```
----
The hidden size is 2000
The step size is 0.00001
The maxiter is 25
The weight matrix is for hidden layer is [[ 0.96722747 -1.22433896  0.92837993 ...  0.28196554 -0.24343172
   0.08098833]
 [ 0.92597504 -0.36148964 -1.16367107 ... -1.03614313 -0.12238883
   2.24190095]
 [-0.95641072 -1.13812116  0.70420113 ... -1.00995313  0.79908254
  -2.47072845]
 ...
 [ 1.3336342  -0.22215507 -0.08462302 ... -0.7132881  -1.7407416
   0.45392718]
 [-0.5350761   0.27802685  0.5141667  ... -0.68167876  0.91358341
   1.6600978 ]
 [ 0.72890096  0.69300872  0.40552612 ... -0.52206768 -0.94191885
  -0.06084582]]
The bias matrix is for hidden layer is [0.34883829 0.56240029 0.36718574 ... 0.16922276 0.82556765 0.25440644]
The weight matrix is for ouput layer is [[-0.04981012  2.31277972  0.8808994  ...  0.32658239 -0.34880361
  -2.34777162]
 [ 0.5066029   0.20728113  0.90828958 ... -1.15458808  2.01738823
   2.03776537]
 [ 0.62252787  0.20555664 -0.83919157 ...  0.6431902   1.51661607
   0.34322461]
 ...
 [ 2.20652551 -0.21216137 -1.3728159  ... -0.03887921 -0.22114047
  -1.98347368]
 [-0.14657094  1.42420755  0.09277716 ...  0.27264976 -2.0651929
  -0.0727649 ]
 [-0.59187702 -0.79732028 -1.50829844 ...  0.26251097 -0.15050089
  -0.71505529]]
The bias matrix is for output layer is [[0.22047224 0.50057682 0.59018688 ... 0.90237113 0.96801751 0.49814727]]
The final accurancy is 0.7239834
```

The highest accuracy I tested is around 0.72, but sometimes it will go to 0.65 based on the data is random. When I tested when hidden size is 50 to 60, it sometimes the accuracy is 0.6 to 0.7.

```
The hidden size is 1000
The step size is 0.0001
The maxiter is 25
The weight matrix is for hidden layer is [[ 0.27739685  0.10390565 -0.35721322 ...  0.32031867  1.26025951
  -0.29407252]
 [ 0.09856034 -0.54601218  1.115201   ...  0.6343659   0.49146731
  -2.08115354]
 [-0.82032914  0.8446175  -2.75084794 ... -1.82495471 -1.58738787
  -1.59507843]
 ...
 [-1.75445445 -0.13801959 -0.65483336 ... -0.36301914 -0.98069632
   0.88369772]
 [ 0.36169287  0.19225286 -0.17341783 ...  1.18187082  0.4885873
   0.63624824]
 [ 0.34471806 -1.73508547 -1.35403748 ... -1.3728102  -1.51862017
  -1.55094781]]
The bias matrix is for hidden layer is [ 6.14472293e-01  4.71973713e-01  4.06476110e-01  8.37660906e-01
  4.43275015e-01  7.08730098e-01  3.49670480e-04  6.40485831e-01

 ......

  5.08052019e-01  7.21542765e-01  9.82671119e-01  8.46040605e-01
  3.59872797e-01  1.05698127e-01  3.19500032e-01  7.96436278e-01
  7.22409484e-01  1.51903906e-01  6.77715346e-01  4.96629013e-01
  3.30847035e-01  8.90366988e-01  5.78160933e-02  6.31683461e-01]
The weight matrix is for ouput layer is [[-0.02192406  0.48467092 -1.53876386 ... -0.22261734  0.07926419
   0.64673952]
 [-0.14698899 -0.44712605  0.15037361 ...  0.21447097  1.28559658
   0.2448122 ]
 [ 0.07837878  0.32894695  1.05682709 ... -1.81671381 -0.99507434
  -0.16544016]
 ...
 [ 2.29130451 -0.8455055   1.04738308 ...  0.00929555 -0.7631708
  -0.9375364 ]
 [ 1.37367308  0.51711812  1.56369915 ... -0.43492859 -0.91648305
  -0.53224174]
 [ 0.43603341  1.78444907 -0.49186224 ... -0.14126679 -0.59759645
   0.36191054]]
The bias matrix is for output layer is [[0.16906669 0.61367378 0.11930732 ... 0.71142404 0.34997059 0.31016408]]
The final accurancy is 0.663733845
```