

Python 3.8.8 (default, Feb 24 2021, 15:54:32) [MSC v.1928 64 bit (AMD64)]

Type "copyright", "credits" or "license" for more information.

IPython 7.18.1 -- An enhanced Interactive Python.

Removing all variables...

```
runfile('D:/PolyU Course/COMP5423 Natural Language  
Processing/2021Feb/Project/ProjectCode/run_race_test.py', wdir='D:/PolyU  
Course/COMP5423 Natural Language Processing/2021Feb/Project/ProjectCode')
```

level middle

After sorted: ./datasets/RACE/train/middle/2.txt

level high

After sorted: ./datasets/RACE/train/high/1.txt

Optimization Step: 7418

Freeze network

bert.embeddings.word_embeddings.weight

bert.embeddings.position_embeddings.weight

bert.embeddings.token_type_embeddings.weight

bert.embeddings.LayerNorm.weight

bert.embeddings.LayerNorm.bias

bert.encoder.layer.0.attention.self.query.weight

bert.encoder.layer.0.attention.self.query.bias
bert.encoder.layer.0.attention.self.key.weight
bert.encoder.layer.0.attention.self.key.bias
bert.encoder.layer.0.attention.self.value.weight
bert.encoder.layer.0.attention.self.value.bias
bert.encoder.layer.0.attention.output.dense.weight
bert.encoder.layer.0.attention.output.dense.bias
bert.encoder.layer.0.attention.output.LayerNorm.weight
bert.encoder.layer.0.attention.output.LayerNorm.bias
bert.encoder.layer.0.intermediate.dense.weight
bert.encoder.layer.0.intermediate.dense.bias
bert.encoder.layer.0.output.dense.weight
bert.encoder.layer.0.output.dense.bias
bert.encoder.layer.0.output.LayerNorm.weight
bert.encoder.layer.0.output.LayerNorm.bias
bert.encoder.layer.1.attention.self.query.weight
bert.encoder.layer.1.attention.self.query.bias
bert.encoder.layer.1.attention.self.key.weight
bert.encoder.layer.1.attention.self.key.bias
bert.encoder.layer.1.attention.self.value.weight
bert.encoder.layer.1.attention.self.value.bias
bert.encoder.layer.1.attention.output.dense.weight
bert.encoder.layer.1.attention.output.dense.bias
bert.encoder.layer.1.attention.output.LayerNorm.weight
bert.encoder.layer.1.attention.output.LayerNorm.bias
bert.encoder.layer.1.intermediate.dense.weight

bert.encoder.layer.1.intermediate.dense.bias
bert.encoder.layer.1.output.dense.weight
bert.encoder.layer.1.output.dense.bias
bert.encoder.layer.1.output.LayerNorm.weight
bert.encoder.layer.1.output.LayerNorm.bias
bert.encoder.layer.2.attention.self.query.weight
bert.encoder.layer.2.attention.self.query.bias
bert.encoder.layer.2.attention.self.key.weight
bert.encoder.layer.2.attention.self.key.bias
bert.encoder.layer.2.attention.self.value.weight
bert.encoder.layer.2.attention.self.value.bias
bert.encoder.layer.2.attention.output.dense.weight
bert.encoder.layer.2.attention.output.dense.bias
bert.encoder.layer.2.attention.output.LayerNorm.weight
bert.encoder.layer.2.attention.output.LayerNorm.bias
bert.encoder.layer.2.intermediate.dense.weight
bert.encoder.layer.2.intermediate.dense.bias
bert.encoder.layer.2.output.dense.weight
bert.encoder.layer.2.output.dense.bias
bert.encoder.layer.2.output.LayerNorm.weight
bert.encoder.layer.2.output.LayerNorm.bias
bert.encoder.layer.3.attention.self.query.weight
bert.encoder.layer.3.attention.self.query.bias
bert.encoder.layer.3.attention.self.key.weight
bert.encoder.layer.3.attention.self.key.bias
bert.encoder.layer.3.attention.self.value.weight

bert.encoder.layer.3.attention.self.value.bias
bert.encoder.layer.3.attention.output.dense.weight
bert.encoder.layer.3.attention.output.dense.bias
bert.encoder.layer.3.attention.output.LayerNorm.weight
bert.encoder.layer.3.attention.output.LayerNorm.bias
bert.encoder.layer.3.intermediate.dense.weight
bert.encoder.layer.3.intermediate.dense.bias
bert.encoder.layer.3.output.dense.weight
bert.encoder.layer.3.output.dense.bias
bert.encoder.layer.3.output.LayerNorm.weight
bert.encoder.layer.3.output.LayerNorm.bias
bert.encoder.layer.4.attention.self.query.weight
bert.encoder.layer.4.attention.self.query.bias
bert.encoder.layer.4.attention.self.key.weight
bert.encoder.layer.4.attention.self.key.bias
bert.encoder.layer.4.attention.self.value.weight
bert.encoder.layer.4.attention.self.value.bias
bert.encoder.layer.4.attention.output.dense.weight
bert.encoder.layer.4.attention.output.dense.bias
bert.encoder.layer.4.attention.output.LayerNorm.weight
bert.encoder.layer.4.attention.output.LayerNorm.bias
bert.encoder.layer.4.intermediate.dense.weight
bert.encoder.layer.4.intermediate.dense.bias
bert.encoder.layer.4.output.dense.weight
bert.encoder.layer.4.output.dense.bias
bert.encoder.layer.4.output.LayerNorm.weight

bert.encoder.layer.4.output.LayerNorm.bias
bert.encoder.layer.5.attention.self.query.weight
bert.encoder.layer.5.attention.self.query.bias
bert.encoder.layer.5.attention.self.key.weight
bert.encoder.layer.5.attention.self.key.bias
bert.encoder.layer.5.attention.self.value.weight
bert.encoder.layer.5.attention.self.value.bias
bert.encoder.layer.5.attention.output.dense.weight
bert.encoder.layer.5.attention.output.dense.bias
bert.encoder.layer.5.attention.output.LayerNorm.weight
bert.encoder.layer.5.attention.output.LayerNorm.bias
bert.encoder.layer.5.intermediate.dense.weight
bert.encoder.layer.5.intermediate.dense.bias
bert.encoder.layer.5.output.dense.weight
bert.encoder.layer.5.output.dense.bias
bert.encoder.layer.5.output.LayerNorm.weight
bert.encoder.layer.5.output.LayerNorm.bias

Length of Example: 22256 race_id: middle2.txt:0, context_sentence: Drinking water is good for your health. There are some scientific ways of drinking water.

1. It is the best medicine to drink two glasses of water in the morning.
2. Drink clean water.
3. Drink the water that has not been boiled.

Many people think boiled water is safe and good to people's health. In fact, it is not true. The boiling point of water is 100degC. By boiling it, most bacteria in water can be killed. In the past, the water was less polluted. So boiling was a good way to make clean water. But heavy metals and other dangerous things in today's water are much more terrible than bacteria. Boiling doesn't fix that problem. And boiling water may give us more of the dangerous things in our glass.

4. Never use soft drinks to take the place of water.

5. Water is also needed in winter.

6. Drink water at the right time.

1) After getting up in the morning, you have less water in your body, because you weren't drinking for the whole night. So you should drink some water to keep your health after getting up in the morning. That can prevent high blood pressure, cerebral hemorrhages and so on.

2) Drinking water at about 10 am helps your body keep enough water.

3) Drinking water at about 3 pm can clean out the wastes in your body.

4) About eight o'clock in the evening is the best time to drink water. Your blood gets thicker when you sleep. Water will make your blood less thick.

Besides, we should drink 2L of water every day. Water is so important for our life. We should drink water often., start_ending: According to the passage, _ in the morning is the best _ ., ending_0: drinking some hot soup; medicine, ending_1: drinking some porridge; breakfast, ending_2: drinking some water; medicine, ending_3: Drinking some soft drinks; medicine, label: 2

*** Example ***

race_id: middle2.txt:0

choice: 0

tokens: [CLS] drinking water is good for your health . there are some scientific ways of drinking water . 1 . it is the best medicine to drink two glasses of water in the morning . 2 . drink clean water . 3 . drink the water that has not been boiled . many people think boiled water is safe and good to people ' s health . in fact , it is not true . the boiling point of water is 100 ##de ##gc . by boiling it , most bacteria in water can be killed . in the past , the water was less poll ##uted . so boiling was a good way to make clean water . but heavy metals and other dangerous things in today ' s water are much more terrible than bacteria . boiling doesn ' t fix that problem . and boiling water may give us more of the dangerous things in our glass . 4 . never use soft drinks to take the place of water . 5 . water is also needed in winter . 6 . drink water at the right time . 1) after getting up in the morning , you have less water in your body , because you weren ' t drinking for the whole night . so you should drink some water to keep your health after getting up in the morning . that can prevent high blood pressure , cerebral hem ##or ##rh ##ages and so on . 2) drinking water at about 10 am helps your body keep enough water . 3) drinking water at about 3 pm can clean out the waste ##s in your body . 4) about eight o ' clock in the evening is the best time to drink water . your blood gets thicker when you sleep . water will make your blood less thick . besides , we should drink 2 ##l of water every day . water is so important for our life . we should drink water often .

[SEP] according to the passage , _ in the morning is the best _ . drinking some hot soup ; medicine [SEP]

[illegible][illegible]

[illegible][illegible][illegible]

choice: 2

tokens: [CLS] drinking water is good for your health . there are some scientific ways of drinking water . 1 . it is the best medicine to drink two glasses of water in the morning . 2 . drink clean water . 3 . drink the water that has not been boiled . many people think boiled water is safe and good to people ' s health . in fact , it is not true . the boiling point of water is 100 ##de ##gc . by boiling it , most bacteria in water can be killed . in the past , the water was less poll ##uted . so boiling was a good way to make clean water . but heavy metals and other dangerous things in today ' s water are much more terrible than bacteria . boiling doesn ' t fix that problem . and boiling water may give us more of the dangerous things in our glass . 4 . never use soft drinks to take the place of water . 5 . water is also needed in winter . 6 . drink water at the right time . 1) after getting up in the morning , you have less water in your body , because you weren ' t drinking for the whole night . so you should drink some water to keep your health after getting up in the morning . that can prevent high blood pressure , cerebral hem ##or ##rh ##ages and so on . 2) drinking water at about 10 am helps your body keep enough water . 3) drinking water at about 3 pm can clean out the waste ##s in your body . 4) about eight o ' clock in the evening is the best time to drink water . your blood gets thicker when you sleep . water will make your blood less thick . besides , we should drink 2 ##l of water every day . water is so important for our life . we should drink water often . [SEP] according to the passage , _ in the morning is the best _ . drinking some water ; medicine [SEP]

input_ids: 101 5948 2300 2003 2204 2005 2115 2740 1012 2045 2024 2070 4045
3971 1997 5948 2300 1012 1015 1012 2009 2003 1996 2190 4200 2000 4392
2048 7877 1997 2300 1999 1996 2851 1012 1016 1012 4392 4550 2300 1012
1017 1012 4392 1996 2300 2008 2038 2025 2042 17020 1012 2116 2111 2228
17020 2300 2003 3647 1998 2204 2000 2111 1005 1055 2740 1012 1999 2755
1010 2009 2003 2025 2995 1012 1996 16018 2391 1997 2300 2003 2531 3207
18195 1012 2011 16018 2009 1010 2087 10327 1999 2300 2064 2022 2730 1012
1999 1996 2627 1010 1996 2300 2001 2625 8554 12926 1012 2061 16018 2001
1037 2204 2126 2000 2191 4550 2300 1012 2021 3082 11970 1998 2060 4795
2477 1999 2651 1005 1055 2300 2024 2172 2062 6659 2084 10327 1012 16018
2987 1005 1056 8081 2008 3291 1012 1998 16018 2300 2089 2507 2149 2062
1997 1996 4795 2477 1999 2256 3221 1012 1018 1012 2196 2224 3730 8974
2000 2202 1996 2173 1997 2300 1012 1019 1012 2300 2003 2036 2734 1999
3467 1012 1020 1012 4392 2300 2012 1996 2157 2051 1012 1015 1007 2044
2893 2039 1999 1996 2851 1010 2017 2031 2625 2300 1999 2115 2303 1010
2138 2017 4694 1005 1056 5948 2005 1996 2878 2305 1012 2061 2017 2323
4392 2070 2300 2000 2562 2115 2740 2044 2893 2039 1999 1996 2851 1012
2008 2064 4652 2152 2668 3778 1010 18439 19610 2953 25032 13923 1998 2061
2006 1012 1016 1007 5948 2300 2012 2055 2184 2572 7126 2115 2303 2562
2438 2300 1012 1017 1007 5948 2300 2012 2055 1017 7610 2064 4550 2041

[illegible][illegible]

label: 2

0

5000

10000

15000

20000

***** Running training *****

Num examples = 22256

Batch size = 2

Num steps = 7418

Iteration: 0%| | 2/11128 [00:01<2:04:00, 1.50it/s]c:\ProgramData\Anaconda3\envs\pytorch16+\lib\site-packages\pytorch_pretrained_bert\optimization.py:275: UserWarning: This overload of add_ is deprecated:

add_(Number alpha, Tensor other)

Consider using one of the following signatures instead:

add_(Tensor other, *, Number alpha) (Triggered internally at ..\torch\csrc\utils\python_arg_parser.cpp:766.)

next_m.mul_(beta1).add_(1 - beta1, grad)

Iteration: 2%|| | 256/11128 [03:17<2:12:47, 1.36it/s]

Num Of Epoch: 0 , Num Of Step: 512 , Loss: 0.233

Iteration: 5%||| | 512/11128 [06:38<2:11:10, 1.35it/s]

Num Of Epoch: 0 , Num Of Step: 1024 , Loss: 0.231

Iteration: 7%||| | 768/11128 [10:01<2:33:28, 1.13it/s]

Num Of Epoch: 0 , Num Of Step: 1536 , Loss: 0.229

Iteration: 9%||| | 1024/11128 [13:24<2:06:25, 1.33it/s]

Num Of Epoch: 0 , Num Of Step: 2048 , Loss: 0.229

Iteration: 12%||| | 1280/11128 [16:51<2:06:01, 1.30it/s]

Num Of Epoch: 0 , Num Of Step: 2560 , Loss: 0.232

Iteration: 14%||| | 1536/11128 [23:24<2:18:11, 1.16it/s]

Num Of Epoch: 0 , Num Of Step: 3072 , Loss: 0.227

Iteration: 16%||| | 1792/11128 [26:43<1:54:45, 1.36it/s]

Num Of Epoch: 0 , Num Of Step: 3584 , Loss: 0.227

Iteration: 18%||| | 2048/11128 [30:08<1:53:48, 1.33it/s]


Num Of Epoch: 0 , Num Of Step: 4096 , Loss: 0.225

Iteration: 21%||| | 2304/11128 [33:33<2:14:35, 1.09it/s]


Num Of Epoch: 0 , Num Of Step: 4608 , Loss: 0.228

Iteration: 23%||| | 2560/11128 [36:57<1:45:50, 1.35it/s]


Num Of Epoch: 0 , Num Of Step: 5120 , Loss: 0.220

Iteration: 25% | 2816/11128 [40:22<1:44:43, 1.32it/s]


Num Of Epoch: 0 , Num Of Step: 5632 , Loss: 0.218

Iteration: 28% | 3072/11128 [43:50<1:58:31, 1.13it/s]


Num Of Epoch: 0 , Num Of Step: 6144 , Loss: 0.218

Iteration: 30% | 3328/11128 [47:14<1:38:12, 1.32it/s]


Num Of Epoch: 0 , Num Of Step: 6656 , Loss: 0.225

Iteration: 32% | 3584/11128 [50:40<1:36:13, 1.31it/s]


Num Of Epoch: 0 , Num Of Step: 7168 , Loss: 0.225

Iteration: 35% | 3840/11128 [54:08<1:48:14, 1.12it/s]


Num Of Epoch: 0 , Num Of Step: 7680 , Loss: 0.219

Iteration: 37% | 4096/11128 [57:33<1:31:45, 1.28it/s]


Num Of Epoch: 0 , Num Of Step: 8192 , Loss: 0.211

Iteration: 39% | 4352/11128 [1:00:59<1:26:29, 1.31it/s]


Num Of Epoch: 0 , Num Of Step: 8704 , Loss: 0.205

Iteration: 41% | 4608/11128 [1:04:26<1:37:40, 1.11it/s]


Num Of Epoch: 0 , Num Of Step: 9216 , Loss: 0.215

Iteration: 44% | 4864/11128 [1:07:52<1:18:58, 1.32it/s]


Num Of Epoch: 0 , Num Of Step: 9728 , Loss: 0.207

Iteration: 46% | 5120/11128 [1:11:19<1:16:36, 1.31it/s]


Num Of Epoch: 0 , Num Of Step: 10240 , Loss: 0.206

Iteration: 48% | 5376/11128 [1:14:46<1:26:05, 1.11it/s]

Num Of Epoch: 0 , Num Of Step: 10752 , Loss: 0.220


Iteration: 51% | 5632/11128 [1:18:12<1:09:28, 1.32it/s]


Num Of Epoch: 0 , Num Of Step: 11264 , Loss: 0.212


Iteration: 53% | 5888/11128 [1:21:39<1:06:59, 1.30it/s]


Num Of Epoch: 0 , Num Of Step: 11776 , Loss: 0.209
Iteration: 55%|██████ | 6144/11128 [1:25:06<1:14:48, 1.11it/s]
Num Of Epoch: 0 , Num Of Step: 12288 , Loss: 0.208
Iteration: 58%|██████ | 6400/11128 [1:28:33<59:49, 1.32it/s]
Num Of Epoch: 0 , Num Of Step: 12800 , Loss: 0.205
Iteration: 60%|██████ | 6656/11128 [1:32:00<56:57, 1.31it/s]
Num Of Epoch: 0 , Num Of Step: 13312 , Loss: 0.204
Iteration: 62%|██████ | 6912/11128 [1:35:27<1:03:14, 1.11it/s]
Num Of Epoch: 0 , Num Of Step: 13824 , Loss: 0.203
Iteration: 64%|██████ | 7168/11128 [1:38:54<50:06, 1.32it/s]
Num Of Epoch: 0 , Num Of Step: 14336 , Loss: 0.207
Iteration: 67%|██████ | 7424/11128 [1:42:21<47:18, 1.30it/s]
Num Of Epoch: 0 , Num Of Step: 14848 , Loss: 0.203
Iteration: 69%|██████ | 7680/11128 [1:45:48<51:37, 1.11it/s]
Num Of Epoch: 0 , Num Of Step: 15360 , Loss: 0.201
Iteration: 71%|██████ | 7936/11128 [1:49:15<40:18, 1.32it/s]
Num Of Epoch: 0 , Num Of Step: 15872 , Loss: 0.211
Iteration: 74%|██████ | 8192/11128 [1:52:42<37:35, 1.30it/s]
Num Of Epoch: 0 , Num Of Step: 16384 , Loss: 0.202
Iteration: 76%|██████ | 8448/11128 [1:56:09<40:10, 1.11it/s]
Num Of Epoch: 0 , Num Of Step: 16896 , Loss: 0.196
Iteration: 78%|██████ | 8704/11128 [1:59:36<30:39, 1.32it/s]
Num Of Epoch: 0 , Num Of Step: 17408 , Loss: 0.206
Iteration: 81%|██████ | 8960/11128 [2:03:03<27:41, 1.30it/s]
Num Of Epoch: 0 , Num Of Step: 17920 , Loss: 0.204
Iteration: 83%|██████ | 9216/11128 [2:06:31<29:36, 1.08it/s]


Num Of Epoch: 0 , Num Of Step: 18432 , Loss: 0.209
Iteration: 85%|██████████ | 9472/11128 [2:10:00<20:53, 1.32it/s]
Num Of Epoch: 0 , Num Of Step: 18944 , Loss: 0.198
Iteration: 87%|██████████ | 9728/11128 [2:13:27<17:51, 1.31it/s]
Num Of Epoch: 0 , Num Of Step: 19456 , Loss: 0.206
Iteration: 90%|██████████ | 9984/11128 [2:16:54<17:05, 1.12it/s]
Num Of Epoch: 0 , Num Of Step: 19968 , Loss: 0.194
Iteration: 92%|██████████ | 10240/11128 [2:20:20<11:12, 1.32it/s]
Num Of Epoch: 0 , Num Of Step: 20480 , Loss: 0.201
Iteration: 94%|██████████ | 10496/11128 [2:23:46<08:02, 1.31it/s]
Num Of Epoch: 0 , Num Of Step: 20992 , Loss: 0.195
Iteration: 97%|██████████ | 10752/11128 [2:27:13<05:37, 1.11it/s]
Num Of Epoch: 0 , Num Of Step: 21504 , Loss: 0.199
Iteration: 99%|██████████ | 11008/11128 [2:30:39<01:30, 1.32it/s]
Num Of Epoch: 0 , Num Of Step: 22016 , Loss: 0.184
Iteration: 100%|██████████ | 11128/11128 [2:32:15<00:00, 1.22it/s]
Iteration: 2%|| | 256/11128 [03:26<2:16:56, 1.32it/s]
Num Of Epoch: 1 , Num Of Step: 512 , Loss: 0.179
Iteration: 5%|| | 512/11128 [06:52<2:14:51, 1.31it/s]
Num Of Epoch: 1 , Num Of Step: 1024 , Loss: 0.166
Iteration: 7%|| | 768/11128 [10:19<2:35:15, 1.11it/s]
Num Of Epoch: 1 , Num Of Step: 1536 , Loss: 0.186
Iteration: 9%|| | 1024/11128 [13:45<2:07:12, 1.32it/s]
Num Of Epoch: 1 , Num Of Step: 2048 , Loss: 0.176
Iteration: 12%|| | 1280/11128 [17:11<2:05:24, 1.31it/s]
Num Of Epoch: 1 , Num Of Step: 2560 , Loss: 0.174


Iteration: 14% | 1536/11128 [20:38<2:23:58, 1.11it/s]
Num Of Epoch: 1 , Num Of Step: 3072 , Loss: 0.176


Iteration: 16% | 1792/11128 [24:04<1:57:55, 1.32it/s]
Num Of Epoch: 1 , Num Of Step: 3584 , Loss: 0.178


Iteration: 18% | 2048/11128 [27:31<1:56:03, 1.30it/s]
Num Of Epoch: 1 , Num Of Step: 4096 , Loss: 0.177


Iteration: 21% | 2304/11128 [30:58<2:12:05, 1.11it/s]
Num Of Epoch: 1 , Num Of Step: 4608 , Loss: 0.177


Iteration: 23% | 2560/11128 [34:25<1:48:12, 1.32it/s]
Num Of Epoch: 1 , Num Of Step: 5120 , Loss: 0.177


Iteration: 25% | 2816/11128 [37:52<1:46:16, 1.30it/s]
Num Of Epoch: 1 , Num Of Step: 5632 , Loss: 0.179


Iteration: 28% | 3072/11128 [41:19<2:00:44, 1.11it/s]
Num Of Epoch: 1 , Num Of Step: 6144 , Loss: 0.180


Iteration: 30% | 3328/11128 [44:46<1:38:53, 1.31it/s]
Num Of Epoch: 1 , Num Of Step: 6656 , Loss: 0.173

Iteration: 32% | 3584/11128 [48:13<1:36:22, 1.30it/s]
Num Of Epoch: 1 , Num Of Step: 7168 , Loss: 0.177

Iteration: 35% | 3840/11128 [51:40<1:49:16, 1.11it/s]
Num Of Epoch: 1 , Num Of Step: 7680 , Loss: 0.170

Iteration: 37% | 4096/11128 [55:07<1:28:54, 1.32it/s]
Num Of Epoch: 1 , Num Of Step: 8192 , Loss: 0.169

Iteration: 39% | 4352/11128 [58:34<1:26:39, 1.30it/s]
Num Of Epoch: 1 , Num Of Step: 8704 , Loss: 0.172

Iteration: 41% | 4608/11128 [1:02:01<1:37:57, 1.11it/s]
Num Of Epoch: 1 , Num Of Step: 9216 , Loss: 0.174

Iteration: 44%|██████ | 4864/11128 [1:05:28<1:19:22, 1.32it/s]
Num Of Epoch: 1 , Num Of Step: 9728 , Loss: 0.170

Iteration: 46%|██████ | 5120/11128 [1:08:55<1:16:42, 1.31it/s]
Num Of Epoch: 1 , Num Of Step: 10240 , Loss: 0.163

Iteration: 48%|██████ | 5376/11128 [1:12:23<1:26:11, 1.11it/s]
Num Of Epoch: 1 , Num Of Step: 10752 , Loss: 0.175

Iteration: 51%|██████ | 5632/11128 [1:15:49<1:09:45, 1.31it/s]
Num Of Epoch: 1 , Num Of Step: 11264 , Loss: 0.173

Iteration: 53%|██████ | 5888/11128 [1:19:16<1:06:51, 1.31it/s]
Num Of Epoch: 1 , Num Of Step: 11776 , Loss: 0.172

Iteration: 55%|██████ | 6144/11128 [1:22:44<1:14:32, 1.11it/s]
Num Of Epoch: 1 , Num Of Step: 12288 , Loss: 0.180

Iteration: 58%|██████ | 6400/11128 [1:26:10<59:42, 1.32it/s]
Num Of Epoch: 1 , Num Of Step: 12800 , Loss: 0.176

Iteration: 60%|██████ | 6656/11128 [1:29:36<56:58, 1.31it/s]
Num Of Epoch: 1 , Num Of Step: 13312 , Loss: 0.162

Iteration: 62%|██████ | 6912/11128 [1:33:03<1:03:01, 1.11it/s]
Num Of Epoch: 1 , Num Of Step: 13824 , Loss: 0.173

Iteration: 64%|██████ | 7168/11128 [1:36:29<49:55, 1.32it/s]
Num Of Epoch: 1 , Num Of Step: 14336 , Loss: 0.169

Iteration: 67%|██████ | 7424/11128 [1:39:56<47:08, 1.31it/s]
Num Of Epoch: 1 , Num Of Step: 14848 , Loss: 0.171

Iteration: 69%|██████ | 7680/11128 [1:43:22<51:33, 1.11it/s]
Num Of Epoch: 1 , Num Of Step: 15360 , Loss: 0.173

Iteration: 71%|██████ | 7936/11128 [1:46:48<40:14, 1.32it/s]
Num Of Epoch: 1 , Num Of Step: 15872 , Loss: 0.161

Iteration: 74%|██████████ | 8192/11128 [1:50:15<37:26, 1.31it/s]
Num Of Epoch: 1 , Num Of Step: 16384 , Loss: 0.165
Iteration: 76%|██████████ | 8448/11128 [1:53:42<40:07, 1.11it/s]
Num Of Epoch: 1 , Num Of Step: 16896 , Loss: 0.169
Iteration: 78%|██████████ | 8704/11128 [1:57:08<30:35, 1.32it/s]
Num Of Epoch: 1 , Num Of Step: 17408 , Loss: 0.171
Iteration: 81%|██████████ | 8960/11128 [2:00:34<27:35, 1.31it/s]
Num Of Epoch: 1 , Num Of Step: 17920 , Loss: 0.173
Iteration: 83%|██████████ | 9216/11128 [2:04:01<28:38, 1.11it/s]
Num Of Epoch: 1 , Num Of Step: 18432 , Loss: 0.164
Iteration: 85%|██████████ | 9472/11128 [2:07:27<20:52, 1.32it/s]
Num Of Epoch: 1 , Num Of Step: 18944 , Loss: 0.163
Iteration: 87%|██████████ | 9728/11128 [2:10:53<17:48, 1.31it/s]
Num Of Epoch: 1 , Num Of Step: 19456 , Loss: 0.167
Iteration: 90%|██████████ | 9984/11128 [2:14:20<17:05, 1.12it/s]
Num Of Epoch: 1 , Num Of Step: 19968 , Loss: 0.164
Iteration: 92%|██████████ | 10240/11128 [2:17:46<11:13, 1.32it/s]
Num Of Epoch: 1 , Num Of Step: 20480 , Loss: 0.173
Iteration: 94%|██████████ | 10496/11128 [2:21:13<08:03, 1.31it/s]
Num Of Epoch: 1 , Num Of Step: 20992 , Loss: 0.167
Iteration: 97%|██████████ | 10752/11128 [2:24:40<05:38, 1.11it/s]
Num Of Epoch: 1 , Num Of Step: 21504 , Loss: 0.166
Iteration: 99%|██████████ | 11008/11128 [2:28:07<01:31, 1.32it/s]
Num Of Epoch: 1 , Num Of Step: 22016 , Loss: 0.167
Iteration: 100%|██████████ | 11128/11128 [2:29:44<00:00, 1.24it/s]

```
runfile('D:/PolyU Course/COMP5423 Natural Language  
Processing/2021Feb/Project/ProjectCode/run_race_test.py', wdir='D:/PolyU  
Course/COMP5423 Natural Language Processing/2021Feb/Project/ProjectCode')
```

Reloaded modules: params, mcqaRACE

level middle

After sorted: ./datasets/RACE/dev/middle/13.txt

level high

After sorted: ./datasets/RACE/dev/high/63.txt

Length of Example: 1481 race_id: middle13.txt:0, context_sentence: What is color?
Why do some of the things around us look red, some green, others blue?

Colors are really made by deflected light. We see color because most of the things
reflect light. In the same way, if something is green, it reflects most of the green
light. If something reflects all light, it is white. If it doesn't reflect any light, it is
black.

Some of the light is reflected and some is taken in and turned into heat .The
darker the color is, the less light is reflected, the more light is taken in. So dark-
colored clothes are warmer in the sun than light-colored clothes., start_ending:
When something reflects light, we can _ ., ending_0: see its color, ending_1: see
its heat, ending_2: not see its color, ending_3: see nothing, label: 0

*** Example ***

race_id: middle13.txt:0

choice: 0

tokens: [CLS] what is color ? why do some of the things around us look red , some
green , others blue ? colors are really made by deflected light . we see color
because most of the things reflect light . in the same way , if something is green , it
reflects most of the green light . if something reflects all light , it is white . if it doesn
' t reflect any light , it is black . some of the light is reflected and some is taken in
and turned into heat . the darker the color is , the less light is reflected , the more
light is taken in . so dark - colored clothes are warmer in the sun than light - colored
clothes . [SEP] when something reflects light , we can _ . see its color [SEP]

input_ids: 101 2054 2003 3609 1029 2339 2079 2070 1997 1996 2477 2105 2149
2298 2417 1010 2070 2665 1010 2500 2630 1029 6087 2024 2428 2081 2011

13366 2571 10985 2422 1012 2057 2156 3609 2138 2087 1997 1996 2477 8339
2422 1012 1999 1996 2168 2126 1010 2065 2242 2003 2665 1010 2009 11138
2087 1997 1996 2665 2422 1012 2065 2242 11138 2035 2422 1010 2009 2003
2317 1012 2065 2009 2987 1005 1056 8339 2151 2422 1010 2009 2003 2304
1012 2070 1997 1996 2422 2003 7686 1998 2070 2003 2579 1999 1998 2357
2046 3684 1012 1996 9904 1996 3609 2003 1010 1996 2625 2422 2003 7686
1010 1996 2062 2422 2003 2579 1999 1012 2061 2601 1011 6910 4253 2024
16676 1999 1996 3103 2084 2422 1011 6910 4253 1012 102 2043 2242 11138
2422 1010 2057 2064 1035 1012 2156 2049 3609 102 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0
0
0
0
0
0
0
0 0

input_mask: 1
1
1
1
0
0
0
0
0
0
0
0
0 0

segment_ids: 0
0
0
0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0
0
0
0
0
0
0
0 0

choice: 1

tokens: [CLS] what is color ? why do some of the things around us look red , some
green , others blue ? colors are really made by def ##le ##cted light . we see color

[illegible]

choice: 2

tokens: [CLS] what is color ? why do some of the things around us look red , some green , others blue ? colors are really made by def ##le ##ctd light . we see color because most of the things reflect light . in the same way , if something is green , it reflects most of the green light . if something reflects all light , it is white . if it doesn ' t reflect any light , it is black . some of the light is reflected and some is taken in and turned into heat . the darker the color is , the less light is reflected , the more light is taken in . so dark - colored clothes are warmer in the sun than light - colored clothes . [SEP] when something reflects light , we can _ . not see its color [SEP]

[illegible][illegible]

[illegible][illegible]

choice: 3

tokens: [CLS] what is color ? why do some of the things around us look red , some green , others blue ? colors are really made by def ##le ##ctd light . we see color because most of the things reflect light . in the same way , if something is green , it reflects most of the green light . if something reflects all light , it is white . if it doesn ' t reflect any light , it is black . some of the light is reflected and some is taken in and turned into heat . the darker the color is , the less light is reflected , the more light is taken in . so dark - colored clothes are warmer in the sun than light - colored clothes . [SEP] when something reflects light , we can . see nothing [SEP]

[illegible]

[illegible][illegible]

label: 0

0

Evaluating: 0%| | 0/494 [00:00<?, ?it/s]

```
***** Running evaluation *****
```

Num examples = 1481

Batch size = 3

```
Evaluating: 100%|██████████| 494/494 [09:36<00:00, 1.17s/it]
```

****Eval Reult****

Evaluate loss 1.110

Accuaracy 54.085%

level middle

After sorted: ./datasets/RACE/test/middle/1.txt

level high

After sorted: ./datasets/RACE/test/high/73.txt

Length of Example: 1454 race_id: middle1.txt:0, context_sentence: Take a class at Dulangkou School, and you'll see lots of things different from other schools, You can see the desks are not in rows and students sit in groups. They put their desks together so they're facing each other. How can they see the blackboard? There are three blackboards on the three walls of the classroom!

The school calls the new way of learning "Tuantuanzuo", meaning sitting in groups. Wei Liying, a Junior 3 teacher, said it was to give students more chances to communicate.

Each group has five or six students, according to Wei, and they play different roles .There is a team leader who takes care of the whole group. There is a "study leader"who makes sure that everyone finishes their homework. And there is a discipline leader who makes sure that nobody chats in class.

Wang Lin is a team leader. The 15-year-old said that having to deal with so many things was tiring.

"I just looked after my own business before,"said Wang. "But now I have to think about my five group members."

But Wang has got used to it and can see the benefits now.

"I used to speak too little. But being a team leader means you have to talk a lot. You could even call me an excellent speaker today."

Zhang Qi, 16, was weak in English. She used to get about 70 in English tests. But in a recent test, Zhang got a grade of more than 80.

"I rarely asked others when I had problems with my English tests. But now I can ask the team leader or study leader. They are really helpful.", start_ending: A discipline leader is supposed to _ ., ending_0: take care of the whole group, ending_1: make sure that everybody finishes homework, ending_2: make sure that nobody chats in class, ending_3: collect all the homework and hand it in to teachers, label: 2

*** Example ***

race_id: middle1.txt:0

choice: 0

tokens: [CLS] take a class at du ##lang ##kou school , and you ' ll see lots of things different from other schools , you can see the desk ##s are not in rows and

students sit in groups . they put their desk ##s together so they ' re facing each other . how can they see the black ##board ? there are three black ##boards on the three walls of the classroom ! the school calls the new way of learning " tu ##ant ##uan ##zu ##o " , meaning sitting in groups . wei li ##ying , a junior 3 teacher , said it was to give students more chances to communicate . each group has five or six students , according to wei , and they play different roles . there is a team leader who takes care of the whole group . there is a " study leader " who makes sure that everyone finishes their homework . and there is a discipline leader who makes sure that nobody chat ##s in class . wang lin is a team leader . the 15 - year - old said that having to deal with so many things was ti ##ring . " i just looked after my own business before , " said wang . " but now i have to think about my five group members . " but wang has got used to it and can see the benefits now . " i used to speak too little . but being a team leader means you have to talk a lot . you could even call me an excellent speaker today . " zhang qi , 16 , was weak in english . she used to get about 70 in english tests . but in a recent test , zhang got a grade of more than 80 . " i rarely asked others when i had problems with my english tests . but now i can ask the team leader or study leader . they are really helpful . "

[SEP] a discipline leader is supposed to _ . take care of the whole group [SEP]

input_ids: 101 2202 1037 2465 2012 4241 25023 24861 2082 1010 1998 2017
1005 2222 2156 7167 1997 2477 2367 2013 2060 2816 1010 2017 2064 2156
1996 4624 2015 2024 2025 1999 10281 1998 2493 4133 1999 2967 1012 2027
2404 2037 4624 2015 2362 2061 2027 1005 2128 5307 2169 2060 1012 2129
2064 2027 2156 1996 2304 6277 1029 2045 2024 2093 2304 15271 2006 1996
2093 3681 1997 1996 9823 999 1996 2082 4455 1996 2047 2126 1997 4083 1000
10722 4630 13860 9759 2080 1000 1010 3574 3564 1999 2967 1012 11417 5622
14147 1010 1037 3502 1017 3836 1010 2056 2009 2001 2000 2507 2493 2062
9592 2000 10639 1012 2169 2177 2038 2274 2030 2416 2493 1010 2429 2000
11417 1010 1998 2027 2377 2367 4395 1012 2045 2003 1037 2136 3003 2040
3138 2729 1997 1996 2878 2177 1012 2045 2003 1037 1000 2817 3003 1000
2040 3084 2469 2008 3071 12321 2037 19453 1012 1998 2045 2003 1037 9009
3003 2040 3084 2469 2008 6343 11834 2015 1999 2465 1012 7418 11409 2003
1037 2136 3003 1012 1996 2321 1011 2095 1011 2214 2056 2008 2383 2000
3066 2007 2061 2116 2477 2001 14841 4892 1012 1000 1045 2074 2246 2044
2026 2219 2449 2077 1010 1000 2056 7418 1012 1000 2021 2085 1045 2031
2000 2228 2055 2026 2274 2177 2372 1012 1000 2021 7418 2038 2288 2109
2000 2009 1998 2064 2156 1996 6666 2085 1012 1000 1045 2109 2000 3713
2205 2210 1012 2021 2108 1037 2136 3003 2965 2017 2031 2000 2831 1037
2843 1012 2017 2071 2130 2655 2033 2019 6581 5882 2651 1012 1000 9327
18816 1010 2385 1010 2001 5410 1999 2394 1012 2016 2109 2000 2131 2055
3963 1999 2394 5852 1012 2021 1999 1037 3522 3231 1010 9327 2288 1037
3694 1997 2062 2084 3770 1012 1000 1045 6524 2356 2500 2043 1045 2018
3471 2007 2026 2394 5852 1012 2021 2085 1045 2064 3198 1996 2136 3003
2030 2817 3003 1012 2027 2024 2428 14044 1012 1000 102 1037 9009 3003

2003 4011 2000 1035 1012 2202 2729 1997 1996 2878 2177 102 0 0 0 0 0 0 0 0
0
0 0

input_mask: 1
1
1
1
1
1
1
1
1
1
0
0
0 0

segment_ids: 0
0
0
0
0
0
0
0
0
0 1
0
0
0 0

choice: 1

tokens: [CLS] take a class at du ##lang ##kou school , and you ' ll see lots of things different from other schools , you can see the desk ##s are not in rows and students sit in groups . they put their desk ##s together so they ' re facing each other . how can they see the black ##board ? there are three black ##boards on the three walls of the classroom ! the school calls the new way of learning " tu ##ant ##uan ##zu ##o " , meaning sitting in groups . wei li ##ying , a junior 3 teacher , said it was to give students more chances to communicate . each group has five or six students , according to wei , and they play different roles . there is a team leader who takes care of the whole group . there is a " study leader " who makes sure that everyone finishes their homework . and there is a discipline leader who makes sure that nobody chat ##s in class . wang lin is a team leader . the 15 - year - old said that having to deal with so many things was ti ##ring . " i just looked after my own business before , " said wang . " but now i have to think about my five group members . " but wang has got used to it and can see the benefits now . " i used to speak too little . but being a team leader means you have to talk a lot . you

[SEP] a discipline leader is supposed to _ . make sure that everybody finishes homework [SEP]

[illegible]


```
2093 3681 1997 1996 9823 999 1996 2082 4455 1996 2047 2126 1997 4083 1000  
10722 4630 13860 9759 2080 1000 1010 3574 3564 1999 2967 1012 11417 5622  
14147 1010 1037 3502 1017 3836 1010 2056 2009 2001 2000 2507 2493 2062  
9592 2000 10639 1012 2169 2177 2038 2274 2030 2416 2493 1010 2429 2000  
11417 1010 1998 2027 2377 2367 4395 1012 2045 2003 1037 2136 3003 2040  
3138 2729 1997 1996 2878 2177 1012 2045 2003 1037 1000 2817 3003 1000  
2040 3084 2469 2008 3071 12321 2037 19453 1012 1998 2045 2003 1037 9009  
3003 2040 3084 2469 2008 6343 11834 2015 1999 2465 1012 7418 11409 2003  
1037 2136 3003 1012 1996 2321 1011 2095 1011 2214 2056 2008 2383 2000  
3066 2007 2061 2116 2477 2001 14841 4892 1012 1000 1045 2074 2246 2044  
2026 2219 2449 2077 1010 1000 2056 7418 1012 1000 2021 2085 1045 2031  
2000 2228 2055 2026 2274 2177 2372 1012 1000 2021 7418 2038 2288 2109  
2000 2009 1998 2064 2156 1996 6666 2085 1012 1000 1045 2109 2000 3713  
2205 2210 1012 2021 2108 1037 2136 3003 2965 2017 2031 2000 2831 1037  
2843 1012 2017 2071 2130 2655 2033 2019 6581 5882 2651 1012 1000 9327  
18816 1010 2385 1010 2001 5410 1999 2394 1012 2016 2109 2000 2131 2055  
3963 1999 2394 5852 1012 2021 1999 1037 3522 3231 1010 9327 2288 1037  
3694 1997 2062 2084 3770 1012 1000 1045 6524 2356 2500 2043 1045 2018  
3471 2007 2026 2394 5852 1012 2021 2085 1045 2064 3198 1996 2136 3003  
2030 2817 3003 1012 2027 2024 2428 14044 1012 1000 102 1037 9009 3003  
2003 4011 2000 1035 1012 2191 2469 2008 6343 11834 2015 1999 2465 102 0 0  
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
```

[illegible][illegible]

2000 2009 1998 2064 2156 1996 6666 2085 1012 1000 1045 2109 2000 3713
2205 2210 1012 2021 2108 1037 2136 3003 2965 2017 2031 2000 2831 1037
2843 1012 2017 2071 2130 2655 2033 2019 6581 5882 2651 1012 1000 9327
18816 1010 2385 1010 2001 5410 1999 2394 1012 2016 2109 2000 2131 2055
3963 1999 2394 5852 1012 2021 1999 1037 3522 3231 1010 9327 2288 1037
3694 1997 2062 2084 3770 1012 1000 1045 6524 2356 2500 2043 1045 2018
3471 2007 2026 2394 5852 1012 2021 2085 1045 2064 3198 1996 2136 3003
2030 2817 3003 1012 2027 2024 2428 14044 1012 1000 102 1037 9009 3003
2003 4011 2000 1035 1012 8145 2035 1996 19453 1998 2192 2009 1999 2000
5089 102 0
0
0 0

input_mask: 1
1
1
1
1
1
1
1
1
1
1
0
0
0 0

segment_ids: 0
0
0
0
0
0
0
0
0
0 1
0
0
0 0

label: 2

0

Testing: 0%| | 0/485 [00:00<?, ?it/s]

***** Running test *****

Num examples = 1454

Batch size = 3

Testing: 100%|██████████| 485/485 [09:31<00:00, 1.18s/it]

****TEST Reult****

Test loss 1.118

Test Accuaracy 54.470%