HPML Lab 2

ar7996

HPC Configuration:

```
#CPU
#SBATCH --nodes=1
#SBATCH --cpus-per-task=16
#SBATCH --mem=6GB

#GPU
#SBATCH --gres=gpu:rtx8000:1
```

C1: Code is present in python files.

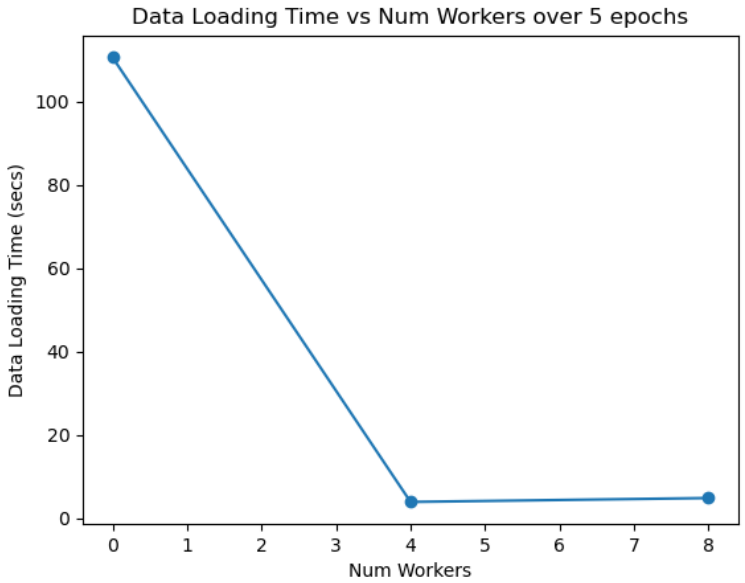
C2: Time Measurement of code in C1

```
C2: Data loading time, Training Time, Running Time
##################################################

Starting Training:
Optimizer: sgd, num_workers: 2, Device: cpu

Epoch: 1/5 Training loss: 1.9029017772211139  Training acc:
0.30729299875171595
Data Loading Time: 0.5640538521111012 secs  Training time:
217.4487524665892 secs  Total Running Time: 218.18879755958915
secs

Epoch: 2/5 Training loss: 1.4780132419922773  Training acc:
0.4559782608543211
Data Loading Time: 0.609431117773056 secs  Training time:
215.38720513880253 secs  Total Running Time: 216.21747229248285
secs

Epoch: 3/5 Training loss: 1.2495761716457279  Training acc:
0.5485693733862904
Data Loading Time: 0.6301647052168846 secs  Training time:
209.70884777978063 secs  Total Running Time: 210.55391976609826
secs

Epoch: 4/5 Training loss: 1.0480180138822102  Training acc:
0.6254076087261404
Data Loading Time: 0.6413500532507896 secs  Training time:
209.04028104245663 secs  Total Running Time: 209.92135556414723
secs
```

```
Epoch: 5/5 Training loss: 0.8970319901585884  Training acc:
0.683843510230179
Data Loading Time: 0.5778227187693119 secs  Training time:
209.88264348357916 secs  Total Running Time: 210.70187001302838
secs
```

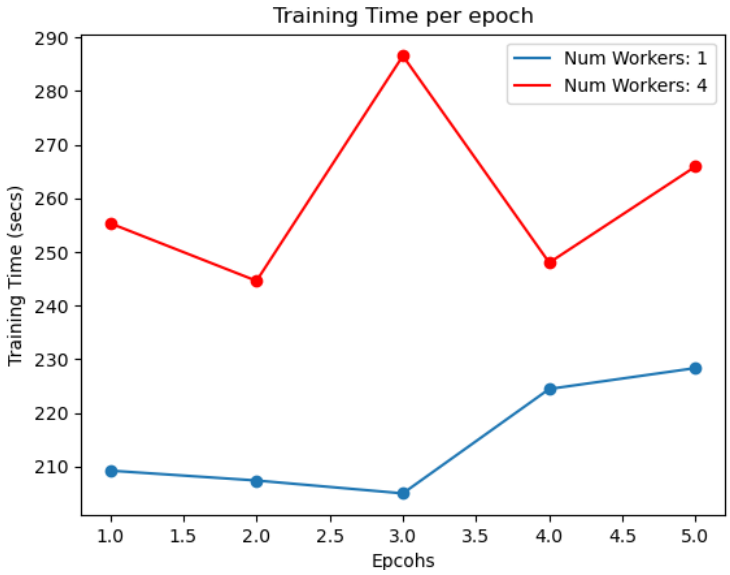
C3: I/O optimization for code in C2

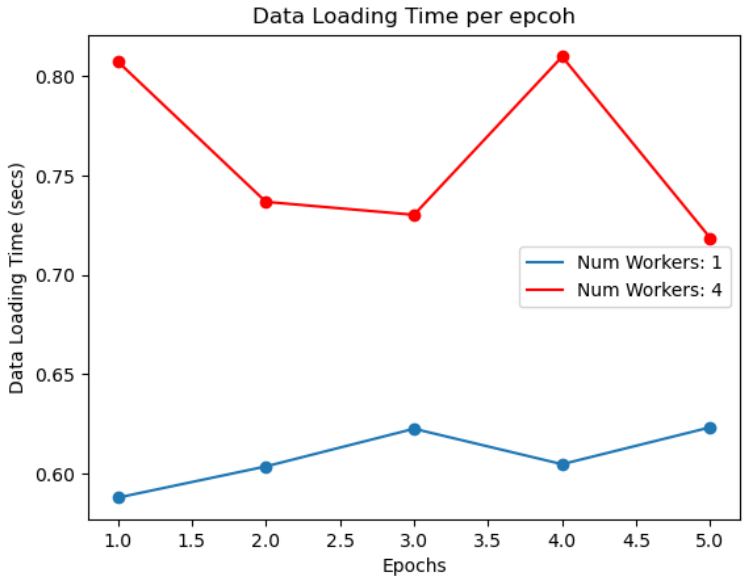
```
C3: I/O Optimization
####################

Starting Training:
Optimizer: sgd, num_workers: 0, Device: cpu

Total Data Time:  110.57924083247781

Starting Training:
Optimizer: sgd, num_workers: 4, Device: cpu

Total Data Time:  3.8983422219753265

Starting Training:
Optimizer: sgd, num_workers: 8, Device: cpu

Total Data Time:  4.829236317425966

Optimal Num Workers:  4
```



C4: Profiling starting from code in C3

```
C4: Profiling with Optimal Number of Workers
##############################################

Starting Training:
Optimizer: sgd, num_workers: 1, Device: cpu

Total Data Loading time using 1 workers: 3.041916858404875
Total Computing time using 1 workers: 1074.4139270149171

Starting Training:
Optimizer: sgd, num_workers: 4, Device: cpu

Total Data Loading time using 4 workers: 3.8029788620769978
Total Computing time using 4 workers: 1300.5269144438207
```





We can see that time required for training and data loading in case of num_workers 1 is less than time required for the optimal num_workers found in C4. This is because, while finding the optimal number we incremented the num_workers value by 4 at a time.

C5: Training in GPUs vs CPUs

```
C5: GPU vs CPU Training
######################

Starting Training:
Optimizer: sgd, num_workers: 4, Device: cpu

Epoch: 1/5 Training loss: 1.9826291281244028  Training acc:
0.299292679028133
Data Loading Time: 0.8232308253645897 secs  Training time:
182.13876163586974 secs  Total Running Time: 183.11436820030212
secs

Epoch: 2/5 Training loss: 1.4747350959826613  Training acc:
0.4564418158567775
Data Loading Time: 0.6421146802604198 secs  Training time:
187.57486278563738 secs  Total Running Time: 188.41183660924435
secs

Epoch: 3/5 Training loss: 1.2001903797964306  Training acc:
0.5668038683169333
Data Loading Time: 0.6230802275240421 secs  Training time:
186.90559655427933 secs  Total Running Time: 187.7243613190949
secs

Epoch: 4/5 Training loss: 0.9941737072547073  Training acc:
0.6475823209108904
Data Loading Time: 0.6245329193770885 secs  Training time:
186.71505769342184 secs  Total Running Time: 187.527556784451
secs

Epoch: 5/5 Training loss: 0.8555821322114266  Training acc:
0.6988610933198953
Data Loading Time: 0.6495090983808041 secs  Training time:
186.42919075489044 secs  Total Running Time: 187.27554327249527
secs

Total Running Time over 5 epochs: 934.0536661855876
Avg Running Time per epoch: 186.81073323711752

Starting Training:
Optimizer: sgd, num_workers: 4, Device: cuda
```

```
Epoch: 1/5 Training loss: 1.8515554822009543  Training acc:
0.3232416880252721
Data Loading Time: 0.6135864965617657 secs  Training time:
21.587373908609152 secs  Total Running Time: 22.44868417829275
secs

Epoch: 2/5 Training loss: 1.368476944811204  Training acc:
0.5003436701681913
Data Loading Time: 0.6163769997656345 secs  Training time:
16.924429159611464 secs  Total Running Time: 17.790607560425997
secs

Epoch: 3/5 Training loss: 1.1090849912380014  Training acc:
0.6047754156619997
Data Loading Time: 0.6803603917360306 secs  Training time:
16.95467720553279 secs  Total Running Time: 17.896385606378317
secs

Epoch: 4/5 Training loss: 0.9413983469728924  Training acc:
0.6647378517233807
Data Loading Time: 0.7282652333378792 secs  Training time:
16.960697140544653 secs  Total Running Time: 17.95354050025344
secs

Epoch: 5/5 Training loss: 0.8369476351591633  Training acc:
0.7043158567470053
Data Loading Time: 0.7205252423882484 secs  Training time:
16.96277939900756 secs  Total Running Time: 17.950579538941383
secs

Total Running Time over 5 epochs: 94.03979738429189
Avg Running Time per epoch: 18.80795947685838
```

C6: Experimenting with different optimizers

```
C6:Various Optimizers
######################

Starting Training:
Optimizer: sgd, num_workers: 4, Device: cuda

Epoch: 1/5 Training loss: 1.8556353149511624  Training acc:
0.320372442455243
```

Data Loading Time: 0.7119207791984081 secs Training time: 16.94483019784093 secs Total Running Time: 17.920144248753786 secs

Epoch: 2/5 Training loss: 1.4158140555062257 Training acc: 0.4768142582815322
Data Loading Time: 0.7270686365664005 secs Training time: 16.954905539751053 secs Total Running Time: 17.94797659665346 secs

Epoch: 3/5 Training loss: 1.139264629167669 Training acc: 0.587923593380872
Data Loading Time: 0.7229722328484058 secs Training time: 16.97421320900321 secs Total Running Time: 17.96436085551977 secs

Epoch: 4/5 Training loss: 0.9657547129389575 Training acc: 0.6560701726342711
Data Loading Time: 0.715088427066803 secs Training time: 16.955111283808947 secs Total Running Time: 17.938485383987427 secs

Epoch: 5/5 Training loss: 0.8252268425948784 Training acc: 0.7108335998052221
Data Loading Time: 0.6940878219902515 secs Training time: 16.99344302341342 secs Total Running Time: 17.9539975784719 secs

Starting Training:
Optimizer: adam, num_workers: 4, Device: cuda

Epoch: 1/5 Training loss: 2.288795213565192 Training acc: 0.22038043479023078
Data Loading Time: 0.7009884119033813 secs Training time: 17.54200639948249 secs Total Running Time: 18.5121387578547 secs

Epoch: 2/5 Training loss: 1.8636413215066465 Training acc: 0.277225863217088
Data Loading Time: 0.723708126693964 secs Training time: 17.35221966728568 secs Total Running Time: 18.341930128633976 secs

Epoch: 3/5 Training loss: 1.8376613852313108 Training acc: 0.2841871803069054

Data Loading Time: 0.7214375101029873 secs Training time: 17.354987543076277 secs Total Running Time: 18.345352310687304 secs

Epoch: 4/5 Training loss: 1.8287574958313457 Training acc: 0.2886069373553976
Data Loading Time: 0.7379802577197552 secs Training time: 17.357664607465267 secs Total Running Time: 18.36416593566537 secs

Epoch: 5/5 Training loss: 1.8152782621286105 Training acc: 0.293901854235193
Data Loading Time: 0.713894572108984 secs Training time: 17.35929213836789 secs Total Running Time: 18.340254120528698 secs

Starting Training:
Optimizer: sgdnestrov, num_workers: 4, Device: cuda

Epoch: 1/5 Training loss: 1.807479586137835 Training acc: 0.3540121483833284
Data Loading Time: 0.7141064964234829 secs Training time: 17.104396399110556 secs Total Running Time: 18.08597193285823 secs

Epoch: 2/5 Training loss: 1.3412869516236092 Training acc: 0.5113091432225064
Data Loading Time: 0.7065048553049564 secs Training time: 17.101064041256905 secs Total Running Time: 18.070765525102615 secs

Epoch: 3/5 Training loss: 1.0614891607133323 Training acc: 0.61967311384123
Data Loading Time: 0.7423919625580311 secs Training time: 17.12551746889949 secs Total Running Time: 18.134448803961277 secs

Epoch: 4/5 Training loss: 0.8999776873747101 Training acc: 0.6811421036415393
Data Loading Time: 0.7124726250767708 secs Training time: 17.09675594046712 secs Total Running Time: 18.07400380447507 secs

Epoch: 5/5 Training loss: 0.7803458518079479 Training acc: 0.7256833439897699

Data Loading Time: 0.7181510291993618 secs Training time: 17.1217682659626 secs Total Running Time: 18.104899939149618 secs

Starting Training:
Optimizer: adagrad, num_workers: 4, Device: cuda

Epoch: 1/5 Training loss: 2.1875609321057645 Training acc: 0.252973145764807
Data Loading Time: 0.6315088234841824 secs Training time: 17.144987165927887 secs Total Running Time: 18.031643632799387 secs

Epoch: 2/5 Training loss: 1.6726551824213598 Training acc: 0.36953324811232974
Data Loading Time: 0.6188275068998337 secs Training time: 17.146590817719698 secs Total Running Time: 18.019006945192814 secs

Epoch: 3/5 Training loss: 1.4233775269954712 Training acc: 0.4725743286749896
Data Loading Time: 0.6320114284753799 secs Training time: 17.146269883960485 secs Total Running Time: 18.03009717538953 secs

Epoch: 4/5 Training loss: 1.1808249462595986 Training acc: 0.568710038424148
Data Loading Time: 0.6331517659127712 secs Training time: 17.139227371662855 secs Total Running Time: 18.025778714567423 secs

Epoch: 5/5 Training loss: 1.0226440159865962 Training acc: 0.6306905371453756
Data Loading Time: 0.6338972002267838 secs Training time: 17.14873433113098 secs Total Running Time: 18.038429740816355 secs

Starting Training:
Optimizer: adadelta, num_workers: 4, Device: cuda

Epoch: 1/5 Training loss: 1.3830789342865615 Training acc: 0.49418957809658004

```
Data Loading Time: 0.6545203290879726 secs  Training time:
17.725560072809458 secs  Total Running Time: 18.6364688500762
secs

Epoch: 2/5 Training loss: 0.8846788744792304  Training acc:
0.6856657609610302
Data Loading Time: 0.5973335318267345 secs  Training time:
17.719937559217215 secs  Total Running Time: 18.563861068338156
secs

Epoch: 3/5 Training loss: 0.6847470732753539  Training acc:
0.7595748082756082
Data Loading Time: 0.6099692918360233 secs  Training time:
17.721197459846735 secs  Total Running Time: 18.5823675096035
secs

Epoch: 4/5 Training loss: 0.5749494282485884  Training acc:
0.7990768862807233
Data Loading Time: 0.6320807747542858 secs  Training time:
17.724840637296438 secs  Total Running Time: 18.60879333689809
secs

Epoch: 5/5 Training loss: 0.5039319515685596  Training acc:
0.8240808823224529
Data Loading Time: 0.6293251998722553 secs  Training time:
17.72560801357031 secs  Total Running Time: 18.607658721506596
secs
```

C7: Experimenting without Batch Norm

```
C7: Without Batch-Normalization
##############################

Starting Training:
Optimizer: sgd, num_workers: 4, Device: cuda

Epoch: 1/5 Training loss: 1.928061533462056  Training acc:
0.27256633634762384

Epoch: 2/5 Training loss: 1.5251085980773886  Training acc:
0.4359734655188782

Epoch: 3/5 Training loss: 1.2849336042428565  Training acc:
0.5408847507308511

Epoch: 4/5 Training loss: 1.10516936638776  Training acc:
0.6119125639386189
```

```
Epoch: 5/5 Training loss: 0.9554623642845836  Training acc:
0.6649936061076192
```

Q1:

Total Convolutional layer in Resnet-18 model (as per model definition in assignment) is **20**.

Detailed Composition:

- **Starting Layer:** 1
- **Residual Layer 1:** 4
- **Residual Layer 2:** 5
- **Residual Layer 3:** 5
- **Residual Layer 4:** 5

Model Architecture is present in resnet18.txt file.

Q2:

Input dimension of last linear layer is **512.**

Q3:

Total trainable parameters in ResNet-18: 11173962
Total gradients in ResNet-18: 11173962

Code is present in q3_lab2.py file.

Q4:

Total trainable parameters in ResNet-18: 11173962
Total gradients in ResNet-18: 11173962