Date: 2023.12.06

The pytorch template may update in the future, but the code in this example is not affected by that.

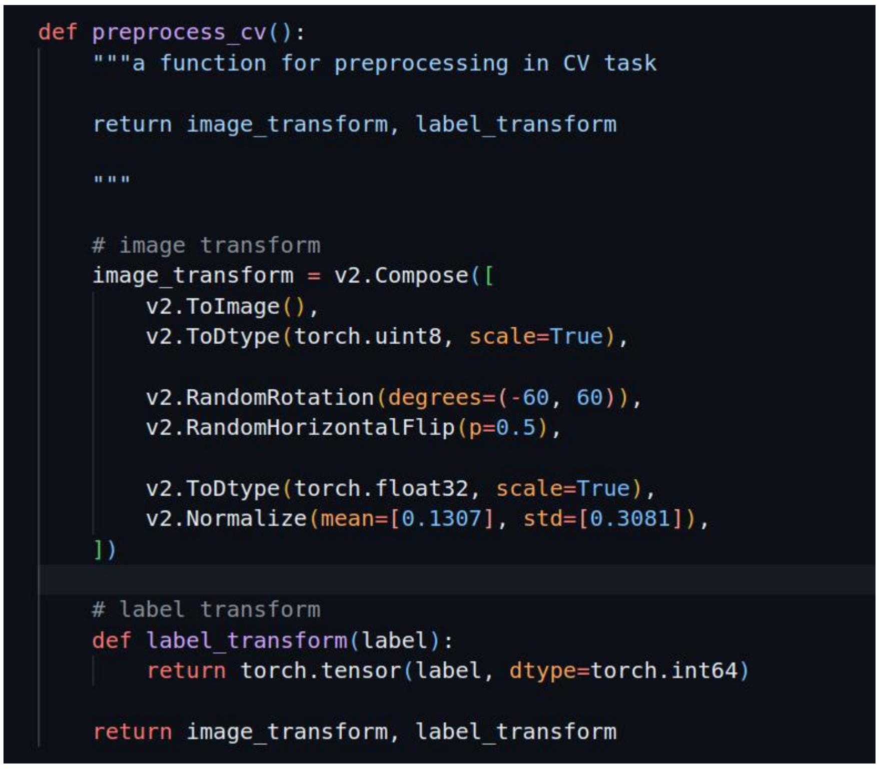
MNIST

This is the record of how to tweak the pytorch template for this project, as well as what the training procedure looks like.

1. tweak code

remove the code not needed, e.g., the code for nlp, and the code about test dataset, as there's only train and valid data

tweak **./preprocess.py**, tweak image transform and implement label transform



```python
def preprocess_cv():
    """a function for preprocessing in CV task

    return image_transform, label_transform

    """

    # image transform
    image_transform = v2.Compose([
        v2.ToImage(),
        v2.ToDtype(torch.uint8, scale=True),

        v2.RandomRotation(degrees=(-60, 60)),
        v2.RandomHorizontalFlip(p=0.5),

        v2.ToDtype(torch.float32, scale=True),
        v2.Normalize(mean=[0.1307], std=[0.3081]),
    ])

    # label transform
    def label_transform(label):
        return torch.tensor(label, dtype=torch.int64)

    return image_transform, label_transform
```

tweak **./main.py**, add code to load data, MNIST dataset will be loaded from Huggingface

```python
# load dataset from huggingface
cache_dir = "./.huggingface"
dataset_path = "mnist"
mnist_dataset = load_dataset(path=dataset_path, cache_dir=cache_dir)

train_data = mnist_dataset['train']
valid_data = mnist_dataset['test']
```

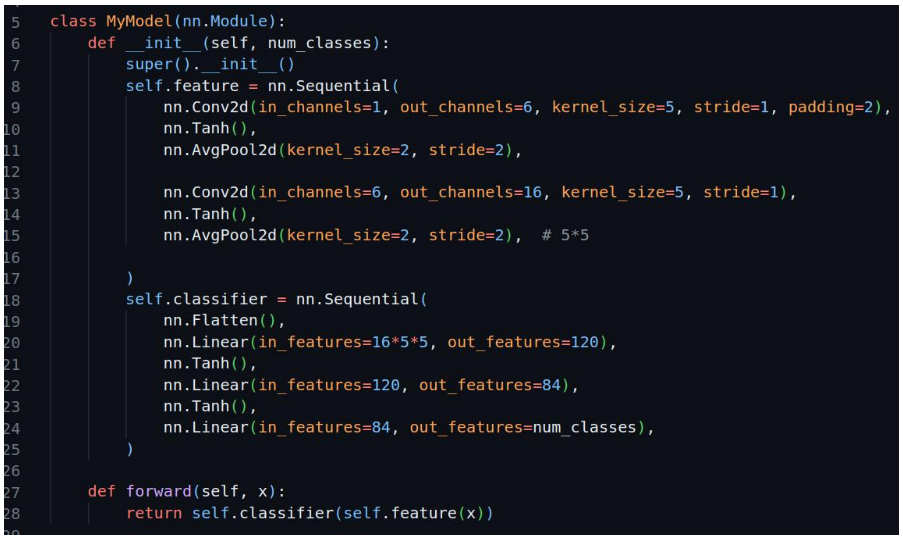
tweak **./dataset.py**, according to the MNIST data format

```python
from torch.utils.data import Dataset
import os


class ImageDataset(Dataset):
    def __init__(self, data, image_transform=None, label_transform=None):
        self.images = data['image']
        self.labels = data['label']
        self.image_transform = image_transform
        self.label_transform = label_transform

    def __len__(self):
        return len(self.labels)

    def __getitem__(self, idx):
        image = self.images[idx]
        label = self.labels[idx]
        if self.image_transform:
            image = self.image_transform(image)
        if self.label_transform:
            label = self.label_transform(label)
        return image, label
```

tweak **./main.py**, as the ImageDataset changed, tweak the part about dataset

```python
# create datasets
train_dataset = ImageDataset(
    data=train_data,
    image_transform=image_transform,
    label_transform=label_transform,
)
valid_dataset = ImageDataset(
    data=valid_data,
    image_transform=image_transform,
    label_transform=label_transform,
)
```

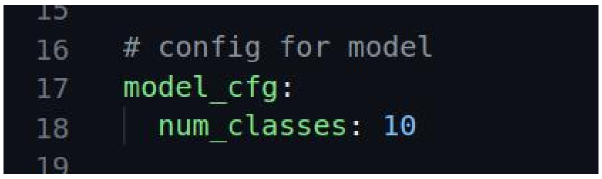
tweak **./model.py**, here I implement LeNet5



```python
class MyModel(nn.Module):
    def __init__(self, num_classes):
        super().__init__()
        self.feature = nn.Sequential(
            nn.Conv2d(in_channels=1, out_channels=6, kernel_size=5, stride=1, padding=2),
            nn.Tanh(),
            nn.AvgPool2d(kernel_size=2, stride=2),

            nn.Conv2d(in_channels=6, out_channels=16, kernel_size=5, stride=1),
            nn.Tanh(),
            nn.AvgPool2d(kernel_size=2, stride=2),  # 5*5

        )
        self.classifier = nn.Sequential(
            nn.Flatten(),
            nn.Linear(in_features=16*5*5, out_features=120),
            nn.Tanh(),
            nn.Linear(in_features=120, out_features=84),
            nn.Tanh(),
            nn.Linear(in_features=84, out_features=num_classes),
        )

    def forward(self, x):
        return self.classifier(self.feature(x))
```

tweak **./config.yaml**

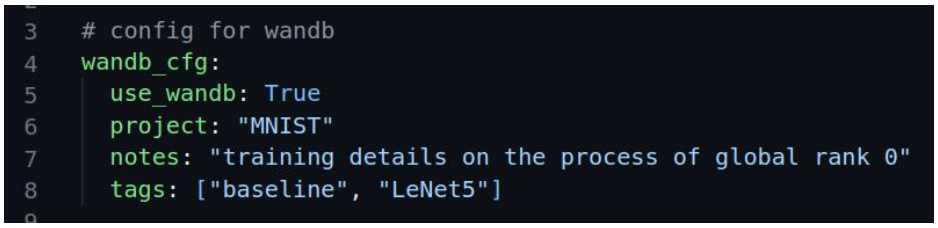
as I just changed the model, so change the model_cfg



```yaml
# config for model
model_cfg:
  num_classes: 10
```

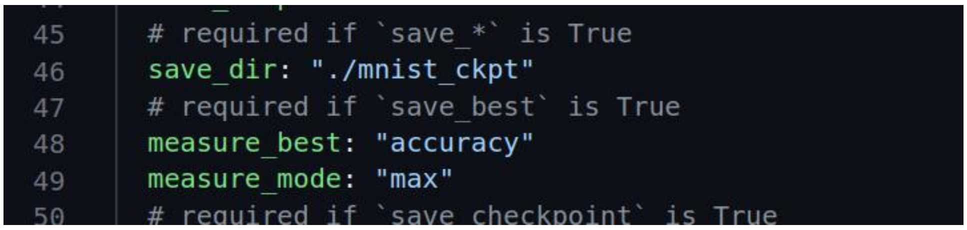
by the way, change other configs

keep using wandb and change some config for it



```yaml
# config for wandb
wandb_cfg:
  use_wandb: True
  project: "MNIST"
  notes: "training details on the process of global rank 0"
  tags: ["baseline", "LeNet5"]
```

for the train_cfg, the default ones are almost ok, I just tweak the part about save dir and the measurement for best model, and remember to create the save directory manually



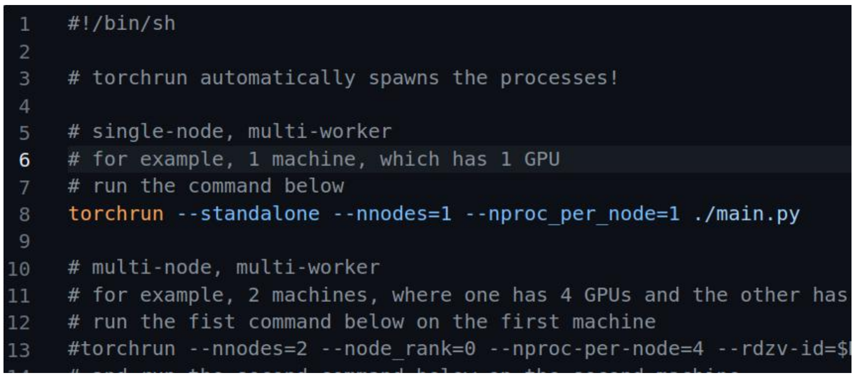
```yaml
  # required if `save_*` is True
  save_dir: "./mnist_ckpt"
  # required if `save_best` is True
  measure_best: "accuracy"
  measure_mode: "max"
  # required if `save_checkpoint` is True
```

by default, the template use CrossEntropyLoss for criterion, AdamW for optimizer, CosineAnnealingWarmRestarts for lr scheduler, it seems ok, so I didn’t change these and the corresponding configs

the Trainer is off-the-shelf, no need to change anything about it

as for the test method, I just want to test accuracy, which is already implemented by template, so I just keep it

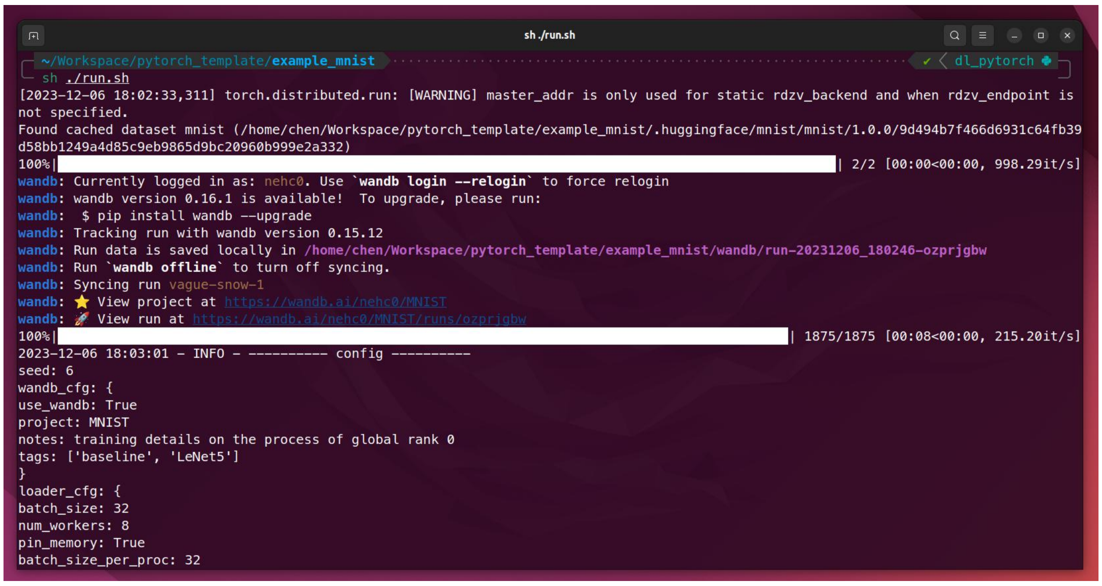
lastly, tweak **./run.sh**, as I will train on my laptop with single gpu



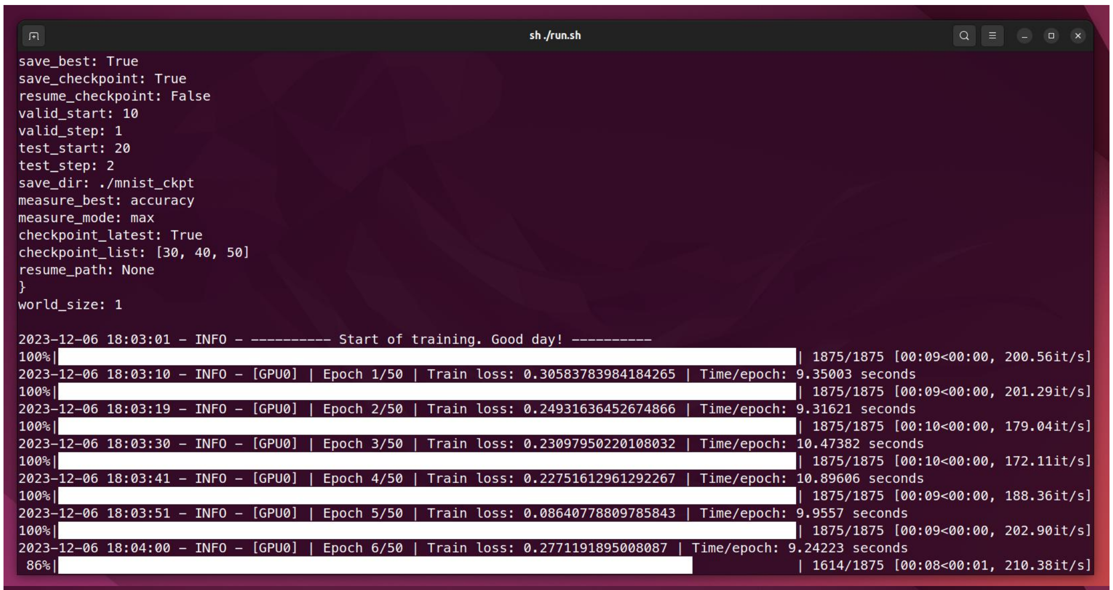
```
#!/bin/sh

# torchrun automatically spawns the processes!

# single-node, multi-worker
# for example, 1 machine, which has 1 GPU
# run the command below
torchrun --standalone --nnodes=1 --nproc_per_node=1 ./main.py

# multi-node, multi-worker
# for example, 2 machines, where one has 4 GPUs and the other has
# run the fist command below on the first machine
#torchrun --nnodes=2 --node_rank=0 --nproc-per-node=4 --rdzv-id=$
```

2. train procedure

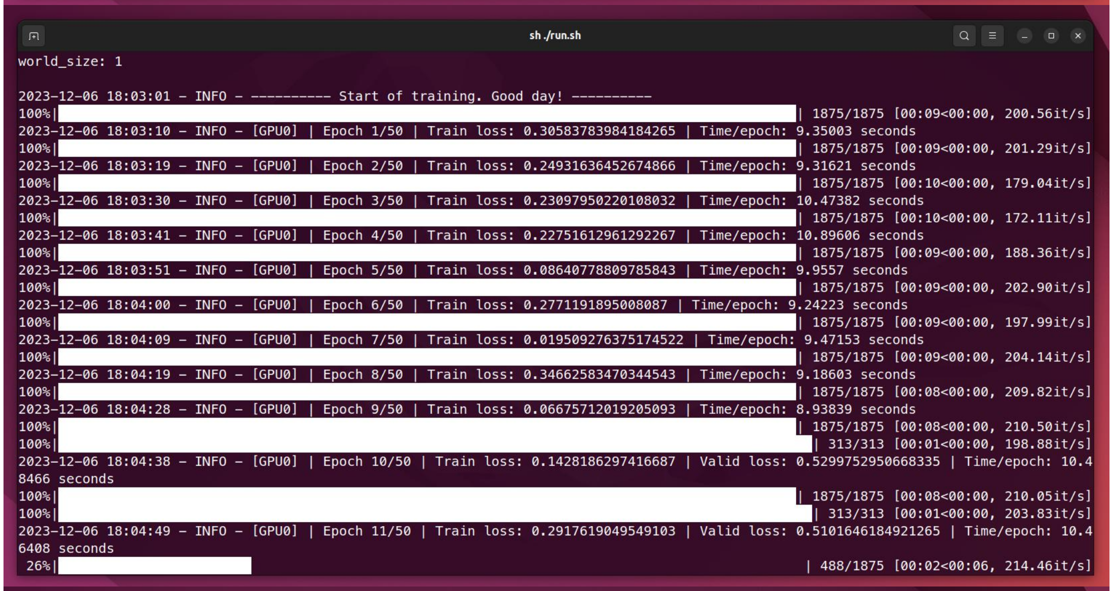
start training, terminal:



```
~/Workspace/pytorch_template/example_mnist ·························· ✔ ❮ dl_pytorch
sh ./run.sh
[2023-12-06 18:02:33,311] torch.distributed.run: [WARNING] master_addr is only used for static rdzv_backend and when rdzv_endpoint is
not specified.
Found cached dataset mnist (/home/chen/Workspace/pytorch_template/example_mnist/.huggingface/mnist/mnist/1.0.0/9d494b7f466d6931c64fb39
d58bb1249a4d85c9eb9865d9bc20960b999e2a332)
100%|██████████████████████████████████████████████████| 2/2 [00:00<00:00, 998.29it/s]
wandb: Currently logged in as: nehc0. Use `wandb login --relogin` to force relogin
wandb: wandb version 0.16.1 is available!  To upgrade, please run:
wandb:  $ pip install wandb --upgrade
wandb: Tracking run with wandb version 0.15.12
wandb: Run data is saved locally in /home/chen/Workspace/pytorch_template/example_mnist/wandb/run-20231206_180246-ozprjgbw
wandb: Run `wandb offline` to turn off syncing.
wandb: Syncing run vague-snow-1
wandb: ⭐ View project at https://wandb.ai/nehc0/MNIST
wandb: 🚀 View run at https://wandb.ai/nehc0/MNIST/runs/ozprjgbw
100%|███████████████████████████████████████████████| 1875/1875 [00:08<00:00, 215.20it/s]
2023-12-06 18:03:01 - INFO - ---------- config ----------
seed: 6
wandb_cfg: {
use_wandb: True
project: MNIST
notes: training details on the process of global rank 0
tags: ['baseline', 'LeNet5']
}
loader_cfg: {
batch_size: 32
num_workers: 8
pin_memory: True
batch_size_per_proc: 32
```

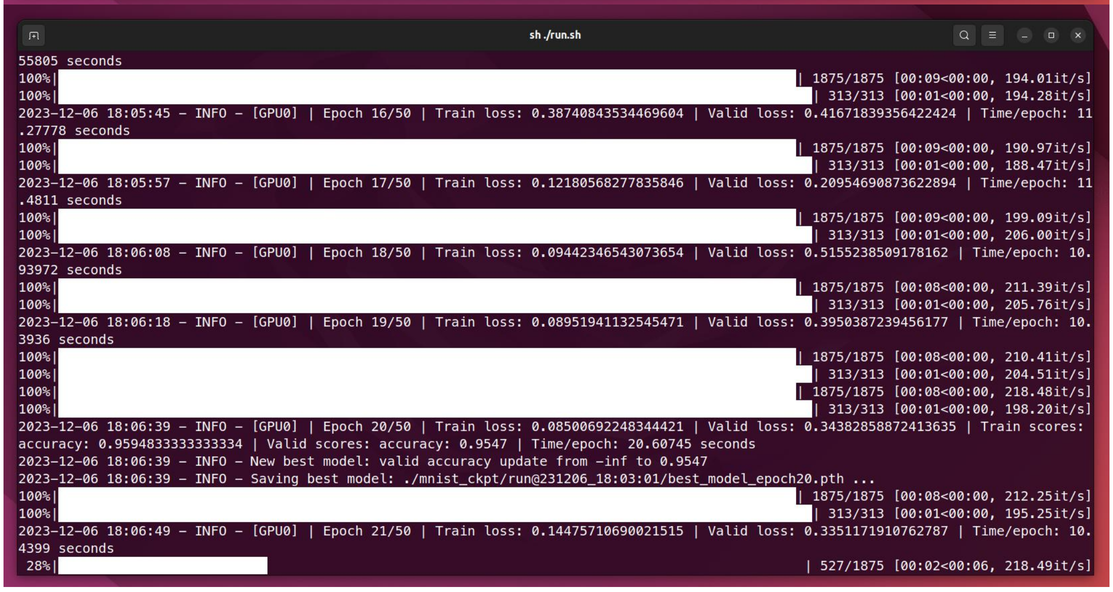


```
save_best: True
save_checkpoint: True
resume_checkpoint: False
valid_start: 10
valid_step: 1
test_start: 20
test_step: 2
save_dir: ./mnist_ckpt
measure_best: accuracy
measure_mode: max
checkpoint_latest: True
checkpoint_list: [30, 40, 50]
resume_path: None
}
world_size: 1

2023-12-06 18:03:01 - INFO - ---------- Start of training. Good day! ----------
100%|████████████████████| 1875/1875 [00:09<00:00, 200.56it/s]
2023-12-06 18:03:10 - INFO - [GPU0] | Epoch 1/50 | Train loss: 0.30583783984184265 | Time/epoch: 9.35003 seconds
100%|████████████████████| 1875/1875 [00:09<00:00, 201.29it/s]
2023-12-06 18:03:19 - INFO - [GPU0] | Epoch 2/50 | Train loss: 0.24931636452674866 | Time/epoch: 9.31621 seconds
100%|████████████████████| 1875/1875 [00:10<00:00, 179.04it/s]
2023-12-06 18:03:30 - INFO - [GPU0] | Epoch 3/50 | Train loss: 0.23097950220108032 | Time/epoch: 10.47382 seconds
100%|████████████████████| 1875/1875 [00:10<00:00, 172.11it/s]
2023-12-06 18:03:41 - INFO - [GPU0] | Epoch 4/50 | Train loss: 0.22751612961292267 | Time/epoch: 10.89606 seconds
100%|████████████████████| 1875/1875 [00:09<00:00, 188.36it/s]
2023-12-06 18:03:51 - INFO - [GPU0] | Epoch 5/50 | Train loss: 0.08640778809785843 | Time/epoch: 9.9557 seconds
100%|████████████████████| 1875/1875 [00:09<00:00, 202.90it/s]
2023-12-06 18:04:00 - INFO - [GPU0] | Epoch 6/50 | Train loss: 0.2771191895008087 | Time/epoch: 9.24223 seconds
 86%|█████████████████   | 1614/1875 [00:08<00:01, 210.38it/s]
```

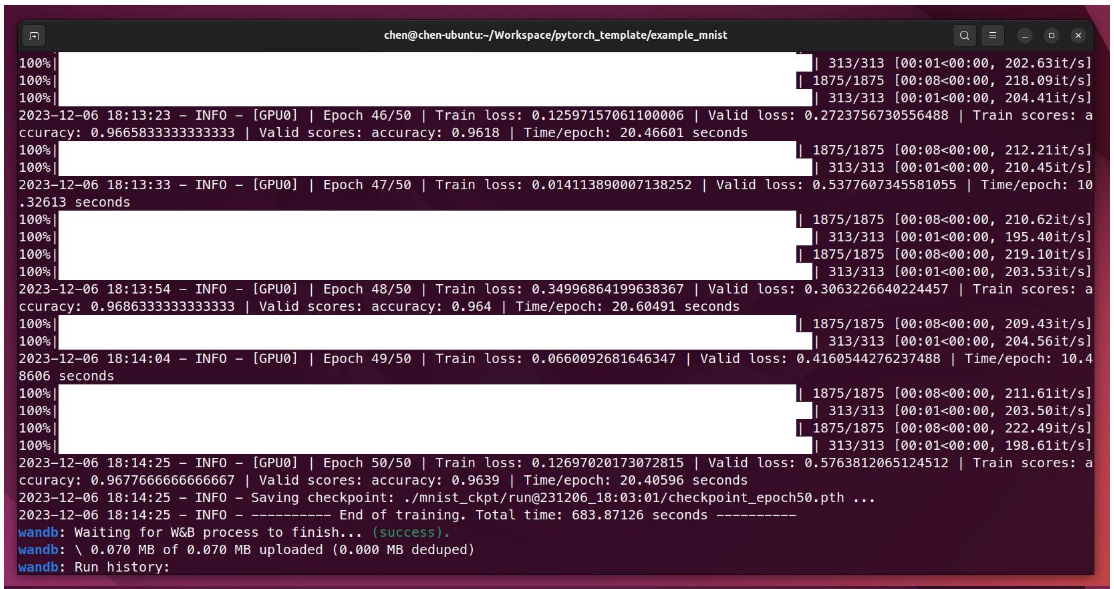


```
world_size: 1

2023-12-06 18:03:01 - INFO - ---------- Start of training. Good day! ----------
100%|████████████████████| 1875/1875 [00:09<00:00, 200.56it/s]
2023-12-06 18:03:10 - INFO - [GPU0] | Epoch 1/50 | Train loss: 0.30583783984184265 | Time/epoch: 9.35003 seconds
100%|████████████████████| 1875/1875 [00:09<00:00, 201.29it/s]
2023-12-06 18:03:19 - INFO - [GPU0] | Epoch 2/50 | Train loss: 0.24931636452674866 | Time/epoch: 9.31621 seconds
100%|████████████████████| 1875/1875 [00:10<00:00, 179.04it/s]
2023-12-06 18:03:30 - INFO - [GPU0] | Epoch 3/50 | Train loss: 0.23097950220108032 | Time/epoch: 10.47382 seconds
100%|████████████████████| 1875/1875 [00:10<00:00, 172.11it/s]
2023-12-06 18:03:41 - INFO - [GPU0] | Epoch 4/50 | Train loss: 0.22751612961292267 | Time/epoch: 10.89606 seconds
100%|████████████████████| 1875/1875 [00:09<00:00, 188.36it/s]
2023-12-06 18:03:51 - INFO - [GPU0] | Epoch 5/50 | Train loss: 0.08640778809785843 | Time/epoch: 9.9557 seconds
100%|████████████████████| 1875/1875 [00:09<00:00, 202.90it/s]
2023-12-06 18:04:00 - INFO - [GPU0] | Epoch 6/50 | Train loss: 0.2771191895008087 | Time/epoch: 9.24223 seconds
```

```
100%|████████████████████| 1875/1875 [00:09<00:00, 197.99it/s]
2023-12-06 18:04:09 - INFO - [GPU0] | Epoch 7/50 | Train loss: 0.019509276375174522 | Time/epoch: 9.47153 seconds
100%|████████████████████| 1875/1875 [00:09<00:00, 204.14it/s]
2023-12-06 18:04:19 - INFO - [GPU0] | Epoch 8/50 | Train loss: 0.34662583470344543 | Time/epoch: 9.18603 seconds
100%|████████████████████| 1875/1875 [00:08<00:00, 209.82it/s]
2023-12-06 18:04:28 - INFO - [GPU0] | Epoch 9/50 | Train loss: 0.06675712019205093 | Time/epoch: 8.93839 seconds
100%|████████████████████| 1875/1875 [00:08<00:00, 210.50it/s]
100%|█████████████████████| 313/313 [00:01<00:00, 198.88it/s]
2023-12-06 18:04:38 - INFO - [GPU0] | Epoch 10/50 | Train loss: 0.1428186297416687 | Valid loss: 0.5299752950668335 | Time/epoch: 10.4
8466 seconds
100%|████████████████████| 1875/1875 [00:08<00:00, 210.05it/s]
100%|█████████████████████| 313/313 [00:01<00:00, 203.83it/s]
2023-12-06 18:04:49 - INFO - [GPU0] | Epoch 11/50 | Train loss: 0.2917619049549103 | Valid loss: 0.5101646184921265 | Time/epoch: 10.4
6408 seconds
 26%|█████               | 488/1875 [00:02<00:06, 214.46it/s]
```

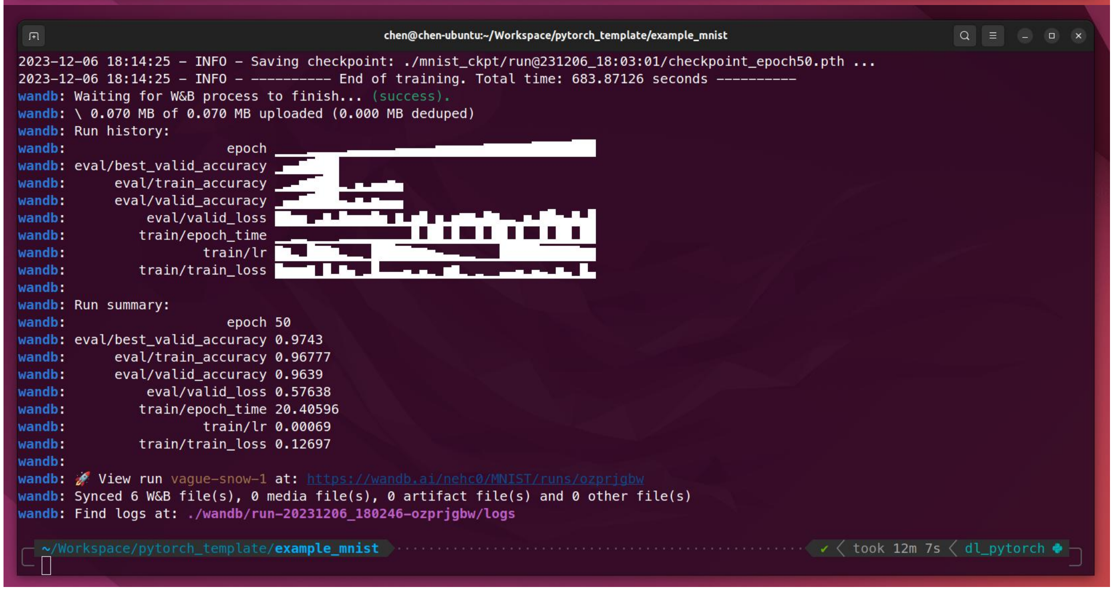


```
55805 seconds
100%|████████████████████| 1875/1875 [00:09<00:00, 194.01it/s]
100%|█████████████████████| 313/313 [00:01<00:00, 194.28it/s]
2023-12-06 18:05:45 - INFO - [GPU0] | Epoch 16/50 | Train loss: 0.38740843534469604 | Valid loss: 0.41671839356422424 | Time/epoch: 11
.27778 seconds
100%|████████████████████| 1875/1875 [00:09<00:00, 190.97it/s]
100%|█████████████████████| 313/313 [00:01<00:00, 188.47it/s]
2023-12-06 18:05:57 - INFO - [GPU0] | Epoch 17/50 | Train loss: 0.12180568277835846 | Valid loss: 0.20954690873622894 | Time/epoch: 11
.4811 seconds
100%|████████████████████| 1875/1875 [00:09<00:00, 199.09it/s]
100%|█████████████████████| 313/313 [00:01<00:00, 206.00it/s]
2023-12-06 18:06:08 - INFO - [GPU0] | Epoch 18/50 | Train loss: 0.09442346543073654 | Valid loss: 0.5155238509178162 | Time/epoch: 10.
93972 seconds
100%|████████████████████| 1875/1875 [00:08<00:00, 211.39it/s]
100%|█████████████████████| 313/313 [00:01<00:00, 205.76it/s]
2023-12-06 18:06:18 - INFO - [GPU0] | Epoch 19/50 | Train loss: 0.08951941132545471 | Valid loss: 0.3950387239456177 | Time/epoch: 10.
3936 seconds
100%|████████████████████| 1875/1875 [00:08<00:00, 210.41it/s]
100%|█████████████████████| 313/313 [00:01<00:00, 204.51it/s]
100%|████████████████████| 1875/1875 [00:08<00:00, 218.48it/s]
100%|█████████████████████| 313/313 [00:01<00:00, 198.20it/s]
2023-12-06 18:06:39 - INFO - [GPU0] | Epoch 20/50 | Train loss: 0.08500692248344421 | Valid loss: 0.34382858872413635 | Train scores:
accuracy: 0.9594833333333334 | Valid scores: accuracy: 0.9547 | Time/epoch: 20.60745 seconds
2023-12-06 18:06:39 - INFO - New best model: valid accuracy update from -inf to 0.9547
2023-12-06 18:06:39 - INFO - Saving best model: ./mnist_ckpt/run@231206_18:03:01/best_model_epoch20.pth ...
100%|████████████████████| 1875/1875 [00:08<00:00, 212.25it/s]
100%|█████████████████████| 313/313 [00:01<00:00, 195.25it/s]
2023-12-06 18:06:49 - INFO - [GPU0] | Epoch 21/50 | Train loss: 0.14475710690021515 | Valid loss: 0.3351171910762787 | Time/epoch: 10.
4399 seconds
 28%|██████              | 527/1875 [00:02<00:06, 218.49it/s]
```



chen@chen-ubuntu:~/Workspace/pytorch_template/example_mnist

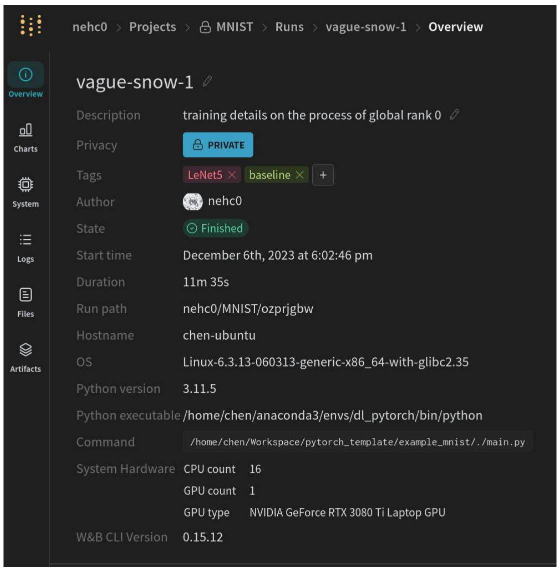
```
100%|                                                    | 313/313 [00:01<00:00, 202.63it/s]
100%|                                                  | 1875/1875 [00:08<00:00, 218.09it/s]
100%|                                                    | 313/313 [00:01<00:00, 204.41it/s]
2023-12-06 18:13:23 - INFO - [GPU0] | Epoch 46/50 | Train loss: 0.12597157061100006 | Valid loss: 0.2723756730556488 | Train scores: a
ccuracy: 0.9665833333333333 | Valid scores: accuracy: 0.9618 | Time/epoch: 20.46601 seconds
100%|                                                  | 1875/1875 [00:08<00:00, 212.21it/s]
100%|                                                    | 313/313 [00:01<00:00, 210.45it/s]
2023-12-06 18:13:33 - INFO - [GPU0] | Epoch 47/50 | Train loss: 0.014113890007138252 | Valid loss: 0.5377607345581055 | Time/epoch: 10
.32613 seconds
100%|                                                  | 1875/1875 [00:08<00:00, 210.62it/s]
100%|                                                    | 313/313 [00:01<00:00, 195.40it/s]
100%|                                                  | 1875/1875 [00:08<00:00, 219.10it/s]
100%|                                                    | 313/313 [00:01<00:00, 203.53it/s]
2023-12-06 18:13:54 - INFO - [GPU0] | Epoch 48/50 | Train loss: 0.34996864199638367 | Valid loss: 0.3063226640224457 | Train scores: a
ccuracy: 0.9686333333333333 | Valid scores: accuracy: 0.964 | Time/epoch: 20.60491 seconds
100%|                                                  | 1875/1875 [00:08<00:00, 209.43it/s]
100%|                                                    | 313/313 [00:01<00:00, 204.56it/s]
2023-12-06 18:14:04 - INFO - [GPU0] | Epoch 49/50 | Train loss: 0.0660092681646347 | Valid loss: 0.4160544276237488 | Time/epoch: 10.4
8606 seconds
100%|                                                  | 1875/1875 [00:08<00:00, 211.61it/s]
100%|                                                    | 313/313 [00:01<00:00, 203.50it/s]
100%|                                                  | 1875/1875 [00:08<00:00, 222.49it/s]
100%|                                                    | 313/313 [00:01<00:00, 198.61it/s]
2023-12-06 18:14:25 - INFO - [GPU0] | Epoch 50/50 | Train loss: 0.12697020173072815 | Valid loss: 0.5763812065124512 | Train scores: a
ccuracy: 0.9677666666666667 | Valid scores: accuracy: 0.9639 | Time/epoch: 20.40596 seconds
2023-12-06 18:14:25 - INFO - Saving checkpoint: ./mnist_ckpt/run@231206_18:03:01/checkpoint_epoch50.pth ...
2023-12-06 18:14:25 - INFO - ---------- End of training. Total time: 683.87126 seconds ----------
wandb: Waiting for W&B process to finish... (success).
wandb: \ 0.070 MB of 0.070 MB uploaded (0.000 MB deduped)
wandb: Run history:
```



chen@chen-ubuntu:~/Workspace/pytorch_template/example_mnist

```
2023-12-06 18:14:25 - INFO - Saving checkpoint: ./mnist_ckpt/run@231206_18:03:01/checkpoint_epoch50.pth ...
2023-12-06 18:14:25 - INFO - ---------- End of training. Total time: 683.87126 seconds ----------
wandb: Waiting for W&B process to finish... (success).
wandb: \ 0.070 MB of 0.070 MB uploaded (0.000 MB deduped)
wandb: Run history:
wandb:                     epoch
wandb: eval/best_valid_accuracy
wandb:      eval/train_accuracy
wandb:      eval/valid_accuracy
wandb:          eval/valid_loss
wandb:         train/epoch_time
wandb:                 train/lr
wandb:         train/train_loss
wandb:
wandb: Run summary:
wandb:                     epoch 50
wandb: eval/best_valid_accuracy 0.9743
wandb:      eval/train_accuracy 0.96777
wandb:      eval/valid_accuracy 0.9639
wandb:          eval/valid_loss 0.57638
wandb:         train/epoch_time 20.40596
wandb:                 train/lr 0.00069
wandb:         train/train_loss 0.12697
wandb:
wandb: 🚀 View run vague-snow-1 at: https://wandb.ai/nehc0/MNIST/runs/ozprjgbw
wandb: Synced 6 W&B file(s), 0 media file(s), 0 artifact file(s) and 0 other file(s)
wandb: Find logs at: ./wandb/run-20231206_180246-ozprjgbw/logs

~/Workspace/pytorch_template/example_mnist ·································· ✓ took 12m 7s dl_pytorch
```

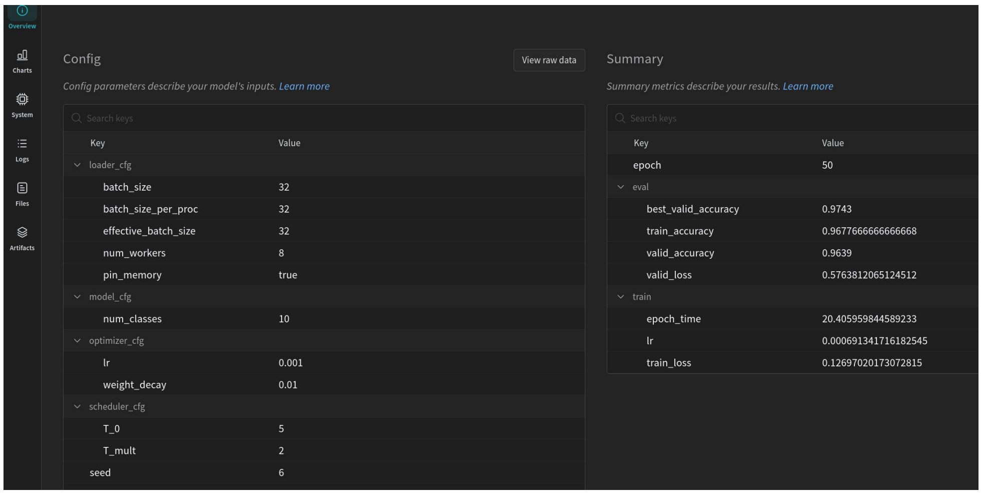
wandb:



nehc0 › Projects › MNIST › Runs › vague-snow-1 › Overview

vague-snow-1

Description	training details on the process of global rank 0
Privacy	PRIVATE
Tags	LeNet5 baseline
Author	nehc0
State	Finished
Start time	December 6th, 2023 at 6:02:46 pm
Duration	11m 35s
Run path	nehc0/MNIST/ozprjgbw
Hostname	chen-ubuntu
OS	Linux-6.3.13-060313-generic-x86_64-with-glibc2.35
Python version	3.11.5
Python executable	/home/chen/anaconda3/envs/dl_pytorch/bin/python
Command	/home/chen/Workspace/pytorch_template/example_mnist/./main.py
System Hardware	CPU count 16
	GPU count 1
	GPU type NVIDIA GeForce RTX 3080 Ti Laptop GPU
W&B CLI Version	0.15.12



Config

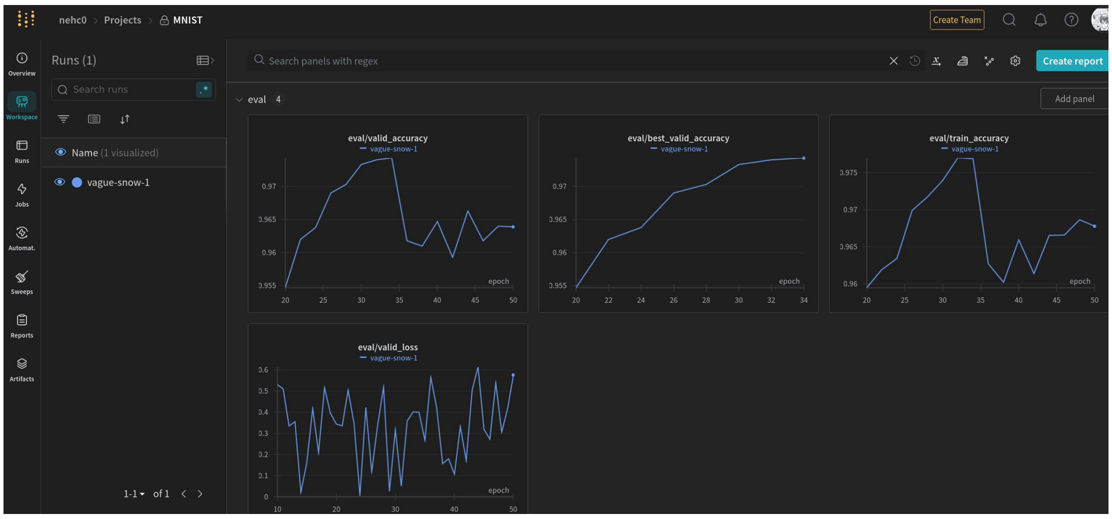
Config parameters describe your model's inputs. Learn more

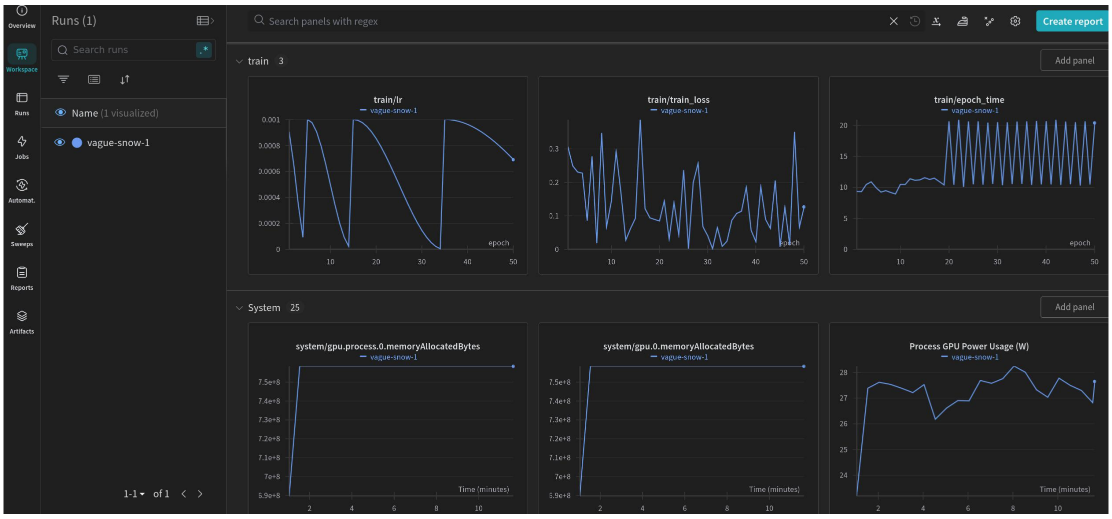
Key	Value
loader_cfg	
batch_size	32
batch_size_per_proc	32
effective_batch_size	32
num_workers	8
pin_memory	true
model_cfg	
num_classes	10
optimizer_cfg	
lr	0.001
weight_decay	0.01
scheduler_cfg	
T_0	5
T_mult	2
seed	6

Summary

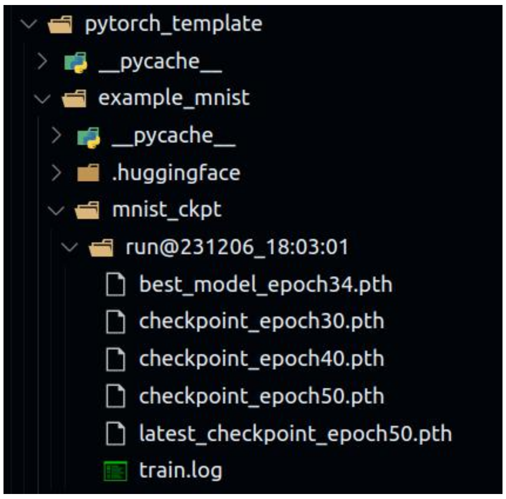
Summary metrics describe your results. Learn more

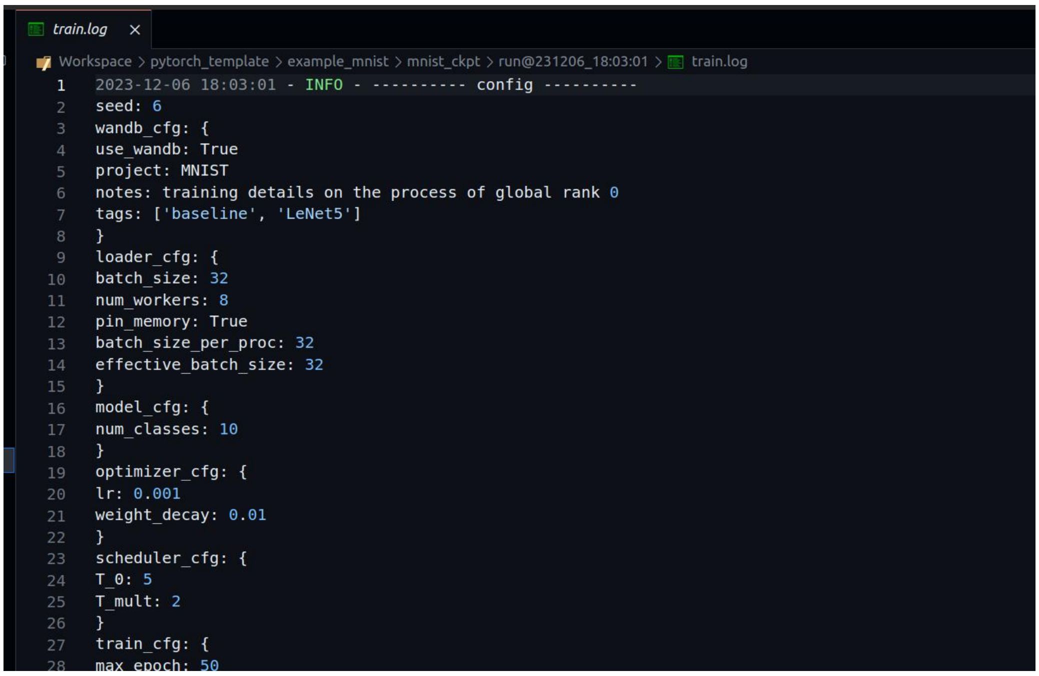
Key	Value
epoch	50
eval	
best_valid_accuracy	0.9743
train_accuracy	0.9677666666666668
valid_accuracy	0.9639
valid_loss	0.5763812065124512
train	
epoch_time	20.405959844589233
lr	0.000691341716182545
train_loss	0.12697020173072815



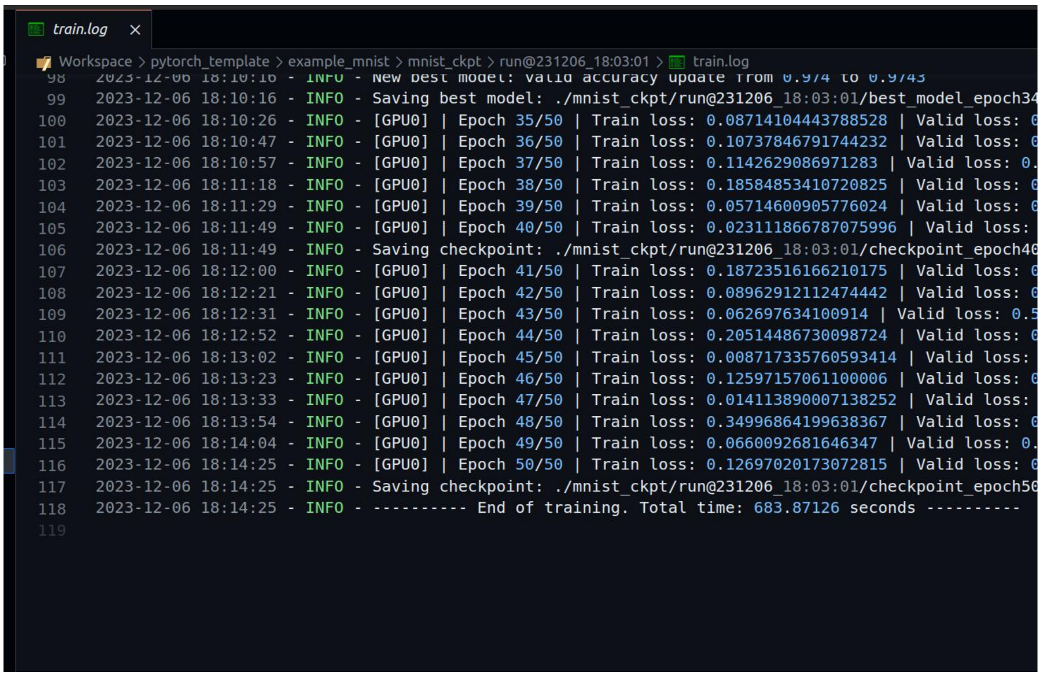


saved logs and checkpoints:


```
pytorch_template
  __pycache__
  example_mnist
    __pycache__
    .huggingface
    mnist_ckpt
      run@231206_18:03:01
        best_model_epoch34.pth
        checkpoint_epoch30.pth
        checkpoint_epoch40.pth
        checkpoint_epoch50.pth
        latest_checkpoint_epoch50.pth
        train.log
```


train.log
Workspace > pytorch_template > example_mnist > mnist_ckpt > run@231206_18:03:01 > train.log

```
2023-12-06 18:03:01 - INFO - ---------- config ----------
seed: 6
wandb_cfg: {
use_wandb: True
project: MNIST
notes: training details on the process of global rank 0
tags: ['baseline', 'LeNet5']
}
loader_cfg: {
batch_size: 32
num_workers: 8
pin_memory: True
batch_size_per_proc: 32
effective_batch_size: 32
}
model_cfg: {
num_classes: 10
}
optimizer_cfg: {
lr: 0.001
weight_decay: 0.01
}
scheduler_cfg: {
T_0: 5
T_mult: 2
}
train_cfg: {
max_epoch: 50
```


train.log
Workspace > pytorch_template > example_mnist > mnist_ckpt > run@231206_18:03:01 > train.log

```
2023-12-06 18:10:16 - INFO - New best model: Valid accuracy update from 0.974 to 0.9743
2023-12-06 18:10:16 - INFO - Saving best model: ./mnist_ckpt/run@231206_18:03:01/best_model_epoch34
2023-12-06 18:10:26 - INFO - [GPU0] | Epoch 35/50 | Train loss: 0.08714104443788528 | Valid loss: 0
2023-12-06 18:10:47 - INFO - [GPU0] | Epoch 36/50 | Train loss: 0.10737846791744232 | Valid loss: 0
2023-12-06 18:10:57 - INFO - [GPU0] | Epoch 37/50 | Train loss: 0.1142629086971283 | Valid loss: 0.
2023-12-06 18:11:18 - INFO - [GPU0] | Epoch 38/50 | Train loss: 0.18584853410720825 | Valid loss: 0
2023-12-06 18:11:29 - INFO - [GPU0] | Epoch 39/50 | Train loss: 0.05714600905776024 | Valid loss: 0
2023-12-06 18:11:49 - INFO - [GPU0] | Epoch 40/50 | Train loss: 0.023111866787075996 | Valid loss:
2023-12-06 18:11:49 - INFO - Saving checkpoint: ./mnist_ckpt/run@231206_18:03:01/checkpoint_epoch40
2023-12-06 18:12:00 - INFO - [GPU0] | Epoch 41/50 | Train loss: 0.18723516166210175 | Valid loss: 0
2023-12-06 18:12:21 - INFO - [GPU0] | Epoch 42/50 | Train loss: 0.08962912112474442 | Valid loss: 0
2023-12-06 18:12:31 - INFO - [GPU0] | Epoch 43/50 | Train loss: 0.062697634100914 | Valid loss: 0.5
2023-12-06 18:12:52 - INFO - [GPU0] | Epoch 44/50 | Train loss: 0.20514486730098724 | Valid loss: 0
2023-12-06 18:13:02 - INFO - [GPU0] | Epoch 45/50 | Train loss: 0.008717335760593414 | Valid loss:
2023-12-06 18:13:23 - INFO - [GPU0] | Epoch 46/50 | Train loss: 0.12597157061100006 | Valid loss: 0
2023-12-06 18:13:33 - INFO - [GPU0] | Epoch 47/50 | Train loss: 0.014113890007138252 | Valid loss:
2023-12-06 18:13:54 - INFO - [GPU0] | Epoch 48/50 | Train loss: 0.34996864199638367 | Valid loss: 0
2023-12-06 18:14:04 - INFO - [GPU0] | Epoch 49/50 | Train loss: 0.0660092681646347 | Valid loss: 0.
2023-12-06 18:14:25 - INFO - [GPU0] | Epoch 50/50 | Train loss: 0.12697020173072815 | Valid loss: 0
2023-12-06 18:14:25 - INFO - Saving checkpoint: ./mnist_ckpt/run@231206_18:03:01/checkpoint_epoch50
2023-12-06 18:14:25 - INFO - ---------- End of training. Total time: 683.87126 seconds ----------
```