

# **Initial report of homework 2 of CSCE970**

By Jiazheng shi

The homework is to design multi-layer neural network trained with backpropagation derived from exponentiated gradient. This initial report presents training and test results with some figures and tables. Totally, there are six datasets in this experience. For convenient comparison among different classifiers, results of each dataset are showed independently. In this initial report, only Table 1-6 present 95% confidence intervals. The further data on confidence intervals will be listed in my final report.

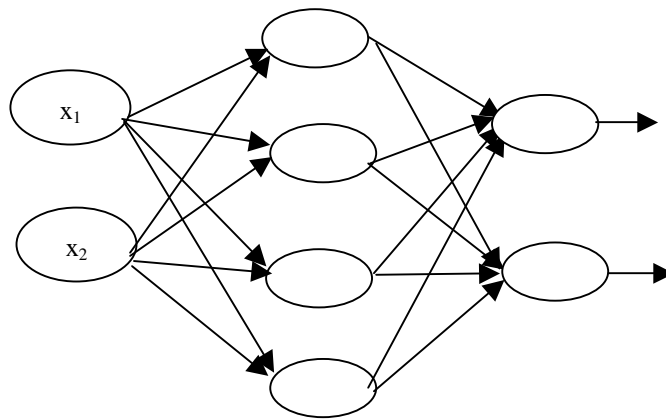
# 1. Dataset one

Table 1 max learning iteration: 10000  
 learning rate ETA: 0.009000  
 sigmoid parameter ALPHA: 6.000000  
 class number: 2

Network structure	Sum of squared error	Network structure versus performance					
		train precision	95% confidence train interval	train time(ms)	test precision	95% confidence test interval	test time (ms)
{2 4 2}	47.266487	0.987	[0.005979, 0.020021]	136677	0.992	[0.005531, 0.010469]	47
{2 4 4 2}	72.085036	0.987	[0.005979, 0.020021]	264452	0.9942	[0.003695, 0.007905]	93

Notes:

network structure {2 4 2} stands for:



network structure {2 4 4 2} stands for:

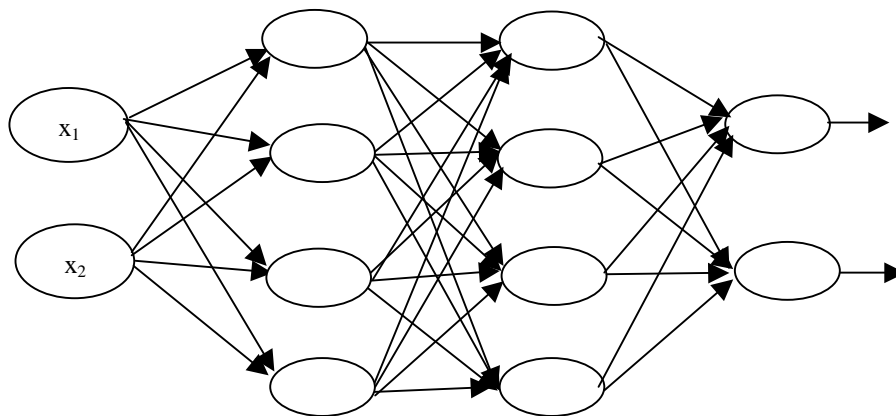


Figure 1-1      learning rate ETA: 0.009000  
sigmoid parameter ALPHA: 6.000000  
class number: 2  
network structure(including input): layerSize[3]={2 4 2 }

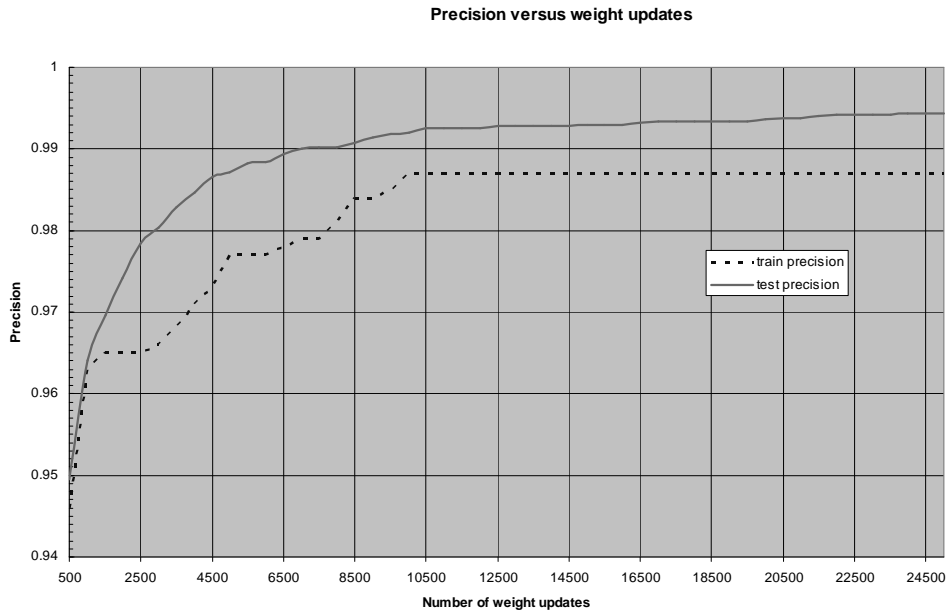


Figure 1-2      learning rate ETA: 0.009000  
sigmoid parameter ALPHA: 6.000000  
class number: 2  
network structure(including input): layerSize[3]={2 4 2 }

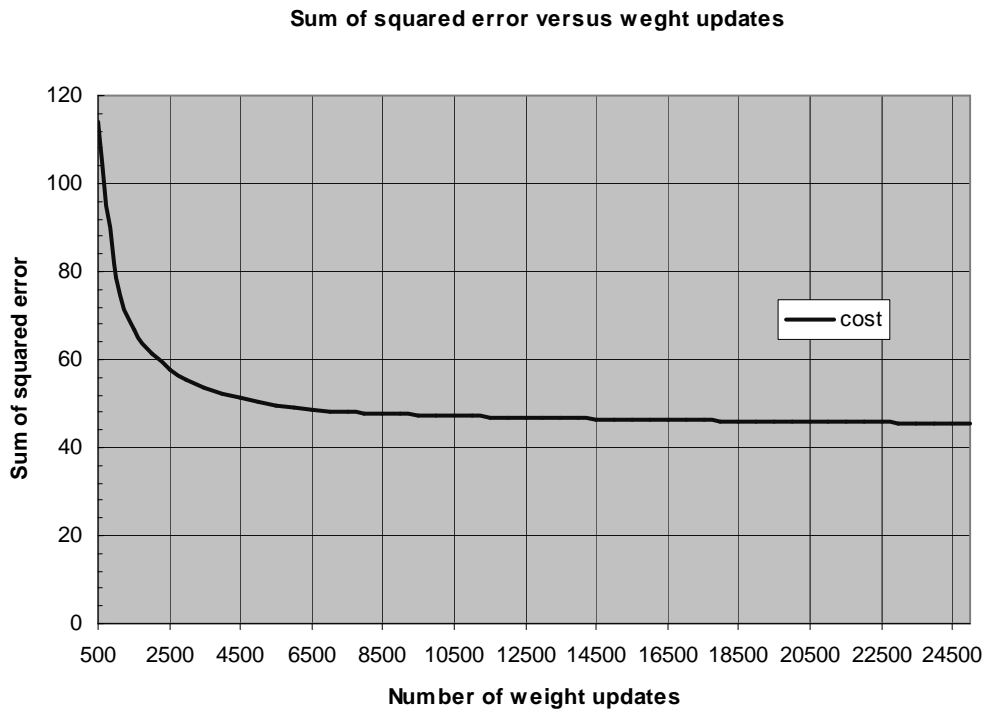


Figure 1-3 max learning iteration: 10000  
sigmoid parameter ALPHA: 6.000000  
class number: 2  
network structure(including input): layerSize[3]={2 4 2 }

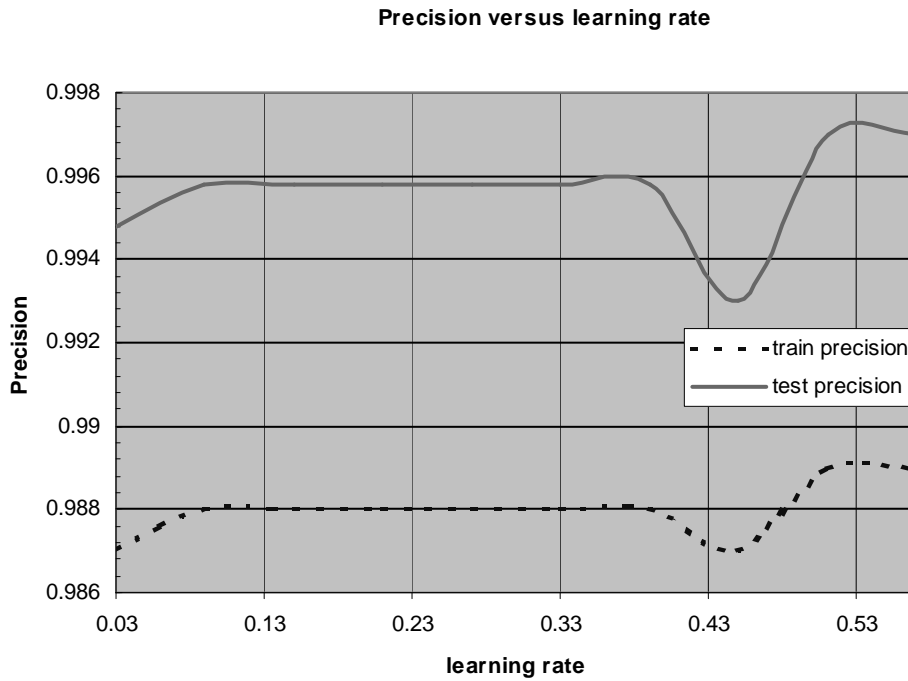


Figure 1-4 max learning iteration: 10000  
learning rate ETA: 0.009000  
class number: 2  
network structure(including input): layerSize[3]={2 4 2 }

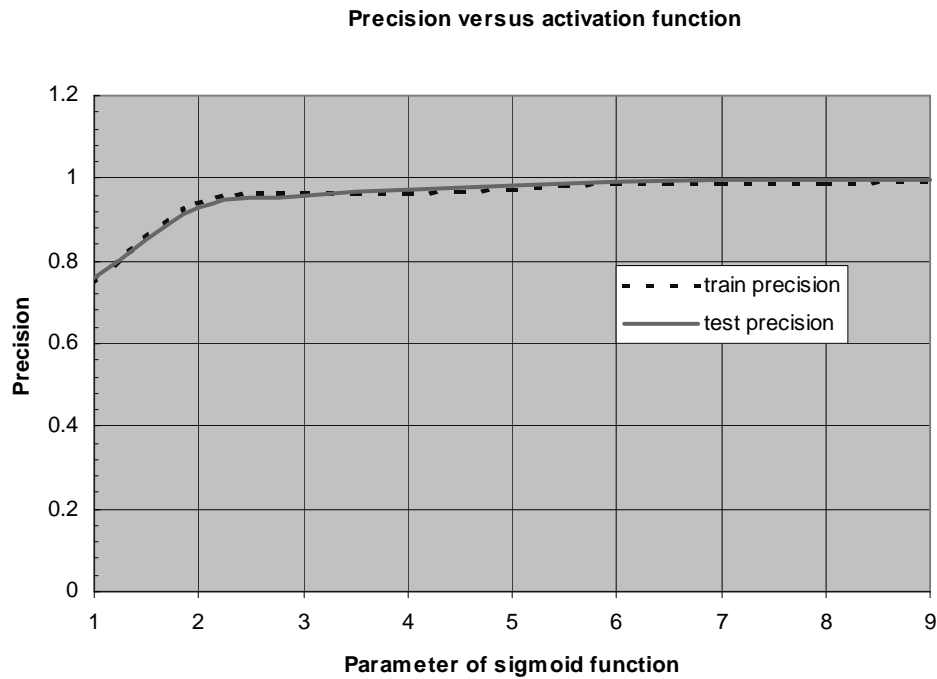


Figure 1-5      max learning iteration: 10000  
 learning rate ETA: 0.009000  
 sigmoid parameter ALPHA: 6.000000  
 class number: 2  
 network structure(including input): layerSize[3]={2 X 2 }

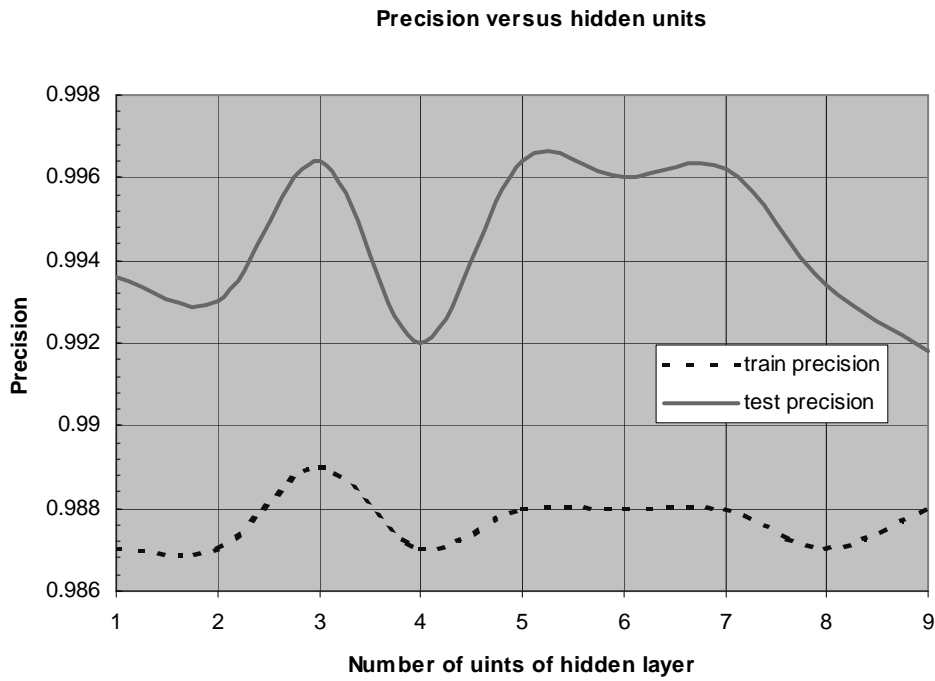
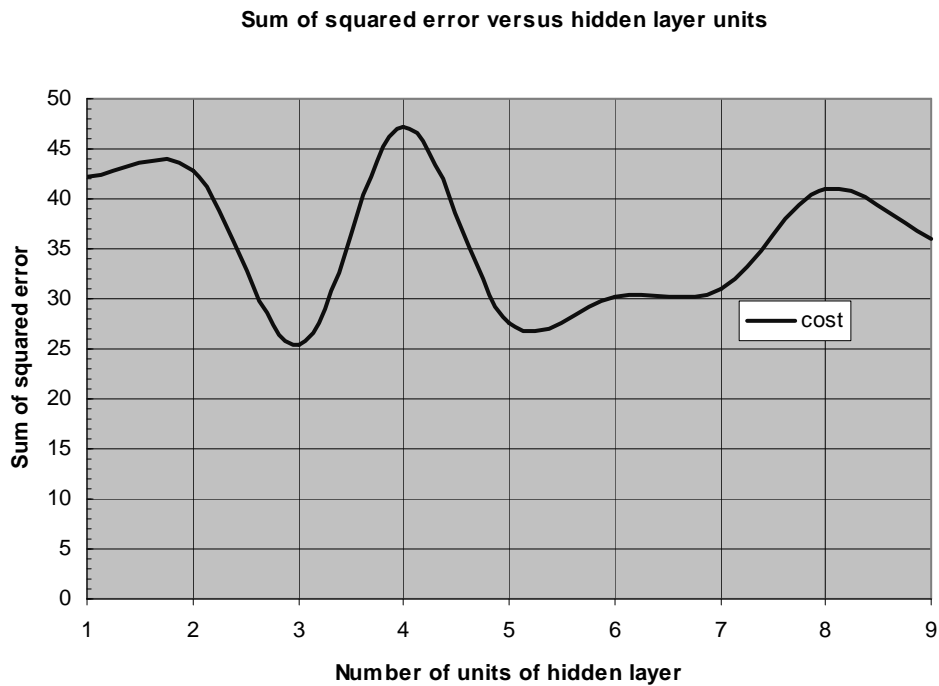


Figure 1-6      max learning iteration: 10000  
 learning rate ETA: 0.009000  
 sigmoid parameter ALPHA: 6.000000  
 class number: 2  
 network structure(including input): layerSize[3]={2 X 2 }



## 2. Dataset two

Table 2 max learning iteration: 10000  
learning rate ETA: 0.009000  
sigmoid parameter ALPHA: 6.000000  
class number: 2

Network structure	Sum of squared error	Network structure versus performance					
		train precision	95% confidence train interval	train time(ms)	test precision	95% confidence test interval	test time (ms)
{2 4 2}	109.376409	0.939	[0.046166, 0.075834]	142380	0.9518	[0.042263, 0.054137]	289
{2 4 4 2}	108.247225	0.94	[0.045280, 0.074720]	272314	0.953	[0.041134, 0.052866]	179

Figure 2-1      learning rate ETA: 0.009000  
sigmoid parameter ALPHA: 6.000000  
class number: 2  
network structure(including input): layerSize[3]={ 2 4 2 }

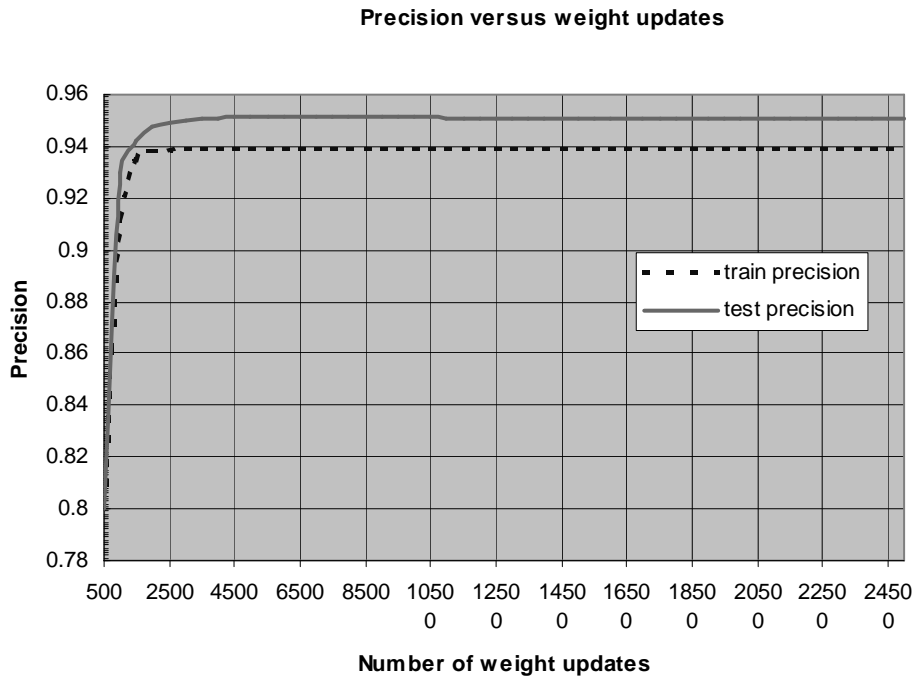


Figure 2-2      learning rate ETA: 0.009000  
sigmoid parameter ALPHA: 6.000000  
class number: 2  
network structure(including input): layerSize[3]={ 2 4 2 }

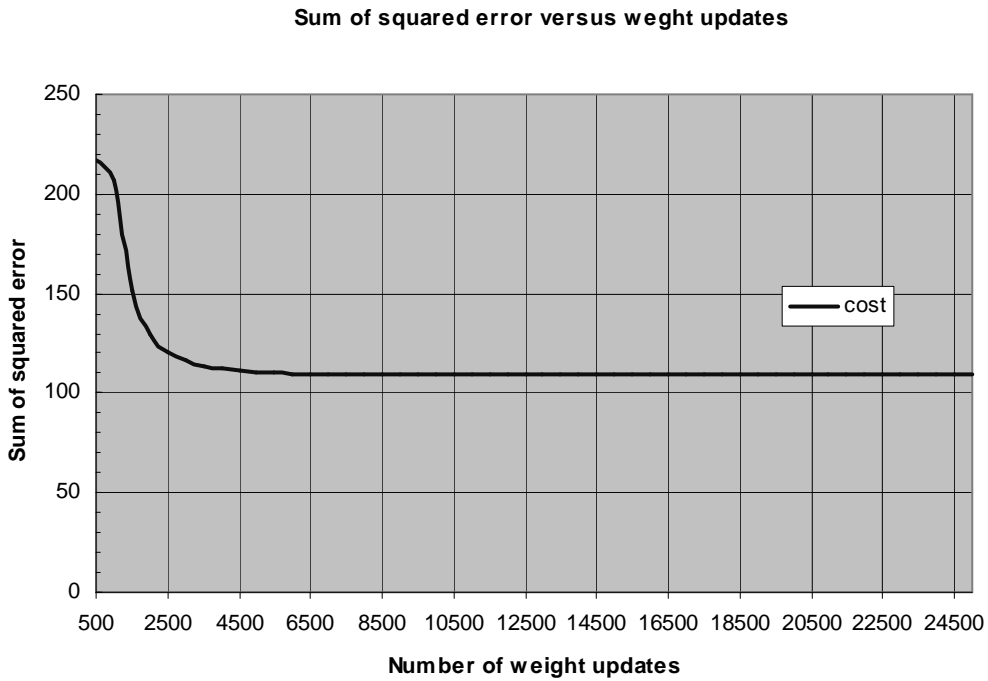


Figure 2-3 max learning iteration: 10000  
sigmoid parameter ALPHA: 6.000000  
class number: 2  
network structure(including input): layerSize[3]={2 4 2 }

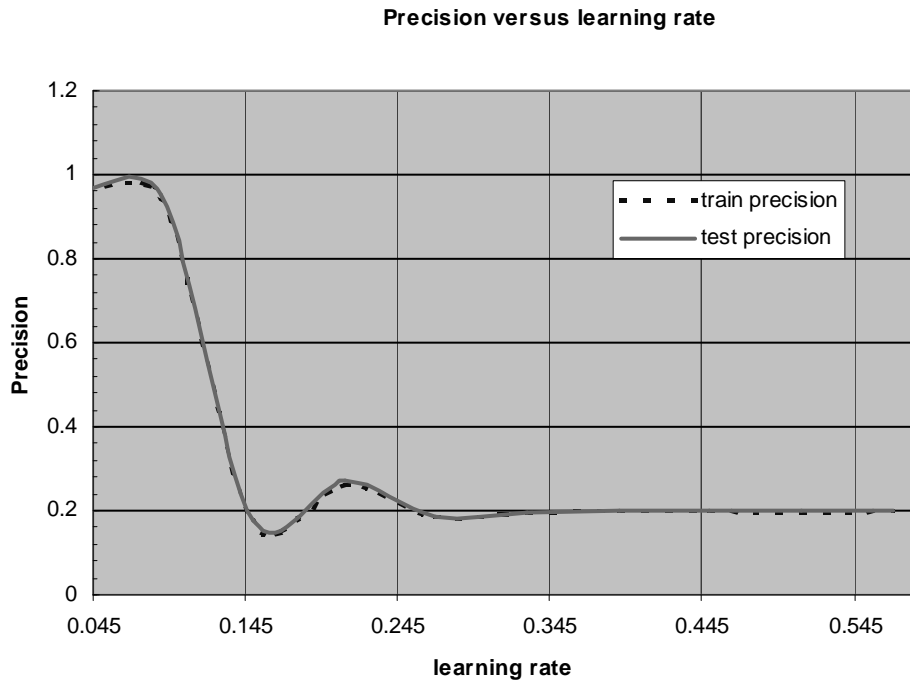


Figure 2-4 max learning iteration: 10000  
learning rate ETA: 0.009000  
class number: 2  
network structure(including input): layerSize[3]={2 4 2 }

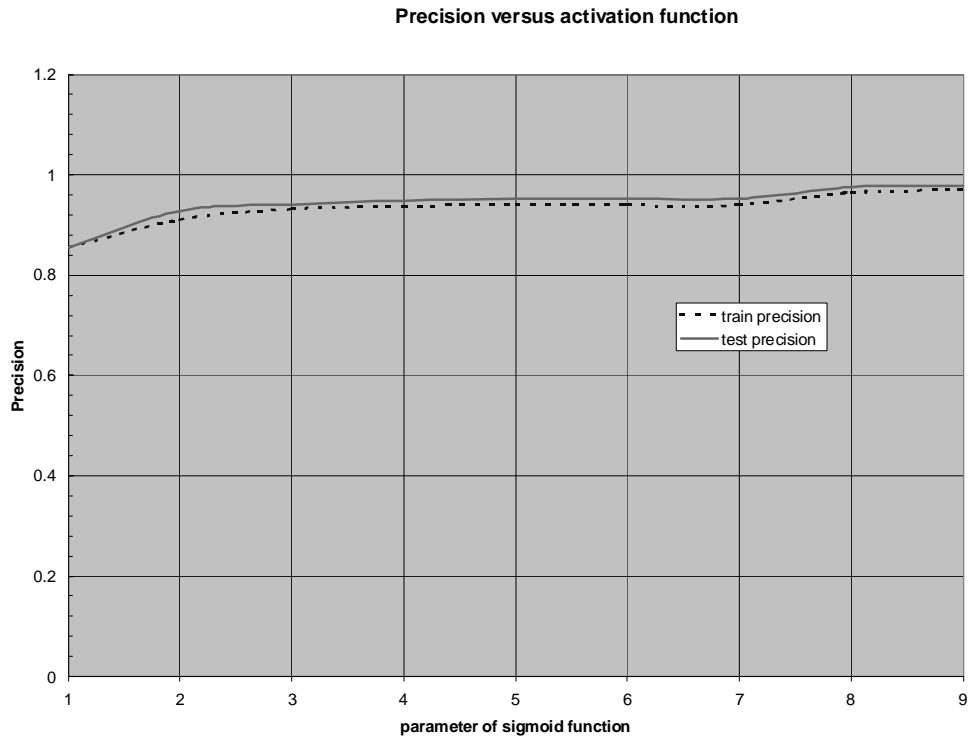




Figure 2-5 max learning iteration: 10000  
learning rate ETA: 0.009000  
sigmoid parameter ALPHA: 6.000000  
class number: 2  
network structure(including input): layerSize[3]={2 X 2 }

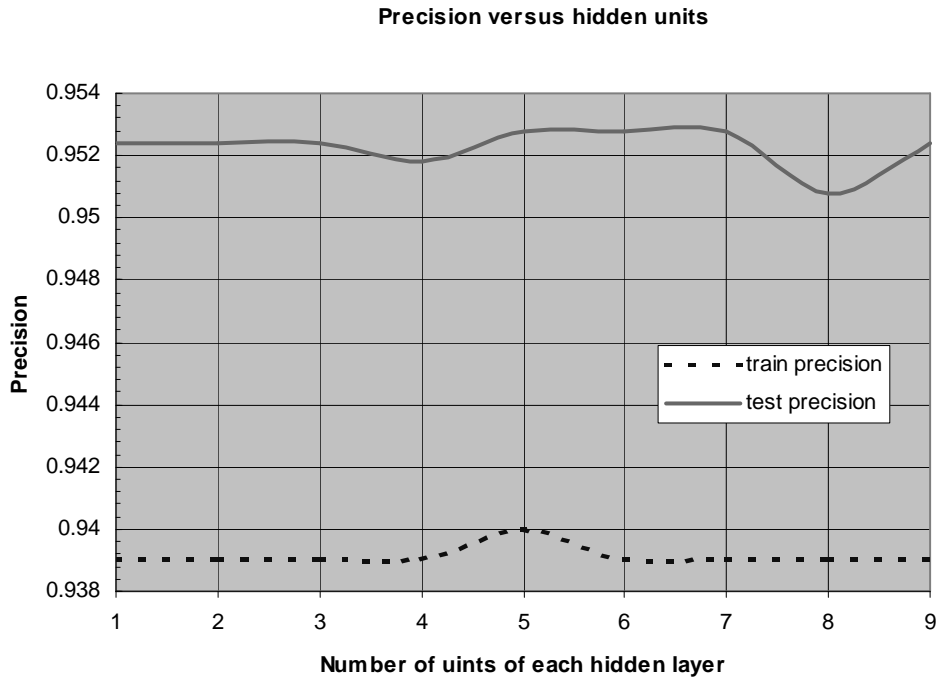
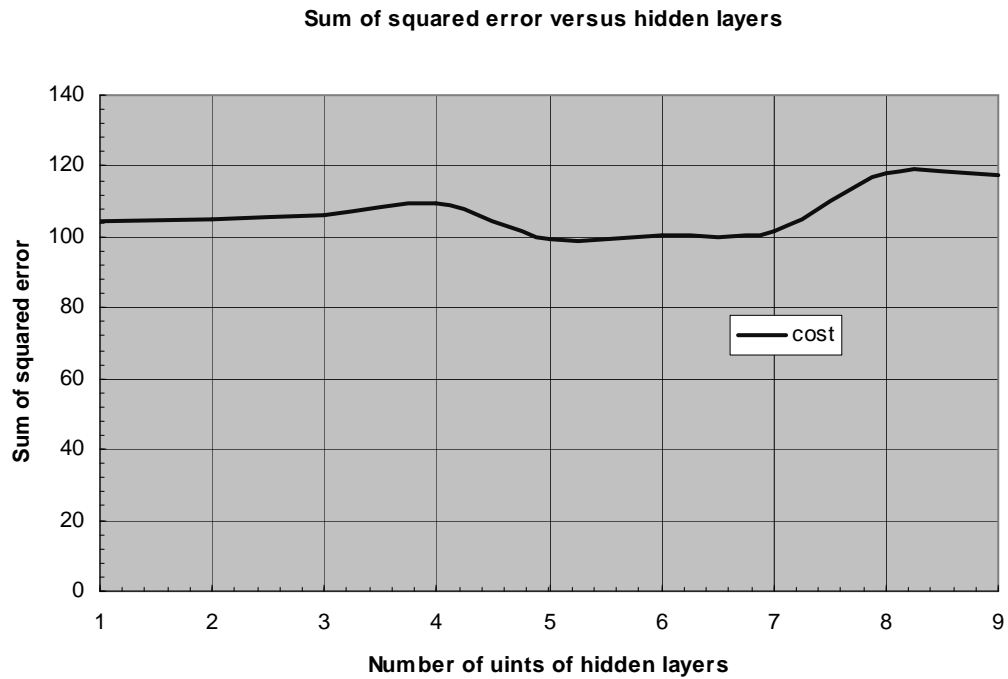


Figure 2-6 max learning iteration: 10000  
learning rate ETA: 0.009000  
sigmoid parameter ALPHA: 6.000000  
class number: 2  
network structure(including input): layerSize[3]={2 X 2 }



### 3. Dataset three

Table 3 max learning iteration: 10000  
learning rate ETA: 0.009000  
sigmoid parameter ALPHA: 6.000000  
class number: 2

Network structure versus performance

Network structure	Sum of squared error	train precision	95% confidence train interval	train time(ms)	test precision	95% confidence test interval	test time (ms)
{2 4 2}	36.336496	0.98	[0.011323, 0.028677]	144712	0.9824	[0.013955, 0.021245]	86
{2 4 4 2}	43.838644	0.978	[0.012908, 0.031092]	281706	0.978	[0.017934, 0.026066]	64

Figure 3-1      learning rate ETA: 0.009000  
sigmoid parameter ALPHA: 6.000000  
class number: 2  
network structure(including input): layerSize[3]={ 2 4 2 }

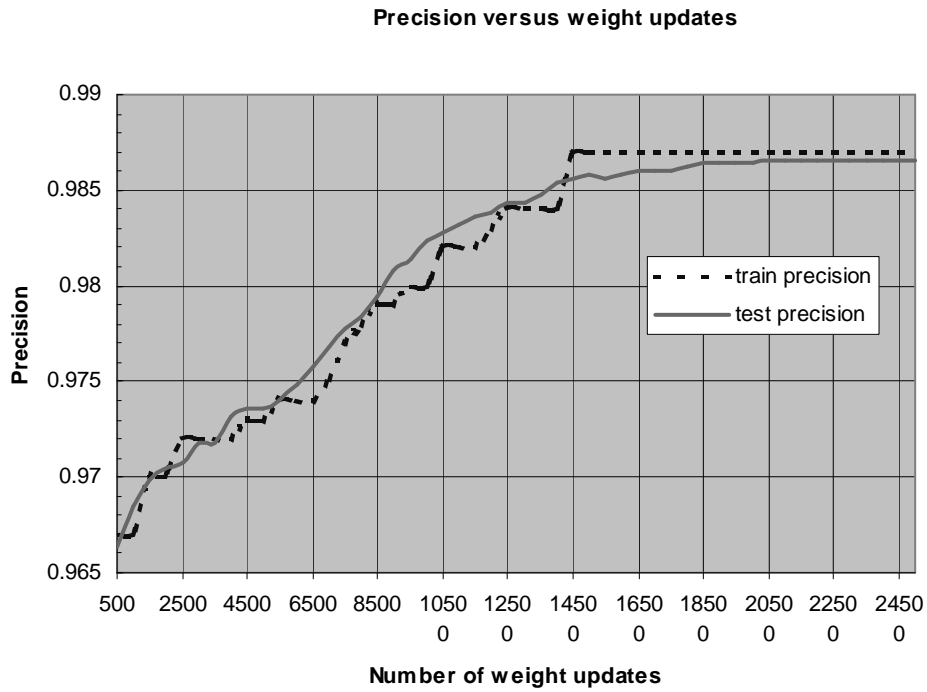


Figure 3-2      learning rate ETA: 0.009000  
sigmoid parameter ALPHA: 6.000000  
class number: 2  
network structure(including input): layerSize[3]={ 2 4 2 }

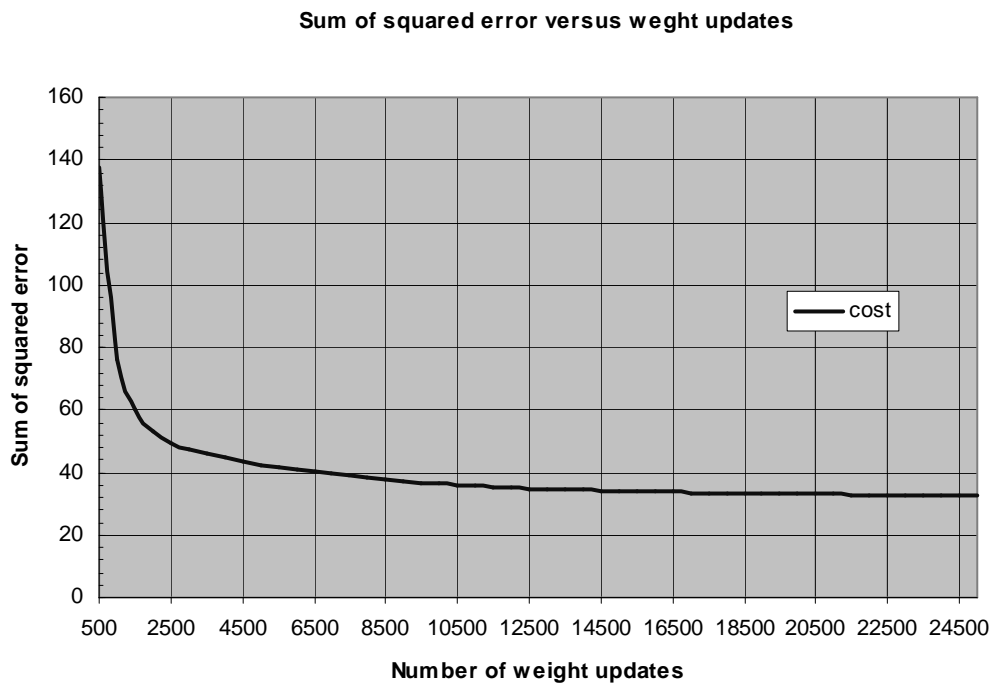


Figure 3-3 max learning iteration: 10000  
 sigmoid parameter ALPHA: 6.000000  
 class number: 2  
 network structure(including input): layerSize[3]={2 4 2 }

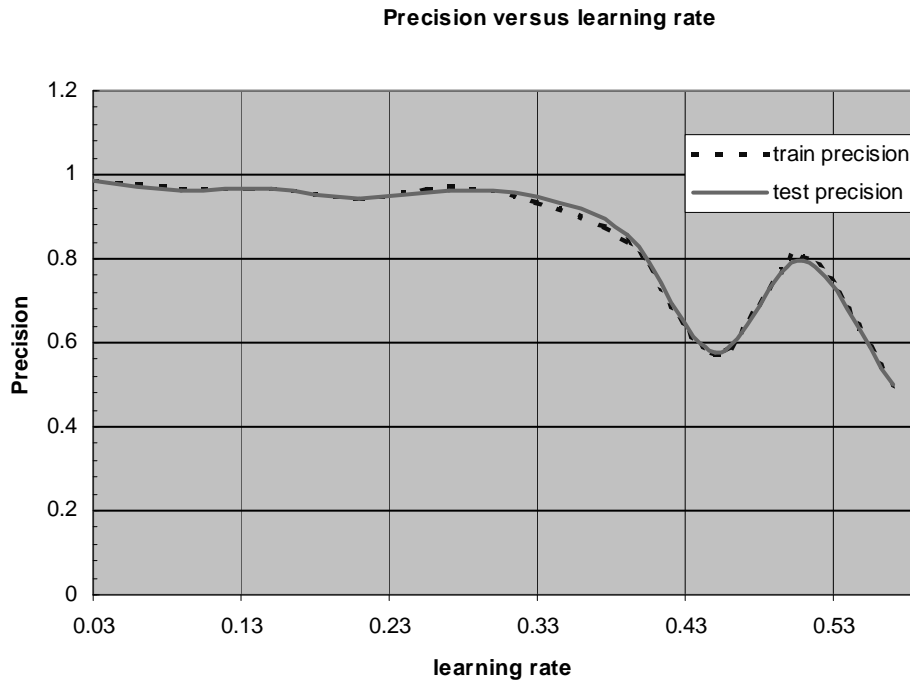


Figure 3-4 max learning iteration: 10000  
 learning rate ETA: 0.009000  
 class number: 2  
 network structure(including input): layerSize[3]={2 4 2 }

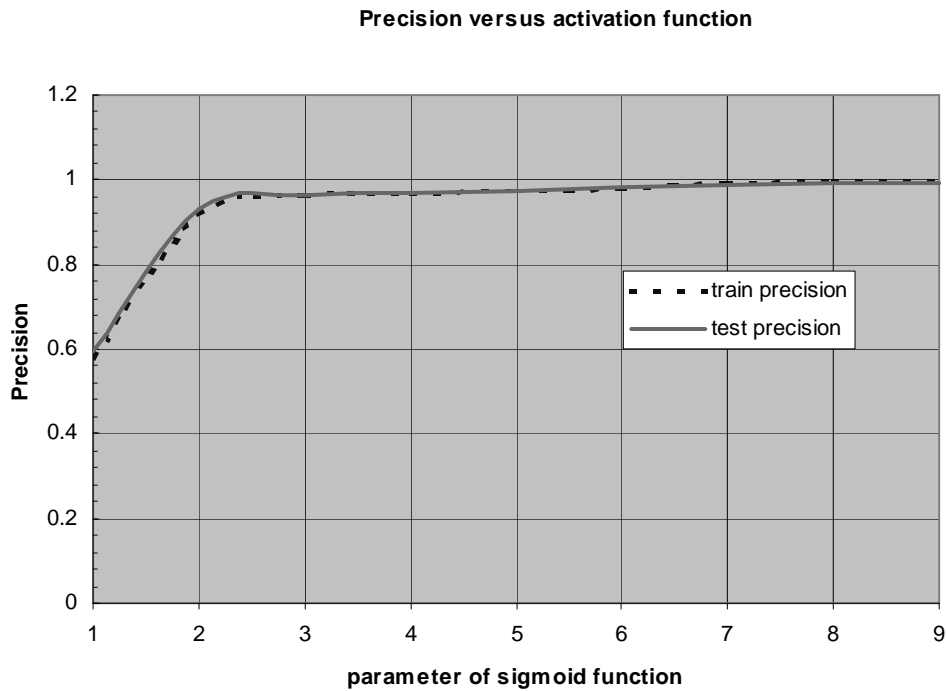


Figure 3-5      max learning iteration: 10000  
 learning rate ETA: 0.009000  
 sigmoid parameter ALPHA: 6.000000  
 class number: 2  
 network structure(including input): layerSize[3]={2 X 2 }

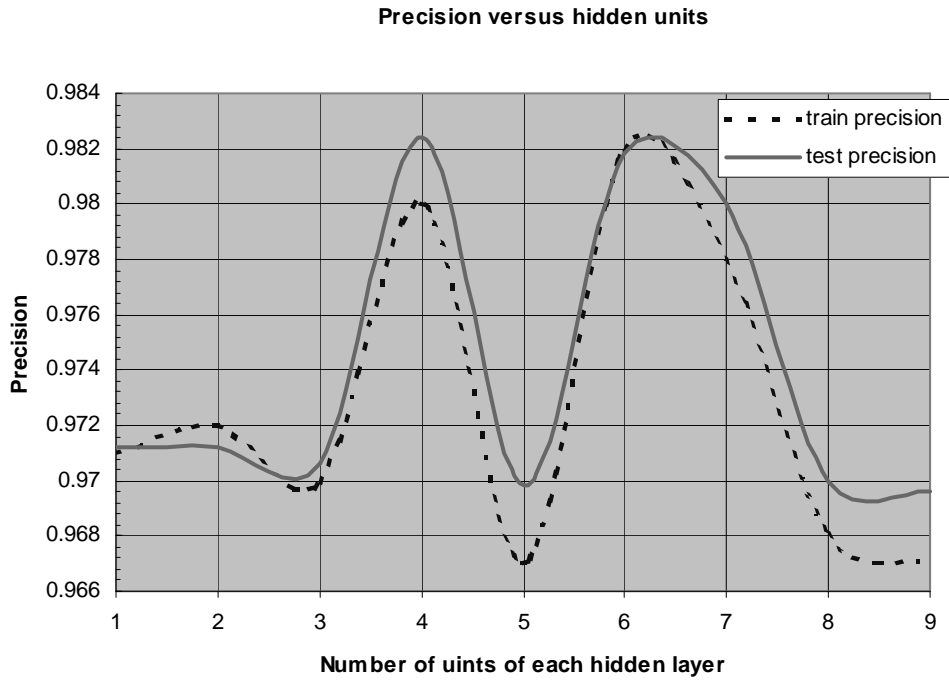
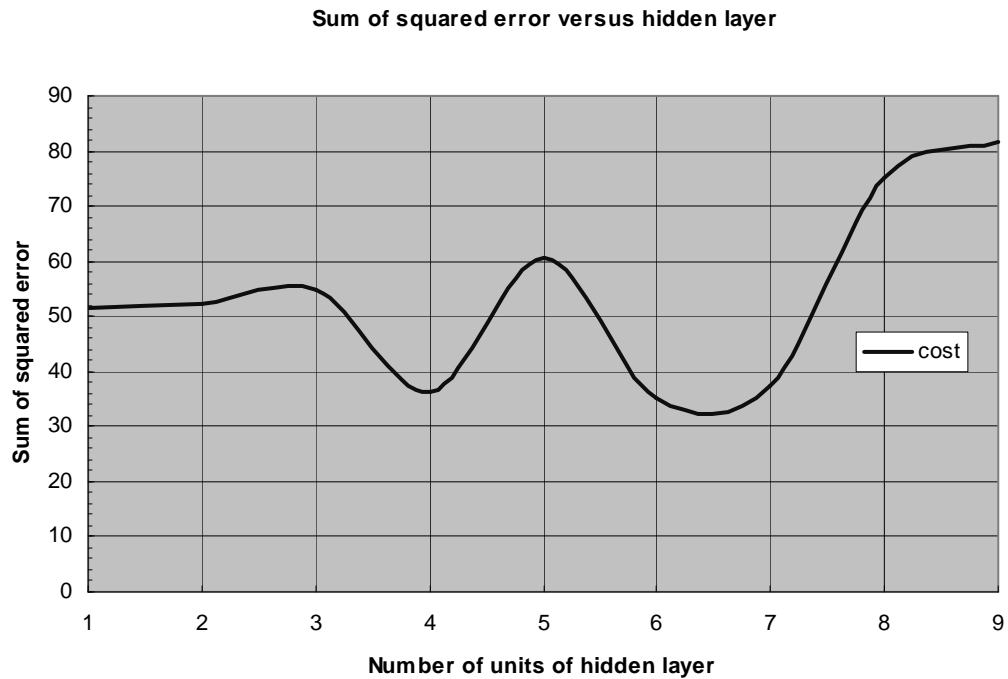


Figure 3-6      max learning iteration: 10000  
 learning rate ETA: 0.009000  
 sigmoid parameter ALPHA: 6.000000  
 class number: 2  
 network structure(including input): layerSize[3]={2 X 2 }



## 4. Dataset four

Table 4 max learning iteration: 10000  
learning rate ETA: 0.009000  
sigmoid parameter ALPHA: 6.000000  
class number: 2

Network structure	Sum of squared error	Network structure versus performance					
		train precision	95% confidence train interval	train time(ms)	test precision	95% confidence test interval	test time (ms)
{2 4 2}	84.033235	0.973	[0.016954, 0.037046]	122624	0.9638	[0.031023, 0.041377]	45
{2 4 4 2}	250.021144	0.4	[0.569636, 0.630364]	217059	0.4	[0.586421, 0.613579]	61

Figure 4-1      learning rate ETA: 0.009000  
sigmoid parameter ALPHA: 6.000000  
class number: 2  
network structure(including input): layerSize[3]={ 2 4 2 }

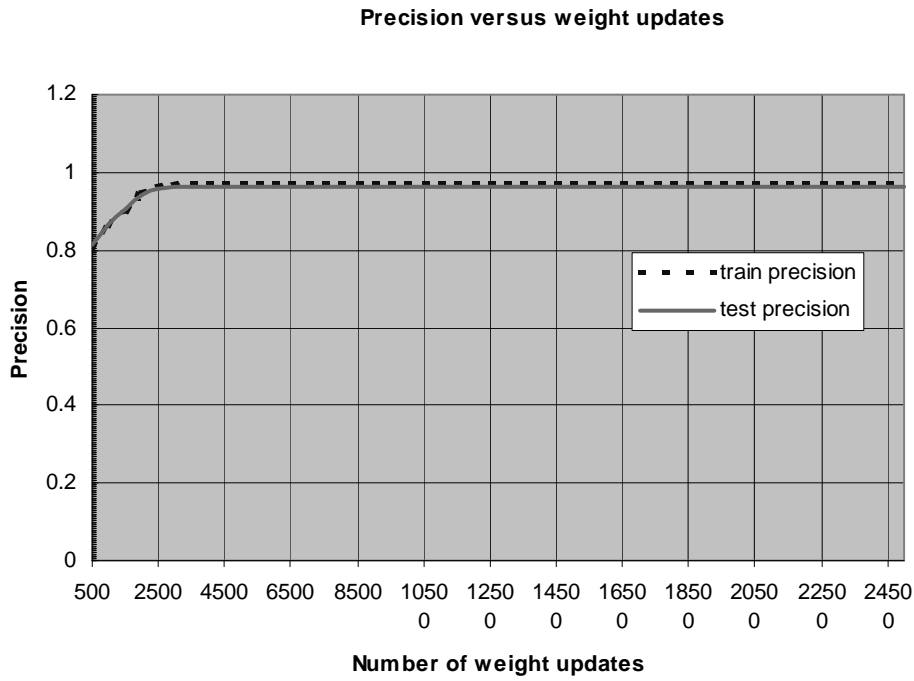


Figure 4-2      learning rate ETA: 0.009000  
sigmoid parameter ALPHA: 6.000000  
class number: 2  
network structure(including input): layerSize[3]={ 2 4 2 }

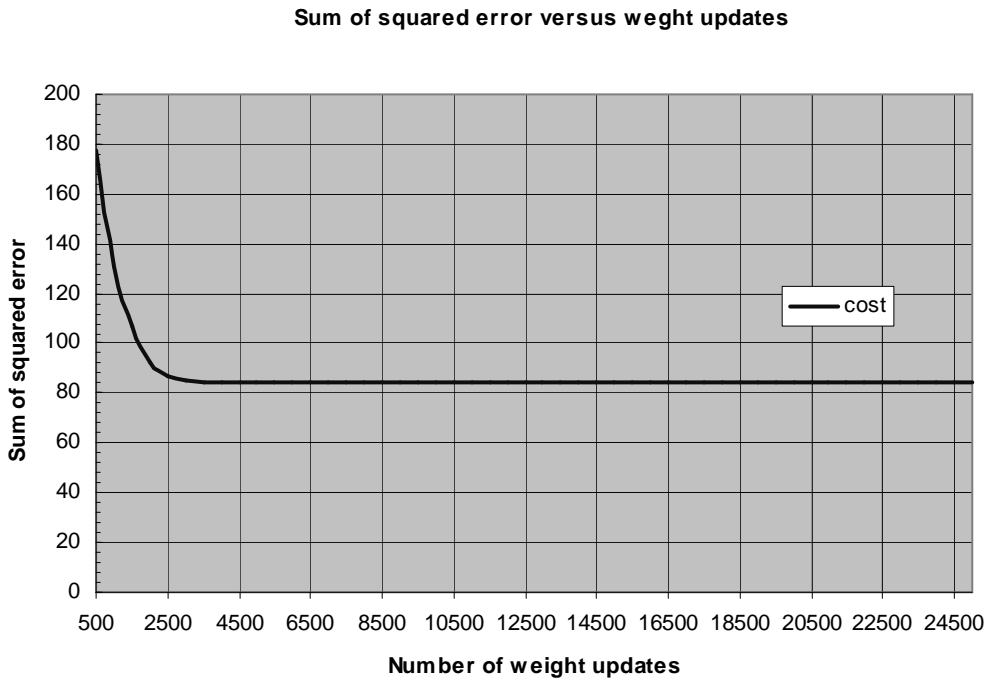


Figure 4-3 max learning iteration: 10000  
sigmoid parameter ALPHA: 6.000000  
class number: 2  
network structure(including input): layerSize[3]={2 4 2 }

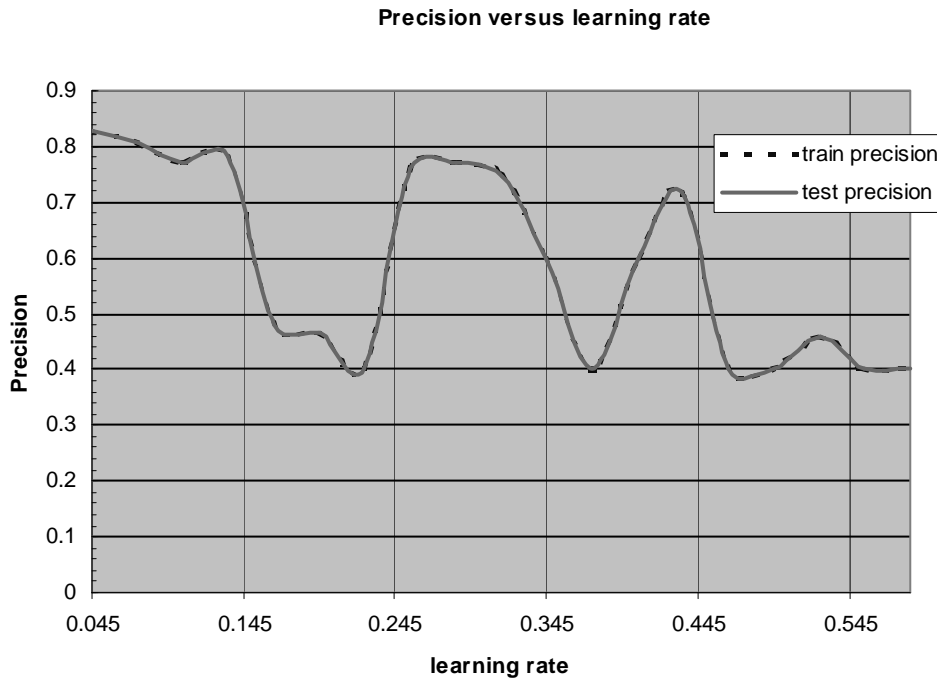


Figure 4-4 max learning iteration: 10000  
learning rate ETA: 0.009000  
class number: 2  
network structure(including input): layerSize[3]={2 4 2 }

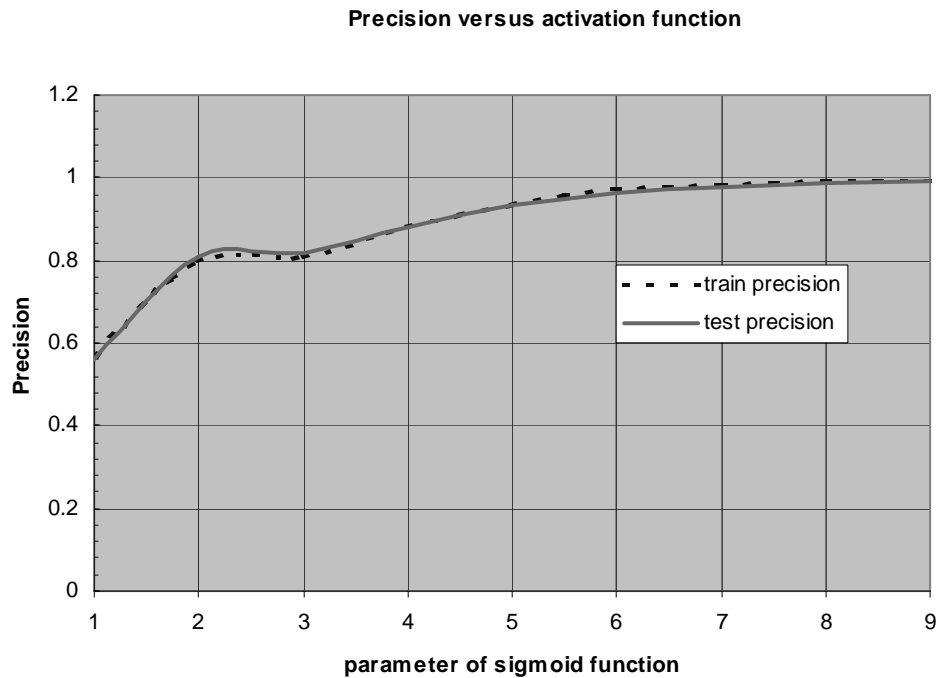




Figure 4-5

max learning iteration: 10000  
learning rate ETA: 0.009000  
sigmoid parameter ALPHA: 6.000000  
class number: 2  
network structure(including input): layerSize[3]={2 X 2 }

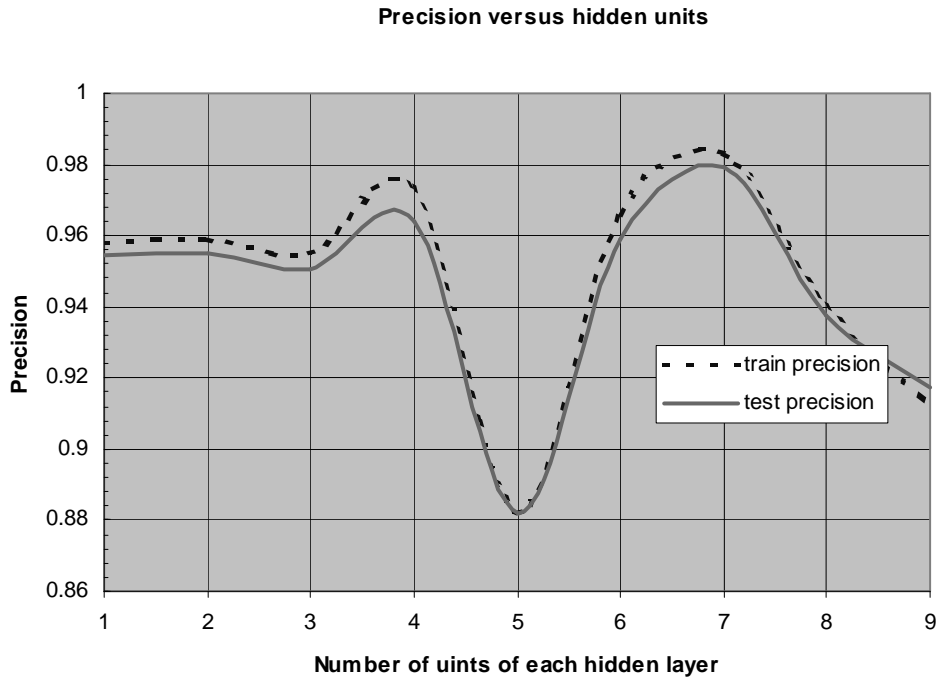
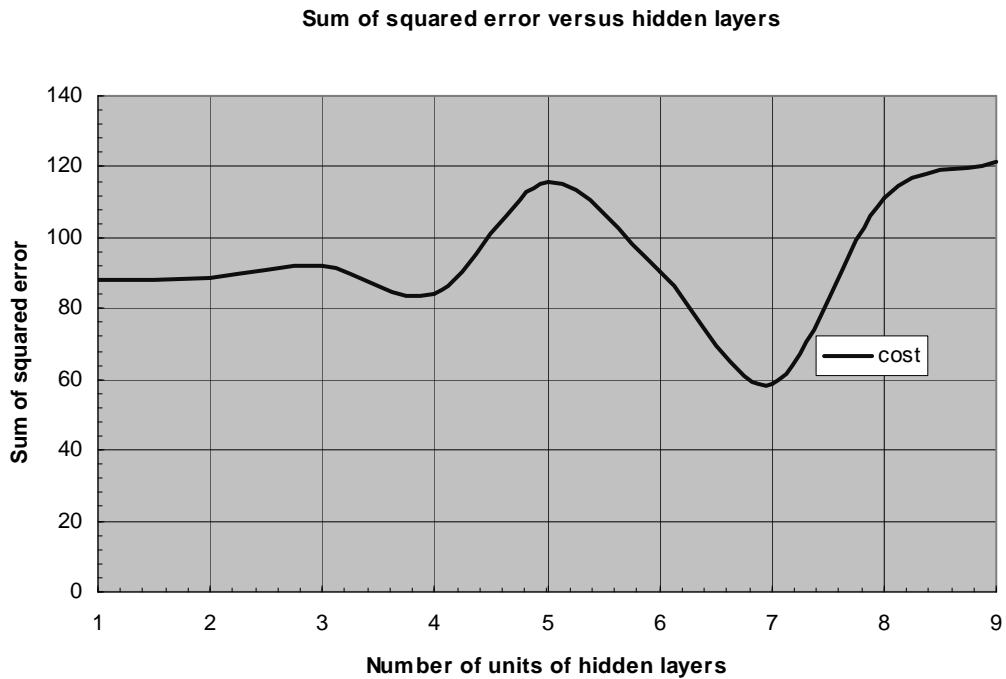


Figure 4-6

max learning iteration: 10000  
learning rate ETA: 0.009000  
sigmoid parameter ALPHA: 6.000000  
class number: 2  
network structure(including input): layerSize[3]={2 X 2 }



## 5. Dataset five

Table 5 max learning iteration: 10000  
learning rate ETA: 0.009000  
sigmoid parameter ALPHA: 6.000000  
class number: 2

Network structure	Sum of squared error	Network structure versus performance					
		train precision	95% confidence train interval	train time(ms)	test precision	95% confidence test interval	test time (ms)
{2 4 2}	196.885561	0.7	[0.271597, 0.328403]	124316	0.7	[0.287298, 0.312702]	45
{2 4 4 2}	249.844284	0.685	[0.286209, 0.343791]	222112	0.684	[0.303113, 0.328887]	62

Figure 5-1      learning rate ETA: 0.009000  
sigmoid parameter ALPHA: 6.000000  
class number: 2  
network structure(including input): layerSize[3]={ 2 4 2 }

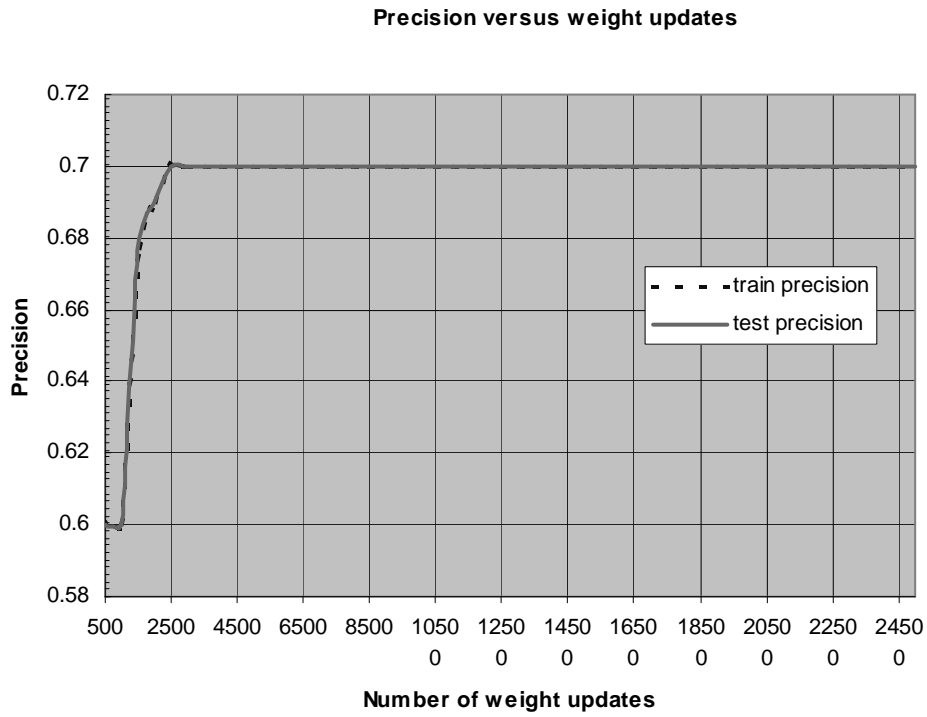


Figure 5-2      learning rate ETA: 0.009000  
sigmoid parameter ALPHA: 6.000000  
class number: 2  
network structure(including input): layerSize[3]={ 2 4 2 }

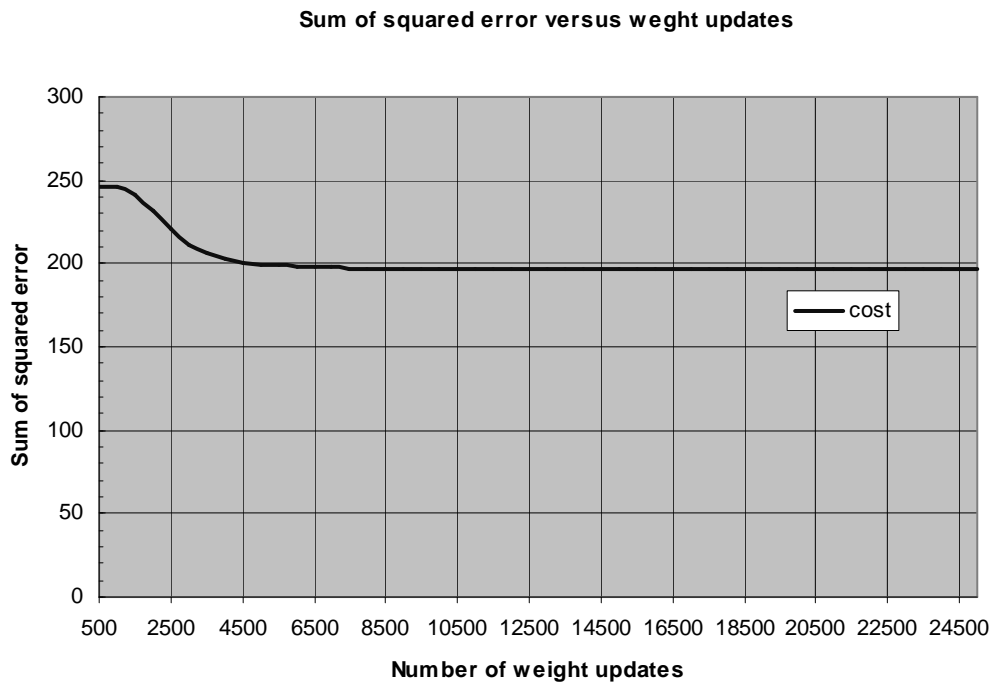


Figure 5-3 max learning iteration: 10000  
 sigmoid parameter ALPHA: 6.000000  
 class number: 2  
 network structure(including input): layerSize[3]={2 4 2 }

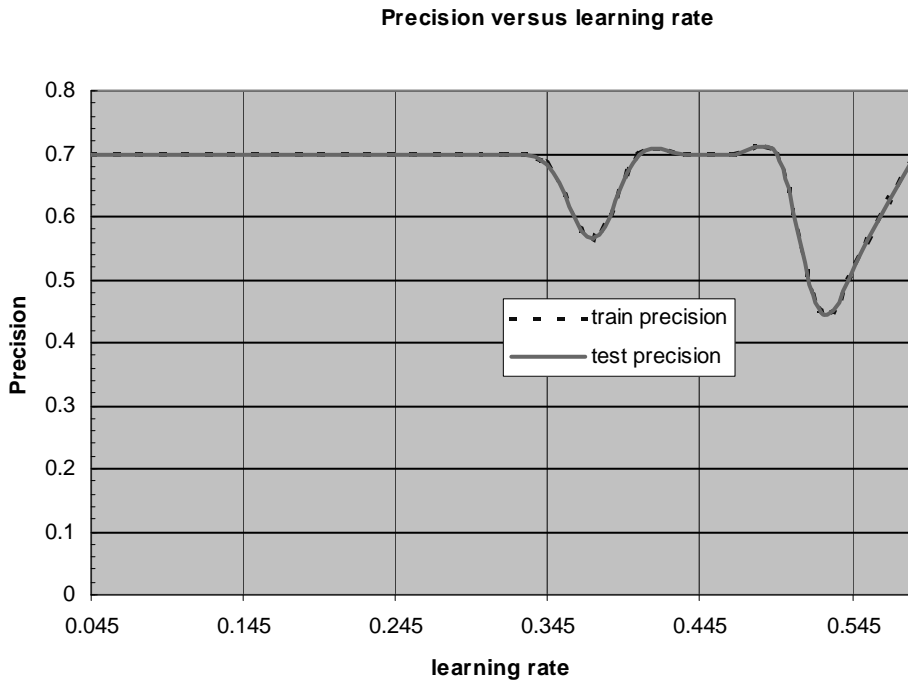


Figure 5-4 max learning iteration: 10000  
 learning rate ETA: 0.009000  
 class number: 2  
 network structure(including input): layerSize[3]={2 4 2 }

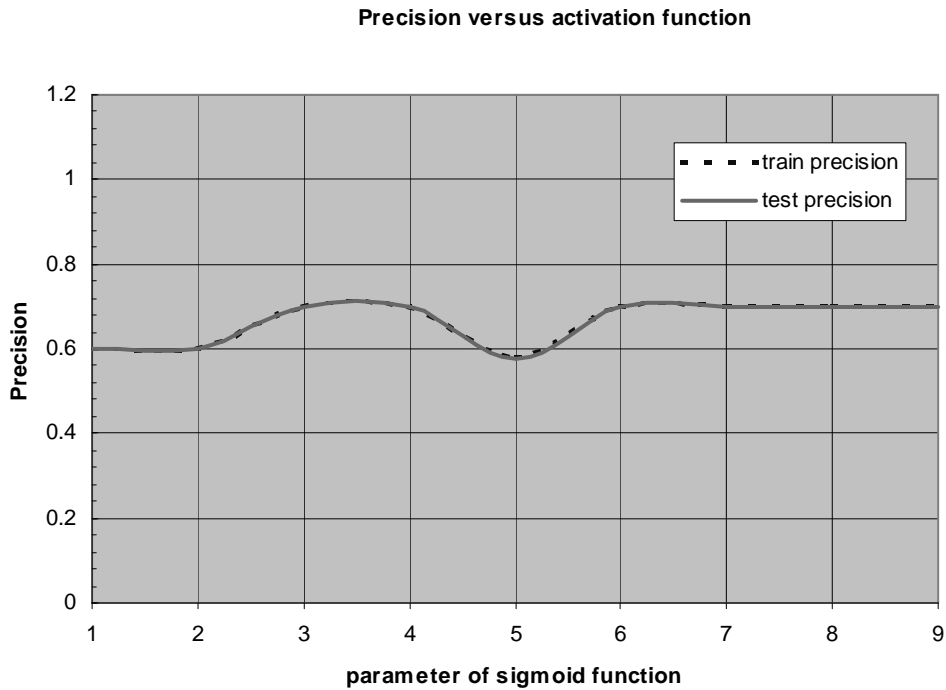


Figure 5-5      max learning iteration: 10000  
learning rate ETA: 0.009000  
sigmoid parameter ALPHA: 6.000000  
class number: 2  
network structure(including input): layerSize[3]={ 2 X 2 }

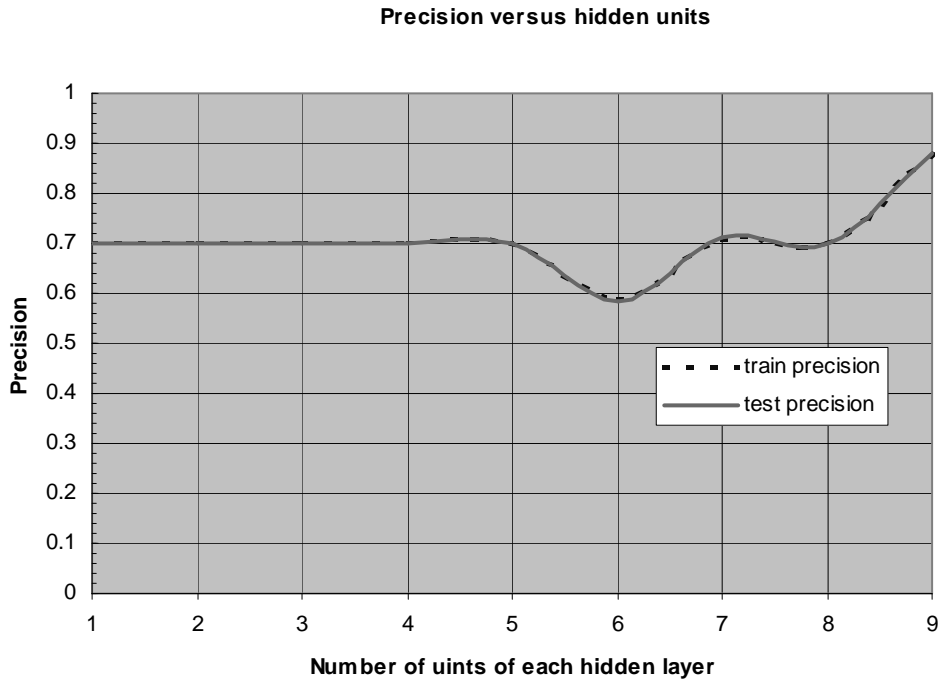
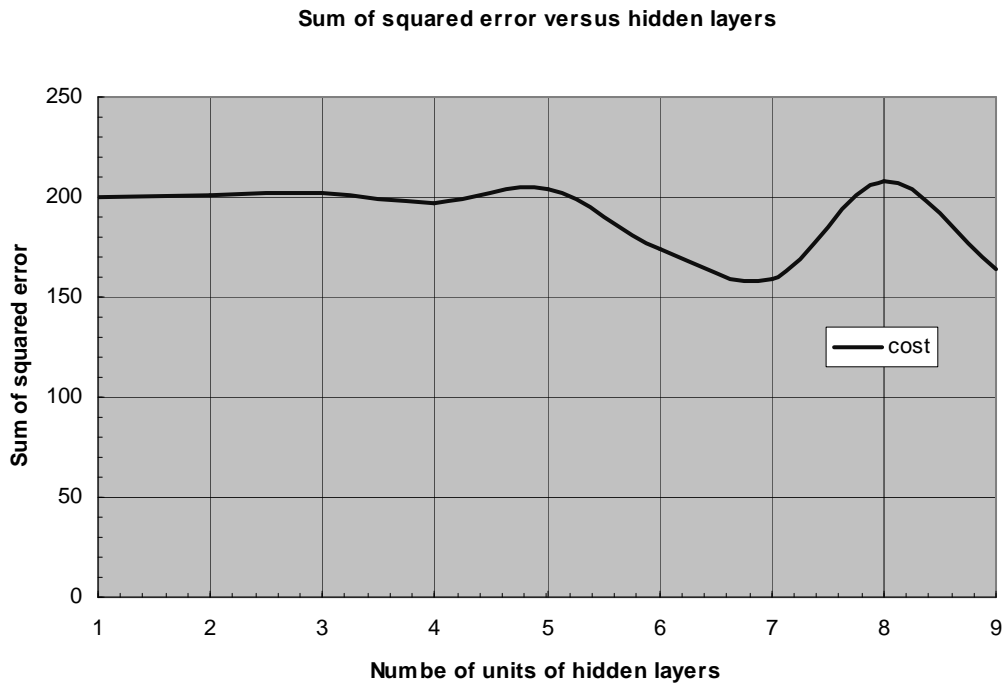


Figure 5-6      max learning iteration: 10000  
learning rate ETA: 0.009000  
sigmoid parameter ALPHA: 6.000000  
class number: 2  
network structure(including input): layerSize[3]={ 2 X 2 }



## 6. Dataset six

Table 6 max learning iteration: 10000  
learning rate ETA: 0.009000  
sigmoid parameter ALPHA: 6.000000  
class number: 4

Network structure	Sum of squared error	Network structure versus performance					
		train precision	95% confidence train interval	train time(ms)	test precision	95% confidence test interval	test time (ms)
{4 4 4}	297.378593	0.514	[0.455022, 0.516978]	188670	0.5102	[0.475944, 0.503656]	706
{4 4 4 4}	334.712839	0.37	[0.600075, 0.659925]	431191	0.37	[0.616617, 0.643383]	1256

Figure 6-1      learning rate ETA: 0.009000  
sigmoid parameter ALPHA: 6.000000  
class number: 4  
network structure(including input): layerSize[3]={4 4 4 }

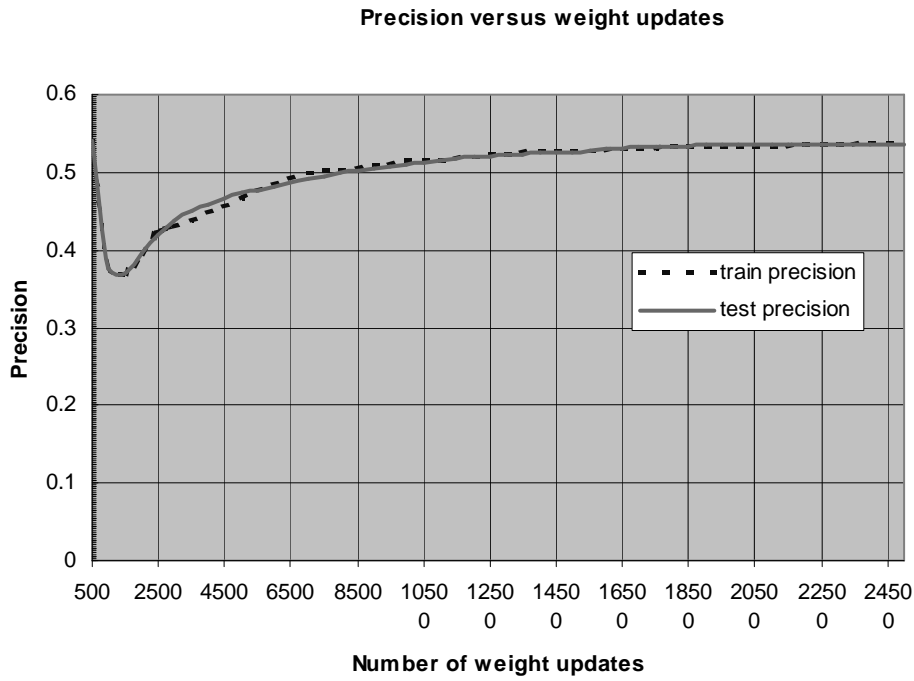


Figure 6-2      learning rate ETA: 0.009000  
sigmoid parameter ALPHA: 6.000000  
class number: 4  
network structure(including input): layerSize[3]={4 4 4 }

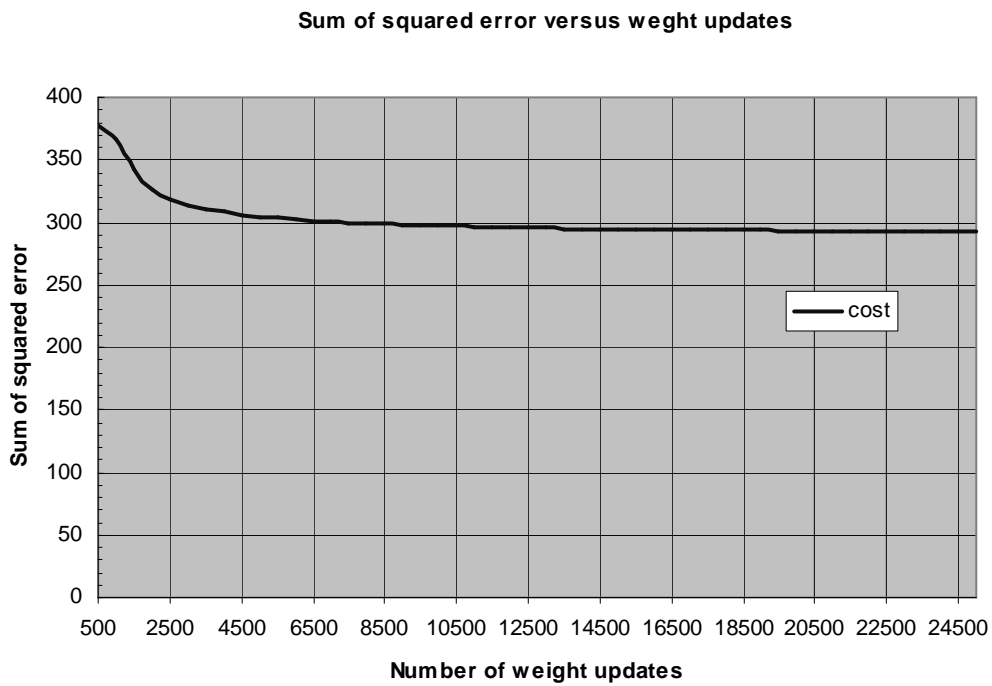


Figure 6-3 max learning iteration: 10000  
sigmoid parameter ALPHA: 6.000000  
class number: 4  
network structure(including input): layerSize[3]= {4 4 4 }

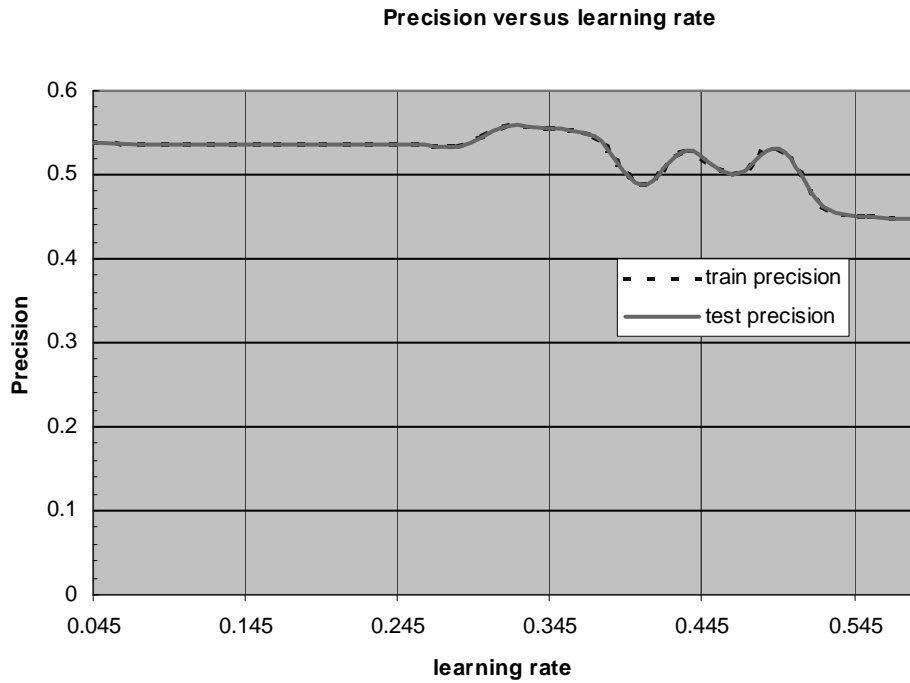


Figure 6-4 max learning iteration: 10000  
learning rate ETA: 0.009000  
class number: 4  
network structure(including input): layerSize[3]= {4 4 4 }

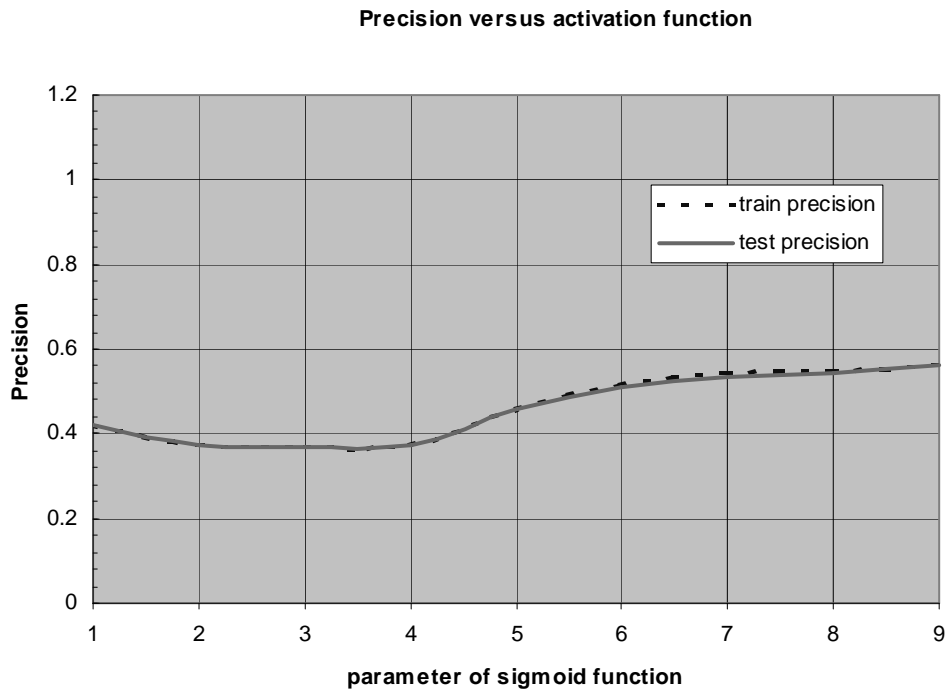




Figure 6-5      max learning iteration: 10000  
learning rate ETA: 0.009000  
sigmoid parameter ALPHA: 6.000000  
class number: 4  
network structure(including input): layerSize[3]={4 X 4 }

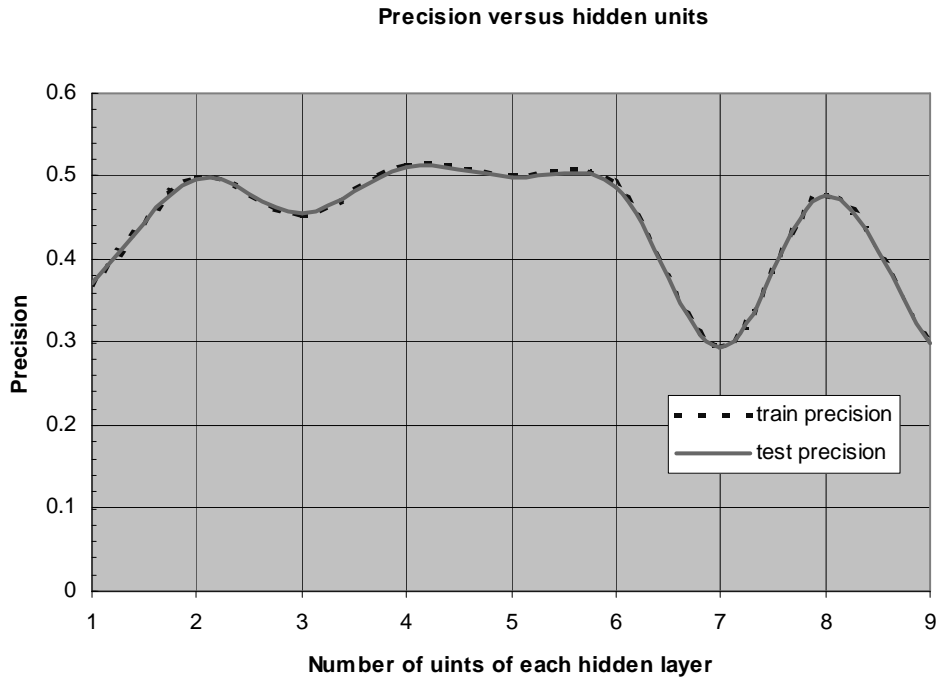


Figure 6-6      max learning iteration: 10000  
learning rate ETA: 0.009000  
sigmoid parameter ALPHA: 6.000000  
class number: 4  
network structure(including input): layerSize[3]={4 X 4 }

