python -u engine.py --corpus=dyda --mode=train --gpu=0,1 --batch\_size=10 --batch\_size\_val=10 --epochs=100 --lr=0.0001 --nlayer=2 --chunk\_size=0 --dropout=0.5 --nfinetune=1 --speaker\_info=emb\_cls --topic\_info=emb\_cls --nclass=4 --emb\_batch=0

Namespace(corpus='dyda', mode='train', nclass=4, batch\_size=10, batch\_size\_val=10, emb\_batch=0, epochs=100, gpu='0,1', lr=0.0001, nlayer=2, chunk\_size=0, dropout=0.5, speaker\_info='emb\_cls', topic\_info='emb\_cls', nfinetune=1, seed=0)

Tokenizing train....

/usr/local/lib/python3.11/dist-packages/huggingface\_hub/file\_download.py:896: FutureWarning: `resume\_download` is deprecated and will be removed in version 1.0.0. Downloads always resume when possible. If you want to force a new download, use `force\_download=True`.

warnings.warn(

Done

Tokenizing val....

/usr/local/lib/python3.11/dist-packages/huggingface\_hub/file\_download.py:896: FutureWarning: `resume\_download` is deprecated and will be removed in version 1.0.0. Downloads always resume when possible. If you want to force a new download, use `force\_download=True`.

warnings.warn(

Done

Tokenizing test....

/usr/local/lib/python3.11/dist-packages/huggingface\_hub/file\_download.py:896: FutureWarning: `resume\_download` is deprecated and will be removed in version 1.0.0. Downloads always resume when possible. If you want to force a new download, use `force\_download=True`.

warnings.warn(

Done

Done

Let's use 1 GPUs!

Initializing model....

/usr/local/lib/python3.11/dist-packages/huggingface\_hub/file\_download.py:896: FutureWarning: `resume\_download` is deprecated and will be removed in version 1.0.0. Downloads always resume when possible. If you want to force a new download, use `force\_download=True`.

warnings.warn(

Some weights of the model checkpoint at roberta-base were not used when initializing RobertaModel: ['lm\_head.dense.weight', 'lm\_head.layer\_norm.weight', 'lm\_head.layer\_norm.bias', 'lm\_head.dense.bias', 'lm\_head.bias']

- This IS expected if you are initializing RobertaModel from the checkpoint of a model trained on another task or with another architecture (e.g. initializing a BertForSequenceClassification model from a BertForPreTraining model).

- This IS NOT expected if you are initializing RobertaModel from the checkpoint of a model that you expect to be exactly identical (initializing a BertForSequenceClassification model from a BertForSequenceClassification model).

2025-05-07 17:01:36.904322: E external/local\_xla/xla/stream\_executor/cuda/cuda\_fft.cc:477] Unable to register cuFFT factory: Attempting to register factory for plugin cuFFT when one has already been registered

WARNING: All log messages before absl::InitializeLog() is called are written to STDERR

E0000 00:00:1746637296.922961 9501 cuda\_dnn.cc:8310] Unable to register cuDNN factory: Attempting to register factory for plugin cuDNN when one has already been registered

E0000 00:00:1746637296.928357 9501 cuda\_blas.cc:1418] Unable to register cuBLAS factory: Attempting to register factory for plugin cuBLAS when one has already been registered

/usr/local/lib/python3.11/dist-packages/transformers/optimization.py:391: FutureWarning: This implementation of AdamW is deprecated and will be removed in a future version. Use the PyTorch implementation torch.optim.AdamW instead, or set `no\_deprecation\_warning=True` to disable this warning

warnings.warn(

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Epoch: 1\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Batch: 1/1112 loss: 1.447 loss\_act:1.447

Batch: 56/1112 loss: 1.273 loss\_act:1.273

Batch: 111/1112 loss: 0.938 loss\_act:0.938

Batch: 166/1112 loss: 0.853 loss\_act:0.853

Batch: 221/1112 loss: 0.643 loss\_act:0.643

Batch: 276/1112 loss: 0.630 loss\_act:0.630

Batch: 331/1112 loss: 0.742 loss\_act:0.742

Batch: 386/1112 loss: 0.558 loss\_act:0.558

Batch: 441/1112 loss: 0.705 loss\_act:0.705

Batch: 496/1112 loss: 0.515 loss\_act:0.515

Batch: 551/1112 loss: 0.351 loss\_act:0.351

Batch: 606/1112 loss: 0.549 loss\_act:0.549

Batch: 661/1112 loss: 0.373 loss\_act:0.373

Batch: 716/1112 loss: 0.297 loss\_act:0.297

Batch: 771/1112 loss: 0.419 loss\_act:0.419

Batch: 826/1112 loss: 0.493 loss\_act:0.493

Batch: 881/1112 loss: 0.293 loss\_act:0.293

Batch: 936/1112 loss: 0.363 loss\_act:0.363

Batch: 991/1112 loss: 0.585 loss\_act:0.585

Batch: 1046/1112 loss: 0.332 loss\_act:0.332

Batch: 1101/1112 loss: 0.271 loss\_act:0.271

Batch: 1112/1112 loss: 0.377 loss\_act:0.377

Epoch 1 Train Loss: 0.571 Val Acc: 0.822 Test Acc: 0.854

Best Epoch: 1 Best Epoch Val Acc: 0.822 Best Epoch Test Acc: 0.854, Best Test Acc: 0.854

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Epoch: 2\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Batch: 1/1112 loss: 0.319 loss\_act:0.319

Batch: 56/1112 loss: 0.432 loss\_act:0.432

Batch: 111/1112 loss: 0.493 loss\_act:0.493

Batch: 166/1112 loss: 0.365 loss\_act:0.365

Batch: 221/1112 loss: 0.442 loss\_act:0.442

Batch: 276/1112 loss: 0.375 loss\_act:0.375

Batch: 331/1112 loss: 0.607 loss\_act:0.607

Batch: 386/1112 loss: 0.219 loss\_act:0.219

Batch: 441/1112 loss: 0.338 loss\_act:0.338

Batch: 496/1112 loss: 0.299 loss\_act:0.299

Batch: 551/1112 loss: 0.472 loss\_act:0.472

Batch: 606/1112 loss: 0.607 loss\_act:0.607

Batch: 661/1112 loss: 0.343 loss\_act:0.343

Batch: 716/1112 loss: 0.342 loss\_act:0.342

Batch: 771/1112 loss: 0.267 loss\_act:0.267

Batch: 826/1112 loss: 0.306 loss\_act:0.306

Batch: 881/1112 loss: 0.414 loss\_act:0.414

Batch: 936/1112 loss: 0.277 loss\_act:0.277

Batch: 991/1112 loss: 0.385 loss\_act:0.385

Batch: 1046/1112 loss: 0.397 loss\_act:0.397

Batch: 1101/1112 loss: 0.319 loss\_act:0.319

Batch: 1112/1112 loss: 0.716 loss\_act:0.716

Epoch 2 Train Loss: 0.411 Val Acc: 0.836 Test Acc: 0.859

Best Epoch: 2 Best Epoch Val Acc: 0.836 Best Epoch Test Acc: 0.859, Best Test Acc: 0.859

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Epoch: 3\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Batch: 1/1112 loss: 0.423 loss\_act:0.423

Batch: 56/1112 loss: 0.324 loss\_act:0.324

Batch: 111/1112 loss: 0.462 loss\_act:0.462

Batch: 166/1112 loss: 0.263 loss\_act:0.263

Batch: 221/1112 loss: 0.359 loss\_act:0.359

Batch: 276/1112 loss: 0.366 loss\_act:0.366

Batch: 331/1112 loss: 0.550 loss\_act:0.550

Batch: 386/1112 loss: 0.547 loss\_act:0.547

Batch: 441/1112 loss: 0.311 loss\_act:0.311

Batch: 496/1112 loss: 0.478 loss\_act:0.478

Batch: 551/1112 loss: 0.328 loss\_act:0.328

Batch: 606/1112 loss: 0.365 loss\_act:0.365

Batch: 661/1112 loss: 0.413 loss\_act:0.413

Batch: 716/1112 loss: 0.408 loss\_act:0.408

Batch: 771/1112 loss: 0.210 loss\_act:0.210

Batch: 826/1112 loss: 0.596 loss\_act:0.596

Batch: 881/1112 loss: 0.287 loss\_act:0.287

Batch: 936/1112 loss: 0.318 loss\_act:0.318

Batch: 991/1112 loss: 0.791 loss\_act:0.791

Batch: 1046/1112 loss: 0.312 loss\_act:0.312

Batch: 1101/1112 loss: 0.510 loss\_act:0.510

Batch: 1112/1112 loss: 0.729 loss\_act:0.729

Epoch 3 Train Loss: 0.387 Val Acc: 0.840 Test Acc: 0.858

Best Epoch: 3 Best Epoch Val Acc: 0.840 Best Epoch Test Acc: 0.858, Best Test Acc: 0.859

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Epoch: 4\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Batch: 1/1112 loss: 0.218 loss\_act:0.218

Batch: 56/1112 loss: 0.296 loss\_act:0.296

Batch: 111/1112 loss: 0.280 loss\_act:0.280

Batch: 166/1112 loss: 0.259 loss\_act:0.259