CSE 5526

Programming Assignment 1

Pragya Arora

arora.170

Two-layer perceptron with the backpropagation algorithm

1. Number of epochs for each η

η	Epochs
0.05	33742
0.10	17479
0.15	11876
0.20	9841
0.25	7995
0.30	6627
0.35	5725
0.40	5614
0.45	5845
0.50	101003

Analysis:

It can be observed that the number of epochs decreases consistently with increase in η from 0.05 to 0.40. This clearly indicates that increasing the learning rate fastens the descent in the gradient descent algorithm. Beyond η = 0.40, an increase in number of epochs can be noticed indicating that the gradient descent algorithm is oscillating from this point. At η=0.50, the maximum number of epochs can be observed which is more than 17 times greater than the preceding epochs for η=0.45. This indicates that the algorithm takes plenty of iterations to converge because of the oscillating weights. Through the various steps in η, the minimum value of epochs is found to be 5614 at η=0.40.

2. Number of epochs for each η with 0.9 momentum

η	Epochs	Effect on speed of training
0.05	3390	9.95 x
0.10	1755	9.96 x
0.15	1200	9.90 x

0.20	981	10.03 x
0.25	827	9.67 x
0.30	687	9.65 x
0.35	590	9.70 x
0.40	561	10.01 x
0.45	5135	1.13 x
0.50	9351	10.80 x

Analysis:

Using momentum equal to 0.9, on average a speed-up of around 9 times is noticed as compared to when momentum is not used in the weight update step of back propagation. This ties up with the expected speed-up of $1/(1-\alpha) \sim 10$ times in down-hill regions of gradient descent. The pattern of change in epochs with respect to η is exactly same as the one noticed without momentum. The number of epochs reduces from $\eta=0.05$ to 0.40. This increase beyond $\eta=0.40$ is due to oscillations in the weight update, as seen before. But, in general, including momentum helps in stabilization of gradient descent algorithm when oscillating and thus converge faster.

3. Comparison of desired and actual output

Desired output: [0 1 1 0 1 0 0 1 1 0 0 1 0 1 1 0]

3.1. Without Momentum

η	Actual Output
0.05	[0.01719589, 0.9702333 , 0.9745374 , 0.01181877, 0.97025145, 0.04999876, 0.01178972, 0.97845215, 0.97455687, 0.01170813, 0.04477065, 0.98243616, 0.01166873, 0.97851944, 0.98265201, 0.00866039]
0.10	[0.01664883, 0.97028034, 0.97526107, 0.01140069, 0.9703005 , 0.0499974 , 0.01141417, 0.97852461, 0.97527777, 0.01109622, 0.0440979 , 0.98297583, 0.0110879 , 0.97867991, 0.98342299, 0.00821976]
0.15	[0.01625382, 0.97030495, 0.97574691, 0.0109469 , 0.9703265 ,0.04999682, 0.01093827, 0.97857496, 0.97576624, 0.01071661,0.04379211, 0.98350223, 0.01068368, 0.9787216 , 0.98389303,0.00780873]

0.20	[0.02125671, 0.96645481, 0.98690973, 0.01586552, 0.96645987, 0.04999918, 0.01587547, 0.97488999, 0.9864596 , 0.01525973, 0.0287422 , 0.99043802, 0.01526035, 0.9752427 , 0.99115899, 0.01252639]
0.25	[0.02126435, 0.96644367, 0.98705269, 0.01595153, 0.96644818, 0.04999595, 0.01599089, 0.97490753, 0.98651189, 0.01522115, 0.02855643, 0.99040339, 0.01524994, 0.9752737 , 0.99113289, 0.01259068]
0.30	[0.02144661, 0.96641511, 0.98651857, 0.01593043, 0.966416 ,0.04999917, 0.0160054 , 0.97491239, 0.98615416, 0.01539949, 0.02880735, 0.99037224, 0.01546754, 0.97517279, 0.99082922, 0.01264795]
0.35	[0.0215768 , 0.96640368, 0.98611406, 0.01582844, 0.96639836, 0.04999414, 0.01593693, 0.9749325 , 0.98591963, 0.01554164, 0.02904883, 0.99045823, 0.01564929, 0.97507621, 0.99063514, 0.01264192]
0.40	[0.02132053, 0.96651086, 0.98741759, 0.01525951, 0.96650022, 0.04999084, 0.01538825, 0.97509064, 0.98732931, 0.01517433, 0.02844411, 0.99137031, 0.01530619, 0.9751241 , 0.99134853, 0.01218827]
0.45	[0.02808136, 0.99998344, 0.97301257, 0.02160198, 0.95670303, 0.02415148, 0.02222652, 0.99998125, 0.99999145, 0.01816547, 0.02403156, 0.95000357, 0.02403288, 0.9750836 , 0.99999281, 0.03312777]
0.50	[0.01041244, 0.9730789 , 0.9730527 , 0.02248735, 0.9948298 , 0.01987797, 0.01936303, 0.97244298, 0.97283164, 0.02191849, 0.02124108, 0.97533872, 0.01942012, 0.97302052, 0.97406162, 0.04999937]

3.2. With 0.9 Momentum

η	Actual Output
0.05	[0.01712228, 0.97025572, 0.97464474, 0.01174231, 0.97027232, 0.04999092, 0.01170487, 0.97844526, 0.97465832, 0.0116369 , 0.04468646, 0.98250813, 0.01158191, 0.97854983, 0.98275028, 0.00858147]
0.10	[0.01658111, 0.9703025 , 0.97539053, 0.01131537, 0.97031966, 0.04999197, 0.01131221, 0.97849425, 0.97539459, 0.01101296, 0.04400416, 0.98303337, 0.01097367, 0.97873359, 0.98355951, 0.00811355]
0.15	[0.01615897, 0.97033879, 0.97597296, 0.01088104, 0.97035362, 0.04998373, 0.01088344, 0.97853119, 0.97596866, 0.01054824, 0.04356744, 0.98353055, 0.01050354, 0.97884592, 0.98412035, 0.0076853]
0.20	[0.02141259, 0.96646722, 0.98696519, 0.01581435, 0.96647252, 0.04995809, 0.01579952, 0.97483694, 0.98644576, 0.01533403, 0.02871084, 0.99029001, 0.01528164, 0.97529518, 0.9911872 , 0.01244888]
0.25	[0.02127991, 0.96644394, 0.98836907, 0.01595879, 0.96645241, 0.04998322, 0.01599184, 0.97484437, 0.98707026, 0.01494536, 0.02788265, 0.99017701, 0.01492607, 0.97558847, 0.99170649, 0.01241088]
0.30	[0.02142226, 0.96642896, 0.98790415, 0.01595247, 0.96643436, 0.04999213, 0.01600744, 0.97482758, 0.98683481, 0.01510595, 0.02805724, 0.99017015, 0.015105 , 0.97548472, 0.99141466, 0.01247336]

0.35	[0.02153075, 0.96646158, 0.98730099, 0.01588109, 0.96646014, 0.04993851, 0.01595198, 0.9748297 , 0.98656988, 0.01529498, 0.02832312, 0.99025066, 0.01531172, 0.97537091, 0.99111577, 0.01251017]
0.40	[0.02145412, 0.96650181, 0.98745168, 0.01565463, 0.96649214, 0.04993218, 0.01576167, 0.97483436, 0.98691896, 0.01526202, 0.02829802, 0.99070279, 0.01531608, 0.97530879, 0.99123686, 0.01239522]
0.45	[0.01414889, 0.99886662, 0.99887061, 0.00653589, 0.99888106, 0.00680884, 0.00670415, 0.97584302, 0.9988977 , 0.00672821, 0.00663 , 0.97583334, 0.0069219 , 0.97579015, 0.97576061, 0.04996308]
0.50	[0.01479887, 0.99954969, 0.99955113, 0.00549542, 0.99955335, 0.0056724 , 0.00564815, 0.97614885, 0.99955988, 0.00564719, 0.00562737, 0.97613675, 0.0058237 , 0.97609116, 0.97605611, 0.04996142]