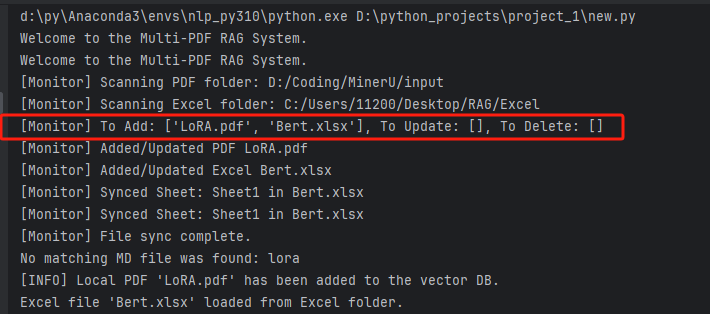
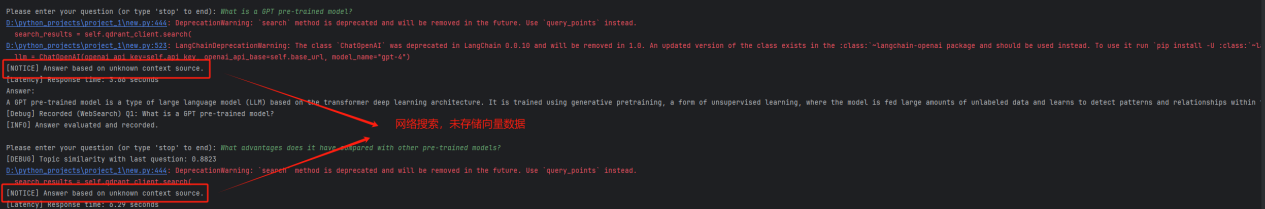
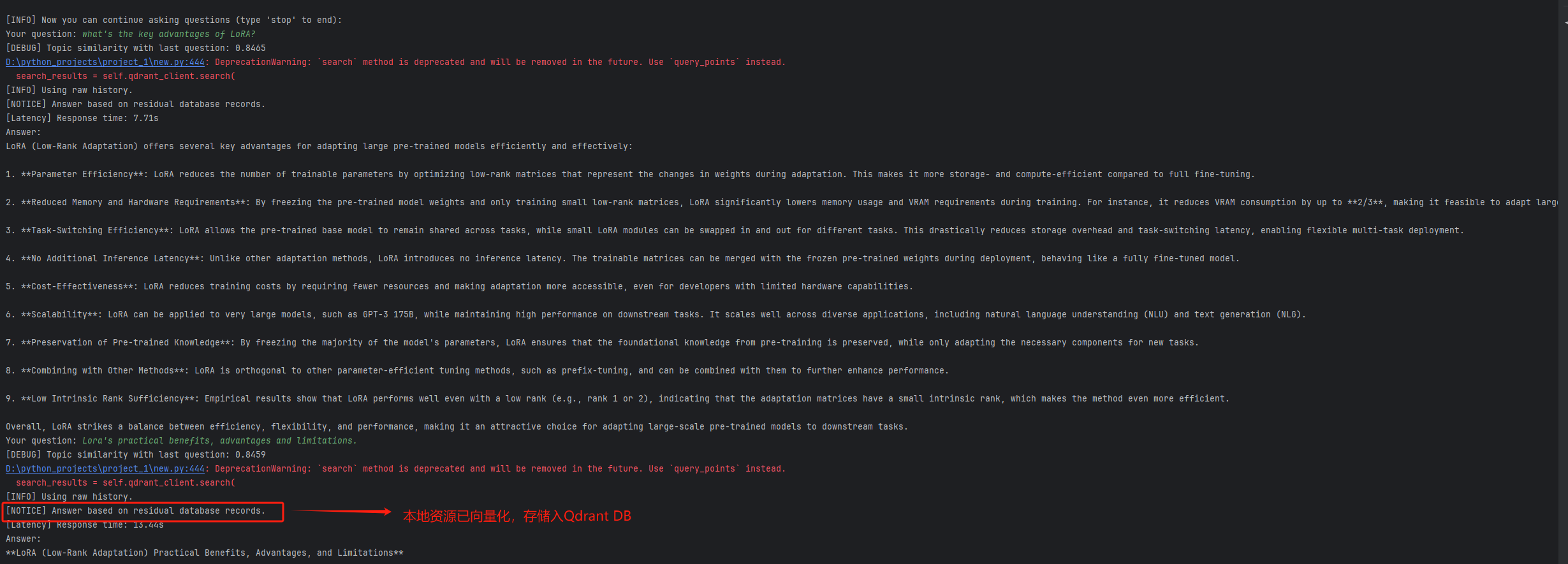
1. **正常运行（MinerU与excel文件夹下都保存有pdf与excel，清除Qdrant内所有历史记录）：**
2. 初始化时，本地资源LoRA.pdf与Bert.xlsx最先进入监管器Add列表。url下载的**Attention is all you need.pdf**也会在**MCP**类加载进监管器Add列表，初次写入processed\_files.json。

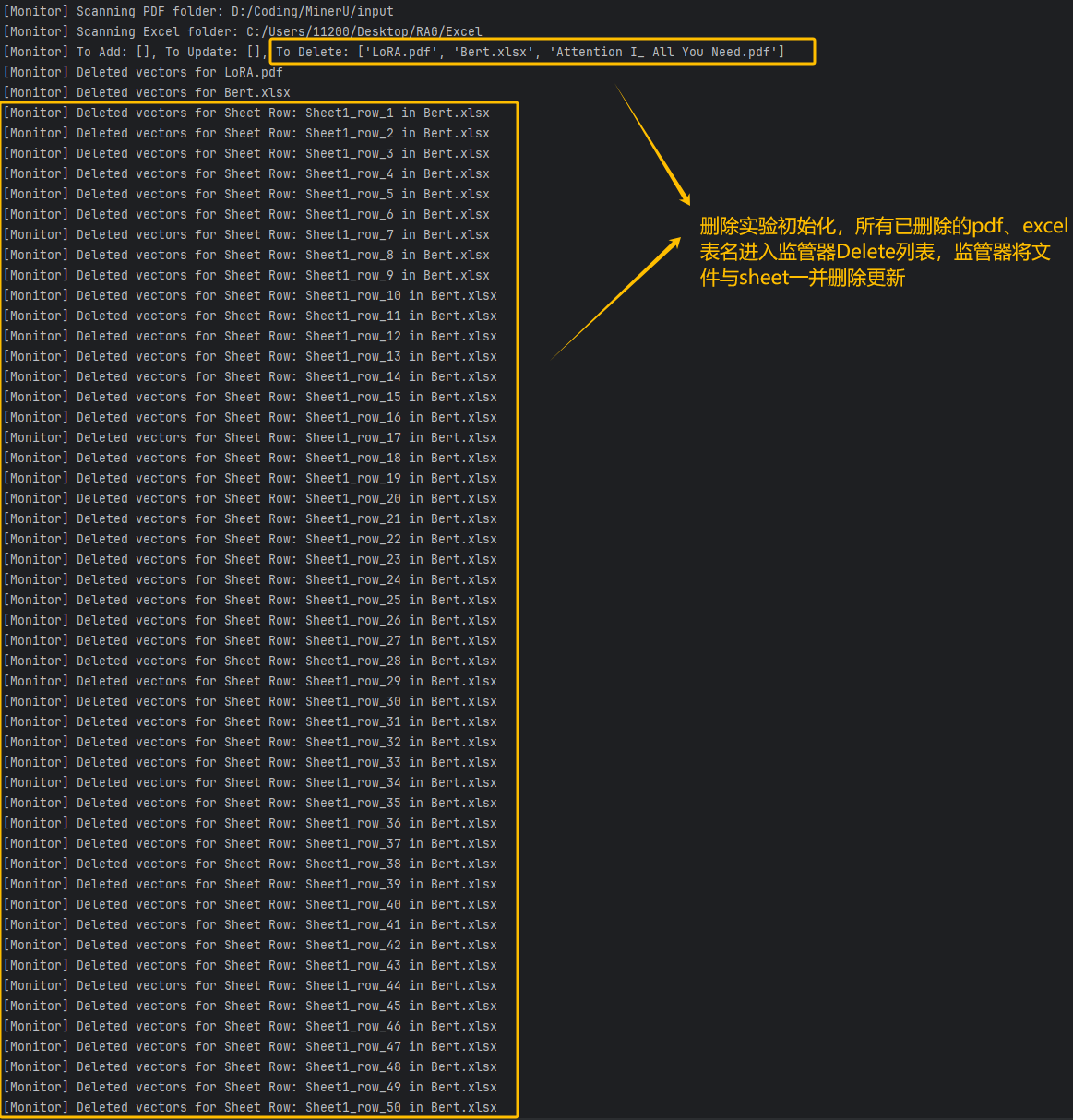
****



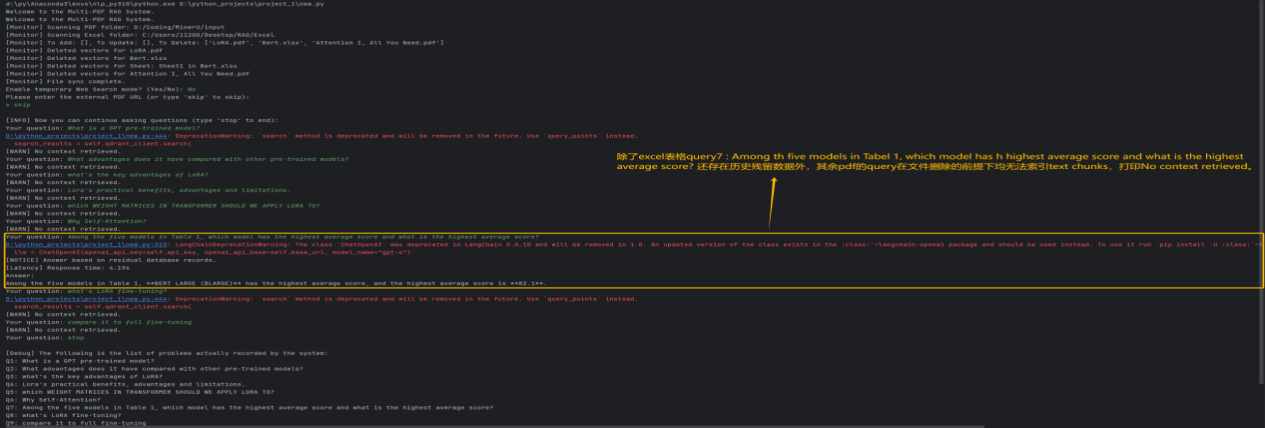
1. 以query 3 : what's the key advantages of LoRA? 为例，LLM面对所有query都能找到text chunks。



1. **删除测试（在第一次正常运行后，将MinerU与excel文件夹下的pdf与excel都清空，停止url下载skip和WebSearch）：**
2. 第2次运行前，清除所有pdf和excel，停止WebSearch和url下载，同步更新processed\_files.json。



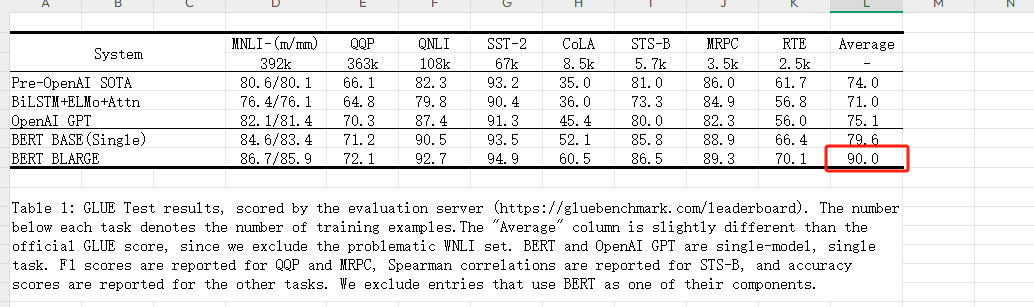
1. 面对Query 7 ，LLM依然能回答，推测Qdrant内仍残留有Sheet级别的历史数据，或者是LLM基于预训练知识进行回答。



1. **修改测试（以Bert.xlsx中shee1页为例，在第一次正常运行后，将Table 1 Bert Blarge的average score从82.1人为修改成90）**
2. 第一次正常运行后query 7 的答案为Bert Blarge和82.1。

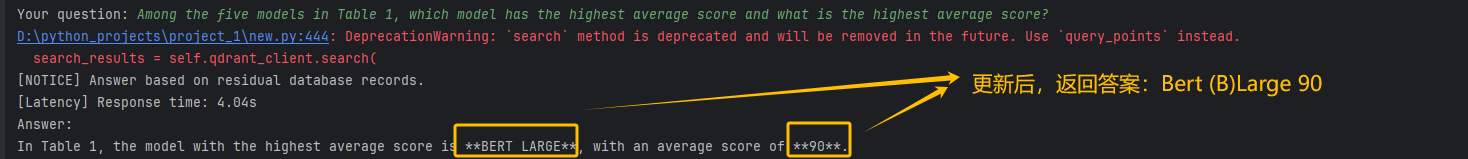


1. 修改Bert.xlsx后，检索精细到以行列为单位，更新processed\_files.json

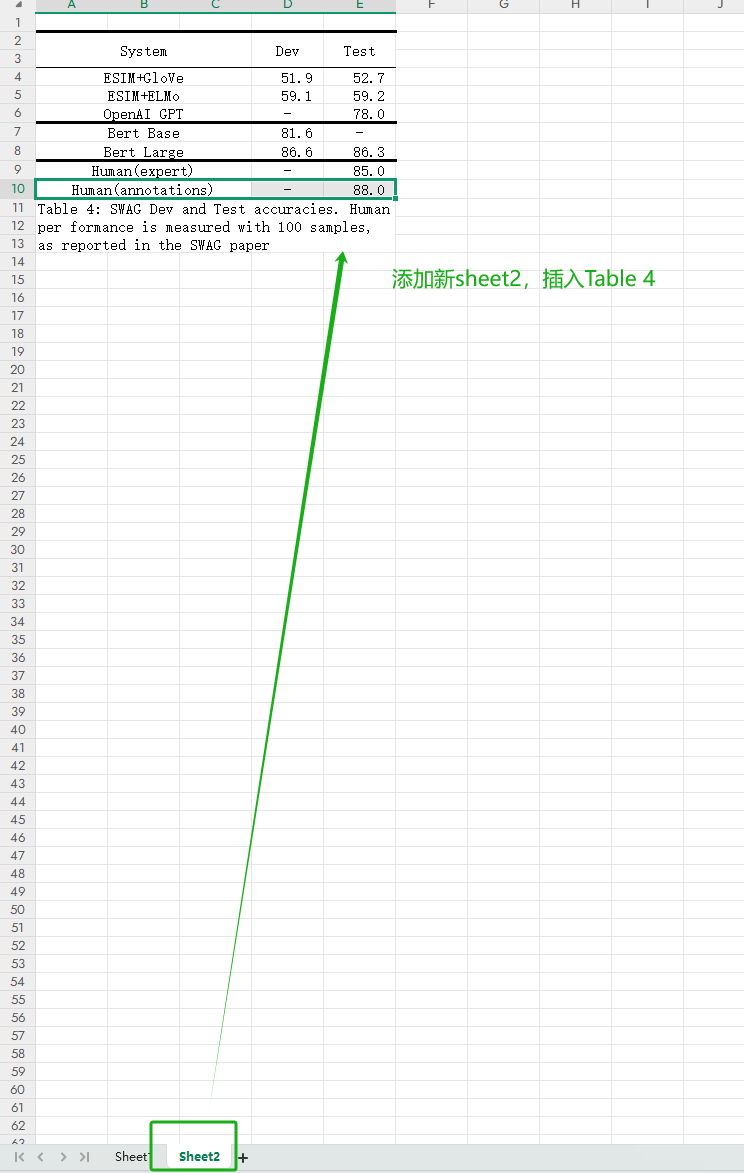


****

1. Qdrant更新后，LLM检索到的答案为Bert Large和90。



1. **新增测试（以Bert.xlsx中新增sheet2页为例，插入Table4）**
2. 文章来源：[BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding](https://aclanthology.org/N19-1423.pdf)



1. 输入新的query 10 : According to Table 4, which model has the highest Dev and Test scores? What are the highest scores of Dev and Test respectively?

正确答案：Bert Large; Dev: 86.6, Test: 88

Fake\_answer: ESIM+ELMo scored the highest on Test, while ESIM+GloVe scored the highest on DEV. The highest score for DEV is 51.9, while that for Test is 52.7.

