



Summary:

Back-propagation is a method to compute gradient of cost function

When we do back-propagation method, we compute each sample and put it in the for loop. Then summarize all of result with another equation, we have the partial derivative of cost function. We should use, the reshape code to implement vectorized.

We need to initialize random parameters. Because it can be learnt the same feature -> redundancy features

We should be check the back-propagation with the numerical partial derivative to make sure that our back-propagation is working correctly. Then stop Numerical method because it is computational expensive

Cost function in neural network is not a convex function. So when we apply gradient descent or advanced optimization, it can stuck at local minimum. But it is not a big problem, and we can accept the results.





