Department of Computer Science and Engineering,
Delhi Technological University

AI509: Artificial Neural Networks
Assignment 1

Max Marks: 100 **Due Date**: 11:59PM Oct 2, 2024

Instructions:

- Keep collaborations among teams at high level discussions.
- Copying/Plagiarism will be dealt with strictly.

Submission Instructions: Submissions will be through Google classroom. Create a single folder ANN_A1_GroupXX.*zip* containing a report **ANN_A1_GroupXX.pdf** and the code files. The report must contain the google drive link of the folder where you upload the plots, output images, saved model etc. Report all your solutions, results and outputs in **ANN_A1_GroupXX.pdf**. (XX denotes your group number)
You are allowed to use *numpy*, *scipy* and *matplotlib* only, unless specified otherwise. In case of any doubt, initiate a discussion on Google classroom or drop an email to anurag@dtu.ac.in with subject line ANN Doubt Assignment 1.

Q1. (**Neural Networks)**
Implement the forward and back propagation algorithms to train an artificial neural network for MNIST dataset (attached) from scratch. You are not permitted to use any external libraries in this part.
a) (*60 points*) The arguments to your function would be the number of layers and the number of nodes in each layer. Assume a sigmoid activation function in each layer and softmax in the output layer. You have to construct the following neural network architectures from the above:

i) 1 hidden layer - [100 units]
ii) 3 hidden layers - [100, 50, 50 units]

Implement forward and backward propagation, which can be used for a general fully-connected neural network with softmax layer as the output. Use this function to train the two architectures and test them using the provided MNIST subset. Report accuracy as the evaluation metric in this part. Save the weights of your best models. Please ensure that your code is general and does not hard code equations.

b) (*20 points*) How would you diagnose if your model has overfitted or underfitted? Show plots/graphs and any other evidence to establish that your model is well-trained and can generalize well. Give proper justification with challenges you faced and counter-measures you took.

c) (*10 points*) Implement ReLU for the hidden layers and repeat part A and B using these activations at every layer (except output). You should still use softmax at the output.

d) (*10 points*) Implement part a) using the sklearn. Report the accuracy, as compared to what you received with your own network. Explain reasons for an observed difference in accuracies, if any.