Comp5211 2019-Fall Final Project

COMP5211

WANG, Sukai 20528424

Problem 3 - NLP:

Evaluating pretrained language model: BERT

Introduction

BERT[1] (Bidirectional Encoder Representations from Transformers) is a state-of-the-art method in the area of NLP task.

Transformer is used in BERT as an attention mechanism to learn the contextual relations between words and sentences. The transformer can be split into two parts, one is encoder and the other is decoder. Encoder can consume text input and decoder can output the predict result in different tasks. It can be considered as a bidirectional text model since it reads the entire sequence at once, rather than read sequence from left to right or from right to left.

Experiment Results:

1. Environment:

   I use docker to create a valid environment. Docker is the

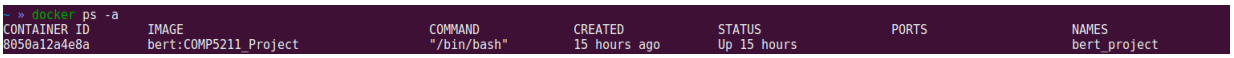
   Docker is a popular tool to make it easier to build, deploy and run applications using containers

   I placed my dockerfile in docker/dockerfile

   The docker image and container created are shown below:

Docker image:

Docker container:



2. Bert server:

I downloaded the model from https://github.com/CodePothunter/COMP5211_Project/blob/master/Bert-as-Service_Install.md and placed it in Model folder.

Then I use

bert-serving-start -model_dir /code/Models/uncased_L-12_H-768_A-12/ -num_worker=1 -pooling_strategy=NONE -show_tokens_to_client

to start the server.

The starting and responding interfaces are shown in below.

```
usage: /miniconda/envs/py37/bin/bert-serving-start -model_dir /code/Models/uncased_L-12_H-768_A-12/ -n
um_worker=1 -pooling_strategy=NONE -show_tokens_to_client
                ARG   VALUE
__________________________________________________
           ckpt_name = bert_model.ckpt
         config_name = bert_config.json
                cors = *
                 cpu = False
          device_map = []
       do_lower_case = True
  fixed_embed_length = False
                fp16 = False
 gpu_memory_fraction = 0.5
      graph_tmp_dir = None
    http_max_connect = 10
           http_port = None
       mask_cls_sep = False
      max_batch_size = 256
         max_seq_len = 25
           model_dir = /code/Models/uncased_L-12_H-768_A-12/
no_position_embeddings = False
    no_special_token = False
          num_worker = 1
       pooling_layer = [-2]
    pooling_strategy = NONE
                port = 5555
            port_out = 5556
       prefetch_size = 10
 priority_batch_size = 16
show_tokens_to_client = True
     tuned_model_dir = None
             verbose = False
                 xla = False

I:VENTILATOR:[__i:__i: 67]:freeze, optimize and export graph, could take a while...
WARNING:tensorflow:From /miniconda/envs/py37/lib/python3.7/site-packages/bert_serving/server/helper.py
:186: The name tf.logging.set_verbosity is deprecated. Please use tf.compat.v1.logging.set_verbosity i
nstead.
```

```
WARNING:tensorflow:From /miniconda/envs/py37/lib/python3.7/site-packages/bert_serving/server/helper.py
:186: The name tf.logging.ERROR is deprecated. Please use tf.compat.v1.logging.ERROR instead.

I:GRAPHOPT:[gra:opt: 52]:model config: /code/Models/uncased_L-12_H-768_A-12/bert_config.json
I:GRAPHOPT:[gra:opt: 55]:checkpoint: /code/Models/uncased_L-12_H-768_A-12/bert_model.ckpt
I:GRAPHOPT:[gra:opt: 59]:build graph...
I:GRAPHOPT:[gra:opt:129]:load parameters from checkpoint...
I:GRAPHOPT:[gra:opt:133]:optimize...
I:GRAPHOPT:[gra:opt:141]:freeze...
I:GRAPHOPT:[gra:opt:146]:write graph to a tmp file: /tmp/tmpce8cb8vl
I:VENTILATOR:[__i:__i: 75]:optimized graph is stored at: /tmp/tmpce8cb8vl
I:VENTILATOR:[__i:_ru:129]:bind all sockets
I:VENTILATOR:[__i:_ru:133]:open 8 ventilator-worker sockets
I:VENTILATOR:[__i:_ru:136]:start the sink
I:SINK:[__i:_ru:306]:ready
I:VENTILATOR:[__i:_ge:222]:get devices
I:VENTILATOR:[__i:_ge:255]:device map:
                worker  0 -> gpu  0
I:WORKER-0:[__i:_ru:531]:use device gpu: 0, load graph from /tmp/tmpce8cb8vl
WARNING:tensorflow:From /miniconda/envs/py37/lib/python3.7/site-packages/bert_serving/server/helper.py
:186: The name tf.logging.set_verbosity is deprecated. Please use tf.compat.v1.logging.set_verbosity i
nstead.

WARNING:tensorflow:From /miniconda/envs/py37/lib/python3.7/site-packages/bert_serving/server/helper.py
:186: The name tf.logging.ERROR is deprecated. Please use tf.compat.v1.logging.ERROR instead.

I:WORKER-0:[__i:gen:559]:ready and listening!
I:VENTILATOR:[__i:_ru:164]:all set, ready to serve request!
I:VENTILATOR:[__i:_ru:180]:new config request   req id: 1       client: b'ffbd2dd2-8e23-4e1a-9887-d994
2dc2c079'
I:SINK:[__i:_ru:348]:send config        client b'ffbd2dd2-8e23-4e1a-9887-d9942dc2c079'
I:VENTILATOR:[__i:_ru:196]:new encode request   req id: 2       size: 2 client: b'ffbd2dd2-8e23-4e1a-9
887-d9942dc2c079'
I:SINK:[__i:_ru:342]:job register       size: 2 job id: b'ffbd2dd2-8e23-4e1a-9887-d9942dc2c079#2'
I:WORKER-0:[__i:gen:568]:new job        socket: 0       size: 2 client: b'ffbd2dd2-8e23-4e1a-9887-d994
2dc2c079#2'
```

3. Q&A

I used example.txt as the evaluation content. The content contains lots of lines, where each line use ||| to split the question and answer.

Our code use BERT to classify each question(encode each question language), and when we input the query sentence, we use BERT to encode query sentence too. Then we can find the closest question sentence, which means this question has the closest meaning with the query problem.

The results are shown below:

Firstly, I output the entire question list.

```
root@8050a12a4e8a:/code# python simple_exam.py
The question list:
What are the differences between research postgraduate programs & taught postgraduate programs?
Can I work and study at the same time as a full-time research postgraduate student?
Will I be provided with University housing?
Am I required to take the English Proficiency Test?
I obtained a master's degree from an institution where the medium of instruction was English. Am I req
uired to take the English Proficiency Test?
What should I do if my IELTS/ TOEFL (or another public examination) result is not yet available?
Is IELTS for UKVI accepted to fulfill the English proficiency requirement?
Am I required to submit GRE/GMAT score for application?
I want to apply for a postgraduate program at HKUST. What should I do?
I am currently a final year undergraduate student, can I apply now?
Can I apply for more than one postgraduate program?
Is there any difference if I apply as a local or non-local applicant?
What documents do I need to upload for application?
Why is "Institution Grading System" required?
Do I need to include a CV/ personal statement in my application?
Am I required to submit a research proposal for my MPhil/PhD application?
Can I make changes to my application after submission?
What if I forget my login password in the Admission System?
I cannot upload the supporting documents. What should I do?
Can I update/ delete supporting documents after submitting the application?
I have submitted my application already. Can I still change my program choice?
How many referees should I nominate to support my application?
I have provided a wrong email address of the referee. What should I do?
When should my referees submit the reference report?
Will the University accept reference report or letter submitted by hard copy or mail?
Why my referees cannot receive the invitation from the University?
How can I know if the University has received the reference report or not?
When will I be notified of an admission offer?
Am I eligible to apply for the HK PhD Fellowship Scheme?
What should I do if IELTS/ TOEFL (or another public examination) result is not yet available?
Are there any special admission requirements for HK PhD Fellowship application?
```

Every query sentence will be found a closest meaning question and output the prepared answer as the query's answer.

```
>>> (input your query):I cannot upload the supporting documents.
If you mean this?:  I cannot upload the supporting documents. What should I do?
The answer is:
 Use another internet browser and disable any pop-up blocker.  If the total file size exceeds the allo
wed size, you may try to reduce the file size (e.g. reduce resolution/ change to black and white versi
on). Each file should not exceed 1 MB.
```

We can see that though the query is not as same as the prepared question, we can get the correct answer successfully. That's mean the BERT has the ability to find the sentences' meaning.

But we can also find some interesting results:

```
>>> (input your query):Am I eligible to apply for HKPFS
If you mean this?:  Am I required to submit GRE/GMAT score for application?
The answer is:
 Only a few programs require applicants to submit GMAT or GRE score for application. You may refer to
the information on "Admission Requirements" tab in the individual program page.
```

When I asked "Am I eligible to apply for HKPFS", I got the closest problem "Am I required to submit GRE/GMAT score for application?". But when I input the query with "?" at the end of it, then we can get what we want.

```
>>> (input your query):Am I eligible to apply for HKPFS?
If you mean this?:  Am I eligible to apply for the HK PhD Fellowship Scheme?
The answer is:
 HKPFS is an internationally recognized award aiming to draw the best PhD students around the globe to
 Hong Kong. You are welcome to apply if you are interested in pursuing PhD studies at HKUST. Please ch
eck the admission requirements here.
```

4. More Evaluations – NOT test:

In this part, we evaluate the BERT model's ability in distinguishing the "NOT" in a sentence.

When we set the sentence list:

A loves B.

A does not love B.

We can test the BERT results:

```
root@8050a12a4e8a:/code# python NOT_test.py
The sentence list:
A loves B.
A does not love B.
>>> (input your query):A likes B.
The cloest meaning sentence is:
A loves B.
>>> (input your query):A dislikes B.
The cloest meaning sentence is:
A does not love B.
>>> (input your query):B dislikes A.
The cloest meaning sentence is:
A does not love B.
>>> (input your query):B likes A.
The cloest meaning sentence is:
A loves B.
```

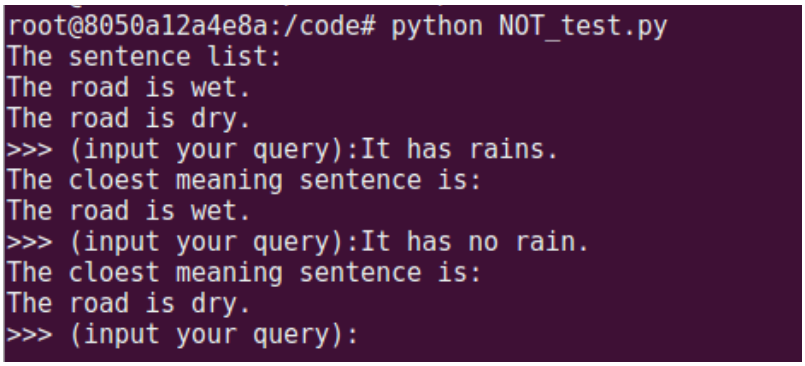
Then we set the sentence list as:

The road is wet.

The road is dry.

So, when we input that “It has rains.” BERT will understand that the road is wet.

And when we input the “NOT”, BERT will think the road is dry.



Lastly we test the “Apple” question.

We set the sentence list as:

We should eat apple.

We should not eat apple.

And when we input that: “Apple is good for our health.”

It will output that: We should eat apple.

And when we input that: ”Apple is bad for our health.”

It will connect the mean of “bad” with “not”, and output “We should not eat apple.”

```
root@8050a12a4e8a:/code# python NOT_test.py
The sentence list:
We should eat apple.
We should not eat apple.
>>> (input your query):Apple is good for our health.
The cloest meaning sentence is:
We should not eat apple.
>>> (input your query):Apple is bad for our health.
The cloest meaning sentence is:
We should not eat apple.
```