

Report

- 1) In this assignment of training and optimizing the neural network, we have used MLP classifier from scikit learn (https://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPClassifier.html)
- 2) We have tried to optimize the code with different values of hyperparameters and their possible combinations. Please see below
 - a. activations = ['logistic', 'tanh', 'relu']
 - b. learning_rate = [0.01, 0.1]
 - c. max_iterations = [100, 200]
 - d. hidden layer number and size = (2,2),(3,3) (2,2,2),(3,3,3)
- 3) We have presented the plots for :
 - a) Number of Hidden layer = 2
 - b) Number of Hidden layer = 3For each of the hidden layer, loss curve has been presented to show the cost/error in training data with number of iterations.
- 4) We have also presented the mean accuracy and MSE of training and test data for each model in a tab separated file. A model represents hyperparameters values used.

Conclusions :

The mean accuracy for test data in case of logistic and Relu activations functions are comparable and better than tanh activation function. Changing the hidden layer number and size gives comparable results for models between number of hidden layer 2 and 3. We see the same behavior when changing the maximum iterations. In one of the case we see accuracy of training data at ~81 % for logistic activation function, number of hidden layer 3 and maximum iteration size as 200. It could be concluded that the choice of activation function (here logistic and Relu) and learning rate plays an important role in the optimization of model.