

**本科生课程论文**



课 程: IT创业项目实践

姓 名: 侯兆晗

学 院: 信息管理学院

专 业: 信息管理与信息系统

学 号: 19122120

任课教师: 朱毅华

2025年1月6日

南京农业大学教务处制

**IT创业项目实践课程个人报告**

**-- 侯兆晗** 19122120

**一、项目背景与意义**

在当前全面从严治党的背景下，党政机关对党内法规的学习、理解和应用需求日益增长，但传统的法规检索、学习和管理方式存在效率低下、更新滞后、教学复杂、研究效率低下和管理难度大等问题。为解决这些痛点，党内法规智能化科技有限公司计划开发一款党内法规大语言模型产品。我们小组期末项目旨在通过先进的AI技术和大数据信息化手段，构建专门面向党内法规领域的生成式对话大语言模型，以提升法规的检索效率与准确性，增强法规的理解与应用能力，并确保法规解释的一致性。通过深度融合和充分利用现有法规资源，打破传统的研究和服务模式，为全球党务工作提供数字化、智能化的解决方案。小组期末项目的长期目标是成为全球领先的党内法规智能化解决方案提供商，助力全面从严治党的深入实施，推动党务工作的现代化转型。

**二、项目实施过程**

在本次小组期末作业中，我作为项目核心成员参加了项目的全部实施过程，以下是我在全部过程中的步骤总结：

1.项目启动与团队组建：确定项目目标为开发党内法规大模型，选址北京中关村，成立党内法规智能化科技有限公司。组建包括产品经理、技术负责人、UI设计师、市场营销经理和运营经理在内的创业团队，明确各成员角色与职责分配。

2.法规数据库构建：收集党内法规数据，识别并处理数据中的缺失值，补充关键条款解释或适用范围说明等信息。标准化数据格式，包括统一日期格式、法规编号格式、文本编码格式等，确保数据的一致性和准确性。



图1 模型数据来源之一

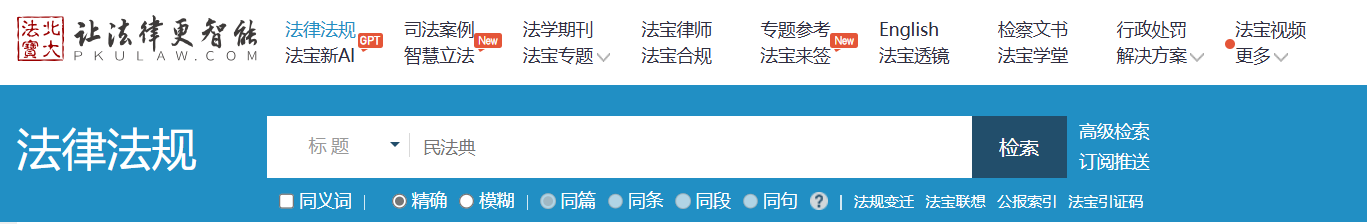


图2 模型数据来源之一

3.数据标注与质量控制：采取人工标注与自动标注相结合的方法，对法规类别和简单条款信息进行初步自动标注，随后由人工审核和修正。建立质量检查和审核机制，定期对标注结果进行随机抽样检查，评估标注的准确性、完整性和一致性，确保标注结果符合项目要求。

4.数据分析与核心功能开发：运用文本分析技术对标注后的党内法规数据进行深入分析，提取关键信息、主题和趋势。基于分析结果，开发法规查询与智能解读的最小可行产品（MVP），聚焦于实现核心功能。

5.模型评测与优化：使用BLEU、chrF、ROUGE-L和BERTScore等指标评估模型性能。进行领域知识评测，通过完成党内法规填空、定义和开放领域知识评测，验证模型的领域知识能力。引入人类反馈进行强化学习，优化模型理解人类偏好的能力。

6.应用集成与平台搭建：构建对话数据模拟法规问答场景，增强模型的交互能力。搭建网页平台，将微调后的模型集成到平台中，提供法规查询等功能。同时开发手机应用程序，提升用户体验。

7.功能扩展与试点推广：在完成初期开发阶段后，进行功能拓展，如增加法规合规性审查模块。推出基础版本进行内部测试，随后发布MVP，并选定3-5个党政机关进行首批试点合作。根据试点反馈进行产品优化和调整。

8.市场推广与品牌建设：制定营销策略，通过线上推广如微信公众号、微博、抖音等平台发布内容，吸引潜在用户关注；线下参加党建会议、党务工作研讨会等活动，展示产品功能和优势。建立官方网站和客服团队，提供产品介绍和技术支持。

9.融资计划与资金使用：制定种子轮、天使轮和A轮的融资计划，涵盖产品原型开发、法规数据库整理、功能模块完善、技术团队扩展和全国市场推广等阶段。融资来源包括政府专项资金、天使投资或风险投资以及党政相关机构或大型国有企业的资金与政策资源支持。

10.风险管理与质量控制：定期进行风险评估，包括信息安全风险和市场风险等，并采取相应的应对措施。建立严格的数据安全管理制度和应急预案，确保数据在传输和存储过程中的安全性。通过持续集成和自动化测试确保每次版本发布的稳定性，提供优质的用户体验。

**三、个人对项目的贡献**

在本次小组期末作业中，我作为项目核心成员参加了项目的全部实施过程，我个人主要负责数据分析与质量控制、模型评价与优化、风险管理与优化等部分文书写作，以下是我对这些内容的具体心得与贡献说明：

1.数据分析与质量控制：作为项目核心成员，我深入运用文本分析技术，对标注后的党内法规数据进行细致分析，提取关键信息、主题和趋势。通过词频统计、主题模型分析等手段，我有效地支持了法规分类和知识图谱构建，为模型优化提供了数据基础。同时，我严格把控数据质量，确保数据的准确性和一致性，对采集到的法规文本和相关资料进行去重、格式检查、语法修正和内容核实，有效避免了数据冗余和错误对模型训练的影响。



图3 数据来源分析

2.模型评价与优化：在基于小组其他成员完成他们任务后，我又负责使用BLEU、chrF、ROUGE-L和BERTScore等指标对模型性能进行客观评估，通过对比不同模型版本的表现，为项目团队提供了模型改进的明确方向。在模型微调与优化阶段，我积极参与参数调整、指令微调等工作，提升了模型在特定任务上的适应性和准确性。此外，我还通过领域知识评测等方式，验证了模型的领域知识能力，为模型的进一步优化提供了有力支持。

文本

描述已自动生成

图4 模型测试

**图形用户界面, 文本, 应用程序, 电子邮件

描述已自动生成**

图5 DEMO测试

3.风险管理与质量控制：同时我对小组项目中的潜在风险进行了全面识别和分析，包括市场风险、技术风险、资金风险等，并制定了针对性的应对措施。在数据安全与隐私保护方面，我强调了加密技术和安全协议的重要性，并推动了权限管理制度的建立，确保了数据在传输和存储过程中的安全性。同时，我还参与了质量检查和审核机制的建立，预计将来项目落地后，实现定期对标注结果、模型性能等进行随机抽样检查，确保项目质量符合既定标准。

4.部分文书写作：作为项目核心成员，我还承担了部分项目文档的写作工作，包括项目计划书部分内容、需求分析报告、风险评估报告等。我清晰、准确地阐述了项目目标、实施步骤、风险评估及应对措施等内容，为项目的顺利推进提供了有力的文字支持。我的文书工作不仅提升了项目的专业性和可信度，还有助于团队成员之间的沟通和协作。

**四、项目未来展望**

在IT创业项目实践课程所学知识基础上，我对党内法规大语言模型项目有了更深刻的体会。通过本次课程期末小组作业项目实践，我深刻认识到，党内法规智能化科技有限公司的项目不仅具有深远的社会意义，还蕴含着巨大的商业潜力。公司致力于通过先进的AI技术和大数据信息化手段，解决党内法规检索困难、更新滞后、教学复杂、研究效率低下以及管理效率不高等问题，旨在提升党内法规的学习、理解和应用效率，助力全面从严治党的深入实施。展望未来，随着模型的持续优化和应用场景的不断拓展，我们项目团队有信心将党内法规大语言模型打造成为全球领先的智能化解决方案，推动党务工作向数字化、智能化转型。

在项目开发与建设过程中，我深刻体会到了团队协作的重要性。从需求调研、系统设计、数据库构建到模型训练、功能开发，每一个环节都需要团队成员之间的紧密配合和高效沟通。通过敏捷开发的方法，我们分阶段、迭代地进行产品开发，确保了产品能够快速适应市场变化，满足用户需求。同时，我也深刻认识到了持续学习和创新的重要性。在技术领域，只有不断跟进最新技术动态，不断优化算法和模型，才能在激烈的市场竞争中保持领先地位。

作为项目核心成员，我还深刻体会到了创业的不易。从项目立项、融资计划到市场推广、用户反馈收集，每一个环节都需要投入大量的精力和时间。在这个过程中，我学会了如何制定详细的计划和时间表，如何合理配置资源，以及如何应对各种风险和挑战。这些经验不仅对我的个人成长具有重要意义，也将对我的未来职业发展产生深远的影响。

通过本次IT创业期末小组项目实践，我不仅收获了宝贵的项目经验和知识，还深刻体会到了团队协作、持续学习和创新以及创业精神的重要性。我相信，在未来的日子里，我将继续秉承这些理念，不断追求卓越，为实现个人价值和社会价值而不懈努力。