

CheckList undecided

1. Data preprocessing
   1. Zero center the data
   2. Subtract mean
   3. PCA or Whitening depending on situation
2. Decide which activation functions, loss functions and parameter update method to use. Usually Relu, Softmax and Adam Update. And decide NN structure (#of layer, layer size etc)
3. Implement neural network’s layers and structure with selected activation function and loss function.
   1. If use Batch Normalization, implement BN model to structure as well
4. Implement back propagation method
5. Implement parameter check method and remember to implement ratio of weight check with verbose option, and also implement running average of parameters method (smooth effect)
6. Implement learning rates decay
7. Implement regularization and inverted dropout
8. Initialize Weight:
   1. If it’s small network, simply use random initialization \* 0.01
   2. If it’s deep network and use non-linearity, use random initialization \* sqrt(2/n)
   3. If use Batch Normalization, no need to initialize weight?
9. Do a forward pass; sanity check whether loss is within desired range
10. Select few data points, turn off regularization and run gradient check
11. Select a small batch, and check whether this model can overfit this batch by running
12. Train the whole training data
13. Implement loss function & iterations diagram and validation accuracy vs training accuracy diagram
14. Hyper parameter Optimization
    1. Randomly initialize hyperparameter ranges from coarse to fine use validation set
    2. run to decide the top optimal hyperparameter sets
    3. Remember to ensemble the top performance models; if computation is too expensive, use different checkpoints of same model to form the ensemble
15. Try different initializations with same model and run step 1-14 again
    1. Run 14C
16. Use different activation function, loss function and parameter update methods and run step 1 – 15 again
    1. Run 14C
17. Use different layer size, number of layers and run step 1-16 again and memorize top optimal models
    1. Run 14C
18. See first layer visualization