



黑马程序员线上品牌

LLM实现金融文本匹配

一样的教育，不一样的品质



目录

Contents

1. 文本匹配任务介绍
2. Prompt设计
3. 代码实现
4. 结果展示

01

文本匹配任务介绍

信息文本匹配介绍

首先，我们构造几个短文本对：：

1. ('股票市场今日大涨，投资者乐观。'， '持续上涨的市场让投资者感到满意。'),
2. ('油价大幅下跌，能源公司面临挑战。'， '未来智能城市的建设趋势愈发明显。'),
3. ('利率上升，影响房地产市场。'， '高利率对房地产有一定冲击。'),

我们期望模型能够帮我们识别出这 3 对句子中，哪几对描述的是相似的语言。

我们期望模型输出的结果为：

[‘相似’， ‘不相似’，
‘相似’]

02

Prompt设计

Prompt设计

对于大模型来讲，prompt 的设计非常重要，一个「明确」的 prompt 能够帮助我们更好从大模型中获得我们想要的结果。



设计要点

向模型解释什么叫作「文本匹配任务」

让模型按照我们指定的格式输出

Prompt设计

为了让模型知道什么叫做「文本匹配」，我们借用 Incontext Learning 的方式，先给模型展示几个正确的例子

- > User: 句子一：公司ABC发布了季度财报，显示盈利增长。\\n句子二：财报披露，公司ABC利润上升
- > Bot: 是
- > User: 句子一：黄金价格下跌，投资者抛售。\\n句子二：外汇市场交易额创下新高
- > Bot: 不是

其中，`User` 代表我们输入给模型的句子，`Bot` 代表模型的回复内容。

注意：上述例子中 `Bot` 的部分也是由人工输入的，其目的是希望看到在看到类似 `User` 中的句子时，模型应当做出类似 `Bot` 的回答。

03

代码实现

代码实现

本章节使用的模型为ChatGLM-6B，参数参数较大（6B），下载到本地大概需要 12G+ 的磁盘空间，请确保磁盘有充足的空间。此外，加载模型大概需要 13G 左右的显存，如果您显存不够，可以进行模型量化加载以缩小模型成本。

本次文本匹配任务实现的主要过程：

构造prompt：对应init_prompts()函数

进行文本匹配：对应inference()函数

代码存放位置：[/Users/**/PycharmProjects/llm/zero-shot/finance_text_matching.py](#)

具体代码实现参考：[pdf文档（附件资料）](#)

04

结果展示

结果展示

```
The dtype of attention mask (torch.int64) is not bool
>>> sentence: ('股票市场今日大涨，投资者乐观。', '持续上涨的市场让投资者感到满意。')
>>> inference answer: 是
>>> sentence: ('油价大幅下跌，能源公司面临挑战。', '未来智能城市的建设趋势愈发明显。')
>>> inference answer: 不是
>>> sentence: ('利率上升，影响房地产市场。', '高利率对房地产有一定冲击。')
>>> inference answer: 是
```



黑马程序员线上品牌



扫码关注博学谷微信公众号

