



湖南电子科技职业学院
HUNAN VOCATIONAL COLLEGE OF ELECTRONIC AND TECHNOLOGY

大数据技术专业

2025届毕业设计整体情况分析报告

专 业 代 码 : 510205

年 级 : 2022级

专 业 负 责 人 : 黄毅伟

编 制 时 间 : 2025年7月

信息工程学院

二〇二五年七月

目 录

一、毕业设计过程总结	1
二、毕业设计选题分析	2
三、毕业设计成绩分析	4
四、毕业设计存在的问题	5
五、毕业设计改进措施	8

一、毕业设计过程总结

自2024年11月到2025年6月，在信息工程学院领导的指导和部署下，在团队各位教师的积极配合下，大数据技术专业2022级毕业生的毕业设计圆满完成。教研室严格按照教学计划的要求，组织和落实完成教学计划中规定的毕业设计任务。本次毕业设计针对的是大数据技术专业2025届（2022级）的学生，共有82名学生参加毕业设计，参与指导的教师5人。

毕业设计过程的具体时间点见表1：

表 1 大数据技术专业毕业设计过程阶段表

阶段	教师要求	学生要求	时间安排
选题指导阶段	指导教师在课题审批后应编写毕业设计任务书，发给学生。任务书包括目的和作用，任务与要求，基本内容，工作程序与时间分配，所需知识、主要参考资料等，对学生完成课题起指导作用。	学生严格遵守毕业设计要求，服从指挥，主动接受指导教师指导并根据自己专业特长选择合适毕业设计课题。	11月1日至12月1日
开题论证阶段	当课题确定后，指导教师应向学生讲清毕业设计题目的意义，提出明确的要求，制定工作计划。	学生收集和查阅文献资料及正确使用工具的方法，确定方案、理论分析、画图、试验处理数据等工作，并拟订毕业设计提纲。	12月1日至12月30日
指导过程阶段	指导教师应抓住关键问题进行指导，因材施教，不能出现原则性错误；要把握学生的工作进度，使全部工作任务保质有序按时完成。	学生应对本人的毕业设计质量负责，必须在规定时间内完成给定的毕业设计各项任务。毕业设计书写格式遵照学院“毕业设计排版规范”有关要求。学生应经常（定期）主动向指导教师汇报工作进度和遇到的疑难问题，争取指导教师的指导和监督。	次年1月1日至3月31日
资料整理阶段	审阅毕业设计初稿，指导学生修改，直到完成定稿。	学生必须独立完成规定的全部工作任务，根据指导教师提出的修改建议，认真完成修改，进一步优化完善毕业设计，直到完成定稿。	次年4月1日至5月15日
成果答辩阶段	指导教师评阅学生毕业设计并写出评语，成立答辩小组，完成答辩	学生答辩前应进行充分准备：如写出提要或汇报提纲、必要的图表、试讲等，锻炼自己的表达表述能力。答辩后，学生应提交相关资料（包括任务书、作品、查重报告等）。	次年5月26日至6月15日

质量监控阶段	对学生的毕业设计进行详细的评阅，给出建设性的意见和建议，帮助学生提高毕业设计的质量。指导学生如何有效地准备答辩，包括提要撰写、PPT制作、试讲练习等方面；参与答辩小组的组建，与其他成员一起制定评审标准和流程；在答辩过程中，公正、客观地评价学生的答辩表现，提出有针对性的问题，引导学生深入思考。	撰写毕业设计的提要或汇报提纲，确保内容清晰、逻辑严密；准备必要的图表、PPT等辅助材料，以直观展示研究成果；进行试讲，邀请同学或导师给予意见，不断改进汇报内容和表达技巧；确保毕业设计作品完整，符合学术规范，包括但不限于查重报告、文献引用等；在答辩过程中，清晰、准确地回答答辩小组提出的问题，展现良好的专业素养；积极听取答辩小组的意见和建议，记录下来以便后续修改和完善毕业设计；答辩结束后，根据答辩小组的意见对毕业设计进行最终修改；提交包括任务书、作品、查重报告在内的所有相关资料，确保资料齐全、格式正确。	整个毕业设计阶段
--------	---	---	----------

二、毕业设计选题分析

大数据技术专业毕业设计选题类别分为方案设计类、产品设计类。从选题方向上来看本届学生的毕业设计选题主要集中在以下几个方向：

数据采集与可视化分析类（占比约40%），此类选题数量最多，涵盖电商平台、招聘网站、文化内容平台、旅游服务等多个行业领域。学生普遍采用Python + Requests/Scrapy + Pandas + Matplotlib/PyEcharts的技术栈，完成从网页爬取、数据清洗、结构化存储到可视化展示的全流程开发。部分题目还结合Flask或Django搭建简易Web平台，实现数据的动态展示，体现出较强的工程整合能力。

DevOps与系统运维部署类（占比约35%），该类选题近年来显著增多，体现出学生对云原生、自动化运维等企业级开发流程的高度关注。技术栈覆盖Docker、Kubernetes、Spark、Hadoop、Maxwell、PostgreSQL、Git/GitHub、Conda环境管理等，涉及容器化部署、集群配置、数据同步、持久化存储、自动化运维等多个关键技术环节。部分题目聚焦国产化操作系统（如欧拉系统），体现出对信创生态的关注。

AI与智能系统开发类（占比约15%），随着大模型技术的普及，学生开始积极探索AI在实际场景中的集成与应用。这类题目反映出学生对生成式AI技术的关注，能够熟练调用主流大模型API，构建具备自然语言处理能力的系统。

虽然多数仍处于“接口调用+功能封装”阶段，尚未深入模型微调或本地部署，但已展现出良好的技术敏感度和创新能力。

系统开发与后端服务设计类（占比约10%），包括跨平台通信系统、数据库服务接口开发等，强调系统架构设计与模块化实现。

部分选题见下表：

表 2 大数据技术专业毕业生选题示例

产品设计类	基于淘宝男装的数据采集分析与可视化实现
	基于前程无忧网站数据可视化设计与实现
	基于千问 API 的翻译程序设计与实现
	跨平台 IT 通信系统设计与实现
	基于 flask 的 mysql 服务设计与实现
	基于大数据技术的招聘信息数据分析系统的设计与实现
	AI 聊天机器人框架设计与实现
	基于 deepseek 的 word 智能体设计与实现
	文轩网电商网站数据采集与分析
	基于 Python 的火锅餐饮行业数据可视化分析与实现
	基于网站流量数据分析及可视化的设计与实现
	基于 Python 的淘宝店铺销售数据采集与可视化
	基于 Python 的天猫服装数据采集分析与实现
	基于 Python 的淘宝电视机数据采集分析与可视化实现
	基于 Python 的校园售货机商品数据分析与可视化实现
	基于 Python 的小说排行榜数据可视化与实现
方案设计类	RockyLinux 部署 PostgreSQL 设计方案
	WSL 下的 docker 配置设计方案
	AKShare 部署设计方案
	基于 docker 的 redis 服务搭建设计方案
	centOS8 下 docker 环境搭建设计方案
	协作式项目 git 运维设计方案
	基于 docker 的 mySQL 搭建设计方案
	github 连接设计方案
	PG 数据库 schema 设计方案
	RL 搭建 k8s 集群设计方案
	SpringBoot 项目云端部署设计方案
	DockerCompose 的开发设计方案
	Docker 多阶段构建设计方案
	基于 Maxwell 的数据同步设计方案
	windows 下 Spark 集群配置设计方案
	欧拉系统配置 k8s 集群设计方案

选题方向覆盖面广、技术领域多元，充分体现了学生在信息技术应用、数据处理分析、系统架构设计以及运维部署实践等多个维度的综合能力。选题内容紧跟当前信息技术发展的主流趋势，紧密结合实际应用场景，既注重技术实

现的可行性，又体现出一定的创新性与挑战性，整体呈现出“技术前沿性强、实践导向明确、工程落地清晰”的特点。

三、毕业设计成绩分析

截止到2025年7月，大数据技术专业2025届毕业生共82人，顺利毕业人数78人，毕业率达到95%，整体毕业情况良好，体现了专业教学工作的有效推进和学生综合素质的稳步提升。

毕业设计难易程度	完成人数	毕业设计成绩评定
较难	8	优秀
中等	32	良好
一般	34	合格
征兵	4	
毕业人数	78	

从学生的毕业设计成绩分布来看，绝大多数学生均达到了毕业设计的基本要求，成果规范、内容完整，具备较强的专业应用能力。其中，8名学生的毕业设计选题具有较高的技术难度和创新性，研究成果突出，数据处理逻辑清晰，系统实现完整，被评定为优秀毕业设计，占总人数的10%。另有41%的学生结合自身技术基础与学习能力，合理评估项目复杂度，选择难度适中、可行性较高的题目，在规定时间内高质量完成了数据采集、清洗、分析、可视化或系统部署等核心任务，体现了良好的工程实践能力和时间管理意识。

通过本次毕业设计的成绩评定可以看出，学生对毕业设计环节高度重视，能够根据自身的技术掌握情况（如Python编程、爬虫技术、数据可视化、大数据平台搭建、AI模型调用等）科学选题，目标明确。部分学生勇于挑战前沿技术方向，如基于大模型API的智能应用开发、实时数据同步方案设计、Kubernetes集群在国产系统上的部署优化等，展现了较强的探索精神和技术攻坚能力。同时，这些学生能够在指导教师的帮助下，按时推进各阶段任务，文档撰写规范，代码结构清晰，整体专业性与完成度较高。

此外，毕业论文的完成质量与学生日常学习习惯、学习态度和技术积累密切相关。成绩优异的学生普遍具备以下特点：主动与指导教师保持沟通，及时汇报进展；对教师提出的修改意见认真对待、积极反馈；具备较强的自主学习能力，能独立查阅文献、调试代码、解决技术难题；注重细节，重视数据准确性、系统稳定性与成果展示的专业性；展现出良好的创新思维，能够在基础功能之上进行拓展优化，如引入机器学习算法进行趋势预测、利用Docker实现多环境一键部署等。

相反，个别学习态度不够端正的学生存在拖延现象，任务推进缓慢，代码质量不高，文档撰写不规范，需多次催促才能提交阶段性成果，导致最终毕业设计成绩偏低，甚至影响正常毕业。

综上所述，大数据技术专业本届毕业生整体表现良好，毕业设计选题贴近行业需求，技术路线清晰，实践性强。学生能够结合大数据领域的核心技术（如数据采集、处理分析、可视化、云原生部署、AI集成等）开展应用型研究，体现出扎实的专业素养和较强的工程实现能力。未来可进一步鼓励学生参与真实项目、企业合作课题或竞赛成果转化，持续提升毕业设计的创新性、深度与社会价值。

四、毕业设计存在的问题

大数据技术专业毕业设计周期较长，技术综合性强，涉及数据采集、处理、分析、可视化、系统开发、运维部署等多个技术领域，对学生的编程能力、工程实践能力和系统思维要求较高。在2025届毕业设计实施过程中，尽管整体完成情况良好，但仍暴露出一些共性问题，现将主要问题汇总如下：

1、选题方面

大部分学生的选题能够紧密结合当前大数据技术的发展趋势和实际应用场景，如电商数据分析、招聘信息挖掘、AI智能应用、容器化部署等，具有较强的实践性和可行性。部分选题还体现出一定的创新性，如基于大模型API的智能体设计、跨平台数据同步方案、K8s集群在国产操作系统上的部署等。

但也有少数学生对专业方向理解不够深入，选题存在以下问题：选题范围过窄或重复性高：例如多个学生选择“基于Python的淘宝XX商品数据采集”类题目，研究内容高度相似，缺乏差异化和深度拓展，仅停留在数据爬取与简单可视化层面，未能体现大数据分析的核心价值；技术难度与自身能力不匹配：个别学生盲目追求“热门技术”（如AI、K8s、Spark集群），但缺乏前置知识储备，导致项目推进困难，最终成果流于表面；应用场景脱离实际：部分题目缺乏真实业务背景支撑，数据来源虚构或样本量不足，分析结论缺乏可信度和实用性。

2、技术工具与平台应用方面

大多数学生能够熟练使用Python、Java、MySQL、Hadoop、Flask、Docker、Git等主流工具完成毕业设计任务，部分学生已掌握Spark、Kubernetes等高阶技术，体现出较强的技术应用能力。

但仍存在以下问题：

工具版本滞后，与行业脱节：部分学生仍在使用较旧版本的Python（如3.7以下）、Docker或数据库系统，未及时跟进当前企业广泛使用的最新稳定版本，影响功能兼容性与性能表现；开发环境配置混乱：如Conda环境管理不当、依赖包冲突、虚拟环境未隔离等问题频发，导致项目迁移困难；平台部署脱离真实场景：部分“云端部署”“K8s集群搭建”类题目仅停留在本地模拟或单机部署阶段，缺乏真实服务器环境验证，成果展示多为截图或理论方案，缺乏可运行的系统支撑；数据合规意识薄弱：在网页爬虫类项目中，部分学生未充分考虑robots协议、反爬机制、用户隐私保护等问题，存在潜在法律风险。

3、指导教师指导方面

① 指导力度不足

由于每位指导教师需同时指导多名学生（平均18人），且教师自身承担教学、科研等多重任务，精力分配有限，难以做到全过程精细化指导。部分学生反映：问题反馈周期长，修改意见滞后；对复杂技术问题（如Docker网络配置、Maxwell数据同步异常、API调用限流处理等）指导不够深入；缺乏阶段性检查机制，导致学生前期松懈、后期赶工，影响整体质量。

此外，部分学生在外实习或远程学习，与指导教师沟通频率低，存在“报进度、不解决问题”的现象，导致设计过程走弯路。

② 指导教师技术领域局限

大数据技术涵盖数据工程、AI应用、系统运维、前端展示等多个方向，而部分指导教师的研究背景偏重某一领域（如数据分析或软件开发），在面对涉及容器编排、K8s集群、Spark调优、PostgreSQL高级运维等专业性较强的课题时，指导能力受限，难以提供有效技术支持，影响学生项目深度推进。

4、学生方面

① 少数学生对毕业设计重视程度不够

部分学生将毕业设计视为“走过场”，认为只要完成即可，投入时间与精力严重不足。尤其在就业、考研、考公等压力下，优先级排序错位，导致：文档撰写敷衍，格式不规范；代码缺乏注释，结构混乱；成果演示仅限于基础功能，无优化与扩展；多次拖延提交，依赖教师催促推进。

② 专业知识与技能掌握不扎实

部分学生在日常学习中基础不牢，具体表现为：Python编程能力薄弱，函数封装、异常处理、文件操作等基本语法掌握不熟练；数据处理依赖现成模板，对pandas数据清洗逻辑理解不清；对数据库设计规范（如范式、索引、schema设计）认知模糊；Docker、Git等工具仅会简单命令，不了解底层原理与最佳实践；在遇到技术报错时缺乏独立排查能力，过度依赖搜索引擎或同学帮助。

③ 缺乏自主学习与创新能力

依赖模板化思维：大量学生照搬往届项目结构或网络开源代码，未进行个性化调整与优化，导致项目同质化严重；

缺乏问题意识：对数据来源的真实性、采集频率的合理性、分析模型的适用性缺乏质疑与验证；

创新意识薄弱：多数项目停留在“实现功能”层面，缺少对性能优化、用户体验、系统可扩展性的深入思考；

行业敏感度低：对当前大数据领域的热点（如数据湖、实时计算、AI Agent、DataOps）了解不深，成果难以体现技术前瞻性。

综上所述，本届大数据技术专业毕业设计整体完成质量较高，但在选题创新性、技术深度、工程规范性、师生互动效率等方面仍存在提升空间。

五、毕业设计改进措施

为使大数据技术专业毕业设计具备一定的理论深度，紧密对接行业技术发展趋势与企业实际需求，全面提升学生的工程实践能力、技术创新能力和就业竞争力，指导教师团队在总结2025届及往届毕业生毕业设计中存在的问题基础上，深入反思选题质量、技术应用、指导机制与学生能力等方面的不足，现提出以下针对性改进措施：

① 严控选题质量，推动“真题真做”，提升创新性与实用性

强化选题审核机制：指导教师须对学生选题进行前置评估，重点审查其技术可行性、数据来源可靠性、工程价值与创新性，杜绝“重复性高、内容空泛、脱离实际”的题目。

建立动态选题库：联合合作企业、实习单位每年更新不少于20个真实项目选题，涵盖电商数据分析、招聘平台数据挖掘、智能客服系统开发、实时数据同步、云原生部署优化、AI应用集成等方向，确保选题紧贴岗位需求和技术前沿。

鼓励跨学科融合与竞赛成果转化：支持学生将参与的学科竞赛（如大数据挑战赛、人工智能大赛）、创新创业项目或实习成果转化为毕业设计内容，提升项目深度与应用价值。

引导差异化选题：避免多个学生集中于“淘宝数据采集”类同质化题目，鼓励在数据采集基础上延伸至预测建模、用户画像构建、可视化交互优化、自动化流程设计等高阶方向。

② 统一技术标准与成果规范，提升工程化水平

制定毕业设计技术规范指南：明确代码编写规范（如PEP8）、项目目录结构、文档撰写格式（含开题报告、设计说明书、测试报告）、数据处理流程图、系统架构图等标准，确保成果的专业性与可读性。

统一开发环境与工具版本要求：建议使用Python 3.13+、Docker 24+、MySQL 8.0+等主流稳定版本，避免因环境差异导致部署失败；鼓励使用虚拟环境（conda/venv）管理依赖。

规范数据使用与伦理要求：明确禁止非法爬取敏感信息，要求所有数据采集项目必须说明来源、是否遵守robots协议、是否进行脱敏处理，增强学生的法律意识与职业素养。

推广Git版本管理：要求所有项目使用GitHub/Gitee等平台进行代码托管，提交记录完整，体现开发过程的可追溯性。

③ 合理分配指导任务，保障指导质量与师生互动效率

控制指导人数上限：每位指导教师指导毕业设计学生人数原则上不超过15人，确保有足够精力对每位学生进行阶段性检查与个性化指导。

实行“阶段性检查+定期反馈”机制：设置开题、中期检查、预答辩三个关键节点，要求学生提交阶段性成果（如数据集样本、爬虫日志、系统界面截图、部署流程文档），教师及时给予书面或面对面反馈。

建立多维指导模式：针对涉及不同技术方向（如AI、大数据平台、DevOps）的学生，组建“导师组”或“联合指导小组”，实现专业互补，提升指导的专业性和深度。

加强远程沟通机制：利用企业微信、钉钉、腾讯会议等工具建立常态化沟通渠道，尤其对在外实习学生，确保每周至少一次有效交流。

④ 持续提升师生专业能力，紧跟技术发展前沿

加强教师技术更新培训：组织教师参加大数据、人工智能、云原生等相关领域的技术培训、企业调研或认证考试（如阿里云ACA/ACP、Docker/K8s认证），提升指导教师的前沿技术指导能力。

推动“以研促教、以赛促学”：鼓励教师将科研项目中的数据处理模块、系统开发任务转化为毕业设计课题；支持学生参与Kaggle、天池、华为ICT大赛等高水平竞赛，反哺毕业设计质量提升。

强化学生自主学习能力培养：在日常教学中增设“工具链实训”“项目实战周”等环节，提升Python编程、Linux操作、Docker部署、Git协作等核心技能；开设“大数据行业前沿讲座”系列课程，邀请企业工程师分享真实项目案例，增强学生对新技术（如DataOps、LLM应用、实时数仓）的认知；鼓励学生撰写技术博客、录制演示视频，提升表达能力与成果展示水平。

⑤ 推动成果落地与成果转化，增强社会价值

优秀项目孵化支持：对完成度高、创新性强的毕业设计项目（如AI聊天机器人框架、自动化部署方案、数据同步系统），推荐申请软件著作权、参与创新创业项目申报或对接企业实际应用场景。

建立优秀案例库：每年遴选10%的优秀毕业设计成果，整理成册或制作展示网页，作为后续学生参考范本，形成良性传承机制。

加强校企协同育人：与大数据企业共建“毕业设计联合指导基地”，由企业工程师与校内教师共同命题、共同指导，提升毕业设计的真实性和实用性。

通过以上改进措施的系统实施，旨在构建一个选题科学、标准统一、指导有力、能力扎实、成果可转化的高质量毕业设计管理体系。未来将持续优化毕业设计全过程管理，切实提升大数据技术专业学生的综合实践能力与职业竞争力，为培养适应数字经济时代需求的高素质技术技能人才提供有力支撑。