**Spring 2023 Introduction to Deep Learning**

**Homework Assignment 4**

**Due date: April 25 2023**

**Problem (Build LeNet for colorful image classification).** In this problem, you are asked to train and test a neural network for ***entire*** CIFAR‐10 colorful image dataset. Some information of the network is as follows:

* Its structure is **modified** **LeNet**. You can check the 4th slide in Lecture 10 for details.
* An incomplete code has been given. You can fill it or re‐write all the codes by yourself.

**Performance Requirement and Submission:**

* The test accuracy should achieve above 50%
* You need to submit **three** results: 1) network without dropout/batch normalization, 2) network with one additional dropout layer and 3) network with one additional batch normalization. Compare the results in your submission.
* Submission should include your source codes and screen snapshot of your train and test accuracy, plus the training time

**Suggestion for hyperparameter setting (not necessary to follow):** Check the default setting in the code. You are allowed to change them

**About dataset loading:** Check the default setting in the code. You are allowed to change them

**Reminding:** You can check PyTorch *torch.nn* to find the packed Batch Normalization and Dropout layer if you would like to use.

**Model w/o dropout and batch normalization :-**

A screen shot of a computer program

Description automatically generated with low confidence

**Model w/o dropout and batch normalization Training:-**

A screen shot of a computer program

Description automatically generated with low confidence

A screen shot of a computer

Description automatically generated with low confidence

To compare and observe the difference a batch normalization and dropout layers make, we first run a base model. We see that with 5 epochs we achieve a accuracy of 54% in about 83 sec.

**Model w/ dropout :-**

A screen shot of a computer program

Description automatically generated with medium confidence

**Model w/ dropout Training :-**

A screen shot of a computer program

Description automatically generated with low confidence

A picture containing text, font, screenshot, black

Description automatically generated

A screen shot of a computer

Description automatically generated with low confidence

After adding one layer of dropout between the last two hidden layers, we observe that training takes longer than base model and we also see a drop in accuracy. A dropout layer turns an input to 0 at probability of p, allowing us to combat overfitting. However, on the other hand this means it takes longer for the model to train to almost the same amount of accuracy as the base model.

**Model w/ batch normalization :-**

A screen shot of a computer program

Description automatically generated with medium confidence

**Model w/ batch normalization Training:-**

A screen shot of a computer

Description automatically generated with medium confidenceA picture containing text, screenshot, font, black

Description automatically generated

When we add only one layer of batch normalization, we observe that we achieve higher accuracy in less time than the base model. Thus, batch normalization after activation of convolution layer can help reduce training time of the model, and even help with the accuracy of the model.