# #1 Inews 遍历

```
Inews 遍历
bert-base
env-hardware:
env-software:
model:
acc = 0.84534782984012783
log detail:
roberta-wwm
env-hardware:
env-software:
model:
acc = 0.8418418418418419
log detail:
```

#### #2 bert-base

#### #3 env-hardware:

- Quadro RTX 8000
- VRAM: 48GiB
- Platinum 8160 96核128G
- 4TB SSD

#### #3 env-software:

- pytorch
- cuda9
- cudnn7 1.0.1
- python 3.6

### #3 model:

chinese\_wwm\_pytorch(bert-base)

### #3 acc = 0.84534782984012783

### #3 log detail:

```
1 (dl) root@1b20aad9b5b6:/data/baselines/classifier_pytorch# sh
   run_classifier_inews.sh
   Start training text classfication on: with bert-base-chinese
   id:10165102211
 4
    data exists
   Finish download dataset.
    Start running...
 8
   n_gpus: 4
9
   01/12/2020 12:21:53 - WARNING - root - Process rank: -1, device: cuda,
10
    n_gpu: 4,
11
    distributed training: False, 16-bits training: False
    01/12/2020 12:21:53 - INFO - transformers.configuration_utils - loading
    configuration
   file https://s3.amazonaws.com/models.huggingface.co/bert/bert-base-
13
    chinese-config.json
   from cache at
14
```

```
/home/ubuntu/.cache/torch/transformers/8a3b1cfe5da58286e12a0f5d7d182b8d6ec
    a88c08e26c332e
16
    e3817548cf7e60a.0c16faba8be66db3f02805c912e4cf94d3c9cffc1f12fa1a39906f9270
17
   01/12/2020 12:21:53 - INFO - transformers.configuration_utils - Model
    config {
18
    "attention_probs_dropout_prob": 0.1,
    "directionality": "bidi",
19
    "finetuning_task": "inews",
20
    "hidden_act": "gelu",
21
    "hidden_dropout_prob": 0.1,
23
    "hidden_size": 768,
    "initializer_range": 0.02,
24
    "intermediate_size": 3072,
25
    "layer_norm_eps": 1e-12,
26
    "max_position_embeddings": 512,
27
    "num_attention_heads": 12,
28
29
    "num_hidden_layers": 12,
    "num_labels": 2,
30
31
    "output_attentions": false,
    "output_hidden_states": false,
32
    "output_past": true,
33
34
     "pooler_fc_size": 768,
    "pooler_num_attention_heads": 12,
35
36
    "pooler_num_fc_layers": 3,
    "pooler_size_per_head": 128,
37
    "pooler_type": "first_token_transform",
38
    "pruned_heads": {},
39
    "torchscript": false,
41
    "type_vocab_size": 2,
    "use_bfloat16": false,
42
    "vocab_size": 21128
43
44
45
   01/12/2020 12:21:54 - INFO - transformers.tokenization_utils - loading
    https://s3.amazonaws.com/models.huggingface.co/bert/bert-base-chinese-
    vocab.txt from
    cache at
47
    /home/ubuntu/.cache/torch/transformers/8a0c070123c1f794c42a29c6904beb7c1b8
    715741e235bee0
49
    4aca2c7636fc83f.9b42061518a39ca00b8b52059fd2bede8daa613f8a8671500e518a8c29
    de8c00
50
    01/12/2020 12:21:55 - INFO - transformers.modeling_utils - loading weights
51
    https://s3.amazonaws.com/models.huggingface.co/bert/bert-base-chinese-
    pytorch_model.bin
52
    from cache at
    /home/ubuntu/.cache/torch/transformers/b1b5e295889f2d0979ede9a78ad9cb5dc6a
    0e25ab7f9417b3
    15f0a2c22f4683d.929717ca66a3ba9eb9ec2f85973c6398c54c38a4faa464636a491d7a70
    5f7eb6
55
    01/12/2020 12:21:58 - INFO - transformers.modeling_utils - Weights of
    BertForSequenceClassification not initialized from pretrained model:
    ['classifier.weight', 'classifier.bias']
58
    01/12/2020 12:21:58 - INFO - transformers.modeling_utils - Weights from
    pretrained
59
    model not used in BertForSequenceClassification: ['cls.predictions.bias',
    'cls.predictions.transform.dense.weight',
    'cls.predictions.transform.dense.bias',
    'cls.predictions.decoder.weight', 'cls.seq_relationship.weight',
61
    'cls.seq_relationship.bias', 'cls.predictions.transform.LayerNorm.weight',
63
    'cls.predictions.transform.LayerNorm.bias']
    01/12/2020 12:22:03 - INFO - root - Training/evaluation parameters
64
    Namespace(adam_epsilon=1e-08, cache_dir='', config_name='',
65
```

```
data_dir='/home/ubuntu/CLUE/baselines/models_pytorch/classifier_pytorch/CL
    UEdatasets/afq
67
    mc/', device=device(type='cuda'), do_eval=True, do_lower_case=True,
    do_predict=False,
    do_train=True, eval_all_checkpoints=False, fp16=False,
    fp16_opt_level='01',
    gradient_accumulation_steps=1, learning_rate=2e-05, local_rank=-1,
    logging_steps=2146,
 70
    max_grad_norm=1.0, max_seq_length=128, max_steps=-1,
    model_name_or_path='bert-base-chinese', model_type='bert', n_gpu=4,
    no_cuda=False,
72
    num_train_epochs=3.0,
    output_dir='/home/ubuntu/CLUE/baselines/models_pytorch/classifier_pytorch/
73
    inews_output/b
74
    ert', output_mode='classification', overwrite_cache=False,
    overwrite_output_dir=True,
75
    per_gpu_eval_batch_size=16, per_gpu_train_batch_size=16,
    predict_checkpoints=0,
    save_steps=2146, seed=42, server_ip='', server_port='', task_name='inews',
76
77
    tokenizer_name='', warmup_proportion=0.1, weight_decay=0.01)
    01/12/2020 12:22:03 - INFO - root - Loading features from cached file
    /home/ubuntu/CLUE/baselines/models_pytorch/classifier_pytorch/CLUEdatasets
    /inews/cached_
80
    train_bert-base-chinese_128_inews
81 01/12/2020 12:22:04 - INFO - root - ***** Running training *****
    01/12/2020 12:22:04 - INFO - root - Num examples = 34334
    01/12/2020 12:22:04 - INFO - root - Num Epochs = 3
    01/12/2020 12:22:04 - INFO - root - Instantaneous batch size per GPU = 16
    01/12/2020 12:22:04 - INFO - root - Total train batch size (w. parallel,
    distributed
    & accumulation) = 64
86
    01/12/2020 12:22:04 - INFO - root - Gradient Accumulation steps = 1
    01/12/2020 12:22:04 - INFO - root - Total optimization steps = 1611
    /home/ubuntu/anaconda3/lib/python3.7/site-
    packages/torch/nn/parallel/_functions.py:61:
    UserWarning: Was asked to gather along dimension 0, but all input tensors
    were scalars;
    will instead unsqueeze and return a vector.
91
    warnings.warn('Was asked to gather along dimension 0, but all '
    0.4132
    94
95
    0.4571
96
    01/12/2020 12:28:37 - INFO - root - global_step = 1611, average loss =
    0.47748216146789374
    01/12/2020 12:28:37 - INFO - root - Saving model checkpoint to
98
99
    /home/ubuntu/CLUE/baselines/models_pytorch/classifier_pytorch/inews_output
    01/12/2020 12:28:37 - INFO - transformers.configuration_utils -
100
    Configuration saved in
101
    /home/ubuntu/CLUE/baselines/models_pytorch/classifier_pytorch/inews_output
    /bert/config.j
102
103
    01/12/2020 12:28:38 - INFO - transformers.modeling_utils - Model weights
    /home/ubuntu/CLUE/baselines/models_pytorch/classifier_pytorch/inews_output
104
    /bert/pytorch_
105
    model.bin
106
    01/12/2020 12:28:38 - INFO - transformers.configuration_utils - loading
    configuration
107
    /home/ubuntu/CLUE/baselines/models_pytorch/classifier_pytorch/inews_output
108
    /bert/config.j
```

```
109 son
110
     01/12/2020 12:28:38 - INFO - transformers.configuration_utils - Model
     "attention_probs_dropout_prob": 0.1,
111
     "directionality": "bidi",
112
     "finetuning_task": "inews",
113
     "hidden_act": "gelu",
114
115
     "hidden_dropout_prob": 0.1,
     "hidden_size": 768,
116
     "initializer_range": 0.02,
117
     "intermediate_size": 3072,
     "layer_norm_eps": 1e-12,
119
     "max_position_embeddings": 512,
120
     "num_attention_heads": 12,
121
122
     "num_hidden_layers": 12,
123
     "num_labels": 2,
     "output_attentions": false,
124
125
     "output_hidden_states": false,
126
     "output_past": true,
127
     "pooler_fc_size": 768,
     "pooler_num_attention_heads": 12,
128
129
     "pooler_num_fc_layers": 3,
130
     "pooler_size_per_head": 128,
     "pooler_type": "first_token_transform",
131
     "pruned_heads": {},
132
     "torchscript": false,
133
     "type_vocab_size": 2,
134
     "use_bfloat16": false,
135
     "vocab_size": 21128
136
137
138
     01/12/2020 12:28:38 - INFO - transformers.modeling_utils - loading weights
     /home/ubuntu/CLUE/baselines/models_pytorch/classifier_pytorch/inews_output
     /bert/pytorch_
140
     model.bin
141
     01/12/2020 12:28:41 - INFO - transformers.tokenization_utils - Model name
     '/home/ubuntu/CLUE/baselines/models_pytorch/classifier_pytorch/inews_outpu
     t/bert' not
143
    found in model shortcut name list (bert-base-uncased, bert-large-uncased,
144 bert-base-cased, bert-large-cased, bert-base-multilingual-uncased,
145
    bert-base-multilingual-cased, bert-base-chinese, bert-base-german-cased,
     bert-large-uncased-whole-word-masking, bert-large-cased-whole-word-
146
     masking,
147
     bert-large-uncased-whole-word-masking-finetuned-squad,
     bert-large-cased-whole-word-masking-finetuned-squad, bert-base-cased-
148
     finetuned-mrpc,
     bert-base-german-dbmdz-cased, bert-base-german-dbmdz-uncased). Assuming
149
     '/home/ubuntu/CLUE/baselines/models_pytorch/classifier_pytorch/inews_outpu
150
     t/bert' is a
151
     path or url to a directory containing tokenizer files.
     01/12/2020 12:28:41 - INFO - transformers.tokenization_utils - loading
153
     /home/ubuntu/CLUE/baselines/models_pytorch/classifier_pytorch/inews_output
     /bert/vocab.tx
154
     01/12/2020 12:28:41 - INFO - transformers.tokenization_utils - loading
155
     /home/ubuntu/CLUE/baselines/models_pytorch/classifier_pytorch/inews_output
156
     /bert/added_to
     kens.json
157
     01/12/2020 12:28:41 - INFO - transformers.tokenization_utils - loading
159
     /home/ubuntu/CLUE/baselines/models_pytorch/classifier_pytorch/inews_output
     /bert/special_
160
     tokens_map.json
```

```
01/12/2020 12:28:41 - INFO - transformers.tokenization_utils - loading
     /home/ubuntu/CLUE/baselines/models_pytorch/classifier_pytorch/inews_output
162
     /bert/tokenize
163
     r_config.json
     01/12/2020 12:28:41 - INFO - transformers.tokenization_utils - Model name
164
165
     '/home/ubuntu/CLUE/baselines/models_pytorch/classifier_pytorch/inews_outpu
     t/bert' not
     found in model shortcut name list (bert-base-uncased, bert-large-uncased,
166
     bert-base-cased, bert-large-cased, bert-base-multilingual-uncased,
167
     bert-base-multilingual-cased, bert-base-chinese, bert-base-german-cased,
169
     bert-large-uncased-whole-word-masking, bert-large-cased-whole-word-
     masking,
170
     bert-large-uncased-whole-word-masking-finetuned-squad,
171
     bert-large-cased-whole-word-masking-finetuned-squad, bert-base-cased-
     finetuned-mrpc,
172
     bert-base-german-dbmdz-cased, bert-base-german-dbmdz-uncased). Assuming
     '/home/ubuntu/CLUE/baselines/models_pytorch/classifier_pytorch/inews_outpu
     t/bert' is a
174
     path or url to a directory containing tokenizer files.
     01/12/2020 12:28:41 - INFO - transformers.tokenization_utils - loading
175
     /home/ubuntu/CLUE/baselines/models_pytorch/classifier_pytorch/inews_output
176
     /bert/vocab.tx
177
178
     01/12/2020 12:28:41 - INFO - transformers.tokenization_utils - loading
179
     /home/ubuntu/CLUE/baselines/models_pytorch/classifier_pytorch/inews_output
     /bert/added_to
180
     kens.json
181
     01/12/2020 12:28:41 - INFO - transformers.tokenization_utils - loading
     /home/ubuntu/CLUE/baselines/models_pytorch/classifier_pytorch/inews_output
182
     /bert/special_
183
     tokens_map.json
184
     01/12/2020 12:28:41 - INFO - transformers.tokenization_utils - loading
185
     /home/ubuntu/CLUE/baselines/models_pytorch/classifier_pytorch/inews_output
     /bert/tokenize
186
    r_config.json
187
     01/12/2020 12:28:41 - INFO - root - Evaluate the following checkpoints:
     ['/home/ubuntu/CLUE/baselines/models_pytorch/classifier_pytorch/inews_outp
188
189
     01/12/2020 12:28:41 - INFO - transformers.configuration_utils - loading
     configuration
190
     file
     /home/ubuntu/CLUE/baselines/models_pytorch/classifier_pytorch/inews_output
191
     /bert/config.j
192
     son
193
     01/12/2020 12:28:41 - INFO - transformers.configuration_utils - Model
     confiq {
194
     "attention_probs_dropout_prob": 0.1,
     "directionality": "bidi",
195
     "finetuning_task": "inews",
197
     "hidden_act": "gelu",
     "hidden_dropout_prob": 0.1,
198
     "hidden_size": 768,
199
     "initializer_range": 0.02,
200
201
     "intermediate_size": 3072,
     "layer_norm_eps": 1e-12,
202
     "max_position_embeddings": 512,
203
204
     "num_attention_heads": 12,
205
     "num_hidden_layers": 12,
206
     "num_labels": 2,
207
     "output_attentions": false,
```

```
"output_hidden_states": false,
208
209
     "output_past": true,
     "pooler_fc_size": 768,
210
     "pooler_num_attention_heads": 12,
211
     "pooler_num_fc_layers": 3,
212
213
     "pooler_size_per_head": 128,
     "pooler_type": "first_token_transform",
214
     "pruned_heads": {},
215
     "torchscript": false,
216
     "type_vocab_size": 2,
217
    "use_bfloat16": false,
218
     "vocab_size": 21128
219
220
221
     01/12/2020 12:28:41 - INFO - transformers.modeling_utils - loading weights
222
     /home/ubuntu/CLUE/baselines/models_pytorch/classifier_pytorch/inews_output
     /bert/pytorch_
223
    model.bin
224
    01/12/2020 12:28:43 - INFO - root - Loading features from cached file
225
    /home/ubuntu/CLUE/baselines/models_pytorch/classifier_pytorch/CLUEdatasets
     /inews/cached_
226
    dev_bert-base-chinese_128_inews
    01/12/2020 12:28:43 - INFO - root - ******** Running evaluation ********
227
228
    [Evaluating] 68/68 [=========== ] 89.1ms/step
01/12/2020 \ 12:28:49 - INFO - root - Num examples = 4316
230 01/12/2020 12:28:49 - INFO - root - Batch size = 64
231 01/12/2020 12:28:49 - INFO - root - ******* Eval results *******
232 01/12/2020 12:28:49 - INFO - root - dev: acc = 0.84534782984012783
```

#### #2 roberta-wwm

### #3 env-hardware:

- Quadro RTX 8000
- VRAM: 48GiB
- Platinum 8160 96核128G
- 4TB SSD

### #3 env-software:

- pytorch
- cuda9
- cudnn7 1.0.1
- python 3.6

#### #3 model:

chinese\_roberta\_wwm\_ext\_pytorch

## #3 acc = 0.8418418418419

# #3 log detail:

```
1 (dl) root@1b20aad9b5b6:/data/baselines/classifier_pytorch# sh
    run_classifier_inews.sh
2 /usr/local/miniconda3/envs/dl/lib/python3.6/site-
    packages/sklearn/externals/joblib/externals/cloudpickle/cloudpickle.py:47:
    DeprecationWarning: the imp module is deprecated in favour of importlib;
    see the module's documentation for alternative uses
```

```
import imp
 4
   01/14/2020 14:53:47 - WARNING - root - Process rank: -1, device: cuda,
    n_gpu: 1, distributed training: False, 16-bits training: False
    01/14/2020 14:53:47 - INFO - transformers.configuration_utils -
    configuration file
    /data/baselines/classifier_pytorch/prev_trained_model/chinese_roberta_wwm_
    ext_pytorch/config.json
    01/14/2020 14:53:47 - INFO - transformers.configuration_utils - Model
    confiq {
 7
      "attention_probs_dropout_prob": 0.1,
      "directionality": "bidi",
 8
 9
      "finetuning_task": "inews",
      "hidden_act": "gelu",
10
      "hidden_dropout_prob": 0.1,
11
12
      "hidden_size": 768,
13
      "initializer_range": 0.02,
14
      "intermediate_size": 3072,
15
      "layer_norm_eps": 1e-12,
      "max_position_embeddings": 512,
16
17
      "num_attention_heads": 12,
18
      "num_hidden_layers": 12,
19
      "num_labels": 3,
20
      "output_attentions": false,
21
      "output_hidden_states": false,
22
      "output_past": true,
      "pooler_fc_size": 768,
23
24
      "pooler_num_attention_heads": 12,
      "pooler_num_fc_layers": 3,
25
      "pooler_size_per_head": 128,
26
27
      "pooler_type": "first_token_transform",
28
      "pruned_heads": {},
29
      "torchscript": false,
      "type_vocab_size": 2,
30
31
      "use_bfloat16": false,
32
      "vocab_size": 21128
    }
33
34
35
    01/14/2020 14:53:47 - INFO - transformers.tokenization_utils - Model
    '/data/baselines/classifier_pytorch/prev_trained_model/chinese_roberta_wwm
    _ext_pytorch' not found in model shortcut name list(bert-base-uncased,
    bert-large-uncased, bert-base-cased, bert-large-cased, bert-base-
    multilingual-uncased, bert-base-multilingual-cased, bert-base-chinese,
    bert-base-german-cased, bert-large-uncased-whole-word-masking, bert-large-
    cased-whole-word-masking, bert-large-uncased-whole-word-masking-finetuned-
    squad, bert-large-cased-whole-word-masking-finetuned-squad, bert-base-
    cased-finetuned-mrpc, bert-base-german-dbmdz-cased, bert-base-german-
    dbmdz-uncased). Assuming
    '/data/baselines/classifier_pytorch/prev_trained_model/chinese_roberta_wwm
    _ext_pytorch' is a path or url to a directory containing tokenizer files.
    01/14/2020 14:53:47 - INFO - transformers.tokenization_utils - Didn't
    find file
    /data/baselines/classifier_pytorch/prev_trained_model/chinese_roberta_wwm_
    ext_pytorch/added_tokens.json. We won't load it.
37
    01/14/2020 14:53:48 - INFO - transformers.tokenization_utils -
    find file
    /data/baselines/classifier_pytorch/prev_trained_model/chinese_roberta_wwm_
    ext_pytorch/special_tokens_map.json. We won'tload it.
38
    01/14/2020 14:53:48 - INFO - transformers.tokenization_utils -
    find file
    /data/baselines/classifier_pytorch/prev_trained_model/chinese_roberta_wwm_
    ext_pytorch/tokenizer_config.json. We won't load it.
```

```
01/14/2020 14:53:48 - INFO - transformers.tokenization_utils - loading
    /data/baselines/classifier_pytorch/prev_trained_model/chinese_roberta_wwm_
    ext_pytorch/vocab.txt
    01/14/2020 14:53:48 - INFO - transformers.tokenization_utils -
                                                                      loading
    01/14/2020 14:53:48 - INFO - transformers.tokenization_utils -
                                                                      loading
    file None
    01/14/2020 14:53:48 - INFO - transformers.tokenization_utils -
                                                                      loading
    01/14/2020 14:53:48 - INFO - transformers.modeling_utils - loading
43
    weights file
    /data/baselines/classifier_pytorch/prev_trained_model/chinese_roberta_wwm_
    ext_pytorch/pytorch_model.bin
    01/14/2020 14:54:12 - INFO - transformers.modeling_utils - Weights of
    BertForSequenceClassification not initialized from pretrained model:
    ['classifier.weight', 'classifier.bias']
    01/14/2020 14:54:12 - INFO - transformers.modeling_utils -
    pretrained model not used in BertForSequenceClassification:
    ['cls.predictions.bias', 'cls.predictions.transform.dense.weight',
    'cls.predictions.transform.dense.bias',
    'cls.predictions.transform.LayerNorm.weight',
    'cls.predictions.transform.LayerNorm.bias',
    'cls.predictions.decoder.weight', 'cls.seq_relationship.weight',
    'cls.seq_relationship.bias']
    01/14/2020 14:54:14 - INFO - root - Training/evaluation parameters
    Namespace(adam_epsilon=1e-08, cache_dir='', config_name='',
    data_dir='/data/baselines/classifier_pytorch/chineseGLUEdatasets/inews/',
    device=device(type='cuda'), do_eval=True, do_lower_case=True,
    do_predict=False, do_train=True, eval_all_checkpoints=False, fp16=False,
    fp16_opt_level='01', gradient_accumulation_steps=1, learning_rate=2e-05,
    local_rank=-1, logging_steps=670, max_grad_norm=1.0, max_seq_length=512,
    max_steps=-1,
    model_name_or_path='/data/baselines/classifier_pytorch/prev_trained_model/
    chinese_roberta_wwm_ext_pytorch',model_type='roberta', n_gpu=1,
    no_cuda=False, num_train_epochs=4.0,
    output_dir='/data/baselines/classifier_pytorch/outputs/inews_output/robert
    a', output_mode='classification', overwrite_cache=False,
    overwrite_output_dir=True, per_gpu_eval_batch_size=8,
    per_qpu_train_batch_size=8, save_steps=670, seed=42, server_ip='',
    server_port='', task_name='inews', tokenizer_name='',
    warmup_proportion=0.1, weight_decay=0.01)
    01/14/2020 14:54:15 - INFO - root - Loading features from cached file
    /data/baselines/classifier_pytorch/chineseGLUEdatasets/inews/cached_train_
    chinese_roberta_wwm_ext_pytorch_512_inews
    01/14/2020 14:54:18 - INFO - root - ***** Running training *****
48
    01/14/2020 14:54:18 - INFO - root - Num examples = 5355 01/14/2020 14:54:18 - INFO - root - Num Epochs = 4 01/14/2020 14:54:18 - INFO - root - Instantaneous batch
50
                                           Instantaneous batch size per GPU =
51
    01/14/2020 14:54:18 - INFO - root - Total train batch size (w.
52
    parallel, distributed & accumulation) = 8
    01/14/2020 14:54:18 - INFO - root - Gradient Accumulation steps = 1
53
    01/14/2020 14:54:18 - INFO - root -
                                           Total optimization steps = 2680
    [Training] 669/670 [==========>.] - ETA: Os loss:
    0.6268 01/14/2020 14:58:09 - INFO - root - Creating features from
    dataset file at
    /data/baselines/classifier_pytorch/chineseGLUEdatasets/inews/
56
   01/14/2020 14:58:10 - INFO - processors.glue - Writing example 0
    01/14/2020 14:58:10 - INFO - processors.glue - *** Example ***
57
    01/14/2020 14:58:10 - INFO - processors.glue - guid: dev-1
```

```
59 01/14/2020 14:58:10 - INFO - processors.glue - input_ids: 101 2526 2360
 872 1962 8024 6435 7309 2769 1762 4397 4263 5381 677 4692 1168 671 702
 2552 811 4638 1996 2023 8024 1377 3221 5381 4991 119 119119 142 2526 2360
 8728 102 4761 6399 772 3326 3221 1920 2157 6963 7478 2382 1068 3800 4638
 6413 75798024 3684 4994 2769 812 6716 1905 754 671 702 6158 2255 2181 1259
 1741 4638 4852 833 8024 6929 720 1762 4761 6399 772 3326 6401 6390 704
 8024 6963 3300 1525 763 1372 966 2533 2769 812 3800 2692 1450 8024 4761
 6399 772 3326 6401 6390 3198 3126 3221 1914 719 8024 8024 4761 6399 772
 3326 4509 6435 1453 3309 3221 1914 719 8024 809 1350 4761 6399 772 3326
 6401 6390 2418 6401 3332 3160 3300 1525 763 1450 102 0 0 0 0 0 0 0 0 0 0 0
 01/14/2020 14:58:10 - INFO - processors.glue - attention_mask: 1 1 1 1 1
 01/14/2020 14:58:10 - INFO - processors.glue - token_type_ids: 0 0 0 0
 62
 01/14/2020 14:58:10 - INFO - processors.glue -
               label: 1 \text{ (id = 1)}
63
 01/14/2020 14:58:10 - INFO - processors.glue - input length: 136
 01/14/2020 14:58:10 - INFO - processors.glue - *** Example ***
64
 01/14/2020 14:58:10 - INFO - processors.glue - guid: dev-2
```

```
66 01/14/2020 14:58:10 - INFO - processors.glue - input_ids: 101 100 6375
  5401 1744 782 1373 2769 8124268 4268 100 8024 704 1744 4522 2110 4495 7770
  1548 1226 4255 3625 6404 8024 3637 6228 5401 1744 782 8024 2471 6759 4197
  1920 3797 102 2537 889 3760 3300 782 898 7479 6347 5393 800 782 889 6560
  2533 2203 7028 8024 738 679 4500 7478 2533 6366 1071 800 1751 2157 935
  7674 4935 5628 8024 2798 5543 6303 3209 671 943 1751 2157 4638 2485 1920
  511 3297 6818 3300 671 7674 704 1751 4522 2119 4495 9796 8565 5836 6167
  5439 2374 4638 3625 3289 872 947 6963 5481 6882 749 1621 8043 3625 3289
  1399 1373 517 704 1751 4268 4268 518 511 3625 3289 6917 3760 3633 2466
  7274 1993 8024 4028 1548 5442 4684 6241 679 6325 1765 2990 1168 8038 6857
  7674 3625 8024 679 1006 6843 5183 12637 8171 10223 8736 4638 7376 1062
  1062 1469 152 8329 8680 9341 738 6843 5183 2792 3300 2552 7027 1403 1168
  5401 1751 4638 4318 3346 6205 1506 7274 7531 2218 1210 2869 2480 2484 8024
  2527 7481 4638 3625 6270 738 3221 3291 1217 1233 4255 119 119100 2300 3817
  2055 1912 3149 1073 2563 4862 6546 1751 3724 3532 6303 4638 3221 872 947
  1292 6717 2235 5607 6210 1164 2563 5412 7953 7531 1506 5587 6303 4638 3221
  872 947 856 5476 678 3706 2864 7679 3977 7519 1762 872 5401 1751 4269 7481
  1184 872 6134 4412 722 1962 6638 4142 7353 1248 1060 7481 676 1143 2894
  1462 2682 6084 1057 1348 6084 1057 679 1168 119 119 119 119 119 3625 6270
  3221 2769 4638 3636 1690 2897 679 6624 2769 4638 3636 1690 872 1372 5543
  4495 3706 5445 2769 877 3302 872 4638 4028 2825 1331 7543 4192 2615 872
  872 2396 5401 1751 782 5476 2484 3633 5412 872 6960 816 704 3122 3080 2769
  2974 6270 722 1713 1341 6295 3706 722 2066 3468 2807 5183 2769 2066 4542
  4534 2769 3221 7919 7797 5436 679 6629 4032 4034 6887 3624 3221 5178 2229
  5439 2094 2595 3419 2523 1361 6847 2682 3791 3300 6210 1765 1430 1366 3706
  2999 4959 872 7990 1051 2094 4031 1960 4638 7481 1072 1391 4478 4638 5408
  6120 69635023 1168 6629 4692 2783 119 119 119 119 119 119 100 3418 3087
  5206 5181 3173 5472 8024 4028 1548 5442 4638 1399 2099 1373 4374 823 1127
  8024 3221 671 1399 5401 1751 4522 2119 4495 511 6857 7674 3625 8024 5476
  6245 4638 3221 1060 943 3299 1184 671 6629 5836 6167 5439 2374 5645 704
  1751 2119 4495 4638 5144 5160 511 671 1147 8024 6917 2533 2537 6929 1842
  1039 2156 3241 3298 6303 6629 119 119 8271 2399 123 3299 8133 3189 8024
  5401 1751 6540 2336 2336 4989 1920 2119 12637 8171 10223 8736 4638 704
  1751 2119 4495 5474 1394 8935 8139 8020 704 1751 2119 5474 8021 5647 6794
  749 671 1842 1039 2156 3241 3298 511 102
  01/14/2020 14:58:10 - INFO - processors.glue -
                              attention_mask: 1 1 1 1 1
  01/14/2020 14:58:10 - INFO - processors.glue - token_type_ids: 0 0 0 0 0
```

```
69 01/14/2020 14:58:10 - INFO - processors.glue -
                          label: 2 (id = 2)
70
 01/14/2020 14:58:10 - INFO - processors.glue -
                          input length: 512
71
 01/14/2020 14:58:10 - INFO - processors.glue - *** Example ***
 01/14/2020 14:58:10 - INFO - processors.glue -
72
                          guid: dev-3
 01/14/2020 14:58:10 - INFO - processors.glue - input_ids: 101 1921 3698
  1359 1265 3193 4761 6887 142 7433 7434 102 1333 3403 7579 8038 1921 3698
  1359 1265 3193 4761 6887 139 156 7252 1790 1344 3698 6496 2229 1169 868
  131 3330 881 4448 5041 1355 8038 865 3817 9160 2399 122 3299 8143 3189
  8110 3198 139 156 7433 7434 7360 3946 3867 2622 139 156 1358 1298 3118
  3553 1469 691 6662 1107 4958 3698 1066 1398 2512 1510 8024 122 3299 8162
  3189 1915 7313 1168 8114 3189 4635 1921 2769 1344 2199 3300 671 3613 7433
  7434 7360 3946 1921 3698 6814 4923 511 7564 6369 1059 1344 7433 1931 7434
  6760 704 7434 8024 704 7770 2255 1765 1277 3300 1920 7434 8024 6956 1146
  1765 1277 3274 7434 511 6814 4923 7360 3717 7030 8038 8108 12078 8278 8039
  6814 4923 7360 3946 125 10709 8320 511 139 156 1072 860 7564 2845 1963 678
  8038 8162 3189 7346 1921 3300 2207 7433 8039 8114 3189 704 1168 1920 7434
  8024 2229 1765 3274 7434 80398176 3189 7433 1931 7434 6760 1914 756 511
  139 156 2990 4850 131 122 119 3315 3613 6814 4923 3633 966 3217 6817 7770
  2292 3309 8024 7433 7434 1921 3698 2190 769 6858 6817 6783 3300 679 1164
  2512 1510 8039 123 119 3698 3946 974 856 8024 886 4500 4209 4125 1357 3265
  3198 2418 3800 2692 6858 7599 8024 6474 7344 671 3709 1265 4823 704 3681
  8039 124 119 1062 830 4294 1166 3221 5439 2483 4567 3655 6206 3800 2692
  7344 2170 924 3265 8024 7564 7344 1461 1429 6887 2697 3381 5023 4565 4567
  8039 125 119 5763 1736 510 6392 3177 1093 689 6206 7023 1357 2974 3177
  7344 2170 7344 1108 511 139 10564 8735 139 156 5356 6782 8038 1155 686 756
  139 156 3413 2190 8038 3742 4851 5526 139 156 2144 3417 8038 7716 5683
  6819 1726 3017 4316 8024 3389 4692 3291 1914 6569 818 5356 6782 8038 102 0
  01/14/2020 14:58:10 - INFO - processors.glue - attention_mask: 1 1 1 1 1
  75
 01/14/2020 14:58:10 - INFO - processors.glue - token_type_ids: 0 0 0 0 0
  76
 01/14/2020 14:58:10 - INFO - processors.glue -
                         label: 1 \text{ (id = 1)}
  01/14/2020 14:58:10 - INFO - processors.glue - input length: 337
77
  01/14/2020 14:58:10 - INFO - processors.glue -
                          *** Example ***
```

```
79 01/14/2020 14:58:10 - INFO - processors.glue - guid: dev-4
01/14/2020 14:58:10 - INFO - processors.glue - input_ids: 101 1298 2123
4638 1351 798 812 8024 3300 6443 4761 6887 1525 7027 1377 809 6133 1215
1912 1765 4638 2797 3322 1305 4919 1220 4638 102 1298 2123 4638 1351 798
812 8024 3300 6443 4761 6887 1525 7027 1377 809 6133 1215 1912 1765 4638
01/14/2020 14:58:10 - INFO - processors.glue - attention_mask: 1 1 1 1 1
82
01/14/2020 14:58:10 - INFO - processors.glue - token_type_ids: 0 0 0 0
83
01/14/2020 14:58:10 - INFO - processors.glue - label: 1 (id = 1)
01/14/2020 14:58:10 - INFO - processors.glue -
84
          input length: 55
01/14/2020 14:58:10 - INFO - processors.glue -
85
          *** Example ***
01/14/2020 14:58:10 - INFO - processors.glue - guid: dev-5
```

```
87 01/14/2020 14:58:10 - INFO - processors.glue - input_ids: 101 3833 1220
  170 782 3152 2110 7368 4263 2552 721 1297 3833 1220 102 4263 2552 721 1297
  4507 2110 3413 1392 2110 7368 4638 2562 2703 704 7339 1469 3413 1730 1999
  2562 2703 5442 2339 868 6956 671 6629 2130 2768 511 782 3152 2110 7368
  1730 2110 5468 2141 6664 6956 6566 6569 3315 3613 782 3152 2110 7368 721
  1297 3833 1220 4638 2458 2245 511 721 1297 4638 1059 6956 4289 1501 4507
  782 3152 2110 7368 4638 1059 860 1398 2110 2935 6615 8024 721 1297 1059
  6956 2792 2533 2100 1057 1398 3845 1920 2110 3136 5509 1355 2245 1825 7032
  833 4638 1398 38457471 2399 4263 2552 1825 7032 683 7555 1825 7032 704 511
  125 3299 8123 3189 680 8131 3189 8024 782 3152 2110 7368 4638 4263 2552
  721 1297 3833 1220 1762 678 3756 2466 2408 1767 7556 1164 2458 2245 511
  136 1762 3833 1220 1184 4638 2935 6615 4289 1501 2519 7415 704 8024 1398
  2110 812 6963 4178 2552 17652935 1139 5632 2346 4638 7312 5390 1501 8024
  1938 4346 5632 2346 4638 4263 2552 511 1728 711 1398 2110 812 4638 3118
  2898 8024 3315 3613 721 1297 2519 7415 4638 4289 1501 3144 7030 2523 1914
  8024 4905 5102 705 2168 511 1762 2173 2147 3517 678 4638 3119 7415 5056
  704 8024 738 3300 1398 2110 7355 5330 2935 6615 6132 4289 8024 3933 3933
  1856 4007 749 3119 7415 5056 511 3315 3613 4263 2552 721 1297 860 4385 749
  782 3152 2110 7368 1398 2110 100 1265 2768 1921 678 100 4638 2658 2577
  8024 860 4385 749 31733198 807 1920 2110 4495 2418 3300 4638 2857 2496 511
  01/14/2020 14:58:10 - INFO - processors.glue - attention_mask: 1 1 1 1 1
88
  89
 01/14/2020 14:58:10 - INFO - processors.glue - token_type_ids: 0 0 0 0 0
  90
 01/14/2020 14:58:10 - INFO - processors.glue -
                         label: 0 \text{ (id = 0)}
  01/14/2020 14:58:10 - INFO - processors.glue - input length: 293
  01/14/2020 14:58:22 - INFO - root - Saving features into cached file
  /data/baselines/classifier_pytorch/chineseGLUEdatasets/inews/cached_dev_ch
  inese_roberta_wwm_ext_pytorch_512_inews
  01/14/2020 14:58:24 - INFO - root - ***** Running evaluation *****
93
 01/14/2020 14:58:24 - INFO - root - Num examples = 999 01/14/2020 14:58:24 - INFO - root - Batch size = 8
94
95
  [Evaluating] 125/125 [===========] 114.5ms/step
```

```
97 01/14/2020 14:58:38 - INFO - root - ***** Eval results *****
    01/14/2020 14:58:38 - INFO - root - acc = 0.8388388388388
98
99
    01/14/2020 14:58:40 - INFO - transformers.configuration_utils -
    Configuration saved in
    /data/baselines/classifier_pytorch/outputs/inews_output/roberta/checkpoint
    -670/config.json
100
    01/14/2020 15:00:23 - INFO - transformers.modeling_utils - Model weights
    saved in
    /data/baselines/classifier_pytorch/outputs/inews_output/roberta/checkpoint
    -670/pytorch_model.bin
    01/14/2020 15:00:23 - INFO - root - Saving model checkpoint to
    /data/baselines/classifier_pytorch/outputs/inews_output/roberta/checkpoint
102
    [Training] 670/670 [============== ] 546.6ms/step loss:
    0.0379
103
    0.1278 01/14/2020 15:04:14 - INFO - root - Loading features from cached
    /data/baselines/classifier_pytorch/chineseGLUEdatasets/inews/cached_dev_ch
    inese_roberta_wwm_ext_pytorch_512_inews
    01/14/2020 15:04:15 - INFO - root - ***** Running evaluation *****
104
    01/14/2020 15:04:15 - INFO - root - Num examples = 999
01/14/2020 15:04:15 - INFO - root - Batch size = 8
105
106
107
    108 01/14/2020 15:04:29 - INFO - root - ***** Eval results *****
    01/14/2020 15:04:29 - INFO - root - acc = 0.8278278278278
109
110 01/14/2020 15:04:31 - INFO - transformers.configuration_utils -
    Configuration saved in
    /data/baselines/classifier_pytorch/outputs/inews_output/roberta/checkpoint
    -1340/config.json
111 01/14/2020 15:06:14 - INFO - transformers.modeling_utils - Model weights
    saved in
    /data/baselines/classifier_pytorch/outputs/inews_output/roberta/checkpoint
    -1340/pytorch_model.bin
112
    01/14/2020 15:06:15 - INFO - root - Saving model checkpoint to
    /data/baselines/classifier_pytorch/outputs/inews_output/roberta/checkpoint
    -1340
    113
    0.2891
    0.0177 01/14/2020 15:10:06 - INFO - root - Loading features from cached
    file
    /data/baselines/classifier_pytorch/chineseGLUEdatasets/inews/cached_dev_ch
    inese_roberta_wwm_ext_pytorch_512_inews
115
    01/14/2020 15:10:07 - INFO - root - ***** Running evaluation *****
   01/14/2020 15:10:07 - INFO - root - Num examples = 999
01/14/2020 15:10:07 - INFO - root - Batch size = 8
116
117
    118
    01/14/2020 15:10:21 - INFO - root - ***** Eval results *****
119
    01/14/2020 15:10:21 - INFO - root - acc = 0.8458458458458459
120
121
    01/14/2020 15:10:23 - INFO - transformers.configuration_utils -
    Configuration saved in
    /data/baselines/classifier_pytorch/outputs/inews_output/roberta/checkpoint
    -2010/config.json
122
    01/14/2020 15:12:04 - INFO - transformers.modeling_utils - Model weights
    saved in
    /data/baselines/classifier_pytorch/outputs/inews_output/roberta/checkpoint
    -2010/pytorch_model.bin
123
    01/14/2020 15:12:05 - INFO - root - Saving model checkpoint to
    /data/baselines/classifier_pytorch/outputs/inews_output/roberta/checkpoint
    124
    0.0023
```

```
0.4203 01/14/2020 15:15:55 - INFO - root - Loading features from cached
     /data/baselines/classifier_pytorch/chineseGLUEdatasets/inews/cached_dev_ch
     inese_roberta_wwm_ext_pytorch_512_inews
     01/14/2020 15:15:57 - INFO - root - ***** Running evaluation *****
126
     01/14/2020 15:15:57 - INFO - root - Num examples = 999
01/14/2020 15:15:57 - INFO - root - Batch size = 8
127
128
129
     [Evaluating] 125/125 [============== ] 115.7ms/step
    01/14/2020 15:16:12 - INFO - root - ***** Eval results *****
130
131 01/14/2020 15:16:12 - INFO - root - acc = 0.8418418418419
132
    01/14/2020 15:16:13 - INFO - transformers.configuration_utils -
     Configuration saved in
     /data/baselines/classifier_pytorch/outputs/inews_output/roberta/checkpoint
     -2680/config.json
     01/14/2020 15:17:55 - INFO - transformers.modeling_utils - Model weights
133
     saved in
     /data/baselines/classifier_pytorch/outputs/inews_output/roberta/checkpoint
     -2680/pytorch_model.bin
     01/14/2020 15:17:56 - INFO - root - Saving model checkpoint to
134
     /data/baselines/classifier_pytorch/outputs/inews_output/roberta/checkpoint
     -2680
     135
     0.0022
136
     01/14/2020 15:17:57 - INFO - root - global_step = 2680, average loss =
     0.31527684231179265
     01/14/2020 15:17:57 - INFO - root - Saving model checkpoint to
137
     /data/baselines/classifier_pytorch/outputs/inews_output/roberta
138
     01/14/2020 15:17:57 - INFO - transformers.configuration_utils -
     Configuration saved in
     /data/baselines/classifier_pytorch/outputs/inews_output/roberta/config.jso
     01/14/2020 15:19:41 - INFO - transformers.modeling_utils - Model weights
     saved in
     /data/baselines/classifier_pytorch/outputs/inews_output/roberta/pytorch_mo
     del.bin
140
     01/14/2020 15:19:44 - INFO - transformers.configuration_utils - loading
     configuration file
     /data/baselines/classifier_pytorch/outputs/inews_output/roberta/config.jso
141
     01/14/2020 15:19:45 - INFO - transformers.configuration_utils - Model
     config {
142
       "attention_probs_dropout_prob": 0.1,
       "directionality": "bidi",
143
       "finetuning_task": "inews",
144
145
       "hidden_act": "gelu",
       "hidden_dropout_prob": 0.1,
146
147
       "hidden_size": 768,
148
       "initializer_range": 0.02,
       "intermediate_size": 3072,
149
150
       "layer_norm_eps": 1e-12,
       "max_position_embeddings": 512,
151
152
       "num_attention_heads": 12,
153
       "num_hidden_layers": 12,
154
       "num_labels": 3,
       "output_attentions": false,
155
       "output_hidden_states": false,
156
       "output_past": true,
157
       "pooler_fc_size": 768,
158
159
       "pooler_num_attention_heads": 12,
160
       "pooler_num_fc_layers": 3,
       "pooler_size_per_head": 128,
161
       "pooler_type": "first_token_transform",
162
       "pruned_heads": {},
163
164
       "torchscript": false,
```

```
166
       "use_bfloat16": false,
167
       "vocab_size": 21128
168
169
170
     01/14/2020 15:19:45 - INFO - transformers.modeling_utils - loading
     weights file
     /data/baselines/classifier_pytorch/outputs/inews_output/roberta/pytorch_mo
     del.bin
    01/14/2020 15:20:13 - INFO - transformers.tokenization_utils -
171
     name '/data/baselines/classifier_pytorch/outputs/inews_output/roberta' not
     found in model shortcut name list (bert-base-uncased, bert-large-uncased,
     bert-base-cased, bert-large-cased, bert-base-multilingual-uncased, bert-
     base-multilingual-cased, bert-base-chinese, bert-base-german-cased, bert-
     large-uncased-whole-word-masking, bert-large-cased-whole-word-masking,
     bert-large-uncased-whole-word-masking-finetuned-squad, bert-large-cased-
     whole-word-masking-finetuned-squad, bert-base-cased-finetuned-mrpc, bert-
     base-german-dbmdz-cased, bert-base-german-dbmdz-uncased). Assuming
     '/data/baselines/classifier_pytorch/outputs/inews_output/roberta' is a
     path or url to a directory containing tokenizer files.
     01/14/2020 15:20:14 - INFO - transformers.tokenization_utils -
172
     /data/baselines/classifier_pytorch/outputs/inews_output/roberta/vocab.txt
173
     01/14/2020 15:20:14 - INFO - transformers.tokenization_utils -
     file
     /data/baselines/classifier_pytorch/outputs/inews_output/roberta/added_toke
     ns.json
174
     01/14/2020 15:20:14 - INFO - transformers.tokenization_utils -
     file
     /data/baselines/classifier_pytorch/outputs/inews_output/roberta/special_to
     kens_map.json
175
     01/14/2020 15:20:14 - INFO - transformers.tokenization_utils -
     /data/baselines/classifier_pytorch/outputs/inews_output/roberta/tokenizer_
     config.json
176
     01/14/2020 15:20:14 - INFO - transformers.tokenization_utils -
     name '/data/baselines/classifier_pytorch/outputs/inews_output/roberta' not
     found in model shortcut name list (bert-base-uncased, bert-large-uncased,
     bert-base-cased, bert-large-cased, bert-base-multilingual-uncased, bert-
     base-multilingual-cased, bert-base-chinese, bert-base-german-cased, bert-
     large-uncased-whole-word-masking, bert-large-cased-whole-word-masking,
     bert-large-uncased-whole-word-masking-finetuned-squad, bert-large-cased-
     whole-word-masking-finetuned-squad, bert-base-cased-finetuned-mrpc, bert-
     base-german-dbmdz-cased, bert-base-german-dbmdz-uncased). Assuming
     '/data/baselines/classifier_pytorch/outputs/inews_output/roberta' is a
     path or url to a directory containing tokenizer files.
     01/14/2020 15:20:14 - INFO - transformers.tokenization_utils -
177
     /data/baselines/classifier_pytorch/outputs/inews_output/roberta/vocab.txt
178
     01/14/2020 15:20:14 - INFO - transformers.tokenization_utils -
     file
     /data/baselines/classifier_pytorch/outputs/inews_output/roberta/added_toke
179
     01/14/2020 15:20:14 - INFO - transformers.tokenization_utils -
     /data/baselines/classifier_pytorch/outputs/inews_output/roberta/special_to
     kens_map.json
     01/14/2020 15:20:14 - INFO - transformers.tokenization_utils -
180
     /data/baselines/classifier_pytorch/outputs/inews_output/roberta/tokenizer_
181
     01/14/2020 15:20:15 - INFO - root - Evaluate the following checkpoints:
     ['/data/baselines/classifier_pytorch/outputs/inews_output/roberta']
```

165

"type\_vocab\_size": 2,

```
01/14/2020 15:20:15 - INFO - transformers.configuration_utils - loading
     configuration file
     /data/baselines/classifier_pytorch/outputs/inews_output/roberta/config.jso
183
     01/14/2020 15:20:15 - INFO - transformers.configuration_utils - Model
     config {
184
       "attention_probs_dropout_prob": 0.1,
       "directionality": "bidi",
185
       "finetuning_task": "inews",
186
       "hidden_act": "gelu",
187
       "hidden_dropout_prob": 0.1,
188
189
       "hidden_size": 768,
190
       "initializer_range": 0.02,
       "intermediate_size": 3072,
191
192
       "layer_norm_eps": 1e-12,
193
       "max_position_embeddings": 512,
194
       "num_attention_heads": 12,
       "num_hidden_layers": 12,
195
       "num_labels": 3,
196
197
       "output_attentions": false,
       "output_hidden_states": false,
198
199
       "output_past": true,
200
       "pooler_fc_size": 768,
201
       "pooler_num_attention_heads": 12,
202
       "pooler_num_fc_layers": 3,
       "pooler_size_per_head": 128,
203
       "pooler_type": "first_token_transform",
204
       "pruned_heads": {},
205
       "torchscript": false,
206
207
       "type_vocab_size": 2,
       "use_bfloat16": false,
208
209
       "vocab_size": 21128
210 }
211
212
     01/14/2020 15:20:15 - INFO - transformers.modeling_utils - loading
     weights file
     /data/baselines/classifier_pytorch/outputs/inews_output/roberta/pytorch_mo
     del.bin
213
    01/14/2020 15:20:41 - INFO - root - Loading features from cached file
     /data/baselines/classifier_pytorch/chineseGLUEdatasets/inews/cached_dev_ch
     inese_roberta_wwm_ext_pytorch_512_inews
214 01/14/2020 15:20:41 - INFO - root - ***** Running evaluation *****
215 01/14/2020 15:20:41 - INFO - root - Num examples = 999 216 01/14/2020 15:20:41 - INFO - root - Batch size = 8
     217
218 01/14/2020 15:20:56 - INFO - root - ***** Eval results *****
219 01/14/2020 15:20:56 - INFO - root -
                                           acc = 0.8418418418418419
```

复现成功。