E0 294: Systems for Machine Learning, Jan'25
Indian Institute of Science, Bangalore
Assignment #3

Deadline: 6th March, 11:59 PM

(The numbers in the braces denote the points)

Please write the solution in latex. Put all your codes and the PDF into a zip file and submit. The submitted code should be properly commented and the name of the variables should be intuitive. The code should be implemented by yourself. Marks will be deducted due to codes without comments, with non-intuitive variable names, with high similarity.

Instruction: Problem 3 is compulsory. Solve (1(a), 1(b) and 1(c)) or (1(a) or 1(b), 2(a) and 2(b))

Problem 1: Data Parallelism. Consider LeNet-5 network proposed in https://ieeexplore.ieee.org/document/726791 . Consider 100 examples from MNIST dataset randomly. Divide the dataset into four parts (i.e. 25 examples in each part). Launch data parallel training on 4 CPU cores for 100 epochs.

(a) Use 'all reduce' to update the gradients after every epoch. **(5)**
(b) Use 'ring all reduce' to update the gradients after every epoch. **(5)**
(c) Clearly describe how did you perform above two exercises. Compare the accuracy of the trained model and the training time. **(15)**

Problem 2: Model parallelism. Consider the same LeNet-5 network as above. Segregate LeNet-5 into five pieces. Consider 100 examples from MNIST dataset and launch model parallel training for 100 epochs.

(a) Compare the convergence of loss and accuracy with respect to data parallel training (any of 1(a) or 1(b)). **(10)**
(b) Comment on the above conversion. Clearly describe your observation and potential reason behind your observations. **(10)**

Problem 3: Eyeriss proposes a row-stationary dataflow. Consider below input feature map and the kernel. Show the convolution output at each cycle considering row-stationary dataflow. Assume both multiplication and addition take one clock cycle. **(25)**

Input feature map

1	3	9
4	5	6
7	4	9

Kernel

-1	2
2	-1