ROB 537
Learning Based Control
HW#1 Neural Networks
TA: Shauharda Khadka

Eric Klinkhammer

October 2, 2017

1 Problem 1

1.1 Number of Hidden Units

Increasing the number of hidden units increases the classification accuracy on the training set, but tends to overfit (decrease the classification accuracy on the test set). As the learning rate and training time increase, more additional hidden units can be incorporated and increase accuracy.

For an example of overfitting, at a learning rate of 0.01 and 100000 training iterations with a sigmoid activation function, an increase in the hidden units increases the training set accuracy (up to 98 percent for 16), but increases the error for the test sets (from 85, 86, and 84 percent respectively to 79, 82, and 77.83 percent from 4 hidden units to 16).

Generally, 8 hidden units seemed to be a sweet spot across learning rates and training times.

Hidden Units	Iterations	Learning Rate	Activation	Train	Test 1	Test 2	Test 3
4	5000	0.005	Sigmoid	74.50%	73.33%	86.50%	58.50%
8	5000	0.005	Sigmoid	90.00%	86.33%	87.17%	82.33%
12	5000	0.005	Sigmoid	92.17%	83.17%	85.33%	79.00%
16	5000	0.005	Sigmoid	91.83%	83.33%	84.00%	80.33%
4	100000	0.010	Sigmoid	94.17%	85.00%	86.33%	84.17%
16	100000	0.010	Sigmoid	98.00%	79.33%	82.00%	77.83%

Table 1: Examples of diminishing returns of hidden units, and overfitting.

1.2 Training Time

Training time increases classification accuracy on the training set (this is the expected result given how the network is trained). However, increased accuracy on the training set suffers when generalizing, and - as with hidden units above - a longer training time overfits.

Hidden Units	Iterations	Learning Rate	Activation	Train	Test 1	Test 2	Test 3
8	1000	0.010	Sigmoid	88.67%	84.50%	83.17%	82.00%
8	5000	0.010	Sigmoid	90.00%	85.33%	85.83%	80.17%
8	10000	0.010	Sigmoid	90.50%	81.67%	84.00%	79.67%
8	100000	0.010	Sigmoid	96.17%	81.50%	84.17%	81.00%

Table 2: Example of effect of training time.

1.3 Learning Rate

The effects of changing the learning rate are more difficult to see. Both a low and a high learning rate appear to produce good results. The low learning rate is effective because the model does not

converge fast enough to overfit, while the high learning rate presumably takes gradient steps that are sufficiently large as to oscillate around the problem over fitting areas.

Hidden Units	Iterations	Learning Rate	Activation	Train	Test 1	Test 2	Test 3
8	10000	0.005	Sigmoid	93.67%	84.33%	85.17%	83.00%
8	10000	0.010	Sigmoid	90.50%	81.67%	84.00%	79.67%
8	10000	0.050	Sigmoid	91.67%	82.17%	82.83%	77.67%
8	10000	0.100	Sigmoid	88.00%	85.17%	81.17%	86.50%

Table 3: Learning rate effect example.

1.4 Other (Activation function)

I compared sigmoid and hyperbolic tangent as activation functions. The choice in activation function had an effect on the classification accuracy. It would appear that tanh generalizes slightly better in some cases, but that sigmoid was more likely to dominate tanh across the test cases.

Additionally, because the classification step used positive numbers only, if the initialization of the random weights was unlucky, tanh would fail to learn (would be a constant). ReLu (results not shown) had a similar problem.

Hidden Units	Iterations	Learning Rate	Activation	Train	Test 1	Test 2	Test 3
8	1000	0.005	Sigmoid	87.50%	84.33%	86.17%	80.00%
8	1000	0.005	Tanh	83.00%	79.17%	**91.33**%	67.83%
16	10000	0.005	Sigmoid	92.33%	83.83%	84.17%	82.00%
16	10000	0.005	Tanh	88.50%	85.83%	81.33%	**86.50**%
4	10000	0.010	Sigmoid	90.17%	**85.50**%	86.33%	83.33%
4	10000	0.010	Tanh	62.00%	61.50%	61.33%	59.67%
16	100000	0.010	Sigmoid	**98.00**%	79.33%	82.00%	77.83%
16	100000	0.010	Tanh	50.00%	50.00%	90.00%	10.00%

Table 4: Best performance on test sets, plus network trained identically except with a different activation function.

1.5 Difference Across Test Sets

The performance differed across the test sets. As a general rule, most networks performed substantially worse on the third test set. That test set is the most dissimilar to the training data, so the the networks generalized the worst.

2 Problem 2

2.1 Number of Hidden Units

Because there were fewer examples of positive test cases the in the second training data set, networks that could fit those cases generalized better, especially to the challenging third test set. For this training set, increasing the number of hidden nodes improved both classification accuracy on the training set and the ability to generalize to the test sets with dramatically different distributions.

Hidden Units	Iterations	Learning Rate	Activation	Train	Test 1	Test 2	Test 3
4	5000	0.005	Sigmoid	94.00%	69.67%	91.67%	37.33%
8	5000	0.005	Sigmoid	94.00%	66.00%	90.83%	38.67%
12	5000	0.005	Sigmoid	95.83%	67.67%	91.00%	37.67%
16	5000	0.005	Sigmoid	96.17%	69.17%	91.33%	38.33%
4	100000	0.005	Sigmoid	96.50%	67.50%	91.33%	40.00%
16	100000	0.005	Sigmoid	99.67%	67.83%	88.83%	41.67%

Table 5: Example of effect of hidden units with second data set.

2.2 Training Time

Increased training time is beneficial to a point. Again, because this training set has fewer positive examples, the model wants to move closer to overfitting those examples. However, this strategy will eventually backfire when the model actually does overfit (as can be seen in the row with 98.17 percent training set accuracy - it performs worse (although still better than the model with 92 percent training set accuracy)).

Hidden Units	Iterations	Learning Rate	Activation	Train	Test 1	Test 2	Test 3
8	1000	0.010	Sigmoid	92.00%	58.17%	90.83%	26.50%
8	5000	0.010	Sigmoid	95.83%	67.33%	91.50%	39.83%
8	10000	0.010	Sigmoid	95.50%	69.00%	90.83%	44.17%
8	100000	0.010	Sigmoid	98.17%	69.67%	90.33%	40.67%

Table 6: Example of effect of increased iterations with second data set.

2.3 Learning Rate

The effect of the learning rate, similar to the effect with the previous training set, is not a straightforward trend. There is no positive or negative correlation across any of the test sets, nor does there seem to be a readily discernible pattern across the test sets.

Hidden Units	Iterations	Learning Rate	Activation	Train	Test 1	Test 2	Test 3
8	1000	0.005	Sigmoid	91.33%	60.17%	91.83%	24.67%
8	1000	0.010	Sigmoid	92.00%	58.17%	90.83%	26.50%
8	1000	0.050	Sigmoid	92.83%	63.83%	92.00%	33.33%
8	1000	0.100	Sigmoid	91.33%	55.67%	91.17%	17.83%

Table 7: Example of effect of learning rate trained with second data set

2.4 Other (Activation function)

The choice of activation function mattered much more for this test set. This difference is because of the range of hyperbolic tangent (it can extend negatively, and, at least for this training set, negative classifications are essentially wrong).

Hidden Units	Iterations	Learning Rate	Activation	Train	Test 1	Test 2	Test 3
16	10000	0.005	Sigmoid	97.17%	71.17%	**92.33**%	39.83%
16	10000	0.005	Tanh	90.33%	50.67%	90.17%	11.00%
16	100000	0.005	Sigmoid	**99.67**%	67.83%	88.83%	41.67%
16	100000	0.005	Tanh	91.83%	54.50%	90.33%	17.67%
16	5000	0.050	Sigmoid	98.50%	**71.33**%	90.67%	37.17%
16	5000	0.050	Tanh	90.00%	50.00%	90.00%	10.00%
4	10000	0.050	Sigmoid	94.50%	70.00%	91.83%	**47.67**%
4	10000	0.050	Tanh	90.00%	50.00%	90.00%	10.00%

Table 8: Networks with best performance, plus networks with different activation function.

2.5 Difference Across Test Sets

There was a large performance across test sets. Test set 3 is populated with classes not suitably represented in the training data, so performance is worse on that test set, for all networks. Performance was excellent on the second test set, but the distribution matched the training set. Indeed, emitting a constant for the training set would score 90 percent on the second test set (this is one reason why the risk of overfitting was reduced for this dataset compared to the first one).

High performance on the first data set required the model to recognize the positive examples and not overfit. The larger (in terms of number of weights) networks tended to perform better on

this test set, as well as the other sets.

2.6 Source of Difference

The performance of the networks with the second training set was different from the first training set (generally worse) because the second training set does not contain a sufficient number of examples for each class that needs to be classified.

This difference in performance could be corrected by increasing the number and diversity of the passing training patterns.

3 Experiment Design

For each set of training data, I trained a network with every permutation of learning rates, training iterations, number of hidden units, and activation layer listed in the table, [0.1, 0.05, 0.01, 0.005], [1000, 5000, 10000, 100000], [4,8,12,16], and [Sigmoid, Tanh] respectively.

For each network, I tested it with training data and the three sets of test data. I report the average of three statistical runs.

3.1 Tables of full experiment results

Hidden Units	Iterations	Learning Rate	Activation	Train	Test 1	Test 2	Test 3
4	1000	0.005	Sigmoid	86.17%	85.00%	84.33%	82.67%
4	1000	0.005	Tanh	85.67%	83.83%	85.17%	79.67%
8	1000	0.005	Sigmoid	87.50%	84.33%	86.17%	80.00%
8	1000	0.005	Tanh	83.00%	79.17%	91.33%	67.83%
12	1000	0.005	Sigmoid	88.00%	84.50%	85.33%	80.67%
12	1000	0.005	Tanh	84.67%	83.00%	84.33%	78.67%
16	1000	0.005	Sigmoid	88.50%	84.50%	86.50%	81.17%
16	1000	0.005	Tanh	83.83%	79.33%	90.83%	70.33%
4	5000	0.005	Sigmoid	74.50%	73.33%	86.50%	58.50%
4	5000	0.005	Tanh	86.67%	81.67%	88.50%	78.00%
8	5000	0.005	Sigmoid	90.00%	86.33%	87.17%	82.33%
8	5000	0.005	Tanh	87.17%	82.83%	86.17%	79.67%
12	5000	0.005	Sigmoid	92.17%	83.17%	85.33%	79.00%
12	5000	0.005	Tanh	87.50%	83.83%	83.67%	81.33%
16	5000	0.005	Sigmoid	91.83%	83.33%	84.00%	80.33%
16	5000	0.005	Tanh	87.83%	85.00%	83.17%	82.50%
4	10000	0.005	Sigmoid	87.50%	84.00%	87.00%	81.50%
4	10000	0.005	Tanh	88.50%	84.00%	87.33%	79.50%
8	10000	0.005	Sigmoid	93.67%	84.33%	85.17%	83.00%
8	10000	0.005	Tanh	88.00%	81.67%	88.83%	76.67%
12	10000	0.005	Sigmoid	92.67%	83.17%	88.17%	77.33%
12	10000	0.005	Tanh	87.83%	83.67%	83.67%	81.50%
16	10000	0.005	Sigmoid	92.33%	83.83%	84.17%	82.00%
16	10000	0.005	Tanh	88.50%	85.83%	81.33%	86.50%
4	100000	0.005	Sigmoid	78.00%	73.17%	88.33%	57.67%
4	100000	0.005	Tanh	89.00%	84.33%	89.50%	80.17%
8	100000	0.005	Sigmoid	96.83%	83.00%	84.00%	81.67%
8	100000	0.005	Tanh	90.00%	83.33%	87.67%	74.33%
12	100000	0.005	Sigmoid	96.83%	80.50%	83.67%	79.33%
12	100000	0.005	Tanh	91.17%	84.33%	84.33%	81.67%
16	100000	0.005	Sigmoid	98.33%	79.50%	81.33%	80.00%
16	100000	0.005	Tanh	92.17%	84.67%	83.33%	81.83%
4	1000	0.010	Sigmoid	86.33%	84.67%	87.00%	80.50%
4	1000	0.010	Tanh	72.17%	73.33%	83.67%	60.33%
8	1000	0.010	Sigmoid	88.67%	84.50%	83.17%	82.00%
8	1000	0.010	Tanh	57.00%	56.83%	90.00%	27.33%

12	1000	0.010	Sigmoid	88.00%	85.33%	84.00%	82.67%
12	1000	0.010	Tanh	50.00%	50.00%	90.00%	10.00%
16	1000	0.010	Sigmoid	88.83%	84.83%	87.50%	78.33%
16	1000	0.010	Tanh	50.00%	50.00%	90.00%	10.00%
4	5000	0.010	Sigmoid	88.33%	84.17%	83.50%	83.67%
4	5000	0.010	Tanh	84.17%	84.00%	85.33%	79.17%
8	5000	0.010	Sigmoid	90.00%	85.33%	85.83%	80.17%
8	5000	0.010	Tanh	50.00%	50.17%	90.17%	10.00%
12	5000	0.010	Sigmoid	94.33%	81.50%	84.50%	78.67%
12	5000	0.010	Tanh	50.00%	50.00%	90.00%	10.00%
16	5000	0.010	Sigmoid	93.33%	83.83%	85.00%	79.67%
16	5000	0.010	Tanh	74.00%	72.17%	88.00%	55.83%
4	10000	0.010	Sigmoid	90.17%	85.50%	86.33%	83.33%
4	10000	0.010	Tanh	62.00%	61.50%	61.33%	59.67%
8	10000	0.010	Sigmoid	90.50%	81.67%	84.00%	79.67%
8	10000	0.010	Tanh	84.17%	83.00%	87.17%	77.33%
12	10000	0.010	Sigmoid	93.17%	82.67%	84.00%	79.50%
12	10000	0.010	Tanh	60.83%	59.83%	90.17%	29.50%
16	10000	0.010	Sigmoid	96.33%	78.83%	82.83%	77.83%
16	10000	0.010	Tanh	73.17%	73.00%	86.33%	56.17%
4	100000	0.010	Sigmoid	94.17%	85.00%	86.33%	84.17%
4	100000	0.010	Tanh	74.17%	72.50%	87.33%	55.33%
8	100000	0.010	Sigmoid	96.17%	81.50%	84.17%	81.00%
8	100000	0.010	Tanh	85.17%	82.83%	86.67%	77.00%
12	100000	0.010	Sigmoid	97.83%	82.33%	83.83%	76.00%
12	100000	0.010	Tanh	50.50%	50.50%	35.67%	65.17%
16	100000	0.010	Sigmoid	98.00%	79.33%	82.00%	77.83%
16	100000	0.010	Tanh	50.00%	50.00%	90.00%	10.00%
4	1000	0.050	Sigmoid	87.00%	83.83%	82.00%	84.17%
4	1000	0.050	Tanh	61.17%	63.67%	87.33%	38.17%
8	1000	0.050	Sigmoid	87.33%	83.50%	82.67%	82.17%
8	1000	0.050	Tanh	61.50%	61.67%	86.33%	35.33%
12	1000	0.050	Sigmoid	90.67%	82.17%	81.67%	81.67%
12	1000	0.050	Tanh	70.67%	68.50%	65.00%	74.50%
16	1000	0.050	Sigmoid	89.17%	81.83%	88.50%	75.33%
16	1000	0.050	Tanh	85.17%	85.00%	84.67%	82.00%
4	5000	0.050	Sigmoid	89.00%	83.17%	81.50%	82.33%
4	5000	0.050	Tanh	71.33%	71.33%	79.17%	61.50%
8	5000	0.050	Sigmoid	89.83%	82.00%	81.83%	82.00%
8	5000	0.050	Tanh	72.50%	72.50%	87.83%	55.17%
12	5000	0.050	Sigmoid	92.17%	81.83%	86.83%	73.33%
12	5000	0.050	Tanh	50.00%	50.00%	90.00%	10.00%
16	5000	0.050	Sigmoid	94.00%	80.50%	84.83%	74.83%
16	5000	0.050	Tanh	61.00%	60.67%	90.50%	30.50%
4	10000	0.050	Sigmoid	89.00%	84.67%	83.50%	83.00%
4	10000	0.050	Tanh	62.00%	61.83%	87.67%	33.83%
8	10000	0.050	Sigmoid	91.67%	82.17%	82.83%	77.67%
8	10000	0.050	Tanh	50.00%	50.00%	90.00%	10.00%
12	10000	0.050	Sigmoid	92.83%	81.83%	80.83%	82.83%
12	10000	0.050	Tanh	61.50%	62.00%	87.67%	34.17%
16	10000	0.050	Sigmoid	95.50%	80.67%	85.67%	77.67%
16	10000	0.050	Tanh	50.00%	50.00%	90.00%	10.00%
4	100000	0.050	Sigmoid	91.33%	80.00%	78.67%	80.83%
4	100000	0.050	Tanh	61.17%	61.67%	84.83%	36.83%
8	100000	0.050	Sigmoid	93.83%	81.17%	81.33%	78.83%
8	100000	0.050	Tanh	61.83%	61.00%	89.00%	32.50%
12	100000	0.050	Sigmoid	97.00%	81.67%	80.33%	80.00%
12	100000	0.050	Tanh	50.00%	50.00%	90.00%	10.00%

16	100000	0.050	Sigmoid	96.83%	79.83%	81.67%	75.00%
16	100000	0.050	Tanh	81.33%	77.50%	87.17%	68.17%
4	1000	0.100	Sigmoid	50.00%	50.00%	90.00%	10.00%
4	1000	0.100	Tanh	55.83%	55.33%	90.83%	19.50%
8	1000	0.100	Sigmoid	61.33%	60.50%	90.00%	30.50%
8	1000	0.100	Tanh	62.33%	62.33%	87.33%	35.33%
12	1000	0.100	Sigmoid	88.83%	83.33%	84.17%	82.00%
12	1000	0.100	Tanh	66.50%	64.33%	91.50%	37.17%
16	1000	0.100	Sigmoid	74.67%	71.83%	84.83%	57.67%
16	1000	0.100	Tanh	55.50%	56.33%	90.33%	19.67%
4	5000	0.100	Sigmoid	87.50%	85.50%	83.50%	86.33%
4	5000	0.100	Tanh	50.00%	50.00%	90.00%	10.00%
8	5000	0.100	Sigmoid	89.00%	82.83%	84.17%	80.00%
8	5000	0.100	Tanh	60.00%	59.33%	89.50%	29.67%
12	5000	0.100	Sigmoid	87.50%	83.00%	83.00%	83.17%
12	5000	0.100	Tanh	50.00%	50.00%	90.00%	10.00%
16	5000	0.100	Sigmoid	90.33%	82.50%	80.83%	81.67%
16	5000	0.100	Tanh	61.33%	60.83%	89.50%	32.00%
4	10000	0.100	Sigmoid	87.50%	84.67%	83.00%	84.33%
4	10000	0.100	Tanh	50.00%	50.00%	90.00%	10.00%
8	10000	0.100	Sigmoid	88.00%	85.17%	81.17%	86.50%
8	10000	0.100	Tanh	61.50%	61.00%	89.17%	31.83%
12	10000	0.100	Sigmoid	89.00%	83.67%	81.83%	83.50%
12	10000	0.100	Tanh	50.00%	50.00%	90.00%	10.00%
16	10000	0.100	Sigmoid	90.67%	81.67%	82.00%	81.33%
16	10000	0.100	Tanh	50.00%	50.00%	90.00%	10.00%
4	100000	0.100	Sigmoid	88.83%	85.17%	83.33%	85.67%
4	100000	0.100	Tanh	62.17%	61.83%	88.17%	34.67%
8	100000	0.100	Sigmoid	89.17%	83.33%	82.00%	84.83%
8	100000	0.100	Tanh	70.17%	68.50%	88.00%	47.67%
12	100000	0.100	Sigmoid	91.67%	82.67%	84.17%	77.50%
12	100000	0.100	Tanh	73.00%	72.00%	88.33%	54.17%
16	100000	0.100	Sigmoid	95.00%	78.00%	84.00%	75.00%
16	100000	0.100	Tanh	63.00%	61.83%	66.00%	59.50%

Table 9: Networks trained with training set 1.

Hidden Units	Iterations	Learning Rate	Activation	Train	Test 1	Test 2	Test 3
4	1000	0.005	Sigmoid	91.67%	58.83%	91.17%	24.33%
4	1000	0.005	Tanh	90.67%	52.83%	90.17%	13.50%
8	1000	0.005	Sigmoid	91.33%	60.17%	91.83%	24.67%
8	1000	0.005	Tanh	90.00%	50.00%	90.00%	10.00%
12	1000	0.005	Sigmoid	93.17%	63.00%	91.33%	32.33%
12	1000	0.005	Tanh	90.50%	51.33%	90.17%	12.83%
16	1000	0.005	Sigmoid	92.83%	63.17%	91.83%	33.67%
16	1000	0.005	Tanh	90.00%	50.00%	90.00%	10.00%
4	5000	0.005	Sigmoid	94.00%	69.67%	91.67%	37.33%
4	5000	0.005	Tanh	90.17%	51.17%	90.17%	11.33%
8	5000	0.005	Sigmoid	94.00%	66.00%	90.83%	38.67%
8	5000	0.005	Tanh	91.50%	54.00%	89.83%	19.67%
12	5000	0.005	Sigmoid	95.83%	67.67%	91.00%	37.67%
12	5000	0.005	Tanh	90.83%	51.83%	90.00%	12.17%
16	5000	0.005	Sigmoid	96.17%	69.17%	91.33%	38.33%
16	5000	0.005	Tanh	91.00%	52.17%	89.50%	17.67%
4	10000	0.005	Sigmoid	94.33%	67.17%	91.50%	39.00%
4	10000	0.005	Tanh	91.67%	56.17%	91.50%	22.17%
8	10000	0.005	Sigmoid	96.83%	68.50%	90.67%	37.83%
8	10000	0.005	Tanh	90.50%	51.83%	90.67%	14.00%

12	10000	0.005	Sigmoid	97.33%	69.83%	91.17%	38.67%
12	10000	0.005	Tanh	90.17%	50.83%	90.00%	10.50%
16	10000	0.005	Sigmoid	97.17%	71.17%	92.33%	39.83%
16	10000	0.005	Tanh	90.33%	50.67%	90.17%	11.00%
4	100000	0.005	Sigmoid	96.50%	67.50%	91.33%	40.00%
4	100000	0.005	Tanh	92.17%	61.00%	91.67%	29.83%
8	100000	0.005	Sigmoid	96.67%	69.50%	90.00%	45.67%
8	100000	0.005	Tanh	90.00%	50.33%	90.00%	10.33%
12	100000	0.005	Sigmoid	99.17%	70.83%	90.50%	37.83%
12	100000	0.005	Tanh	90.17%	50.67%	90.50%	11.50%
16	100000	0.005	Sigmoid	99.67%	67.83%	88.83%	41.67%
16	100000	0.005	Tanh	91.83%	54.50%	90.33%	17.67%
4	1000	0.010	Sigmoid	91.83%	62.33%	91.50%	32.50%
4	1000	0.010	Tanh	90.00%	50.00%	90.00%	10.00%
8	1000	0.010	Sigmoid	92.00%	58.17%	90.83%	26.50%
8	1000	0.010	Tanh	90.00%	50.00%	90.00%	10.00%
12	1000	0.010	Sigmoid	93.17%	63.00%	92.17%	35.33%
12	1000	0.010	Tanh	90.00%	50.00%	90.00%	10.00%
16	1000	0.010	Sigmoid	93.33%	65.17%	91.67%	37.33%
16	1000	0.010	Tanh	90.00%	50.00%	90.00%	10.00%
4	5000	0.010	Sigmoid	93.67%	64.00%	91.50%	36.83%
4	5000	0.010	Tanh	90.00%	50.00%	90.00%	10.00%
8	5000	0.010	Sigmoid	95.83%	67.33%	91.50%	39.83%
8	5000	0.010	Tanh	91.00%	53.83%	90.67%	17.50%
12	5000	0.010	Sigmoid	96.17%	68.83%	91.33%	35.83%
12	5000	0.010	Tanh	90.00%	50.00%	90.00%	10.00%
16	5000	0.010	Sigmoid	97.17%	69.33%	91.50%	38.50%
16	5000	0.010	Tanh	90.00%	50.17%	90.00%	10.17%
4	10000	0.010	Sigmoid	95.33%	64.83%	90.67%	37.33%
4	10000	0.010	Tanh	90.00%	50.00%	90.00%	10.00%
8	10000	0.010	Sigmoid	95.50%	69.00%	90.83%	44.17%
8	10000	0.010	Tanh	90.00%	50.00%	90.00%	10.00%
12	10000	0.010	Sigmoid	98.33%	69.50%	91.17%	38.83%
12	10000	0.010	Tanh	90.00%	50.00%	90.00%	10.00%
16	10000	0.010	Sigmoid	98.17%	70.33%	91.00%	38.50%
16	10000	0.010	Tanh	90.00%	50.00%	90.00%	10.00%
4	100000	0.010	Sigmoid	95.67%	66.33%	90.67%	36.83%
4	100000	0.010	Tanh	90.00%	50.00%	90.00%	10.00%
8	100000	0.010	Sigmoid	98.17%	69.67%	90.33%	40.67%
8	100000	0.010	Tanh	90.00%	50.00%	90.00%	10.00%
12	100000	0.010	Sigmoid	99.33%	67.17%	89.33%	39.50%
12	100000	0.010	Tanh	90.00%	50.00%	90.00%	10.00%
16	100000	0.010	Sigmoid	99.50%	69.83%	89.83%	40.67%
16	100000	0.010	Tanh	87.00%	59.17%	87.67%	33.67%
4	1000	0.050	Sigmoid	92.00%	58.17%	90.50%	24.83%
4	1000	0.050	Tanh	90.00%	50.00%	90.00%	10.00%
8	1000	0.050	Sigmoid	92.83%	63.83%	92.00%	33.33%
8	1000	0.050	Tanh	90.00%	50.00%	90.00%	10.00%
12	1000	0.050	Sigmoid	93.83%	65.17%	92.00%	32.67%
12	1000	0.050	Tanh	90.00%	50.00%	90.00%	10.00%
16	1000	0.050	Sigmoid	91.33%	54.50%	90.17%	16.67%
16	1000	0.050	Tanh	90.00%	50.00%	90.00%	10.00%
4	5000	0.050	Sigmoid	92.67%	66.50%	91.50%	41.17%
4	5000	0.050	Tanh	90.00%	50.00%	90.00%	10.00%
8	5000	0.050	Sigmoid	97.33%	68.50%	90.33%	38.00%
8	5000	0.050	Tanh	86.00%	49.00%	87.00%	11.00%
12	5000	0.050	Sigmoid	97.17%	68.00%	91.33%	40.83%
12	5000	0.050	Tanh	90.00%	50.00%	90.00%	10.00%

16	5000	0.050	Sigmoid	98.50%	71.33%	90.67%	37.17%
16	5000	0.050	Tanh	90.00%	50.00%	90.00%	10.00%
4	10000	0.050	Sigmoid	94.50%	70.00%	91.83%	47.67%
4	10000	0.050	Tanh	90.00%	50.00%	90.00%	10.00%
8	10000	0.050	Sigmoid	97.67%	69.17%	91.50%	39.33%
8	10000	0.050	Tanh	90.00%	51.83%	90.33%	12.67%
12	10000	0.050	Sigmoid	97.67%	67.33%	91.67%	39.00%
12	10000	0.050	Tanh	90.00%	50.00%	90.00%	10.00%
16	10000	0.050	Sigmoid	96.00%	63.00%	90.00%	30.67%
16	10000	0.050	Tanh	88.67%	53.83%	90.17%	17.50%
4	100000	0.050	Sigmoid	95.83%	69.83%	90.83%	42.00%
4	100000	0.050	Tanh	80.17%	58.33%	79.17%	37.33%
8	100000	0.050	Sigmoid	99.00%	66.67%	89.17%	41.17%
8	100000	0.050	Tanh	90.00%	50.17%	90.00%	10.00%
12	100000	0.050	Sigmoid	99.67%	67.17%	88.00%	37.67%
12	100000	0.050	Tanh	90.00%	50.00%	90.00%	10.00%
16	100000	0.050	Sigmoid	99.67%	68.17%	88.83%	42.50%
16	100000	0.050	Tanh	90.00%	50.00%	90.00%	10.00%
4	1000	0.100	Sigmoid	90.00%	50.00%	90.00%	10.00%
4	1000	0.100	Tanh	89.50%	57.67%	90.83%	24.83%
8	1000	0.100	Sigmoid	91.33%	55.67%	91.17%	17.83%
8	1000	0.100	Tanh	90.00%	50.00%	90.00%	10.00%
12	1000	0.100	Sigmoid	91.00%	62.33%	91.17%	34.83%
12	1000	0.100	Tanh	90.00%	50.00%	90.00%	10.00%
16	1000	0.100	Sigmoid	90.00%	50.00%	90.00%	10.00%
16	1000	0.100	Tanh	90.00%	50.00%	90.00%	10.00%
4	5000	0.100	Sigmoid	92.83%	62.50%	90.83%	33.17%
4	5000	0.100	Tanh	90.00%	50.00%	90.00%	10.00%
8	5000	0.100	Sigmoid	95.00%	67.00%	91.33%	43.00%
8	5000	0.100	Tanh	90.00%	50.00%	90.00%	10.00%
12	5000	0.100	Sigmoid	94.00%	60.50%	90.50%	26.17%
12	5000	0.100	Tanh	90.00%	50.00%	90.00%	10.00%
16	5000	0.100	Sigmoid	92.83%	57.17%	90.17%	19.83%
16	5000	0.100	Tanh	90.00%	50.00%	90.00%	10.00%
4	10000	0.100	Sigmoid	92.17%	57.33%	90.83%	21.83%
4	10000	0.100	Tanh	90.00%	50.00%	90.00%	10.00%
8	10000	0.100	Sigmoid	94.50%	62.50%	91.50%	31.67%
8	10000	0.100	Tanh	90.00%	50.83%	90.33%	11.17%
12	10000	0.100	Sigmoid	95.00%	63.33%	90.67%	29.17%
12	10000	0.100	Tanh	90.00%	50.00%	90.00%	10.00%
16	10000	0.100	Sigmoid	94.50%	60.83%	90.67%	27.33%
16	10000	0.100	Tanh	90.00%	50.00%	90.00%	10.00%
4	100000	0.100	Sigmoid	91.50%	54.83%	90.17%	18.33%
4	100000	0.100	Tanh	90.00%	50.00%	90.00%	10.00%
8	100000	0.100	Sigmoid	96.83%	67.67%	90.67%	45.00%
8	100000	0.100	Tanh	86.00%	62.00%	86.17%	36.17%
12	100000	0.100	Sigmoid	97.83%	66.83%	91.50%	40.50%
12	100000	0.100	Tanh	85.00%	51.33%	82.50%	18.00%
16	100000	0.100	Sigmoid	96.17%	62.33%	90.67%	30.50%
16	100000	0.100	Tanh	90.00%	50.00%	90.00%	10.00%

Table 10: Networks trained with training set 2.