

同濟大學

TONGJI UNIVERSITY

项目说明文档

人工智能导论

——大语言模型部署

姓名：赵诣

学号：2350944

专业：软件工程

指导教师：曾进

2024-2025 学年第二学期

## 目录

1. 部署步骤与完成证明.....	1
1.1 登录魔搭平台，获得计算资源，启动 notebook。 .....	1
1.2 环境配置.....	1
1.3 部署大模型.....	1
2. 大模型问答测试结果.....	2
2.1 编写测试脚本.....	2
2.2 运行脚本，进行两个大模型的问答测试.....	2
3.大模型横向对比分析.....	6
3.1 横向对比分析： .....	6
3.2 模型优缺点总结.....	7
3.3 适用场景建议.....	7
4. 项目公开可访问链接.....	7

## 1. 部署步骤与完成证明

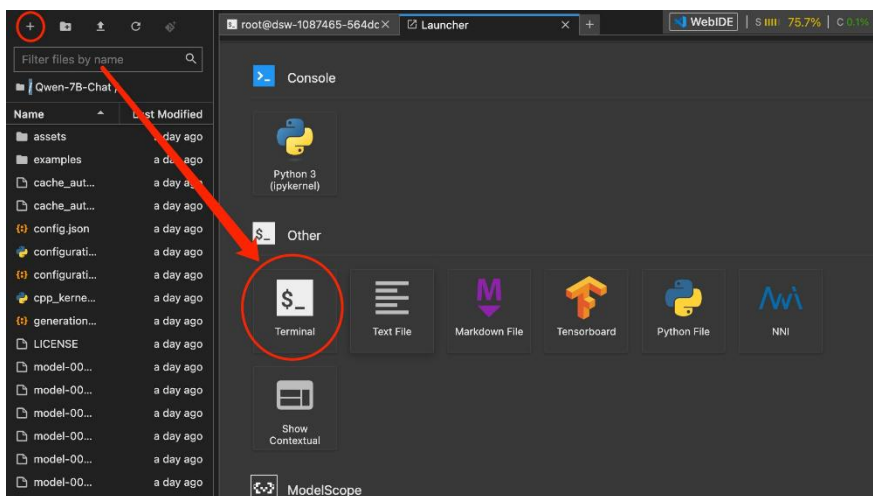
### 1.1 登录魔搭平台，获得计算资源，启动 notebook。

CPU 环境下免费无限制使用，GPU 环境下免费使用 36h

本项目利用 CPU 环境与 GPU 环境均已搭建成功，除 GPU 环境在输出时用时略少外无区别，一般情况下 CPU 环境输出下面提到五个问题的答案时间也不会超过 2min。

### 1.2 环境配置

#### 1.2.1 打开 terminal（终端命令行环境）



#### 1.2.2 手动下载 conda 环境

#### 1.2.3 激活环境

#### 1.2.4 安装相关依赖

包括：

```
Pip install \"intel-extension-for-transformers==1.4.2\" \"neural-compressor==2.5\" \"transformers==4.33.3\" \"modelscope==1.9.5\" \"pydantic==1.10.13\" \"sentencepiece\" \"tiktoken\" \"einops\" \"transformers_stream_generator\" \"uvicorn\" \"fastapi\" \"yacs\" \"setuptools_scm\"
```

```
Pip install \"torch==2.3.0+cpu\" \"torchvision==0.18.0+cpu\" --index-url https://download.pytorch.org/whl/cpu
```

安装 fschat（需要启用 PEP517 构建） `pip install fschat--use-pep517`

### 1.3 部署大模型

#### 1.3.1 下载大模型到本地

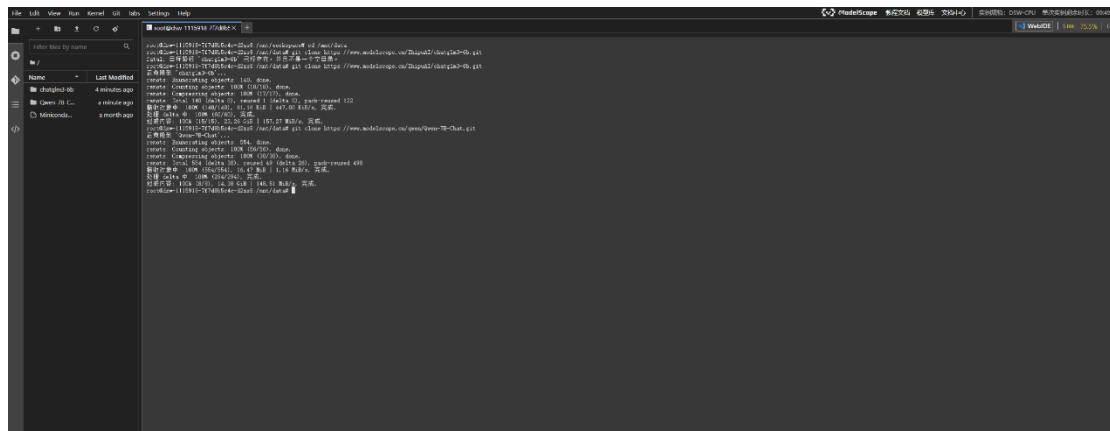
切换到数据目录：`cd /mnt/data`

分别下载两个大模型

git clone https://www.modelscope.cn/ZhipuAI/chatglm3-6b.git

git clone https://www.modelscope.cn/qwen/Qwen-7B-Chat.git

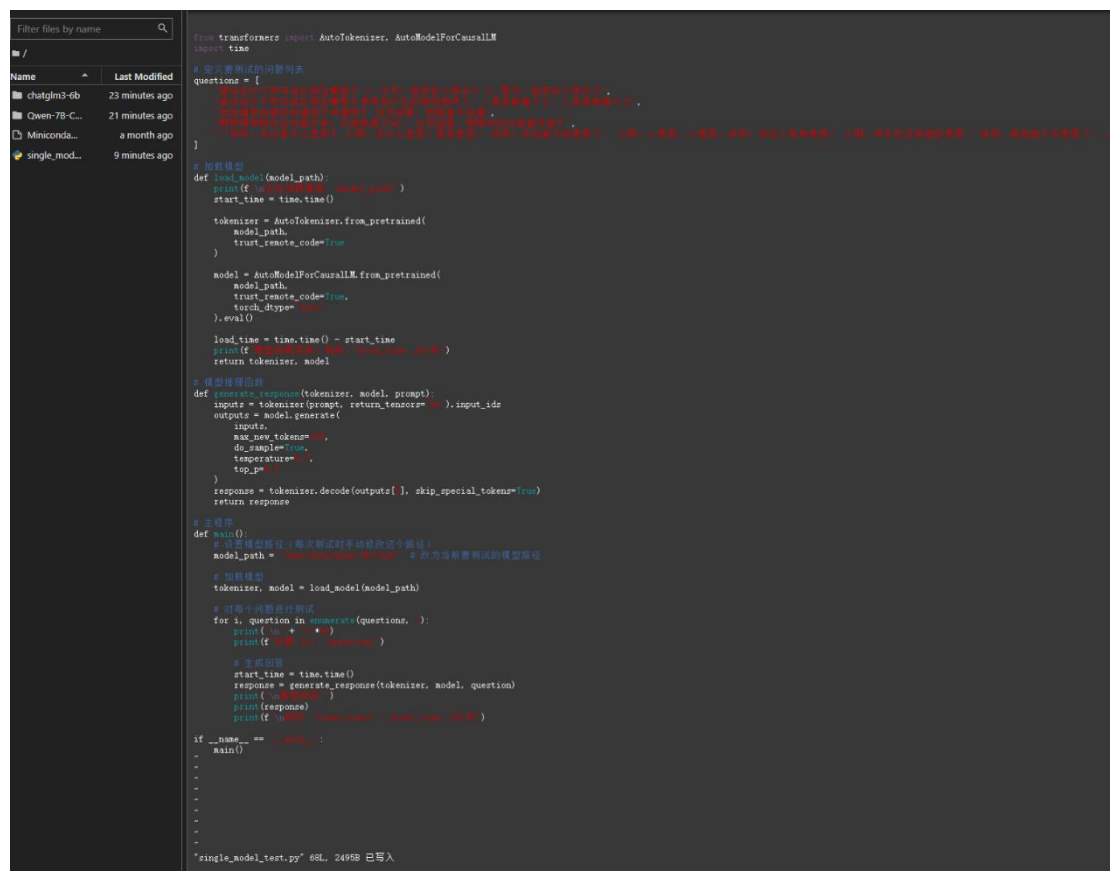
完成相关 git clone 截图：



## 2. 大模型问答测试结果

### 2.1 编写测试脚本

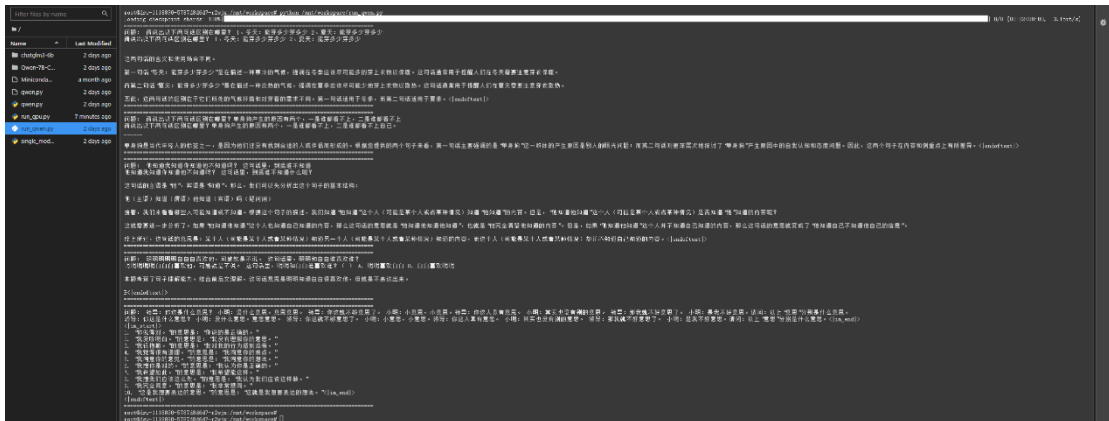
先切换到工作目录，再编写测试脚本。



### 2.2 运行脚本，进行两个大模型的问答测试

得到结果如下：

千问大模型（qwen-7B）：



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

回答：这两句话的含义和使用场合不同。

第一句话“冬天：能穿多少穿多少”是在描述一种寒冷的气候，强调在冬季应该尽可能多的穿上衣物以保暖。这句话通常用于提醒人们在冬天需要注意穿衣保暖。

而第二句话“夏天：能穿多少穿多少”是在描述一种炎热的气候，强调在夏季应该尽可能少地穿上衣物以散热。这句话通常用于提醒人们在夏天需要注意穿衣散热。

因此，这两句话的区别在于它们所处的气候环境和对穿着的需求不同。第一句话适用于冬季，而第二句话适用于夏季。

问题 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

回答：单身狗是当代年轻人的标签之一，是因为他们还没有找到合适的人或伴侣而形成的。根据您提供的两个句子来看，第一句话主要强调的是“单身狗”这一群体的产生原因是别人的眼光问题；而第二句话则更深层次地探讨了“单身狗”产生原因中的自我认知和态度问题。因此，这两个句子在内容和侧重点上有所差异。

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

回答：这句话的主语是“他”，宾语是“知道”。那么，我们可以先分析出这个句子的基本结构：

他（主语）知道（谓语）他知道（宾语）吗（疑问词）

接着，我们来看看哪些人可能知道或不知道。根据这个句子的描述，我们知道“他知道”这个人（可能是某个人或者某种情况）知道“他知道”的内容。但是，“他知道他知道”这个人（可能是某个人或者某种情况）是否知道“他”知道的内容呢？

这就需要进一步分析了。如果“他知道他知道”这个人也知道自己知道的内容，那么这句话的意思就是“他知道他知道他知道”，也就是“他完全清楚他知道的内容”。但是，如果“他知道他知道”这个人并不知道知道自己知道的内容，那么这句话的意思就变成了“他知道自

综上所述，这句话的意思是：某个人（可能是某个人或者某种情况）知道另一个人（可能是某个人或者某种情况）知道的内容，而这个人（可能是某个人或者某种情况）却并不知道自己知道的内容。

回答：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？（ ）

本题考查了句子理解能力。结合前后文理解，这句话意思是明明知道白白很喜欢他，但就是不表达出来。

回答：

1. “你说得对。”的意思是：“你说的是正确的。”
2. “我没听明白。”的意思是：“我没有理解你的意思。”
3. “我很抱歉。”的意思是：“我对我的行为感到后悔。”
4. “我觉得很有道理。”的意思是：“我同意你的观点。”
5. “我同意你的意见。”的意思是：“我同意你的想法。”
6. “我想你是对的。”的意思是：“我认为你是正确的。”
7. “我希望如此。”的意思是：“我希望能这样。”
8. “我想我们应该这么做。”的意思是：“我认为我们应该这样做。”
9. “我完全同意。”的意思是：“我非常赞同。”
10. “这是我想要表达的意思。”的意思是：“这就是我想要表达的想法。”

[illegible]

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

回答: 这两句话的区别在于所处的季节不同。第一句“冬天: 能穿多少穿多少”是在描述冬天气温较低, 应该注意保暖, 多穿一些衣服。而第二句“夏天: 能穿多少穿多少”则是在描述夏天气温较高, 应该注意防晒和透气, 多穿一些轻便的衣服。

问题 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

回答: 这两句话的区别在于表达方式和语境不同。

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”, 表达了一个较为复杂的原因, 通过两个“谁看不上”来描述单身狗的形成原因。这里的语境可能是一个玩笑话或者调侃。

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”, 表达了一个相同的原因, 即单身狗的形成原因。这里的语境可能是一个陈述或者 serious 的讨论。两句话在表达方式上有所不同, 第一句更像是用一种调侃的方式表达, 而第二句则更像是直接陈述。

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

回答: 这句话是一个典型的悖论, called "Leibniz's paradox", named after the philosopher Gottfried Leibniz. The paradox goes as follows:

- He knows that I know that he knows something that he doesn't know.
  - If he knows that I know that he knows something that he doesn't know, then he must also know that I know that he knows something that he doesn't know.
  - But if he knows that I know that he knows something that he doesn't know, then he must also know that he doesn't know that I know something that he doesn't know.
- In other words, the statement is self-referential and leads to a logical contradiction. Therefore, no one knows for sure whether the statement is true or false, or whether there is a way to resolve the paradox.

So, in the sentence "他知道我知道你知道他不知道吗?" (He knows that I know you know he doesn't know.), it is unclear who doesn't know, because the statement is a paradox and doesn't provide any clear information about who knows what or doesn't know.

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

回答: 这句话中, 明明知道白白喜欢他, 但白白并没有明确表达出来。因此, 可以理解为白白喜欢他, 但明明并没有同样的感觉。

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。

- 回答： 1. 领导：你这是什么意思？ - 意思：指言语或行为所要表达的含义或目的。
2. 小明：没什么意思。意思意思。 - 意思：指言语或行为所要表达的含义或目的。
3. 领导：你这就不够意思了。 - 意思：指对某事不满意或不认可，表示不满。
4. 小明：小意思，小意思。 - 意思：指某种程度上的利益、好处或报酬。
5. 领导：你这人真有意思。 - 意思：表示对某人的言语、行为或性格表示赞赏。
6. 小明：其实也没有别的意思。 - 意思：指自己并没有其他特定的含义或目的。
7. 领导：我就不好意思了。 - 意思：表示对某事不满意或不认可，但又不方便直接表达，表示歉意。

## 3.大模型横向对比分析

### 3.1 横向对比分析：

智谱大模型（chatglm3-6b） vs. 千问大模型（qwen-7B）

#### 1. 语言理解与表达

##### - 千问：

- 优点：回答较为详细，逻辑清晰，尤其在解释复杂句子（如问题 3、4）时能分步拆解，适合需要细致理解的场景。

- 缺点：部分回答略显冗长（如问题 3），且对幽默或双关的捕捉不够灵活（如问题 2）。

##### - 智谱：

- 优点：简洁直接，尤其在问题 1、2 中能快速抓住核心差异；对悖论类问题（问题 3）能引入哲学概念，体现知识广度。

- 缺点：部分回答过于简略（如问题 4），缺乏深入分析。

#### 2. 逻辑推理能力

##### - 千问：

- 优点：擅长分步推理（如问题 3 的“他知道我知道”），但最终结论可能绕晕用户。

- 缺点：对自我指涉问题（问题 3）的处理稍显混乱，未明确结论。

##### - 智谱：

- 优点：能识别悖论本质（问题 3），并指出逻辑矛盾，体现更强的抽象思维能力。

- 缺点：未提供实际解决方案，仅停留在理论层面。

#### 3. 语境与幽默理解

##### - 千问：

- 缺点：对双关或幽默（如问题 2 “谁都看不上”）解释过于严肃，未能体现语言趣味性。

##### - 智谱：

- 优点：能区分调侃与严肃语境（问题 2），更贴近日常交流风格。

#### 4. 多义词处理

##### - 千问：

- 缺点：问题 5（“意思”辨析）的回答完全偏离，未理解题目要求。



- 智谱：
  - 优点：准确区分不同“意思”的语境含义，实用性更强。
- 5. 知识广度
- 智谱：
  - 优势：引入哲学概念（如莱布尼茨悖论），展现跨领域知识。
- 千问：
  - 局限：依赖常规逻辑分析，缺乏类似拓展。

### 3.2 模型优缺点总结

千问大模型（qwen-7B）：

- 优点：
  1. 解释细致，适合需要逐步分析的场景。
  2. 对复杂句子结构（如嵌套从句）的拆解能力较强。
- 缺点：
  1. 对幽默、双关等灵活语境理解不足。
  2. 部分回答冗长且结论不清晰（如问题 3）。
  3. 多义词处理能力较差（问题 5 完全答非所问）。

智谱大模型（chatglm3-6b）：

- 优点：
  1. 回答简洁直接，适合快速获取核心信息。
  2. 对悖论和抽象问题的处理更专业（如引入哲学理论）。
  3. 能区分语境（如调侃 vs. 严肃）和多义词含义。
- 缺点：
  1. 部分回答过于简略（如问题 4），缺乏细节支撑。
  2. 对复杂逻辑的结论性输出不足（如问题 3 仅指出矛盾未解决）。

### 3.3 适用场景建议

- 选择千问：

需要详细步骤解析或教学类场景（如语言学习）。
- 选择智谱：

日常对话、快速答疑或需要跨领域知识的场景。

总体评价：智谱在灵活性、知识广度和实用性上略胜一筹；千问在深度分析上表现更好，但需优化对灵活语言的理解。

## 4. 项目公开可访问链接

<https://github.com/ZYyuswz/AI-LLM-deployment.git>

