CSCI-B 555 – Machine Learning – Programming Assignment

Subramanian Sivaraman | 2000114337 | ssivaram@umail.iu.edu

Implement a two-layer perceptron with the backpropagation algorithm to solve the parity problem. The desired output for the parity problem is 1 if an input pattern contains an odd number of 1's and 0 otherwise. Follow the algorithm introduced in class and consult textbooks and other references. Use a network with 4 binary input elements, 4 hidden units for the first layer, and one output unit for the second layer. The learning procedure is stopped when an absolute error (difference) of 0.05 is reached for every input pattern. Other implementation details are:

- *Initialize all weights and biases to random numbers between -1 and 1.*
- *Use a logistic sigmoid with a = 1 as the activation function for all units.*

General Information:

- The algorithm was implemented in R (v) using R-Studio.
- Weights were initialized randomly from a uniform distribution of values between -1 and 1.
- A logistic sigmoid with a = 1 in $\phi(x) = 1/(1 + e^{-ax})$ was used.

The Code:

- The code is commented with appropriate explanations.
- Two sets of codes are attached – with and without the momentum term.

Results:

a) Without the momentum term:

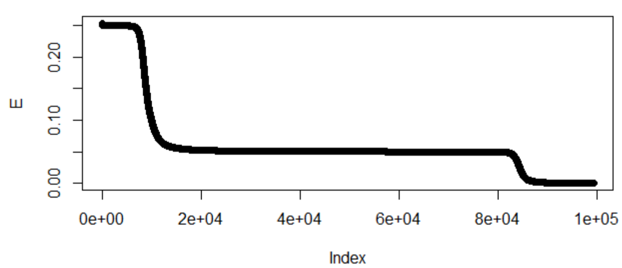
η	Epochs	MSE
0.05	99249	0.000472285
0.10	48052	0.000475731
0.15	31038	0.000479191
0.20	22564	0.000482849
0.25	17503	0.000486611
0.30	14146	0.000490312
0.35	11757	0.000494457
0.40	9970	0.000498813
0.45	8580	0.000503279
0.50	7460	0.000508537

- The above table shows the number of epochs for each run of the algorithm with each value of η.
- An increase in η corresponds to a significant decrease in the number of epochs.
- The MSE column indicates the Mean Squared Errors for all runs of the algorithm.
- The absolute errors for each input sample under each run of η are provided below as proof of the algorithm's convergence. These results can be reproduced by simply running the code.

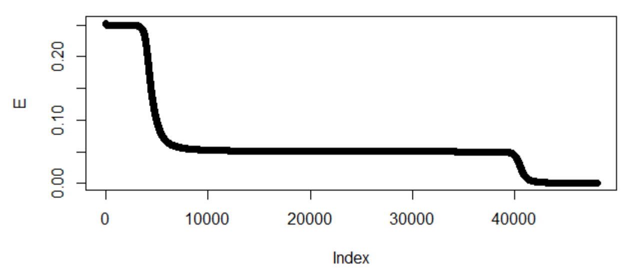
	η = 0.05	η = 0.10	η = 0.15	η = 0.20	η = 0.25	η = 0.30	η = 0.35	η = 0.40	η = 0.45	η = 0.50
input sample 1	-0.019751751	-0.01988343	-0.020013976	-0.020146433	-0.020279666	-0.0204115	-0.0205505	-0.0206938	-0.0208413	-0.0210058
input sample 2	0.033551696	0.033578056	0.033602165	0.033631124	0.033660479	0.03368273	0.03371671	0.03375083	0.03377746	0.03381779
input sample 3	0.010708228	0.010974908	0.011243263	0.011514457	0.011789896	0.01207144	0.01236254	0.01266764	0.01299386	0.01335376
input sample 4	-0.011351236	-0.011447669	-0.011544549	-0.011643521	-0.011744203	-0.0118458	-0.011953	-0.0120651	-0.0121832	-0.0123155
input sample 5	0.011020001	0.011297684	0.011576922	0.011858939	0.012145127	0.0124373	0.01273915	0.01305511	0.01339231	0.01376377
input sample 6	-0.011182351	-0.011274877	-0.011367874	-0.011462927	-0.011559682	-0.0116574	-0.0117606	-0.0118686	-0.0119825	-0.0121104
input sample 7	-0.018965923	-0.019286767	-0.019612314	-0.019944307	-0.020283009	-0.0206288	-0.0209881	-0.0213636	-0.021762	-0.0222003
input sample 8	0.009367707	0.009581349	0.009793843	0.010006235	0.010219496	0.01043475	0.01065504	0.01088316	0.01112367	0.01138593
input sample 9	0.033549098	0.033575227	0.033599097	0.033627811	0.033656912	0.0336789	0.03371261	0.03374645	0.03377278	0.03381279
input sample 10	-0.049998821	-0.049998515	-0.049993981	-0.049996389	-0.049998473	-0.0499881	-0.0499946	-0.0499989	-0.0499878	-0.0499931
input sample 11	-0.011437176	-0.011535647	-0.011634542	-0.011735534	-0.011838229	-0.0119418	-0.012051	-0.0121652	-0.0122853	-0.0124199
input sample 12	0.023133501	0.023178184	0.023221892	0.023269174	0.023317317	0.02336174	0.02341463	0.02346909	0.023521	0.02358496
input sample 13	-0.011265446	-0.011359895	-0.011454794	-0.011551754	-0.011650408	-0.01175	-0.0118551	-0.011965	-0.0120809	-0.0122108
input sample 14	0.023053821	0.023099734	0.02314471	0.023193244	0.023242659	0.02328843	0.02334261	0.02339842	0.02345179	0.0235172
input sample 15	0.00935796	0.009570637	0.009782125	0.009993466	0.010205626	0.01041973	0.01063879	0.01086559	0.01110466	0.01136529
input sample 16	-0.008117585	-0.008193534	-0.008269738	-0.008347209	-0.008425832	-0.0085053	-0.0085885	-0.0086754	-0.0087671	-0.0088693

- The Error function follows pretty much the same pattern for all values of η. What is interesting is that as the value of η increases, the period of zero MSE change decreases. The graphs of error functions for all values of η are provided below. The labels for the graphs are summarized in this table:

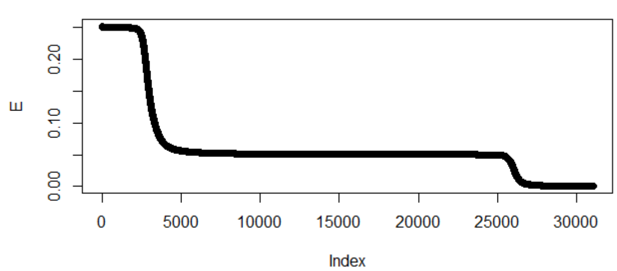
η = 0.05



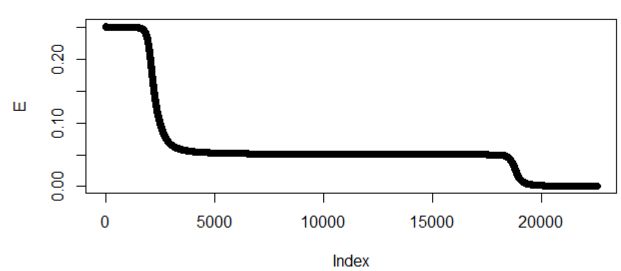
η = 0.10

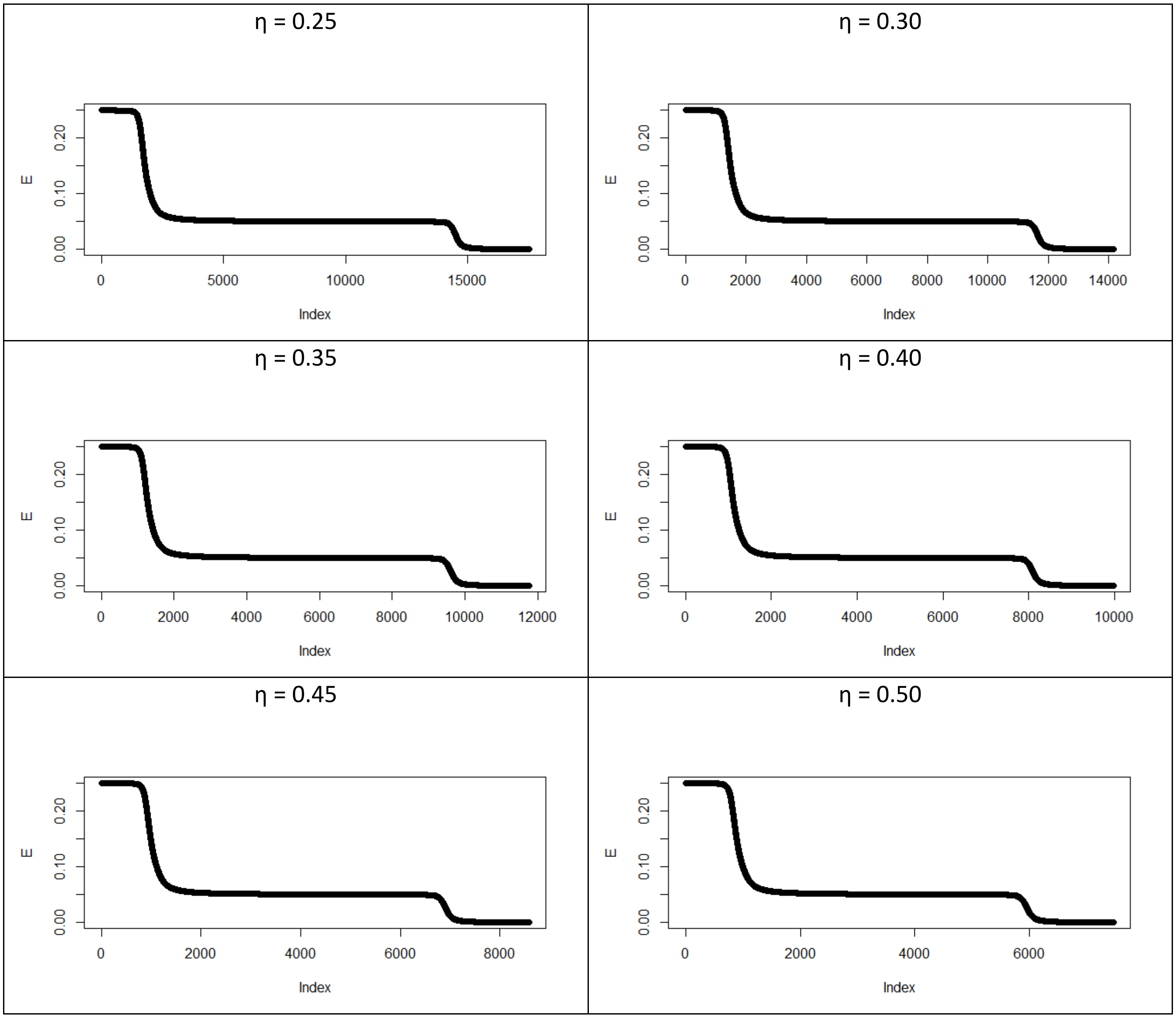


η = 0.15



η = 0.20





- The above graphs show decrease in the MSE values with respect to the number of epochs for various values of η.

b) With the momentum term (α):

η	Epochs	MSE
0.05	48055	0.000481513
0.10	21563	0.000495077
0.15	13045	0.000509009
0.20	8955	0.000523577
0.25	6600	0.000538688
0.30	5094	0.000554957
0.35	4103	0.000569756
0.40	2678	0.000584025
0.45	2612	0.00056438
0.50	2632	0.000564811

- With the added momentum term (α = 0.9), the number of epochs for each value of η decrease even further. This confirms that the momentum term eases the oscillating weights by adding some stability to their changes over each iteration. The acceleration of learning is given by 1/(1-α) and it helps smooth the weight change.

	η = 0.05	η = 0.10	η = 0.15	η = 0.20	η = 0.25	η = 0.30	η = 0.35	η = 0.40	η = 0.45	η = 0.50
input sample 1	-0.0200976	-0.0205704	-0.0210202	-0.0214562	-0.0218805	-0.0223031	-0.0226608	-0.0217509	-0.0211266	-0.0210868
input sample 2	0.03362189	0.0337227	0.0338212	0.03392325	0.0340173	0.0341224	0.03421026	0.03363363	0.03353606	0.03352441
input sample 3	0.0114129	0.01240379	0.01338631	0.01437793	0.01539765	0.01645656	0.01741445	0.01565827	0.01413904	0.01421851
input sample 4	-0.0116067	-0.0119681	-0.0123263	-0.0126875	-0.0130557	-0.0134378	-0.0137769	-0.0160969	-0.0154762	-0.0154775
input sample 5	0.01175337	0.01278199	0.01379759	0.01481705	0.01585811	0.0169302	0.01788853	0.01564316	0.0137591	0.01386682
input sample 6	-0.0114275	-0.011775	-0.0121206	-0.0124706	-0.0128294	-0.0132044	-0.0135398	-0.0161084	-0.0163097	-0.0163664
input sample 7	-0.0198201	-0.0210396	-0.0222416	-0.023429	-0.0246117	-0.0257997	-0.0268329	-0.0304766	-0.0287901	-0.0288065
input sample 8	0.00992699	0.01068598	0.01140908	0.01211205	0.01280878	0.01350769	0.01411581	0.01035791	0.00908728	0.00915452
input sample 9	0.03361866	0.03371856	0.03381616	0.03391738	0.03401072	0.03411529	0.03420286	0.03363435	0.03355127	0.03353968
input sample 10	-0.0499981	-0.0499975	-0.0499938	-0.0499951	-0.0499812	-0.0499815	-0.0499767	-0.0499903	-0.0499929	-0.0499789
input sample 11	-0.0116979	-0.0120664	-0.0124309	-0.0127973	-0.0131696	-0.0135544	-0.0138946	-0.0160928	-0.0154946	-0.0154908
input sample 12	0.02325246	0.02342288	0.02359065	0.02375926	0.02392182	0.02409027	0.02423309	0.02518889	0.02485228	0.02485103
input sample 13	-0.0115156	-0.0118698	-0.0122213	-0.0125763	-0.012939	-0.0133166	-0.0136531	-0.0161038	-0.0163158	-0.0163679
input sample 14	0.02317606	0.02335105	0.02352308	0.02369535	0.02386086	0.02403113	0.02417474	0.02522542	0.02516648	0.0251406
input sample 15	0.00991462	0.01066956	0.01138829	0.01208654	0.01277811	0.01347129	0.0140739	0.01044815	0.01006315	0.01008863
input sample 16	-0.0083183	-0.0086	-0.0088772	-0.0091551	-0.0094383	-0.0097322	-0.0099942	-0.0129194	-0.0130141	-0.0130717