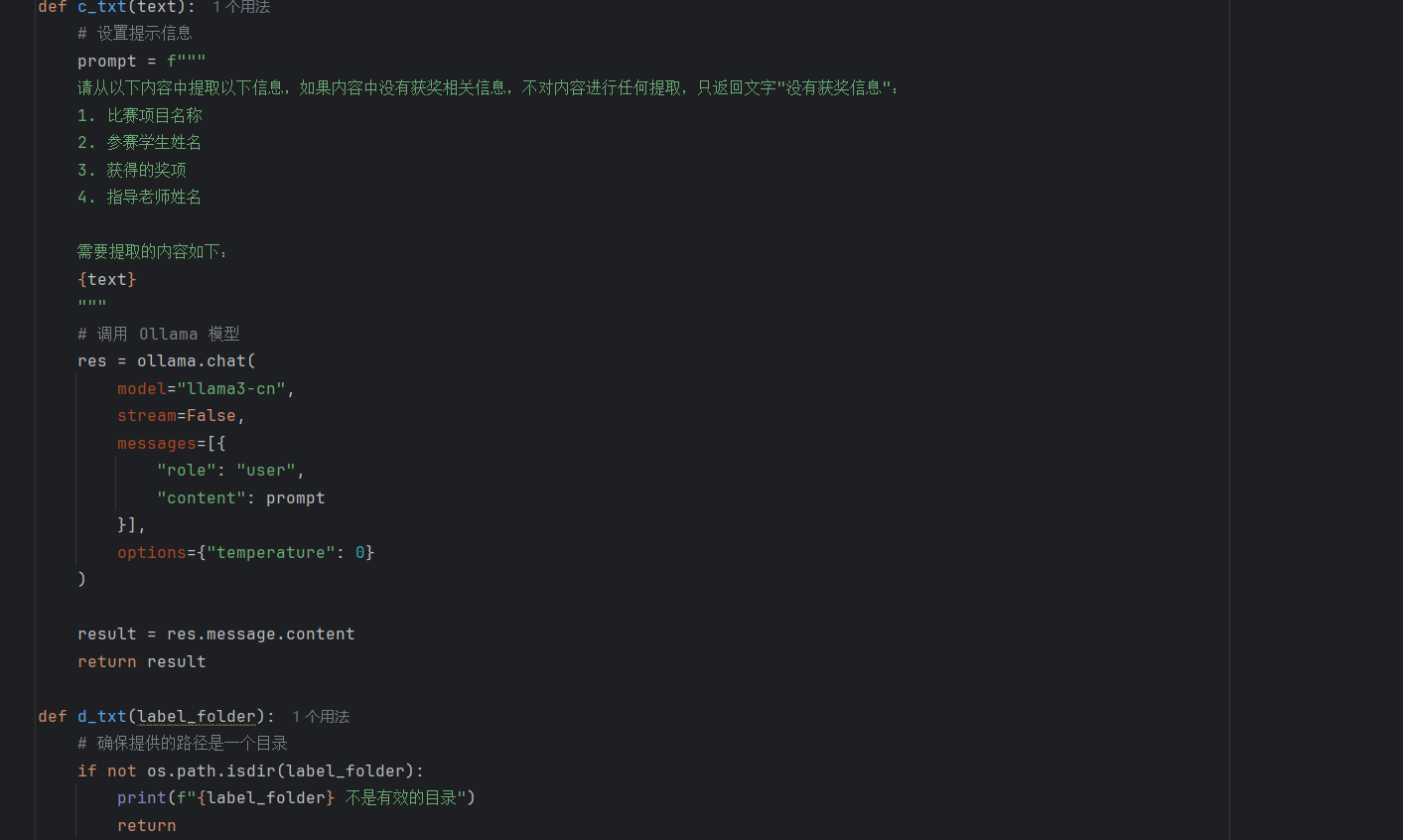
使用ollama对大模型进行本地部署，大型选用llama3-8b的中文调试并量化的版本llama3-8b-chinese-chat-8bit，然后在本地用ollama库实现对信息内容的提取

代码



处理效果

