CDA5125 Programming Assignment No. 4: Parallelizing Deep Neural Network Code with Message Passing Interface (MPI)

(Due: April 8)

Purpose:

- Practice MPI programming
- Practice parallelizing an application with MPI

Statement of work:

This is a group assignment. Each group can have 2 people. In this assignment, you will improve the deep neural network code that you developed in Assignment 2 by using MPI to explore parallel execution of the Deep neural network code across multiple processes (nodes). Your program should be able to utilize at least 16 processes to be considered correct.

Due dates:

The assignment is due on April 8, 11:59pm. Put all related source code, the makefile, and a README file in a tar file and submit the tar file. In the README file, you must describe (1) how to compile and run the program, (2) whether and how the parallelized program achieves high accuracy (parallelized code should still be correct), and (3) report the speedups over the sequential code. What is described in the README file must be repeatable with your submitted files.

Grading:

1. Submission has all components (all related source code, makefile, README file); the executable can be successfully produced with a 'make' command in the directory; a deep neural network for handwriting digit recognition with the MNIST dataset is built; MPI is used to parallelize the program (30 points).
2. The README file describes the information as required and the program is correct (5 points).
3. Points in 3), 4), and 5) can be obtained only after all points in 1) and 2) are obtained. One can follow the description in the README file to repeat the claims (5 points).
4. The MPI program running on multiple processes is at least 2.5 times faster than the single thread code. (10 points).

Hint: This program is much harder than any of the earlier assignments. Start now!