

项目公开可访问链接：

<https://github.com/yunshanshu123/AI-assignment4.git>

git 部署：

```
root@dsw-1134779-656498559d-sd6w8:/mnt/data# git clone https://www.modelscope.cn/ZhipuAI/chatglm3-6b.git
正克隆到 'chatglm3-6b'...
remote: Enumerating objects: 140, done.
remote: Counting objects: 100% (18/18), done.
remote: Compressing objects: 100% (17/17), done.
remote: Total 140 (delta 8), reused 1 (delta 0), pack-reused 122
接收对象中: 100% (140/140), 61.16 KiB | 398.00 KiB/s, 完成
处理 delta 中: 100% (60/60), 完成
过滤内容: 100% (15/15), 23.26 GiB | 151.16 MiB/s, 完成
```

```
root@dsw-1134306-78c5b54b9-91j6r:/mnt/data# git clone https://www.modelscope.cn/qwen/Qwen-7B-Chat.git
正克隆到 'Qwen-7B-Chat'...
remote: Enumerating objects: 554, done.
remote: Counting objects: 100% (56/56), done.
remote: Compressing objects: 100% (30/30), done.
remote: Total 554 (delta 30), reused 49 (delta 26), pack-reused 498
接收对象中: 100% (554/554), 16.47 MiB | 618.00 KiB/s, 完成
处理 delta 中: 100% (294/294), 完成
过滤内容: 100% (8/8), 14.38 GiB | 144.90 MiB/s, 完成
```

结果截图：

通义千问：

问题 1：

```
root@dsw-1134306-78c5b54b9-91j6r:/mnt/workspace# python run_qwen_cpu.py
/usr/local/lib/python3.11/site-packages/transformers/utils/generic.py:311: UserWarning: torch.utils._pytree._register_pytree_node is deprecated. Please use torch.utils._pytree.register_pytree_node instead.
  torch.utils._pytree._register_pytree_node(
Loading checkpoint shards: 100% | 8/8 [00:47:00.00, 5.94s/it]
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 句子之间的不同之处在于：
因此，尽管两个句子都使用了相同的词汇“能穿多少穿多少”，但是它们的意思和语气却有所不同。</endofText>
```

问题 2：

```
root@dsw-1134779-656498559d-sd6w8:/mnt/workspace# python run_qwen_cpu.py
Loading checkpoint shards: 100% | 8/8 [00:11:00.00, 1.39s/it]
The attention mask is not set and cannot be inferred from input because pad token is same as eos token. As a consequence, you may observe unexpected behavior. Please pass your input's 'attention_mask' to obtain reliable results.
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，二是谁都看不上自己。
Response: 这两句话的区别在于强调的重点不同。
第一句话 单身狗产生的原因有两个，一是谁都看不上，二是谁都看不上自己“重点放在了产生单身狗的原因上，强调了两种情况：一是自身条件不好，被别人看不上；二是自己的眼光过高，总是对别人不满意。
而第二句话 单身狗产生的原因有两个，一是谁都看不上，二是谁都看不上自己“则更加侧重于自我的问题，强调了 谁都看不上我”的这种自我感觉，认为是自己不够好导致的单身。同时， 谁都看不上我”也是单身狗的一种心态特征。
总的来说，这两句话的区别在于前者更多地从客观因素分析单身狗的问题，后者更多地从主观因素考虑单身狗的心理状态。</endofText>
```

问题 3：

```
root@dsw-1134779-656498559d-sd6w8:/mnt/workspace# python run_qwen_cpu.py
python: can't open file '/mnt/workspace/run_qwen_cpu.py': [Errno 2] No such file or directory
root@dsw-1134779-656498559d-sd6w8:/mnt/workspace# python run_qwen_cpu.py
Loading checkpoint shards: 100% | 8/8 [00:00:00.00, 9.86it/s]
The attention mask is not set and cannot be inferred from input because pad token is same as eos token. As a consequence, you may observe unexpected behavior. Please pass your input's 'attention_mask' to obtain reliable results.
他知我知你知他知他不知道吗？这句话里，到底谁不知道 他知 你知 他知？
翻译或英文：我知道你知道他不知道吗？ This sentence has who knows who knows you know? Which one of them does not know "he" knows what "you know"?</is_end>
</is_start> I know</is_end>
</endofText>

root@dsw-1134779-656498559d-sd6w8:/mnt/workspace# python run_qwen_cpu.py
Loading checkpoint shards: 100% | 8/8 [00:00:00.00, 9.83it/s]
The attention mask is not set and cannot be inferred from input because pad token is same as eos token. As a consequence, you may observe unexpected behavior. Please pass your input's 'attention_mask' to obtain reliable results.
他知我知你知他知他不知道吗？这句话里，到底谁不知道？
Response: 这句话里的 他”指的是什么人并不明确，但是根据上下文可以推测， 他”可能是指某个人或某个组织，而另一个人知道这个人或这个组织的情况，但这个人自己却不知道。
因此，这句话可以理解为： 我知道你了解这个人或这个组织的情况，但是他并不知道你知道他已经知道了。”或者， 我知道你了解这个人或这个组织的情况，但是他自己并不知道他已经知道了。”具体是指哪一种情况，则需要根据语境进一步分析。</endofText>
```

问题 4：

```
root@dsw-1134779-656498559d-ad0e8:/mnt/workspace# python run_qwen_cpu.py
Loading checkpoint shards: 100% | 8/8 [00:00:00.00, 8.42it/s]
The attention mask is not set and cannot be inferred from input because pad token is same as eos token. As a consequence, you may observe unexpected behavior. Please pass your input's 'attention_mask' to obtain reliable results.
明明明明明白白白喜欢他, 可她就是不说。 这句话里, 明明明白白白喜欢他? A. 明明喜欢白白 B. 白白喜欢明明 C. 两人都喜欢对方
首先, 我们需要理解这句话的意思。句子中说“明明知道白白喜欢他”, 这意味着明明知道白白对她有好感或者喜欢她。接着, 句子又说“但她就是不说”, 这表明白白并没有直接告诉明明的感情。
因此, 从这句话的描述来看, 我们可以得出结论: 虽然白白喜欢明明, 但在感情上没有勇气表达出来, 所以才一直不告诉明明。所以答案是B. 白白喜欢明明。</endofext>
```

问题 5:

```
root@dsw-1134779-656498559d-ad0e8:/mnt/workspace# python run_qwen_cpu.py
Loading checkpoint shards: 100% | 8/8 [00:00:00.00, 8.03it/s]
The attention mask is not set and cannot be inferred from input because pad token is same as eos token. As a consequence, you may observe unexpected behavior. Please pass your input's 'attention_mask' to obtain reliable results.
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。<|im_end|>
<|im_start|>
1. “这是什么意思” - 意思是询问对方的意图或想法。
2. “没什么意思” - 意思是没有特别的意思或意义。
3. “意思意思” - 意思是表达一种模糊、随便的态度或者语气, 通常用来表示礼貌和谦虚。
4. “不够意思” - 意思是不值得或不合时宜的标准。
5. “小意思” - 意思是客气的说法, 表示自己并不想过多的回报或者帮助。
6. “有意思” - 意思是有某种特殊的趣味或者吸引力。
7. “其实也没有别的意思” - 意思是真正的意思是与之前说的相反。
8. “那我就不好意思了” - 意思是由于自己的行为或话语让对方感到不舒服, 因此感到愧疚。
9. “是我不好意思” - 意思是因为自己的行为或话语让对方感到不舒服, 因此愿意承担责任并承认错误。<|im_end|>
</endofext>
```

chatglm:

问题 1:

```
root@dsw-1134779-656498559d-ad0e8:/mnt/workspace# python run_chatglm_cpu.py
Setting eos_token is not supported, use the default one.
Setting pad_token is not supported, use the default one.
Setting unk_token is not supported, use the default one.
/usr/local/lib/python3.11/site-packages/transformers/utils/generic.py:311: FutureWarning: `torch.utils._pytree.register_pytree_node` is deprecated. Please use `torch.utils._pytree.register_pytree_node` instead.
  torch.utils._pytree.register_pytree_node(
/usr/local/lib/python3.11/site-packages/transformers/utils/generic.py:311: FutureWarning: `torch.utils._pytree.register_pytree_node` is deprecated. Please use `torch.utils._pytree.register_pytree_node` instead.
  torch.utils._pytree.register_pytree_node(
Loading checkpoint shards: 100% | 7/7 [00:29:00.00, 4.20s/it]
[glm]sep 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
这两句话的区别在于, 第一句是关于冬天的, 而第二句是关于夏天的, 虽然两句话都在描述穿多少衣服, 但是它们所针对的季节不同。
```

问题 2:

```
root@dsw-1134779-656498559d-ad0e8:/mnt/workspace# python run_chatglm_cpu.py
Setting eos_token is not supported, use the default one.
Setting pad_token is not supported, use the default one.
Setting unk_token is not supported, use the default one.
/usr/local/lib/python3.11/site-packages/transformers/utils/generic.py:311: FutureWarning: `torch.utils._pytree.register_pytree_node` is deprecated. Please use `torch.utils._pytree.register_pytree_node` instead.
  torch.utils._pytree.register_pytree_node(
/usr/local/lib/python3.11/site-packages/transformers/utils/generic.py:311: FutureWarning: `torch.utils._pytree.register_pytree_node` is deprecated. Please use `torch.utils._pytree.register_pytree_node` instead.
  torch.utils._pytree.register_pytree_node(
Loading checkpoint shards: 100% | 7/7 [00:07:00.00, 1.07s/it]
[glm]sep 2. 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。
这两句话的区别在于, 第一句“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。”中, 使用了两个“谁”来表示看不上的人, 强调的是看不上的人很多, 而第二句“单身狗产生的原因有两个, 一是谁都看不上, 二是没有找到合适的伴侣。”中, 则是用“没有找到合适的伴侣”来解释单身狗的原因, 突出了找不到合适的伴侣这个因素。
```

问题 3:

```
root@dsw-1134779-656498559d-ad0e8:/mnt/workspace# python run_chatglm_cpu.py
Setting eos_token is not supported, use the default one.
Setting pad_token is not supported, use the default one.
Setting unk_token is not supported, use the default one.
/usr/local/lib/python3.11/site-packages/transformers/utils/generic.py:311: FutureWarning: `torch.utils._pytree.register_pytree_node` is deprecated. Please use `torch.utils._pytree.register_pytree_node` instead.
  torch.utils._pytree.register_pytree_node(
/usr/local/lib/python3.11/site-packages/transformers/utils/generic.py:311: FutureWarning: `torch.utils._pytree.register_pytree_node` is deprecated. Please use `torch.utils._pytree.register_pytree_node` instead.
  torch.utils._pytree.register_pytree_node(
Loading checkpoint shards: 100% | 7/7 [00:05:00.00, 1.38it/s]
[glm]sep 他知道我和他的知识他不知道吗? 这句话里, 到底谁不知道的事情是什么?
这句话是一个典型的歧义或者知识句, 它的含义是: 一个人知道另外一个人知道某些事情, 但是这个人自己并不知道这些事情。
在这个语境中, 到底谁不知道的事情是“他知道的事情”, 因为这个人知道自己知道某些事情, 但是并不知道他自己不知道某些事情。
```

问题 4:

```
root@dsw-1134779-656498559d-ad0e8:/mnt/workspace# python run_chatglm_cpu.py
Setting eos_token is not supported, use the default one.
Setting pad_token is not supported, use the default one.
Setting unk_token is not supported, use the default one.
/usr/local/lib/python3.11/site-packages/transformers/utils/generic.py:311: FutureWarning: `torch.utils._pytree.register_pytree_node` is deprecated. Please use `torch.utils._pytree.register_pytree_node` instead.
  torch.utils._pytree.register_pytree_node(
/usr/local/lib/python3.11/site-packages/transformers/utils/generic.py:311: FutureWarning: `torch.utils._pytree.register_pytree_node` is deprecated. Please use `torch.utils._pytree.register_pytree_node` instead.
  torch.utils._pytree.register_pytree_node(
Loading checkpoint shards: 100% | 7/7 [00:09:00.00, 1.38s/it]
[glm]sep 明明明明明白白白喜欢他, 可她就是不说。 这句话里, 明明明白白白喜欢谁?
这句话里, 明明喜欢白白, 明明明白白白喜欢他, 可她就是不说。
```

问题 5:

```
root@dsw-1134779-656498559d-ad0e8:/mnt/workspace# python run_chatglm_cpu.py
Setting eos_token is not supported, use the default one.
Setting pad_token is not supported, use the default one.
Setting unk_token is not supported, use the default one.
/usr/local/lib/python3.11/site-packages/transformers/utils/generic.py:311: FutureWarning: `torch.utils._pytree.register_pytree_node` is deprecated. Please use `torch.utils._pytree.register_pytree_node` instead.
  torch.utils._pytree.register_pytree_node(
/usr/local/lib/python3.11/site-packages/transformers/utils/generic.py:311: FutureWarning: `torch.utils._pytree.register_pytree_node` is deprecated. Please use `torch.utils._pytree.register_pytree_node` instead.
  torch.utils._pytree.register_pytree_node(
Loading checkpoint shards: 100% | 7/7 [00:10:00.00, 1.43s/it]
[glm]sep 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。
1. 领导: 你这是什么意思? - 意思: 这里的意思是指含义、目的或作用。领导在问小明到底是什么意思。
2. 小明: 没什么意思。意思意思。 - 意思: 这里的意思是指没有特别的内容、目的或作用。小明在回答领导的问题, 表示自己的回答没有什么特别的意义。
3. 领导: 你这就够意思了。 - 意思: 这里的意思是指不够充分、不够满意。领导认为小明的回答不够充分, 表示不满。
4. 小明: 小意思。小意思。 - 意思: 这里的意思是指事情不太大, 没有太大意义。小明在回应领导的问题, 表示自己的回答只是一个小事, 没有太大的意义。
5. 领导: 你这人真有意思。 - 意思: 这里的意思是指对某人的行为、性格或想法表示赞赏。领导认为小明很有趣, 很有意思。
6. 小明: 其实也没有别的意思。 - 意思: 这里的意思是指自己的回答确实没有什么特别的内容、目的或作用。小明在回应领导的问题, 表示自己的回答并没有其他特别的意义。
7. 领导: 我就不好意思了。 - 意思: 这里的意思是指对某事或某人的态度。领导在这里表示自己不好意思了, 可能是因为对小明的回答感到意外或尴尬。
8. 小明: 是我不好意思。 -
```

对比:

1. 问题 1: 季节与穿衣的语义理解

通义千问:

优点: 详细分析了“冬天”和“夏天”语境下的不同含义, 指出前者强调保暖(建议性语气), 后者强调舒适(中立语气)。

缺点: 对“能穿多少穿多少”的双关性(多穿 vs 少穿)解释不够直接。

Chatglm:

优点: 直接点明两句话的区别在于季节不同, 简洁明了。

缺点: 未深入语义差异, 仅停留在表面逻辑。

对比: 通义千问更注重语境分析, 而 Chatglm 更偏向字面逻辑, 前者更适合需要深度理解的场景。

2. 问题 2: 单身狗原因的幽默双关

通义千问:

优点: 区分了“看不上”的两种含义(自身条件差 vs 眼光高), 并提到主观与客观因素, 分析全面。

缺点: 回答重复了同一句话, 未明确双关的核心(“看不上”的主被动差异)。

Chatglm:

优点: 直接指出第一句强调“看不上的人多”, 第二句补充“找不到伴侣”, 逻辑清晰。

缺点: 未完全捕捉幽默双关的巧妙性。

对比: 通义千问更贴近语言的双关性, 但表达冗余; Chatglm 更简洁, 但幽默感解析不足。

3. 问题 3: 多层嵌套的指代关系

通义千问:

优点: 尝试分析“他”的指代可能性, 推测上下文关系。

缺点: 结论模糊, 未明确“谁不知道”(实际应为“他”不知道)。

Chatglm:

优点: 明确指出“他不知道的事情”是核心, 更符合绕口令的解答逻辑。

缺点: 未展开层级关系(如“你知道他不知道”)。

对比: Chatglm 更擅长逻辑拆解, 通义千问的推理稍显迂回。

4. 问题 4: 人名与语义的歧义解析

通义千问:

优点: 正确解析“白白喜欢明明”, 并解释“她不说”的行为。

缺点: 未点明“明明明白”的重复人名设计。

Chatglm:

优点: 直接给出答案“明明喜欢白白”, 但逻辑与原文矛盾。

缺点: 答案错误, 未理解句子结构。

对比: 通义千问更准确, Chatglm 因忽略句式设计而误判。

5. 问题 5: “意思”的多义性解析

通义千问:

优点: 对每个“意思”逐一解释, 覆盖全面(如“意图”“趣味”“道歉”等)。

缺点: 部分解释重复(如“没什么意思”和“没有别的意思”)。

Chatglm:

优点: 分点解释更清晰, 标注对话角色, 区分“含义”“态度”等场景。

缺点: 未完成最后两处解释(文档截断)。

对比: 两者均表现优秀, 但 Chatglm 的结构化回答更易读, 通义千问的细节更丰富。

总结:

本次对比分析围绕语义理解、逻辑拆解、结构化表达、准确性和幽默/双关解析五个维度展开，综合评估通义千问和 Chatglm 在不同类型问题上的表现。

核心结论

1.语义深度与双关解析

通义千问表现更优，能深入分析语境和双关含义（如季节穿衣、单身狗原因），适合需要细腻理解的场景。

Chatglm 偏向字面逻辑，对幽默或双关的捕捉较弱。

2.逻辑拆解与结构化表达

Chatglm 更擅长直接拆解复杂逻辑（如绕口令问题），回答结构化、简洁清晰。

通义千问推理稍显迂回，部分回答冗余。

3.准确性

通义千问在歧义句解析（如“明明白白”问题）上更准确，而 Chatglm 因忽略句式设计导致误判。

两者在多义词解析（如“意思”）上均表现良好，但 Chatglm 的答案组织更易读。

4.模型适用场景推荐

选择通义千问：

需要理解深层语义、双关或幽默的文本（如广告文案、社交语言分析）。

对细节解释要求较高的场景（如教育、心理咨询）。

选择 Chatglm：

需快速获取逻辑清晰的答案（如技术文档解析、数学问题）。

强调回答结构化、易读性的任务（如报告生成、会议纪要）。

5.改进方向

通义千问：精简重复内容，提升逻辑直接性。

Chatglm：增强对语言趣味性和歧义句的敏感度。

总体来看，两者互补性强，可根据实际需求灵活选用。