vocab\_size = len(train\_data)

embed\_dim = 300

hidden\_dim = 128

num\_layers = 2

bidirec = True,

dropout = 0.5

num\_heads = 4

0%| | 0/30000 [00:00<?, ?it/s]/usr/local/lib/python3.7/dist-packages/ipykernel\_launcher.py:9: VisibleDeprecationWarning: Creating an ndarray from ragged nested sequences (which is a list-or-tuple of lists-or-tuples-or ndarrays with different lengths or shapes) is deprecated. If you meant to do this, you must specify 'dtype=object' when creating the ndarray.

if \_\_name\_\_ == '\_\_main\_\_':

3%|▎ | 999/30000 [02:59<1:26:11, 5.61it/s] 1000th iters >> loss = 4.2685, start\_acc = 6.4969, end\_acc = 6.8672, acc = 1.8016

3%|▎ | 1001/30000 [03:20<37:33:02, 4.66s/it]validation score : {'exact\_match': 5.851063829787234, 'f1': 12.006872870549497}

7%|▋ | 1999/30000 [06:20<1:25:07, 5.48it/s] 2000th iters >> loss = 3.8182, start\_acc = 8.9797, end\_acc = 9.4094, acc = 3.3438

7%|▋ | 2001/30000 [06:41<35:33:07, 4.57s/it]validation score : {'exact\_match': 6.372110841527263, 'f1': 12.08961789333993}

10%|▉ | 2999/30000 [09:39<1:20:36, 5.58it/s] 3000th iters >> loss = 3.6980, start\_acc = 9.8016, end\_acc = 10.5906, acc = 4.3453

10%|█ | 3001/30000 [09:59<32:07:43, 4.28s/it]validation score : {'exact\_match': 7.725427481855086, 'f1': 12.693159378621997}

13%|█▎ | 3999/30000 [12:58<1:14:48, 5.79it/s] 4000th iters >> loss = 3.5742, start\_acc = 11.6359, end\_acc = 12.7406, acc = 5.8234

13%|█▎ | 4001/30000 [13:18<31:38:24, 4.38s/it]validation score : {'exact\_match': 10.432378079436903, 'f1': 16.70940890798954}

17%|█▋ | 4999/30000 [16:16<1:30:29, 4.60it/s] 5000th iters >> loss = 3.3015, start\_acc = 16.4344, end\_acc = 17.8641, acc = 9.6125

17%|█▋ | 5000/30000 [16:37<45:59:51, 6.62s/it]validation score : {'exact\_match': 15.119204504612435, 'f1': 22.27923233370541}

20%|█▉ | 5999/30000 [19:37<1:07:20, 5.94it/s] 6000th iters >> loss = 3.0857, start\_acc = 20.4625, end\_acc = 21.9109, acc = 12.9594

20%|██ | 6001/30000 [19:56<27:05:14, 4.06s/it]validation score : {'exact\_match': 19.780743565300288, 'f1': 27.363742659149157}

23%|██▎ | 6999/30000 [22:56<1:06:38, 5.75it/s] 7000th iters >> loss = 2.9201, start\_acc = 23.5703, end\_acc = 25.2563, acc = 15.7000

23%|██▎ | 7001/30000 [23:15<26:20:35, 4.12s/it]validation score : {'exact\_match': 21.38941716075624, 'f1': 28.889769814852304}

27%|██▋ | 7999/30000 [26:14<1:03:05, 5.81it/s] 8000th iters >> loss = 2.7809, start\_acc = 26.5406, end\_acc = 28.6484, acc = 18.4187

27%|██▋ | 8001/30000 [26:32<24:10:45, 3.96s/it]validation score : {'exact\_match': 22.84652383034543, 'f1': 31.192816941272213}

30%|██▉ | 8999/30000 [29:32<1:04:49, 5.40it/s] 9000th iters >> loss = 2.6762, start\_acc = 28.6953, end\_acc = 31.0344, acc = 20.5281

30%|███ | 9001/30000 [29:51<23:30:09, 4.03s/it]validation score : {'exact\_match': 24.095482766003, 'f1': 32.2042747475392}

33%|███▎ | 9999/30000 [32:51<1:12:18, 4.61it/s] 10000th iters >> loss = 2.5846, start\_acc = 30.8656, end\_acc = 33.4047, acc = 22.3047

33%|███▎ | 10001/30000 [33:09<21:42:52, 3.91s/it]validation score : {'exact\_match': 26.255903821382567, 'f1': 35.823855579085254}

37%|███▋ | 10999/30000 [36:07<56:02, 5.65it/s] 11000th iters >> loss = 2.4827, start\_acc = 33.3328, end\_acc = 35.6469, acc = 24.3094

37%|███▋ | 11001/30000 [36:25<21:15:21, 4.03s/it]validation score : {'exact\_match': 28.37967401725791, 'f1': 37.94064652127457}

40%|███▉ | 11999/30000 [39:22<54:05, 5.55it/s] 12000th iters >> loss = 2.3843, start\_acc = 35.2375, end\_acc = 38.2563, acc = 26.3891

40%|████ | 12001/30000 [39:39<17:46:36, 3.56s/it]validation score : {'exact\_match': 30.877595742460606, 'f1': 40.69044451767763}

43%|████▎ | 12999/30000 [42:36<47:39, 5.94it/s] 13000th iters >> loss = 2.2860, start\_acc = 37.6437, end\_acc = 41.0391, acc = 28.5187

43%|████▎ | 13001/30000 [42:54<18:01:56, 3.82s/it]validation score : {'exact\_match': 33.99958359358734, 'f1': 44.99175083091344}

47%|████▋ | 13999/30000 [45:53<41:59, 6.35it/s] 14000th iters >> loss = 2.1930, start\_acc = 39.9422, end\_acc = 43.3531, acc = 30.6578

47%|████▋ | 14001/30000 [46:09<15:46:29, 3.55s/it]validation score : {'exact\_match': 34.80194435825835, 'f1': 45.08140534927864}

50%|████▉ | 14999/30000 [49:08<40:52, 6.12it/s] 15000th iters >> loss = 2.1155, start\_acc = 41.9062, end\_acc = 45.1984, acc = 32.4578

50%|█████ | 15001/30000 [49:25<15:28:48, 3.72s/it]validation score : {'exact\_match': 36.73281505728314, 'f1': 47.30188950872747}

53%|█████▎ | 15999/30000 [52:23<46:59, 4.96it/s] 16000th iters >> loss = 2.0259, start\_acc = 43.5719, end\_acc = 47.4141, acc = 34.0250

53%|█████▎ | 16001/30000 [52:39<13:43:23, 3.53s/it]validation score : {'exact\_match': 39.491161745172164, 'f1': 50.55054888655386}

57%|█████▋ | 16999/30000 [55:35<40:27, 5.35it/s] 17000th iters >> loss = 1.9715, start\_acc = 44.8656, end\_acc = 48.5359, acc = 35.0516

57%|█████▋ | 17000/30000 [55:53<19:30:27, 5.40s/it]validation score : {'exact\_match': 40.21453147136207, 'f1': 51.72248950192802}

60%|█████▉ | 17999/30000 [58:58<34:36, 5.78it/s] 18000th iters >> loss = 1.9048, start\_acc = 46.6578, end\_acc = 50.5313, acc = 36.7609

60%|██████ | 18001/30000 [59:16<12:40:32, 3.80s/it]validation score : {'exact\_match': 39.80739989863152, 'f1': 51.30696013766146}

63%|██████▎ | 18999/30000 [1:02:22<36:36, 5.01it/s] 19000th iters >> loss = 1.8594, start\_acc = 47.4875, end\_acc = 51.3922, acc = 37.4062

63%|██████▎ | 19001/30000 [1:02:39<11:26:11, 3.74s/it]validation score : {'exact\_match': 40.892531876138435, 'f1': 52.361558240521326}

67%|██████▋ | 19999/30000 [1:05:47<32:04, 5.20it/s] 20000th iters >> loss = 1.8076, start\_acc = 48.9000, end\_acc = 52.8750, acc = 38.9266

67%|██████▋ | 20001/30000 [1:06:05<10:49:54, 3.90s/it]validation score : {'exact\_match': 42.269603880355696, 'f1': 54.02299365946645}

70%|██████▉ | 20999/30000 [1:09:07<28:08, 5.33it/s] 21000th iters >> loss = 1.7649, start\_acc = 49.6906, end\_acc = 53.4484, acc = 39.6203

70%|███████ | 21001/30000 [1:09:24<8:59:39, 3.60s/it] validation score : {'exact\_match': 43.44890143116307, 'f1': 54.99166850290268}

73%|███████▎ | 21999/30000 [1:12:30<25:02, 5.32it/s] 22000th iters >> loss = 1.7204, start\_acc = 50.5656, end\_acc = 55.1422, acc = 40.6547

73%|███████▎ | 22000/30000 [1:12:47<11:36:44, 5.23s/it]validation score : {'exact\_match': 45.4417811653796, 'f1': 57.12540160537197}

77%|███████▋ | 22999/30000 [1:15:48<19:14, 6.06it/s] 23000th iters >> loss = 1.6869, start\_acc = 51.5094, end\_acc = 55.5531, acc = 41.3172

77%|███████▋ | 23001/30000 [1:16:05<6:54:59, 3.56s/it]validation score : {'exact\_match': 44.47485806974858, 'f1': 55.75770814456726}

80%|███████▉ | 23999/30000 [1:19:11<17:38, 5.67it/s] 24000th iters >> loss = 1.6374, start\_acc = 52.6281, end\_acc = 56.5500, acc = 42.3391

80%|████████ | 24001/30000 [1:19:28<6:20:20, 3.80s/it]validation score : {'exact\_match': 45.30591775325978, 'f1': 57.658797539177016}

83%|████████▎ | 24999/30000 [1:22:34<15:11, 5.49it/s] 25000th iters >> loss = 1.6102, start\_acc = 53.1875, end\_acc = 57.5125, acc = 42.7656

83%|████████▎ | 25001/30000 [1:22:50<4:57:19, 3.57s/it]validation score : {'exact\_match': 45.66596194503171, 'f1': 57.733438263109356}

87%|████████▋ | 25999/30000 [1:25:57<12:14, 5.45it/s] 26000th iters >> loss = 1.5756, start\_acc = 53.9219, end\_acc = 58.1828, acc = 43.5266

87%|████████▋ | 26001/30000 [1:26:14<4:10:53, 3.76s/it]validation score : {'exact\_match': 45.524706825699106, 'f1': 58.06802818060594}

90%|████████▉ | 26999/30000 [1:29:20<08:12, 6.09it/s] 27000th iters >> loss = 1.5343, start\_acc = 54.7812, end\_acc = 59.0734, acc = 44.3969

90%|█████████ | 27001/30000 [1:29:37<3:08:00, 3.76s/it]validation score : {'exact\_match': 46.71039354187689, 'f1': 58.4075319472214}

93%|█████████▎| 27999/30000 [1:32:45<06:01, 5.53it/s] 28000th iters >> loss = 1.5178, start\_acc = 55.1141, end\_acc = 59.6031, acc = 44.7516

93%|█████████▎| 28001/30000 [1:33:01<2:01:07, 3.64s/it]validation score : {'exact\_match': 47.190675433353256, 'f1': 58.89770429235121}

97%|█████████▋| 28999/30000 [1:36:09<03:08, 5.31it/s] 29000th iters >> loss = 1.4878, start\_acc = 55.7984, end\_acc = 60.3938, acc = 45.5984

97%|█████████▋| 29001/30000 [1:36:27<1:04:20, 3.86s/it]validation score : {'exact\_match': 46.814044213263976, 'f1': 59.632483423897526}

100%|█████████▉| 29999/30000 [1:39:36<00:00, 5.17it/s] 30000th iters >> loss = 1.4683, start\_acc = 56.6266, end\_acc = 60.7547, acc = 46.1109

100%|██████████| 30000/30000 [1:39:53<00:00, 5.01it/s]validation score : {'exact\_match': 47.24244539742695, 'f1': 58.91200908719829}

vocab\_size = len(train\_data)

embed\_dim = 300

hidden\_dim = 128

num\_layers = 3

bidirec = True,

dropout = 0.5

0%| | 0/30000 [00:00<?, ?it/s]/usr/local/lib/python3.7/dist-packages/ipykernel\_launcher.py:9: VisibleDeprecationWarning: Creating an ndarray from ragged nested sequences (which is a list-or-tuple of lists-or-tuples-or ndarrays with different lengths or shapes) is deprecated. If you meant to do this, you must specify 'dtype=object' when creating the ndarray.

if \_\_name\_\_ == '\_\_main\_\_':

3%|▎ | 999/30000 [03:46<1:44:02, 4.65it/s] 1000th iters >> loss = 4.6038, start\_acc = 4.1953, end\_acc = 4.7000, acc = 1.0406

3%|▎ | 1000/30000 [04:10<58:12:08, 7.23s/it]validation score : {'exact\_match': 4.652428241055639, 'f1': 10.947140332348866}

7%|▋ | 1999/30000 [07:53<1:44:38, 4.46it/s] 2000th iters >> loss = 3.9428, start\_acc = 8.2234, end\_acc = 8.9203, acc = 2.9766

7%|▋ | 2000/30000 [08:15<53:09:03, 6.83s/it]validation score : {'exact\_match': 6.2635264162953534, 'f1': 11.867341633805369}

10%|▉ | 2999/30000 [12:01<1:40:32, 4.48it/s] 3000th iters >> loss = 3.7901, start\_acc = 9.6062, end\_acc = 9.9969, acc = 3.9172

10%|█ | 3000/30000 [12:26<55:39:54, 7.42s/it]validation score : {'exact\_match': 6.312165263963275, 'f1': 12.148304890601054}

13%|█▎ | 3999/30000 [16:10<1:34:11, 4.60it/s] 4000th iters >> loss = 3.7034, start\_acc = 10.1125, end\_acc = 10.8594, acc = 4.5797

13%|█▎ | 4000/30000 [16:32<49:21:46, 6.83s/it]validation score : {'exact\_match': 6.708522264852924, 'f1': 12.168353090590745}

17%|█▋ | 4999/30000 [20:16<1:48:55, 3.83it/s] 5000th iters >> loss = 3.6198, start\_acc = 11.0203, end\_acc = 11.6312, acc = 5.4656

17%|█▋ | 5000/30000 [20:36<44:08:54, 6.36s/it]validation score : {'exact\_match': 7.4634087335188095, 'f1': 12.584792146261256}

20%|█▉ | 5999/30000 [24:17<1:30:48, 4.41it/s] 6000th iters >> loss = 3.5614, start\_acc = 11.5875, end\_acc = 12.0625, acc = 6.0391

20%|██ | 6000/30000 [24:38<41:58:24, 6.30s/it]validation score : {'exact\_match': 7.439055038954511, 'f1': 12.527405855843558}

23%|██▎ | 6999/30000 [28:22<1:23:17, 4.60it/s] 7000th iters >> loss = 3.5085, start\_acc = 12.2469, end\_acc = 12.7312, acc = 6.4062

23%|██▎ | 7000/30000 [28:44<42:58:11, 6.73s/it]validation score : {'exact\_match': 8.132421734436848, 'f1': 13.023621822913006}

27%|██▋ | 7999/30000 [32:30<1:24:29, 4.34it/s] 8000th iters >> loss = 3.3363, start\_acc = 15.6156, end\_acc = 16.7969, acc = 9.0984

27%|██▋ | 8000/30000 [32:51<39:41:08, 6.49s/it]validation score : {'exact\_match': 12.536302032913843, 'f1': 19.3288769715269}

30%|██▉ | 8999/30000 [36:40<1:19:24, 4.41it/s] 9000th iters >> loss = 3.1307, start\_acc = 19.2922, end\_acc = 20.7859, acc = 12.3781

30%|███ | 9000/30000 [37:00<35:59:36, 6.17s/it]validation score : {'exact\_match': 15.711009174311927, 'f1': 22.213225549824053}

33%|███▎ | 9999/30000 [40:43<1:14:33, 4.47it/s] 10000th iters >> loss = 3.0061, start\_acc = 21.2125, end\_acc = 23.0953, acc = 14.0563

33%|███▎ | 10000/30000 [41:05<37:02:51, 6.67s/it]validation score : {'exact\_match': 16.365075774711602, 'f1': 22.926733951155725}

37%|███▋ | 10999/30000 [44:52<1:14:20, 4.26it/s] 11000th iters >> loss = 2.8963, start\_acc = 23.5813, end\_acc = 25.5656, acc = 16.2594

37%|███▋ | 11001/30000 [45:13<23:03:48, 4.37s/it]validation score : {'exact\_match': 18.452448585388023, 'f1': 25.726219441786707}

40%|███▉ | 11999/30000 [48:57<1:00:00, 5.00it/s] 12000th iters >> loss = 2.7902, start\_acc = 25.6719, end\_acc = 27.4813, acc = 17.9922

40%|████ | 12000/30000 [49:17<30:03:47, 6.01s/it]validation score : {'exact\_match': 19.712480940971467, 'f1': 27.208086345914896}

43%|████▎ | 12999/30000 [53:01<1:05:28, 4.33it/s] 13000th iters >> loss = 2.7017, start\_acc = 27.5844, end\_acc = 29.5609, acc = 19.8703

43%|████▎ | 13000/30000 [53:21<30:07:44, 6.38s/it]validation score : {'exact\_match': 20.18815716657443, 'f1': 27.85368968674052}

47%|████▋ | 13999/30000 [57:06<59:35, 4.48it/s] 14000th iters >> loss = 2.6257, start\_acc = 29.1109, end\_acc = 31.5453, acc = 21.3391

47%|████▋ | 14001/30000 [57:25<18:21:52, 4.13s/it]validation score : {'exact\_match': 20.910064239828692, 'f1': 28.526007764187455}

50%|████▉ | 14999/30000 [1:01:09<55:09, 4.53it/s] 15000th iters >> loss = 2.5666, start\_acc = 30.3766, end\_acc = 32.9625, acc = 22.6172

50%|█████ | 15001/30000 [1:01:29<17:34:46, 4.22s/it]validation score : {'exact\_match': 21.959061193869896, 'f1': 30.249025371674318}

53%|█████▎ | 15999/30000 [1:05:11<56:36, 4.12it/s] 16000th iters >> loss = 2.4965, start\_acc = 31.6266, end\_acc = 34.5047, acc = 23.7969

53%|█████▎ | 16000/30000 [1:05:32<24:35:03, 6.32s/it]validation score : {'exact\_match': 22.16565913670591, 'f1': 30.182788071888098}

57%|█████▋ | 16999/30000 [1:09:16<42:39, 5.08it/s] 17000th iters >> loss = 2.4253, start\_acc = 33.2797, end\_acc = 35.9734, acc = 25.2266

57%|█████▋ | 17000/30000 [1:09:34<20:15:45, 5.61s/it]validation score : {'exact\_match': 22.921300709670586, 'f1': 31.230435878443142}

60%|█████▉ | 17999/30000 [1:13:17<39:55, 5.01it/s] 18000th iters >> loss = 2.3775, start\_acc = 34.8969, end\_acc = 37.3203, acc = 26.4391

60%|██████ | 18001/30000 [1:13:35<13:14:18, 3.97s/it]validation score : {'exact\_match': 23.27820808038518, 'f1': 32.088325080968545}

63%|██████▎ | 18999/30000 [1:17:18<38:28, 4.76it/s] 19000th iters >> loss = 2.3298, start\_acc = 35.5828, end\_acc = 38.2063, acc = 26.9625

63%|██████▎ | 19001/30000 [1:17:37<12:27:21, 4.08s/it]validation score : {'exact\_match': 24.338790931989923, 'f1': 32.91055391142063}

67%|██████▋ | 19999/30000 [1:21:18<40:25, 4.12it/s] 20000th iters >> loss = 2.2726, start\_acc = 36.6531, end\_acc = 39.6188, acc = 28.1156

67%|██████▋ | 20000/30000 [1:21:38<16:47:04, 6.04s/it]validation score : {'exact\_match': 23.856277418007103, 'f1': 32.455286247446516}

70%|██████▉ | 20999/30000 [1:25:18<36:01, 4.16it/s] 21000th iters >> loss = 2.2295, start\_acc = 37.5984, end\_acc = 40.7984, acc = 29.0203

70%|███████ | 21001/30000 [1:25:37<9:59:07, 3.99s/it] validation score : {'exact\_match': 24.289699349106314, 'f1': 32.86457577485915}

73%|███████▎ | 21999/30000 [1:29:21<30:48, 4.33it/s] 22000th iters >> loss = 2.1739, start\_acc = 39.2172, end\_acc = 42.3391, acc = 30.6500

73%|███████▎ | 22000/30000 [1:29:40<13:19:24, 6.00s/it]validation score : {'exact\_match': 26.436781609195403, 'f1': 35.25149563163167}

77%|███████▋ | 22999/30000 [1:33:24<28:10, 4.14it/s] 23000th iters >> loss = 2.1347, start\_acc = 39.8047, end\_acc = 43.3156, acc = 30.9234

77%|███████▋ | 23000/30000 [1:33:42<10:50:15, 5.57s/it]validation score : {'exact\_match': 26.606446314925066, 'f1': 36.00870335461944}

80%|███████▉ | 23999/30000 [1:37:25<20:45, 4.82it/s] 24000th iters >> loss = 2.0843, start\_acc = 41.3078, end\_acc = 44.1922, acc = 32.3438

80%|████████ | 24000/30000 [1:37:43<9:14:17, 5.54s/it]validation score : {'exact\_match': 27.65313576124387, 'f1': 36.79396246184639}

83%|████████▎ | 24999/30000 [1:41:28<19:54, 4.19it/s] 25000th iters >> loss = 2.0291, start\_acc = 42.6219, end\_acc = 45.8109, acc = 33.5047

83%|████████▎ | 25000/30000 [1:41:47<8:17:59, 5.98s/it]validation score : {'exact\_match': 28.8406390823433, 'f1': 38.014863549372116}

87%|████████▋ | 25999/30000 [1:45:30<13:32, 4.92it/s] 26000th iters >> loss = 1.9801, start\_acc = 43.7047, end\_acc = 47.2672, acc = 34.4313

87%|████████▋ | 26001/30000 [1:45:48<4:18:07, 3.87s/it]validation score : {'exact\_match': 29.724469937519206, 'f1': 39.346814043457584}

90%|████████▉ | 26999/30000 [1:49:28<10:58, 4.56it/s] 27000th iters >> loss = 1.9326, start\_acc = 44.9828, end\_acc = 48.5641, acc = 35.8359

90%|█████████ | 27000/30000 [1:49:45<4:33:13, 5.46s/it]validation score : {'exact\_match': 30.943589743589744, 'f1': 41.17474733624275}

93%|█████████▎| 27999/30000 [1:53:26<06:55, 4.81it/s] 28000th iters >> loss = 1.8919, start\_acc = 45.8234, end\_acc = 49.3500, acc = 36.5672

93%|█████████▎| 28000/30000 [1:53:46<3:20:34, 6.02s/it]validation score : {'exact\_match': 31.479229989868287, 'f1': 41.49158889030132}

97%|█████████▋| 28999/30000 [1:57:29<03:55, 4.26it/s] 29000th iters >> loss = 1.8629, start\_acc = 46.3234, end\_acc = 50.0203, acc = 36.9844

97%|█████████▋| 29000/30000 [1:57:46<1:31:05, 5.47s/it]validation score : {'exact\_match': 32.247888470540346, 'f1': 42.35050881971879}

100%|█████████▉| 29999/30000 [2:01:29<00:00, 4.23it/s] 30000th iters >> loss = 1.8190, start\_acc = 47.7578, end\_acc = 51.3922, acc = 38.2359

100%|██████████| 30000/30000 [2:01:47<00:00, 4.11it/s]validation score : {'exact\_match': 32.83551967709384, 'f1': 43.04048122995652}

LSTM\_attn2(

(embedding): Embedding(87599, 300)

(lstm): LSTM(300, 128, num\_layers=3, batch\_first=True, dropout=0.5, bidirectional=True)

(lin1): Linear(in\_features=256, out\_features=128, bias=True)

(attention): Attention(

(k\_proj): Linear(in\_features=128, out\_features=128, bias=True)

(q\_proj): Linear(in\_features=128, out\_features=128, bias=True)

(v\_proj): Linear(in\_features=128, out\_features=128, bias=True)

(out\_proj): Linear(in\_features=128, out\_features=128, bias=True)

(softmax): Softmax(dim=-1)

)

(lstm2): LSTM(128, 128, num\_layers=3, batch\_first=True, dropout=0.5, bidirectional=True)

(relu): ReLU()

(out\_lin): Linear(in\_features=256, out\_features=2, bias=True)

)

embed\_dim = 100

hidden\_dim = 128

num\_layers = 3

bidirec = True,

dropout = 0.5

0%| | 0/30000 [00:00<?, ?it/s]/usr/local/lib/python3.7/dist-packages/ipykernel\_launcher.py:9: VisibleDeprecationWarning: Creating an ndarray from ragged nested sequences (which is a list-or-tuple of lists-or-tuples-or ndarrays with different lengths or shapes) is deprecated. If you meant to do this, you must specify 'dtype=object' when creating the ndarray.

if \_\_name\_\_ == '\_\_main\_\_':

3%|▎ | 999/30000 [03:41<1:55:47, 4.17it/s] 1000th iters >> loss = 4.7144, start\_acc = 3.5172, end\_acc = 3.9672, acc = 0.6406

3%|▎ | 1000/30000 [04:08<65:46:43, 8.17s/it]validation score : {'exact\_match': 3.916211293260474, 'f1': 10.431236350240951}

7%|▋ | 1999/30000 [07:47<2:12:17, 3.53it/s] 2000th iters >> loss = 4.0839, start\_acc = 7.1547, end\_acc = 8.0516, acc = 2.4219

7%|▋ | 2000/30000 [08:13<63:30:11, 8.16s/it]validation score : {'exact\_match': 5.2746127256433235, 'f1': 11.25206557254239}

10%|▉ | 2999/30000 [11:54<1:37:29, 4.62it/s] 3000th iters >> loss = 3.8833, start\_acc = 9.2359, end\_acc = 10.2547, acc = 3.6328

10%|█ | 3000/30000 [12:17<53:51:24, 7.18s/it]validation score : {'exact\_match': 7.506211586242971, 'f1': 14.32626276837127}

13%|█▎ | 3999/30000 [15:58<1:34:07, 4.60it/s] 4000th iters >> loss = 3.6586, start\_acc = 12.0984, end\_acc = 13.1938, acc = 5.5922

13%|█▎ | 4000/30000 [16:22<52:17:40, 7.24s/it]validation score : {'exact\_match': 11.112496882015465, 'f1': 19.38682628475112}

17%|█▋ | 4999/30000 [19:59<1:20:34, 5.17it/s] 5000th iters >> loss = 3.4833, start\_acc = 14.7109, end\_acc = 16.0094, acc = 7.8359

17%|█▋ | 5000/30000 [20:20<46:00:33, 6.63s/it]validation score : {'exact\_match': 13.961239353166276, 'f1': 21.82578029180385}

20%|█▉ | 5999/30000 [24:01<2:14:28, 2.97it/s] 6000th iters >> loss = 3.3596, start\_acc = 16.5609, end\_acc = 17.9531, acc = 9.6297

20%|██ | 6000/30000 [24:23<45:58:46, 6.90s/it]validation score : {'exact\_match': 16.510296088704763, 'f1': 24.295734109461673}

23%|██▎ | 6999/30000 [28:01<1:25:35, 4.48it/s] 7000th iters >> loss = 3.2553, start\_acc = 17.9438, end\_acc = 19.7750, acc = 11.1859

23%|██▎ | 7001/30000 [28:22<29:05:43, 4.55s/it]validation score : {'exact\_match': 16.790648246546226, 'f1': 25.334439066343148}

27%|██▋ | 7999/30000 [31:59<1:25:39, 4.28it/s] 8000th iters >> loss = 3.1661, start\_acc = 19.7437, end\_acc = 21.3891, acc = 12.4891

27%|██▋ | 8000/30000 [32:21<40:16:19, 6.59s/it]validation score : {'exact\_match': 19.07394609536973, 'f1': 27.042842756458338}

30%|██▉ | 8999/30000 [35:59<1:14:14, 4.71it/s] 9000th iters >> loss = 3.0941, start\_acc = 21.1406, end\_acc = 22.6875, acc = 13.9125

30%|███ | 9000/30000 [36:20<37:30:08, 6.43s/it]validation score : {'exact\_match': 19.458777663112027, 'f1': 27.416858111337753}

33%|███▎ | 9999/30000 [40:00<1:22:47, 4.03it/s] 10000th iters >> loss = 3.0215, start\_acc = 22.0703, end\_acc = 24.2875, acc = 15.0641

33%|███▎ | 10000/30000 [40:21<35:53:08, 6.46s/it]validation score : {'exact\_match': 20.716546112115733, 'f1': 29.134077452743195}

37%|███▋ | 10999/30000 [44:01<1:12:19, 4.38it/s] 11000th iters >> loss = 2.9649, start\_acc = 23.0953, end\_acc = 25.0156, acc = 15.6766

37%|███▋ | 11000/30000 [44:21<32:26:44, 6.15s/it]validation score : {'exact\_match': 21.844079591171035, 'f1': 30.510365146558495}

40%|███▉ | 11999/30000 [47:58<1:00:16, 4.98it/s] 12000th iters >> loss = 2.8878, start\_acc = 24.2984, end\_acc = 26.7453, acc = 16.9391

40%|████ | 12000/30000 [48:18<30:53:39, 6.18s/it]validation score : {'exact\_match': 22.34308713151311, 'f1': 30.519760808066437}

43%|████▎ | 12999/30000 [51:55<57:33, 4.92it/s] 13000th iters >> loss = 2.8192, start\_acc = 25.9859, end\_acc = 28.2734, acc = 18.5547

43%|████▎ | 13000/30000 [52:14<28:04:28, 5.95s/it]validation score : {'exact\_match': 23.349436392914654, 'f1': 31.821270251191148}

47%|████▋ | 13999/30000 [55:53<57:28, 4.64it/s] 14000th iters >> loss = 2.7599, start\_acc = 27.0453, end\_acc = 29.4766, acc = 19.4406

47%|████▋ | 14001/30000 [56:13<18:48:59, 4.23s/it]validation score : {'exact\_match': 23.971861471861473, 'f1': 32.750130742671494}

50%|████▉ | 14999/30000 [59:50<49:16, 5.07it/s] 15000th iters >> loss = 2.7187, start\_acc = 27.7203, end\_acc = 30.1641, acc = 20.0125

50%|█████ | 15000/30000 [1:00:08<23:18:36, 5.59s/it]validation score : {'exact\_match': 24.08627284317892, 'f1': 32.081395715908634}

53%|█████▎ | 15999/30000 [1:03:46<52:01, 4.49it/s] 16000th iters >> loss = 2.6580, start\_acc = 29.3625, end\_acc = 31.8906, acc = 21.4031

53%|█████▎ | 16000/30000 [1:04:06<23:56:06, 6.15s/it]validation score : {'exact\_match': 25.027191646726127, 'f1': 33.895853363133504}

57%|█████▋ | 16999/30000 [1:07:46<47:05, 4.60it/s] 17000th iters >> loss = 2.6045, start\_acc = 30.3438, end\_acc = 32.9547, acc = 22.3516

57%|█████▋ | 17000/30000 [1:08:04<20:38:18, 5.72s/it]validation score : {'exact\_match': 25.7008847670824, 'f1': 35.293099931140034}

60%|█████▉ | 17999/30000 [1:11:42<40:45, 4.91it/s] 18000th iters >> loss = 2.5559, start\_acc = 31.4344, end\_acc = 34.1641, acc = 23.2750

60%|██████ | 18000/30000 [1:12:02<20:11:06, 6.06s/it]validation score : {'exact\_match': 27.342598441777216, 'f1': 36.491267937037264}

63%|██████▎ | 18999/30000 [1:15:40<38:36, 4.75it/s] 19000th iters >> loss = 2.5053, start\_acc = 32.4484, end\_acc = 35.4766, acc = 24.2891

63%|██████▎ | 19000/30000 [1:15:58<17:02:10, 5.58s/it]validation score : {'exact\_match': 27.580407820054656, 'f1': 37.15360633081101}

67%|██████▋ | 19999/30000 [1:19:35<38:11, 4.36it/s] 20000th iters >> loss = 2.4440, start\_acc = 34.0406, end\_acc = 36.9766, acc = 25.5547

67%|██████▋ | 20000/30000 [1:19:54<16:05:23, 5.79s/it]validation score : {'exact\_match': 29.001353743621785, 'f1': 38.24605275389647}

70%|██████▉ | 20999/30000 [1:23:29<33:40, 4.45it/s] 21000th iters >> loss = 2.3947, start\_acc = 35.1844, end\_acc = 38.2203, acc = 26.5750

70%|███████ | 21000/30000 [1:23:46<13:29:43, 5.40s/it]validation score : {'exact\_match': 31.406930898526664, 'f1': 40.74676197457522}

73%|███████▎ | 21999/30000 [1:27:23<27:22, 4.87it/s] 22000th iters >> loss = 2.3533, start\_acc = 36.4438, end\_acc = 39.6297, acc = 27.8453

73%|███████▎ | 22001/30000 [1:27:42<9:01:30, 4.06s/it] validation score : {'exact\_match': 31.540847983453983, 'f1': 41.19369125818356}

77%|███████▋ | 22999/30000 [1:31:17<23:47, 4.90it/s] 23000th iters >> loss = 2.3174, start\_acc = 37.3297, end\_acc = 40.1547, acc = 28.3031

77%|███████▋ | 23001/30000 [1:31:35<7:34:09, 3.89s/it] validation score : {'exact\_match': 32.54017305315204, 'f1': 43.06319880664363}

80%|███████▉ | 23999/30000 [1:35:10<24:22, 4.10it/s] 24000th iters >> loss = 2.2794, start\_acc = 37.8469, end\_acc = 41.1438, acc = 28.9672

80%|████████ | 24000/30000 [1:35:29<9:34:07, 5.74s/it]validation score : {'exact\_match': 33.27870531598894, 'f1': 43.76165173913021}

83%|████████▎ | 24999/30000 [1:39:06<17:26, 4.78it/s] 25000th iters >> loss = 2.2345, start\_acc = 38.8875, end\_acc = 42.3672, acc = 29.9953

83%|████████▎ | 25000/30000 [1:39:24<7:59:48, 5.76s/it]validation score : {'exact\_match': 33.509020879789176, 'f1': 44.17459596951358}

87%|████████▋ | 25999/30000 [1:43:01<14:38, 4.56it/s] 26000th iters >> loss = 2.2046, start\_acc = 39.3484, end\_acc = 42.8312, acc = 30.4250

87%|████████▋ | 26001/30000 [1:43:18<4:16:17, 3.85s/it]validation score : {'exact\_match': 34.173814198615275, 'f1': 44.37011594015369}

90%|████████▉ | 26999/30000 [1:46:55<09:56, 5.03it/s] 27000th iters >> loss = 2.1651, start\_acc = 39.9984, end\_acc = 43.6656, acc = 30.9234

90%|█████████ | 27000/30000 [1:47:14<4:43:31, 5.67s/it]validation score : {'exact\_match': 34.15406620351451, 'f1': 45.2323065555367}

93%|█████████▎| 27999/30000 [1:50:47<06:59, 4.78it/s] 28000th iters >> loss = 2.1406, start\_acc = 40.7750, end\_acc = 44.1234, acc = 31.5984

93%|█████████▎| 28000/30000 [1:51:05<3:00:52, 5.43s/it]validation score : {'exact\_match': 34.165732341249615, 'f1': 44.993544406687576}

97%|█████████▋| 28999/30000 [1:54:41<03:19, 5.02it/s] 29000th iters >> loss = 2.0990, start\_acc = 41.3672, end\_acc = 45.2828, acc = 32.3906

97%|█████████▋| 29000/30000 [1:54:59<1:33:17, 5.60s/it]validation score : {'exact\_match': 34.94492044063647, 'f1': 45.59037061944873}

100%|█████████▉| 29999/30000 [1:58:38<00:00, 5.24it/s] 30000th iters >> loss = 2.1123, start\_acc = 41.2375, end\_acc = 44.9094, acc = 32.1250

100%|██████████| 30000/30000 [1:58:56<00:00, 4.20it/s]validation score : {'exact\_match': 34.54415225621611, 'f1': 45.21415692111223}

LSTM\_attn2(

(embedding): Embedding(87599, 100)

(lstm): LSTM(100, 128, num\_layers=3, batch\_first=True, dropout=0.5, bidirectional=True)

(lin1): Linear(in\_features=256, out\_features=128, bias=True)

(attention): Attention(

(k\_proj): Linear(in\_features=128, out\_features=128, bias=True)

(q\_proj): Linear(in\_features=128, out\_features=128, bias=True)

(v\_proj): Linear(in\_features=128, out\_features=128, bias=True)

(out\_proj): Linear(in\_features=128, out\_features=128, bias=True)

(softmax): Softmax(dim=-1)

)

(lstm2): LSTM(128, 128, num\_layers=3, batch\_first=True, dropout=0.5, bidirectional=True)

(relu): ReLU()

(out\_lin): Linear(in\_features=256, out\_features=2, bias=True)

)

vocab\_size = len(train\_data)

embed\_dim = 300

hidden\_dim = 256

num\_layers = 3

bidirec = True,

dropout = 0.5

0%| | 0/30000 [00:00<?, ?it/s]/usr/local/lib/python3.7/dist-packages/ipykernel\_launcher.py:9: VisibleDeprecationWarning: Creating an ndarray from ragged nested sequences (which is a list-or-tuple of lists-or-tuples-or ndarrays with different lengths or shapes) is deprecated. If you meant to do this, you must specify 'dtype=object' when creating the ndarray.

if \_\_name\_\_ == '\_\_main\_\_':

/usr/local/lib/python3.7/dist-packages/ipykernel\_launcher.py:59: UserWarning: Creating a tensor from a list of numpy.ndarrays is extremely slow. Please consider converting the list to a single numpy.ndarray with numpy.array() before converting to a tensor. (Triggered internally at ../torch/csrc/utils/tensor\_new.cpp:201.)

3%|▎ | 999/30000 [09:18<4:47:51, 1.68it/s] 1000th iters >> loss = 4.5469, start\_acc = 4.6734, end\_acc = 5.1734, acc = 1.1516

3%|▎ | 1000/30000 [09:55<93:42:14, 11.63s/it]validation score : {'exact\_match': 4.593341761483354, 'f1': 11.321476606298397}

7%|▋ | 1999/30000 [19:04<4:26:15, 1.75it/s] 2000th iters >> loss = 3.9190, start\_acc = 8.4469, end\_acc = 9.1734, acc = 3.1250

7%|▋ | 2000/30000 [19:39<85:52:55, 11.04s/it]validation score : {'exact\_match': 6.619208192831272, 'f1': 12.110064767141148}

10%|▉ | 2999/30000 [28:48<4:08:56, 1.81it/s] 3000th iters >> loss = 3.7732, start\_acc = 9.6547, end\_acc = 10.0891, acc = 3.9937

10%|█ | 3000/30000 [29:24<83:08:06, 11.08s/it]validation score : {'exact\_match': 6.620292636123771, 'f1': 12.282805866539244}

13%|█▎ | 3999/30000 [38:32<3:58:34, 1.82it/s] 4000th iters >> loss = 3.6577, start\_acc = 10.5953, end\_acc = 11.0719, acc = 4.8906

13%|█▎ | 4000/30000 [39:05<75:45:38, 10.49s/it]validation score : {'exact\_match': 7.4037754351556755, 'f1': 12.515326119831721}

17%|█▋ | 4999/30000 [48:15<3:35:16, 1.94it/s] 5000th iters >> loss = 3.4975, start\_acc = 13.5781, end\_acc = 14.4266, acc = 7.1984

17%|█▋ | 5000/30000 [48:53<81:38:37, 11.76s/it]validation score : {'exact\_match': 9.963873674396924, 'f1': 16.603673468300485}

20%|█▉ | 5999/30000 [58:12<3:21:42, 1.98it/s] 6000th iters >> loss = 3.2375, start\_acc = 17.5594, end\_acc = 18.9125, acc = 10.6063

20%|██ | 6000/30000 [58:47<72:07:42, 10.82s/it]validation score : {'exact\_match': 14.850553071310896, 'f1': 21.925091226606355}

23%|██▎ | 6999/30000 [1:07:55<3:32:51, 1.80it/s] 7000th iters >> loss = 3.0715, start\_acc = 20.4094, end\_acc = 21.9672, acc = 13.2438

23%|██▎ | 7000/30000 [1:08:28<65:50:14, 10.30s/it]validation score : {'exact\_match': 18.043141461089263, 'f1': 24.941622015330044}

27%|██▋ | 7999/30000 [1:17:42<3:48:51, 1.60it/s] 8000th iters >> loss = 2.9334, start\_acc = 23.2078, end\_acc = 24.8844, acc = 15.7125

27%|██▋ | 8000/30000 [1:18:13<60:20:20, 9.87s/it]validation score : {'exact\_match': 19.94057821963204, 'f1': 26.955849804474646}

30%|██▉ | 8999/30000 [1:27:21<3:10:49, 1.83it/s] 9000th iters >> loss = 2.7669, start\_acc = 26.2562, end\_acc = 28.3109, acc = 18.6422

30%|███ | 9000/30000 [1:27:54<59:58:51, 10.28s/it]validation score : {'exact\_match': 20.15965819653699, 'f1': 27.848634940398725}

33%|███▎ | 9999/30000 [1:37:09<3:01:13, 1.84it/s] 10000th iters >> loss = 2.6442, start\_acc = 29.4078, end\_acc = 31.5953, acc = 21.4828

33%|███▎ | 10000/30000 [1:37:41<55:49:09, 10.05s/it]validation score : {'exact\_match': 22.168021680216803, 'f1': 30.099433614855126}

37%|███▋ | 10999/30000 [1:46:54<2:51:35, 1.85it/s] 11000th iters >> loss = 2.5168, start\_acc = 32.1453, end\_acc = 34.1672, acc = 23.9750

37%|███▋ | 11000/30000 [1:47:26<51:42:14, 9.80s/it]validation score : {'exact\_match': 24.324612920448477, 'f1': 32.41778243123324}

40%|███▉ | 11999/30000 [1:56:31<2:45:20, 1.81it/s] 12000th iters >> loss = 2.4050, start\_acc = 34.4703, end\_acc = 37.0656, acc = 26.2328

40%|████ | 12000/30000 [1:57:03<49:47:48, 9.96s/it]validation score : {'exact\_match': 23.509620439491115, 'f1': 31.885507068888007}

43%|████▎ | 12999/30000 [2:06:17<2:28:51, 1.90it/s] 13000th iters >> loss = 2.2949, start\_acc = 37.1328, end\_acc = 39.8406, acc = 28.8453

43%|████▎ | 13000/30000 [2:06:49<46:42:19, 9.89s/it]validation score : {'exact\_match': 24.587079238971356, 'f1': 33.25933236595166}

47%|████▋ | 13999/30000 [2:16:03<2:24:42, 1.84it/s] 14000th iters >> loss = 2.1829, start\_acc = 39.5031, end\_acc = 42.5422, acc = 30.9453

47%|████▋ | 14000/30000 [2:16:35<44:34:07, 10.03s/it]validation score : {'exact\_match': 25.34111030101031, 'f1': 34.633272920275516}

50%|████▉ | 14999/30000 [2:25:49<2:14:26, 1.86it/s] 15000th iters >> loss = 2.0963, start\_acc = 41.3906, end\_acc = 44.5781, acc = 32.7297

50%|█████ | 15000/30000 [2:26:21<41:27:38, 9.95s/it]validation score : {'exact\_match': 27.060296265387024, 'f1': 37.03914187152087}

53%|█████▎ | 15999/30000 [2:35:41<2:00:33, 1.94it/s] 16000th iters >> loss = 2.0051, start\_acc = 43.7984, end\_acc = 47.0641, acc = 34.9484

53%|█████▎ | 16000/30000 [2:36:13<38:48:04, 9.98s/it]validation score : {'exact\_match': 28.004523491312842, 'f1': 37.91177507436138}

57%|█████▋ | 16999/30000 [2:45:24<1:51:36, 1.94it/s] 17000th iters >> loss = 1.9011, start\_acc = 45.9609, end\_acc = 49.3984, acc = 36.8188

57%|█████▋ | 17000/30000 [2:45:55<34:50:47, 9.65s/it]validation score : {'exact\_match': 29.361832685620783, 'f1': 39.18222869365798}

60%|█████▉ | 17999/30000 [2:55:12<1:47:24, 1.86it/s] 18000th iters >> loss = 1.8716, start\_acc = 46.4172, end\_acc = 50.5078, acc = 37.7016

60%|██████ | 18000/30000 [2:55:43<32:08:27, 9.64s/it]validation score : {'exact\_match': 31.55275718492942, 'f1': 41.17979529289139}

63%|██████▎ | 18999/30000 [3:04:54<2:10:56, 1.40it/s] 19000th iters >> loss = 1.8055, start\_acc = 48.5281, end\_acc = 52.2984, acc = 39.4297

63%|██████▎ | 19000/30000 [3:05:25<29:57:05, 9.80s/it]validation score : {'exact\_match': 30.75979138971265, 'f1': 40.77522737581976}

67%|██████▋ | 19999/30000 [3:14:36<1:23:46, 1.99it/s] 20000th iters >> loss = 1.7444, start\_acc = 49.3672, end\_acc = 53.5406, acc = 40.3813

67%|██████▋ | 20000/30000 [3:15:07<27:19:17, 9.84s/it]validation score : {'exact\_match': 31.832535397779363, 'f1': 41.993728349519046}

70%|██████▉ | 20999/30000 [3:24:22<1:18:52, 1.90it/s] 21000th iters >> loss = 1.6743, start\_acc = 51.2938, end\_acc = 55.2875, acc = 42.0500

70%|███████ | 21000/30000 [3:24:52<23:43:24, 9.49s/it]validation score : {'exact\_match': 31.46788990825688, 'f1': 41.15038903531966}

73%|███████▎ | 21999/30000 [3:33:58<1:16:10, 1.75it/s] 22000th iters >> loss = 1.7215, start\_acc = 50.0609, end\_acc = 54.1516, acc = 40.7578

73%|███████▎ | 22000/30000 [3:34:30<21:39:40, 9.75s/it]validation score : {'exact\_match': 31.798356831321634, 'f1': 42.19686287367456}

77%|███████▋ | 22999/30000 [3:43:41<1:10:06, 1.66it/s] 23000th iters >> loss = 1.7257, start\_acc = 49.9578, end\_acc = 54.0219, acc = 40.6109

77%|███████▋ | 23000/30000 [3:44:12<18:53:18, 9.71s/it]validation score : {'exact\_match': 32.48401177545427, 'f1': 42.559410331814775}

80%|███████▉ | 23999/30000 [3:53:26<52:30, 1.90it/s] 24000th iters >> loss = 1.7590, start\_acc = 48.8891, end\_acc = 53.0297, acc = 39.6422

80%|████████ | 24000/30000 [3:53:57<16:01:05, 9.61s/it]validation score : {'exact\_match': 31.707564388562158, 'f1': 42.1302822226528}

83%|████████▎ | 24999/30000 [4:03:04<45:44, 1.82it/s] 25000th iters >> loss = 3.4864, start\_acc = 21.5344, end\_acc = 22.8094, acc = 15.3484

83%|████████▎ | 25000/30000 [4:03:39<15:06:17, 10.88s/it]validation score : {'exact\_match': 0.03817522427944264, 'f1': 8.66436026902789}

87%|████████▋ | 25999/30000 [4:12:54<39:17, 1.70it/s] 26000th iters >> loss = 5.1510, start\_acc = 2.5375, end\_acc = 0.9578, acc = 0.0516

87%|████████▋ | 26000/30000 [4:13:27<11:25:20, 10.28s/it]validation score : {'exact\_match': 0.2195075396067952, 'f1': 8.023649974266867}

90%|████████▉ | 26999/30000 [4:22:37<25:25, 1.97it/s] 27000th iters >> loss = 5.1259, start\_acc = 2.5531, end\_acc = 1.0063, acc = 0.0281

90%|█████████ | 27000/30000 [4:23:13<9:27:45, 11.36s/it]validation score : {'exact\_match': 0.15270089711777057, 'f1': 8.266144824244176}

93%|█████████▎| 27999/30000 [4:32:28<18:48, 1.77it/s] 28000th iters >> loss = 5.1208, start\_acc = 2.5625, end\_acc = 1.0437, acc = 0.0328

93%|█████████▎| 28000/30000 [4:33:08<6:45:40, 12.17s/it]validation score : {'exact\_match': 0.01908943399828195, 'f1': 8.909000321487243}

97%|█████████▋| 28999/30000 [4:42:21<09:14, 1.81it/s] 29000th iters >> loss = 5.0883, start\_acc = 2.4734, end\_acc = 1.1172, acc = 0.0344

97%|█████████▋| 29000/30000 [4:42:55<3:00:06, 10.81s/it]validation score : {'exact\_match': 0.028631418209581982, 'f1': 8.471574778204907}

100%|█████████▉| 29999/30000 [4:52:08<00:00, 1.75it/s] 30000th iters >> loss = 5.0706, start\_acc = 2.6766, end\_acc = 1.0141, acc = 0.0578

100%|██████████| 30000/30000 [4:52:44<00:00, 1.71it/s]validation score : {'exact\_match': 0.1717885092574919, 'f1': 8.41463484881983}

LSTM\_attn2(

(embedding): Embedding(87599, 300)

(lstm): LSTM(300, 256, num\_layers=3, batch\_first=True, dropout=0.5, bidirectional=True)

(lin1): Linear(in\_features=512, out\_features=256, bias=True)

(attention): Attention(

(k\_proj): Linear(in\_features=256, out\_features=256, bias=True)

(q\_proj): Linear(in\_features=256, out\_features=256, bias=True)

(v\_proj): Linear(in\_features=256, out\_features=256, bias=True)

(out\_proj): Linear(in\_features=256, out\_features=256, bias=True)

(softmax): Softmax(dim=-1)

)

(lstm2): LSTM(256, 256, num\_layers=3, batch\_first=True, dropout=0.5, bidirectional=True)

(relu): ReLU()

(out\_lin): Linear(in\_features=512, out\_features=2, bias=True)

)

vocab\_size = len(word2id)

embed\_dim = 300

hidden\_dim = 128

num\_layers = 2

bidirec = True,

dropout = 0.5

print(vocab\_size)

model = LSTM\_attn2(vocab\_size=vocab\_size,embed\_dim=embed\_dim,hidden\_dim=hidden\_dim,num\_layers=num\_layers,bidirec=bidirec,dropout=dropout)

heads = 3

0%| | 0/30000 [00:00<?, ?it/s]/usr/local/lib/python3.7/dist-packages/ipykernel\_launcher.py:9: VisibleDeprecationWarning: Creating an ndarray from ragged nested sequences (which is a list-or-tuple of lists-or-tuples-or ndarrays with different lengths or shapes) is deprecated. If you meant to do this, you must specify 'dtype=object' when creating the ndarray.

if \_\_name\_\_ == '\_\_main\_\_':

3%|▎ | 998/30000 [01:48<54:00, 8.95it/s] 1000th iters >> loss = 4.3284, start\_acc = 6.0438, end\_acc = 6.2969, acc = 1.6125

3%|▎ | 1001/30000 [02:11<32:47:24, 4.07s/it]validation score : {'exact\_match': 5.611006204478015, 'f1': 11.399516836510395}

7%|▋ | 1999/30000 [04:04<54:10, 8.62it/s] 2000th iters >> loss = 3.8899, start\_acc = 8.6062, end\_acc = 8.8781, acc = 3.0750

7%|▋ | 2001/30000 [04:28<39:36:03, 5.09s/it]validation score : {'exact\_match': 6.534194806778008, 'f1': 12.237907241538537}

10%|▉ | 2998/30000 [06:19<47:05, 9.56it/s] 3000th iters >> loss = 3.7878, start\_acc = 9.1594, end\_acc = 9.7312, acc = 3.5937

10%|█ | 3001/30000 [06:41<21:56:46, 2.93s/it]validation score : {'exact\_match': 6.649266716380811, 'f1': 12.355696931140773}

13%|█▎ | 3999/30000 [08:33<49:15, 8.80it/s] 4000th iters >> loss = 3.7040, start\_acc = 10.1750, end\_acc = 10.4875, acc = 4.5219

13%|█▎ | 4001/30000 [08:54<27:53:44, 3.86s/it]validation score : {'exact\_match': 6.860229574611749, 'f1': 12.593706505161775}

17%|█▋ | 4998/30000 [10:45<43:12, 9.65it/s] 5000th iters >> loss = 3.6669, start\_acc = 10.3250, end\_acc = 10.9969, acc = 4.7313

17%|█▋ | 5000/30000 [11:08<27:30:19, 3.96s/it]validation score : {'exact\_match': 6.783860311467674, 'f1': 11.917138037335324}

20%|█▉ | 5999/30000 [13:01<49:41, 8.05it/s] 6000th iters >> loss = 3.5758, start\_acc = 11.5781, end\_acc = 12.4094, acc = 6.0344

20%|██ | 6001/30000 [13:23<29:53:35, 4.48s/it]validation score : {'exact\_match': 10.821209570758686, 'f1': 16.70472767864729}

23%|██▎ | 6999/30000 [15:14<39:31, 9.70it/s] 7000th iters >> loss = 3.3324, start\_acc = 16.2156, end\_acc = 17.6000, acc = 9.7156

23%|██▎ | 7001/30000 [15:35<23:54:16, 3.74s/it]validation score : {'exact\_match': 15.287221570926143, 'f1': 22.15095947766638}

27%|██▋ | 7999/30000 [17:26<37:40, 9.73it/s] 8000th iters >> loss = 3.1876, start\_acc = 18.3219, end\_acc = 20.2469, acc = 11.6188

27%|██▋ | 8001/30000 [17:47<20:26:08, 3.34s/it]validation score : {'exact\_match': 17.916865742952698, 'f1': 25.270531569878763}

30%|██▉ | 8999/30000 [19:39<36:03, 9.70it/s] 9000th iters >> loss = 3.0822, start\_acc = 20.8094, end\_acc = 22.3813, acc = 13.5563

30%|███ | 9001/30000 [19:59<18:55:48, 3.25s/it]validation score : {'exact\_match': 17.547793249940995, 'f1': 25.727476802631895}

33%|███▎ | 9998/30000 [21:49<37:18, 8.93it/s] 10000th iters >> loss = 3.0069, start\_acc = 22.0062, end\_acc = 23.8281, acc = 14.3750

33%|███▎ | 10001/30000 [22:10<17:48:25, 3.21s/it]validation score : {'exact\_match': 19.627591052816488, 'f1': 27.983121165904834}

37%|███▋ | 10998/30000 [23:59<30:04, 10.53it/s] 11000th iters >> loss = 2.9280, start\_acc = 23.6594, end\_acc = 25.4406, acc = 16.0875

37%|███▋ | 11001/30000 [24:20<14:54:37, 2.83s/it]validation score : {'exact\_match': 20.455873758036237, 'f1': 28.714773097657353}

40%|███▉ | 11998/30000 [26:11<31:25, 9.55it/s] 12000th iters >> loss = 2.8649, start\_acc = 24.8344, end\_acc = 27.2313, acc = 17.5063

40%|████ | 12001/30000 [26:31<16:20:51, 3.27s/it]validation score : {'exact\_match': 21.81591758179554, 'f1': 30.04216619154175}

43%|████▎ | 12999/30000 [28:23<29:03, 9.75it/s] 13000th iters >> loss = 2.7946, start\_acc = 26.1719, end\_acc = 28.5938, acc = 18.5250

43%|████▎ | 13001/30000 [28:41<15:11:11, 3.22s/it]validation score : {'exact\_match': 22.69091705927468, 'f1': 30.850262074546194}

47%|████▋ | 13999/30000 [30:32<28:26, 9.38it/s] 14000th iters >> loss = 2.7843, start\_acc = 26.7875, end\_acc = 29.2250, acc = 19.0344

47%|████▋ | 14001/30000 [30:51<14:48:41, 3.33s/it]validation score : {'exact\_match': 23.451083966105426, 'f1': 31.277332422443145}

50%|████▉ | 14999/30000 [32:42<27:53, 8.97it/s] 15000th iters >> loss = 2.6948, start\_acc = 28.1438, end\_acc = 30.6750, acc = 20.2406

50%|█████ | 15001/30000 [33:01<14:32:29, 3.49s/it]validation score : {'exact\_match': 24.06879929886065, 'f1': 32.75164034106435}

53%|█████▎ | 15999/30000 [34:52<24:31, 9.51it/s] 16000th iters >> loss = 2.6696, start\_acc = 28.9312, end\_acc = 31.1594, acc = 20.8500

53%|█████▎ | 16001/30000 [35:12<15:26:05, 3.97s/it]validation score : {'exact\_match': 23.504320244996173, 'f1': 31.651816805975496}

57%|█████▋ | 16999/30000 [37:03<24:52, 8.71it/s] 17000th iters >> loss = 2.6241, start\_acc = 29.6750, end\_acc = 32.1781, acc = 21.7844

57%|█████▋ | 17001/30000 [37:21<13:15:45, 3.67s/it]validation score : {'exact\_match': 24.655888136333843, 'f1': 33.486055839365584}

60%|█████▉ | 17999/30000 [39:14<23:01, 8.69it/s] 18000th iters >> loss = 2.6178, start\_acc = 29.7938, end\_acc = 31.9906, acc = 21.6062

60%|██████ | 18001/30000 [39:33<13:03:14, 3.92s/it]validation score : {'exact\_match': 24.989182172219817, 'f1': 33.362492826521645}

63%|██████▎ | 18999/30000 [41:25<25:24, 7.22it/s] 19000th iters >> loss = 2.5750, start\_acc = 31.0781, end\_acc = 33.5063, acc = 22.8313

63%|██████▎ | 19001/30000 [41:44<11:50:40, 3.88s/it]validation score : {'exact\_match': 24.986623863028356, 'f1': 33.608475178106}

67%|██████▋ | 19999/30000 [43:34<19:58, 8.34it/s] 20000th iters >> loss = 2.5444, start\_acc = 31.6250, end\_acc = 33.7750, acc = 23.1562

67%|██████▋ | 20000/30000 [43:55<16:55:26, 6.09s/it]validation score : {'exact\_match': 25.715823466092573, 'f1': 34.884653539341755}

70%|██████▉ | 20999/30000 [45:46<17:27, 8.60it/s] 21000th iters >> loss = 2.5286, start\_acc = 32.1062, end\_acc = 34.5562, acc = 23.5969

70%|███████ | 21001/30000 [46:04<8:51:30, 3.54s/it] validation score : {'exact\_match': 26.150069377735083, 'f1': 35.33153740699517}

73%|███████▎ | 21999/30000 [47:55<15:28, 8.62it/s] 22000th iters >> loss = 2.5056, start\_acc = 32.4906, end\_acc = 34.9438, acc = 24.0688

73%|███████▎ | 22001/30000 [48:15<8:14:23, 3.71s/it] validation score : {'exact\_match': 26.77114641477029, 'f1': 36.119906976567805}

77%|███████▋ | 22998/30000 [50:05<12:28, 9.36it/s] 23000th iters >> loss = 2.4904, start\_acc = 33.0562, end\_acc = 35.4719, acc = 24.4562

77%|███████▋ | 23001/30000 [50:23<5:36:13, 2.88s/it]validation score : {'exact\_match': 26.700010641694156, 'f1': 36.21097358434735}

80%|███████▉ | 23999/30000 [52:15<11:21, 8.81it/s] 24000th iters >> loss = 2.4861, start\_acc = 32.9500, end\_acc = 35.6125, acc = 24.3750

80%|████████ | 24000/30000 [52:34<7:38:17, 4.58s/it]validation score : {'exact\_match': 26.962566844919785, 'f1': 36.08754739626627}

83%|████████▎ | 24999/30000 [54:27<09:22, 8.89it/s] 25000th iters >> loss = 2.5332, start\_acc = 31.5750, end\_acc = 34.0594, acc = 23.1656

83%|████████▎ | 25001/30000 [54:46<4:53:52, 3.53s/it]validation score : {'exact\_match': 27.38246505717916, 'f1': 36.34899630097603}

87%|████████▋ | 25999/30000 [56:38<07:37, 8.75it/s] 26000th iters >> loss = 3.5761, start\_acc = 15.5375, end\_acc = 16.6344, acc = 9.5562

87%|████████▋ | 26001/30000 [56:59<4:31:10, 4.07s/it]validation score : {'exact\_match': 4.388248419486798, 'f1': 11.323439760351018}

90%|████████▉ | 26999/30000 [58:51<05:19, 9.40it/s] 27000th iters >> loss = 4.7670, start\_acc = 3.6281, end\_acc = 3.9125, acc = 0.8375

90%|█████████ | 27001/30000 [59:11<3:15:46, 3.92s/it]validation score : {'exact\_match': 0.0763577359931278, 'f1': 8.229606306704584}

93%|█████████▎| 27999/30000 [1:01:01<03:57, 8.44it/s] 28000th iters >> loss = 5.3858, start\_acc = 2.5000, end\_acc = 0.8625, acc = 0.2312

93%|█████████▎| 28001/30000 [1:01:18<1:51:04, 3.33s/it]validation score : {'exact\_match': 0.591715976331361, 'f1': 5.9744684919269675}

97%|█████████▋| 28999/30000 [1:03:07<01:42, 9.76it/s] 29000th iters >> loss = 5.1908, start\_acc = 2.4938, end\_acc = 1.0125, acc = 0.2281

97%|█████████▋| 29000/30000 [1:03:28<1:19:31, 4.77s/it]validation score : {'exact\_match': 0.2328062857697158, 'f1': 7.565274545549647}

100%|█████████▉| 29999/30000 [1:05:19<00:00, 9.21it/s] 30000th iters >> loss = 4.9555, start\_acc = 2.6469, end\_acc = 0.9875, acc = 0.2031

100%|██████████| 30000/30000 [1:05:35<00:00, 7.62it/s]validation score : {'exact\_match': 0.7253292613094102, 'f1': 6.058316828090047}

LSTM\_attn2(

(embedding): Embedding(79423, 300)

(lstm): LSTM(300, 128, num\_layers=2, batch\_first=True, dropout=0.5, bidirectional=True)

(lin1): Linear(in\_features=256, out\_features=128, bias=True)

(attention): Attention(

(k\_proj): Linear(in\_features=128, out\_features=128, bias=True)

(q\_proj): Linear(in\_features=128, out\_features=128, bias=True)

(v\_proj): Linear(in\_features=128, out\_features=128, bias=True)

(out\_proj): Linear(in\_features=128, out\_features=128, bias=True)

(softmax): Softmax(dim=-1)

)

(lstm2): LSTM(128, 128, num\_layers=2, batch\_first=True, dropout=0.5, bidirectional=True)

(relu): ReLU()

(out\_lin): Linear(in\_features=256, out\_features=2, bias=True)

español

vocab\_size = len(word2id)

embed\_dim = 300

hidden\_dim = 128

num\_layers = 2

bidirec = True,

dropout = 0.5

4 heads

0%| | 0/30000 [00:00<?, ?it/s]/usr/local/lib/python3.7/dist-packages/ipykernel\_launcher.py:9: VisibleDeprecationWarning: Creating an ndarray from ragged nested sequences (which is a list-or-tuple of lists-or-tuples-or ndarrays with different lengths or shapes) is deprecated. If you meant to do this, you must specify 'dtype=object' when creating the ndarray.

if \_\_name\_\_ == '\_\_main\_\_':

/usr/local/lib/python3.7/dist-packages/ipykernel\_launcher.py:61: UserWarning: Creating a tensor from a list of numpy.ndarrays is extremely slow. Please consider converting the list to a single numpy.ndarray with numpy.array() before converting to a tensor. (Triggered internally at ../torch/csrc/utils/tensor\_new.cpp:201.)

3%|▎ | 999/30000 [03:05<1:28:17, 5.47it/s] 1000th iters >> loss = 4.3556, start\_acc = 6.0047, end\_acc = 6.8078, acc = 1.7750

3%|▎ | 1000/30000 [03:26<52:53:42, 6.57s/it]validation score : {'exact\_match': 5.540933647319573, 'f1': 11.918024324991771}

7%|▋ | 1999/30000 [06:31<1:32:03, 5.07it/s] 2000th iters >> loss = 3.9629, start\_acc = 8.7844, end\_acc = 9.5719, acc = 3.5562

7%|▋ | 2001/30000 [06:52<35:07:14, 4.52s/it]validation score : {'exact\_match': 5.191748580922055, 'f1': 12.58425376984798}

10%|▉ | 2999/30000 [09:58<1:19:02, 5.69it/s] 3000th iters >> loss = 3.8104, start\_acc = 10.5594, end\_acc = 11.7000, acc = 4.9859

10%|█ | 3001/30000 [10:18<32:48:42, 4.38s/it]validation score : {'exact\_match': 8.879657276385824, 'f1': 16.97443544769878}

13%|█▎ | 3999/30000 [13:24<1:19:45, 5.43it/s] 4000th iters >> loss = 3.5386, start\_acc = 15.4016, end\_acc = 16.9672, acc = 8.5750

13%|█▎ | 4000/30000 [13:44<43:29:14, 6.02s/it]validation score : {'exact\_match': 13.24110671936759, 'f1': 22.35447525800842}

17%|█▋ | 4999/30000 [16:51<1:28:01, 4.73it/s] 5000th iters >> loss = 3.3697, start\_acc = 18.0297, end\_acc = 20.1391, acc = 10.9281

17%|█▋ | 5000/30000 [17:12<43:21:44, 6.24s/it]validation score : {'exact\_match': 15.207982688146188, 'f1': 25.183738715200917}

20%|█▉ | 5999/30000 [20:21<1:05:58, 6.06it/s] 6000th iters >> loss = 3.2476, start\_acc = 20.0609, end\_acc = 22.4219, acc = 12.6828

20%|██ | 6000/30000 [20:39<38:24:39, 5.76s/it]validation score : {'exact\_match': 16.44160151622838, 'f1': 26.237885844973636}

23%|██▎ | 6999/30000 [23:46<1:12:36, 5.28it/s] 7000th iters >> loss = 3.1482, start\_acc = 21.8875, end\_acc = 24.2906, acc = 14.2375

23%|██▎ | 7001/30000 [24:06<27:31:51, 4.31s/it]validation score : {'exact\_match': 17.963721474546517, 'f1': 28.58029568931576}

27%|██▋ | 7999/30000 [27:10<1:09:01, 5.31it/s] 8000th iters >> loss = 3.0387, start\_acc = 23.8531, end\_acc = 26.3922, acc = 15.9609

27%|██▋ | 8001/30000 [27:28<23:49:27, 3.90s/it]validation score : {'exact\_match': 18.736818736818737, 'f1': 29.08747065124157}

30%|██▉ | 8999/30000 [30:34<1:13:22, 4.77it/s] 9000th iters >> loss = 2.9642, start\_acc = 25.4922, end\_acc = 28.4625, acc = 17.3063

30%|███ | 9001/30000 [30:54<24:30:52, 4.20s/it]validation score : {'exact\_match': 18.686924697793057, 'f1': 29.891950299642144}

33%|███▎ | 9999/30000 [34:00<1:01:26, 5.42it/s] 10000th iters >> loss = 2.8775, start\_acc = 26.9797, end\_acc = 30.0594, acc = 18.8266

33%|███▎ | 10000/30000 [34:18<31:24:34, 5.65s/it]validation score : {'exact\_match': 18.690095846645367, 'f1': 30.83104460272911}

37%|███▋ | 10999/30000 [37:22<55:22, 5.72it/s] 11000th iters >> loss = 2.7989, start\_acc = 28.4688, end\_acc = 31.6531, acc = 19.9031

37%|███▋ | 11001/30000 [37:41<21:33:11, 4.08s/it]validation score : {'exact\_match': 19.682886823400764, 'f1': 31.018039482702452}

40%|███▉ | 11999/30000 [40:46<1:01:31, 4.88it/s] 12000th iters >> loss = 2.7400, start\_acc = 29.7188, end\_acc = 33.3516, acc = 21.1719

40%|████ | 12001/30000 [41:04<19:03:07, 3.81s/it]validation score : {'exact\_match': 20.063036626453645, 'f1': 31.43098024645888}

43%|████▎ | 12999/30000 [44:09<53:06, 5.33it/s] 13000th iters >> loss = 2.6794, start\_acc = 30.7391, end\_acc = 34.6703, acc = 21.9656

43%|████▎ | 13000/30000 [44:28<27:52:53, 5.90s/it]validation score : {'exact\_match': 20.739705408020644, 'f1': 32.408300348743154}

47%|████▋ | 13999/30000 [47:35<50:45, 5.25it/s] 14000th iters >> loss = 2.6239, start\_acc = 31.8813, end\_acc = 36.0328, acc = 23.0578

47%|████▋ | 14001/30000 [47:52<16:49:00, 3.78s/it]validation score : {'exact\_match': 21.734463883224215, 'f1': 33.553360478636726}

50%|████▉ | 14999/30000 [50:59<48:26, 5.16it/s] 15000th iters >> loss = 2.5420, start\_acc = 33.5922, end\_acc = 37.6672, acc = 24.4875

50%|█████ | 15001/30000 [51:17<16:29:23, 3.96s/it]validation score : {'exact\_match': 22.523892267593396, 'f1': 35.46176452898489}

53%|█████▎ | 15999/30000 [54:27<41:27, 5.63it/s] 16000th iters >> loss = 2.4821, start\_acc = 34.8062, end\_acc = 39.2172, acc = 25.6016

53%|█████▎ | 16001/30000 [54:44<14:37:53, 3.76s/it]validation score : {'exact\_match': 24.03035413153457, 'f1': 36.52533000084613}

57%|█████▋ | 16999/30000 [57:53<42:22, 5.11it/s] 17000th iters >> loss = 2.4084, start\_acc = 36.6000, end\_acc = 40.9969, acc = 27.1047

57%|█████▋ | 17001/30000 [58:11<13:57:02, 3.86s/it]validation score : {'exact\_match': 24.249061326658325, 'f1': 36.92250479717108}

60%|█████▉ | 17999/30000 [1:01:18<37:41, 5.31it/s] 18000th iters >> loss = 2.3756, start\_acc = 37.1078, end\_acc = 41.7422, acc = 27.4500

60%|██████ | 18000/30000 [1:01:36<18:23:30, 5.52s/it]validation score : {'exact\_match': 25.441991840150642, 'f1': 39.000139828907564}

63%|██████▎ | 18999/30000 [1:04:45<41:07, 4.46it/s] 19000th iters >> loss = 2.3248, start\_acc = 38.2063, end\_acc = 43.1297, acc = 28.4859

63%|██████▎ | 19001/30000 [1:05:02<11:21:33, 3.72s/it]validation score : {'exact\_match': 26.008537886873, 'f1': 39.70754315650823}

67%|██████▋ | 19999/30000 [1:08:11<34:01, 4.90it/s] 20000th iters >> loss = 2.3093, start\_acc = 38.3813, end\_acc = 43.1937, acc = 28.5219

67%|██████▋ | 20000/30000 [1:08:29<15:33:58, 5.60s/it]validation score : {'exact\_match': 25.80206918173268, 'f1': 39.183253680569955}

70%|██████▉ | 20999/30000 [1:11:38<31:12, 4.81it/s] 21000th iters >> loss = 2.2821, start\_acc = 39.1281, end\_acc = 44.1109, acc = 29.4875

70%|███████ | 21001/30000 [1:11:54<8:59:44, 3.60s/it] validation score : {'exact\_match': 27.22648951629645, 'f1': 40.52006959547451}

73%|███████▎ | 21999/30000 [1:15:01<25:35, 5.21it/s] 22000th iters >> loss = 2.2311, start\_acc = 39.9609, end\_acc = 44.8813, acc = 29.9094

73%|███████▎ | 22001/30000 [1:15:19<8:28:34, 3.81s/it] validation score : {'exact\_match': 27.620528771384137, 'f1': 41.68119009364039}

77%|███████▋ | 22999/30000 [1:18:26<22:01, 5.30it/s] 23000th iters >> loss = 2.1683, start\_acc = 41.2328, end\_acc = 46.5859, acc = 31.3141

77%|███████▋ | 23001/30000 [1:18:43<6:58:27, 3.59s/it]validation score : {'exact\_match': 28.132754342431763, 'f1': 42.28743546981523}

80%|███████▉ | 23999/30000 [1:21:50<18:01, 5.55it/s] 24000th iters >> loss = 2.1486, start\_acc = 41.9859, end\_acc = 46.9891, acc = 31.7000

80%|████████ | 24001/30000 [1:22:08<6:33:17, 3.93s/it]validation score : {'exact\_match': 27.85558667911609, 'f1': 42.45329027113317}

83%|████████▎ | 24999/30000 [1:25:14<14:48, 5.63it/s] 25000th iters >> loss = 2.1189, start\_acc = 42.1031, end\_acc = 47.5141, acc = 31.8734

83%|████████▎ | 25000/30000 [1:25:32<7:27:26, 5.37s/it]validation score : {'exact\_match': 27.80255107331743, 'f1': 42.885721139925145}

87%|████████▋ | 25999/30000 [1:28:37<11:44, 5.68it/s] 26000th iters >> loss = 2.1016, start\_acc = 42.6250, end\_acc = 47.6828, acc = 32.3172

87%|████████▋ | 26000/30000 [1:28:55<6:04:48, 5.47s/it]validation score : {'exact\_match': 29.857333470183722, 'f1': 44.88793554005786}

90%|████████▉ | 26999/30000 [1:32:01<08:42, 5.75it/s] 27000th iters >> loss = 2.0745, start\_acc = 43.0734, end\_acc = 48.4734, acc = 32.9047

90%|█████████ | 27001/30000 [1:32:19<3:11:50, 3.84s/it]validation score : {'exact\_match': 29.153313346428202, 'f1': 44.58212764312744}

93%|█████████▎| 27999/30000 [1:35:27<05:58, 5.58it/s] 28000th iters >> loss = 2.0889, start\_acc = 43.0156, end\_acc = 48.1359, acc = 32.5891

93%|█████████▎| 28001/30000 [1:35:43<1:58:02, 3.54s/it]validation score : {'exact\_match': 29.995888157894736, 'f1': 44.598636403436615}

97%|█████████▋| 28999/30000 [1:38:50<02:58, 5.61it/s] 29000th iters >> loss = 2.0469, start\_acc = 43.7672, end\_acc = 49.0125, acc = 33.3891

97%|█████████▋| 29001/30000 [1:39:07<1:02:17, 3.74s/it]validation score : {'exact\_match': 30.22000407414952, 'f1': 45.2402331221328}

100%|█████████▉| 29999/30000 [1:42:15<00:00, 4.86it/s] 30000th iters >> loss = 2.0301, start\_acc = 43.8937, end\_acc = 49.6016, acc = 33.5625

100%|██████████| 30000/30000 [1:42:32<00:00, 4.88it/s]validation score : {'exact\_match': 29.451632778804683, 'f1': 44.97207177522891}

LSTM\_attn2(

(embedding): Embedding(90281, 300)

(lstm): LSTM(300, 128, num\_layers=2, batch\_first=True, dropout=0.5, bidirectional=True)

(lin1): Linear(in\_features=256, out\_features=128, bias=True)

(attention): Attention(

(k\_proj): Linear(in\_features=128, out\_features=128, bias=True)

(q\_proj): Linear(in\_features=128, out\_features=128, bias=True)

(v\_proj): Linear(in\_features=128, out\_features=128, bias=True)

(out\_proj): Linear(in\_features=128, out\_features=128, bias=True)

(softmax): Softmax(dim=-1)

)

(lstm2): LSTM(128, 128, num\_layers=2, batch\_first=True, dropout=0.5, bidirectional=True)

(relu): ReLU()

(out\_lin): Linear(in\_features=256, out\_features=2, bias=True)

)