A Short Paper on Feedforward Neural Networking Showing Best Approach for Training a Network that would Yield Accurate Results in the Minimum of Training Steps

University of the People

Department of Computer Science

Bachelor of Science Degree Program

(Student name: Shehu Saliu Ibrahim)

Data Mining and Machine Learning Unit 7 Programming Assignment

Introduction

The key aim and objective of this paper is to show the feedforward neural network (FNN) alternatives that were tested to determine the best approach for training a network that would yield accurate results in the minimum of training steps.

Network design used and number of iterations

Four different network configurations categories were tested for this purpose, namely (see appendix):

N: Network with no hidden layer or nodes (iteration range: infinity)

N1: Network with one hidden layer consisting of five nodes (iteration range: 3 to 6 million)

N2: Network with one hidden layer consisting of ten nodes (iteration range: 1-2 million)

N3: Network with two hidden layers consisting of ten nodes (iteration range: 200 to 400 thousand)

Results

Networks	Convergence	Average per pattern error threshold	Average training steps (millions)
N	No	>0.05	infinite
N1	yes	0.004188	4.5
N2	yes	0.009905	1.5
N3	yes	0.004937	0.3

Tested alternatives

An average of about three instances where tested under each network categories in order to determine the average per pattern error and average training steps values. Under these instances, I typically kept the weight range constant at [-1 -1] and learning rate at 0.3 while modifying both the momentum and learning steps in accordance with my observation of the error progress. Training and testing were carried out intermittently at some points where the average per pattern error threshold is approaching 0.05. The training phases basically show the network sample pattern input and the correct classifications while the training phases were meant to verify that all classifications are correctly represented as shown by the pattern results in the console (see appendix).

Observations

I observed that the entire purpose of training and testing is to assign the right weights to the links between layers. By entering new patterns, these weights can correctly determine the output classification correctly when the saved weights are loaded without need for any training. In this case, the weights that minimize the errors are then considered to be a solution (not the only solution) to the learning problem.

Conclusion

In conclusion, the best approach for training a network such as been tested, appears to be to start with a relatively low learning rate value with the possibility of decreasing the value as the training proceeds and nearing the approved or acceptable average per pattern error threshold value of less than 0.05 to avoid "overfitting" the pattern at the end of the training. Initially starting with a low learning rate help avoid oscillation in pattern signal and improve convergence stability, as well as lessen the possibility of destroying correct classification of the outputs.

Aiming to create a neural network with the "right" number of hidden nodes that will lead to a good solution to a problem would intuitively imply a higher number of epochs or iterations. In other words, the more the epochs or iterations the better the actual output. The learning steps in relation to this would either help reach accuracy quickly or slowly with possible consequence of overfitting or overtraining. In this respect the learning steps was used to help stem or prevent overfitting or overtraining. In these cases, the network with the highest number of hidden nodes reached accuracy with minimum training steps of 0.3 million steps on average.

The initial convergence and the final optimization of the results had been aided by chosen the right momentum. Momentum in this regard had helped with weight updates and acts as an optimization tool. Furthermore, it prevented the learning process from settling in a local minimum by “over stepping” the small “hill” (i.e. local minimum).

Reference

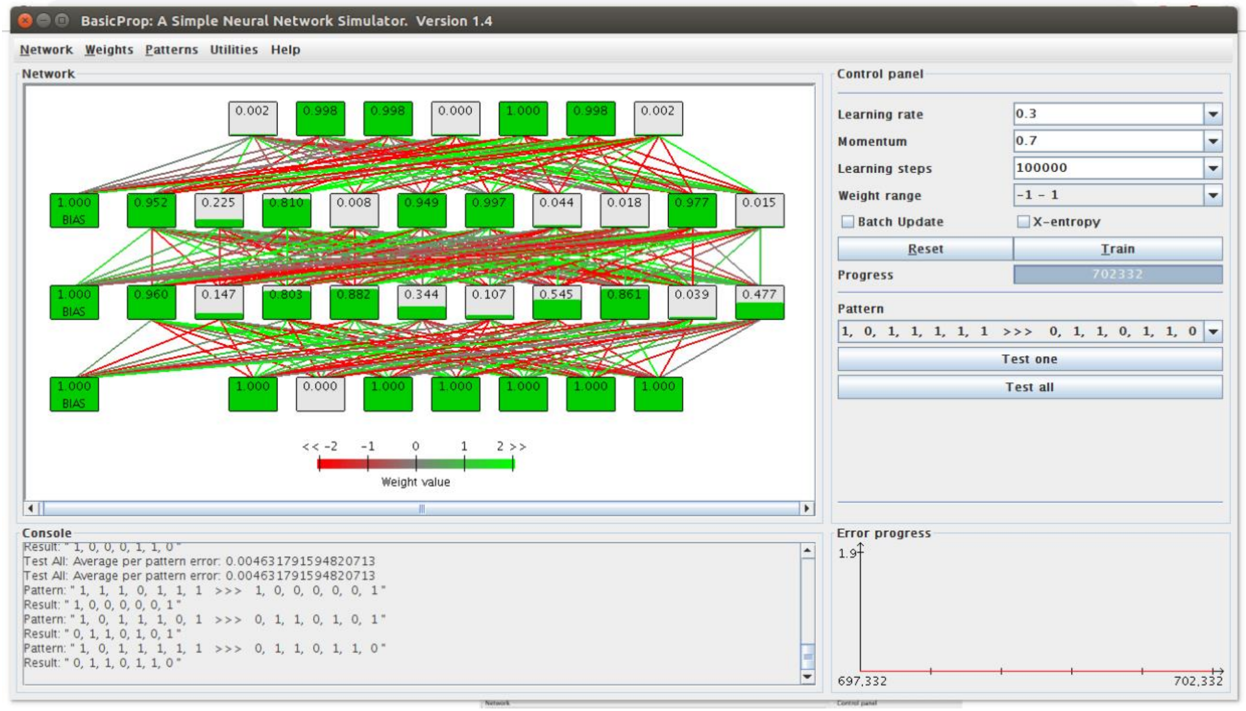
Ahire, J. B. (2018). Perceptron and Backpropagation. Available from: https://medium.com/@jayeshbahire/perceptron-and-backpropagation-970d752f4e44

Kan, A. (2017). Lecture 7: Multiple Perceptron. Backpropagation. COMP90051 Statistical Machine Learning. The University of Melbourne. Available from: https://trevorcohn.github.io/comp90051-2017/slides/07_backpropagation.pdf

https://www.cse.unsw.edu.au/~cs9417ml/MLP2/BackPropagation.html

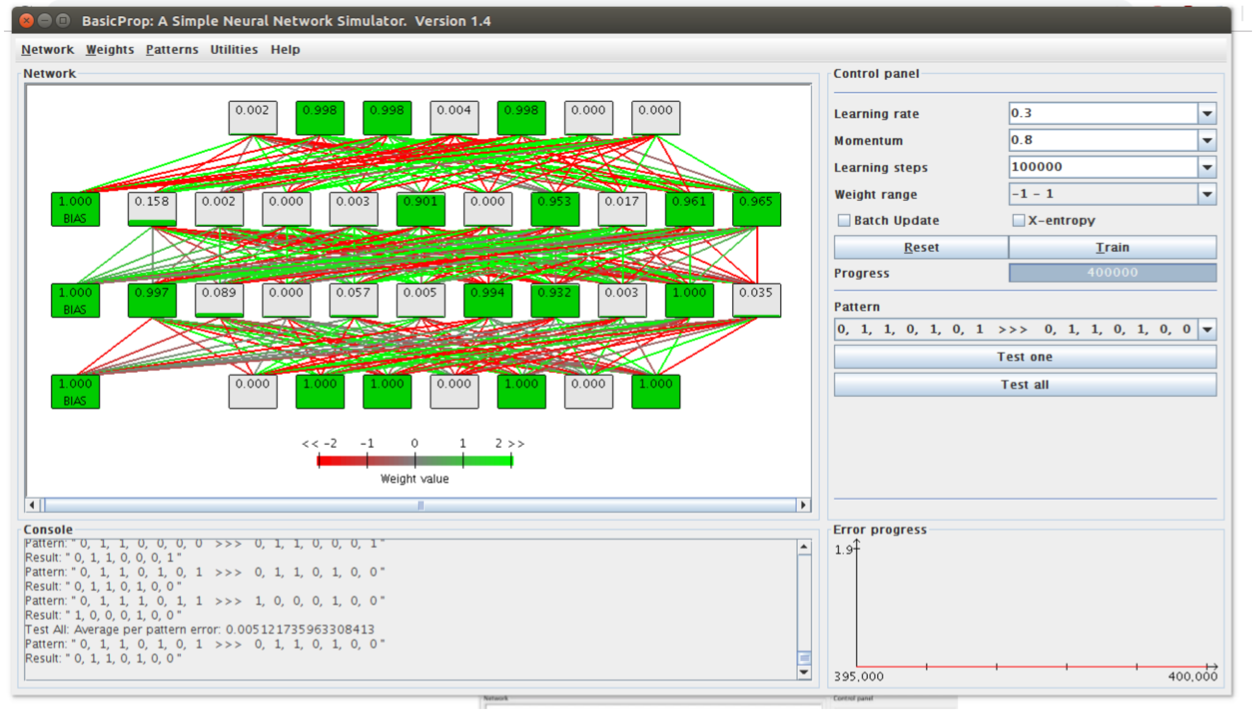
Appendix

N3:


BasicProp: A Simple Neural Network Simulator. Version 1.4
Network Weights Patterns Utilities Help
Network
<< -2 -1 0 1 2 >>
Weight value
Reset
Train
Test one
Test all
Error progress

Learning rate 0.3
Momentum 0.7
Learning steps 100000
Weight range -1 - 1
Batch Update
X-entropy
Progress 702332
Pattern 1, 0, 1, 1, 1, 1, 1 >>> 0, 1, 1, 0, 1, 1, 0

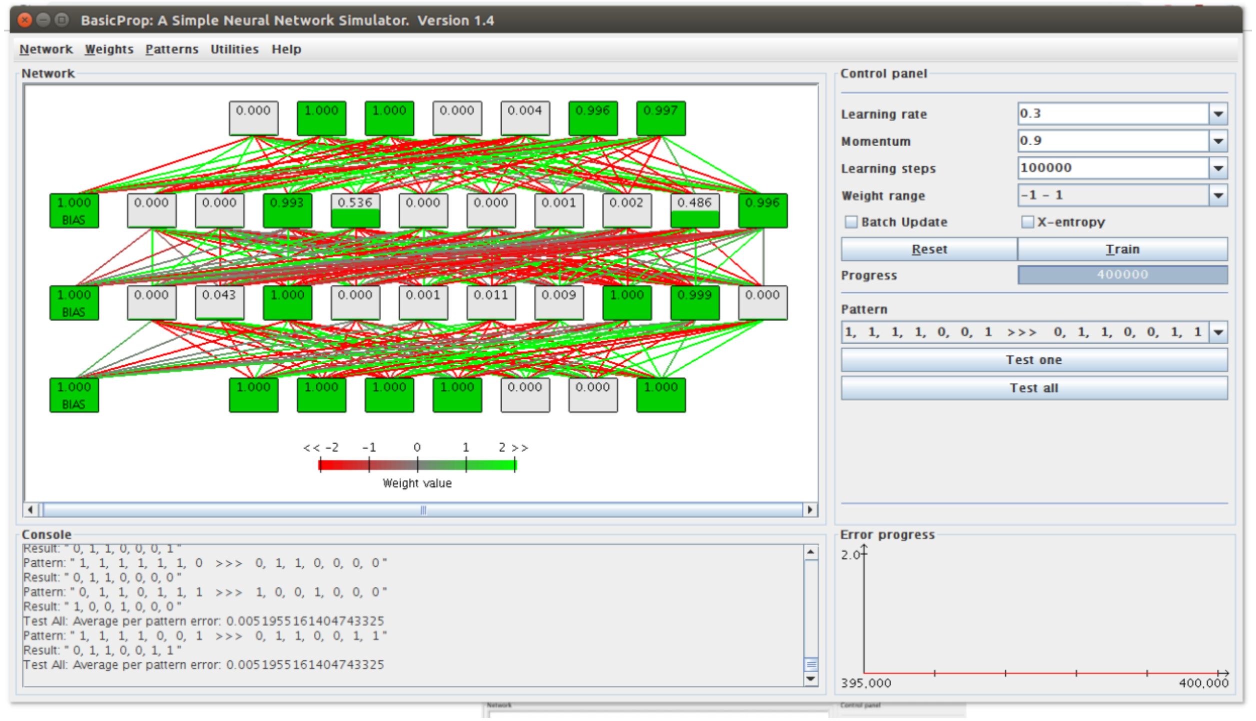
```
Console
Result: " 1, 0, 0, 0, 1, 1, 0 "
Test All: Average per pattern error: 0.004631791594820713
Test All: Average per pattern error: 0.004631791594820713
Pattern: " 1, 1, 1, 0, 1, 1, 1 >>> 1, 0, 0, 0, 0, 0, 1 "
Result: " 1, 0, 0, 0, 0, 0, 1 "
Pattern: " 1, 0, 1, 1, 1, 0, 1 >>> 0, 1, 1, 0, 1, 0, 1 "
Result: " 0, 1, 1, 0, 1, 0, 1 "
Pattern: " 1, 0, 1, 1, 1, 1, 1 >>> 0, 1, 1, 0, 1, 1, 0 "
Result: " 0, 1, 1, 0, 1, 1, 0 "
```


BasicProp: A Simple Neural Network Simulator. Version 1.4
Network Weights Patterns Utilities Help
Network
<< -2 -1 0 1 2 >>
Weight value
Reset
Train
Test one
Test all
Error progress

Learning rate 0.3
Momentum 0.8
Learning steps 100000
Weight range -1 - 1
Batch Update
X-entropy
Progress 400000
Pattern 0, 1, 1, 0, 1, 0, 1 >>> 0, 1, 1, 0, 1, 0, 0

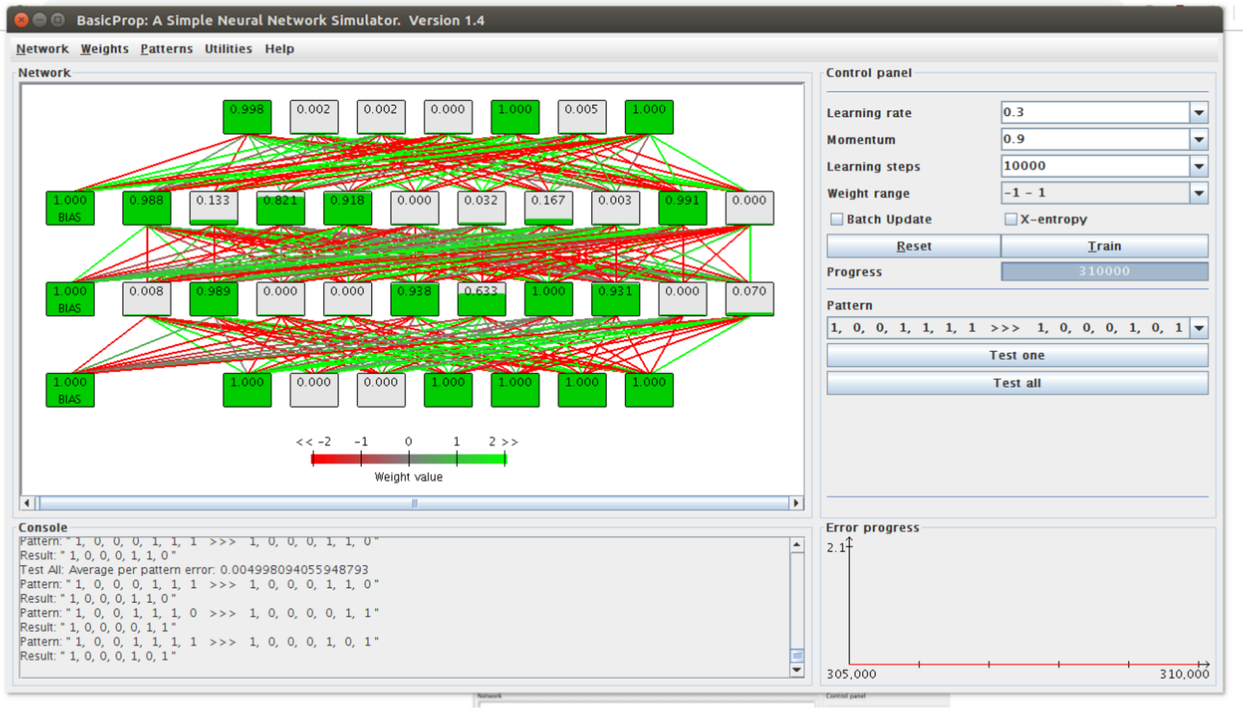
```
Console
Pattern: " 0, 1, 1, 0, 0, 0, 0 >>> 0, 1, 1, 0, 0, 0, 1 "
Result: " 0, 1, 1, 0, 0, 0, 1 "
Pattern: " 0, 1, 1, 0, 1, 0, 1 >>> 0, 1, 1, 0, 1, 0, 0 "
Result: " 0, 1, 1, 0, 1, 0, 0 "
Pattern: " 0, 1, 1, 1, 0, 1, 1 >>> 1, 0, 0, 0, 1, 0, 0 "
Result: " 1, 0, 0, 0, 1, 0, 0 "
Test All: Average per pattern error: 0.005121735963308413
Pattern: " 0, 1, 1, 0, 1, 0, 1 >>> 0, 1, 1, 0, 1, 0, 0 "
Result: " 0, 1, 1, 0, 1, 0, 0 "
```

Data Mining and Machine Learning Unit 7 Programming Assignment


BasicProp: A Simple Neural Network Simulator. Version 1.4
Network Weights Patterns Utilities Help
Network
0.000 1.000 1.000 0.000 0.004 0.996 0.997
1.000 BIAS 0.000 0.000 0.993 0.536 0.000 0.000 0.001 0.002 0.486 0.996
1.000 BIAS 0.000 0.043 1.000 0.000 0.001 0.011 0.009 1.000 0.999 0.000
1.000 BIAS 1.000 1.000 1.000 1.000 0.000 0.000 1.000
<< -2 -1 0 1 2 >>
Weight value
Control panel
Reset
Train
Test one
Test all
Error progress
2.0
395,000
400,000

Learning rate 0.3
Momentum 0.9
Learning steps 100000
Weight range -1 - 1
Batch Update
X-entropy
Progress 400000
Pattern 1, 1, 1, 1, 0, 0, 1 >>> 0, 1, 1, 0, 0, 1, 1

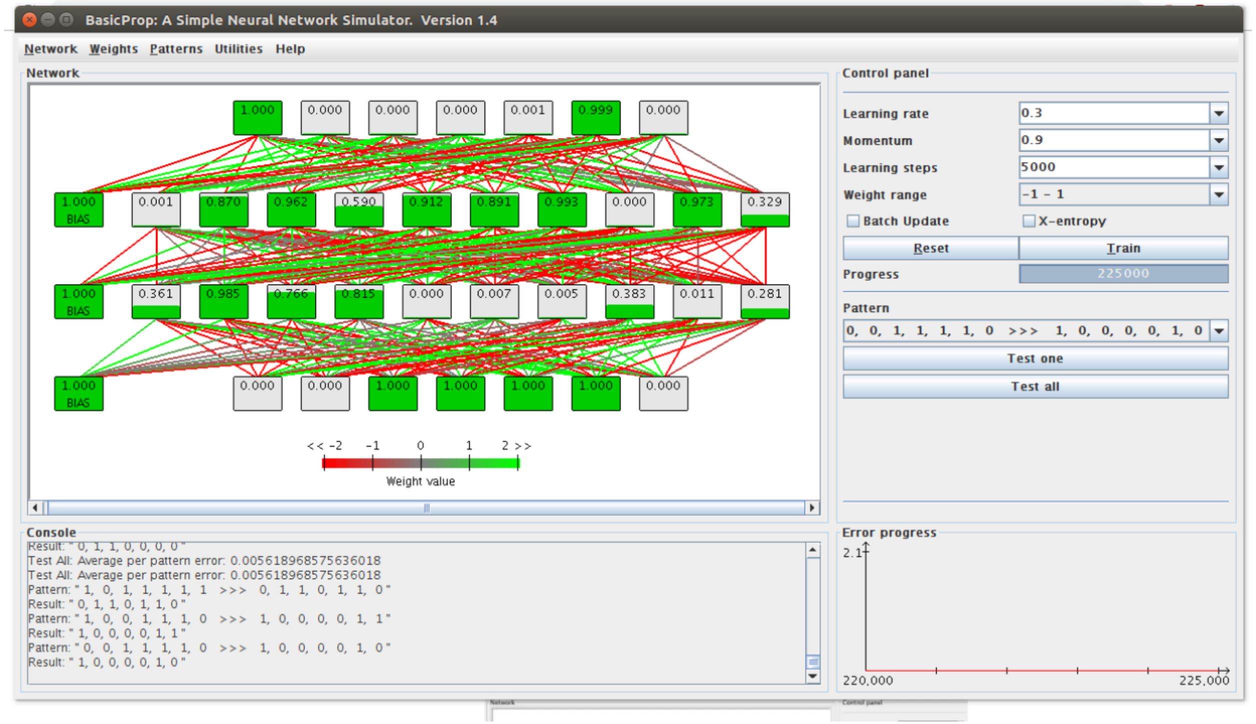
Console
```
Result: " 0, 1, 1, 0, 0, 0, 1 "
Pattern: " 1,  1,  1,  1,  1,  1,  0   >>>  0,  1,  1,  0,  0,  0,  0 "
Result: " 0, 1, 1, 0, 0, 0, 0 "
Pattern: " 0,  1,  1,  0,  1,  1,  1   >>>  1,  0,  0,  1,  0,  0,  0 "
Result: " 1, 0, 0, 1, 0, 0, 0 "
Test All: Average per pattern error: 0.005195516140474332 5
Pattern: " 1,  1,  1,  1,  0,  0,  1   >>>  0,  1,  1,  0,  0,  1,  1 "
Result: " 0, 1, 1, 0, 0, 1, 1 "
Test All: Average per pattern error: 0.005195516140474332 5
```


BasicProp: A Simple Neural Network Simulator. Version 1.4
Network Weights Patterns Utilities Help
Network
0.998 0.002 0.002 0.000 1.000 0.005 1.000
1.000 BIAS 0.988 0.133 0.821 0.918 0.000 0.032 0.167 0.003 0.991 0.000
1.000 BIAS 0.008 0.989 0.000 0.000 0.938 0.633 1.000 0.931 0.000 0.070
1.000 BIAS 1.000 0.000 0.000 1.000 1.000 1.000 1.000
<< -2 -1 0 1 2 >>
Weight value
Control panel
Reset
Train
Test one
Test all
Error progress
2.1
305,000
310,000

Learning rate 0.3
Momentum 0.9
Learning steps 10000
Weight range -1 - 1
Batch Update
X-entropy
Progress 310000
Pattern 1, 0, 0, 1, 1, 1, 1 >>> 1, 0, 0, 0, 1, 0, 1

Console
```
Pattern: " 1,  0,  0,  0,  1,  1,  1   >>>  1,  0,  0,  0,  1,  1,  0 "
Result: " 1, 0, 0, 0, 1, 1, 0 "
Test All: Average per pattern error: 0.004998094055948793
Pattern: " 1,  0,  0,  0,  1,  1,  1   >>>  1,  0,  0,  0,  1,  1,  0 "
Result: " 1, 0, 0, 0, 1, 1, 0 "
Pattern: " 1,  0,  0,  1,  1,  1,  0   >>>  1,  0,  0,  0,  0,  1,  1 "
Result: " 1, 0, 0, 0, 0, 1, 1 "
Pattern: " 1,  0,  0,  1,  1,  1,  1   >>>  1,  0,  0,  0,  1,  0,  1 "
Result: " 1, 0, 0, 0, 1, 0, 1 "
```

Data Mining and Machine Learning Unit 7 Programming Assignment

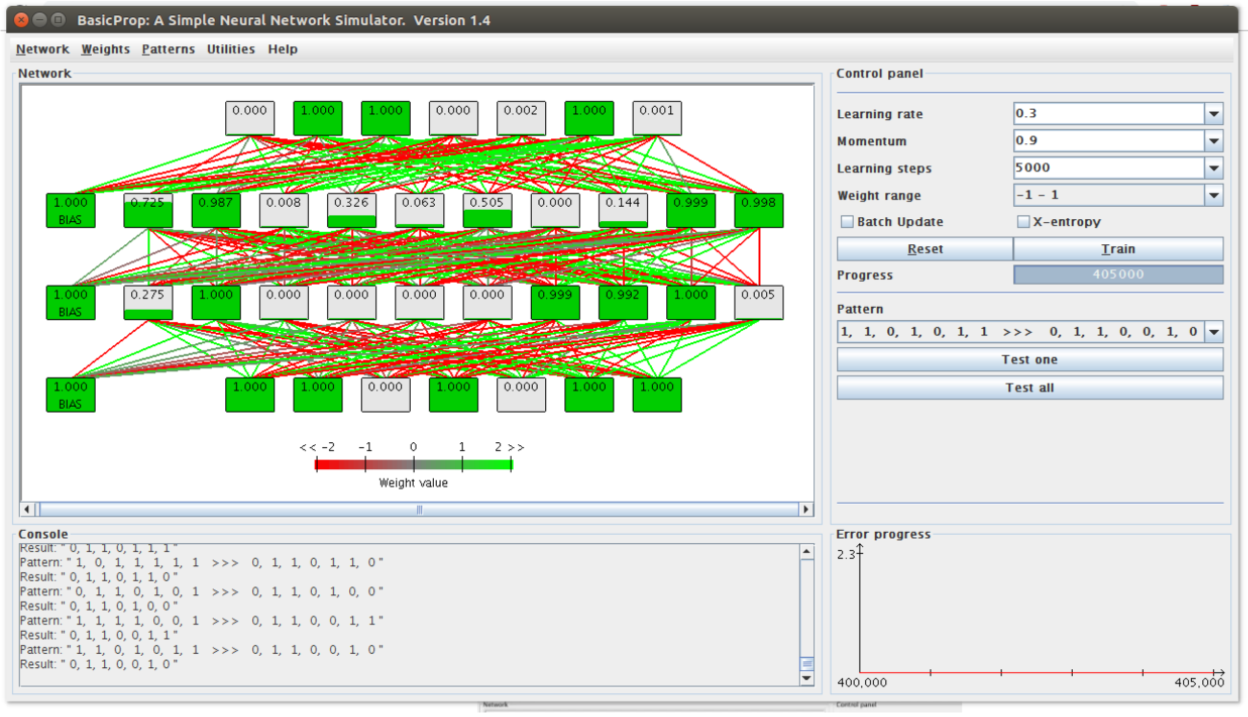

BasicProp: A Simple Neural Network Simulator. Version 1.4
Network Weights Patterns Utilities Help
Network
1.000 0.000 0.000 0.000 0.001 0.999 0.000
1.000 BIAS 0.001 0.870 0.962 0.590 0.912 0.891 0.993 0.000 0.973 0.329
1.000 BIAS 0.361 0.985 0.766 0.815 0.000 0.007 0.005 0.383 0.011 0.281
1.000 BIAS 0.000 0.000 1.000 1.000 1.000 1.000 0.000
<< -2 -1 0 1 2 >>
Weight value
Reset
Train
Test one
Test all
Error progress
2.1
220,000
225,000

Control panel

Learning rate 0.3
Momentum 0.9
Learning steps 5000
Weight range -1 - 1
Batch Update
X-entropy
Progress 225000
Pattern 0, 0, 1, 1, 1, 1, 0 >>> 1, 0, 0, 0, 0, 1, 0

Console

```
Result: " 0, 1, 1, 0, 0, 0, 0 "
Test All: Average per pattern error: 0.005618968575636018
Test All: Average per pattern error: 0.005618968575636018
Pattern: " 1,  0,  1,  1,  1,  1,  1   >>>   0,  1,  1,  0,  1,  1,  0 "
Result: " 0, 1, 1, 0, 1, 1, 0 "
Pattern: " 1,  0,  0,  1,  1,  1,  0   >>>   1,  0,  0,  0,  0,  1,  1 "
Result: " 1, 0, 0, 0, 0, 1, 1 "
Pattern: " 0,  0,  1,  1,  1,  1,  0   >>>   1,  0,  0,  0,  0,  1,  0 "
Result: " 1, 0, 0, 0, 0, 1, 0 "
```


BasicProp: A Simple Neural Network Simulator. Version 1.4
Network Weights Patterns Utilities Help
Network
0.000 1.000 1.000 0.000 0.002 1.000 0.001
1.000 BIAS 0.725 0.987 0.008 0.326 0.063 0.505 0.000 0.144 0.999 0.998
1.000 BIAS 0.275 1.000 0.000 0.000 0.000 0.000 0.999 0.992 1.000 0.005
1.000 BIAS 1.000 1.000 0.000 1.000 0.000 1.000 1.000
<< -2 -1 0 1 2 >>
Weight value
Reset
Train
Test one
Test all
Error progress
2.3
400,000
405,000

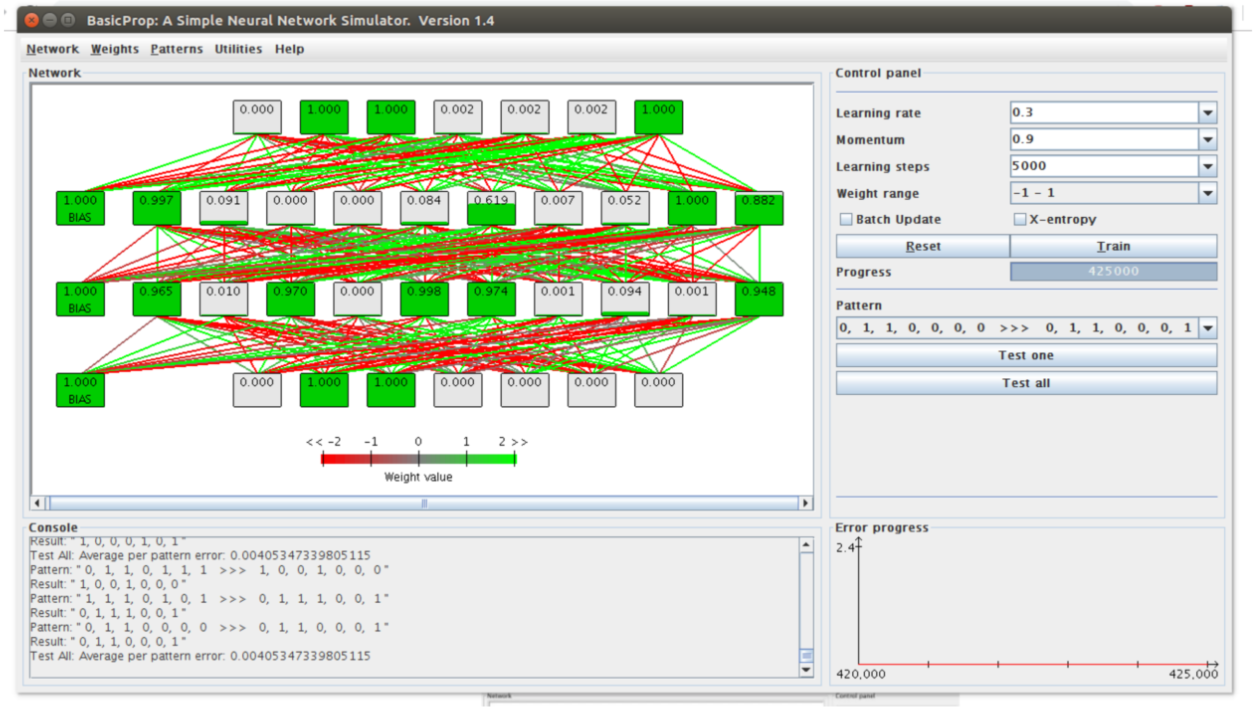
Control panel

Learning rate 0.3
Momentum 0.9
Learning steps 5000
Weight range -1 - 1
Batch Update
X-entropy
Progress 405000
Pattern 1, 1, 0, 1, 0, 1, 1 >>> 0, 1, 1, 0, 0, 1, 0

Console

```
Result: " 0, 1, 1, 0, 1, 1, 1 "
Pattern: " 1,  0,  1,  1,  1,  1,  1   >>>   0,  1,  1,  0,  1,  1,  0 "
Result: " 0, 1, 1, 0, 1, 1, 0 "
Pattern: " 0,  1,  1,  0,  1,  0,  1   >>>   0,  1,  1,  0,  1,  0,  0 "
Result: " 0, 1, 1, 0, 1, 0, 0 "
Pattern: " 1,  1,  1,  1,  0,  0,  1   >>>   0,  1,  1,  0,  0,  1,  1 "
Result: " 0, 1, 1, 0, 0, 1, 1 "
Pattern: " 1,  1,  0,  1,  0,  1,  1   >>>   0,  1,  1,  0,  0,  1,  0 "
Result: " 0, 1, 1, 0, 0, 1, 0 "
```

Data Mining and Machine Learning Unit 7 Programming Assignment

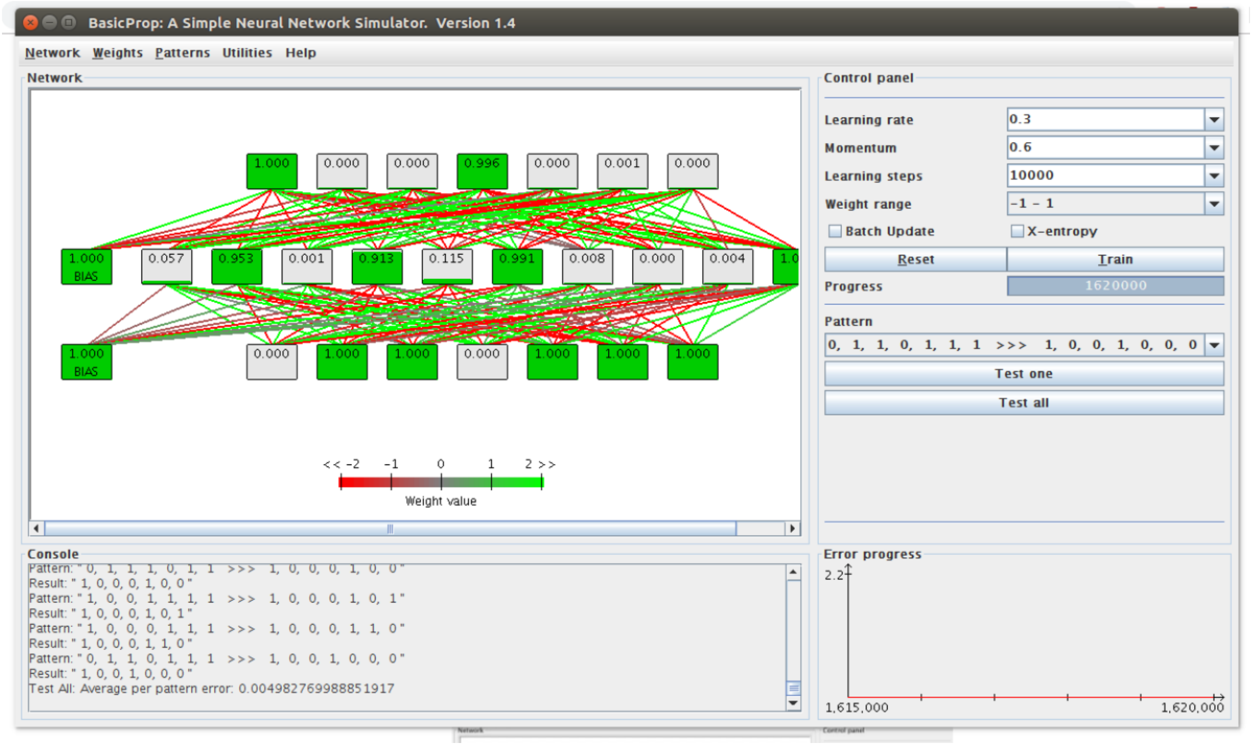


Control panel

Learning rate 0.3
Momentum 0.9
Learning steps 5000
Weight range -1 – 1
Batch Update
X-entropy
Progress 425000
Pattern 0, 1, 1, 0, 0, 0, 0 >>> 0, 1, 1, 0, 0, 0, 1

Console
```
Result: " 1, 0, 0, 0, 1, 0, 1 "
Test All: Average per pattern error: 0.00405347339805115
Pattern: " 0,  1,  1,  0,  1,  1,  1   >>>   1,  0,  0,  1,  0,  0,  0 "
Result: " 1, 0, 0, 1, 0, 0, 0 "
Pattern: " 1,  1,  1,  0,  1,  0,  1   >>>   0,  1,  1,  1,  0,  0,  1 "
Result: " 0, 1, 1, 1, 0, 0, 1 "
Pattern: " 0,  1,  1,  0,  0,  0,  0   >>>   0,  1,  1,  0,  0,  0,  1 "
Result: " 0, 1, 1, 0, 0, 0, 1 "
Test All: Average per pattern error: 0.00405347339805115
```

N2:

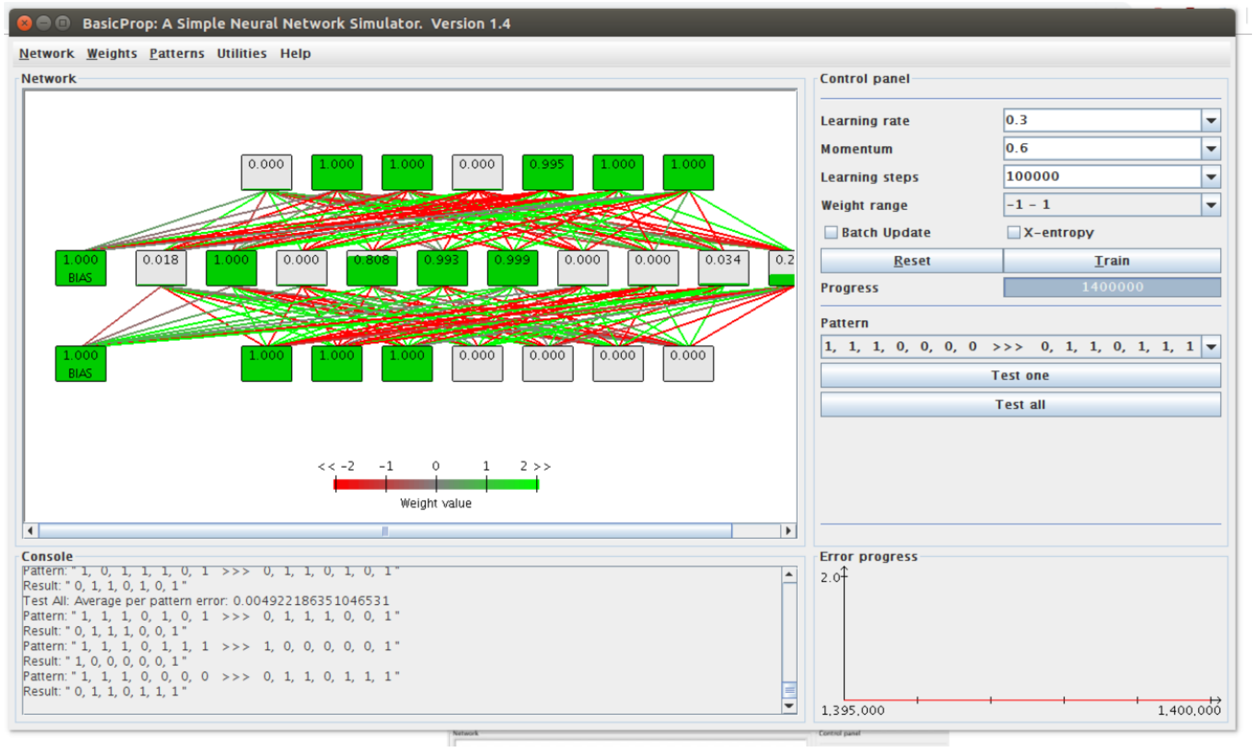


Control panel

Learning rate 0.3
Momentum 0.6
Learning steps 10000
Weight range -1 – 1
Batch Update
X-entropy
Progress 1620000
Pattern 0, 1, 1, 0, 1, 1, 1 >>> 1, 0, 0, 1, 0, 0, 0

Console
```
Pattern: " 0,  1,  1,  1,  0,  1,  1   >>>   1,  0,  0,  0,  1,  0,  0 "
Result: " 1, 0, 0, 0, 1, 0, 0 "
Pattern: " 1,  0,  0,  1,  1,  1,  1   >>>   1,  0,  0,  0,  1,  0,  1 "
Result: " 1, 0, 0, 0, 1, 0, 1 "
Pattern: " 1,  0,  0,  0,  1,  1,  1   >>>   1,  0,  0,  0,  1,  1,  0 "
Result: " 1, 0, 0, 0, 1, 1, 0 "
Pattern: " 0,  1,  1,  0,  1,  1,  1   >>>   1,  0,  0,  1,  0,  0,  0 "
Result: " 1, 0, 0, 1, 0, 0, 0 "
Test All: Average per pattern error: 0.004982769988851917
```

Data Mining and Machine Learning Unit 7 Programming Assignment


BasicProp: A Simple Neural Network Simulator. Version 1.4
Network Weights Patterns Utilities Help
Network
0.000 1.000 1.000 0.000 0.995 1.000 1.000
1.000 BIAS 0.018 1.000 0.000 0.808 0.993 0.999 0.000 0.000 0.034 0.2
1.000 BIAS 1.000 1.000 1.000 0.000 0.000 0.000 0.000
<< -2 -1 0 1 2 >>
Weight value
Reset
Train
Test one
Test all
Error progress
2.0
1,395,000
1,400,000

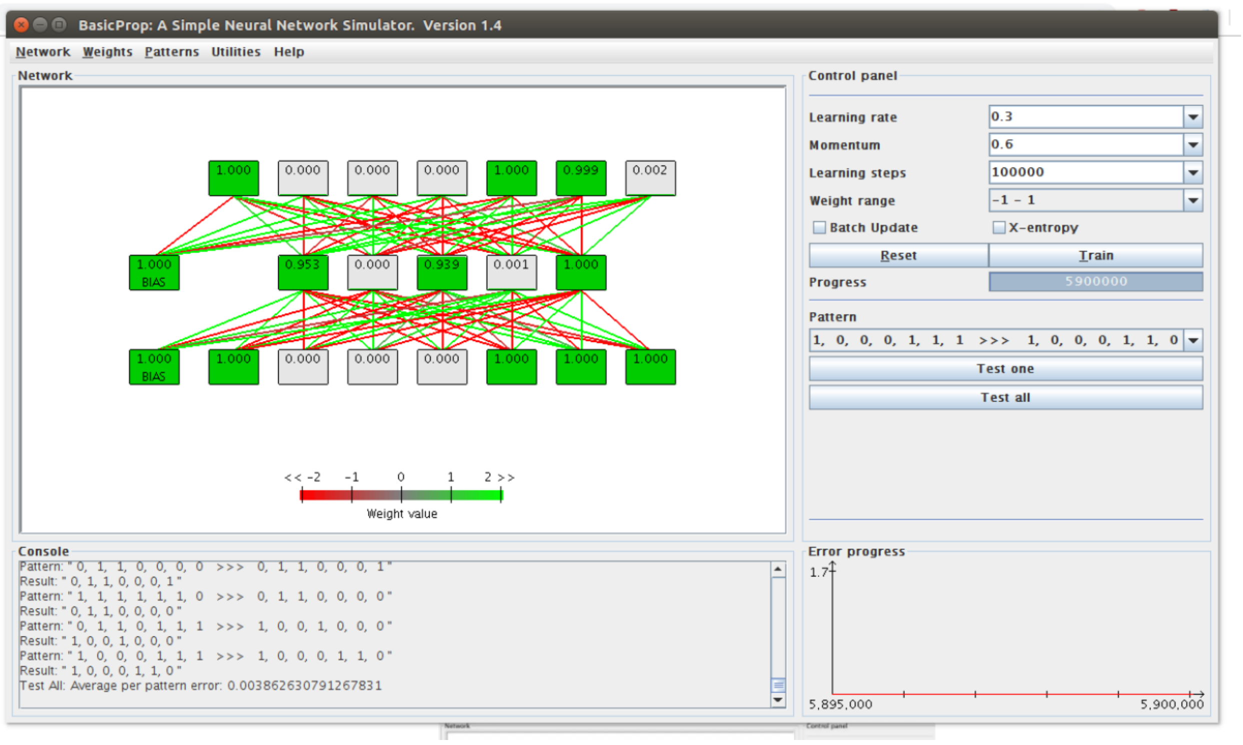
Control panel

Learning rate 0.3
Momentum 0.6
Learning steps 100000
Weight range -1 - 1
Batch Update
X-entropy
Progress 1400000
Pattern 1, 1, 1, 0, 0, 0, 0 >>> 0, 1, 1, 0, 1, 1, 1

Console

```
Pattern: " 1, 0, 1, 1, 1, 0, 1 >>> 0, 1, 1, 0, 1, 0, 1"
Result: " 0, 1, 1, 0, 1, 0, 1 "
Test All: Average per pattern error: 0.004922186351046531
Pattern: " 1, 1, 1, 0, 1, 0, 1 >>> 0, 1, 1, 1, 0, 0, 1"
Result: " 0, 1, 1, 1, 0, 0, 1 "
Pattern: " 1, 1, 1, 0, 1, 1, 1 >>> 1, 0, 0, 0, 0, 0, 1"
Result: " 1, 0, 0, 0, 0, 0, 1 "
Pattern: " 1, 1, 1, 0, 0, 0, 0 >>> 0, 1, 1, 0, 1, 1, 1"
Result: " 0, 1, 1, 0, 1, 1, 1 "
```

N1:


BasicProp: A Simple Neural Network Simulator. Version 1.4
Network Weights Patterns Utilities Help
Network
1.000 0.000 0.000 0.000 1.000 0.999 0.002
1.000 BIAS 0.953 0.000 0.939 0.001 1.000
1.000 BIAS 1.000 0.000 0.000 0.000 1.000 1.000 1.000
<< -2 -1 0 1 2 >>
Weight value
Reset
Train
Test one
Test all
Error progress
1.7
5,895,000
5,900,000

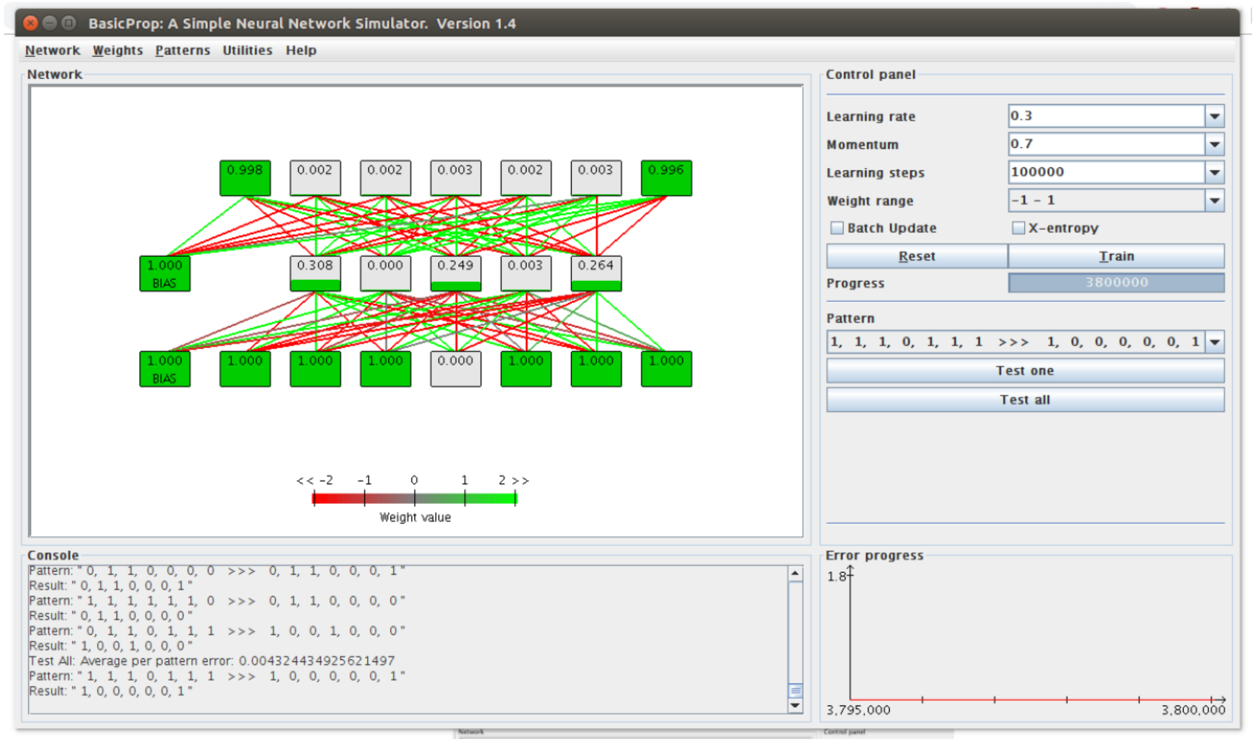
Control panel

Learning rate 0.3
Momentum 0.6
Learning steps 100000
Weight range -1 - 1
Batch Update
X-entropy
Progress 5900000
Pattern 1, 0, 0, 0, 1, 1, 1 >>> 1, 0, 0, 0, 1, 1, 0

Console

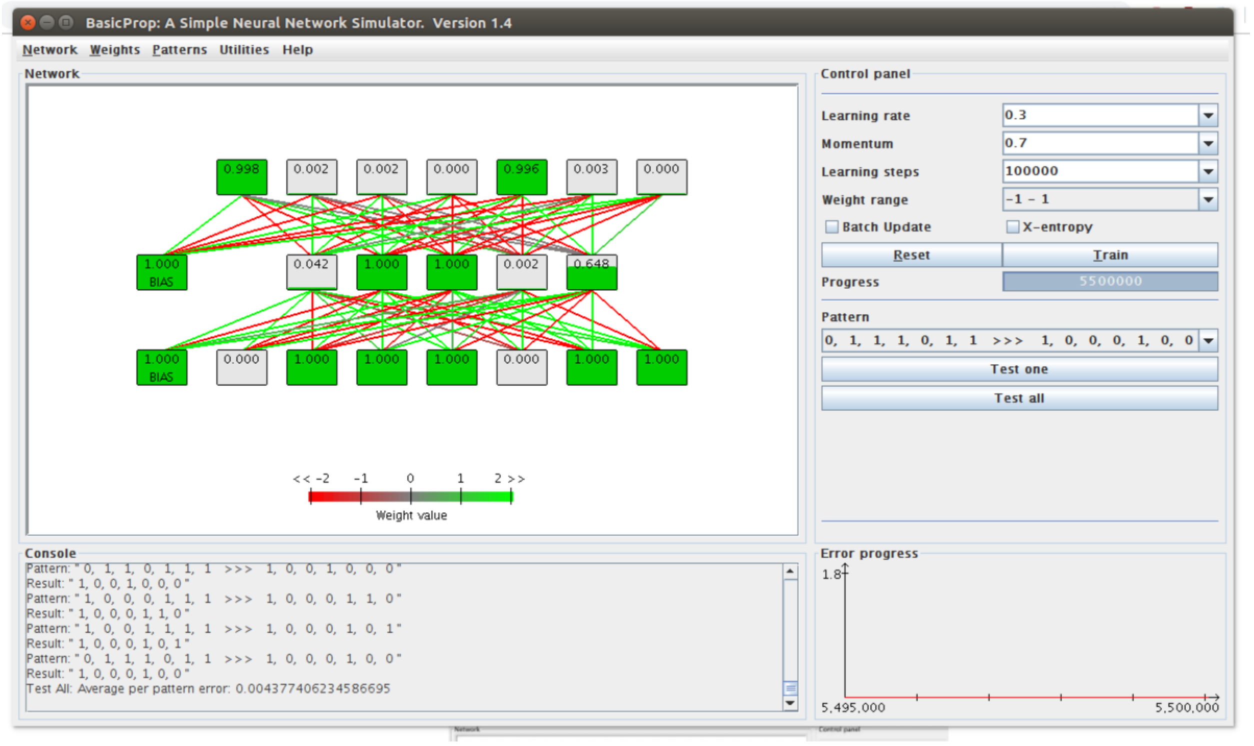
```
Pattern: " 0, 1, 1, 0, 0, 0, 0 >>> 0, 1, 1, 0, 0, 0, 1"
Result: " 0, 1, 1, 0, 0, 0, 1 "
Pattern: " 1, 1, 1, 1, 1, 1, 0 >>> 0, 1, 1, 0, 0, 0, 0"
Result: " 0, 1, 1, 0, 0, 0, 0 "
Pattern: " 0, 1, 1, 0, 1, 1, 1 >>> 1, 0, 0, 1, 0, 0, 0"
Result: " 1, 0, 0, 1, 0, 0, 0 "
Pattern: " 1, 0, 0, 0, 1, 1, 1 >>> 1, 0, 0, 0, 1, 1, 0"
Result: " 1, 0, 0, 0, 1, 1, 0 "
Test All: Average per pattern error: 0.003862630791267831
```

Data Mining and Machine Learning Unit 7 Programming Assignment


BasicProp: A Simple Neural Network Simulator. Version 1.4
Network Weights Patterns Utilities Help
Network
0.998 0.002 0.002 0.003 0.002 0.003 0.996
1.000 BIAS 0.308 0.000 0.249 0.003 0.264
1.000 BIAS 1.000 1.000 1.000 0.000 1.000 1.000 1.000
<< -2 -1 0 1 2 >>
Weight value
Control panel
Reset
Train
Test one
Test all
Error progress
1.8
3,795,000
3,800,000

Learning rate 0.3
Momentum 0.7
Learning steps 100000
Weight range -1 - 1
Batch Update
X-entropy
Progress 3800000
Pattern
1, 1, 1, 0, 1, 1, 1 >>> 1, 0, 0, 0, 0, 0, 1

Console
```
Pattern: " 0,  1,  1,  0,  0,  0,  0   >>>   0,  1,  1,  0,  0,  0,  1 "
Result: " 0, 1, 1, 0, 0, 0, 1 "
Pattern: " 1,  1,  1,  1,  1,  1,  0   >>>   0,  1,  1,  0,  0,  0,  0 "
Result: " 0, 1, 1, 0, 0, 0, 0 "
Pattern: " 0,  1,  1,  0,  1,  1,  1   >>>   1,  0,  0,  1,  0,  0,  0 "
Result: " 1, 0, 0, 1, 0, 0, 0 "
Test All: Average per pattern error: 0.004324434925621497
Pattern: " 1,  1,  1,  0,  1,  1,  1   >>>   1,  0,  0,  0,  0,  0,  1 "
Result: " 1, 0, 0, 0, 0, 0, 1 "
```

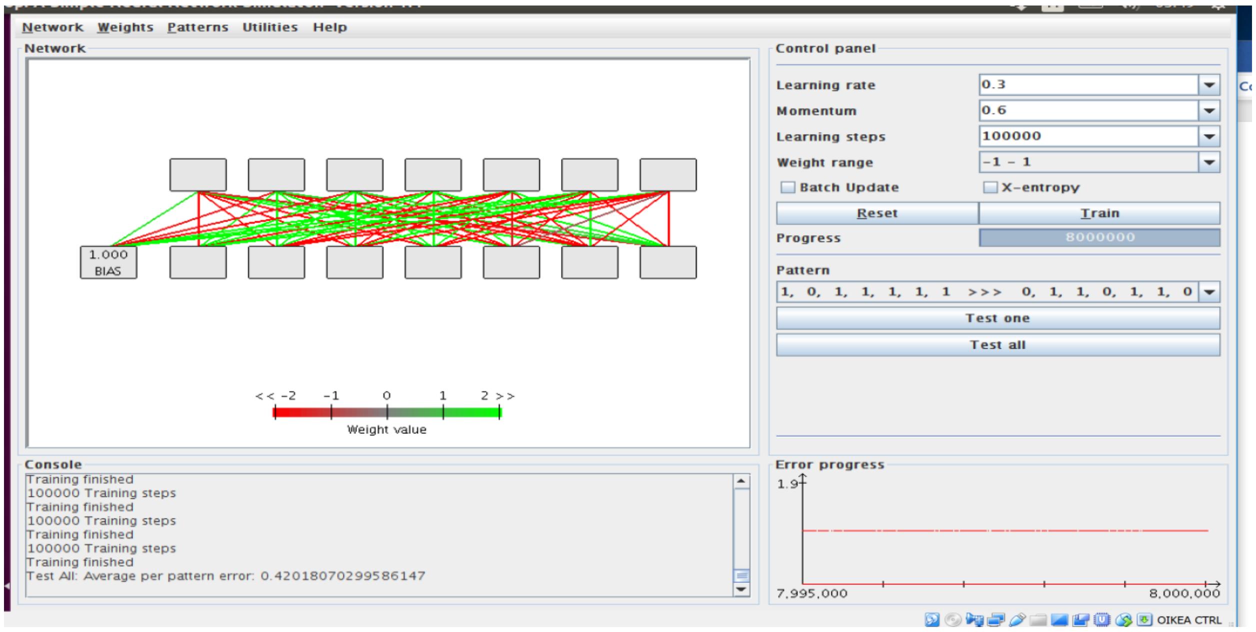

BasicProp: A Simple Neural Network Simulator. Version 1.4
Network Weights Patterns Utilities Help
Network
0.998 0.002 0.002 0.000 0.996 0.003 0.000
1.000 BIAS 0.042 1.000 1.000 0.002 0.648
1.000 BIAS 0.000 1.000 1.000 1.000 0.000 1.000 1.000
<< -2 -1 0 1 2 >>
Weight value
Control panel
Reset
Train
Test one
Test all
Error progress
1.8
5,495,000
5,500,000

Learning rate 0.3
Momentum 0.7
Learning steps 100000
Weight range -1 - 1
Batch Update
X-entropy
Progress 5500000
Pattern
0, 1, 1, 1, 0, 1, 1 >>> 1, 0, 0, 0, 1, 0, 0

Console
```
Pattern: " 0,  1,  1,  0,  1,  1,  1   >>>   1,  0,  0,  1,  0,  0,  0 "
Result: " 1, 0, 0, 1, 0, 0, 0 "
Pattern: " 1,  0,  0,  0,  1,  1,  1   >>>   1,  0,  0,  0,  1,  1,  0 "
Result: " 1, 0, 0, 0, 1, 1, 0 "
Pattern: " 1,  0,  0,  1,  1,  1,  1   >>>   1,  0,  0,  0,  1,  0,  1 "
Result: " 1, 0, 0, 0, 1, 0, 1 "
Pattern: " 0,  1,  1,  1,  0,  1,  1   >>>   1,  0,  0,  0,  1,  0,  0 "
Result: " 1, 0, 0, 0, 1, 0, 0 "
Test All: Average per pattern error: 0.004377406234586695
```

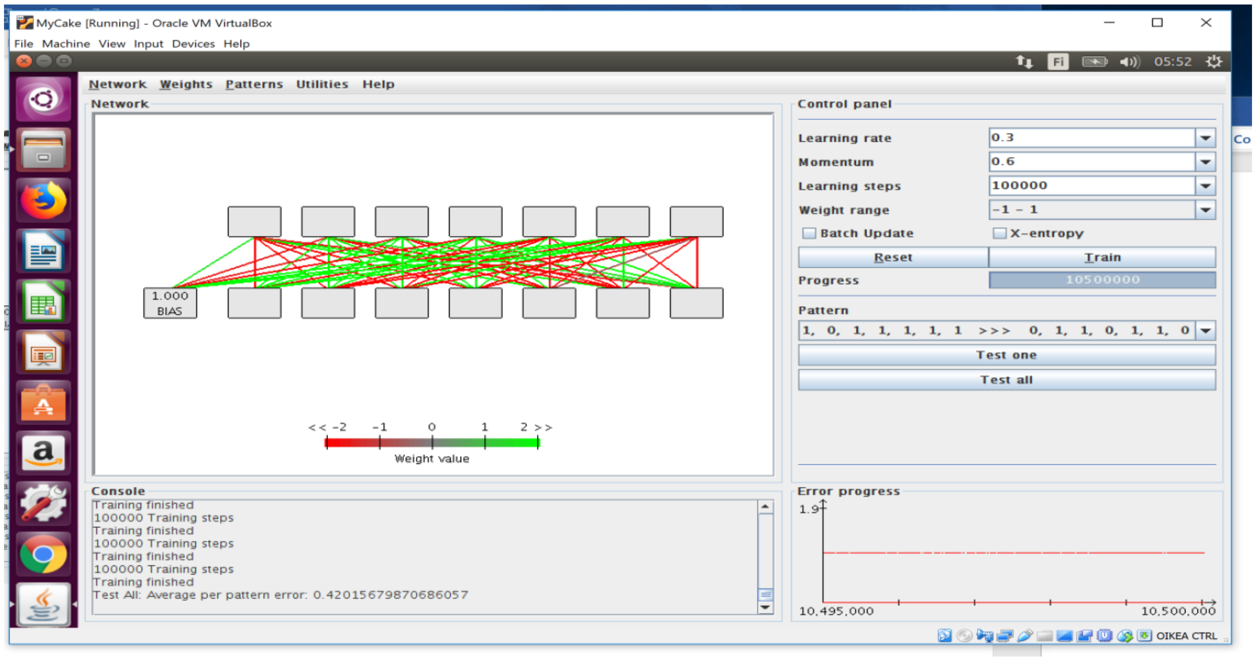
N:

Data Mining and Machine Learning Unit 7 Programming Assignment



Learning rate 0.3
Momentum 0.6
Learning steps 100000
Weight range -1 - 1
Batch Update
X-entropy
Progress 8000000
Pattern 1, 0, 1, 1, 1, 1, 1 >>> 0, 1, 1, 0, 1, 1, 0

Console
```
Training finished
100000 Training steps
Training finished
100000 Training steps
Training finished
100000 Training steps
Training finished
Test All: Average per pattern error: 0.42018070299586147
```



Learning rate 0.3
Momentum 0.6
Learning steps 100000
Weight range -1 - 1
Batch Update
X-entropy
Progress 10500000
Pattern 1, 0, 1, 1, 1, 1, 1 >>> 0, 1, 1, 0, 1, 1, 0

Console
```
Training finished
100000 Training steps
Training finished
100000 Training steps
Training finished
100000 Training steps
Training finished
Test All: Average per pattern error: 0.42015679870686057
```