



漯河职业技术学院

LUO HE VOCATIONAL TECHNOLOGY COLLEGE

漯河职业技术学院（群内）专业 人才培养方案 (2025版)

专业名称: 大数据技术专业

专业代码: 510205

专业大类: 电子与信息大类

所属学院: 人工智能学院

所属专业群: 大数据技术专业群

二〇二五年八月

目 录

一、专业描述	1
二、职业面向	1
(一) 职业面向岗位	1
(二) 职业发展路径及职业能力分析	1
三、培养目标与培养规格	4
(一) 培养目标	4
(二) 培养规格	5
四、人才培养模式	7
五、课程设置及要求	7
六、教学进程总体安排	18
七、实施保障	18
(一) 师资队伍	18
(二) 教学设施	19
(三) 教学资源	20
(四) 教学方法	21
(五) 学习评价	21
(六) 质量管理	22
八、毕业要求	22
(一) 学分条件	22
(二) 证书	22
附录一 大数据技术专业教学进程表	24
附录二 学时与学分分配表	26
编制说明	27

漯河职业技术学院大数据技术专业(群内)人才培养方案

(2025 版)

一、专业描述

专业名称：大数据技术

专业代码：510205

入学要求：中等职业学校毕业、普通高级中学毕业或具备同等学力

基本修业年限：三年

教育类型：高等职业教育

学历层次：专科

所属专业群名称：大数据技术专业群

二、职业面向

(一) 职业面向岗位

表 1 大数据技术专业职业面向岗位一览表

所属专业大类 (代码)	所属专业类 (代码)	对应行业 (代码)	主要职业 类别 (代码)	主要岗位群或 技术领域	职业资格证书和技能 等级证书
电子与信息大 类 (51)	计算机类 (5102)	互联网和相 关服务 (64)、 软件和信息 技术服务业 (65)	大数据工程技术人员 S (2-02-38-03)、数据 分析处理工程技术人员 S (2-02-30-09)、 信息系统运行维护工 程技术人员 S (2-02-10-08)	大数据实施与 运维、数据采集 与处理、大数据 分析与可视化、 大数据平台管 理、大数据技术 服务、大数据产 品运营	大数据分析与应用、大 数据应用开发 (Python)、大数据工 程化处理与应用

(二) 职业发展路径及职业能力分析

表 2 大数据技术专业职业岗位及其岗位能力分析

序号	岗位群	岗位类别		岗位任务描述与核心能力要求	
		入职岗位	发展岗位	岗位任务描述	核心能力要求
1	大数 据运 维	初级系 统运维人 员	大数 据系 统运维工 程师	1. 日常监控与巡检： 负责 7x24 小时监控报警的响 应, 每日对大数据集群进行健康 巡检, 检查磁盘空间、服务状态、 资源使用率等。 2. 基础故障处理： 处理常见的集群告警和故障, 如	1. 熟悉 Linux 操作和网络知识, 熟 练使用常用命令, 具备系统性能分 析能力。 2. 精通 Hadoop (HDFS, YARN)、 Spark 的核心原理、架构和运维。 3. 熟练使用 Shell 进行脚本编写; 掌握 Python 脚本编写, 能够进行

				<p>节点宕机、磁盘空间不足、服务进程异常等。分析问题原因，并执行标准的恢复操作。</p> <p>3. 运维脚本编写与执行： 编写和维护 Shell/Python 脚本，自动化执行日志清理、数据备份、服务重启等重复性任务。</p> <p>4. 文档维护： 更新运维操作手册、故障处理清单等文档；记录每次故障处理的详细过程和原因。</p>	<p>自动化工具和平台运维。</p> <p>4. 熟练撰写并提交日常运维报告。</p>
2	数据采集	初级数据采集工程师	中级数据采集工程师	<p>1. 基础数据采集开发与配置： 使用成熟的采集工具进行数据同步任务的配置、开发和测试；编写和维护数据采集任务的配置脚本；负责从常见数据源到大数据平台（HDFS, Hive, Kafka）的定时批处理同步。</p> <p>2. 日志数据采集： 部署和配置 Logstash 等日志采集 Agent，实现应用日志、Nginx 访问日志等数据的实时或准实时收集；理解日志格式，能进行简单的数据解析和过滤。</p> <p>3. 网站数据采集。</p> <p>4. 任务监控与基础运维： 监控数据采集任务的运行状态、速度和数据质量；处理任务失败、延迟等常见告警，执行任务重跑等基础恢复操作；维护采集任务的运行日志和文档。</p>	<p>1. 熟练掌握 SQL，了解一种或多种关系型数据库（如 MySQL）的基本原理和操作。</p> <p>2. 了解 HDFS、Hive、Kafka 等大数据组件的基本概念，至少熟练使用一种主流的数据采集工具（如 Sqoop, DataX, Logstash, Flume）。</p> <p>3. 能够熟练配置网站数据抓取规则；掌握 Shell 脚本，能进行简单的 Python 脚本编写。</p> <p>4. 熟练使用 Linux 常用命令。</p> <p>5. 能与数据开发、业务方进行有效沟通，能严谨地完成配置和测试。</p>
3	数据标注	数据标注员、标注组长	数据标注工程师	<p>1. 对文本/图像/语音等数据进行标注分类。</p> <p>2. 数据清洗与质量检查。</p> <p>3. 使用标注工具完成标注任务。</p> <p>4. 标注规则培训与质量把控。</p> <p>5. 复杂数据标注方案设计。</p> <p>6. 标注工具开发优化。</p> <p>7. 数据脱敏与格式转换。</p>	<p>1. 具备计算机基础操作能力，以及细心耐心与责任心。</p> <p>2. 熟悉机器学习基础。</p> <p>3. 掌握数据结构/数据库基础知识。</p> <p>4. 熟练掌握 Python 编程。</p> <p>5. 根据详细的标注规则和规范，使用标注工具对原始数据（如图像、文本、语音、视频）进行准确的标注。</p> <p>6. 编写脚本（如 Python）或开发小工具，自动化处理重复性任务。</p>

4	数据预处理	初级数据清洗工程师	中级数据清洗工程师	<p>1. 数据采集与接入: 从多种数据源提取数据;编写脚本或使用工具定期/实时地拉取或接收数据。</p> <p>2. 数据清洗与修正: 处理缺失值:识别缺失数据,并采用适当策略处理; 处理异常值:检测并处理不符合预期的异常数据点,分析其产生原因并决定修正或剔除;格式标准化:统一数据格式,例如日期格式、单位、字符编码等;纠正错误:修正明显的逻辑错误或输入错误。</p> <p>3. 数据转换与集成: 数据规约:对数据进行归一化、标准化,使其适合模型处理;特征工程:进行特征构建、特征变换等;数据融合:将来自不同来源的数据进行关联和合并,形成宽表或统一的数据视图。</p> <p>4. 数据质量监控与验证: 建立数据质量校验规则,定期检查数据的完整性、准确性、唯一性和一致性; 编写自动化脚本或使用数据质量工具对数据管道进行监控和告警。</p> <p>5. 文档编写与协作: 详细记录数据预处理流程、规则、逻辑和数据字典;与数据分析师和业务人员紧密沟通,理解数据需求和数据背后的业务含义。</p>	<p>1. 精通 Python 的 Pandas, NumPy 等库) 和 SQL, 用于数据提取和转换。</p> <p>2. 熟悉至少一种大数据处理工具,如 Spark、Apache Flink 或 Hadoop, 以处理海量数据。</p> <p>3. 熟练操作一种或多种关系型数据库(如 MySQL, PostgreSQL) 和 NoSQL 数据库(如 MongoDB)。</p> <p>4. 了解或使用过 ETL/ELT 调度工具,如 Apache Airflow, DolphinScheduler, Kettle 等。</p> <p>5. 版本控制:熟练使用 Git 进行代码管理。</p> <p>6. 具备基本的 Linux 系统操作能力。</p>
5	数据可视化	初级大数据可视化工程师	中级大数据可视化工程师	<p>1. 看板开发与维护:根据业务需求和分析师的设计原型,使用主流 BI 工具开发和迭代常规数据看板。</p> <p>2. 数据准备与建模:在 BI 工具内进行数据连接、关联、清洗和简单的数据模型构建(如创建计算字段、层次结构)。</p> <p>3. 组件配置与美化:配置图表属性、颜色、滤镜、交互动作,确</p>	<p>1. 熟练掌握至少一门行业标准工具,如 Tableau 或 Power BI, 理解其核心概念(如 LOD 表达式、上下文过滤器、数据混合等)。</p> <p>2. 能够编写高效的 SQL 查询,为可视化准备和聚合数据。</p> <p>3. 熟练掌握 HTML5、CSS3、JavaScript 以及前端框架(如 Vue)。</p> <p>4. 熟练使用 ECharts、AntV 等高级</p>

				<p>保可视化作品的可用性和美观性。</p> <p>4. 定制化可视化开发: 使用前端技术栈 (如 D3. js, ECharts, Three. js) 开发 BI 工具无法满足的、高度定制化的交互式可视化图表和应用。</p>	<p>库。</p> <p>2. 具有数据分析及编写数据分析报告的能力, 以及图形效果的可视化、科学化、美观化的能力。</p>
6	大数据应用开发	初级大数据应用开发 (Python) 工程师	中级大数据应用开发 (Python) 工程师	<p>1. 数据处理脚本: 使用 PySpark 或 Pandas 编写数据处理脚本, 完成数据抽取、转换、加载等任务; 开发和维护批处理或微批处理的数据计算任务, 并部署到 YARN 或 K8s 集群上稳定运行。</p> <p>2. 数据 API 与服务开发: 使用 Flask、Django 等框架, 开发简单的 RESTful API, 为前端应用或业务系统提供数据查询服务; 实现基于数据库或缓存的数据查询接口。</p> <p>3. 数据产品功能模块实现: 在资深工程师的架构设计下, 完成数据产品中具体功能模块的编码和部署。</p> <p>4. 任务监控与运维: 监控自己开发的应用和作业的运行状态, 处理常见异常。</p> <p>5. 撰写相关技术文档和 API 文档。</p>	<p>1. 熟练掌握 Python 编程, 包括面向对象、常用数据结构、异常处理等。熟悉核心库如 requests, json 等。</p> <p>2. 熟练掌握 PySpark 的 DataFrame API 和基本概念, 了解其运行原理。熟悉 Pandas 进行中小规模数据处理。</p> <p>3. 掌握至少一种 Python Web 框架 (如 Flask/FastAPI) 进行 API 开发。</p> <p>4. 熟练使用 SQL, 能操作一种以上数据库。</p>

三、培养目标与培养规格

(一) 培养目标

本专业依托河南省软件与信息技术服务产业集群、区域大数据产业园区行业优势, 紧跟大数据技术与数字经济融合发展趋势, 在专业中设立“数据标注”微专业方向, 与河南的重点产业领域直接对应。培养能够践行社会主义核心价值观, 传承技能文明, 德智体美劳全面发展, 具有一定的科学文化水平, 良好的人文素养、科学素养、数字素养、职业道德、创新意识, 爱岗敬业的职业精神和精益求精的工匠精神, 较强的就业创业能力和可持续发展的能力, 掌握本专业知识和技术技能, 具备职业综合素质和行动能力, 面向软件和信息技术服务、互联网和相关服务等行业的[大数据工程技术人员](#)、[数据分析处理工程技术人员](#)、[信息系统运行维护工程技术人员](#)等职业,

能够从事大数据实施与运维、数据采集与处理、大数据分析与可视化、大数据平台管理、大数据技术服务、大数据产品运营等工作的高技能人才。

（二）培养规格

本专业学生应在系统学习本专业知识并完成有关实习实训基础上，全面提升知识、能力、素质，掌握并实际运用岗位需要的专业核心技术技能，实现德智体美劳全面发展，总体上须达到以下要求：

1. 素质

（1）坚定拥护中国共产党领导和中国特色社会主义制度，以习近平新时代中国特色社会主义思想为指导，践行社会主义核心价值观，具有坚定的理想信念、深厚的爱国情感和中华民族自豪感；

（2）掌握与本专业对应职业活动相关的国家法律、行业规定，掌握绿色生产、环境保护、安全防护、质量管理等相关知识与技能，了解相关行业文化，具有爱岗敬业的职业精神，遵守职业道德准则和行为规范，具备社会责任感和担当精神；

（3）具有质量意识、环保意识、安全意识、信息素养、工匠精神、创新思维，具有学以致用、爱岗敬业的职业理念和服务“数字化转型”，支撑智慧城市、工业互联网等国家战略的职业理想；

（4）掌握身体运动的基本知识和至少1项体育运动技能，达到国家大学生体质健康测试合格标准，养成良好的运动习惯、卫生习惯和行为习惯；具备一定的心理调适能力；

（5）掌握必备的美育知识，具有一定的文化修养、审美能力，形成至少1项艺术特长或爱好；

（6）树立正确的劳动观，尊重劳动，热爱劳动，具备与本专业职业发展相适应的劳动素养，弘扬劳模精神、劳动精神、工匠精神，弘扬劳动光荣、技能宝贵、创造伟大的时代风尚。

2. 知识

（1）掌握支撑本专业学习和可持续发展必备的数学、外语、信息技术等文化基础知识，具有良好的人文素养与科学素养，具备职业生涯规划能力；

（2）具有良好的语言表达能力、文字表达能力、沟通合作能力，具有较强的集体意识和团队合作意识，学习1门外语并结合本专业加以运用；

（3）掌握必备的军事理论知识、心理健康知识、创新创业知识、职业发展与就业指导知识、大数据技术专业素养知识；

（4）掌握数据库基本原理、程序设计、操作系统原理、计算机网络、云计算等方面的专业基础理论知识；

- (5) 掌握大数据采集与大数据预处理技术技能，具有数据采集、抽取、清洗、转换与加载等数据预处理能力；
- (6) 掌握数据分析和数据挖掘应用技术技能，具有面向业务需求，基于大数据分析平台进行数据的批量、实时、分布式计算，基础特征工程处理以及机器学习算法应用等大数据分析挖掘实践能力；
- (7) 具有数据可视化设计和数据分析报告撰写能力，具有开发应用程序进行数据可视化展示、撰写数据可视化结果分析报告等实践能力；
- (8) 掌握大数据平台搭建与部署、大数据平台运维、数据库开发与管理等技术技能，具有大数据平台部署与运维、数据库管理与应用、大数据技术服务、大数据产品运营、大数据平台管理等实践能力；
- (9) 掌握信息技术基础知识，具有适应本行业数字化和智能化发展需求的数字技能；
- (10) 具有探究学习、终身学习和可持续发展的能力，具有整合知识和综合运用知识分析问题和解决问题的能力。

3. 能力

- (1) 具有探究学习、终身学习、分析问题和解决问题的能力；
- (2) 具有良好的语言、文字表达能力和沟通能力；
- (3) 具有文字、表格、图像等计算机处理能力，本专业必需的信息技术应用能力；
- (4) 具备良好的团队协作与抗压能力；
- (5) 具备较强的创新创业能力；
- (6) 能够阅读并正确理解大数据项目需求分析报告和项目建设方案；
- (7) 具备 Hadoop 平台下的各种应用类型数据的抓取和收集能力；
- (8) 具备对数据进行整理、计算，对计算结果进行综合分析的能力；
- (9) 具备整合不同数据源、不同结构类型数据的能力和探索数据背后价值的能力；
- (10) 具备 SQL 的计算与存储过程调优，并具有严密的逻辑分析能力；
- (11) 能运用 Hadoop 及其生态圈产品搭建大数据平台的能力；
- (12) 具备将分析数据进行数据展示的能力；
- (13) 具备大数据集群环境的管理和优化的能力；
- (14) 具备大数据挖掘、建立基本模型或方法的能力；
- (15) 具备大数据采集与预处理能力。

4. 职业态度

- (1) 严守《数据安全法》，不泄露隐私与核心数据，确保数据采集处理合法，守好行业伦理底线；
- (2) 严谨负责：全流程把控数据质量，排查异常、监控运维，失误后主动担责复盘，保障服务稳定；
- (3) 持续学习，紧跟技术迭代，学新工具算法，探究原理，避免技术滞后；
- (4) 业务导向，懂业务需求，用通俗语言解读数据，输出落地建议，赋能业务；
- (5) 高效协作，跨部门清晰沟通，团队内分享经验，打破协作壁垒；
- (6) 创新破局，遇瓶颈不局限传统，试新方法，缺数据时找替代方案；
- (7) 细节把控，重视格式、逻辑，不忽视小细节，防结果偏差；
- (8) 服务保障，预判风险做预案，及时响应需求，保数据服务高效。

四、人才培养模式

根据专业人才培养目标，本专业采用“三能进阶、四链融合”人才培养模式。按照学生的认知规律和职业成长规律，以“基础能力（掌握 SQL、Python 等大数据基础技术）、核心能力（精通数据采集清洗、大数据平台运维等岗位技能）、创新能力（具备数据建模优化、业务解决方案设计能力）”为培养目标，构建“基础-核心-创新”的“三能进阶”递进式人才培养的能力体系。以漯河及豫中南大数据产业需求对接专业建设，以大数据运维、数据分析等职业岗位要求设置课程体系，以企业真实业务流程（如数据处理项目全流程）设计教学过程及实践环节，将工作任务（如客户行为数据分析、集群故障排查）对接教学内容，深化“岗课赛证”融合，联合本地数字产业企业、行业协会推进“产业链、教育链、人才链、创新链”四链融合，实现人才培养与岗位需求精准适配。

五、课程设置及要求

主要包括公共基础课程和专业（技能）课程。

1. 公共基础课程

见大数据技术专业群公共基础课程内容。

2. 专业（技能）课程

(1) 专业群共享及专业基础课

本专业开设的专业群共享共 4 门，包括程序设计基础、Web 前端技术、数据库技术、Python 编程基础，见大数据技术专业群共享课程内容。

本专业基础课程为 Linux 操作系统。

表 3 专业基础课程描述

课程代码	课程名称	课程目标	主要内容	教学要求
023519	Linux 操作系统	<p>1. 掌握 Linux 操作系统的基本架构、核心概念及其在大数据技术栈中的基础地位。深入理解 Linux 文件系统、权限管理、进程管理和网络配置原理。</p> <p>2. 能够熟练使用命令行界面 (CLI) 高效地进行系统操作、日志分析和故障排查。</p> <p>3. 具备编写复杂 Shell 脚本的能力，以自动化大数据环境下的日常维护任务。</p> <p>4. 能够独立完成 Linux 服务器的系统配置、网络设置、软件安装，为部署 Hadoop、Spark、Hive 等大数据组件打下坚实基础。</p> <p>5. 具备初步的多节点集群管理意识和能力。</p> <p>6. 培养通过命令行解决复杂问题的思维方式和严谨的系统管理习惯。</p> <p>7. 建立良好的网络安全和权限管理意识，理解大数据平台安全的基础。</p>	<p>1. Linux 基础与大数据环境概述：Linux 与开源文化，服务器版本 Linux 安装及最小化原则。</p> <p>2. 高级命令行操作与高效文本处理：强大的命令行工具：grep, awk, sed 等，重点讲解其在日志分析、数据清洗中的应用；高级 I/O 重定向、管道和命令组合技巧。</p> <p>3. 用户、权限与安全：大数据集群中服务账户的管理；SSH 密钥对生成与无密码登录配置；文件权限、访问控制列表 (ACL) 和 SELinux 基础概念。</p> <p>4. 系统管理与进程调度：系统服务管理，用于启动/停止 HDFS、YARN 等服务。进程管理、资源监控及性能瓶颈初步分析；定时任务配置，用于定时执行数据采集、ETL 任务等。</p> <p>5. 网络配置与软件包管理：网络配置、防火墙管理，确保集群节点间通信畅通；使用 YUM/APT 管理 Java 环境、Python 环境及各类依赖库。</p> <p>6. Shell 脚本编程进阶：脚本调试技巧；函数、数组、复杂条件判断和循环。</p> <p>7. 大数据平台基础实践：在单机/多台虚拟机上配置 Java 环境；部署并启动一个伪分布式 Hadoop 集群，验证 Linux 基础技能的掌握情况。</p>	<p>1. 在实训室实施“教、学、做”一体化教学模式；采用启发式、任务驱动式、项目式等教学方法实施教学；采取过程性考核与终结性考核相结合的方式进行考核评价，突出对学生的大数据平台部署与运维能力的考核。</p> <p>2. 教学案例应紧密结合大数据应用场景，如使用 Linux 命令分析模拟的 Web 日志、服务器性能数据等。</p> <p>3. 引导学生树立“Linux 是大数据技术的基石”这一核心观念，避免理论与应用脱节。</p> <p>4. 提供接近生产环境的实验任务，培养学生的实战能力。</p>

(2) 专业核心课

本专业开设专业核心课程 8 门，包括：大数据平台部署与运维、数据预处理技术、Spark 应用

技术、大数据分析技术应用、数据采集技术、数据挖掘应用、数据可视化技术与应用、大数据项目管理。

表 4 专业核心课程描述

课程代码	课程名称	课程目标	主要内容	教学要求
023523	大数据平台部署与运维	<p>1. 集群全生命周期管理能力：掌握 Hadoop 集群的规划、部署、配置、启停、扩缩容及版本升级的核心原理与实操技能，能够独立完成集群环境搭建与资源调配。</p> <p>2. 运维监控与健康保障能力：具备对集群硬件资源 (CPU、内存、磁盘、网络) 及核心服务 (HDFS、YARN 等) 运行状态的监控能力，能通过监控图表分析系统性能并提出优化建议。</p> <p>3. 故障诊断与数据恢复能力：掌握日志分析、命令行工具使用及故障场景模拟方法，能够快速定位并解决常见组件级故障 (如 DataNode 宕机、服务启动失败)，并制定数据备份与恢复预案。</p> <p>4. 运维自动化与效率提升能力：能够运用 Shell/Python 脚本实现日志清理、服务巡检、批量配置等重复任务的自动化，培养智能化运维思维。</p> <p>5. 数据安全意识：在运维过程中严格遵守数据安全规范，具备集群权限管理、敏感</p>	<p>1. 集群规划与部署实践：系统环境配置 (虚拟化平台、操作系统、依赖库)；Hadoop 生态组件 (HDFS、YARN、MapReduce 等) 部署与参数调优；集群节点扩缩容与版本升级流程实战。</p> <p>2. 监控体系搭建与运维工具应用：集群健康度监控 (Prometheus/Grafana 部署与指标采集)；服务级监控 (HDFS 存储平衡、YARN 任务队列状态告警)；监控数据解读与性能优化报告撰写。</p> <p>3. 故障场景库与恢复机制：常见故障分类 (节点宕机、数据块损坏、服务不可用)；日志分析工具 (ELK 栈) 与诊断命令 (如 hdfs dfsadmin、yarn application)；数据备份策略 (快照、异地容灾) 与恢复演练。</p> <p>4. 自动化运维脚本开发：Shell/Python 脚本编写规范 (参数化、错误处理)；自动化任务案例如日志轮转、服务健康巡检、配置批量下发；脚本集成与调度 (Crontab、Airflow 基础)。</p> <p>5. 数据安全规范：集群权限管理 (Kerberos 认证、HDFS ACL)；审计日志分析与合规性检查。</p>	<p>1. 能独立完成 3 节点 Hadoop 集群的部署与调试，服务启动一次成功率 $\geq 95\%$。</p> <p>2. 能熟练搭建 Hive 数据仓库，完成复杂 SQL 查询；能配置 Sqoop 实现 MySQL 与 HDFS 数据互通，调试 Flume 完成日志采集，操作熟练度达 90% 以上。</p> <p>3. 能独立完成离线数仓项目的部署测试，可定位并解决项目中 80% 以上的常见问题，具备团队协作完成项目的能力。</p> <p>4. 根据模拟故障场景 (如 DataNode 离线)，在 30 分钟内定位原因并恢复服务，解决率 $\geq 80\%$。</p> <p>5. 自动化脚本：编写 Python 脚本实现每日自动巡检并输出报告，代码规范性与功能完成</p>

		操作审计的意识和能力。		度 $\geq 90\%$ 。
023524	数据预处理技术	<p>1. 熟悉数据 ETL 基础概念与流程，了解多源数据特性及质量评估标准。</p> <p>2. 熟练掌握 ETL 工具配置、数据清洗方法、多源数据抽取加载与迁移装载技能，初步具备数据格式转换与整合优化能力。</p> <p>3. 形成数据质量管控意识，养成适配业务场景选择预处理方案的思维。</p>	<p>1. ETL 基础和 Python：核心概念、三阶段流程、多源数据（文本、网页、数据库）特性分析。</p> <p>2. 工具实操：Kettle 和 Python 等工具的安装配置与基础功能操作。</p> <p>3. 数据清洗：缺失值、重复值、不一致数值的识别方法与针对性处理方案。</p> <p>4. 抽取装载：文本/网页/数据库数据的抽取技巧，多数据源迁移与目标存储装载流程。</p> <p>5. 转换整合：常见数据格式转换实操，多源数据匹配与整合优化方法。</p>	<p>1. 能准确描述 ETL 流程与多源数据特性，区分工具适用场景，理论准确率大于 95%。</p> <p>2. 能独立完成 2 种工具跨系统配置，操作熟练度大于 90%。</p> <p>3. 能精准识别数据质量问题，处理后数据完整性与一致性大于 90%。</p> <p>4. 能独立完成三类数据抽取（完整率$\geq 98\%$）与迁移装载。</p> <p>5. 能完成 3 种以上格式转换，整合 2 个数据源，独立解决全流项目 80% 以上问题。</p>
023525	Spark 应用技术	<p>1. 掌握 Spark 核心架构、运行原理及核心技术概念（RDD, DataFrame/Dataset）。</p> <p>2. 理解 Structured Streaming 流处理的基本原理。</p> <p>3. 具备初步的性能监控、数据倾斜处理与参数调优能力。</p> <p>4. 形成分布式系统思维和高性能编程规范意识。</p>	<p>1. Spark 核心概念、运行模式、RDD 编程模型（创建、转换/行动操作、持久化）。</p> <p>2. Spark SQL 与结构化数据处理。</p> <p>3. Spark 流处理基本概念以及与批处理的区别。</p> <p>4. Scala 语言基础知识。</p> <p>5. Spark 集群架构与运行模式。</p> <p>6. Spark 任务的三种提交模式。</p> <p>7. 基于 Spark 的数据分析综合应用。</p>	<p>1. 能使用 Scala/Python 独立编写、调试 Spark 分布式应用程序。</p> <p>2. 能针对典型业务场景（如日志分析、数据清洗）设计 Spark 解决方案。</p> <p>3. 能独立安装部署 Spark 集群。</p> <p>4. 熟练使用 Spark 任务的三种提交模式。</p>

023526	大数据分析技术应用	<p>1. 理解大数据分析的流程（数据采集—清洗—建模—可视化），掌握聚类、回归、分类等核心算法原理，熟悉 Hadoop、Spark 等分析框架的应用场景。</p> <p>2. 能运用 Python、Spark SQL 处理海量数据，实现数据可视化，具备基于分析结果提出业务决策建议的能力。</p> <p>3. 培养数据敏感度与逻辑分析能力，提升跨场景问题转化能力，强化严谨的数据分析伦理意识。</p>	<p>1. 涵盖大数据分析概述、数据类型（结构化/非结构化）及分析流程，讲解描述性/预测性/诊断性分析的差异与适用场景。</p> <p>2. Python 数据处理（如清洗、集成、转换等）、经典算法（K-Means 聚类、线性回归）应用，介绍 Spark 离线分析与 Flink 实时分析的实现。</p> <p>3. 讲解数据可视化原则，实操制作仪表盘，结合电商销量预测、用户行为分析等案例落地分析方案。</p> <p>4. 包含数据集实操、完整项目实战（从需求拆解到报告输出）。</p>	<p>1. 采用案例驱动（如用外卖平台数据讲解用户画像）、算法可视化演示，引入大模型辅助分析等前沿趋势。</p> <p>2. 提供电商、金融等领域真实数据集，明确代码规范与分析报告标准，加强工具实操指导，组织项目成果展示与点评。</p>
023527	数据采集技术	<p>围绕数据采集工程师岗位职责，培养学生具备多源数据采集的方案设计、工具部署、流程实施及合规管理能力。</p> <p>1. 数据采集方案设计能力：能够根据业务需求（如舆情监控、用户行为分析）制定数据采集方案，包括数据源识别、技术选型（爬虫/Flume 等）、合规性评估。</p> <p>2. 工具部署与配置能力：掌握 Python 爬虫框架（Scrapy）、日志采集工具（Flume）、数据同步工具（Sqoop/DataX）的部署、核心参数配置及高可用架构搭建。</p> <p>3. 多源数据采集实施能力：能够从网页、</p>	<p>1. 数据采集基础与法规数据采集概念/流程（全量采集、自动化优先）；相关法律法规与伦理规范；需求分析文档编写。</p> <p>2. 网络爬虫技术：Python 爬虫基础（Requests/BeautifulSoup）；动态页面采集（Selenium）；反爬机制应对（代理轮询、UA 模拟）；Scrapy 框架与分布式爬虫。</p> <p>3. 日志与实时数据采集 – Flume 架构与组件（Source/Channel/Sink）；负载均衡/故障转移配置；日志解析与 Kafka 集成。</p> <p>4. 数据库与增量数据同步 Sqoop/DataX 数据迁移（RDBMS→HDFS/Hive）；Canal 解析 MySQL binlog 实现增量同步。</p>	<p>独立完成 Flume 三节点高可用集群部署，配置一次成功率≥95%。</p> <p>编写 Python 爬虫采集指定电商网站评论数据，数据采集完整性≥90%。</p> <p>使用 Sqoop 完成 MySQL 至 Hive 的定时同步任务；Flume 收集点击日志，完成率>90%。</p> <p>能进行数据采集中的故障处理，如根据模拟的 Flume 通道阻塞场景，15 分钟内定位问题并恢复数据流。</p> <p>小组协作完成</p>

		<p>日志文件、数据库、API 等来源完成结构化/半结构化数据的采集、清洗和存储(如HDFS、Hive)。</p> <p>4. 质量监控与故障处理能力：具备采集任务监控、数据质量校验、常见故障(如爬虫封禁、Flume 通道阻塞)的诊断与恢复能力。</p> <p>5. 合规与安全意识：遵循数据安全法律法规(如个人信息保护法)，在采集过程中实施数据脱敏、权限控制和审计追踪。</p>	<p>5. 综合项目实战案例：电商用户行为数据采集(爬虫抓取商品评论+Flume 收集点击日志+Sqoop 同步订单数据)。</p>	<p>综合项目，分工明确、文档规范(编写项目报告中需包含数据来源合规性说明，违规操作次数为0)，团队互评满意度$\geq 85\%$。</p>
023528	数据挖掘应用	<p>1. 掌握机器学习的基础概念、原理、算法和模型。</p> <p>2. 能够将机器学习技术应用于实际项目中，实现产品的智能化升级和性能提升。</p> <p>3. 结合业务背景，对数据进行探索。</p> <p>4. 评估挖掘需求并选择合适方法对数据进行特征工程处理。</p> <p>5. 调用经典算法进行模型训练。</p> <p>6. 根据合适评价指标对模型进行验证和测试。</p> <p>7. 培养将数据挖掘技术应用于解决现实世界问题的能力。</p>	<p>1. 熟悉数据特征管理的基础知识。</p> <p>2. 熟悉监督学习、无监督学习、半监督学习的概念及应用。</p> <p>3. 熟悉回归、分类、关联、聚类等算法原理及应用；</p> <p>4. 掌握训练集、验证集、测试集的基本应用。</p> <p>5. 了解机器学习算法基础应用经典模型的原理及过程。</p> <p>6. 了解模型性能的计算和评价方法。</p>	<p>充分利用 AI 工具，采用任务驱动式、项目式、案例式等教学方法实施教学。采取过程性考核与终结性考核相结合的方式进行考核评价，突出对学生程序设计能力的考核。</p> <p>能够针对一个具体的业务问题，独立完成从数据理解到模型构建与评估的全过程。</p> <p>能结合数据背景、模型评估等对挖掘结果进行分析。</p>
023529	数据可视化技术与应用	<p>1. 可视化方案设计能力：能够根据业务需求(如销售分析、用户行为洞察)制定数</p>	<p>1. 可视化基础与设计原则：数据可视化概念与价值(如“文不如表，表不如图”)、视觉编码原则</p>	<p>1. 能独立使用 Echarts 配置包含5种以上图表类型的交互式</p>

		<p>据可视化方案，包括数据源整合、图表类型选择、交互设计及色彩搭配等。</p> <p>2. 工具掌握与配置能力：熟练使用 Echarts、PowerBI 等主流可视化工具，完成数据连接、图表配置、仪表板设计及发布分享。</p> <p>3. 多维度数据展示能力：能够针对结构化数据（如销售报表）、地理空间数据（如区域销售分布）及时间序列数据（如月度趋势）设计合适的可视化图表。</p> <p>4. 可视化交互与叙事能力：具备添加交互元素（如筛选、下钻）、设计动态图表及编写数据故事的能力，提升可视化成果的沟通效果。</p> <p>5. 可视化规范与审美素养：遵循可视化设计原则（如简洁性、准确性），在设计中合理运用色彩、布局。</p>	<p>（位置、长度、颜色等）、常见图表类型适用场景（柱状图对比、折线图趋势、饼图占比）。</p> <p>2. Echarts 可视化开发-Echarts 基础配置（option 对象结构）、数据绑定（dataset 使用）；常见二维/三维图表绘制（柱状图、折线图、散点图、饼图）；</p> <p>交互功能实现（提示框、图例筛选、数据区域缩放）；</p> <p>地理空间可视化（地图坐标系、热力图）。</p> <p>3. PowerBI 商业智能分析. PowerBI 数据连接（Excel/SQL/API）、数据建模（关系建立、度量值创建）；可视化图表制作（堆积柱状图、树状图、卡片图）；仪表板设计（页面布局、主题应用）与发布分享。</p> <p>4. 项目案例：电商销售数据分析仪表板（Echarts 实现商品销量排行、PowerBI 制作区域销售分布与用户行为下钻分析）。</p>	<p>可视化页面，功能完整率\geq95%。</p> <p>2. 能使用 PowerBI 完成多数据源连接并发布交互式仪表板，完成度$>90\%$。</p> <p>3. 能根据给定数据集（如销售数据），选择合适的图表类型完成可视化，图表选择合理率$\geq90\%$。</p> <p>4. 能为可视化项目添加筛选、下钻等交互功能，交互流畅度达标率$\geq85\%$。</p>
--	--	---	--	---

023530	大数据项目管理	<p>1. 熟悉大数据项目各阶段核心流程，了解需求分析、数据处理、计算架构、存储管理及可视化的关键技术标准与方法论。</p> <p>2. 能结合行业案例拆解业务需求并完成技术方案设计；能实操数据采集预处理、分布式计算、数据仓库搭建及可视化呈现全流程；能优化项目关键环节的实施效率。</p> <p>3. 养成以业务价值为导向的项目规划思维，强化跨环节协同的团队合作意识，树立数据项目质量与风险管理理念。</p>	<p>1. 需求分析与方案设计：通过行业案例进行业务需求拆解，完成技术可行性评估与架构设计。</p> <p>2. 数据采集与预处理：实践实时数据采集、离线数据同步，完成数据清洗与特征工程。</p> <p>3. 分布式计算实践：进行离线批处理开发，构建实时计算链路，实现架构应用。</p> <p>4. 数据仓库存储与管理：对比 HDFS/HBase 等存储系统的应用场景，完成数据分层设计与分区策略优化。</p> <p>5. 数据可视化，利用可视化工具进行数据可视化，将分析结果以直观的方式呈现出来。</p>	<p>1. 能结合行业案例完成需求拆解，撰写技术可行性报告，设计架构蓝图，方案通过率\geq85%。</p> <p>2. 能独立完成数据清洗与特征工程，处理后数据适配性\geq90%。</p> <p>3. 能准确区分存储系统适用场景，完成数据分层设计，优化分区策略。</p> <p>4. 能使用工具设计 3 类以上图表，制作交互式仪表盘，呈现逻辑清晰度\geq90%；能独立完成项目全流程，解决 80% 以上常见问题，具备团队协作能力。</p>
--------	---------	--	---	---

(3) 专业拓展课

本专业开设 6 门专业拓展课程，包括体现行业发展新技术的 NoSQL 数据库技术应用、云计算平台技术应用、人工智能导论、Flink 应用技术、大数据安全技术、大数据产品运营 6 门课程，学生任选其中 3 门。

表 5 专业拓展课程描述

课程代码	课程名称	课程目标	主要内容	教学要求
023531	NoSQL 数据库技术应用	<p>1. 建立非结构化数据管理的系统性思维，理解分布式存储场景下的 CAP 理论权衡。</p> <p>2. 理解最终一致性、分区容错等 BASE 理论的核心概念。</p> <p>3. 熟悉列式存储/文</p>	<p>1. NoSQL 产生的原因以及优势；</p> <p>2. 熟悉键值对存储数据库 Redis 等。</p> <p>3. 文档存储数据库 MongoDB 分布式数据库技术的原理和实践。</p> <p>4. 列式存储数据库 HBase 分布式数据库技术及其核心技术原理与实践。</p>	<p>1. 借助 AI 工具，采用任务驱动式、小组讨论、情境教学等教学法。</p> <p>2. 能够根据</p>

		<p>档存储/键值对存储等数据模型。</p> <p>4. 熟悉 MongoDB 完成地理空间数据查询；基于 Redis 实现缓存策略。</p>	<p>5. 图形存储数据库 neo4j 的原理与应用。</p> <p>6. HBase 和 MongoDB 数据库技术的综合实验。</p>	业务场景选择合适的数 据存储方 案。3. 了解 设计跨数据 库系统的数 据同步方 案；优化非 结构化数据 的读写性 能。
023532	云计算平 台技术应 用	<p>1. 熟悉系统分布式架构、ArkUI 框架核心概念，了解应用开发全流程规范与多设备适配原理。</p> <p>2. 能运用 ArkUI 完成界面开发；能实现组件交互、页面跳转等基础功能；能适配多设备场景并处理分布式数据同步；能完成应用调试与打包发布。</p> <p>3. 养成组件化、模块化的开发思维，强化跨设备适配的兼容性意识，树立以用户体验为核心的产品设计理念。</p>	<p>1. 开发基础：系统架构（分布式软总线、数据管理等）、开发环境（DevEco Studio）安装配置、ArkUI 框架语法与组件库详解。</p> <p>2. 界面开发：布局组件使用、常用控件配置、样式与动画效果实现，多分辨率适配技巧。</p> <p>3. 功能开发实践：页面路由与参数传递、事件处理（点击/滑动等）、数据存储）操作，网络请求实现。</p> <p>4. 分布式能力适配：多设备互联原理、设备管理与服务调度、分布式数据同步实操，跨设备应用流转实现方案。</p> <p>5. 调试与发布：常见 Bug 定位与修复方法、应用签名配置、多设备兼容性测试，华为应用市场发布流程。</p>	<p>1. 能准确描述系统架构与 ArkUI 核心原理（准确率$\geq 95\%$），独立完成开发环境配置（成功率 100%），掌握组件使用方法。</p> <p>2. 能设计 3 类以上典型界面（布局合理性$\geq 90\%$），实现页面跳转与数据交互（功能完整性$\geq 95\%$），动画效果流畅。</p> <p>3. 能独立开发 1 个包含 UI 界面、核心功能与多设备适配的完整应用，通过测试验收，具备团队协作开发中小型项目的能力。</p>

023533	人工智能导论	<p>1. 培养学生适应管理、服务等一线岗位需求的能力。</p> <p>2. 使学生掌握数据标注与处理的理论基础知识，具备实际操作技能。</p> <p>3. 培养学生为数据素材打上不同标签的能力，以满足多样化人工智能应用的需求。</p> <p>4. 提升学生的AI素养与职业责任感，强化岗位适配能力，使其成为符合行业需求的高技能人才。</p>	<p>1. 熟知人工智能原理技术。</p> <p>2. 探究人工智能支撑技术。</p> <p>3. 跟踪 AIGC 技术应用。</p> <p>4. 图像标注：学习图像分类、标框、描点等标注方法。</p> <p>5. 文本标注：掌握文本数据的分类、标签标注等操作。</p> <p>6. 语音标注：学习语音数据的标注与处理技术。</p> <p>7. 视频标注：掌握视频素材的分类、关键帧标注等操作。</p> <p>8. 数据标注与处理的应用：了解数据标注在交通、金融、医疗、安防等行业的实际应用场景。</p>	<p>1. 掌握数据标注与处理的基础知识；</p> <p>2. 熟练完成图像、视频、语音、文本等素材的分类、标框、描点等标注操作，能为数据素材打上符合不同人工智能应用需求的标签；</p> <p>3. 采用情境教学法组织知识点和教学内容，根据场景或需要完成的任务逐渐展开教学内容，提高课堂教学效果。</p>
023534	Flink 应用技术	<p>1. 实时计算方案设计能力：能够根据业务需求（如实时监控、实时 ETL）制定 Flink 流处理方案，包括数据源接入、流处理逻辑设计及结果输出方式选择。</p> <p>2. Flink 应用开发能力：掌握 Flink 核心 API（DataStream/Table API）的使用，能够编写实时数据处理程序，实现数据转换、窗口聚合、状态管理等功能。</p> <p>3. 集群部署与运维能力：能够独立编写 Flink 程序实现实时数据过滤、聚合及结果输出，代码规范性与功能完成度$\geq 90\%$。</p>	<p>Flink 架构与核心组件（JobManager/TaskManager）；流处理基础概念（事件时间、处理时间、状态）；实时计算应用场景（实时 ETL、实时监控）。数据源接入（Socket/Kafka）；转换算子（map/filter/keyBy）；窗口操作（滚动/滑动/会话窗口）；状态管理与容错机制（Checkpoint）。集群部署与运维。集群部署模式（Standalone/YARN）；作业提交与监控；性能调优（并行度/内存配置）。</p> <p>综合项目案例：实时用户行为分析系统（Flink 消费 Kafka 日志流，实时计算 PV/UV，结果写入</p>	<p>2. 完成 Flink 集群部署，成功提交作业并监控运行状态，操作熟练度$\geq 95\%$。</p> <p>3. 根据模拟的作业失败</p>

		<p>力：能够独立完成 Flink 集群的部署、配置及日常监控，具备作业提交、性能调优及常见故障处理的能力。</p> <p>4. 实时数据管道构建能力：能够使用 Flink 整合 Kafka、MySQL 等数据源，构建端到端的实时数据采集与处理管道。</p>	<p>MySQL 供可视化展示）。</p>	<p>场景，20 分钟内定位原因并恢复，解决率$\geq 80\%$。</p> <p>4. 小组协作完成综合项目，分工明确、文档规范，团队互评满意度$\geq 85\%$。</p>
023535	大数据安全技术	<p>1. 认知大数据安全生命周期安全威胁，掌握密码学、访问控制等核心技术原理，熟悉《数据安全法》等法规标准。</p> <p>2. 能分析业务安全需求并设计防护方案，熟练使用工具，具备安全事件应急响应能力。</p> <p>3. 培养安全防范意识与职业道德，提升团队协作及跨专业沟通能力。</p>	<p>1. 涵盖大数据安全概述、全生命周期威胁（如数据泄露、网络攻击）及相关法规标准。</p> <p>2. 讲解 AES、RSA 等加密算法，RBAC/ABAC 访问控制模型，静态/动态数据加密，多方式身份认证及安全审计技术。</p> <p>3. 围绕采集、传输、存储、处理、共享环节，介绍数据源鉴别、SSL/TLS 协议、HDFS 安全机制、联邦学习、差分隐私等防护手段；</p> <p>4. 安全工具实操与项目实践，加密数据、构建防护体系。</p>	<p>1. 采用案例分析、小组讨论等方式，结合数据泄露案例授课，及时引入区块链等新技术应用。</p> <p>2. 会制定详细计划与考核标准，提供实验室资源，加强过程指导，组织项目汇报交流。</p>
023536	大数据产品运营	<p>1. 理解大数据产品（如数据中台、BI 工具）的运营逻辑，掌握用户分层、数据指标、商业化变现等核心概念，熟悉行业运营规范。</p> <p>2. 能通过数据洞察用户需求，制定留存策略，熟练使用工具分析运营数据，具备落地运营方案的能力。</p>	<p>1. 涵盖大数据产品类型与运营定位，讲解用户生命周期（获客—激活—留存—变现—推荐）及核心运营指标体系。</p> <p>2. 拆解用户运营（分层运营、社群运营）、内容运营（数据可视化呈现）、商业化运营（广告/订阅/增值服务）的落地方法；</p> <p>3. 可视化分析，通过案例讲解如何用数据优化运营策略（如根据留存数据调整产品功能）。</p> <p>4. 包含模拟运营（如为 BI 工具设</p>	<p>1. 能够为不同类型的大数据产品设计关键绩效指标（KPI）与数据指标体系。</p> <p>2. 能根据真实运营数据集，制定清晰的实操考核标准，加</p>

		<p>3. 培养数据驱动思维与用户同理心，提升跨部门（技术、市场）协作效率，强化结果导向与问题解决意识。</p>	<p>计拉新方案）、真实项目复盘（分析大厂大数据产品运营案例）。</p>	<p>强工具使用指导，组织运营方案答辩与互评。</p> <p>3. 能够运用数据分析方法（如漏斗分析、留存分析、用户分群）定位产品问题、发现增长机会。</p>
--	--	--	--------------------------------------	---

六、教学进程总体安排

见附录一：大数据技术专业教学进程表；附录二：学时与学分分配表。

七、实施保障

（一）师资队伍

表 6 师资队伍结构与配置表

类别	数量	具体要求
师资队伍结构	16	学生数与本专业专兼任教师数比例为 18:1，双师素质教师占专业教师比为 75%。
专业负责人	1	教授，较好地把握国内外大数据技术行业、专业发展，广泛联系行业企业，了解行业企业对大数据技术专业人才的实际需求，教学设计、专业研究能力强，组织开展教科研工作能力强，在大数据技术领域具有一定的专业影响力。
专任教师	11	具有高校教师资格和本专业领域有关证书； 有理想信念、有道德情操、有扎实学识、有仁爱之心； 具有计算机等相关专业研究生及以上学历； 具有扎实的本专业相关理论功底和实践能力； 具有较强的信息化教学能力，能够开展课程教学改革和科学研究； 每 5 年累计不少于 6 个月的企业实践经历。
兼职教师	4	主要从大数据等相关企业、机构聘任，具备良好的思想政治素质、职业道德和工匠精神，具有扎实的大数据技术专业知识和丰富的实际工作经验，具有中级及以上相关专业职称，能承担专业课程教学、实习实训指导和学生职业发展规划指导等教学任务。

（二）教学设施

主要包括能够满足正常的课程教学、实习实训所需的专业教室、实训室和实训基地。

1. 专业教室基本条件

具有利用信息化手段开展混合式教学的条件。具有黑板、多媒体计算机、投影设备、音响设备，具有互联网接入或无线网络环境及网络安全防护措施。安装有应急照明装置，状态良好，符合紧急疏散要求，安防标志明显，保持逃生通道畅通无阻。

2. 校内实训室基本要求

（1）大数据平台部署与运维实训室

具有计算机（云桌面）、管理节点服务器、计算节点服务器、交换机、无线 AP、网络机柜、多媒体中控台、投影仪、电脑桌椅等设备，安装操作系统、办公软件、基础开发软件（Java、Python、Web 前端）、数据预处理软件、数据可视化软件、大数据平台部署与运维实训系统，用于大数据平台部署与运维、数据预处理、数据可视化技术等实训教学。

（2）大数据采集与分析实训室

具有计算机、服务器、交换机、网络机柜、多媒体中控台、投影仪、投影幕、电脑桌椅、交互式电子白板等设备，安装操作系统软件、办公软件、基础开发软件（Java、Python、Web 前端）、数据采集软件、数据存储软件、数据预处理软件、数据分析软件、数据挖掘软件、大数据分析与挖掘平台，用于数据采集技术、数据预处理、大数据分析技术应用、数据挖掘应用、数据可视化技术与应用、基于行业应用的大数据分析项目实践等实训教学。

（3）大数据可视化实训室

具有计算机、实训系统服务器、交换机、网络机柜、多媒体中控台、投影仪、投影幕、电脑桌椅、交互式电子白板等设备，安装操作系统软件、办公软件、数据可视化开发软件、数据可视化实训系统软件、可视化项目软件、行业数据资源包，用于数据采集技术、数据预处理、大数据分析技术应用、数据挖掘应用、数据可视化技术与应用等实训教学。

表 7 校内实践教学条件配置

序号	实验室或实训室名称	实验实训项目名称	主要实验实训仪器设备	备注
1	大数据平台部署与运维实训室	大数据平台部署与运维、数据预处理等。	具有云桌面电脑、服务器、交换机、多媒体中控台、投影仪、投影幕、电脑桌椅等设备，并安装了相关软件。	相关配套器材

2	大数据采集与分析实训室	数据采集技术、大数据分析技术应用、数据挖掘应用等。	具有计算机、服务器、交换机、多媒体中控台、投影仪、投影幕、电脑桌椅等设备，并安装了相关软件。	相关配套器材
3	大数据可视化实训室	数据可视化技术与应用、大数据平台部署与运维、数据预处理等。	具有计算机、服务器、交换机、多媒体中控台、投影仪、投影幕、电脑桌椅等设备，并安装了相关软件。	相关配套器材

表 8 校外实践教学条件配置

序号	实习实训基地名称	实习实训项目名称	备注
1	漯河大数据运营有限公司	大数据可视化、大数据综合项目、大数据平台部署与运维、数据预处理等。	
2	漯河精工数字科技有限公司	数据采集	

（三）教学资源

1. 教材选用

教材选用按照国家规定及学校教材选用制度，优先选用国家级规划教材、行业优秀教材及“十四五”职业教育国家规划教材，如《大数据平台运维实务》《Hadoop 生态系统应用》等，联合企业共同开发活页式、工作手册式校本教材；建立教材审核机制，对教材的思想性、专业性、适用性严格把关，禁止不合格教材进入课堂，确保教材内容贴合“岗课赛证融通”需求。

2. 图书文献配置

图书文献配备围绕大数据技术专业核心能力培养，配置涵盖大数据基础理论、平台运维、数据分析、机器学习等领域的图书文献，包括《大数据技术原理与应用》《Spark 大数据处理技术》等专业书籍，以及《中国大数据发展报告》等行业前沿期刊；同时补充证书培训、技能竞赛相关参考资料，依托学校图书馆线上线下资源，满足学生课内外学习与教师教学科研需求，年均更新专业图书占比不低于 15%。

3. 数字资源配置

搭建“专业 + 产业”双维度数字资源库：一是引入国家高等教育智慧教育平台、超星学习通等平台的优质课程资源（如大数据技术国家级精品在线开放课程）；二是联合合作企业（如河南数字产业有限公司、江苏知途教育科技有限公司）共建企业真实项目案例库（含离线数仓开发、大数据平台故障处理等实操视频）、虚拟仿真实训资源（如 Hadoop 集群虚拟运维环境）；三是

配置数据库技术、大数据分析工具等在线实训软件，实现数字资源与课程教学、实训实操的深度融合，支持学生自主学习与个性化训练。

（四）教学方法

结合大数据技术专业“实践性强、技术迭代快”的特点，采用多元化教学方法，推动“教、学、做”一体化：

项目驱动教学法：以企业真实项目（如电商用户行为数据分析、政务大数据平台运维）为载体，将教学内容拆解为项目任务，引导学生分组完成“需求分析→方案设计→实操落地→成果验收”全流程，提升岗位适配能力。

案例教学法：选取全国职业院校技能大赛典型案例、企业故障处理案例（如 HDFS 集群宕机恢复），通过案例剖析讲解技术原理与实操要点，帮助学生理解理论与实践的关联。

理实一体化教学法：在实训教室同步开展理论讲授与实操训练，如讲解 Hive 组件时，边演示数据仓库搭建流程，边让学生同步操作，即时消化知识点。

线上线下混合教学法：利用数字资源库推送预习视频、拓展资料，课堂聚焦重难点讲解与实操指导，课后通过在线实训平台布置作业、答疑，延伸学习链条。

（五）学习评价

构建“过程性评价 + 终结性评价 + 职业素养评价”三位一体的多元学习评价体系，注重评价的科学性与实用性。

过程性评价（占比 60%）：涵盖课堂表现（如项目讨论参与度）、实操任务（如 Hadoop 组件配置、数据采集实验报告）、阶段性测试（如 SQL 数据分析测验）、证书培训成果，通过多次、多维度评价反映学生学习过程。

终结性评价（占比 30%）：以综合项目考核为主，要求学生独立或小组完成完整项目（如离线数仓开发），从“技术规范性、成果完整性、问题解决能力”等维度评分，替代传统理论考试，凸显实操能力。

职业素养评价（占比 10%）：结合大数据岗位要求，评价学生数据合规意识（如数据保密执行情况）、团队协作表现（如项目分工配合度）、学习态度（如自主学习打卡率）、AI 素养能力，采用教师评价、学生互评、企业导师评价相结合的方式，全面衡量职业素养达标情况。

评价结果应用：建立评价反馈机制，针对评价中发现的问题（如 Hadoop 组件配置操作薄弱），调整教学内容与方法，同时将评价结果作为学生奖学金评定、企业实习推荐的重要依据，激励学生主动提升能力。

（六）质量管理

1. 已经建立专业建设和教学质量诊断与改进机制、专业教学质量监控管理制度，制定了课堂教学、教学评价、实习实训、毕业设计以及专业调研、人才培养方案更新、资源建设等方面标准，通过教学实施、过程监控、质量评价和持续改进，实现人才培养规格。
2. 已经建立教学管理机制，加强日常教学组织运行与管理，定期开展课程建设水平和教学质量诊断与改进，建立了巡课、听课、评教、评学等制度，已建立与企业联动的实践教学环节督导制度，严明教学纪律，强化教学组织功能，定期开展公开课、示范课等教研活动。
3. 已经建立毕业生跟踪反馈机制及社会评价机制，并对生源情况、在校生学业水平、毕业生就业情况等进行分析，定期评价人才培养质量和培养目标达成情况。
4. 专业教研室充分利用评价分析结果，有效改进专业教学，持续提高人才培养质量。

八、毕业要求

本专业学生毕业时应达到培养目标及培养规格的素质、知识和能力等方面要求，同时满足以下条件。

（一）学分条件

本专业学生在毕业前总学分须取得 148 个学分，最低学分要求及所包括内容如下表。

表 9 最低学分要求

课程类别		最低学分
公共基础及素质教育课程	必修课	38
	限选课	8
	任选课	4
	合计	50
专业课程	专业群共享及专业基础课	22
	专业核心课	32
	专业拓展课	12
	合计	66
岗位实习及单列实习实训		32
总计		148

（二）证书

学生在校期间，应考取必要的基本能力证书及职业资格证书，鼓励学生考取多项职（执）业资格证书。

表 10 考取证书一览表

证书类别	证书名称	考证等级要求	备注
基本能力证书	英语等级证书	B 级以上	任选其中（一）项
	普通话证书	乙级以上	
职（执）业资格证书	大数据平台运维	初级	任选其中（一）项
	大数据应用开发（Python）	初级	
	大数据平台运维	中级	
	大数据应用开发（Python）	中级	
	全国计算机等级证书	二级	

附录一 大数据技术专业教学进程表

课程类别	序号	课程名称	课程代码	学时		学分	开课学期与周学时						开课单位	考核方式	
				理论	实践		一	二	三	四	五	六			
公共基础及素质教育课程	必修课	1	思想道德与法治	161010	44	4	3	4/12						马院	考试
		2	毛泽东思想和中国特色社会主义理论体系概论	180100 13	32	4	2		2						考试
		3	习近平新时代中国特色社会主义思想概论	161008	46	8	3			4/14					考试
		4	※形势与政策(一)	161004	8	0	0.25								考查
		5	※形势与政策(二)	161005	8	0	0.25								考查
		6	※形势与政策(三)	161006	8	0	0.25								考查
		7	※形势与政策(四)	161007	8	0	0.25								考查
		8	中国共产党历史	161011	16	0	1		1						考试
		9	※军事理论	231001	36	0	2		2					学工部	考查
		10	军事技能	231006	0	112	2	2周							考查
		11	劳动教育	231003	6	30	2	1	1						考查
		12	※大学生心理健康	231005	36	0	2	2						公共教学部	考查
		13	大学体育(一)	101001	10	26	2	2							考试
		14	大学体育(二)	101002	10	26	2		2					公共体育部	考试
		15	大学体育(三)	101003	10	26	2			2					考试
		16	大学英语(一)	201001	64	0	4	4						公共教学部	考试
		17	大学英语(二)	201002	72	0	4		4						考查
		18	职业生涯规划	181001	18	4	1	1						招生就业处	考查
		19	创新创业教育	181002	16	16	2		1						考查
		20	大学生就业指导	181003	12	4	1				1				考查
		21	※实验室安全教育	141001	8	8	1	1						教务处	考查
		22	国家安全教育	161012	18	0	1			1					马院 考查
		小计			486	268	38	11	11	6	1				
	限选课	23	※美学和艺术史论		16	0	0.5	0.5						公共艺术部	考查
		24	※艺术鉴赏和评论		16	0	0.5	0.5							考查
		25	艺术体验和实践		0	16	1		1						考查
		26	高等数学	101007	72	0	4		4					公共教学部	考试
		27	※职业人文素养		36	0	2	2							考查

任选课	小计			140	16	8	0	5					
	28	公共任选课程		64	0	4						教务处	考查
	小计			64	0	4							
专业群共享及公共基础课程	29	程序设计基础	023518	36	60	6	6					人工智能学院	考试
	30	Linux 操作系统	023519	34	30	4	4						考试
	31	Web 前端技术	023520	20	44	4	4						考查
	32	数据库技术	023521	36	36	4		4					考试
	33	Python 编程基础	023522	18	54	4			4				考试
	小计			144	224	22	14	4	4				
专业技能课程	34	大数据平台部署与运维	023523	18	54	4		4				人工智能学院	考试
	35	数据预处理技术	023524	18	54	4			4				考查
	36	Spark 应用技术	023525	18	54	4			4				考试
	37	大数据分析技术应用	023526	18	54	4				4			考查
	38	数据采集技术	023527	18	54	4				4			考试
	39	数据挖掘应用	023528	18	54	4				4			考查
	40	数据可视化技术与应用	023529	18	54	4				4			考试
	41	大数据项目管理	023530	18	54	4				4			考查
	小计			144	432	32	0	4	8	20			
专业拓展课	42	NoSQL 数据库技术应用	023531	36	36	4			4			人工智能学院	考查
	43	云计算平台技术应用	023532	36	36	4			4				考查
	44	人工智能导论	023533	36	36	4				4			考查
	45	Flink 应用技术	023534	36	36	4				4			考查
	46	大数据安全技术	023535	36	36	4				4			考查
	47	大数据产品运营	023536	36	36	4				4			考查
	小计			108	108	12	0	0	8	4			
岗位实习及单列实习实训	48	毕业设计	024121	0	108	6				18/6		人工智能学院	考查
	49	岗位实习(一)	024122	0	180	10				18/10			考查
	50	岗位实习(二)	024123	0	288	16					18/16		考查
	小计			0	576	32							
教学计划总计				1086	1624	148	25	24	26	25			

备注：1.※表示线上教学课程，课时数不计入周学时，计入总学时，☆表示线上、线下混合教学课程，公共任选课程每学期初由教务处提供公共任选课程目录，学生自由选择。

2.每学期安排 20 周的教学活动，其中第 19、20 周为复习考试时间。

3.美学和艺术史论类含《美术欣赏》《音乐欣赏》2 门课程，学生任选 1 门；艺术鉴赏和评论类含《书法鉴赏》、《影视鉴赏》、《艺术导论》、《舞蹈鉴赏》、《戏剧鉴赏》、《戏曲鉴赏》6 门课程，学生任选 1 门；艺术体验和实践类含《手工剪纸》《硬笔书法》《手机摄影》《手工编织》《戏剧教育》《现代舞》《歌曲演唱》《大学美育》8 门课程，学生任选 1 门。

附录二 学时与学分分配表

课程类型	总学分	总学时	占总学时百分比 (%)	实践学时	占总学时百分比 (%)	选修课学时	占总学时百分比 (%)
公共基础及素质教育课程	50	974	35.94	284	10.48	220	8.11
专业（技能）课程	66	1160	42.80	764	28.19	216	7.97
岗位实习及单列实习实训	32	576	21.25	576	21.25	0	0
总计	148	2710	100	1624	59.92	436	16.08

编制说明

本专业人才培养方案适用于三年全日制高职大数据技术专业，由河职业技术学院商人工智能学院专业建设委员会组织专业教师，由漯河职业技术学院人工智能学院专业建设委员会组织专业教师，与漯河市大数据运营有限公司、漯河精工数字科技有限公司等合作企业的专家共同制订，经学校党委会审定，批准从 2025 级大数据技术学生开始实施。

主要编制人员一览表

序号	姓 名	所在单位	职称/职务	签 名
1	李会凯	漯河职业技术学院	副教授/人工智能学院院长	李会凯
2	王鸿飞	漯河职业技术学院	副教授/人工智能学院副院长	王鸿飞
3	李娜	漯河职业技术学院	教授	李娜
4	林丽丹	漯河职业技术学院	讲师	林丽丹
5	李真真	漯河职业技术学院	助教	李真真
6	曹冰许	漯河职业技术学院	助教	曹冰许
7	韩红宇	漯河职业技术学院	副教授	韩红宇
8	庞建成	漯河职业技术学院	助教	庞建成
9	王宇光	漯河市大数据运营有限公司	工程师/部门经理	王宇光
10	张云龙	漯河精工数字科技有限公司	工程师/部门经理	张云龙

专业负责人: 李娜

复核人: 王鸿飞

人工智能学院院长: 李会凯

漯河职业技术学院
专业人才培养方案论证与审定意见表

专业建设指导委员会成员	姓名	单位	职务/职称	签名
	李会凯	漯河职业技术学院	人工智能学院院长/副教授	李会凯
	王鸿飞	漯河职业技术学院	人工智能学院副院长/副教授	王鸿飞
	赵永乐	漯河职业技术学院	秘书/副教授	赵永乐
	程学军	漯河职业技术学院	教研室主任/教授	程学军
	李娜	漯河职业技术学院	教研室主任/教授	李娜
	王红纪	漯河职业技术学院	教研室主任/副教授	王红纪
	武世栋	中国移动漯河分公司	部门经理/高级工程师	武世栋
	吴铭烨	中国电信漯河分公司	云中台总师/高级工程师	吴铭烨
	王宇光	漯河市大数据运营有限公司	部门经理/工程师	王宇光

论证意见：

本专业群人才培养方案编制规范，科学合理，符合《国家职业教育改革实施方案》《河南省职业教育改革实施方案》《职业教育专业教学标准(2025 版)》文件要求，能够满足三年全日制高职大数据技术专业培养需要，同意从 2025 级大数据技术专业学生开始实施。

专业建设指导委员会主任签名：李会凯

2025年 8月 19日

审定意见：

中共漯河职业技术学院委员会（签章）

