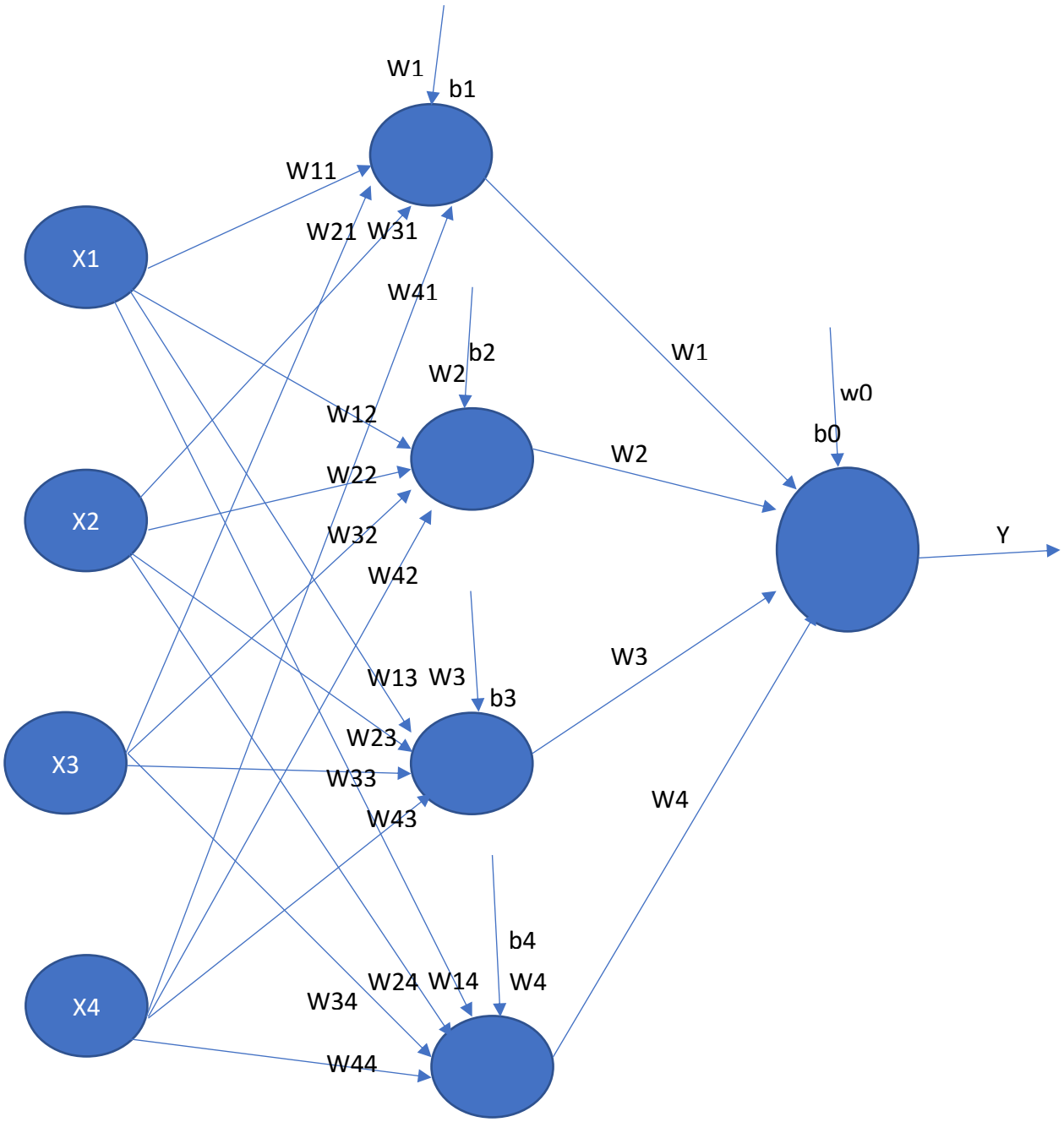
Saheli Saha
sasaha@iu.edu
Machine Learning (CSCI-B 555)

Programming Assignment

Implement a two-layer perceptron with the backpropagation algorithm to solve the parity problem: -

As per the given problem out perceptron should look like the following: -



Saheli Saha
sasaha@iu.edu
Machine Learning (CSCI-B 555)

For input X (X1, X2, X3 and X4) I have taken a list of binary numbers 0-16.
Expected Output Y should be 0 or 1.

As per my code the output is coming as follows: -

Epoch 155056

Output	Desired
0.107533168342	0
0.950002057309	1
0.915276013809	1
0.914470455766	1
0.893277922647	1
0.0483810477537	0
0.0492718702971	0
0.0415243560055	0
0.0884719364765	0
0.0490727177503	0
0.0494123872454	0
0.0049026370655	1
0.964243165618	1
0.978772408795	1
0.96405009256	1
0.0569161855757	0

Many times, my epoch is varying widely in the range of 60,000 to 2,00,000. I am terminating my model when perceptron is learning sufficiently to predict the dependent variable and when the error rate is less than 0.05.

1. While varying η from 0.05 to 0.5

η = 0.05
Epoch 224422
Output
0.12686274007, 0.816585324115, 0.970690951878, 0.0201626817081, 0.893952685572, 0.190512987309, 0.028180460982, 0.950007991252, 0.970007833282, 0.013892660728, 0.991516279916, 0.953305964382, 0.0325632814263, 0.953275565125, 0.905596907899,
0.106801839234.

Saheli Saha
sasaha@iu.edu
Machine Learning (CSCI-B 555)

η = 0.10
Epoch 119606
0.1292143111, 0.818592278805, 0.970076636413, 0.0204314464696, 0.893889788098,
0.189876709064, 0.0339655054648, 0.950001125864,
0.969742999471,0.0168061292316, 0.992744765486, 0.955783188668,
0.0361503105332, 0.951985776121, 0.898999111728, 0.106773483529.

η = 0.15
Epoch 116842
0.125378200962, 0.974761759978, 0.82907481773, 0.046383110415, 0.906131540436,
0.0546891769525, 0.0297437860923, 0.962223154167, 0.829052334148,
0.0464798224231, 0.625863527202, 0.559929170298, 0.0298272132411,
0.962314386807, 0.95001272622,
0.0197318008182.

η = 0.20
Epoch: 151823
0.0782931389173, 0.974505662612, 0.920570339242, 0.0263122730979,
0.961522316594, 0.0123829493079, 0.00966341535101, 0.972261508993,
0.918506902253, 0.57658448127, 0.058279892398, 0.568461703168, 0.0144714140752,
0.950000462652, 0.995335804832, 0.579263216371.

η = 0.25
Epoch 331440
0.107775104706, 0.950296640684, 0.924831708607, 0.0125039741303, 0.9500008839,
0.0583051892458, 0.0122969779445, 0.797036574829, 0.926209498726,
0.0165432491625, 2.34232493175e-05, 0.961888570717, 0.016297405748,
0.774332097754, 0.960427226098, 0.708247155407.

η = 0.30
Epoch 280880
0.107672826283, 0.950371675179, 0.927165522501, 0.0133634840637,
0.950000810201, 0.0602251379488, 0.0131906782599, 0.797711209162,
0.928034931194, 0.016410496443, 2.45357341202e-05, 0.960831216554,
0.0161208875311, 0.780783947349, 0.958995735046, 0.70886609164.

η = 0.35
Epoch 238016
0.10845718947, 0.950443253264, 0.928706524707, 0.0140063642177, 0.950000396517,
0.0619886830223, 0.0138153240977, 0.797960472957, 0.929388157153,
0.016595694012, 2.64984409686e-05, 0.961080575946, 0.016251115195,
0.782701707057, 0.958949424064, 0.709634870908.

Saheli Saha
sasaha@iu.edu
Machine Learning (CSCI-B 555)

η = 0.40
Epoch 204048
0.109561857978, 0.950516257927, 0.92997700654, 0.0145887277792, 0.950000181235,
0.0636959886852, 0.0143701381567, 0.797904970985, 0.93055804725,
0.0168843957674, 2.89397457281e-05, 0.961783666104, 0.0164822932023,
0.783247343235, 0.959389358401, 0.710477718017.

η = 0.45
Epoch 176944
0.110831005773, 0.95059185229, 0.931111818306, 0.0151349846703, 0.950000933863,
0.0653719273176, 0.0148847154783, 0.797671518736, 0.931634145142,
0.0172379154177, 3.18111988754e-05, 0.962710093415, 0.0167767091842,
0.783103361833, 0.96007807125, 0.711360271964.

For all the cases my model is converging. As we can see the results epoch is minimum while $\eta = 0.15$ and maximum when $\eta = 0.25$. Prediction wise all are close. But for convergence rate, when η is too small learning converges slowly but for large η value learning does not converge at all.

2. I have added momentum by dividing learning rate(η) by (1-α). For low value of η, momentum fasten the execution but for high value, it slows down. But in this process, it smooths the weight change and stabilize the oscillation.

η = 0.05
Epoch 155056
0.112186507738, 0.950668847073, 0.932175790217, 0.0156564146097,0.95000144882
0.0670322179378, 0.0153718898492, 0.797347935185, 0.932662551209,
0.0176372974092, 3.51154978352e-05, 0.963741589974, 0.0171155388166,
0.782588786388, 0.960893873293, 0.712263523691.

η = 0.10
Epoch 59056
0.126642602113, 0.951483755325, 0.941622159191, 0.0194234423683,
0.950004047618, 0.0833975807267, 0.018650729656, 0.797965664714,
0.942703608084, 0.023704180746, 0.000103157041052, 0.974164875887,
0.0225004996706, 0.767497397332, 0.970181452459, 0.721340900223.

Saheli Saha
sasaha@iu.edu
Machine Learning (CSCI-B 555)

η = 0.15
Epoch 203893
0.0167307187598, 0.930802188716, 0.930787800276, 0.590948394708, 0.92598939334
0.0499800098488, 0.0499996198293, 0.572885502623, 0.995917377195,
0.00388707814008, 0.00388921002219, 0.859836656589, 0.0129135680802,
0.982168717075, 0.982175181967, 0.605904660038.

As we can see the outputs are quite irregular. For η more than 0.15, execution rate becomes really slow and outputs are irregular. For more higher value of η the process is going to infinite loop. Hence, momentum works well with low value of η.