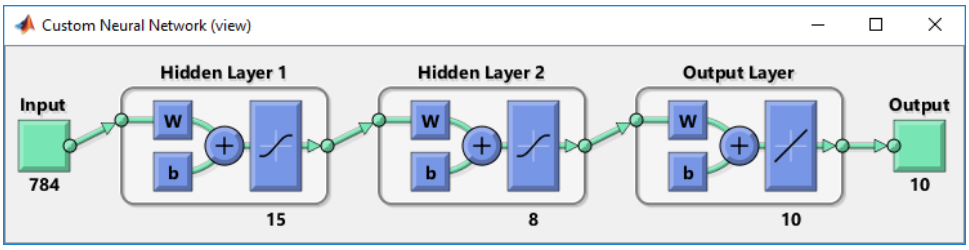
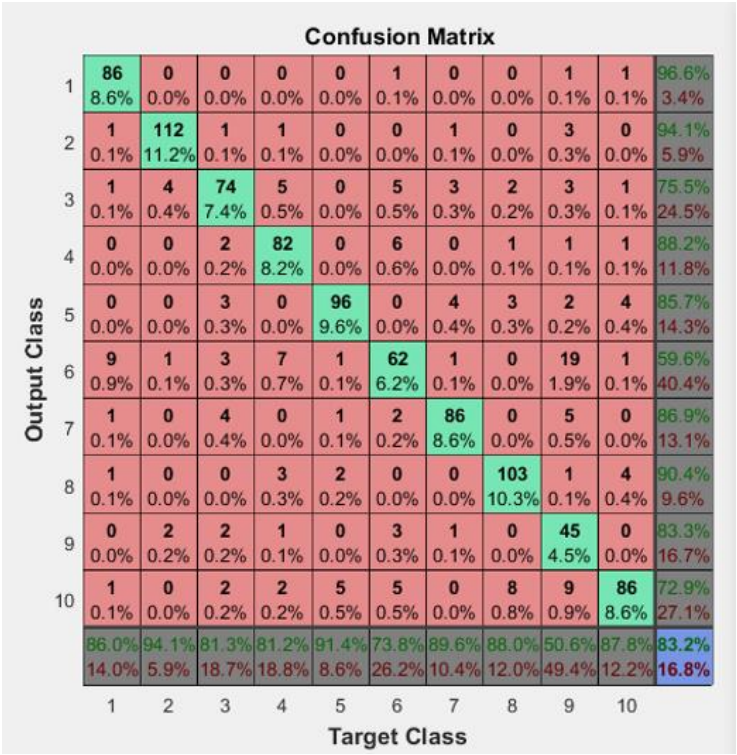
1) Neural Network structure.

This network has 2 hidden layers with 15 and 8 neurons respectively, and the transfer function is sigmoid function, the output layer function is linear function.



This is one of the variable networks, in this network, parameters are set as:
Hidden layers: [15, 8]
Goal: 0.01
Epochs: 1500
Learning rate: 0.05
Max_fail: 5
And the rest of the parameters are default sets.

2) Confusion matrix:



Row:
Output class, also the actual output, these 10 classes represent number 0-9(1000000000 - 0000000001), and it is the test result obtained by putting test data into the network.

Column:
Target class, is the expected result specified by test_label, represents label 1000000000 – 0000000001, and referred from 'label'.

Cell:
Integers in the cell: Number of results that are expected to be target A but actually output B, for example, for target class 1 and output class 10, the integer is 1, this means there is one data with label 1000000000 was recognized to be the data with label 0000000001.
The red cells are wrong results and green cells are right results.
Percentage below the integers: ratio of integer in this cell and sum of all integers.

Gray cells:

The green percentages are rate of correctness, vertical column for each row's correctness and horizontal row for each column's correctness.
The red percentages are rate of falsity.
The bottom right corner cell colored blue is the overall accuracy and the overall error rate.

In this matrix, the overall accuracy is 83.2%, most of the data are evaluated correctly by the network, which are those in the green cells, and some data are evaluated falsely as in the red cells.
I adopted this result, because from the experiment below, it has the highest accuracy.

3) Experiment

This is a controlled experiment, for each time, only one item is evaluated by controlling the other items the same.

Test item	No.	Goal	Epoch	Learning rate	Max fail	Neurons	Hidden layers	Time	Accuracy
Goal	1	0.05	300	0.05	5	[15, 8]	2	0:06:30	67.8%
	2	0.01	300	0.05	5	[15, 8]	2	0:20:04	72.3%
	3	Default	Default	Default	Default	[15, 8]	2	0:25:08	71.2%
Epoch	4	0.01	1500	0.05	5	[15, 8]	2	0:21:40	83.2%
Neuron (hidden layers)	5	0.01	1500	0.05	5	[23]	1	0:33:31	79.6%
	6	0.01	1500	0.05	5	[7, 7, 9]	3	0:07:13	72.2%
	7	0.01	1500	0.05	5	[5, 8, 5, 5]	4	0:03:21	47.4%
	8	0.01	1500	0.05	5	[5, 3, 6, 4, 5]	5	0:02:24	38.4%
Learning rate	9	0.01	1500	0.01	5	[15, 8]	2	0:16:44	72.8%
	10	0.01	1500	0.08	5	[15, 8]	2	0:18:19	71.5%
Max fail	11	0.01	1500	0.05	2	[15, 8]	2	0:10:57	80.2%
	12	0.01	1500	0.05	7	[15, 8]	2	0:21:25	81.7%
Hidden layers	13	0.01	1500	0.05	5	[5, 8]	2	0:03:21	64.9%
	14	0.01	1500	0.05	5	[24, 24]	2	0:30:08	81.7%

1. Test item 'Goal'
   Keep all the other items the same, by changing Goal, it shows that the accuracy is higher when goal is smaller, and mark it red (No.2).
2. Test item 'Epoch'
   Take the best 'goal' from above, keep all the other items the same, compare No.2 and 4, No.4 is better, and mark it red.
3. Test item 'Neuron'
   Take the best 'goal' and 'epoch' obtained from above, keep all the other items the same, keep the total number of neurons the same and change the arrangement of neurons by assigning them into different layers, compare No.4, 5, 6, 7 and 8, No.4 is the best, then mark it red.
4. Test item 'Learning rate'
   Take the best 'goal', 'epoch' and 'neuron', keep all the other items the same, compare No.4, 9 and 10, still No.4 has the best accuracy
5. Test item 'Max_fail'
   Compare No.4, 11 and 12, No.4 has the best accuracy.
6. Test item 'Hidden layers'
   Compare No.4, 13 and 14, No.4 has the best accuracy.

The result shows that parameters like epoch, learning rate and max fail do not affect the result mush, the accuracies by adjusting these parameters are roughly speaking close. However, it is the structure of the network (neuron and hidden layers) that has a larger effect.