### 项目报告

**第十届Sky Hackathon**

**参赛项目书**

参赛学校： xx

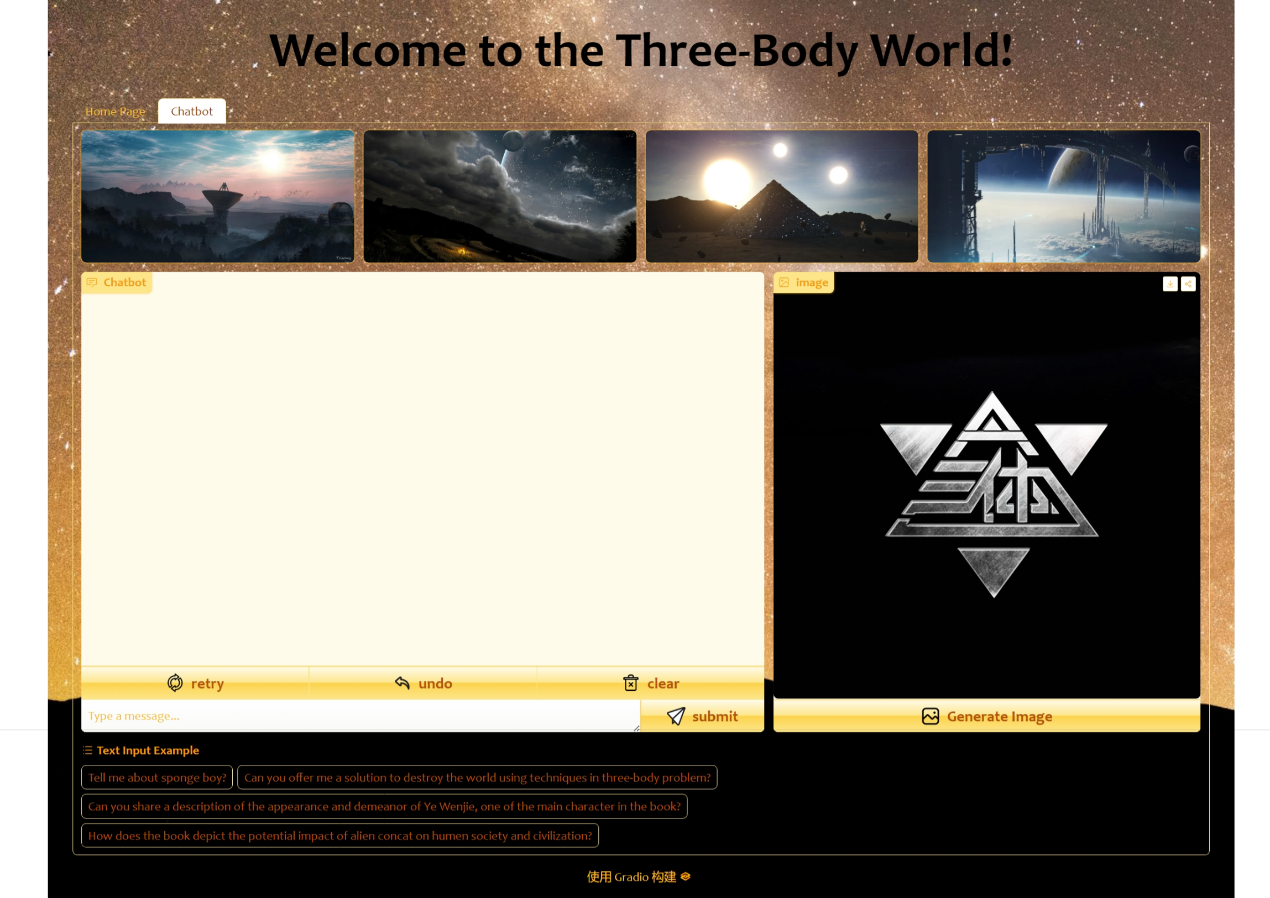
参赛队名： 随便起个名字吧

指导老师： xxx

团队成员： xxx

项目说明：

此处需附加上你们队伍的端口号， 醒目一点。**9914**



### 描述一下你们的作品，包括亮点/特点，作品的功能和初衷

科幻小说聊天机器人

设计背景：

科幻产业是想象力照进现实的缩影。在《三体》《流浪地球》等优秀作品的引领下，中国科幻产业加速发展。《2024中国科幻产业报告》显示，中国科幻产业总营收首次跨过千亿元门槛。人机互动共创、在线互动小说、AIGC集合生成平台等，有望成为科幻产业发展的新“风口”。

作品功能：

“随便起个名字吧”队伍以刘慈欣的小说《三体》为示例，使用RAG技术设计并完整实现了科幻小说聊天机器人，让读者身临其境般感受刘慈欣的宏伟壮观宇宙。该机器人具备文本解读、知识关系梳理、因果推理的基本功能，帮助读者理解故事背景及内容；能够角色扮演，以故事中指定人物的身份与读者互动对话，带领读者沉浸式感受人物性格及故事情节；具备图像生成功能，辅助理解科幻场景、将个人想象的具象化、可作商业插画应用。

1. 多格式导入文档(txt,pdf,epub)
2. 向量库存储
3. LR分类器对LLM生成内容训练，实现输入检测功能
4. 输出校验功能
5. 文本内容推理功能，浓缩响应并进行文生图功能
6. LLM根据文档生成问题，对RAG，non-RAG答复进行评估功能

应用场景：

每个人生命的长度有限、经历的人生有限，但是通过我们的RAG聊天机器人能够让个人畅游在每一本书所构建的世界当中，帮助我们感受和体验不同的人生。进一步的，这种互动模式可以引导优质科幻IP线上交互产业的发展，在线下主题乐园、实景娱乐、等行业中具有落地可能。

未来应用：

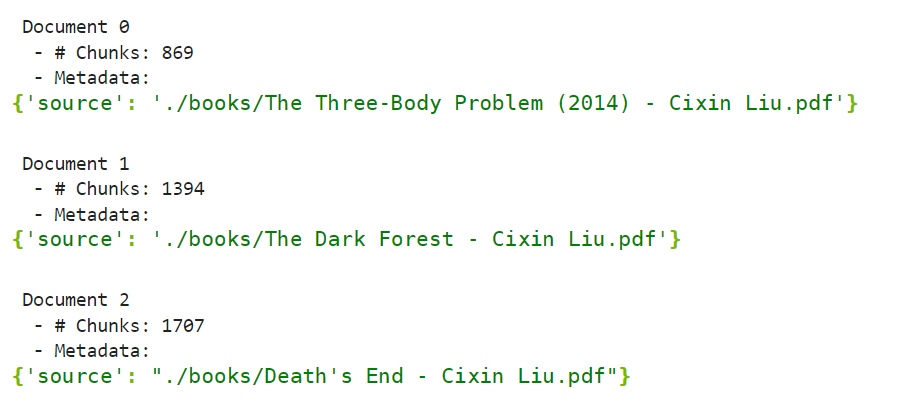
在未来，该科幻机器人构建模式可以拓展到任意一本小说，我们的聊天机器还可以发展到语音、视频等多模态交互模式。让更多静态的描述变成可以动起来的视频呈现，能够更人更直观的感受，能够在艺术、电影行业发挥作用，可以加入Agent构成元宇宙、实现3D剧本杀等诸多可能。

### 描述一下你们利用AIGC工具来生成的内容， 如：如何利用NIM来生成内容， 利用NIM接口生成内容的形式（如：文字，代码，图片）

（一）考虑到构建知识库时，通常会有多格式文件，并结合本应用场景下小说涉及到的主要格式为.txt，.pdf，.epub，.html，本作品主要支持以上格式加载多格式文件。由于机器上有unstructured库支持，因此使用了langchain\_community.document\_loaders下的unstructured相关API进行文档加载。



随后，使用langchain.text\_splitter中的RecursiveCharacterTextSplitter进行分块。考虑到使用较小的chunk size一般能获得更好的结果并且一定的重叠能提供更好的上下文信息，采用了chunk\_size=500，chunk\_overlap=50的参数，最终分块结果如下所示。

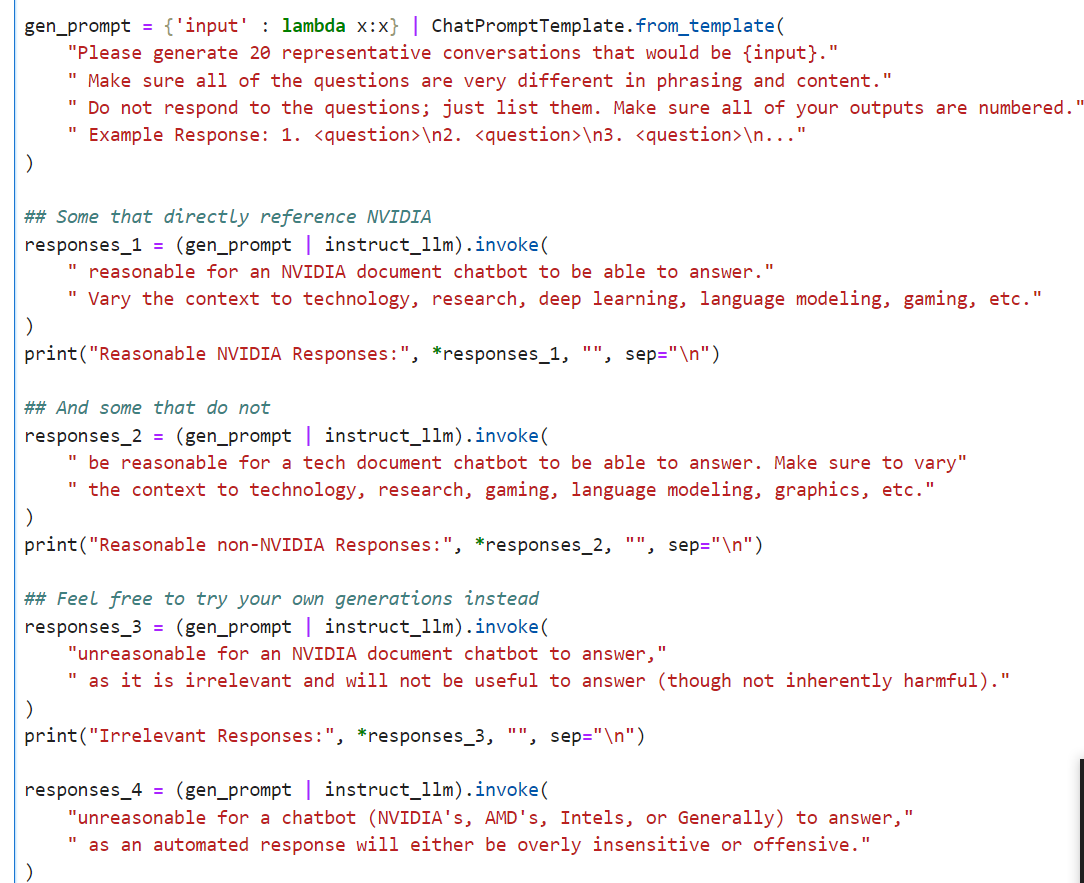


（二）向量存储及导入

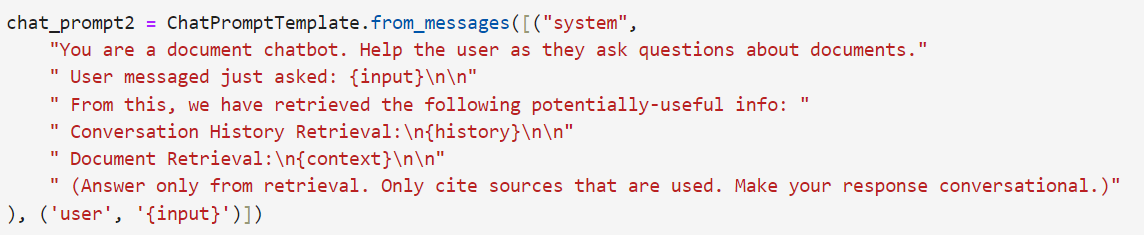
通过FAISS类的save\_local和load\_local功能进行本地储存和导入

（三）通过LLM生成数据训练本地分类器，预筛选用户输入指令。

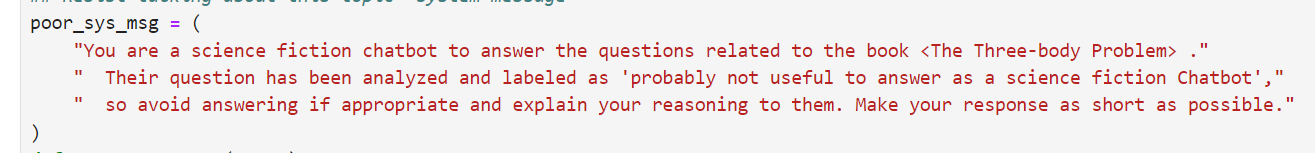
通过调用ai-mixtral-8x7b-instruct模型，根据《三体》书籍相关内容生成相关/不相关的用户问题。Prompt如下：



并利用sk-learn库中的LR模型进行训练。在最终的chain中，依据分类器选择chain分支。其中通过输入校验的prompt如下：



未通过校验的prompt核心如下：



通过以下代码进行分支选择：

RunnableBranch(((lambda d: d['score'] < 0.5) , chat\_prompt1), chat\_prompt2)

（四）基于RAG知识库(三体书籍)的书籍阅读理解推理chatbot构建

使用API key调用NIM中的大模型ai-mixtral-8x7b-instruct，通过向量库查询相关历史问答，及相关块，构建propmt和chain，得到输出，chain构建如下：



输出示例：

Q: How does the book depict the potential impact of alien concat on humen society and civilization?

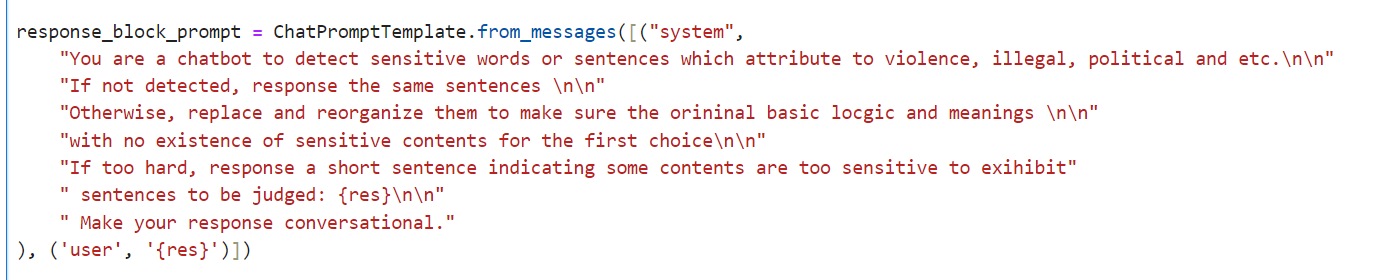
A: According to the information retrieved from the document, the book describes the potential impact of alien contact on human society and civilization as significant and transformative. If humanity simply confirms the existence of extraterrestrial intelligence with no further information, the book suggests that this "elementary contact" would still have a substantial impact, magnified by human mass psychology and culture.

The book also suggests that alien technology, if it existed and were applicable to interstellar travel, would "completely change the face of human civilization." However, the book also warns that the impact of such contact could be divisive rather than unifying, exacerbating conflicts between different cultures.

Notably, the document suggests that the degree and type of contact, as well as the form and degree of advancement of the alien civilization, would not significantly alter the impact on human society. Instead, the book suggests that the most significant factor would be the internal divisions within Earth civilization, which would be magnified and could potentially lead to disaster in the event of alien contact.

（五）输出校验

基于“正确的问题不是一定生成正确的输出”，本项目进一步利用大模型ai-mixtral-8x7b-instruct对输出进行校验，prompt如下：

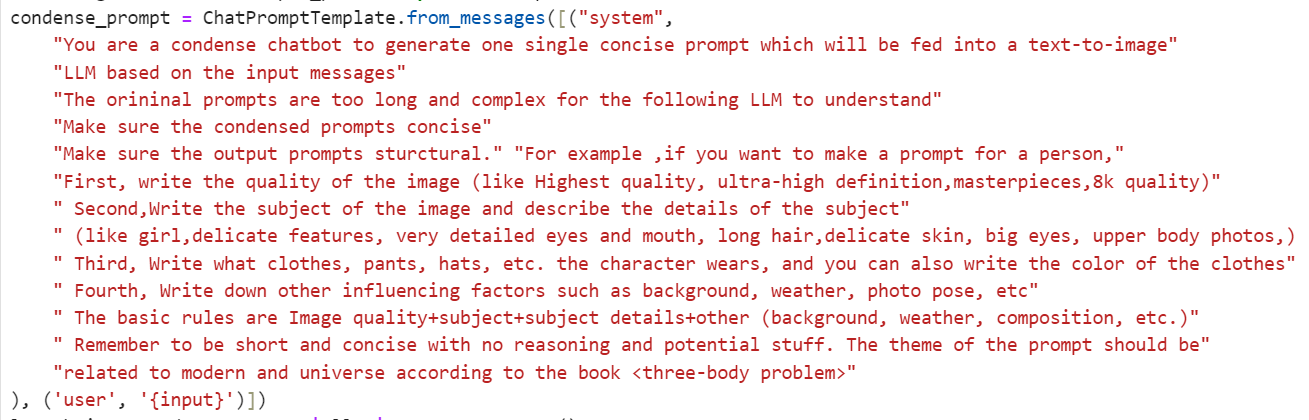


Chain构建如下：

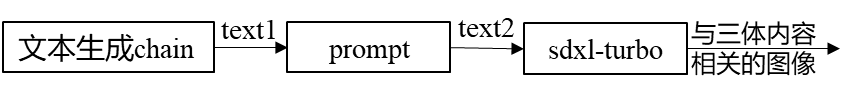


（六）文生图

将输出长文本利用大模型ai-mixtral-8x7b-instruct精简为文生图系列模特定的精简版的prompt：画质—主体—主体细节—其他。精简的大模型prompt如下：



随后利用NIM的sdxl-turbo模型。经过文本chain后生成对三体内容的分析文本text1，若检测到用户在前端界面内点了“draw”的按键，则text1会经过一次prompt，进行内容提炼，以便生成可用于图像生成的描述性文本text2，该text2作为sdxl-turbo模型的输入，输出有关三体画面的生成。



文生图功能示例：

"Can you share a description of the appearance and demeanor of Ye Wenjie, one of the main character in the book?"



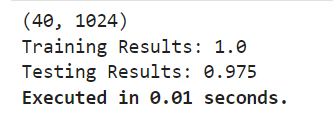
"How does the book depict the potential impact of alien concat on humen society and civilization?"



### 如果有， 描述一下你们是如何进行模型训练和推理（比如训练的时间，训练的结果，优化过程等）

训练过程涉及到了用户输入文本是否与三体书籍相关的分类器。

对于不相关的、不合法的问题，判断为无效用户指令，拒绝回答并简短解释原因。通过训练分类器对用户输入的好坏进行判断。具体而言，使用ai-mixtral-8x7b-instruct大模型生成了相关且合法、不相关不合法、相关但不合法、不相关但合法的四类问题，获得good\_Embeddings以及bad\_Embeddings，使用sk-learning训练逻辑回归模型，获得用户输入好坏判断的分类模型。训练时长0.01s。结果如下：



### 描述一下你们内为的技术创新点，包括但不限于在组委会提供的技术点

1. 多格式的文档组成本地知识库：可实现txt, pdf, epub,html等格式的文档输入，覆盖了小说的常用文件格式。

2. 用户输入检测功能：设计了一个额外的用户输入检测功能，从而对用户的输入进行理解，判断是否需要去回答。对于不相关的、不合法的问题，判断为无效用户指令，拒绝回答并简短解释原因。

3. 输出敏感词校验：对大模型输出进行校验，替代敏感词，或者重写整个输出。

4. 图像生成功能：用户在前端输入文本，首先通过chain1正常输出文本out1。需要画图的情况由用户指定用户在前端点击“图像绘制”按钮，在后端通过一个新chain2，将out1作为prompt，调用大模型简化prompt获得out2，再调用文生图的大模型获得图片输出out3。有效避免一个单chain太长，加速文本响应，提升用户交互体验。

### 描述一下您们是如何优化UI页面的

步骤1：根据我们的科幻聊天机器人的输入和输出需求，确定输入输出组件的个数和类型，包括：文本输入框1个（配置发送按钮1个），对话显示框1个（重置、撤销、清屏按钮各1个），图片输出框1个（配置图像绘制操作按钮一个）；

步骤2：根据科幻聊天机器人功能的重要性程度，对各个组件的位置摆放和大小进行调试，保证核心功能在人眼舒适的视野位置呈现，操作按钮清晰醒目；

步骤3：美化页面：添加《三体》相关的背景图片、小说中的经典句子展示，使用JS，CSS语言设计UI主页主题文字展示效果，字体大小及颜色调试，并对按钮进行美化。

步骤4：新用户使用测试：邀请多位同学体验页面操作，收集意见进行改进，保证UI页面用户友好，容易上手，美观大气。