Command to run NAR training:

(pytorch_p36) ubuntu@ip-10-0-0-148:~/fairseq$ fairseq-train   data-bin/iwslt14.tokenized.de-en   --save-dir checkpoints   --ddp-backend=legacy_ddp   --task translation_lev   --criterion nat_loss   --arch nacrf_transformer   --noise full_mask   --share-all-embeddings   --optimizer adam --adam-betas '(0.9,0.98)'   --lr 0.0005 --lr-scheduler inverse_sqrt   --stop-min-lr '1e-09' --warmup-updates 10000   --warmup-init-lr '1e-07' --label-smoothing 0.1   --dropout 0.3 --weight-decay 0.01   --decoder-learned-pos   --encoder-learned-pos   --pred-length-offset   --length-loss-factor 0.1   --word-ins-loss-factor 0.5   --crf-lowrank-approx 32   --crf-beam-approx 64   --apply-bert-init   --log-format 'simple' --log-interval 100   --fixed-validation-seed 7   --max-tokens 8000   --save-interval-updates 10000   --max-update 300000

Output:

```
2021-03-27 19:59:39 | INFO | fairseq_cli.train | task: TranslationLevenshteinTask
2021-03-27 19:59:39 | INFO | fairseq_cli.train | model: NACRFTransformerModel
2021-03-27 19:59:39 | INFO | fairseq_cli.train | criterion: LabelSmoothedDualImitationCriterion
2021-03-27 19:59:39 | INFO | fairseq_cli.train | num. model params: 51,169,792 (num. trained: 51,169,792)
2021-03-27 20:00:52 | INFO | fairseq.trainer | detected shared parameter: encoder.embed_tokens.weight <- decoder.embed_tokens.weight
2021-03-27 20:00:52 | INFO | fairseq.trainer | detected shared parameter: encoder.embed_tokens.weight <- decoder.output_projection.weight
2021-03-27 20:00:52 | INFO | fairseq.utils | ***********************CUDA enviroments for all 1 workers***********************
2021-03-27 20:00:52 | INFO | fairseq.utils | rank   0: capabilities =  7.0  ; total memory = 15.782 GB ; name = Tesla V100-SXM2-16GB
2021-03-27 20:00:52 | INFO | fairseq.utils | ***********************CUDA enviroments for all 1 workers***********************
2021-03-27 20:00:52 | INFO | fairseq_cli.train | training on 1 devices (GPUs/TPUs)
2021-03-27 20:00:52 | INFO | fairseq_cli.train | max tokens per device = 8000 and max sentences per device = None
2021-03-27 20:00:52 | INFO | fairseq.trainer | Preparing to load checkpoint checkpoints/checkpoint_last.pt
2021-03-27 20:03:07 | INFO | fairseq.trainer | NOTE: your device may support faster training with --fp16
2021-03-27 20:03:07 | INFO | fairseq.trainer | Loaded checkpoint checkpoints/checkpoint_last.pt (epoch 5 @ 2352 updates)
2021-03-27 20:03:07 | INFO | fairseq.trainer | loading train data for epoch 5
2021-03-27 20:03:09 | INFO | fairseq.data.data_utils | loaded 160,239 examples from: data-bin/iwslt14.tokenized.de-en/train.de-en.de
2021-03-27 20:03:11 | INFO | fairseq.data.data_utils | loaded 160,239 examples from: data-bin/iwslt14.tokenized.de-en/train.de-en.en
2021-03-27 20:03:11 | INFO | fairseq.tasks.translation | data-bin/iwslt14.tokenized.de-en train de-en 160239 examples
2021-03-27 20:03:12 | INFO | fairseq.trainer | begin training epoch 5
2021-03-27 20:03:12 | INFO | fairseq_cli.train | Start iterating over samples
/home/ubuntu/fairseq/fairseq/utils.py:360: UserWarning: amp_C fused kernels unavailable, disabling multi_tensor_l2norm; you may get better performance by installing NVIDIA's apex library
  "amp_C fused kernels unavailable, disabling multi_tensor_l2norm; "
2021-03-27 20:03:48 | INFO | train_inner | epoch 005:    48 / 588 loss=17.336, nll_loss=8.741, word_ins=9.39, word_crf=12.216, length=4.257, ppl=165502, wps=20218.9, ups=2.89, wpb=7007.5, bsz=286.7, num_updates=2400, lr=0.000120076, gnorm=7.483, train_wall=36, gb_free=10.2, wall=176
2021-03-27 20:04:23 | INFO | train_inner | epoch 005:   148 / 588 loss=17.325, nll_loss=8.823, word_ins=9.457, word_crf=12.163, length=4.331, ppl=164192, wps=20188.5, ups=2.89, wpb=6993.7, bsz=275.4, num_updates=2500, lr=0.000125075, gnorm=7.097, train_wall=34, gb_free=9.9, wall=210
2021-03-27 20:04:57 | INFO | train_inner | epoch 005:   248 / 588 loss=17.153, nll_loss=8.807, word_ins=9.44, word_crf=12.001, length=4.323, ppl=145784, wps=19923.7, ups=2.88, wpb=6906.2, bsz=265.2, num_updates=2600, lr=0.000130074, gnorm=6.332, train_wall=34, gb_free=9.3, wall=245
2021-03-27 20:05:32 | INFO | train_inner | epoch 005:   348 / 588 loss=16.687, nll_loss=8.646, word_ins=9.299, word_crf=11.607, length=4.3, ppl=105479, wps=20036.2, ups=2.9, wpb=6901.6, bsz=280.9, num_updates=2700, lr=0.000135073, gnorm=6.221, train_wall=34, gb_free=9.4, wall=280
2021-03-27 20:06:07 | INFO | train_inner | epoch 005:   448 / 588 loss=16.585, nll_loss=8.694, word_ins=9.338, word_crf=11.491, length=4.252, ppl=98334.7, wps=20243.4, ups=2.87, wpb=7065, bsz=275.4, num_updates=2800, lr=0.000140072, gnorm=6.242, train_wall=35, gb_free=9.6, wall=314
2021-03-27 20:06:42 | INFO | train_inner | epoch 005:   548 / 588 loss=16.394, nll_loss=8.681, word_ins=9.325, word_crf=11.301, length=4.3, ppl=86097.5, wps=19851.3, ups=2.84, wpb=6999.4, bsz=259.8, num_updates=2900, lr=0.000145071, gnorm=5.584, train_wall=35, gb_free=9.6, wall=350
2021-03-27 20:06:56 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 20:06:59 | INFO | valid | epoch 005 | valid on 'valid' subset | loss 15.5 | nll_loss 8.455 | word_ins 9.15 | word_crf 10.487 | length 4.38 | ppl 46345.8 | wps 60132 | wpb 5311.6 | bsz 208.1 | num_updates 2940 | best_loss 15.5
2021-03-27 20:06:59 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 5 @ 2940 updates
2021-03-27 20:06:59 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint5.pt
2021-03-27 20:07:00 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint5.pt
2021-03-27 20:08:00 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint5.pt (epoch 5 @ 2940 updates, score 15.5) (writing took 61.415995563999786 seconds)
2021-03-27 20:08:00 | INFO | fairseq_cli.train | end of epoch 5 (average epoch stats below)
2021-03-27 20:08:00 | INFO | train | epoch 005 | loss 16.823 | nll_loss 8.723 | word_ins 9.366 | word_crf 11.711 | length 4.291 | ppl 115964 | wps 15269.7 | ups 2.19 | wpb 6988.7 | bsz 272.5 | num_updates 2940 | lr 0.000147071 | gnorm 6.379 | train_wall 222 | gb_free 9.7 | wall 428
2021-03-27 20:08:00 | INFO | fairseq.trainer | begin training epoch 6
2021-03-27 20:08:00 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 20:08:22 | INFO | train_inner | epoch 006:    60 / 588 loss=16.12, nll_loss=8.631, word_ins=9.28, word_crf=11.065, length=4.155, ppl=71236.7, wps=7147.6, ups=1, wpb=7131, bsz=260, num_updates=3000, lr=0.00015007, gnorm=6.232, train_wall=35, gb_free=9.4, wall=449
2021-03-27 20:08:57 | INFO | train_inner | epoch 006:   160 / 588 loss=15.91, nll_loss=8.59, word_ins=9.244, word_crf=10.853, length=4.353, ppl=61578.8, wps=19735.2, ups=2.83, wpb=6979.9, bsz=249.4, num_updates=3100, lr=0.000155069, gnorm=5.852, train_wall=35, gb_free=9.4, wall=485
2021-03-27 20:09:32 | INFO | train_inner | epoch 006:   260 / 588 loss=15.27, nll_loss=8.289, word_ins=8.981, word_crf=10.362, length=4.175, ppl=39500.4, wps=20090.6, ups=2.88, wpb=6972.6, bsz=295.5, num_updates=3200, lr=0.000160068, gnorm=5.765, train_wall=34, gb_free=9.8, wall=520
2021-03-27 20:10:07 | INFO | train_inner | epoch 006:   360 / 588 loss=15.347, nll_loss=8.415, word_ins=9.089, word_crf=10.372, length=4.298, ppl=41675.7, wps=19768.5, ups=2.85, wpb=6948.2, bsz=263, num_updates=3300, lr=0.000165067, gnorm=6.185, train_wall=35, gb_free=9.7, wall=555
2021-03-27 20:10:42 | INFO | train_inner | epoch 006:   460 / 588 loss=14.945, nll_loss=8.276, word_ins=8.968, word_crf=10.044, length=4.172, ppl=31548.4, wps=19854.2, ups=2.86, wpb=6943.6, bsz=275.6, num_updates=3400, lr=0.000170066, gnorm=6.033, train_wall=35, gb_free=10.2, wall=590
2021-03-27 20:11:17 | INFO | train_inner | epoch 006:   560 / 588 loss=14.598, nll_loss=8.148, word_ins=8.855, word_crf=9.76, length=4.108, ppl=24800, wps=20060.8, ups=2.87, wpb=6997.4, bsz=290.9, num_updates=3500, lr=0.000175065, gnorm=5.524, train_wall=35, gb_free=9.3, wall=625
2021-03-27 20:11:27 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 20:11:30 | INFO | valid | epoch 006 | valid on 'valid' subset | loss 14.186 | nll_loss 8.146 | word_ins 8.87 | word_crf 9.324 | length 4.268 | ppl 18635.6 | wps 59633.7 | wpb 5311.6 | bsz 208.1 | num_updates 3528 | best_loss 14.186
2021-03-27 20:11:30 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 6 @ 3528 updates
```

2021-03-27 20:11:30 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint6.pt
2021-03-27 20:11:31 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint6.pt
2021-03-27 20:12:32 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint6.pt (epoch 6 @ 3528 updates, score 14.186) (writing took 62.36921142599999 seconds)
2021-03-27 20:12:32 | INFO | fairseq_cli.train | end of epoch 6 (average epoch stats below)
2021-03-27 20:12:32 | INFO | train | epoch 006 | loss 15.277 | nll_loss 8.369 | word_ins 9.05 | word_crf 10.331 | length 4.21 | ppl 39695.4 | wps 15112.9 | ups 2.16 | wpb 6988.7 | bsz 272.5 | num_updates 3528 | lr 0.000176465 | gnorm 5.94 | train_wall 204 | gb_free 9.7 | wall 700
2021-03-27 20:12:32 | INFO | fairseq.trainer | begin training epoch 7
2021-03-27 20:12:32 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 20:12:58 | INFO | train_inner | epoch 007:   72 / 588 loss=14.638, nll_loss=8.267, word_ins=8.959, word_crf=9.726, length=4.324, ppl=25490.5, wps=6840.1, ups=0.99, wpb=6929.7, bsz=259.5, num_updates=3600, lr=0.000180064, gnorm=5.849, train_wall=35, gb_free=9.4, wall=726
2021-03-27 20:13:33 | INFO | train_inner | epoch 007:   172 / 588 loss=14.194, nll_loss=8.066, word_ins=8.782, word_crf=9.376, length=4.261, ppl=18736.5, wps=19862.4, ups=2.86, wpb=6942, bsz=285.4, num_updates=3700, lr=0.000185063, gnorm=5.589, train_wall=35, gb_free=9.3, wall=761
2021-03-27 20:14:08 | INFO | train_inner | epoch 007:   272 / 588 loss=14.025, nll_loss=8.063, word_ins=8.78, word_crf=9.223, length=4.121, ppl=16665.5, wps=20167, ups=2.85, wpb=7073.6, bsz=279, num_updates=3800, lr=0.000190062, gnorm=5.479, train_wall=35, gb_free=9.5, wall=796
2021-03-27 20:14:43 | INFO | train_inner | epoch 007:   372 / 588 loss=13.772, nll_loss=8, word_ins=8.724, word_crf=9.004, length=4.065, ppl=13990, wps=19997.8, ups=2.85, wpb=7006.6, bsz=272.7, num_updates=3900, lr=0.000195061, gnorm=5.022, train_wall=35, gb_free=10.2, wall=831
2021-03-27 20:15:18 | INFO | train_inner | epoch 007:   472 / 588 loss=13.791, nll_loss=8.116, word_ins=8.823, word_crf=8.958, length=4.212, ppl=14172.3, wps=20044.7, ups=2.83, wpb=7089.2, bsz=253.1, num_updates=4000, lr=0.00020006, gnorm=5.558, train_wall=35, gb_free=9.7, wall=866
2021-03-27 20:15:54 | INFO | train_inner | epoch 007:   572 / 588 loss=13.295, nll_loss=7.856, word_ins=8.597, word_crf=8.574, length=4.231, ppl=10052.5, wps=19470.2, ups=2.83, wpb=6869.9, bsz=279.7, num_updates=4100, lr=0.000205059, gnorm=4.906, train_wall=35, gb_free=9.5, wall=902
2021-03-27 20:15:59 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 20:16:03 | INFO | valid | epoch 007 | valid on 'valid' subset | loss 13.138 | nll_loss 7.824 | word_ins 8.591 | word_crf 8.413 | length 4.291 | ppl 9015.23 | wps 59065.1 | wpb 5311.6 | bsz 208.1 | num_updates 4116 | best_loss 13.138
2021-03-27 20:16:03 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 7 @ 4116 updates
2021-03-27 20:16:03 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint7.pt
2021-03-27 20:16:03 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint7.pt
2021-03-27 20:17:06 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint7.pt (epoch 7 @ 4116 updates, score 13.138) (writing took 62.996597479 seconds)
2021-03-27 20:17:06 | INFO | fairseq_cli.train | end of epoch 7 (average epoch stats below)
2021-03-27 20:17:06 | INFO | train | epoch 007 | loss 13.895 | nll_loss 8.043 | word_ins 8.761 | word_crf 9.095 | length 4.191 | ppl 15231.6 | wps 15033.7 | ups 2.15 | wpb 6988.7 | bsz 272.5 | num_updates 4116 | lr 0.000205859 | gnorm 5.346 | train_wall 205 | gb_free 9.5 | wall 973
2021-03-27 20:17:06 | INFO | fairseq.trainer | begin training epoch 8
2021-03-27 20:17:06 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 20:17:35 | INFO | train_inner | epoch 008:   84 / 588 loss=12.9, nll_loss=7.669, word_ins=8.434, word_crf=8.28, length=4.031, ppl=7642.94, wps=6993.9, ups=0.99, wpb=7087, bsz=299.8, num_updates=4200, lr=0.000210058, gnorm=5.027, train_wall=35, gb_free=9.5, wall=1003
2021-03-27 20:18:10 | INFO | train_inner | epoch 008:   184 / 588 loss=12.875, nll_loss=7.735, word_ins=8.491, word_crf=8.219, length=4.116, ppl=7514.67, wps=19685.2, ups=2.87, wpb=6855.4, bsz=276.1, num_updates=4300, lr=0.000215057, gnorm=4.352, train_wall=35, gb_free=9.8, wall=1038
2021-03-27 20:18:45 | INFO | train_inner | epoch 008:   284 / 588 loss=12.665, nll_loss=7.711, word_ins=8.469, word_crf=8.024, length=4.061, ppl=6494.29, wps=20198.7, ups=2.87, wpb=7045.1, bsz=279.9, num_updates=4400, lr=0.000220056, gnorm=4.604, train_wall=35, gb_free=9.9, wall=1073
2021-03-27 20:19:20 | INFO | train_inner | epoch 008:   384 / 588 loss=12.658, nll_loss=7.807, word_ins=8.551, word_crf=7.956, length=4.265, ppl=6463.66, wps=19889.8, ups=2.84, wpb=6992.9, bsz=262.2, num_updates=4500, lr=0.000225055, gnorm=4.657, train_wall=35, gb_free=9.9, wall=1108
2021-03-27 20:19:55 | INFO | train_inner | epoch 008:   484 / 588 loss=12.228, nll_loss=7.582, word_ins=8.356, word_crf=7.654, length=3.964, ppl=4798.19, wps=19936.4, ups=2.84, wpb=7031.5, bsz=289.2, num_updates=4600, lr=0.000230054, gnorm=3.92, train_wall=35, gb_free=9.6, wall=1143
2021-03-27 20:20:31 | INFO | train_inner | epoch 008:   584 / 588 loss=12.504, nll_loss=7.866, word_ins=8.602, word_crf=7.769, length=4.332, ppl=5807.08, wps=19532.4, ups=2.8, wpb=6968.8, bsz=233.8, num_updates=4700, lr=0.000235053, gnorm=3.876, train_wall=35, gb_free=9.3, wall=1179
2021-03-27 20:20:32 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 20:20:36 | INFO | valid | epoch 008 | valid on 'valid' subset | loss 12.206 | nll_loss 7.635 | word_ins 8.431 | word_crf 7.573 | length 4.172 | ppl 4723.47 | wps 59696.7 | wpb 5311.6 | bsz 208.1 | num_updates 4704 | best_loss 12.206
2021-03-27 20:20:36 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 8 @ 4704 updates
2021-03-27 20:20:36 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint8.pt
2021-03-27 20:20:36 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint8.pt
2021-03-27 20:21:06 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint8.pt (epoch 8 @ 4704 updates, score 12.206) (writing took 30.763496154999984 seconds)
2021-03-27 20:21:06 | INFO | fairseq_cli.train | end of epoch 8 (average epoch stats below)
2021-03-27 20:21:06 | INFO | train | epoch 008 | loss 12.626 | nll_loss 7.73 | word_ins 8.485 | word_crf 7.97 | length 4.139 | ppl 6322.59 | wps 17065.8 | ups 2.44 | wpb 6988.7 | bsz 272.5 | num_updates 4704 | lr 0.000235253 | gnorm 4.401 | train_wall 205 | gb_free 9.4 | wall 1214
2021-03-27 20:21:06 | INFO | fairseq.trainer | begin training epoch 9
2021-03-27 20:21:06 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 20:21:39 | INFO | train_inner | epoch 009:   96 / 588 loss=11.752, nll_loss=7.382, word_ins=8.181, word_crf=7.258, length=4.033, ppl=3449.42, wps=10337.4, ups=1.46, wpb=7097.9, bsz=287.4, num_updates=4800, lr=0.000240052, gnorm=4.624, train_wall=34, gb_free=9.4, wall=1247
2021-03-27 20:22:14 | INFO | train_inner | epoch 009:   196 / 588 loss=11.831, nll_loss=7.538, word_ins=8.316, word_crf=7.261, length=4.119, ppl=3642.55, wps=20098, ups=2.86, wpb=7018.5, bsz=265, num_updates=4900, lr=0.000245051, gnorm=3.837, train_wall=35, gb_free=9.5, wall=1282
2021-03-27 20:22:50 | INFO | train_inner | epoch 009:   296 / 588 loss=11.673, nll_loss=7.5, word_ins=8.284, word_crf=7.114, length=4.164, ppl=3264.23, wps=19206.5, ups=2.78, wpb=6902.5, bsz=273.5, num_updates=5000, lr=0.00025005, gnorm=4.174, train_wall=36, gb_free=9.4, wall=1318
2021-03-27 20:23:26 | INFO | train_inner | epoch 009:   396 / 588 loss=11.3, nll_loss=7.321, word_ins=8.127, word_crf=6.827, length=4.09, ppl=2521.45, wps=19832.7, ups=2.84, wpb=6971.8, bsz=288.9, num_updates=5100, lr=0.000255049, gnorm=3.574, train_wall=35, gb_free=10, wall=1353
2021-03-27 20:24:01 | INFO | train_inner | epoch 009:   496 / 588 loss=11.322, nll_loss=7.406, word_ins=8.201, word_crf=6.805, length=4.164, ppl=2560.52, wps=19821.4, ups=2.84, wpb=6983.6, bsz=261.8, num_updates=5200, lr=0.000260048, gnorm=3.394, train_wall=35, gb_free=9.2, wall=1389
2021-03-27 20:24:33 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 20:24:37 | INFO | valid | epoch 009 | valid on 'valid' subset | loss 11.006 | nll_loss 7.245 | word_ins 8.081 | word_crf 6.542 | length 4.235 | ppl 2056.83 | wps 59461 | wpb 5311.6 | bsz 208.1 | num_updates 5292 | best_loss 11.006
2021-03-27 20:24:37 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 9 @ 5292 updates
2021-03-27 20:24:37 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint9.pt
2021-03-27 20:24:38 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint9.pt
2021-03-27 20:24:47 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint9.pt (epoch 9 @ 5292 updates, score 11.006) (writing took 10.146494974999769 seconds)
2021-03-27 20:24:47 | INFO | fairseq_cli.train | end of epoch 9 (average epoch stats below)
2021-03-27 20:24:47 | INFO | train | epoch 009 | loss 11.523 | nll_loss 7.435 | word_ins 8.226 | word_crf 6.998 | length 4.117 | ppl 2942.12 | wps 18629.3 | ups 2.67 | wpb 6988.7 | bsz 272.5 | num_updates 5292 | lr 0.000264647 | gnorm 3.879 | train_wall 205 | gb_free 9.6 | wall 1435
2021-03-27 20:24:47 | INFO | fairseq.trainer | begin training epoch 10
2021-03-27 20:24:47 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 20:24:50 | INFO | train_inner | epoch 010:   8 / 588 loss=11.307, nll_loss=7.507, word_ins=8.288, word_crf=6.742, length=4.202, ppl=2533.35, wps=14064.6, ups=2.03, wpb=6926.2, bsz=250.2, num_updates=5300, lr=0.000265047, gnorm=3.711, train_wall=35, gb_free=9.5, wall=1438
2021-03-27 20:25:26 | INFO | train_inner | epoch 010:   108 / 588 loss=11.151, nll_loss=7.438, word_ins=8.229, word_crf=6.622, length=4.145, ppl=2273.33, wps=19115, ups=2.77, wpb=6890.7, bsz=248.8, num_updates=5400, lr=0.000270046, gnorm=3.865, train_wall=36, gb_free=9.3, wall=1474
2021-03-27 20:26:01 | INFO | train_inner | epoch 010:   208 / 588 loss=10.84, nll_loss=7.274, word_ins=8.086, word_crf=6.385, length=4.114, ppl=1832.66, wps=19628, ups=2.82, wpb=6951, bsz=268.5, num_updates=5500, lr=0.000275045, gnorm=3.457, train_wall=35, gb_free=9.9, wall=1509
2021-03-27 20:26:36 | INFO | train_inner | epoch 010:   308 / 588 loss=10.515, nll_loss=7.079, word_ins=7.916, word_crf=6.152, length=4.046, ppl=1463.32, wps=20474.1, ups=2.88, wpb=7115.6, bsz=285.6, num_updates=5600, lr=0.000280044, gnorm=2.986, train_wall=34, gb_free=9.5, wall=1544
2021-03-27 20:27:11 | INFO | train_inner | epoch 010:   408 / 588 loss=10.335, nll_loss=7.021, word_ins=7.864, word_crf=6.007, length=3.962, ppl=1291.59, wps=20355.7, ups=2.87, wpb=7095.9, bsz=301.4, num_updates=5700, lr=0.000285043, gnorm=3.595, train_wall=35, gb_free=9.4, wall=1579

2021-03-27 20:27:47 | INFO | train_inner | epoch 010:   508 / 588 loss=10.601, nll_loss=7.304, word_ins=8.112, word_crf=6.121, length=4.239, ppl=1553.36, wps=19476.4, ups=2.8, wpb=6946.9, bsz=246, num_updates=5800, lr=0.000290042, gnorm=2.874, train_wall=35, gb_free=10, wall=1615
2021-03-27 20:28:14 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 20:28:18 | INFO | valid | epoch 010 | valid on 'valid' subset | loss 10.305 | nll_loss 7.059 | word_ins 7.925 | word_crf 5.927 | length 4.159 | ppl 1265.15 | wps 59787.1 | wpb 5311.6 | bsz 208.1 | num_updates 5880 | best_loss 10.305
2021-03-27 20:28:18 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 10 @ 5880 updates
2021-03-27 20:28:18 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint10.pt
2021-03-27 20:28:18 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint10.pt
2021-03-27 20:28:24 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint10.pt (epoch 10 @ 5880 updates, score 10.305) (writing took 6.612083792999783 seconds)
2021-03-27 20:28:24 | INFO | fairseq_cli.train | end of epoch 10 (average epoch stats below)
2021-03-27 20:28:24 | INFO | train | epoch 010 | loss 10.611 | nll_loss 7.187 | word_ins 8.01 | word_crf 6.195 | length 4.106 | ppl 1563.48 | wps 18910 | ups 2.71 | wpb 6988.7 | bsz 272.5 | num_updates 5880 | lr 0.000294041 | gnorm 3.323 | train_wall 205 | gb_free 9.9 | wall 1652
2021-03-27 20:28:24 | INFO | fairseq.trainer | begin training epoch 11
2021-03-27 20:28:24 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 20:28:31 | INFO | train_inner | epoch 011:   20 / 588 loss=10.055, nll_loss=6.929, word_ins=7.786, word_crf=5.755, length=4.067, ppl=1063.55, wps=15563.1, ups=2.23, wpb=6965.6, bsz=290.1, num_updates=5900, lr=0.000295041, gnorm=3.125, train_wall=34, gb_free=9.8, wall=1659
2021-03-27 20:29:07 | INFO | train_inner | epoch 011:   120 / 588 loss=10.014, nll_loss=6.934, word_ins=7.79, word_crf=5.716, length=4.023, ppl=1033.89, wps=19659.5, ups=2.82, wpb=6979.2, bsz=281.4, num_updates=6000, lr=0.00030004, gnorm=2.835, train_wall=35, gb_free=9.7, wall=1695
2021-03-27 20:29:42 | INFO | train_inner | epoch 011:   220 / 588 loss=10.016, nll_loss=7.001, word_ins=7.849, word_crf=5.688, length=4.035, ppl=1035.31, wps=19774.3, ups=2.85, wpb=6927.6, bsz=268.6, num_updates=6100, lr=0.000305039, gnorm=3.194, train_wall=35, gb_free=9.4, wall=1730
2021-03-27 20:30:17 | INFO | train_inner | epoch 011:   320 / 588 loss=9.94, nll_loss=7.017, word_ins=7.863, word_crf=5.593, length=4.153, ppl=982.09, wps=19746.5, ups=2.82, wpb=6996.6, bsz=259.9, num_updates=6200, lr=0.000310038, gnorm=2.671, train_wall=35, gb_free=9.4, wall=1765
2021-03-27 20:30:52 | INFO | train_inner | epoch 011:   420 / 588 loss=9.637, nll_loss=6.825, word_ins=7.696, word_crf=5.375, length=4.142, ppl=796.12, wps=19970.8, ups=2.86, wpb=6981.6, bsz=283.9, num_updates=6300, lr=0.000315037, gnorm=2.485, train_wall=35, gb_free=9.5, wall=1800
2021-03-27 20:31:28 | INFO | train_inner | epoch 011:   520 / 588 loss=9.58, nll_loss=6.831, word_ins=7.701, word_crf=5.322, length=4.081, ppl=765.57, wps=19892.7, ups=2.84, wpb=7005.2, bsz=280.6, num_updates=6400, lr=0.000320036, gnorm=3.599, train_wall=35, gb_free=9.3, wall=1836
2021-03-27 20:31:52 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 20:31:55 | INFO | valid | epoch 011 | valid on 'valid' subset | loss 9.501 | nll_loss 6.71 | word_ins 7.627 | word_crf 5.262 | length 4.252 | ppl 724.43 | wps 59466.6 | wpb 5311.6 | bsz 208.1 | num_updates 6468 | best_loss 9.501
2021-03-27 20:31:55 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 11 @ 6468 updates
2021-03-27 20:31:55 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint11.pt
2021-03-27 20:31:56 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint11.pt
2021-03-27 20:32:02 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint11.pt (epoch 11 @ 6468 updates, score 9.501) (writing took 6.633191413000077 seconds)
2021-03-27 20:32:02 | INFO | fairseq_cli.train | end of epoch 11 (average epoch stats below)
2021-03-27 20:32:02 | INFO | train | epoch 011 | loss 9.823 | nll_loss 6.923 | word_ins 7.781 | word_crf 5.523 | length 4.087 | ppl 905.54 | wps 18904.8 | ups 2.71 | wpb 6988.7 | bsz 272.5 | num_updates 6468 | lr 0.000323435 | gnorm 2.948 | train_wall 205 | gb_free 9.2 | wall 1869
2021-03-27 20:32:02 | INFO | fairseq.trainer | begin training epoch 12
2021-03-27 20:32:02 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 20:32:13 | INFO | train_inner | epoch 012:   32 / 588 loss=9.545, nll_loss=6.849, word_ins=7.718, word_crf=5.283, length=4.033, ppl=747.02, wps=15234, ups=2.2, wpb=6937.7, bsz=265.8, num_updates=6500, lr=0.000325035, gnorm=2.872, train_wall=35, gb_free=9.4, wall=1881
2021-03-27 20:32:49 | INFO | train_inner | epoch 012:   132 / 588 loss=9.442, nll_loss=6.819, word_ins=7.692, word_crf=5.191, length=4.05, ppl=695.41, wps=19852.7, ups=2.81, wpb=7062.5, bsz=260.5, num_updates=6600, lr=0.000330034, gnorm=2.666, train_wall=35, gb_free=10.6, wall=1917
2021-03-27 20:33:24 | INFO | train_inner | epoch 012:   232 / 588 loss=9.242, nll_loss=6.677, word_ins=7.568, word_crf=5.056, length=4.024, ppl=605.67, wps=20156.4, ups=2.86, wpb=7054.8, bsz=279.8, num_updates=6700, lr=0.000335033, gnorm=2.86, train_wall=35, gb_free=9.8, wall=1952
2021-03-27 20:33:59 | INFO | train_inner | epoch 012:   332 / 588 loss=9.353, nll_loss=6.847, word_ins=7.717, word_crf=5.079, length=4.157, ppl=654.03, wps=19559.1, ups=2.8, wpb=6982.7, bsz=242.3, num_updates=6800, lr=0.000340032, gnorm=2.569, train_wall=35, gb_free=10.7, wall=1987
2021-03-27 20:34:34 | INFO | train_inner | epoch 012:   432 / 588 loss=8.741, nll_loss=6.351, word_ins=7.286, word_crf=4.703, length=3.951, ppl=427.71, wps=20274, ups=2.9, wpb=6985.2, bsz=321, num_updates=6900, lr=0.000345031, gnorm=2.385, train_wall=34, gb_free=9.5, wall=2022
2021-03-27 20:35:09 | INFO | train_inner | epoch 012:   532 / 588 loss=9.099, nll_loss=6.743, word_ins=7.626, word_crf=4.866, length=4.192, ppl=548.21, wps=19545.3, ups=2.83, wpb=6905.8, bsz=264.8, num_updates=7000, lr=0.00035003, gnorm=2.41, train_wall=35, gb_free=9.1, wall=2057
2021-03-27 20:35:29 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 20:35:32 | INFO | valid | epoch 012 | valid on 'valid' subset | loss 8.857 | nll_loss 6.503 | word_ins 7.456 | word_crf 4.712 | length 4.169 | ppl 463.68 | wps 59483.1 | wpb 5311.6 | bsz 208.1 | num_updates 7056 | best_loss 8.857
2021-03-27 20:35:32 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 12 @ 7056 updates
2021-03-27 20:35:32 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint12.pt
2021-03-27 20:35:33 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint12.pt
2021-03-27 20:35:39 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint12.pt (epoch 12 @ 7056 updates, score 8.857) (writing took 6.613705517000199 seconds)
2021-03-27 20:35:39 | INFO | fairseq_cli.train | end of epoch 12 (average epoch stats below)
2021-03-27 20:35:39 | INFO | train | epoch 012 | loss 9.178 | nll_loss 6.692 | word_ins 7.582 | word_crf 4.981 | length 4.056 | ppl 579.06 | wps 18905.1 | ups 2.71 | wpb 6988.7 | bsz 272.5 | num_updates 7056 | lr 0.000352829 | gnorm 2.574 | train_wall 205 | gb_free 9.5 | wall 2087
2021-03-27 20:35:39 | INFO | fairseq.trainer | begin training epoch 13
2021-03-27 20:35:39 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 20:35:55 | INFO | train_inner | epoch 013:   44 / 588 loss=8.906, nll_loss=6.585, word_ins=7.489, word_crf=4.764, length=3.974, ppl=479.6, wps=15436.8, ups=2.21, wpb=6993.4, bsz=274, num_updates=7100, lr=0.000355029, gnorm=2.605, train_wall=35, gb_free=9.4, wall=2102
2021-03-27 20:36:30 | INFO | train_inner | epoch 013:   144 / 588 loss=9.045, nll_loss=6.786, word_ins=7.666, word_crf=4.809, length=4.031, ppl=528.31, wps=19602.4, ups=2.8, wpb=6989.9, bsz=246.3, num_updates=7200, lr=0.000360028, gnorm=3.534, train_wall=35, gb_free=9.4, wall=2138
2021-03-27 20:37:05 | INFO | train_inner | epoch 013:   244 / 588 loss=8.626, nll_loss=6.438, word_ins=7.362, word_crf=4.543, length=4.017, ppl=394.98, wps=19935.7, ups=2.85, wpb=6987, bsz=283.4, num_updates=7300, lr=0.000365027, gnorm=2.355, train_wall=35, gb_free=9.3, wall=2173
2021-03-27 20:37:40 | INFO | train_inner | epoch 013:   344 / 588 loss=8.593, nll_loss=6.409, word_ins=7.337, word_crf=4.508, length=4.16, ppl=386.06, wps=19850.9, ups=2.88, wpb=6886.1, bsz=283.5, num_updates=7400, lr=0.000370026, gnorm=2.248, train_wall=34, gb_free=9.6, wall=2208
2021-03-27 20:38:15 | INFO | train_inner | epoch 013:   444 / 588 loss=8.588, nll_loss=6.48, word_ins=7.399, word_crf=4.485, length=4.033, ppl=384.73, wps=20150.2, ups=2.85, wpb=7080, bsz=270, num_updates=7500, lr=0.000375025, gnorm=2.584, train_wall=35, gb_free=9.6, wall=2243
2021-03-27 20:38:50 | INFO | train_inner | epoch 013:   544 / 588 loss=8.595, nll_loss=6.559, word_ins=7.468, word_crf=4.457, length=4.038, ppl=386.58, wps=19697.4, ups=2.83, wpb=6953.5, bsz=264.2, num_updates=7600, lr=0.000380024, gnorm=2.539, train_wall=35, gb_free=9.9, wall=2278
2021-03-27 20:39:06 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 20:39:09 | INFO | valid | epoch 013 | valid on 'valid' subset | loss 8.544 | nll_loss 6.358 | word_ins 7.328 | word_crf 4.462 | length 4.183 | ppl 373.33 | wps 59599.9 | wpb 5311.6 | bsz 208.1 | num_updates 7644 | best_loss 8.544
2021-03-27 20:39:09 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 13 @ 7644 updates
2021-03-27 20:39:09 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint13.pt
2021-03-27 20:39:10 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint13.pt
2021-03-27 20:39:16 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint13.pt (epoch 13 @ 7644 updates, score 8.544) (writing took 6.611765787999957 seconds)
2021-03-27 20:39:16 | INFO | fairseq_cli.train | end of epoch 13 (average epoch stats below)
2021-03-27 20:39:16 | INFO | train | epoch 013 | loss 8.672 | nll_loss 6.519 | word_ins 7.433 | word_crf 4.551 | length 4.045 | ppl 407.9 | wps 18942.3 | ups 2.71 | wpb 6988.7 | bsz 272.5 | num_updates 7644 | lr 0.000382224 | gnorm 2.65 | train_wall 205 | gb_free 9.3 | wall 2304
2021-03-27 20:39:16 | INFO | fairseq.trainer | begin training epoch 14
2021-03-27 20:39:16 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 20:39:36 | INFO | train_inner | epoch 014:   56 / 588 loss=8.561, nll_loss=6.54, word_ins=7.453, word_crf=4.433, length=4.017, ppl=377.77, wps=15543.9, ups=2.19, wpb=7110.3, bsz=256.5, num_updates=7700, lr=0.000385023, gnorm=2.448, train_wall=35, gb_free=9.2, wall=2324

2021-03-27 20:40:11 | INFO | train_inner | epoch 014:   156 / 588 loss=8.252, nll_loss=6.289, word_ins=7.234, word_crf=4.229, length=4.054, ppl=304.83, wps=19522.4, ups=2.83, wpb=6898.7, bsz=286.2, num_updates=7800, lr=0.000390022, gnorm=2.287, train_wall=35, gb_free=9.9, wall=2359
2021-03-27 20:40:46 | INFO | train_inner | epoch 014:   256 / 588 loss=8.212, nll_loss=6.267, word_ins=7.216, word_crf=4.191, length=4.13, ppl=296.44, wps=19956.8, ups=2.88, wpb=6925.5, bsz=275.9, num_updates=7900, lr=0.000395021, gnorm=2.245, train_wall=34, gb_free=9.9, wall=2394
2021-03-27 20:41:21 | INFO | train_inner | epoch 014:   356 / 588 loss=8.183, nll_loss=6.306, word_ins=7.25, word_crf=4.168, length=3.9, ppl=290.65, wps=20263.6, ups=2.84, wpb=7144.2, bsz=278.6, num_updates=8000, lr=0.00040002, gnorm=2.24, train_wall=35, gb_free=9.6, wall=2429
2021-03-27 20:41:57 | INFO | train_inner | epoch 014:   456 / 588 loss=8.409, nll_loss=6.56, word_ins=7.47, word_crf=4.25, length=4.242, ppl=339.96, wps=19319.5, ups=2.81, wpb=6870.6, bsz=245, num_updates=8100, lr=0.000405019, gnorm=2.382, train_wall=35, gb_free=9.5, wall=2465
2021-03-27 20:42:32 | INFO | train_inner | epoch 014:   556 / 588 loss=8.069, nll_loss=6.25, word_ins=7.202, word_crf=4.072, length=3.964, ppl=268.54, wps=20160.1, ups=2.85, wpb=7069.6, bsz=279.4, num_updates=8200, lr=0.000410018, gnorm=2.139, train_wall=35, gb_free=10, wall=2500
2021-03-27 20:42:43 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 20:42:46 | INFO | valid | epoch 014 | valid on 'valid' subset | loss 8.204 | nll_loss 6.222 | word_ins 7.219 | word_crf 4.177 | length 4.175 | ppl 294.83 | wps 59493.6 | wpb 5311.6 | bsz 208.1 | num_updates 8232 | best_loss 8.204
2021-03-27 20:42:46 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 14 @ 8232 updates
2021-03-27 20:42:46 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint14.pt
2021-03-27 20:42:47 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint14.pt
2021-03-27 20:42:53 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint14.pt (epoch 14 @ 8232 updates, score 8.204) (writing took 6.602070447999722 seconds)
2021-03-27 20:42:53 | INFO | fairseq_cli.train | end of epoch 14 (average epoch stats below)
2021-03-27 20:42:53 | INFO | train | epoch 014 | loss 8.237 | nll_loss 6.341 | word_ins 7.28 | word_crf 4.192 | length 4.048 | ppl 301.63 | wps 18943.1 | ups 2.71 | wpb 6988.7 | bsz 272.5 | num_updates 8232 | lr 0.000411618 | gnorm 2.322 | train_wall 205 | gb_free 9.6 | wall 2521
2021-03-27 20:42:53 | INFO | fairseq.trainer | begin training epoch 15
2021-03-27 20:42:53 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 20:43:17 | INFO | train_inner | epoch 015:   68 / 588 loss=7.871, nll_loss=6.097, word_ins=7.069, word_crf=3.934, length=4.034, ppl=234.17, wps=15274.6, ups=2.22, wpb=6867.4, bsz=289.6, num_updates=8300, lr=0.000415017, gnorm=2.556, train_wall=35, gb_free=10.6, wall=2545
2021-03-27 20:43:52 | INFO | train_inner | epoch 015:   168 / 588 loss=7.893, nll_loss=6.119, word_ins=7.091, word_crf=3.955, length=3.928, ppl=237.76, wps=20068, ups=2.85, wpb=7040.8, bsz=279.8, num_updates=8400, lr=0.000420016, gnorm=2.082, train_wall=35, gb_free=9.6, wall=2580
2021-03-27 20:44:27 | INFO | train_inner | epoch 015:   268 / 588 loss=7.811, nll_loss=6.088, word_ins=7.063, word_crf=3.88, length=3.992, ppl=224.59, wps=20043.9, ups=2.85, wpb=7025.6, bsz=284.4, num_updates=8500, lr=0.000425015, gnorm=2.163, train_wall=35, gb_free=9.6, wall=2615
2021-03-27 20:45:03 | INFO | train_inner | epoch 015:   368 / 588 loss=7.778, nll_loss=6.115, word_ins=7.087, word_crf=3.851, length=3.833, ppl=219.45, wps=20279.3, ups=2.83, wpb=7169.4, bsz=284.5, num_updates=8600, lr=0.000430014, gnorm=1.851, train_wall=35, gb_free=9.7, wall=2650
2021-03-27 20:45:38 | INFO | train_inner | epoch 015:   468 / 588 loss=7.774, nll_loss=6.124, word_ins=7.093, word_crf=3.823, length=4.039, ppl=218.87, wps=19855.4, ups=2.84, wpb=6997, bsz=280.6, num_updates=8700, lr=0.000435013, gnorm=2.446, train_wall=35, gb_free=9.3, wall=2686
2021-03-27 20:46:13 | INFO | train_inner | epoch 015:   568 / 588 loss=8.014, nll_loss=6.411, word_ins=7.344, word_crf=3.923, length=4.183, ppl=258.41, wps=19280.8, ups=2.81, wpb=6863, bsz=242.5, num_updates=8800, lr=0.000440012, gnorm=2.083, train_wall=35, gb_free=10.3, wall=2721
2021-03-27 20:46:20 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 20:46:24 | INFO | valid | epoch 015 | valid on 'valid' subset | loss 7.868 | nll_loss 6.142 | word_ins 7.138 | word_crf 3.879 | length 4.203 | ppl 233.67 | wps 59405.2 | wpb 5311.6 | bsz 208.1 | num_updates 8820 | best_loss 7.868
2021-03-27 20:46:24 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 15 @ 8820 updates
2021-03-27 20:46:24 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint15.pt
2021-03-27 20:46:25 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint15.pt
2021-03-27 20:46:30 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint15.pt (epoch 15 @ 8820 updates, score 7.868) (writing took 6.643401391000225 seconds)
2021-03-27 20:46:30 | INFO | fairseq_cli.train | end of epoch 15 (average epoch stats below)
2021-03-27 20:46:30 | INFO | train | epoch 015 | loss 7.872 | nll_loss 6.179 | word_ins 7.142 | word_crf 3.899 | length 4.019 | ppl 234.22 | wps 18897 | ups 2.7 | wpb 6988.7 | bsz 272.5 | num_updates 8820 | lr 0.000441012 | gnorm 2.143 | train_wall 206 | gb_free 9.9 | wall 2738
2021-03-27 20:46:30 | INFO | fairseq.trainer | begin training epoch 16
2021-03-27 20:46:30 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 20:46:59 | INFO | train_inner | epoch 016:   80 / 588 loss=7.698, nll_loss=6.073, word_ins=7.052, word_crf=3.759, length=4.127, ppl=207.67, wps=15221.8, ups=2.21, wpb=6901.2, bsz=259.4, num_updates=8900, lr=0.000445011, gnorm=1.867, train_wall=35, gb_free=9.7, wall=2767
2021-03-27 20:47:34 | INFO | train_inner | epoch 016:   180 / 588 loss=7.703, nll_loss=6.128, word_ins=7.1, word_crf=3.753, length=3.996, ppl=208.33, wps=19444.4, ups=2.82, wpb=6903.5, bsz=266.3, num_updates=9000, lr=0.00045001, gnorm=2.382, train_wall=35, gb_free=9.4, wall=2802
2021-03-27 20:48:09 | INFO | train_inner | epoch 016:   280 / 588 loss=7.467, nll_loss=5.928, word_ins=6.926, word_crf=3.605, length=3.993, ppl=176.93, wps=20466.6, ups=2.88, wpb=7100.7, bsz=292.9, num_updates=9100, lr=0.000455009, gnorm=2.027, train_wall=34, gb_free=9.6, wall=2837
2021-03-27 20:48:44 | INFO | train_inner | epoch 016:   380 / 588 loss=7.668, nll_loss=6.153, word_ins=7.121, word_crf=3.695, length=4.122, ppl=203.34, wps=19543.5, ups=2.82, wpb=6934, bsz=254.9, num_updates=9200, lr=0.000460008, gnorm=1.884, train_wall=35, gb_free=9.4, wall=2872
2021-03-27 20:49:20 | INFO | train_inner | epoch 016:   480 / 588 loss=7.583, nll_loss=6.104, word_ins=7.079, word_crf=3.641, length=4.021, ppl=191.75, wps=19665.7, ups=2.8, wpb=7018.9, bsz=269.8, num_updates=9300, lr=0.000465007, gnorm=2.128, train_wall=35, gb_free=10, wall=2908
2021-03-27 20:49:55 | INFO | train_inner | epoch 016:   580 / 588 loss=7.452, nll_loss=5.998, word_ins=6.985, word_crf=3.567, length=3.92, ppl=175.09, wps=20086, ups=2.87, wpb=7003.5, bsz=283.7, num_updates=9400, lr=0.000470006, gnorm=2.248, train_wall=35, gb_free=9.8, wall=2943
2021-03-27 20:49:58 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 20:50:01 | INFO | valid | epoch 016 | valid on 'valid' subset | loss 7.738 | nll_loss 6.207 | word_ins 7.191 | word_crf 3.701 | length 4.408 | ppl 213.42 | wps 59624 | wpb 5311.6 | bsz 208.1 | num_updates 9408 | best_loss 7.738
2021-03-27 20:50:01 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 16 @ 9408 updates
2021-03-27 20:50:01 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint16.pt
2021-03-27 20:50:02 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint16.pt
2021-03-27 20:50:08 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint16.pt (epoch 16 @ 9408 updates, score 7.738) (writing took 6.996303532000184 seconds)
2021-03-27 20:50:08 | INFO | fairseq_cli.train | end of epoch 16 (average epoch stats below)
2021-03-27 20:50:08 | INFO | train | epoch 016 | loss 7.58 | nll_loss 6.053 | word_ins 7.034 | word_crf 3.662 | length 4.01 | ppl 191.41 | wps 18875.8 | ups 2.7 | wpb 6988.7 | bsz 272.5 | num_updates 9408 | lr 0.000470406 | gnorm 2.105 | train_wall 205 | gb_free 9.6 | wall 2956
2021-03-27 20:50:08 | INFO | fairseq.trainer | begin training epoch 17
2021-03-27 20:50:08 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 20:50:40 | INFO | train_inner | epoch 017:   92 / 588 loss=7.231, nll_loss=5.785, word_ins=6.803, word_crf=3.435, length=3.937, ppl=150.19, wps=15441.7, ups=2.2, wpb=7031.1, bsz=307.3, num_updates=9500, lr=0.000475005, gnorm=2.354, train_wall=35, gb_free=9.9, wall=2988
2021-03-27 20:51:15 | INFO | train_inner | epoch 017:   192 / 588 loss=7.08, nll_loss=5.622, word_ins=6.662, word_crf=3.374, length=3.751, ppl=135.31, wps=20499.7, ups=2.9, wpb=7067.4, bsz=308.9, num_updates=9600, lr=0.000480004, gnorm=1.771, train_wall=34, gb_free=9.3, wall=3023
2021-03-27 20:51:50 | INFO | train_inner | epoch 017:   292 / 588 loss=7.54, nll_loss=6.108, word_ins=7.085, word_crf=3.577, length=4.2, ppl=186.09, wps=19494.6, ups=2.83, wpb=6895.2, bsz=240.1, num_updates=9700, lr=0.000485003, gnorm=1.689, train_wall=35, gb_free=9.4, wall=3058
2021-03-27 20:52:26 | INFO | train_inner | epoch 017:   392 / 588 loss=7.457, nll_loss=6.089, word_ins=7.066, word_crf=3.516, length=4.079, ppl=175.75, wps=19792.7, ups=2.81, wpb=7035.3, bsz=260.7, num_updates=9800, lr=0.000490002, gnorm=1.915, train_wall=35, gb_free=9.9, wall=3094
2021-03-27 20:53:01 | INFO | train_inner | epoch 017:   492 / 588 loss=7.454, nll_loss=6.09, word_ins=7.069, word_crf=3.517, length=4.032, ppl=175.39, wps=19634.2, ups=2.81, wpb=6992.6, bsz=251.5, num_updates=9900, lr=0.000495001, gnorm=1.821, train_wall=35, gb_free=9.4, wall=3129
2021-03-27 20:53:35 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 20:53:39 | INFO | valid | epoch 017 | valid on 'valid' subset | loss 7.446 | nll_loss 5.919 | word_ins 6.952 | word_crf 3.556 | length 4.135 | ppl 174.33 | wps 59309.9 | wpb 5311.6 | bsz 208.1 | num_updates 9996 | best_loss 7.446
2021-03-27 20:53:39 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 17 @ 9996 updates
2021-03-27 20:53:39 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint17.pt
2021-03-27 20:53:39 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint17.pt
2021-03-27 20:53:45 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint17.pt (epoch 17 @ 9996 updates, score 7.446) (writing took 6.617650905000119 seconds)
2021-03-27 20:53:45 | INFO | fairseq_cli.train | end of epoch 17 (average epoch stats below)

2021-03-27 20:53:45 | INFO | train | epoch 017 | loss 7.353 | nll_loss 5.953 | word_ins 6.949 | word_crf 3.477 | length 4.021 | ppl 163.49 | wps 18917.1 | ups 2.71 | wpb 6988.7 | bsz 272.5 | num_updates 9996 | lr 0.0004998 | gnorm 2.025 | train_wall 205 | gb_free 9.7 | wall 3173
2021-03-27 20:53:45 | INFO | fairseq.trainer | begin training epoch 18
2021-03-27 20:53:45 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 20:53:47 | INFO | train_inner | epoch 018:   4 / 588 loss=7.342, nll_loss=6.004, word_ins=6.992, word_crf=3.439, length=4.066, ppl=162.26, wps=15304.7, ups=2.2, wpb=6953.7, bsz=271.3, num_updates=10000, lr=0.0005, gnorm=2.622, train_wall=35, gb_free=10, wall=3175
2021-03-27 20:53:47 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 20:53:50 | INFO | valid | epoch 018 | valid on 'valid' subset | loss 7.468 | nll_loss 5.961 | word_ins 6.996 | word_crf 3.553 | length 4.162 | ppl 176.99 | wps 59370 | wpb 5311.6 | bsz 208.1 | num_updates 10000 | best_loss 7.446
2021-03-27 20:53:50 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 18 @ 10000 updates
2021-03-27 20:53:50 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint_18_10000.pt
2021-03-27 20:53:51 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint_18_10000.pt
2021-03-27 20:53:54 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint_18_10000.pt (epoch 18 @ 10000 updates, score 7.468) (writing took 4.0222673530001885 seconds)
2021-03-27 20:54:30 | INFO | train_inner | epoch 018: 104 / 588 loss=7.087, nll_loss=5.73, word_ins=6.759, word_crf=3.317, length=3.91, ppl=135.97, wps=16638.1, ups=2.34, wpb=7101.4, bsz=280.2, num_updates=10100, lr=0.000497519, gnorm=1.683, train_wall=35, gb_free=9.4, wall=3217
2021-03-27 20:55:05 | INFO | train_inner | epoch 018: 204 / 588 loss=7.096, nll_loss=5.735, word_ins=6.763, word_crf=3.306, length=4.089, ppl=136.81, wps=19456.9, ups=2.83, wpb=6864.3, bsz=285.4, num_updates=10200, lr=0.000495074, gnorm=1.822, train_wall=35, gb_free=9.6, wall=3253
2021-03-27 20:55:40 | INFO | train_inner | epoch 018: 304 / 588 loss=7.361, nll_loss=6.064, word_ins=7.047, word_crf=3.428, length=4.101, ppl=164.43, wps=19511.1, ups=2.82, wpb=6914.8, bsz=244.6, num_updates=10300, lr=0.000492665, gnorm=3.091, train_wall=35, gb_free=9.2, wall=3288
2021-03-27 20:56:16 | INFO | train_inner | epoch 018: 404 / 588 loss=7.268, nll_loss=5.995, word_ins=6.987, word_crf=3.366, length=4.083, ppl=154.1, wps=19368.6, ups=2.8, wpb=6927.1, bsz=257.8, num_updates=10400, lr=0.00049029, gnorm=2.198, train_wall=35, gb_free=9.3, wall=3324
2021-03-27 20:56:51 | INFO | train_inner | epoch 018: 504 / 588 loss=6.867, nll_loss=5.589, word_ins=6.635, word_crf=3.166, length=3.838, ppl=116.76, wps=20170.2, ups=2.89, wpb=6984.9, bsz=308.5, num_updates=10500, lr=0.00048795, gnorm=1.799, train_wall=34, gb_free=9.7, wall=3359
2021-03-27 20:57:20 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 20:57:23 | INFO | valid | epoch 018 | valid on 'valid' subset | loss 7.3 | nll_loss 5.839 | word_ins 6.898 | word_crf 3.43 | length 4.206 | ppl 157.61 | wps 59542.4 | wpb 5311.6 | bsz 208.1 | num_updates 10584 | best_loss 7.3
2021-03-27 20:57:23 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 18 @ 10584 updates
2021-03-27 20:57:23 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint18.pt
2021-03-27 20:57:24 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint18.pt
2021-03-27 20:57:30 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint18.pt (epoch 18 @ 10584 updates, score 7.3) (writing took 6.609782521999932 seconds)
2021-03-27 20:57:30 | INFO | fairseq_cli.train | end of epoch 18 (average epoch stats below)
2021-03-27 20:57:30 | INFO | train | epoch 018 | loss 7.139 | nll_loss 5.834 | word_ins 6.848 | word_crf 3.315 | length 3.993 | ppl 140.9 | wps 18276.1 | ups 2.62 | wpb 6988.7 | bsz 272.5 | num_updates 10584 | lr 0.00048601 | gnorm 2.111 | train_wall 206 | gb_free 9.2 | wall 3398
2021-03-27 20:57:30 | INFO | fairseq.trainer | begin training epoch 19
2021-03-27 20:57:30 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 20:57:36 | INFO | train_inner | epoch 019:   16 / 588 loss=7.069, nll_loss=5.815, word_ins=6.833, word_crf=3.263, length=3.901, ppl=134.28, wps=15910, ups=2.22, wpb=7162.8, bsz=264.4, num_updates=10600, lr=0.000485643, gnorm=1.983, train_wall=35, gb_free=9.4, wall=3404
2021-03-27 20:58:10 | INFO | train_inner | epoch 019: 116 / 588 loss=6.754, nll_loss=5.461, word_ins=6.526, word_crf=3.089, length=4.024, ppl=107.94, wps=19752.7, ups=2.88, wpb=6864.7, bsz=309.3, num_updates=10700, lr=0.000483368, gnorm=1.895, train_wall=34, gb_free=9.8, wall=3438
2021-03-27 20:58:46 | INFO | train_inner | epoch 019: 216 / 588 loss=7.072, nll_loss=5.844, word_ins=6.859, word_crf=3.254, length=3.885, ppl=134.59, wps=19394.3, ups=2.78, wpb=6973.6, bsz=266.7, num_updates=10800, lr=0.000481125, gnorm=2.902, train_wall=36, gb_free=9.8, wall=3474
2021-03-27 20:59:22 | INFO | train_inner | epoch 019: 316 / 588 loss=7.14, nll_loss=5.904, word_ins=6.913, word_crf=3.273, length=4.113, ppl=141.09, wps=19817.6, ups=2.84, wpb=6978.9, bsz=240.6, num_updates=10900, lr=0.000478913, gnorm=1.636, train_wall=35, gb_free=9.1, wall=3510
2021-03-27 20:59:57 | INFO | train_inner | epoch 019: 416 / 588 loss=6.886, nll_loss=5.656, word_ins=6.696, word_crf=3.156, length=3.818, ppl=118.28, wps=20162.9, ups=2.86, wpb=7038, bsz=275.1, num_updates=11000, lr=0.000476731, gnorm=1.597, train_wall=35, gb_free=9.8, wall=3544
2021-03-27 21:00:31 | INFO | train_inner | epoch 019: 516 / 588 loss=6.764, nll_loss=5.524, word_ins=6.582, word_crf=3.092, length=3.818, ppl=108.71, wps=20776.5, ups=2.9, wpb=7152.8, bsz=291.4, num_updates=11100, lr=0.000474579, gnorm=1.578, train_wall=34, gb_free=9.3, wall=3579
2021-03-27 21:00:56 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 21:00:59 | INFO | valid | epoch 019 | valid on 'valid' subset | loss 7.242 | nll_loss 5.767 | word_ins 6.841 | word_crf 3.405 | length 4.171 | ppl 151.39 | wps 60171.8 | wpb 5311.6 | bsz 208.1 | num_updates 11172 | best_loss 7.242
2021-03-27 21:00:59 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 19 @ 11172 updates
2021-03-27 21:00:59 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint19.pt
2021-03-27 21:01:00 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint19.pt
2021-03-27 21:01:06 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint19.pt (epoch 19 @ 11172 updates, score 7.242) (writing took 6.616271652999785 seconds)
2021-03-27 21:01:06 | INFO | fairseq_cli.train | end of epoch 19 (average epoch stats below)
2021-03-27 21:01:06 | INFO | train | epoch 019 | loss 6.944 | nll_loss 5.703 | word_ins 6.737 | word_crf 3.179 | length 3.96 | ppl 123.09 | wps 19021.5 | ups 2.72 | wpb 6988.7 | bsz 272.5 | num_updates 11172 | lr 0.000473047 | gnorm 1.872 | train_wall 204 | gb_free 9.3 | wall 3614
2021-03-27 21:01:06 | INFO | fairseq.trainer | begin training epoch 20
2021-03-27 21:01:06 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 21:01:16 | INFO | train_inner | epoch 020:   28 / 588 loss=7.017, nll_loss=5.814, word_ins=6.833, word_crf=3.186, length=4.143, ppl=129.47, wps=15289.9, ups=2.22, wpb=6890.2, bsz=252, num_updates=11200, lr=0.000472456, gnorm=1.691, train_wall=35, gb_free=9.3, wall=3624
2021-03-27 21:01:50 | INFO | train_inner | epoch 020: 128 / 588 loss=6.822, nll_loss=5.585, word_ins=6.638, word_crf=3.106, length=3.975, ppl=113.13, wps=20456.9, ups=2.9, wpb=7046.6, bsz=265.9, num_updates=11300, lr=0.00047036, gnorm=1.835, train_wall=34, gb_free=9.1, wall=3658
2021-03-27 21:02:25 | INFO | train_inner | epoch 020: 228 / 588 loss=6.783, nll_loss=5.559, word_ins=6.615, word_crf=3.076, length=3.993, ppl=110.13, wps=19892, ups=2.86, wpb=6948.6, bsz=266.2, num_updates=11400, lr=0.000468293, gnorm=1.588, train_wall=35, gb_free=9.4, wall=3693
2021-03-27 21:03:00 | INFO | train_inner | epoch 020: 328 / 588 loss=6.701, nll_loss=5.493, word_ins=6.556, word_crf=3.033, length=3.904, ppl=104.07, wps=19922.9, ups=2.87, wpb=6940.2, bsz=280.6, num_updates=11500, lr=0.000466252, gnorm=1.782, train_wall=35, gb_free=10, wall=3728
2021-03-27 21:03:35 | INFO | train_inner | epoch 020: 428 / 588 loss=6.739, nll_loss=5.538, word_ins=6.597, word_crf=3.051, length=3.897, ppl=106.82, wps=20146.6, ups=2.85, wpb=7057, bsz=273.7, num_updates=11600, lr=0.000464238, gnorm=1.461, train_wall=35, gb_free=9.4, wall=3763
2021-03-27 21:04:11 | INFO | train_inner | epoch 020: 528 / 588 loss=6.897, nll_loss=5.741, word_ins=6.77, word_crf=3.104, length=4.081, ppl=119.17, wps=19378.9, ups=2.8, wpb=6923.5, bsz=267.7, num_updates=11700, lr=0.00046225, gnorm=1.593, train_wall=35, gb_free=9.3, wall=3799
2021-03-27 21:04:32 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 21:04:35 | INFO | valid | epoch 020 | valid on 'valid' subset | loss 7.127 | nll_loss 5.688 | word_ins 6.77 | word_crf 3.334 | length 4.086 | ppl 139.8 | wps 59745.2 | wpb 5311.6 | bsz 208.1 | num_updates 11760 | best_loss 7.127
2021-03-27 21:04:35 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 20 @ 11760 updates
2021-03-27 21:04:35 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint20.pt
2021-03-27 21:04:36 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint20.pt
2021-03-27 21:04:42 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint20.pt (epoch 20 @ 11760 updates, score 7.127) (writing took 6.666855917000248 seconds)
2021-03-27 21:04:42 | INFO | fairseq_cli.train | end of epoch 20 (average epoch stats below)
2021-03-27 21:04:42 | INFO | train | epoch 020 | loss 6.782 | nll_loss 5.58 | word_ins 6.632 | word_crf 3.071 | length 3.955 | ppl 110.09 | wps 19043.5 | ups 2.72 | wpb 6988.7 | bsz 272.5 | num_updates 11760 | lr 0.000461069 | gnorm 1.667 | train_wall 204 | gb_free 9.4 | wall 3830
2021-03-27 21:04:42 | INFO | fairseq.trainer | begin training epoch 21
2021-03-27 21:04:42 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 21:04:56 | INFO | train_inner | epoch 021:   40 / 588 loss=6.722, nll_loss=5.543, word_ins=6.6, word_crf=3.032, length=3.89, ppl=105.54, wps=15701.5, ups=2.23, wpb=7045.2, bsz=270.7, num_updates=11800, lr=0.000460287, gnorm=1.621, train_wall=34, gb_free=9.9, wall=3844
2021-03-27 21:05:31 | INFO | train_inner | epoch 021: 140 / 588 loss=6.534, nll_loss=5.353, word_ins=6.437, word_crf=2.943, length=3.723, ppl=92.66, wps=19905, ups=2.85, wpb=6983.6, bsz=291.2, num_updates=11900, lr=0.000458349, gnorm=1.689, train_wall=35, gb_free=9.8, wall=3879

2021-03-27 21:06:06 | INFO | train_inner | epoch 021:   240 / 588 loss=6.525, nll_loss=5.329, word_ins=6.418, word_crf=2.929, length=3.872, ppl=92.09, wps=20306.2, ups=2.88, wpb=7040.3, bsz=274.4, num_updates=12000, lr=0.000456435, gnorm=1.453, train_wall=34, gb_free=9.3, wall=3914
2021-03-27 21:06:41 | INFO | train_inner | epoch 021:   340 / 588 loss=6.67, nll_loss=5.484, word_ins=6.552, word_crf=2.993, length=4.016, ppl=101.85, wps=19630.2, ups=2.84, wpb=6918, bsz=268.3, num_updates=12100, lr=0.000454545, gnorm=1.43, train_wall=35, gb_free=9.5, wall=3949
2021-03-27 21:07:16 | INFO | train_inner | epoch 021:   440 / 588 loss=6.748, nll_loss=5.609, word_ins=6.658, word_crf=3.023, length=3.956, ppl=107.48, wps=19833.5, ups=2.83, wpb=7009.2, bsz=258.6, num_updates=12200, lr=0.000452679, gnorm=1.539, train_wall=35, gb_free=9.2, wall=3984
2021-03-27 21:07:51 | INFO | train_inner | epoch 021:   540 / 588 loss=6.613, nll_loss=5.464, word_ins=6.533, word_crf=2.953, length=3.937, ppl=97.87, wps=19969.9, ups=2.84, wpb=7023.6, bsz=281.1, num_updates=12300, lr=0.000450835, gnorm=1.413, train_wall=35, gb_free=9.2, wall=4019
2021-03-27 21:08:09 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 21:08:12 | INFO | valid | epoch 021 | valid on 'valid' subset | loss 7.076 | nll_loss 5.715 | word_ins 6.797 | word_crf 3.257 | length 4.205 | ppl 134.94 | wps 58666.4 | wpb 5311.6 | bsz 208.1 | num_updates 12348 | best_loss 7.076
2021-03-27 21:08:12 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 21 @ 12348 updates
2021-03-27 21:08:12 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint21.pt
2021-03-27 21:08:13 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint21.pt
2021-03-27 21:08:19 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint21.pt (epoch 21 @ 12348 updates, score 7.076) (writing took 6.682698700000401 seconds)
2021-03-27 21:08:19 | INFO | fairseq_cli.train | end of epoch 21 (average epoch stats below)
2021-03-27 21:08:19 | INFO | train | epoch 021 | loss 6.628 | nll_loss 5.457 | word_ins 6.528 | word_crf 2.972 | length 3.922 | ppl 98.91 | wps 18954.6 | ups 2.71 | wpb 6988.7 | bsz 272.5 | num_updates 12348 | lr 0.000449958 | gnorm 1.498 | train_wall 205 | gb_free 10.5 | wall 4047
2021-03-27 21:08:19 | INFO | fairseq.trainer | begin training epoch 22
2021-03-27 21:08:19 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 21:08:37 | INFO | train_inner | epoch 022:   52 / 588 loss=6.61, nll_loss=5.451, word_ins=6.524, word_crf=2.947, length=4.012, ppl=97.71, wps=15195.6, ups=2.2, wpb=6922.7, bsz=264.6, num_updates=12400, lr=0.000449013, gnorm=1.749, train_wall=35, gb_free=9.1, wall=4065
2021-03-27 21:09:12 | INFO | train_inner | epoch 022:   152 / 588 loss=6.606, nll_loss=5.462, word_ins=6.536, word_crf=2.946, length=3.922, ppl=97.42, wps=20317.6, ups=2.87, wpb=7068.4, bsz=257.5, num_updates=12500, lr=0.000447214, gnorm=1.872, train_wall=35, gb_free=10.4, wall=4100
2021-03-27 21:09:47 | INFO | train_inner | epoch 022:   252 / 588 loss=6.408, nll_loss=5.266, word_ins=6.364, word_crf=2.86, length=3.657, ppl=84.94, wps=20660.3, ups=2.87, wpb=7187.7, bsz=289.9, num_updates=12600, lr=0.000445435, gnorm=1.632, train_wall=35, gb_free=9.6, wall=4134
2021-03-27 21:10:21 | INFO | train_inner | epoch 022:   352 / 588 loss=6.468, nll_loss=5.299, word_ins=6.392, word_crf=2.879, length=3.933, ppl=88.54, wps=19881.5, ups=2.91, wpb=6829.1, bsz=276.9, num_updates=12700, lr=0.000443678, gnorm=1.482, train_wall=34, gb_free=9.7, wall=4169
2021-03-27 21:10:56 | INFO | train_inner | epoch 022:   452 / 588 loss=6.435, nll_loss=5.299, word_ins=6.391, word_crf=2.856, length=3.838, ppl=86.53, wps=19782.7, ups=2.85, wpb=6942.8, bsz=284.7, num_updates=12800, lr=0.000441942, gnorm=1.738, train_wall=35, gb_free=9.8, wall=4204
2021-03-27 21:11:31 | INFO | train_inner | epoch 022:   552 / 588 loss=6.703, nll_loss=5.592, word_ins=6.647, word_crf=2.978, length=4.023, ppl=104.22, wps=19628, ups=2.82, wpb=6960.9, bsz=241, num_updates=12900, lr=0.000440225, gnorm=1.295, train_wall=35, gb_free=10.8, wall=4239
2021-03-27 21:11:44 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 21:11:47 | INFO | valid | epoch 022 | valid on 'valid' subset | loss 7.026 | nll_loss 5.594 | word_ins 6.696 | word_crf 3.264 | length 4.143 | ppl 130.36 | wps 60066.2 | wpb 5311.6 | bsz 208.1 | num_updates 12936 | best_loss 7.026
2021-03-27 21:11:47 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 22 @ 12936 updates
2021-03-27 21:11:47 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint22.pt
2021-03-27 21:11:48 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint22.pt
2021-03-27 21:11:54 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint22.pt (epoch 22 @ 12936 updates, score 7.026) (writing took 6.6073480539998855 seconds)
2021-03-27 21:11:54 | INFO | fairseq_cli.train | end of epoch 22 (average epoch stats below)
2021-03-27 21:11:54 | INFO | train | epoch 022 | loss 6.515 | nll_loss 5.373 | word_ins 6.457 | word_crf 2.898 | length 3.883 | ppl 91.44 | wps 19095.3 | ups 2.73 | wpb 6988.7 | bsz 272.5 | num_updates 12936 | lr 0.000439612 | gnorm 1.641 | train_wall 204 | gb_free 9.9 | wall 4262
2021-03-27 21:11:54 | INFO | fairseq.trainer | begin training epoch 23
2021-03-27 21:11:54 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 21:12:17 | INFO | train_inner | epoch 023:   64 / 588 loss=6.301, nll_loss=5.172, word_ins=6.282, word_crf=2.775, length=3.85, ppl=78.83, wps=15435, ups=2.22, wpb=6966.2, bsz=292.6, num_updates=13000, lr=0.000438529, gnorm=1.523, train_wall=35, gb_free=9.1, wall=4284
2021-03-27 21:12:52 | INFO | train_inner | epoch 023:   164 / 588 loss=6.455, nll_loss=5.327, word_ins=6.421, word_crf=2.856, length=3.888, ppl=87.75, wps=19944.3, ups=2.85, wpb=7005, bsz=266.8, num_updates=13100, lr=0.000436852, gnorm=1.545, train_wall=35, gb_free=9.7, wall=4320
2021-03-27 21:13:28 | INFO | train_inner | epoch 023:   264 / 588 loss=6.535, nll_loss=5.422, word_ins=6.501, word_crf=2.884, length=3.995, ppl=92.71, wps=19242.5, ups=2.77, wpb=6937.4, bsz=258, num_updates=13200, lr=0.000435194, gnorm=1.575, train_wall=36, gb_free=9.6, wall=4356
2021-03-27 21:14:03 | INFO | train_inner | epoch 023:   364 / 588 loss=6.523, nll_loss=5.394, word_ins=6.477, word_crf=2.889, length=3.952, ppl=91.95, wps=19612, ups=2.85, wpb=6871.6, bsz=257.6, num_updates=13300, lr=0.000433555, gnorm=1.719, train_wall=35, gb_free=9.4, wall=4391
2021-03-27 21:14:37 | INFO | train_inner | epoch 023:   464 / 588 loss=6.18, nll_loss=5.05, word_ins=6.179, word_crf=2.707, length=3.832, ppl=72.51, wps=20636.4, ups=2.92, wpb=7076.8, bsz=299.8, num_updates=13400, lr=0.000431934, gnorm=1.214, train_wall=34, gb_free=9.9, wall=4425
2021-03-27 21:15:12 | INFO | train_inner | epoch 023:   564 / 588 loss=6.38, nll_loss=5.263, word_ins=6.364, word_crf=2.814, length=3.844, ppl=83.31, wps=20162.5, ups=2.86, wpb=7041.4, bsz=271.3, num_updates=13500, lr=0.000430331, gnorm=1.264, train_wall=35, gb_free=9.4, wall=4460
2021-03-27 21:15:20 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 21:15:24 | INFO | valid | epoch 023 | valid on 'valid' subset | loss 6.999 | nll_loss 5.511 | word_ins 6.644 | word_crf 3.259 | length 4.185 | ppl 127.93 | wps 60149.5 | wpb 5311.6 | bsz 208.1 | num_updates 13524 | best_loss 6.999
2021-03-27 21:15:24 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 23 @ 13524 updates
2021-03-27 21:15:24 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint23.pt
2021-03-27 21:15:24 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint23.pt
2021-03-27 21:15:30 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint23.pt (epoch 23 @ 13524 updates, score 6.999) (writing took 6.702953353999874 seconds)
2021-03-27 21:15:30 | INFO | fairseq_cli.train | end of epoch 23 (average epoch stats below)
2021-03-27 21:15:30 | INFO | train | epoch 023 | loss 6.391 | nll_loss 5.267 | word_ins 6.368 | word_crf 2.818 | length 3.893 | ppl 83.91 | wps 18994.8 | ups 2.72 | wpb 6988.7 | bsz 272.5 | num_updates 13524 | lr 0.000429949 | gnorm 1.471 | train_wall 205 | gb_free 9.8 | wall 4478
2021-03-27 21:15:30 | INFO | fairseq.trainer | begin training epoch 24
2021-03-27 21:15:30 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 21:15:57 | INFO | train_inner | epoch 024:   76 / 588 loss=6.222, nll_loss=5.097, word_ins=6.222, word_crf=2.719, length=3.918, ppl=74.66, wps=15485.1, ups=2.24, wpb=6920.2, bsz=266.2, num_updates=13600, lr=0.000428746, gnorm=1.457, train_wall=34, gb_free=9.9, wall=4505
2021-03-27 21:16:31 | INFO | train_inner | epoch 024:   176 / 588 loss=6.219, nll_loss=5.113, word_ins=6.237, word_crf=2.724, length=3.757, ppl=74.48, wps=20257.3, ups=2.88, wpb=7034.3, bsz=277.6, num_updates=13700, lr=0.000427179, gnorm=1.708, train_wall=34, gb_free=9.5, wall=4539
2021-03-27 21:17:06 | INFO | train_inner | epoch 024:   276 / 588 loss=6.226, nll_loss=5.106, word_ins=6.23, word_crf=2.732, length=3.796, ppl=74.87, wps=19946.6, ups=2.87, wpb=6940.3, bsz=284, num_updates=13800, lr=0.000425628, gnorm=1.454, train_wall=35, gb_free=9.9, wall=4574
2021-03-27 21:17:41 | INFO | train_inner | epoch 024:   376 / 588 loss=6.207, nll_loss=5.093, word_ins=6.218, word_crf=2.726, length=3.717, ppl=73.88, wps=20420.4, ups=2.85, wpb=7161, bsz=291.8, num_updates=13900, lr=0.000424094, gnorm=1.307, train_wall=35, gb_free=9.2, wall=4609
2021-03-27 21:18:17 | INFO | train_inner | epoch 024:   476 / 588 loss=6.471, nll_loss=5.368, word_ins=6.457, word_crf=2.842, length=4.012, ppl=88.74, wps=19551.5, ups=2.84, wpb=6892.3, bsz=248.7, num_updates=14000, lr=0.000422577, gnorm=1.421, train_wall=35, gb_free=9.2, wall=4644
2021-03-27 21:18:52 | INFO | train_inner | epoch 024:   576 / 588 loss=6.375, nll_loss=5.285, word_ins=6.383, word_crf=2.791, length=3.932, ppl=83.02, wps=19952.9, ups=2.86, wpb=6980.9, bsz=267.8, num_updates=14100, lr=0.000421076, gnorm=1.711, train_wall=35, gb_free=9.4, wall=4679
2021-03-27 21:18:56 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 21:18:59 | INFO | valid | epoch 024 | valid on 'valid' subset | loss 6.892 | nll_loss 5.474 | word_ins 6.601 | word_crf 3.147 | length 4.451 | ppl 118.8 | wps 60284.7 | wpb 5311.6 | bsz 208.1 | num_updates 14112 | best_loss 6.892
2021-03-27 21:18:59 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 24 @ 14112 updates
2021-03-27 21:18:59 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint24.pt
2021-03-27 21:19:00 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint24.pt
2021-03-27 21:19:06 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint24.pt (epoch 24 @ 14112 updates, score 6.892) (writing took 6.642967477999264 seconds)
2021-03-27 21:19:06 | INFO | fairseq_cli.train | end of epoch 24 (average epoch stats below)

2021-03-27 21:19:06 | INFO | train | epoch 024 | loss 6.283 | nll_loss 5.174 | word_ins 6.288 | word_crf 2.754 | length 3.852 | ppl 77.87 | wps 19088.4 | ups 2.73 | wpb 6988.7 | bsz 272.5 | num_updates 14112 | lr 0.000420897 | gnorm 1.499 | train_wall 204 | gb_free 10.5 | wall 4693
2021-03-27 21:19:06 | INFO | fairseq.trainer | begin training epoch 25
2021-03-27 21:19:06 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 21:19:37 | INFO | train_inner | epoch 025:   88 / 588 loss=6.168, nll_loss=5.085, word_ins=6.214, word_crf=2.681, length=3.794, ppl=71.88, wps=15685.6, ups=2.22, wpb=7066.1, bsz=271.6, num_updates=14200, lr=0.000419591, gnorm=1.529, train_wall=35, gb_free=9.6, wall=4724
2021-03-27 21:20:11 | INFO | train_inner | epoch 025:   188 / 588 loss=6.2, nll_loss=5.09, word_ins=6.219, word_crf=2.696, length=3.943, ppl=73.54, wps=19653.3, ups=2.89, wpb=6802.3, bsz=259.4, num_updates=14300, lr=0.000418121, gnorm=1.377, train_wall=34, gb_free=10.6, wall=4759
2021-03-27 21:20:46 | INFO | train_inner | epoch 025:   288 / 588 loss=6.345, nll_loss=5.238, word_ins=6.35, word_crf=2.779, length=3.912, ppl=81.29, wps=19802.6, ups=2.85, wpb=6941.2, bsz=243.8, num_updates=14400, lr=0.000416667, gnorm=1.486, train_wall=35, gb_free=10.1, wall=4794
2021-03-27 21:21:21 | INFO | train_inner | epoch 025:   388 / 588 loss=6.146, nll_loss=5.067, word_ins=6.198, word_crf=2.674, length=3.736, ppl=70.81, wps=20737.9, ups=2.89, wpb=7180.5, bsz=280.5, num_updates=14500, lr=0.000415227, gnorm=1.415, train_wall=34, gb_free=9.5, wall=4829
2021-03-27 21:21:55 | INFO | train_inner | epoch 025:   488 / 588 loss=6.027, nll_loss=4.937, word_ins=6.085, word_crf=2.613, length=3.715, ppl=65.23, wps=20357.8, ups=2.9, wpb=7021.7, bsz=291.2, num_updates=14600, lr=0.000413803, gnorm=1.222, train_wall=34, gb_free=9.4, wall=4863
2021-03-27 21:22:30 | INFO | train_inner | epoch 025:   588 / 588 loss=6.24, nll_loss=5.159, word_ins=6.275, word_crf=2.717, length=3.856, ppl=75.59, wps=19800.9, ups=2.85, wpb=6935.6, bsz=286, num_updates=14700, lr=0.000412393, gnorm=1.5, train_wall=35, gb_free=10, wall=4898
2021-03-27 21:22:30 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 21:22:34 | INFO | valid | epoch 025 | valid on 'valid' subset | loss 6.937 | nll_loss 5.498 | word_ins 6.624 | word_crf 3.211 | length 4.138 | ppl 122.5 | wps 60297.1 | wpb 5311.6 | bsz 208.1 | num_updates 14700 | best_loss 6.892
2021-03-27 21:22:34 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 25 @ 14700 updates
2021-03-27 21:22:34 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint25.pt
2021-03-27 21:22:34 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint25.pt
2021-03-27 21:22:37 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint25.pt (epoch 25 @ 14700 updates, score 6.937) (writing took 3.6378525350000928 seconds)
2021-03-27 21:22:37 | INFO | fairseq_cli.train | end of epoch 25 (average epoch stats below)
2021-03-27 21:22:37 | INFO | train | epoch 025 | loss 6.186 | nll_loss 5.094 | word_ins 6.222 | word_crf 2.693 | length 3.825 | ppl 72.8 | wps 19406.4 | ups 2.78 | wpb 6988.7 | bsz 272.5 | num_updates 14700 | lr 0.000412393 | gnorm 1.427 | train_wall 203 | gb_free 10 | wall 4905
2021-03-27 21:22:37 | INFO | fairseq.trainer | begin training epoch 26
2021-03-27 21:22:37 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 21:23:12 | INFO | train_inner | epoch 026:   100 / 588 loss=5.956, nll_loss=4.898, word_ins=6.055, word_crf=2.571, length=3.579, ppl=62.09, wps=17345.2, ups=2.39, wpb=7264.3, bsz=287, num_updates=14800, lr=0.000410997, gnorm=1.461, train_wall=35, gb_free=9.7, wall=4940
2021-03-27 21:23:47 | INFO | train_inner | epoch 026:   200 / 588 loss=6.051, nll_loss=4.957, word_ins=6.107, word_crf=2.617, length=3.802, ppl=66.3, wps=20517.5, ups=2.89, wpb=7108.4, bsz=270.9, num_updates=14900, lr=0.000409616, gnorm=1.481, train_wall=34, gb_free=10, wall=4975
2021-03-27 21:24:23 | INFO | train_inner | epoch 026:   300 / 588 loss=6.392, nll_loss=5.317, word_ins=6.419, word_crf=2.784, length=3.978, ppl=83.95, wps=19099, ups=2.81, wpb=6803.3, bsz=227.2, num_updates=15000, lr=0.000408248, gnorm=1.316, train_wall=35, gb_free=9.3, wall=5010
2021-03-27 21:24:58 | INFO | train_inner | epoch 026:   400 / 588 loss=6.151, nll_loss=5.068, word_ins=6.2, word_crf=2.665, length=3.857, ppl=71.08, wps=20041.7, ups=2.85, wpb=7029.9, bsz=269.4, num_updates=15100, lr=0.000406894, gnorm=1.315, train_wall=35, gb_free=10.1, wall=5045
2021-03-27 21:25:32 | INFO | train_inner | epoch 026:   500 / 588 loss=6.008, nll_loss=4.918, word_ins=6.072, word_crf=2.595, length=3.772, ppl=64.37, wps=19766.1, ups=2.88, wpb=6858.2, bsz=284, num_updates=15200, lr=0.000405554, gnorm=1.441, train_wall=34, gb_free=9.2, wall=5080
2021-03-27 21:26:02 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 21:26:06 | INFO | valid | epoch 026 | valid on 'valid' subset | loss 7.068 | nll_loss 5.491 | word_ins 6.642 | word_crf 3.317 | length 4.3 | ppl 134.16 | wps 60208.8 | wpb 5311.6 | bsz 208.1 | num_updates 15288 | best_loss 6.892
2021-03-27 21:26:06 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 26 @ 15288 updates
2021-03-27 21:26:06 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint26.pt
2021-03-27 21:26:07 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint26.pt
2021-03-27 21:26:09 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint26.pt (epoch 26 @ 15288 updates, score 7.068) (writing took 3.6307868730000337 seconds)
2021-03-27 21:26:09 | INFO | fairseq_cli.train | end of epoch 26 (average epoch stats below)
2021-03-27 21:26:09 | INFO | train | epoch 026 | loss 6.079 | nll_loss 4.997 | word_ins 6.14 | word_crf 2.63 | length 3.79 | ppl 67.6 | wps 19378.3 | ups 2.77 | wpb 6988.7 | bsz 272.5 | num_updates 15288 | lr 0.000404385 | gnorm 1.381 | train_wall 203 | gb_free 9.6 | wall 5117
2021-03-27 21:26:09 | INFO | fairseq.trainer | begin training epoch 27
2021-03-27 21:26:09 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 21:26:14 | INFO | train_inner | epoch 027:   12 / 588 loss=5.858, nll_loss=4.768, word_ins=5.94, word_crf=2.512, length=3.763, ppl=58, wps=16577.3, ups=2.42, wpb=6851.8, bsz=302.5, num_updates=15300, lr=0.000404226, gnorm=1.28, train_wall=34, gb_free=9.2, wall=5122
2021-03-27 21:26:48 | INFO | train_inner | epoch 027:   112 / 588 loss=5.921, nll_loss=4.854, word_ins=6.02, word_crf=2.535, length=3.758, ppl=60.58, wps=20150.1, ups=2.89, wpb=6982.3, bsz=278.8, num_updates=15400, lr=0.000402911, gnorm=1.563, train_wall=34, gb_free=10.6, wall=5156
2021-03-27 21:27:23 | INFO | train_inner | epoch 027:   212 / 588 loss=5.917, nll_loss=4.849, word_ins=6.016, word_crf=2.546, length=3.635, ppl=60.43, wps=20128.4, ups=2.85, wpb=7056.9, bsz=280.9, num_updates=15500, lr=0.00040161, gnorm=1.373, train_wall=35, gb_free=9.8, wall=5191
2021-03-27 21:27:58 | INFO | train_inner | epoch 027:   312 / 588 loss=6.05, nll_loss=4.983, word_ins=6.129, word_crf=2.615, length=3.704, ppl=66.26, wps=20409.3, ups=2.89, wpb=7066.6, bsz=275.8, num_updates=15600, lr=0.00040032, gnorm=1.677, train_wall=34, gb_free=9.7, wall=5226
2021-03-27 21:28:34 | INFO | train_inner | epoch 027:   412 / 588 loss=6.104, nll_loss=5.031, word_ins=6.172, word_crf=2.635, length=3.833, ppl=68.8, wps=19577.3, ups=2.8, wpb=6979.7, bsz=258.9, num_updates=15700, lr=0.000399043, gnorm=1.361, train_wall=35, gb_free=9.7, wall=5262
2021-03-27 21:29:08 | INFO | train_inner | epoch 027:   512 / 588 loss=6.032, nll_loss=4.955, word_ins=6.105, word_crf=2.601, length=3.791, ppl=65.45, wps=19856.8, ups=2.88, wpb=6892.9, bsz=274.8, num_updates=15800, lr=0.000397779, gnorm=1.421, train_wall=34, gb_free=9.6, wall=5296
2021-03-27 21:29:35 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 21:29:38 | INFO | valid | epoch 027 | valid on 'valid' subset | loss 6.994 | nll_loss 5.457 | word_ins 6.614 | word_crf 3.258 | length 4.286 | ppl 127.48 | wps 60051 | wpb 5311.6 | bsz 208.1 | num_updates 15876 | best_loss 6.892
2021-03-27 21:29:38 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 27 @ 15876 updates
2021-03-27 21:29:38 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint27.pt
2021-03-27 21:29:39 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint27.pt
2021-03-27 21:29:42 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint27.pt (epoch 27 @ 15876 updates, score 6.994) (writing took 3.654849474999537 seconds)
2021-03-27 21:29:42 | INFO | fairseq_cli.train | end of epoch 27 (average epoch stats below)
2021-03-27 21:29:42 | INFO | train | epoch 027 | loss 6.016 | nll_loss 4.945 | word_ins 6.097 | word_crf 2.591 | length 3.762 | ppl 64.7 | wps 19336 | ups 2.77 | wpb 6988.7 | bsz 272.5 | num_updates 15876 | lr 0.000396825 | gnorm 1.468 | train_wall 204 | gb_free 9.3 | wall 5330
2021-03-27 21:29:42 | INFO | fairseq.trainer | begin training epoch 28
2021-03-27 21:29:42 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 21:29:51 | INFO | train_inner | epoch 028:   24 / 588 loss=6.087, nll_loss=5.029, word_ins=6.17, word_crf=2.623, length=3.792, ppl=68, wps=16601.8, ups=2.37, wpb=7017.6, bsz=258.7, num_updates=15900, lr=0.000396526, gnorm=1.364, train_wall=35, gb_free=9.6, wall=5338
2021-03-27 21:30:26 | INFO | train_inner | epoch 028:   124 / 588 loss=5.996, nll_loss=4.948, word_ins=6.104, word_crf=2.577, length=3.666, ppl=63.81, wps=19823.3, ups=2.85, wpb=6964.2, bsz=258.1, num_updates=16000, lr=0.000395285, gnorm=1.615, train_wall=35, gb_free=9.6, wall=5374
2021-03-27 21:31:00 | INFO | train_inner | epoch 028:   224 / 588 loss=5.716, nll_loss=4.654, word_ins=5.847, word_crf=2.429, length=3.637, ppl=52.57, wps=20293.7, ups=2.9, wpb=6995.7, bsz=309.2, num_updates=16100, lr=0.000394055, gnorm=1.426, train_wall=34, gb_free=9.7, wall=5408
2021-03-27 21:31:35 | INFO | train_inner | epoch 028:   324 / 588 loss=5.959, nll_loss=4.891, word_ins=6.055, word_crf=2.553, length=3.781, ppl=62.2, wps=20179.1, ups=2.86, wpb=7050, bsz=259, num_updates=16200, lr=0.000392837, gnorm=1.296, train_wall=35, gb_free=9.5, wall=5443
2021-03-27 21:32:10 | INFO | train_inner | epoch 028:   424 / 588 loss=5.967, nll_loss=4.908, word_ins=6.067, word_crf=2.562, length=3.722, ppl=62.56, wps=20157.6, ups=2.85, wpb=7061, bsz=275.6, num_updates=16300, lr=0.00039163, gnorm=1.42, train_wall=35, gb_free=9.5, wall=5478
2021-03-27 21:32:45 | INFO | train_inner | epoch 028:   524 / 588 loss=6.006, nll_loss=4.936, word_ins=6.091, word_crf=2.574, length=3.863, ppl=64.27, wps=19819.8, ups=2.87, wpb=6906.1, bsz=259.6, num_updates=16400, lr=0.000390434, gnorm=1.401, train_wall=35, gb_free=10.6, wall=5513
2021-03-27 21:33:07 | INFO | fairseq_cli.train | begin validation on "valid" subset

2021-03-27 21:33:11 | INFO | valid | epoch 028 | valid on 'valid' subset | loss 7.005 | nll_loss 5.441 | word_ins 6.601 | word_crf 3.282 | length 4.221 | ppl 128.45 | wps 60036.5 | wpb 5311.6 | bsz 208.1 | num_updates 16464 | best_loss 6.892
2021-03-27 21:33:11 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 28 @ 16464 updates
2021-03-27 21:33:11 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint28.pt
2021-03-27 21:33:12 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint28.pt
2021-03-27 21:33:14 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint28.pt (epoch 28 @ 16464 updates, score 7.005) (writing took 3.6403963160000785 seconds)
2021-03-27 21:33:14 | INFO | fairseq_cli.train | end of epoch 28 (average epoch stats below)
2021-03-27 21:33:14 | INFO | train | epoch 028 | loss 5.924 | nll_loss 4.863 | word_ins 6.029 | word_crf 2.536 | length 3.739 | ppl 60.73 | wps 19344.3 | ups 2.77 | wpb 6988.7 | bsz 272.5 | num_updates 16464 | lr 0.000389675 | gnorm 1.416 | train_wall 204 | gb_free 9.3 | wall 5542
2021-03-27 21:33:14 | INFO | fairseq.trainer | begin training epoch 29
2021-03-27 21:33:14 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 21:33:27 | INFO | train_inner | epoch 029:   36 / 588 loss=5.917, nll_loss=4.852, word_ins=6.019, word_crf=2.523, length=3.843, ppl=60.4, wps=16449.1, ups=2.38, wpb=6921.3, bsz=268.6, num_updates=16500, lr=0.000389249, gnorm=1.492, train_wall=35, gb_free=10.5, wall=5555
2021-03-27 21:34:02 | INFO | train_inner | epoch 029:   136 / 588 loss=5.752, nll_loss=4.705, word_ins=5.895, word_crf=2.443, length=3.616, ppl=53.88, wps=20403, ups=2.89, wpb=7051.1, bsz=278.4, num_updates=16600, lr=0.000388075, gnorm=1.478, train_wall=34, gb_free=9.2, wall=5590
2021-03-27 21:34:37 | INFO | train_inner | epoch 029:   236 / 588 loss=5.824, nll_loss=4.766, word_ins=5.948, word_crf=2.475, length=3.749, ppl=56.65, wps=19614.4, ups=2.82, wpb=6955.4, bsz=280.4, num_updates=16700, lr=0.000386912, gnorm=1.586, train_wall=35, gb_free=9.6, wall=5625
2021-03-27 21:35:12 | INFO | train_inner | epoch 029:   336 / 588 loss=5.768, nll_loss=4.706, word_ins=5.895, word_crf=2.448, length=3.726, ppl=54.48, wps=19965.1, ups=2.89, wpb=6916.7, bsz=282.6, num_updates=16800, lr=0.000385758, gnorm=1.435, train_wall=34, gb_free=9.3, wall=5660
2021-03-27 21:35:46 | INFO | train_inner | epoch 029:   436 / 588 loss=5.82, nll_loss=4.757, word_ins=5.938, word_crf=2.483, length=3.681, ppl=56.49, wps=19982.5, ups=2.88, wpb=6940.4, bsz=285.4, num_updates=16900, lr=0.000384615, gnorm=1.467, train_wall=34, gb_free=9.4, wall=5694
2021-03-27 21:36:22 | INFO | train_inner | epoch 029:   536 / 588 loss=6.132, nll_loss=5.104, word_ins=6.238, word_crf=2.644, length=3.683, ppl=70.12, wps=19645.6, ups=2.8, wpb=7021.3, bsz=247.1, num_updates=17000, lr=0.000383482, gnorm=1.485, train_wall=35, gb_free=9.8, wall=5730
2021-03-27 21:36:40 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 21:36:43 | INFO | valid | epoch 029 | valid on 'valid' subset | loss 6.867 | nll_loss 5.445 | word_ins 6.59 | word_crf 3.144 | length 4.276 | ppl 116.7 | wps 60130.8 | wpb 5311.6 | bsz 208.1 | num_updates 17052 | best_loss 6.867
2021-03-27 21:36:43 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 29 @ 17052 updates
2021-03-27 21:36:43 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint29.pt
2021-03-27 21:36:44 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint29.pt
2021-03-27 21:36:50 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint29.pt (epoch 29 @ 17052 updates, score 6.867) (writing took 6.623916104000273 seconds)
2021-03-27 21:36:50 | INFO | fairseq_cli.train | end of epoch 29 (average epoch stats below)
2021-03-27 21:36:50 | INFO | train | epoch 029 | loss 5.877 | nll_loss 4.824 | word_ins 5.996 | word_crf 2.507 | length 3.719 | ppl 58.77 | wps 19056.3 | ups 2.73 | wpb 6988.7 | bsz 272.5 | num_updates 17052 | lr 0.000382897 | gnorm 1.516 | train_wall 204 | gb_free 9.2 | wall 5758
2021-03-27 21:36:50 | INFO | fairseq.trainer | begin training epoch 30
2021-03-27 21:36:50 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 21:37:07 | INFO | train_inner | epoch 030:   48 / 588 loss=5.915, nll_loss=4.862, word_ins=6.031, word_crf=2.521, length=3.783, ppl=60.34, wps=15676.5, ups=2.23, wpb=7027.5, bsz=266.2, num_updates=17100, lr=0.00038236, gnorm=1.651, train_wall=34, gb_free=10.1, wall=5775
2021-03-27 21:37:42 | INFO | train_inner | epoch 030:   148 / 588 loss=5.535, nll_loss=4.495, word_ins=5.715, word_crf=2.322, length=3.56, ppl=46.38, wps=20796.6, ups=2.89, wpb=7186.3, bsz=298.6, num_updates=17200, lr=0.000381246, gnorm=1.229, train_wall=34, gb_free=9.2, wall=5810
2021-03-27 21:38:17 | INFO | train_inner | epoch 030:   248 / 588 loss=5.991, nll_loss=4.941, word_ins=6.104, word_crf=2.557, length=3.823, ppl=63.59, wps=19561.9, ups=2.85, wpb=6870.6, bsz=230.5, num_updates=17300, lr=0.000380143, gnorm=1.445, train_wall=35, gb_free=9.6, wall=5845
2021-03-27 21:38:51 | INFO | train_inner | epoch 030:   348 / 588 loss=5.696, nll_loss=4.651, word_ins=5.849, word_crf=2.405, length=3.668, ppl=51.85, wps=19947.2, ups=2.89, wpb=6891.2, bsz=289, num_updates=17400, lr=0.000379049, gnorm=1.329, train_wall=34, gb_free=9.7, wall=5879
2021-03-27 21:39:26 | INFO | train_inner | epoch 030:   448 / 588 loss=5.812, nll_loss=4.764, word_ins=5.948, word_crf=2.473, length=3.657, ppl=56.19, wps=20231.6, ups=2.86, wpb=7070, bsz=267, num_updates=17500, lr=0.000377964, gnorm=1.299, train_wall=35, gb_free=9.5, wall=5914
2021-03-27 21:40:01 | INFO | train_inner | epoch 030:   548 / 588 loss=5.855, nll_loss=4.815, word_ins=5.99, word_crf=2.486, length=3.731, ppl=57.87, wps=20126.8, ups=2.87, wpb=7001.9, bsz=269.1, num_updates=17600, lr=0.000376889, gnorm=1.354, train_wall=35, gb_free=9.5, wall=5949
2021-03-27 21:40:15 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 21:40:18 | INFO | valid | epoch 030 | valid on 'valid' subset | loss 6.96 | nll_loss 5.407 | word_ins 6.576 | word_crf 3.248 | length 4.248 | ppl 124.52 | wps 60061.8 | wpb 5311.6 | bsz 208.1 | num_updates 17640 | best_loss 6.867
2021-03-27 21:40:18 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 30 @ 17640 updates
2021-03-27 21:40:18 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint30.pt
2021-03-27 21:40:19 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint30.pt
2021-03-27 21:40:22 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint30.pt (epoch 30 @ 17640 updates, score 6.96) (writing took 3.6386014529998647 seconds)
2021-03-27 21:40:22 | INFO | fairseq_cli.train | end of epoch 30 (average epoch stats below)
2021-03-27 21:40:22 | INFO | train | epoch 030 | loss 5.773 | nll_loss 4.73 | word_ins 5.918 | word_crf 2.445 | length 3.686 | ppl 54.68 | wps 19394.4 | ups 2.78 | wpb 6988.7 | bsz 272.5 | num_updates 17640 | lr 0.000376462 | gnorm 1.361 | train_wall 203 | gb_free 9.8 | wall 5970
2021-03-27 21:40:22 | INFO | fairseq.trainer | begin training epoch 31
2021-03-27 21:40:22 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 21:40:43 | INFO | train_inner | epoch 031:   60 / 588 loss=5.719, nll_loss=4.685, word_ins=5.879, word_crf=2.411, length=3.679, ppl=52.66, wps=16632, ups=2.37, wpb=7022.9, bsz=280.2, num_updates=17700, lr=0.000375823, gnorm=1.56, train_wall=35, gb_free=9.5, wall=5991
2021-03-27 21:41:18 | INFO | train_inner | epoch 031:   160 / 588 loss=5.541, nll_loss=4.502, word_ins=5.724, word_crf=2.317, length=3.62, ppl=46.57, wps=20292.1, ups=2.91, wpb=6975.6, bsz=286.6, num_updates=17800, lr=0.000374766, gnorm=1.364, train_wall=34, gb_free=9.4, wall=6026
2021-03-27 21:41:53 | INFO | train_inner | epoch 031:   260 / 588 loss=5.666, nll_loss=4.63, word_ins=5.834, word_crf=2.394, length=3.55, ppl=50.77, wps=19816.2, ups=2.87, wpb=6915.4, bsz=278.9, num_updates=17900, lr=0.000373718, gnorm=1.399, train_wall=35, gb_free=9.9, wall=6060
2021-03-27 21:42:28 | INFO | train_inner | epoch 031:   360 / 588 loss=5.762, nll_loss=4.733, word_ins=5.923, word_crf=2.437, length=3.64, ppl=54.28, wps=19892.3, ups=2.85, wpb=6986.6, bsz=273.8, num_updates=18000, lr=0.000372678, gnorm=1.439, train_wall=35, gb_free=9.3, wall=6096
2021-03-27 21:43:03 | INFO | train_inner | epoch 031:   460 / 588 loss=5.871, nll_loss=4.84, word_ins=6.016, word_crf=2.493, length=3.697, ppl=58.54, wps=19827.5, ups=2.85, wpb=6955.3, bsz=247.4, num_updates=18100, lr=0.000371647, gnorm=1.408, train_wall=35, gb_free=9.5, wall=6131
2021-03-27 21:43:37 | INFO | train_inner | epoch 031:   560 / 588 loss=5.762, nll_loss=4.716, word_ins=5.907, word_crf=2.433, length=3.756, ppl=54.27, wps=20346.9, ups=2.91, wpb=7002.5, bsz=269.6, num_updates=18200, lr=0.000370625, gnorm=1.439, train_wall=34, gb_free=10.1, wall=6165
2021-03-27 21:43:47 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 21:43:50 | INFO | valid | epoch 031 | valid on 'valid' subset | loss 7.088 | nll_loss 5.422 | word_ins 6.599 | word_crf 3.356 | length 4.324 | ppl 136.04 | wps 60210.9 | wpb 5311.6 | bsz 208.1 | num_updates 18228 | best_loss 6.867
2021-03-27 21:43:50 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 31 @ 18228 updates
2021-03-27 21:43:50 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint31.pt
2021-03-27 21:43:51 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint31.pt
2021-03-27 21:43:53 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint31.pt (epoch 31 @ 18228 updates, score 7.088) (writing took 3.643357536999247 seconds)
2021-03-27 21:43:53 | INFO | fairseq_cli.train | end of epoch 31 (average epoch stats below)
2021-03-27 21:43:53 | INFO | train | epoch 031 | loss 5.718 | nll_loss 4.682 | word_ins 5.879 | word_crf 2.412 | length 3.659 | ppl 52.62 | wps 19415.7 | ups 2.78 | wpb 6988.7 | bsz 272.5 | num_updates 18228 | lr 0.00037034 | gnorm 1.433 | train_wall 203 | gb_free 9.6 | wall 6181
2021-03-27 21:43:54 | INFO | fairseq.trainer | begin training epoch 32
2021-03-27 21:43:54 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 21:44:19 | INFO | train_inner | epoch 032:   72 / 588 loss=5.571, nll_loss=4.546, word_ins=5.763, word_crf=2.324, length=3.652, ppl=47.54, wps=16872.9, ups=2.4, wpb=7019.1, bsz=278.9, num_updates=18300, lr=0.000369611, gnorm=1.343, train_wall=34, gb_free=9.3, wall=6207
2021-03-27 21:44:53 | INFO | train_inner | epoch 032:   172 / 588 loss=5.582, nll_loss=4.555, word_ins=5.772, word_crf=2.334, length=3.623, ppl=47.9, wps=20052, ups=2.88, wpb=6955.7, bsz=273.9, num_updates=18400, lr=0.000368605, gnorm=1.405, train_wall=34, gb_free=10, wall=6241

2021-03-27 21:45:29 | INFO | train_inner | epoch 032:   272 / 588 loss=5.718, nll_loss=4.697, word_ins=5.896, word_crf=2.41, length=3.607, ppl=52.65, wps=19838, ups=2.85, wpb=6957.8, bsz=256.9, num_updates=18500, lr=0.000367607, gnorm=1.396, train_wall=35, gb_free=9.8, wall=6276
2021-03-27 21:46:03 | INFO | train_inner | epoch 032:   372 / 588 loss=5.695, nll_loss=4.664, word_ins=5.866, word_crf=2.399, length=3.625, ppl=51.79, wps=19687.6, ups=2.86, wpb=6886.9, bsz=256, num_updates=18600, lr=0.000366618, gnorm=1.333, train_wall=35, gb_free=9.3, wall=6311
2021-03-27 21:46:39 | INFO | train_inner | epoch 032:   472 / 588 loss=5.737, nll_loss=4.707, word_ins=5.901, word_crf=2.421, length=3.658, ppl=53.34, wps=19737, ups=2.85, wpb=6919.4, bsz=278.5, num_updates=18700, lr=0.000365636, gnorm=1.666, train_wall=35, gb_free=9.4, wall=6346
2021-03-27 21:47:13 | INFO | train_inner | epoch 032:   572 / 588 loss=5.493, nll_loss=4.457, word_ins=5.685, word_crf=2.291, length=3.589, ppl=45.02, wps=20953.7, ups=2.92, wpb=7176.2, bsz=296.7, num_updates=18800, lr=0.000364662, gnorm=1.16, train_wall=34, gb_free=9.3, wall=6381
2021-03-27 21:47:19 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 21:47:22 | INFO | valid | epoch 032 | valid on 'valid' subset | loss 7.096 | nll_loss 5.436 | word_ins 6.621 | word_crf 3.356 | length 4.299 | ppl 136.8 | wps 60041.3 | wpb 5311.6 | bsz 208.1 | num_updates 18816 | best_loss 6.867
2021-03-27 21:47:22 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 32 @ 18816 updates
2021-03-27 21:47:22 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint32.pt
2021-03-27 21:47:23 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint32.pt
2021-03-27 21:47:25 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint32.pt (epoch 32 @ 18816 updates, score 7.096) (writing took 3.6426531659999455 seconds)
2021-03-27 21:47:25 | INFO | fairseq_cli.train | end of epoch 32 (average epoch stats below)
2021-03-27 21:47:25 | INFO | train | epoch 032 | loss 5.638 | nll_loss 4.611 | word_ins 5.819 | word_crf 2.367 | length 3.615 | ppl 49.79 | wps 19387.3 | ups 2.77 | wpb 6988.7 | bsz 272.5 | num_updates 18816 | lr 0.000364507 | gnorm 1.387 | train_wall 203 | gb_free 9.8 | wall 6393
2021-03-27 21:47:26 | INFO | fairseq.trainer | begin training epoch 33
2021-03-27 21:47:26 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 21:47:55 | INFO | train_inner | epoch 033:   84 / 588 loss=5.534, nll_loss=4.514, word_ins=5.737, word_crf=2.306, length=3.596, ppl=46.35, wps=16872.8, ups=2.39, wpb=7050.6, bsz=282.2, num_updates=18900, lr=0.000363696, gnorm=1.521, train_wall=34, gb_free=9.9, wall=6422
2021-03-27 21:48:29 | INFO | train_inner | epoch 033:   184 / 588 loss=5.456, nll_loss=4.442, word_ins=5.675, word_crf=2.261, length=3.576, ppl=43.9, wps=20145.6, ups=2.89, wpb=6979.6, bsz=288.2, num_updates=19000, lr=0.000362738, gnorm=1.371, train_wall=34, gb_free=9.8, wall=6457
2021-03-27 21:49:04 | INFO | train_inner | epoch 033:   284 / 588 loss=5.564, nll_loss=4.543, word_ins=5.764, word_crf=2.327, length=3.553, ppl=47.31, wps=20016.4, ups=2.86, wpb=6992.9, bsz=269.8, num_updates=19100, lr=0.000361787, gnorm=1.468, train_wall=35, gb_free=9.4, wall=6492
2021-03-27 21:49:40 | INFO | train_inner | epoch 033:   384 / 588 loss=5.698, nll_loss=4.687, word_ins=5.886, word_crf=2.407, length=3.49, ppl=51.93, wps=19378.2, ups=2.78, wpb=6965.8, bsz=269.2, num_updates=19200, lr=0.000360844, gnorm=1.494, train_wall=36, gb_free=10.3, wall=6528
2021-03-27 21:50:15 | INFO | train_inner | epoch 033:   484 / 588 loss=5.723, nll_loss=4.693, word_ins=5.893, word_crf=2.411, length=3.66, ppl=52.83, wps=20136, ups=2.87, wpb=7014.1, bsz=251.5, num_updates=19300, lr=0.000359908, gnorm=1.441, train_wall=35, gb_free=9.3, wall=6563
2021-03-27 21:50:50 | INFO | train_inner | epoch 033:   584 / 588 loss=5.584, nll_loss=4.556, word_ins=5.773, word_crf=2.332, length=3.657, ppl=47.97, wps=19953.4, ups=2.87, wpb=6959.8, bsz=275.9, num_updates=19400, lr=0.000358979, gnorm=1.378, train_wall=35, gb_free=9.3, wall=6598
2021-03-27 21:50:51 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 21:50:54 | INFO | valid | epoch 033 | valid on 'valid' subset | loss 7.152 | nll_loss 5.421 | word_ins 6.626 | word_crf 3.407 | length 4.328 | ppl 142.25 | wps 59975.6 | wpb 5311.6 | bsz 208.1 | num_updates 19404 | best_loss 6.867
2021-03-27 21:50:54 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 33 @ 19404 updates
2021-03-27 21:50:54 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint33.pt
2021-03-27 21:50:55 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint33.pt
2021-03-27 21:50:58 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint33.pt (epoch 33 @ 19404 updates, score 7.152) (writing took 3.6411032819996763 seconds)
2021-03-27 21:50:58 | INFO | fairseq_cli.train | end of epoch 33 (average epoch stats below)
2021-03-27 21:50:58 | INFO | train | epoch 033 | loss 5.588 | nll_loss 4.567 | word_ins 5.783 | word_crf 2.337 | length 3.593 | ppl 48.1 | wps 19322.2 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 19404 | lr 0.000358942 | gnorm 1.449 | train_wall 204 | gb_free 10.1 | wall 6606
2021-03-27 21:50:58 | INFO | fairseq.trainer | begin training epoch 34
2021-03-27 21:50:58 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 21:51:32 | INFO | train_inner | epoch 034:   96 / 588 loss=5.552, nll_loss=4.548, word_ins=5.77, word_crf=2.306, length=3.606, ppl=46.9, wps=16333, ups=2.35, wpb=6941.8, bsz=252.8, num_updates=19500, lr=0.000358057, gnorm=1.492, train_wall=35, gb_free=9.4, wall=6640
2021-03-27 21:52:07 | INFO | train_inner | epoch 034:   196 / 588 loss=5.428, nll_loss=4.41, word_ins=5.653, word_crf=2.252, length=3.501, ppl=43.06, wps=20664.8, ups=2.91, wpb=7093.8, bsz=279, num_updates=19600, lr=0.000357143, gnorm=1.606, train_wall=34, gb_free=9.9, wall=6675
2021-03-27 21:52:41 | INFO | train_inner | epoch 034:   296 / 588 loss=5.385, nll_loss=4.372, word_ins=5.617, word_crf=2.227, length=3.487, ppl=41.78, wps=20262.7, ups=2.87, wpb=7057.2, bsz=300.6, num_updates=19700, lr=0.000356235, gnorm=1.404, train_wall=35, gb_free=10.4, wall=6709
2021-03-27 21:53:17 | INFO | train_inner | epoch 034:   396 / 588 loss=5.63, nll_loss=4.611, word_ins=5.824, word_crf=2.356, length=3.62, ppl=49.53, wps=20016.8, ups=2.85, wpb=7017.2, bsz=261.4, num_updates=19800, lr=0.000355335, gnorm=1.296, train_wall=35, gb_free=9.3, wall=6744
2021-03-27 21:53:51 | INFO | train_inner | epoch 034:   496 / 588 loss=5.541, nll_loss=4.519, word_ins=5.743, word_crf=2.306, length=3.64, ppl=46.56, wps=19838.2, ups=2.89, wpb=6860.7, bsz=268.4, num_updates=19900, lr=0.000354441, gnorm=1.494, train_wall=34, gb_free=9.4, wall=6779
2021-03-27 21:54:23 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 21:54:27 | INFO | valid | epoch 034 | valid on 'valid' subset | loss 7.057 | nll_loss 5.389 | word_ins 6.582 | word_crf 3.33 | length 4.366 | ppl 133.19 | wps 59440.5 | wpb 5311.6 | bsz 208.1 | num_updates 19992 | best_loss 6.867
2021-03-27 21:54:27 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 34 @ 19992 updates
2021-03-27 21:54:27 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint34.pt
2021-03-27 21:54:27 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint34.pt
2021-03-27 21:54:30 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint34.pt (epoch 34 @ 19992 updates, score 7.057) (writing took 3.6598553800004083 seconds)
2021-03-27 21:54:30 | INFO | fairseq_cli.train | end of epoch 34 (average epoch stats below)
2021-03-27 21:54:30 | INFO | train | epoch 034 | loss 5.534 | nll_loss 4.517 | word_ins 5.742 | word_crf 2.304 | length 3.583 | ppl 46.33 | wps 19368.3 | ups 2.77 | wpb 6988.7 | bsz 272.5 | num_updates 19992 | lr 0.000353624 | gnorm 1.466 | train_wall 203 | gb_free 9.5 | wall 6818
2021-03-27 21:54:30 | INFO | fairseq.trainer | begin training epoch 35
2021-03-27 21:54:30 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 21:54:33 | INFO | train_inner | epoch 035:   8 / 588 loss=5.641, nll_loss=4.617, word_ins=5.827, word_crf=2.363, length=3.643, ppl=49.89, wps=16548.5, ups=2.37, wpb=6968.4, bsz=274.7, num_updates=20000, lr=0.000353553, gnorm=1.493, train_wall=35, gb_free=9.3, wall=6821
2021-03-27 21:54:33 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 21:54:37 | INFO | valid | epoch 035 | valid on 'valid' subset | loss 7.168 | nll_loss 5.405 | word_ins 6.618 | word_crf 3.422 | length 4.375 | ppl 143.81 | wps 59890.4 | wpb 5311.6 | bsz 208.1 | num_updates 20000 | best_loss 6.867
2021-03-27 21:54:37 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 35 @ 20000 updates
2021-03-27 21:54:37 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint_35_20000.pt
2021-03-27 21:54:38 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint_35_20000.pt
2021-03-27 21:54:41 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint_35_20000.pt (epoch 35 @ 20000 updates, score 7.168) (writing took 4.00380122200022 seconds)
2021-03-27 21:55:16 | INFO | train_inner | epoch 035:   108 / 588 loss=5.487, nll_loss=4.502, word_ins=5.732, word_crf=2.288, length=3.332, ppl=44.85, wps=16566.7, ups=2.35, wpb=7059.9, bsz=267.7, num_updates=20100, lr=0.000352673, gnorm=1.473, train_wall=35, gb_free=9.7, wall=6864
2021-03-27 21:55:51 | INFO | train_inner | epoch 035:   208 / 588 loss=5.397, nll_loss=4.379, word_ins=5.626, word_crf=2.219, length=3.654, ppl=42.15, wps=19863.7, ups=2.88, wpb=6903.1, bsz=271.8, num_updates=20200, lr=0.000351799, gnorm=1.405, train_wall=34, gb_free=9.2, wall=6899
2021-03-27 21:56:25 | INFO | train_inner | epoch 035:   308 / 588 loss=5.329, nll_loss=4.322, word_ins=5.575, word_crf=2.188, length=3.538, ppl=40.2, wps=20204, ups=2.88, wpb=7021.3, bsz=292.2, num_updates=20300, lr=0.000350931, gnorm=1.39, train_wall=34, gb_free=10.1, wall=6933
2021-03-27 21:57:01 | INFO | train_inner | epoch 035:   408 / 588 loss=5.688, nll_loss=4.688, word_ins=5.894, word_crf=2.387, length=3.542, ppl=51.56, wps=19378.8, ups=2.82, wpb=6861.3, bsz=234, num_updates=20400, lr=0.00035007, gnorm=1.493, train_wall=35, gb_free=9.8, wall=6969
2021-03-27 21:57:36 | INFO | train_inner | epoch 035:   508 / 588 loss=5.416, nll_loss=4.4, word_ins=5.641, word_crf=2.237, length=3.589, ppl=42.7, wps=19831, ups=2.87, wpb=6919.2, bsz=290.6, num_updates=20500, lr=0.000349215, gnorm=1.358, train_wall=35, gb_free=9.9, wall=7004
2021-03-27 21:58:04 | INFO | fairseq_cli.train | begin validation on "valid" subset

2021-03-27 21:58:07 | INFO | valid | epoch 035 | valid on 'valid' subset | loss 7.118 | nll_loss 5.431 | word_ins 6.626 | word_crf 3.364 | length 4.408 | ppl 138.88 | wps 59940.1 | wpb 5311.6 | bsz 208.1 | num_updates 20580 | best_loss 6.867
2021-03-27 21:58:07 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 35 @ 20580 updates
2021-03-27 21:58:07 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint35.pt
2021-03-27 21:58:08 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint35.pt
2021-03-27 21:58:11 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint35.pt (epoch 35 @ 20580 updates, score 7.118) (writing took 3.645485195999754 seconds)
2021-03-27 21:58:11 | INFO | fairseq_cli.train | end of epoch 35 (average epoch stats below)
2021-03-27 21:58:11 | INFO | train | epoch 035 | loss 5.47 | nll_loss 4.463 | word_ins 5.698 | word_crf 2.268 | length 3.532 | ppl 44.34 | wps 18645.3 | ups 2.67 | wpb 6988.7 | bsz 272.5 | num_updates 20580 | lr 0.000348536 | gnorm 1.436 | train_wall 204 | gb_free 9.9 | wall 7039
2021-03-27 21:58:11 | INFO | fairseq.trainer | begin training epoch 36
2021-03-27 21:58:11 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 21:58:18 | INFO | train_inner | epoch 036:   20 / 588 loss=5.494, nll_loss=4.487, word_ins=5.716, word_crf=2.283, length=3.526, ppl=45.07, wps=16758.6, ups=2.37, wpb=7065.8, bsz=273.1, num_updates=20600, lr=0.000348367, gnorm=1.516, train_wall=35, gb_free=9.9, wall=7046
2021-03-27 21:58:52 | INFO | train_inner | epoch 036:   120 / 588 loss=5.241, nll_loss=4.23, word_ins=5.501, word_crf=2.138, length=3.523, ppl=37.81, wps=20561.4, ups=2.9, wpb=7091.5, bsz=281.9, num_updates=20700, lr=0.000347524, gnorm=1.448, train_wall=34, gb_free=9.4, wall=7080
2021-03-27 21:59:28 | INFO | train_inner | epoch 036:   220 / 588 loss=5.408, nll_loss=4.417, word_ins=5.662, word_crf=2.233, length=3.438, ppl=42.45, wps=19703.8, ups=2.84, wpb=6940.5, bsz=264.1, num_updates=20800, lr=0.000346688, gnorm=1.469, train_wall=35, gb_free=9.6, wall=7115
2021-03-27 22:00:03 | INFO | train_inner | epoch 036:   320 / 588 loss=5.459, nll_loss=4.463, word_ins=5.7, word_crf=2.259, length=3.503, ppl=43.99, wps=19500.8, ups=2.82, wpb=6916, bsz=270.1, num_updates=20900, lr=0.000345857, gnorm=1.48, train_wall=35, gb_free=9.2, wall=7151
2021-03-27 22:00:38 | INFO | train_inner | epoch 036:   420 / 588 loss=5.427, nll_loss=4.432, word_ins=5.673, word_crf=2.253, length=3.376, ppl=43.03, wps=19920.3, ups=2.85, wpb=6995.2, bsz=275.8, num_updates=21000, lr=0.000345033, gnorm=1.411, train_wall=35, gb_free=9.6, wall=7186
2021-03-27 22:01:13 | INFO | train_inner | epoch 036:   520 / 588 loss=5.372, nll_loss=4.35, word_ins=5.6, word_crf=2.21, length=3.629, ppl=41.43, wps=20329, ups=2.89, wpb=7027.3, bsz=289.1, num_updates=21100, lr=0.000344214, gnorm=1.485, train_wall=34, gb_free=9.4, wall=7221
2021-03-27 22:01:37 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 22:01:40 | INFO | valid | epoch 036 | valid on 'valid' subset | loss 7.184 | nll_loss 5.415 | word_ins 6.623 | word_crf 3.438 | length 4.34 | ppl 145.38 | wps 60057.6 | wpb 5311.6 | bsz 208.1 | num_updates 21168 | best_loss 6.867
2021-03-27 22:01:40 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 36 @ 21168 updates
2021-03-27 22:01:40 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint36.pt
2021-03-27 22:01:41 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint36.pt
2021-03-27 22:01:44 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint36.pt (epoch 36 @ 21168 updates, score 7.184) (writing took 3.632457632000296 seconds)
2021-03-27 22:01:44 | INFO | fairseq_cli.train | end of epoch 36 (average epoch stats below)
2021-03-27 22:01:44 | INFO | train | epoch 036 | loss 5.412 | nll_loss 4.411 | word_ins 5.655 | word_crf 2.235 | length 3.501 | ppl 42.59 | wps 19299.1 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 21168 | lr 0.000343661 | gnorm 1.471 | train_wall 204 | gb_free 9.2 | wall 7251
2021-03-27 22:01:44 | INFO | fairseq.trainer | begin training epoch 37
2021-03-27 22:01:44 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 22:01:55 | INFO | train_inner | epoch 037:   32 / 588 loss=5.558, nll_loss=4.573, word_ins=5.794, word_crf=2.313, length=3.481, ppl=47.12, wps=16630.2, ups=2.35, wpb=7061.7, bsz=256.6, num_updates=21200, lr=0.000343401, gnorm=1.571, train_wall=35, gb_free=9.4, wall=7263
2021-03-27 22:02:31 | INFO | train_inner | epoch 037:   132 / 588 loss=5.418, nll_loss=4.44, word_ins=5.684, word_crf=2.241, length=3.349, ppl=42.75, wps=19875.6, ups=2.82, wpb=7044.3, bsz=254.7, num_updates=21300, lr=0.000342594, gnorm=1.778, train_wall=35, gb_free=9.2, wall=7298
2021-03-27 22:03:06 | INFO | train_inner | epoch 037:   232 / 588 loss=5.311, nll_loss=4.318, word_ins=5.577, word_crf=2.183, length=3.404, ppl=39.71, wps=19738, ups=2.84, wpb=6938.9, bsz=279.6, num_updates=21400, lr=0.000341793, gnorm=1.403, train_wall=35, gb_free=10, wall=7334
2021-03-27 22:03:41 | INFO | train_inner | epoch 037:   332 / 588 loss=5.346, nll_loss=4.346, word_ins=5.598, word_crf=2.198, length=3.489, ppl=40.68, wps=20072.4, ups=2.87, wpb=6990.4, bsz=291.4, num_updates=21500, lr=0.000340997, gnorm=1.59, train_wall=35, gb_free=9.6, wall=7368
2021-03-27 22:04:16 | INFO | train_inner | epoch 037:   432 / 588 loss=5.521, nll_loss=4.52, word_ins=5.75, word_crf=2.295, length=3.51, ppl=45.93, wps=19855.9, ups=2.84, wpb=6983.6, bsz=263.8, num_updates=21600, lr=0.000340207, gnorm=1.556, train_wall=35, gb_free=10.2, wall=7404
2021-03-27 22:04:51 | INFO | train_inner | epoch 037:   532 / 588 loss=5.403, nll_loss=4.394, word_ins=5.641, word_crf=2.219, length=3.63, ppl=42.31, wps=20055.8, ups=2.87, wpb=6986.2, bsz=264.7, num_updates=21700, lr=0.000339422, gnorm=1.336, train_wall=35, gb_free=10.6, wall=7438
2021-03-27 22:05:10 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 22:05:13 | INFO | valid | epoch 037 | valid on 'valid' subset | loss 7.044 | nll_loss 5.375 | word_ins 6.576 | word_crf 3.322 | length 4.348 | ppl 132 | wps 59556.7 | wpb 5311.6 | bsz 208.1 | num_updates 21756 | best_loss 6.867
2021-03-27 22:05:13 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 37 @ 21756 updates
2021-03-27 22:05:13 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint37.pt
2021-03-27 22:05:14 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint37.pt
2021-03-27 22:05:17 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint37.pt (epoch 37 @ 21756 updates, score 7.044) (writing took 3.637799907000044 seconds)
2021-03-27 22:05:17 | INFO | fairseq_cli.train | end of epoch 37 (average epoch stats below)
2021-03-27 22:05:17 | INFO | train | epoch 037 | loss 5.387 | nll_loss 4.392 | word_ins 5.639 | word_crf 2.22 | length 3.468 | ppl 41.84 | wps 19267.3 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 21756 | lr 0.000338985 | gnorm 1.531 | train_wall 204 | gb_free 10.1 | wall 7465
2021-03-27 22:05:17 | INFO | fairseq.trainer | begin training epoch 38
2021-03-27 22:05:17 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 22:05:33 | INFO | train_inner | epoch 038:   44 / 588 loss=5.263, nll_loss=4.272, word_ins=5.534, word_crf=2.161, length=3.356, ppl=38.4, wps=16563.2, ups=2.36, wpb=7013.2, bsz=303, num_updates=21800, lr=0.000338643, gnorm=1.593, train_wall=35, gb_free=9.3, wall=7481
2021-03-27 22:06:08 | INFO | train_inner | epoch 038:   144 / 588 loss=5.275, nll_loss=4.295, word_ins=5.562, word_crf=2.15, length=3.444, ppl=38.73, wps=20453.5, ups=2.84, wpb=7192.4, bsz=262.1, num_updates=21900, lr=0.000337869, gnorm=1.398, train_wall=35, gb_free=9.2, wall=7516
2021-03-27 22:06:43 | INFO | train_inner | epoch 038:   244 / 588 loss=5.319, nll_loss=4.324, word_ins=5.584, word_crf=2.176, length=3.506, ppl=39.91, wps=20165.2, ups=2.86, wpb=7045.1, bsz=265.3, num_updates=22000, lr=0.0003371, gnorm=1.422, train_wall=35, gb_free=9.6, wall=7551
2021-03-27 22:07:18 | INFO | train_inner | epoch 038:   344 / 588 loss=5.231, nll_loss=4.236, word_ins=5.509, word_crf=2.129, length=3.469, ppl=37.55, wps=19641.1, ups=2.89, wpb=6800.9, bsz=274.8, num_updates=22100, lr=0.000336336, gnorm=1.441, train_wall=34, gb_free=9.7, wall=7586
2021-03-27 22:07:52 | INFO | train_inner | epoch 038:   444 / 588 loss=5.348, nll_loss=4.348, word_ins=5.603, word_crf=2.202, length=3.448, ppl=40.74, wps=20330.2, ups=2.87, wpb=7073.2, bsz=279.4, num_updates=22200, lr=0.000335578, gnorm=1.517, train_wall=35, gb_free=9.5, wall=7620
2021-03-27 22:08:28 | INFO | train_inner | epoch 038:   544 / 588 loss=5.512, nll_loss=4.529, word_ins=5.761, word_crf=2.282, length=3.494, ppl=45.62, wps=19326.4, ups=2.82, wpb=6850, bsz=241.3, num_updates=22300, lr=0.000334825, gnorm=1.373, train_wall=35, gb_free=9.1, wall=7656
2021-03-27 22:08:43 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 22:08:47 | INFO | valid | epoch 038 | valid on 'valid' subset | loss 7.169 | nll_loss 5.414 | word_ins 6.62 | word_crf 3.413 | length 4.461 | ppl 143.9 | wps 59865.7 | wpb 5311.6 | bsz 208.1 | num_updates 22344 | best_loss 6.867
2021-03-27 22:08:47 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 38 @ 22344 updates
2021-03-27 22:08:47 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint38.pt
2021-03-27 22:08:47 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint38.pt
2021-03-27 22:08:50 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint38.pt (epoch 38 @ 22344 updates, score 7.169) (writing took 3.6485973260005267 seconds)
2021-03-27 22:08:50 | INFO | fairseq_cli.train | end of epoch 38 (average epoch stats below)
2021-03-27 22:08:50 | INFO | train | epoch 038 | loss 5.318 | nll_loss 4.328 | word_ins 5.587 | word_crf 2.18 | length 3.445 | ppl 39.89 | wps 19267.2 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 22344 | lr 0.000334495 | gnorm 1.461 | train_wall 204 | gb_free 10.1 | wall 7678
2021-03-27 22:08:50 | INFO | fairseq.trainer | begin training epoch 39
2021-03-27 22:08:50 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 22:09:10 | INFO | train_inner | epoch 039:   56 / 588 loss=5.167, nll_loss=4.191, word_ins=5.466, word_crf=2.097, length=3.373, ppl=35.92, wps=16414.2, ups=2.38, wpb=6907.3, bsz=303, num_updates=22400, lr=0.000334077, gnorm=1.568, train_wall=35, gb_free=9.3, wall=7698
2021-03-27 22:09:45 | INFO | train_inner | epoch 039:   156 / 588 loss=5.191, nll_loss=4.211, word_ins=5.489, word_crf=2.101, length=3.449, ppl=36.53, wps=19959.8, ups=2.87, wpb=6950.4, bsz=264.8, num_updates=22500, lr=0.000333333, gnorm=1.607, train_wall=35, gb_free=9.5, wall=7733

2021-03-27 22:10:20 | INFO | train_inner | epoch 039:   256 / 588 loss=5.093, nll_loss=4.107, word_ins=5.399, word_crf=2.06, length=3.339, ppl=34.13, wps=19874.9, ups=2.85, wpb=6980.1, bsz=294.1, num_updates=22600, lr=0.000332595, gnorm=1.479, train_wall=35, gb_free=9.7, wall=7768
2021-03-27 22:10:55 | INFO | train_inner | epoch 039:   356 / 588 loss=5.191, nll_loss=4.189, word_ins=5.469, word_crf=2.104, length=3.528, ppl=36.54, wps=20257, ups=2.87, wpb=7052.7, bsz=279.4, num_updates=22700, lr=0.000331862, gnorm=1.361, train_wall=35, gb_free=9.4, wall=7803
2021-03-27 22:11:30 | INFO | train_inner | epoch 039:   456 / 588 loss=5.374, nll_loss=4.396, word_ins=5.647, word_crf=2.227, length=3.239, ppl=41.47, wps=19553, ups=2.8, wpb=6984.7, bsz=273.1, num_updates=22800, lr=0.000331133, gnorm=1.437, train_wall=35, gb_free=9.5, wall=7838
2021-03-27 22:12:06 | INFO | train_inner | epoch 039:   556 / 588 loss=5.554, nll_loss=4.564, word_ins=5.792, word_crf=2.304, length=3.537, ppl=46.98, wps=19608.5, ups=2.81, wpb=6985, bsz=241, num_updates=22900, lr=0.000330409, gnorm=1.637, train_wall=35, gb_free=9.7, wall=7874
2021-03-27 22:12:17 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 22:12:21 | INFO | valid | epoch 039 | valid on 'valid' subset | loss 7.157 | nll_loss 5.384 | word_ins 6.595 | word_crf 3.407 | length 4.536 | ppl 142.76 | wps 59819.3 | wpb 5311.6 | bsz 208.1 | num_updates 22932 | best_loss 6.867
2021-03-27 22:12:21 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 39 @ 22932 updates
2021-03-27 22:12:21 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint39.pt
2021-03-27 22:12:21 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint39.pt
2021-03-27 22:12:24 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint39.pt (epoch 39 @ 22932 updates, score 7.157) (writing took 3.6373173569991195 seconds)
2021-03-27 22:12:24 | INFO | fairseq_cli.train | end of epoch 39 (average epoch stats below)
2021-03-27 22:12:24 | INFO | train | epoch 039 | loss 5.279 | nll_loss 4.295 | word_ins 5.56 | word_crf 2.158 | length 3.411 | ppl 38.82 | wps 19195.7 | ups 2.75 | wpb 6988.7 | bsz 272.5 | num_updates 22932 | lr 0.000330179 | gnorm 1.518 | train_wall 205 | gb_free 9.4 | wall 7892
2021-03-27 22:12:24 | INFO | fairseq.trainer | begin training epoch 40
2021-03-27 22:12:24 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 22:12:49 | INFO | train_inner | epoch 040:   68 / 588 loss=5.406, nll_loss=4.452, word_ins=5.697, word_crf=2.224, length=3.345, ppl=42.41, wps=16299.1, ups=2.34, wpb=6971.6, bsz=248.2, num_updates=23000, lr=0.00032969, gnorm=1.452, train_wall=35, gb_free=9.2, wall=7917
2021-03-27 22:13:24 | INFO | train_inner | epoch 040:   168 / 588 loss=5.128, nll_loss=4.155, word_ins=5.442, word_crf=2.074, length=3.325, ppl=34.96, wps=20158.2, ups=2.87, wpb=7034.1, bsz=282.6, num_updates=23100, lr=0.000328976, gnorm=1.558, train_wall=35, gb_free=9.2, wall=7952
2021-03-27 22:13:59 | INFO | train_inner | epoch 040:   268 / 588 loss=5.263, nll_loss=4.292, word_ins=5.56, word_crf=2.153, length=3.3, ppl=38.4, wps=19517.6, ups=2.84, wpb=6880.3, bsz=263.8, num_updates=23200, lr=0.000328266, gnorm=1.49, train_wall=35, gb_free=9.6, wall=7987
2021-03-27 22:14:34 | INFO | train_inner | epoch 040:   368 / 588 loss=5.288, nll_loss=4.321, word_ins=5.585, word_crf=2.166, length=3.298, ppl=39.08, wps=20123.9, ups=2.86, wpb=7036.6, bsz=264.1, num_updates=23300, lr=0.000327561, gnorm=1.504, train_wall=35, gb_free=9.3, wall=8022
2021-03-27 22:15:09 | INFO | train_inner | epoch 040:   468 / 588 loss=5.113, nll_loss=4.12, word_ins=5.41, word_crf=2.068, length=3.401, ppl=34.62, wps=20052.2, ups=2.87, wpb=6995.4, bsz=286.7, num_updates=23400, lr=0.00032686, gnorm=1.359, train_wall=35, gb_free=9.4, wall=8057
2021-03-27 22:15:44 | INFO | train_inner | epoch 040:   568 / 588 loss=5.187, nll_loss=4.191, word_ins=5.47, word_crf=2.101, length=3.513, ppl=36.44, wps=20177.5, ups=2.87, wpb=7019.2, bsz=287.8, num_updates=23500, lr=0.000326164, gnorm=1.274, train_wall=35, gb_free=9.8, wall=8092
2021-03-27 22:15:50 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 22:15:54 | INFO | valid | epoch 040 | valid on 'valid' subset | loss 7.227 | nll_loss 5.408 | word_ins 6.631 | word_crf 3.461 | length 4.508 | ppl 149.85 | wps 60077 | wpb 5311.6 | bsz 208.1 | num_updates 23520 | best_loss 6.867
2021-03-27 22:15:54 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 40 @ 23520 updates
2021-03-27 22:15:54 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint40.pt
2021-03-27 22:15:55 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint40.pt
2021-03-27 22:15:57 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint40.pt (epoch 40 @ 23520 updates, score 7.227) (writing took 3.6334887749999325 seconds)
2021-03-27 22:15:57 | INFO | fairseq_cli.train | end of epoch 40 (average epoch stats below)
2021-03-27 22:15:57 | INFO | train | epoch 040 | loss 5.213 | nll_loss 4.236 | word_ins 5.511 | word_crf 2.121 | length 3.373 | ppl 37.1 | wps 19284 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 23520 | lr 0.000326025 | gnorm 1.425 | train_wall 204 | gb_free 9.9 | wall 8105
2021-03-27 22:15:57 | INFO | fairseq.trainer | begin training epoch 41
2021-03-27 22:15:57 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 22:16:26 | INFO | train_inner | epoch 041:   80 / 588 loss=5.087, nll_loss=4.119, word_ins=5.412, word_crf=2.045, length=3.359, ppl=33.98, wps=16676.1, ups=2.37, wpb=7038.6, bsz=280.6, num_updates=23600, lr=0.000325472, gnorm=1.389, train_wall=35, gb_free=9.9, wall=8134
2021-03-27 22:17:01 | INFO | train_inner | epoch 041:   180 / 588 loss=5.082, nll_loss=4.107, word_ins=5.402, word_crf=2.047, length=3.347, ppl=33.88, wps=20058.1, ups=2.88, wpb=6964.4, bsz=283.4, num_updates=23700, lr=0.000324785, gnorm=1.526, train_wall=34, gb_free=9.4, wall=8168
2021-03-27 22:17:36 | INFO | train_inner | epoch 041:   280 / 588 loss=5.1, nll_loss=4.124, word_ins=5.417, word_crf=2.067, length=3.246, ppl=34.29, wps=19678, ups=2.86, wpb=6883.9, bsz=285.3, num_updates=23800, lr=0.000324102, gnorm=1.457, train_wall=35, gb_free=9.9, wall=8203
2021-03-27 22:18:11 | INFO | train_inner | epoch 041:   380 / 588 loss=5.272, nll_loss=4.293, word_ins=5.564, word_crf=2.151, length=3.388, ppl=38.63, wps=19919.2, ups=2.84, wpb=7005.4, bsz=250.5, num_updates=23900, lr=0.000323423, gnorm=1.488, train_wall=35, gb_free=9.7, wall=8239
2021-03-27 22:18:46 | INFO | train_inner | epoch 041:   480 / 588 loss=5.314, nll_loss=4.331, word_ins=5.595, word_crf=2.17, length=3.462, ppl=39.78, wps=20118.3, ups=2.84, wpb=7092.5, bsz=258.4, num_updates=24000, lr=0.000322749, gnorm=1.327, train_wall=35, gb_free=9.2, wall=8274
2021-03-27 22:19:21 | INFO | train_inner | epoch 041:   580 / 588 loss=5.271, nll_loss=4.294, word_ins=5.561, word_crf=2.156, length=3.349, ppl=38.62, wps=19837.4, ups=2.86, wpb=6948.3, bsz=270.6, num_updates=24100, lr=0.000322078, gnorm=1.558, train_wall=35, gb_free=9.8, wall=8309
2021-03-27 22:19:24 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 22:19:27 | INFO | valid | epoch 041 | valid on 'valid' subset | loss 7.113 | nll_loss 5.376 | word_ins 6.587 | word_crf 3.365 | length 4.554 | ppl 138.47 | wps 59472.6 | wpb 5311.6 | bsz 208.1 | num_updates 24108 | best_loss 6.867
2021-03-27 22:19:27 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 41 @ 24108 updates
2021-03-27 22:19:27 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint41.pt
2021-03-27 22:19:28 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint41.pt
2021-03-27 22:19:31 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint41.pt (epoch 41 @ 24108 updates, score 7.113) (writing took 3.6916772829990805 seconds)
2021-03-27 22:19:31 | INFO | fairseq_cli.train | end of epoch 41 (average epoch stats below)
2021-03-27 22:19:31 | INFO | train | epoch 041 | loss 5.183 | nll_loss 4.208 | word_ins 5.489 | word_crf 2.104 | length 3.345 | ppl 36.33 | wps 19260.2 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 24108 | lr 0.000322025 | gnorm 1.467 | train_wall 204 | gb_free 9.6 | wall 8319
2021-03-27 22:19:31 | INFO | fairseq.trainer | begin training epoch 42
2021-03-27 22:19:31 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 22:20:03 | INFO | train_inner | epoch 042:   92 / 588 loss=5.116, nll_loss=4.155, word_ins=5.444, word_crf=2.054, length=3.396, ppl=34.67, wps=16593.2, ups=2.36, wpb=7016.8, bsz=271.2, num_updates=24200, lr=0.000321412, gnorm=1.625, train_wall=35, gb_free=10.4, wall=8351
gggggggggggiuuuuuuuuuuuuuuuuuuuuuuuuuuuuufv2021-03-27 22:20:38 | INFO | train_inner | epoch 042:   192 / 588 loss=5.066, nll_loss=4.101, word_ins=5.397, word_crf=2.038, length=3.293, ppl=33.5, wps=20247.4, ups=2.84, wpb=7119.4, bsz=283.6, num_updates=24300, lr=0.00032075, gnorm=1.541, train_wall=35, gb_free=10.4, wall=8386
2021-03-27 22:21:13 | INFO | train_inner | epoch 042:   292 / 588 loss=5.049, nll_loss=4.07, word_ins=5.37, word_crf=2.032, length=3.314, ppl=33.11, wps=20410.3, ups=2.87, wpb=7115.6, bsz=300.1, num_updates=24400, lr=0.000320092, gnorm=1.387, train_wall=35, gb_free=10.1, wall=8421
2021-03-27 22:21:48 | INFO | train_inner | epoch 042:   392 / 588 loss=5.137, nll_loss=4.155, word_ins=5.444, word_crf=2.088, length=3.269, ppl=35.19, wps=20109, ups=2.87, wpb=7014.7, bsz=278.6, num_updates=24500, lr=0.000319438, gnorm=1.655, train_wall=35, gb_free=9.4, wall=8456
2021-03-27 22:22:24 | INFO | train_inner | epoch 042:   492 / 588 loss=5.249, nll_loss=4.28, word_ins=5.552, word_crf=2.143, length=3.302, ppl=38.02, wps=19299.8, ups=2.82, wpb=6833.3, bsz=253.8, num_updates=24600, lr=0.000318788, gnorm=1.536, train_wall=35, gb_free=9.3, wall=8491
2021-03-27 22:22:57 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 22:23:01 | INFO | valid | epoch 042 | valid on 'valid' subset | loss 7.223 | nll_loss 5.424 | word_ins 6.638 | word_crf 3.447 | length 4.569 | ppl 149.43 | wps 59561.2 | wpb 5311.6 | bsz 208.1 | num_updates 24696 | best_loss 6.867
2021-03-27 22:23:01 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 42 @ 24696 updates
2021-03-27 22:23:01 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint42.pt
2021-03-27 22:23:01 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint42.pt
2021-03-27 22:23:04 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint42.pt (epoch 42 @ 24696 updates, score 7.223) (writing took 3.654281445000379 seconds)
2021-03-27 22:23:04 | INFO | fairseq_cli.train | end of epoch 42 (average epoch stats below)
2021-03-27 22:23:04 | INFO | train | epoch 042 | loss 5.157 | nll_loss 4.186 | word_ins 5.471 | word_crf 2.088 | length 3.333 | ppl 35.67 | wps 19242.5 | ups 2.75 | wpb 6988.7 | bsz 272.5 | num_updates 24696 | lr 0.000318168 | gnorm 1.547 | train_wall 205 | gb_free 9.7 | wall 8532

2021-03-27 22:23:04 | INFO | fairseq.trainer | begin training epoch 43
2021-03-27 22:23:04 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 22:23:06 | INFO | train_inner | epoch 043:   4 / 588 loss=5.273, nll_loss=4.301, word_ins=5.57, word_crf=2.146, length=3.413, ppl=38.67, wps=16180.4, ups=2.37, wpb=6836.7, bsz=254, num_updates=24700, lr=0.000318142, gnorm=1.529, train_wall=35, gb_free=10.2, wall=8534
2021-03-27 22:23:41 | INFO | train_inner | epoch 043: 104 / 588 loss=5.11, nll_loss=4.169, word_ins=5.461, word_crf=2.062, length=3.182, ppl=34.54, wps=19948.6, ups=2.81, wpb=7096.4, bsz=257.5, num_updates=24800, lr=0.0003175, gnorm=1.514, train_wall=35, gb_free=9.6, wall=8569
2021-03-27 22:24:16 | INFO | train_inner | epoch 043: 204 / 588 loss=5.109, nll_loss=4.145, word_ins=5.44, word_crf=2.059, length=3.303, ppl=34.52, wps=19900.6, ups=2.86, wpb=6962.9, bsz=265, num_updates=24900, lr=0.000316862, gnorm=1.783, train_wall=35, gb_free=9.7, wall=8604
2021-03-27 22:24:51 | INFO | train_inner | epoch 043: 304 / 588 loss=5.043, nll_loss=4.071, word_ins=5.372, word_crf=2.024, length=3.323, ppl=32.96, wps=19877.8, ups=2.86, wpb=6960.9, bsz=287.4, num_updates=25000, lr=0.000316228, gnorm=1.507, train_wall=35, gb_free=9.2, wall=8639
2021-03-27 22:25:26 | INFO | train_inner | epoch 043: 404 / 588 loss=5.044, nll_loss=4.067, word_ins=5.37, word_crf=2.024, length=3.346, ppl=32.99, wps=19991.6, ups=2.86, wpb=6988.9, bsz=278.6, num_updates=25100, lr=0.000315597, gnorm=1.411, train_wall=35, gb_free=9.1, wall=8674
2021-03-27 22:26:01 | INFO | train_inner | epoch 043: 504 / 588 loss=5.078, nll_loss=4.114, word_ins=5.41, word_crf=2.046, length=3.27, ppl=33.77, wps=19840.3, ups=2.87, wpb=6916.9, bsz=279.4, num_updates=25200, lr=0.00031497, gnorm=1.457, train_wall=35, gb_free=10, wall=8709
2021-03-27 22:26:31 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 22:26:34 | INFO | valid | epoch 043 | valid on 'valid' subset | loss 7.283 | nll_loss 5.417 | word_ins 6.649 | word_crf 3.494 | length 4.641 | ppl 155.72 | wps 60001.9 | wpb 5311.6 | bsz 208.1 | num_updates 25284 | best_loss 6.867
2021-03-27 22:26:34 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 43 @ 25284 updates
2021-03-27 22:26:34 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint43.pt
2021-03-27 22:26:35 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint43.pt
2021-03-27 22:26:38 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint43.pt (epoch 43 @ 25284 updates, score 7.283) (writing took 3.7112448360003327 seconds)
2021-03-27 22:26:38 | INFO | fairseq_cli.train | end of epoch 43 (average epoch stats below)
2021-03-27 22:26:38 | INFO | train | epoch 043 | loss 5.095 | nll_loss 4.132 | word_ins 5.426 | word_crf 2.053 | length 3.291 | ppl 34.18 | wps 19254.2 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 25284 | lr 0.000314447 | gnorm 1.513 | train_wall 205 | gb_free 9.6 | wall 8746
2021-03-27 22:26:38 | INFO | fairseq.trainer | begin training epoch 44
2021-03-27 22:26:38 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 22:26:44 | INFO | train_inner | epoch 044:   16 / 588 loss=5.224, nll_loss=4.272, word_ins=5.546, word_crf=2.122, length=3.29, ppl=37.37, wps=16475.4, ups=2.35, wpb=6996.1, bsz=263.5, num_updates=25300, lr=0.000314347, gnorm=1.384, train_wall=35, gb_free=9.3, wall=8752
2021-03-27 22:27:19 | INFO | train_inner | epoch 044: 116 / 588 loss=4.998, nll_loss=4.05, word_ins=5.36, word_crf=2, length=3.186, ppl=31.96, wps=20453.6, ups=2.86, wpb=7156.3, bsz=269.8, num_updates=25400, lr=0.000313728, gnorm=1.496, train_wall=35, gb_free=10.1, wall=8787
2021-03-27 22:27:54 | INFO | train_inner | epoch 044: 216 / 588 loss=5.071, nll_loss=4.105, word_ins=5.408, word_crf=2.028, length=3.388, ppl=33.61, wps=20034.1, ups=2.86, wpb=7013.8, bsz=257.6, num_updates=25500, lr=0.000313112, gnorm=1.48, train_wall=35, gb_free=9.4, wall=8822
2021-03-27 22:28:29 | INFO | train_inner | epoch 044: 316 / 588 loss=5.036, nll_loss=4.084, word_ins=5.387, word_crf=2.021, length=3.216, ppl=32.81, wps=19629, ups=2.86, wpb=6870.1, bsz=277, num_updates=25600, lr=0.0003125, gnorm=1.547, train_wall=35, gb_free=9.1, wall=8857
2021-03-27 22:29:04 | INFO | train_inner | epoch 044: 416 / 588 loss=5.136, nll_loss=4.167, word_ins=5.458, word_crf=2.07, length=3.369, ppl=35.16, wps=20058.6, ups=2.87, wpb=6981.5, bsz=259.9, num_updates=25700, lr=0.000311891, gnorm=1.547, train_wall=35, gb_free=9.6, wall=8891
2021-03-27 22:29:38 | INFO | train_inner | epoch 044: 516 / 588 loss=4.947, nll_loss=3.972, word_ins=5.289, word_crf=1.984, length=3.189, ppl=30.85, wps=19882.7, ups=2.86, wpb=6943.5, bsz=300.2, num_updates=25800, lr=0.000311286, gnorm=1.549, train_wall=35, gb_free=9.3, wall=8926
2021-03-27 22:30:04 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 22:30:07 | INFO | valid | epoch 044 | valid on 'valid' subset | loss 7.306 | nll_loss 5.43 | word_ins 6.659 | word_crf 3.505 | length 4.711 | ppl 158.23 | wps 60105.4 | wpb 5311.6 | bsz 208.1 | num_updates 25872 | best_loss 6.867
2021-03-27 22:30:07 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 44 @ 25872 updates
2021-03-27 22:30:07 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint44.pt
2021-03-27 22:30:08 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint44.pt
2021-03-27 22:30:11 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint44.pt (epoch 44 @ 25872 updates, score 7.306) (writing took 3.6500188170011825 seconds)
2021-03-27 22:30:11 | INFO | fairseq_cli.train | end of epoch 44 (average epoch stats below)
2021-03-27 22:30:11 | INFO | train | epoch 044 | loss 5.053 | nll_loss 4.092 | word_ins 5.393 | word_crf 2.029 | length 3.267 | ppl 33.19 | wps 19293.5 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 25872 | lr 0.000310853 | gnorm 1.506 | train_wall 204 | gb_free 9.3 | wall 8959
2021-03-27 22:30:11 | INFO | fairseq.trainer | begin training epoch 45
2021-03-27 22:30:11 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 22:30:21 | INFO | train_inner | epoch 045:   28 / 588 loss=5.016, nll_loss=4.046, word_ins=5.352, word_crf=2.008, length=3.313, ppl=32.35, wps=16530.5, ups=2.37, wpb=6968.2, bsz=282.3, num_updates=25900, lr=0.000310685, gnorm=1.402, train_wall=35, gb_free=9.2, wall=8969
2021-03-27 22:30:55 | INFO | train_inner | epoch 045: 128 / 588 loss=4.945, nll_loss=4, word_ins=5.319, word_crf=1.955, length=3.304, ppl=30.8, wps=20199.1, ups=2.87, wpb=7047.4, bsz=265.4, num_updates=26000, lr=0.000310087, gnorm=1.45, train_wall=35, gb_free=9.3, wall=9003
2021-03-27 22:31:30 | INFO | train_inner | epoch 045: 228 / 588 loss=5.033, nll_loss=4.064, word_ins=5.372, word_crf=2.019, length=3.284, ppl=32.74, wps=20244.9, ups=2.86, wpb=7078.2, bsz=275.2, num_updates=26100, lr=0.000309492, gnorm=1.671, train_wall=35, gb_free=9.5, wall=9038
2021-03-27 22:32:05 | INFO | train_inner | epoch 045: 328 / 588 loss=4.955, nll_loss=4.001, word_ins=5.317, word_crf=1.974, length=3.222, ppl=31.01, wps=19931.1, ups=2.88, wpb=6922.5, bsz=281.6, num_updates=26200, lr=0.000308901, gnorm=1.432, train_wall=34, gb_free=9.3, wall=9073
2021-03-27 22:32:41 | INFO | train_inner | epoch 045: 428 / 588 loss=5.13, nll_loss=4.184, word_ins=5.476, word_crf=2.075, length=3.169, ppl=35.01, wps=19139.3, ups=2.81, wpb=6813.2, bsz=248.4, num_updates=26300, lr=0.000308313, gnorm=1.534, train_wall=35, gb_free=9.6, wall=9109
2021-03-27 22:33:16 | INFO | train_inner | epoch 045: 528 / 588 loss=5.054, nll_loss=4.09, word_ins=5.392, word_crf=2.036, length=3.223, ppl=33.22, wps=20043.8, ups=2.85, wpb=7021.4, bsz=281.9, num_updates=26400, lr=0.000307729, gnorm=1.531, train_wall=35, gb_free=9.2, wall=9144
2021-03-27 22:33:37 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 22:33:40 | INFO | valid | epoch 045 | valid on 'valid' subset | loss 7.336 | nll_loss 5.453 | word_ins 6.675 | word_crf 3.536 | length 4.619 | ppl 161.55 | wps 59663.3 | wpb 5311.6 | bsz 208.1 | num_updates 26460 | best_loss 6.867
2021-03-27 22:33:40 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 45 @ 26460 updates
2021-03-27 22:33:40 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint45.pt
2021-03-27 22:33:41 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint45.pt
2021-03-27 22:33:43 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint45.pt (epoch 45 @ 26460 updates, score 7.336) (writing took 3.638574361000792 seconds)
2021-03-27 22:33:43 | INFO | fairseq_cli.train | end of epoch 45 (average epoch stats below)
2021-03-27 22:33:43 | INFO | train | epoch 045 | loss 5.008 | nll_loss 4.05 | word_ins 5.359 | word_crf 2.003 | length 3.262 | ppl 32.19 | wps 19312.2 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 26460 | lr 0.00030738 | gnorm 1.495 | train_wall 204 | gb_free 9.7 | wall 9171
2021-03-27 22:33:44 | INFO | fairseq.trainer | begin training epoch 46
2021-03-27 22:33:44 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 22:33:58 | INFO | train_inner | epoch 046:   40 / 588 loss=4.932, nll_loss=3.976, word_ins=5.295, word_crf=1.965, length=3.185, ppl=30.52, wps=16818.4, ups=2.38, wpb=7073, bsz=284.6, num_updates=26500, lr=0.000307148, gnorm=1.397, train_wall=35, gb_free=9.5, wall=9186
2021-03-27 22:34:33 | INFO | train_inner | epoch 046: 140 / 588 loss=4.97, nll_loss=4.017, word_ins=5.293, word_crf=1.942, length=3.182, ppl=30, wps=19837.9, ups=2.87, wpb=6911.6, bsz=269.4, num_updates=26600, lr=0.00030657, gnorm=1.588, train_wall=35, gb_free=9.4, wall=9221
2021-03-27 22:35:08 | INFO | train_inner | epoch 046: 240 / 588 loss=5.068, nll_loss=4.123, word_ins=5.425, word_crf=2.033, length=3.222, ppl=33.53, wps=19720.6, ups=2.84, wpb=6941.8, bsz=255, num_updates=26700, lr=0.000305995, gnorm=1.533, train_wall=35, gb_free=9.4, wall=9256
2021-03-27 22:35:43 | INFO | train_inner | epoch 046: 340 / 588 loss=4.951, nll_loss=3.993, word_ins=5.312, word_crf=1.972, length=3.231, ppl=30.93, wps=20169, ups=2.88, wpb=6991.8, bsz=275.6, num_updates=26800, lr=0.000305424, gnorm=1.456, train_wall=34, gb_free=9.1, wall=9290
2021-03-27 22:36:18 | INFO | train_inner | epoch 046: 440 / 588 loss=4.951, nll_loss=3.997, word_ins=5.314, word_crf=1.986, length=3.082, ppl=30.94, wps=19834.9, ups=2.82, wpb=7024.9, bsz=287.2, num_updates=26900, lr=0.000304855, gnorm=1.502, train_wall=35, gb_free=9.4, wall=9326
2021-03-27 22:36:53 | INFO | train_inner | epoch 046: 540 / 588 loss=5.12, nll_loss=4.143, word_ins=5.441, word_crf=2.045, length=3.539, ppl=34.77, wps=20144.1, ups=2.88, wpb=7000.8, bsz=242.2, num_updates=27000, lr=0.00030429, gnorm=1.481, train_wall=34, gb_free=9.4, wall=9361
2021-03-27 22:37:09 | INFO | fairseq_cli.train | begin validation on "valid" subset

2021-03-27 22:37:13 | INFO | valid | epoch 046 | valid on 'valid' subset | loss 7.501 | nll_loss 5.515 | word_ins 6.756 | word_crf 3.661 | length 4.625 | ppl 181.19 | wps 59838.1 | wpb 5311.6 | bsz 208.1 | num_updates 27048 | best_loss 6.867
2021-03-27 22:37:13 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 46 @ 27048 updates
2021-03-27 22:37:13 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint46.pt
2021-03-27 22:37:14 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint46.pt
2021-03-27 22:37:16 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint46.pt (epoch 46 @ 27048 updates, score 7.501) (writing took 3.6407320849993994 seconds)
2021-03-27 22:37:16 | INFO | fairseq_cli.train | end of epoch 46 (average epoch stats below)
2021-03-27 22:37:16 | INFO | train | epoch 046 | loss 4.97 | nll_loss 4.017 | word_ins 5.332 | word_crf 1.982 | length 3.223 | ppl 31.35 | wps 19307.7 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 27048 | lr 0.00030402 | gnorm 1.496 | train_wall 204 | gb_free 9.7 | wall 9384
2021-03-27 22:37:16 | INFO | fairseq.trainer | begin training epoch 47
2021-03-27 22:37:16 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 22:37:35 | INFO | train_inner | epoch 047:   52 / 588 loss=4.907, nll_loss=3.957, word_ins=5.279, word_crf=1.952, length=3.151, ppl=30, wps=16684.7, ups=2.38, wpb=7006.8, bsz=287.5, num_updates=27100, lr=0.000303728, gnorm=1.585, train_wall=35, gb_free=9.6, wall=9403
2021-03-27 22:38:10 | INFO | train_inner | epoch 047:   152 / 588 loss=4.919, nll_loss=3.97, word_ins=5.293, word_crf=1.954, length=3.182, ppl=30.25, wps=19959.6, ups=2.85, wpb=6995.9, bsz=281, num_updates=27200, lr=0.00030317, gnorm=1.706, train_wall=35, gb_free=10, wall=9438
2021-03-27 22:38:45 | INFO | train_inner | epoch 047:   252 / 588 loss=4.878, nll_loss=3.944, word_ins=5.274, word_crf=1.925, length=3.164, ppl=29.41, wps=19667.2, ups=2.86, wpb=6886.1, bsz=260.1, num_updates=27300, lr=0.000302614, gnorm=1.471, train_wall=35, gb_free=9.8, wall=9473
2021-03-27 22:39:20 | INFO | train_inner | epoch 047:   352 / 588 loss=4.844, nll_loss=3.893, word_ins=5.227, word_crf=1.902, length=3.285, ppl=28.72, wps=19699.6, ups=2.87, wpb=6858.7, bsz=274, num_updates=27400, lr=0.000302061, gnorm=1.396, train_wall=35, gb_free=9.4, wall=9508
2021-03-27 22:39:55 | INFO | train_inner | epoch 047:   452 / 588 loss=4.986, nll_loss=4.033, word_ins=5.347, word_crf=2.002, length=3.094, ppl=31.68, wps=20042.6, ups=2.84, wpb=7057.1, bsz=276.1, num_updates=27500, lr=0.000301511, gnorm=1.444, train_wall=35, gb_free=9.3, wall=9543
2021-03-27 22:40:30 | INFO | train_inner | epoch 047:   552 / 588 loss=4.874, nll_loss=3.904, word_ins=5.233, word_crf=1.931, length=3.259, ppl=29.32, wps=20517.6, ups=2.87, wpb=7143.4, bsz=294.4, num_updates=27600, lr=0.000300965, gnorm=1.354, train_wall=35, gb_free=9.7, wall=9578
2021-03-27 22:40:42 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 22:40:46 | INFO | valid | epoch 047 | valid on 'valid' subset | loss 7.402 | nll_loss 5.465 | word_ins 6.698 | word_crf 3.579 | length 4.736 | ppl 169.15 | wps 59894.5 | wpb 5311.6 | bsz 208.1 | num_updates 27636 | best_loss 6.867
2021-03-27 22:40:46 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 47 @ 27636 updates
2021-03-27 22:40:46 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint47.pt
2021-03-27 22:40:47 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint47.pt
2021-03-27 22:40:49 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint47.pt (epoch 47 @ 27636 updates, score 7.402) (writing took 3.6364891749999515 seconds)
2021-03-27 22:40:49 | INFO | fairseq_cli.train | end of epoch 47 (average epoch stats below)
2021-03-27 22:40:49 | INFO | train | epoch 047 | loss 4.934 | nll_loss 3.985 | word_ins 5.306 | word_crf 1.962 | length 3.191 | ppl 30.57 | wps 19282.1 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 27636 | lr 0.000300769 | gnorm 1.504 | train_wall 204 | gb_free 9.4 | wall 9597
2021-03-27 22:40:50 | INFO | fairseq.trainer | begin training epoch 48
2021-03-27 22:40:50 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 22:41:12 | INFO | train_inner | epoch 048:   64 / 588 loss=4.971, nll_loss=4.039, word_ins=5.354, word_crf=1.974, length=3.195, ppl=31.36, wps=16338.2, ups=2.35, wpb=6955.2, bsz=257, num_updates=27700, lr=0.000300421, gnorm=1.467, train_wall=35, gb_free=10.6, wall=9620
2021-03-27 22:41:48 | INFO | train_inner | epoch 048:   164 / 588 loss=4.941, nll_loss=4.011, word_ins=5.333, word_crf=1.958, length=3.168, ppl=30.72, wps=19692.6, ups=2.83, wpb=6958.2, bsz=255.3, num_updates=27800, lr=0.00029988, gnorm=1.596, train_wall=35, gb_free=10.5, wall=9655
2021-03-27 22:42:23 | INFO | train_inner | epoch 048:   264 / 588 loss=5.167, nll_loss=4.237, word_ins=5.526, word_crf=2.092, length=3.114, ppl=35.92, wps=19496.4, ups=2.81, wpb=6941.3, bsz=233.2, num_updates=27900, lr=0.000299342, gnorm=1.693, train_wall=35, gb_free=9.9, wall=9691
2021-03-27 22:42:58 | INFO | train_inner | epoch 048:   364 / 588 loss=4.861, nll_loss=3.907, word_ins=5.24, word_crf=1.926, length=3.155, ppl=29.06, wps=20146, ups=2.85, wpb=7070.1, bsz=283.8, num_updates=28000, lr=0.000298807, gnorm=1.558, train_wall=35, gb_free=9.2, wall=9726
2021-03-27 22:43:33 | INFO | train_inner | epoch 048:   464 / 588 loss=4.753, nll_loss=3.787, word_ins=5.136, word_crf=1.86, length=3.247, ppl=26.96, wps=19925.5, ups=2.89, wpb=6904.2, bsz=293.3, num_updates=28100, lr=0.000298275, gnorm=1.483, train_wall=34, gb_free=10.1, wall=9761
2021-03-27 22:44:07 | INFO | train_inner | epoch 048:   564 / 588 loss=4.77, nll_loss=3.804, word_ins=5.147, word_crf=1.878, length=3.183, ppl=27.29, wps=20982.5, ups=2.93, wpb=7165, bsz=314.6, num_updates=28200, lr=0.000297746, gnorm=1.457, train_wall=35, gb_free=9.7, wall=9795
2021-03-27 22:44:16 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 22:44:19 | INFO | valid | epoch 048 | valid on 'valid' subset | loss 7.369 | nll_loss 5.451 | word_ins 6.683 | word_crf 3.558 | length 4.699 | ppl 165.33 | wps 59568.5 | wpb 5311.6 | bsz 208.1 | num_updates 28224 | best_loss 6.867
2021-03-27 22:44:19 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 48 @ 28224 updates
2021-03-27 22:44:19 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint48.pt
2021-03-27 22:44:20 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint48.pt
2021-03-27 22:44:23 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint48.pt (epoch 48 @ 28224 updates, score 7.369) (writing took 3.6523637139998755 seconds)
2021-03-27 22:44:23 | INFO | fairseq_cli.train | end of epoch 48 (average epoch stats below)
2021-03-27 22:44:23 | INFO | train | epoch 048 | loss 4.901 | nll_loss 3.956 | word_ins 5.282 | word_crf 1.942 | length 3.176 | ppl 29.88 | wps 19284.6 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 28224 | lr 0.000297619 | gnorm 1.545 | train_wall 204 | gb_free 9.4 | wall 9810
2021-03-27 22:44:23 | INFO | fairseq.trainer | begin training epoch 49
2021-03-27 22:44:23 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 22:44:50 | INFO | train_inner | epoch 049:   76 / 588 loss=4.999, nll_loss=4.088, word_ins=5.401, word_crf=1.992, length=3.058, ppl=31.97, wps=16152.6, ups=2.34, wpb=6915.5, bsz=228.7, num_updates=28300, lr=0.000297219, gnorm=1.488, train_wall=35, gb_free=9.4, wall=9838
2021-03-27 22:45:25 | INFO | train_inner | epoch 049:   176 / 588 loss=4.782, nll_loss=3.84, word_ins=5.183, word_crf=1.879, length=3.112, ppl=27.51, wps=20326, ups=2.86, wpb=7097.3, bsz=288.6, num_updates=28400, lr=0.000296695, gnorm=1.624, train_wall=35, gb_free=9.6, wall=9873
2021-03-27 22:45:59 | INFO | train_inner | epoch 049:   276 / 588 loss=4.795, nll_loss=3.857, word_ins=5.199, word_crf=1.883, length=3.121, ppl=27.76, wps=20007.1, ups=2.89, wpb=6933.6, bsz=276.9, num_updates=28500, lr=0.000296174, gnorm=1.494, train_wall=34, gb_free=9.4, wall=9907
2021-03-27 22:46:34 | INFO | train_inner | epoch 049:   376 / 588 loss=4.808, nll_loss=3.856, word_ins=5.198, word_crf=1.883, length=3.259, ppl=28.01, wps=20287.8, ups=2.87, wpb=7065.8, bsz=285.4, num_updates=28600, lr=0.000295656, gnorm=1.353, train_wall=35, gb_free=9.9, wall=9942
2021-03-27 22:47:10 | INFO | train_inner | epoch 049:   476 / 588 loss=5.052, nll_loss=4.116, word_ins=5.422, word_crf=2.028, length=3.125, ppl=33.17, wps=19630.7, ups=2.84, wpb=6920, bsz=247.1, num_updates=28700, lr=0.000295141, gnorm=1.508, train_wall=35, gb_free=9.1, wall=9977
2021-03-27 22:47:44 | INFO | train_inner | epoch 049:   576 / 588 loss=4.777, nll_loss=3.812, word_ins=5.159, word_crf=1.874, length=3.243, ppl=27.42, wps=20246.1, ups=2.89, wpb=6996.2, bsz=289.7, num_updates=28800, lr=0.000294628, gnorm=1.419, train_wall=34, gb_free=9.4, wall=10012
2021-03-27 22:47:48 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 22:47:52 | INFO | valid | epoch 049 | valid on 'valid' subset | loss 7.451 | nll_loss 5.474 | word_ins 6.714 | word_crf 3.627 | length 4.675 | ppl 175.02 | wps 59789.6 | wpb 5311.6 | bsz 208.1 | num_updates 28812 | best_loss 6.867
2021-03-27 22:47:52 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 49 @ 28812 updates
2021-03-27 22:47:52 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint49.pt
2021-03-27 22:47:52 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint49.pt
2021-03-27 22:47:55 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint49.pt (epoch 49 @ 28812 updates, score 7.451) (writing took 3.6403295090003667 seconds)
2021-03-27 22:47:55 | INFO | fairseq_cli.train | end of epoch 49 (average epoch stats below)
2021-03-27 22:47:55 | INFO | train | epoch 049 | loss 4.848 | nll_loss 3.907 | word_ins 5.242 | word_crf 1.912 | length 3.15 | ppl 28.81 | wps 19320.1 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 28812 | lr 0.000294566 | gnorm 1.48 | train_wall 204 | gb_free 9.8 | wall 10023
2021-03-27 22:47:55 | INFO | fairseq.trainer | begin training epoch 50
2021-03-27 22:47:55 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 22:48:26 | INFO | train_inner | epoch 050:   88 / 588 loss=4.708, nll_loss=3.77, word_ins=5.127, word_crf=1.827, length=3.17, ppl=26.13, wps=16778.9, ups=2.38, wpb=7038.1, bsz=279.2, num_updates=28900, lr=0.000294118, gnorm=1.444, train_wall=35, gb_free=10.5, wall=10054
2021-03-27 22:49:01 | INFO | train_inner | epoch 050:   188 / 588 loss=4.716, nll_loss=3.772, word_ins=5.129, word_crf=1.832, length=3.197, ppl=26.28, wps=20539.3, ups=2.87, wpb=7150.1, bsz=279.4, num_updates=29000, lr=0.00029361, gnorm=1.441, train_wall=35, gb_free=10.6, wall=10089

2021-03-27 22:49:36 | INFO | train_inner | epoch 050:   288 / 588 loss=4.883, nll_loss=3.935, word_ins=5.27, word_crf=1.927, length=3.219, ppl=29.51, wps=19826, ups=2.85, wpb=6957.3, bsz=253.8, num_updates=29100, lr=0.000293105, gnorm=1.543, train_wall=35, gb_free=9.4, wall=10124
2021-03-27 22:50:11 | INFO | train_inner | epoch 050:   388 / 588 loss=4.788, nll_loss=3.839, word_ins=5.183, word_crf=1.889, length=3.068, ppl=27.62, wps=19978.5, ups=2.86, wpb=6979.1, bsz=296.6, num_updates=29200, lr=0.000292603, gnorm=1.636, train_wall=35, gb_free=9.5, wall=10159
2021-03-27 22:50:46 | INFO | train_inner | epoch 050:   488 / 588 loss=5.017, nll_loss=4.08, word_ins=5.391, word_crf=2.002, length=3.189, ppl=32.38, wps=19604.5, ups=2.84, wpb=6891.7, bsz=243.7, num_updates=29300, lr=0.000292103, gnorm=1.559, train_wall=35, gb_free=9.3, wall=10194
2021-03-27 22:51:21 | INFO | train_inner | epoch 050:   588 / 588 loss=4.846, nll_loss=3.911, word_ins=5.244, word_crf=1.921, length=3.033, ppl=28.76, wps=19530, ups=2.83, wpb=6903.2, bsz=285.4, num_updates=29400, lr=0.000291606, gnorm=1.53, train_wall=35, gb_free=9.7, wall=10229
2021-03-27 22:51:21 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 22:51:25 | INFO | valid | epoch 050 | valid on 'valid' subset | loss 7.441 | nll_loss 5.488 | word_ins 6.719 | word_crf 3.612 | length 4.693 | ppl 173.79 | wps 59790.5 | wpb 5311.6 | bsz 208.1 | num_updates 29400 | best_loss 6.867
2021-03-27 22:51:25 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 50 @ 29400 updates
2021-03-27 22:51:25 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint50.pt
2021-03-27 22:51:26 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint50.pt
2021-03-27 22:51:28 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint50.pt (epoch 50 @ 29400 updates, score 7.441) (writing took 3.657989863999319 seconds)
2021-03-27 22:51:28 | INFO | fairseq_cli.train | end of epoch 50 (average epoch stats below)
2021-03-27 22:51:28 | INFO | train | epoch 050 | loss 4.831 | nll_loss 3.89 | word_ins 5.228 | word_crf 1.902 | length 3.146 | ppl 28.47 | wps 19282.6 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 29400 | lr 0.000291606 | gnorm 1.531 | train_wall 204 | gb_free 9.7 | wall 10236
2021-03-27 22:51:28 | INFO | fairseq.trainer | begin training epoch 51
2021-03-27 22:51:28 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 22:52:04 | INFO | train_inner | epoch 051:   100 / 588 loss=4.845, nll_loss=3.929, word_ins=5.265, word_crf=1.907, length=3.047, ppl=28.73, wps=16407.4, ups=2.34, wpb=6997.9, bsz=255.4, num_updates=29500, lr=0.000291111, gnorm=1.676, train_wall=35, gb_free=10.4, wall=10272
2021-03-27 22:52:39 | INFO | train_inner | epoch 051:   200 / 588 loss=4.777, nll_loss=3.857, word_ins=5.202, word_crf=1.874, length=3.022, ppl=27.42, wps=19346, ups=2.86, wpb=6773.2, bsz=268.1, num_updates=29600, lr=0.000290619, gnorm=1.7, train_wall=35, gb_free=10.3, wall=10307
2021-03-27 22:53:14 | INFO | train_inner | epoch 051:   300 / 588 loss=4.642, nll_loss=3.694, word_ins=5.061, word_crf=1.798, length=3.139, ppl=24.97, wps=20335.4, ups=2.9, wpb=7023.7, bsz=301.1, num_updates=29700, lr=0.000290129, gnorm=1.492, train_wall=34, gb_free=9.6, wall=10341
2021-03-27 22:53:48 | INFO | train_inner | epoch 051:   400 / 588 loss=4.797, nll_loss=3.863, word_ins=5.208, word_crf=1.886, length=3.071, ppl=27.79, wps=20060.4, ups=2.87, wpb=6999.5, bsz=261.4, num_updates=29800, lr=0.000289642, gnorm=1.498, train_wall=35, gb_free=9.2, wall=10376
2021-03-27 22:54:24 | INFO | train_inner | epoch 051:   500 / 588 loss=4.888, nll_loss=3.942, word_ins=5.274, word_crf=1.943, length=3.083, ppl=29.61, wps=20258.7, ups=2.85, wpb=7100.6, bsz=274.5, num_updates=29900, lr=0.000289157, gnorm=1.547, train_wall=35, gb_free=9.2, wall=10411
2021-03-27 22:54:54 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 22:54:58 | INFO | valid | epoch 051 | valid on 'valid' subset | loss 7.52 | nll_loss 5.507 | word_ins 6.758 | word_crf 3.676 | length 4.66 | ppl 183.6 | wps 59812.9 | wpb 5311.6 | bsz 208.1 | num_updates 29988 | best_loss 6.867
2021-03-27 22:54:58 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 51 @ 29988 updates
2021-03-27 22:54:58 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint51.pt
2021-03-27 22:54:58 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint51.pt
2021-03-27 22:55:01 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint51.pt (epoch 51 @ 29988 updates, score 7.52) (writing took 3.6306021899999905 seconds)
2021-03-27 22:55:01 | INFO | fairseq_cli.train | end of epoch 51 (average epoch stats below)
2021-03-27 22:55:01 | INFO | train | epoch 051 | loss 4.802 | nll_loss 3.866 | word_ins 5.209 | word_crf 1.887 | length 3.104 | ppl 27.9 | wps 19309.3 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 29988 | lr 0.000288733 | gnorm 1.573 | train_wall 204 | gb_free 9.2 | wall 10449
2021-03-27 22:55:01 | INFO | fairseq.trainer | begin training epoch 52
2021-03-27 22:55:01 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 22:55:06 | INFO | train_inner | epoch 052:   12 / 588 loss=4.856, nll_loss=3.905, word_ins=5.239, word_crf=1.914, length=3.234, ppl=28.97, wps=16765.7, ups=2.38, wpb=7046.1, bsz=275.4, num_updates=30000, lr=0.000288675, gnorm=1.511, train_wall=35, gb_free=9.6, wall=10453
2021-03-27 22:55:06 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 22:55:09 | INFO | valid | epoch 052 | valid on 'valid' subset | loss 7.576 | nll_loss 5.524 | word_ins 6.773 | word_crf 3.723 | length 4.66 | ppl 190.78 | wps 59702.8 | wpb 5311.6 | bsz 208.1 | num_updates 30000 | best_loss 6.867
2021-03-27 22:55:09 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 52 @ 30000 updates
2021-03-27 22:55:09 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint_52_30000.pt
2021-03-27 22:55:10 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint_52_30000.pt
2021-03-27 22:55:13 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint_52_30000.pt (epoch 52 @ 30000 updates, score 7.576) (writing took 3.9641098669999337 seconds)
2021-03-27 22:55:48 | INFO | train_inner | epoch 052:   112 / 588 loss=4.785, nll_loss=3.865, word_ins=5.213, word_crf=1.87, length=3.086, ppl=27.56, wps=16568.8, ups=2.36, wpb=7023.3, bsz=255.3, num_updates=30100, lr=0.000288195, gnorm=1.556, train_wall=35, gb_free=9.5, wall=10496
2021-03-27 22:56:23 | INFO | train_inner | epoch 052:   212 / 588 loss=4.725, nll_loss=3.816, word_ins=5.169, word_crf=1.842, length=2.986, ppl=26.45, wps=19426.9, ups=2.84, wpb=6833.9, bsz=262.5, num_updates=30200, lr=0.000287718, gnorm=1.66, train_wall=35, gb_free=9.6, wall=10531
2021-03-27 22:56:58 | INFO | train_inner | epoch 052:   312 / 588 loss=4.799, nll_loss=3.857, word_ins=5.203, word_crf=1.884, length=3.129, ppl=27.83, wps=20042.6, ups=2.84, wpb=7062.5, bsz=268.3, num_updates=30300, lr=0.000287242, gnorm=1.462, train_wall=35, gb_free=9.4, wall=10566
2021-03-27 22:57:33 | INFO | train_inner | epoch 052:   412 / 588 loss=4.763, nll_loss=3.814, word_ins=5.164, word_crf=1.865, length=3.153, ppl=27.15, wps=19952.5, ups=2.86, wpb=6978, bsz=285.5, num_updates=30400, lr=0.00028677, gnorm=1.513, train_wall=35, gb_free=9.3, wall=10601
2021-03-27 22:58:08 | INFO | train_inner | epoch 052:   512 / 588 loss=4.83, nll_loss=3.896, word_ins=5.236, word_crf=1.901, length=3.108, ppl=28.43, wps=19562.9, ups=2.85, wpb=6857, bsz=269.3, num_updates=30500, lr=0.000286299, gnorm=1.7, train_wall=35, gb_free=9.8, wall=10636
2021-03-27 22:58:35 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 22:58:38 | INFO | valid | epoch 052 | valid on 'valid' subset | loss 7.493 | nll_loss 5.49 | word_ins 6.733 | word_crf 3.654 | length 4.721 | ppl 180.12 | wps 59820.1 | wpb 5311.6 | bsz 208.1 | num_updates 30576 | best_loss 6.867
2021-03-27 22:58:38 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 52 @ 30576 updates
2021-03-27 22:58:38 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint52.pt
2021-03-27 22:58:39 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint52.pt
2021-03-27 22:58:41 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint52.pt (epoch 52 @ 30576 updates, score 7.493) (writing took 3.63967112499995 seconds)
2021-03-27 22:58:41 | INFO | fairseq_cli.train | end of epoch 52 (average epoch stats below)
2021-03-27 22:58:41 | INFO | train | epoch 052 | loss 4.771 | nll_loss 3.837 | word_ins 5.186 | word_crf 1.869 | length 3.097 | ppl 27.31 | wps 18649 | ups 2.67 | wpb 6988.7 | bsz 272.5 | num_updates 30576 | lr 0.000285943 | gnorm 1.567 | train_wall 204 | gb_free 9.2 | wall 10669
2021-03-27 22:58:42 | INFO | fairseq.trainer | begin training epoch 53
2021-03-27 22:58:42 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 22:58:50 | INFO | train_inner | epoch 053:   24 / 588 loss=4.702, nll_loss=3.755, word_ins=5.113, word_crf=1.837, length=3.082, ppl=26.03, wps=17294.6, ups=2.39, wpb=7242.9, bsz=293.6, num_updates=30600, lr=0.000285831, gnorm=1.472, train_wall=34, gb_free=9.4, wall=10678
2021-03-27 22:59:25 | INFO | train_inner | epoch 053:   124 / 588 loss=4.705, nll_loss=3.79, word_ins=5.148, word_crf=1.82, length=3.109, ppl=26.08, wps=19905.3, ups=2.85, wpb=6981.5, bsz=264.8, num_updates=30700, lr=0.000285365, gnorm=1.63, train_wall=35, gb_free=9.2, wall=10713
2021-03-27 23:00:00 | INFO | train_inner | epoch 053:   224 / 588 loss=4.714, nll_loss=3.789, word_ins=5.147, word_crf=1.842, length=2.978, ppl=26.24, wps=20033.7, ups=2.86, wpb=6994.9, bsz=271.4, num_updates=30800, lr=0.000284901, gnorm=1.61, train_wall=35, gb_free=9.5, wall=10748
2021-03-27 23:00:35 | INFO | train_inner | epoch 053:   324 / 588 loss=4.689, nll_loss=3.761, word_ins=5.123, word_crf=1.824, length=3.034, ppl=25.79, wps=20178.6, ups=2.87, wpb=7028.2, bsz=277.7, num_updates=30900, lr=0.00028444, gnorm=1.521, train_wall=35, gb_free=9.8, wall=10783
2021-03-27 23:01:10 | INFO | train_inner | epoch 053:   424 / 588 loss=4.891, nll_loss=3.939, word_ins=5.275, word_crf=1.927, length=3.265, ppl=29.68, wps=20291.3, ups=2.85, wpb=7113.7, bsz=262.1, num_updates=31000, lr=0.000283981, gnorm=1.476, train_wall=35, gb_free=10.1, wall=10818
2021-03-27 23:01:45 | INFO | train_inner | epoch 053:   524 / 588 loss=4.761, nll_loss=3.821, word_ins=5.173, word_crf=1.866, length=3.084, ppl=27.11, wps=19654.4, ups=2.85, wpb=6886.4, bsz=271.3, num_updates=31100, lr=0.000283524, gnorm=1.534, train_wall=35, gb_free=9.7, wall=10853
2021-03-27 23:02:07 | INFO | fairseq_cli.train | begin validation on "valid" subset

2021-03-27 23:02:10 | INFO | valid | epoch 053 | valid on 'valid' subset | loss 7.605 | nll_loss 5.534 | word_ins 6.785 | word_crf 3.735 | length 4.779 | ppl 194.73 | wps 60008.9 | wpb 5311.6 | bsz 208.1 | num_updates 31164 | best_loss 6.867
2021-03-27 23:02:10 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 53 @ 31164 updates
2021-03-27 23:02:10 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint53.pt
2021-03-27 23:02:11 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint53.pt
2021-03-27 23:02:14 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint53.pt (epoch 53 @ 31164 updates, score 7.605) (writing took 3.633895193999706 seconds)
2021-03-27 23:02:14 | INFO | fairseq_cli.train | end of epoch 53 (average epoch stats below)
2021-03-27 23:02:14 | INFO | train | epoch 053 | loss 4.727 | nll_loss 3.795 | word_ins 5.151 | word_crf 1.843 | length 3.085 | ppl 26.48 | wps 19336.3 | ups 2.77 | wpb 6988.7 | bsz 272.5 | num_updates 31164 | lr 0.000283233 | gnorm 1.531 | train_wall 204 | gb_free 9.4 | wall 10882
2021-03-27 23:02:14 | INFO | fairseq.trainer | begin training epoch 54
2021-03-27 23:02:14 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 23:02:27 | INFO | train_inner | epoch 054:   36 / 588 loss=4.506, nll_loss=3.575, word_ins=4.961, word_crf=1.715, length=3.109, ppl=22.72, wps=16511.2, ups=2.41, wpb=6837.4, bsz=289.3, num_updates=31200, lr=0.000283069, gnorm=1.422, train_wall=34, gb_free=9.3, wall=10894
2021-03-27 23:03:02 | INFO | train_inner | epoch 054: 136 / 588 loss=4.677, nll_loss=3.78, word_ins=5.141, word_crf=1.823, length=2.832, ppl=25.57, wps=19551.1, ups=2.83, wpb=6896.8, bsz=271.9, num_updates=31300, lr=0.000282617, gnorm=1.551, train_wall=35, gb_free=9.4, wall=10930
2021-03-27 23:03:37 | INFO | train_inner | epoch 054: 236 / 588 loss=4.744, nll_loss=3.82, word_ins=5.175, word_crf=1.848, length=3.087, ppl=26.8, wps=19902.8, ups=2.86, wpb=6962.9, bsz=265, num_updates=31400, lr=0.000282166, gnorm=1.532, train_wall=35, gb_free=9.7, wall=10965
2021-03-27 23:04:12 | INFO | train_inner | epoch 054: 336 / 588 loss=4.726, nll_loss=3.789, word_ins=5.149, word_crf=1.836, length=3.151, ppl=26.46, wps=20312.6, ups=2.85, wpb=7130.1, bsz=262.5, num_updates=31500, lr=0.000281718, gnorm=1.507, train_wall=35, gb_free=10.5, wall=11000
2021-03-27 23:04:47 | INFO | train_inner | epoch 054: 436 / 588 loss=4.739, nll_loss=3.803, word_ins=5.158, word_crf=1.85, length=3.099, ppl=26.7, wps=20002.9, ups=2.88, wpb=6950.9, bsz=269, num_updates=31600, lr=0.000281272, gnorm=1.621, train_wall=34, gb_free=9.4, wall=11035
2021-03-27 23:05:22 | INFO | train_inner | epoch 054: 536 / 588 loss=4.592, nll_loss=3.642, word_ins=5.019, word_crf=1.777, length=3.055, ppl=24.11, wps=20413.6, ups=2.87, wpb=7115.9, bsz=304.1, num_updates=31700, lr=0.000280828, gnorm=1.463, train_wall=35, gb_free=9.9, wall=11069
2021-03-27 23:05:40 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 23:05:43 | INFO | valid | epoch 054 | valid on 'valid' subset | loss 7.691 | nll_loss 5.581 | word_ins 6.833 | word_crf 3.8 | length 4.746 | ppl 206.71 | wps 59860.2 | wpb 5311.6 | bsz 208.1 | num_updates 31752 | best_loss 6.867
2021-03-27 23:05:43 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 54 @ 31752 updates
2021-03-27 23:05:43 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint54.pt
2021-03-27 23:05:44 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint54.pt
2021-03-27 23:05:47 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint54.pt (epoch 54 @ 31752 updates, score 7.691) (writing took 3.647782422000091 seconds)
2021-03-27 23:05:47 | INFO | fairseq_cli.train | end of epoch 54 (average epoch stats below)
2021-03-27 23:05:47 | INFO | train | epoch 054 | loss 4.688 | nll_loss 3.76 | word_ins 5.122 | word_crf 1.821 | length 3.059 | ppl 25.78 | wps 19337.7 | ups 2.77 | wpb 6988.7 | bsz 272.5 | num_updates 31752 | lr 0.000280598 | gnorm 1.526 | train_wall 204 | gb_free 10 | wall 11094
2021-03-27 23:05:47 | INFO | fairseq.trainer | begin training epoch 55
2021-03-27 23:05:47 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 23:06:04 | INFO | train_inner | epoch 055:   48 / 588 loss=4.7, nll_loss=3.786, word_ins=5.145, word_crf=1.823, length=3.048, ppl=25.99, wps=16430.6, ups=2.37, wpb=6923.2, bsz=260.8, num_updates=31800, lr=0.000280386, gnorm=1.469, train_wall=35, gb_free=9.6, wall=11112
2021-03-27 23:06:39 | INFO | train_inner | epoch 055: 148 / 588 loss=4.755, nll_loss=3.859, word_ins=5.21, word_crf=1.856, length=2.944, ppl=27.01, wps=19390.7, ups=2.81, wpb=6898.8, bsz=246.5, num_updates=31900, lr=0.000279946, gnorm=1.699, train_wall=35, gb_free=9.4, wall=11147
2021-03-27 23:07:15 | INFO | train_inner | epoch 055: 248 / 588 loss=4.663, nll_loss=3.73, word_ins=5.097, word_crf=1.812, length=3.022, ppl=25.33, wps=20142.5, ups=2.83, wpb=7123.8, bsz=291, num_updates=32000, lr=0.000279508, gnorm=1.549, train_wall=35, gb_free=9.5, wall=11183
2021-03-27 23:07:50 | INFO | train_inner | epoch 055: 348 / 588 loss=4.581, nll_loss=3.647, word_ins=5.027, word_crf=1.759, length=3.086, ppl=23.94, wps=19909, ups=2.86, wpb=6950.9, bsz=277.9, num_updates=32100, lr=0.000279073, gnorm=1.633, train_wall=35, gb_free=9.8, wall=11217
2021-03-27 23:08:25 | INFO | train_inner | epoch 055: 448 / 588 loss=4.721, nll_loss=3.795, word_ins=5.152, word_crf=1.841, length=3.036, ppl=26.37, wps=19584.6, ups=2.84, wpb=6904.2, bsz=269.4, num_updates=32200, lr=0.000278639, gnorm=1.521, train_wall=35, gb_free=10.1, wall=11253
2021-03-27 23:08:59 | INFO | train_inner | epoch 055: 548 / 588 loss=4.659, nll_loss=3.706, word_ins=5.077, word_crf=1.803, length=3.17, ppl=25.26, wps=20227.5, ups=2.89, wpb=6997.7, bsz=274.3, num_updates=32300, lr=0.000278207, gnorm=1.528, train_wall=34, gb_free=9.2, wall=11287
2021-03-27 23:09:13 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 23:09:17 | INFO | valid | epoch 055 | valid on 'valid' subset | loss 7.661 | nll_loss 5.569 | word_ins 6.82 | word_crf 3.776 | length 4.758 | ppl 202.46 | wps 59742.5 | wpb 5311.6 | bsz 208.1 | num_updates 32340 | best_loss 6.867
2021-03-27 23:09:17 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 55 @ 32340 updates
2021-03-27 23:09:17 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint55.pt
2021-03-27 23:09:18 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint55.pt
2021-03-27 23:09:35 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint55.pt (epoch 55 @ 32340 updates, score 7.661) (writing took 17.93628745800015 seconds)
2021-03-27 23:09:35 | INFO | fairseq_cli.train | end of epoch 55 (average epoch stats below)
2021-03-27 23:09:35 | INFO | train | epoch 055 | loss 4.672 | nll_loss 3.745 | word_ins 5.11 | word_crf 1.812 | length 3.047 | ppl 25.49 | wps 18012.1 | ups 2.58 | wpb 6988.7 | bsz 272.5 | num_updates 32340 | lr 0.000278035 | gnorm 1.566 | train_wall 205 | gb_free 9.7 | wall 11323
2021-03-27 23:09:35 | INFO | fairseq.trainer | begin training epoch 56
2021-03-27 23:09:35 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 23:09:56 | INFO | train_inner | epoch 056:   60 / 588 loss=4.544, nll_loss=3.626, word_ins=5.007, word_crf=1.743, length=2.985, ppl=23.33, wps=12389.3, ups=1.77, wpb=7007.3, bsz=289.3, num_updates=32400, lr=0.000277778, gnorm=1.457, train_wall=35, gb_free=9.7, wall=11344
2021-03-27 23:10:31 | INFO | train_inner | epoch 056: 160 / 588 loss=4.413, nll_loss=3.496, word_ins=4.898, word_crf=1.672, length=2.914, ppl=21.3, wps=20413.8, ups=2.88, wpb=7085, bsz=302.5, num_updates=32500, lr=0.00027735, gnorm=1.556, train_wall=34, gb_free=9.2, wall=11379
2021-03-27 23:11:06 | INFO | train_inner | epoch 056: 260 / 588 loss=4.689, nll_loss=3.78, word_ins=5.145, word_crf=1.818, length=2.99, ppl=25.8, wps=19515.6, ups=2.83, wpb=6904.8, bsz=251.5, num_updates=32600, lr=0.000276924, gnorm=1.512, train_wall=35, gb_free=9.5, wall=11414
2021-03-27 23:11:41 | INFO | train_inner | epoch 056: 360 / 588 loss=4.638, nll_loss=3.721, word_ins=5.093, word_crf=1.794, length=2.974, ppl=24.89, wps=19874.2, ups=2.86, wpb=6953.6, bsz=266.2, num_updates=32700, lr=0.000276501, gnorm=1.451, train_wall=35, gb_free=9.9, wall=11449
2021-03-27 23:12:16 | INFO | train_inner | epoch 056: 460 / 588 loss=4.785, nll_loss=3.841, word_ins=5.192, word_crf=1.875, length=3.135, ppl=27.57, wps=20347, ups=2.86, wpb=7110.1, bsz=268.4, num_updates=32800, lr=0.000276079, gnorm=1.68, train_wall=35, gb_free=9.9, wall=11484
2021-03-27 23:12:51 | INFO | train_inner | epoch 056: 560 / 588 loss=4.796, nll_loss=3.866, word_ins=5.215, word_crf=1.878, length=3.105, ppl=27.77, wps=19610.7, ups=2.84, wpb=6915.3, bsz=250.5, num_updates=32900, lr=0.000275659, gnorm=1.605, train_wall=35, gb_free=9.4, wall=11519
2021-03-27 23:13:01 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 23:13:04 | INFO | valid | epoch 056 | valid on 'valid' subset | loss 7.725 | nll_loss 5.602 | word_ins 6.853 | word_crf 3.824 | length 4.742 | ppl 211.55 | wps 59948.8 | wpb 5311.6 | bsz 208.1 | num_updates 32928 | best_loss 6.867
2021-03-27 23:13:04 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 56 @ 32928 updates
2021-03-27 23:13:04 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint56.pt
2021-03-27 23:13:05 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint56.pt
2021-03-27 23:13:08 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint56.pt (epoch 56 @ 32928 updates, score 7.725) (writing took 3.640582918000291 seconds)
2021-03-27 23:13:08 | INFO | fairseq_cli.train | end of epoch 56 (average epoch stats below)
2021-03-27 23:13:08 | INFO | train | epoch 056 | loss 4.634 | nll_loss 3.711 | word_ins 5.083 | word_crf 1.791 | length 3.016 | ppl 24.82 | wps 19286.7 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 32928 | lr 0.000275542 | gnorm 1.541 | train_wall 204 | gb_free 9.5 | wall 11536
2021-03-27 23:13:08 | INFO | fairseq.trainer | begin training epoch 57
2021-03-27 23:13:08 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 23:13:34 | INFO | train_inner | epoch 057:   72 / 588 loss=4.69, nll_loss=3.785, word_ins=5.146, word_crf=1.831, length=2.855, ppl=25.81, wps=16204.4, ups=2.35, wpb=6892.3, bsz=276.9, num_updates=33000, lr=0.000275241, gnorm=1.82, train_wall=35, gb_free=9.3, wall=11562
2021-03-27 23:14:09 | INFO | train_inner | epoch 057: 172 / 588 loss=4.477, nll_loss=3.559, word_ins=4.953, word_crf=1.693, length=3.071, ppl=22.26, wps=20043.1, ups=2.87, wpb=6979.1, bsz=280, num_updates=33100, lr=0.000274825, gnorm=1.398, train_wall=35, gb_free=9.7, wall=11596

2021-03-27 23:14:44 | INFO | train_inner | epoch 057:   272 / 588 loss=4.688, nll_loss=3.773, word_ins=5.143, word_crf=1.809, length=3.072, ppl=25.77, wps=20074.9, ups=2.85, wpb=7040.4, bsz=239.3, num_updates=33200, lr=0.000274411, gnorm=1.439, train_wall=35, gb_free=9.4, wall=11632
2021-03-27 23:15:19 | INFO | train_inner | epoch 057:   372 / 588 loss=4.635, nll_loss=3.712, word_ins=5.084, word_crf=1.796, length=2.975, ppl=24.85, wps=20021.2, ups=2.86, wpb=7005.9, bsz=277.1, num_updates=33300, lr=0.000273998, gnorm=1.643, train_wall=35, gb_free=9.5, wall=11667
2021-03-27 23:15:53 | INFO | train_inner | epoch 057:   472 / 588 loss=4.495, nll_loss=3.559, word_ins=4.95, word_crf=1.724, length=2.957, ppl=22.55, wps=20244.6, ups=2.88, wpb=7021.2, bsz=308.4, num_updates=33400, lr=0.000273588, gnorm=1.593, train_wall=34, gb_free=9.5, wall=11701
2021-03-27 23:16:28 | INFO | train_inner | epoch 057:   572 / 588 loss=4.725, nll_loss=3.788, word_ins=5.149, word_crf=1.838, length=3.125, ppl=26.45, wps=19957.8, ups=2.87, wpb=6942.2, bsz=257.8, num_updates=33500, lr=0.000273179, gnorm=1.578, train_wall=35, gb_free=9.3, wall=11736
2021-03-27 23:16:34 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 23:16:37 | INFO | valid | epoch 057 | valid on 'valid' subset | loss 7.656 | nll_loss 5.574 | word_ins 6.819 | word_crf 3.771 | length 4.751 | ppl 201.72 | wps 59895.7 | wpb 5311.6 | bsz 208.1 | num_updates 33516 | best_loss 6.867
2021-03-27 23:16:37 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 57 @ 33516 updates
2021-03-27 23:16:37 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint57.pt
2021-03-27 23:16:38 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint57.pt
2021-03-27 23:16:41 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint57.pt (epoch 57 @ 33516 updates, score 7.656) (writing took 3.645455341998968 seconds)
2021-03-27 23:16:41 | INFO | fairseq_cli.train | end of epoch 57 (average epoch stats below)
2021-03-27 23:16:41 | INFO | train | epoch 057 | loss 4.619 | nll_loss 3.697 | word_ins 5.072 | word_crf 1.783 | length 3.005 | ppl 24.57 | wps 19297.5 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 33516 | lr 0.000273114 | gnorm 1.579 | train_wall 204 | gb_free 9.6 | wall 11749
2021-03-27 23:16:41 | INFO | fairseq.trainer | begin training epoch 58
2021-03-27 23:16:41 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 23:17:11 | INFO | train_inner | epoch 058:   84 / 588 loss=4.47, nll_loss=3.582, word_ins=4.974, word_crf=1.701, length=2.824, ppl=22.16, wps=16407.8, ups=2.36, wpb=6963.1, bsz=281, num_updates=33600, lr=0.000272772, gnorm=1.41, train_wall=35, gb_free=9.4, wall=11778
2021-03-27 23:17:46 | INFO | train_inner | epoch 058:   184 / 588 loss=4.514, nll_loss=3.592, word_ins=4.982, word_crf=1.726, length=2.963, ppl=22.84, wps=20160, ups=2.86, wpb=7046, bsz=291.3, num_updates=33700, lr=0.000272367, gnorm=1.669, train_wall=35, gb_free=9.3, wall=11813
2021-03-27 23:18:21 | INFO | train_inner | epoch 058:   284 / 588 loss=4.662, nll_loss=3.751, word_ins=5.12, word_crf=1.806, length=2.957, ppl=25.32, wps=19782.3, ups=2.85, wpb=6947.8, bsz=254.8, num_updates=33800, lr=0.000271964, gnorm=1.734, train_wall=35, gb_free=9.4, wall=11849
2021-03-27 23:18:56 | INFO | train_inner | epoch 058:   384 / 588 loss=4.534, nll_loss=3.619, word_ins=5.005, word_crf=1.733, length=2.982, ppl=23.17, wps=19671.2, ups=2.86, wpb=6870.6, bsz=281.8, num_updates=33900, lr=0.000271563, gnorm=1.583, train_wall=35, gb_free=10.1, wall=11883
2021-03-27 23:19:31 | INFO | train_inner | epoch 058:   484 / 588 loss=4.669, nll_loss=3.74, word_ins=5.11, word_crf=1.813, length=3.014, ppl=25.44, wps=20419.7, ups=2.84, wpb=7177.7, bsz=264.7, num_updates=34000, lr=0.000271163, gnorm=1.47, train_wall=35, gb_free=9.2, wall=11919
2021-03-27 23:20:06 | INFO | train_inner | epoch 058:   584 / 588 loss=4.607, nll_loss=3.667, word_ins=5.046, word_crf=1.77, length=3.134, ppl=24.37, wps=19957.9, ups=2.87, wpb=6944.1, bsz=267.9, num_updates=34100, lr=0.000270765, gnorm=1.482, train_wall=35, gb_free=9.4, wall=11953
2021-03-27 23:20:07 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 23:20:10 | INFO | valid | epoch 058 | valid on 'valid' subset | loss 7.68 | nll_loss 5.563 | word_ins 6.813 | word_crf 3.791 | length 4.828 | ppl 205.12 | wps 59616.7 | wpb 5311.6 | bsz 208.1 | num_updates 34104 | best_loss 6.867
2021-03-27 23:20:10 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 58 @ 34104 updates
2021-03-27 23:20:10 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint58.pt
2021-03-27 23:20:11 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint58.pt
2021-03-27 23:20:14 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint58.pt (epoch 58 @ 34104 updates, score 7.68) (writing took 3.641904075000639 seconds)
2021-03-27 23:20:14 | INFO | fairseq_cli.train | end of epoch 58 (average epoch stats below)
2021-03-27 23:20:14 | INFO | train | epoch 058 | loss 4.578 | nll_loss 3.661 | word_ins 5.042 | word_crf 1.759 | length 2.979 | ppl 23.89 | wps 19275.8 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 34104 | lr 0.000270749 | gnorm 1.564 | train_wall 204 | gb_free 9.5 | wall 11962
2021-03-27 23:20:14 | INFO | fairseq.trainer | begin training epoch 59
2021-03-27 23:20:14 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 23:20:48 | INFO | train_inner | epoch 059:   96 / 588 loss=4.492, nll_loss=3.598, word_ins=4.991, word_crf=1.706, length=2.91, ppl=22.5, wps=16475.8, ups=2.37, wpb=6962.9, bsz=262.7, num_updates=34200, lr=0.000270369, gnorm=1.575, train_wall=35, gb_free=9.4, wall=11996
2021-03-27 23:21:23 | INFO | train_inner | epoch 059:   196 / 588 loss=4.46, nll_loss=3.538, word_ins=4.938, word_crf=1.698, length=2.929, ppl=22.01, wps=20416.3, ups=2.88, wpb=7096.3, bsz=291.5, num_updates=34300, lr=0.000269975, gnorm=1.668, train_wall=34, gb_free=10.3, wall=12030
2021-03-27 23:21:58 | INFO | train_inner | epoch 059:   296 / 588 loss=4.602, nll_loss=3.695, word_ins=5.075, word_crf=1.766, length=2.98, ppl=24.28, wps=19765.1, ups=2.85, wpb=6943.8, bsz=247, num_updates=34400, lr=0.000269582, gnorm=1.594, train_wall=35, gb_free=9.5, wall=12066
2021-03-27 23:22:33 | INFO | train_inner | epoch 059:   396 / 588 loss=4.556, nll_loss=3.634, word_ins=5.018, word_crf=1.756, length=2.906, ppl=23.52, wps=19807.5, ups=2.82, wpb=7016, bsz=288.9, num_updates=34500, lr=0.000269191, gnorm=1.474, train_wall=35, gb_free=9.7, wall=12101
2021-03-27 23:23:08 | INFO | train_inner | epoch 059:   496 / 588 loss=4.689, nll_loss=3.765, word_ins=5.133, word_crf=1.814, length=3.091, ppl=25.8, wps=19975.9, ups=2.86, wpb=6989.1, bsz=253.3, num_updates=34600, lr=0.000268802, gnorm=1.575, train_wall=35, gb_free=9.7, wall=12136
2021-03-27 23:23:40 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 23:23:43 | INFO | valid | epoch 059 | valid on 'valid' subset | loss 7.627 | nll_loss 5.557 | word_ins 6.798 | word_crf 3.741 | length 4.87 | ppl 197.74 | wps 59877.7 | wpb 5311.6 | bsz 208.1 | num_updates 34692 | best_loss 6.867
2021-03-27 23:23:43 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 59 @ 34692 updates
2021-03-27 23:23:43 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint59.pt
2021-03-27 23:23:44 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint59.pt
2021-03-27 23:23:47 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint59.pt (epoch 59 @ 34692 updates, score 7.627) (writing took 3.640680857000916 seconds)
2021-03-27 23:23:47 | INFO | fairseq_cli.train | end of epoch 59 (average epoch stats below)
2021-03-27 23:23:47 | INFO | train | epoch 059 | loss 4.55 | nll_loss 3.633 | word_ins 5.019 | word_crf 1.742 | length 2.982 | ppl 23.43 | wps 19288.7 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 34692 | lr 0.000268445 | gnorm 1.568 | train_wall 204 | gb_free 9.2 | wall 12175
2021-03-27 23:23:47 | INFO | fairseq.trainer | begin training epoch 60
2021-03-27 23:23:47 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 23:23:50 | INFO | train_inner | epoch 060:   8 / 588 loss=4.469, nll_loss=3.533, word_ins=4.929, word_crf=1.694, length=3.105, ppl=22.15, wps=16597.2, ups=2.39, wpb=6934.2, bsz=298.1, num_updates=34700, lr=0.000268414, gnorm=1.507, train_wall=34, gb_free=10.4, wall=12178
2021-03-27 23:24:25 | INFO | train_inner | epoch 060:   108 / 588 loss=4.412, nll_loss=3.53, word_ins=4.932, word_crf=1.661, length=2.856, ppl=21.29, wps=19555.3, ups=2.88, wpb=6793.3, bsz=268.9, num_updates=34800, lr=0.000268028, gnorm=1.684, train_wall=34, gb_free=9.9, wall=12212
2021-03-27 23:24:59 | INFO | train_inner | epoch 060:   208 / 588 loss=4.489, nll_loss=3.563, word_ins=4.962, word_crf=1.703, length=3.043, ppl=22.45, wps=20580.5, ups=2.88, wpb=7158.4, bsz=274.1, num_updates=34900, lr=0.000267644, gnorm=1.751, train_wall=35, gb_free=9.5, wall=12247
2021-03-27 23:25:35 | INFO | train_inner | epoch 060:   308 / 588 loss=4.632, nll_loss=3.731, word_ins=5.105, word_crf=1.79, length=2.895, ppl=24.79, wps=19559.2, ups=2.8, wpb=6987.1, bsz=254.6, num_updates=35000, lr=0.000267261, gnorm=1.48, train_wall=35, gb_free=9.4, wall=12283
2021-03-27 23:26:10 | INFO | train_inner | epoch 060:   408 / 588 loss=4.638, nll_loss=3.726, word_ins=5.101, word_crf=1.788, length=2.997, ppl=24.9, wps=19550.7, ups=2.83, wpb=6912.6, bsz=254.7, num_updates=35100, lr=0.00026688, gnorm=1.572, train_wall=35, gb_free=9.5, wall=12318
2021-03-27 23:26:45 | INFO | train_inner | epoch 060:   508 / 588 loss=4.576, nll_loss=3.644, word_ins=5.029, word_crf=1.763, length=2.98, ppl=23.84, wps=20180.7, ups=2.85, wpb=7069.2, bsz=281, num_updates=35200, lr=0.000266501, gnorm=1.646, train_wall=35, gb_free=9.7, wall=12353
2021-03-27 23:27:13 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 23:27:16 | INFO | valid | epoch 060 | valid on 'valid' subset | loss 7.751 | nll_loss 5.589 | word_ins 6.848 | word_crf 3.838 | length 4.898 | ppl 215.49 | wps 59796.8 | wpb 5311.6 | bsz 208.1 | num_updates 35280 | best_loss 6.867
2021-03-27 23:27:16 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 60 @ 35280 updates
2021-03-27 23:27:16 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint60.pt
2021-03-27 23:27:17 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint60.pt
2021-03-27 23:27:20 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint60.pt (epoch 60 @ 35280 updates, score 7.751) (writing took 3.6377237389988295 seconds)
2021-03-27 23:27:20 | INFO | fairseq_cli.train | end of epoch 60 (average epoch stats below)
2021-03-27 23:27:20 | INFO | train | epoch 060 | loss 4.526 | nll_loss 3.613 | word_ins 5.003 | word_crf 1.729 | length 2.963 | ppl 23.05 | wps 19276 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 35280 | lr 0.000266199 | gnorm 1.6 | train_wall 204 | gb_free 9.5 | wall 12388

2021-03-27 23:27:20 | INFO | fairseq.trainer | begin training epoch 61
2021-03-27 23:27:20 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 23:27:27 | INFO | train_inner | epoch 061:   20 / 588 loss=4.377, nll_loss=3.453, word_ins=4.862, word_crf=1.649, length=2.972, ppl=20.78, wps=16854.6, ups=2.4, wpb=7032, bsz=300.2, num_updates=35300, lr=0.000266123, gnorm=1.455, train_wall=34, gb_free=9.3, wall=12395
2021-03-27 23:28:02 | INFO | train_inner | epoch 061:   120 / 588 loss=4.447, nll_loss=3.545, word_ins=4.945, word_crf=1.689, length=2.851, ppl=21.81, wps=19981.8, ups=2.85, wpb=7008.1, bsz=282.4, num_updates=35400, lr=0.000265747, gnorm=1.74, train_wall=35, gb_free=9.2, wall=12430
2021-03-27 23:28:37 | INFO | train_inner | epoch 061:   220 / 588 loss=4.485, nll_loss=3.595, word_ins=4.992, word_crf=1.693, length=2.955, ppl=22.39, wps=19628.5, ups=2.87, wpb=6841.4, bsz=249.3, num_updates=35500, lr=0.000265372, gnorm=1.523, train_wall=35, gb_free=9.5, wall=12465
2021-03-27 23:29:12 | INFO | train_inner | epoch 061:   320 / 588 loss=4.374, nll_loss=3.454, word_ins=4.87, word_crf=1.638, length=3.009, ppl=20.73, wps=20311.8, ups=2.89, wpb=7037.8, bsz=279.8, num_updates=35600, lr=0.000264999, gnorm=1.46, train_wall=34, gb_free=10.1, wall=12500
2021-03-27 23:29:47 | INFO | train_inner | epoch 061:   420 / 588 loss=4.698, nll_loss=3.782, word_ins=5.147, word_crf=1.835, length=2.897, ppl=25.96, wps=19837.9, ups=2.82, wpb=7022.8, bsz=262.2, num_updates=35700, lr=0.000264628, gnorm=1.796, train_wall=35, gb_free=9.6, wall=12535
2021-03-27 23:30:22 | INFO | train_inner | epoch 061:   520 / 588 loss=4.607, nll_loss=3.691, word_ins=5.069, word_crf=1.782, length=2.907, ppl=24.37, wps=19881.4, ups=2.85, wpb=6973.3, bsz=272.1, num_updates=35800, lr=0.000264258, gnorm=1.587, train_wall=35, gb_free=9.4, wall=12570
2021-03-27 23:30:46 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 23:30:49 | INFO | valid | epoch 061 | valid on 'valid' subset | loss 7.711 | nll_loss 5.564 | word_ins 6.819 | word_crf 3.813 | length 4.882 | ppl 209.57 | wps 59996.7 | wpb 5311.6 | bsz 208.1 | num_updates 35868 | best_loss 6.867
2021-03-27 23:30:49 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 61 @ 35868 updates
2021-03-27 23:30:49 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint61.pt
2021-03-27 23:30:50 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint61.pt
2021-03-27 23:30:53 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint61.pt (epoch 61 @ 35868 updates, score 7.711) (writing took 3.6418758799991338 seconds)
2021-03-27 23:30:53 | INFO | fairseq_cli.train | end of epoch 61 (average epoch stats below)
2021-03-27 23:30:53 | INFO | train | epoch 061 | loss 4.502 | nll_loss 3.592 | word_ins 4.986 | word_crf 1.717 | length 2.921 | ppl 22.66 | wps 19297.5 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 35868 | lr 0.000264008 | gnorm 1.596 | train_wall 204 | gb_free 9.5 | wall 12601
2021-03-27 23:30:53 | INFO | fairseq.trainer | begin training epoch 62
2021-03-27 23:30:53 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 23:31:05 | INFO | train_inner | epoch 062:   32 / 588 loss=4.406, nll_loss=3.493, word_ins=4.9, word_crf=1.672, length=2.842, ppl=21.2, wps=16778.6, ups=2.36, wpb=7102, bsz=292.7, num_updates=35900, lr=0.00026389, gnorm=1.426, train_wall=35, gb_free=9.6, wall=12612
2021-03-27 23:31:40 | INFO | train_inner | epoch 062:   132 / 588 loss=4.435, nll_loss=3.544, word_ins=4.946, word_crf=1.676, length=2.856, ppl=21.63, wps=19480.4, ups=2.84, wpb=6857.2, bsz=269, num_updates=36000, lr=0.000263523, gnorm=1.798, train_wall=35, gb_free=9.3, wall=12648
2021-03-27 23:32:14 | INFO | train_inner | epoch 062:   232 / 588 loss=4.393, nll_loss=3.48, word_ins=4.892, word_crf=1.655, length=2.918, ppl=21, wps=20130.5, ups=2.89, wpb=6962.9, bsz=272.4, num_updates=36100, lr=0.000263158, gnorm=1.659, train_wall=34, gb_free=9.5, wall=12682
2021-03-27 23:32:50 | INFO | train_inner | epoch 062:   332 / 588 loss=4.591, nll_loss=3.69, word_ins=5.072, word_crf=1.768, length=2.868, ppl=24.1, wps=19944.3, ups=2.83, wpb=7038.7, bsz=263.1, num_updates=36200, lr=0.000262794, gnorm=1.58, train_wall=35, gb_free=9.8, wall=12718
2021-03-27 23:33:24 | INFO | train_inner | epoch 062:   432 / 588 loss=4.451, nll_loss=3.538, word_ins=4.94, word_crf=1.69, length=2.912, ppl=21.88, wps=20144.9, ups=2.87, wpb=7014.2, bsz=279.6, num_updates=36300, lr=0.000262432, gnorm=1.659, train_wall=35, gb_free=10.7, wall=12752
2021-03-27 23:33:59 | INFO | train_inner | epoch 062:   532 / 588 loss=4.503, nll_loss=3.572, word_ins=4.969, word_crf=1.71, length=3.09, ppl=22.68, wps=20252.5, ups=2.9, wpb=6993.8, bsz=272.3, num_updates=36400, lr=0.000262071, gnorm=1.537, train_wall=34, gb_free=9.2, wall=12787
2021-03-27 23:34:19 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 23:34:22 | INFO | valid | epoch 062 | valid on 'valid' subset | loss 7.79 | nll_loss 5.617 | word_ins 6.868 | word_crf 3.863 | length 4.93 | ppl 221.32 | wps 59915.6 | wpb 5311.6 | bsz 208.1 | num_updates 36456 | best_loss 6.867
2021-03-27 23:34:22 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 62 @ 36456 updates
2021-03-27 23:34:22 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint62.pt
2021-03-27 23:34:23 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint62.pt
2021-03-27 23:34:26 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint62.pt (epoch 62 @ 36456 updates, score 7.79) (writing took 3.6517018529993948 seconds)
2021-03-27 23:34:26 | INFO | fairseq_cli.train | end of epoch 62 (average epoch stats below)
2021-03-27 23:34:26 | INFO | train | epoch 062 | loss 4.476 | nll_loss 3.568 | word_ins 4.966 | word_crf 1.702 | length 2.913 | ppl 22.26 | wps 19313.5 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 36456 | lr 0.00026187 | gnorm 1.622 | train_wall 204 | gb_free 9.8 | wall 12814
2021-03-27 23:34:26 | INFO | fairseq.trainer | begin training epoch 63
2021-03-27 23:34:26 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 23:34:41 | INFO | train_inner | epoch 063:   44 / 588 loss=4.466, nll_loss=3.572, word_ins=4.968, word_crf=1.698, length=2.837, ppl=22.11, wps=16617.3, ups=2.36, wpb=7029.6, bsz=275.2, num_updates=36500, lr=0.000261712, gnorm=1.569, train_wall=35, gb_free=9.7, wall=12829
2021-03-27 23:35:17 | INFO | train_inner | epoch 063:   144 / 588 loss=4.439, nll_loss=3.551, word_ins=4.957, word_crf=1.67, length=2.906, ppl=21.69, wps=20326.1, ups=2.84, wpb=7151.3, bsz=252, num_updates=36600, lr=0.000261354, gnorm=1.534, train_wall=35, gb_free=9.4, wall=12864
2021-03-27 23:35:51 | INFO | train_inner | epoch 063:   244 / 588 loss=4.4, nll_loss=3.483, word_ins=4.895, word_crf=1.664, length=2.886, ppl=21.11, wps=19936.1, ups=2.9, wpb=6882.9, bsz=284.4, num_updates=36700, lr=0.000260998, gnorm=1.865, train_wall=34, gb_free=9.5, wall=12899
2021-03-27 23:36:26 | INFO | train_inner | epoch 063:   344 / 588 loss=4.387, nll_loss=3.479, word_ins=4.889, word_crf=1.649, length=2.932, ppl=20.92, wps=20271.1, ups=2.87, wpb=7060.3, bsz=287.8, num_updates=36800, lr=0.000260643, gnorm=1.46, train_wall=35, gb_free=9.4, wall=12934
2021-03-27 23:37:02 | INFO | train_inner | epoch 063:   444 / 588 loss=4.686, nll_loss=3.794, word_ins=5.162, word_crf=1.83, length=2.751, ppl=25.74, wps=18948.6, ups=2.78, wpb=6815.2, bsz=236.8, num_updates=36900, lr=0.00026029, gnorm=1.697, train_wall=36, gb_free=9.3, wall=12970
2021-03-27 23:37:36 | INFO | train_inner | epoch 063:   544 / 588 loss=4.423, nll_loss=3.489, word_ins=4.898, word_crf=1.668, length=3.055, ppl=21.45, wps=20408.2, ups=2.89, wpb=7056.3, bsz=284.6, num_updates=37000, lr=0.000259938, gnorm=1.52, train_wall=34, gb_free=9.5, wall=13004
2021-03-27 23:37:52 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 23:37:55 | INFO | valid | epoch 063 | valid on 'valid' subset | loss 7.782 | nll_loss 5.598 | word_ins 6.852 | word_crf 3.846 | length 5.093 | ppl 220.05 | wps 59645.8 | wpb 5311.6 | bsz 208.1 | num_updates 37044 | best_loss 6.867
2021-03-27 23:37:55 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 63 @ 37044 updates
2021-03-27 23:37:55 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint63.pt
2021-03-27 23:37:56 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint63.pt
2021-03-27 23:37:59 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint63.pt (epoch 63 @ 37044 updates, score 7.782) (writing took 3.643135030000849 seconds)
2021-03-27 23:37:59 | INFO | fairseq_cli.train | end of epoch 63 (average epoch stats below)
2021-03-27 23:37:59 | INFO | train | epoch 063 | loss 4.455 | nll_loss 3.549 | word_ins 4.951 | word_crf 1.69 | length 2.895 | ppl 21.93 | wps 19293.9 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 37044 | lr 0.000259783 | gnorm 1.616 | train_wall 204 | gb_free 9.3 | wall 13027
2021-03-27 23:37:59 | INFO | fairseq.trainer | begin training epoch 64
2021-03-27 23:37:59 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 23:38:18 | INFO | train_inner | epoch 064:   56 / 588 loss=4.342, nll_loss=3.42, word_ins=4.835, word_crf=1.635, length=2.901, ppl=20.28, wps=16549.6, ups=2.38, wpb=6939.8, bsz=310, num_updates=37100, lr=0.000259587, gnorm=1.89, train_wall=35, gb_free=9.4, wall=13046
2021-03-27 23:38:54 | INFO | train_inner | epoch 064:   156 / 588 loss=4.397, nll_loss=3.502, word_ins=4.912, word_crf=1.657, length=2.841, ppl=21.06, wps=19925.7, ups=2.83, wpb=7029.1, bsz=273.4, num_updates=37200, lr=0.000259238, gnorm=1.692, train_wall=35, gb_free=9.6, wall=13082
2021-03-27 23:39:28 | INFO | train_inner | epoch 064:   256 / 588 loss=4.43, nll_loss=3.526, word_ins=4.932, word_crf=1.662, length=3.016, ppl=21.55, wps=19918.2, ups=2.87, wpb=6944, bsz=259, num_updates=37300, lr=0.00025889, gnorm=1.586, train_wall=35, gb_free=9.2, wall=13116
2021-03-27 23:40:04 | INFO | train_inner | epoch 064:   356 / 588 loss=4.548, nll_loss=3.638, word_ins=5.029, word_crf=1.747, length=2.86, ppl=23.39, wps=19972.1, ups=2.85, wpb=7018.2, bsz=263.7, num_updates=37400, lr=0.000258544, gnorm=1.879, train_wall=35, gb_free=9.2, wall=13152
2021-03-27 23:40:39 | INFO | train_inner | epoch 064:   456 / 588 loss=4.424, nll_loss=3.524, word_ins=4.929, word_crf=1.678, length=2.81, ppl=21.46, wps=19992.4, ups=2.86, wpb=6982.7, bsz=285.4, num_updates=37500, lr=0.000258199, gnorm=1.554, train_wall=35, gb_free=9.3, wall=13186
2021-03-27 23:41:14 | INFO | train_inner | epoch 064:   556 / 588 loss=4.53, nll_loss=3.626, word_ins=5.02, word_crf=1.728, length=2.916, ppl=23.1, wps=19850.8, ups=2.85, wpb=6968.9, bsz=253.2, num_updates=37600, lr=0.000257855, gnorm=1.491, train_wall=35, gb_free=9.5, wall=13222
2021-03-27 23:41:25 | INFO | fairseq_cli.train | begin validation on "valid" subset

2021-03-27 23:41:28 | INFO | valid | epoch 064 | valid on 'valid' subset | loss 7.755 | nll_loss 5.586 | word_ins 6.846 | word_crf 3.839 | length 4.926 | ppl 215.96 | wps 59757.6 | wpb 5311.6 | bsz 208.1 | num_updates 37632 | best_loss 6.867
2021-03-27 23:41:28 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 64 @ 37632 updates
2021-03-27 23:41:28 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint64.pt
2021-03-27 23:41:29 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint64.pt
2021-03-27 23:41:32 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint64.pt (epoch 64 @ 37632 updates, score 7.755) (writing took 3.6329897809991962 seconds)
2021-03-27 23:41:32 | INFO | fairseq_cli.train | end of epoch 64 (average epoch stats below)
2021-03-27 23:41:32 | INFO | train | epoch 064 | loss 4.44 | nll_loss 3.535 | word_ins 4.939 | word_crf 1.68 | length 2.894 | ppl 21.7 | wps 19307.4 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 37632 | lr 0.000257746 | gnorm 1.67 | train_wall 204 | gb_free 9.8 | wall 13240
2021-03-27 23:41:32 | INFO | fairseq.trainer | begin training epoch 65
2021-03-27 23:41:32 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 23:41:56 | INFO | train_inner | epoch 065:   68 / 588 loss=4.463, nll_loss=3.554, word_ins=4.955, word_crf=1.694, length=2.916, ppl=22.06, wps=16771.7, ups=2.36, wpb=7098.6, bsz=273.6, num_updates=37700, lr=0.000257513, gnorm=1.646, train_wall=35, gb_free=9.4, wall=13264
2021-03-27 23:42:31 | INFO | train_inner | epoch 065:   168 / 588 loss=4.45, nll_loss=3.572, word_ins=4.973, word_crf=1.691, length=2.725, ppl=21.85, wps=19442.8, ups=2.82, wpb=6902.1, bsz=255, num_updates=37800, lr=0.000257172, gnorm=1.733, train_wall=35, gb_free=9.5, wall=13299
2021-03-27 23:43:07 | INFO | train_inner | epoch 065:   268 / 588 loss=4.298, nll_loss=3.391, word_ins=4.817, word_crf=1.603, length=2.861, ppl=19.67, wps=20185.4, ups=2.83, wpb=7133.6, bsz=299.6, num_updates=37900, lr=0.000256833, gnorm=1.396, train_wall=35, gb_free=9.5, wall=13335
2021-03-27 23:43:42 | INFO | train_inner | epoch 065:   368 / 588 loss=4.431, nll_loss=3.521, word_ins=4.931, word_crf=1.668, length=2.976, ppl=21.56, wps=19906.4, ups=2.86, wpb=6964.8, bsz=263.2, num_updates=38000, lr=0.000256495, gnorm=1.658, train_wall=35, gb_free=9.6, wall=13370
2021-03-27 23:44:17 | INFO | train_inner | epoch 065:   468 / 588 loss=4.34, nll_loss=3.427, word_ins=4.846, word_crf=1.626, length=2.906, ppl=20.25, wps=20097.6, ups=2.87, wpb=6993.4, bsz=285.4, num_updates=38100, lr=0.000256158, gnorm=1.604, train_wall=35, gb_free=9.6, wall=13405
2021-03-27 23:44:52 | INFO | train_inner | epoch 065:   568 / 588 loss=4.544, nll_loss=3.639, word_ins=5.029, word_crf=1.74, length=2.893, ppl=23.33, wps=19486.9, ups=2.83, wpb=6888.3, bsz=257.1, num_updates=38200, lr=0.000255822, gnorm=1.771, train_wall=35, gb_free=9.8, wall=13440
2021-03-27 23:44:59 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 23:45:02 | INFO | valid | epoch 065 | valid on 'valid' subset | loss 7.839 | nll_loss 5.628 | word_ins 6.889 | word_crf 3.898 | length 4.968 | ppl 228.97 | wps 59568.5 | wpb 5311.6 | bsz 208.1 | num_updates 38220 | best_loss 6.867
2021-03-27 23:45:02 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 65 @ 38220 updates
2021-03-27 23:45:02 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint65.pt
2021-03-27 23:45:03 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint65.pt
2021-03-27 23:45:06 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint65.pt (epoch 65 @ 38220 updates, score 7.839) (writing took 3.662112211999556 seconds)
2021-03-27 23:45:06 | INFO | fairseq_cli.train | end of epoch 65 (average epoch stats below)
2021-03-27 23:45:06 | INFO | train | epoch 065 | loss 4.418 | nll_loss 3.515 | word_ins 4.923 | word_crf 1.669 | length 2.875 | ppl 21.38 | wps 19190.1 | ups 2.75 | wpb 6988.7 | bsz 272.5 | num_updates 38220 | lr 0.000255755 | gnorm 1.643 | train_wall 205 | gb_free 9.8 | wall 13454
2021-03-27 23:45:06 | INFO | fairseq.trainer | begin training epoch 66
2021-03-27 23:45:06 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 23:45:34 | INFO | train_inner | epoch 066:   80 / 588 loss=4.157, nll_loss=3.261, word_ins=4.706, word_crf=1.522, length=2.819, ppl=17.84, wps=16701.8, ups=2.39, wpb=7002.2, bsz=303.1, num_updates=38300, lr=0.000255488, gnorm=1.508, train_wall=34, gb_free=9.9, wall=13482
2021-03-27 23:46:09 | INFO | train_inner | epoch 066:   180 / 588 loss=4.362, nll_loss=3.457, word_ins=4.876, word_crf=1.626, length=2.978, ppl=20.56, wps=19981.6, ups=2.86, wpb=6985.9, bsz=271.7, num_updates=38400, lr=0.000255155, gnorm=1.631, train_wall=35, gb_free=9.9, wall=13517
2021-03-27 23:46:44 | INFO | train_inner | epoch 066:   280 / 588 loss=4.474, nll_loss=3.588, word_ins=4.989, word_crf=1.705, length=2.749, ppl=22.23, wps=19914.9, ups=2.85, wpb=6981.9, bsz=258.5, num_updates=38500, lr=0.000254824, gnorm=1.697, train_wall=35, gb_free=10, wall=13552
2021-03-27 23:47:19 | INFO | train_inner | epoch 066:   380 / 588 loss=4.339, nll_loss=3.448, word_ins=4.867, word_crf=1.621, length=2.845, ppl=20.24, wps=19946.5, ups=2.88, wpb=6929.7, bsz=274.8, num_updates=38600, lr=0.000254493, gnorm=1.524, train_wall=34, gb_free=9.4, wall=13587
2021-03-27 23:47:54 | INFO | train_inner | epoch 066:   480 / 588 loss=4.544, nll_loss=3.661, word_ins=5.051, word_crf=1.74, length=2.781, ppl=23.33, wps=19413.4, ups=2.84, wpb=6826.4, bsz=242.2, num_updates=38700, lr=0.000254164, gnorm=1.743, train_wall=35, gb_free=10, wall=13622
2021-03-27 23:48:29 | INFO | train_inner | epoch 066:   580 / 588 loss=4.375, nll_loss=3.447, word_ins=4.865, word_crf=1.648, length=2.943, ppl=20.75, wps=20381.4, ups=2.86, wpb=7137, bsz=285.4, num_updates=38800, lr=0.000253837, gnorm=1.526, train_wall=35, gb_free=9.9, wall=13657
2021-03-27 23:48:32 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 23:48:35 | INFO | valid | epoch 066 | valid on 'valid' subset | loss 7.884 | nll_loss 5.652 | word_ins 6.909 | word_crf 3.928 | length 5.006 | ppl 236.18 | wps 59946.4 | wpb 5311.6 | bsz 208.1 | num_updates 38808 | best_loss 6.867
2021-03-27 23:48:35 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 66 @ 38808 updates
2021-03-27 23:48:35 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint66.pt
2021-03-27 23:48:36 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint66.pt
2021-03-27 23:48:39 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint66.pt (epoch 66 @ 38808 updates, score 7.884) (writing took 3.6473776590009948 seconds)
2021-03-27 23:48:39 | INFO | fairseq_cli.train | end of epoch 66 (average epoch stats below)
2021-03-27 23:48:39 | INFO | train | epoch 066 | loss 4.381 | nll_loss 3.482 | word_ins 4.896 | word_crf 1.647 | length 2.856 | ppl 20.83 | wps 19316.3 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 38808 | lr 0.00025381 | gnorm 1.605 | train_wall 204 | gb_free 9.5 | wall 13666
2021-03-27 23:48:39 | INFO | fairseq.trainer | begin training epoch 67
2021-03-27 23:48:39 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 23:49:11 | INFO | train_inner | epoch 067:   92 / 588 loss=4.454, nll_loss=3.575, word_ins=4.977, word_crf=1.692, length=2.725, ppl=21.91, wps=16524.7, ups=2.36, wpb=7013.6, bsz=254.9, num_updates=38900, lr=0.00025351, gnorm=1.769, train_wall=35, gb_free=9.4, wall=13699
2021-03-27 23:49:46 | INFO | train_inner | epoch 067:   192 / 588 loss=4.419, nll_loss=3.516, word_ins=4.927, word_crf=1.66, length=2.954, ppl=21.39, wps=20123.7, ups=2.86, wpb=7035.9, bsz=256.6, num_updates=39000, lr=0.000253185, gnorm=1.731, train_wall=35, gb_free=10, wall=13734
2021-03-27 23:50:21 | INFO | train_inner | epoch 067:   292 / 588 loss=4.261, nll_loss=3.361, word_ins=4.795, word_crf=1.584, length=2.795, ppl=19.17, wps=20085.1, ups=2.87, wpb=7008.9, bsz=292.3, num_updates=39100, lr=0.000252861, gnorm=1.487, train_wall=35, gb_free=9.9, wall=13769
2021-03-27 23:50:56 | INFO | train_inner | epoch 067:   392 / 588 loss=4.285, nll_loss=3.394, word_ins=4.822, word_crf=1.589, length=2.849, ppl=19.49, wps=19947.6, ups=2.88, wpb=6931.8, bsz=274.7, num_updates=39200, lr=0.000252538, gnorm=1.578, train_wall=35, gb_free=9.4, wall=13804
2021-03-27 23:51:31 | INFO | train_inner | epoch 067:   492 / 588 loss=4.375, nll_loss=3.458, word_ins=4.874, word_crf=1.649, length=2.892, ppl=20.75, wps=20062.4, ups=2.85, wpb=7034.3, bsz=288.9, num_updates=39300, lr=0.000252217, gnorm=1.58, train_wall=35, gb_free=9.9, wall=13839
2021-03-27 23:52:04 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 23:52:08 | INFO | valid | epoch 067 | valid on 'valid' subset | loss 7.88 | nll_loss 5.643 | word_ins 6.905 | word_crf 3.93 | length 4.968 | ppl 235.53 | wps 60035.5 | wpb 5311.6 | bsz 208.1 | num_updates 39396 | best_loss 6.867
2021-03-27 23:52:08 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 67 @ 39396 updates
2021-03-27 23:52:08 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint67.pt
2021-03-27 23:52:08 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint67.pt
2021-03-27 23:52:11 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint67.pt (epoch 67 @ 39396 updates, score 7.88) (writing took 3.6450610809988575 seconds)
2021-03-27 23:52:11 | INFO | fairseq_cli.train | end of epoch 67 (average epoch stats below)
2021-03-27 23:52:11 | INFO | train | epoch 067 | loss 4.356 | nll_loss 3.458 | word_ins 4.877 | word_crf 1.633 | length 2.847 | ppl 20.47 | wps 19324.5 | ups 2.77 | wpb 6988.7 | bsz 272.5 | num_updates 39396 | lr 0.000251909 | gnorm 1.6 | train_wall 204 | gb_free 9.5 | wall 13879
2021-03-27 23:52:11 | INFO | fairseq.trainer | begin training epoch 68
2021-03-27 23:52:11 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 23:52:13 | INFO | train_inner | epoch 068:   4 / 588 loss=4.325, nll_loss=3.426, word_ins=4.849, word_crf=1.611, length=2.89, ppl=20.04, wps=16596, ups=2.39, wpb=6934, bsz=271.2, num_updates=39400, lr=0.000251896, gnorm=1.423, train_wall=34, gb_free=9.5, wall=13881
2021-03-27 23:52:48 | INFO | train_inner | epoch 068:   104 / 588 loss=4.314, nll_loss=3.443, word_ins=4.863, word_crf=1.61, length=2.725, ppl=19.89, wps=19523.6, ups=2.84, wpb=6880.3, bsz=272.8, num_updates=39500, lr=0.000251577, gnorm=1.613, train_wall=35, gb_free=9.8, wall=13916
2021-03-27 23:53:23 | INFO | train_inner | epoch 068:   204 / 588 loss=4.179, nll_loss=3.27, word_ins=4.717, word_crf=1.533, length=2.875, ppl=18.11, wps=20451.2, ups=2.88, wpb=7100.6, bsz=300.6, num_updates=39600, lr=0.000251259, gnorm=1.616, train_wall=34, gb_free=9.2, wall=13951

2021-03-27 23:53:58 | INFO | train_inner | epoch 068:   304 / 588 loss=4.248, nll_loss=3.338, word_ins=4.774, word_crf=1.569, length=2.918, ppl=19, wps=20135.2, ups=2.85, wpb=7058.1, bsz=287.6, num_updates=39700, lr=0.000250943, gnorm=1.418, train_wall=35, gb_free=9.2, wall=13986
2021-03-27 23:54:33 | INFO | train_inner | epoch 068:   404 / 588 loss=4.496, nll_loss=3.629, word_ins=5.026, word_crf=1.723, length=2.603, ppl=22.57, wps=19420, ups=2.82, wpb=6898.4, bsz=247.6, num_updates=39800, lr=0.000250627, gnorm=1.726, train_wall=35, gb_free=9.6, wall=14021
2021-03-27 23:55:08 | INFO | train_inner | epoch 068:   504 / 588 loss=4.359, nll_loss=3.459, word_ins=4.88, word_crf=1.63, length=2.89, ppl=20.52, wps=20084.9, ups=2.87, wpb=6999.8, bsz=257.1, num_updates=39900, lr=0.000250313, gnorm=1.557, train_wall=35, gb_free=9.5, wall=14056
2021-03-27 23:55:38 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 23:55:41 | INFO | valid | epoch 068 | valid on 'valid' subset | loss 7.93 | nll_loss 5.664 | word_ins 6.928 | word_crf 3.964 | length 5.015 | ppl 243.83 | wps 59919.4 | wpb 5311.6 | bsz 208.1 | num_updates 39984 | best_loss 6.867
2021-03-27 23:55:41 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 68 @ 39984 updates
2021-03-27 23:55:41 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint68.pt
2021-03-27 23:55:42 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint68.pt
2021-03-27 23:55:45 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint68.pt (epoch 68 @ 39984 updates, score 7.93) (writing took 3.642890843000714 seconds)
2021-03-27 23:55:45 | INFO | fairseq_cli.train | end of epoch 68 (average epoch stats below)
2021-03-27 23:55:45 | INFO | train | epoch 068 | loss 4.325 | nll_loss 3.431 | word_ins 4.855 | word_crf 1.616 | length 2.823 | ppl 20.05 | wps 19256.4 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 39984 | lr 0.00025005 | gnorm 1.569 | train_wall 205 | gb_free 10 | wall 14092
2021-03-27 23:55:45 | INFO | fairseq.trainer | begin training epoch 69
2021-03-27 23:55:45 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 23:55:50 | INFO | train_inner | epoch 069:   16 / 588 loss=4.319, nll_loss=3.414, word_ins=4.839, word_crf=1.611, length=2.877, ppl=19.96, wps=16654.4, ups=2.38, wpb=7011.8, bsz=272.6, num_updates=40000, lr=0.00025, gnorm=1.472, train_wall=35, gb_free=9.2, wall=14098
2021-03-27 23:55:50 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 23:55:54 | INFO | valid | epoch 069 | valid on 'valid' subset | loss 8.019 | nll_loss 5.703 | word_ins 6.968 | word_crf 4.013 | length 5.216 | ppl 259.34 | wps 59692 | wpb 5311.6 | bsz 208.1 | num_updates 40000 | best_loss 6.867
2021-03-27 23:55:54 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 69 @ 40000 updates
2021-03-27 23:55:54 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint_69_40000.pt
2021-03-27 23:55:55 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint_69_40000.pt
2021-03-27 23:55:57 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint_69_40000.pt (epoch 69 @ 40000 updates, score 8.019) (writing took 3.963417537999703 seconds)
2021-03-27 23:56:33 | INFO | train_inner | epoch 069:   116 / 588 loss=4.304, nll_loss=3.436, word_ins=4.859, word_crf=1.606, length=2.68, ppl=19.75, wps=16396.9, ups=2.36, wpb=6961.8, bsz=273.1, num_updates=40100, lr=0.000249688, gnorm=1.696, train_wall=35, gb_free=9.5, wall=14141
2021-03-27 23:57:07 | INFO | train_inner | epoch 069:   216 / 588 loss=4.204, nll_loss=3.32, word_ins=4.763, word_crf=1.543, length=2.794, ppl=18.43, wps=20198.6, ups=2.89, wpb=6987.8, bsz=270.1, num_updates=40200, lr=0.000249377, gnorm=1.608, train_wall=34, gb_free=9.4, wall=14175
2021-03-27 23:57:42 | INFO | train_inner | epoch 069:   316 / 588 loss=4.349, nll_loss=3.458, word_ins=4.879, word_crf=1.627, length=2.824, ppl=20.37, wps=19698.9, ups=2.87, wpb=6868, bsz=261.8, num_updates=40300, lr=0.000249068, gnorm=1.719, train_wall=35, gb_free=10.8, wall=14210
2021-03-27 23:58:17 | INFO | train_inner | epoch 069:   416 / 588 loss=4.326, nll_loss=3.42, word_ins=4.846, word_crf=1.614, length=2.896, ppl=20.06, wps=19887.3, ups=2.86, wpb=6958.1, bsz=272.3, num_updates=40400, lr=0.000248759, gnorm=1.575, train_wall=35, gb_free=9.7, wall=14245
2021-03-27 23:58:53 | INFO | train_inner | epoch 069:   516 / 588 loss=4.399, nll_loss=3.5, word_ins=4.915, word_crf=1.664, length=2.769, ppl=21.09, wps=19879, ups=2.83, wpb=7034, bsz=264.6, num_updates=40500, lr=0.000248452, gnorm=1.672, train_wall=35, gb_free=9.5, wall=14280
2021-03-27 23:59:18 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-27 23:59:21 | INFO | valid | epoch 069 | valid on 'valid' subset | loss 8.002 | nll_loss 5.706 | word_ins 6.965 | word_crf 4.007 | length 5.118 | ppl 256.27 | wps 59380.2 | wpb 5311.6 | bsz 208.1 | num_updates 40572 | best_loss 6.867
2021-03-27 23:59:21 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 69 @ 40572 updates
2021-03-27 23:59:21 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint69.pt
2021-03-27 23:59:22 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint69.pt
2021-03-27 23:59:25 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint69.pt (epoch 69 @ 40572 updates, score 8.002) (writing took 3.6454859229997965 seconds)
2021-03-27 23:59:25 | INFO | fairseq_cli.train | end of epoch 69 (average epoch stats below)
2021-03-27 23:59:25 | INFO | train | epoch 069 | loss 4.308 | nll_loss 3.415 | word_ins 4.842 | word_crf 1.606 | length 2.803 | ppl 19.8 | wps 18671.7 | ups 2.67 | wpb 6988.7 | bsz 272.5 | num_updates 40572 | lr 0.000248231 | gnorm 1.621 | train_wall 204 | gb_free 9.3 | wall 14313
2021-03-27 23:59:25 | INFO | fairseq.trainer | begin training epoch 70
2021-03-27 23:59:25 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-27 23:59:35 | INFO | train_inner | epoch 070:   28 / 588 loss=4.341, nll_loss=3.433, word_ins=4.856, word_crf=1.619, length=2.94, ppl=20.27, wps=16778.3, ups=2.37, wpb=7074.4, bsz=276, num_updates=40600, lr=0.000248146, gnorm=1.505, train_wall=35, gb_free=10.1, wall=14323
2021-03-28 00:00:10 | INFO | train_inner | epoch 070:   128 / 588 loss=4.202, nll_loss=3.319, word_ins=4.762, word_crf=1.536, length=2.853, ppl=18.4, wps=20174.7, ups=2.87, wpb=7039.1, bsz=267.7, num_updates=40700, lr=0.000247841, gnorm=1.504, train_wall=35, gb_free=9.8, wall=14357
2021-03-28 00:00:45 | INFO | train_inner | epoch 070:   228 / 588 loss=4.368, nll_loss=3.485, word_ins=4.903, word_crf=1.647, length=2.694, ppl=20.65, wps=19657.1, ups=2.83, wpb=6944.8, bsz=261, num_updates=40800, lr=0.000247537, gnorm=1.866, train_wall=35, gb_free=9.4, wall=14393
2021-03-28 00:01:20 | INFO | train_inner | epoch 070:   328 / 588 loss=4.176, nll_loss=3.289, word_ins=4.736, word_crf=1.522, length=2.862, ppl=18.08, wps=19941.9, ups=2.89, wpb=6911.8, bsz=280.2, num_updates=40900, lr=0.000247234, gnorm=1.482, train_wall=34, gb_free=9.7, wall=14427
2021-03-28 00:01:54 | INFO | train_inner | epoch 070:   428 / 588 loss=4.244, nll_loss=3.346, word_ins=4.783, word_crf=1.572, length=2.804, ppl=18.95, wps=20518.8, ups=2.88, wpb=7123.4, bsz=284.2, num_updates=41000, lr=0.000246932, gnorm=1.666, train_wall=34, gb_free=9.4, wall=14462
2021-03-28 00:02:30 | INFO | train_inner | epoch 070:   528 / 588 loss=4.304, nll_loss=3.413, word_ins=4.838, word_crf=1.62, length=2.656, ppl=19.76, wps=19533.2, ups=2.83, wpb=6913.7, bsz=288.9, num_updates=41100, lr=0.000246632, gnorm=1.635, train_wall=35, gb_free=9.5, wall=14498
2021-03-28 00:02:51 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 00:02:54 | INFO | valid | epoch 070 | valid on 'valid' subset | loss 8.017 | nll_loss 5.724 | word_ins 6.981 | word_crf 4.023 | length 5.03 | ppl 258.97 | wps 59386.7 | wpb 5311.6 | bsz 208.1 | num_updates 41160 | best_loss 6.867
2021-03-28 00:02:54 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 70 @ 41160 updates
2021-03-28 00:02:54 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint70.pt
2021-03-28 00:02:55 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint70.pt
2021-03-28 00:02:58 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint70.pt (epoch 70 @ 41160 updates, score 8.017) (writing took 3.665831997999703 seconds)
2021-03-28 00:02:58 | INFO | fairseq_cli.train | end of epoch 70 (average epoch stats below)
2021-03-28 00:02:58 | INFO | train | epoch 070 | loss 4.282 | nll_loss 3.391 | word_ins 4.823 | word_crf 1.591 | length 2.796 | ppl 19.45 | wps 19274.3 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 41160 | lr 0.000246452 | gnorm 1.626 | train_wall 204 | gb_free 9.9 | wall 14526
2021-03-28 00:02:58 | INFO | fairseq.trainer | begin training epoch 71
2021-03-28 00:02:58 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 00:03:12 | INFO | train_inner | epoch 071:   40 / 588 loss=4.371, nll_loss=3.475, word_ins=4.895, word_crf=1.635, length=2.883, ppl=20.7, wps=16697.1, ups=2.36, wpb=7072.3, bsz=257.8, num_updates=41200, lr=0.000246332, gnorm=1.623, train_wall=35, gb_free=9.7, wall=14540
2021-03-28 00:03:48 | INFO | train_inner | epoch 071:   140 / 588 loss=4.271, nll_loss=3.414, word_ins=4.844, word_crf=1.59, length=2.589, ppl=19.3, wps=19299.7, ups=2.81, wpb=6871.4, bsz=261.3, num_updates=41300, lr=0.000246034, gnorm=1.544, train_wall=35, gb_free=9.5, wall=14576
2021-03-28 00:04:22 | INFO | train_inner | epoch 071:   240 / 588 loss=4.252, nll_loss=3.351, word_ins=4.792, word_crf=1.562, length=2.941, ppl=19.05, wps=20174.8, ups=2.87, wpb=7030.1, bsz=265.7, num_updates=41400, lr=0.000245737, gnorm=1.699, train_wall=35, gb_free=9.7, wall=14610
2021-03-28 00:04:57 | INFO | train_inner | epoch 071:   340 / 588 loss=4.191, nll_loss=3.301, word_ins=4.747, word_crf=1.544, length=2.73, ppl=18.26, wps=20217.1, ups=2.88, wpb=7029, bsz=287, num_updates=41500, lr=0.00024544, gnorm=1.612, train_wall=34, gb_free=9.5, wall=14645
2021-03-28 00:05:32 | INFO | train_inner | epoch 071:   440 / 588 loss=4.38, nll_loss=3.485, word_ins=4.904, word_crf=1.652, length=2.759, ppl=20.82, wps=19837.7, ups=2.85, wpb=6951.4, bsz=259.6, num_updates=41600, lr=0.000245145, gnorm=1.661, train_wall=35, gb_free=9.3, wall=14680
2021-03-28 00:06:07 | INFO | train_inner | epoch 071:   540 / 588 loss=4.243, nll_loss=3.341, word_ins=4.776, word_crf=1.57, length=2.853, ppl=18.93, wps=20009.8, ups=2.86, wpb=7002.4, bsz=288.5, num_updates=41700, lr=0.000244851, gnorm=1.556, train_wall=35, gb_free=9.1, wall=14715
2021-03-28 00:06:24 | INFO | fairseq_cli.train | begin validation on "valid" subset

2021-03-28 00:06:27 | INFO | valid | epoch 071 | valid on 'valid' subset | loss 7.983 | nll_loss 5.696 | word_ins 6.955 | word_crf 3.99 | length 5.151 | ppl 252.95 | wps 59806.9 | wpb 5311.6 | bsz 208.1 | num_updates 41748 | best_loss 6.867
2021-03-28 00:06:27 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 71 @ 41748 updates
2021-03-28 00:06:27 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint71.pt
2021-03-28 00:06:28 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint71.pt
2021-03-28 00:06:31 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint71.pt (epoch 71 @ 41748 updates, score 7.983) (writing took 3.6495065740000427 seconds)
2021-03-28 00:06:31 | INFO | fairseq_cli.train | end of epoch 71 (average epoch stats below)
2021-03-28 00:06:31 | INFO | train | epoch 071 | loss 4.26 | nll_loss 3.372 | word_ins 4.807 | word_crf 1.578 | length 2.782 | ppl 19.15 | wps 19323.2 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 41748 | lr 0.00024471 | gnorm 1.617 | train_wall 204 | gb_free 9.5 | wall 14738
2021-03-28 00:06:31 | INFO | fairseq.trainer | begin training epoch 72
2021-03-28 00:06:31 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 00:06:49 | INFO | train_inner | epoch 072:   52 / 588 loss=4.127, nll_loss=3.239, word_ins=4.691, word_crf=1.503, length=2.777, ppl=17.47, wps=16769.5, ups=2.4, wpb=7000.9, bsz=294.7, num_updates=41800, lr=0.000244558, gnorm=1.699, train_wall=34, gb_free=9.4, wall=14757
2021-03-28 00:07:24 | INFO | train_inner | epoch 072:   152 / 588 loss=4.114, nll_loss=3.236, word_ins=4.691, word_crf=1.493, length=2.757, ppl=17.32, wps=19897.6, ups=2.88, wpb=6915.6, bsz=282.7, num_updates=41900, lr=0.000244266, gnorm=1.763, train_wall=34, gb_free=9.3, wall=14792
2021-03-28 00:07:59 | INFO | train_inner | epoch 072:   252 / 588 loss=4.297, nll_loss=3.43, word_ins=4.858, word_crf=1.599, length=2.689, ppl=19.65, wps=19724.8, ups=2.84, wpb=6953.1, bsz=262.1, num_updates=42000, lr=0.000243975, gnorm=1.601, train_wall=35, gb_free=9.7, wall=14827
2021-03-28 00:08:34 | INFO | train_inner | epoch 072:   352 / 588 loss=4.255, nll_loss=3.369, word_ins=4.806, word_crf=1.573, length=2.788, ppl=19.1, wps=20061.7, ups=2.85, wpb=7029.3, bsz=269.7, num_updates=42100, lr=0.000243685, gnorm=1.541, train_wall=35, gb_free=9.8, wall=14862
2021-03-28 00:09:09 | INFO | train_inner | epoch 072:   452 / 588 loss=4.23, nll_loss=3.325, word_ins=4.767, word_crf=1.569, length=2.779, ppl=18.77, wps=20106.8, ups=2.84, wpb=7070.9, bsz=281.7, num_updates=42200, lr=0.000243396, gnorm=1.598, train_wall=35, gb_free=9.3, wall=14897
2021-03-28 00:09:45 | INFO | train_inner | epoch 072:   552 / 588 loss=4.362, nll_loss=3.474, word_ins=4.895, word_crf=1.634, length=2.806, ppl=20.56, wps=19651.7, ups=2.83, wpb=6938, bsz=251.3, num_updates=42300, lr=0.000243108, gnorm=1.536, train_wall=35, gb_free=9.7, wall=14932
2021-03-28 00:09:57 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 00:10:00 | INFO | valid | epoch 072 | valid on 'valid' subset | loss 8.02 | nll_loss 5.722 | word_ins 6.984 | word_crf 4.028 | length 5.005 | ppl 259.59 | wps 59811.8 | wpb 5311.6 | bsz 208.1 | num_updates 42336 | best_loss 6.867
2021-03-28 00:10:00 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 72 @ 42336 updates
2021-03-28 00:10:00 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint72.pt
2021-03-28 00:10:01 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint72.pt
2021-03-28 00:10:04 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint72.pt (epoch 72 @ 42336 updates, score 8.02) (writing took 3.658453907999501 seconds)
2021-03-28 00:10:04 | INFO | fairseq_cli.train | end of epoch 72 (average epoch stats below)
2021-03-28 00:10:04 | INFO | train | epoch 072 | loss 4.237 | nll_loss 3.351 | word_ins 4.79 | word_crf 1.565 | length 2.777 | ppl 18.86 | wps 19244.7 | ups 2.75 | wpb 6988.7 | bsz 272.5 | num_updates 42336 | lr 0.000243005 | gnorm 1.617 | train_wall 205 | gb_free 10.6 | wall 14952
2021-03-28 00:10:04 | INFO | fairseq.trainer | begin training epoch 73
2021-03-28 00:10:04 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 00:10:27 | INFO | train_inner | epoch 073:   64 / 588 loss=4.163, nll_loss=3.274, word_ins=4.724, word_crf=1.514, length=2.871, ppl=17.92, wps=17020.4, ups=2.38, wpb=7144.6, bsz=271.1, num_updates=42400, lr=0.000242821, gnorm=1.605, train_wall=35, gb_free=9.6, wall=14974
2021-03-28 00:11:01 | INFO | train_inner | epoch 073:   164 / 588 loss=4.195, nll_loss=3.308, word_ins=4.756, word_crf=1.528, length=2.891, ppl=18.31, wps=19934.4, ups=2.86, wpb=6966.7, bsz=260.9, num_updates=42500, lr=0.000242536, gnorm=1.812, train_wall=35, gb_free=9.9, wall=15009
2021-03-28 00:11:37 | INFO | train_inner | epoch 073:   264 / 588 loss=4.176, nll_loss=3.295, word_ins=4.744, word_crf=1.529, length=2.751, ppl=18.08, wps=20019.2, ups=2.84, wpb=7040, bsz=272.2, num_updates=42600, lr=0.000242251, gnorm=1.674, train_wall=35, gb_free=9.6, wall=15045
2021-03-28 00:12:12 | INFO | train_inner | epoch 073:   364 / 588 loss=4.181, nll_loss=3.293, word_ins=4.741, word_crf=1.539, length=2.722, ppl=18.14, wps=19709.4, ups=2.84, wpb=6929.5, bsz=284.5, num_updates=42700, lr=0.000241967, gnorm=1.582, train_wall=35, gb_free=9.5, wall=15080
2021-03-28 00:12:47 | INFO | train_inner | epoch 073:   464 / 588 loss=4.264, nll_loss=3.364, word_ins=4.798, word_crf=1.577, length=2.873, ppl=19.21, wps=20133.4, ups=2.86, wpb=7037.9, bsz=275.8, num_updates=42800, lr=0.000241684, gnorm=1.648, train_wall=35, gb_free=10.1, wall=15115
2021-03-28 00:13:22 | INFO | train_inner | epoch 073:   564 / 588 loss=4.304, nll_loss=3.422, word_ins=4.848, word_crf=1.621, length=2.588, ppl=19.75, wps=19552, ups=2.84, wpb=6896.5, bsz=280.9, num_updates=42900, lr=0.000241402, gnorm=1.675, train_wall=35, gb_free=10.2, wall=15150
2021-03-28 00:13:31 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 00:13:34 | INFO | valid | epoch 073 | valid on 'valid' subset | loss 8.054 | nll_loss 5.731 | word_ins 6.997 | word_crf 4.044 | length 5.122 | ppl 265.83 | wps 59861.8 | wpb 5311.6 | bsz 208.1 | num_updates 42924 | best_loss 6.867
2021-03-28 00:13:34 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 73 @ 42924 updates
2021-03-28 00:13:34 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint73.pt
2021-03-28 00:13:35 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint73.pt
2021-03-28 00:13:38 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint73.pt (epoch 73 @ 42924 updates, score 8.054) (writing took 3.6498728030001075 seconds)
2021-03-28 00:13:38 | INFO | fairseq_cli.train | end of epoch 73 (average epoch stats below)
2021-03-28 00:13:38 | INFO | train | epoch 073 | loss 4.226 | nll_loss 3.34 | word_ins 4.781 | word_crf 1.559 | length 2.762 | ppl 18.71 | wps 19251 | ups 2.75 | wpb 6988.7 | bsz 272.5 | num_updates 42924 | lr 0.000241335 | gnorm 1.687 | train_wall 205 | gb_free 9.5 | wall 15165
2021-03-28 00:13:38 | INFO | fairseq.trainer | begin training epoch 74
2021-03-28 00:13:38 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 00:14:04 | INFO | train_inner | epoch 074:   76 / 588 loss=4.268, nll_loss=3.393, word_ins=4.826, word_crf=1.581, length=2.738, ppl=19.27, wps=16718.2, ups=2.35, wpb=7099.2, bsz=265.3, num_updates=43000, lr=0.000241121, gnorm=1.685, train_wall=35, gb_free=10, wall=15192
2021-03-28 00:14:39 | INFO | train_inner | epoch 074:   176 / 588 loss=4.06, nll_loss=3.204, word_ins=4.666, word_crf=1.474, length=2.522, ppl=16.68, wps=19685.2, ups=2.88, wpb=6834.2, bsz=282.8, num_updates=43100, lr=0.000240842, gnorm=1.736, train_wall=34, gb_free=9.3, wall=15227
2021-03-28 00:15:14 | INFO | train_inner | epoch 074:   276 / 588 loss=4.157, nll_loss=3.253, word_ins=4.707, word_crf=1.516, length=2.87, ppl=17.84, wps=20533.4, ups=2.89, wpb=7108.2, bsz=284.3, num_updates=43200, lr=0.000240563, gnorm=1.614, train_wall=34, gb_free=9.5, wall=15262
2021-03-28 00:15:49 | INFO | train_inner | epoch 074:   376 / 588 loss=4.245, nll_loss=3.378, word_ins=4.813, word_crf=1.58, length=2.588, ppl=18.97, wps=20056.1, ups=2.82, wpb=7101.3, bsz=271, num_updates=43300, lr=0.000240285, gnorm=1.58, train_wall=35, gb_free=9.7, wall=15297
2021-03-28 00:16:24 | INFO | train_inner | epoch 074:   476 / 588 loss=4.212, nll_loss=3.302, word_ins=4.748, word_crf=1.546, length=2.914, ppl=18.53, wps=19901.1, ups=2.87, wpb=6932.9, bsz=276.5, num_updates=43400, lr=0.000240008, gnorm=1.58, train_wall=35, gb_free=9.2, wall=15332
2021-03-28 00:16:59 | INFO | train_inner | epoch 074:   576 / 588 loss=4.439, nll_loss=3.561, word_ins=4.971, word_crf=1.677, length=2.764, ppl=21.69, wps=19327.8, ups=2.83, wpb=6827.9, bsz=238, num_updates=43500, lr=0.000239732, gnorm=1.797, train_wall=35, gb_free=9.2, wall=15367
2021-03-28 00:17:03 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 00:17:07 | INFO | valid | epoch 074 | valid on 'valid' subset | loss 7.91 | nll_loss 5.652 | word_ins 6.911 | word_crf 3.937 | length 5.17 | ppl 240.49 | wps 59525.7 | wpb 5311.6 | bsz 208.1 | num_updates 43512 | best_loss 6.867
2021-03-28 00:17:07 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 74 @ 43512 updates
2021-03-28 00:17:07 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint74.pt
2021-03-28 00:17:08 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint74.pt
2021-03-28 00:17:10 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint74.pt (epoch 74 @ 43512 updates, score 7.91) (writing took 3.659542056999271 seconds)
2021-03-28 00:17:10 | INFO | fairseq_cli.train | end of epoch 74 (average epoch stats below)
2021-03-28 00:17:10 | INFO | train | epoch 074 | loss 4.21 | nll_loss 3.326 | word_ins 4.77 | word_crf 1.551 | length 2.737 | ppl 18.5 | wps 19303.1 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 43512 | lr 0.000239699 | gnorm 1.663 | train_wall 204 | gb_free 9.6 | wall 15378
2021-03-28 00:17:11 | INFO | fairseq.trainer | begin training epoch 75
2021-03-28 00:17:11 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 00:17:42 | INFO | train_inner | epoch 075:   88 / 588 loss=4.023, nll_loss=3.13, word_ins=4.598, word_crf=1.456, length=2.681, ppl=16.26, wps=16695, ups=2.38, wpb=7027.5, bsz=321.8, num_updates=43600, lr=0.000239457, gnorm=1.88, train_wall=35, gb_free=10.3, wall=15409
2021-03-28 00:18:16 | INFO | train_inner | epoch 075:   188 / 588 loss=4.088, nll_loss=3.21, word_ins=4.671, word_crf=1.478, length=2.751, ppl=17.01, wps=20218.7, ups=2.86, wpb=7061.5, bsz=287.4, num_updates=43700, lr=0.000239182, gnorm=1.619, train_wall=35, gb_free=9.5, wall=15444

2021-03-28 00:18:52 | INFO | train_inner | epoch 075:   288 / 588 loss=4.168, nll_loss=3.308, word_ins=4.757, word_crf=1.531, length=2.591, ppl=17.98, wps=19634.8, ups=2.85, wpb=6890.9, bsz=260.7, num_updates=43800, lr=0.000238909, gnorm=1.691, train_wall=35, gb_free=9.6, wall=15479
2021-03-28 00:19:27 | INFO | train_inner | epoch 075:   388 / 588 loss=4.263, nll_loss=3.376, word_ins=4.813, word_crf=1.573, length=2.828, ppl=19.2, wps=20182.9, ups=2.86, wpb=7059.8, bsz=261.8, num_updates=43900, lr=0.000238637, gnorm=1.658, train_wall=35, gb_free=9.5, wall=15514
2021-03-28 00:20:02 | INFO | train_inner | epoch 075:   488 / 588 loss=4.222, nll_loss=3.348, word_ins=4.792, word_crf=1.557, length=2.686, ppl=18.66, wps=19162.5, ups=2.82, wpb=6798.3, bsz=250.9, num_updates=44000, lr=0.000238366, gnorm=1.624, train_wall=35, gb_free=9.4, wall=15550
2021-03-28 00:20:37 | INFO | train_inner | epoch 075:   588 / 588 loss=4.347, nll_loss=3.456, word_ins=4.88, word_crf=1.626, length=2.808, ppl=20.35, wps=20349.4, ups=2.87, wpb=7081, bsz=260.9, num_updates=44100, lr=0.000238095, gnorm=1.765, train_wall=35, gb_free=9.3, wall=15585
2021-03-28 00:20:37 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 00:20:40 | INFO | valid | epoch 075 | valid on 'valid' subset | loss 8.053 | nll_loss 5.722 | word_ins 6.985 | word_crf 4.049 | length 5.111 | ppl 265.53 | wps 60103.7 | wpb 5311.6 | bsz 208.1 | num_updates 44100 | best_loss 6.867
2021-03-28 00:20:40 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 75 @ 44100 updates
2021-03-28 00:20:40 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint75.pt
2021-03-28 00:20:41 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint75.pt
2021-03-28 00:20:44 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint75.pt (epoch 75 @ 44100 updates, score 8.053) (writing took 3.645926936000251 seconds)
2021-03-28 00:20:44 | INFO | fairseq_cli.train | end of epoch 75 (average epoch stats below)
2021-03-28 00:20:44 | INFO | train | epoch 075 | loss 4.192 | nll_loss 3.312 | word_ins 4.758 | word_crf 1.54 | length 2.723 | ppl 18.27 | wps 19266.6 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 44100 | lr 0.000238095 | gnorm 1.703 | train_wall 204 | gb_free 9.3 | wall 15592
2021-03-28 00:20:44 | INFO | fairseq.trainer | begin training epoch 76
2021-03-28 00:20:44 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 00:21:19 | INFO | train_inner | epoch 076:   100 / 588 loss=4.14, nll_loss=3.279, word_ins=4.729, word_crf=1.517, length=2.583, ppl=17.63, wps=16541.2, ups=2.35, wpb=7023.9, bsz=284.2, num_updates=44200, lr=0.000237826, gnorm=1.71, train_wall=35, gb_free=9.2, wall=15627
2021-03-28 00:21:54 | INFO | train_inner | epoch 076:   200 / 588 loss=4.019, nll_loss=3.145, word_ins=4.619, word_crf=1.44, length=2.694, ppl=16.22, wps=20033, ups=2.87, wpb=6980.5, bsz=278.6, num_updates=44300, lr=0.000237557, gnorm=1.533, train_wall=35, gb_free=9.7, wall=15662
2021-03-28 00:22:29 | INFO | train_inner | epoch 076:   300 / 588 loss=4.114, nll_loss=3.234, word_ins=4.691, word_crf=1.502, length=2.669, ppl=17.32, wps=20417.5, ups=2.86, wpb=7129.3, bsz=289.9, num_updates=44400, lr=0.000237289, gnorm=1.732, train_wall=35, gb_free=10.3, wall=15697
2021-03-28 00:23:05 | INFO | train_inner | epoch 076:   400 / 588 loss=4.3, nll_loss=3.428, word_ins=4.861, word_crf=1.602, length=2.683, ppl=19.7, wps=19467.5, ups=2.81, wpb=6924.8, bsz=246.7, num_updates=44500, lr=0.000237023, gnorm=1.613, train_wall=35, gb_free=9.5, wall=15732
2021-03-28 00:23:39 | INFO | train_inner | epoch 076:   500 / 588 loss=4.192, nll_loss=3.292, word_ins=4.741, word_crf=1.531, length=2.905, ppl=18.28, wps=20116.5, ups=2.87, wpb=7005.7, bsz=272.4, num_updates=44600, lr=0.000236757, gnorm=1.673, train_wall=35, gb_free=9.5, wall=15767
2021-03-28 00:24:10 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 00:24:13 | INFO | valid | epoch 076 | valid on 'valid' subset | loss 8.088 | nll_loss 5.738 | word_ins 7.007 | word_crf 4.074 | length 5.105 | ppl 272.18 | wps 59819.2 | wpb 5311.6 | bsz 208.1 | num_updates 44688 | best_loss 6.867
2021-03-28 00:24:13 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 76 @ 44688 updates
2021-03-28 00:24:13 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint76.pt
2021-03-28 00:24:14 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint76.pt
2021-03-28 00:24:17 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint76.pt (epoch 76 @ 44688 updates, score 8.088) (writing took 3.6393612360006955 seconds)
2021-03-28 00:24:17 | INFO | fairseq_cli.train | end of epoch 76 (average epoch stats below)
2021-03-28 00:24:17 | INFO | train | epoch 076 | loss 4.169 | nll_loss 3.291 | word_ins 4.741 | word_crf 1.527 | length 2.711 | ppl 17.99 | wps 19254.1 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 44688 | lr 0.000236524 | gnorm 1.685 | train_wall 205 | gb_free 9.5 | wall 15805
2021-03-28 00:24:17 | INFO | fairseq.trainer | begin training epoch 77
2021-03-28 00:24:17 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 00:24:21 | INFO | train_inner | epoch 077:   12 / 588 loss=4.168, nll_loss=3.296, word_ins=4.743, word_crf=1.525, length=2.71, ppl=17.98, wps=16263.3, ups=2.38, wpb=6828.8, bsz=267.5, num_updates=44700, lr=0.000236492, gnorm=1.805, train_wall=35, gb_free=9.9, wall=15809
2021-03-28 00:24:57 | INFO | train_inner | epoch 077:   112 / 588 loss=4.059, nll_loss=3.197, word_ins=4.664, word_crf=1.458, length=2.697, ppl=16.67, wps=20131.7, ups=2.84, wpb=7098.1, bsz=270.1, num_updates=44800, lr=0.000236228, gnorm=1.63, train_wall=35, gb_free=9.2, wall=15845
2021-03-28 00:25:32 | INFO | train_inner | epoch 077:   212 / 588 loss=4.154, nll_loss=3.293, word_ins=4.745, word_crf=1.514, length=2.677, ppl=17.81, wps=19969.9, ups=2.84, wpb=7019.9, bsz=260.2, num_updates=44900, lr=0.000235965, gnorm=1.731, train_wall=35, gb_free=9.3, wall=15880
2021-03-28 00:26:06 | INFO | train_inner | epoch 077:   312 / 588 loss=4.063, nll_loss=3.197, word_ins=4.661, word_crf=1.47, length=2.623, ppl=16.71, wps=20021.5, ups=2.89, wpb=6922.7, bsz=281.2, num_updates=45000, lr=0.000235702, gnorm=1.587, train_wall=34, gb_free=10.4, wall=15914
2021-03-28 00:26:42 | INFO | train_inner | epoch 077:   412 / 588 loss=4.329, nll_loss=3.438, word_ins=4.865, word_crf=1.63, length=2.667, ppl=20.09, wps=19645, ups=2.83, wpb=6941.8, bsz=278.8, num_updates=45100, lr=0.000235441, gnorm=1.864, train_wall=35, gb_free=9.5, wall=15950
2021-03-28 00:27:17 | INFO | train_inner | epoch 077:   512 / 588 loss=4.235, nll_loss=3.365, word_ins=4.807, word_crf=1.562, length=2.697, ppl=18.83, wps=19633.7, ups=2.86, wpb=6875.3, bsz=246.2, num_updates=45200, lr=0.00023518, gnorm=1.619, train_wall=35, gb_free=9.7, wall=15985
2021-03-28 00:27:43 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 00:27:46 | INFO | valid | epoch 077 | valid on 'valid' subset | loss 8.115 | nll_loss 5.754 | word_ins 7.023 | word_crf 4.091 | length 5.126 | ppl 277.31 | wps 59915.3 | wpb 5311.6 | bsz 208.1 | num_updates 45276 | best_loss 6.867
2021-03-28 00:27:46 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 77 @ 45276 updates
2021-03-28 00:27:46 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint77.pt
2021-03-28 00:27:47 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint77.pt
2021-03-28 00:27:50 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint77.pt (epoch 77 @ 45276 updates, score 8.115) (writing took 3.637051290999807 seconds)
2021-03-28 00:27:50 | INFO | fairseq_cli.train | end of epoch 77 (average epoch stats below)
2021-03-28 00:27:50 | INFO | train | epoch 077 | loss 4.142 | nll_loss 3.267 | word_ins 4.721 | word_crf 1.511 | length 2.698 | ppl 17.65 | wps 19310.5 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 45276 | lr 0.000234983 | gnorm 1.652 | train_wall 204 | gb_free 9.3 | wall 16018
2021-03-28 00:27:50 | INFO | fairseq.trainer | begin training epoch 78
2021-03-28 00:27:50 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 00:27:58 | INFO | train_inner | epoch 078:   24 / 588 loss=4.041, nll_loss=3.134, word_ins=4.604, word_crf=1.451, length=2.877, ppl=16.46, wps=17097.8, ups=2.4, wpb=7135.2, bsz=300.7, num_updates=45300, lr=0.00023492, gnorm=1.493, train_wall=34, gb_free=9.4, wall=16026
2021-03-28 00:28:34 | INFO | train_inner | epoch 078:   124 / 588 loss=4.166, nll_loss=3.306, word_ins=4.756, word_crf=1.531, length=2.572, ppl=17.95, wps=19614, ups=2.8, wpb=7003.3, bsz=269.8, num_updates=45400, lr=0.000234662, gnorm=1.902, train_wall=35, gb_free=9.7, wall=16062
2021-03-28 00:30:54 | INFO | train_inner | epoch 078:   524 / 588 loss=4.07, nll_loss=3.209, word_ins=4.671, word_crf=1.475, length=2.588, ppl=16.79, wps=19750.3, ups=2.89, wpb=6835.6, bsz=281.6, num_updates=45800, lr=0.000233635, gnorm=1.759, train_wall=34, gb_free=10, wall=16202
2021-03-28 00:31:17 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 00:31:20 | INFO | valid | epoch 078 | valid on 'valid' subset | loss 8.085 | nll_loss 5.726 | word_ins 6.993 | word_crf 4.068 | length 5.207 | ppl 271.57 | wps 59862.6 | wpb 5311.6 | bsz 208.1 | num_updates 45864 | best_loss 6.867
2021-03-28 00:31:20 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 78 @ 45864 updates
2021-03-28 00:31:20 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint78.pt
2021-03-28 00:31:21 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint78.pt
2021-03-28 00:31:24 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint78.pt (epoch 78 @ 45864 updates, score 8.085) (writing took 3.6482280359996366 seconds)
2021-03-28 00:31:24 | INFO | fairseq_cli.train | end of epoch 78 (average epoch stats below)
2021-03-28 00:31:24 | INFO | train | epoch 078 | loss 4.128 | nll_loss 3.255 | word_ins 4.712 | word_crf 1.504 | length 2.676 | ppl 17.48 | wps 19208.2 | ups 2.75 | wpb 6988.7 | bsz 272.5 | num_updates 45864 | lr 0.000233472 | gnorm 1.674 | train_wall 205 | gb_free 9.4 | wall 16232
2021-03-28 00:31:24 | INFO | fairseq.trainer | begin training epoch 79
2021-03-28 00:31:24 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 00:31:37 | INFO | train_inner | epoch 079:   36 / 588 loss=4.075, nll_loss=3.188, word_ins=4.652, word_crf=1.48, length=2.686, ppl=16.85, wps=16705.4, ups=2.36, wpb=7089.6, bsz=300.9, num_updates=45900, lr=0.00023338, gnorm=1.597, train_wall=35, gb_free=9.3, wall=16245
2021-03-28 00:32:12 | INFO | train_inner | epoch 079:   136 / 588 loss=4.1, nll_loss=3.241, word_ins=4.702, word_crf=1.477, length=2.717, ppl=17.15, wps=19959.5, ups=2.84, wpb=7037.2, bsz=267, num_updates=46000, lr=0.000233126, gnorm=1.625, train_wall=35, gb_free=9.9, wall=16280

2021-03-28 00:32:47 | INFO | train_inner | epoch 079: 236 / 588 loss=4.18, nll_loss=3.312, word_ins=4.764, word_crf=1.53, length=2.682, ppl=18.13, wps=20021.6, ups=2.84, wpb=7055.7, bsz=254.7, num_updates=46100, lr=0.000232873, gnorm=1.706, train_wall=35, gb_free=9.3, wall=16315
2021-03-28 00:33:22 | INFO | train_inner | epoch 079: 336 / 588 loss=4.067, nll_loss=3.2, word_ins=4.665, word_crf=1.476, length=2.586, ppl=16.76, wps=19704.8, ups=2.85, wpb=6911.4, bsz=277.8, num_updates=46200, lr=0.000232621, gnorm=1.669, train_wall=35, gb_free=9.7, wall=16350
2021-03-28 00:33:57 | INFO | train_inner | epoch 079: 436 / 588 loss=4.105, nll_loss=3.236, word_ins=4.697, word_crf=1.49, length=2.666, ppl=17.21, wps=19821, ups=2.89, wpb=6869, bsz=269.4, num_updates=46300, lr=0.00023237, gnorm=1.67, train_wall=34, gb_free=10, wall=16385
2021-03-28 00:34:32 | INFO | train_inner | epoch 079: 536 / 588 loss=4.121, nll_loss=3.238, word_ins=4.698, word_crf=1.498, length=2.741, ppl=17.39, wps=20032.7, ups=2.86, wpb=7009.1, bsz=271, num_updates=46400, lr=0.000232119, gnorm=1.617, train_wall=35, gb_free=9.5, wall=16420
2021-03-28 00:34:50 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 00:34:54 | INFO | valid | epoch 079 | valid on 'valid' subset | loss 8.154 | nll_loss 5.769 | word_ins 7.035 | word_crf 4.119 | length 5.18 | ppl 284.82 | wps 59311.3 | wpb 5311.6 | bsz 208.1 | num_updates 46452 | best_loss 6.867
2021-03-28 00:34:54 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 79 @ 46452 updates
2021-03-28 00:34:54 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint79.pt
2021-03-28 00:34:54 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint79.pt
2021-03-28 00:34:57 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint79.pt (epoch 79 @ 46452 updates, score 8.154) (writing took 3.6617801239990513 seconds)
2021-03-28 00:34:57 | INFO | fairseq_cli.train | end of epoch 79 (average epoch stats below)
2021-03-28 00:34:57 | INFO | train | epoch 079 | loss 4.099 | nll_loss 3.228 | word_ins 4.69 | word_crf 1.486 | length 2.677 | ppl 17.13 | wps 19258.2 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 46452 | lr 0.000231989 | gnorm 1.643 | train_wall 205 | gb_free 9.7 | wall 16445
2021-03-28 00:34:57 | INFO | fairseq.trainer | begin training epoch 80
2021-03-28 00:34:57 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 00:35:14 | INFO | train_inner | epoch 080: 48 / 588 loss=4.063, nll_loss=3.197, word_ins=4.664, word_crf=1.469, length=2.616, ppl=16.71, wps=16386.8, ups=2.37, wpb=6922.9, bsz=273.4, num_updates=46500, lr=0.000231869, gnorm=1.596, train_wall=35, gb_free=9.6, wall=16462
2021-03-28 00:35:50 | INFO | train_inner | epoch 080: 148 / 588 loss=4.095, nll_loss=3.218, word_ins=4.68, word_crf=1.481, length=2.747, ppl=17.09, wps=20066.9, ups=2.8, wpb=7163.9, bsz=277.5, num_updates=46600, lr=0.000231621, gnorm=1.816, train_wall=35, gb_free=10.3, wall=16498
2021-03-28 00:36:25 | INFO | train_inner | epoch 080: 248 / 588 loss=4.169, nll_loss=3.297, word_ins=4.75, word_crf=1.525, length=2.693, ppl=17.99, wps=19988.5, ups=2.83, wpb=7065.3, bsz=263.4, num_updates=46700, lr=0.000231372, gnorm=1.749, train_wall=35, gb_free=9.5, wall=16533
2021-03-28 00:37:00 | INFO | train_inner | epoch 080: 348 / 588 loss=4.01, nll_loss=3.142, word_ins=4.616, word_crf=1.438, length=2.646, ppl=16.12, wps=19977.8, ups=2.88, wpb=6943.4, bsz=283.6, num_updates=46800, lr=0.000231125, gnorm=1.642, train_wall=34, gb_free=9.5, wall=16568
2021-03-28 00:37:35 | INFO | train_inner | epoch 080: 448 / 588 loss=4.088, nll_loss=3.23, word_ins=4.693, word_crf=1.481, length=2.605, ppl=17.01, wps=19747, ups=2.87, wpb=6879.6, bsz=262.8, num_updates=46900, lr=0.000230879, gnorm=1.642, train_wall=35, gb_free=9.4, wall=16603
2021-03-28 00:38:10 | INFO | train_inner | epoch 080: 548 / 588 loss=4.172, nll_loss=3.29, word_ins=4.742, word_crf=1.538, length=2.632, ppl=18.02, wps=19832.2, ups=2.84, wpb=6978.2, bsz=274.2, num_updates=47000, lr=0.000230633, gnorm=1.684, train_wall=35, gb_free=9.2, wall=16638
2021-03-28 00:38:24 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 00:38:27 | INFO | valid | epoch 080 | valid on 'valid' subset | loss 8.237 | nll_loss 5.808 | word_ins 7.078 | word_crf 4.173 | length 5.257 | ppl 301.77 | wps 59776.7 | wpb 5311.6 | bsz 208.1 | num_updates 47040 | best_loss 6.867
2021-03-28 00:38:27 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 80 @ 47040 updates
2021-03-28 00:38:27 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint80.pt
2021-03-28 00:38:28 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint80.pt
2021-03-28 00:38:31 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint80.pt (epoch 80 @ 47040 updates, score 8.237) (writing took 3.6441418829999748 seconds)
2021-03-28 00:38:31 | INFO | fairseq_cli.train | end of epoch 80 (average epoch stats below)
2021-03-28 00:38:31 | INFO | train | epoch 080 | loss 4.093 | nll_loss 3.223 | word_ins 4.686 | word_crf 1.484 | length 2.652 | ppl 17.06 | wps 19232.3 | ups 2.75 | wpb 6988.7 | bsz 272.5 | num_updates 47040 | lr 0.000230535 | gnorm 1.702 | train_wall 205 | gb_free 9.6 | wall 16659
2021-03-28 00:38:31 | INFO | fairseq.trainer | begin training epoch 81
2021-03-28 00:38:31 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 00:38:52 | INFO | train_inner | epoch 081: 60 / 588 loss=4.065, nll_loss=3.202, word_ins=4.669, word_crf=1.463, length=2.672, ppl=16.73, wps=16600.2, ups=2.37, wpb=6989.7, bsz=265.6, num_updates=47100, lr=0.000230388, gnorm=1.719, train_wall=35, gb_free=9.3, wall=16680
2021-03-28 00:39:27 | INFO | train_inner | epoch 081: 160 / 588 loss=4.088, nll_loss=3.242, word_ins=4.705, word_crf=1.479, length=2.571, ppl=17.01, wps=19654.5, ups=2.85, wpb=6902.6, bsz=248.6, num_updates=47200, lr=0.000230144, gnorm=1.719, train_wall=35, gb_free=9.8, wall=16715
2021-03-28 00:40:02 | INFO | train_inner | epoch 081: 260 / 588 loss=3.879, nll_loss=3.007, word_ins=4.501, word_crf=1.365, length=2.638, ppl=14.71, wps=20517.3, ups=2.91, wpb=7057.6, bsz=307.9, num_updates=47300, lr=0.0002299, gnorm=1.624, train_wall=34, gb_free=9.4, wall=16750
2021-03-28 00:40:37 | INFO | train_inner | epoch 081: 360 / 588 loss=4.166, nll_loss=3.303, word_ins=4.755, word_crf=1.525, length=2.635, ppl=17.96, wps=19475.8, ups=2.81, wpb=6921.8, bsz=261.3, num_updates=47400, lr=0.000229658, gnorm=1.635, train_wall=35, gb_free=9.4, wall=16785
2021-03-28 00:41:12 | INFO | train_inner | epoch 081: 460 / 588 loss=4.077, nll_loss=3.196, word_ins=4.663, word_crf=1.474, length=2.711, ppl=16.87, wps=20160.3, ups=2.86, wpb=7057.2, bsz=272.9, num_updates=47500, lr=0.000229416, gnorm=1.711, train_wall=35, gb_free=9.2, wall=16820
2021-03-28 00:41:47 | INFO | train_inner | epoch 081: 560 / 588 loss=4.133, nll_loss=3.247, word_ins=4.704, word_crf=1.508, length=2.728, ppl=17.54, wps=20209.9, ups=2.86, wpb=7064.2, bsz=281.2, num_updates=47600, lr=0.000229175, gnorm=1.67, train_wall=35, gb_free=9.9, wall=16855
2021-03-28 00:41:57 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 00:42:00 | INFO | valid | epoch 081 | valid on 'valid' subset | loss 8.072 | nll_loss 5.733 | word_ins 6.996 | word_crf 4.054 | length 5.206 | ppl 269.12 | wps 59571.3 | wpb 5311.6 | bsz 208.1 | num_updates 47628 | best_loss 6.867
2021-03-28 00:42:00 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 81 @ 47628 updates
2021-03-28 00:42:00 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint81.pt
2021-03-28 00:42:01 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint81.pt
2021-03-28 00:42:04 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint81.pt (epoch 81 @ 47628 updates, score 8.072) (writing took 3.674210809000215 seconds)
2021-03-28 00:42:04 | INFO | fairseq_cli.train | end of epoch 81 (average epoch stats below)
2021-03-28 00:42:04 | INFO | train | epoch 081 | loss 4.074 | nll_loss 3.205 | word_ins 4.671 | word_crf 1.473 | length 2.657 | ppl 16.84 | wps 19306.8 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 47628 | lr 0.000229107 | gnorm 1.701 | train_wall 204 | gb_free 9.8 | wall 16872
2021-03-28 00:42:04 | INFO | fairseq.trainer | begin training epoch 82
2021-03-28 00:42:04 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 00:42:29 | INFO | train_inner | epoch 082: 72 / 588 loss=4.09, nll_loss=3.242, word_ins=4.704, word_crf=1.48, length=2.581, ppl=17.03, wps=16193.9, ups=2.37, wpb=6831, bsz=253.6, num_updates=47700, lr=0.000228934, gnorm=1.832, train_wall=35, gb_free=9.7, wall=16897
2021-03-28 00:43:04 | INFO | train_inner | epoch 082: 172 / 588 loss=4.066, nll_loss=3.223, word_ins=4.69, word_crf=1.466, length=2.556, ppl=16.75, wps=20068.2, ups=2.86, wpb=7014.7, bsz=258.7, num_updates=47800, lr=0.000228695, gnorm=1.714, train_wall=35, gb_free=9.8, wall=16932
2021-03-28 00:43:39 | INFO | train_inner | epoch 082: 272 / 588 loss=4.044, nll_loss=3.174, word_ins=4.644, word_crf=1.466, length=2.554, ppl=16.49, wps=20057.6, ups=2.86, wpb=7002.6, bsz=285.7, num_updates=47900, lr=0.000228456, gnorm=1.828, train_wall=35, gb_free=9.9, wall=16967
2021-03-28 00:44:14 | INFO | train_inner | epoch 082: 372 / 588 loss=4.022, nll_loss=3.153, word_ins=4.627, word_crf=1.437, length=2.716, ppl=16.25, wps=20330.6, ups=2.9, wpb=7001.3, bsz=279.2, num_updates=48000, lr=0.000228218, gnorm=1.657, train_wall=34, gb_free=9.4, wall=17002
2021-03-28 00:44:48 | INFO | train_inner | epoch 082: 472 / 588 loss=4.124, nll_loss=3.237, word_ins=4.699, word_crf=1.501, length=2.739, ppl=17.43, wps=20559.5, ups=2.9, wpb=7096.1, bsz=273.4, num_updates=48100, lr=0.00022798, gnorm=1.69, train_wall=34, gb_free=9.5, wall=17036
2021-03-28 00:45:24 | INFO | train_inner | epoch 082: 572 / 588 loss=4.052, nll_loss=3.175, word_ins=4.644, word_crf=1.466, length=2.643, ppl=16.59, wps=19389.6, ups=2.8, wpb=6913.7, bsz=273.7, num_updates=48200, lr=0.000227744, gnorm=1.684, train_wall=35, gb_free=9.5, wall=17072
2021-03-28 00:45:29 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 00:45:33 | INFO | valid | epoch 082 | valid on 'valid' subset | loss 8.171 | nll_loss 5.786 | word_ins 7.044 | word_crf 4.13 | length 5.193 | ppl 288.29 | wps 60105.2 | wpb 5311.6 | bsz 208.1 | num_updates 48216 | best_loss 6.867
2021-03-28 00:45:33 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 82 @ 48216 updates
2021-03-28 00:45:33 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint82.pt
2021-03-28 00:45:33 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint82.pt
2021-03-28 00:45:36 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint82.pt (epoch 82 @ 48216 updates, score 8.171) (writing took 3.6482901410017803 seconds)
2021-03-28 00:45:36 | INFO | fairseq_cli.train | end of epoch 82 (average epoch stats below)

2021-03-28 00:45:36 | INFO | train | epoch 082 | loss 4.057 | nll_loss 3.191 | word_ins 4.66 | word_crf 1.464 | length 2.639 | ppl 16.65 | wps 19337.2 | ups 2.77 | wpb 6988.7 | bsz 272.5 | num_updates 48216 | lr 0.000227706 | gnorm 1.715 | train_wall 204 | gb_free 9.6 | wall 17084
2021-03-28 00:45:36 | INFO | fairseq.trainer | begin training epoch 83
2021-03-28 00:45:36 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 00:46:06 | INFO | train_inner | epoch 083:   84 / 588 loss=4.128, nll_loss=3.282, word_ins=4.738, word_crf=1.495, length=2.631, ppl=17.48, wps=16786, ups=2.37, wpb=7081.4, bsz=254.6, num_updates=48300, lr=0.000227508, gnorm=1.739, train_wall=35, gb_free=10, wall=17114
2021-03-28 00:46:41 | INFO | train_inner | epoch 083:   184 / 588 loss=4.008, nll_loss=3.15, word_ins=4.629, word_crf=1.433, length=2.605, ppl=16.09, wps=19968.9, ups=2.86, wpb=6975.2, bsz=264.9, num_updates=48400, lr=0.000227273, gnorm=1.628, train_wall=35, gb_free=9.7, wall=17149
2021-03-28 00:47:16 | INFO | train_inner | epoch 083:   284 / 588 loss=3.995, nll_loss=3.124, word_ins=4.601, word_crf=1.437, length=2.574, ppl=15.94, wps=20204.6, ups=2.86, wpb=7062.1, bsz=291.9, num_updates=48500, lr=0.000227038, gnorm=1.814, train_wall=35, gb_free=9.2, wall=17184
2021-03-28 00:47:51 | INFO | train_inner | epoch 083:   384 / 588 loss=3.959, nll_loss=3.099, word_ins=4.579, word_crf=1.415, length=2.543, ppl=15.55, wps=19504.6, ups=2.84, wpb=6860.8, bsz=289.2, num_updates=48600, lr=0.000226805, gnorm=1.721, train_wall=35, gb_free=9.7, wall=17219
2021-03-28 00:48:26 | INFO | train_inner | epoch 083:   484 / 588 loss=4.095, nll_loss=3.204, word_ins=4.67, word_crf=1.482, length=2.779, ppl=17.09, wps=20319.4, ups=2.88, wpb=7062.8, bsz=276.2, num_updates=48700, lr=0.000226572, gnorm=1.824, train_wall=34, gb_free=9.4, wall=17254
2021-03-28 00:49:00 | INFO | train_inner | epoch 083:   584 / 588 loss=4.088, nll_loss=3.218, word_ins=4.682, word_crf=1.48, length=2.666, ppl=17.01, wps=19841.5, ups=2.88, wpb=6881.5, bsz=266.3, num_updates=48800, lr=0.000226339, gnorm=1.722, train_wall=34, gb_free=9.6, wall=17288
2021-03-28 00:49:02 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 00:49:05 | INFO | valid | epoch 083 | valid on 'valid' subset | loss 8.217 | nll_loss 5.803 | word_ins 7.07 | word_crf 4.159 | length 5.222 | ppl 297.46 | wps 59156.5 | wpb 5311.6 | bsz 208.1 | num_updates 48804 | best_loss 6.867
2021-03-28 00:49:05 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 83 @ 48804 updates
2021-03-28 00:49:05 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint83.pt
2021-03-28 00:49:06 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint83.pt
2021-03-28 00:49:09 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint83.pt (epoch 83 @ 48804 updates, score 8.217) (writing took 3.6793228999995335 seconds)
2021-03-28 00:49:09 | INFO | fairseq_cli.train | end of epoch 83 (average epoch stats below)
2021-03-28 00:49:09 | INFO | train | epoch 083 | loss 4.048 | nll_loss 3.183 | word_ins 4.653 | word_crf 1.459 | length 2.627 | ppl 16.54 | wps 19323.3 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 48804 | lr 0.00022633 | gnorm 1.74 | train_wall 204 | gb_free 9.4 | wall 17297
2021-03-28 00:49:09 | INFO | fairseq.trainer | begin training epoch 84
2021-03-28 00:49:09 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 00:49:44 | INFO | train_inner | epoch 084:   96 / 588 loss=4.013, nll_loss=3.18, word_ins=4.655, word_crf=1.431, length=2.549, ppl=16.15, wps=15979.2, ups=2.32, wpb=6877.5, bsz=249.3, num_updates=48900, lr=0.000226108, gnorm=1.688, train_wall=36, gb_free=9.3, wall=17331
2021-03-28 00:50:18 | INFO | train_inner | epoch 084:   196 / 588 loss=3.774, nll_loss=2.9, word_ins=4.412, word_crf=1.297, length=2.712, ppl=13.68, wps=20527.8, ups=2.91, wpb=7047.9, bsz=308.9, num_updates=49000, lr=0.000225877, gnorm=1.547, train_wall=34, gb_free=9.5, wall=17366
2021-03-28 00:50:53 | INFO | train_inner | epoch 084:   296 / 588 loss=4.142, nll_loss=3.285, word_ins=4.744, word_crf=1.515, length=2.557, ppl=17.66, wps=19626.4, ups=2.83, wpb=6933.3, bsz=255.1, num_updates=49100, lr=0.000225647, gnorm=1.744, train_wall=35, gb_free=10.1, wall=17401
2021-03-28 00:51:28 | INFO | train_inner | epoch 084:   396 / 588 loss=4.039, nll_loss=3.167, word_ins=4.641, word_crf=1.44, length=2.784, ppl=16.44, wps=20346.5, ups=2.86, wpb=7123.4, bsz=265.3, num_updates=49200, lr=0.000225417, gnorm=1.575, train_wall=35, gb_free=9.7, wall=17436
2021-03-28 00:52:03 | INFO | train_inner | epoch 084:   496 / 588 loss=4.002, nll_loss=3.146, word_ins=4.62, word_crf=1.45, length=2.425, ppl=16.02, wps=19666.8, ups=2.84, wpb=6926.9, bsz=289.4, num_updates=49300, lr=0.000225189, gnorm=1.631, train_wall=35, gb_free=9.4, wall=17471
2021-03-28 00:52:36 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 00:52:39 | INFO | valid | epoch 084 | valid on 'valid' subset | loss 8.255 | nll_loss 5.825 | word_ins 7.093 | word_crf 4.186 | length 5.225 | ppl 305.57 | wps 59981.9 | wpb 5311.6 | bsz 208.1 | num_updates 49392 | best_loss 6.867
2021-03-28 00:52:39 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 84 @ 49392 updates
2021-03-28 00:52:39 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint84.pt
2021-03-28 00:52:40 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint84.pt
2021-03-28 00:52:43 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint84.pt (epoch 84 @ 49392 updates, score 8.255) (writing took 3.6416564750034013 seconds)
2021-03-28 00:52:43 | INFO | fairseq_cli.train | end of epoch 84 (average epoch stats below)
2021-03-28 00:52:43 | INFO | train | epoch 084 | loss 4.008 | nll_loss 3.147 | word_ins 4.624 | word_crf 1.434 | length 2.614 | ppl 16.09 | wps 19236 | ups 2.75 | wpb 6988.7 | bsz 272.5 | num_updates 49392 | lr 0.000224979 | gnorm 1.63 | train_wall 205 | gb_free 9.4 | wall 17510
2021-03-28 00:52:43 | INFO | fairseq.trainer | begin training epoch 85
2021-03-28 00:52:43 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 00:52:46 | INFO | train_inner | epoch 085:   8 / 588 loss=4.032, nll_loss=3.161, word_ins=4.634, word_crf=1.451, length=2.647, ppl=16.36, wps=16636.8, ups=2.37, wpb=7005.1, bsz=273.8, num_updates=49400, lr=0.000224961, gnorm=1.569, train_wall=35, gb_free=9.8, wall=17513
2021-03-28 00:53:20 | INFO | train_inner | epoch 085:   108 / 588 loss=3.951, nll_loss=3.098, word_ins=4.583, word_crf=1.401, length=2.584, ppl=15.46, wps=20310.3, ups=2.88, wpb=7053.1, bsz=274.4, num_updates=49500, lr=0.000224733, gnorm=1.746, train_wall=34, gb_free=9.3, wall=17548
2021-03-28 00:53:55 | INFO | train_inner | epoch 085:   208 / 588 loss=4.034, nll_loss=3.19, word_ins=4.663, word_crf=1.444, length=2.583, ppl=16.38, wps=20135, ups=2.88, wpb=6993.7, bsz=255.5, num_updates=49600, lr=0.000224507, gnorm=1.733, train_wall=34, gb_free=9.6, wall=17583
2021-03-28 00:54:30 | INFO | train_inner | epoch 085:   308 / 588 loss=3.911, nll_loss=3.05, word_ins=4.543, word_crf=1.372, length=2.678, ppl=15.05, wps=20000.4, ups=2.88, wpb=6950.7, bsz=272.5, num_updates=49700, lr=0.000224281, gnorm=1.566, train_wall=34, gb_free=9.6, wall=17618
2021-03-28 00:55:05 | INFO | train_inner | epoch 085:   408 / 588 loss=4.069, nll_loss=3.21, word_ins=4.677, word_crf=1.481, length=2.504, ppl=16.79, wps=19513.6, ups=2.83, wpb=6885.8, bsz=273.3, num_updates=49800, lr=0.000224055, gnorm=1.94, train_wall=35, gb_free=9.8, wall=17653
2021-03-28 00:55:41 | INFO | train_inner | epoch 085:   508 / 588 loss=4.102, nll_loss=3.239, word_ins=4.701, word_crf=1.494, length=2.581, ppl=17.18, wps=19844.1, ups=2.81, wpb=7050.5, bsz=266.3, num_updates=49900, lr=0.000223831, gnorm=1.698, train_wall=35, gb_free=9.7, wall=17688
2021-03-28 00:56:09 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 00:56:12 | INFO | valid | epoch 085 | valid on 'valid' subset | loss 8.297 | nll_loss 5.841 | word_ins 7.11 | word_crf 4.217 | length 5.246 | ppl 314.44 | wps 59613 | wpb 5311.6 | bsz 208.1 | num_updates 49980 | best_loss 6.867
2021-03-28 00:56:12 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 85 @ 49980 updates
2021-03-28 00:56:12 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint85.pt
2021-03-28 00:56:13 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint85.pt
2021-03-28 00:56:16 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint85.pt (epoch 85 @ 49980 updates, score 8.297) (writing took 3.6836506470026507 seconds)
2021-03-28 00:56:16 | INFO | fairseq_cli.train | end of epoch 85 (average epoch stats below)
2021-03-28 00:56:16 | INFO | train | epoch 085 | loss 3.999 | nll_loss 3.141 | word_ins 4.619 | word_crf 1.431 | length 2.589 | ppl 15.99 | wps 19285.1 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 49980 | lr 0.000223652 | gnorm 1.701 | train_wall 204 | gb_free 9.9 | wall 17724
2021-03-28 00:56:16 | INFO | fairseq.trainer | begin training epoch 86
2021-03-28 00:56:16 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 00:56:23 | INFO | train_inner | epoch 086:   20 / 588 loss=3.92, nll_loss=3.049, word_ins=4.538, word_crf=1.387, length=2.642, ppl=15.13, wps=16821.2, ups=2.37, wpb=7102, bsz=295, num_updates=50000, lr=0.000223607, gnorm=1.504, train_wall=35, gb_free=9.7, wall=17731
2021-03-28 00:56:23 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 00:56:26 | INFO | valid | epoch 086 | valid on 'valid' subset | loss 8.273 | nll_loss 5.828 | word_ins 7.096 | word_crf 4.199 | length 5.255 | ppl 309.27 | wps 59754.2 | wpb 5311.6 | bsz 208.1 | num_updates 50000 | best_loss 6.867
2021-03-28 00:56:26 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 86 @ 50000 updates
2021-03-28 00:56:26 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint_86_50000.pt
2021-03-28 00:56:27 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint_86_50000.pt
2021-03-28 00:56:30 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint_86_50000.pt (epoch 86 @ 50000 updates, score 8.273) (writing took 3.9880910119973123 seconds)
2021-03-28 00:57:05 | INFO | train_inner | epoch 086:   120 / 588 loss=4.014, nll_loss=3.17, word_ins=4.645, word_crf=1.439, length=2.517, ppl=16.15, wps=16527.7, ups=2.36, wpb=7002.4, bsz=265, num_updates=50100, lr=0.000223384, gnorm=1.834, train_wall=35, gb_free=10.3, wall=17773
2021-03-28 00:57:40 | INFO | train_inner | epoch 086:   220 / 588 loss=3.967, nll_loss=3.095, word_ins=4.583, word_crf=1.399, length=2.762, ppl=15.63, wps=20258, ups=2.89, wpb=7020.8, bsz=263.9, num_updates=50200, lr=0.000223161, gnorm=1.626, train_wall=34, gb_free=9.4, wall=17808

2021-03-28 00:58:14 | INFO | train_inner | epoch 086:   320 / 588 loss=3.846, nll_loss=2.993, word_ins=4.491, word_crf=1.349, length=2.512, ppl=14.38, wps=20055.6, ups=2.9, wpb=6913.9, bsz=292.5, num_updates=50300, lr=0.000222939, gnorm=1.691, train_wall=34, gb_free=9.5, wall=17842
2021-03-28 00:58:50 | INFO | train_inner | epoch 086:   420 / 588 loss=4.176, nll_loss=3.335, word_ins=4.787, word_crf=1.536, length=2.462, ppl=18.08, wps=19529.2, ups=2.83, wpb=6897.8, bsz=241.4, num_updates=50400, lr=0.000222718, gnorm=1.845, train_wall=35, gb_free=9.2, wall=17878
2021-03-28 00:59:25 | INFO | train_inner | epoch 086:   520 / 588 loss=3.994, nll_loss=3.139, word_ins=4.617, word_crf=1.429, length=2.558, ppl=15.93, wps=19989.9, ups=2.86, wpb=6986, bsz=277.2, num_updates=50500, lr=0.000222497, gnorm=1.58, train_wall=35, gb_free=9.9, wall=17912
2021-03-28 00:59:49 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 00:59:52 | INFO | valid | epoch 086 | valid on 'valid' subset | loss 8.295 | nll_loss 5.838 | word_ins 7.108 | word_crf 4.208 | length 5.324 | ppl 313.98 | wps 59540.4 | wpb 5311.6 | bsz 208.1 | num_updates 50568 | best_loss 6.867
2021-03-28 00:59:52 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 86 @ 50568 updates
2021-03-28 00:59:52 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint86.pt
2021-03-28 00:59:53 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint86.pt
2021-03-28 00:59:56 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint86.pt (epoch 86 @ 50568 updates, score 8.295) (writing took 3.671152594000887 seconds)
2021-03-28 00:59:56 | INFO | fairseq_cli.train | end of epoch 86 (average epoch stats below)
2021-03-28 00:59:56 | INFO | train | epoch 086 | loss 3.994 | nll_loss 3.136 | word_ins 4.615 | word_crf 1.428 | length 2.581 | ppl 15.93 | wps 18638 | ups 2.67 | wpb 6988.7 | bsz 272.5 | num_updates 50568 | lr 0.000222347 | gnorm 1.729 | train_wall 204 | gb_free 9.3 | wall 17944
2021-03-28 00:59:56 | INFO | fairseq.trainer | begin training epoch 87
2021-03-28 00:59:56 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 01:00:08 | INFO | train_inner | epoch 087:   32 / 588 loss=4.035, nll_loss=3.18, word_ins=4.65, word_crf=1.451, length=2.596, ppl=16.4, wps=16088.2, ups=2.31, wpb=6954.3, bsz=273.1, num_updates=50600, lr=0.000222277, gnorm=1.853, train_wall=36, gb_free=10, wall=17956
2021-03-28 01:00:43 | INFO | train_inner | epoch 087:   132 / 588 loss=3.957, nll_loss=3.118, word_ins=4.602, word_crf=1.414, length=2.412, ppl=15.53, wps=19632.5, ups=2.82, wpb=6953.4, bsz=272.5, num_updates=50700, lr=0.000222058, gnorm=1.788, train_wall=35, gb_free=9.3, wall=17991
2021-03-28 01:01:18 | INFO | train_inner | epoch 087:   232 / 588 loss=3.863, nll_loss=3.003, word_ins=4.503, word_crf=1.346, length=2.655, ppl=14.55, wps=19951.8, ups=2.87, wpb=6955.2, bsz=277.2, num_updates=50800, lr=0.000221839, gnorm=1.619, train_wall=35, gb_free=9.5, wall=18026
2021-03-28 01:01:54 | INFO | train_inner | epoch 087:   332 / 588 loss=3.961, nll_loss=3.076, word_ins=4.564, word_crf=1.409, length=2.699, ppl=15.58, wps=19850.5, ups=2.8, wpb=7087.3, bsz=287.3, num_updates=50900, lr=0.000221621, gnorm=1.795, train_wall=35, gb_free=9.5, wall=18062
2021-03-28 01:02:28 | INFO | train_inner | epoch 087:   432 / 588 loss=3.95, nll_loss=3.05, word_ins=4.538, word_crf=1.377, length=2.532, ppl=14.92, wps=20212.3, ups=2.89, wpb=7001.3, bsz=285.7, num_updates=51000, lr=0.000221404, gnorm=1.72, train_wall=34, gb_free=9.9, wall=18096
2021-03-28 01:03:04 | INFO | train_inner | epoch 087:   532 / 588 loss=4.144, nll_loss=3.28, word_ins=4.74, word_crf=1.507, length=2.669, ppl=17.68, wps=19895.5, ups=2.84, wpb=7007.4, bsz=244.4, num_updates=51100, lr=0.000221187, gnorm=1.75, train_wall=35, gb_free=9.4, wall=18132
2021-03-28 01:03:24 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 01:03:28 | INFO | valid | epoch 087 | valid on 'valid' subset | loss 8.269 | nll_loss 5.822 | word_ins 7.092 | word_crf 4.195 | length 5.281 | ppl 308.45 | wps 59586.6 | wpb 5311.6 | bsz 208.1 | num_updates 51156 | best_loss 6.867
2021-03-28 01:03:28 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 87 @ 51156 updates
2021-03-28 01:03:28 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint87.pt
2021-03-28 01:03:28 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint87.pt
2021-03-28 01:03:31 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint87.pt (epoch 87 @ 51156 updates, score 8.269) (writing took 3.6518860979995225 seconds)
2021-03-28 01:03:31 | INFO | fairseq_cli.train | end of epoch 87 (average epoch stats below)
2021-03-28 01:03:31 | INFO | train | epoch 087 | loss 3.976 | nll_loss 3.12 | word_ins 4.601 | word_crf 1.418 | length 2.578 | ppl 15.74 | wps 19107.7 | ups 2.73 | wpb 6988.7 | bsz 272.5 | num_updates 51156 | lr 0.000221066 | gnorm 1.741 | train_wall 206 | gb_free 9.6 | wall 18159
2021-03-28 01:03:31 | INFO | fairseq.trainer | begin training epoch 88
2021-03-28 01:03:31 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 01:03:47 | INFO | train_inner | epoch 088:   44 / 588 loss=3.936, nll_loss=3.089, word_ins=4.574, word_crf=1.401, length=2.478, ppl=15.3, wps=16247.4, ups=2.32, wpb=7016.3, bsz=281.8, num_updates=51200, lr=0.000220971, gnorm=1.718, train_wall=35, gb_free=9.4, wall=18175
2021-03-28 01:04:22 | INFO | train_inner | epoch 088:   144 / 588 loss=3.931, nll_loss=3.087, word_ins=4.575, word_crf=1.391, length=2.526, ppl=15.25, wps=20024.7, ups=2.84, wpb=7043.3, bsz=271, num_updates=51300, lr=0.000220755, gnorm=1.691, train_wall=35, gb_free=9.7, wall=18210
2021-03-28 01:04:57 | INFO | train_inner | epoch 088:   244 / 588 loss=3.856, nll_loss=2.996, word_ins=4.495, word_crf=1.347, length=2.612, ppl=14.48, wps=19955.9, ups=2.87, wpb=6948.8, bsz=292, num_updates=51400, lr=0.000220541, gnorm=1.595, train_wall=35, gb_free=9.5, wall=18245
2021-03-28 01:05:31 | INFO | train_inner | epoch 088:   344 / 588 loss=3.851, nll_loss=3.006, word_ins=4.504, word_crf=1.35, length=2.481, ppl=14.43, wps=20100.3, ups=2.89, wpb=6954, bsz=293, num_updates=51500, lr=0.000220326, gnorm=1.681, train_wall=34, gb_free=9.3, wall=18279
2021-03-28 01:06:07 | INFO | train_inner | epoch 088:   444 / 588 loss=4.183, nll_loss=3.308, word_ins=4.763, word_crf=1.535, length=2.669, ppl=18.17, wps=19786.1, ups=2.84, wpb=6968, bsz=249.2, num_updates=51600, lr=0.000220113, gnorm=2.108, train_wall=35, gb_free=9.6, wall=18315
2021-03-28 01:06:42 | INFO | train_inner | epoch 088:   544 / 588 loss=4.035, nll_loss=3.167, word_ins=4.641, word_crf=1.453, length=2.618, ppl=16.4, wps=20171.8, ups=2.85, wpb=7071.1, bsz=270.7, num_updates=51700, lr=0.0002199, gnorm=1.706, train_wall=35, gb_free=9.5, wall=18350
2021-03-28 01:06:57 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 01:07:00 | INFO | valid | epoch 088 | valid on 'valid' subset | loss 8.287 | nll_loss 5.834 | word_ins 7.104 | word_crf 4.213 | length 5.222 | ppl 312.42 | wps 59971.2 | wpb 5311.6 | bsz 208.1 | num_updates 51744 | best_loss 6.867
2021-03-28 01:07:00 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 88 @ 51744 updates
2021-03-28 01:07:00 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint88.pt
2021-03-28 01:07:01 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint88.pt
2021-03-28 01:07:04 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint88.pt (epoch 88 @ 51744 updates, score 8.287) (writing took 3.6744139910006197 seconds)
2021-03-28 01:07:04 | INFO | fairseq_cli.train | end of epoch 88 (average epoch stats below)
2021-03-28 01:07:04 | INFO | train | epoch 088 | loss 3.962 | nll_loss 3.108 | word_ins 4.592 | word_crf 1.41 | length 2.562 | ppl 15.59 | wps 19300.6 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 51744 | lr 0.000219806 | gnorm 1.744 | train_wall 204 | gb_free 9.9 | wall 18372
2021-03-28 01:07:04 | INFO | fairseq.trainer | begin training epoch 89
2021-03-28 01:07:04 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 01:07:24 | INFO | train_inner | epoch 089:   56 / 588 loss=3.936, nll_loss=3.11, word_ins=4.598, word_crf=1.391, length=2.46, ppl=15.3, wps=16452.4, ups=2.38, wpb=6926.8, bsz=247.6, num_updates=51800, lr=0.000219687, gnorm=1.778, train_wall=35, gb_free=10.5, wall=18392
2021-03-28 01:07:59 | INFO | train_inner | epoch 089:   156 / 588 loss=3.953, nll_loss=3.126, word_ins=4.611, word_crf=1.399, length=2.48, ppl=15.48, wps=19908, ups=2.84, wpb=7006.8, bsz=260, num_updates=51900, lr=0.000219476, gnorm=1.664, train_wall=35, gb_free=9.4, wall=18427
2021-03-28 01:08:33 | INFO | train_inner | epoch 089:   256 / 588 loss=3.892, nll_loss=3.044, word_ins=4.539, word_crf=1.368, length=2.544, ppl=14.84, wps=19965.5, ups=2.9, wpb=6892.2, bsz=274.1, num_updates=52000, lr=0.000219265, gnorm=1.878, train_wall=34, gb_free=9.4, wall=18461
2021-03-28 01:09:09 | INFO | train_inner | epoch 089:   356 / 588 loss=3.969, nll_loss=3.119, word_ins=4.602, word_crf=1.425, length=2.428, ppl=15.66, wps=19534.2, ups=2.82, wpb=6918.4, bsz=275.7, num_updates=52100, lr=0.000219054, gnorm=1.852, train_wall=35, gb_free=10, wall=18497
2021-03-28 01:09:44 | INFO | train_inner | epoch 089:   456 / 588 loss=4.06, nll_loss=3.173, word_ins=4.646, word_crf=1.461, length=2.769, ppl=16.68, wps=20309.8, ups=2.85, wpb=7132.2, bsz=276.2, num_updates=52200, lr=0.000218844, gnorm=1.895, train_wall=35, gb_free=9.2, wall=18532
2021-03-28 01:10:19 | INFO | train_inner | epoch 089:   556 / 588 loss=3.907, nll_loss=3.045, word_ins=4.537, word_crf=1.377, length=2.616, ppl=15, wps=20180.3, ups=2.87, wpb=7030.2, bsz=279.8, num_updates=52300, lr=0.000218635, gnorm=1.633, train_wall=35, gb_free=9.4, wall=18567
2021-03-28 01:10:30 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 01:10:33 | INFO | valid | epoch 089 | valid on 'valid' subset | loss 8.338 | nll_loss 5.859 | word_ins 7.127 | word_crf 4.24 | length 5.35 | ppl 323.64 | wps 59774.3 | wpb 5311.6 | bsz 208.1 | num_updates 52332 | best_loss 6.867
2021-03-28 01:10:33 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 89 @ 52332 updates
2021-03-28 01:10:33 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint89.pt
2021-03-28 01:10:34 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint89.pt
2021-03-28 01:10:37 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint89.pt (epoch 89 @ 52332 updates, score 8.338) (writing took 3.6375518050008395 seconds)
2021-03-28 01:10:37 | INFO | fairseq_cli.train | end of epoch 89 (average epoch stats below)
2021-03-28 01:10:37 | INFO | train | epoch 089 | loss 3.947 | nll_loss 3.095 | word_ins 4.582 | word_crf 1.401 | length 2.544 | ppl 15.42 | wps 19303.5 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 52332 | lr 0.000218568 | gnorm 1.784 | train_wall 204 | gb_free 9.4 | wall 18585

2021-03-28 01:10:37 | INFO | fairseq.trainer | begin training epoch 90
2021-03-28 01:10:37 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 01:11:01 | INFO | train_inner | epoch 090:   68 / 588 loss=3.824, nll_loss=2.987, word_ins=4.487, word_crf=1.341, length=2.397, ppl=14.16, wps=16415.8, ups=2.38, wpb=6902.3, bsz=296.2, num_updates=52400, lr=0.000218426, gnorm=1.699, train_wall=35, gb_free=9.9, wall=18609
2021-03-28 01:11:36 | INFO | train_inner | epoch 090:   168 / 588 loss=3.954, nll_loss=3.095, word_ins=4.583, word_crf=1.402, length=2.6, ppl=15.49, wps=20329.7, ups=2.88, wpb=7063.1, bsz=273.1, num_updates=52500, lr=0.000218218, gnorm=1.8, train_wall=34, gb_free=9.6, wall=18644
2021-03-28 01:12:11 | INFO | train_inner | epoch 090:   268 / 588 loss=3.984, nll_loss=3.136, word_ins=4.618, word_crf=1.42, length=2.544, ppl=15.82, wps=19685.8, ups=2.84, wpb=6939.1, bsz=259.5, num_updates=52600, lr=0.00021801, gnorm=1.708, train_wall=35, gb_free=9.6, wall=18679
2021-03-28 01:12:46 | INFO | train_inner | epoch 090:   368 / 588 loss=3.983, nll_loss=3.128, word_ins=4.613, word_crf=1.414, length=2.625, ppl=15.81, wps=20292.9, ups=2.86, wpb=7090.4, bsz=255.7, num_updates=52700, lr=0.000217803, gnorm=1.717, train_wall=35, gb_free=10.4, wall=18714
2021-03-28 01:13:22 | INFO | train_inner | epoch 090:   468 / 588 loss=3.991, nll_loss=3.159, word_ins=4.639, word_crf=1.427, length=2.445, ppl=15.9, wps=19275.2, ups=2.78, wpb=6939.3, bsz=249.2, num_updates=52800, lr=0.000217597, gnorm=1.741, train_wall=36, gb_free=9.2, wall=18750
2021-03-28 01:13:56 | INFO | train_inner | epoch 090:   568 / 588 loss=3.769, nll_loss=2.895, word_ins=4.406, word_crf=1.304, length=2.614, ppl=13.63, wps=20188.6, ups=2.89, wpb=6982.6, bsz=313.8, num_updates=52900, lr=0.000217391, gnorm=1.516, train_wall=34, gb_free=9.1, wall=18784
2021-03-28 01:14:03 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 01:14:07 | INFO | valid | epoch 090 | valid on 'valid' subset | loss 8.349 | nll_loss 5.867 | word_ins 7.13 | word_crf 4.241 | length 5.42 | ppl 325.97 | wps 59381.8 | wpb 5311.6 | bsz 208.1 | num_updates 52920 | best_loss 6.867
2021-03-28 01:14:07 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 90 @ 52920 updates
2021-03-28 01:14:07 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint90.pt
2021-03-28 01:14:08 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint90.pt
2021-03-28 01:14:10 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint90.pt (epoch 90 @ 52920 updates, score 8.349) (writing took 3.7062933330016676 seconds)
2021-03-28 01:14:10 | INFO | fairseq_cli.train | end of epoch 90 (average epoch stats below)
2021-03-28 01:14:10 | INFO | train | epoch 090 | loss 3.924 | nll_loss 3.075 | word_ins 4.565 | word_crf 1.388 | length 2.543 | ppl 15.18 | wps 19253.4 | ups 2.75 | wpb 6988.7 | bsz 272.5 | num_updates 52920 | lr 0.00021735 | gnorm 1.698 | train_wall 205 | gb_free 9.4 | wall 18798
2021-03-28 01:14:11 | INFO | fairseq.trainer | begin training epoch 91
2021-03-28 01:14:11 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 01:14:39 | INFO | train_inner | epoch 091:   80 / 588 loss=3.872, nll_loss=3.047, word_ins=4.543, word_crf=1.357, length=2.441, ppl=14.64, wps=16449, ups=2.38, wpb=6920.5, bsz=272.2, num_updates=53000, lr=0.000217186, gnorm=1.774, train_wall=35, gb_free=9.8, wall=18826
2021-03-28 01:15:14 | INFO | train_inner | epoch 091:   180 / 588 loss=3.861, nll_loss=3.003, word_ins=4.506, word_crf=1.349, length=2.598, ppl=14.53, wps=19818.4, ups=2.85, wpb=6965, bsz=276.9, num_updates=53100, lr=0.000216982, gnorm=1.781, train_wall=35, gb_free=9.4, wall=18862
2021-03-28 01:15:49 | INFO | train_inner | epoch 091:   280 / 588 loss=3.855, nll_loss=3.006, word_ins=4.507, word_crf=1.346, length=2.549, ppl=14.47, wps=20142.1, ups=2.86, wpb=7052.2, bsz=275, num_updates=53200, lr=0.000216777, gnorm=1.703, train_wall=35, gb_free=9.6, wall=18897
2021-03-28 01:16:24 | INFO | train_inner | epoch 091:   380 / 588 loss=4.039, nll_loss=3.189, word_ins=4.662, word_crf=1.46, length=2.475, ppl=16.44, wps=19687.3, ups=2.82, wpb=6984.3, bsz=261.8, num_updates=53300, lr=0.000216574, gnorm=1.859, train_wall=35, gb_free=9.1, wall=18932
2021-03-28 01:16:59 | INFO | train_inner | epoch 091:   480 / 588 loss=3.989, nll_loss=3.146, word_ins=4.625, word_crf=1.425, length=2.511, ppl=15.88, wps=19515.1, ups=2.83, wpb=6893.9, bsz=265, num_updates=53400, lr=0.000216371, gnorm=1.676, train_wall=35, gb_free=9.3, wall=18967
2021-03-28 01:17:35 | INFO | train_inner | epoch 091:   580 / 588 loss=3.947, nll_loss=3.084, word_ins=4.574, word_crf=1.401, length=2.59, ppl=15.42, wps=19994.4, ups=2.81, wpb=7109.7, bsz=271.7, num_updates=53500, lr=0.000216169, gnorm=1.791, train_wall=35, gb_free=9.8, wall=19003
2021-03-28 01:17:38 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 01:17:41 | INFO | valid | epoch 091 | valid on 'valid' subset | loss 8.368 | nll_loss 5.879 | word_ins 7.137 | word_crf 4.262 | length 5.374 | ppl 330.28 | wps 59760.4 | wpb 5311.6 | bsz 208.1 | num_updates 53508 | best_loss 6.867
2021-03-28 01:17:41 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 91 @ 53508 updates
2021-03-28 01:17:41 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint91.pt
2021-03-28 01:17:42 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint91.pt
2021-03-28 01:17:45 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint91.pt (epoch 91 @ 53508 updates, score 8.368) (writing took 3.6450149350021093 seconds)
2021-03-28 01:17:45 | INFO | fairseq_cli.train | end of epoch 91 (average epoch stats below)
2021-03-28 01:17:45 | INFO | train | epoch 091 | loss 3.917 | nll_loss 3.068 | word_ins 4.56 | word_crf 1.385 | length 2.529 | ppl 15.11 | wps 19173.9 | ups 2.74 | wpb 6988.7 | bsz 272.5 | num_updates 53508 | lr 0.000216153 | gnorm 1.772 | train_wall 205 | gb_free 9.8 | wall 19013
2021-03-28 01:17:45 | INFO | fairseq.trainer | begin training epoch 92
2021-03-28 01:17:45 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 01:18:17 | INFO | train_inner | epoch 092:   92 / 588 loss=3.755, nll_loss=2.922, word_ins=4.435, word_crf=1.294, length=2.441, ppl=13.51, wps=16680.2, ups=2.37, wpb=7024.6, bsz=285.4, num_updates=53600, lr=0.000215967, gnorm=1.806, train_wall=35, gb_free=9.4, wall=19045
2021-03-28 01:18:52 | INFO | train_inner | epoch 092:   192 / 588 loss=3.863, nll_loss=3.006, word_ins=4.506, word_crf=1.348, length=2.611, ppl=14.55, wps=19964.6, ups=2.85, wpb=6998.3, bsz=284.6, num_updates=53700, lr=0.000215766, gnorm=1.797, train_wall=35, gb_free=9.3, wall=19080
2021-03-28 01:19:27 | INFO | train_inner | epoch 092:   292 / 588 loss=4, nll_loss=3.172, word_ins=4.652, word_crf=1.434, length=2.4, ppl=16, wps=19851.9, ups=2.83, wpb=7004.9, bsz=248.6, num_updates=53800, lr=0.000215565, gnorm=1.78, train_wall=35, gb_free=9.6, wall=19115
2021-03-28 01:20:03 | INFO | train_inner | epoch 092:   392 / 588 loss=3.88, nll_loss=3.019, word_ins=4.518, word_crf=1.365, length=2.563, ppl=14.72, wps=19809.4, ups=2.83, wpb=6988.2, bsz=287.3, num_updates=53900, lr=0.000215365, gnorm=1.774, train_wall=35, gb_free=9.7, wall=19151
2021-03-28 01:20:38 | INFO | train_inner | epoch 092:   492 / 588 loss=3.983, nll_loss=3.142, word_ins=4.621, word_crf=1.423, length=2.496, ppl=15.81, wps=19354.8, ups=2.81, wpb=6897.1, bsz=269, num_updates=54000, lr=0.000215166, gnorm=1.711, train_wall=35, gb_free=9.6, wall=19186
2021-03-28 01:21:12 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 01:21:15 | INFO | valid | epoch 092 | valid on 'valid' subset | loss 8.28 | nll_loss 5.829 | word_ins 7.089 | word_crf 4.193 | length 5.427 | ppl 310.83 | wps 59742.1 | wpb 5311.6 | bsz 208.1 | num_updates 54096 | best_loss 6.867
2021-03-28 01:21:15 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 92 @ 54096 updates
2021-03-28 01:21:15 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint92.pt
2021-03-28 01:21:16 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint92.pt
2021-03-28 01:21:19 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint92.pt (epoch 92 @ 54096 updates, score 8.28) (writing took 3.651706602999184 seconds)
2021-03-28 01:21:19 | INFO | fairseq_cli.train | end of epoch 92 (average epoch stats below)
2021-03-28 01:21:19 | INFO | train | epoch 092 | loss 3.904 | nll_loss 3.058 | word_ins 4.551 | word_crf 1.377 | length 2.511 | ppl 14.97 | wps 19199.9 | ups 2.75 | wpb 6988.7 | bsz 272.5 | num_updates 54096 | lr 0.000214975 | gnorm 1.777 | train_wall 205 | gb_free 9.2 | wall 19227
2021-03-28 01:21:19 | INFO | fairseq.trainer | begin training epoch 93
2021-03-28 01:21:19 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 01:21:20 | INFO | train_inner | epoch 093:   4 / 588 loss=3.918, nll_loss=3.055, word_ins=4.548, word_crf=1.388, length=2.561, ppl=15.12, wps=16706.3, ups=2.38, wpb=7016, bsz=270.6, num_updates=54100, lr=0.000214967, gnorm=1.836, train_wall=35, gb_free=9.5, wall=19228
2021-03-28 01:21:55 | INFO | train_inner | epoch 093:   104 / 588 loss=3.9, nll_loss=3.073, word_ins=4.566, word_crf=1.368, length=2.49, ppl=14.93, wps=19770.8, ups=2.86, wpb=6909.3, bsz=261.2, num_updates=54200, lr=0.000214768, gnorm=1.824, train_wall=35, gb_free=10.4, wall=19263
2021-03-28 01:22:30 | INFO | train_inner | epoch 093:   204 / 588 loss=3.9, nll_loss=3.054, word_ins=4.549, word_crf=1.378, length=2.478, ppl=14.93, wps=20059.1, ups=2.85, wpb=7037, bsz=274, num_updates=54300, lr=0.000214571, gnorm=1.887, train_wall=35, gb_free=9.4, wall=19298
2021-03-28 01:23:06 | INFO | train_inner | epoch 093:   304 / 588 loss=3.912, nll_loss=3.1, word_ins=4.588, word_crf=1.392, length=2.26, ppl=15.05, wps=19495.4, ups=2.85, wpb=6844.6, bsz=261.8, num_updates=54400, lr=0.000214373, gnorm=1.839, train_wall=35, gb_free=10, wall=19333
2021-03-28 01:23:40 | INFO | train_inner | epoch 093:   404 / 588 loss=3.817, nll_loss=2.973, word_ins=4.478, word_crf=1.331, length=2.471, ppl=14.09, wps=20459.5, ups=2.9, wpb=7047, bsz=291, num_updates=54500, lr=0.000214176, gnorm=1.561, train_wall=34, gb_free=9.3, wall=19368
2021-03-28 01:24:15 | INFO | train_inner | epoch 093:   504 / 588 loss=3.953, nll_loss=3.091, word_ins=4.58, word_crf=1.405, length=2.579, ppl=15.48, wps=20163.2, ups=2.87, wpb=7030.8, bsz=267.9, num_updates=54600, lr=0.00021398, gnorm=1.916, train_wall=35, gb_free=9.8, wall=19403
2021-03-28 01:24:44 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 01:24:47 | INFO | valid | epoch 093 | valid on 'valid' subset | loss 8.339 | nll_loss 5.86 | word_ins 7.125 | word_crf 4.239 | length 5.374 | ppl 323.76 | wps 59893.6 | wpb 5311.6 | bsz 208.1 | num_updates 54684 | best_loss 6.867
2021-03-28 01:24:47 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 93 @ 54684 updates

```
2021-03-28 01:24:47 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint93.pt
2021-03-28 01:24:48 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint93.pt
2021-03-28 01:24:51 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint93.pt (epoch 93 @ 54684 updates, score 8.339) (writing took 3.646518472000025 seconds)
2021-03-28 01:24:51 | INFO | fairseq_cli.train | end of epoch 93 (average epoch stats below)
2021-03-28 01:24:51 | INFO | train | epoch 093 | loss 3.888 | nll_loss 3.045 | word_ins 4.541 | word_crf 1.369 | length 2.491 | ppl 14.81 | wps 19377.6 | ups 2.77 | wpb 6988.7 | bsz 272.5 |
num_updates 54684 | lr 0.000213816 | gnorm 1.771 | train_wall 203 | gb_free 9.5 | wall 19439
2021-03-28 01:24:51 | INFO | fairseq.trainer | begin training epoch 94
2021-03-28 01:24:51 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 01:24:56 | INFO | train_inner | epoch 094:   16 / 588 loss=3.802, nll_loss=2.944, word_ins=4.452, word_crf=1.316, length=2.604, ppl=13.95, wps=16677.6, ups=2.41, wpb=6933.5,
bsz=282.8, num_updates=54700, lr=0.000213785, gnorm=1.582, train_wall=34, gb_free=10.6, wall=19444
2021-03-28 01:25:32 | INFO | train_inner | epoch 094:  116 / 588 loss=3.892, nll_loss=3.049, word_ins=4.548, word_crf=1.364, length=2.543, ppl=14.84, wps=20282.1, ups=2.82,
wpb=7184.9, bsz=256.5, num_updates=54800, lr=0.000213589, gnorm=1.695, train_wall=35, gb_free=9.3, wall=19480
2021-03-28 01:26:07 | INFO | train_inner | epoch 094:  216 / 588 loss=3.838, nll_loss=3.005, word_ins=4.507, word_crf=1.345, length=2.39, ppl=14.3, wps=19735.5, ups=2.86, wpb=6910.9,
bsz=280.3, num_updates=54900, lr=0.000213395, gnorm=1.826, train_wall=35, gb_free=10.1, wall=19515
2021-03-28 01:26:42 | INFO | train_inner | epoch 094:  316 / 588 loss=3.992, nll_loss=3.168, word_ins=4.651, word_crf=1.427, length=2.399, ppl=15.91, wps=19682.8, ups=2.85,
wpb=6906.7, bsz=239.8, num_updates=55000, lr=0.000213201, gnorm=1.849, train_wall=35, gb_free=9.6, wall=19550
2021-03-28 01:27:16 | INFO | train_inner | epoch 094:  416 / 588 loss=3.708, nll_loss=2.837, word_ins=4.361, word_crf=1.255, length=2.73, ppl=13.07, wps=20607.4, ups=2.93, wpb=7037.6,
bsz=305, num_updates=55100, lr=0.000213007, gnorm=1.507, train_wall=34, gb_free=9.4, wall=19584
2021-03-28 01:27:51 | INFO | train_inner | epoch 094:  516 / 588 loss=4.024, nll_loss=3.188, word_ins=4.666, word_crf=1.443, length=2.48, ppl=16.27, wps=19907.2, ups=2.83, wpb=7040.9,
bsz=246.6, num_updates=55200, lr=0.000212814, gnorm=1.638, train_wall=35, gb_free=9.5, wall=19619
2021-03-28 01:28:16 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 01:28:20 | INFO | valid | epoch 094 | valid on 'valid' subset | loss 8.398 | nll_loss 5.895 | word_ins 7.159 | word_crf 4.279 | length 5.393 | ppl 337.28 | wps 59842 | wpb 5311.6 |
bsz 208.1 | num_updates 55272 | best_loss 6.867
2021-03-28 01:28:20 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 94 @ 55272 updates
2021-03-28 01:28:20 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint94.pt
2021-03-28 01:28:21 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint94.pt
2021-03-28 01:28:23 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint94.pt (epoch 94 @ 55272 updates, score 8.398) (writing took 3.6476277360015956 seconds)
2021-03-28 01:28:23 | INFO | fairseq_cli.train | end of epoch 94 (average epoch stats below)
2021-03-28 01:28:23 | INFO | train | epoch 094 | loss 3.863 | nll_loss 3.021 | word_ins 4.521 | word_crf 1.353 | length 2.494 | ppl 14.55 | wps 19331.4 | ups 2.77 | wpb 6988.7 | bsz 272.5 |
num_updates 55272 | lr 0.000212675 | gnorm 1.699 | train_wall 204 | gb_free 9.2 | wall 19651
2021-03-28 01:28:24 | INFO | fairseq.trainer | begin training epoch 95
2021-03-28 01:28:24 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 01:28:34 | INFO | train_inner | epoch 095:   28 / 588 loss=3.766, nll_loss=2.918, word_ins=4.431, word_crf=1.306, length=2.447, ppl=13.61, wps=16806.6, ups=2.38, wpb=7070.7,
bsz=300.9, num_updates=55300, lr=0.000212622, gnorm=1.693, train_wall=35, gb_free=9.2, wall=19661
2021-03-28 01:29:08 | INFO | train_inner | epoch 095:  128 / 588 loss=3.74, nll_loss=2.904, word_ins=4.424, word_crf=1.273, length=2.554, ppl=13.36, wps=20174.2, ups=2.9, wpb=6965.3,
bsz=270.6, num_updates=55400, lr=0.00021243, gnorm=1.786, train_wall=34, gb_free=9.9, wall=19696
2021-03-28 01:29:43 | INFO | train_inner | epoch 095:  228 / 588 loss=3.81, nll_loss=2.988, word_ins=4.496, word_crf=1.317, length=2.444, ppl=14.03, wps=19832.3, ups=2.89, wpb=6860.6,
bsz=261.8, num_updates=55500, lr=0.000212238, gnorm=1.676, train_wall=34, gb_free=9.4, wall=19731
2021-03-28 01:30:18 | INFO | train_inner | epoch 095:  328 / 588 loss=3.871, nll_loss=3.022, word_ins=4.523, word_crf=1.359, length=2.501, ppl=14.63, wps=19714.4, ups=2.84,
wpb=6947.4, bsz=273.3, num_updates=55600, lr=0.000212047, gnorm=1.773, train_wall=35, gb_free=9.4, wall=19766
2021-03-28 01:30:53 | INFO | train_inner | epoch 095:  428 / 588 loss=4.007, nll_loss=3.183, word_ins=4.66, word_crf=1.448, length=2.285, ppl=16.07, wps=19566.6, ups=2.84, wpb=6895.7,
bsz=253.6, num_updates=55700, lr=0.000211857, gnorm=1.777, train_wall=35, gb_free=10, wall=19801
2021-03-28 01:31:28 | INFO | train_inner | epoch 095:  528 / 588 loss=3.81, nll_loss=2.949, word_ins=4.456, word_crf=1.329, length=2.53, ppl=14.03, wps=20287.8, ups=2.84, wpb=7142.8,
bsz=306.4, num_updates=55800, lr=0.000211667, gnorm=1.653, train_wall=35, gb_free=9.1, wall=19836
2021-03-28 01:31:50 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 01:31:53 | INFO | valid | epoch 095 | valid on 'valid' subset | loss 8.441 | nll_loss 5.922 | word_ins 7.19 | word_crf 4.31 | length 5.362 | ppl 347.63 | wps 59591.9 | wpb 5311.6 |
bsz 208.1 | num_updates 55860 | best_loss 6.867
2021-03-28 01:31:53 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 95 @ 55860 updates
2021-03-28 01:31:53 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint95.pt
2021-03-28 01:31:54 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint95.pt
2021-03-28 01:31:56 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint95.pt (epoch 95 @ 55860 updates, score 8.441) (writing took 3.6585408960017958 seconds)
2021-03-28 01:31:56 | INFO | fairseq_cli.train | end of epoch 95 (average epoch stats below)
2021-03-28 01:31:56 | INFO | train | epoch 095 | loss 3.857 | nll_loss 3.016 | word_ins 4.517 | word_crf 1.35 | length 2.479 | ppl 14.49 | wps 19288.7 | ups 2.76 | wpb 6988.7 | bsz 272.5 |
num_updates 55860 | lr 0.000211553 | gnorm 1.74 | train_wall 204 | gb_free 9.3 | wall 19864
2021-03-28 01:31:57 | INFO | fairseq.trainer | begin training epoch 96
2021-03-28 01:31:57 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 01:32:11 | INFO | train_inner | epoch 096:   40 / 588 loss=3.892, nll_loss=3.046, word_ins=4.543, word_crf=1.367, length=2.539, ppl=14.85, wps=16560, ups=2.35, wpb=7058.6,
bsz=263.9, num_updates=55900, lr=0.000211477, gnorm=1.771, train_wall=35, gb_free=9.3, wall=19879
2021-03-28 01:32:46 | INFO | train_inner | epoch 096:  140 / 588 loss=3.803, nll_loss=2.982, word_ins=4.49, word_crf=1.321, length=2.372, ppl=13.96, wps=19755.5, ups=2.82, wpb=6995.3,
bsz=269.8, num_updates=56000, lr=0.000211289, gnorm=1.674, train_wall=35, gb_free=9.2, wall=19914
2021-03-28 01:33:21 | INFO | train_inner | epoch 096:  240 / 588 loss=3.658, nll_loss=2.819, word_ins=4.345, word_crf=1.248, length=2.374, ppl=12.63, wps=19945.9, ups=2.88,
wpb=6934.4, bsz=319, num_updates=56100, lr=0.0002111, gnorm=1.866, train_wall=34, gb_free=9.5, wall=19949
2021-03-28 01:33:56 | INFO | train_inner | epoch 096:  340 / 588 loss=3.787, nll_loss=2.93, word_ins=4.446, word_crf=1.301, length=2.631, ppl=13.8, wps=20153.8, ups=2.86, wpb=7040.9,
bsz=271, num_updates=56200, lr=0.000210912, gnorm=1.668, train_wall=35, gb_free=10, wall=19984
2021-03-28 01:34:31 | INFO | train_inner | epoch 096:  440 / 588 loss=3.912, nll_loss=3.057, word_ins=4.554, word_crf=1.376, length=2.594, ppl=15.06, wps=20088.4, ups=2.84,
wpb=7080.5, bsz=264.3, num_updates=56300, lr=0.000210725, gnorm=1.784, train_wall=35, gb_free=9.4, wall=20019
2021-03-28 01:35:07 | INFO | train_inner | epoch 096:  540 / 588 loss=4.007, nll_loss=3.18, word_ins=4.658, word_crf=1.436, length=2.429, ppl=16.08, wps=19574.4, ups=2.82, wpb=6934.2,
bsz=244.8, num_updates=56400, lr=0.000210538, gnorm=1.629, train_wall=35, gb_free=9.9, wall=20055
2021-03-28 01:35:23 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 01:35:27 | INFO | valid | epoch 096 | valid on 'valid' subset | loss 8.4 | nll_loss 5.887 | word_ins 7.153 | word_crf 4.268 | length 5.561 | ppl 337.9 | wps 59579.1 | wpb 5311.6 |
bsz 208.1 | num_updates 56448 | best_loss 6.867
2021-03-28 01:35:27 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 96 @ 56448 updates
2021-03-28 01:35:27 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint96.pt
2021-03-28 01:35:28 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint96.pt
2021-03-28 01:35:30 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint96.pt (epoch 96 @ 56448 updates, score 8.4) (writing took 3.644671695998113 seconds)
2021-03-28 01:35:30 | INFO | fairseq_cli.train | end of epoch 96 (average epoch stats below)
2021-03-28 01:35:30 | INFO | train | epoch 096 | loss 3.84 | nll_loss 3.002 | word_ins 4.506 | word_crf 1.34 | length 2.465 | ppl 14.32 | wps 19206.9 | ups 2.75 | wpb 6988.7 | bsz 272.5 |
num_updates 56448 | lr 0.000210448 | gnorm 1.737 | train_wall 205 | gb_free 9.7 | wall 20078
2021-03-28 01:35:31 | INFO | fairseq.trainer | begin training epoch 97
2021-03-28 01:35:31 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 01:35:49 | INFO | train_inner | epoch 097:   52 / 588 loss=3.819, nll_loss=2.989, word_ins=4.495, word_crf=1.326, length=2.456, ppl=14.11, wps=16388, ups=2.38, wpb=6887.8,
bsz=262.8, num_updates=56500, lr=0.000210352, gnorm=1.783, train_wall=35, gb_free=9.2, wall=20097
2021-03-28 01:36:24 | INFO | train_inner | epoch 097:  152 / 588 loss=3.84, nll_loss=3.006, word_ins=4.512, word_crf=1.334, length=2.506, ppl=14.32, wps=19991.3, ups=2.84, wpb=7042.6,
bsz=264.5, num_updates=56600, lr=0.000210166, gnorm=1.704, train_wall=35, gb_free=9.3, wall=20132
2021-03-28 01:36:59 | INFO | train_inner | epoch 097:  252 / 588 loss=3.857, nll_loss=3.042, word_ins=4.544, word_crf=1.346, length=2.396, ppl=14.49, wps=19846.8, ups=2.85,
wpb=6972.3, bsz=248.2, num_updates=56700, lr=0.00020998, gnorm=1.652, train_wall=35, gb_free=10, wall=20167
2021-03-28 01:37:34 | INFO | train_inner | epoch 097:  352 / 588 loss=3.856, nll_loss=3.021, word_ins=4.522, word_crf=1.359, length=2.354, ppl=14.48, wps=19893.1, ups=2.84,
wpb=7013.1, bsz=280.2, num_updates=56800, lr=0.000209795, gnorm=1.767, train_wall=35, gb_free=9.4, wall=20202
```

2021-03-28 01:38:09 | INFO | train_inner | epoch 097: 452 / 588 loss=3.812, nll_loss=2.955, word_ins=4.465, word_crf=1.324, length=2.556, ppl=14.05, wps=20417, ups=2.89, wpb=7071.8, bsz=287.3, num_updates=56900, lr=0.000209611, gnorm=1.839, train_wall=34, gb_free=9.5, wall=20237
2021-03-28 01:38:44 | INFO | train_inner | epoch 097: 552 / 588 loss=3.785, nll_loss=2.932, word_ins=4.442, word_crf=1.312, length=2.515, ppl=13.78, wps=20230.8, ups=2.89, wpb=6997.4, bsz=295.8, num_updates=57000, lr=0.000209427, gnorm=1.681, train_wall=34, gb_free=9.4, wall=20272
2021-03-28 01:38:56 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 01:39:00 | INFO | valid | epoch 097 | valid on 'valid' subset | loss 8.461 | nll_loss 5.92 | word_ins 7.185 | word_crf 4.32 | length 5.48 | ppl 352.27 | wps 59833.2 | wpb 5311.6 | bsz 208.1 | num_updates 57036 | best_loss 6.867
2021-03-28 01:39:00 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 97 @ 57036 updates
2021-03-28 01:39:00 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint97.pt
2021-03-28 01:39:01 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint97.pt
2021-03-28 01:39:03 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint97.pt (epoch 97 @ 57036 updates, score 8.461) (writing took 3.651389262999146 seconds)
2021-03-28 01:39:03 | INFO | fairseq_cli.train | end of epoch 97 (average epoch stats below)
2021-03-28 01:39:03 | INFO | train | epoch 097 | loss 3.823 | nll_loss 2.986 | word_ins 4.492 | word_crf 1.33 | length 2.464 | ppl 14.15 | wps 19304.3 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 57036 | lr 0.000209361 | gnorm 1.73 | train_wall 204 | gb_free 9.3 | wall 20291
2021-03-28 01:39:03 | INFO | fairseq.trainer | begin training epoch 98
2021-03-28 01:39:03 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 01:39:26 | INFO | train_inner | epoch 098: 64 / 588 loss=3.792, nll_loss=2.977, word_ins=4.482, word_crf=1.326, length=2.257, ppl=13.86, wps=16076.4, ups=2.35, wpb=6843.3, bsz=284.4, num_updates=57100, lr=0.000209243, gnorm=1.819, train_wall=35, gb_free=9, wall=20314
2021-03-28 01:40:02 | INFO | train_inner | epoch 098: 164 / 588 loss=3.88, nll_loss=3.056, word_ins=4.555, word_crf=1.362, length=2.409, ppl=14.72, wps=19692.8, ups=2.82, wpb=6976.7, bsz=254, num_updates=57200, lr=0.000209061, gnorm=1.823, train_wall=35, gb_free=9.3, wall=20350
2021-03-28 01:40:37 | INFO | train_inner | epoch 098: 264 / 588 loss=3.848, nll_loss=3.021, word_ins=4.526, word_crf=1.34, length=2.445, ppl=14.4, wps=20004.2, ups=2.85, wpb=7014.9, bsz=256.9, num_updates=57300, lr=0.000208878, gnorm=1.68, train_wall=35, gb_free=9.7, wall=20385
2021-03-28 01:41:12 | INFO | train_inner | epoch 098: 364 / 588 loss=3.827, nll_loss=3.008, word_ins=4.515, word_crf=1.339, length=2.306, ppl=14.19, wps=19420.9, ups=2.85, wpb=6814.6, bsz=253.8, num_updates=57400, lr=0.000208696, gnorm=1.787, train_wall=35, gb_free=10.6, wall=20420
2021-03-28 01:41:46 | INFO | train_inner | epoch 098: 464 / 588 loss=3.687, nll_loss=2.823, word_ins=4.353, word_crf=1.238, length=2.731, ppl=12.88, wps=20544.1, ups=2.91, wpb=7050.5, bsz=286.7, num_updates=57500, lr=0.000208514, gnorm=1.594, train_wall=34, gb_free=9.8, wall=20454
2021-03-28 01:42:21 | INFO | train_inner | epoch 098: 564 / 588 loss=3.843, nll_loss=2.996, word_ins=4.5, word_crf=1.346, length=2.475, ppl=14.35, wps=20344.8, ups=2.87, wpb=7085, bsz=281, num_updates=57600, lr=0.000208333, gnorm=1.662, train_wall=35, gb_free=9.5, wall=20489
2021-03-28 01:42:29 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 01:42:32 | INFO | valid | epoch 098 | valid on 'valid' subset | loss 8.479 | nll_loss 5.937 | word_ins 7.203 | word_crf 4.335 | length 5.422 | ppl 356.72 | wps 59676.7 | wpb 5311.6 | bsz 208.1 | num_updates 57624 | best_loss 6.867
2021-03-28 01:42:32 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 98 @ 57624 updates
2021-03-28 01:42:32 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint98.pt
2021-03-28 01:42:33 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint98.pt
2021-03-28 01:42:36 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint98.pt (epoch 98 @ 57624 updates, score 8.479) (writing took 3.638825256999553 seconds)
2021-03-28 01:42:36 | INFO | fairseq_cli.train | end of epoch 98 (average epoch stats below)
2021-03-28 01:42:36 | INFO | train | epoch 098 | loss 3.802 | nll_loss 2.969 | word_ins 4.479 | word_crf 1.319 | length 2.436 | ppl 13.95 | wps 19307 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 57624 | lr 0.00020829 | gnorm 1.718 | train_wall 204 | gb_free 9.5 | wall 20504
2021-03-28 01:42:36 | INFO | fairseq.trainer | begin training epoch 99
2021-03-28 01:42:36 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 01:43:03 | INFO | train_inner | epoch 099: 76 / 588 loss=3.728, nll_loss=2.896, word_ins=4.415, word_crf=1.272, length=2.482, ppl=13.25, wps=16956.2, ups=2.4, wpb=7058.8, bsz=277.8, num_updates=57700, lr=0.000208153, gnorm=1.856, train_wall=34, gb_free=10.1, wall=20530
2021-03-28 01:43:38 | INFO | train_inner | epoch 099: 176 / 588 loss=3.783, nll_loss=2.951, word_ins=4.463, word_crf=1.315, length=2.366, ppl=13.76, wps=19860.7, ups=2.85, wpb=6969.2, bsz=282.2, num_updates=57800, lr=0.000207973, gnorm=1.834, train_wall=35, gb_free=9.5, wall=20566
2021-03-28 01:44:13 | INFO | train_inner | epoch 099: 276 / 588 loss=3.8, nll_loss=2.964, word_ins=4.474, word_crf=1.323, length=2.397, ppl=13.93, wps=20046.2, ups=2.85, wpb=7034.4, bsz=280.7, num_updates=57900, lr=0.000207793, gnorm=1.779, train_wall=35, gb_free=9.9, wall=20601
2021-03-28 01:44:48 | INFO | train_inner | epoch 099: 376 / 588 loss=3.85, nll_loss=3.034, word_ins=4.537, word_crf=1.345, length=2.373, ppl=14.42, wps=19924.7, ups=2.87, wpb=6952.1, bsz=248.3, num_updates=58000, lr=0.000207614, gnorm=1.844, train_wall=35, gb_free=9.7, wall=20636
2021-03-28 01:45:23 | INFO | train_inner | epoch 099: 476 / 588 loss=3.953, nll_loss=3.116, word_ins=4.604, word_crf=1.402, length=2.49, ppl=15.49, wps=19725.2, ups=2.82, wpb=6999, bsz=257.7, num_updates=58100, lr=0.000207435, gnorm=1.79, train_wall=35, gb_free=9.3, wall=20671
2021-03-28 01:45:57 | INFO | train_inner | epoch 099: 576 / 588 loss=3.645, nll_loss=2.786, word_ins=4.319, word_crf=1.229, length=2.563, ppl=12.51, wps=20541.9, ups=2.93, wpb=7005, bsz=305.7, num_updates=58200, lr=0.000207257, gnorm=1.655, train_wall=34, gb_free=9.5, wall=20705
2021-03-28 01:46:02 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 01:46:05 | INFO | valid | epoch 099 | valid on 'valid' subset | loss 8.532 | nll_loss 5.963 | word_ins 7.234 | word_crf 4.374 | length 5.411 | ppl 370.17 | wps 59921.6 | wpb 5311.6 | bsz 208.1 | num_updates 58212 | best_loss 6.867
2021-03-28 01:46:05 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 99 @ 58212 updates
2021-03-28 01:46:05 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint99.pt
2021-03-28 01:46:06 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint99.pt
2021-03-28 01:46:08 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint99.pt (epoch 99 @ 58212 updates, score 8.532) (writing took 3.638975202000438 seconds)
2021-03-28 01:46:08 | INFO | fairseq_cli.train | end of epoch 99 (average epoch stats below)
2021-03-28 01:46:08 | INFO | train | epoch 099 | loss 3.804 | nll_loss 2.969 | word_ins 4.479 | word_crf 1.32 | length 2.45 | ppl 13.97 | wps 19355.8 | ups 2.77 | wpb 6988.7 | bsz 272.5 | num_updates 58212 | lr 0.000207235 | gnorm 1.797 | train_wall 204 | gb_free 9.4 | wall 20716
2021-03-28 01:46:09 | INFO | fairseq.trainer | begin training epoch 100
2021-03-28 01:46:09 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 01:46:39 | INFO | train_inner | epoch 100: 88 / 588 loss=3.726, nll_loss=2.88, word_ins=4.4, word_crf=1.273, length=2.532, ppl=13.23, wps=16739.4, ups=2.39, wpb=7012.6, bsz=292.7, num_updates=58300, lr=0.000207079, gnorm=1.872, train_wall=34, gb_free=9.8, wall=20747
2021-03-28 01:47:14 | INFO | train_inner | epoch 100: 188 / 588 loss=3.788, nll_loss=2.959, word_ins=4.472, word_crf=1.317, length=2.353, ppl=13.81, wps=20003.5, ups=2.85, wpb=7020.7, bsz=270, num_updates=58400, lr=0.000206901, gnorm=1.999, train_wall=35, gb_free=9.3, wall=20782
2021-03-28 01:47:49 | INFO | train_inner | epoch 100: 288 / 588 loss=3.72, nll_loss=2.88, word_ins=4.401, word_crf=1.269, length=2.5, ppl=13.18, wps=20397.5, ups=2.9, wpb=7028.2, bsz=283.4, num_updates=58500, lr=0.000206725, gnorm=1.91, train_wall=34, gb_free=9.5, wall=20817
2021-03-28 01:48:24 | INFO | train_inner | epoch 100: 388 / 588 loss=3.828, nll_loss=2.994, word_ins=4.501, word_crf=1.33, length=2.473, ppl=14.2, wps=20150.9, ups=2.87, wpb=7019.1, bsz=266.8, num_updates=58600, lr=0.000206548, gnorm=1.68, train_wall=35, gb_free=9.5, wall=20851
2021-03-28 01:48:59 | INFO | train_inner | epoch 100: 488 / 588 loss=3.888, nll_loss=3.063, word_ins=4.56, word_crf=1.375, length=2.335, ppl=14.81, wps=19451.9, ups=2.83, wpb=6883.6, bsz=263.4, num_updates=58700, lr=0.000206372, gnorm=1.85, train_wall=35, gb_free=9.4, wall=20887
2021-03-28 01:49:34 | INFO | train_inner | epoch 100: 588 / 588 loss=3.903, nll_loss=3.086, word_ins=4.58, word_crf=1.381, length=2.316, ppl=14.95, wps=19630.6, ups=2.84, wpb=6900.5, bsz=252.7, num_updates=58800, lr=0.000206197, gnorm=1.657, train_wall=35, gb_free=9.8, wall=20922
2021-03-28 01:49:34 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 01:49:37 | INFO | valid | epoch 100 | valid on 'valid' subset | loss 8.529 | nll_loss 5.963 | word_ins 7.228 | word_crf 4.363 | length 5.514 | ppl 369.29 | wps 59998.3 | wpb 5311.6 | bsz 208.1 | num_updates 58800 | best_loss 6.867
2021-03-28 01:49:37 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 100 @ 58800 updates
2021-03-28 01:49:37 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint100.pt
2021-03-28 01:49:38 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint100.pt
2021-03-28 01:49:41 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint100.pt (epoch 100 @ 58800 updates, score 8.529) (writing took 3.646880718999455 seconds)
2021-03-28 01:49:41 | INFO | fairseq_cli.train | end of epoch 100 (average epoch stats below)
2021-03-28 01:49:41 | INFO | train | epoch 100 | loss 3.804 | nll_loss 2.972 | word_ins 4.481 | word_crf 1.322 | length 2.419 | ppl 13.97 | wps 19335.4 | ups 2.77 | wpb 6988.7 | bsz 272.5 | num_updates 58800 | lr 0.000206197 | gnorm 1.832 | train_wall 204 | gb_free 9.8 | wall 20929
2021-03-28 01:49:41 | INFO | fairseq.trainer | begin training epoch 101

2021-03-28 01:49:41 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 01:50:17 | INFO | train_inner | epoch 101:   100 / 588 loss=3.787, nll_loss=2.98, word_ins=4.491, word_crf=1.315, length=2.272, ppl=13.81, wps=16050.6, ups=2.33, wpb=6887.4, bsz=263.5, num_updates=58900, lr=0.000206021, gnorm=1.789, train_wall=36, gb_free=9.9, wall=20965
2021-03-28 01:50:52 | INFO | train_inner | epoch 101:   200 / 588 loss=3.733, nll_loss=2.923, word_ins=4.443, word_crf=1.279, length=2.326, ppl=13.29, wps=19876.9, ups=2.84, wpb=6992.2, bsz=262.9, num_updates=59000, lr=0.000205847, gnorm=1.827, train_wall=35, gb_free=9.4, wall=21000
2021-03-28 01:51:28 | INFO | train_inner | epoch 101:   300 / 588 loss=3.835, nll_loss=3.008, word_ins=4.512, word_crf=1.346, length=2.32, ppl=14.27, wps=19256.5, ups=2.81, wpb=6843.5, bsz=270.7, num_updates=59100, lr=0.000205673, gnorm=1.94, train_wall=35, gb_free=9.1, wall=21036
2021-03-28 01:52:02 | INFO | train_inner | epoch 101:   400 / 588 loss=3.735, nll_loss=2.89, word_ins=4.41, word_crf=1.277, length=2.528, ppl=13.32, wps=20192.4, ups=2.89, wpb=6977.6, bsz=281.1, num_updates=59200, lr=0.000205499, gnorm=1.773, train_wall=34, gb_free=9.6, wall=21070
2021-03-28 01:52:37 | INFO | train_inner | epoch 101:   500 / 588 loss=3.857, nll_loss=3.03, word_ins=4.532, word_crf=1.349, length=2.425, ppl=14.49, wps=20078.3, ups=2.85, wpb=7038.4, bsz=257.9, num_updates=59300, lr=0.000205325, gnorm=1.686, train_wall=35, gb_free=9, wall=21105
2021-03-28 01:53:07 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 01:53:11 | INFO | valid | epoch 101 | valid on 'valid' subset | loss 8.502 | nll_loss 5.934 | word_ins 7.204 | word_crf 4.342 | length 5.583 | ppl 362.57 | wps 59392 | wpb 5311.6 | bsz 208.1 | num_updates 59388 | best_loss 6.867
2021-03-28 01:53:11 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 101 @ 59388 updates
2021-03-28 01:53:11 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint101.pt
2021-03-28 01:53:12 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint101.pt
2021-03-28 01:53:14 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint101.pt (epoch 101 @ 59388 updates, score 8.502) (writing took 3.670162133003032 seconds)
2021-03-28 01:53:14 | INFO | fairseq_cli.train | end of epoch 101 (average epoch stats below)
2021-03-28 01:53:14 | INFO | train | epoch 101 | loss 3.778 | nll_loss 2.949 | word_ins 4.463 | word_crf 1.306 | length 2.402 | ppl 13.72 | wps 19258.5 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 59388 | lr 0.000205173 | gnorm 1.792 | train_wall 205 | gb_free 9.3 | wall 21142
2021-03-28 01:53:14 | INFO | fairseq.trainer | begin training epoch 102
2021-03-28 01:53:14 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 01:53:19 | INFO | train_inner | epoch 102:   12 / 588 loss=3.651, nll_loss=2.792, word_ins=4.325, word_crf=1.233, length=2.556, ppl=12.56, wps=17410.3, ups=2.41, wpb=7213.1, bsz=312.2, num_updates=59400, lr=0.000205152, gnorm=1.683, train_wall=34, gb_free=10.6, wall=21147
2021-03-28 01:53:53 | INFO | train_inner | epoch 102:   112 / 588 loss=3.629, nll_loss=2.799, word_ins=4.338, word_crf=1.212, length=2.478, ppl=12.37, wps=20464, ups=2.88, wpb=7109.5, bsz=277.7, num_updates=59500, lr=0.00020498, gnorm=1.694, train_wall=34, gb_free=9.4, wall=21181
2021-03-28 01:54:29 | INFO | train_inner | epoch 102:   212 / 588 loss=3.861, nll_loss=3.041, word_ins=4.544, word_crf=1.349, length=2.397, ppl=14.53, wps=19859.1, ups=2.85, wpb=6976, bsz=242.2, num_updates=59600, lr=0.000204808, gnorm=2.132, train_wall=35, gb_free=9.5, wall=21216
2021-03-28 01:55:03 | INFO | train_inner | epoch 102:   312 / 588 loss=3.762, nll_loss=2.943, word_ins=4.457, word_crf=1.302, length=2.315, ppl=13.57, wps=19934, ups=2.87, wpb=6939.3, bsz=280.1, num_updates=59700, lr=0.000204636, gnorm=1.813, train_wall=35, gb_free=9.4, wall=21251
2021-03-28 01:55:39 | INFO | train_inner | epoch 102:   412 / 588 loss=3.847, nll_loss=3.026, word_ins=4.528, word_crf=1.35, length=2.325, ppl=14.39, wps=19386.3, ups=2.83, wpb=6847, bsz=264.8, num_updates=59800, lr=0.000204465, gnorm=1.784, train_wall=35, gb_free=9.3, wall=21287
2021-03-28 01:56:14 | INFO | train_inner | epoch 102:   512 / 588 loss=3.807, nll_loss=2.971, word_ins=4.482, word_crf=1.317, length=2.491, ppl=14, wps=20024.5, ups=2.85, wpb=7022.3, bsz=275.4, num_updates=59900, lr=0.000204294, gnorm=1.704, train_wall=35, gb_free=9.4, wall=21322
2021-03-28 01:56:40 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 01:56:44 | INFO | valid | epoch 102 | valid on 'valid' subset | loss 8.561 | nll_loss 5.988 | word_ins 7.253 | word_crf 4.385 | length 5.498 | ppl 377.78 | wps 59660.5 | wpb 5311.6 | bsz 208.1 | num_updates 59976 | best_loss 6.867
2021-03-28 01:56:44 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 102 @ 59976 updates
2021-03-28 01:56:44 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint102.pt
2021-03-28 01:56:44 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint102.pt
2021-03-28 01:56:47 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint102.pt (epoch 102 @ 59976 updates, score 8.561) (writing took 3.644681236000906 seconds)
2021-03-28 01:56:47 | INFO | fairseq_cli.train | end of epoch 102 (average epoch stats below)
2021-03-28 01:56:47 | INFO | train | epoch 102 | loss 3.766 | nll_loss 2.937 | word_ins 4.453 | word_crf 1.299 | length 2.413 | ppl 13.61 | wps 19308.6 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 59976 | lr 0.000204165 | gnorm 1.806 | train_wall 204 | gb_free 9.5 | wall 21355
2021-03-28 01:56:47 | INFO | fairseq.trainer | begin training epoch 103
2021-03-28 01:56:47 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 01:56:56 | INFO | train_inner | epoch 103:   24 / 588 loss=3.764, nll_loss=2.927, word_ins=4.443, word_crf=1.297, length=2.46, ppl=13.59, wps=16822.6, ups=2.38, wpb=7072.1, bsz=272.2, num_updates=60000, lr=0.000204124, gnorm=1.728, train_wall=34, gb_free=9.7, wall=21364
2021-03-28 01:56:56 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 01:56:59 | INFO | valid | epoch 103 | valid on 'valid' subset | loss 8.575 | nll_loss 5.978 | word_ins 7.248 | word_crf 4.394 | length 5.572 | ppl 381.3 | wps 59842.5 | wpb 5311.6 | bsz 208.1 | num_updates 60000 | best_loss 6.867
2021-03-28 01:56:59 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 103 @ 60000 updates
2021-03-28 01:56:59 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint_103_60000.pt
2021-03-28 01:57:00 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint_103_60000.pt
2021-03-28 01:57:03 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint_103_60000.pt (epoch 103 @ 60000 updates, score 8.575) (writing took 3.9977693149994593 seconds)
2021-03-28 01:57:38 | INFO | train_inner | epoch 103:   124 / 588 loss=3.466, nll_loss=2.64, word_ins=4.2, word_crf=1.127, length=2.395, ppl=11.05, wps=16697.2, ups=2.39, wpb=6972.1, bsz=301.6, num_updates=60100, lr=0.000203954, gnorm=1.58, train_wall=34, gb_free=9.2, wall=21405
2021-03-28 01:58:12 | INFO | train_inner | epoch 103:   224 / 588 loss=3.696, nll_loss=2.858, word_ins=4.384, word_crf=1.26, length=2.447, ppl=12.96, wps=20472, ups=2.9, wpb=7058.4, bsz=289.7, num_updates=60200, lr=0.000203785, gnorm=1.87, train_wall=34, gb_free=9.5, wall=21440
2021-03-28 01:58:47 | INFO | train_inner | epoch 103:   324 / 588 loss=3.72, nll_loss=2.925, word_ins=4.446, word_crf=1.269, length=2.273, ppl=13.17, wps=19443.2, ups=2.86, wpb=6797, bsz=253.5, num_updates=60300, lr=0.000203616, gnorm=1.699, train_wall=35, gb_free=9.5, wall=21475
2021-03-28 01:59:23 | INFO | train_inner | epoch 103:   424 / 588 loss=3.997, nll_loss=3.169, word_ins=4.652, word_crf=1.434, length=2.368, ppl=15.96, wps=19803.1, ups=2.81, wpb=7041.4, bsz=243.4, num_updates=60400, lr=0.000203447, gnorm=1.835, train_wall=35, gb_free=9.3, wall=21510
2021-03-28 01:59:58 | INFO | train_inner | epoch 103:   524 / 588 loss=3.814, nll_loss=2.981, word_ins=4.488, word_crf=1.325, length=2.455, ppl=14.07, wps=20118.4, ups=2.86, wpb=7037.6, bsz=269.5, num_updates=60500, lr=0.000203279, gnorm=1.852, train_wall=35, gb_free=9.3, wall=21545
2021-03-28 02:00:21 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 02:00:24 | INFO | valid | epoch 103 | valid on 'valid' subset | loss 8.512 | nll_loss 5.95 | word_ins 7.213 | word_crf 4.351 | length 5.544 | ppl 365.03 | wps 59724.1 | wpb 5311.6 | bsz 208.1 | num_updates 60564 | best_loss 6.867
2021-03-28 02:00:24 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 103 @ 60564 updates
2021-03-28 02:00:24 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint103.pt
2021-03-28 02:00:25 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint103.pt
2021-03-28 02:00:28 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint103.pt (epoch 103 @ 60564 updates, score 8.512) (writing took 3.696486354998342 seconds)
2021-03-28 02:00:28 | INFO | fairseq_cli.train | end of epoch 103 (average epoch stats below)
2021-03-28 02:00:28 | INFO | train | epoch 103 | loss 3.753 | nll_loss 2.927 | word_ins 4.445 | word_crf 1.292 | length 2.384 | ppl 13.48 | wps 18615.6 | ups 2.66 | wpb 6988.7 | bsz 272.5 | num_updates 60564 | lr 0.000203171 | gnorm 1.787 | train_wall 205 | gb_free 9.7 | wall 21576
2021-03-28 02:00:28 | INFO | fairseq.trainer | begin training epoch 104
2021-03-28 02:00:28 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 02:00:41 | INFO | train_inner | epoch 104:   36 / 588 loss=3.818, nll_loss=2.983, word_ins=4.488, word_crf=1.344, length=2.291, ppl=14.1, wps=16012.3, ups=2.3, wpb=6959.4, bsz=288.1, num_updates=60600, lr=0.000203111, gnorm=1.921, train_wall=36, gb_free=9.4, wall=21589
2021-03-28 02:01:16 | INFO | train_inner | epoch 104:   136 / 588 loss=3.605, nll_loss=2.788, word_ins=4.328, word_crf=1.203, length=2.387, ppl=12.17, wps=20558.8, ups=2.9, wpb=7100.2, bsz=279.9, num_updates=60700, lr=0.000202944, gnorm=1.57, train_wall=34, gb_free=9.5, wall=21623
2021-03-28 02:01:50 | INFO | train_inner | epoch 104:   236 / 588 loss=3.698, nll_loss=2.872, word_ins=4.399, word_crf=1.259, length=2.399, ppl=12.98, wps=20548.6, ups=2.88, wpb=7133.5, bsz=282.1, num_updates=60800, lr=0.000202777, gnorm=1.806, train_wall=34, gb_free=9.7, wall=21658

2021-03-28 02:02:25 | INFO | train_inner | epoch 104: 336 / 588 loss=3.79, nll_loss=2.981, word_ins=4.494, word_crf=1.313, length=2.297, ppl=13.83, wps=19441, ups=2.85, wpb=6814.7, bsz=252.7, num_updates=60900, lr=0.00020261, gnorm=1.794, train_wall=35, gb_free=9.3, wall=21693
2021-03-28 02:03:00 | INFO | train_inner | epoch 104: 436 / 588 loss=3.704, nll_loss=2.879, word_ins=4.406, word_crf=1.253, length=2.475, ppl=13.03, wps=19946, ups=2.87, wpb=6957.8, bsz=266.4, num_updates=61000, lr=0.000202444, gnorm=1.531, train_wall=35, gb_free=10, wall=21728
2021-03-28 02:03:35 | INFO | train_inner | epoch 104: 536 / 588 loss=3.815, nll_loss=2.981, word_ins=4.489, word_crf=1.329, length=2.421, ppl=14.08, wps=20094.6, ups=2.85, wpb=7043.3, bsz=267.5, num_updates=61100, lr=0.000202278, gnorm=1.723, train_wall=35, gb_free=9.8, wall=21763
2021-03-28 02:03:53 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 02:03:56 | INFO | valid | epoch 104 | valid on 'valid' subset | loss 8.581 | nll_loss 5.985 | word_ins 7.245 | word_crf 4.4 | length 5.58 | ppl 382.87 | wps 60242.2 | wpb 5311.6 | bsz 208.1 | num_updates 61152 | best_loss 6.867
2021-03-28 02:03:56 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 104 @ 61152 updates
2021-03-28 02:03:56 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint104.pt
2021-03-28 02:03:57 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint104.pt
2021-03-28 02:04:00 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint104.pt (epoch 104 @ 61152 updates, score 8.581) (writing took 3.6381222210002306 seconds)
2021-03-28 02:04:00 | INFO | fairseq_cli.train | end of epoch 104 (average epoch stats below)
2021-03-28 02:04:00 | INFO | train | epoch 104 | loss 3.723 | nll_loss 2.901 | word_ins 4.423 | word_crf 1.275 | length 2.372 | ppl 13.21 | wps 19383.1 | ups 2.77 | wpb 6988.7 | bsz 272.5 | num_updates 61152 | lr 0.000202192 | gnorm 1.705 | train_wall 203 | gb_free 9.5 | wall 21788
2021-03-28 02:04:00 | INFO | fairseq.trainer | begin training epoch 105
2021-03-28 02:04:00 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 02:04:17 | INFO | train_inner | epoch 105: 48 / 588 loss=3.667, nll_loss=2.849, word_ins=4.373, word_crf=1.256, length=2.252, ppl=12.71, wps=16393.1, ups=2.42, wpb=6784.5, bsz=289.8, num_updates=61200, lr=0.000202113, gnorm=1.841, train_wall=34, gb_free=9.5, wall=21805
2021-03-28 02:04:52 | INFO | train_inner | epoch 105: 148 / 588 loss=3.661, nll_loss=2.859, word_ins=4.388, word_crf=1.24, length=2.273, ppl=12.65, wps=19490.6, ups=2.86, wpb=6821.5, bsz=268.9, num_updates=61300, lr=0.000201948, gnorm=1.755, train_wall=35, gb_free=9.4, wall=21840
2021-03-28 02:05:28 | INFO | train_inner | epoch 105: 248 / 588 loss=3.905, nll_loss=3.085, word_ins=4.584, word_crf=1.374, length=2.394, ppl=14.98, wps=19546.8, ups=2.78, wpb=7038.8, bsz=245.7, num_updates=61400, lr=0.000201784, gnorm=1.671, train_wall=36, gb_free=9.1, wall=21876
2021-03-28 02:06:02 | INFO | train_inner | epoch 105: 348 / 588 loss=3.551, nll_loss=2.709, word_ins=4.257, word_crf=1.186, length=2.37, ppl=11.72, wps=20682.3, ups=2.89, wpb=7153.3, bsz=318.8, num_updates=61500, lr=0.000201619, gnorm=1.633, train_wall=34, gb_free=9.7, wall=21910
2021-03-28 02:06:37 | INFO | train_inner | epoch 105: 448 / 588 loss=3.84, nll_loss=3.021, word_ins=4.526, word_crf=1.34, length=2.372, ppl=14.32, wps=19852.6, ups=2.84, wpb=6992, bsz=256.3, num_updates=61600, lr=0.000201456, gnorm=2.005, train_wall=35, gb_free=9.7, wall=21945
2021-03-28 02:07:12 | INFO | train_inner | epoch 105: 548 / 588 loss=3.643, nll_loss=2.813, word_ins=4.347, word_crf=1.224, length=2.452, ppl=12.49, wps=20212.6, ups=2.89, wpb=6996.9, bsz=277.7, num_updates=61700, lr=0.000201292, gnorm=1.634, train_wall=34, gb_free=9.9, wall=21980
2021-03-28 02:07:26 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 02:07:29 | INFO | valid | epoch 105 | valid on 'valid' subset | loss 8.524 | nll_loss 5.956 | word_ins 7.22 | word_crf 4.363 | length 5.511 | ppl 368.16 | wps 59763.1 | wpb 5311.6 | bsz 208.1 | num_updates 61740 | best_loss 6.867
2021-03-28 02:07:29 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 105 @ 61740 updates
2021-03-28 02:07:29 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint105.pt
2021-03-28 02:07:30 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint105.pt
2021-03-28 02:07:33 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint105.pt (epoch 105 @ 61740 updates, score 8.524) (writing took 3.6457533269967826 seconds)
2021-03-28 02:07:33 | INFO | fairseq_cli.train | end of epoch 105 (average epoch stats below)
2021-03-28 02:07:33 | INFO | train | epoch 105 | loss 3.721 | nll_loss 2.9 | word_ins 4.422 | word_crf 1.274 | length 2.364 | ppl 13.19 | wps 19280.6 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 61740 | lr 0.000201227 | gnorm 1.75 | train_wall 204 | gb_free 9.4 | wall 22001
2021-03-28 02:07:33 | INFO | fairseq.trainer | begin training epoch 106
2021-03-28 02:07:33 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 02:07:54 | INFO | train_inner | epoch 106: 60 / 588 loss=3.742, nll_loss=2.927, word_ins=4.447, word_crf=1.285, length=2.336, ppl=13.38, wps=16852.3, ups=2.37, wpb=7117.4, bsz=263, num_updates=61800, lr=0.000201129, gnorm=1.786, train_wall=35, gb_free=10.1, wall=22022
2021-03-28 02:08:29 | INFO | train_inner | epoch 106: 160 / 588 loss=3.713, nll_loss=2.899, word_ins=4.422, word_crf=1.268, length=2.339, ppl=13.11, wps=20165, ups=2.87, wpb=7034.1, bsz=274.9, num_updates=61900, lr=0.000200967, gnorm=1.825, train_wall=35, gb_free=10.1, wall=22057
2021-03-28 02:09:04 | INFO | train_inner | epoch 106: 260 / 588 loss=3.631, nll_loss=2.799, word_ins=4.338, word_crf=1.225, length=2.37, ppl=12.39, wps=19912.5, ups=2.87, wpb=6937.8, bsz=278.3, num_updates=62000, lr=0.000200805, gnorm=1.826, train_wall=35, gb_free=9.6, wall=22092
2021-03-28 02:09:39 | INFO | train_inner | epoch 106: 360 / 588 loss=3.718, nll_loss=2.926, word_ins=4.448, word_crf=1.271, length=2.224, ppl=13.16, wps=19323.6, ups=2.82, wpb=6842.9, bsz=251.7, num_updates=62100, lr=0.000200643, gnorm=1.669, train_wall=35, gb_free=9.6, wall=22127
2021-03-28 02:10:15 | INFO | train_inner | epoch 106: 460 / 588 loss=3.621, nll_loss=2.782, word_ins=4.321, word_crf=1.209, length=2.517, ppl=12.3, wps=19825.4, ups=2.85, wpb=6953.5, bsz=285.8, num_updates=62200, lr=0.000200482, gnorm=1.738, train_wall=35, gb_free=10.3, wall=22162
2021-03-28 02:10:50 | INFO | train_inner | epoch 106: 560 / 588 loss=3.825, nll_loss=2.999, word_ins=4.502, word_crf=1.338, length=2.358, ppl=14.17, wps=20148.9, ups=2.85, wpb=7071.4, bsz=277.7, num_updates=62300, lr=0.000200321, gnorm=1.771, train_wall=35, gb_free=9.5, wall=22198
2021-03-28 02:10:59 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 02:11:03 | INFO | valid | epoch 106 | valid on 'valid' subset | loss 8.545 | nll_loss 5.964 | word_ins 7.226 | word_crf 4.366 | length 5.655 | ppl 373.51 | wps 59720.2 | wpb 5311.6 | bsz 208.1 | num_updates 62328 | best_loss 6.867
2021-03-28 02:11:03 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 106 @ 62328 updates
2021-03-28 02:11:03 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint106.pt
2021-03-28 02:11:03 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint106.pt
2021-03-28 02:11:06 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint106.pt (epoch 106 @ 62328 updates, score 8.545) (writing took 3.642266743001528 seconds)
2021-03-28 02:11:06 | INFO | fairseq_cli.train | end of epoch 106 (average epoch stats below)
2021-03-28 02:11:06 | INFO | train | epoch 106 | loss 3.705 | nll_loss 2.886 | word_ins 4.411 | word_crf 1.265 | length 2.351 | ppl 13.05 | wps 19272.9 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 62328 | lr 0.000200276 | gnorm 1.786 | train_wall 204 | gb_free 9.6 | wall 22214
2021-03-28 02:11:06 | INFO | fairseq.trainer | begin training epoch 107
2021-03-28 02:11:06 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 02:11:32 | INFO | train_inner | epoch 107: 72 / 588 loss=3.707, nll_loss=2.907, word_ins=4.429, word_crf=1.273, length=2.202, ppl=13.06, wps=16255.8, ups=2.37, wpb=6854.1, bsz=259.8, num_updates=62400, lr=0.00020016, gnorm=1.918, train_wall=35, gb_free=10.7, wall=22240
2021-03-28 02:12:07 | INFO | train_inner | epoch 107: 172 / 588 loss=3.672, nll_loss=2.85, word_ins=4.381, word_crf=1.246, length=2.349, ppl=12.75, wps=19958.3, ups=2.83, wpb=7054.2, bsz=279.9, num_updates=62500, lr=0.0002, gnorm=1.903, train_wall=35, gb_free=9.5, wall=22275
2021-03-28 02:12:42 | INFO | train_inner | epoch 107: 272 / 588 loss=3.643, nll_loss=2.831, word_ins=4.363, word_crf=1.237, length=2.245, ppl=12.49, wps=19687.4, ups=2.85, wpb=6898.8, bsz=282.8, num_updates=62600, lr=0.00019984, gnorm=1.852, train_wall=35, gb_free=9.5, wall=22310
2021-03-28 02:13:17 | INFO | train_inner | epoch 107: 372 / 588 loss=3.721, nll_loss=2.893, word_ins=4.418, word_crf=1.268, length=2.439, ppl=13.19, wps=20072.2, ups=2.85, wpb=7039.5, bsz=263.9, num_updates=62700, lr=0.000199681, gnorm=1.88, train_wall=35, gb_free=10, wall=22345
2021-03-28 02:13:52 | INFO | train_inner | epoch 107: 472 / 588 loss=3.72, nll_loss=2.894, word_ins=4.417, word_crf=1.274, length=2.378, ppl=13.18, wps=20131.3, ups=2.87, wpb=7009.1, bsz=277.6, num_updates=62800, lr=0.000199522, gnorm=1.817, train_wall=35, gb_free=9.4, wall=22380
2021-03-28 02:14:27 | INFO | train_inner | epoch 107: 572 / 588 loss=3.809, nll_loss=2.982, word_ins=4.493, word_crf=1.321, length=2.409, ppl=14.01, wps=20262.1, ups=2.85, wpb=7107.3, bsz=263.9, num_updates=62900, lr=0.000199363, gnorm=1.736, train_wall=35, gb_free=9.4, wall=22415
2021-03-28 02:14:33 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 02:14:36 | INFO | valid | epoch 107 | valid on 'valid' subset | loss 8.475 | nll_loss 5.919 | word_ins 7.183 | word_crf 4.32 | length 5.636 | ppl 355.89 | wps 59662.6 | wpb 5311.6 | bsz 208.1 | num_updates 62916 | best_loss 6.867
2021-03-28 02:14:36 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 107 @ 62916 updates
2021-03-28 02:14:36 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint107.pt
2021-03-28 02:14:37 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint107.pt

2021-03-28 02:14:40 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint107.pt (epoch 107 @ 62916 updates, score 8.475) (writing took 3.6438957619975554 seconds)
2021-03-28 02:14:40 | INFO | fairseq_cli.train | end of epoch 107 (average epoch stats below)
2021-03-28 02:14:40 | INFO | train | epoch 107 | loss 3.708 | nll_loss 2.888 | word_ins 4.413 | word_crf 1.267 | length 2.342 | ppl 13.06 | wps 19263.3 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 62916 | lr 0.000199338 | gnorm 1.839 | train_wall 205 | gb_free 9.5 | wall 22428
2021-03-28 02:14:40 | INFO | fairseq.trainer | begin training epoch 108
2021-03-28 02:14:40 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 02:15:09 | INFO | train_inner | epoch 108:   84 / 588 loss=3.697, nll_loss=2.899, word_ins=4.422, word_crf=1.264, length=2.224, ppl=12.97, wps=16667.2, ups=2.37, wpb=7044.7, bsz=268.7, num_updates=63000, lr=0.000199205, gnorm=1.828, train_wall=35, gb_free=9.2, wall=22457
2021-03-28 02:15:44 | INFO | train_inner | epoch 108:   184 / 588 loss=3.545, nll_loss=2.726, word_ins=4.275, word_crf=1.169, length=2.383, ppl=11.67, wps=20427.9, ups=2.89, wpb=7071.2, bsz=291.4, num_updates=63100, lr=0.000199047, gnorm=1.646, train_wall=34, gb_free=9.6, wall=22492
2021-03-28 02:16:19 | INFO | train_inner | epoch 108:   284 / 588 loss=3.732, nll_loss=2.916, word_ins=4.439, word_crf=1.28, length=2.325, ppl=13.29, wps=19744.8, ups=2.85, wpb=6939.1, bsz=262.5, num_updates=63200, lr=0.000198889, gnorm=1.89, train_wall=35, gb_free=9.6, wall=22527
2021-03-28 02:16:54 | INFO | train_inner | epoch 108:   384 / 588 loss=3.747, nll_loss=2.945, word_ins=4.462, word_crf=1.292, length=2.245, ppl=13.43, wps=19882.9, ups=2.87, wpb=6932.3, bsz=264.6, num_updates=63300, lr=0.000198732, gnorm=1.875, train_wall=35, gb_free=9.8, wall=22562
2021-03-28 02:17:29 | INFO | train_inner | epoch 108:   484 / 588 loss=3.596, nll_loss=2.764, word_ins=4.306, word_crf=1.19, length=2.525, ppl=12.09, wps=20505.5, ups=2.9, wpb=7080.2, bsz=285.8, num_updates=63400, lr=0.000198575, gnorm=1.598, train_wall=34, gb_free=9.5, wall=22596
2021-03-28 02:18:04 | INFO | train_inner | epoch 108:   584 / 588 loss=3.794, nll_loss=2.968, word_ins=4.481, word_crf=1.327, length=2.265, ppl=13.87, wps=19504.2, ups=2.84, wpb=6867.6, bsz=264.5, num_updates=63500, lr=0.000198419, gnorm=1.742, train_wall=35, gb_free=9.8, wall=22632
2021-03-28 02:18:05 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 02:18:08 | INFO | valid | epoch 108 | valid on 'valid' subset | loss 8.629 | nll_loss 6.01 | word_ins 7.276 | word_crf 4.424 | length 5.663 | ppl 395.77 | wps 59735.2 | wpb 5311.6 | bsz 208.1 | num_updates 63504 | best_loss 6.867
2021-03-28 02:18:08 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 108 @ 63504 updates
2021-03-28 02:18:08 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint108.pt
2021-03-28 02:18:09 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint108.pt
2021-03-28 02:18:12 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint108.pt (==epoch 108 @ 63504 updates, score 8.629==) (writing took 3.635412161002023 seconds)
2021-03-28 02:18:12 | INFO | fairseq_cli.train | end of epoch 108 (average epoch stats below)
2021-03-28 02:18:12 | INFO | train | epoch 108 | loss 3.681 | nll_loss 2.866 | word_ins 4.394 | word_crf 1.251 | length 2.328 | ppl 12.83 | wps 19340.8 | ups 2.77 | wpb 6988.7 | bsz 272.5 | num_updates 63504 | lr 0.000198413 | gnorm 1.758 | train_wall 204 | gb_free 10.6 | wall 22640
2021-03-28 02:18:12 | INFO | fairseq.trainer | begin training epoch 109
2021-03-28 02:18:12 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 02:18:46 | INFO | train_inner | epoch 109:   96 / 588 loss=3.562, nll_loss=2.752, word_ins=4.298, word_crf=1.178, length=2.354, ppl=11.81, wps=16735.9, ups=2.37, wpb=7054.5, bsz=281.1, num_updates=63600, lr=0.000198263, gnorm=1.712, train_wall=35, gb_free=9.4, wall=22674
2021-03-28 02:19:21 | INFO | train_inner | epoch 109:   196 / 588 loss=3.615, nll_loss=2.787, word_ins=4.327, word_crf=1.212, length=2.391, ppl=12.25, wps=20231, ups=2.85, wpb=7090.7, bsz=289, num_updates=63700, lr=0.000198107, gnorm=1.721, train_wall=35, gb_free=9.5, wall=22709
2021-03-28 02:19:56 | INFO | train_inner | epoch 109:   296 / 588 loss=3.619, nll_loss=2.797, word_ins=4.337, word_crf=1.213, length=2.379, ppl=12.29, wps=19895.8, ups=2.84, wpb=7007.4, bsz=273.4, num_updates=63800, lr=0.000197952, gnorm=1.834, train_wall=35, gb_free=9.9, wall=22744
2021-03-28 02:20:32 | INFO | train_inner | epoch 109:   396 / 588 loss=3.847, nll_loss=3.051, word_ins=4.553, word_crf=1.357, length=2.13, ppl=14.39, wps=19152.6, ups=2.81, wpb=6807.6, bsz=250.6, num_updates=63900, lr=0.000197797, gnorm=2.247, train_wall=35, gb_free=10, wall=22780
2021-03-28 02:21:07 | INFO | train_inner | epoch 109:   496 / 588 loss=3.657, nll_loss=2.847, word_ins=4.337, word_crf=1.239, length=2.29, ppl=12.62, wps=19491.1, ups=2.85, wpb=6843.8, bsz=269.5, num_updates=64000, lr=0.000197642, gnorm=1.614, train_wall=35, gb_free=10.3, wall=22815
2021-03-28 02:21:39 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 02:21:42 | INFO | valid | epoch 109 | valid on 'valid' subset | loss 8.619 | nll_loss 6.011 | word_ins 7.266 | word_crf 4.418 | length 5.687 | ppl 393.21 | wps 59546 | wpb 5311.6 | bsz 208.1 | num_updates 64092 | best_loss 6.867
2021-03-28 02:21:42 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 109 @ 64092 updates
2021-03-28 02:21:42 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint109.pt
2021-03-28 02:21:43 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint109.pt
2021-03-28 02:21:46 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint109.pt (epoch 109 @ 64092 updates, score 8.619) (writing took 3.64186216500093 seconds)
2021-03-28 02:21:46 | INFO | fairseq_cli.train | end of epoch 109 (average epoch stats below)
2021-03-28 02:21:46 | INFO | train | epoch 109 | loss 3.676 | nll_loss 2.861 | word_ins 4.39 | word_crf 1.249 | length 2.323 | ppl 12.78 | wps 19213.9 | ups 2.75 | wpb 6988.7 | bsz 272.5 | num_updates 64092 | lr 0.0001975 | gnorm 1.824 | train_wall 205 | gb_free 9.5 | wall 22854
2021-03-28 02:21:46 | INFO | fairseq.trainer | begin training epoch 110
2021-03-28 02:21:46 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 02:21:49 | INFO | train_inner | epoch 110:   8 / 588 loss=3.734, nll_loss=2.904, word_ins=4.426, word_crf=1.279, length=2.422, ppl=13.31, wps=16845.2, ups=2.37, wpb=7094.8, bsz=273.3, num_updates=64100, lr=0.000197488, gnorm=1.825, train_wall=35, gb_free=9.5, wall=22857
2021-03-28 02:22:24 | INFO | train_inner | epoch 110:   108 / 588 loss=3.472, nll_loss=2.678, word_ins=4.233, word_crf=1.136, length=2.195, ppl=11.1, wps=20006.7, ups=2.89, wpb=6919.9, bsz=293.4, num_updates=64200, lr=0.000197334, gnorm=1.716, train_wall=34, gb_free=9.5, wall=22891
2021-03-28 02:22:59 | INFO | train_inner | epoch 110:   208 / 588 loss=3.718, nll_loss=2.919, word_ins=4.443, word_crf=1.27, length=2.262, ppl=13.16, wps=19864.7, ups=2.86, wpb=6951.1, bsz=252.6, num_updates=64300, lr=0.000197181, gnorm=1.901, train_wall=35, gb_free=9.4, wall=22926
2021-03-28 02:23:34 | INFO | train_inner | epoch 110:   308 / 588 loss=3.66, nll_loss=2.842, word_ins=4.376, word_crf=1.239, length=2.34, ppl=12.64, wps=19746.2, ups=2.8, wpb=7050.5, bsz=271.5, num_updates=64400, lr=0.000197028, gnorm=1.667, train_wall=35, gb_free=9.9, wall=22962
2021-03-28 02:24:09 | INFO | train_inner | epoch 110:   408 / 588 loss=3.685, nll_loss=2.843, word_ins=4.376, word_crf=1.253, length=2.442, ppl=12.87, wps=20142.7, ups=2.85, wpb=7076.7, bsz=281.9, num_updates=64500, lr=0.000196875, gnorm=1.876, train_wall=35, gb_free=9.6, wall=22997
2021-03-28 02:24:44 | INFO | train_inner | epoch 110:   508 / 588 loss=3.733, nll_loss=2.923, word_ins=4.441, word_crf=1.285, length=2.271, ppl=13.3, wps=20050.3, ups=2.86, wpb=7008.1, bsz=272.2, num_updates=64600, lr=0.000196722, gnorm=1.831, train_wall=35, gb_free=9.4, wall=23032
2021-03-28 02:25:12 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 02:25:16 | INFO | valid | epoch 110 | valid on 'valid' subset | loss 8.729 | nll_loss 6.063 | word_ins 7.327 | word_crf 4.496 | length 5.697 | ppl 424.38 | wps 59885.3 | wpb 5311.6 | bsz 208.1 | num_updates 64680 | best_loss 6.867
2021-03-28 02:25:16 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 110 @ 64680 updates
2021-03-28 02:25:16 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint110.pt
2021-03-28 02:25:16 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint110.pt
2021-03-28 02:25:19 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint110.pt (epoch 110 @ 64680 updates, score 8.729) (writing took 3.6352346950006904 seconds)
2021-03-28 02:25:19 | INFO | fairseq_cli.train | end of epoch 110 (average epoch stats below)
2021-03-28 02:25:19 | INFO | train | epoch 110 | loss 3.662 | nll_loss 2.849 | word_ins 4.381 | word_crf 1.241 | length 2.311 | ppl 12.66 | wps 19266.2 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 64680 | lr 0.000196601 | gnorm 1.779 | train_wall 204 | gb_free 9.5 | wall 23067
2021-03-28 02:25:19 | INFO | fairseq.trainer | begin training epoch 111
2021-03-28 02:25:19 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 02:25:26 | INFO | train_inner | epoch 111:   20 / 588 loss=3.646, nll_loss=2.842, word_ins=4.374, word_crf=1.23, length=2.287, ppl=12.52, wps=16427.3, ups=2.38, wpb=6904.8, bsz=274.5, num_updates=64700, lr=0.00019657, gnorm=1.681, train_wall=35, gb_free=9.8, wall=23074
2021-03-28 02:26:02 | INFO | train_inner | epoch 111:   120 / 588 loss=3.571, nll_loss=2.769, word_ins=4.315, word_crf=1.185, length=2.291, ppl=11.89, wps=20088.8, ups=2.84, wpb=7067, bsz=273.5, num_updates=64800, lr=0.000196419, gnorm=1.711, train_wall=35, gb_free=9.3, wall=23109
2021-03-28 02:26:37 | INFO | train_inner | epoch 111:   220 / 588 loss=3.547, nll_loss=2.736, word_ins=4.285, word_crf=1.173, length=2.312, ppl=11.69, wps=20074.3, ups=2.86, wpb=7029.2, bsz=288.6, num_updates=64900, lr=0.000196267, gnorm=1.669, train_wall=35, gb_free=9.3, wall=23144
2021-03-28 02:27:11 | INFO | train_inner | epoch 111:   320 / 588 loss=3.618, nll_loss=2.79, word_ins=4.33, word_crf=1.207, length=2.466, ppl=12.28, wps=20017.4, ups=2.9, wpb=6914.2, bsz=276.2, num_updates=65000, lr=0.000196116, gnorm=1.846, train_wall=34, gb_free=9.6, wall=23179

2021-03-28 02:27:46 | INFO | train_inner | epoch 111:   420 / 588 loss=3.687, nll_loss=2.875, word_ins=4.405, word_crf=1.254, length=2.301, ppl=12.88, wps=20274.5, ups=2.88, wpb=7042.3, bsz=258.9, num_updates=65100, lr=0.000195965, gnorm=1.777, train_wall=34, gb_free=9.4, wall=23214
2021-03-28 02:28:21 | INFO | train_inner | epoch 111:   520 / 588 loss=3.795, nll_loss=2.972, word_ins=4.483, word_crf=1.323, length=2.304, ppl=13.88, wps=19730.4, ups=2.82, wpb=6993.9, bsz=269.9, num_updates=65200, lr=0.000195815, gnorm=1.761, train_wall=35, gb_free=10.9, wall=23249
2021-03-28 02:28:45 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 02:28:49 | INFO | valid | epoch 111 | valid on 'valid' subset | loss 8.681 | nll_loss 6.031 | word_ins 7.295 | word_crf 4.464 | length 5.698 | ppl 410.55 | wps 59636.4 | wpb 5311.6 | bsz 208.1 | num_updates 65268 | best_loss 6.867
2021-03-28 02:28:49 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 111 @ 65268 updates
2021-03-28 02:28:49 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint111.pt
2021-03-28 02:28:50 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint111.pt
2021-03-28 02:28:52 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint111.pt (epoch 111 @ 65268 updates, score 8.681) (writing took 3.637092125998606 seconds)
2021-03-28 02:28:52 | INFO | fairseq_cli.train | end of epoch 111 (average epoch stats below)
2021-03-28 02:28:52 | INFO | train | epoch 111 | loss 3.645 | nll_loss 2.834 | word_ins 4.368 | word_crf 1.23 | length 2.307 | ppl 12.51 | wps 19277.3 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 65268 | lr 0.000195713 | gnorm 1.761 | train_wall 204 | gb_free 9.4 | wall 23280
2021-03-28 02:28:53 | INFO | fairseq.trainer | begin training epoch 112
2021-03-28 02:28:53 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 02:29:04 | INFO | train_inner | epoch 112:   32 / 588 loss=3.635, nll_loss=2.839, word_ins=4.373, word_crf=1.226, length=2.224, ppl=12.42, wps=16428.4, ups=2.36, wpb=6966.1, bsz=264.6, num_updates=65300, lr=0.000195665, gnorm=1.69, train_wall=35, gb_free=9.8, wall=23292
2021-03-28 02:29:38 | INFO | train_inner | epoch 112:   132 / 588 loss=3.452, nll_loss=2.636, word_ins=4.197, word_crf=1.122, length=2.304, ppl=10.94, wps=19999.6, ups=2.89, wpb=6918.5, bsz=294.9, num_updates=65400, lr=0.000195515, gnorm=1.793, train_wall=34, gb_free=9.7, wall=23326
2021-03-28 02:30:13 | INFO | train_inner | epoch 112:   232 / 588 loss=3.577, nll_loss=2.773, word_ins=4.318, word_crf=1.194, length=2.244, ppl=11.94, wps=19499.1, ups=2.85, wpb=6837.8, bsz=282, num_updates=65500, lr=0.000195366, gnorm=1.684, train_wall=35, gb_free=9.9, wall=23361
2021-03-28 02:30:49 | INFO | train_inner | epoch 112:   332 / 588 loss=3.767, nll_loss=2.972, word_ins=4.489, word_crf=1.298, length=2.252, ppl=13.62, wps=19388.3, ups=2.81, wpb=6911.7, bsz=243.4, num_updates=65600, lr=0.000195217, gnorm=1.765, train_wall=35, gb_free=9.3, wall=23397
2021-03-28 02:31:24 | INFO | train_inner | epoch 112:   432 / 588 loss=3.686, nll_loss=2.879, word_ins=4.41, word_crf=1.252, length=2.287, ppl=12.87, wps=20151.6, ups=2.87, wpb=7023.1, bsz=259.4, num_updates=65700, lr=0.000195069, gnorm=1.84, train_wall=35, gb_free=9.6, wall=23432
2021-03-28 02:31:59 | INFO | train_inner | epoch 112:   532 / 588 loss=3.666, nll_loss=2.842, word_ins=4.376, word_crf=1.232, length=2.456, ppl=12.69, wps=20581.8, ups=2.87, wpb=7174.2, bsz=272.8, num_updates=65800, lr=0.00019492, gnorm=1.622, train_wall=35, gb_free=10.1, wall=23467
2021-03-28 02:32:19 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 02:32:22 | INFO | valid | epoch 112 | valid on 'valid' subset | loss 8.663 | nll_loss 6.024 | word_ins 7.284 | word_crf 4.452 | length 5.693 | ppl 405.34 | wps 59370.5 | wpb 5311.6 | bsz 208.1 | num_updates 65856 | best_loss 6.867
2021-03-28 02:32:22 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 112 @ 65856 updates
2021-03-28 02:32:22 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint112.pt
2021-03-28 02:32:23 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint112.pt
2021-03-28 02:32:26 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint112.pt (epoch 112 @ 65856 updates, score 8.663) (writing took 3.640914179999527 seconds)
2021-03-28 02:32:26 | INFO | fairseq_cli.train | end of epoch 112 (average epoch stats below)
2021-03-28 02:32:26 | INFO | train | epoch 112 | loss 3.627 | nll_loss 2.817 | word_ins 4.355 | word_crf 1.22 | length 2.302 | ppl 12.36 | wps 19273.3 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 65856 | lr 0.000194837 | gnorm 1.719 | train_wall 204 | gb_free 9.4 | wall 23494
2021-03-28 02:32:26 | INFO | fairseq.trainer | begin training epoch 113
2021-03-28 02:32:26 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 02:32:41 | INFO | train_inner | epoch 113:   44 / 588 loss=3.664, nll_loss=2.859, word_ins=4.389, word_crf=1.245, length=2.237, ppl=12.67, wps=16348, ups=2.34, wpb=6972.2, bsz=268.9, num_updates=65900, lr=0.000194772, gnorm=1.761, train_wall=35, gb_free=9.8, wall=23509
2021-03-28 02:33:16 | INFO | train_inner | epoch 113:   144 / 588 loss=3.438, nll_loss=2.613, word_ins=4.179, word_crf=1.106, length=2.418, ppl=10.83, wps=20146.6, ups=2.88, wpb=6992.4, bsz=303.8, num_updates=66000, lr=0.000194625, gnorm=1.753, train_wall=34, gb_free=9.4, wall=23544
2021-03-28 02:33:51 | INFO | train_inner | epoch 113:   244 / 588 loss=3.626, nll_loss=2.837, word_ins=4.374, word_crf=1.225, length=2.137, ppl=12.34, wps=19682.1, ups=2.87, wpb=6868.7, bsz=260.3, num_updates=66100, lr=0.000194477, gnorm=1.851, train_wall=35, gb_free=9.3, wall=23579
2021-03-28 02:34:26 | INFO | train_inner | epoch 113:   344 / 588 loss=3.682, nll_loss=2.853, word_ins=4.385, word_crf=1.24, length=2.498, ppl=12.84, wps=20588.5, ups=2.86, wpb=7190.7, bsz=265.9, num_updates=66200, lr=0.000194331, gnorm=1.784, train_wall=35, gb_free=10.1, wall=23614
2021-03-28 02:35:01 | INFO | train_inner | epoch 113:   444 / 588 loss=3.601, nll_loss=2.797, word_ins=4.336, word_crf=1.216, length=2.169, ppl=12.13, wps=19869.4, ups=2.88, wpb=6910.2, bsz=281, num_updates=66300, lr=0.000194184, gnorm=1.92, train_wall=34, gb_free=9.3, wall=23649
2021-03-28 02:35:36 | INFO | train_inner | epoch 113:   544 / 588 loss=3.675, nll_loss=2.864, word_ins=4.393, word_crf=1.255, length=2.239, ppl=12.77, wps=20108.6, ups=2.86, wpb=7038, bsz=277.5, num_updates=66400, lr=0.000194038, gnorm=1.893, train_wall=35, gb_free=9.7, wall=23684
2021-03-28 02:35:52 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 02:35:55 | INFO | valid | epoch 113 | valid on 'valid' subset | loss 8.654 | nll_loss 6.02 | word_ins 7.284 | word_crf 4.45 | length 5.616 | ppl 402.83 | wps 59802.3 | wpb 5311.6 | bsz 208.1 | num_updates 66444 | best_loss 6.867
2021-03-28 02:35:55 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 113 @ 66444 updates
2021-03-28 02:35:55 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint113.pt
2021-03-28 02:35:56 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint113.pt
2021-03-28 02:35:59 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint113.pt (epoch 113 @ 66444 updates, score 8.654) (writing took 3.6389639010012615 seconds)
2021-03-28 02:35:59 | INFO | fairseq_cli.train | end of epoch 113 (average epoch stats below)
2021-03-28 02:35:59 | INFO | train | epoch 113 | loss 3.639 | nll_loss 2.829 | word_ins 4.365 | word_crf 1.228 | length 2.285 | ppl 12.46 | wps 19282.6 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 66444 | lr 0.000193973 | gnorm 1.86 | train_wall 204 | gb_free 10.7 | wall 23707
2021-03-28 02:35:59 | INFO | fairseq.trainer | begin training epoch 114
2021-03-28 02:35:59 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 02:36:19 | INFO | train_inner | epoch 114:   56 / 588 loss=3.819, nll_loss=3.008, word_ins=4.517, word_crf=1.331, length=2.293, ppl=14.11, wps=16247.9, ups=2.32, wpb=7004.3, bsz=256.6, num_updates=66500, lr=0.000193892, gnorm=1.925, train_wall=36, gb_free=10.6, wall=23727
2021-03-28 02:36:54 | INFO | train_inner | epoch 114:   156 / 588 loss=3.639, nll_loss=2.878, word_ins=4.41, word_crf=1.228, length=2.055, ppl=12.46, wps=19702.4, ups=2.86, wpb=6898.1, bsz=239.4, num_updates=66600, lr=0.000193746, gnorm=1.871, train_wall=35, gb_free=9.2, wall=23762
2021-03-28 02:37:29 | INFO | train_inner | epoch 114:   256 / 588 loss=3.615, nll_loss=2.808, word_ins=4.345, word_crf=1.221, length=2.223, ppl=12.26, wps=20048.9, ups=2.85, wpb=7025.3, bsz=287.9, num_updates=66700, lr=0.000193601, gnorm=1.915, train_wall=35, gb_free=9.2, wall=23797
2021-03-28 02:38:04 | INFO | train_inner | epoch 114:   356 / 588 loss=3.685, nll_loss=2.865, word_ins=4.398, word_crf=1.247, length=2.39, ppl=12.86, wps=20043.9, ups=2.85, wpb=7043, bsz=258.5, num_updates=66800, lr=0.000193456, gnorm=1.67, train_wall=35, gb_free=10.2, wall=23832
2021-03-28 02:38:39 | INFO | train_inner | epoch 114:   456 / 588 loss=3.515, nll_loss=2.696, word_ins=4.25, word_crf=1.148, length=2.413, ppl=11.43, wps=19968.2, ups=2.87, wpb=6950.1, bsz=286.8, num_updates=66900, lr=0.000193311, gnorm=1.667, train_wall=35, gb_free=9.6, wall=23867
2021-03-28 02:39:14 | INFO | train_inner | epoch 114:   556 / 588 loss=3.667, nll_loss=2.843, word_ins=4.375, word_crf=1.249, length=2.303, ppl=12.71, wps=20188.8, ups=2.86, wpb=7070.4, bsz=277.5, num_updates=67000, lr=0.000193167, gnorm=1.913, train_wall=35, gb_free=9.3, wall=23902
2021-03-28 02:39:25 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 02:39:28 | INFO | valid | epoch 114 | valid on 'valid' subset | loss 8.644 | nll_loss 6.016 | word_ins 7.277 | word_crf 4.443 | length 5.625 | ppl 399.96 | wps 59621.7 | wpb 5311.6 | bsz 208.1 | num_updates 67032 | best_loss 6.867
2021-03-28 02:39:28 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 114 @ 67032 updates
2021-03-28 02:39:28 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint114.pt
2021-03-28 02:39:29 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint114.pt
2021-03-28 02:39:32 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint114.pt (epoch 114 @ 67032 updates, score 8.644) (writing took 3.663510379003128 seconds)
2021-03-28 02:39:32 | INFO | fairseq_cli.train | end of epoch 114 (average epoch stats below)

2021-03-28 02:39:32 | INFO | train | epoch 114 | loss 3.623 | nll_loss 2.815 | word_ins 4.353 | word_crf 1.218 | length 2.284 | ppl 12.32 | wps 19283.1 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 67032 | lr 0.000193121 | gnorm 1.819 | train_wall 204 | gb_free 9.7 | wall 23920
2021-03-28 02:39:32 | INFO | fairseq.trainer | begin training epoch 115
2021-03-28 02:39:32 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 02:39:56 | INFO | train_inner | epoch 115:   68 / 588 loss=3.63, nll_loss=2.827, word_ins=4.361, word_crf=1.233, length=2.168, ppl=12.38, wps=16272.6, ups=2.36, wpb=6891.6, bsz=286.9, num_updates=67100, lr=0.000193023, gnorm=1.837, train_wall=35, gb_free=9.4, wall=23944
2021-03-28 02:40:31 | INFO | train_inner | epoch 115:   168 / 588 loss=3.566, nll_loss=2.764, word_ins=4.309, word_crf=1.182, length=2.293, ppl=11.84, wps=20254.5, ups=2.86, wpb=7086, bsz=274.6, num_updates=67200, lr=0.000192879, gnorm=1.7, train_wall=35, gb_free=9.3, wall=23979
2021-03-28 02:41:06 | INFO | train_inner | epoch 115:   268 / 588 loss=3.68, nll_loss=2.892, word_ins=4.423, word_crf=1.251, length=2.175, ppl=12.82, wps=19666.1, ups=2.84, wpb=6935.2, bsz=245.5, num_updates=67300, lr=0.000192736, gnorm=1.768, train_wall=35, gb_free=9.3, wall=24014
2021-03-28 02:41:42 | INFO | train_inner | epoch 115:   368 / 588 loss=3.617, nll_loss=2.824, word_ins=4.363, word_crf=1.211, length=2.249, ppl=12.27, wpb=6983.1, bsz=259, num_updates=67400, lr=0.000192593, gnorm=1.682, train_wall=35, gb_free=9.6, wall=24050
2021-03-28 02:42:16 | INFO | train_inner | epoch 115:   468 / 588 loss=3.499, nll_loss=2.687, word_ins=4.243, word_crf=1.148, length=2.287, ppl=11.3, wps=19855.8, ups=2.89, wpb=6865.8, bsz=288.9, num_updates=67500, lr=0.00019245, gnorm=1.978, train_wall=34, gb_free=9, wall=24084
2021-03-28 02:42:51 | INFO | train_inner | epoch 115:   568 / 588 loss=3.616, nll_loss=2.789, word_ins=4.329, word_crf=1.214, length=2.379, ppl=12.26, wps=20412.1, ups=2.86, wpb=7138.1, bsz=285.2, num_updates=67600, lr=0.000192308, gnorm=1.706, train_wall=35, gb_free=9.3, wall=24119
2021-03-28 02:42:58 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 02:43:01 | INFO | valid | epoch 115 | valid on 'valid' subset | loss 8.84 | nll_loss 6.116 | word_ins 7.381 | word_crf 4.567 | length 5.819 | ppl 458.12 | wps 59347.9 | wpb 5311.6 | bsz 208.1 | num_updates 67620 | best_loss 6.867
2021-03-28 02:43:01 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 115 @ 67620 updates
2021-03-28 02:43:01 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint115.pt
2021-03-28 02:43:02 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint115.pt
2021-03-28 02:43:05 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint115.pt (epoch 115 @ 67620 updates, score 8.84) (writing took 3.931844024002203 seconds)
2021-03-28 02:43:05 | INFO | fairseq_cli.train | end of epoch 115 (average epoch stats below)
2021-03-28 02:43:05 | INFO | train | epoch 115 | loss 3.6 | nll_loss 2.797 | word_ins 4.338 | word_crf 1.206 | length 2.258 | ppl 12.13 | wps 19244.6 | ups 2.75 | wpb 6988.7 | bsz 272.5 | num_updates 67620 | lr 0.000192279 | gnorm 1.768 | train_wall 204 | gb_free 9.8 | wall 24133
2021-03-28 02:43:06 | INFO | fairseq.trainer | begin training epoch 116
2021-03-28 02:43:06 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 02:43:34 | INFO | train_inner | epoch 116:   80 / 588 loss=3.559, nll_loss=2.773, word_ins=4.319, word_crf=1.176, length=2.235, ppl=11.78, wps=16355.8, ups=2.36, wpb=6942.6, bsz=266.1, num_updates=67700, lr=0.000192166, gnorm=1.722, train_wall=35, gb_free=9.5, wall=24162
2021-03-28 02:44:09 | INFO | train_inner | epoch 116:   180 / 588 loss=3.562, nll_loss=2.752, word_ins=4.302, word_crf=1.185, length=2.264, ppl=11.81, wps=19760.5, ups=2.87, wpb=6897.2, bsz=273.8, num_updates=67800, lr=0.000192024, gnorm=1.989, train_wall=35, gb_free=9.8, wall=24196
2021-03-28 02:44:44 | INFO | train_inner | epoch 116:   280 / 588 loss=3.552, nll_loss=2.75, word_ins=4.297, word_crf=1.185, length=2.187, ppl=11.73, wps=19916, ups=2.85, wpb=6977.4, bsz=290.4, num_updates=67900, lr=0.000191882, gnorm=1.815, train_wall=35, gb_free=9.2, wall=24231
2021-03-28 02:45:19 | INFO | train_inner | epoch 116:   380 / 588 loss=3.794, nll_loss=3.015, word_ins=4.525, word_crf=1.327, length=2.048, ppl=13.88, wps=19417.9, ups=2.79, wpb=6950.5, bsz=248.4, num_updates=68000, lr=0.000191741, gnorm=1.931, train_wall=36, gb_free=9.4, wall=24267
2021-03-28 02:45:54 | INFO | train_inner | epoch 116:   480 / 588 loss=3.613, nll_loss=2.804, word_ins=4.343, word_crf=1.211, length=2.306, ppl=12.24, wps=20140.3, ups=2.85, wpb=7059.1, bsz=269.4, num_updates=68100, lr=0.0001916, gnorm=1.75, train_wall=35, gb_free=10, wall=24302
2021-03-28 02:46:29 | INFO | train_inner | epoch 116:   580 / 588 loss=3.519, nll_loss=2.685, word_ins=4.241, word_crf=1.146, length=2.526, ppl=11.47, wps=20524.9, ups=2.9, wpb=7088.7, bsz=285.8, num_updates=68200, lr=0.00019146, gnorm=1.714, train_wall=34, gb_free=9.4, wall=24337
2021-03-28 02:46:32 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 02:46:35 | INFO | valid | epoch 116 | valid on 'valid' subset | loss 8.709 | nll_loss 6.054 | word_ins 7.317 | word_crf 4.48 | length 5.705 | ppl 418.58 | wps 59375.7 | wpb 5311.6 | bsz 208.1 | num_updates 68208 | best_loss 6.867
2021-03-28 02:46:35 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 116 @ 68208 updates
2021-03-28 02:46:35 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint116.pt
2021-03-28 02:46:36 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint116.pt
2021-03-28 02:46:39 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint116.pt (epoch 116 @ 68208 updates, score 8.709) (writing took 3.638189679997595 seconds)
2021-03-28 02:46:39 | INFO | fairseq_cli.train | end of epoch 116 (average epoch stats below)
2021-03-28 02:46:39 | INFO | train | epoch 116 | loss 3.597 | nll_loss 2.793 | word_ins 4.335 | word_crf 1.204 | length 2.261 | ppl 12.1 | wps 19262.1 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 68208 | lr 0.000191449 | gnorm 1.823 | train_wall 205 | gb_free 9.5 | wall 24347
2021-03-28 02:46:39 | INFO | fairseq.trainer | begin training epoch 117
2021-03-28 02:46:39 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 02:47:11 | INFO | train_inner | epoch 117:   92 / 588 loss=3.55, nll_loss=2.741, word_ins=4.291, word_crf=1.174, length=2.307, ppl=11.71, wps=16927.6, ups=2.36, wpb=7160.6, bsz=279.6, num_updates=68300, lr=0.00019132, gnorm=1.865, train_wall=35, gb_free=9.5, wall=24379
2021-03-28 02:47:47 | INFO | train_inner | epoch 117:   192 / 588 loss=3.681, nll_loss=2.895, word_ins=4.423, word_crf=1.259, length=2.107, ppl=12.83, wps=19523.2, ups=2.81, wpb=6948.3, bsz=260.1, num_updates=68400, lr=0.00019118, gnorm=2.013, train_wall=35, gb_free=9.4, wall=24415
2021-03-28 02:48:21 | INFO | train_inner | epoch 117:   292 / 588 loss=3.358, nll_loss=2.525, word_ins=4.104, word_crf=1.064, length=2.428, ppl=10.25, wps=20608.4, ups=2.89, wpb=7121.2, bsz=316.6, num_updates=68500, lr=0.00019104, gnorm=1.703, train_wall=34, gb_free=9.6, wall=24449
2021-03-28 02:48:56 | INFO | train_inner | epoch 117:   392 / 588 loss=3.613, nll_loss=2.801, word_ins=4.34, word_crf=1.217, length=2.258, ppl=12.24, wps=20053, ups=2.87, wpb=6990.2, bsz=278.3, num_updates=68600, lr=0.000190901, gnorm=1.922, train_wall=35, gb_free=9.4, wall=24484
2021-03-28 02:49:31 | INFO | train_inner | epoch 117:   492 / 588 loss=3.652, nll_loss=2.869, word_ins=4.399, word_crf=1.244, length=2.083, ppl=12.57, wps=19263.6, ups=2.85, wpb=6753.5, bsz=263.5, num_updates=68700, lr=0.000190762, gnorm=1.844, train_wall=35, gb_free=9.7, wall=24519
2021-03-28 02:50:05 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 02:50:08 | INFO | valid | epoch 117 | valid on 'valid' subset | loss 8.764 | nll_loss 6.076 | word_ins 7.341 | word_crf 4.511 | length 5.822 | ppl 434.69 | wps 59527.2 | wpb 5311.6 | bsz 208.1 | num_updates 68796 | best_loss 6.867
2021-03-28 02:50:08 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 117 @ 68796 updates
2021-03-28 02:50:08 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint117.pt
2021-03-28 02:50:09 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint117.pt
2021-03-28 02:50:12 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint117.pt (epoch 117 @ 68796 updates, score 8.764) (writing took 3.6661441850010306 seconds)
2021-03-28 02:50:12 | INFO | fairseq_cli.train | end of epoch 117 (average epoch stats below)
2021-03-28 02:50:12 | INFO | train | epoch 117 | loss 3.594 | nll_loss 2.793 | word_ins 4.334 | word_crf 1.203 | length 2.239 | ppl 12.08 | wps 19259.1 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 68796 | lr 0.000190629 | gnorm 1.851 | train_wall 205 | gb_free 9.3 | wall 24560
2021-03-28 02:50:12 | INFO | fairseq.trainer | begin training epoch 118
2021-03-28 02:50:12 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 02:50:14 | INFO | train_inner | epoch 118:   4 / 588 loss=3.72, nll_loss=2.929, word_ins=4.453, word_crf=1.266, length=2.273, ppl=13.17, wps=16397.8, ups=2.36, wpb=6962.8, bsz=236.5, num_updates=68800, lr=0.000190623, gnorm=1.764, train_wall=35, gb_free=9.9, wall=24562
2021-03-28 02:50:49 | INFO | train_inner | epoch 118:   104 / 588 loss=3.552, nll_loss=2.774, word_ins=4.319, word_crf=1.192, length=2.008, ppl=11.73, wps=19527, ups=2.85, wpb=6854, bsz=276.1, num_updates=68900, lr=0.000190485, gnorm=1.945, train_wall=35, gb_free=9.9, wall=24597
2021-03-28 02:51:24 | INFO | train_inner | epoch 118:   204 / 588 loss=3.615, nll_loss=2.82, word_ins=4.361, word_crf=1.206, length=2.282, ppl=12.25, wps=19946.5, ups=2.87, wpb=6950.4, bsz=252.1, num_updates=69000, lr=0.000190347, gnorm=2.168, train_wall=35, gb_free=9.5, wall=24632
2021-03-28 02:51:58 | INFO | train_inner | epoch 118:   304 / 588 loss=3.436, nll_loss=2.62, word_ins=4.185, word_crf=1.112, length=2.322, ppl=10.83, wps=20245.3, ups=2.9, wpb=6974.3, bsz=300.9, num_updates=69100, lr=0.000190209, gnorm=1.744, train_wall=34, gb_free=9.1, wall=24666
2021-03-28 02:52:34 | INFO | train_inner | epoch 118:   404 / 588 loss=3.649, nll_loss=2.845, word_ins=4.382, word_crf=1.229, length=2.289, ppl=12.54, wps=19689.7, ups=2.82, wpb=6981.3, bsz=258.2, num_updates=69200, lr=0.000190071, gnorm=1.745, train_wall=35, gb_free=9.9, wall=24702
2021-03-28 02:53:09 | INFO | train_inner | epoch 118:   504 / 588 loss=3.634, nll_loss=2.835, word_ins=4.37, word_crf=1.228, length=2.208, ppl=12.41, wps=20261.3, ups=2.85, wpb=7100.9, bsz=269.1, num_updates=69300, lr=0.000189934, gnorm=1.814, train_wall=35, gb_free=9.5, wall=24737

2021-03-28 02:53:38 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 02:53:41 | INFO | valid | epoch 118 | valid on 'valid' subset | loss 8.709 | nll_loss 6.04 | word_ins 7.306 | word_crf 4.48 | length 5.756 | ppl 418.37 | wps 59869.4 | wpb 5311.6 | bsz 208.1 | num_updates 69384 | best_loss 6.867
2021-03-28 02:53:41 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 118 @ 69384 updates
2021-03-28 02:53:41 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint118.pt
2021-03-28 02:53:42 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint118.pt
2021-03-28 02:53:45 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint118.pt (epoch 118 @ 69384 updates, score 8.709) (writing took 3.634912516998156 seconds)
2021-03-28 02:53:45 | INFO | fairseq_cli.train | end of epoch 118 (average epoch stats below)
2021-03-28 02:53:45 | INFO | train | epoch 118 | loss 3.586 | nll_loss 2.785 | word_ins 4.328 | word_crf 1.198 | length 2.231 | ppl 12.01 | wps 19315.2 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 69384 | lr 0.000189819 | gnorm 1.868 | train_wall 204 | gb_free 9.8 | wall 24773
2021-03-28 02:53:45 | INFO | fairseq.trainer | begin training epoch 119
2021-03-28 02:53:45 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 02:53:51 | INFO | train_inner | epoch 119:   16 / 588 loss=3.554, nll_loss=2.74, word_ins=4.288, word_crf=1.18, length=2.295, ppl=11.74, wps=16930.2, ups=2.39, wpb=7095, bsz=288.9, num_updates=69400, lr=0.000189797, gnorm=1.727, train_wall=34, gb_free=9.4, wall=24778
2021-03-28 02:54:26 | INFO | train_inner | epoch 119:   116 / 588 loss=3.595, nll_loss=2.832, word_ins=4.374, word_crf=1.205, length=2.026, ppl=12.08, wps=19582.6, ups=2.85, wpb=6867.1, bsz=239.3, num_updates=69500, lr=0.000189661, gnorm=2.03, train_wall=35, gb_free=9.6, wall=24814
2021-03-28 02:55:01 | INFO | train_inner | epoch 119:   216 / 588 loss=3.625, nll_loss=2.832, word_ins=4.37, word_crf=1.225, length=2.157, ppl=12.34, wps=19323.2, ups=2.81, wpb=6868.9, bsz=262.2, num_updates=69600, lr=0.000189525, gnorm=2.01, train_wall=35, gb_free=9.6, wall=24849
2021-03-28 02:55:36 | INFO | train_inner | epoch 119:   316 / 588 loss=3.42, nll_loss=2.606, word_ins=4.172, word_crf=1.109, length=2.247, ppl=10.7, wps=20439.9, ups=2.9, wpb=7054.3, bsz=315.9, num_updates=69700, lr=0.000189389, gnorm=1.605, train_wall=34, gb_free=10.4, wall=24884
2021-03-28 02:56:10 | INFO | train_inner | epoch 119:   416 / 588 loss=3.472, nll_loss=2.667, word_ins=4.226, word_crf=1.128, length=2.307, ppl=11.09, wps=20168.3, ups=2.91, wpb=6931.9, bsz=289.2, num_updates=69800, lr=0.000189253, gnorm=1.638, train_wall=34, gb_free=9.2, wall=24918
2021-03-28 02:56:45 | INFO | train_inner | epoch 119:   516 / 588 loss=3.769, nll_loss=2.964, word_ins=4.483, word_crf=1.297, length=2.304, ppl=13.63, wps=20000.5, ups=2.85, wpb=7029, bsz=244.9, num_updates=69900, lr=0.000189117, gnorm=2.071, train_wall=35, gb_free=9.4, wall=24953
2021-03-28 02:57:11 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 02:57:14 | INFO | valid | epoch 119 | valid on 'valid' subset | loss 8.79 | nll_loss 6.099 | word_ins 7.361 | word_crf 4.537 | length 5.728 | ppl 442.56 | wps 59865.8 | wpb 5311.6 | bsz 208.1 | num_updates 69972 | best_loss 6.867
2021-03-28 02:57:14 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 119 @ 69972 updates
2021-03-28 02:57:14 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint119.pt
2021-03-28 02:57:15 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint119.pt
2021-03-28 02:57:17 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint119.pt (epoch 119 @ 69972 updates, score 8.79) (writing took 3.6369356159993913 seconds)
2021-03-28 02:57:17 | INFO | fairseq_cli.train | end of epoch 119 (average epoch stats below)
2021-03-28 02:57:17 | INFO | train | epoch 119 | loss 3.572 | nll_loss 2.773 | word_ins 4.318 | word_crf 1.19 | length 2.229 | ppl 11.89 | wps 19327 | ups 2.77 | wpb 6988.7 | bsz 272.5 | num_updates 69972 | lr 0.00018902 | gnorm 1.852 | train_wall 204 | gb_free 9.7 | wall 24985
2021-03-28 02:57:18 | INFO | fairseq.trainer | begin training epoch 120
2021-03-28 02:57:18 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 02:57:28 | INFO | train_inner | epoch 120:   28 / 588 loss=3.582, nll_loss=2.769, word_ins=4.314, word_crf=1.191, length=2.339, ppl=11.97, wps=16817.9, ups=2.36, wpb=7129.4, bsz=272.2, num_updates=70000, lr=0.000188982, gnorm=1.758, train_wall=35, gb_free=9.3, wall=24995
2021-03-28 02:57:28 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 02:57:31 | INFO | valid | epoch 120 | valid on 'valid' subset | loss 8.897 | nll_loss 6.149 | word_ins 7.419 | word_crf 4.602 | length 5.857 | ppl 476.74 | wps 59839.4 | wpb 5311.6 | bsz 208.1 | num_updates 70000 | best_loss 6.867
2021-03-28 02:57:31 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 120 @ 70000 updates
2021-03-28 02:57:31 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint_120_70000.pt
2021-03-28 02:57:32 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint_120_70000.pt
2021-03-28 02:57:35 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint_120_70000.pt (epoch 120 @ 70000 updates, score 8.897) (writing took 3.979297835001489 seconds)
2021-03-28 02:58:09 | INFO | train_inner | epoch 120:   128 / 588 loss=3.408, nll_loss=2.608, word_ins=4.179, word_crf=1.095, length=2.234, ppl=10.61, wps=16615.9, ups=2.39, wpb=6946.6, bsz=288.2, num_updates=70100, lr=0.000188847, gnorm=1.776, train_wall=34, gb_free=9.2, wall=25037
2021-03-28 02:58:45 | INFO | train_inner | epoch 120:   228 / 588 loss=3.56, nll_loss=2.773, word_ins=4.322, word_crf=1.177, length=2.221, ppl=11.79, wps=20118.3, ups=2.85, wpb=7066.6, bsz=265.6, num_updates=70200, lr=0.000188713, gnorm=1.773, train_wall=35, gb_free=9.5, wall=25072
2021-03-28 02:59:19 | INFO | train_inner | epoch 120:   328 / 588 loss=3.557, nll_loss=2.769, word_ins=4.317, word_crf=1.18, length=2.189, ppl=11.77, wps=19937.6, ups=2.89, wpb=6897.1, bsz=262.2, num_updates=70300, lr=0.000188579, gnorm=1.746, train_wall=34, gb_free=9.6, wall=25107
2021-03-28 02:59:54 | INFO | train_inner | epoch 120:   428 / 588 loss=3.662, nll_loss=2.86, word_ins=4.392, word_crf=1.25, length=2.153, ppl=12.65, wps=19878.5, ups=2.84, wpb=7000.4, bsz=267.4, num_updates=70400, lr=0.000188445, gnorm=1.956, train_wall=35, gb_free=9.4, wall=25142
2021-03-28 03:00:29 | INFO | train_inner | epoch 120:   528 / 588 loss=3.62, nll_loss=2.811, word_ins=4.352, word_crf=1.213, length=2.316, ppl=12.3, wps=20231, ups=2.87, wpb=7045.5, bsz=265.3, num_updates=70500, lr=0.000188311, gnorm=1.786, train_wall=35, gb_free=10.1, wall=25177
2021-03-28 03:00:50 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 03:00:53 | INFO | valid | epoch 120 | valid on 'valid' subset | loss 8.752 | nll_loss 6.068 | word_ins 7.331 | word_crf 4.501 | length 5.846 | ppl 431 | wps 59801.6 | wpb 5311.6 | bsz 208.1 | num_updates 70560 | best_loss 6.867
2021-03-28 03:00:53 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 120 @ 70560 updates
2021-03-28 03:00:53 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint120.pt
2021-03-28 03:00:54 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint120.pt
2021-03-28 03:00:57 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint120.pt (epoch 120 @ 70560 updates, score 8.752) (writing took 3.63940273100161 seconds)
2021-03-28 03:00:57 | INFO | fairseq_cli.train | end of epoch 120 (average epoch stats below)
2021-03-28 03:00:57 | INFO | train | epoch 120 | loss 3.551 | nll_loss 2.754 | word_ins 4.303 | word_crf 1.177 | length 2.222 | ppl 11.72 | wps 18734.5 | ups 2.68 | wpb 6988.7 | bsz 272.5 | num_updates 70560 | lr 0.000188231 | gnorm 1.784 | train_wall 203 | gb_free 9.5 | wall 25205
2021-03-28 03:00:57 | INFO | fairseq.trainer | begin training epoch 121
2021-03-28 03:00:57 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 03:01:11 | INFO | train_inner | epoch 121:   40 / 588 loss=3.451, nll_loss=2.649, word_ins=4.211, word_crf=1.124, length=2.219, ppl=10.94, wps=16842.3, ups=2.4, wpb=7031.5, bsz=302, num_updates=70600, lr=0.000188177, gnorm=1.696, train_wall=34, gb_free=9.6, wall=25219
2021-03-28 03:01:46 | INFO | train_inner | epoch 121:   140 / 588 loss=3.542, nll_loss=2.734, word_ins=4.286, word_crf=1.163, length=2.356, ppl=11.65, wps=20361.5, ups=2.85, wpb=7133.1, bsz=271.2, num_updates=70700, lr=0.000188044, gnorm=1.899, train_wall=35, gb_free=9.9, wall=25254
2021-03-28 03:02:21 | INFO | train_inner | epoch 121:   240 / 588 loss=3.407, nll_loss=2.615, word_ins=4.184, word_crf=1.099, length=2.166, ppl=10.61, wps=19895.8, ups=2.81, wpb=7071.5, bsz=293, num_updates=70800, lr=0.000187912, gnorm=1.636, train_wall=35, gb_free=9.5, wall=25289
2021-03-28 03:02:57 | INFO | train_inner | epoch 121:   340 / 588 loss=3.639, nll_loss=2.865, word_ins=4.4, word_crf=1.232, length=2.066, ppl=12.46, wps=19326.8, ups=2.83, wpb=6838.1, bsz=247.9, num_updates=70900, lr=0.000187779, gnorm=1.787, train_wall=35, gb_free=9.7, wall=25325
2021-03-28 03:03:33 | INFO | train_inner | epoch 121:   440 / 588 loss=3.634, nll_loss=2.837, word_ins=4.375, word_crf=1.227, length=2.196, ppl=12.42, wps=19349.4, ups=2.76, wpb=7004.7, bsz=261.4, num_updates=71000, lr=0.000187647, gnorm=1.724, train_wall=36, gb_free=9.6, wall=25361
2021-03-28 03:04:08 | INFO | train_inner | epoch 121:   540 / 588 loss=3.571, nll_loss=2.776, word_ins=4.32, word_crf=1.194, length=2.172, ppl=11.88, wps=19964.5, ups=2.88, wpb=6934.2, bsz=273.4, num_updates=71100, lr=0.000187515, gnorm=1.808, train_wall=34, gb_free=9.7, wall=25396
2021-03-28 03:04:24 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 03:04:27 | INFO | valid | epoch 121 | valid on 'valid' subset | loss 8.808 | nll_loss 6.1 | word_ins 7.363 | word_crf 4.538 | length 5.887 | ppl 448.22 | wps 60079.8 | wpb 5311.6 | bsz 208.1 | num_updates 71148 | best_loss 6.867
2021-03-28 03:04:27 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 121 @ 71148 updates
2021-03-28 03:04:27 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint121.pt
2021-03-28 03:04:28 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint121.pt

2021-03-28 03:04:31 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint121.pt (epoch 121 @ 71148 updates, score 8.808) (writing took 3.6425897460030683 seconds)
2021-03-28 03:04:31 | INFO | fairseq_cli.train | end of epoch 121 (average epoch stats below)
2021-03-28 03:04:31 | INFO | train | epoch 121 | loss 3.534 | nll_loss 2.74 | word_ins 4.291 | word_crf 1.168 | length 2.207 | ppl 11.59 | wps 19196.7 | ups 2.75 | wpb 6988.7 | bsz 272.5 | num_updates 71148 | lr 0.000187451 | gnorm 1.762 | train_wall 205 | gb_free 9.2 | wall 25419
2021-03-28 03:04:31 | INFO | fairseq.trainer | begin training epoch 122
2021-03-28 03:04:31 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 03:04:49 | INFO | train_inner | epoch 122:   52 / 588 loss=3.476, nll_loss=2.691, word_ins=4.248, word_crf=1.136, length=2.165, ppl=11.13, wps=16439.2, ups=2.41, wpb=6819.5, bsz=275.2, num_updates=71200, lr=0.000187383, gnorm=1.835, train_wall=34, gb_free=9, wall=25437
2021-03-28 03:05:24 | INFO | train_inner | epoch 122:   152 / 588 loss=3.503, nll_loss=2.72, word_ins=4.275, word_crf=1.153, length=2.122, ppl=11.33, wps=20202.2, ups=2.89, wpb=6986.4, bsz=266.1, num_updates=71300, lr=0.000187251, gnorm=1.926, train_wall=34, gb_free=9.4, wall=25472
2021-03-28 03:05:58 | INFO | train_inner | epoch 122:   252 / 588 loss=3.514, nll_loss=2.739, word_ins=4.291, word_crf=1.155, length=2.137, ppl=11.43, wps=20125.4, ups=2.89, wpb=6957.4, bsz=264.9, num_updates=71400, lr=0.00018712, gnorm=1.673, train_wall=34, gb_free=9.4, wall=25506
2021-03-28 03:06:33 | INFO | train_inner | epoch 122:   352 / 588 loss=3.444, nll_loss=2.629, word_ins=4.195, word_crf=1.112, length=2.348, ppl=10.88, wps=20207.8, ups=2.89, wpb=6997.8, bsz=297.2, num_updates=71500, lr=0.000186989, gnorm=1.703, train_wall=34, gb_free=9.9, wall=25541
2021-03-28 03:07:08 | INFO | train_inner | epoch 122:   452 / 588 loss=3.587, nll_loss=2.788, word_ins=4.334, word_crf=1.202, length=2.178, ppl=12.01, wps=20343, ups=2.85, wpb=7127.5, bsz=272.7, num_updates=71600, lr=0.000186859, gnorm=1.763, train_wall=35, gb_free=9.9, wall=25576
2021-03-28 03:07:43 | INFO | train_inner | epoch 122:   552 / 588 loss=3.666, nll_loss=2.877, word_ins=4.409, word_crf=1.245, length=2.157, ppl=12.69, wps=19899.2, ups=2.85, wpb=6974.8, bsz=248.4, num_updates=71700, lr=0.000186728, gnorm=1.925, train_wall=35, gb_free=9.9, wall=25611
2021-03-28 03:07:56 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 03:07:59 | INFO | valid | epoch 122 | valid on 'valid' subset | loss 8.647 | nll_loss 6.028 | word_ins 7.279 | word_crf 4.424 | length 5.84 | ppl 400.93 | wps 59723.5 | wpb 5311.6 | bsz 208.1 | num_updates 71736 | best_loss 6.867
2021-03-28 03:07:59 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 122 @ 71736 updates
2021-03-28 03:07:59 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint122.pt
2021-03-28 03:08:00 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint122.pt
2021-03-28 03:08:03 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint122.pt (epoch 122 @ 71736 updates, score 8.647) (writing took 3.7049459200025012 seconds)
2021-03-28 03:08:03 | INFO | fairseq_cli.train | end of epoch 122 (average epoch stats below)
2021-03-28 03:08:03 | INFO | train | epoch 122 | loss 3.54 | nll_loss 2.747 | word_ins 4.297 | word_crf 1.174 | length 2.174 | ppl 11.63 | wps 19392.5 | ups 2.77 | wpb 6988.7 | bsz 272.5 | num_updates 71736 | lr 0.000186682 | gnorm 1.847 | train_wall 203 | gb_free 9.4 | wall 25631
2021-03-28 03:08:03 | INFO | fairseq.trainer | begin training epoch 123
2021-03-28 03:08:03 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 03:08:25 | INFO | train_inner | epoch 123:   64 / 588 loss=3.424, nll_loss=2.617, word_ins=4.183, word_crf=1.112, length=2.2, ppl=10.73, wps=16602.5, ups=2.39, wpb=6953.2, bsz=294.9, num_updates=71800, lr=0.000186598, gnorm=2.109, train_wall=34, gb_free=9.5, wall=25653
2021-03-28 03:09:00 | INFO | train_inner | epoch 123:   164 / 588 loss=3.48, nll_loss=2.692, word_ins=4.251, word_crf=1.137, length=2.17, ppl=11.16, wps=19797.8, ups=2.86, wpb=6920.8, bsz=275.7, num_updates=71900, lr=0.000186469, gnorm=1.747, train_wall=35, gb_free=9.4, wall=25688
2021-03-28 03:09:35 | INFO | train_inner | epoch 123:   264 / 588 loss=3.582, nll_loss=2.793, word_ins=4.339, word_crf=1.195, length=2.174, ppl=11.97, wps=20092.9, ups=2.86, wpb=7033.9, bsz=261.5, num_updates=72000, lr=0.000186339, gnorm=1.833, train_wall=35, gb_free=9.5, wall=25723
2021-03-28 03:10:10 | INFO | train_inner | epoch 123:   364 / 588 loss=3.48, nll_loss=2.68, word_ins=4.242, word_crf=1.125, length=2.338, ppl=11.16, wps=20359.8, ups=2.88, wpb=7074.9, bsz=264.6, num_updates=72100, lr=0.00018621, gnorm=1.713, train_wall=34, gb_free=9.6, wall=25758
2021-03-28 03:10:45 | INFO | train_inner | epoch 123:   464 / 588 loss=3.598, nll_loss=2.801, word_ins=4.341, word_crf=1.217, length=2.11, ppl=12.11, wps=19792.7, ups=2.84, wpb=6966.3, bsz=282.7, num_updates=72200, lr=0.000186081, gnorm=2.019, train_wall=35, gb_free=9.8, wall=25793
2021-03-28 03:11:20 | INFO | train_inner | epoch 123:   564 / 588 loss=3.586, nll_loss=2.793, word_ins=4.335, word_crf=1.2, length=2.187, ppl=12.01, wps=20067.3, ups=2.86, wpb=7018.3, bsz=272.6, num_updates=72300, lr=0.000185952, gnorm=1.618, train_wall=35, gb_free=9.3, wall=25828
2021-03-28 03:11:28 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 03:11:31 | INFO | valid | epoch 123 | valid on 'valid' subset | loss 8.873 | nll_loss 6.135 | word_ins 7.399 | word_crf 4.59 | length 5.833 | ppl 468.86 | wps 60224.1 | wpb 5311.6 | bsz 208.1 | num_updates 72324 | best_loss 6.867
2021-03-28 03:11:31 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 123 @ 72324 updates
2021-03-28 03:11:31 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint123.pt
2021-03-28 03:11:32 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint123.pt
2021-03-28 03:11:35 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint123.pt (epoch 123 @ 72324 updates, score 8.873) (writing took 3.6639297879992228 seconds)
2021-03-28 03:11:35 | INFO | fairseq_cli.train | end of epoch 123 (average epoch stats below)
2021-03-28 03:11:35 | INFO | train | epoch 123 | loss 3.526 | nll_loss 2.733 | word_ins 4.285 | word_crf 1.164 | length 2.192 | ppl 11.52 | wps 19358.8 | ups 2.77 | wpb 6988.7 | bsz 272.5 | num_updates 72324 | lr 0.000185921 | gnorm 1.809 | train_wall 204 | gb_free 9.7 | wall 25843
2021-03-28 03:11:35 | INFO | fairseq.trainer | begin training epoch 124
2021-03-28 03:11:35 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 03:12:02 | INFO | train_inner | epoch 124:   76 / 588 loss=3.475, nll_loss=2.691, word_ins=4.252, word_crf=1.136, length=2.137, ppl=11.12, wps=16622.5, ups=2.39, wpb=6963.9, bsz=265.1, num_updates=72400, lr=0.000185824, gnorm=1.796, train_wall=34, gb_free=9.5, wall=25870
2021-03-28 03:12:36 | INFO | train_inner | epoch 124:   176 / 588 loss=3.423, nll_loss=2.636, word_ins=4.203, word_crf=1.108, length=2.132, ppl=10.73, wps=20082.7, ups=2.9, wpb=6927.4, bsz=282.5, num_updates=72500, lr=0.000185695, gnorm=1.94, train_wall=34, gb_free=10, wall=25904
2021-03-28 03:13:11 | INFO | train_inner | epoch 124:   276 / 588 loss=3.564, nll_loss=2.768, word_ins=4.316, word_crf=1.188, length=2.185, ppl=11.83, wps=20088, ups=2.88, wpb=6986.9, bsz=270.7, num_updates=72600, lr=0.000185567, gnorm=1.842, train_wall=35, gb_free=9.2, wall=25939
2021-03-28 03:13:45 | INFO | train_inner | epoch 124:   376 / 588 loss=3.347, nll_loss=2.562, word_ins=4.142, word_crf=1.06, length=2.163, ppl=10.18, wps=20256.2, ups=2.95, wpb=6877.5, bsz=287.5, num_updates=72700, lr=0.00018544, gnorm=1.616, train_wall=34, gb_free=9.7, wall=25973
2021-03-28 03:14:20 | INFO | train_inner | epoch 124:   476 / 588 loss=3.719, nll_loss=2.948, word_ins=4.473, word_crf=1.275, length=2.079, ppl=13.17, wps=19686.2, ups=2.83, wpb=6966.7, bsz=237.4, num_updates=72800, lr=0.000185312, gnorm=1.931, train_wall=35, gb_free=9.1, wall=26008
2021-03-28 03:14:55 | INFO | train_inner | epoch 124:   576 / 588 loss=3.523, nll_loss=2.719, word_ins=4.271, word_crf=1.164, length=2.237, ppl=11.5, wps=20435.8, ups=2.87, wpb=7117, bsz=288.1, num_updates=72900, lr=0.000185185, gnorm=1.734, train_wall=35, gb_free=9.1, wall=26043
2021-03-28 03:14:59 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 03:15:03 | INFO | valid | epoch 124 | valid on 'valid' subset | loss 8.862 | nll_loss 6.124 | word_ins 7.386 | word_crf 4.575 | length 5.938 | ppl 465.28 | wps 59653.6 | wpb 5311.6 | bsz 208.1 | num_updates 72912 | best_loss 6.867
2021-03-28 03:15:03 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 124 @ 72912 updates
2021-03-28 03:15:03 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint124.pt
2021-03-28 03:15:04 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint124.pt
2021-03-28 03:15:06 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint124.pt (epoch 124 @ 72912 updates, score 8.862) (writing took 3.645078977999219 seconds)
2021-03-28 03:15:06 | INFO | fairseq_cli.train | end of epoch 124 (average epoch stats below)
2021-03-28 03:15:06 | INFO | train | epoch 124 | loss 3.508 | nll_loss 2.718 | word_ins 4.274 | word_crf 1.154 | length 2.168 | ppl 11.37 | wps 19448.6 | ups 2.78 | wpb 6988.7 | bsz 272.5 | num_updates 72912 | lr 0.00018517 | gnorm 1.806 | train_wall 203 | gb_free 9.5 | wall 26054
2021-03-28 03:15:07 | INFO | fairseq.trainer | begin training epoch 125
2021-03-28 03:15:07 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 03:15:38 | INFO | train_inner | epoch 125:   88 / 588 loss=3.528, nll_loss=2.744, word_ins=4.297, word_crf=1.159, length=2.208, ppl=11.54, wps=16597, ups=2.36, wpb=7043.2, bsz=257.9, num_updates=73000, lr=0.000185058, gnorm=1.875, train_wall=35, gb_free=9.9, wall=26086
2021-03-28 03:16:13 | INFO | train_inner | epoch 125:   188 / 588 loss=3.456, nll_loss=2.672, word_ins=4.234, word_crf=1.128, length=2.115, ppl=10.98, wps=19744.5, ups=2.82, wpb=6997.4, bsz=282.5, num_updates=73100, lr=0.000184932, gnorm=1.753, train_wall=35, gb_free=9.7, wall=26121
2021-03-28 03:16:48 | INFO | train_inner | epoch 125:   288 / 588 loss=3.496, nll_loss=2.693, word_ins=4.251, word_crf=1.144, length=2.273, ppl=11.28, wps=19939.4, ups=2.83, wpb=7036.9, bsz=279, num_updates=73200, lr=0.000184805, gnorm=1.897, train_wall=35, gb_free=9.4, wall=26156

2021-03-28 03:17:25 | INFO | train_inner | epoch 125:   388 / 588 loss=3.598, nll_loss=2.82, word_ins=4.362, word_crf=1.212, length=2.045, ppl=12.11, wps=19218.8, ups=2.76, wpb=6955.4, bsz=260.2, num_updates=73300, lr=0.000184679, gnorm=1.721, train_wall=36, gb_free=10.4, wall=26193
2021-03-28 03:18:00 | INFO | train_inner | epoch 125:   488 / 588 loss=3.425, nll_loss=2.634, word_ins=4.201, word_crf=1.112, length=2.128, ppl=10.74, wps=19649.9, ups=2.85, wpb=6888, bsz=282.5, num_updates=73400, lr=0.000184553, gnorm=1.768, train_wall=35, gb_free=9.4, wall=26228
2021-03-28 03:18:34 | INFO | train_inner | epoch 125:   588 / 588 loss=3.475, nll_loss=2.681, word_ins=4.243, word_crf=1.129, length=2.246, ppl=11.12, wps=20486.4, ups=2.9, wpb=7062.3, bsz=274.2, num_updates=73500, lr=0.000184428, gnorm=1.646, train_wall=34, gb_free=9.6, wall=26262
2021-03-28 03:18:34 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 03:18:37 | INFO | valid | epoch 125 | valid on 'valid' subset | loss 8.828 | nll_loss 6.09 | word_ins 7.35 | word_crf 4.535 | length 6.179 | ppl 454.32 | wps 59376.3 | wpb 5311.6 | bsz 208.1 | num_updates 73500 | best_loss 6.867
2021-03-28 03:18:37 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 125 @ 73500 updates
2021-03-28 03:18:37 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint125.pt
2021-03-28 03:18:38 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint125.pt
2021-03-28 03:18:41 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint125.pt (epoch 125 @ 73500 updates, score 8.828) (writing took 3.6408715919969836 seconds)
2021-03-28 03:18:41 | INFO | fairseq_cli.train | end of epoch 125 (average epoch stats below)
2021-03-28 03:18:41 | INFO | train | epoch 125 | loss 3.496 | nll_loss 2.708 | word_ins 4.265 | word_crf 1.147 | length 2.162 | ppl 11.28 | wps 19135.5 | ups 2.74 | wpb 6988.7 | bsz 272.5 | num_updates 73500 | lr 0.000184428 | gnorm 1.781 | train_wall 206 | gb_free 9.6 | wall 26269
2021-03-28 03:18:41 | INFO | fairseq.trainer | begin training epoch 126
2021-03-28 03:18:41 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 03:19:16 | INFO | train_inner | epoch 126:   100 / 588 loss=3.457, nll_loss=2.684, word_ins=4.245, word_crf=1.123, length=2.113, ppl=10.98, wps=16766.8, ups=2.38, wpb=7049.4, bsz=273.2, num_updates=73600, lr=0.000184302, gnorm=1.783, train_wall=35, gb_free=9.7, wall=26304
2021-03-28 03:19:51 | INFO | train_inner | epoch 126:   200 / 588 loss=3.462, nll_loss=2.675, word_ins=4.235, word_crf=1.13, length=2.148, ppl=11.02, wps=19859.9, ups=2.88, wpb=6905.5, bsz=276.9, num_updates=73700, lr=0.000184177, gnorm=1.783, train_wall=35, gb_free=9.4, wall=26339
2021-03-28 03:20:26 | INFO | train_inner | epoch 126:   300 / 588 loss=3.398, nll_loss=2.604, word_ins=4.177, word_crf=1.09, length=2.196, ppl=10.54, wps=20222.7, ups=2.89, wpb=7005.6, bsz=285.8, num_updates=73800, lr=0.000184053, gnorm=1.668, train_wall=34, gb_free=9.7, wall=26374
2021-03-28 03:21:01 | INFO | train_inner | epoch 126:   400 / 588 loss=3.564, nll_loss=2.777, word_ins=4.322, word_crf=1.192, length=2.104, ppl=11.83, wps=19978.7, ups=2.85, wpb=6999.7, bsz=272, num_updates=73900, lr=0.000183928, gnorm=1.964, train_wall=35, gb_free=9.4, wall=26409
2021-03-28 03:21:36 | INFO | train_inner | epoch 126:   500 / 588 loss=3.553, nll_loss=2.77, word_ins=4.319, word_crf=1.178, length=2.155, ppl=11.74, wps=19358.6, ups=2.79, wpb=6935.2, bsz=257.9, num_updates=74000, lr=0.000183804, gnorm=1.859, train_wall=36, gb_free=9.7, wall=26444
2021-03-28 03:22:07 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 03:22:11 | INFO | valid | epoch 126 | valid on 'valid' subset | loss 8.836 | nll_loss 6.126 | word_ins 7.379 | word_crf 4.552 | length 5.944 | ppl 457.12 | wps 60069.5 | wpb 5311.6 | bsz 208.1 | num_updates 74088 | best_loss 6.867
2021-03-28 03:22:11 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 126 @ 74088 updates
2021-03-28 03:22:11 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint126.pt
2021-03-28 03:22:11 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint126.pt
2021-03-28 03:22:14 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint126.pt (epoch 126 @ 74088 updates, score 8.836) (writing took 3.633899645999918 seconds)
2021-03-28 03:22:14 | INFO | fairseq_cli.train | end of epoch 126 (average epoch stats below)
2021-03-28 03:22:14 | INFO | train | epoch 126 | loss 3.494 | nll_loss 2.708 | word_ins 4.265 | word_crf 1.147 | length 2.148 | ppl 11.27 | wps 19288.8 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 74088 | lr 0.000183694 | gnorm 1.832 | train_wall 204 | gb_free 9.7 | wall 26482
2021-03-28 03:22:14 | INFO | fairseq.trainer | begin training epoch 127
2021-03-28 03:22:14 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 03:22:19 | INFO | train_inner | epoch 127:   12 / 588 loss=3.556, nll_loss=2.76, word_ins=4.308, word_crf=1.185, length=2.162, ppl=11.76, wps=16717.4, ups=2.37, wpb=7062.9, bsz=271, num_updates=74100, lr=0.00018368, gnorm=1.917, train_wall=35, gb_free=9.6, wall=26487
2021-03-28 03:22:53 | INFO | train_inner | epoch 127:   112 / 588 loss=3.359, nll_loss=2.587, word_ins=4.162, word_crf=1.071, length=2.066, ppl=10.26, wps=20164.3, ups=2.89, wpb=6989.1, bsz=284.6, num_updates=74200, lr=0.000183556, gnorm=1.694, train_wall=34, gb_free=9.4, wall=26521
2021-03-28 03:23:28 | INFO | train_inner | epoch 127:   212 / 588 loss=3.402, nll_loss=2.605, word_ins=4.178, word_crf=1.093, length=2.209, ppl=10.57, wps=20454, ups=2.89, wpb=7068.5, bsz=275.8, num_updates=74300, lr=0.000183432, gnorm=1.864, train_wall=34, gb_free=9.5, wall=26556
2021-03-28 03:24:03 | INFO | train_inner | epoch 127:   312 / 588 loss=3.573, nll_loss=2.797, word_ins=4.344, word_crf=1.188, length=2.136, ppl=11.9, wps=20225.5, ups=2.86, wpb=7083.1, bsz=251.2, num_updates=74400, lr=0.000183309, gnorm=1.931, train_wall=35, gb_free=9.4, wall=26591
2021-03-28 03:24:38 | INFO | train_inner | epoch 127:   412 / 588 loss=3.484, nll_loss=2.716, word_ins=4.274, word_crf=1.15, length=1.975, ppl=11.19, wps=19582.5, ups=2.86, wpb=6843.7, bsz=270.6, num_updates=74500, lr=0.000183186, gnorm=1.739, train_wall=35, gb_free=9.5, wall=26626
2021-03-28 03:25:12 | INFO | train_inner | epoch 127:   512 / 588 loss=3.593, nll_loss=2.802, word_ins=4.343, word_crf=1.203, length=2.184, ppl=12.07, wps=20078.5, ups=2.89, wpb=6939.9, bsz=271.6, num_updates=74600, lr=0.000183063, gnorm=2.015, train_wall=34, gb_free=9.7, wall=26660
2021-03-28 03:25:39 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 03:25:42 | INFO | valid | epoch 127 | valid on 'valid' subset | loss 8.85 | nll_loss 6.13 | word_ins 7.387 | word_crf 4.553 | length 6.028 | ppl 461.3 | wps 59660.5 | wpb 5311.6 | bsz 208.1 | num_updates 74676 | best_loss 6.867
2021-03-28 03:25:42 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 127 @ 74676 updates
2021-03-28 03:25:42 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint127.pt
2021-03-28 03:25:43 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint127.pt
2021-03-28 03:25:46 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint127.pt (epoch 127 @ 74676 updates, score 8.85) (writing took 3.6577043830002367 seconds)
2021-03-28 03:25:46 | INFO | fairseq_cli.train | end of epoch 127 (average epoch stats below)
2021-03-28 03:25:46 | INFO | train | epoch 127 | loss 3.483 | nll_loss 2.7 | word_ins 4.258 | word_crf 1.142 | length 2.126 | ppl 11.18 | wps 19415.7 | ups 2.78 | wpb 6988.7 | bsz 272.5 | num_updates 74676 | lr 0.00018297 | gnorm 1.849 | train_wall 203 | gb_free 9.6 | wall 26694
2021-03-28 03:25:46 | INFO | fairseq.trainer | begin training epoch 128
2021-03-28 03:25:46 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 03:25:54 | INFO | train_inner | epoch 128:   24 / 588 loss=3.469, nll_loss=2.676, word_ins=4.236, word_crf=1.13, length=2.209, ppl=11.07, wps=16619.2, ups=2.39, wpb=6965.6, bsz=277, num_updates=74700, lr=0.00018294, gnorm=1.857, train_wall=34, gb_free=9.7, wall=26702
2021-03-28 03:26:29 | INFO | train_inner | epoch 128:   124 / 588 loss=3.477, nll_loss=2.714, word_ins=4.273, word_crf=1.134, length=2.058, ppl=11.13, wps=19696.2, ups=2.85, wpb=6911, bsz=255.4, num_updates=74800, lr=0.000182818, gnorm=1.942, train_wall=35, gb_free=9.5, wall=26737
2021-03-28 03:27:04 | INFO | train_inner | epoch 128:   224 / 588 loss=3.457, nll_loss=2.675, word_ins=4.237, word_crf=1.128, length=2.101, ppl=10.98, wps=20149.4, ups=2.89, wpb=6983.3, bsz=283.3, num_updates=74900, lr=0.000182696, gnorm=2.086, train_wall=34, gb_free=9.3, wall=26772
2021-03-28 03:27:39 | INFO | train_inner | epoch 128:   324 / 588 loss=3.547, nll_loss=2.761, word_ins=4.31, word_crf=1.18, length=2.119, ppl=11.69, wps=19971.3, ups=2.85, wpb=7004.9, bsz=270, num_updates=75000, lr=0.000182574, gnorm=2.015, train_wall=35, gb_free=9.5, wall=26807
2021-03-28 03:28:15 | INFO | train_inner | epoch 128:   424 / 588 loss=3.612, nll_loss=2.85, word_ins=4.389, word_crf=1.22, length=1.975, ppl=12.23, wps=19660.5, ups=2.83, wpb=6947.8, bsz=241.3, num_updates=75100, lr=0.000182453, gnorm=1.85, train_wall=35, gb_free=9.8, wall=26842
2021-03-28 03:28:49 | INFO | train_inner | epoch 128:   524 / 588 loss=3.415, nll_loss=2.604, word_ins=4.174, word_crf=1.104, length=2.241, ppl=10.67, wps=20296.2, ups=2.9, wpb=7010.6, bsz=300.3, num_updates=75200, lr=0.000182331, gnorm=1.827, train_wall=34, gb_free=9.7, wall=26877
2021-03-28 03:29:11 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 03:29:15 | INFO | valid | epoch 128 | valid on 'valid' subset | loss 8.872 | nll_loss 6.13 | word_ins 7.39 | word_crf 4.574 | length 6.03 | ppl 468.59 | wps 59927.6 | wpb 5311.6 | bsz 208.1 | num_updates 75264 | best_loss 6.867
2021-03-28 03:29:15 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 128 @ 75264 updates
2021-03-28 03:29:15 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint128.pt
2021-03-28 03:29:15 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint128.pt
2021-03-28 03:29:18 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint128.pt (epoch 128 @ 75264 updates, score 8.872) (writing took 3.6453784149998683 seconds)
2021-03-28 03:29:18 | INFO | fairseq_cli.train | end of epoch 128 (average epoch stats below)

2021-03-28 03:29:18 | INFO | train | epoch 128 | loss 3.492 | nll_loss 2.707 | word_ins 4.264 | word_crf 1.146 | length 2.137 | ppl 11.25 | wps 19351.1 | ups 2.77 | wpb 6988.7 | bsz 272.5 | num_updates 75264 | lr 0.000182254 | gnorm 1.922 | train_wall 204 | gb_free 9.7 | wall 26906
2021-03-28 03:29:18 | INFO | fairseq.trainer | begin training epoch 129
2021-03-28 03:29:18 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 03:29:31 | INFO | train_inner | epoch 129:   36 / 588 loss=3.35, nll_loss=2.549, word_ins=4.127, word_crf=1.058, length=2.279, ppl=10.2, wps=17081.3, ups=2.4, wpb=7116.2, bsz=296.2, num_updates=75300, lr=0.00018221, gnorm=1.76, train_wall=34, gb_free=9.4, wall=26919
2021-03-28 03:30:05 | INFO | train_inner | epoch 129: 136 / 588 loss=3.413, nll_loss=2.626, word_ins=4.193, word_crf=1.104, length=2.125, ppl=10.65, wps=20732.6, ups=2.9, wpb=7153.5, bsz=298.1, num_updates=75400, lr=0.000182089, gnorm=1.823, train_wall=34, gb_free=9.9, wall=26953
2021-03-28 03:30:40 | INFO | train_inner | epoch 129: 236 / 588 loss=3.461, nll_loss=2.687, word_ins=4.248, word_crf=1.128, length=2.087, ppl=11.01, wps=19784.6, ups=2.88, wpb=6867.7, bsz=269.5, num_updates=75500, lr=0.000181969, gnorm=1.95, train_wall=34, gb_free=9.4, wall=26988
2021-03-28 03:31:15 | INFO | train_inner | epoch 129: 336 / 588 loss=3.546, nll_loss=2.772, word_ins=4.321, word_crf=1.176, length=2.092, ppl=11.68, wps=19634.8, ups=2.82, wpb=6953.7, bsz=259.2, num_updates=75600, lr=0.000181848, gnorm=1.983, train_wall=35, gb_free=9.4, wall=27023
2021-03-28 03:31:50 | INFO | train_inner | epoch 129: 436 / 588 loss=3.249, nll_loss=2.451, word_ins=4.046, word_crf=0.997, length=2.285, ppl=9.51, wps=19897.1, ups=2.86, wpb=6961.2, bsz=302.1, num_updates=75700, lr=0.000181728, gnorm=1.632, train_wall=35, gb_free=9.9, wall=27058
2021-03-28 03:32:26 | INFO | train_inner | epoch 129: 536 / 588 loss=3.686, nll_loss=2.905, word_ins=4.436, word_crf=1.258, length=2.106, ppl=12.87, wps=19589.6, ups=2.8, wpb=7001.6, bsz=236.2, num_updates=75800, lr=0.000181608, gnorm=1.815, train_wall=35, gb_free=10.1, wall=27094
2021-03-28 03:32:45 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 03:32:48 | INFO | valid | epoch 129 | valid on 'valid' subset | loss 8.888 | nll_loss 6.136 | word_ins 7.396 | word_crf 4.592 | length 5.976 | ppl 473.73 | wps 59526.4 | wpb 5311.6 | bsz 208.1 | num_updates 75852 | best_loss 6.867
2021-03-28 03:32:48 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 129 @ 75852 updates
2021-03-28 03:32:48 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint129.pt
2021-03-28 03:32:49 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint129.pt
2021-03-28 03:32:52 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint129.pt (epoch 129 @ 75852 updates, score 8.888) (writing took 3.6460389049971127 seconds)
2021-03-28 03:32:52 | INFO | fairseq_cli.train | end of epoch 129 (average epoch stats below)
2021-03-28 03:32:52 | INFO | train | epoch 129 | loss 3.469 | nll_loss 2.687 | word_ins 4.248 | word_crf 1.132 | length 2.131 | ppl 11.08 | wps 19257.3 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 75852 | lr 0.000181546 | gnorm 1.856 | train_wall 205 | gb_free 9.9 | wall 27119
2021-03-28 03:32:52 | INFO | fairseq.trainer | begin training epoch 130
2021-03-28 03:32:52 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 03:33:09 | INFO | train_inner | epoch 130:   48 / 588 loss=3.52, nll_loss=2.735, word_ins=4.29, word_crf=1.166, length=2.091, ppl=11.47, wps=16368, ups=2.34, wpb=6981.8, bsz=264.5, num_updates=75900, lr=0.000181489, gnorm=2.013, train_wall=35, gb_free=9.4, wall=27137
2021-03-28 03:33:43 | INFO | train_inner | epoch 130: 148 / 588 loss=3.39, nll_loss=2.609, word_ins=4.182, word_crf=1.081, length=2.19, ppl=10.49, wps=20369.5, ups=2.88, wpb=7064.3, bsz=273.4, num_updates=76000, lr=0.000181369, gnorm=1.827, train_wall=34, gb_free=9.5, wall=27171
2021-03-28 03:34:18 | INFO | train_inner | epoch 130: 248 / 588 loss=3.507, nll_loss=2.744, word_ins=4.299, word_crf=1.153, length=2.047, ppl=11.37, wps=19711, ups=2.88, wpb=6848, bsz=249.5, num_updates=76100, lr=0.00018125, gnorm=2.056, train_wall=34, gb_free=9.5, wall=27206
2021-03-28 03:34:54 | INFO | train_inner | epoch 130: 348 / 588 loss=3.464, nll_loss=2.686, word_ins=4.247, word_crf=1.13, length=2.104, ppl=11.03, wps=19731, ups=2.79, wpb=7072.2, bsz=274.3, num_updates=76200, lr=0.000181131, gnorm=1.795, train_wall=36, gb_free=9.7, wall=27242
2021-03-28 03:35:29 | INFO | train_inner | epoch 130: 448 / 588 loss=3.442, nll_loss=2.641, word_ins=4.206, word_crf=1.118, length=2.214, ppl=10.87, wps=20376.1, ups=2.89, wpb=7053.7, bsz=291.8, num_updates=76300, lr=0.000181012, gnorm=1.789, train_wall=34, gb_free=9.1, wall=27277
2021-03-28 03:36:04 | INFO | train_inner | epoch 130: 548 / 588 loss=3.542, nll_loss=2.76, word_ins=4.31, word_crf=1.184, length=2.024, ppl=11.65, wps=19843.7, ups=2.86, wpb=6950.4, bsz=271.8, num_updates=76400, lr=0.000180894, gnorm=1.977, train_wall=35, gb_free=9.5, wall=27312
2021-03-28 03:36:18 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 03:36:21 | INFO | valid | epoch 130 | valid on 'valid' subset | loss 8.994 | nll_loss 6.222 | word_ins 7.479 | word_crf 4.665 | length 5.888 | ppl 509.79 | wps 59939.3 | wpb 5311.6 | bsz 208.1 | num_updates 76440 | best_loss 6.867
2021-03-28 03:36:21 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 130 @ 76440 updates
2021-03-28 03:36:21 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint130.pt
2021-03-28 03:36:22 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint130.pt
2021-03-28 03:36:25 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint130.pt (epoch 130 @ 76440 updates, score 8.994) (writing took 3.6331691929990484 seconds)
2021-03-28 03:36:25 | INFO | fairseq_cli.train | end of epoch 130 (average epoch stats below)
2021-03-28 03:36:25 | INFO | train | epoch 130 | loss 3.466 | nll_loss 2.685 | word_ins 4.246 | word_crf 1.131 | length 2.113 | ppl 11.05 | wps 19298.1 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 76440 | lr 0.000180846 | gnorm 1.895 | train_wall 204 | gb_free 9.5 | wall 27332
2021-03-28 03:36:25 | INFO | fairseq.trainer | begin training epoch 131
2021-03-28 03:36:25 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 03:36:46 | INFO | train_inner | epoch 131:   60 / 588 loss=3.423, nll_loss=2.657, word_ins=4.22, word_crf=1.106, length=2.077, ppl=10.73, wps=16668.2, ups=2.39, wpb=6978.1, bsz=278, num_updates=76500, lr=0.000180775, gnorm=1.969, train_wall=34, gb_free=9.3, wall=27353
2021-03-28 03:37:21 | INFO | train_inner | epoch 131: 160 / 588 loss=3.494, nll_loss=2.728, word_ins=4.285, word_crf=1.148, length=2.03, ppl=11.27, wps=19897.2, ups=2.86, wpb=6962.7, bsz=256.4, num_updates=76600, lr=0.000180657, gnorm=1.918, train_wall=35, gb_free=10.1, wall=27388
2021-03-28 03:37:55 | INFO | train_inner | epoch 131: 260 / 588 loss=3.452, nll_loss=2.682, word_ins=4.244, word_crf=1.124, length=2.059, ppl=10.94, wps=19966.2, ups=2.89, wpb=6902.2, bsz=274.5, num_updates=76700, lr=0.00018054, gnorm=1.774, train_wall=34, gb_free=9.8, wall=27423
2021-03-28 03:38:29 | INFO | train_inner | epoch 131: 360 / 588 loss=3.296, nll_loss=2.516, word_ins=4.103, word_crf=1.027, length=2.176, ppl=9.82, wps=20409, ups=2.93, wpb=6962.5, bsz=281.3, num_updates=76800, lr=0.000180422, gnorm=1.726, train_wall=34, gb_free=9.1, wall=27457
2021-03-28 03:39:04 | INFO | train_inner | epoch 131: 460 / 588 loss=3.525, nll_loss=2.741, word_ins=4.295, word_crf=1.17, length=2.081, ppl=11.51, wps=20078.7, ups=2.87, wpb=6995.4, bsz=262.3, num_updates=76900, lr=0.000180305, gnorm=1.974, train_wall=35, gb_free=9.4, wall=27492
2021-03-28 03:39:39 | INFO | train_inner | epoch 131: 560 / 588 loss=3.469, nll_loss=2.667, word_ins=4.229, word_crf=1.137, length=2.183, ppl=11.08, wps=20265.8, ups=2.87, wpb=7065.3, bsz=287.7, num_updates=77000, lr=0.000180187, gnorm=1.886, train_wall=35, gb_free=9.4, wall=27527
2021-03-28 03:39:49 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 03:39:52 | INFO | valid | epoch 131 | valid on 'valid' subset | loss 8.989 | nll_loss 6.199 | word_ins 7.457 | word_crf 4.652 | length 6.08 | ppl 508.08 | wps 59656.9 | wpb 5311.6 | bsz 208.1 | num_updates 77028 | best_loss 6.867
2021-03-28 03:39:52 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 131 @ 77028 updates
2021-03-28 03:39:52 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint131.pt
2021-03-28 03:39:53 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint131.pt
2021-03-28 03:39:56 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint131.pt (epoch 131 @ 77028 updates, score 8.989) (writing took 3.6611635490007757 seconds)
2021-03-28 03:39:56 | INFO | fairseq_cli.train | end of epoch 131 (average epoch stats below)
2021-03-28 03:39:56 | INFO | train | epoch 131 | loss 3.45 | nll_loss 2.673 | word_ins 4.236 | word_crf 1.123 | length 2.088 | ppl 10.93 | wps 19428.7 | ups 2.78 | wpb 6988.7 | bsz 272.5 | num_updates 77028 | lr 0.000180155 | gnorm 1.87 | train_wall 203 | gb_free 9.3 | wall 27544
2021-03-28 03:39:56 | INFO | fairseq.trainer | begin training epoch 132
2021-03-28 03:39:56 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 03:40:21 | INFO | train_inner | epoch 132:   72 / 588 loss=3.367, nll_loss=2.607, word_ins=4.181, word_crf=1.076, length=2, ppl=10.32, wps=16673.6, ups=2.37, wpb=7039, bsz=268.8, num_updates=77100, lr=0.000180071, gnorm=1.695, train_wall=35, gb_free=9.8, wall=27569
2021-03-28 03:40:56 | INFO | train_inner | epoch 132: 172 / 588 loss=3.528, nll_loss=2.75, word_ins=4.303, word_crf=1.164, length=2.123, ppl=11.53, wps=20344, ups=2.86, wpb=7119.8, bsz=265.9, num_updates=77200, lr=0.000179954, gnorm=2.134, train_wall=35, gb_free=9.8, wall=27604
2021-03-28 03:41:31 | INFO | train_inner | epoch 132: 272 / 588 loss=3.353, nll_loss=2.585, word_ins=4.16, word_crf=1.069, length=2.046, ppl=10.22, wps=19644.8, ups=2.89, wpb=6790.2, bsz=290.3, num_updates=77300, lr=0.000179838, gnorm=1.791, train_wall=34, gb_free=9.3, wall=27639
2021-03-28 03:42:06 | INFO | train_inner | epoch 132: 372 / 588 loss=3.442, nll_loss=2.675, word_ins=4.238, word_crf=1.124, length=1.99, ppl=10.87, wps=19990, ups=2.85, wpb=7015.4, bsz=270.5, num_updates=77400, lr=0.000179721, gnorm=1.744, train_wall=35, gb_free=9.8, wall=27674

2021-03-28 03:42:40 | INFO | train_inner | epoch 132: 472 / 588 loss=3.456, nll_loss=2.659, word_ins=4.226, word_crf=1.116, length=2.266, ppl=10.97, wps=20508.6, ups=2.89, wpb=7086.8, bsz=266.4, num_updates=77500, lr=0.000179605, gnorm=1.722, train_wall=34, gb_free=9.4, wall=27708
2021-03-28 03:43:16 | INFO | train_inner | epoch 132: 572 / 588 loss=3.56, nll_loss=2.78, word_ins=4.329, word_crf=1.187, length=2.08, ppl=11.79, wps=19521.5, ups=2.83, wpb=6897.4, bsz=256.1, num_updates=77600, lr=0.00017949, gnorm=1.935, train_wall=35, gb_free=9.1, wall=27744
2021-03-28 03:43:21 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 03:43:25 | INFO | valid | epoch 132 | valid on 'valid' subset | loss 9.098 | nll_loss 6.27 | word_ins 7.525 | word_crf 4.732 | length 6.039 | ppl 548.1 | wps 60292.2 | wpb 5311.6 | bsz 208.1 | num_updates 77616 | best_loss 6.867
2021-03-28 03:43:25 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 132 @ 77616 updates
2021-03-28 03:43:25 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint132.pt
2021-03-28 03:43:25 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint132.pt
2021-03-28 03:43:28 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint132.pt (epoch 132 @ 77616 updates, score 9.098) (writing took 3.6440849719983817 seconds)
2021-03-28 03:43:28 | INFO | fairseq_cli.train | end of epoch 132 (average epoch stats below)
2021-03-28 03:43:28 | INFO | train | epoch 132 | loss 3.439 | nll_loss 2.662 | word_ins 4.227 | word_crf 1.116 | length 2.095 | ppl 10.84 | wps 19368.5 | ups 2.77 | wpb 6988.7 | bsz 272.5 | num_updates 77616 | lr 0.000179471 | gnorm 1.836 | train_wall 203 | gb_free 9.7 | wall 27756
2021-03-28 03:43:28 | INFO | fairseq.trainer | begin training epoch 133
2021-03-28 03:43:28 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 03:43:58 | INFO | train_inner | epoch 133: 84 / 588 loss=3.489, nll_loss=2.715, word_ins=4.27, word_crf=1.151, length=2.026, ppl=11.22, wps=16645.2, ups=2.38, wpb=7007.6, bsz=274.8, num_updates=77700, lr=0.000179374, gnorm=2.037, train_wall=35, gb_free=9.4, wall=27786
2021-03-28 03:44:32 | INFO | train_inner | epoch 133: 184 / 588 loss=3.352, nll_loss=2.56, word_ins=4.139, word_crf=1.065, length=2.175, ppl=10.21, wps=20183.8, ups=2.89, wpb=6972.9, bsz=286.6, num_updates=77800, lr=0.000179259, gnorm=1.843, train_wall=34, gb_free=9.6, wall=27820
2021-03-28 03:45:07 | INFO | train_inner | epoch 133: 284 / 588 loss=3.456, nll_loss=2.698, word_ins=4.261, word_crf=1.126, length=1.994, ppl=10.97, wps=19752.9, ups=2.87, wpb=6888.9, bsz=256.4, num_updates=77900, lr=0.000179144, gnorm=1.934, train_wall=35, gb_free=9.5, wall=27855
2021-03-28 03:45:42 | INFO | train_inner | epoch 133: 384 / 588 loss=3.457, nll_loss=2.679, word_ins=4.243, word_crf=1.122, length=2.139, ppl=10.98, wps=20494.2, ups=2.87, wpb=7133.9, bsz=265, num_updates=78000, lr=0.000179029, gnorm=1.833, train_wall=35, gb_free=9.1, wall=27890
2021-03-28 03:46:18 | INFO | train_inner | epoch 133: 484 / 588 loss=3.524, nll_loss=2.74, word_ins=4.295, word_crf=1.159, length=2.166, ppl=11.5, wps=18949.2, ups=2.77, wpb=6833.7, bsz=256, num_updates=78100, lr=0.000178914, gnorm=1.854, train_wall=36, gb_free=9.7, wall=27926
2021-03-28 03:46:53 | INFO | train_inner | epoch 133: 584 / 588 loss=3.304, nll_loss=2.51, word_ins=4.094, word_crf=1.045, length=2.124, ppl=9.88, wps=20602.5, ups=2.9, wpb=7101.4, bsz=312.7, num_updates=78200, lr=0.0001788, gnorm=1.714, train_wall=34, gb_free=9.5, wall=27960
2021-03-28 03:46:54 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 03:46:57 | INFO | valid | epoch 133 | valid on 'valid' subset | loss 8.914 | nll_loss 6.148 | word_ins 7.404 | word_crf 4.602 | length 6.099 | ppl 482.45 | wps 59612.4 | wpb 5311.6 | bsz 208.1 | num_updates 78204 | best_loss 6.867
2021-03-28 03:46:57 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 133 @ 78204 updates
2021-03-28 03:46:57 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint133.pt
2021-03-28 03:46:58 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint133.pt
2021-03-28 03:47:01 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint133.pt (epoch 133 @ 78204 updates, score 8.914) (writing took 3.638637377000123 seconds)
2021-03-28 03:47:01 | INFO | fairseq_cli.train | end of epoch 133 (average epoch stats below)
2021-03-28 03:47:01 | INFO | train | epoch 133 | loss 3.438 | nll_loss 2.66 | word_ins 4.226 | word_crf 1.116 | length 2.099 | ppl 10.84 | wps 19310.6 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 78204 | lr 0.000178795 | gnorm 1.879 | train_wall 204 | gb_free 9.1 | wall 27969
2021-03-28 03:47:01 | INFO | fairseq.trainer | begin training epoch 134
2021-03-28 03:47:01 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 03:47:35 | INFO | train_inner | epoch 134: 96 / 588 loss=3.494, nll_loss=2.739, word_ins=4.295, word_crf=1.153, length=1.935, ppl=11.27, wps=16435.8, ups=2.34, wpb=7028.8, bsz=257.7, num_updates=78300, lr=0.000178685, gnorm=1.847, train_wall=35, gb_free=9.5, wall=28003
2021-03-28 03:48:10 | INFO | train_inner | epoch 134: 196 / 588 loss=3.33, nll_loss=2.548, word_ins=4.13, word_crf=1.054, length=2.109, ppl=10.06, wps=19829.9, ups=2.85, wpb=6953.8, bsz=287.3, num_updates=78400, lr=0.000178571, gnorm=1.759, train_wall=35, gb_free=9.5, wall=28038
2021-03-28 03:48:45 | INFO | train_inner | epoch 134: 296 / 588 loss=3.384, nll_loss=2.618, word_ins=4.188, word_crf=1.09, length=2.004, ppl=10.44, wps=19881.8, ups=2.86, wpb=6941.5, bsz=286.2, num_updates=78500, lr=0.000178458, gnorm=1.962, train_wall=35, gb_free=9.7, wall=28073
2021-03-28 03:49:20 | INFO | train_inner | epoch 134: 396 / 588 loss=3.443, nll_loss=2.667, word_ins=4.233, word_crf=1.112, length=2.142, ppl=10.88, wps=20047.3, ups=2.85, wpb=7022.4, bsz=260.7, num_updates=78600, lr=0.000178344, gnorm=1.804, train_wall=35, gb_free=9.2, wall=28108
2021-03-28 03:49:55 | INFO | train_inner | epoch 134: 496 / 588 loss=3.497, nll_loss=2.723, word_ins=4.281, word_crf=1.152, length=2.049, ppl=11.29, wps=19782.5, ups=2.85, wpb=6944.1, bsz=265.6, num_updates=78700, lr=0.000178231, gnorm=1.957, train_wall=35, gb_free=9.4, wall=28143
2021-03-28 03:50:27 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 03:50:30 | INFO | valid | epoch 134 | valid on 'valid' subset | loss 8.983 | nll_loss 6.195 | word_ins 7.454 | word_crf 4.653 | length 6.027 | ppl 505.9 | wps 59556.4 | wpb 5311.6 | bsz 208.1 | num_updates 78792 | best_loss 6.867
2021-03-28 03:50:30 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 134 @ 78792 updates
2021-03-28 03:50:30 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint134.pt
2021-03-28 03:50:31 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint134.pt
2021-03-28 03:50:34 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint134.pt (epoch 134 @ 78792 updates, score 8.983) (writing took 3.643852589000744 seconds)
2021-03-28 03:50:34 | INFO | fairseq_cli.train | end of epoch 134 (average epoch stats below)
2021-03-28 03:50:34 | INFO | train | epoch 134 | loss 3.42 | nll_loss 2.646 | word_ins 4.214 | word_crf 1.105 | length 2.078 | ppl 10.7 | wps 19292.6 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 78792 | lr 0.000178127 | gnorm 1.837 | train_wall 204 | gb_free 9.9 | wall 28182
2021-03-28 03:50:34 | INFO | fairseq.trainer | begin training epoch 135
2021-03-28 03:50:34 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 03:50:37 | INFO | train_inner | epoch 135: 8 / 588 loss=3.402, nll_loss=2.619, word_ins=4.191, word_crf=1.087, length=2.196, ppl=10.57, wps=16818.5, ups=2.4, wpb=7006, bsz=269.7, num_updates=78800, lr=0.000178118, gnorm=1.713, train_wall=34, gb_free=9.2, wall=28185
2021-03-28 03:51:12 | INFO | train_inner | epoch 135: 108 / 588 loss=3.246, nll_loss=2.489, word_ins=4.079, word_crf=1.013, length=1.94, ppl=9.49, wps=20292.2, ups=2.89, wpb=7016.5, bsz=296.3, num_updates=78900, lr=0.000178005, gnorm=1.805, train_wall=34, gb_free=9.3, wall=28220
2021-03-28 03:51:47 | INFO | train_inner | epoch 135: 208 / 588 loss=3.376, nll_loss=2.629, word_ins=4.202, word_crf=1.079, length=1.966, ppl=10.38, wps=19381, ups=2.84, wpb=6813.1, bsz=262.2, num_updates=79000, lr=0.000177892, gnorm=1.854, train_wall=35, gb_free=10.1, wall=28255
2021-03-28 03:52:22 | INFO | train_inner | epoch 135: 308 / 588 loss=3.479, nll_loss=2.707, word_ins=4.267, word_crf=1.138, length=2.07, ppl=11.15, wps=19882.5, ups=2.84, wpb=6990.7, bsz=262.5, num_updates=79100, lr=0.00017778, gnorm=1.926, train_wall=35, gb_free=9.8, wall=28290
2021-03-28 03:52:57 | INFO | train_inner | epoch 135: 408 / 588 loss=3.452, nll_loss=2.634, word_ins=4.202, word_crf=1.121, length=2.292, ppl=10.94, wps=20324.7, ups=2.87, wpb=7087.6, bsz=287.9, num_updates=79200, lr=0.000177667, gnorm=1.934, train_wall=35, gb_free=9.4, wall=28325
2021-03-28 03:53:32 | INFO | train_inner | epoch 135: 508 / 588 loss=3.42, nll_loss=2.65, word_ins=4.219, word_crf=1.099, length=2.107, ppl=10.7, wps=20109.2, ups=2.85, wpb=7059.3, bsz=256.6, num_updates=79300, lr=0.000177555, gnorm=1.761, train_wall=35, gb_free=9.6, wall=28360
2021-03-28 03:54:00 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 03:54:04 | INFO | valid | epoch 135 | valid on 'valid' subset | loss 8.998 | nll_loss 6.197 | word_ins 7.455 | word_crf 4.66 | length 6.108 | ppl 511.38 | wps 59524 | wpb 5311.6 | bsz 208.1 | num_updates 79380 | best_loss 6.867
2021-03-28 03:54:04 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 135 @ 79380 updates
2021-03-28 03:54:04 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint135.pt
2021-03-28 03:54:05 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint135.pt
2021-03-28 03:54:07 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint135.pt (epoch 135 @ 79380 updates, score 8.998) (writing took 3.670523497999966 seconds)
2021-03-28 03:54:07 | INFO | fairseq_cli.train | end of epoch 135 (average epoch stats below)
2021-03-28 03:54:07 | INFO | train | epoch 135 | loss 3.413 | nll_loss 2.641 | word_ins 4.209 | word_crf 1.102 | length 2.063 | ppl 10.65 | wps 19265.2 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 79380 | lr 0.000177466 | gnorm 1.849 | train_wall 204 | gb_free 9.7 | wall 28395

2021-03-28 03:54:07 | INFO | fairseq.trainer | begin training epoch 136
2021-03-28 03:54:07 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 03:54:14 | INFO | train_inner | epoch 136:   20 / 588 loss=3.421, nll_loss=2.639, word_ins=4.206, word_crf=1.111, length=2.069, ppl=10.71, wps=16510.5, ups=2.37, wpb=6976.1, bsz=283, num_updates=79400, lr=0.000177443, gnorm=1.735, train_wall=35, gb_free=9.9, wall=28402
2021-03-28 03:54:50 | INFO | train_inner | epoch 136:   120 / 588 loss=3.301, nll_loss=2.537, word_ins=4.122, word_crf=1.038, length=2.022, ppl=9.86, wps=19614.7, ups=2.81, wpb=6980.6, bsz=277.8, num_updates=79500, lr=0.000177332, gnorm=1.782, train_wall=35, gb_free=9.5, wall=28438
2021-03-28 03:55:25 | INFO | train_inner | epoch 136:   220 / 588 loss=3.447, nll_loss=2.688, word_ins=4.254, word_crf=1.116, length=2.048, ppl=10.91, wps=19841.4, ups=2.86, wpb=6942.1, bsz=255.3, num_updates=79600, lr=0.00017722, gnorm=1.869, train_wall=35, gb_free=10.3, wall=28473
2021-03-28 03:56:00 | INFO | train_inner | epoch 136:   320 / 588 loss=3.469, nll_loss=2.713, word_ins=4.275, word_crf=1.136, length=1.957, ppl=11.07, wps=19741.4, ups=2.86, wpb=6899.1, bsz=257.2, num_updates=79700, lr=0.000177109, gnorm=1.996, train_wall=35, gb_free=9.4, wall=28508
2021-03-28 03:56:35 | INFO | train_inner | epoch 136:   420 / 588 loss=3.521, nll_loss=2.751, word_ins=4.304, word_crf=1.165, length=2.034, ppl=11.48, wps=19676.4, ups=2.83, wpb=6951.7, bsz=256.2, num_updates=79800, lr=0.000176998, gnorm=1.842, train_wall=35, gb_free=9.7, wall=28543
2021-03-28 03:57:09 | INFO | train_inner | epoch 136:   520 / 588 loss=3.264, nll_loss=2.456, word_ins=4.048, word_crf=1.016, length=2.235, ppl=9.6, wps=20707.7, ups=2.91, wpb=7110.1, bsz=316.1, num_updates=79900, lr=0.000176887, gnorm=1.838, train_wall=35, gb_free=9.9, wall=28577
2021-03-28 03:57:33 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 03:57:36 | INFO | valid | epoch 136 | valid on 'valid' subset | loss 8.987 | nll_loss 6.192 | word_ins 7.448 | word_crf 4.649 | length 6.138 | ppl 507.4 | wps 60233.9 | wpb 5311.6 | bsz 208.1 | num_updates 79968 | best_loss 6.867
2021-03-28 03:57:36 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 136 @ 79968 updates
2021-03-28 03:57:36 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint136.pt
2021-03-28 03:57:37 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint136.pt
2021-03-28 03:57:40 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint136.pt (epoch 136 @ 79968 updates, score 8.987) (writing took 3.6470541179987777 seconds)
2021-03-28 03:57:40 | INFO | fairseq_cli.train | end of epoch 136 (average epoch stats below)
2021-03-28 03:57:40 | INFO | train | epoch 136 | loss 3.407 | nll_loss 2.634 | word_ins 4.204 | word_crf 1.099 | length 2.061 | ppl 10.6 | wps 19310.7 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 79968 | lr 0.000176812 | gnorm 1.876 | train_wall 204 | gb_free 9.5 | wall 28608
2021-03-28 03:57:40 | INFO | fairseq.trainer | begin training epoch 137
2021-03-28 03:57:40 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 03:57:51 | INFO | train_inner | epoch 137:   32 / 588 loss=3.475, nll_loss=2.702, word_ins=4.261, word_crf=1.145, length=1.988, ppl=11.12, wps=16998.4, ups=2.38, wpb=7129.6, bsz=269, num_updates=80000, lr=0.000176777, gnorm=2.069, train_wall=35, gb_free=9.6, wall=28619
2021-03-28 03:57:51 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 03:57:55 | INFO | valid | epoch 137 | valid on 'valid' subset | loss 9.03 | nll_loss 6.218 | word_ins 7.478 | word_crf 4.682 | length 6.094 | ppl 522.88 | wps 60201.1 | wpb 5311.6 | bsz 208.1 | num_updates 80000 | best_loss 6.867
2021-03-28 03:57:55 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 137 @ 80000 updates
2021-03-28 03:57:55 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint_137_80000.pt
2021-03-28 03:57:56 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint_137_80000.pt
2021-03-28 03:57:59 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint_137_80000.pt (epoch 137 @ 80000 updates, score 9.03) (writing took 3.9976997329977166 seconds)
2021-03-28 03:58:33 | INFO | train_inner | epoch 137:   132 / 588 loss=3.28, nll_loss=2.5, word_ins=4.087, word_crf=1.023, length=2.13, ppl=9.71, wps=16783.3, ups=2.39, wpb=7017.6, bsz=296.6, num_updates=80100, lr=0.000176666, gnorm=1.814, train_wall=34, gb_free=9.3, wall=28661
2021-03-28 03:59:08 | INFO | train_inner | epoch 137:   232 / 588 loss=3.277, nll_loss=2.504, word_ins=4.093, word_crf=1.022, length=2.083, ppl=9.69, wps=20305.5, ups=2.9, wpb=6990.1, bsz=289.4, num_updates=80200, lr=0.000176556, gnorm=1.722, train_wall=34, gb_free=9.5, wall=28695
2021-03-28 03:59:43 | INFO | train_inner | epoch 137:   332 / 588 loss=3.485, nll_loss=2.721, word_ins=4.28, word_crf=1.145, length=2.009, ppl=11.2, wps=20051.8, ups=2.84, wpb=7050.4, bsz=259.7, num_updates=80300, lr=0.000176446, gnorm=2.01, train_wall=35, gb_free=9.2, wall=28731
2021-03-28 04:00:18 | INFO | train_inner | epoch 137:   432 / 588 loss=3.566, nll_loss=2.821, word_ins=4.365, word_crf=1.198, length=1.854, ppl=11.85, wps=19435.2, ups=2.81, wpb=6920.6, bsz=246.2, num_updates=80400, lr=0.000176336, gnorm=1.974, train_wall=35, gb_free=9.4, wall=28766
2021-03-28 04:00:53 | INFO | train_inner | epoch 137:   532 / 588 loss=3.394, nll_loss=2.622, word_ins=4.195, word_crf=1.087, length=2.099, ppl=10.52, wps=20004.8, ups=2.87, wpb=6959, bsz=271.4, num_updates=80500, lr=0.000176227, gnorm=1.713, train_wall=35, gb_free=10.6, wall=28801
2021-03-28 04:01:13 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 04:01:16 | INFO | valid | epoch 137 | valid on 'valid' subset | loss 8.984 | nll_loss 6.19 | word_ins 7.447 | word_crf 4.648 | length 6.116 | ppl 506.2 | wps 59767 | wpb 5311.6 | bsz 208.1 | num_updates 80556 | best_loss 6.867
2021-03-28 04:01:16 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 137 @ 80556 updates
2021-03-28 04:01:16 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint137.pt
2021-03-28 04:01:17 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint137.pt
2021-03-28 04:01:20 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint137.pt (epoch 137 @ 80556 updates, score 8.984) (writing took 3.652825641001982 seconds)
2021-03-28 04:01:20 | INFO | fairseq_cli.train | end of epoch 137 (average epoch stats below)
2021-03-28 04:01:20 | INFO | train | epoch 137 | loss 3.394 | nll_loss 2.625 | word_ins 4.197 | word_crf 1.091 | length 2.043 | ppl 10.51 | wps 18716.1 | ups 2.68 | wpb 6988.7 | bsz 272.5 | num_updates 80556 | lr 0.000176166 | gnorm 1.858 | train_wall 204 | gb_free 10.1 | wall 28828
2021-03-28 04:01:20 | INFO | fairseq.trainer | begin training epoch 138
2021-03-28 04:01:20 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 04:01:36 | INFO | train_inner | epoch 138:   44 / 588 loss=3.449, nll_loss=2.678, word_ins=4.243, word_crf=1.123, length=2.053, ppl=10.92, wps=16260.3, ups=2.35, wpb=6921.1, bsz=253.8, num_updates=80600, lr=0.000176117, gnorm=1.862, train_wall=35, gb_free=9.8, wall=28844
2021-03-28 04:02:11 | INFO | train_inner | epoch 138:   144 / 588 loss=3.342, nll_loss=2.591, word_ins=4.168, word_crf=1.061, length=1.974, ppl=10.14, wps=19981.5, ups=2.85, wpb=7019.1, bsz=270.6, num_updates=80700, lr=0.000176008, gnorm=1.823, train_wall=34, gb_free=10.1, wall=28879
2021-03-28 04:02:46 | INFO | train_inner | epoch 138:   244 / 588 loss=3.424, nll_loss=2.666, word_ins=4.234, word_crf=1.112, length=1.953, ppl=10.74, wps=19779.4, ups=2.84, wpb=6975.3, bsz=258.4, num_updates=80800, lr=0.000175899, gnorm=1.94, train_wall=35, gb_free=9.1, wall=28914
2021-03-28 04:03:21 | INFO | train_inner | epoch 138:   344 / 588 loss=3.304, nll_loss=2.542, word_ins=4.125, word_crf=1.046, length=1.956, ppl=9.88, wps=19905.7, ups=2.9, wpb=6858.7, bsz=284.3, num_updates=80900, lr=0.000175791, gnorm=1.987, train_wall=34, gb_free=10.3, wall=28948
2021-03-28 04:03:55 | INFO | train_inner | epoch 138:   444 / 588 loss=3.405, nll_loss=2.631, word_ins=4.202, word_crf=1.089, length=2.159, ppl=10.6, wps=20642.2, ups=2.9, wpb=7115.7, bsz=272.1, num_updates=81000, lr=0.000175682, gnorm=1.749, train_wall=34, gb_free=10, wall=28983
2021-03-28 04:04:30 | INFO | train_inner | epoch 138:   544 / 588 loss=3.354, nll_loss=2.565, word_ins=4.143, word_crf=1.075, length=2.077, ppl=10.22, wps=20239.6, ups=2.9, wpb=6979.3, bsz=294.4, num_updates=81100, lr=0.000175574, gnorm=1.884, train_wall=34, gb_free=10.3, wall=29017
2021-03-28 04:04:46 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 04:04:49 | INFO | valid | epoch 138 | valid on 'valid' subset | loss 9.086 | nll_loss 6.254 | word_ins 7.51 | word_crf 4.719 | length 6.126 | ppl 543.58 | wps 60030 | wpb 5311.6 | bsz 208.1 | num_updates 81144 | best_loss 6.867
2021-03-28 04:04:49 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 138 @ 81144 updates
2021-03-28 04:04:49 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint138.pt
2021-03-28 04:04:50 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint138.pt
2021-03-28 04:04:53 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint138.pt (epoch 138 @ 81144 updates, score 9.086) (writing took 3.642399918000592 seconds)
2021-03-28 04:04:53 | INFO | fairseq_cli.train | end of epoch 138 (average epoch stats below)
2021-03-28 04:04:53 | INFO | train | epoch 138 | loss 3.392 | nll_loss 2.624 | word_ins 4.196 | word_crf 1.091 | length 2.024 | ppl 10.5 | wps 19302.9 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 81144 | lr 0.000175526 | gnorm 1.888 | train_wall 204 | gb_free 9.3 | wall 29040
2021-03-28 04:04:53 | INFO | fairseq.trainer | begin training epoch 139
2021-03-28 04:04:53 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 04:05:12 | INFO | train_inner | epoch 139:   56 / 588 loss=3.394, nll_loss=2.614, word_ins=4.186, word_crf=1.091, length=2.096, ppl=10.51, wps=16761, ups=2.36, wpb=7112.9, bsz=281.8, num_updates=81200, lr=0.000175466, gnorm=1.943, train_wall=35, gb_free=9.3, wall=29060

2021-03-28 04:05:47 | INFO | train_inner | epoch 139: 156 / 588 loss=3.367, nll_loss=2.615, word_ins=4.188, word_crf=1.079, length=1.941, ppl=10.32, wps=19815.5, ups=2.85, wpb=6950.4, bsz=269, num_updates=81300, lr=0.000175358, gnorm=1.959, train_wall=35, gb_free=9.6, wall=29095
2021-03-28 04:06:22 | INFO | train_inner | epoch 139: 256 / 588 loss=3.347, nll_loss=2.577, word_ins=4.156, word_crf=1.062, length=2.076, ppl=10.18, wps=19962.5, ups=2.87, wpb=6954.3, bsz=277.7, num_updates=81400, lr=0.00017525, gnorm=1.911, train_wall=35, gb_free=9.5, wall=29130
2021-03-28 04:06:57 | INFO | train_inner | epoch 139: 356 / 588 loss=3.416, nll_loss=2.653, word_ins=4.223, word_crf=1.104, length=2.005, ppl=10.67, wps=19651.7, ups=2.82, wpb=6973.1, bsz=260.4, num_updates=81500, lr=0.000175142, gnorm=1.846, train_wall=35, gb_free=9.5, wall=29165
2021-03-28 04:07:33 | INFO | train_inner | epoch 139: 456 / 588 loss=3.419, nll_loss=2.655, word_ins=4.221, word_crf=1.112, length=1.963, ppl=10.69, wps=19732.4, ups=2.83, wpb=6969.6, bsz=270.5, num_updates=81600, lr=0.000175035, gnorm=2.046, train_wall=35, gb_free=9.2, wall=29201
2021-03-28 04:08:08 | INFO | train_inner | epoch 139: 556 / 588 loss=3.504, nll_loss=2.721, word_ins=4.276, word_crf=1.156, length=2.099, ppl=11.34, wps=19720, ups=2.81, wpb=7028.4, bsz=270.2, num_updates=81700, lr=0.000174928, gnorm=1.923, train_wall=35, gb_free=9.5, wall=29236
2021-03-28 04:08:20 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 04:08:23 | INFO | valid | epoch 139 | valid on 'valid' subset | loss 9.051 | nll_loss 6.227 | word_ins 7.485 | word_crf 4.692 | length 6.163 | ppl 530.3 | wps 59249.3 | wpb 5311.6 | bsz 208.1 | num_updates 81732 | best_loss 6.867
2021-03-28 04:08:23 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 139 @ 81732 updates
2021-03-28 04:08:23 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint139.pt
2021-03-28 04:08:24 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint139.pt
2021-03-28 04:08:27 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint139.pt (epoch 139 @ 81732 updates, score 9.051) (writing took 3.6759935219997715 seconds)
2021-03-28 04:08:27 | INFO | fairseq_cli.train | end of epoch 139 (average epoch stats below)
2021-03-28 04:08:27 | INFO | train | epoch 139 | loss 3.39 | nll_loss 2.623 | word_ins 4.194 | word_crf 1.09 | length 2.028 | ppl 10.48 | wps 19201.9 | ups 2.75 | wpb 6988.7 | bsz 272.5 | num_updates 81732 | lr 0.000174894 | gnorm 1.917 | train_wall 205 | gb_free 9.9 | wall 29254
2021-03-28 04:08:27 | INFO | fairseq.trainer | begin training epoch 140
2021-03-28 04:08:27 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 04:08:51 | INFO | train_inner | epoch 140:   68 / 588 loss=3.371, nll_loss=2.629, word_ins=4.202, word_crf=1.077, length=1.935, ppl=10.35, wps=16241.8, ups=2.37, wpb=6850.5, bsz=252.9, num_updates=81800, lr=0.000174821, gnorm=1.887, train_wall=35, gb_free=9.5, wall=29278
2021-03-28 04:09:26 | INFO | train_inner | epoch 140: 168 / 588 loss=3.397, nll_loss=2.658, word_ins=4.228, word_crf=1.089, length=1.944, ppl=10.54, wps=19545.2, ups=2.83, wpb=6906.8, bsz=243.5, num_updates=81900, lr=0.000174714, gnorm=1.79, train_wall=35, gb_free=10.4, wall=29314
2021-03-28 04:10:01 | INFO | train_inner | epoch 140: 268 / 588 loss=3.321, nll_loss=2.557, word_ins=4.138, word_crf=1.062, length=1.901, ppl=9.99, wps=19901.4, ups=2.84, wpb=7018.5, bsz=292.6, num_updates=82000, lr=0.000174608, gnorm=1.783, train_wall=35, gb_free=9.5, wall=29349
2021-03-28 04:10:36 | INFO | train_inner | epoch 140: 368 / 588 loss=3.317, nll_loss=2.538, word_ins=4.122, word_crf=1.042, length=2.133, ppl=9.96, wps=20313.6, ups=2.87, wpb=7065.8, bsz=286.9, num_updates=82100, lr=0.000174501, gnorm=1.942, train_wall=34, gb_free=9.7, wall=29384
2021-03-28 04:11:11 | INFO | train_inner | epoch 140: 468 / 588 loss=3.39, nll_loss=2.618, word_ins=4.191, word_crf=1.087, length=2.07, ppl=10.48, wps=19890.6, ups=2.85, wpb=6979.6, bsz=271.2, num_updates=82200, lr=0.000174395, gnorm=1.778, train_wall=35, gb_free=9.8, wall=29419
2021-03-28 04:11:46 | INFO | train_inner | epoch 140: 568 / 588 loss=3.432, nll_loss=2.67, word_ins=4.236, word_crf=1.114, length=2.001, ppl=10.79, wps=19669, ups=2.82, wpb=6983.6, bsz=262.1, num_updates=82300, lr=0.000174289, gnorm=1.787, train_wall=35, gb_free=9.5, wall=29454
2021-03-28 04:11:53 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 04:11:57 | INFO | valid | epoch 140 | valid on 'valid' subset | loss 9.019 | nll_loss 6.212 | word_ins 7.467 | word_crf 4.671 | length 6.144 | ppl 518.92 | wps 59380.2 | wpb 5311.6 | bsz 208.1 | num_updates 82320 | best_loss 6.867
2021-03-28 04:11:57 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 140 @ 82320 updates
2021-03-28 04:11:57 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint140.pt
2021-03-28 04:11:57 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint140.pt
2021-03-28 04:12:00 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint140.pt (epoch 140 @ 82320 updates, score 9.019) (writing took 3.655462702001387 seconds)
2021-03-28 04:12:00 | INFO | fairseq_cli.train | end of epoch 140 (average epoch stats below)
2021-03-28 04:12:00 | INFO | train | epoch 140 | loss 3.363 | nll_loss 2.6 | word_ins 4.176 | word_crf 1.074 | length 2.01 | ppl 10.29 | wps 19227.3 | ups 2.75 | wpb 6988.7 | bsz 272.5 | num_updates 82320 | lr 0.000174268 | gnorm 1.823 | train_wall 205 | gb_free 9.9 | wall 29468
2021-03-28 04:12:00 | INFO | fairseq.trainer | begin training epoch 141
2021-03-28 04:12:00 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 04:12:29 | INFO | train_inner | epoch 141:   80 / 588 loss=3.162, nll_loss=2.38, word_ins=3.986, word_crf=0.954, length=2.151, ppl=8.95, wps=16715, ups=2.37, wpb=7038.8, bsz=304.6, num_updates=82400, lr=0.000174183, gnorm=1.674, train_wall=35, gb_free=9.5, wall=29496
2021-03-28 04:13:04 | INFO | train_inner | epoch 141: 180 / 588 loss=3.363, nll_loss=2.612, word_ins=4.189, word_crf=1.068, length=2.006, ppl=10.29, wps=19667.6, ups=2.84, wpb=6914.6, bsz=260.3, num_updates=82500, lr=0.000174078, gnorm=1.846, train_wall=35, gb_free=9.2, wall=29532
2021-03-28 04:13:39 | INFO | train_inner | epoch 141: 280 / 588 loss=3.253, nll_loss=2.475, word_ins=4.071, word_crf=1.01, length=2.078, ppl=9.53, wps=20143, ups=2.87, wpb=7022.8, bsz=290.1, num_updates=82600, lr=0.000173972, gnorm=1.802, train_wall=35, gb_free=10.4, wall=29567
2021-03-28 04:14:14 | INFO | train_inner | epoch 141: 380 / 588 loss=3.441, nll_loss=2.668, word_ins=4.23, word_crf=1.127, length=1.991, ppl=10.86, wps=19911.3, ups=2.83, wpb=7044.1, bsz=279.5, num_updates=82700, lr=0.000173867, gnorm=2.099, train_wall=35, gb_free=10.1, wall=29602
2021-03-28 04:14:49 | INFO | train_inner | epoch 141: 480 / 588 loss=3.49, nll_loss=2.724, word_ins=4.282, word_crf=1.135, length=2.143, ppl=11.24, wps=19944.7, ups=2.82, wpb=7061.9, bsz=253, num_updates=82800, lr=0.000173762, gnorm=1.783, train_wall=35, gb_free=9.7, wall=29637
2021-03-28 04:15:25 | INFO | train_inner | epoch 141: 580 / 588 loss=3.442, nll_loss=2.685, word_ins=4.247, word_crf=1.131, length=1.884, ppl=10.87, wps=19499, ups=2.82, wpb=6902.7, bsz=266.9, num_updates=82900, lr=0.000173657, gnorm=2.038, train_wall=35, gb_free=9.7, wall=29673
2021-03-28 04:15:28 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 04:15:31 | INFO | valid | epoch 141 | valid on 'valid' subset | loss 9.058 | nll_loss 6.227 | word_ins 7.485 | word_crf 4.696 | length 6.198 | ppl 533.13 | wps 59666.9 | wpb 5311.6 | bsz 208.1 | num_updates 82908 | best_loss 6.867
2021-03-28 04:15:31 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 141 @ 82908 updates
2021-03-28 04:15:31 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint141.pt
2021-03-28 04:15:32 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint141.pt
2021-03-28 04:15:34 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint141.pt (epoch 141 @ 82908 updates, score 9.058) (writing took 3.6441096249982365 seconds)
2021-03-28 04:15:34 | INFO | fairseq_cli.train | end of epoch 141 (average epoch stats below)
2021-03-28 04:15:34 | INFO | train | epoch 141 | loss 3.366 | nll_loss 2.601 | word_ins 4.177 | word_crf 1.075 | length 2.024 | ppl 10.31 | wps 19186.8 | ups 2.75 | wpb 6988.7 | bsz 272.5 | num_updates 82908 | lr 0.000173649 | gnorm 1.882 | train_wall 205 | gb_free 9.5 | wall 29682
2021-03-28 04:15:35 | INFO | fairseq.trainer | begin training epoch 142
2021-03-28 04:15:35 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 04:16:06 | INFO | train_inner | epoch 142:   92 / 588 loss=3.274, nll_loss=2.514, word_ins=4.103, word_crf=1.017, length=2.058, ppl=9.67, wps=16948.1, ups=2.41, wpb=7028.8, bsz=276.4, num_updates=83000, lr=0.000173553, gnorm=1.721, train_wall=34, gb_free=9.4, wall=29714
2021-03-28 04:16:42 | INFO | train_inner | epoch 142: 192 / 588 loss=3.455, nll_loss=2.722, word_ins=4.284, word_crf=1.133, length=1.81, ppl=10.97, wps=19518.4, ups=2.82, wpb=6917.3, bsz=242.9, num_updates=83100, lr=0.000173448, gnorm=2.167, train_wall=35, gb_free=9.4, wall=29750
2021-03-28 04:17:17 | INFO | train_inner | epoch 142: 292 / 588 loss=3.268, nll_loss=2.502, word_ins=4.094, word_crf=1.011, length=2.09, ppl=9.63, wps=20103.7, ups=2.87, wpb=7002.9, bsz=274.5, num_updates=83200, lr=0.000173344, gnorm=1.614, train_wall=35, gb_free=9.6, wall=29784
2021-03-28 04:17:51 | INFO | train_inner | epoch 142: 392 / 588 loss=3.303, nll_loss=2.546, word_ins=4.131, word_crf=1.041, length=1.969, ppl=9.87, wps=20055.9, ups=2.89, wpb=6945.5, bsz=273.7, num_updates=83300, lr=0.00017324, gnorm=1.813, train_wall=34, gb_free=10.1, wall=29819
2021-03-28 04:18:27 | INFO | train_inner | epoch 142: 492 / 588 loss=3.425, nll_loss=2.655, word_ins=4.22, word_crf=1.116, length=1.996, ppl=10.74, wps=19878.1, ups=2.81, wpb=7063.2, bsz=284.9, num_updates=83400, lr=0.000173136, gnorm=1.813, train_wall=35, gb_free=9.5, wall=29855
2021-03-28 04:19:00 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 04:19:04 | INFO | valid | epoch 142 | valid on 'valid' subset | loss 8.999 | nll_loss 6.195 | word_ins 7.445 | word_crf 4.652 | length 6.246 | ppl 511.72 | wps 59514.1 | wpb 5311.6 | bsz 208.1 | num_updates 83496 | best_loss 6.867
2021-03-28 04:19:04 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 142 @ 83496 updates
2021-03-28 04:19:04 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint142.pt

2021-03-28 04:19:05 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint142.pt
2021-03-28 04:19:07 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint142.pt (epoch 142 @ 83496 updates, score 8.999) (writing took 3.6318372530004126 seconds)
2021-03-28 04:19:07 | INFO | fairseq_cli.train | end of epoch 142 (average epoch stats below)
2021-03-28 04:19:07 | INFO | train | epoch 142 | loss 3.353 | nll_loss 2.59 | word_ins 4.168 | word_crf 1.069 | length 2.008 | ppl 10.22 | wps 19294.8 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 83496 | lr 0.000173036 | gnorm 1.852 | train_wall 204 | gb_free 9.3 | wall 29895
2021-03-28 04:19:08 | INFO | fairseq.trainer | begin training epoch 143
2021-03-28 04:19:08 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 04:19:09 | INFO | train_inner | epoch 143:   4 / 588 loss=3.368, nll_loss=2.569, word_ins=4.146, word_crf=1.077, length=2.172, ppl=10.32, wps=16554.9, ups=2.37, wpb=6999.7, bsz=288.2, num_updates=83500, lr=0.000173032, gnorm=1.957, train_wall=35, gb_free=10, wall=29897
2021-03-28 04:19:44 | INFO | train_inner | epoch 143: 104 / 588 loss=3.339, nll_loss=2.577, word_ins=4.156, word_crf=1.061, length=2.003, ppl=10.12, wps=20030, ups=2.85, wpb=7034.9, bsz=275.1, num_updates=83600, lr=0.000172929, gnorm=2.049, train_wall=35, gb_free=9.4, wall=29932
2021-03-28 04:20:19 | INFO | train_inner | epoch 143: 204 / 588 loss=3.349, nll_loss=2.601, word_ins=4.179, word_crf=1.064, length=1.951, ppl=10.19, wps=20266.6, ups=2.85, wpb=7104.6, bsz=261.1, num_updates=83700, lr=0.000172825, gnorm=1.94, train_wall=35, gb_free=9.3, wall=29967
2021-03-28 04:20:55 | INFO | train_inner | epoch 143: 304 / 588 loss=3.461, nll_loss=2.703, word_ins=4.264, word_crf=1.139, length=1.906, ppl=11.01, wps=19556.9, ups=2.81, wpb=6971, bsz=264.5, num_updates=83800, lr=0.000172722, gnorm=2.036, train_wall=35, gb_free=10.4, wall=30003
2021-03-28 04:21:30 | INFO | train_inner | epoch 143: 404 / 588 loss=3.396, nll_loss=2.648, word_ins=4.219, word_crf=1.089, length=1.975, ppl=10.52, wps=19508.1, ups=2.83, wpb=6904.8, bsz=251.3, num_updates=83900, lr=0.000172619, gnorm=1.863, train_wall=35, gb_free=9.6, wall=30038
2021-03-28 04:22:05 | INFO | train_inner | epoch 143: 504 / 588 loss=3.276, nll_loss=2.512, word_ins=4.099, word_crf=1.03, length=1.968, ppl=9.69, wps=20097.8, ups=2.87, wpb=7009.2, bsz=294.8, num_updates=84000, lr=0.000172516, gnorm=1.77, train_wall=35, gb_free=9.9, wall=30073
2021-03-28 04:22:34 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 04:22:37 | INFO | valid | epoch 143 | valid on 'valid' subset | loss 9.092 | nll_loss 6.254 | word_ins 7.509 | word_crf 4.72 | length 6.172 | ppl 545.77 | wps 59847.1 | wpb 5311.6 | bsz 208.1 | num_updates 84084 | best_loss 6.867
2021-03-28 04:22:37 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 143 @ 84084 updates
2021-03-28 04:22:37 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint143.pt
2021-03-28 04:22:38 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint143.pt
2021-03-28 04:22:41 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint143.pt (epoch 143 @ 84084 updates, score 9.092) (writing took 3.6331035229995905 seconds)
2021-03-28 04:22:41 | INFO | fairseq_cli.train | end of epoch 143 (average epoch stats below)
2021-03-28 04:22:41 | INFO | train | epoch 143 | loss 3.36 | nll_loss 2.598 | word_ins 4.174 | word_crf 1.073 | length 1.997 | ppl 10.27 | wps 19231.5 | ups 2.75 | wpb 6988.7 | bsz 272.5 | num_updates 84084 | lr 0.00017243 | gnorm 1.94 | train_wall 205 | gb_free 9.9 | wall 30109
2021-03-28 04:22:41 | INFO | fairseq.trainer | begin training epoch 144
2021-03-28 04:22:41 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 04:22:47 | INFO | train_inner | epoch 144:   16 / 588 loss=3.318, nll_loss=2.541, word_ins=4.122, word_crf=1.049, length=2.08, ppl=9.98, wps=16440.6, ups=2.4, wpb=6863.6, bsz=285.4, num_updates=84100, lr=0.000172414, gnorm=1.953, train_wall=34, gb_free=10.5, wall=30115
2021-03-28 04:23:22 | INFO | train_inner | epoch 144: 116 / 588 loss=3.381, nll_loss=2.618, word_ins=4.192, word_crf=1.085, length=2.006, ppl=10.42, wps=20225.4, ups=2.84, wpb=7109.2, bsz=266.3, num_updates=84200, lr=0.000172311, gnorm=1.985, train_wall=35, gb_free=9.4, wall=30150
2021-03-28 04:23:57 | INFO | train_inner | epoch 144: 216 / 588 loss=3.392, nll_loss=2.641, word_ins=4.215, word_crf=1.085, length=1.996, ppl=10.5, wps=19902.6, ups=2.83, wpb=7021.7, bsz=254.6, num_updates=84300, lr=0.000172209, gnorm=1.84, train_wall=35, gb_free=10.1, wall=30185
2021-03-28 04:24:32 | INFO | train_inner | epoch 144: 316 / 588 loss=3.38, nll_loss=2.631, word_ins=4.201, word_crf=1.086, length=1.929, ppl=10.41, wps=19858.3, ups=2.88, wpb=6899.5, bsz=265, num_updates=84400, lr=0.000172107, gnorm=2.158, train_wall=34, gb_free=9.7, wall=30220
2021-03-28 04:25:07 | INFO | train_inner | epoch 144: 416 / 588 loss=3.324, nll_loss=2.541, word_ins=4.121, word_crf=1.06, length=2.035, ppl=10.01, wps=19677.2, ups=2.84, wpb=6926.1, bsz=302.7, num_updates=84500, lr=0.000172005, gnorm=1.873, train_wall=35, gb_free=10, wall=30255
2021-03-28 04:25:42 | INFO | train_inner | epoch 144: 516 / 588 loss=3.333, nll_loss=2.577, word_ins=4.16, word_crf=1.053, length=2.003, ppl=10.08, wps=19960.1, ups=2.88, wpb=6932.9, bsz=257.4, num_updates=84600, lr=0.000171904, gnorm=1.862, train_wall=34, gb_free=9.6, wall=30290
2021-03-28 04:26:07 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 04:26:10 | INFO | valid | epoch 144 | valid on 'valid' subset | loss 9.079 | nll_loss 6.248 | word_ins 7.502 | word_crf 4.713 | length 6.143 | ppl 540.65 | wps 59811.9 | wpb 5311.6 | bsz 208.1 | num_updates 84672 | best_loss 6.867
2021-03-28 04:26:10 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 144 @ 84672 updates
2021-03-28 04:26:10 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint144.pt
2021-03-28 04:26:11 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint144.pt
2021-03-28 04:26:14 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint144.pt (epoch 144 @ 84672 updates, score 9.079) (writing took 3.645042341999215 seconds)
2021-03-28 04:26:14 | INFO | fairseq_cli.train | end of epoch 144 (average epoch stats below)
2021-03-28 04:26:14 | INFO | train | epoch 144 | loss 3.346 | nll_loss 2.586 | word_ins 4.164 | word_crf 1.065 | length 1.993 | ppl 10.17 | wps 19319 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 84672 | lr 0.00017183 | gnorm 1.898 | train_wall 204 | gb_free 9.7 | wall 30322
2021-03-28 04:26:14 | INFO | fairseq.trainer | begin training epoch 145
2021-03-28 04:26:14 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 04:26:24 | INFO | train_inner | epoch 145:   28 / 588 loss=3.259, nll_loss=2.495, word_ins=4.085, word_crf=1.011, length=2.062, ppl=9.57, wps=17035.5, ups=2.39, wpb=7112.9, bsz=290.2, num_updates=84700, lr=0.000171802, gnorm=1.662, train_wall=34, gb_free=9.3, wall=30332
2021-03-28 04:26:59 | INFO | train_inner | epoch 145: 128 / 588 loss=3.289, nll_loss=2.52, word_ins=4.11, word_crf=1.025, length=2.083, ppl=9.77, wps=20350.5, ups=2.87, wpb=7101.3, bsz=275.4, num_updates=84800, lr=0.000171701, gnorm=1.904, train_wall=35, gb_free=9.4, wall=30367
2021-03-28 04:27:34 | INFO | train_inner | epoch 145: 228 / 588 loss=3.317, nll_loss=2.558, word_ins=4.14, word_crf=1.053, length=1.944, ppl=9.97, wps=19705.6, ups=2.84, wpb=6927.8, bsz=273.3, num_updates=84900, lr=0.0001716, gnorm=2.075, train_wall=35, gb_free=9.5, wall=30402
2021-03-28 04:28:09 | INFO | train_inner | epoch 145: 328 / 588 loss=3.445, nll_loss=2.689, word_ins=4.253, word_crf=1.132, length=1.871, ppl=10.89, wps=19351.6, ups=2.8, wpb=6915.3, bsz=269.6, num_updates=85000, lr=0.000171499, gnorm=1.952, train_wall=35, gb_free=9.3, wall=30437
2021-03-28 04:28:44 | INFO | train_inner | epoch 145: 428 / 588 loss=3.298, nll_loss=2.54, word_ins=4.123, word_crf=1.038, length=1.99, ppl=9.84, wps=20055.6, ups=2.86, wpb=7003.7, bsz=275.3, num_updates=85100, lr=0.000171398, gnorm=1.778, train_wall=35, gb_free=9.5, wall=30472
2021-03-28 04:29:19 | INFO | train_inner | epoch 145: 528 / 588 loss=3.318, nll_loss=2.567, word_ins=4.148, word_crf=1.054, length=1.903, ppl=9.97, wps=20005.4, ups=2.88, wpb=6956.2, bsz=273.8, num_updates=85200, lr=0.000171297, gnorm=1.939, train_wall=34, gb_free=10.3, wall=30507
2021-03-28 04:29:40 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 04:29:43 | INFO | valid | epoch 145 | valid on 'valid' subset | loss 9.146 | nll_loss 6.282 | word_ins 7.54 | word_crf 4.757 | length 6.188 | ppl 566.48 | wps 59274 | wpb 5311.6 | bsz 208.1 | num_updates 85260 | best_loss 6.867
2021-03-28 04:29:43 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 145 @ 85260 updates
2021-03-28 04:29:43 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint145.pt
2021-03-28 04:29:44 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint145.pt
2021-03-28 04:29:47 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint145.pt (epoch 145 @ 85260 updates, score 9.146) (writing took 3.6445545890019275 seconds)
2021-03-28 04:29:47 | INFO | fairseq_cli.train | end of epoch 145 (average epoch stats below)
2021-03-28 04:29:47 | INFO | train | epoch 145 | loss 3.338 | nll_loss 2.58 | word_ins 4.159 | word_crf 1.061 | length 1.974 | ppl 10.11 | wps 19278.4 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 85260 | lr 0.000171237 | gnorm 1.91 | train_wall 204 | gb_free 9.6 | wall 30535
2021-03-28 04:29:47 | INFO | fairseq.trainer | begin training epoch 146
2021-03-28 04:29:47 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 04:30:01 | INFO | train_inner | epoch 146:   40 / 588 loss=3.294, nll_loss=2.534, word_ins=4.12, word_crf=1.032, length=2.018, ppl=9.81, wps=16797.5, ups=2.4, wpb=7008.1, bsz=278.6, num_updates=85300, lr=0.000171197, gnorm=1.787, train_wall=34, gb_free=9.4, wall=30549
2021-03-28 04:30:36 | INFO | train_inner | epoch 146: 140 / 588 loss=3.331, nll_loss=2.597, word_ins=4.18, word_crf=1.05, length=1.912, ppl=10.06, wps=19841.5, ups=2.83, wpb=7009.7, bsz=247.8, num_updates=85400, lr=0.000171096, gnorm=1.781, train_wall=35, gb_free=9.6, wall=30584

2021-03-28 04:31:12 | INFO | train_inner | epoch 146:   240 / 588 loss=3.374, nll_loss=2.604, word_ins=4.179, word_crf=1.086, length=1.98, ppl=10.36, wps=19770.6, ups=2.83, wpb=6994.5, bsz=273.6, num_updates=85500, lr=0.000170996, gnorm=2.082, train_wall=35, gb_free=9.8, wall=30620
2021-03-28 04:31:46 | INFO | train_inner | epoch 146:   340 / 588 loss=3.334, nll_loss=2.57, word_ins=4.152, word_crf=1.057, length=2.017, ppl=10.09, wps=19996.9, ups=2.89, wpb=6921.4, bsz=267.8, num_updates=85600, lr=0.000170896, gnorm=2.022, train_wall=34, gb_free=9.6, wall=30654
2021-03-28 04:32:21 | INFO | train_inner | epoch 146:   440 / 588 loss=3.378, nll_loss=2.617, word_ins=4.191, word_crf=1.082, length=2.003, ppl=10.39, wps=20159.3, ups=2.88, wpb=6994.8, bsz=270.8, num_updates=85700, lr=0.000170797, gnorm=1.918, train_wall=34, gb_free=9.2, wall=30689
2021-03-28 04:32:56 | INFO | train_inner | epoch 146:   540 / 588 loss=3.406, nll_loss=2.66, word_ins=4.225, word_crf=1.108, length=1.855, ppl=10.6, wps=19808.3, ups=2.84, wpb=6967.1, bsz=272.4, num_updates=85800, lr=0.000170697, gnorm=1.729, train_wall=35, gb_free=9.5, wall=30724
2021-03-28 04:33:12 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 04:33:16 | INFO | valid | epoch 146 | valid on 'valid' subset | loss 9.079 | nll_loss 6.245 | word_ins 7.499 | word_crf 4.714 | length 6.15 | ppl 540.7 | wps 59887.5 | wpb 5311.6 | bsz 208.1 | num_updates 85848 | best_loss 6.867
2021-03-28 04:33:16 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 146 @ 85848 updates
2021-03-28 04:33:16 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint146.pt
2021-03-28 04:33:17 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint146.pt
2021-03-28 04:33:19 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint146.pt (epoch 146 @ 85848 updates, score 9.079) (writing took 3.6467816100011987 seconds)
2021-03-28 04:33:19 | INFO | fairseq_cli.train | end of epoch 146 (average epoch stats below)
2021-03-28 04:33:19 | INFO | train | epoch 146 | loss 3.328 | nll_loss 2.571 | word_ins 4.152 | word_crf 1.055 | length 1.963 | ppl 10.04 | wps 19347.1 | ups 2.77 | wpb 6988.7 | bsz 272.5 | num_updates 85848 | lr 0.000170649 | gnorm 1.885 | train_wall 204 | gb_free 9.2 | wall 30747
2021-03-28 04:33:20 | INFO | fairseq.trainer | begin training epoch 147
2021-03-28 04:33:20 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 04:33:38 | INFO | train_inner | epoch 147:   52 / 588 loss=3.271, nll_loss=2.51, word_ins=4.097, word_crf=1.023, length=1.996, ppl=9.65, wps=16978.6, ups=2.39, wpb=7095.3, bsz=286.2, num_updates=85900, lr=0.000170598, gnorm=1.907, train_wall=34, gb_free=9.7, wall=30766
2021-03-28 04:34:13 | INFO | train_inner | epoch 147:   152 / 588 loss=3.264, nll_loss=2.518, word_ins=4.107, word_crf=1.024, length=1.863, ppl=9.6, wps=19672.4, ups=2.88, wpb=6841.9, bsz=279.4, num_updates=86000, lr=0.000170499, gnorm=1.82, train_wall=35, gb_free=9.7, wall=30801
2021-03-28 04:34:47 | INFO | train_inner | epoch 147:   252 / 588 loss=3.234, nll_loss=2.474, word_ins=4.068, word_crf=1, length=1.998, ppl=9.41, wps=20344.6, ups=2.88, wpb=7064.1, bsz=285.2, num_updates=86100, lr=0.0001704, gnorm=1.762, train_wall=34, gb_free=9.3, wall=30835
2021-03-28 04:35:22 | INFO | train_inner | epoch 147:   352 / 588 loss=3.25, nll_loss=2.492, word_ins=4.086, word_crf=1.005, length=2.024, ppl=9.51, wps=20371.5, ups=2.91, wpb=7009.8, bsz=273.2, num_updates=86200, lr=0.000170301, gnorm=1.748, train_wall=34, gb_free=10.9, wall=30870
2021-03-28 04:35:57 | INFO | train_inner | epoch 147:   452 / 588 loss=3.464, nll_loss=2.709, word_ins=4.271, word_crf=1.133, length=1.958, ppl=11.04, wps=19791.1, ups=2.83, wpb=6992.6, bsz=255.3, num_updates=86300, lr=0.000170202, gnorm=1.998, train_wall=35, gb_free=9.5, wall=30905
2021-03-28 04:36:32 | INFO | train_inner | epoch 147:   552 / 588 loss=3.364, nll_loss=2.602, word_ins=4.178, word_crf=1.071, length=2.044, ppl=10.3, wps=19825.1, ups=2.86, wpb=6943, bsz=264.7, num_updates=86400, lr=0.000170103, gnorm=1.838, train_wall=35, gb_free=9.2, wall=30940
2021-03-28 04:36:44 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 04:36:48 | INFO | valid | epoch 147 | valid on 'valid' subset | loss 9.136 | nll_loss 6.277 | word_ins 7.533 | word_crf 4.749 | length 6.206 | ppl 562.55 | wps 59624.2 | wpb 5311.6 | bsz 208.1 | num_updates 86436 | best_loss 6.867
2021-03-28 04:36:48 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 147 @ 86436 updates
2021-03-28 04:36:48 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint147.pt
2021-03-28 04:36:49 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint147.pt
2021-03-28 04:36:51 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint147.pt (epoch 147 @ 86436 updates, score 9.136) (writing took 3.717002674999094 seconds)
2021-03-28 04:36:51 | INFO | fairseq_cli.train | end of epoch 147 (average epoch stats below)
2021-03-28 04:36:51 | INFO | train | epoch 147 | loss 3.314 | nll_loss 2.559 | word_ins 4.142 | word_crf 1.047 | length 1.966 | ppl 9.95 | wps 19382.2 | ups 2.77 | wpb 6988.7 | bsz 272.5 | num_updates 86436 | lr 0.000170068 | gnorm 1.849 | train_wall 203 | gb_free 9.4 | wall 30959
2021-03-28 04:36:52 | INFO | fairseq.trainer | begin training epoch 148
2021-03-28 04:36:52 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 04:37:15 | INFO | train_inner | epoch 148:   64 / 588 loss=3.292, nll_loss=2.557, word_ins=4.143, word_crf=1.038, length=1.829, ppl=9.79, wps=16135, ups=2.34, wpb=6890.2, bsz=264.5, num_updates=86500, lr=0.000170005, gnorm=1.925, train_wall=35, gb_free=9.4, wall=30983
2021-03-28 04:37:49 | INFO | train_inner | epoch 148:   164 / 588 loss=3.216, nll_loss=2.459, word_ins=4.055, word_crf=0.994, length=1.948, ppl=9.29, wps=20154.4, ups=2.89, wpb=6970.7, bsz=297.2, num_updates=86600, lr=0.000169907, gnorm=1.825, train_wall=34, gb_free=10.5, wall=31017
2021-03-28 04:38:25 | INFO | train_inner | epoch 148:   264 / 588 loss=3.392, nll_loss=2.654, word_ins=4.225, word_crf=1.1, length=1.801, ppl=10.5, wps=19850.1, ups=2.84, wpb=6987.4, bsz=255.4, num_updates=86700, lr=0.000169809, gnorm=2.082, train_wall=35, gb_free=9.5, wall=31053
2021-03-28 04:38:59 | INFO | train_inner | epoch 148:   364 / 588 loss=3.235, nll_loss=2.472, word_ins=4.066, word_crf=1.001, length=2.014, ppl=9.42, wps=20685.5, ups=2.9, wpb=7123, bsz=285.2, num_updates=86800, lr=0.000169711, gnorm=1.767, train_wall=34, gb_free=9.9, wall=31087
2021-03-28 04:39:34 | INFO | train_inner | epoch 148:   464 / 588 loss=3.34, nll_loss=2.578, word_ins=4.158, word_crf=1.056, length=2.049, ppl=10.13, wps=19865.5, ups=2.84, wpb=6996.1, bsz=272.2, num_updates=86900, lr=0.000169613, gnorm=1.845, train_wall=35, gb_free=9.5, wall=31122
2021-03-28 04:40:09 | INFO | train_inner | epoch 148:   564 / 588 loss=3.408, nll_loss=2.655, word_ins=4.221, word_crf=1.109, length=1.886, ppl=10.62, wps=19962.6, ups=2.87, wpb=6957.9, bsz=262.5, num_updates=87000, lr=0.000169516, gnorm=2.182, train_wall=35, gb_free=9.5, wall=31157
2021-03-28 04:40:17 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 04:40:21 | INFO | valid | epoch 148 | valid on 'valid' subset | loss 9.14 | nll_loss 6.258 | word_ins 7.516 | word_crf 4.731 | length 6.508 | ppl 564.26 | wps 59933.6 | wpb 5311.6 | bsz 208.1 | num_updates 87024 | best_loss 6.867
2021-03-28 04:40:21 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 148 @ 87024 updates
2021-03-28 04:40:21 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint148.pt
2021-03-28 04:40:22 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint148.pt
2021-03-28 04:40:24 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint148.pt (epoch 148 @ 87024 updates, score 9.14) (writing took 3.6378305860016553 seconds)
2021-03-28 04:40:24 | INFO | fairseq_cli.train | end of epoch 148 (average epoch stats below)
2021-03-28 04:40:24 | INFO | train | epoch 148 | loss 3.315 | nll_loss 2.561 | word_ins 4.143 | word_crf 1.049 | length 1.942 | ppl 9.95 | wps 19303 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 87024 | lr 0.000169492 | gnorm 1.934 | train_wall 204 | gb_free 9.4 | wall 31172
2021-03-28 04:40:24 | INFO | fairseq.trainer | begin training epoch 149
2021-03-28 04:40:24 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 04:40:51 | INFO | train_inner | epoch 149:   76 / 588 loss=3.189, nll_loss=2.436, word_ins=4.038, word_crf=0.972, length=1.983, ppl=9.12, wps=16516.7, ups=2.37, wpb=6956.1, bsz=277.8, num_updates=87100, lr=0.000169419, gnorm=1.726, train_wall=35, gb_free=9.9, wall=31199
2021-03-28 04:41:26 | INFO | train_inner | epoch 149:   176 / 588 loss=3.226, nll_loss=2.485, word_ins=4.079, word_crf=1.007, length=1.798, ppl=9.36, wps=19602.7, ups=2.87, wpb=6832.2, bsz=284.8, num_updates=87200, lr=0.000169321, gnorm=1.892, train_wall=35, gb_free=10, wall=31234
2021-03-28 04:42:02 | INFO | train_inner | epoch 149:   276 / 588 loss=3.423, nll_loss=2.688, word_ins=4.255, word_crf=1.111, length=1.845, ppl=10.72, wps=19473.6, ups=2.83, wpb=6891.8, bsz=249.6, num_updates=87300, lr=0.000169224, gnorm=2.071, train_wall=35, gb_free=9.3, wall=31269
2021-03-28 04:42:36 | INFO | train_inner | epoch 149:   376 / 588 loss=3.304, nll_loss=2.55, word_ins=4.135, word_crf=1.042, length=1.942, ppl=9.88, wps=20076.1, ups=2.87, wpb=6996.5, bsz=269.6, num_updates=87400, lr=0.000169128, gnorm=1.882, train_wall=35, gb_free=9.5, wall=31304
2021-03-28 04:43:11 | INFO | train_inner | epoch 149:   476 / 588 loss=3.251, nll_loss=2.482, word_ins=4.075, word_crf=1.013, length=2.009, ppl=9.52, wps=20561, ups=2.87, wpb=7155.6, bsz=291.7, num_updates=87500, lr=0.000169031, gnorm=1.819, train_wall=35, gb_free=9.6, wall=31339
2021-03-28 04:43:46 | INFO | train_inner | epoch 149:   576 / 588 loss=3.323, nll_loss=2.562, word_ins=4.144, word_crf=1.049, length=2.019, ppl=10, wps=20305.4, ups=2.87, wpb=7085.9, bsz=273, num_updates=87600, lr=0.000168934, gnorm=1.732, train_wall=35, gb_free=9.6, wall=31374
2021-03-28 04:43:50 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 04:43:54 | INFO | valid | epoch 149 | valid on 'valid' subset | loss 9.173 | nll_loss 6.289 | word_ins 7.541 | word_crf 4.767 | length 6.35 | ppl 577.07 | wps 59758.4 | wpb 5311.6 | bsz 208.1 | num_updates 87612 | best_loss 6.867
2021-03-28 04:43:54 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 149 @ 87612 updates
2021-03-28 04:43:54 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint149.pt

2021-03-28 04:43:54 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint149.pt
2021-03-28 04:43:57 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint149.pt (epoch 149 @ 87612 updates, score 9.173) (writing took 3.642993966001086 seconds)
2021-03-28 04:43:57 | INFO | fairseq_cli.train | end of epoch 149 (average epoch stats below)
2021-03-28 04:43:57 | INFO | train | epoch 149 | loss 3.294 | nll_loss 2.544 | word_ins 4.13 | word_crf 1.037 | length 1.917 | ppl 9.81 | wps 19302.1 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 87612 | lr 0.000168923 | gnorm 1.856 | train_wall 204 | gb_free 9.5 | wall 31385
2021-03-28 04:43:57 | INFO | fairseq.trainer | begin training epoch 150
2021-03-28 04:43:57 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 04:44:28 | INFO | train_inner | epoch 150:   88 / 588 loss=3.322, nll_loss=2.58, word_ins=4.161, word_crf=1.052, length=1.888, ppl=10, wps=16521.5, ups=2.36, wpb=7002.2, bsz=262.7, num_updates=87700, lr=0.000168838, gnorm=1.933, train_wall=35, gb_free=9.8, wall=31416
2021-03-28 04:45:03 | INFO | train_inner | epoch 150:   188 / 588 loss=3.215, nll_loss=2.469, word_ins=4.065, word_crf=0.99, length=1.921, ppl=9.29, wps=20156.8, ups=2.89, wpb=6967.8, bsz=279, num_updates=87800, lr=0.000168742, gnorm=1.918, train_wall=34, gb_free=9.9, wall=31451
2021-03-28 04:45:38 | INFO | train_inner | epoch 150:   288 / 588 loss=3.294, nll_loss=2.558, word_ins=4.144, word_crf=1.032, length=1.898, ppl=9.81, wps=20059.2, ups=2.86, wpb=7001.6, bsz=257.5, num_updates=87900, lr=0.000168646, gnorm=1.796, train_wall=35, gb_free=9.4, wall=31486
2021-03-28 04:46:13 | INFO | train_inner | epoch 150:   388 / 588 loss=3.359, nll_loss=2.573, word_ins=4.153, word_crf=1.079, length=2.037, ppl=10.26, wps=19844.5, ups=2.85, wpb=6961.2, bsz=287.1, num_updates=88000, lr=0.00016855, gnorm=2.181, train_wall=35, gb_free=9.3, wall=31521
2021-03-28 04:46:48 | INFO | train_inner | epoch 150:   488 / 588 loss=3.282, nll_loss=2.545, word_ins=4.13, word_crf=1.03, length=1.87, ppl=9.73, wps=19893, ups=2.85, wpb=6973.6, bsz=271.3, num_updates=88100, lr=0.000168454, gnorm=1.795, train_wall=35, gb_free=9.5, wall=31556
2021-03-28 04:47:24 | INFO | train_inner | epoch 150:   588 / 588 loss=3.354, nll_loss=2.596, word_ins=4.172, word_crf=1.077, length=1.912, ppl=10.23, wps=19926.4, ups=2.82, wpb=7072.6, bsz=275.9, num_updates=88200, lr=0.000168359, gnorm=1.994, train_wall=35, gb_free=9.3, wall=31591
2021-03-28 04:47:24 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 04:47:27 | INFO | valid | epoch 150 | valid on 'valid' subset | loss 9.1 | nll_loss 6.252 | word_ins 7.499 | word_crf 4.722 | length 6.277 | ppl 548.63 | wps 59664.3 | wpb 5311.6 | bsz 208.1 | num_updates 88200 | best_loss 6.867
2021-03-28 04:47:27 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 150 @ 88200 updates
2021-03-28 04:47:27 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint150.pt
2021-03-28 04:47:28 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint150.pt
2021-03-28 04:47:31 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint150.pt (epoch 150 @ 88200 updates, score 9.1) (writing took 3.63895304000107 seconds)
2021-03-28 04:47:31 | INFO | fairseq_cli.train | end of epoch 150 (average epoch stats below)
2021-03-28 04:47:31 | INFO | train | epoch 150 | loss 3.301 | nll_loss 2.551 | word_ins 4.135 | word_crf 1.042 | length 1.918 | ppl 9.86 | wps 19264.5 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 88200 | lr 0.000168359 | gnorm 1.942 | train_wall 205 | gb_free 9.3 | wall 31598
2021-03-28 04:47:31 | INFO | fairseq.trainer | begin training epoch 151
2021-03-28 04:47:31 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 04:48:06 | INFO | train_inner | epoch 151:   100 / 588 loss=3.256, nll_loss=2.517, word_ins=4.107, word_crf=1.014, length=1.878, ppl=9.55, wps=16767.4, ups=2.37, wpb=7082.7, bsz=269.8, num_updates=88300, lr=0.000168263, gnorm=1.966, train_wall=35, gb_free=9.4, wall=31634
2021-03-28 04:48:41 | INFO | train_inner | epoch 151:   200 / 588 loss=3.27, nll_loss=2.532, word_ins=4.118, word_crf=1.028, length=1.835, ppl=9.65, wps=19844.7, ups=2.86, wpb=6938.3, bsz=275, num_updates=88400, lr=0.000168168, gnorm=1.868, train_wall=35, gb_free=9.4, wall=31669
2021-03-28 04:49:16 | INFO | train_inner | epoch 151:   300 / 588 loss=3.292, nll_loss=2.535, word_ins=4.12, word_crf=1.035, length=1.964, ppl=9.79, wps=20000.2, ups=2.87, wpb=6971.8, bsz=276.1, num_updates=88500, lr=0.000168073, gnorm=1.909, train_wall=35, gb_free=10.5, wall=31704
2021-03-28 04:49:51 | INFO | train_inner | epoch 151:   400 / 588 loss=3.243, nll_loss=2.494, word_ins=4.089, word_crf=1, length=1.988, ppl=9.47, wps=20214.8, ups=2.87, wpb=7051.4, bsz=270.2, num_updates=88600, lr=0.000167978, gnorm=1.822, train_wall=35, gb_free=9.5, wall=31738
2021-03-28 04:50:26 | INFO | train_inner | epoch 151:   500 / 588 loss=3.349, nll_loss=2.595, word_ins=4.173, word_crf=1.069, length=1.928, ppl=10.19, wps=19739.6, ups=2.83, wpb=6981.7, bsz=274.1, num_updates=88700, lr=0.000167884, gnorm=1.824, train_wall=35, gb_free=9.4, wall=31774
2021-03-28 04:50:57 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 04:51:00 | INFO | valid | epoch 151 | valid on 'valid' subset | loss 9.195 | nll_loss 6.306 | word_ins 7.561 | word_crf 4.786 | length 6.275 | ppl 585.97 | wps 59604.5 | wpb 5311.6 | bsz 208.1 | num_updates 88788 | best_loss 6.867
2021-03-28 04:51:00 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 151 @ 88788 updates
2021-03-28 04:51:00 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint151.pt
2021-03-28 04:51:01 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint151.pt
2021-03-28 04:51:04 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint151.pt (epoch 151 @ 88788 updates, score 9.195) (writing took 3.66724277600224 seconds)
2021-03-28 04:51:04 | INFO | fairseq_cli.train | end of epoch 151 (average epoch stats below)
2021-03-28 04:51:04 | INFO | train | epoch 151 | loss 3.287 | nll_loss 2.537 | word_ins 4.124 | word_crf 1.032 | length 1.928 | ppl 9.76 | wps 19290.8 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 88788 | lr 0.0001678 | gnorm 1.875 | train_wall 204 | gb_free 10.1 | wall 31811
2021-03-28 04:51:04 | INFO | fairseq.trainer | begin training epoch 152
2021-03-28 04:51:04 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 04:51:08 | INFO | train_inner | epoch 152:   12 / 588 loss=3.323, nll_loss=2.573, word_ins=4.155, word_crf=1.054, length=1.92, ppl=10.01, wps=16373.8, ups=2.37, wpb=6898.5, bsz=262, num_updates=88800, lr=0.000167789, gnorm=1.86, train_wall=35, gb_free=9.4, wall=31816
2021-03-28 04:51:43 | INFO | train_inner | epoch 152:   112 / 588 loss=3.177, nll_loss=2.412, word_ins=4.016, word_crf=0.968, length=2.011, ppl=9.04, wps=20071, ups=2.88, wpb=6967.8, bsz=285, num_updates=88900, lr=0.000167695, gnorm=2.069, train_wall=34, gb_free=9.6, wall=31851
2021-03-28 04:52:18 | INFO | train_inner | epoch 152:   212 / 588 loss=3.295, nll_loss=2.561, word_ins=4.145, word_crf=1.038, length=1.838, ppl=9.81, wps=19898.2, ups=2.86, wpb=6949.5, bsz=263.3, num_updates=89000, lr=0.0001676, gnorm=1.905, train_wall=35, gb_free=9.4, wall=31886
2021-03-28 04:52:53 | INFO | train_inner | epoch 152:   312 / 588 loss=3.255, nll_loss=2.509, word_ins=4.102, word_crf=1.013, length=1.904, ppl=9.54, wps=20021.6, ups=2.86, wpb=6993.9, bsz=274.9, num_updates=89100, lr=0.000167506, gnorm=1.805, train_wall=35, gb_free=9.7, wall=31920
2021-03-28 04:53:28 | INFO | train_inner | epoch 152:   412 / 588 loss=3.355, nll_loss=2.611, word_ins=4.186, word_crf=1.076, length=1.86, ppl=10.23, wps=19941.1, ups=2.86, wpb=6981.4, bsz=262.8, num_updates=89200, lr=0.000167412, gnorm=2.103, train_wall=35, gb_free=9.5, wall=31955
2021-03-28 04:54:03 | INFO | train_inner | epoch 152:   512 / 588 loss=3.39, nll_loss=2.642, word_ins=4.214, word_crf=1.095, length=1.882, ppl=10.49, wps=19724.2, ups=2.83, wpb=6977.2, bsz=264, num_updates=89300, lr=0.000167319, gnorm=1.972, train_wall=35, gb_free=9.7, wall=31991
2021-03-28 04:54:29 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 04:54:32 | INFO | valid | epoch 152 | valid on 'valid' subset | loss 9.159 | nll_loss 6.287 | word_ins 7.539 | word_crf 4.754 | length 6.355 | ppl 571.71 | wps 59275.3 | wpb 5311.6 | bsz 208.1 | num_updates 89376 | best_loss 6.867
2021-03-28 04:54:32 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 152 @ 89376 updates
2021-03-28 04:54:32 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint152.pt
2021-03-28 04:54:33 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint152.pt
2021-03-28 04:54:36 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint152.pt (epoch 152 @ 89376 updates, score 9.159) (writing took 3.6388409020000836 seconds)
2021-03-28 04:54:36 | INFO | fairseq_cli.train | end of epoch 152 (average epoch stats below)
2021-03-28 04:54:36 | INFO | train | epoch 152 | loss 3.286 | nll_loss 2.537 | word_ins 4.124 | word_crf 1.033 | length 1.908 | ppl 9.76 | wps 19329.8 | ups 2.77 | wpb 6988.7 | bsz 272.5 | num_updates 89376 | lr 0.000167247 | gnorm 1.957 | train_wall 204 | gb_free 9.7 | wall 32024
2021-03-28 04:54:36 | INFO | fairseq.trainer | begin training epoch 153
2021-03-28 04:54:36 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 04:54:45 | INFO | train_inner | epoch 153:   24 / 588 loss=3.255, nll_loss=2.493, word_ins=4.085, word_crf=1.012, length=2.007, ppl=9.54, wps=16912.3, ups=2.39, wpb=7090.7, bsz=281.8, num_updates=89400, lr=0.000167225, gnorm=1.863, train_wall=35, gb_free=9.4, wall=32033
2021-03-28 04:55:20 | INFO | train_inner | epoch 153:   124 / 588 loss=3.279, nll_loss=2.541, word_ins=4.129, word_crf=1.03, length=1.842, ppl=9.7, wps=19449.3, ups=2.83, wpb=6866, bsz=270, num_updates=89500, lr=0.000167132, gnorm=1.805, train_wall=35, gb_free=9.5, wall=32068
2021-03-28 04:55:55 | INFO | train_inner | epoch 153:   224 / 588 loss=3.388, nll_loss=2.663, word_ins=4.235, word_crf=1.093, length=1.783, ppl=10.47, wps=19950.8, ups=2.83, wpb=7037.7, bsz=243.9, num_updates=89600, lr=0.000167038, gnorm=2.021, train_wall=35, gb_free=9.9, wall=32103
2021-03-28 04:56:30 | INFO | train_inner | epoch 153:   324 / 588 loss=3.28, nll_loss=2.533, word_ins=4.123, word_crf=1.028, length=1.91, ppl=9.71, wps=19809.7, ups=2.87, wpb=6900.4, bsz=267.4, num_updates=89700, lr=0.000166945, gnorm=2.007, train_wall=35, gb_free=9.6, wall=32138

2021-03-28 04:57:05 | INFO | train_inner | epoch 153:   424 / 588 loss=3.125, nll_loss=2.346, word_ins=3.956, word_crf=0.943, length=2.039, ppl=8.72, wps=20449.7, ups=2.9, wpb=7060.6, bsz=319.8, num_updates=89800, lr=0.000166852, gnorm=1.931, train_wall=34, gb_free=9.5, wall=32173
2021-03-28 04:57:40 | INFO | train_inner | epoch 153:   524 / 588 loss=3.386, nll_loss=2.645, word_ins=4.217, word_crf=1.082, length=1.944, ppl=10.45, wps=19948.9, ups=2.84, wpb=7012.9, bsz=246.8, num_updates=89900, lr=0.000166759, gnorm=1.899, train_wall=35, gb_free=9.4, wall=32208
2021-03-28 04:58:02 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 04:58:05 | INFO | valid | epoch 153 | valid on 'valid' subset | loss 9.162 | nll_loss 6.286 | word_ins 7.539 | word_crf 4.776 | length 6.164 | ppl 572.66 | wps 59686.9 | wpb 5311.6 | bsz 208.1 | num_updates 89964 | best_loss 6.867
2021-03-28 04:58:05 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 153 @ 89964 updates
2021-03-28 04:58:05 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint153.pt
2021-03-28 04:58:06 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint153.pt
2021-03-28 04:58:09 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint153.pt (epoch 153 @ 89964 updates, score 9.162) (writing took 3.654911274003098 seconds)
2021-03-28 04:58:09 | INFO | fairseq_cli.train | end of epoch 153 (average epoch stats below)
2021-03-28 04:58:09 | INFO | train | epoch 153 | loss 3.275 | nll_loss 2.526 | word_ins 4.115 | word_crf 1.025 | length 1.925 | ppl 9.68 | wps 19308.4 | ups 2.76 | wpb 6988.7 | bsz 272.5 | num_updates 89964 | lr 0.0001667 | gnorm 1.891 | train_wall 204 | gb_free 9.3 | wall 32237
2021-03-28 04:58:09 | INFO | fairseq.trainer | begin training epoch 154
2021-03-28 04:58:09 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 04:58:22 | INFO | train_inner | epoch 154:   36 / 588 loss=3.145, nll_loss=2.395, word_ins=4.002, word_crf=0.951, length=1.931, ppl=8.84, wps=16717.1, ups=2.4, wpb=6965.6, bsz=290.9, num_updates=90000, lr=0.000166667, gnorm=1.664, train_wall=34, gb_free=9.7, wall=32250
2021-03-28 04:58:22 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 04:58:25 | INFO | valid | epoch 154 | valid on 'valid' subset | loss 9.269 | nll_loss 6.341 | word_ins 7.598 | word_crf 4.823 | length 6.474 | ppl 616.94 | wps 59850.1 | wpb 5311.6 | bsz 208.1 | num_updates 90000 | best_loss 6.867
2021-03-28 04:58:25 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 154 @ 90000 updates
2021-03-28 04:58:25 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint_154_90000.pt
2021-03-28 04:58:26 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint_154_90000.pt
2021-03-28 04:58:29 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint_154_90000.pt (epoch 154 @ 90000 updates, score 9.269) (writing took 3.9683552100032102 seconds)
2021-03-28 04:59:04 | INFO | train_inner | epoch 154:   136 / 588 loss=3.326, nll_loss=2.604, word_ins=4.184, word_crf=1.054, length=1.799, ppl=10.03, wps=16614.3, ups=2.35, wpb=7066.1, bsz=254, num_updates=90100, lr=0.000166574, gnorm=1.83, train_wall=35, gb_free=9.4, wall=32292
2021-03-28 04:59:39 | INFO | train_inner | epoch 154:   236 / 588 loss=3.294, nll_loss=2.516, word_ins=4.105, word_crf=1.031, length=2.104, ppl=9.81, wps=19921.5, ups=2.84, wpb=7005.4, bsz=280.3, num_updates=90200, lr=0.000166482, gnorm=2.031, train_wall=35, gb_free=9.4, wall=32327
2021-03-28 05:00:14 | INFO | train_inner | epoch 154:   336 / 588 loss=3.296, nll_loss=2.563, word_ins=4.149, word_crf=1.035, length=1.867, ppl=9.82, wps=20343, ups=2.87, wpb=7089.9, bsz=255.3, num_updates=90300, lr=0.00016639, gnorm=1.738, train_wall=35, gb_free=10, wall=32362
2021-03-28 05:00:49 | INFO | train_inner | epoch 154:   436 / 588 loss=3.375, nll_loss=2.628, word_ins=4.199, word_crf=1.091, length=1.841, ppl=10.37, wps=19789, ups=2.84, wpb=6971, bsz=266.5, num_updates=90400, lr=0.000166298, gnorm=2.199, train_wall=35, gb_free=9.9, wall=32397
2021-03-28 05:01:24 | INFO | train_inner | epoch 154:   536 / 588 loss=3.119, nll_loss=2.355, word_ins=3.965, word_crf=0.936, length=2.006, ppl=8.69, wps=20390.2, ups=2.92, wpb=6986.4, bsz=307.4, num_updates=90500, lr=0.000166206, gnorm=1.731, train_wall=34, gb_free=9.5, wall=32432
2021-03-28 05:01:42 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 05:01:45 | INFO | valid | epoch 154 | valid on 'valid' subset | loss 9.184 | nll_loss 6.295 | word_ins 7.552 | word_crf 4.778 | length 6.304 | ppl 581.66 | wps 59775.1 | wpb 5311.6 | bsz 208.1 | num_updates 90552 | best_loss 6.867
2021-03-28 05:01:45 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 154 @ 90552 updates
2021-03-28 05:01:45 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint154.pt
2021-03-28 05:01:46 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint154.pt
2021-03-28 05:01:49 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint154.pt (epoch 154 @ 90552 updates, score 9.184) (writing took 3.6412021860014647 seconds)
2021-03-28 05:01:49 | INFO | fairseq_cli.train | end of epoch 154 (average epoch stats below)
2021-03-28 05:01:49 | INFO | train | epoch 154 | loss 3.267 | nll_loss 2.521 | word_ins 4.11 | word_crf 1.022 | length 1.904 | ppl 9.63 | wps 18696.5 | ups 2.68 | wpb 6988.7 | bsz 272.5 | num_updates 90552 | lr 0.000166158 | gnorm 1.884 | train_wall 204 | gb_free 10 | wall 32457
2021-03-28 05:01:49 | INFO | fairseq.trainer | begin training epoch 155
2021-03-28 05:01:49 | INFO | fairseq_cli.train | Start iterating over samples
2021-03-28 05:02:05 | INFO | train_inner | epoch 155:   48 / 588 loss=3.112, nll_loss=2.361, word_ins=3.973, word_crf=0.937, length=1.895, ppl=8.65, wps=16450.1, ups=2.39, wpb=6875.9, bsz=290.3, num_updates=90600, lr=0.000166114, gnorm=1.724, train_wall=34, gb_free=9.8, wall=32473
2021-03-28 05:02:40 | INFO | train_inner | epoch 155:   148 / 588 loss=3.213, nll_loss=2.474, word_ins=4.073, word_crf=0.991, length=1.859, ppl=9.27, wps=20482.2, ups=2.88, wpb=7114.3, bsz=270.3, num_updates=90700, lr=0.000166022, gnorm=1.981, train_wall=34, gb_free=9.3, wall=32508
2021-03-28 05:03:15 | INFO | train_inner | epoch 155:   248 / 588 loss=3.178, nll_loss=2.436, word_ins=4.034, word_crf=0.973, length=1.885, ppl=9.05, wps=19700.3, ups=2.89, wpb=6815.3, bsz=290.6, num_updates=90800, lr=0.000165931, gnorm=1.854, train_wall=34, gb_free=9.4, wall=32543
2021-03-28 05:03:50 | INFO | train_inner | epoch 155:   348 / 588 loss=3.258, nll_loss=2.502, word_ins=4.093, word_crf=1.015, length=1.958, ppl=9.56, wps=20138.7, ups=2.88, wpb=6987, bsz=274.7, num_updates=90900, lr=0.00016584, gnorm=1.962, train_wall=34, gb_free=10.7, wall=32577
2021-03-28 05:04:24 | INFO | train_inner | epoch 155:   448 / 588 loss=3.226, nll_loss=2.491, word_ins=4.086, word_crf=1, length=1.833, ppl=9.36, wps=20053.2, ups=2.91, wpb=6899.9, bsz=267.1, num_updates=91000, lr=0.000165748, gnorm=1.764, train_wall=34, gb_free=9.8, wall=32612
2021-03-28 05:04:59 | INFO | train_inner | epoch 155:   548 / 588 loss=3.395, nll_loss=2.641, word_ins=4.213, word_crf=1.092, length=1.964, ppl=10.52, wps=20093.4, ups=2.82, wpb=7119.7, bsz=258.5, num_updates=91100, lr=0.000165657, gnorm=1.971, train_wall=35, gb_free=10, wall=32647
2021-03-28 05:05:14 | INFO | fairseq_cli.train | begin validation on "valid" subset
2021-03-28 05:05:17 | INFO | valid | epoch 155 | valid on 'valid' subset | loss 9.26 | nll_loss 6.341 | word_ins 7.595 | word_crf 4.826 | length 6.367 | ppl 612.99 | wps 59260.3 | wpb 5311.6 | bsz 208.1 | num_updates 91140 | best_loss 6.867
2021-03-28 05:05:17 | INFO | fairseq.checkpoint_utils | Preparing to save checkpoint for epoch 155 @ 91140 updates
2021-03-28 05:05:17 | INFO | fairseq.trainer | Saving checkpoint to checkpoints/checkpoint155.pt
2021-03-28 05:05:18 | INFO | fairseq.trainer | Finished saving checkpoint to checkpoints/checkpoint155.pt
2021-03-28 05:05:21 | INFO | fairseq.checkpoint_utils | Saved checkpoint checkpoints/checkpoint155.pt (epoch 155 @ 91140 updates, score 9.26) (writing took 3.6332514240057208 seconds)
2021-03-28 05:05:21 | INFO | fairseq_cli.train | end of epoch 155 (average epoch stats below)
2021-03-28 05:05:21 | INFO | train | epoch 155 | loss 3.259 | nll_loss 2.515 | word_ins 4.106 | word_crf 1.018 | length 1.881 | ppl 9.57 | wps 19361.8 | ups 2.77 | wpb 6988.7 | bsz 272.5 | num_updates 91140 | lr 0.000165621 | gnorm 1.888 | train_wall 204 | gb_free 9.8 | wall 32669