Screenshots of printed output as model is training:

```
)
[2024-11-19 17:05:12.660807][INFO][trainer.py:809] - Startup time: 12.0371
                Training Legend

| abbr         | desc                                 |
|--------------|--------------------------------------|
| step         | Current training iteration           |
| loss         | Loss value                           |
| dt           | Elapsed time per training step       |
| dtf          | Elapsed time per forward step        |
| dtb          | Elapsed time per backward step       |
| sps          | Samples per second                   |
| sps_per_gpu  | Samples per second (per GPU)         |
| tps          | Tokens per second                    |
| tps_per_gpu  | Tokens per second (per GPU)          |
| mfu          | Model flops utilization              |

[2024-11-19 17:05:14.086439][INFO][trainer.py:827] - ['prompt']: 'What is an LLM?'
[2024-11-19 17:05:14.088144][INFO][trainer.py:831] - ['response']:

What is an LLM?NZNZcReeIRyRNzzzft?DJEFFtIRFUUlI;IIIIBooVIA?oIGDDII rcJIzIIRL IIPRZQIIRDJHqFIIIqtcR  etgIIII UIUv..AIImIIIIIFDBgVttlFeZ-RJyZ?agmo.zzQFtttttle
RrmaR A3URtggyIImlDt RFtzwKKZMHJtReImRor  FUIBUNIrmIIRJ cIIQRmiIUJRoaFIRoU.IIqtIrhtID?RRggIRBw!JYf;tPrrxRRRRR rRvI3
```

```
513e+06 tps_per_gpu=223141 mfu=43.633
[2024-11-19 17:06:23.084972][INFO][trainer.py:827] - ['prompt']: 'What is an LLM?'
[2024-11-19 17:06:23.088048][INFO][trainer.py:831] - ['response']:

What is an LLM?
MANA:
Thy, shat amy my indare mellile
Whink m'd han h worsord oursthe thean foutlld t thary fouiris y at f;
Whe me wen waslyof tourist asour ald shamethofe al or t m
the, hthind windurear t
Toreenel mpin wan tittes y des bar hen oo y s mes tulin thinoruso
[2024-11-19 17:07:01.484626][INFO][trainer.py:762] - Saving checkpoint to: /home/khalil_nor/outputs/runs/pytorch/DDP/2024-11-19/17-05-01
[2024-11-19 17:07:01.488188][INFO][trainer.py:763] - Saving model to: /home/khalil_nor/outputs/runs/pytorch/DDP/2024-11-19/17-05-01/model.pth
[2024-11-19 17:07:03.744789][INFO][configs.py:141] - Appending /home/khalil_nor/outputs/runs/pytorch/DDP/2024-11-19/17-05-01 to /home/khalil_nor
```

```
37e+06 tps_per_gpu=223796 mfu=43.5145
[2024-11-19 17:07:34.949230][INFO][trainer.py:827] - ['prompt']: 'What is an LLM?'
[2024-11-19 17:07:34.951992][INFO][trainer.py:831] - ['response']:

What is an LLM?

COMPERMPENCELI:
And hinher whers aviciond thred fronth.

Me therronive macht, thou shastengine ingay,
I sath as himend thand to wepore frorsed.

Ay CLARIFFOLIND:
Thall's why sor tinee peele,
And ast wore, I'll to offorter.
ISis sperille dort ong hour the
[2024-11-19 17:08:13.184374][INFO][trainer.py:762] - Saving checkpoint to: /home/khalil_nor/outputs/runs/pytorch/DDP/2024-11-19/17-05-01
```

```
[2024-11-19 17:08:47.450423][INFO][trainer.py:831] - ['response']:

What is an LLM?

LUCIO:
O, no for so, and my lord.

DUKE VINCENTIO:
Most none, but he
spent your till a laves adied
the kneep of the earth is in soon ass one with
title appart of your desild, the know they
pring.

LUCIO:
Here it be in prisoner.

DUKE VINCENTIO:
But as if
[2024-11-19 17:09:25.509870][INFO][trainer.py:762] - Saving checkpoint to: /home/khalil_nor/outputs/runs/pytorch/DDP/202
```

```
[2024-11-19 17:09:59.619171][INFO][trainer.py:831] - ['response']:

What is an LLM?

ROMEO:
She take it of many slew and cares set
Then stay may be forth dear: and true scarried
Should no crim a ject in my eye.

MERCUTIO:
When it is the day in a mind.

ROMEO:
A more of man for mioney; then cert me well me
with mortal be child. He was he
[2024-11-19 17:10:38.184526][INFO][trainer.py:762] - Saving checkpoint to: /h
```

```
2754e+06 tps_per_gpu=215943 mfu=43.8601
[2024-11-19 17:11:08.206408][INFO][trainer.py:892] - step=490 loss=0.70227 dt=0.300995 dtf=0.0064693 dtb=0.0176405 sps=26.5785 sps_per_gpu=3.32232 tps=1.
85e+06 tps_per_gpu=217731 mfu=43.8236
[2024-11-19 17:11:11.207686][INFO][trainer.py:892] - step=500 loss=0.574171 dt=0.307926 dtf=0.00750012 dtb=0.0179673 sps=25.9802 sps_per_gpu=3.24753 tps=
0264e+06 tps_per_gpu=212830 mfu=43.6929
[2024-11-19 17:11:12.347990][INFO][trainer.py:827] - ['prompt']: 'What is an LLM?'
[2024-11-19 17:11:12.350371][INFO][trainer.py:831] - ['response']:

What is an LLM?

MOPSA:
Sir, thou dost at my forces, being them lasthes;
And therefore shall tender hear?

Clown:
What shall think you?

AUTOLYCUS:
Neither to me?

Clown:
Neither. I know I will hear should she she should
melt here; for her fife fourself none
Desneed year
[2024-11-19 17:11:50.547380][INFO][trainer.py:762] - Saving checkpoint to: /home/khalil_nor/outputs/runs/pytorch/DDP/2024-11-19/17-05-01
[2024-11-19 17:11:50.549099][INFO][trainer.py:763] - Saving model to: /home/khalil_nor/outputs/runs/pytorch/DDP/2024-11-19/17-05-01/model.pth
[2024-11-19 17:11:53.317587][INFO][configs.py:141] - Appending /home/khalil_nor/outputs/runs/pytorch/DDP/2024-11-19/17-05-01 to /home/khalil_nor/wordplay
```

```
[2024-11-19 17:12:24.615854][INFO][trainer.py:827] - ['prompt']: 'What is an LLM?'
[2024-11-19 17:12:24.617332][INFO][trainer.py:831] - ['response']:

What is an LLM?

GLOUCESTER:
Then know, my lord, the fourth is entertain?

QUEEN MARGARET:
Who is your may tongue to be princely me?
And all the serperate more than the beast!
The break of your foes adversons! yet direfends!
You have scorns well have impress'd your house
[2024-11-19 17:13:02.427013][INFO][trainer.py:762] - Saving checkpoint to: /home/khalil_nor/outputs/runs
[2024-11-19 17:13:02.428807][INFO][trainer.py:763] - Saving model to: /home/khalil_nor/outputs/runs/pyto
[2024-11-19 17:13:05.152270][INFO][configs.py:141] - Appending /home/khalil_nor/outputs/runs/pytorch/DDP
c/ckpts/checkpoints.log
```

```
[2024-11-19 17:13:36.446207][INFO][trainer.py:827] - ['prompt']: 'What is an LLM?'
[2024-11-19 17:13:36.448447][INFO][trainer.py:831] - ['response']:

What is an LLM?

NORTHUMBERLAND:
Nay, nothing come, have you are banish'd;
And you mercy to have strength upon your heads,
Your crown sweet with your grace and by some prophetance;
The young was admit your action head,
Your son's exile is a rapition;
And the loss I have
[2024-11-19 17:14:14.433147][INFO][trainer.py:762] - Saving checkpoint to: /home/khalil_nor/o
[2024-11-19 17:14:14.434952][INFO][trainer.py:763] - Saving model to: /home/khalil_nor/output
[2024-11-19 17:14:17.257493][INFO][configs.py:141] - Appending /home/khalil_nor/outputs/runs/
```

```
1694e+06 tps_per_gpu=214618 mfu=43.3817
[2024-11-19 17:16:01.014484][INFO][trainer.py:827] - ['prompt']: 'What is an LLM?'
[2024-11-19 17:16:01.016004][INFO][trainer.py:831] - ['response']:

What is an LLM?

POLIXENES:
Young private good night:
I know not where you were let Rome meet hours,
To the carterous of the Duke of Norfolk.

KING EDWARD IV:
Then Oas much in Hereford here sox here in heaven.

HENRY BOLINGBROKE:
Yet as not post to the King.

KING EDWARD
[2024-11-19 17:16:38.849457][INFO][trainer.py:762] - Saving checkpoint to: /home/khalil_nor/(
[2024-11-19 17:16:38.853418][INFO][trainer.py:763] - Saving model to: /home/khalil_nor/output
```

```
1457e+06 tps_per_gpu=214522 mfu=43.4033
[2024-11-19 17:14:48.550933][INFO][trainer.py:827] - ['prompt']: 'What is an LLM?'
[2024-11-19 17:14:48.553596][INFO][trainer.py:831] - ['response']:

What is an LLM?

CLARENCE:
Then cursed stake the cause of thy words,
Thou hast thou hast the one that would death be heaven
I might have been as those son Gloucent's eyes,
When we Lokss in the seat that I did weep from it offence
The house of Lancaster hath been as the o
[2024-11-19 17:15:26.927638][INFO][trainer.py:762] - Saving checkpoint to: /home/khalil_nor/outputs/runs/pytorch/DDf
[2024-11-19 17:15:26.929538][INFO][trainer.py:763] - Saving model to: /home/khalil_nor/outputs/runs/pytorch/DDP/2024
[2024-11-19 17:15:29.689475][INFO][configs.py:141] - Appending /home/khalil_nor/outputs/runs/pytorch/DDP/2024-11-19/
c/ckpts/checkpoints.log
```

[2024-11-19 17:17:12.922100][**INFO**][__main__.py:120] — ['*response*']:

What is an LLM?

CLARENCE:
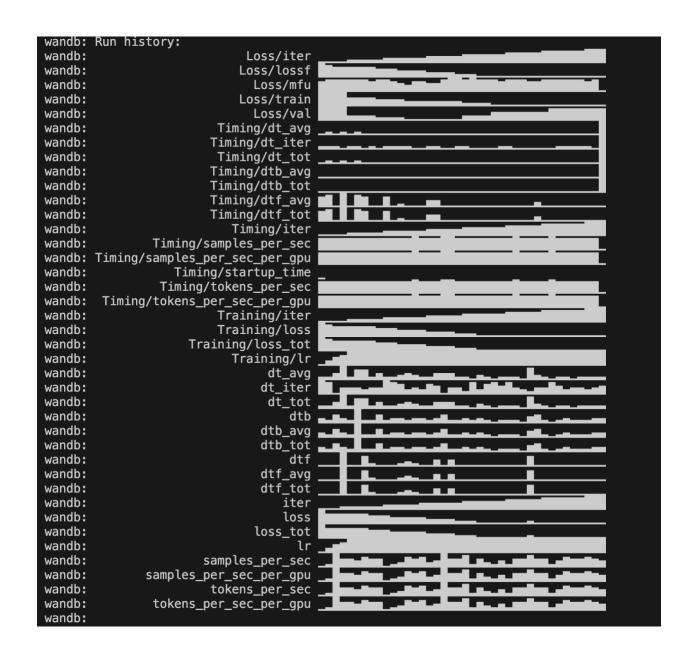Repetiment me to speak, my lord.

WARWICK:
The duke of Norfolk, gentle Warwick; there be I'll place.

CLARENCE:
For may to the storm, petual of this from me:
The duke is to king, which now the harm of the Tower.

WARWICK:
Why then the King of F

[2024-11-19 17:17:12.924298][**INFO**][*trainer.py*:762] — Saving checkpoint to: /h
[2024-11-19 17:17:12.924879][**INFO**][*trainer.py*:763] — Saving model to: /home/k

```
wandb: Run history:
wandb:                        Loss/iter  ▁▁▁▁▁▁▁▁▁▁██████████████████
wandb:                       Loss/lossf  ▁▁▁▁▁▁▁▁▁▁██████████████████
wandb:                         Loss/mfu  ▁▁▁▁▁██████████████████████
wandb:                       Loss/train  ▁▁▁▁▁██████████████████████
wandb:                         Loss/val  ▁▁▁▁▁██████████████████████
wandb:                    Timing/dt_avg  ▁▁▁▁▁██████████████████████
wandb:                   Timing/dt_iter  ▁▁▁▁▁▁▁▁██████████████████
wandb:                    Timing/dt_tot  ▁▁▁▁▁▁██████████████████████
wandb:                   Timing/dtb_avg  ███████████████████████████
wandb:                   Timing/dtb_tot  ███████████████████████████
wandb:                   Timing/dtf_avg  ▁█▁█▁█▁▁▁█▁▁▁▁▁▁▁▁▁█▁▁▁▁▁▁
wandb:                   Timing/dtf_tot  ▁█▁█▁█▁▁▁█▁▁▁▁▁▁▁▁▁█▁▁▁▁▁▁
wandb:                     Timing/iter  ▁▁▁▁▁▁▁▁▁▁██████████████████
wandb:        Timing/samples_per_sec  █████████████████████████████
wandb:    Timing/samples_per_sec_per_gpu  █████████████████████████
wandb:            Timing/startup_time  ▁
wandb:          Timing/tokens_per_sec  █████████████████████████████
wandb:   Timing/tokens_per_sec_per_gpu  ████████████████████████████
wandb:                    Training/iter  ▁▁▁▁▁▁▁▁▁▁██████████████████
wandb:                    Training/loss  ▁▁▁▁▁██████████████████████
wandb:               Training/loss_tot  ▁▁▁▁▁██████████████████████
wandb:                     Training/lr  ▁▁▁▁▁▁▁▁▁▁██████████████████
wandb:                           dt_avg  ▁█▁█▁█▁█▁█▁█▁█▁█▁█▁█▁█▁█▁█
wandb:                          dt_iter  ▁█▁█▁█▁█▁█▁█▁█▁█▁█▁█▁█▁█▁█
wandb:                           dt_tot  ▁█▁█▁█▁█▁█▁█▁█▁█▁█▁█▁█▁█▁█
wandb:                              dtb  ▁██▁█▁█▁█▁█▁█▁█▁█▁█▁█▁█▁█
wandb:                          dtb_avg  ▁██▁█▁█▁█▁█▁█▁█▁█▁█▁█▁█▁█
wandb:                          dtb_tot  ▁██▁█▁█▁█▁█▁█▁█▁█▁█▁█▁█▁█
wandb:                              dtf  ▁█▁▁█▁▁█▁▁█▁▁█▁▁█▁▁█▁▁█▁▁
wandb:                          dtf_avg  ▁█▁▁█▁▁█▁▁█▁▁█▁▁█▁▁█▁▁█▁▁
wandb:                          dtf_tot  ▁█▁▁█▁▁█▁▁█▁▁█▁▁█▁▁█▁▁█▁▁
wandb:                             iter  ▁▁▁▁▁▁▁▁▁▁██████████████████
wandb:                             loss  ▁▁▁▁▁██████████████████████
wandb:                         loss_tot  ▁▁▁▁▁██████████████████████
wandb:                               lr  ▁▁▁▁▁▁▁▁▁▁██████████████████
wandb:                  samples_per_sec  ▁█▁█▁█▁█▁█▁█▁█▁█▁█▁█▁█▁█▁█
wandb:          samples_per_sec_per_gpu  ▁█▁█▁█▁█▁█▁█▁█▁█▁█▁█▁█▁█▁█
wandb:                   tokens_per_sec  ▁█▁█▁█▁█▁█▁█▁█▁█▁█▁█▁█▁█▁█
wandb:           tokens_per_sec_per_gpu  ▁█▁█▁█▁█▁█▁█▁█▁█▁█▁█▁█▁█▁█
wandb:
```

```
wandb:
wandb: Run summary:
wandb:                               Loss/iter 1000
wandb:                              Loss/lossf 0.0322
wandb:                                Loss/mfu 41.73478
wandb:                              Loss/train 0.03877
wandb:                                Loss/val 3.83037
wandb:                           Timing/dt_avg 0.137
wandb:                          Timing/dt_iter 0.55362
wandb:                           Timing/dt_tot 0.27399
wandb:                          Timing/dtb_avg 0.26726
wandb:                          Timing/dtb_tot 0.26726
wandb:                          Timing/dtf_avg 0.00673
wandb:                          Timing/dtf_tot 0.00673
wandb:                             Timing/iter 999
wandb:                 Timing/samples_per_sec 14.45026
wandb:         Timing/samples_per_sec_per_gpu 1.80628
wandb:                    Timing/startup_time 12.03712
wandb:                  Timing/tokens_per_sec 947012.00893
wandb:          Timing/tokens_per_sec_per_gpu 118376.50112
wandb:                           Training/iter 999
wandb:                           Training/loss 0.0322
wandb:                       Training/loss_tot 0.0322
wandb:                             Training/lr 0.0006
wandb:                                  dt_avg 0.137
wandb:                                 dt_iter 0.55362
wandb:                                  dt_tot 0.27399
wandb:                                     dtb 0.26726
wandb:                                 dtb_avg 0.26726
wandb:                                 dtb_tot 0.26726
wandb:                                     dtf 0.00673
wandb:                                 dtf_avg 0.00673
wandb:                                 dtf_tot 0.00673
wandb:                                    iter 999
wandb:                                    loss 0.0322
wandb:                                loss_tot 0.0322
wandb:                                      lr 0.0006
wandb:                         samples_per_sec 14.45026
wandb:                 samples_per_sec_per_gpu 1.80628
wandb:                          tokens_per_sec 947012.00893
wandb:                  tokens_per_sec_per_gpu 118376.50112
```

View project at: https://wandb.ai/khalil-nor-northeastern-university/WordPlay