Spring 2019 Energy-efficient Machine Learning System

Final Project Description

Overview

The Course Project is an opportunity for you to practice what you have learned from this course and apply to some problems that you are interested in. There are three types of projects, Type-1, Type-A and Type-B.

- **Type-1 Project:** This type of project targets students who have no prior background on python or machine learning. The aim of this project is to provide the opportunity to work on some classical type of deep neural networks.
- **Type-A Project:** This type of project is about small-size neural network on small-size dataset. To finish this type of project, you do not need a powerful GPU as computing platform. The aim of this project is to help you have a deeper understanding the basic principle of deep neural network and how it works.
- **Type-B Project:** This type of project is about some advanced topics of neural network and its applications. This type of work might be challenging since it may require train and test large-size neural network on large-size dataset. You may need GPU for your experiment.

Topics of Type-1, Type-A and Type-B Projects

The following is some example topics you can use for your Type-1, Type-A and Type-B.

Available Type-1 Project Topics

- Use other types of optimization method, such as AdaGrad etc., to train and test a fully-connected neural network. The forward and backward procedure should be manually implemented.
- Use other types of optimization method, such as dropout and batch normalization etc., to train and test a fully-connected neural network. The forward and backward procedure should be manually implemented.

Available Type-A Project Topics

- Manually implement a LeNet-5 convolutional NN using Numpy (no automatic gradient). Train and test on MNIST dataset. The high test accuracy should be achieved.
- Manually implement a Recurrent NN using Numpy (no automatic gradient). Train and test on a speech or text dataset. The high test accuracy should be achieved.
- Prune a LeNet-5 convolutional NN on MNIST dataset with high compression ratio and negligible performance loss.
- Perform low-bit quantization of a LeNet-5 convolutional NN on MNIST dataset with high compression ratio and negligible performance loss.

Example Type-B Projects

- Implement a DNN model on mobile phones or IoT devices (e.g. Raspberry PI)
- Apply deep neural network to a domain problem: e.g. plate number recognition, image caption, blood cell recognition, digital communication, transportation etc.
- Implement an advanced DNN model compression approach, such as circulant transform, permutation-based pruning, binarization/tenarization
- Hardware architecture design of a CNN or RNN model using Verilog

Name Your Own Type-B Project

- You can propose a topic for your Type-B project. The complexity of your proposed Type-B project should be higher than Type-1/Type-A project.
- A one-page project proposal is needed if you would like to name your own Type-B project. The project proposal should include: the motivation of this topic, the planned used data, the state-of-the-art approaches and your planned selection, evaluation metrics.
- The project proposal should be submitted and approved by instructor.

Collaboration Policy

- Type-1 project should be done individually.
- Type-A/B project can be done in a team of 2 persons.

Honor Code

- You may refer to public and online resource when preparing your project as long as sufficient and clear citation and notation is presented.
- Incorporating other students' codes into your project is forbidden.

Submission Guideline

- The due date of final submission is **May 22nd.**
- The final submission includes final report and source codes.
- The final submission should be submitted in the ZIP/RAR format.
- Your final report should include abstract, introduction, related work, data description, method description, model description, experimental procedure and results, conclusion and reference.
- At least one Type-1, Type-A or Type-B project should be done and submitted. If multiple projects are submitted, the highest score among those submissions will be counted to your course score.