

## **ASSIGNMENT 2**

### **Problem 1:**

a) MNIST is a handwritten Digit Classification dataset that contains 60,000 training and 10,000 test images. Using this dataset, build a three-layer feedforward neural network. Use 10K images from the training set as validation data. The two hidden layers have a dimension of 500 each. Using a suitable loss function, train the network for 250 epochs, and report the classification accuracy on test data. Do not use regularizations. Plot the error and classification accuracy on both training and validation data over the epochs. Justify the loss used to train the network. Also, normalize the data appropriately.

b) Repeat the above experiment, but train the network with the following regularizations:

- l2 regularization
- Dropout
- Early stopping

Compare with the results in the previous experiment and detail your observations after adding the regularizations. Note: No need to implement them on your own; the software framework (Tensorflow/Pytorch) typically provides implementations for l2 regularization and dropout. Early stopping is employed during training, so you need to tune your training code Accordingly.

### **Problem 2:**

a) Train a CNN network with ResNet-18 as a backbone from scratch with CIFAR-10 and note down the performance.

b) Initialize the ResNet-18 network with pre-trained weights from ImageNet and then try to use these weights to improve the training for the CIFAR-10 dataset. Try to come up with different ways of using these weights to improve the performance and play with the hyper-parameters to get the best performance. Document the results of your experiments.

(c) Train the network from scratch with the Tiny-CIFAR-10 dataset. Try using as many data augmentation techniques as you can think of to try to improve the performance. Try dropout after different layers and with different dropout rates. Document the results of your experiments.

d) Visualize the activations of the CNN for a few test examples in each of the above cases. How are the activation in the first few layers different from the later layers? Note: For the Tiny-CIFAR-10 dataset, take 500 images per class from CIFAR-10 for training. Use the same 10,000 images for testing as per the CIFAR-10 dataset. It is recommended to use Pytorch for this problem.

Suggested Frameworks: Tensorflow/Pytorch.

You can use GPUs available from IITD-HPC or Google Colab.

- Request for GPU on HPC:

[http://www.cc.iitd.ac.in/CSC/index.php?option=com\\_content&view=article&id=153:hpc-cluster&catid=1:latest-news](http://www.cc.iitd.ac.in/CSC/index.php?option=com_content&view=article&id=153:hpc-cluster&catid=1:latest-news)

- Setup HPC:

<http://supercomputing.iitd.ac.in/?access>

- How to submit jobs:

<http://supercomputing.iitd.ac.in/?pbs>

Do not submit jobs on login nodes. The hpc account will be blocked if you

Run your jobs on login nodes.

- You can use colab as well: <https://colab.research.google.com>

**NOTE:** Submissions will be done on moodle in the form of a pdf report and a zip containing source code(with compiled source) and all input files considered and output files generated. Only one of the group members will upload the assignment.

The deadline for this assignment is 11:59 pm @ 02 March, 2022.