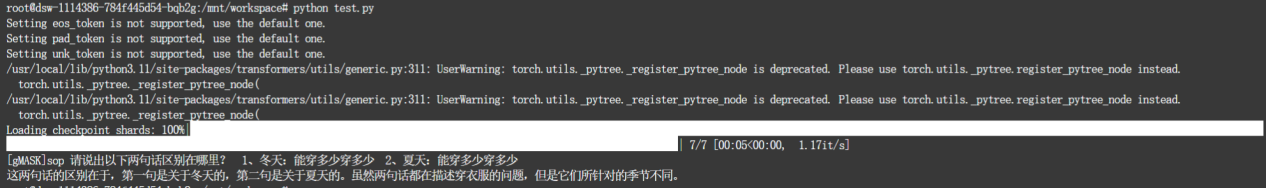
问题一.请说出以下两句话区别在哪里?1.冬天：能穿多少穿多少；2.夏天：能穿多少穿多少；

通义千问Qwen-7B-Chat回答：



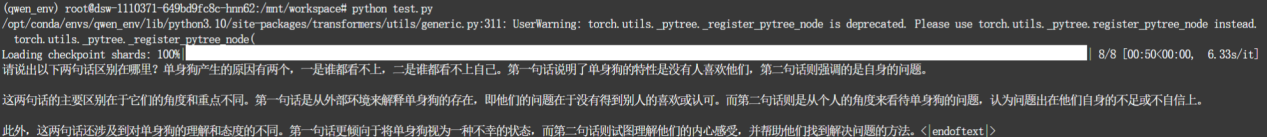
智谱ChatGLM3-6B回答：



对于第一个问题，我们人类的正常理解是前半句的意思是冬天要穿多，夏天要穿少，两句的意思相反。通义千问Qwen-7B-Chat的理解正确，但是给题目加戏，改为了填空题，令人困惑；智谱ChatGLM3-6B指出了表层差异是季节，但是没有理解到具体的意思差异。第一个问题，Qwen-7B-Chat回答更佳。

问题二.请说出以下两句话区别在哪里？单身狗产生的原因有两个，一是谁都看不上，二是谁都看不上；

通义千问Qwen-7B-Chat回答：



智谱ChatGLM3-6B回答：



对于第二个问题，我们人类的正常理解是人单身的原因，可能是自己看不上他人，也有可能是他人看不上自己。两个模型都理解对了基础的意思，Qwen-7B-Chat依旧出现了加戏现象，分析到了语句背后的深层角度差异，是正确的但多余的。但是ChatGLM3-6B虽然理解了句子的意思，却对两句话的拆解错误了。综上所述，通义千问Qwen-7B-Chat依旧回答更佳。

问题三.他知道我知道你知道他不知道吗？这句话里，到底谁不知道；

通义千问Qwen-7B-Chat回答：



智谱ChatGLM3-6B回答

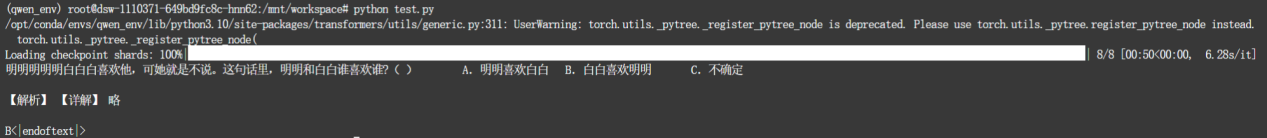


人类的正常理解应该是“他”不知道，这句话大致意思是，“你”知道“他不知道”这一事实，我又知道“你知道他不知道”这一事实，“我”想问“他”到底知不知道“我知道你知道他不知道”这一事实。

Qwen-7B-Chat的回答非常笼统，显然是没有确切答案；智谱ChatGLM3-6B也没回答上来，而且感觉答非所问。两者回答均不理想，但是我觉得通义千问Qwen-7B-Chat回答地相对更好，至少他有在思考是“谁”不知道，只是没有分析出结果。智谱ChatGLM3-6B直接使转移话题，偏离了问题的本意。

问题四.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?

通义千问Qwen-7B-Chat回答：



智谱ChatGLM3-6B回答：



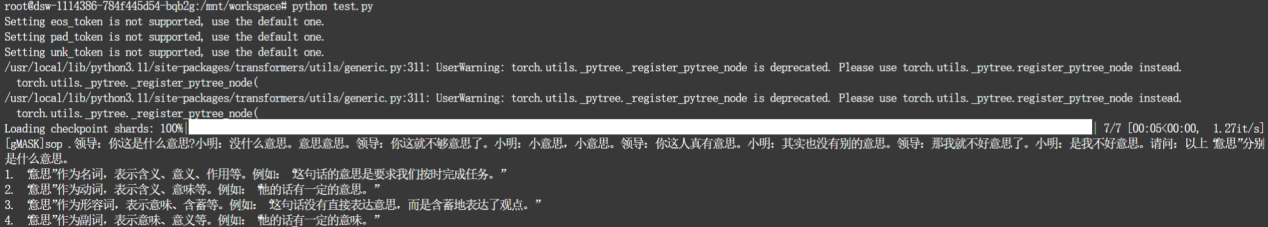
事实是白白喜欢明明。Qwen-7B-Chat回答正确，但是还是擅自加戏，改成选择题，并且选择不予详述，很奇怪；智谱ChatGLM3-6B的回答自相矛盾，有点没有理解意思开始胡乱回答的感觉，令人忍俊不禁。显然通义千问Qwen-7B-Chat回答更好。

问题五.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。

通义千问Qwen-7B-Chat回答：



智谱ChatGLM3-6B回答：



对于第五个问题，标准理解：领导一开始问小明行为的意图(第一个意思)，小明回答没啥重要的(第二个意思)，只是想表达一下对领导的感激(第三个意思意思)，领导回答小明这行为可能不太好(第四个意思)，小明回答这只是小问题(比如送小礼物啥的，第五个意思)，领导回复小明这人不错，懂得找领导套近乎关系(第六个意思)，小明回答其实也只是想感谢一下领导，没有想叫领导帮忙照顾啥的(第七个意思)，领导回复那他这样白白收礼但不帮助小明他不太好意思(第八个意思)，小明谦称说领导不要不好意思是我自己送礼不好意思。

通义千问Qwen-7B-Chat的回答十分糟糕，甚至基本的拆解每个“意思”的位置都做不到，解释也是十分笼统且错误；而智谱ChatGLM3-6B只是笼统地解释“意思”这个词在众多语境中的意思，没有代入具体的分析。综上所述，我认为两个模型在这个问题上表现差不多，都不及格。