Homework

Using the Graphcore testbed

In [ ]:
```
# Initial hyperparameters for the model

learning_rate = 0.03
epochs = 10
batch_size = 8
test_batch_size = 80
```

```
jrpenabastidas@jpb2288: ~
(poptorch33_env) (base) jrpenabastidas@gc-poplar-04:~/graphcore/examples/tutorials/simple_applications/pytorch/mnist$ /opt/slurm/bin/srun --ipus=1 p
ython mnist_poptorch.py
srun: job 27627 queued and waiting for resources
srun: job 27627 has been allocated resources
100%|██████████| 9912422/9912422 [00:00<00:00, 54576413.21it/s]
100%|██████████| 28881/28881 [00:00<00:00, 1562396.09it/s]
100%|██████████| 1648877/1648877 [00:00<00:00, 16187292.91it/s]
100%|██████████| 4542/4542 [00:00<00:00, 29581566.41it/s]
Epochs:   0%|          | 0/10 [00:00<?,[03:09:09.486] [poptorch:cpp] [warning] [DISPATCHER] Type coerced from Long to Int for tensor id 10
  0%|          | 0/150 [00:00<?, ?it/s]                2024-11-17T03:09:09.934872Z PL:POPLIN    3459706.3459706 W: poplin::preplanConvolution() is d
eprecated! Use poplin::preplan() instead
Graph compilation:   3%|         | 3/100 [00:00<00:03]2024-11-17T03:09:13.389114Z PL:POPLIN    3459706.3459706 W: poplin::preplanMatMuls() is depre
cated! Use poplin::preplan() instead
Graph compilation: 100%|██████████| 100/100 [00:22<00:00]2024-11-17T03:09:32.942505Z popart:session 3459706.3459706 W: Rng state buffer was not seri
alized.You did not load poplar Engine.Remember that if you would like to run the model using the model runtime then you have to create your own buff
er and callback in your model runtime application for rngStateTensor.

Epochs: 100%|██████████| 10/10 [01:51<00:00, 11.12s/it]
  0%|          | 0/125 [00:00<?, ?it/s]                2024-11-17T03:11:01.087179Z PL:POPLIN    3459706.3459706 W: poplin::preplanConvolution() is d
eprecated! Use poplin::preplan() instead0 [00:00<00:03]
                                                        2024-11-17T03:11:03.275614Z PL:POPLIN    3459706.3459706 W: poplin::preplanMatMuls() is depre
cated! Use poplin::preplan() instead6/100 [00:02<00:33]
Graph compilation: 100%|██████████| 100/100 [00:14<00:00]
 95%|███████Downloading http://yann.lecun.com/exdb/mnist/train-images-idx3-ubyte.gz
Failed to download (trying next):
HTTP Error 403: Forbidden

Downloading https://ossci-datasets.s3.amazonaws.com/mnist/train-images-idx3-ubyte.gz
Downloading https://ossci-datasets.s3.amazonaws.com/mnist/train-images-idx3-ubyte.gz to /home/jrpenabastidas/.torch/datasets/MNIST/raw/train-images-
idx3-ubyte.gz
Extracting /home/jrpenabastidas/.torch/datasets/MNIST/raw/train-images-idx3-ubyte.gz to /home/jrpenabastidas/.torch/datasets/MNIST/raw

Downloading http://yann.lecun.com/exdb/mnist/train-labels-idx1-ubyte.gz
Failed to download (trying next):
HTTP Error 403: Forbidden

Downloading https://ossci-datasets.s3.amazonaws.com/mnist/train-labels-idx1-ubyte.gz
Downloading https://ossci-datasets.s3.amazonaws.com/mnist/train-labels-idx1-ubyte.gz to /home/jrpenabastidas/.torch/datasets/MNIST/raw/train-labels-
idx1-ubyte.gz
Extracting /home/jrpenabastidas/.torch/datasets/MNIST/raw/train-labels-idx1-ubyte.gz to /home/jrpenabastidas/.torch/datasets/MNIST/raw

Downloading http://yann.lecun.com/exdb/mnist/t10k-images-idx3-ubyte.gz
Failed to download (trying next):
HTTP Error 403: Forbidden

Downloading https://ossci-datasets.s3.amazonaws.com/mnist/t10k-images-idx3-ubyte.gz
Downloading https://ossci-datasets.s3.amazonaws.com/mnist/t10k-images-idx3-ubyte.gz to /home/jrpenabastidas/.torch/datasets/MNIST/raw/t10k-images-id
x3-ubyte.gz
Extracting /home/jrpenabastidas/.torch/datasets/MNIST/raw/t10k-images-idx3-ubyte.gz to /home/jrpenabastidas/.torch/datasets/MNIST/raw
```

```
jrpenabastidas@jpb2288: ~
Downloading http://yann.lecun.com/exdb/mnist/train-labels-idx1-ubyte.gz
Failed to download (trying next):
HTTP Error 403: Forbidden

Downloading https://ossci-datasets.s3.amazonaws.com/mnist/train-labels-idx1-ubyte.gz
Downloading https://ossci-datasets.s3.amazonaws.com/mnist/train-labels-idx1-ubyte.gz to /home/jrpenabastidas/.torch/datasets/MNIST/raw/train-labels-
idx1-ubyte.gz
Extracting /home/jrpenabastidas/.torch/datasets/MNIST/raw/train-labels-idx1-ubyte.gz to /home/jrpenabastidas/.torch/datasets/MNIST/raw

Downloading http://yann.lecun.com/exdb/mnist/t10k-images-idx3-ubyte.gz
Failed to download (trying next):
HTTP Error 403: Forbidden

Downloading https://ossci-datasets.s3.amazonaws.com/mnist/t10k-images-idx3-ubyte.gz
Downloading https://ossci-datasets.s3.amazonaws.com/mnist/t10k-images-idx3-ubyte.gz to /home/jrpenabastidas/.torch/datasets/MNIST/raw/t10k-images-id
x3-ubyte.gz
Extracting /home/jrpenabastidas/.torch/datasets/MNIST/raw/t10k-images-idx3-ubyte.gz to /home/jrpenabastidas/.torch/datasets/MNIST/raw

Downloading http://yann.lecun.com/exdb/mnist/t10k-labels-idx1-ubyte.gz
Failed to download (trying next):
HTTP Error 403: Forbidden

Downloading https://ossci-datasets.s3.amazonaws.com/mnist/t10k-labels-idx1-ubyte.gz
Downloading https://ossci-datasets.s3.amazonaws.com/mnist/t10k-labels-idx1-ubyte.gz to /home/jrpenabastidas/.torch/datasets/MNIST/raw/t10k-labels-id
x1-ubyte.gz
Extracting /home/jrpenabastidas/.torch/datasets/MNIST/raw/t10k-labels-idx1-ubyte.gz to /home/jrpenabastidas/.torch/datasets/MNIST/raw

TrainingModelWithLoss(
  (model): Network(
    (layer1): Block(
      (conv): Conv2d(1, 32, kernel_size=(3, 3), stride=(1, 1))
      (pool): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
      (relu): ReLU()
    )
    (layer2): Block(
      (conv): Conv2d(32, 64, kernel_size=(3, 3), stride=(1, 1))
      (pool): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
      (relu): ReLU()
    )
    (layer3): Linear(in_features=1600, out_features=128, bias=True)
    (layer3_act): ReLU()
    (layer3_dropout): Dropout(p=0.5, inplace=False)
    (layer4): Linear(in_features=128, out_features=10, bias=True)
    (softmax): Softmax(dim=1)
  )
  (loss): CrossEntropyLoss()
)
Accuracy on test set: 98.39%
(poptorch33_env) (base) jrpenabastidas@gc-poplar-04:~/graphcore/examples/tutorials/simple_applications/pytorch/mnist$
```

This is the baseline output for comparison. It took longer time to run as some downloads are required. The accuracy is 98.39% with this set of parameters.

First modification

In [ ]:
```
# First modification to the hyperparameters

learning_rate = 0.1
epochs = 15
batch_size = 80
test_batch_size = 80
```

```
(poptorch33_env) (base) jrpenabastidas@gc-poplar-04:~/graphcore/examples/tutorials/simple_applications/pytorch/mnist$ vim mnist_poptorch.py
(poptorch33_env) (base) jrpenabastidas@gc-poplar-04:~/graphcore/examples/tutorials/simple_applications/pytorch/mnist$ /opt/slurm/bin/srun --ipus=1 python mnist_poptorch.py
srun: job 27709 queued and waiting for resources
srun: job 27709 has been allocated resources
Epochs:   0%|          | 0/10 [00:00<?,[17:22:27.262] [poptorch:cpp] [warning] [DISPATCHER] Type coerced from Long to Int for tensor id 10
Graph compilation: 100%|██████████| 100/100 [00:21<00:00]
Epochs: 100%|██████████| 10/10 [01:54<00:00, 11.42s/it]
Graph compilation: 100%|██████████| 100/100 [00:14<00:00]
 83%|████████  | 104/125 [00:17<TrainingModelWithLoss(0]
  (model): Network(
    (layer1): Block(
      (conv): Conv2d(1, 32, kernel_size=(3, 3), stride=(1, 1))
      (pool): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
      (relu): ReLU()
    )
    (layer2): Block(
      (conv): Conv2d(32, 64, kernel_size=(3, 3), stride=(1, 1))
      (pool): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
      (relu): ReLU()
    )
    (layer3): Linear(in_features=1600, out_features=128, bias=True)
    (layer3_act): ReLU()
    (layer3_dropout): Dropout(p=0.5, inplace=False)
    (layer4): Linear(in_features=128, out_features=10, bias=True)
    (softmax): Softmax(dim=1)
  )
  (loss): CrossEntropyLoss()
)
Accuracy on test set: 97.65%
```

With bigger learning rate the accuracy is reduced to 97.65% even with a higher number of epochs

Second modification

In [ ]:
```
# Second modification to the hyperparameters

learning_rate = 0.001
epochs = 20
batch_size = 8
test_batch_size = 80
```

```
(poptorch33_env) (base) jrpenabastidas@gc-poplar-04:~/graphcore/examples/tutorials/simple_applications/pytorch/mnist$ vim mnist_poptorch.py
(poptorch33_env) (base) jrpenabastidas@gc-poplar-04:~/graphcore/examples/tutorials/simple_applications/pytorch/mnist$ /opt/slurm/bin/srun --ipus=1 python mnist_poptorch.py
srun: job 27710 queued and waiting for resources
srun: job 27710 has been allocated resources
Epochs:   0%|          | 0/10 [00:00<?,[17:28:44.982] [poptorch:cpp] [warning] [DISPATCHER] Type coerced from Long to Int for tensor id 10
Graph compilation: 100%|██████████| 100/100 [00:22<00:00]
Epochs: 100%|██████████| 10/10 [01:50<00:00, 11.02s/it]
Graph compilation: 100%|██████████| 100/100 [00:14<00:00]
 94%|████████▍TrainingModelWithLoss(1, 14.19it/s]<00:00]
  (model): Network(
    (layer1): Block(
      (conv): Conv2d(1, 32, kernel_size=(3, 3), stride=(1, 1))
      (pool): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
      (relu): ReLU()
    )
    (layer2): Block(
      (conv): Conv2d(32, 64, kernel_size=(3, 3), stride=(1, 1))
      (pool): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
      (relu): ReLU()
    )
    (layer3): Linear(in_features=1600, out_features=128, bias=True)
    (layer3_act): ReLU()
    (layer3_dropout): Dropout(p=0.5, inplace=False)
    (layer4): Linear(in_features=128, out_features=10, bias=True)
    (softmax): Softmax(dim=1)
  )
  (loss): CrossEntropyLoss()
)
Accuracy on test set: 98.49%
```

With a smaller learning rate the accuracy returns to values above 98 %. Using lower number of epochs

Third modification

In [ ]:
```
# Third modification to the hyperparameters

learning_rate = 0.0001
epochs = 20
batch_size = 16
test_batch_size = 80
```

```
(poptorch33_env) (base) jrpenabastidas@gc-poplar-04:~/graphcore/examples/tutorials/simple_applications/pytorch/mnist$ vim mnist_poptorch.py
(poptorch33_env) (base) jrpenabastidas@gc-poplar-04:~/graphcore/examples/tutorials/simple_applications/pytorch/mnist$ /opt/slurm/bin/srun --ipus=1 python mnist_poptorch.py
srun: job 27712 queued and waiting for resources
srun: job 27712 has been allocated resources
Epochs:   0%|          | 0/10 [00:00<?,[17:35:58.243] [poptorch:cpp] [warning] [DISPATCHER] Type coerced from Long to Int for tensor id 10
Graph compilation: 100%|██████████| 100/100 [00:20<00:00]
Epochs: 100%|██████████| 10/10 [01:51<00:00, 11.14s/it]
Graph compilation: 100%|██████████| 100/100 [00:14<00:00]
 85%|████████  | 106/125 [00:17<00:0TrainingModelWithLoss(
  (model): Network(
    (layer1): Block(
      (conv): Conv2d(1, 32, kernel_size=(3, 3), stride=(1, 1))
      (pool): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
      (relu): ReLU()
    )
    (layer2): Block(
      (conv): Conv2d(32, 64, kernel_size=(3, 3), stride=(1, 1))
      (pool): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
      (relu): ReLU()
    )
    (layer3): Linear(in_features=1600, out_features=128, bias=True)
    (layer3_act): ReLU()
    (layer3_dropout): Dropout(p=0.5, inplace=False)
    (layer4): Linear(in_features=128, out_features=10, bias=True)
    (softmax): Softmax(dim=1)
  )
  (loss): CrossEntropyLoss()
)
Accuracy on test set: 98.52%
```

In this case the learning rate is the smallest and a high number of epoch is used, the accuracy is high but not sustancially different from simulation less intensive in parameters.