

Natural Language Processing

课程项目说明

袁彩霞

yuancx@bupt.edu.cn

人工智能学院 智能科学与技术中心

大纲

- 题目说明
- 分组要求
- 关键节点
- 注意事项

题目说明

- 总体要求：
 - 可以选择老师指定的题目，也可以自命题
- 指定题目：
 - 大模型检索增强生成，具体说明和要求详见：
https://github.com/ZhaoFangkun1/NLP_RAG_Demo.git
- 自命题目：
 - 鼓励选择自命题目，但不得与其它课程使用相同（含高度相关）题目
 - 须在开题前找老师讨论，评估工作量和考察边界
 - 自命题目至少涵盖本课程学到的三个知识点
 - 例如指定题目涵盖了检索增强生成、检索策略优化、大模型预训练/微调

分组要求

- 原则上，每组一共2-3人，其中组长一人，如果人数多于3人或少于2人，须向老师说明原因。
- 组员须明确分工，每位组员的工作检查点须清晰可度量。
- 过程中的每个检查节点，需由不同的组员主要负责（主笔撰写文档或参加口头答辩）。

关键节点

- 开题：
 - 工作内容：
 - 组队
 - 选题
 - 调研相关工作
 - 制定项目目标
 - 确定技术选型
 - 检查时间：5月21日
 - 检查形式：书面报告或口头报告二择一，书面报告页数不超过4页，口头报告需准备幻灯片到课堂讲解

关键节点

- 开题：

- 报告须包括：题目、任务描述、预期目标、相关工作、技术方案、开发环境、组员分工及检查点（简要示例如下）、时间安排等
- 提交物：上述书面报告或答辩幻灯片，需在5月21日前完成并提交

姓名	任务分工	检查点1	检查点	检查点3	检查点4
张甲	过程管理、xx技术方案制定、xx代码编写、xx前端实现.....	每周简报	xx技术方案（对应开题报告xx部分、中期报告xx部分..）	xx代码	xx前端
王乙	数据处理、cc技术方案制定.....				
李丙					

关键节点

- 中期：
 - 工作内容：
 - 系统概要设计（系统架构、模块划分、接口设计、总体及各部分技术方案等）
 - 原型系统开发
 - 检查时间：6月4日
 - 检查形式：书面报告或口头报告二择一，书面报告页数不超过6页，口头报告需准备幻灯片到课堂讲解

关键节点

- 中期：
 - 报告须包括：题目、详细设计、已完成工作、初步结论、问题及可能的解决方案、后续工作计划等
 - 提交物：上述书面报告或答辩幻灯片，需在6月4日前完成并提交

姓名	任务分工	检查点1	检查点	检查点3	检查点4
张甲	过程管理、xx技术方案制定、xx代码编写、xx前端实现.....	每周简报	xx技术方案（对应开题报告xx部分、中期报告xx部分..）	xx代码	xx前端
王乙	数据处理、cc技术方案制定.....				
李丙					

关键节点

- 结题：
 - 任务：
 - 系统详细设计
 - 编码实现
 - 实验结果及分析
 - 系统调试测试
 - 检查时间：6月18/25日
 - 检查形式：项目总结报告及2-5分钟的演示系统视频，课堂验收
 - 演示视频须包括：任务描述、功能演示、其它关键特征（创新性能力或表现）等

关键节点

- 结题：

- 提交物：

- 上述短视频及说明性文档（如有）
 - 实验报告：实验数据、实验环境、所用的方法或模型、实验结果、实验结果分析、演示系统说明、总结
 - 提交物须在6月18/25日前完成并提交

姓名	任务分工	检查点1	检查点	检查点3	检查点4
张甲	过程管理、xx技术方案制定、xx代码编写、xx前端实现.....	每周简报	xx技术方案（对应开题报告xx部分、中期报告xx部分..）	xx代码	xx前端
王乙	数据处理、cc技术方案制定.....				
李丙					

注意事项

- 1. 请遵守各环节时间约定，按时进行并提交相应文档，如有特殊原因不能按时进行的务必提前说明。
- 2. 学院提供的计算资源有限，所以在设计时需考虑资源约束，鼓励利用外部资源，无论是计算资源、数据资源还是加工资源。
- 3. 可以利用ChatGPT等大模型的开放调用来增强自己系统的能力，但是仍然需要有自己实现的功能，而不能完全靠调用它们来完成所有功能。
- 4. 其它事宜请随时交流。

谢 谢！