**HW5 Report**

Zhongze Tang (zt67)

**Q1 Neural Network**

I add a bias unit to the input layer and the hidden layer respectively. So, the network looks like:



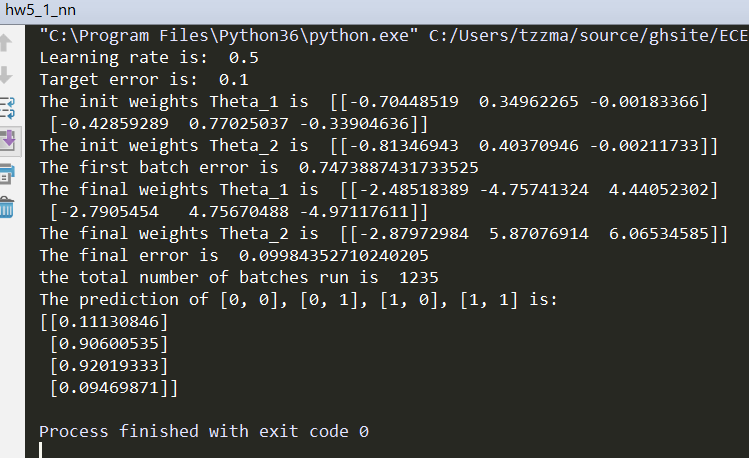
Note that sometimes the program won’t give you the result because the weights (theta1 and theta2) are initialized randomly, which may cause the cost function J won’t converge.

**You can change the learning\_rate and target\_error in Line 97 and 98.**

**Results**

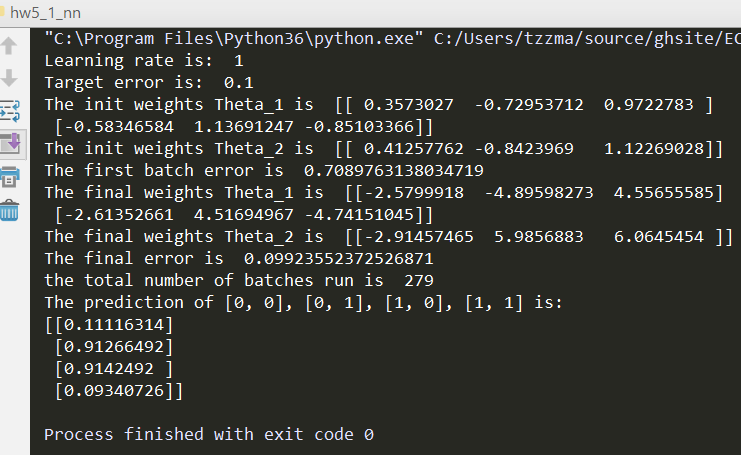
**Learning rate is: 0.5**

**Target error is: 0.1**



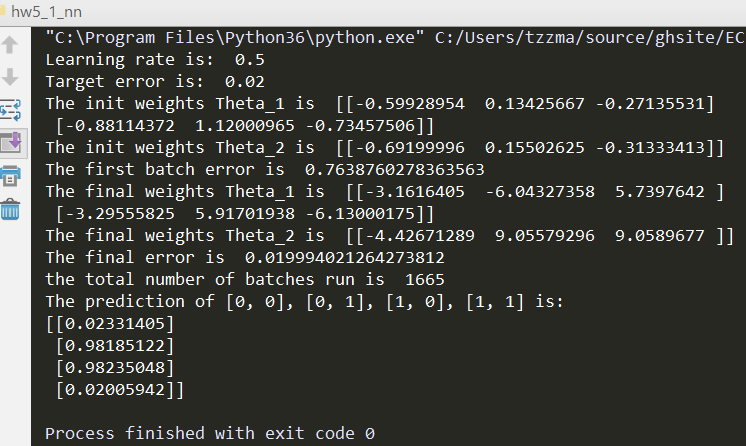
**Learning rate is: 1**

**Target error is: 0.1**



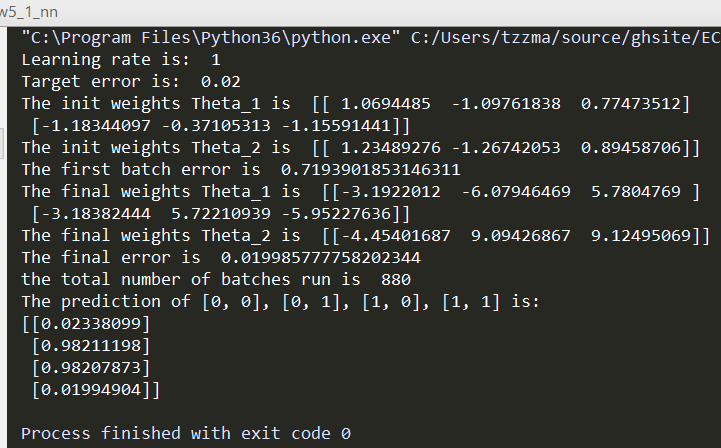
**Learning rate is: 0.5**

**Target error is: 0.02**



**Learning rate is: 1**

**Target error is: 0.02**



**Learning rate in [0.05, 0.1, 0.3, 0.5, 0.75, 1]**

**Target error = 0.1**

The init weights Theta\_1 is [[-0.38404047 -0.06897211 1.12252285]

[ 0.99525219 -0.78971388 0.35975224]]

The init weights Theta\_2 is [[-0.44345173 -0.07964214 1.29617774]]

**Learning rate is: 0.05**

**Target error is: 0.1**

The first batch error is 0.7174708075735095

The final weights Theta\_1 is [[-2.27973707 5.47942452 5.48520361]

[ 6.06182938 -4.05718349 -4.05844979]]

The final weights Theta\_2 is [[-8.85899284 6.16701676 6.06162017]]

The final error is 0.09998439355685633

**the total number of batches run is 12035**

**Learning rate is: 0.1**

**Target error is: 0.1**

The first batch error is 0.7174708075735095

The final weights Theta\_1 is [[-2.27984016 5.47958173 5.4853568 ]

[ 6.06207454 -4.05734265 -4.05860811]]

The final weights Theta\_2 is [[-8.85945217 6.16729371 6.06195859]]

The final error is 0.09996820523917618

**the total number of batches run is 6019**

**Learning rate is: 0.3**

**Target error is: 0.1**

The first batch error is 0.7174708075735095

The final weights Theta\_1 is [[-2.27994594 5.4797204 5.4854824 ]

[ 6.06225125 -4.05745214 -4.05871494]]

The final weights Theta\_2 is [[-8.85998035 6.16759386 6.06237466]]

The final error is 0.09995153130685704

**the total number of batches run is 2008**

**Learning rate is: 0.5**

**Target error is: 0.1**

The first batch error is 0.7174708075735095

The final weights Theta\_1 is [[-2.28032037 5.48028715 5.48603387]

[ 6.06313856 -4.05802736 -4.05928699]]

The final weights Theta\_2 is [[-8.86165711 6.16860239 6.0636132 ]]

The final error is 0.09989264894379402

**the total number of batches run is 1206**

**Learning rate is: 0.75**

**Target error is: 0.1**

The first batch error is 0.7174708075735095

The final weights Theta\_1 is [[-2.28074798 5.4809295 5.48665797]

[ 6.06414795 -4.05868079 -4.05993664]]

The final weights Theta\_2 is [[-8.86358181 6.16975724 6.06503851]]

The final error is 0.09982531382072798

**the total number of batches run is 805**

**Learning rate is: 1**

**Target error is: 0.1**

The first batch error is 0.7174708075735095

The final weights Theta\_1 is [[-2.27965468 5.47914069 5.4848656 ]

[ 6.06116554 -4.05671692 -4.05797233]]

The final weights Theta\_2 is [[-8.85901294 6.16690594 6.06181359]]

The final error is 0.09999633955642356

**the total number of batches run is 604**

**Target error = 0.02**

**Learning rate in [0.05, 0.1, 0.3, 0.5, 0.75, 1]**

The init weights Theta\_1 is [[ 0.76198743 -0.86487841 -0.12342132]

[ 0.21423064 1.09255548 -0.09787427]]

The init weights Theta\_2 is [[ 0.87986729 -0.99867556 -0.14251467]]

**Learning rate is: 0.05**

**Target error is: 0.02**

The first batch error is 0.7006059421176762

The final weights Theta\_1 is [[ 3.32277218 -6.6011102 6.86931383]

[ 3.35944904 6.94771762 -6.67466712]]

The final weights Theta\_2 is [[12.72540872 -8.67063999 -8.66164773]]

The final error is 0.019999581783178462

**the total number of batches run is 25519**

The prediction of [0, 0], [0, 1], [1, 0], [1, 1] is:

[[0.01773647]

[0.9770558 ]

[0.97700279]

[0.01550396]]

**Learning rate is: 0.1**

**Target error is: 0.02**

The first batch error is 0.7006059421176762

The final weights Theta\_1 is [[ 3.32280619 -6.60119189 6.86941465]

[ 3.35950467 6.94781707 -6.67477509]]

The final weights Theta\_2 is [[12.72544304 -8.6706606 -8.66166802]]

The final error is 0.019999016339520934

the total number of batches run is 12760

The prediction of [0, 0], [0, 1], [1, 0], [1, 1] is:

[[0.01773593]

[0.97705644]

[0.97700341]

[0.01550353]]

**Learning rate is: 0.3**

**Target error is: 0.02**

The first batch error is 0.7006059421176762

The final weights Theta\_1 is [[ 3.32294204 -6.60151839 6.86981776]

[ 3.35972755 6.9482154 -6.67520759]]

The final weights Theta\_2 is [[12.7255791 -8.67074224 -8.66174832]]

The final error is 0.019996761445964212

**the total number of batches run is 4254**

The prediction of [0, 0], [0, 1], [1, 0], [1, 1] is:

[[0.0177338 ]

[0.97705896]

[0.97700589]

[0.01550184]]

**Learning rate is: 0.5**

**Target error is: 0.02**

The first batch error is 0.7006059421176762

The final weights Theta\_1 is [[ 3.3231197 -6.60193097 6.870308 ]

[ 3.35999171 6.94869873 -6.67572486]]

The final weights Theta\_2 is [[12.7260305 -8.67103321 -8.66203819]]

The final error is 0.019992295063982615

**the total number of batches run is 2553**

The prediction of [0, 0], [0, 1], [1, 0], [1, 1] is:

[[0.01772969]

[0.97706399]

[0.97701088]

[0.01549845]]

**Learning rate is: 0.75**

**Target error is: 0.02**

The first batch error is 0.7006059421176762

The final weights Theta\_1 is [[ 3.32318306 -6.60212089 6.87059229]

[ 3.36017017 6.94898903 -6.67605914]]

The final weights Theta\_2 is [[12.72540025 -8.67060431 -8.66160618]]

The final error is 0.019995074671481136

**the total number of batches run is 1702**

The prediction of [0, 0], [0, 1], [1, 0], [1, 1] is:

[[0.017732 ]

[0.97706084]

[0.97700764]

[0.01550064]]

**Learning rate is: 1**

**Target error is: 0.02**

The first batch error is 0.7006059421176762

The final weights Theta\_1 is [[ 3.32345632 -6.60274244 6.87131279]

[ 3.36055335 6.94970157 -6.67681418]]

The final weights Theta\_2 is [[12.72635268 -8.67122584 -8.66222624]]

The final error is 0.01998676355932863

**the total number of batches run is 1277**

The prediction of [0, 0], [0, 1], [1, 0], [1, 1] is:

[[0.01772445]

[0.9770702 ]

[0.97701696]

[0.01549431]]

We can safely conclude that 1 is the best choice in [0.05, 0.1, 0.3, 0.5, 0.75, 1] for learning rate because the total number of running times of it is the least.

**Q2 A Web Page**

Just designed as the homework required, but more details. For example, the table will reset itself when the shape, radius or height changes, and the height input textbox will be disabled when the user chooses sphere.