

CMPE 442 Assignment #2

Due Friday, May 19, 23:59.

Note: The assignment is to be done individually. You are going to submit the report and code separately. Please do not include codes in the report. Your report should be self-contained, i.e. explain what you are doing, put the resulting graphs, the hyperparameters that you use, and write your observations.

Note: This assignment worth 10% of overall course grade.

In this assignment you are going to implement a neural network with one hidden layer. First, you will implement a NeuralNetwork class and then use that network to learn digits dataset.

- 1- **[60 pts]** Implement a NeuralNetwork class according to the template that is provided to you (NN_template.ipynb). In this template you will find a class NeuralNetwork with a constructor and 4 methods. Some of the methods are partially filled. Feel free to add you own code where you think appropriate.
 - a) **[30pts]** Fill the feedforward() method
 - b) **[30pts]** Fill the backprop() method. Ignore the regularization part in the algorithm given in slides (i.e. $\lambda=0$).
- 2- **[40 pts]** Train your network on digits dataset. It is a toy dataset in sklearn. The template shows the way of fetching the data and splitting it into three sets: train, validation, and test sets. **Note:** Do not touch test set until the last step of your assignment.

Start with default number of units in the hidden layer as 64 ($sl2=64$). **Fix the number of epochs to 100**, i.e. $iterNo=100$, throughout your assignment.

- a. **[15 pts]** Try different learning rates = {0.01, 0.1, 0.5, 0.9}, This means that you train four networks and for each network that you learn test it on validation set. Report on the learning rate that give you the best performance on validation set.
- b. **[15pts]** Try different values for $sl2 = \{16, 32, 64, 128\}$. The learning rate is fixed to the one that gave you the best result in (a). As in the previous case, train four networks and for each network that you learn test it on validation set. Report on the number of hidden units that give you the best performance on validation set.
- c. **[10pts]** Re-train the network combining **train set and validation set, using the best hyperparameters that you found in (a) and (b)**. Finally, test the model on the test set. Report on the accuracy and confusion matrix on test set.