浙江理工大学 2024 年“挑战杯”大学生创业

计划竞赛作品申报书

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 学院名称 (盖章)：计算机科学与技术学院（人工智能学院） | | | | | | | | | |
| 学院联系人： | | 朱诗威 | | | 联系方式： | | 86843316 | | |
| 项目名称 | | 政务先锋——基于检索增强的便民问答系统 | | | | | | | |
| 本科生数量 | | 8 | | | 研究生数量 | | 0 | | |
| 项目分组 | | ( ) 1.科技创新和未来产业  ( ) 2.乡村振兴和脱贫攻坚  (√) 3.城市治理和社会服务  ( ) 4.生态环保和可持续发展  ( ) 5 文化创意和区域合作 | | | | | | | |
| 团队成员  (最多10人) | | 姓名 | 性别 | 学院 | 年级、 专业 | 手机 | | 备注  (负责人) | |
| 徐赫 | 男 | 计算机科学与技术（人工智能学院） | 2022级  计算机科学与技术全英文班 | 18968683833 | | 负责人 | |
| 方嘉瑜 | 女 | 启新学院 | 2022级  启新卓越实验班 （3）班 | 18958520352 | |  | |
| 金坤东 | 男 | 计算机科学与技术（人工智能学院） | 2022级  计算机科学与技术全英文班 | 19357157043 | |  | |
| 蔡琰 | 女 | 法政学院、史量才新闻与传播学院 | 2021级  传播学  （1）班 | 18767681593 | |  | |
| 夏宇航 | 男 | 法政学院、史量才新闻与传播学院 | 2021级  法学  （2）班 | 15888638787 | |  | |
| 郭奇 | 男 | 计算机科学与技术（人工智能学院） | 2022级  计算机科学与技术  （2）班 | 17751732690 | |  | |
| 杨烁 | 男 | 法政学院、史量才新闻与传播学院 | 2021级  法学  （2）班 | 19883923239 | |  | |
| 黄朵儿 | 女 | 经济管理学院 | 2022级  会计学  （1）班 | 15868114510 | |  | |
| 指导教师 (最多3人) | | 姓名 | 性别 | 学院 | 职称 | 职务 | | 手机 | |
| 罗志一 | 女 | 计算机科学与技术（人工智能学院） | 讲师 | 无 | | 15000200709 | |
| 骆淑云 | 女 | 计算机科学与技术（人工智能学院） | 讲师 | 无 | | 13070163712 | |
| 项目简介 (500字以内) | | 减少人力成本、提升咨询效率和准确性是政府便民服务的重要发展目标。传统方案需要耗费大量的人力物力，并且每次的询问的准确度和效率都不能得到稳定的保证，也难以得到最新的回答。检索增强的语言模型以强大的信息相似度检索能力和感知能力，辅之以可以不断更新维护的向量文本数据库，使得它成为突破当前便民服务的痛点难点的重要手段。然而处理语言模型的回复范围限制，对于安全性的提升，以及如何使得文本近似度检索效果更好更快都是本项目探索的重点，包括后续的前后端开发工作也是本项目涵盖的内容。本项目采用渐进式的设计：1）模型、数据库测试部署，选择合适的模型；2）收集构建完整数据集，模型微调测试；3）文本近似度检索技术探索，优化部分效果；4）前后端开发，测试之后投入开放使用。最终本项目会实现一套完整、高效、稳定的问答系统，有力地提升了便民服务的用户体验效果、稳定回答质量，在基于检索增强的语言模型对其他领域的广泛的使用提供了经验和参考。 | | | | | | | |
| 社会价值 (500字以内) | | 我国《“十四五”规划纲要》中明确指出将深度学习框架算法平台构建和自然语言处理等领域创新列为重点科技前沿领域发展目标，足见国家对相关研究应用到国防民生的重视程度，基于深度学习技术的问答研究具有十分重要的战略意义。  作为国内政务服务领域大模型应用探索，北京市政务服务大模型更是在国内外探索大模型助力政务管理实践方面具有极强的示范带动意义。  本项目立足于“讲民生政策触手可及”的理念，服务于人民群众对民生政策的知悉需求，创设一站式便民问答系统，把政策解读给群众，以技术服务政务，对提高人民的幸福感，推动数字中国建设具有深刻的现实意义。 | | | | | | | |
| 实践过程 (500字以内) | | 项目负责人参加阿里云的云栖大会，调研现在技术趋势和其他公司的进展，观看Elastic公司关于模型私有化部署的演讲分析，学习借鉴经验并且调节自己项目的工作路线，研究他们的开源框架。根据他们为企业服务的计划，加以分析政府公务需求的基础上，不断改进项目实施计划，保证项目的安全性和可靠性。  除此之外，项目成员和浙江天玑星数据科技公司了解如何高效与chatGPT为代表的模型交互，了解他们公司主要的提供稳定的API调用服务业务。  本项目主要分成语言模型部署测试和选择（已完成），数据的收集和处理构建数据集（进行中），语言模型的微调和优化，文本切分和近似度检索技术的选择测试优化（进行中），推广项目寻找商业合作方，根据需求部署开发后端API和设计做出前端UI界面，按照后续的需求和用户反馈不断优化提升项目。  本项目语言模型尝试了chatGLM2，llama2等，最后部署使用chatglm2-6b模型。项目技术人员已经在收集和处理数据的过程中，同步也在进行选择向量数据库切分储存文本技术的探索，并且定期采用小数据进行微调尝试。后续会进行传统的前后端开发，打包交付环境等。 | | | | | | | |
| 创新意义 (500字以内) | | 本项目的创新之处主要体现在如下几点：1）使用检索加生成的新问答范式。这种范式结合了检索和生成技术的优势，提高了回答问题的准确性、全面性和灵活性。它可以根据用户的需求，从知识库中快速找到答案，也可以生成高质量、详实的回答。2）可更新、持续迭代。语言模型可以通过不断地学习和优化，实现自身的持续迭代和更新。这使得它可以随着时间和环境的变化，不断地提高回答问题的效果和用户体验。3）可迁移、扩展到其他领域，这种技术具有广泛适用性。它可以轻松地迁移和扩展到其他领域，如智能家居、金融服务、医疗保健等，为不同行业提供创新解决方案。 4）提高用户粘性。通过提供准确、快速、个性化的回答，语言模型可以提高用户满意度，增加用户粘性，从而提高品牌的声誉和市场份额。5）促进产业融合。语言模型在便民领域的应用，有助于推动便民行业与其他领域的融合发展，例如人工智能、物联网、智能交通等，从而创造出更多的商业机遇和发展空间。6）具有无监督学习的优势。通过大规模无监督学习，语言模型可以自动地从原始文本数据中提取语义信息和知识，从而不断地优化和扩展自身的知识库，提高回答问题的准确性和全面性。 | | | | | | | |
| 发展前景 (500字以内) | | 数字化转型浪潮正在以前所未有的高度、广度与深度塑造政府的便民服务，由语言模型驱动的智能问答系统可以为用户提供有效的、准确的关于便民服务问题的解答，在增强人民幸福感的同时，解决政府便民服务供给智能化的需要，加快政府数字化转型。  未来，我们的项目有望成为便民服务中广泛应用的智能工具，帮助普通用户解答便民服务相关问题，提升用户体验感和满意度；帮助政府提升便民服务效率和数字化治理能力，加快建设智慧政府；帮助AI提高数据收集和学习能力，增强其在专业领域应用的探索；帮助国家在语言模型这个领域的发展，加快建设网络强国、数字中国。  人工智能驱动与赋能便民服务将会成为数字时代构筑便民服务新趋势、新局面和新优势的重要支撑。AI语言模型在未来的发展前景广大，一定会在未来走进千家万户！ | | | | | | | |
| 团队协作 (500字以内) | | 本项目旨在将语言模型应用于各类便民服务，例如12345便民服务平台电话接听。为了实现这一目标，需要涉及到数据采集与清洗、模型训练与优化、文本切分技术与文本相似度检索技术等多个模块的有效协作。  数据采集和清洗模块涉及到对各类政策法规相关的文本数据进行收集和预处理。这个过程需要利用数据下载、数据去重清洗、数据格式转换等多项技术进行。在模型训练和优化模块中，需要使用大量数据来训练语言模型，并不断优化其精度和效率。这个过程需要运用深度学习和自然语言处理等技术。文本切分技术能有效将较长的文本切分成短句子，方便模型理解和处理；而文本相似度检索技术可以对用户提问进行相似度匹配，以找到最相关的答案。  在以上各个模块分工中，我们团队充分发挥每个人的专业优势，注重合作与创新，建立和谐的团队氛围，通过定期召开讨论会议，分享进展和提出疑问等方式，建立高效的沟通机制，确保各个模块间的信息畅通。同时，我们还将建立良好的项目管理机制，包括需求分析、数据评估和模型评估等环节，以实现对各类便民服务问题的准确和高效解释。  总之，在本项目进行过程中，我们团队将在各个模块部分紧密合作，发挥各自优势，确保项目顺利进行。 | | | | | | | |
| 项目介绍材料 | | 项目背景 2022年年底及2023年年初，GPT(Generative Pre-trained Transformer)盛行于各个领域，它的成熟带来了一系列生产力的提升，GPT具备了强大的自然语言处理和对话生成能力，能够理解和生成人类语言。  政务服务需要处理大量的信息和用户需求  政务服务需要处理大量的信息和用户需求 旧方法的问题  * + 搜索引擎可能会通过一些无关的网页——无法根据用户的具体需求和背景提供个性化的服务；   + 问答系统的表现参差不齐 ——无法满足基本咨询，甚至提供错误信息；   + 刚性和有限的回答——对于复杂或个性化的问题，往往无法给出准确或满意的答案；   + 人工处理和延迟——需要一定的时间和人力资源。这会导致服务的延迟和效率低下；   + 更新不及时——由于政策和法规经常变化，传统系统可能无法及时更新信息，导致提供的答案可能过时或不准确。  项目目标  * 简洁且精准的回答 * 提供完整的事项办理要素和便民服务功能 * 增强交互性 * 提高服务的效率和响应速度，减少用户等待时间  项目简介 基于检索增强的语言模型凭借强大的信息检索能力和相似度感知能力，辅之以可以不断更新维护的向量文本数据库，使得它成为突破当前便民服务的痛点难点的重要手段。 探索重点  * + 处理语言模型的回复范围限制；   + 提升安全性；   + 使得文本近似度检索效果；   + 前后端开发工作   以技术为核心，搭建政务便民服务领域的知识库问答系统，提升政府服务的效率和质量，增强政府与公众之间的互动和沟通，推动数字化转型和智慧政务建设。   社会价值 我国《“十四五”规划纲要》中明确指出将**深度学习框架算法平台构建和自然语言处理**等领域创新列为重点科技前沿领域发展目标，足见国家对相关研究应用到国防民生的重视程度，基于深度学习技术的问答研究具有十分重要的战略意义。  作为国内首个政务服务领域大模型应用探索，北京市政务服务大模型更是在国内外探索大模型助力政务管理实践方面具有极强的示范带动意义。  本项目立足于**“讲民生政策触手可及”**的理念，服务于人民群众对民生政策的知悉需求，创设**一站式便民问答系统**，把政策解读给群众，以技术服务政务，对提高人民的幸福感，推动数字中国建设具有深刻的现实意义。 实践过程 本项目语言模型尝试了chatGLM2、llama2等，最后选择chatglm2-6b模型，  成功部署并且测试微调方式通过。  项目技术人员已经在收集和处理数据的过程中，同步也在进行选择向量数据库切分储存文本技术的探索优化，并且定期采用小数据进行微调尝试，本项目由专业的指导老师进行建议咨询方面，以学校的实验室设备作为项目未来研发的坚实保障，拥有极高的落地可行性。       * + 本项目主要分成语言模型部署测试和选择（已完成）   + 数据的收集和处理构建数据集（进行中）   + 语言模型的微调和优化   + 文本切分和近似度检索技术的选择测试优化（进行中）   + 推广项目寻找商业合作方   + 根据需求部署开发后端API和设计做出前端UI界面   + 按照后续的需求和用户反馈不断优化提升项目。  特色和创新之处  1. 使用检索加生成的新问答范式   这种范式结合了检索和生成技术的优势，提高了回答问题的准确性、全面性和灵活性。它可以根据用户的需求，从知识库中快速找到答案，也可以生成高质量、详实的回答。  2.可更新、持续迭代  语言模型可以通过不断地学习和优化，实现自身的持续迭代和更新。这使得它可以随着时间和环境的变化，不断地提高回答问题的效果和用户体验。  3.可迁移、扩展到其他领域，具有广泛适用性  可以轻松地迁移和扩展到其他领域，如智能家居、金融服务、医疗保健等，为不同行业提供创新解决方案。  4.提高用户粘性  通过提供准确、快速、个性化的回答，语言模型可以提高用户满意度，增加用户粘性，从而提高品牌的声誉和市场份额。  5.促进产业融合  语言模型在便民领域的应用，有助于推动便民行业与其他领域的融合发展，例如人工智能、物联网、智能交通等，从而创造出更多的商业机遇和发展空间。  6.具有无监督学习的优势  通过大规模无监督学习，语言模型可以自动地从原始文本数据中提取语义信息和知识，从而不断地优化和扩展自身的知识库，提高回答问题的准确性和全面性。 创新意义   本项目  准确率：85%  选择模型语料：中文开源  问答效率：高  特点：实时对话，可追溯；  方便灵活，信息可靠；  传统系统  准确率：80%左右  选择模型语料：英文  问答效率：低  特点：主观性强，难以证实；  生成结果质量参差不齐； 发展前景 数字化转型浪潮正在以前所未有的高度、广度与深度塑造政府的便民服务，由语言模型驱动的智能问答系统可以为用户提供有效的、准确的关于便民服务问题的解答。  在增强人民幸福感的同时，解决政府便民服务供给智能化的需要，加快政府数字化转型，加速智慧政务的建设。      对用户：帮助普通用户解答便民服务相关问题，提升用户体验感和满意度；  对政府：帮助政府提升便民服务效率和数字化治理能力，加快建设智慧政府；帮助国家在语言模型这个领域的发展，加快建设网络强国、数字中国；  对技术发展：帮助AI提高数据收集和学习能力，增强其在专业领域应用的探索。 团队介绍   徐赫 负责人兼技术员   * + 2022届浙江理工大学大学生程序设计竞赛一等奖   + 蓝桥杯C/C++组省赛二等奖   + 曾获校级二等奖学金     方嘉瑜 产品经理   * + 社会积极分子   + 优秀志愿者   + 曾获校级一等奖学金     金坤东 技术员   * + 蓝桥杯C组省赛3等奖   + mathorcup杯3等奖     郭奇 技术员    夏宇航 法务   * + 全国高校第十七届“金法槌杯”模拟法庭大赛三等奖   + 浙江理工大学第八届互联网+创新创业大赛院级一等奖   + 浙江理工大学2021年本科生创新计划一般立项   + 浙江理工大学2023年“挑战杯”大学生课外学术科技作品竞赛校级三等奖     蔡琰 宣传   * + 浙江省高校首届心理情景剧短视频大赛校级三等奖   + 首届全国高校心理微电影优秀作品校级三等奖   + 曾获校级三等奖学金     黄朵儿 财务    杨烁 法务 团队协作 本项目旨在将语言模型应用于各类便民服务。  为了实现这一目标，需要涉及到多项技术、多个模块的有效协作。   数据采集和清洗 数据采集和清洗模块涉及到对各类政策法规相关的文本数据进行抓取和预处理。这个过程需要利用数据下载、数据去重清洗、数据格式转换等多项技术进行。 模型训练和优化 这个过程需要运用深度学习和自然语言处理等技术，使用大量数据来训练语言模型，并不断优化其精度和效率。 文本切分和相似度检索 将较长的文本切分成短句子，方便模型理解和处理；而文本相似度检索技术可以对用户提问进行相似度匹配，以找到最相关的答案。 | | | | | | | |
| 其他相关证明 材料 | |  | | | | | | | |