

Homework5 ZY2203114 王彪

本作业均选用Huggingface transformers的开源模型进行比较分析。

一、任务1：文本摘要

英文语段：

```
1 The only survivor of a shipwreck was washed up on a small uninhabited island.
2 He prayed feverishly for God to rescue him, and every day he scanned the
  horizon for help, but none seemed forthcoming.
3 Exhausted, he eventually managed to build a little hut out of driftwood to
  protect him from the elements, and to store his few possessions.
4 But then one day, after scavenging for food, he arrived home to find his
  little hut in flames, the smoke rolling up to the sky. The worst had
  happened; everything was lost. He was stunned with grief and anger. "God how
  could you do this to me!" he cried.
5 Early the next day, however, he was awakened by the sound of a ship that was
  approaching the island. It had come to rescue him.
6 "How did you know I was here?" asked the weary man of his rescuers.
7 "We saw your smoke signal," they replied.
8 It is easy to get discouraged when things are going bad. But we shouldn't lose
  heart, because God is at work in our lives, even in the midst of pain and
  suffering.
```

中文语段：

```
1 海难中唯一的幸存者被冲上了一个无人居住的小岛。
2 他狂热地祈求上帝拯救他，每天他都扫视着地平线寻求帮助，但似乎没有人会出现。
3 筋疲力尽的他最终设法用浮木建造了一间小屋，以保护他免受恶劣天气的影响，并存放他为数不多的财
  产。
4 可有一天，他在觅食后回到家，发现他的小屋着火了，浓烟冲天。 最坏的情况发生了；一切都丢失了。
  他因悲伤和愤怒而震惊。“天啊，你怎么可以这样对我！”他哭了。
5 然而，第二天一早，他被一艘靠近该岛的船只的声音吵醒了。它是来救他的。
6 “你怎么知道我在这里？”疲倦的人问他的救援人员。
7 “我们看到了你的烟雾信号，”他们回答道。
8 当事情变坏时很容易灰心。但我们不应该灰心，因为上帝在我们的生活中工作，即使在痛苦和苦难中。
```

1.1 bart-large-cnn-samsum

[philschmid/bart-large-cnn-samsum · Hugging Face](https://huggingface.co/philschmid/bart-large-cnn-samsum)

```
1 The only survivor of a shipwreck was washed up on a small uninhabited island.
  He built a little hut out of driftwood to protect him from the elements and
  store his few possessions. One day he arrived home to find his hut in flames.
  The next day he was rescued by a ship that was approaching the island.
```

该模型不支持中文任务，摘要比较精简，能够保持故事原意。

1.2 bigbird-pegasus-large-bigpatent

[google/bigbird-pegasus-large-bigpatent · Hugging Face](#)

- 1 A man washed up on an island saves most of his possessions and builds a little hut out of driftwood.

该模型不支持中文任务，但摘要更精简，但是一定程度上破坏了故事原意。

1.3 mT5_multilingual_XLSum

[csebuatnlp/mT5_multilingual_XLSum · Hugging Face](#)

- 1 when a shipwreck landed on the Isle of Wight in July last year, one man was stunned with grief and anger. But he couldn't lose heart, because God is at work in our lives, even in the midst of pain and suffering.

- 1 在南卡罗来纳斯群岛(Kangaroo Islands)的海域,只有一个幸存者被冲上了一个无人居住的小岛。

该模型是一个多语言模型，用它测试了中英文，英文效果很好，能够把握句子中心思想，而且存在原语段中未见过的词汇。中文语段虽然没有对语段有整体的分析，但是惊奇的补充出了原语段中不存在的群岛名称以及对应的英文注解。

所以，在文本摘要任务，MT5>=BART>BIGBIRD

二、任务2：文本生成

- 1 My name is wangbiao,like singing, dancing, rap, basketball

- 1 我叫王彪，喜欢唱、跳、rap、篮球。

2.1 GPT2

[gpt2 · Hugging Face](#)

- 1 My name is wangbiao,like singing, dancing, rap, basketball, martial arts, and so on. My hobby is studying magic and mysticism. I play all day long at events and tournaments, like Grand Prix with all my friends.

2.2 gpt2-xl

[gpt2-xl · Hugging Face](#)

- 1 My name is wangbiao,like singing, dancing, rap, basketball, skiing, swimming, water-skiing, skiing, bicycling..." (5)

2.3 gpt2-chinese-cluecorpussmall

[uer/gpt2-chinese-cluecorpussmall · Hugging Face](#)

1 | 我叫王彪，喜欢唱、跳、rap、篮球。还有我们的校队也特别棒，都是世界各地的佼佼者，我们总有一天会成全我们自己的节目，走向世界。而我也非常喜欢唱歌跳舞，这些优点能使我成为很多人追梦梦想的舞台。我喜欢唱歌跳舞。林海音说，希望她能在校园里开一间我所喜爱的歌舞演艺学校，为她以后的梦想铺路。梦想只是一种借口。林海音很小的时候便开始唱歌跳舞，但他从未放弃过，坚持到现在，这是他的信念和骄傲。梦的成功，说明了所谓的梦想。这些因素，让林海音和王彪都非常享受生活，他说：梦想并不是一个词儿，一旦我们决定要走远，就会觉得更累。而现在的我们，只要心情平和，可以很自由的活出自我、健康、快乐，我们就会更加自信。在林海音眼里，梦想能在家

这一部分选用了3个不同参数量支持不同语言的GPT2模型，三个模型的整体表现都比较好，但是没有经过在我们训练集的微调，生成的文本会逐渐偏离原语境，比如中文GPT2后面的主语已经变为林海音了。

三、任务3：问答

1 | what is my age

1 | My name is Wang Biao. I was born in 1999. I will be admitted to Beihang University in 2022.

1 | 我今年多少岁

1 | 我叫王彪，出生于1999年，于2022年考入北京航空航天大学

3.1 tinyroberta-squad2

[deepset/tinyroberta-squad2 · Hugging Face](#)

1 | I was born in 1999

能够分辨出出生信息，但是没有计算出age

3.2 longformer-large-4096-finetuned-triviaqa

[allenai/longformer-large-4096-finetuned-triviaqa · Hugging Face](#)

1 | 2022

这个模型直接把提示文本的句子搞混了，认为我2022年出生，而且回答我是2022岁，虽然参数量少，但效果不佳。

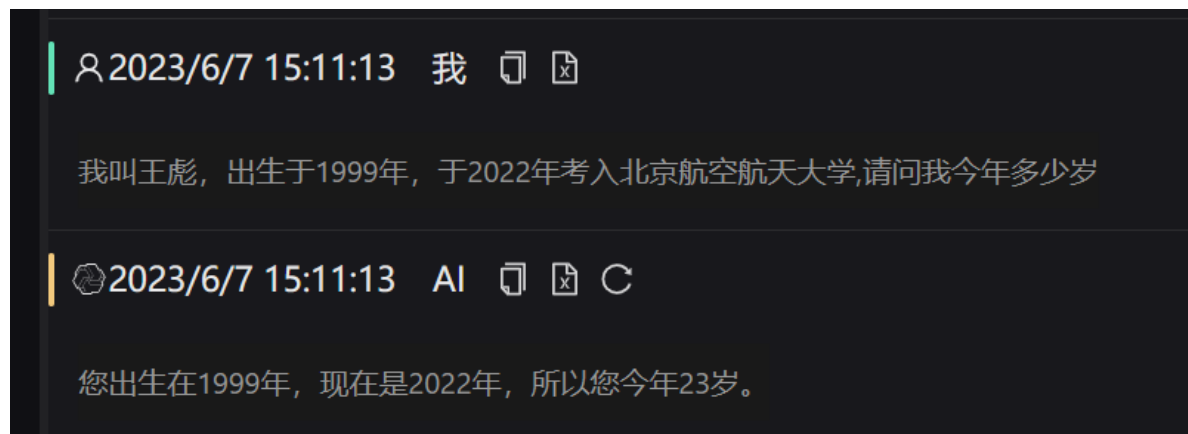
3.3 chinese_pretrain_mrc_roberta_wwm_ext_large

[luhua/chinese_pretrain_mrc_roberta_wwm_ext_large · Hugging Face](#)

1 | 2022年考入北京航空航天大学

这个中文模型也是与上面一样，答非所问了

3.4 GPT4.0



ChatGPT效果很好，有推理过程，能够计算出实际的年龄。不过没有联网的话，预训练模型是2022年训练的，所以会有一些偏差。

下面实验联网后的效果：



联网后就非常正确了，且有称呼还会发表情。

4. 总结:

本文实验了10个模型对于文本摘要、文本生成、文本问答任务的处理水平差异。整体来看，参数量越大的模型表现出的效果越佳。对于摘要或生成任务，也比较依赖于对预训练模型的微调。