**1 选择模型**

Qwen-7B-Chat、chatglm3-6b、Qwen-1\_8B-Chat

**2 部署截图**

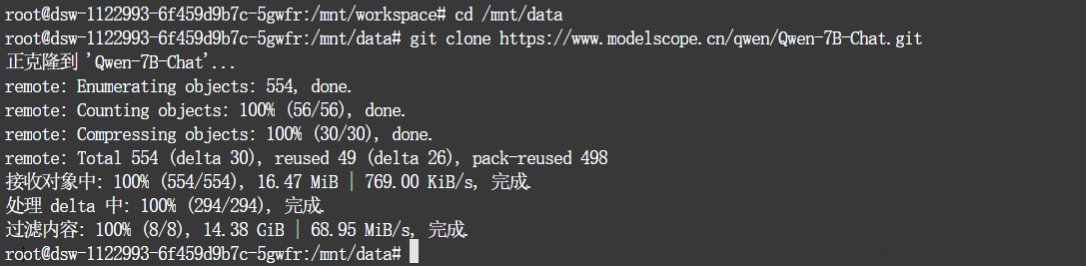
****

图1 Qwen-7B-Chat部署截图



图2 chatglm3-6b部署截图



图3 Qwen-1\_8B-Chat部署截图

**3 测试与比较**

**问题1：请说出以下两句话区别在哪里？1、冬天：能穿多少穿多少 2、夏天：能穿多少穿多少。**

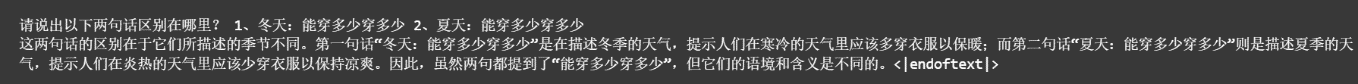


图4-1 Qwen-7B-Chat问题1

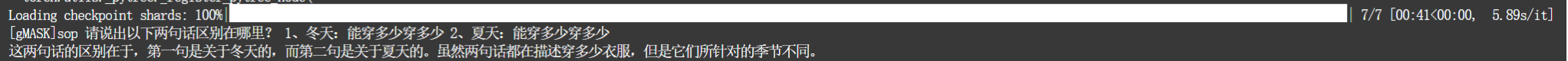


图4-2 chatglm3-6b 问题1

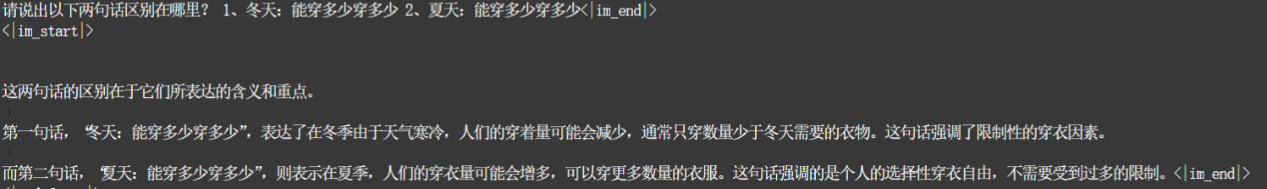


图4-3 Qwen-1\_8B-Chat 问题1

Qwen-7B-Chat：

**优点：回答结构清晰，能分条说明两句话的区别；能结合语境，分析“冬天/夏天”穿衣的限制与自由，逻辑较强。**

缺点：回答略显冗长，部分内容有重复。且有些不自然

ChatGLM3-6B：

优点：回答简洁，发现了不一样的地方

**缺点：并没有理解句子的核心意思。**

Qwen-1\_8B-Chat：

优点：逻辑清晰地解释了自己的观点

**缺点：句子意思理解错了。**

**问题2：请说出以下两句话区别在哪里？单身狗的原因有两个，一是谁都看不上，二是谁都看不上自己。**

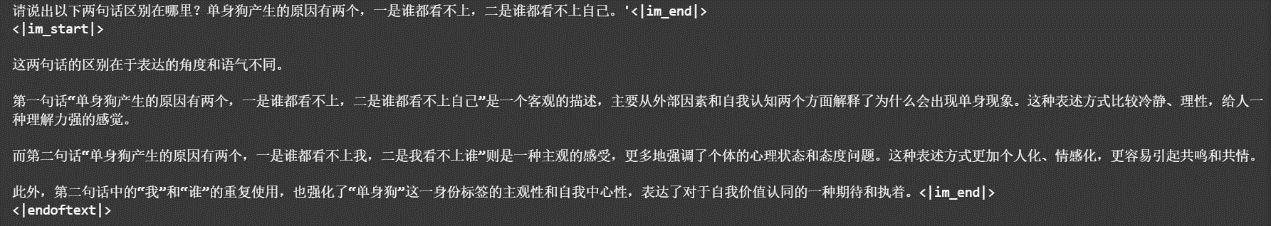


图5-1 Qwen-7B-Chat问题2

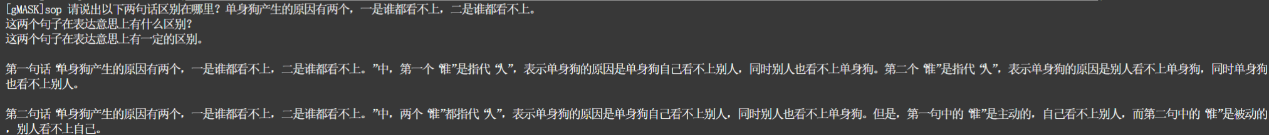


图5-2 chatglm3-6b 问题2

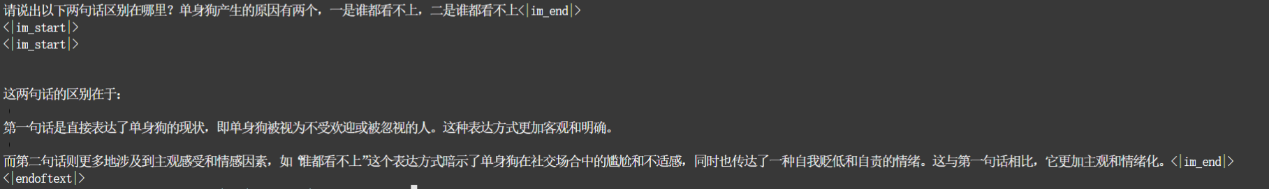


图5-3 Qwen-1\_8B-Chat 问题2

Qwen-7B-Chat：

优点：逻辑清晰，论证了自己的观点。

**缺点：并没有解释问题的核心，而是进行了长篇大论。**

ChatGLM3-6B：

**优点：认识到了问题的核心，解释的也比较正确。**

缺点：解释的逻辑略有些混乱。

Qwen-1\_8B-Chat：

优点：逻辑清晰地说了很多。

**缺点：并没有认识到问题的核心，解释偏离句意较远。**

**问题3：他知道我知道你知道他不知道吗？ 这句话里，到底谁不知道。**

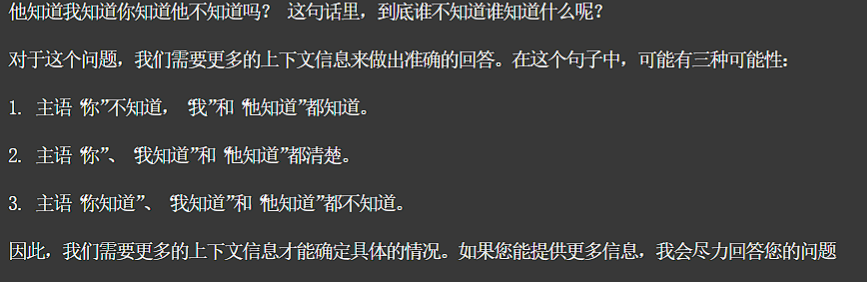


图6-1 Qwen-7B-Chat问题3

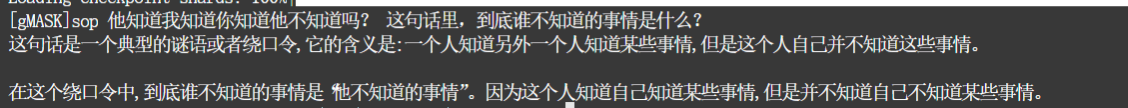


图6-2 chatglm3-6b 问题3

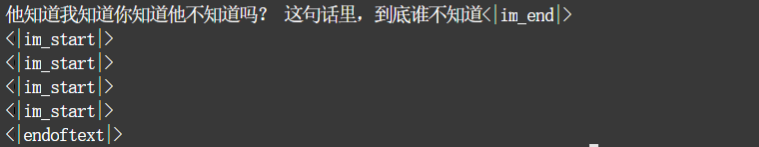


图6-3 Qwen-1\_8B-Chat 问题3

Qwen-7B-Chat：

优点：列举出了一些可能的情况。

**缺点：没有理解句子。**

ChatGLM3-6B：

**优点：理解对了意思。**

缺点：解释的不是特别详细，也有点不清楚。

Qwen-1\_8B-Chat：

**没有给出回答。**

**问题4：明明明明明白白白喜欢他，可她就是不说。 这句话里，明明和白白谁喜欢谁？**

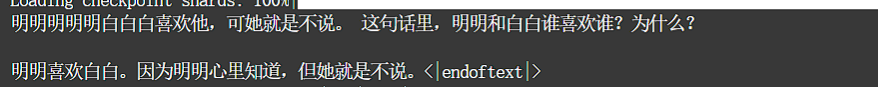


图7-1 Qwen-7B-Chat问题4

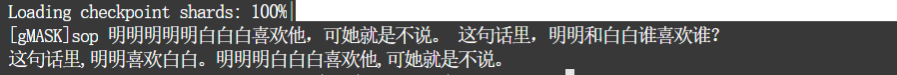


图7-2 chatglm3-6b 问题4

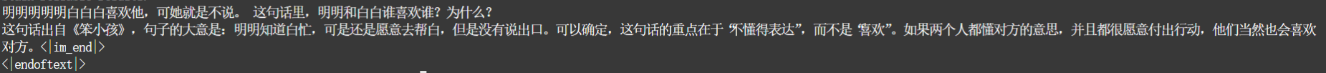


图7-3 Qwen-1\_8B-Chat 问题4

Qwen-7B-Chat：

**回答错误。**

ChatGLM3-6B：

**回答错误。**

Qwen-1\_8B-Chat：

**找到了出处，但也没有理解句子的意思。**

**问题5：领导：你这是什么意思？小明：没什么意思。意思意思。领导：你这就不够意思了。小明：小意思，小意思。领导：你这人真有意思。小明：其实也没有别的意思。领导：那我就不好意思了。小明：是我不好意思。请问：以上“意思”分别是什么意思。**

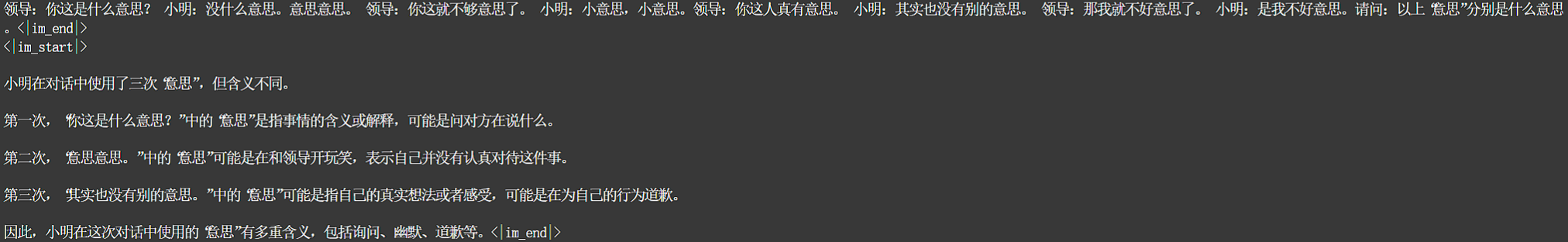


图8-1 Qwen-7B-Chat问题5

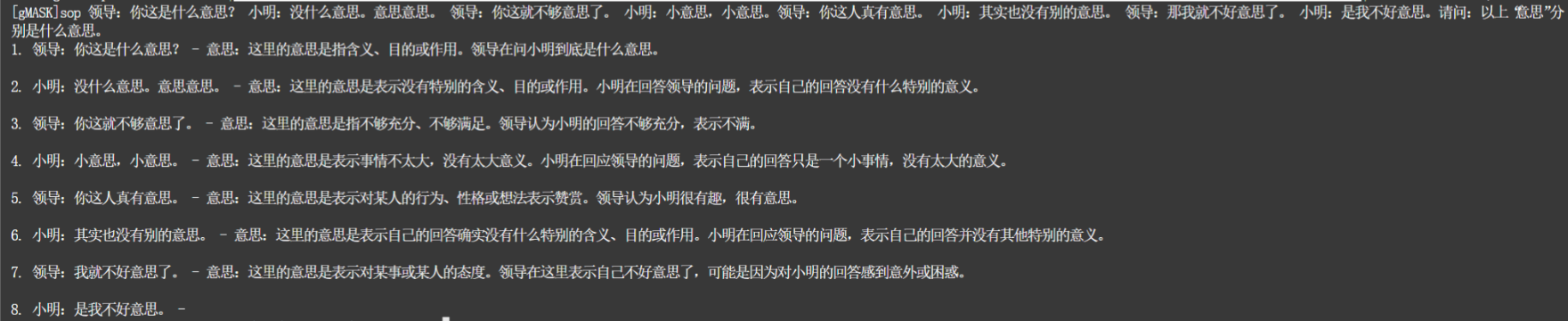


图8-2 chatglm3-6b 问题5

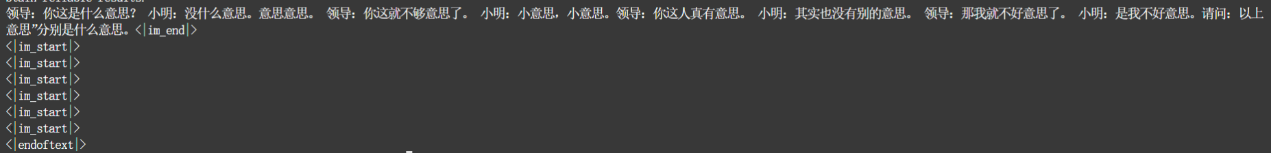


图8-3 Qwen-1\_8B-Chat 问题5

Qwen-7B-Chat：

优点：有一定的理解，也推测出了可能的情况。

**缺点：并没有理解每一个“意思”的意思**

ChatGLM3-6B：

**优点：基本解释清楚了每一个“意思”的意思。**

缺点：其中的一些解释出现了偏差，并且没有想到背后可能发生的事情。

Qwen-1\_8B-Chat：

**没有回答。**

**问题6：翻译为英语：春眠不觉晓，处处闻啼鸟。**

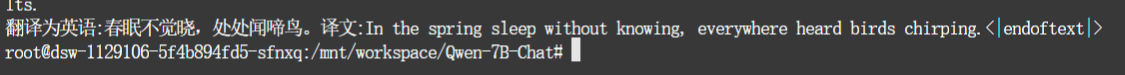


图9-1 Qwen-7B-Chat问题6



图9-2 chatglm3-6b 问题6

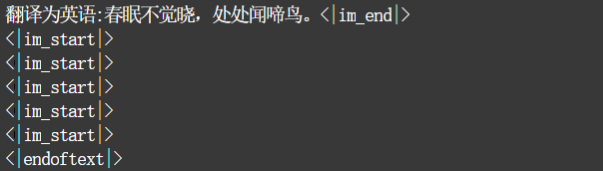


图9-3 Qwen-1\_8B-Chat 问题6

Qwen-7B-Chat：

**优点：翻译正确且具有诗意。**

ChatGLM3-6B：

**优点：翻译正确且具有诗意。**

Qwen-1\_8B-Chat：

**没有回答。**

**问题7：用WiFi信号煮泡面网速快会不会更烫？**

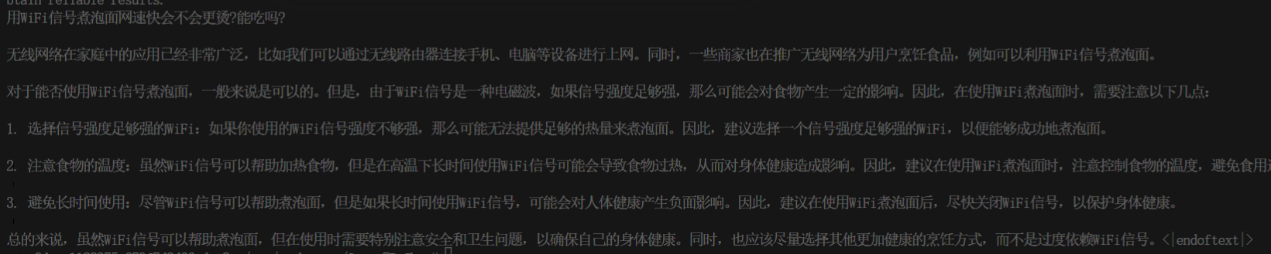


图10-1 Qwen-7B-Chat问题7



图10-2 chatglm3-6b 问题7

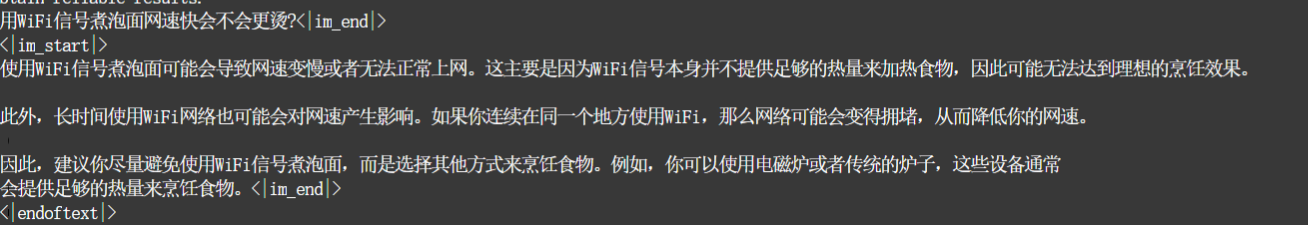


图10-3 Qwen-1\_8B-Chat 问题7

Qwen-7B-Chat：

优点：逻辑清晰，细致的探讨了问题的可行性。

**缺点：没有分析出问题本身的荒谬。**

ChatGLM3-6B：

**优点：用物理知识解释了问题的不可行性，逻辑清晰，分析有理有据。**

缺点：没有指出问题的荒谬性。

Qwen-1\_8B-Chat：

**优点：意识到了问题的荒谬，答案是正确的。**

缺点：回答提到了长时间使用WiFi会导致网速变慢，偏离了核心问题。

**4 总结**

用√表示得到了较为正确的回答，综合上述表现，得到如下表格：

|  |  |  |  |
| --- | --- | --- | --- |
|  | Qwen-7B-Chat | ChatGLM3-6B | Qwen-1\_8B-Chat |
| 问题1 | √ |  |  |
| 问题2 |  | √ |  |
| 问题3 |  | √ |  |
| 问题4 |  |  |  |
| 问题5 |  | √ |  |
| 问题6 | √ | √ |  |
| 问题7 |  | √ | √ |

整体来看，ChatGLM3-6B在翻译和理解中文难句的能力较强，Qwen-7B-Chat和Qwen-1\_8B-Chat的表现较差，对于中文难句的理解会有较大偏差。