Tensorflow at MSI

March 2, 2018
Overview

- Submitting Batch Jobs
- Volume Data Example
- Language Processing Example
Submitting a Batch Job

- ssh username@login.msi.umn.edu
- ssh mesabi
- wget z.umn.edu/tensorflow1
- tar xvfz tensorflow1
- cd tutorial_3_2_2018_mnist
- qsub run_tensorflow_examples.pbs
Submitting a Batch Job

-1 nodes=1:ppn=24:gpus=2
-q k40
Submitting a Batch Job

module load caffe/1.0
(GPU-compiled Caffe)

module load caffe/1.0-cpu
(CPU-compiled Caffe)

module load tensorflow/1.4_gpu_python3
(GPU-compiled TensorFlow)
Volume Data Example

- 3D Volume data: voxels
- Illustrates how to:
  • Feed data from HPC resources to DL training
  • Manage & use 3D volume data
    - Python & TensorFlow arrays
    - 3D convolution layer
  • Ingest high res. volume data into DL training
Goal

- Given voxel data: union of tube interiors, identify
  - Surfaces: area in each voxel
  - Edges: arc length (where tubes intersect)
  - Corners: where edges meet.
- Train a neural net (NN) from voxel data generated by a app. driven by the training script.
Intersecting Tubes: Volume

Random tubes based on a pseudo random number generator and

input seed $\Rightarrow$ set of tubes

Voxel data of tube interiors fed as input to training
Intersecting Tubes: Surface

Surfaces (aqua) generated with volume data

Voxel data of tube surfaces fed as target for training

Loss is means square difference between target and NN model
Intersecting Tubes: Edges & Corners

Edges (red) & corners (yellow)

Also generated with volume data.

Voxel data of tube corners, or edges, fed as target for training.
Run Volume Example

- ssh username@login.msi.umn.edu
- ssh mesabi
- cp -r /home/dhp/public/deep_learning/tensorflow/vsec .
- cd vsec
- qsub run_vsec.pbs
Language Processing Example

Goal

- Read a single book
- Optimize word embeddings using Noise Contrastive Estimation (NCE)
- Identify words used in a similar context
- Identify analogous words
Language Processing Example

- ssh username@login.msi.umn.edu
- ssh mesabi
- wget z.umn.edu/tensorflow2
- tar xvfz tensorflow2
- cd tutorial_3_2_2018_text
- qsub text_example.pbs
Language Processing Example

- Embedding matrix is 8x5000 in this example
- Each word is represented by a vector of 8 values
- The Noise-Contrastive Estimation algorithm is used to calculate loss
- **Beware of over-fitting.** This example is optimizing 40k variables using a book containing only 122k words
Tensorboard

- Run Tensorboard on your local machine
- Mount your MSI home directory using `sshfs` or similar tool

```
tensorboard --logdir /Volumes/msi
```