Chongye Wang
HW2 Report

Test Accuracy : 0.9763

Implementation:

The code implements a neural network using sigmoid function as activation function and includes one hidden layer. The code first implements one hot encoding used for multiclass classification. Then select the batch size and use the mini-batch gradient descent method. Define the forward function using sigmoid function as activation function and the cost function using cross entropy and the backward function to do backpropagation process.
The hyperparameters defined are as following:
number of hidden layers : 1
batch_size : 128
epochs : 20
learning rate : 1.0
units of hidden layer: 64
For each epoch, run through all the batches. For each batch, use the forward function to get the product and use the backward function to get the gradient. Also, a momentum is added for keeping averaging gradients. Update the weight and bias vector using w = w – learning_rate * gradient_w and b = b – learning_rate * gradient_b. For testing, compare the predictions to real labels and the final test accuracy is 0.9763.