Roberta mrda  
  
python -u engine.py --corpus=mrda --mode=train --gpu=0,1 --batch\_size=4 --batch\_size\_val=2 --epochs=100 --lr=0.0001 --nlayer=1 --chunk\_size=350 --dropout=0.5 --nfinetune=1 --speaker\_info=none --topic\_info=none --nclass=5 --emb\_batch=256

Namespace(corpus='mrda', mode='train', nclass=5, batch\_size=4, batch\_size\_val=2, emb\_batch=256, epochs=100, gpu='0,1', lr=0.0001, nlayer=1, chunk\_size=350, dropout=0.5, speaker\_info='none', topic\_info='none', 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(

tokenizer\_config.json: 100% 25.0/25.0 [00:00<00:00, 119kB/s]

config.json: 100% 481/481 [00:00<00:00, 3.74MB/s]

vocab.json: 100% 899k/899k [00:00<00:00, 14.7MB/s]

merges.txt: 100% 456k/456k [00:00<00:00, 6.58MB/s]

tokenizer.json: 100% 1.36M/1.36M [00:00<00:00, 12.9MB/s]

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(

/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(

Xet Storage is enabled for this repo, but the 'hf\_xet' package is not installed. Falling back to regular HTTP download. For better performance, install the package with: `pip install huggingface\_hub[hf\_xet]` or `pip install hf\_xet`

model.safetensors: 100% 499M/499M [00:01<00:00, 257MB/s]

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

/usr/local/lib/python3.11/dist-packages/torch/nn/modules/rnn.py:123: UserWarning: dropout option adds dropout after all but last recurrent layer, so non-zero dropout expects num\_layers greater than 1, but got dropout=0.5 and num\_layers=1

warnings.warn(

2025-05-07 16:30:07.545396: 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 00:00:1746635407.563652 1512 cuda\_dnn.cc:8310] Unable to register cuDNN factory: Attempting to register factory for plugin cuDNN when one has already been registered

E0000 00:00:1746635407.568978 1512 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: 1/60 loss: 1.604 loss\_act:1.604

Batch: 4/60 loss: 1.324 loss\_act:1.324

Batch: 7/60 loss: 1.223 loss\_act:1.223

Batch: 10/60 loss: 1.332 loss\_act:1.332

Batch: 13/60 loss: 1.168 loss\_act:1.168

Batch: 16/60 loss: 1.134 loss\_act:1.134

Batch: 19/60 loss: 1.201 loss\_act:1.201

Batch: 22/60 loss: 1.142 loss\_act:1.142

Batch: 25/60 loss: 1.107 loss\_act:1.107

Batch: 28/60 loss: 1.012 loss\_act:1.012

Batch: 31/60 loss: 0.949 loss\_act:0.949

Batch: 34/60 loss: 0.972 loss\_act:0.972

Batch: 37/60 loss: 0.937 loss\_act:0.937

Batch: 40/60 loss: 0.858 loss\_act:0.858

Batch: 43/60 loss: 0.842 loss\_act:0.842

Batch: 46/60 loss: 0.751 loss\_act:0.751

Batch: 49/60 loss: 0.687 loss\_act:0.687

Batch: 52/60 loss: 0.631 loss\_act:0.631

Batch: 55/60 loss: 0.540 loss\_act:0.540

Batch: 58/60 loss: 0.590 loss\_act:0.590

Batch: 60/60 loss: 0.542 loss\_act:0.542

Epoch 1 Train Loss: 0.963 Val Acc: 0.836 Test Acc: 0.855

Best Epoch: 1 Best Epoch Val Acc: 0.836 Best Epoch Test Acc: 0.855, Best Test Acc: 0.855

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

Batch: 1/60 loss: 0.504 loss\_act:0.504

Batch: 4/60 loss: 0.368 loss\_act:0.368

Batch: 7/60 loss: 0.328 loss\_act:0.328

Batch: 10/60 loss: 0.417 loss\_act:0.417

Batch: 13/60 loss: 0.352 loss\_act:0.352

Batch: 16/60 loss: 0.375 loss\_act:0.375

Batch: 19/60 loss: 0.343 loss\_act:0.343

Batch: 22/60 loss: 0.312 loss\_act:0.312

Batch: 25/60 loss: 0.348 loss\_act:0.348

Batch: 28/60 loss: 0.365 loss\_act:0.365

Batch: 31/60 loss: 0.373 loss\_act:0.373

Batch: 34/60 loss: 0.302 loss\_act:0.302

Batch: 37/60 loss: 0.370 loss\_act:0.370

Batch: 40/60 loss: 0.300 loss\_act:0.300

Batch: 43/60 loss: 0.393 loss\_act:0.393

Batch: 46/60 loss: 0.391 loss\_act:0.391

Batch: 49/60 loss: 0.369 loss\_act:0.369

Batch: 52/60 loss: 0.335 loss\_act:0.335

Batch: 55/60 loss: 0.290 loss\_act:0.290

Batch: 58/60 loss: 0.377 loss\_act:0.377

Batch: 60/60 loss: 0.314 loss\_act:0.314

Epoch 2 Train Loss: 0.358 Val Acc: 0.883 Test Acc: 0.895

Best Epoch: 2 Best Epoch Val Acc: 0.883 Best Epoch Test Acc: 0.895, Best Test Acc: 0.895

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

Batch: 1/60 loss: 0.308 loss\_act:0.308

Batch: 4/60 loss: 0.274 loss\_act:0.274

Batch: 7/60 loss: 0.286 loss\_act:0.286

Batch: 10/60 loss: 0.317 loss\_act:0.317

Batch: 13/60 loss: 0.292 loss\_act:0.292

Batch: 16/60 loss: 0.318 loss\_act:0.318

Batch: 19/60 loss: 0.292 loss\_act:0.292

Batch: 22/60 loss: 0.337 loss\_act:0.337

Batch: 25/60 loss: 0.216 loss\_act:0.216

Batch: 28/60 loss: 0.324 loss\_act:0.324

Batch: 31/60 loss: 0.253 loss\_act:0.253

Batch: 34/60 loss: 0.274 loss\_act:0.274

Batch: 37/60 loss: 0.250 loss\_act:0.250

Batch: 40/60 loss: 0.286 loss\_act:0.286

Batch: 43/60 loss: 0.340 loss\_act:0.340

Batch: 46/60 loss: 0.314 loss\_act:0.314

Batch: 49/60 loss: 0.254 loss\_act:0.254

Batch: 52/60 loss: 0.299 loss\_act:0.299

Batch: 55/60 loss: 0.329 loss\_act:0.329

Batch: 58/60 loss: 0.276 loss\_act:0.276

Batch: 60/60 loss: 0.298 loss\_act:0.298

Epoch 3 Train Loss: 0.301 Val Acc: 0.883 Test Acc: 0.899

Best Epoch: 2 Best Epoch Val Acc: 0.883 Best Epoch Test Acc: 0.895, Best Test Acc: 0.899

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

Batch: 1/60 loss: 0.256 loss\_act:0.256

Batch: 4/60 loss: 0.290 loss\_act:0.290

Batch: 7/60 loss: 0.254 loss\_act:0.254

Batch: 10/60 loss: 0.268 loss\_act:0.268

Batch: 13/60 loss: 0.250 loss\_act:0.250

Batch: 16/60 loss: 0.322 loss\_act:0.322

Batch: 19/60 loss: 0.312 loss\_act:0.312

Batch: 22/60 loss: 0.293 loss\_act:0.293

Batch: 25/60 loss: 0.259 loss\_act:0.259

Batch: 28/60 loss: 0.291 loss\_act:0.291

Batch: 31/60 loss: 0.236 loss\_act:0.236

Batch: 34/60 loss: 0.285 loss\_act:0.285

Batch: 37/60 loss: 0.274 loss\_act:0.274

Batch: 40/60 loss: 0.247 loss\_act:0.247

Batch: 43/60 loss: 0.319 loss\_act:0.319

Batch: 46/60 loss: 0.307 loss\_act:0.307

Batch: 49/60 loss: 0.266 loss\_act:0.266

Batch: 52/60 loss: 0.335 loss\_act:0.335

Batch: 55/60 loss: 0.321 loss\_act:0.321

Batch: 58/60 loss: 0.309 loss\_act:0.309

Batch: 60/60 loss: 0.284 loss\_act:0.284

Epoch 4 Train Loss: 0.284 Val Acc: 0.883 Test Acc: 0.900

Best Epoch: 4 Best Epoch Val Acc: 0.883 Best Epoch Test Acc: 0.900, Best Test Acc: 0.900

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Epoch: 5\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Batch: 1/60 loss: 0.339 loss\_act:0.339

Batch: 4/60 loss: 0.182 loss\_act:0.182

Batch: 7/60 loss: 0.253 loss\_act:0.253

Batch: 10/60 loss: 0.281 loss\_act:0.281

Batch: 13/60 loss: 0.330 loss\_act:0.330

Batch: 16/60 loss: 0.262 loss\_act:0.262

Batch: 19/60 loss: 0.286 loss\_act:0.286

Batch: 22/60 loss: 0.347 loss\_act:0.347

Batch: 25/60 loss: 0.293 loss\_act:0.293

Batch: 28/60 loss: 0.254 loss\_act:0.254

Batch: 31/60 loss: 0.272 loss\_act:0.272

Batch: 34/60 loss: 0.277 loss\_act:0.277

Batch: 37/60 loss: 0.327 loss\_act:0.327

Batch: 40/60 loss: 0.297 loss\_act:0.297

Batch: 43/60 loss: 0.258 loss\_act:0.258

Batch: 46/60 loss: 0.275 loss\_act:0.275

Batch: 49/60 loss: 0.244 loss\_act:0.244

Batch: 52/60 loss: 0.238 loss\_act:0.238

Batch: 55/60 loss: 0.290 loss\_act:0.290

Batch: 58/60 loss: 0.273 loss\_act:0.273

Batch: 60/60 loss: 0.305 loss\_act:0.305

Epoch 5 Train Loss: 0.277 Val Acc: 0.884 Test Acc: 0.900

Best Epoch: 5 Best Epoch Val Acc: 0.884 Best Epoch Test Acc: 0.900, Best Test Acc: 0.900

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Epoch: 6\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Batch: 1/60 loss: 0.215 loss\_act:0.215

Batch: 4/60 loss: 0.305 loss\_act:0.305

Batch: 7/60 loss: 0.239 loss\_act:0.239

Batch: 10/60 loss: 0.298 loss\_act:0.298

Batch: 13/60 loss: 0.288 loss\_act:0.288

Batch: 16/60 loss: 0.279 loss\_act:0.279

Batch: 19/60 loss: 0.276 loss\_act:0.276

Batch: 22/60 loss: 0.310 loss\_act:0.310

Batch: 25/60 loss: 0.291 loss\_act:0.291

Batch: 28/60 loss: 0.295 loss\_act:0.295

Batch: 31/60 loss: 0.275 loss\_act:0.275

Batch: 34/60 loss: 0.281 loss\_act:0.281

Batch: 37/60 loss: 0.257 loss\_act:0.257

Batch: 40/60 loss: 0.236 loss\_act:0.236

Batch: 43/60 loss: 0.278 loss\_act:0.278

Batch: 46/60 loss: 0.298 loss\_act:0.298

Batch: 49/60 loss: 0.341 loss\_act:0.341

Batch: 52/60 loss: 0.241 loss\_act:0.241

Batch: 55/60 loss: 0.249 loss\_act:0.249

Batch: 58/60 loss: 0.252 loss\_act:0.252

Batch: 60/60 loss: 0.274 loss\_act:0.274

Epoch 6 Train Loss: 0.270 Val Acc: 0.885 Test Acc: 0.903

Best Epoch: 6 Best Epoch Val Acc: 0.885 Best Epoch Test Acc: 0.903, Best Test Acc: 0.903

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Epoch: 7\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Batch: 1/60 loss: 0.282 loss\_act:0.282

Batch: 4/60 loss: 0.278 loss\_act:0.278

Batch: 7/60 loss: 0.195 loss\_act:0.195

Batch: 10/60 loss: 0.319 loss\_act:0.319

Batch: 13/60 loss: 0.282 loss\_act:0.282

Batch: 16/60 loss: 0.269 loss\_act:0.269

Batch: 19/60 loss: 0.306 loss\_act:0.306

Batch: 22/60 loss: 0.258 loss\_act:0.258

Batch: 25/60 loss: 0.286 loss\_act:0.286

Batch: 28/60 loss: 0.329 loss\_act:0.329

Batch: 31/60 loss: 0.266 loss\_act:0.266

Batch: 34/60 loss: 0.216 loss\_act:0.216

Batch: 37/60 loss: 0.222 loss\_act:0.222

Batch: 40/60 loss: 0.261 loss\_act:0.261

Batch: 43/60 loss: 0.269 loss\_act:0.269

Batch: 46/60 loss: 0.280 loss\_act:0.280

Batch: 49/60 loss: 0.252 loss\_act:0.252

Batch: 52/60 loss: 0.278 loss\_act:0.278

Batch: 55/60 loss: 0.244 loss\_act:0.244

Batch: 58/60 loss: 0.274 loss\_act:0.274

Batch: 60/60 loss: 0.252 loss\_act:0.252

Epoch 7 Train Loss: 0.265 Val Acc: 0.884 Test Acc: 0.900

Best Epoch: 6 Best Epoch Val Acc: 0.885 Best Epoch Test Acc: 0.903, Best Test Acc: 0.903

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Epoch: 8\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Batch: 1/60 loss: 0.272 loss\_act:0.272

Batch: 4/60 loss: 0.249 loss\_act:0.249

Batch: 7/60 loss: 0.269 loss\_act:0.269

Batch: 10/60 loss: 0.272 loss\_act:0.272

Batch: 13/60 loss: 0.261 loss\_act:0.261

Batch: 16/60 loss: 0.303 loss\_act:0.303

Batch: 19/60 loss: 0.292 loss\_act:0.292

Batch: 22/60 loss: 0.295 loss\_act:0.295

Batch: 25/60 loss: 0.267 loss\_act:0.267

Batch: 28/60 loss: 0.268 loss\_act:0.268

Batch: 31/60 loss: 0.241 loss\_act:0.241

Batch: 34/60 loss: 0.307 loss\_act:0.307

Batch: 37/60 loss: 0.229 loss\_act:0.229

Batch: 40/60 loss: 0.279 loss\_act:0.279

Batch: 43/60 loss: 0.291 loss\_act:0.291

Batch: 46/60 loss: 0.257 loss\_act:0.257

Batch: 49/60 loss: 0.264 loss\_act:0.264

Batch: 52/60 loss: 0.255 loss\_act:0.255

Batch: 55/60 loss: 0.265 loss\_act:0.265

Batch: 58/60 loss: 0.209 loss\_act:0.209

Batch: 60/60 loss: 0.245 loss\_act:0.245

Epoch 8 Train Loss: 0.263 Val Acc: 0.886 Test Acc: 0.902

Best Epoch: 8 Best Epoch Val Acc: 0.886 Best Epoch Test Acc: 0.902, Best Test Acc: 0.903

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Epoch: 9\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Batch: 1/60 loss: 0.206 loss\_act:0.206

Batch: 4/60 loss: 0.283 loss\_act:0.283

Batch: 7/60 loss: 0.201 loss\_act:0.201

Batch: 10/60 loss: 0.287 loss\_act:0.287

Batch: 13/60 loss: 0.269 loss\_act:0.269

Batch: 16/60 loss: 0.255 loss\_act:0.255

Batch: 19/60 loss: 0.270 loss\_act:0.270

Batch: 22/60 loss: 0.305 loss\_act:0.305

Batch: 25/60 loss: 0.249 loss\_act:0.249

Batch: 28/60 loss: 0.309 loss\_act:0.309

Batch: 31/60 loss: 0.289 loss\_act:0.289

Batch: 34/60 loss: 0.253 loss\_act:0.253

Batch: 37/60 loss: 0.262 loss\_act:0.262

Batch: 40/60 loss: 0.239 loss\_act:0.239

Batch: 43/60 loss: 0.268 loss\_act:0.268

Batch: 46/60 loss: 0.318 loss\_act:0.318

Batch: 49/60 loss: 0.268 loss\_act:0.268

Batch: 52/60 loss: 0.258 loss\_act:0.258

Batch: 55/60 loss: 0.210 loss\_act:0.210

Batch: 58/60 loss: 0.306 loss\_act:0.306

Batch: 60/60 loss: 0.239 loss\_act:0.239

Epoch 9 Train Loss: 0.261 Val Acc: 0.884 Test Acc: 0.902

Best Epoch: 8 Best Epoch Val Acc: 0.886 Best Epoch Test Acc: 0.902, Best Test Acc: 0.903

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Epoch: 10\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Batch: 1/60 loss: 0.249 loss\_act:0.249

Batch: 4/60 loss: 0.264 loss\_act:0.264

Batch: 7/60 loss: 0.236 loss\_act:0.236

Batch: 10/60 loss: 0.236 loss\_act:0.236

Batch: 13/60 loss: 0.242 loss\_act:0.242

Batch: 16/60 loss: 0.248 loss\_act:0.248

Batch: 19/60 loss: 0.284 loss\_act:0.284

Batch: 22/60 loss: 0.272 loss\_act:0.272

Batch: 25/60 loss: 0.232 loss\_act:0.232

Batch: 28/60 loss: 0.254 loss\_act:0.254

Batch: 31/60 loss: 0.246 loss\_act:0.246

Batch: 34/60 loss: 0.231 loss\_act:0.231

Batch: 37/60 loss: 0.233 loss\_act:0.233

Batch: 40/60 loss: 0.203 loss\_act:0.203

Batch: 43/60 loss: 0.250 loss\_act:0.250

Batch: 46/60 loss: 0.260 loss\_act:0.260

Batch: 49/60 loss: 0.223 loss\_act:0.223

Batch: 52/60 loss: 0.250 loss\_act:0.250

Batch: 55/60 loss: 0.313 loss\_act:0.313

Batch: 58/60 loss: 0.271 loss\_act:0.271

Batch: 60/60 loss: 0.208 loss\_act:0.208

Epoch 10 Train Loss: 0.259 Val Acc: 0.880 Test Acc: 0.900

Best Epoch: 8 Best Epoch Val Acc: 0.886 Best Epoch Test Acc: 0.902, Best Test Acc: 0.903

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Epoch: 11\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Batch: 1/60 loss: 0.310 loss\_act:0.310

Batch: 4/60 loss: 0.246 loss\_act:0.246

Batch: 7/60 loss: 0.232 loss\_act:0.232

Batch: 10/60 loss: 0.217 loss\_act:0.217

Batch: 13/60 loss: 0.190 loss\_act:0.190

Batch: 16/60 loss: 0.341 loss\_act:0.341

Batch: 19/60 loss: 0.304 loss\_act:0.304

Batch: 22/60 loss: 0.261 loss\_act:0.261

Batch: 25/60 loss: 0.221 loss\_act:0.221

Batch: 28/60 loss: 0.234 loss\_act:0.234

Batch: 31/60 loss: 0.279 loss\_act:0.279

Batch: 34/60 loss: 0.283 loss\_act:0.283

Batch: 37/60 loss: 0.287 loss\_act:0.287

Batch: 40/60 loss: 0.285 loss\_act:0.285

Batch: 43/60 loss: 0.231 loss\_act:0.231

Batch: 46/60 loss: 0.232 loss\_act:0.232

Batch: 49/60 loss: 0.262 loss\_act:0.262

Batch: 52/60 loss: 0.232 loss\_act:0.232

Batch: 55/60 loss: 0.261 loss\_act:0.261

Batch: 58/60 loss: 0.304 loss\_act:0.304

Batch: 60/60 loss: 0.260 loss\_act:0.260

Epoch 11 Train Loss: 0.256 Val Acc: 0.889 Test Acc: 0.906

Best Epoch: 11 Best Epoch Val Acc: 0.889 Best Epoch Test Acc: 0.906, Best Test Acc: 0.906

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Epoch: 12\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Batch: 1/60 loss: 0.247 loss\_act:0.247

Batch: 4/60 loss: 0.232 loss\_act:0.232

Batch: 7/60 loss: 0.253 loss\_act:0.253

Batch: 10/60 loss: 0.282 loss\_act:0.282

Batch: 13/60 loss: 0.258 loss\_act:0.258

Batch: 16/60 loss: 0.248 loss\_act:0.248

Batch: 19/60 loss: 0.245 loss\_act:0.245

Batch: 22/60 loss: 0.215 loss\_act:0.215

Batch: 25/60 loss: 0.277 loss\_act:0.277

Batch: 28/60 loss: 0.213 loss\_act:0.213

Batch: 31/60 loss: 0.274 loss\_act:0.274

Batch: 34/60 loss: 0.314 loss\_act:0.314

Batch: 37/60 loss: 0.219 loss\_act:0.219

Colab paid products - Cancel contracts here