

HW5 Report

Haocong Wang
mw814

Q1

For input layer and hidden layer, I added one bias unit for each. The initial weights are given randomly, so sometimes the cost function J may not converge, which means the program won't give you the result. At this time, you need to stop the program and run it again.

Target error = 0.1

The learning rate is 0.05

The final weights are [[2.58859527 5.49133266 -5.14581325]

[2.78219305 -5.42066051 5.63337457]] and [[8.0261596 -5.54250516 -5.53944416]]

The final error is 0.09999235587719088

The program runs 9723 times to complete computing

The prediction of X is [[0.08741735]

[0.88984872]

[0.89243619]

[0.07502514]]

Target error = 0.1

The learning rate is 0.1

The final weights are [[2.58863493 5.49145784 -5.14591007]

[2.78226347 -5.42080998 5.63351576]] and [[8.02610897 -5.54247492 -5.53940115]]

The final error is 0.09999101513726133

The program runs 4862 times to complete computing

The prediction of X is [[0.08741594]

[0.88985023]

[0.89243709]

[0.07502412]]

Target error = 0.1

The learning rate is 0.25

The final weights are [[2.58914276 5.49256843 -5.14697443]

[2.78278329 -5.42194164 5.63464015]] and [[8.02801724 -5.54376668 -5.54063143]]

The final error is 0.09991763317060672

The program runs 1946 times to complete computing

The prediction of X is [[0.08735272]

[0.88992757]

[0.8925084]

[0.07497098]]

Target error = 0.1

The learning rate is 0.5

The final weights are [[2.58997443 5.49439949 -5.14872311]

[2.78365119 -5.42382582 5.63650843]] and [[8.03115032 -5.54588781 -5.54265056]]

The final error is 0.09979697780014846

The program runs 974 times to complete computing

The prediction of X is [[0.08724875]

[0.89005463]

[0.89262575]

[0.07488359]]

Target error = 0.1

The learning rate is 0.8

The final weights are [[2.58989768 5.49458736 -5.14869366]

[2.78387568 -5.42426415 5.63687025]] and [[8.02929327 -5.54466474 -5.54136677]]

The final error is 0.09984090615871015

The program runs 609 times to complete computing

The prediction of X is [[0.08728453]

[0.89000873]

[0.89258018]

[0.07491495]]

Target error = 0.1

The learning rate is 1

The final weights are [[2.59156895 5.49796923 -5.15210277]

[2.78539556 -5.42758811 5.64022075]] and [[8.03719813 -5.54998384 -5.54654412]]

The final error is 0.09956312480910731

The program runs 488 times to complete computing

The prediction of X is [[0.08704721]

[0.89030043]

[0.89285365]

[0.07471408]]

Target error = 0.08

The learning rate is 0.05

The final weights are [[-6.42066504 4.13888617 4.13856591]

[1.85414026 -5.29482775 -5.29382969]] and [[3.21579345 -6.49528276 -6.93726178]]

The final error is 0.07997797716591737

The program runs 8232 times to complete computing

The prediction of X is [[0.05771427]

[0.91672268]

[0.91672524]
[0.08292628]]

Target error = 0.08

The learning rate is 0.1

The final weights are [[-6.42066504 4.13888617 4.13856591]
[1.85414026 -5.29482775 -5.29382969]] and [[3.21579345 -6.49528276 -6.93726178]]
The final error is 0.07997797716591737
The program runs 8232 times to complete computing
The prediction of X is [[0.05771427]
[0.91672268]
[0.91672524]
[0.08292628]]

Target error = 0.08

The learning rate is 0.25

The final weights are [[-6.42066504 4.13888617 4.13856591]
[1.85414026 -5.29482775 -5.29382969]] and [[3.21579345 -6.49528276 -6.93726178]]
The final error is 0.07997797716591737
The program runs 8232 times to complete computing
The prediction of X is [[0.05771427]
[0.91672268]
[0.91672524]
[0.08292628]]

Target error = 0.08

The learning rate is 0.5

The final weights are [[-6.4278811 4.14370831 4.14341481]
[1.8561634 -5.2990211 -5.29811376]] and [[3.21674638 -6.49728991 -6.93947353]]
The final error is 0.07978461684558971
The program runs 825 times to complete computing
The prediction of X is [[0.05757689]
[0.9169104]
[0.91691262]
[0.08272587]]

Target error = 0.08

The learning rate is 0.8

The final weights are [[-6.43055109 4.14557832 4.14530203]
[1.85658856 -5.30063853 -5.29979157]] and [[3.21524727 -6.4939508 -6.93733363]]
The final error is 0.07981521883522533

The program runs 516 times to complete computing

The prediction of X is [[0.05757906]

[0.9168771]

[0.91687908]

[0.08276916]]

Target error = 0.08

The learning rate is 1

The final weights are [[-6.43234431 4.14684442 4.14657971]

[1.8568606 -5.30177343 -5.30096735]] and [[3.21413258 -6.49145591 -6.93574756]]

The final error is 0.07984127745562086

The program runs 413 times to complete computing

The prediction of X is [[0.05758229]

[0.91684916]

[0.91685096]

[0.08280554]]

We can tell from the numbers that learning rate = 1 is the best choice because it takes the least times to converge.

Q2

Designed as the requirement requires.