

北京市昌平职业学校
大数据技术专业 2023 级人才培养方案（5+2）

专业类别（高职）：5102 计算机类

专业名称（高职）：大数据技术

专业代码（高职）：510205

修订负责人：_____ 陕娟娟 _____

主要成员：_____ 王钰坤 朱凯丽 陆少雄 _____

系主任（签字）：_____ 方荣卫 _____

主管部门主任（签字）：_____ 周林娥 _____

教学主管领导（签字）：_____ 郑艳秋 _____

校长、党委书记（签字）：_____ 段福生 _____

2023 年 6 月

人才培养方案修订说明

为贯彻落实《国家职业教育改革实施方案（国发〔2019〕4号）》《关于推动现代职业教育高质量发展的意见》《教育部关于职业院校专业人才培养方案制订与实施工作的指导意见（教职成〔2019〕13号）》《北京市职业院校教学管理通则》《北京市昌平职业学校专业人才培养方案制（修）订工作方案》等系列文件精神与要求，保障专业建设的科学性与规范性。为使人才培养目标定位更精准，课程结构更科学，人才培养跟上产业发展的变化，毕业生更适合就业市场对人才的要求，服务区域经济发展，以及满足对口学校对升学学生的学业要求，本专业通过深入开展人才需求调研、职业分析、课程转换等工作，与高职院校进行研讨，在2022级第三版的基础上，修订2023级大数据技术专业（5+2）人才培养方案第一版。方案具体修订说明如下：

一、职业面向

根据大数据技术专业发展新趋势和行业产业契合度调研、人才培养调研的工作，对职业面向及主要岗位进行调整及更新。

二、培养目标和培养规格

根据国家统一制定的专业教学标准、北京市职业院校教学管理通则等系列文件精神与要求，依托人才培养调研和职业分析会的成果对培养目标和培养规格进行了调整及更新。

三、课程设置及要求

1. 公共基础课：按照教育部最新的公共基础课程标准，调整思想政治、语文、数学、英语、艺术课程的课时及开设学期，按照大纲要求开设了创新创业课程，充足开设了历史和劳动教育课程。

2. 专业课程：根据北京市职业院校教学管理通则等系列文件精神与要求，依据人才培养调研和职业分析会的课程转化成果对专业基础课、专业核心课、综合应用课、专业拓展课进行了调整及更新。

四、实施保障

校内、校外实训基地要求根据最新的校企合作情况进行了更新。

北京市昌平职业学校

大数据技术专业 2023 级人才培养方案（5+2）

一、专业名称（专业代码）

（一）专业名称

高职：大数据技术

本科：大数据工程技术

（二）专业代码

高职：510205

本科：310205

二、入学要求

初中毕业生或具有同等学力者。

三、基本修业年限

学制：五年

学历：高职

四、职业面向

类别	高职	本科
所属专业大类 (代码)	电子与信息大类（51）	电子与信息大类（31）
所属专业类 (代码)	计算机类（5102）	计算机类（3102）
对应行业（代码）	软件和信息技术服务业（65）	软件和信息技术服务业（65）
主要职业类别 (代码)	1. 信息系统分析工程技术人员（2-02-10-05） 2. 信息系统运行维护工程技术人员（2-02-10-08） 3. 大数据工程技术人员（2-01-10-11） 4. 计算机软件工程技术人员（2-02-10-03）	1. 信息系统分析工程技术人员（2-02-10-05） 2. 信息系统运行维护工程技术人员（2-02-10-08） 3. 大数据工程技术人员（2-01-10-11） 4. 计算机软件工程技术人员（2-02-10-03） 5. 数据库运行管理员（40405-04） 6. 数字化解决方案设计师（40404-05）
主要岗位（群）或 技术领域举例	1. 数据采集与处理工程师 2. 大数据 ETL 工程师 3. 大数据平台实施工程师	1. 数仓工程师 2. 大数据分析工程师 3. 大数据开发工程师

	4. 大数据测试工程师 5. 大数据运维工程 6. 大数据开发工程师 7. 大数据分析工程师 8. 数据挖掘工程师	4. 数据挖掘工程师 5. 数据治理工程师
职业类证书举例	1. 数据采集职业技能等级证书(中级)(必考) 2. 大数据分析与应用职业技能等级证书(初级)(选考) 3. 数据应用开发与服务职业技能等级证书(初级)(必考) 4. 华为 HCIA-BigData 大数据认证证书(选考)	1. 大数据分析与应用职业技能等级证书(中级)(选考) 2. 数据应用开发与服务职业技能等级证书(必考) 3. 大数据工程化处理与应用职业技能等级证书(选考) 4. 华为 HCIA-BigData 大数据认证证书(选考)

五、培养目标和培养规格

(一) 培养目标

本专业坚持立德树人的根本任务，培养能够践行社会主义核心价值观，德、智、体、美、劳全面发展，具有良好的科学与人文素养、职业道德和精益求精的工匠精神，具备扎实的文化基础知识、较强的就业创业能力和学习能力，掌握本专业知识和技术技能，面向软件和信息技术服务行业的信息系统分析及运行维护工程技术等职业，能够从事数据采集、数据清洗与预处理、大数据分析、大数据可视化、大数据ETL、大数据组件安装部署、大数据集群运维监控、大数据应用开发、数据仓库建设、数据挖掘等工作的高素质技术技能人才。

(二) 培养规格

学生在系统学习本专业知识和完成有关实习实训基础上，全面提升素质、知识、能力，掌握并实际运用数据采集与处理工程师、大数据平台实施及运维工程师、大数据分析工程师、大数据应用开发工程师、大数据 ETL 工程师、数仓工程师、数据挖掘工程师等岗位群需要的专业核心技术技能，总体上须达到以下要求。

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，坚定维护国家荣誉与形象；

2. 掌握支撑本专业学习和可持续发展必备的政治、语文、历史、数学、英语、信息技术、体育与健康、物理、化学、生物和地理等文化基础知识，初步形成系统化的文化知识体系，具有较为广博、深入的科学与人文素养储备，具备良好的自我管理能力，拥有较为清晰的职业与人生目标以及较好的职业生涯规划能力；

3. 理解信息技术及大数据前沿领域产业文化，能够熟练掌握与本专业从事职业活动相关

的国家法律、行业规定，掌握项目管理与执行、质量管理、数据思维以及安全意识等相关知识与技能，遵守职业道德准则和行为规范，具备社会责任感、社会参与意识和担当精神；

4. 较全面深入的了解新一代信息技术产业相关知识，能够建立起新一代信息技术产业与技术发展的全景，理解各专业领域之间的底层关系与交叉融合知识，深入理解大数据在产业中的定位与意义；

5. 具备良好的主动学习意识与能力，始终保持对新一代信息技术以及大数据领域最新进展与技术趋势的兴趣并主动学习；正确认识专业理论在职业工作中的价值，在培养实践能力同时，有意识拓展理论知识体系，强化深层次职业能力与问题解决水平，养成终身学习与可持续发展的能力；

6. 具备较强的信息安全意识，较全面掌握信息网络与数据安全相关政策、法律法规，清晰知晓数据采集与处理中的法律与政策规范，熟练掌握在各工作环节的安全规范与要求，在数据收集与预处理、大数据交付与运维、大数据分析、大数据应用开发与测试、数仓建设以及数据挖掘等工作中自觉遵守安全规定与法律要求，能够从技术、客户需求等各角度预判可能的安全问题或隐患，并采取主动行动或技术措施予以规避或解决；

7. 具备良好的思维逻辑性与严谨性，在推动工作与解决问题时，能够有意识的主动全面的收集相关信息，充分的利用各类数据，并运用逻辑思维进行分析与预判，形成严谨有效的工作计划规划以及问题解决思路；

8. 养成良好的数据思维，能够较为深入的理解典型行业场景的业务逻辑，能够基于场景的业务逻辑设计数据指标体系，并且能够透过数据结果形成对业务场景以及相关业务问题的较为清晰的洞察理解；

9. 具备较成熟全面的沟通表达能力，包括言语表达、文字表达等，能够清晰流畅的通过言语或文字表达思想与观点，发挥专业影响力；熟练掌握一门外语并结合专业加以深入运用；掌握大数据以及各细分领域相关专业工作中所必须的术语及表达规范、文档撰写规范以及相应的专业与通用工具，能够在工作中准确理解他人的沟通意图，并可以有效的发挥自身的专业影响力，在工作中做到与各方面人员顺畅沟通交流与协作，推动工作进展；

10. 深入认知团队协作对于工作质量以及自身职业成长的重要性，掌握与他人协作的基本原则与方法，通过学习掌握本专业相关岗位及其关联岗位的整体工作流程，建立整体观念与流程协作基本意识与能力；通过学习掌握项目管理等具体工具方法，强化系统协作与协作管理能力；

11. 具备较好的吃苦耐劳的品质与抗压能力，对于较为单一甚至枯燥的工作能够有正确认知与态度，能够适应兼具体力与脑力双重要求的较高压力弹性工作制，能够在较大工作量的

情况下保持积极的心态与稳定的工作成效；当工作需要反复调整优化，或者在故障问题反复出现等情况下，能够正确的对待各方面意见并积极工作，输出高质量工作结果；

12. 掌握计算机学科、大数据技术应用学科的基本理论，包括通用网络技术、信息技术设备组装与维护、Excel 数据分析技术、程序设计、数据库技术、前端设计与开发、Linux 系统应用基础等方面的专业基础理论知识、基本概念与应用场景，可以在大数据技术应用领域进一步深入学习；

13. 了解市面上常用的数据库，了解数据源提供的各类数据类型以及相应的数据协议格式，能够分析数据源的数据结构，明确数据采集需求，规划数据埋点，完成数据采集顶层设计并搭建采集环境；

14. 熟悉 HTML、图片识别 OCR，能够进行视频流解码与取图；掌握一定的反爬技术，能够根据不同的站点采取不同的策略，完成数据采集开发；能够设计数据入库字段、入库接口调用及开发、不同数据库备份的分布式设计等，完成数据存储；

15. 熟悉 Linux 命令、Linux 发行版、组件发行版，熟悉各组件的配置方法、通信机制、能够下载安装各组件并更改各组件配置，确保其协调工作；能够完成各组件关联关系正常调试，确保正常运行；

16. 掌握 Sqoop 工具，能够使用 Sqoop 对接数据库，应用 Sqoop 的 SQL 方式、增量方式实现数抽数据库，并且能够蹈入数据到数据库，完成基础数据的维护，并且能够应用 Furm、Keylan 提取或同步数据，维护数据的完整性；

17. 掌握 Python/Hive/SQL 等工具，能够完成脏数据填充或删除；掌握 Pandas、Scikit-Learn、Numpy 等工具包，在数据量纲不统一的情况下将数据按比例缩放，使之标准化；在数据单位不统一的情况下，利用工具包统一数据衡量标准，使之归一化，完成数据清晰与预处理工作；

18. 掌握 SQL/Python/Shell 脚本编写方法，以及 Kettle 工具、日志采集工具等，能够从源数据、关系型数据库、业务平台日志等完成数据抽取；熟悉 Hive、Spark、Sqoop 等组件，能够将数据根据业务需求转化为目的数据格式并装载至目的数据源，并确保数据质量；

19. 掌握对接业务需求的方法流程，能够使用 Excel、Python 等工具完成数据的分组、聚合以及数学运算，完成描述性分析、趋势性分析、相关性分析等不同分析工作，得到分析发现和结论；并且能够使用 Excel、Tableau 等工具选择合适的可视化方法输出结构化分析报告；

20. 熟悉 Linux 基础环境与网络配置相关知识，能够对于项目的各种载体机器进行实时监控，对网端和接口进行轮循监控，确保可靠性；能够理解业务的场景与逻辑，建立良好的告警即时反馈实现告警监控；能够根据业务日志做好监控和统计，分析业务系统的日常运行状

态，分析调整项目开发质量；

21. 熟悉 Linux 指令、各组件常见报错日志类型，了解组件参数的意义，能够根据平台报错信息分类报错类型；熟悉 Yarn Web UI 各功能，能够查看任务运行情况、调整任务运行参数、调整 Yarn 队列参数等，提高集群运行效率；

22. 了解大数据平台数据的开发流程、大数据平台数据处理逻辑和处理方法，掌握大数据于开发语言与框架、工具，实现功能模块的开发，并能够使用大数据相关组件和框架完成功能的上线部署；熟悉各种功能测试方法，能够根据功能模块实现各类测试，确保功能模块的正常和准确；

23. 掌握基于业务需求进行数据仓库选型的方法，形成数据仓库系统设计方案，能够根据字典进行 ODS 建模，制定 ETL 策略，从源数据系统抽取形成 ODS 层；能够对 ODS 层进行清洗与整合，形成 DWD 层；掌握常用的数据分析手段，进行数据转化形成 ADS 应用数据层；能够完成数据仓库系统平台的日常管理与维护；

24. 掌握 Spark、Hive 等工具，能够根据业务需求进行特征计算，并能够根据需要利用一定的方法对特征进行降维，保证特征模型最优；能够利用工具包进行数据集和样本划分，并基于业务场景与数据表现形式选取合适的算法模型进行训练测试，选取合适的指标对模型进行优化，保证数据挖掘价值；

25. 掌握较为全面的身体运动知识和至少 1 项体育运动技能，养成良好的运动习惯、卫生习惯和行为习惯；具备较好的稳定的心理调适能力，积极健康的面对未来的工作和生活；

26. 掌握必备的美育知识，具有较为广博的文化修养、具有自我风格的审美能力，形成至少 1 项艺术特长或爱好；

27. 践行并弘扬劳动光荣、技能宝贵、创造伟大的时代精神，热爱劳动人民、珍惜劳动成果、树立劳动观念、积极投身劳动，具备适应并深度参与本专业相关职业领域发展的劳动素养、劳动技能。

六、培养模式

（一）专业群说明

信息技术专业群包括计算机网络技术、大数据技术和人工智能技术及应用三个专业，如图 6-1 所示。三个专业技术领域相近，包含计算机网络搭建技术、程序设计技术、软硬件资源及系统部署技术，软件运维服务技术等。三个专业面向岗位群相通，围绕“建-管-用”三个环节，均可以从事智能终端设备服务、系统部署运维、数据处理分析等岗位。专业群以云计算为基础，大数据技术支撑，服务人工智能应用，紧密对接首都新一代信息技术产业链，面

向软件与信息技术服务产业链中下游，信息系统集成服务、信息技术咨询服务、运行维护服务、互联网数据服务和应用软件开发等，全力打造服务于昌平区乃至北京市智慧教育、智慧农业、智慧城市建设的新一代信息技术技能人才培养和创新服务平台。

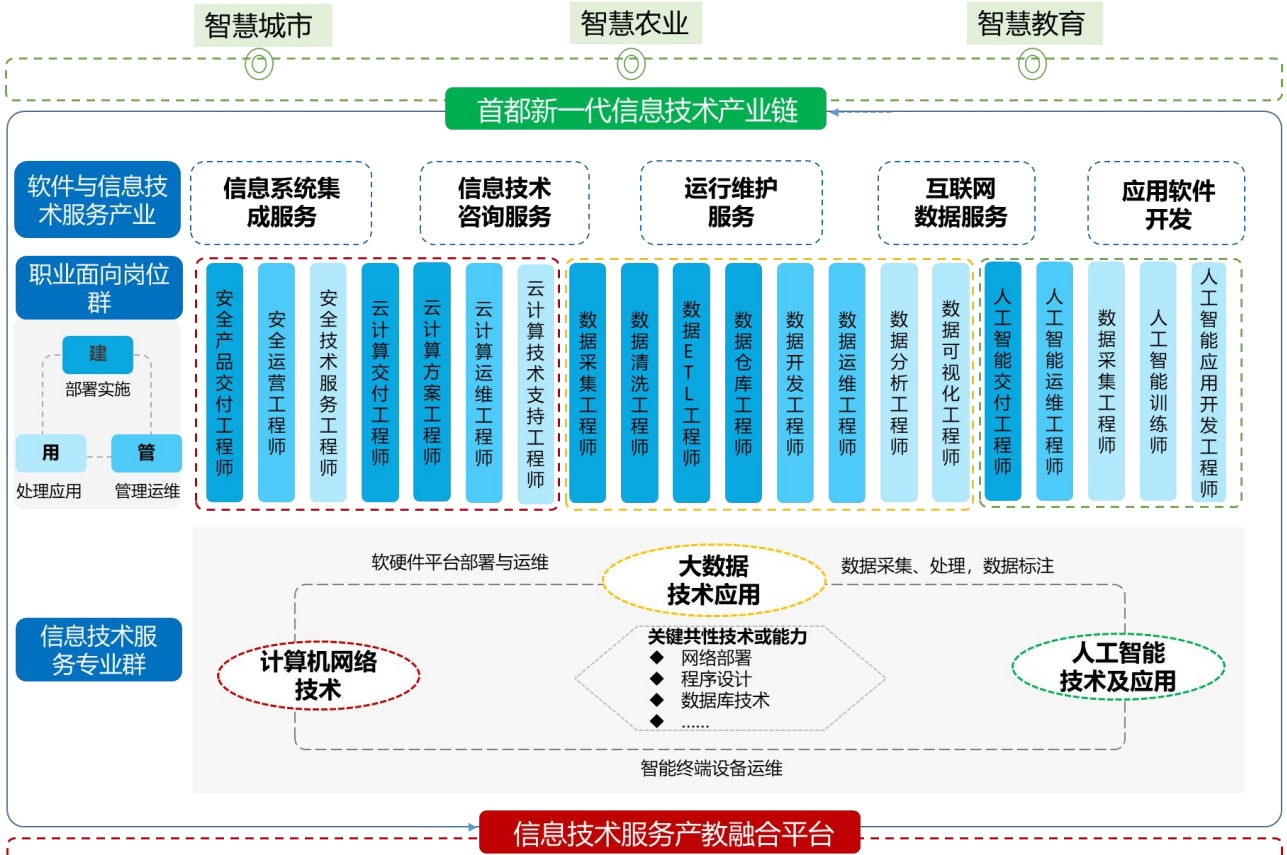


图 6-1 信息技术服务专业群组群逻辑图

（二）人才培养模式

信息技术服务专业群依托北京市特色高水平实训基地联想工程师学院建设项目，深入开展校企合作，构建形成专业群“校企双元，项目引领，四阶四真”人才培养模式，共同为首都新一代信息技术产业培养高素质技术技能人才。校企双元，联想工程师入校与专任教师组建“混编”师资团队，形成专业群建设工作组，保障教学实施；项目引领，将人才链建在智慧城市和智慧社区建设的服务链上，通过“职业启蒙-职业唤醒-职业养成-职业助行”四个阶段，在“真场景、真任务、真过程”条件下完成学习任务和生产任务，产生“真成果”；工学一体，将新一代信息技术产业行业岗位标准、职业技能等级标准证书标准，职业院校技能大赛标准与课程标准融为一体，形成职业素养养成与综合职业能力教育并重的工学一体化课程体系，融入人才培养的全过程，提升育人质量。

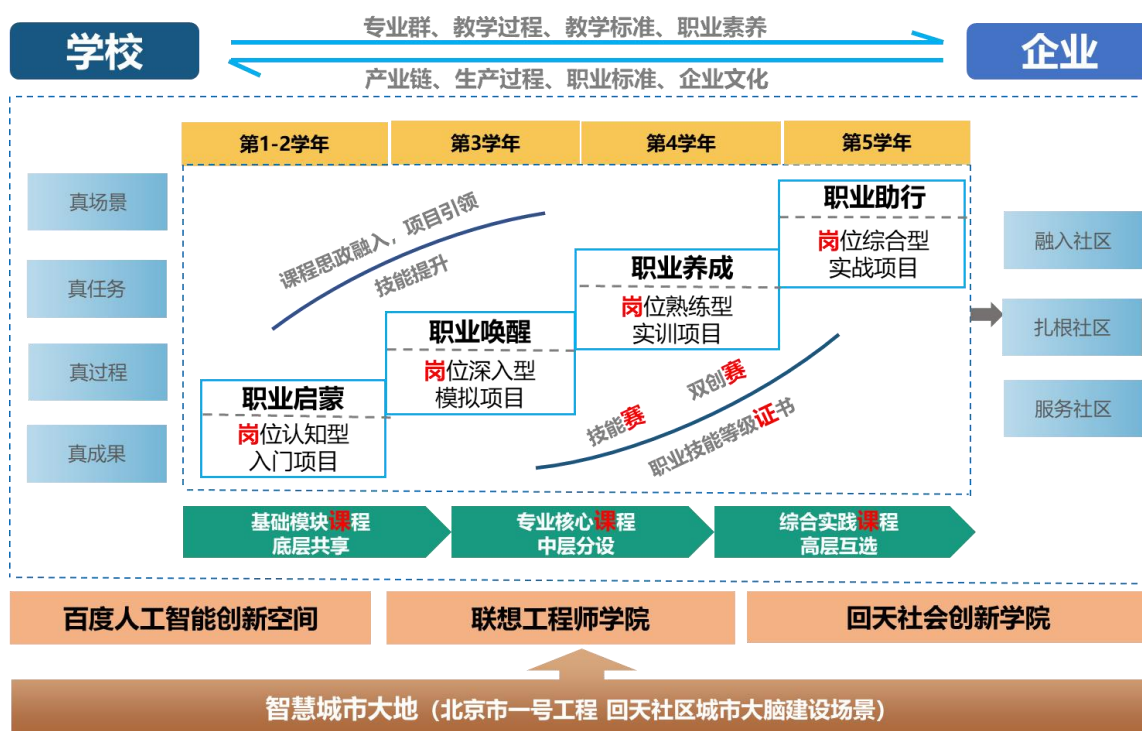


图 6-2 信息技术服务专业群四阶四真四融培养模式

1. 模式内涵释义

大数据技术专业基于联想工程师学院产教融合实训基地构建四真的育人环境（真实工作场景、真实工作任务、真实工作过程、真实工作成果），搭建了由“职业启蒙、职业唤醒、职业养成、职业助行”四个阶段组成的现代学徒制人才培养路径，以技能+素养双线融合培养贯穿学生的培养全过程，实现岗、课、赛、证融通。培育能适应新一代信息技术产业岗位技能迅速迭代，“厚基础，懂技术，能运管，会挖掘”的复合型技术技能人才。

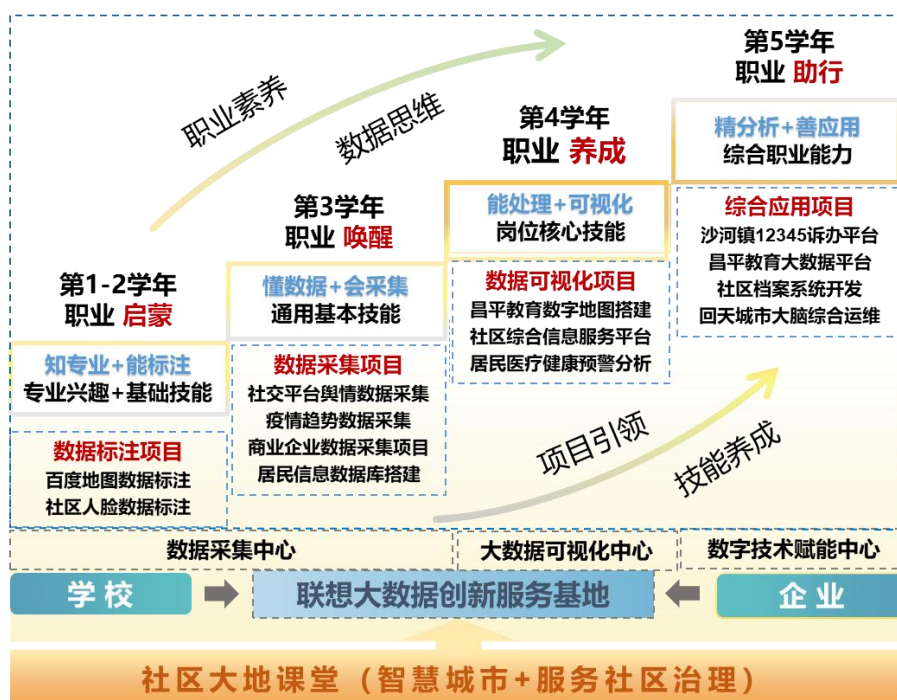


图 6-3 大数据技术专业四阶四真四融培养模式

2. 模式具体运行

职业启蒙阶段：为学生入学后的 1-2 学年，此阶段注重培养学生科学文化素养，企业通过一线专家讲座、企业参观考察等形式帮助学徒了解企业文化、增强岗位认知，通过基础的数据标注项目，激发学生的专业兴趣与专业认知，培养“知专业，能标注”的职业人。

职业唤醒阶段：为学生入学后的 3 学年，加强企业文化的渗透，通过职业唤醒阶段转换学生阶段与社会职场的思维模式，注重培养学生的通用基本技能。通过数据采集项目，让学生加深对数据的理解，培养“懂数据，会采集”的职业人。

职业养成阶段：为学生在校内的 4 学年，学校专任教师和企业师资共同组成分工协作的模块化课程教学团队开展专业课程授课，专业技能与职业技能并重，促进学生综合职业能力的养成，成长为“准职业人”。在技能养成上，企业委派多名一线工程师深入课堂进行授课，引入企业真实项目，并集中开展以周为单位的阶段性数据可视化生产实训。

职业助行阶段：为第 6 学年，每周 3 天的校内集中学习由企业培训师进行授课，主要完成综合项目实践训练；每周 2 天的企业实践由企业负责，根据企业实际生产需要，指导学生参与企业实际生产工作，在企业专家的指导下进一步强化所学知识，熟练岗位核心技能。通过助行强化综合职业能力形成，帮助学生更快成为“合格职业人”，在求职中获得更优质的职位与工作。

七、课程设置与要求

（一）课程体系

1. 专业群课程体系

专业	A 计算机网络技术				B 大数据技术				C 人工智能技术应用						
Z 综合应用课程	Z1-智慧 校园信息 系统运维 服务		A2 C3 A4 B7	Z2-区域 车位引导 系统部署 与应用		A2 C5 A8 C6 B5 C7	Z3-镇街 接诉即办 平台分析 运维		A8 B1 B5 A9 B2 C6 A11 B3 C7	Z4-智慧 楼门长 语音助 手管理		B1 C6 C4 C7 C5	Z5-回天 城市大 脑综合 运维		A9 C3 A11 C8 B9 C10
专业方向课程	A1-Windows 服务器部署 (Z) A2-Linux服务器部署 (Z) A3-中小型局域网搭建 (Z) A4-企业网搭建与应用 (Z) A5-企业网故障诊断与排除 (Z) A6-自动化运维脚本编制 (Z) A7-私有云平台部署交付 (Z) A8-云平台应用部署交付 A9-云计算巡检与运维监控 A10-容器云架构部署 A11-云应用平台规划实施				B1-互联网数据采集 (Z) B2-数据清洗与预处理(Z) B3-大数据ETL B4-数据可视化(Z) B5-大数据分析(Z) B6-大数据组件安装部署(Z) B7-大数据集群运维监控(G) B8-大数据集群运行排查调优(G) B9-大数据应用软件开发与测试 B10-数仓建设 B11-数据挖掘				C1-数据标注 (Z) C2-数据采集 (Z) C3-图像数据预处理 (Z) C4-智能机器人交付与运维 (Z) C5-产品设计 C6-人工智能应用部署 (Z) C7-人工智能应用后端开发 C8-人工智能算法基础 C9-算法模型训练 (Z) C10-算法模型部署						
	P 群平台课程				P1-程序设计基础 P4- 前端设计与开发				P2-数据库技术应用 P5-Linux系统应用基础				P3-通用网络技术 P6-信息技术设备组装与维护		
G 公共基础课程	G1-思想政治 G2-语文 G3-历史 G4-数学 G5-英语 G6-信息技术 G7-体育与健康 G8-艺术 G9-劳动教育 G10-职业素养 G11-中华优秀传统文化 G12-创新创业 G13-高等数学 G14-大学英语 A15-马克思主义哲学 G16-毛泽东思想和中国特色社会主义概论 G17-大学生心理健康教育 G18-形势与政策														

图 7-1 信息技术服务专业群课程体系图

信息技术专业群遵循“成果导向、能力递进、个性发展”的原则，聚焦学生信息素养和工程思维提升，以“岗课赛证融通”为核心，构建“平台为基、核心分设、拓展纵深”的课程体系。以岗位典型工作任务为课程设置出发点，融入职业技能等级证书标准和赛项标准，课程划分为G-公共基础课、P-群平台课、A、B、C-专业方向核心课、Z-综合应用课四大类课程，以课程搭建“岗、课、赛、证”的桥梁，实现学历与技能并重，德技并修。如图7-1所示。

2. 专业课程体系

岗课赛证		职业方向			赛项			证书	
		数据分析	数据平台部署与运维	大数据开发	大数据技术与应用赛项	机器学习赛项	数据分析与可视化赛项	数据采集证书	大数据应用开发与服务
模块化课程组合		A+B+C1+C2+E1+E3	A+B+C3+E2+E4	A+B+C4+E4+E5	C1+C3	C4	C2	C1	C4
E	综合应用课程	E1-区域车位引导系统部署与应用		E2-智慧校园信息系统运维服务	E3-镇街接诉即办平台分析运维	E4-智慧社区轻应用开发运维	E5-回天城市大脑综合运维		
D	专业选修课程	D1-信息安全 D5-数据标注 D9-数据应用开发与服务证书技能训练		D2-常用办公软件 D6-物联网技术	D3-计算机专业英语 D7-数据治理 D10-初级信息系统项目管理	D4-人工智能基础 D8-数据采集证书技能训练			
C	专业核心课程	C1 数据采集模块 1-互联网数据采集(ZZ) 2-数据清洗与预处理(ZZ) 3-大数据ETL(GZ)		C2 数据分析模块 1-大数据分析(ZZ) 2-数据可视化(ZZ)	C3 数据运维模块 1-大数据组件安装部署(ZZ) 2-大数据集群运维监控(ZZ) 3-大数据集群运行排查调优(ZZ)	C4 数据平台开发模块 1-大数据应用软件开发与测试(GZ) 2-数据仓库建设(GZ) 3-数据挖掘(GZ)			
B	专业基础课程	B1-大数据基础 B4-Linux系统应用基础 B8-Excel数据分析技术		B2-通用网络技术 B5-程序设计基础与应用	B3-信息技术设备组装与维护 B6-前端设计与开发 B7-数据库技术应用				
A	公共基础课程	必修 A1-思想政治 A2-语文 A3-历史 A4-数学 A5-英语 A6-信息技术 A7-体育与健康 A8-艺术 A9-劳动教育 A10-地理 A11-生物 A12-物理 A13-化学 A14-高等数学 A15-马克思主义哲学 A16-大学英语 A17-毛泽东思想和中国特色社会主义理论概论					选修	AX1-职业素养 AX2-创新创业 AX3-中华优秀传统文化 AX4-大学生心理健康教育 AX5-文献检索 AX6-科学思维训练 AX7-知识产权保护 AX8-形势与政策	

图 7-2 大数据技术专业课程体系图

本专业课程基于企业岗位典型工作任务进行深化改革。以项目化教学形式，按照人才培养目标与岗位需求对接、课程教学内容与岗位工作任务内容对接的原则，本专业根据人才需求调研报告围绕数据采集与清洗、数据分析、数据运维、数据开发等4个模块11个典型职业活动按照PGSD能力分析模型开展职业分析，分析目标岗位的典型工作任务、工作内容和PGSD能力要求并进行了课程转化和能力对应，确定了互联网数据采集、数据清洗与预处理、大数据分析、数据可视化、大数据组件安装部署、大数据集群运维监控、大数据集群运行排查调优、大数据应用软件开发与测试、数据仓库建设、数据挖掘、大数据ETL共11个专业核心课程。

以职业院校技能大赛、行业大赛为标准，选定了与岗位和专业核心课程对应密切的数据分析与可视化赛项、大数据技术与应用赛项、机器学习等，并将赛项标准分别与模块化课程相融合。同时将数据采集、大数据应用开发与服务等级证书对应的标准与专业核心课程模块融合，建立岗课赛证融合的专业课程体系。

（二）课程设置

主要包括公共基础课程和专业课程。

1. 公共基础课程

包括必修课和选修课。注重学生知识储备、文化素质、科学素养、综合职业能力和可持续发展能力培养，为学生实现更高质量就业和职业生涯更好发展奠定基础。思政教育、素质教育、创新创业教育融入人才培养全过程。

(1) 必修课

序号	课程名称	课程代码	主要教学内容和要求	学时/ 学分	备注
1	思想政治	000000Z0 1	本课程包含六个模块，分别是中国特色社会主义、经济与社会、政治与法治、哲学与文华、法律与生活、逻辑与思维，要求学生掌握马克思主义中国化成果特别是习近平新时代中国特色社会主义思想，理解中国特色社会主义进入新时代的历史方位，了解新时代中国特色社会主义经济、政治、文化、社会、生态文明建设和党的建设进程。引导学生通过自主思考、合作探究和社会实践，深刻领悟辩证唯物主义和历史唯物主义的基本观点和方法，培育政治认同、科学精神、法治意识和公共参与等核心素养，逐步树立共产主义远大理想和中国特色社会主义共同理想，坚定中国特色社会主义道路自信、理论自信、制度自信、文化自信，初步形成正确的世界观、人生观、价值观。该课程能够与中职其他科目的教学和德育工作相互配合，共同完成思想政治教育立德树人的任务。	216/1 2	
2	语文	000000Z0 2	本课程的任务是进一步提高学生正确理解和运用祖国语言文字的水平，使他们具有适应实际需要的现代文阅读能力、写作能力和口语交际能力，具有初步的文学鉴赏能力和阅读浅易文言文的能力；掌握语文学习的基本方法，养成自学语文的习惯，培养发现、探究、解决问题的能力，为继续学习和终身发展打好基础。进一步培养学生热爱祖国语言文字、热爱中华民族优秀文化的感情，培养社会主义思想道德和爱国主义精神，培养高尚的审美情趣和一定的审美能力，发展健康个性，形成健全人格。	432/2 4	
3	历史	000000Z0 3	本课程的任务是在义务教育历史课程的基础上，以唯物史观为指导，促进中等职业学校学生进一步了解人类社会形态从低级到高级发展的基本脉络、基本规律和优秀文化成果；从红色经典影视剧历史题材的角度了解和思考人与人、人与社会、人与自然的关系，增强历史使命感和社会责任感；进一步弘扬以爱国主义为核心的民族精神和以改革创新为核心	72/4	

			的时代精神，培育和践行社会主义核心价值观；树立正确的历史观、民族观、国家观和文化观；塑造健全的人格，养成职业精神，培养德智体美劳全面发展的社会主义建设者和接班人。		
4	数学	000000Z0 4	本课程的任务是进一步提高学生所必要的数学素养，以满足个人发展与社会进步的需要。具体包括：获得必要的数学基础知识和基本技能，理解基本的数学概念、数学结论的本质，了解概念、结论等产生的背景、应用，体会其中所蕴涵的数学思想和方法，以及它们在后续学习中的作用；提高空间想像、抽象概括、推理论证、运算求解、数据处理等基本能力；提高数学地提出、分析和解决问题（包括简单的实际问题）的能力，数学表达和交流的能力，发展独立获取数学知识的能力；发展数学应用意识和创新意识，力求对现实世界中蕴涵的一些数学模式进行思考和作出判断；具有一定的数学视野，逐步认识数学的科学价值、应用价值和文化价值，形成批判性的思维习惯，崇尚数学的理性精神，体会数学的美学意义。	576/3 2	
5	英语	000000Z0 5	本课程的任务是使学生在义务教育阶段英语学习的基础上，进一步明确英语学习的目的，培养自主学习和合作学习的能力；形成有效的英语学习策略；发展综合语言运用能力。能就熟悉的话题交流信息，提出问题并陈述自己的意见和建议。能读懂供高中学生阅读的英文原著简写本及英语报刊。具有初步的实用写作能力。能就熟悉的话题与讲英语的人士进行比较自然的交流。能就口头或书面语言材料的内容发表评价性见解。了解交际中的文化内涵和背景，对异国文化能采取尊重和包容的态度。能听懂有关熟悉话题的演讲、讨论、辩论和报告的主要内容。能就国内外普遍关心的问题表明自己的态度和观点。能自主开拓学习渠道，丰富学习资源。能关注时事，具有较强的世界意识。	576/3 2	
6	信息技术	000000Z0 6	本课程任务是帮助学生认识信息技术对当今人类生产、生活的重要作用，理解信息技术、信息社会等概念和信息社会特征与规范，掌握信息技术设备与系统操作、网络应用、图文编辑、数据处理、程序设计、数字媒体技术应用、信息安全和人工智能等相关知识与技能，综合应用信息技术解决生产、生活和学习情境中各种问题；在数字化学习与创新过程中培养独立思考和主动探究能力，不断强化认知、合作、	108/6	

			创新能力，为职业能力的提升奠定基础。		
7	体育与健康	000000Z0 7	本课程的任务是全面锻炼学生的身体，增强体质。掌握体育基本知识、技术、技能，培养终身体育锻炼的能力和习惯。提高学生的体育运动能力，培养运动爱好的专长，使学生养成终身体育锻炼的习惯，形成健康的行为与生活方式，健全人格，强健体魄，具备身心健康和职业生涯发展必备的体育与健康学科核心素养，了解体育锻炼的原则与方法，常见运动损伤的预防与处理，具有一定的体育文化欣赏能力。增强体质，促进身心健康，培养吃苦耐劳、勇敢顽强的意志品质。养成终身体育锻炼的能力和习惯，健康体质测试标准合格。	324/1 8	
8	艺术 (美术)	000000Z0 8	本课程的任务是以培养学生的美术审美和实践能力，提升其美术品位为目的的美术活动。学生通过观察、体验、赏析、评判等活动，学习美术知识和技能，欣赏美术作品，了解作品主题，感悟作品情感，理解作品内涵，认识美术的基本功能与作用，提高审美情趣和美术实践能力。	18/1	
9	艺术 (音乐)	000000Z0 8	本课程的任务是以培养学生的音乐审美和实践能力，提升其音乐品位为目的的音乐活动。学生通过聆听中外经典音乐作品，参与音乐实践活动，学习有关知识和技能，认识音乐的基本功能与作用，获得精神愉悦，提高审美情趣和音乐实践能力。	18/1	
10	地理	000000Z0 25	学生能够具备人地协调观、综合思维、区域认知、地理实践力等地理学科核心素养，学会从地理视角认识和欣赏自然与人文环境，懂得人与自然和谐共生的道理，提高生活品位和精神境界，为培养德智体美劳全面发展的社会主义建设者和接班人奠定基础。	36/2	
11	生物	000000Z0 14	学生既能获得基础的生物学知识，又能领悟生物学家在研究过程中所持有的观点以及解决问题的思路和方法。学生主动地参与学习，在亲历提出问题、获取信息、寻找证据、检验假设、发现规律等过程中习得生物学知识，养成科学思维的习惯，形成积极的科学态度，发展终身学习及创新实践能力。	36/2	
12	物理	000000Z1 0	学生能够从物理学的视角认识自然，理解自然，建构关于自然界的物理图景；学生经历科学探究过程，体会科学研究方法，养成科学思维习惯，增强创新意识和实践能力；引	324/1 8	

			领学生认识科学的本质以及科学·技术·社会·环境（STSE）的关系，形成科学态度、科学世界观和正确的价值观，为做有社会责感的公民奠定基础。		
13	化学	000000Z1 1	学生能够有学习化学的兴趣，提高学生的科学素养，引导学生认识化学对促进社会进步和提高人类生活质量方面的重要影响，理解科学、技术与社会的相互作用，形成科学的价值观和实事求是的科学态度；培养学生的合作精神，激发学生的创新潜能，提高学生的实践能力。	36/2	
14	劳动教育	000000Z0 9	本课程的任务是结合专业特点，增强职业荣誉感和责任感，提高职业劳动技能水平，培育积极向上的劳动精神和认真负责的劳动态度的基础上，持续开展日常生活劳动，能够促进学生自我管理生活，提高劳动自立自强的意识和能力；运用专业技能为社会、为他人提供相关公益服务，培育社会公德，厚植爱国爱民的情怀；依托实习实训，参与真实的生产劳动和服务性劳动，增强职业认同感和劳动自豪感，提升创意物化能力，培育不断探索、精益求精、追求卓越的工匠精神和爱岗敬业的劳动态度，坚信“三百六十行，行行出状元”，体认劳动不分贵贱，任何职业都很光荣，都能出彩。	18/1	
15	高等数学	000000G0 7	学生能够具备应用数学知识解决实际问题的能力，通过了解微积分的背景思想，较系统地掌握高等数学的基础知识、必需的基本理论和常用的运算技能，了解基本的数学建模方法，为学习后继课程、专业课程和分析实际问题奠定基础。	144/8	
16	大学英语	000000G0 8	学生能够掌握一定的英语基础知识和基本技能，具有一定的英语语言综合应用能力，即一定的听、说、读、写、译的能力，从而能借助词典阅读和翻译有关英语教学案例方面的资料，在日常活动和业务活动中进行简单的口头和书面交流，并为进一步提高英语水平打下较好的基础。通过学习，学生应能够具备通过高等学校应用能力考试 A 级或 B 级的水平。	144/8	
17	马克思主义哲学	000000G0 1	学生能够掌握马克思主义哲学的基本观点、立场和方法；掌握马克思主义的世界观和方法论，树立正确的人生观和价值观；具备运用马克思主义哲学的观点和方法去分析问题、解决问题的能力，提高政治理论素质和思维水平；为正确理解马克思主义，确立社会主义信念自觉坚持党的路线、方针和政策打下坚实的基础。	36/2	

18	毛泽东思想和中国特色社会主义概论	000000G02	学生能够系统地接受思想政治教育，掌握以中国化的马克思主义为主题，以马克思主义中国化为主线，以建设中国特色社会主义理论与实践为重点，着眼于马克思主义理论的应用，着眼于对实际问题的思考，着眼于新的实践和新的发展的马克思主义思想政治教育路线。	36/2	
----	------------------	-----------	--	------	--

(2) 选修课

序号	课程名称	课程代码	主要教学内容和要求	学时/学分	备注
1	职业素养	000000Z13	学生能够提升职业意识，规范职业行为，养成优秀职业品质；学生能够以一个准职业人的身份要求自己，成为崇尚劳动、敬业守信、创新务实的社会好公民；成为立足岗位、服务群众的大数据工程师；成为德才兼备、创新进取、精益求精的优秀工匠。	72/4	第五学期任选5学分，第六、七、八学期任选2学分，共任选11学分。
2	中华优秀传统文化	000000Z12	学生能够了解中国传统文化思想文明成就，中华民族创造历史，融入社会文化的意涵，学生具备爱国主义情操，开阔文化视野。例如：琴棋书画、传统文学、饮食厨艺、民间工艺等。	18/1	
3	创新创业	000000Z16	学生能够了解和掌握基本的创新、创业方法和规范，培养主动创新的意识，激发创业激情，提升创新能力和创业能力。通过对大量创新创业案例的分析与讨论，帮助学生深刻地认识创新的重要性，树立正确的创业成败观，有利于培养学生善于思考、勇于探索的创新精神和敢于承担风险、挑战自我的进取意识。	36/2	
4	文献检索	000000Z15	本课程的任务是让学生在学习文献组织和信息交流的基本规律过程中，通过上机实习实际检索和利用与专业有关的文献信息，掌握国内外本专业主要手工检索工具、电子数据库使用方法以及互联网信息查找和利用的方法，增强独立学习和研究的能力，为终身学习打下更坚实的基础。	36/2	
5	知识产权保护	000000Z17	本课程的任务是帮助学生了解我国现行知识产权保护的法律法规，明确知识产权保护的国内外现状，理解知识产权保护的目的，掌握正确应用的方法。	36/2	
6	科学思维训练	000000Z18	本课程的任务是通过对学生进行多模块多角度的思维训练，提高学生的思维能力和思维兴趣，培养学生创新意识，使得学生在面对工作、任务、问题时，知道怎样思考，找到好的方法，做好工作、完成任务、解决问题。为将来的职业生涯奠定一个较好的基础。	36/2	
7	大学生心理健康教	000000G06	学生能够综合运用马克思主义的基本观点和方法，从自身面临和关心的实际问题出发，进行思想品德和社会主义法制学习，提高自身修养和素质。	36/2	

	育				
8	形势与政策	000000G03	学生通过学习马克思列宁主义、毛泽东思想、邓小平理论和“三个代表”重要思想，紧密结合国内外形势，系统掌握党的路线、方针和政策教育的思想政治理论。适时地学习进行形势政策、世界政治经济与国际关系基本知识，开阔自身视野，及时了解和正确对待国内外重大时事，能够在改革开放的环境下有坚定的立场、有较强的分析能力和适应能力。	36/2	

2. 专业课程

包括专业基础课程、专业核心课程、综合应用课程和专业拓展课程等。

专业基础课程：设置通用网络技术、信息技术设备组装与维护、大数据基础、Excel 数据分析技术、程序设计基础与应用、数据库技术应用、前端设计与开发、Linux 系统应用基础共 8 门课程，意在为专业学生对大数据技术有一定的了解，为专业核心技能和岗位典型任务的实施奠定基础。

专业核心课程：面向岗位群中各岗位的核心能力要求，对应设置互联网数据采集、大数据组件安装部署、数据清洗与预处理、数据可视化、大数据分析、大数据 ETL、大数据应用软件开发与测试、大数据集群运维监控、大数据集群运行问题排查调优、数据仓库建设和数据挖掘等 11 门课程，依据课程标准，强化岗位核心技能培养，确保学生达到岗位任职要求。

综合应用课程：设置智慧校园信息系统运维服务、镇街接诉即办平台分析运维、回天城市大脑综合运维等综合应用课程，通过综合性项目训练，提高学生分析问题、解决问题的能力，为就业打下良好的基础。

专业选修课程：基于专业群拓展纵深，专业选修课程来源于专业群互选课程。设置信息安全、常用办公软件、计算机专业英语、人工智能基础、数据标注、物联网技术、数据治理、数据采集证书技能训练、数据应用开发与服务证书技能训练、大数据平台运维证书技能训练、初级信息系统项目管理等课程，主要提升学生专业群内的岗位综合能力和岗位迁移能力。

(1) 专业基础课程

专业基础课程是必修课程，包括：通用网络技术、信息技术设备组装与维护、大数据基础、Excel 数据分析技术、程序设计基础、数据库技术应用、前端设计与开发、Linux 系统应用基础、程序设计进阶等。

序号	课程名称	课程代码	主要教学内容和要求	学时/学分	备注
1	通用网络技术	7102Z0B01	本课程的主要任务是在帮助学生了解网络基本构成，包括计算机网络体系结构、计算	36/2	▲ 群平台课

			机网络硬件、网络规划与布线、网络操作系统、组建局域网、网络服务构建、因特网应用、网络安全、网络维护等。培养学生的大数据平台网络规划能力，帮助学生具备大数据平台规划所需要的网络知识与技能，以及适应相应岗位的基本职业能力。		
2	信息技术设备组装与维护	7102Z0B02	本课程的任务是帮助学生通过企业岗位实训，熟悉计算机信息技术设备的组成、各部件的结构及基本工作原理、操作系统基础知识、网络基础知识等，能够熟练完成部分信息技术设备的组装与维护操作。	36/2	▲ 群 平 台 课
3	大数据基础	710205Z0B03	本课程的主要任务是引导学生步入大数据时代，积极投身于大数据的变革浪潮中；了解大数据概念，培养大数据思维；认识大数据伦理，努力规范自己的言行，符合大数据时代数据安全等要求；熟悉大数据应用，探寻大数据与生产行业企业应用结合点；熟练应用大数据的初级工具，完成完整的数据采集、处理、分析和展示环节。	36/2	
4	Excel 数据分析技术	710205Z0B04	本课程的主要任务是使学生具备数据处理分析技术应用专业人员所必需的基础知识，建立起数据处理分析的知识技能体系，具备数据处理分析思维，掌握数据处理分析基本原理，了解Excel 中数据处理分析的相关技术应用。	36/2	
5	程序设计基础与应用	7102ZG0B01	通过学习该课程，掌握 Python、JAVA 程序设计语言的基本知识和使用编程语言进行软件开发的思想和基本方法，掌握模块化、结构化编程的基本思想，掌握数据类型、语句格式和函数结构等主要内容，进而掌握程序设计的基本步骤和通用方法，提高通过编写程序解决实际问题的能力，为今后进一步使用数据采集和分析等大数据及人工智能方面的运用打好基础。	180/10	▲ 群 平 台 课
6	数据库技术应用	7102Z0B05	本课程以帮助学生建立数据库概念、掌握数据库基本原理，掌握数据库的基本操作，提高数据库的应用能力和分析问题能力为教学目标。涉及数据库系统概述，数据库设计，关系数据库标	72/4	▲ 群 平 台 课

			准语言 SQL, MySQL 数据库系统中表、查询、窗体、报表、宏、VBA 编程等对象的基本操作和应用等。		
7	前端设计与开发	7102Z0B06	本课程要求学生了解 web 编程技术的产生和发展过程, 会使用 HTML 编写网页, 会用 CSS 对网页样式进行设计, 会使用 JavaScript 实现一些客户端的数据验证及网页特效, 达到一般企业用人标准。	36/2	▲ 群 平 台 课
8	Linux 系统应用基础	7102Z0B07	本课程的主要任务是提高学生对 Linux 操作系统的认识, 并通过案例教学和项目实训培养学生综合运用知识的初步能力。通过 Linux 操作系统的使用, 包括文本界面的常用 Shell 命令、图形界面的多种实用程序以及 Linux 提供的多种 Internet 服务功能, 让学生比较全面地了解 Linux 操作系统提供的功能和服务。	72/4	▲ 群 平 台 课

(2) 专业核心课程

专业核心课程是必修课程, 包括: 互联网数据采集、大数据组件安装部署、数据清洗与预处理、数据可视化、大数据分析、大数据ETL、大数据应用软件开发与测试、大数据集群运维监控、大数据集群运行问题排查调优、数据仓库建设、数据挖掘。

序号	课程名称	课程代码	主要教学内容和要求	学时/学分	备注
1	互联网数据采集	710205Z1B01	本课程的主要任务是让学生掌握 Web 前端、计算机网络的基础知识, 理解网页爬取、数据抽取、数据解析、数据存储等互联网数据采集流程, 能够使用常用的数据采集工具完成简单的数据采集任务, 具备一定的数采实践能力。	72/4	书证融通
2	大数据组件安装部署	710205Z1B02	本课程主要是让学生能够掌握大数据组件的下载、安装方法, 能够根据需求修改各组件配置文件, 使其协调工作, 并开发程序测试确保各组件正常运行。	36/2	书证融通

3	数据清洗与预处理	710205Z1B03	本课程的主要任务是让学生能够使用数据清洗工具（如 Excel、Kettle 等）以及 Python 编程语言对采集的数据进行数据处理，具备数据清洗和预处理的能力。	72/4	书证融通
4	数据可视化	510205G0B01	本课程的主要任务是让学生能够掌握可视化软件的安装、配置方法，并能熟练使用可视化软件完成不同数据源和数据导入导出操作，并能根据业务需求绘制可视化图表；掌握可视化软件相关函数的使用以及常用操作。	72/4	
5	大数据分析	510205G0B02	本课程的主要任务是让学生能够掌握 Python 数据分析工具库的基础知识，熟练使用数据分析工具库完成数据的导入、处理、可视化、分析、存储等任务，并通过一系列的 Python 语言数据分析训练项目让学生具有一定的数据分析实践能力。	108/6	
6	大数据 ETL	510205G0B03	本课程的主要任务是让学生熟练掌握 ETL 工具（如 Kettle）完成数据抽取、清洗和存储等任务，能够设置定时任务，运行数据转换工程，熟练使用源数据获取、记录处理、字段处理、高级转换、迁移和装载等数据转换操作方法，从而具备一定的 ETL 实践能力。	72/4	
7	大数据应用软件开发与测试	510205G0B04	本课程的主要任务是以 Java 编程语言为开发工具帮助学生掌握大数据平台数据开发流程，能够根据业务需求确定解决方案，开发功能模块，开展各类测试，针对功能模块生产的数据进行验证，以确保功能模块正常运	108/6	书证融通

			行，且结果准确可靠。		
8	大数据集群 运维监控	510205G1B01	本课程主要是帮助学生掌握监控基础网络环境的知识，能够设置监控指标完成监控和统计，能够分析业务系统的日常运行状态，查看告警提示，分析问题及原因，并即时调整或反馈，具备一定的集群运维监控能力。	72/4	书证融通
9	大数据集群 运行问题排 查、调优	510205G1B02	本课程主要是帮助学生掌握大数据集群常见运行问题的排查和解决方法，能够通过 Yarn Web UI 查看任务运行情况，适当调整任务运行参数进行任务调优；能够调整 Yarn 队列参数满足业务需求；能够根据组件选型、任务占用资源分类统计情况调整集群参数，具备一定的问题排查和调优能力。	72/4	书证融通
10	数据仓库建 设	510205G0B05	本课程的主要任务是帮助学生掌握数据库选型、完成数据仓库系统设计的方法，能够根据业务需求设计数据模型，完成数据库构建，熟悉业界主流的数据仓库工具，掌握数据仓库工具的使用方法，具备一定的数据仓库设计能力。	72/4	
11	数据挖掘	510205G0B06	本课程的主要任务是帮助学生掌握数据挖掘的全流程，包括数据清洗、特征工程、模型构建、模型评价、模型预测等，掌握常见的数据清洗方法、特征工程技巧，能够根据业务场景和数据类型选取合适的算法模型，熟悉模型的训练和测试过程，具备一定的使用数据挖掘算法模型解决实际问题的能力。	72/4	书证融通

（3）综合应用课程

综合应用课程是必修课程，包括智慧校园信息系统运维服务、镇街接诉即办平台分析运维、回天城市大脑综合运维等。

序号	课程名称	课程代码	主要教学内容和要求	学时/学分	备注
1	智慧校园信息系统运维服务	710205G0B07	本课程的主要任务是基于智慧校园信息系统运维项目，帮助学生掌握网络搭建、服务器部署、信息技术设备安装与维护等知识，能够对校园信息系统的技术设备进行日常维护，能够应用网络设备管理与维护的基本方法和测试工具进行校园网络排障，具备一定的系统运维服务能力。	108/6	
2	镇街接诉即办平台分析运维	710205G0B08	本课程的主要任务是基于镇街接诉即办平台运维项目，让学生掌握网络搭建、服务器部署、数据采集与处理等相关知识，梳理完成网络集成、服务部署、应用开发等分析运维任务，具备一定的综合实践能力。	144/8	
3	回天城市大脑综合运维	710205G0B09	本课程的主要任务是基于回天城市大脑综合运维项目，让学生了解回天城市大脑项目的基础知识、运行架构，能够根据需求完成系统硬件日常维护、系统集群运行日常监控等运维任务，能够理解集群、系统监控指标；能够分析系统告警提示，及时记录并反馈解决方案，具备一定的项目综合运维能力。	216/12	

（4）专业拓展课程

专业拓展课程，为选修课程，包括：信息安全、常用办公软件、计算机专业英语、人工智能基础、数据标注、物联网技术、数据治理、数据采集证书技能训练、数据应用开发与服

务证书技能训练、大数据平台运维证书技能训练、初级信息系统项目管理等。

序号	课程名称	课程代码	主要教学内容和要求	学时/学分	备注
1	信息安全	710205ZG0X01	通过本课程的学习，培养学生具有信息安全管理岗位所需的信息安全风险评估、信息系统保障、信息安全管理、信息安全工程实施等方面的能力，以及分析问题、解决问题的能力，并养成良好的职业道德，为后续课程学习打下坚实的基础。	36/2	第六学期任选2学分，第七、八、九、十学期各任选4学分，共任选18学分。
2	常用办公软件	710205ZG0X02	本课程的主要任务是培养学生了解计算机在办公自动化领域中的相关知识，掌握办公软件的使用，培养学生使用常用办公软件进行日常办公事务处理、数据处理、信息管理的综合能力，同时培养学生的方法能力、社会能力及职业素质，提高学生的就业竞争力。	36/2	
3	计算机专业英语	710205ZG0X03	本课程的任务和目的是通过对专业英语的学习，掌握大数据相关的英语词汇，具有阅读大数据专业资料的能力，具有一定的专业翻译能力，能够用英语表达专业意义并实现业务沟通。	36/2	
4	人工智能基础	710205ZG0X04	通过本课程对人工智能从整体上有一个较清晰全面的系统了解；使学生掌握人工智能的基本概念、基本原理和基本方法；了解人工智能研究与应用的最新进展和发展方向；开阔学生知识视野、提高解决问题的能力，为将来接触使用人工智能的相关问题奠定初步基础。通过本课程系统了解人工智能科技目前的基本实力，加强学生对我国及国际人工智能的认识，提高学生对自己专业科技学习和判断能力，掌握实用人工智能运用主要领域，培养较强的人工智能接受和判断能力。	36/2	
5	数据标注	710205ZG0X05	通过本课程，学生可以了解数据标注行业的新技术和发展趋势，理解数据标注的基本原理、技术和方法，拓宽数据标注	36/2	

			行业的知识面,掌握数据标注的基本技能,使学生具备数据标注技术应用专业人员所必需的基础知识,建立起数据标注的知识技能体系,具备数据标注思维。		
6	物联网技术	710205ZG0X06	本课程的任务是掌握物联网技术相关的基础知识,能完成物联网典型应用智能家居系统的安装,调试,管理,服务等工作任务。	36/2	
7	数据治理	710205ZG0X07	通过本课程,学生可以了解数据治理体系建立、数据标准制定和管理方法;了解数据资产管理制度与流程的制定方法;了解对数据质量进行监测评价和持续管控等的方法。	36/2	
8	数据采集证书技能训练	710205ZG0X08	本课程的主要任务是通过技能训练帮助学生进一步掌握数据采集的知识,助力学生通过数据采集证书认证考试。	36/2	
9	数据应用开发与服务证书技能训练	710205ZG0X09	本课程的主要任务是通过技能训练帮助学生进一步掌握大数据专业知识,助力学生通过数据应用开发与服务职业技能等级认证考试。	36/2	
10	初级信息系统项目管理	710205ZG0X10	本课程的任务是帮助学生发现、分析、研究、解决初级信息系统项目管理实际问题,具有初级信息系统项目管理知识,具有进行企业项目管理的能力,具有从事企业信息项目管理的初步能力。	36/2	

八、教学进程总体安排

(一) 课程结构比例表

表 8-1 大数据技术专业课程结构比例表

总学时	总学分	公共基础课学时	公共基础课学时占比>=33%	实践课学时	实践课学时占比>=50%	选修课学时	选修课学时占比>=10%
5472	304	3348	61.18%	2847	52.03%	522	10%

(二) 教学活动周进程安排表(单位: 周)

表 8-2 大数据技术专业教学活动周进程安排表

项目 学期	军事 训练	理实一 体教学	社会 实践	认识 实习	劳动 教育	岗位 实习	考证	考试	总周数	备注
一	1	18	0.5					0.5	20	
二		18	0.5		1			0.5	20	
三		18	0.5	1				0.5	20	
四		18	0.5		1			0.5	20	
五		18	0.5		1			0.5	20	
六		18	0.5	1				0.5	20	
七		18	0.5		1			0.5	20	
八		14	0.5			4	1	0.5	20	
九		12	0.5			6	1	0.5	20	
十		9	0.5			10		0.5	20	
合 计	1	160	5	2	4	20	2	5	200	

(三) 实践教学环节安排表(单位：周)

表 8-3 大数据技术专业实践教学环节安排表

序号	名称	总周 数	第一学年		第二学年		第三学年		第四学年		第五学年		备注
			1	2	3	4	5	6	7	8	9	10	
1	专业认识和岗 位认知实践	6		1	1	1	1	1	1				
2	1+X职业技能 等级证书考证 专周实训	2								1	1		
3	生产性综合实 训	20								4	6	10	
4	专业实践	5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	

(四) 职业资格证书考取安排表

表 8-4 大数据技术专业职业资格证书考取安排表

序号	证书名称及等级（必 考/选考）	拟考学 期	对应课程	开设学期	证书类型
1	大数据采集职业技 能等级证书（中级） （必考）	8	1. 互联网数据采集 2. 数据清洗与预处理	7、8	职业技能等级证 书

2	数据应用开发与 服务职业技能等级 证书初级（必考）	9	1. 大数据分析 2. 大数据应用软件开发与 测试 3. 数据应用开发与服务职 业技能等级证书技能训练	8	
3	华 为 HCIA-BigData 大数据认证证书（选 考）	10	华为HCIA大数据证书技能 训练	10	行业证书

（五）专业教育活动设置与教育时间安排进程表

如表 8-5 所示，在专业教学的同时，进一步创新学校“五育”并举综合培养途径，提升中职学校学生培养质量，促进学生德智体美劳全面发展。提出了“五育并举、以一带四”的发展理念，以学校“爱国爱党”、“修德修身”、“铸魂铸匠”三大特色主题教育活动为路径，紧扣育人总目标，凸显出昌职育人的特色和风貌。

表 8-5 大数据技术专业教育活动设置与教育时间安排进程表

教育系列	课程名称	教育主题	课程性质	学期课时安排									
				有模有样			有思有责			有勇有智			
				一	二	三	四	五	六	七	八	九	十
爱国爱党系列	主题班会	勿忘历史，以国为荣	必修	1	1					1	1		
		我骄傲我是中国人	必修			1	1						
		我是学校代言人	必修					1	1				
	思政教育活动	“勿忘历史，以国为荣”主题系列活动	选修	√	√	√				√	√		
		“我骄傲我是中国人”主题系列活动	选修				√	√	√				
		“我的梦中国梦”主题系列活动	选修					√		√			
		爱国爱党系列观影会	必修	√	√	√	√	√	√	√	√	√	√
		爱国主义歌曲合唱、古诗朗诵比赛	必修	2	2	2	2	2	2	2	2	2	
		国防教育综合实践活动	必修	2W						2W			
		“新媒体人”思政大讲堂	选修	3	3	3	3	3		3	3	3	
		中国电影博物馆参观	必修	√						√			
		“我奉献，我快乐”志愿服务实践活动	选修	1	1	1	1	1	1	1	1	1	1
修德修身系列	主题班会	知礼懂礼我做起	必修	1						1			
		我的责任我担当	必修			1				1			
		勇于对不良现象说不	必修					1		1			
		关心他人，把握界限	必修					1		1			
		明辨是非、谁是英雄	必修				1			1			
		你会爱自己吗？	必修	1						1			
		做最美的自己	必修	1						1			
		如果我是你	必修				1						
	五育教育活动	“知礼守法我做起”主题系列活动	必修	√	√					√	√		
		“知恩懂孝我担当”主题系列活动	选修			√	√					√	√
		“勇于对不良现象说不”主题系列	选修					√					

		活动											
		入学教育	必修	10						10			
		艺术节-年度教育教学成果展	必修	2W		2W		2W			2W		2W
		体育节-运动会	必修	2W	2W	2W	2W	2W	2W	2W	2W	2W	2W
		社团活动	选修	28	28	28	28	28	28	28	28	28	28
		付冬梅主题教育	必修	√		√		√					
		IT 吉尼斯	选修		1		1			1		1	
		青春风采 show	选修	1		1		1			1		1
铸匠铸魂系列	主题班会	带着目标上路	必修		1					1			
		从菜鸟到达人	必修				1			1			
		铸就非凡匠心	必修					1			1		
		做自己人生的掌舵者	必修		1					1			
		世界这么大，带你去看看	必修			1					1		
		用心经营生活	必修						1				1
		未来扑面而来	必修					1				1	
	五育教育活动	“带着目标上路”主题系列活动	选修	√	√					√	√		
		“从菜鸟到达人”主题系列活动	选修			√	√			√	√		
		“铸就非凡匠心”主题系列活动	选修					√		√		√	
		科技节-IT 类科技竞赛活动	必修	√		√		√		√		√	
		双创节-学生创意项目孵化及展示活动	必修	√	√	√	√	√	√	√	√	√	√
		技能大比拼	选修	1	1	1	1	1	1	1	1	1	1

说明：

1. 主题班会使用周一“主题班会课”学时，“爱国主义歌曲合唱、史诗朗诵比赛”使用周四“团活动”学时，其他教育活动结合实际安排，使用“主题班会课”学时、“团活动”学时、早读晚自习时间及课余时间。
2. 表中“W”表示“周”、“√”表示在相应学期内自行安排时间，不限定学时量。

（六）教学进程安排表（见附录 1）

九、实施保障

（一）师资队伍

按照“四有好老师”“四个相统一”“四个引路人”的要求建设专业教师队伍，将师德师风作为教师队伍建设的第一标准。

1. 队伍结构

专任教师队伍的数量、学历和职称符合国家有关规定，形成合理的梯队结构。本专业学生数与专任教师数比例不高于 16:1，专任教师中具有高级专业技术职务人数不低于 20%。本专业学生数与思政教师数比例不高于 200:1。双师型教师占专业教师比不低于 90%。兼职教师应占专任教师总数的 20%。

2. 专任教师

具有中职及以上教师资格和软件和信息技术服务领域有关证书；具有教师资格证书；中

职及以上教师资格和软件和信息技术服务领域有关证书；具有计算机应用技术、大数据技术、等相关专业学历；具有本专业理论和实践能力；能够落实课程思政要求，挖掘专业课程中的思政教育元素和资源；能够运用信息技术开展混合式教学等教法改革；具有较强的信息化教学能力，能胜任专业课程的教学，能够开展课程教学改革和科学研究；能够跟踪新经济、新技术发展前沿，开展社会服务；专业教师每年至少 1 个月在企业或实训基地实训，每 5 年累计不少于 6 个月的企业实践经历。

3. 专业带头人

专业带头人原则上应具有副高及以上职称，能够较好地把握国内外行业、专业发展方向和趋势，能广泛联系行业企业，了解新一代信息技术产业行业以及大数据相关企业对本专业人才的需求实际，教学设计、专业研究能力强，组织开展教科研工作能力强，具有企业服务的能力，在本区域或本领域具有一定的专业影响力，在专业改革发展中起到引领作用。

4. 兼职教师

主要从相关行业企业聘任，在行业企业一线从事管理、技术工作两年以上，尤其是合作企业的行业大师可以作为兼职教师，具有相关工作的丰富经历和较强的专业技能，熟悉工作流程。具备良好的思想政治素质、职业道德和工匠精神，具备具有扎实的专业知识和丰富的实际工作经验，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，参与专业建设、课程开发等活动。从本专业相关行业企业的高技术技能人才中聘任，应具有扎实的专业知识和丰富的实际工作经验，原则上应具有中级及以上相关专业技术职称，了解教育教学规律，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专业教学任务。建有专门针对兼职教师聘任与管理的具体实施办法。

（二）教学设施

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、校内实训室和校外实训实习基地等。

1. 专业教室要求

配备黑（白）板、多媒体计算机、投影设备、音响设备，互联网接入或无线网络环境，并具有网络安全防护措施，具备利用信息化手段开展混合式教学的条件。安装应急照明装置并保持良好状态，符合紧急疏散要求、标志明显、保持逃生通道畅通无阻等。

2. 校内实训室要求

校内实训基地建设依托校企合作平台，参考实际大数据或者互联网相关企业实际项目制作流程、布局，实训场所的搭建；还需要考虑项目、场景教学等多种教学方法的应用，教学

设备的技术选择须有一定的前瞻性，能够体现“以校为主、企业为辅，师生参与、项目引入”的实践教学硬件、软件建设和教学方式，能承担专业实践课程、各类综合应用课程等教学任务。

表9-1 大数据技术专业校内实训场地一览表

序号	实训室名称	实训室功能	使用课程	设备	数量 (台/套)	场地面积 (m ²)
1	大数据基础教学实训室	完成专业基础课程教学和实训	大数据基础、程序设计基础、数据库技术应用、前端设计与开发、Linux 系统技术应用、Excel 数据分析	计算机	45 套	120
				路由器	4 套	
				交换机	4 套	
				小机柜	1	
2	大数据创新实践中心	大数据综合实践项目应用、互联网+创新创业大赛	综合实践课程和项目课程、创新创业、职业素养	可视化大屏	6 套	120
				笔记本电脑	45 套	
3	大数据分析 与可视化实训室	数据分析社会培训，数据分析实战训练	数据分析、数据可视化、数据标注	可视化大屏	20 套	120
				计算机	40 台	
4	数据清洗与处理实训室	数据整理、数据清洗实战训练	数据清洗与预处理、数据 ETL	计算机	40 台	120
5	大数据开发测试实训室	数据采集、大数据系统开发与测试	离线数据采集、实时数据采集、大数据系统开发与测试	计算机	40 台	120
6	大数据平台运维实训室	大数据监控运维实战训练，社会培训、职业技能等级证书培训及考试	大数据集群运维监控、实时数据采集、大数据集群运行问题排查、调优	高性能服务器	8 台	120
				计算机	40 台	
				交换机等设备	2 台	

3. 校外实训基地基本要求

具有稳定的校外实习基地；能提供数据采集、数据分析和大数据平台运维相关实习岗位，能涵盖当前相关产业发展的主流技术，可接纳 50 名学生的阶段实习和 200 名学生的顶岗实习；配备相应数量的指导教师对学生实习进行指导和管理；制定保证实习生日常工作、学习、生活的规章制度，签订安全协议，具有安全、保险保障。

目前专业与联想集团、浪潮（优派）教育有限公司共建人才培训基地，具有 2 家稳定的

校外实训基地。

（三）教学资源

1. 教材选用机制和要求

选用的教材均符合课程标准的基本要求，具有思想性、科学性、先进性和适用性。相同课程标准的同一门课程选用一种教材，若因教学需要的辅助教材，任课教师可提出拟选用教材，须经各教学实施部位审议通过方可使用。

中职阶段思想政治、语文、历史三科，必须使用国家统编教材。高职阶段必须使用国家统编的思想政治理论课教材、马克思主义理论研究和建设工程重点教材；中高职阶段其他公共基础课教材均选用国家规划教材。

中职阶段专业课程和高职阶段公共基础课程教材优先从国家和省级规划教材中选用，或从学校教材信息库中选用，或选用校企合作开发的工作手册式校本教材；不以岗位培训教材取代专业课程教材。

2. 图书文献配备要求

在数据采集处理、大数据分析、大数据部署运维等方向配备相关经典图书、国内外优秀实践案例资源文献，每个模块不低于10本教材和10册相关案例库，供学生拓展学习。

3. 数字化教学资源配置要求

为保障教学，采用智慧职教平台进行线上线下混合式的方式进行授课，时刻关注学生吸收和理解知识的程度，不断提高课堂质量和效率。推进信息化教学，为课程建设线上线下混合式立体化教学资源，结合北京市职业院校大数据专业教学资源库，包括专业文献、视频音频资料、电子教材、教辅材料、教学课件、案例库、行业政策、职业考评、就业创业信息等，形式多样、使用便捷、动态更新。同时，拥有合作企业——联想、百度、神州数码、京东等企业强大的资源平台，包括获得这些公司的项目开发与策划、技术支持、教学素材以及企业文化等的支撑。

（四）教学方法改革

基于学校整体的“三有课堂”教学策略，根据信息技术系专业特点，推行职业素养和职业能力高度融合的“三段十八式”教学模式。该模式涵盖教学设计、教学实施、教学反思三个教学阶段，教学设计阶段包括：确定教学主题、开展学情分析、确定教学目标及重难点、确定教学方法、准备教学资源、设计教学过程和设计课堂评价七个环节；教学实施阶段包括：课前准备、激趣导入、建立联结、明确目标、探究新知、联系反馈、实践应用、总结评价和课后巩固九个环节；教学反思阶段包括：教学评估到复盘反思两个环节，共十八个环节的教

学模式。

专业课程依据“**课证融通，项目引领、任务驱动、突出实训**”的思想进行教学设计，课程整体依据工作过程系统化的思想进行学习单元设计，按课次应用任务驱动式教学方法，每2/4学时选择一个完整工作任务，依据任务驱动式教学方法的要求进行教学设计。课程内容与职业能力等级证书、行业企业岗位认证证书的考核内容融合，形成对应的证书课程，提升学生学习课程的主动性，在专业课程的教授中注重职业素养教育，把课堂教学当作培养学生的主要场所，打造学生能力的基本保障，提升育人质量的重要途径。在培养学生技术技能的同时，把培养学生的综合职业能力落实到课前、课中、课后的教学全过程，以专业能力培养为载体，明确每个教学环节的综合职业能力培养点。围绕每个环节的教学目标及重难点，选用合适的教学方法和手段，使综合职业能力得以在每个教学环节中得到有效训练。同时课堂教学按照完整工作过程展开，让学生体验综合职业能力在真实工作情境中的需求和应用。

制定“行动导向教学法+信息化资源”课堂策略，提升教学过程有趣性。针对学生学习兴趣低下的现状采用行动导向教学法，针对信息时代下学生特点来开发形式多样的信息化资源，激发学习兴趣，提高学习效率。应用如“旋转木马”法、“拓展小组”法、“自由市场”法等多种教学法有侧重点的挖掘学生潜能。为学生创设线上线下结合的学习环境，专业引进知网平台、LMS网络学习平台、四叶草课程平台，设计并开发出大量的课程配套信息化资源，利用雨课堂、UMU、蓝墨云班课等学习APP使学生时时可学、处处可学。根据每门课程的性质不同，选取合适的教学方法及手段提高学生的参与度，以学生为主体，实现做中学，做中教。

（五）教学评价改革

严格落实培养目标和培养规格要求，依据学校提出的“两维三阶”的学业评价体系内容，在本课程的考核评价中，开展“专业能力+非专业能力”两维评价，开展“课堂评价、课程评价”两阶评价。评价过程中，注重发挥教师、学生、企业等评价主体的重要作用，进行多元化评价包括，学生自评，教师评价，学生互评，企业教师评价等多方式综合评价学生的学习效果。过程性评价与结果性评价相结合，过程评价强调平时成绩评价、实践技能评价，强化实践性教学环节的全过程管理与考核评价，而结果性评价将理论测评和实践结果评价相结合，符合教学诊断和质量监控要求，完善学生学习过程监测、评价与反馈机制，加强成果导向的评价引导，以评促学，引导学生自我管理、主动学习，提高学习效率。

1. 课堂评价

建立课堂评价机制，通过对学生学习成果、学习过程、学习状态的评价，采用学生自评、学生互评、教师评、第三方评的方式，检查学生课堂学习的有效性，专业能力和综合素养的培养确实落实到课堂上。

专业课平时成绩的评定突出过程评价，在每次课评价中既关注了学生的专业技能，又关注了学生的职业素养及方法能力等的培养，有利于学生综合职业能力的形成。任课教师可根据具体情况进行适当调整。

2. 课程评价

学期末，采用理论测试和综合项目测试的方式对学生进行考核，检查学生的专业能力和职业素养的成长确实达到或超过预期。

理论采用试卷测试的方式进行，采用定量的测试，检查学生对专业知识的掌握程度。综合项目测试由教师、企业专家共同形成考官，选取源自企业的真实工作任务作为考核项目，组织学生以小组合作的方式在规定的时间内完成项目任务，考官全程进行测评，可根据项目的难易对评价项目及权重进行适当调整。

除了对学生的专业能力进行测试外，在整体项目完成过程中还要考察学生体现出的主观性、创新性、使用价值、流程性、社会接受度等综合素养，每学期形成每个学生综合素养提升分析报告。

3. 专业实践评价

学生在专业实践期间，校企双方共同对专业实践学生进行评价，确定学生的专业能力和职业素养确实达到企业标准和要求。学校和企业要共同开发针对专业实践学生的评价标准，共同从不同的维度，有针对性的对学生进行综合测评，确实对学生的真实情况有一个客观而公正的认识。专业要善于从总的评价结果中发现专业教学中的优势与不足，不断改进专业教学，提升教学质量。

学生学习效果采用综合职业能力评价的手段进行。综合职业能力指岗位专业知识、技能和能力素养以外最基本、最关键的职业意识、职业，职业精神、态度和职业能力等基本职业素养集合。结合学校的实际情况，每学期从文化基础、职业能力、职业核心素养、企业实践能力四个维度，依据学生学习、实训过程中的表现和效果，对学生进行全方面的综合评价。

（六）质量诊断与改进

1. 建立专业建设和教学质量诊断与改进机制。健全专业教学质量监控管理制度，完善课堂教学、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、资源建设等方面质量标准建设，通过教学实施、过程监控、质量评价和持续改进，达成人才培养规格。

2. 完善教学管理机制。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，建立健全巡课、听课、评教、评学等制度，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. 建立毕业生跟踪反馈机制及社会评价机制。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. 建立专业教研活动机制。成立专业教研组，定期进行教学评估，针对重点、难点问题开展教研，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。

5. 建立人才培养方案实施的监管体系及修订机制。加强对人才培养方案实施情况的检查指导和必要的质量监测。形成人才培养方案修订机制，紧跟区域产业行业发展变化，结合人才培养质量，每年调整、更新专业人才培养方案，推进专业人才培养质量持续提升。

十、毕业要求

（一）学分要求和学分免修条件

中职阶段至少修满192学分。

高职阶段至少修满112学分。

学分免修条件：

1. 获得比赛荣誉

教育部组织的国家级比赛获奖，可抵扣2学分，学会组织的抵1学分；北京市一等奖以上奖项抵1学分。

2. 获得其他荣誉

获得市级“三好学生”“优秀学生干部”等同等荣誉，可酌情抵扣1-2学分。

（二）文化水平要求

达到中高职阶段学校公共基础课课程标准基本规定要求，并考核合格。

（三）思想道德要求

坚持正确的政治方向，爱国拥党，理想信念坚定，思想道德高尚，行为习惯良好，无违规违纪；参与学校“三路十八湾”德育体系要求的十八项核心教育内容学习，五年评价成绩合格；学生综合职业能力评价和大数据专业职业素养评价成绩合格。

（四）获取的职业证书要求

取得数据采集职业技能等级证书中级、数据应用开发与服务职业技能等级证书初级。

（五）综合实践要求

参与1项以上综合实践项目，并通过项目考核。

符合以上要求，并且须通过本专业人才培养方案规定的全部教学环节，并考核合格，可授予本专业高职学历毕业证书。

*（六）转段要求

转段升学工作参照北京市“5+2 中高职衔接”相关规定执行。学生完成前 5 年学业且成绩合格者，由中高职学校颁发高职毕业证书；学生完成全部学业且成绩合格者，由合作院校颁发本科毕业证书。

中高职阶段学生必须具有大数据技术专业 5+2 资格，需报考北京联合大学大数据工程技术专业 2 年制本科，转段考试成绩构成为：

学生转段总成绩=高职阶段总评成绩×50%+面试成绩×50%

高职阶段总评成绩按照百分制核算，科目均在 60 分（含）及以上者；面试成绩包括语言素养、专业素养两部分。两项成绩及格即可转段。

说明：

满足下述条件之一，可直接认定面试合格。

1. 获得市级三好学生、市级优秀学生干部、市级优秀团干部称号且没有处分者。
2. 在各类市级赛项中获得市级一等奖。
3. 在全国技能大赛中取得三等奖以上。
4. 若在中高职阶段取得 1 个或以上职业资格证书或职业技能等级证书。

附件 1: 教学进程安排表

[illegible][illegible]

	基础选修课	4	文献检索	000000Z15	2	36	18	18	选修	考查										第六、七、八学期任选2学分，共任选11学分。		
		5	知识产权保护	000000Z17	2	36	18	18	选修	考查												
		6	科学思维训练	000000Z18	2	36	18	18	选修	考查												
		7	大学生心理健康教育	000000G06	2	36	18	18	选修	考查												
		8	形势与政策	000000G03	2	36	18	18	必修	考试												
		公共基础选修课汇总				11	198	99	99	-	-	0	0	0	0	5	2	2	2	0	0	
		公共基础课汇总				186	3348	1833	1515	-	-	30	30	30	30	20	16	10	8	8	4	
专业课程	专业基础课	1	通用网络技术	7102Z0B01	2	36	18	18	必修	考试	2										▲群平台课	
		2	信息技术设备组装与维护	7102Z0B02	2	36	18	18	必修	考试		2									▲群平台课	
		3	大数据基础	710205Z0B03	2	36	18	18	必修	考试			2									
		4	Excel 数据分析技术	710205Z0B04	2	36	18	18	必修	考试				2								
		5	程序设计基础与应用	7102ZG0B01	10	180	90	90	必修	考试					6		4				▲群平台课	
		6	数据库技术应用	7102Z0B05	4	72	36	36	必修	考试						4					▲群平台课	
		7	前端设计与开发	7102Z0B06	2	36	18	18	必修	考试					2						▲群平台课	
		8	Linux 系统应用基础	7102Z0B07	4	72	36	36	必修	考试						4					▲群平台课	
		专业基础课汇总				28	504	252	252	-	-	2	2	2	2	8	8	4	0	0	0	
	专业核心课	1	互联网数据采集	710205Z1B01	4	72	18	54	必修	考试					4						书证融通	
		2	大数据组件安装部署	710205Z1B02	2	36	18	18	必修	考试						2					书证融通	
		3	数据清洗与预处理	710205Z1B03	4	72	36	36	必修	考试						4					书证融通	
		4	数据可视化	510205G0B01	4	72	36	36	必修	考试							4					

[illegible]

	8	数据采集证书技能训练	710205ZG0X08	2	36	18	18	选修	考查											
	9	数据应用开发与服 务证书技能训练	710205ZG0X09	2	36	18	18	选修	考查											
	10	初级信息系统项目 管理	710205ZG0X010	2	36	18	18	选修	考查											
	专业拓展课汇总			18	288	162	162	—	—	0	0	0	0	0	2	4	4	4	4	
	专业课汇总			118	2088	792	1332	—	—	2	2	2	2	12	16	18	20	20	22	
合计	选修课合计			29	522	261	261	—	—					5	4	6	6	4	4	
	必修课合计			275	4950	2364	2586	—	—	32	32	32	32	27	28	22	22	24	24	
	总计			304	5472	2625	2847	—	—	32	32	32	32	32	32	28	28	28	28	

附件 2：职业分析及课程转化表

1. 典型职业活动和工作任务分析表

编号	典型工作任务	典型工作任务描述	典型工作任务工作过程分析					
			1	2	3	4	5	6
1	离线数据采集	根据项目需求，分析数据源的数据结构，设计离线采集工作规划，并做任务分工，根据不同文档部署搭建采集环境、开发采集系统，完成数据入库等。	采集需求分析	数据采集设计	采集环境搭建	数据采集开发	数据存储	
2	实时数据采集	根据项目需求，对项目进行分解，设计实时采集工作规划，并做任务分工，根据不同文档部署搭建采集环境、开发采集系统，根据埋点组件采集数据，完成数据入库等。	采集需求分析	数据采集设计	采集环境搭建	数据采集开发	数据存储	埋点数据采集设计
3	数据清洗与预处理	利用工具包对数据进行填充、删除、缩放等处理，实现数据标准化、归一化等，保证数据有效、可对比。	维护基础数据	维护数据完整性	数据清洗			
4	大数据分析	了解业务背景，分析数据层面的需求，设计解决方案，使用工具进行统计分析及相关性分析，总结关键性发现和结论，撰写分析报告，给出合理建议。	业务需求对接	业务需求分析	分析报告撰写	数据指标设计与监控	数据实验	

5	大数据可视化	分析数据呈现需求，规划相应内容的展示形式，设计功能原型，选取工具，整合实现完整设计。	分析呈现需求	原型设计	选取工具，整合实现			
6	大数据 ETL	用相应的工具对数据进行抽取，对抽取的数据进行标准化转换，对转换后的数据加载到数据仓库，基于数据库，为业务提供数据服务。	数据抽取	数据转换	数据加载	数据服务		
7	大数据组件安装部署	下载安装大数据各组件，能够更改各组件配置，使其协调工作，开发程序并测试，确保各组件关联关系正常，使其正常运行。	各组件部署安装	组件配置以及联调	测试组件			
8	大数据集群运维监控	监控基础网络环境，设置监控指标并监控平台运作，做好监控和统计，分析业务系统的日常运行状态，查看告警提示，分析问题及原因，即时调整或反馈。	监控基础网络环境与平台	查看告警提示	描述告警提示问题与原因	输出分析报告		
9	大数据集群运行问题排查、调优	根据日志报错信息排查问题，通过 Yarn Web UI 查看任务运行情况，适当调整任务运行参数并多次尝试；调整 Yarn 队列参数，使其适应业务需求；根据组件选型、任务占用资源分类统计情况调整集群参数。	组件日志排查	任务程序跟踪排查	任务级别调优	队列级别	集群级别调优	
10	大数据应用软件开发与测试	熟悉业务流程，了解大数据平台数据开发流程、大数据平台数据处理逻辑和处理方法，针对业务需求进行梳理，分析合理性和可行性，确定需求解决方案，开发功能模块，开展各类测试，针对功能模块生产的数据进行验证，以确保功能模块正常运行，且结果准确可靠。	业务需求分析	根据业务需求开发功能模块	上线实施	基于新功能进行测试	数据准确性验证	
11	数据仓库建设	根据业务需求，进行数据仓库选型，形成数据仓库系统设计方案，设计数据仓库数据模型，构建数据仓库。	数据仓库设计	数据仓库构建	数据仓库管理			
12	数据挖掘	用一系列工程化的方式从原始数据中筛选出更好的数据特征，以提升模型的训练效果，包括数据预处理、特征选择、降维等；根据业务场景和数据表现形式选取合适的算法模型，进行训练、测试。	特征工程	模型构建	模型部署			

13	数据治理	建立完善数据治理体系，负责数据标准的制定和管理；制定数据资产管理制度与流程，维护数据资产目录；对数据质量进行监测评价和持续管控，定期形成数据治理分析报告；开展数据认责、数据模型、元数据管理、共享管理、安全管理等数据管理各领域工作，制定工作方案，推动实施。	数据标准制定与管理	数据资产管理	数据质量管理	数据管理推进		
----	------	---	-----------	--------	--------	--------	--	--

2. 职业能力分析表

典型职业活动名称:	典型职业活动序号:	工作任务及编号	1. 采集需求分析	2. 数据采集设计	3. 采集环境搭建	4. 数据采集开发
离线数据采集	A		5. 数据存储			
典型职业活动描述	1. 活动性质: 该典型职业活动是大数据行业的入门职业活动, 是数据采集工程师需要具备的基础工作技能。 2. 职业活动: 根据项目需求, 分析数据源的数据结构, 设计离线采集工作规划, 并做任务分工, 根据不同文档部署搭建采集环境、开发采集系统, 完成数据入库等。					
工作要求与考核标准	1. 掌握市面上常用数据库工具的使用, 能够分析数据源的数据结构。					
	2. 掌握阅读需求文档和原理的能力、系统设计的能力, 能够为采集工作做顶层设计。					
	3. 能够根据不同文档去部署搭建采集环境。					
	4. 掌握一种数据平台或者具备开发采集数据平台的能力。					
	5. 能够设计数据入库字段、入库接口调用及开发、不同数据库备份的分布式设计等。					
能力类别	编号	能力描述				学习程度
职业能力	P-A-1.1	了解市面上常用的数据库(如MySQL、Oracle、SQL Server、LDAP、MongoDB、Hadoop等), 对于数据源提供的各种数据库类别有所掌握, 能够分析数据源的数据结构, 了解相应数据源的数据协议格式, 完成数据采集需求分析工作。				L3应用
	P-A-2.1	能够根据网页配置出字段规则, Xpath或CSS规则。				L3应用
	P-A-2.2	能够使用Redis数据库对网址进行过滤, 去重。				L3应用
	P-A-2.3	掌握阅读需求文档和原理的能力、系统设计的能力, 能够为采集工作做顶层设计。				L3应用
	P-A-3.1	具备数据展示、容错的分析能力, 具备良好的动手能力, 能够根据不同文档去部署采集环境。				L3应用
	P-A-4.1	熟悉各种常见的代码框架、编程语言、插值逻辑、增删改查逻辑等, 了解相应业务中的数据字典标准, 制定代码映射逻辑进行数据代码映射。				L3应用
	P-A-4.2	掌握HTML, 图片识别OCR, 能够进行视屏流解码和取图。				L3应用
	P-A-4.3	掌握反爬技术, 能够根据不同的站点, 采取不同策略, 完成数据采集开发部署任务。				L3应用
	P-A-5.1	具备数据的入库字段设计能力, 能够在一款数据库中增删改查等。				L3应用
	P-A-5.2	掌握Http, Tcp等入库接口调用及开发的能力, 掌握数据结构知识, 能够分析入库时数据的复杂度, 索引等。				L3应用
	P-A-5.3	掌握不同结点, 不同数据库备份的分布式设计能力。				L3应用
通用能力	G-A-1	自学能力: 能够主动通过自学了解软件与信息技术行业信息, 掌握大数据行业最新技术, 熟悉数据采集处理等岗位相关技能。				L3应用
	G-A-2	沟通能力: 具备良好的沟通理解能力, 能够明确项目意图, 在遇到问题时, 能够主动沟通, 确保各方共识项目意图。				L3应用

	G-A-3	协作能力：具备专业工作语言能力，正确使用专有名词，规范专业术语与行业用语，能够顺畅地与团队沟通交流。	L3应用
	G-A-4	社会人文能力：能够理解和尊重不同区域的文化与当地的风土人情。	L3应用
	G-A-5	身心承受力：具备强健的体魄与良好的心理素质，适应兼具脑力与体力双重劳动的弹性工作制与加班熬夜情况。	L3应用
社会能力	S-A-1	合作意识：善于与团队协作、沟通解决问题；善于外部交流，有效跨部门推进工作进程。	L3应用
	S-A-2	法律安全意识：充分了解大数据行业法规法则，拥有版权意识，自觉维护行业良好风气。	L3应用
	S-A-3	社会责任感：具有社会责任感，能够传递正确的社会价值，时刻维护国家的荣誉与形象。	L3应用
	S-A-4	职业道德：树立正确的职业观，注意不侵犯他人肖像权、保障他人隐私，爱岗敬业、德技并修。	L3应用
发展能力	D-A-1	创业意识：具备一定的创业意识，培养立足岗位创新创业能力。	L3应用
	D-A-2	前瞻眼光：关注行业发展动态了解行业现状与热门资讯，分析行业未来发展趋势与底层逻辑，洞察行业发展走向，随时调整自我职业规划，顺应时代发展。	L3应用
	D-A-3	自我发展：有关关注行业最新前沿技术的意识，不断进行自我知识更新与学习。	L3应用
	D-A-4	纵深发展：善于把握发展实际工作技能，切忌眼高手低、泛而不专，不断在行业内深耕技能，向纵深发展。	L3应用

典型职业活动名称：	典型职业活动序号：	工作任务及编号	1. 采集需求分析	2. 数据采集设计	3. 采集环境搭建	4. 数据采集开发
实时数据采集	B		5. 数据存储	6. 埋点数据采集设计		
典型职业活动描述	1. 活动性质：该典型职业活动是大数据行业的进阶职业活动，是数据采集工程师需要具备的进阶工作技能。 2. 职业活动：根据项目需求，对项目进行分解，设计实时采集工作规划，并做任务分工，根据不同文档部署搭建采集环境、开发采集系统，根据埋点组件采集数据，完成数据入库等。					
工作要求与考核标准	1. 掌握市面上常用数据库工具的使用，能够分析数据源的数据结构。					
	2. 掌握阅读需求文档和原理的能力、系统设计的能力，能够为采集工作做顶层设计。					
	3. 能够根据不同文档去部署搭建采集环境。					
	4. 掌握一种数据平台或者具备开发采集数据平台的能力。					
	5. 能够设计数据入库字段、入库接口调用及开发、不同数据库备份的分布式设计等。					
	6. 能够评估需求中需要的数据及埋点的来源、触发机制，并进行埋点数据采集设计。					
能力类别	编号	能力描述				学习程度
职业能力	P-B-1.1	了解市面上常用的数据库（如MySQL、Oracle、SQL Server、LDAP、MongoDB、Hadoop等），对于数据源提供的各种数据库类别有所掌握，能够分析数据源的数据结构，了解相应数据源的数据协议				L3应用

		格式，完成数据采集需求分析工作。	
	P-B-2.1	能够根据网页配置出字段规则，Xpath或CSS规则。	L3应用
	P-B-2.2	能够使用Redis数据库对网址进行过滤，去重。	L3应用
	P-B-2.3	掌握阅读需求文档和原理的能力、系统设计的能力，能够为采集工作做顶层设计。	L3应用
	P-B-3.1	具备数据展示、容错的分析能力，具备良好的动手能力，能够根据不同文档去部署采集环境。	L3应用
	P-B-4.1	熟悉各种常见的代码框架、编程语言、插值逻辑、增删改查逻辑等，了解相应业务中的数据字典标准，制定代码映射逻辑进行数据代码映射。	L3应用
	P-B-4.2	掌握HTML，图片识别OCR，能够进行视屏流解码和取图。	L3应用
	P-B-4.3	掌握反爬技术，能够根据不同的站点，采取不同策略，完成数据采集开发部署任务。	L3应用
	P-B-5.1	具备数据的入库字段设计能力，能够在一款数据库中增删改查等。	L3应用
	P-B-5.2	掌握http，tcp等入库接口调用及开发的能力，掌握数据结构知识，能够分析入库时数据的复杂度，索引等。	L3应用
	P-B-5.3	掌握不同结点，不同数据库备份的分布式设计能力。	L3应用
	P-B-6.1	在需要如用户行为数据的情况下，利用对业务的理解 and 数据敏感性，评估需求中需要的数据及埋点的来源、触发机制等，完成数据需求评估。	L3应用
	P-B-6.2	根据数据需要，设计埋点组件进行数据埋点，理清用户生命周期和行为路径，抽象出每个路径的关键指标，确定数据上报方式，完成数据埋点设计。	L3应用
通用能力	G-B-1	自学能力：能够主动通过自学了解软件与信息技术行业信息，掌握大数据行业最新技术，熟悉数据采集处理等岗位相关技能。	L3应用
	G-B-2	沟通能力：具备良好的沟通理解能力，能够明确项目意图，在遇到问题时，能够主动沟通，确保各方共识项目意图。	L3应用
	G-B-3	协作能力：具备专业工作语言能力，正确使用专有名词，规范专业术语与行业用语，能够顺畅地与团队沟通交流。	L3应用
	G-B-4	社会人文能力：能够理解和尊重不同区域的文化与当地的风土人情。	L3应用
	G-B-5	身心承受力：具备强健的体魄与良好的心理素质，适应兼具脑力与体力双重劳动的弹性工作制与加班熬夜情况。	L3应用
社会能力	S-B-1	合作意识：善于与团队协作、沟通解决问题；善于外部交流，有效跨部门推进工作进程。	L3应用
	S-B-2	法律安全意识：充分了解大数据行业法规法则，拥有版权意识，自觉维护行业良好风气。	L3应用
	S-B-3	社会责任感：具有社会责任感，能够传递正确的社会价值，时刻维护国家的荣誉与形象。	L3应用
	S-B-4	职业道德：树立正确的职业观，注意不侵犯他人肖像权、保障他人隐私，爱岗敬业、德技并修。	L3应用
发展能力	D-B-1	创业意识：具备一定的创业意识，培养立足岗位创新创业能力。	L3应用
	D-B-2	前瞻眼光：关注行业发展动态了解行业现状与热门资讯，分析行业未来发展趋势与底层逻辑，洞察行业发展走向，随时调整自我职业规划，顺应时代发展。	L3应用
	D-B-3	自我发展：有关关注行业最新前沿技术的意识，不断进行自我知识更新与学习。	L3应用
	D-B-4	纵深发展：善于把握发展实际工作技能，切忌眼高手低、泛而不专，不断在行业内深耕技能，向纵深发展。	L3应用

典型职业活动名称:	典型职业活动序号:	工作任务及编号	1. 维护基础数据	2. 维护数据完整性	3. 数据清洗	
数据清洗与预处理	C					
典型职业活动描述	1. 活动性质: 该典型职业活动是大数据行业的基础职业活动, 是数据处理工程师需要具备的基础工作技能。 2. 职业活动: 利用工具包对数据进行填充、删除、缩放等处理, 实现数据标准化、归一化等, 保证数据有效、可对比。					
工作要求与考核标准	1. 能使用Sqoop维护基础数据。					
	2. 能使用furm、keylan 提取或同步数据, 维护数据的完整性。					
	3. 能够利用python/hive/sql等工具, 去除数据中的脏数据, 如乱码、语句不通顺、内容不规范等数据。					
	4. 能够在不改变原数据分布的前提下, 将数据按比例缩放, 使之落入一个限定的区间, 使数据之间具有可比性。					
能力类别	编号	能力描述				学习程度
职业能力	P-C-1.1	使用Sqoop对接数据库, 应用sqoop的sql方式, 增量方式实现数抽数据库, 并且能使用sqoop导入数据到数据库。				L3应用
	P-C-2.1	使用furm、keylan 提取或同步数据到hadoop或hive或hbase平台。				L3应用
	P-C-3.1	在数据脏、乱、差情况下, 利用python/hive/sql等工具将脏数据替换、删除等操作, 以保证数据准确性（即数据中无噪声）的标准完成数据清洗任务。				L3应用
	P-C-3.2	在不完整的情况下, 利用python/hive/sql等工具将脏数据填充或删除等操作, 以保证数据完整性的标准完成数据清洗任务。				L3应用
	P-C-3.3	在数据不规范的情况下, 利用人工检验或脚本校准处理数据, 保证数据内容的符合行业标准。				L3应用
	P-C-3.4	在数据量纲不统一情况下, 利用工具包（Pandas、scikit-learn、Numpy等）将数据按比例缩放, 使之落入一个限定的区间, 使数据之间具有可比性（标准化）。				L3应用
	P-C-3.5	在数据单位不统一情况下, 利用工具包（Pandas、scikit-learn、Numpy等）对数据集进行区间缩放, 把有单位的数据转化为没有单位的数据, 即统一数据的衡量标准, 消除单位的影响（归一化）。				L3应用
通用能力	G-C-1	自学能力: 能够主动通过自学了解软件与信息技术行业信息, 掌握大数据行业最新技术, 熟悉数据采集处理等岗位相关技能。				L3应用
	G-C-2	沟通能力: 具备良好的沟通理解能力, 能够明确项目意图, 在遇到问题时, 能够主动沟通, 确保各方共识项目意图。				L3应用
	G-C-3	协作能力: 具备专业工作语言能力, 正确使用专有名词, 规范专业术语与行业用语, 能够顺畅地与团队沟通交流。				L3应用
	G-C-4	社会人文能力: 能够理解和尊重不同区域的文化与当地的风土人情。				L3应用
	G-C-5	身心承受力: 具备强健的体魄与良好的心理素质, 适应兼具脑力与体力双重劳动的弹性工作制与加班熬夜情况。				L3应用
社会能力	S-C-1	合作意识: 善于与团队协作、沟通解决问题; 善于外部交流, 有效跨部门推进工作进程。				L3应用
	S-C-2	法律安全意识: 充分了解大数据行业法规规则, 拥有版权意识, 自觉维护行业良好风气。				L3应用

	S-C-3	社会责任感：具有社会责任感，能够传递正确的社会价值，时刻维护国家的荣誉与形象。	L3应用
	S-C-4	职业道德：树立正确的职业观，注意不侵犯他人肖像权、保障他人隐私，爱岗敬业、德技并修。	L3应用
发展能力	D-C-1	创业意识：具备一定的创业意识，培养立足岗位创新创业能力。	L3应用
	D-C-2	前瞻眼光：关注行业发展动态了解行业现状与热门资讯，分析行业未来发展趋势与底层逻辑，洞察行业发展走向，随时调整自我职业规划，顺应时代发展。	L3应用
	D-C-3	自我发展：有关关注行业最新前沿技术的意识，不断进行自我知识更新与学习。	L3应用
	D-C-4	纵深发展：善于把握发展实际工作技能，切忌眼高手低、泛而不专，不断在行业内深耕技能，向纵深发展。	L3应用

典型职业活动名称：	典型职业活动序号：	工作任务及编号	1. 业务需求对接	2. 业务需求分析	3. 分析报告撰写	4. 数据指标设计与监控
大数据分析	D		5. 数据实验			
典型职业活动描述	1. 活动性质：该典型职业活动是大数据行业的进阶职业活动，是数据分析工程师需要具备的工作技能。 2. 职业活动：了解业务背景，分析数据层面的需求，设计解决方案，使用工具进行统计分析与相关性分析，总结关键性发现和结论，撰写分析报告，给出合理建议。					
工作要求与考核标准	1. 能够与业务对接需求，了解需求产生的背景，讨论问题可能产生的数据原因。					
	2. 能够梳理业务流程，通过指标拆解，确定分析目标，运用MECE原则、AARR模型、人货场模型等方法指标体系构建，产出合理的分析思路和解决方案。					
	3. 了解常用数据库，对于数据源提供的各种数据库类别都有所掌握，能够分析数据源的数据结构，撰写SQL语句从数据库提取数据。					
	4. 能够使用Excel、Python等工具对数据进行分析。					
	5. 能够使用Excel、Tableau等工具选择合适的可视化图表类型展示分析结果。					
	6. 能够总结关键性发现和结论，并给予合理建议，输出结构化分析报告。					
	7. 能够与研发团队沟通，协调开发资源，下钻可能影响变动的数据维度，定位数据波动原因，脱离出具体的数据分析，主动设计分析业务指标。					
能力类别	编号	能力描述				学习程度
职业能力	P-D-1.1	与业务对接需求，了解需求产生的背景，讨论问题可能产生的数据原因，沟通讨论业务需求的必要性、可行性。				L2理解
	P-D-1.2	梳理业务流程，通过指标拆解，确定分析目标，运用MECE原则、AARR模型、人货场模型等方法指标体系构建，产出合理的分析思路和解决方案。				L2理解
	P-D-2.1	了解常用数据库（如MySQL、Oracle、SQL Server、LDAP、MongoDB、Hive等），对于数据源提供的各种数据库类别都有所掌握，能够分析数据源的数据结构，撰写SQL语句从数据库提取数据。				L3应用
	P-D-2.2	使用Excel、Python等工具对数据进行分组、聚合以及数学运算，完成描述性统计分析，使用对比分析、漏斗分析、用户分群、趋势分析、相关性分析等方法进行分析，得到分析发现和结论。				

	P-D-3.1	使用Excel、Tableau等工具选择合适的可视化图表类型展示分析结果，总结关键性发现和结论，并给予合理建议，输出结构化分析报告，并合理、清晰地阐述分析结论。	L3应用
	P-D-4.1	在核心数据例行监控需求下，规划监控看板，展示需要监控的数据，在确定规划方案情况下，判断看板报表的合理性。	L3应用
	P-D-4.2	与研发团队沟通，协调开发资源，完成监控看板搭建，下钻可能影响变动的数据维度，定位数据波动原因，能够脱离出具体的数据分析，主动设计分析业务指标。	L3应用
	P-D-5.1	在业务方的要求下，通过确定实验样本量、实验周期等参数完成A/B实验设计，对接研发团队，确定实验对象，上线A/B实验，运用统计学方法分析实验结果，选择最优解决方案。	L3应用
通用能力	G-D-1	自学能力：能够主动通过自学了解软件与信息技术行业信息，掌握大数据行业最新技术，熟悉数据采集处理等岗位相关技能。	L3应用
	G-D-2	沟通能力：具备良好的沟通理解能力，能够明确项目意图，在遇到问题时，能够主动沟通，确保各方共识项目意图。	L3应用
	G-D-3	协作能力：具备专业工作语言能力，正确使用专有名词，规范专业术语与行业用语，能够顺畅地与团队沟通交流。	L3应用
	G-D-4	社会人文能力：能够理解和尊重不同区域的文化与当地的风土人情。	L3应用
	G-D-5	身心承受力：具备强健的体魄与良好的心理素质，适应兼具脑力与体力双重劳动的弹性工作制与加班熬夜情况。	L3应用
社会能力	S-D-1	合作意识：善于与团队协作、沟通解决问题；善于外部交流，有效跨部门推进工作进程。	L3应用
	S-D-2	法律安全意识：充分了解大数据行业法规规则，拥有版权意识，自觉维护行业良好风气。	L3应用
	S-D-3	社会责任感：具有社会责任感，能够传递正确的社会价值，时刻维护国家的荣誉与形象。	L3应用
	S-D-4	职业道德：树立正确的职业观，注意不侵犯他人肖像权、保障他人隐私，爱岗敬业、德技并修。	L3应用
发展能力	D-D-1	创业意识：具备一定的创业意识，培养立足岗位创新创业能力。	L3应用
	D-D-2	前瞻眼光：关注行业发展动态了解行业现状与热门资讯，分析行业未来发展趋势与底层逻辑，洞察行业发展走向，随时调整自我职业规划，顺应时代发展。	L3应用
	D-D-3	自我发展：有关关注行业最新前沿技术的意识，不断进行自我知识更新与学习。	L3应用
	D-D-4	纵深发展：善于把握发展实际工作技能，切忌眼高手低、泛而不专，不断在行业内深耕技能，向纵深发展。	L3应用

典型职业活动名称：	典型职业活动序号：	工作任务及编号	1. 分析呈现需求	2. 原型设计	3. 选取工具，整合实现	
大数据可视化	E					
典型职业活动描述	1. 活动性质：该典型职业活动是大数据行业的进阶职业活动，是数据分析工程师需要具备的工作技能。 2. 职业活动：分析数据呈现需求，规划相应内容的展示形式，设计功能原型，选取工具，整合实现完整设计。					
工作要	1. 熟练使用echarts，data等工具，掌握不同类型对应的常规展示形式，能够规划相应内容的					

求与考核标准	展示形式。		
	2. 熟练使用原型制作工具，如Axure、mockplus等，按照设计内容出具相应的功能原型。		
	3. 能够使用工具完成原型完整设计。		
能力类别	编号	能力描述	学习程度
职业能力	P-E-1.1	熟练使用echarts，data等工具，掌握不同类型对应的常规展示形式，熟悉本次设计内容相应的设计使用场景，规划相应内容的展示形式，完成呈现需求分析。	L3应用
	P-E-2.1	熟练使用原型制作工具，如Axure、mockplus等，按照设计内容出具相应的功能原型。	L3应用
	P-E-3.1	能够完成原型完整设计（UI），重点关注动态内容设计。	L3应用
	P-E-3.2	能够使用工具完成相应设计实现（接口衔接）。	L3应用
	P-E-3.3	能够完成关键元素的复盘（交互实现，动态数据的实现）。	L3应用
通用能力	G-E-1	自学能力：能够主动通过自学了解软件与信息技术行业信息，掌握大数据行业最新技术，熟悉数据采集处理等岗位相关技能。	L3应用
	G-E-2	沟通能力：具备良好的沟通理解能力，能够明确项目意图，在遇到问题时，能够主动沟通，确保各方共识项目意图。	L3应用
	G-E-3	协作能力：具备专业工作语言能力，正确使用专有名词，规范专业术语与行业用语，能够顺畅地与团队沟通交流。	L3应用
	G-E-4	社会人文能力：能够理解和尊重不同区域的文化与当地的风土人情。	L3应用
	G-E-5	身心承受力：具备强健的体魄与良好的心理素质，适应兼具脑力与体力双重劳动的弹性工作制与加班熬夜情况。	L3应用
社会能力	S-E-1	合作意识：善于与团队协作、沟通解决问题；善于外部交流，有效跨部门推进工作进程。	L3应用
	S-E-2	法律安全意识：充分了解大数据行业法规法则，拥有版权意识，自觉维护行业良好风气。	L3应用
	S-E-3	社会责任感：具有社会责任感，能够传递正确的社会价值，时刻维护国家的荣誉与形象。	L3应用
	S-E-4	职业道德：树立正确的职业观，注意不侵犯他人肖像权、保障他人隐私，爱岗敬业、德技并修。	L3应用
发展能力	D-E-1	创业意识：具备一定的创业意识，培养立足岗位创新创业能力。	L3应用
	D-E-2	前瞻眼光：关注行业发展动态了解行业现状与热门资讯，分析行业未来发展趋势与底层逻辑，洞察行业发展走向，随时调整自我职业规划，顺应时代发展。	L3应用
	D-E-3	自我发展：有关关注行业最新前沿技术的意识，不断进行自我知识更新与学习。	L3应用
	D-E-4	纵深发展：善于把握发展实际工作技能，切忌眼高手低、泛而不专，不断在行业内深耕技能，向纵深发展。	L3应用

典型职业活动名称：	典型职业活动序号：	工作任务及编号	1. 数据抽取	2. 数据转换	3. 数据加载	4. 数据服务
大 数 据 ETL	F					

典型职业活动描述	1. 活动性质：该典型职业活动是大数据行业的基础职业活动，是大数据ETL工程师需要具备的工作技能。 2. 职业活动：用相应的工具对数据进行抽取，对抽取的数据进行标准化转换，对转换后的数据加载到数据仓库，基于数据库，为业务提供数据服务。		
工作要求与考核标准	1. 能够对接上游数据接口，参与制定ETL策略，使用ETL工具将业务数据导入数据仓库。 2. 能够使用相应的工具对数据进行抽取。 3. 能够使用相应的工具对抽取的数据进行标准化转换。 4. 能够使用相应的工具对转换后的数据加载到数据仓库。 5. 能够参与数据仓库ETL流程优化及解决ETL相关技术问题。		
能力类别	编号	能力描述	学习程度
职业能力	P-F-1.1	对接上游数据接口，参与制定ETL策略，使用ETL工具将业务数据导入数据仓库。	L3应用
	P-F-1.2	编写SQL/Python/Shell脚本从源数据系统抽取数据；使用Kettle工具完成关系型数据库数据抽取；使用日志采集工具完成业务平台日志数据抽取。	L3应用
	P-F-2.1	熟悉Hive、Spark等组件，将数据按照业务需求转换为目的数据源要求的格式，分析数据格式，按照要求清洗业务数据，完成数据转换。	L3应用
	P-F-3.1	熟悉Sqoop、Hive、Spark等组件，将转换结果装载至目的数据源，对输出数据质量进行核验，对不符合要求的数据进行纠偏。	L3应用
	P-F-4.1	参与数据仓库ETL流程优化及解决ETL相关技术问题，根据业务需要，提供面向业务的报表、数据提取等数据服务，能够对数据仓库有关数据问题进行监控并提供解决方案。	L3应用
通用能力	G-F-1	自学能力：能够主动通过自学了解软件与信息技术行业信息，掌握大数据行业最新技术，熟悉数据采集处理等岗位相关技能。	L3应用
	G-F-2	沟通能力：具备良好的沟通理解能力，能够明确项目意图，在遇到问题时，能够主动沟通，确保各方共识项目意图。	L3应用
	G-F-3	协作能力：具备专业工作语言能力，正确使用专有名词，规范专业术语与行业用语，能够顺畅地与团队沟通交流。	L3应用
	G-F-4	社会人文能力：能够理解和尊重不同区域的文化与当地的风土人情。	L3应用
	G-F-5	身心承受力：具备强健的体魄与良好的心理素质，适应兼具脑力与体力双重劳动的弹性工作制与加班熬夜情况。	L3应用
社会能力	S-F-1	合作意识：善于与团队协作、沟通解决问题；善于外部交流，有效跨部门推进工作进程。	L3应用
	S-F-2	法律安全意识：充分了解大数据行业法规法则，拥有版权意识，自觉维护行业良好风气。	L3应用
	S-F-3	社会责任感：具有社会责任感，能够传递正确的社会价值，时刻维护国家的荣誉与形象。	L3应用
	S-F-4	职业道德：树立正确的职业观，注意不侵犯他人肖像权、保障他人隐私，爱岗敬业、德技并修。	L3应用
发展能力	D-F-1	创业意识：具备一定的创业意识，培养立足岗位创新创业能力。	L3应用
	D-F-2	前瞻眼光：关注行业发展动态了解行业现状与热门资讯，分析行业未来发展趋势与底层逻辑，洞察行业发展走向，随时调整自我职业规划，顺应时代发展。	L3应用
	D-F-3	自我发展：有关关注行业最新前沿技术的意识，不断进行自我知识	L3应用

		更新与学习。	
	D-F-4	纵深发展：善于把握发展实际工作技能，切忌眼高手低、泛而不专，不断在行业内深耕技能，向纵深发展。	L3应用

典型职业活动名称：	典型职业活动序号：	工作任务及编号	1. 各组件部署安装	2. 组件配置以及联调	3. 测试组件	
大数据组件安装部署	G					
典型职业活动描述	1. 活动性质：该典型职业活动是大数据行业的基础职业活动，是大数据运维工程师需要具备的工作技能。 2. 职业活动：下载安装大数据各组件，能够更改各组件配置，使其协调工作，开发程序并测试，确保各组件关联关系正常，使其正常运行。					
工作要求与考核标准	1. 熟悉 Linux 命令、Linux 发行版、组件发行版，能够下载安装各组件。 2. 熟悉各组件配置方法、通信机制，更改各组件配置，使其协调工作。 3. 开发程序并测试，测试各组件关联关系正常，使其正常运行。					
能力类别	编号	能力描述				学习程度
职业能力	P-G-1.1	熟悉 Linux 命令、Linux 发行版、组件发行版，能够下载安装各组件。				L3应用
	P-G-2.1	熟悉各组件配置方法、通信机制，更改各组件配置，使其协调工作。				L3应用
	P-G-3.1	开发程序并测试，测试各组件关联关系正常，使其正常运行。				L3应用
通用能力	G-G-1	自学能力：能够主动通过自学了解软件与信息技术行业信息，掌握大数据行业最新技术，熟悉数据采集处理等岗位相关技能。				L3应用
	G-G-2	沟通能力：具备良好的沟通理解能力，能够明确项目意图，在遇到问题时，能够主动沟通，确保各方共识项目意图。				L3应用
	G-G-3	协作能力：具备专业工作语言能力，正确使用专有名词，规范专业术语与行业用语，能够顺畅地与团队沟通交流。				L3应用
	G-G-4	社会人文能力：能够理解和尊重不同区域的文化与当地的风土人情。				L3应用
	G-G-5	身心承受力：具备强健的体魄与良好的心理素质，适应兼具脑力与体力双重劳动的弹性工作制与加班熬夜情况。				L3应用
社会能力	S-G-1	合作意识：善于与团队协作、沟通解决问题；善于外部交流，有效跨部门推进工作进程。				L3应用
	S-G-2	法律安全意识：充分了解大数据行业法规法则，拥有版权意识，自觉维护行业良好风气。				L3应用
	S-G-3	社会责任感：具有社会责任感，能够传递正确的社会价值，时刻维护国家的荣誉与形象。				L3应用
	S-G-4	职业道德：树立正确的职业观，注意不侵犯他人肖像权、保障他人隐私，爱岗敬业、德技并修。				L3应用
发展能力	D-G-1	创业意识：具备一定的创业意识，培养立足岗位创新创业能力。				L3应用
	D-G-2	前瞻眼光：关注行业发展动态了解行业现状与热门资讯，分析行业未来发展趋势与底层逻辑，洞察行业发展走向，随时调整自我职业规划，顺应时代发展。				L3应用
	D-G-3	自我发展：有关关注行业最新前沿技术的意识，不断进行自我知识				L3应用

		更新与学习。	
	D-G-4	纵深发展：善于把握发展实际工作技能，切忌眼高手低、泛而不专，不断在行业内深耕技能，向纵深发展。	L3应用

典型职业活动名称：	典型职业活动序号：	工作任务及编号	1. 监控基础网络环境与平台	2. 查看告警提示	3. 描述告警提示问题与原因	4. 输出分析报告
大数据集群运维监控	H					
典型职业活动描述	1. 活动性质：该典型职业活动是大数据行业的基础职业活动，是大数据运维工程师需要具备的工作技能。 2. 职业活动：监控基础网络环境，设置监控指标并监控平台运作，做好监控和统计，分析业务系统的日常运行状态，查看告警提示，分析问题及原因，即时调整或反馈。					
工作要求与考核标准	1. 熟悉 Linux 基础环境，对项目的各载体机器实时监控（CPU、内存、磁盘等状态）。 2. 熟悉项目需求的网络配置，对网端和接口，轮循监控，保持其可靠性。 3. 熟悉业务数据需求，对已收集数据做提取分析，保障收集数据干净可靠；了解项目开发逻辑，对产生结果进行同比，环比等分析操作，随时监控产出结果是否达到预期。 4. 能够建立良好的告警（电话，短信，邮件）即时反馈；建立监控的 web 展示平台，清晰的展示出监控收集数据分布，以供后续项目改进。 5. 对数据库各访问 SQL 的性能监控，持续收集各业务中 SQL 操作，即时调整，减轻数据库压力。 6. 对大数据项目-站点流量监控，发展各热点数据，对热点数据合理规划，加强后台的承受能力；硬件和网络配置。 7. 针对业务日志，做好监控和统计，分析业务系统的日常运行状态，分析调整项目开发质量。					
能力类别	编号	能力描述				学习程度
职业能力	P-H-1.1	熟悉 Linux 基础环境，对项目的各载体机器实时监控（CPU、内存、磁盘等状态）。				L3应用
	P-H-1.2	熟悉项目需求的网络配置，对网端和接口，轮循监控，保持其可靠性。				L3应用
	P-H-1.3	熟悉业务数据需求，对已收集数据做提取分析，保障收集数据干净可靠；了解项目开发逻辑，对产生结果进行同比，环比等分析操作，随时监控产出结果是否达到预期。				L3应用
	P-H-2.1	建立良好的告警（电话，短信，邮件）即时反馈；建立监控的 web 展示平台，清晰的展示出监控收集数据分布，以供后续项目改进。				L3应用
	P-H-3.1	对数据库各访问 SQL 的性能监控，持续收集各业务中 SQL 操作，即时调整，减轻数据库压力。				L3应用
	P-H-3.2	对大数据项目-站点流量监控，发展各热点数据，对热点数据合理规划，加强后台的承受能力。				L3应用
	P-H-3.3	针对业务日志，做好监控和统计，分析业务系统的日常运行状态，分析调整项目开发质量。				L3应用
	P-H-4.1	具备一定的写作能力，要求语句通畅，给客户提交满意的报告。				L3应用
通用能力	G-H-1	自学能力：能够主动通过自学了解软件与信息技术行业信息，掌握大数据行业最新技术，熟悉数据采集处理等岗位相关技能。				L3应用
	G-H-2	沟通能力：具备良好的沟通理解能力，能够明确项目意图，在遇				L3应用

		到问题时，能够主动沟通，确保各方共识项目意图。	
	G-H-3	协作能力：具备专业工作语言能力，正确使用专有名词，规范专业术语与行业用语，能够顺畅地与团队沟通交流。	L3应用
	G-H-4	社会人文能力：能够理解和尊重不同区域的文化与当地的风土人情。	L3应用
	G-H-5	身心承受力：具备强健的体魄与良好的心理素质，适应兼具脑力与体力双重劳动的弹性工作制与加班熬夜情况。	L3应用
社会能力	S-H-1	合作意识：善于与团队协作、沟通解决问题；善于外部交流，有效跨部门推进工作进程。	L3应用
	S-H-2	法律安全意识：充分了解大数据行业法规法则，拥有版权意识，自觉维护行业良好风气。	L3应用
	S-H-3	社会责任感：具有社会责任感，能够传递正确的社会价值，时刻维护国家的荣誉与形象。	L3应用
	S-H-4	职业道德：树立正确的职业观，注意不侵犯他人肖像权、保障他人隐私，爱岗敬业、德技并修。	L3应用
发展能力	D-H-1	创业意识：具备一定的创业意识，培养立足岗位创新创业能力。	L3应用
	D-H-2	前瞻眼光：关注行业发展动态了解行业现状与热门资讯，分析行业未来发展趋势与底层逻辑，洞察行业发展走向，随时调整自我职业规划，顺应时代发展。	L3应用
	D-H-3	自我发展：有关关注行业最新前沿技术的意识，不断进行自我知识更新与学习。	L3应用
	D-H-4	纵深发展：善于把握发展实际工作技能，切忌眼高手低、泛而不专，不断在行业内深耕技能，向纵深发展。	L3应用

典型职业活动名称：	典型职业活动序号：		1. 组件日志排查	2. 任务程序跟踪排查	3. 任务级别调优	4. 队列级别
大数据集群运行问题排查、调优	I	工作任务及编号	5. 集群级别调优			
典型职业活动描述	1. 活动性质：该典型职业活动是大数据行业的基础职业活动，是大数据运维工程师需要具备的工作技能。 2. 职业活动：根据日志报错信息排查问题，通过Yarn Web UI查看任务运行情况，适当调整任务运行参数并多次尝试；调整Yarn队列参数，使其适应业务需求；根据组件选型、任务占用资源分类统计情况调整集群参数。					
工作要求与考核标准	1. 熟悉 Linux 命令、熟悉各组件常见报错日志类型，了解组件参数意义，根据报错信息分类报错类型。					
	2. 能够通过 Yarn Web UI 查看任务运行情况。					
	3. 熟悉 Yarn Web UI 各功能，了解各组件运行原理，能够根据任务程序运行情况，适当调整任务运行参数并多次尝试。					
	4. 能够调整 Yarn 队列参数，使其适应业务需求。					
	5. 涉及到组件选型、任务占用资源分类统计，并根据情况调整集群参数。					
能力类别	编号	能力描述				学习程度
职业	P-I-1.1	熟悉 Linux 命令、熟悉各组件常见报错日志类型，了解组件参数				L3应用

能力		意义，根据报错信息分类报错类型。	
	P-I-2.1	能够通过 Yarn Web UI 查看任务运行情况。	L3应用
	P-I-3.1	熟悉 Yarn Web UI 各功能，了解各组件运行原理，能够根据任务程序运行情况，适当调整任务运行参数并多次尝试。	L3应用
	P-I-4.1	能够调整 Yarn 队列参数，使其适应业务需求。	L3应用
	P-I-5.1	涉及到组件选型、任务占用资源分类统计，并根据情况调整集群参数。	L3应用
通用能力	G-I-1	自学能力：能够主动通过自学了解软件与信息技术行业信息，掌握大数据行业最新技术，熟悉数据采集处理等岗位相关技能。	L3应用
	G-I-2	沟通能力：具备良好的沟通理解能力，能够明确项目意图，在遇到问题时，能够主动沟通，确保各方共识项目意图。	L3应用
	G-I-3	协作能力：具备专业工作语言能力，正确使用专有名词，规范专业术语与行业用语，能够顺畅地与团队沟通交流。	L3应用
	G-I-4	社会人文能力：能够理解和尊重不同区域的文化与当地的风土人情。	L3应用
	G-I-5	身心承受力：具备强健的体魄与良好的心理素质，适应兼具脑力与体力双重劳动的弹性工作制与加班熬夜情况。	L3应用
社会能力	S-I-1	合作意识：善于与团队协作、沟通解决问题；善于外部交流，有效跨部门推进工作进程。	L3应用
	S-I-2	法律安全意识：充分了解大数据行业法规法则，拥有版权意识，自觉维护行业良好风气。	L3应用
	S-I-3	社会责任感：具有社会责任感，能够传递正确的社会价值，时刻维护国家的荣誉与形象。	L3应用
	S-I-4	职业道德：树立正确的职业观，注意不侵犯他人肖像权、保障他人隐私，爱岗敬业、德技并修。	L3应用
发展能力	D-I-1	创业意识：具备一定的创业意识，培养立足岗位创新创业能力。	L3应用
	D-I-2	前瞻眼光：关注行业发展动态了解行业现状与热门资讯，分析行业未来发展趋势与底层逻辑，洞察行业发展走向，随时调整自我职业规划，顺应时代发展。	L3应用
	D-I-3	自我发展：有关关注行业最新前沿技术的意识，不断进行自我知识更新与学习。	L3应用
	D-I-4	纵深发展：善于把握发展实际工作技能，切忌眼高手低、泛而不专，不断在行业内深耕技能，向纵深发展。	L3应用

典型职业活动名称：	典型职业活动序号：	工作任务及编号	1. 业务需求分析	2. 根据业务需求开发功能模块	3. 上线实施	4. 基于新开发功能进行测试
大数据应用软件开发与测试	J		5. 数据准确性验证			
典型职业活动描述	1. 活动性质：该典型职业活动是大数据行业的进阶职业活动，是大数据开发工程师需要具备的工作技能。 2. 职业活动：熟悉业务流程，了解大数据平台数据开发流程、大数据平台数据处理逻辑和处理方法，针对业务需求进行梳理， 分析合理性和可行性，确定需求解决方案，开发功能模块，开展各类测试，针对功能模块生产的数据进行验证，以确保功能模块正常运行，且结果准确可靠。					
工作要求与考	1. 熟悉业务流程，了解大数据平台数据开发流程、大数据平台数据处理逻辑和处理方法，针对业务需求进行梳理， 分析合理性和可行性，确定需求解决方案。					

核标准	2. 熟悉大数据开发语言和开发框架、工具等，熟悉大数据相关组件和框架的使用方法、数据处理逻辑和数据处理方法，能够根据需求解决方案，进行功能开发实现功能模块。		
	3. 熟悉大数据相关组件和框架的使用方法，能够将开发完毕的功能进行上线实施。		
	4. 熟悉各种测试方法，了解待测试功能模块的具体功能点和逻辑，针对新开发的功能，根据功能模块实现功能点进行各类测试，以确保功能模块正常和准确。		
	5. 熟悉各种数据验证方法、数据交叉验证方案，能够针对功能模块生产的数据进行验证，以确保功能模块正常运行，且结果准确可靠。		
能力类别	编号	能力描述	学习程度
职业能力	P-J-1.1	熟悉业务流程，了解大数据平台数据开发流程、大数据平台数据处理逻辑和数据处理方法，针对业务需求进行梳理，分析合理性和可行性，确定需求解决方案。	L3应用
	P-J-2.1	熟悉大数据开发语言和开发框架、工具等，熟悉大数据相关组件和框架的使用方法、数据处理逻辑和数据处理方法，能够根据需求解决方案，进行功能开发实现功能模块。	L3应用
	P-J-3.1	熟悉大数据相关组件和框架的使用方法，能够将开发完毕的功能进行上线实施。	L3应用
	P-J-4.1	熟悉各种测试方法，了解待测试功能模块的具体功能点和逻辑，针对新开发的功能，根据功能模块实现功能点进行各类测试，以确保功能模块正常和准确。	L3应用
	P-J-5.1	熟悉各种数据验证方法、数据交叉验证方案，能够针对功能模块生产的数据进行验证，以确保功能模块正常运行，且结果准确可靠。	L3应用
通用能力	G-J-1	自学能力：能够主动通过自学了解软件与信息技术行业信息，掌握大数据行业最新技术，熟悉数据采集处理等岗位相关技能。	L3应用
	G-J-2	沟通能力：具备良好的沟通理解能力，能够明确项目意图，在遇到问题时，能够主动沟通，确保各方共识项目意图。	L3应用
	G-J-3	协作能力：具备专业工作语言能力，正确使用专有名词，规范专业术语与行业用语，能够顺畅地与团队沟通交流。	L3应用
	G-J-4	社会人文能力：能够理解和尊重不同区域的文化与当地的风土人情。	L3应用
	G-J-5	身心承受力：具备强健的体魄与良好的心理素质，适应兼具脑力与体力双重劳动的弹性工作制与加班熬夜情况。	L3应用
社会能力	S-J-1	合作意识：善于与团队协作、沟通解决问题；善于外部交流，有效跨部门推进工作进程。	L3应用
	S-J-2	法律安全意识：充分了解大数据行业法规法则，拥有版权意识，自觉维护行业良好风气。	L3应用
	S-J-3	社会责任感：具有社会责任感，能够传递正确的社会价值，时刻维护国家的荣誉与形象。	L3应用
	S-J-4	职业道德：树立正确的职业观，注意不侵犯他人肖像权、保障他人隐私，爱岗敬业、德技并修。	L3应用
发展能力	D-J-1	创业意识：具备一定的创业意识，培养立足岗位创新创业能力。	L3应用
	D-J-2	前瞻眼光：关注行业发展动态了解行业现状与热门资讯，分析行业未来发展趋势与底层逻辑，洞察行业发展走向，随时调整自我职业规划，顺应时代发展。	L3应用
	D-J-3	自我发展：有关关注行业最新前沿技术的意识，不断进行自我知识更新与学习。	L3应用
	D-J-4	纵深发展：善于把握发展实际工作技能，切忌眼高手低、泛而不专，不断在行业内深耕技能，向纵深发展。	L3应用

典型职业活动名称：	典型职业活动序号：	工作任务及编号	1. 数据仓库设计	2. 数据仓库构建	3. 数据仓库管理	
数据仓库建设	K					
典型职业活动描述	1. 活动性质：该典型职业活动是大数据行业的进阶职业活动，是数仓工程师需要具备的工作技能。 2. 职业活动：根据业务需求，进行数据仓库选型，形成数据仓库系统设计方案，设计数据仓库数据模型，构建数据仓库。					
工作要求与考核标准	1. 能够根据业务需求，进行数据仓库选型，形成数据仓库系统设计方案，设计数据仓库数据模型。 2. 能够根据方案，构建数据仓库。 3. 能够进行数据仓库系统平台的日常管理与维护。					
能力类别	编号	能力描述	学习程度			
职业能力	P-K-1.1	根据业务需求，进行数据仓库选型，形成数据仓库系统设计方案。	L3应用			
	P-K-1.2	根据数据特点，设计数据仓库数据模型。	L3应用			
	P-K-2.1	根据数据字典进行 ODS 建模，制定 ETL 策略，从源数据系统抽取数据形成数据仓库 ODS 层。	L3应用			
	P-K-2.2	对 ODS 层数据进行清洗与整合，主题汇总，形成数据仓库 DWD 层和 DWS 层。	L3应用			
	P-K-2.3	使用数据分析手段，进行数据转化，形成 ADS 应用数据层。	L3应用			
	P-K-2.4	基于业务需求，开发 ADS 层数据应用，制作数据报表等，支持后续数据分析、建模等任务。	L3应用			
	P-K-3.1	能够进行数据仓库系统平台的日常管理与维护。	L3应用			
	P-K-3.2	根据业务数据需求，为其他业务团队提供数据支持服务。	L3应用			
通用能力	G-K-1	自学能力：能够主动通过自学了解软件与信息技术行业信息，掌握大数据行业最新技术，熟悉数据采集处理等岗位相关技能。	L3应用			
	G-K-2	沟通能力：具备良好的沟通理解能力，能够明确项目意图，在遇到问题时，能够主动沟通，确保各方共识项目意图。	L3应用			
	G-K-3	协作能力：具备专业工作语言能力，正确使用专有名词，规范专业术语与行业用语，能够顺畅地与团队沟通交流。	L3应用			
	G-K-4	社会人文能力：能够理解和尊重不同区域的文化与当地的风土人情。	L3应用			
	G-K-5	身心承受力：具备强健的体魄与良好的心理素质，适应兼具脑力与体力双重劳动的弹性工作制与加班熬夜情况。	L3应用			
社会能力	S-K-1	合作意识：善于与团队协作、沟通解决问题；善于外部交流，有效跨部门推进工作进程。	L3应用			
	S-K-2	法律安全意识：充分了解大数据行业法规法则，拥有版权意识，自觉维护行业良好风气。	L3应用			
	S-K-3	社会责任感：具有社会责任感，能够传递正确的社会价值，时刻维护国家的荣誉与形象。	L3应用			
	S-K-4	职业道德：树立正确的职业观，注意不侵犯他人肖像权、保障他人隐私，爱岗敬业、德技并修。	L3应用			
发展能力	D-K-1	创业意识：具备一定的创业意识，培养立足岗位创新创业能力。	L3应用			
	D-K-2	前瞻眼光：关注行业发展动态了解行业现状与热门资讯，分析行	L3应用			

		业未来发展趋势与底层逻辑，洞察行业发展走向，随时调整自我职业规划，顺应时代发展。	
	D-K-3	自我发展：有关关注行业最新前沿技术的意识，不断进行自我知识更新与学习。	L3应用
	D-K-4	纵深发展：善于把握发展实际工作技能，切忌眼高手低、泛而不专，不断在行业内深耕技能，向纵深发展。	L3应用

典型职业活动名称：	典型职业活动序号：	工作任务及编号	1. 特征工程	2. 模型构建	3. 模型部署	
数据挖掘	L					
典型职业活动描述	1. 活动性质：该典型职业活动是大数据行业的进阶职业活动，是数据挖掘工程师需要具备的工作技能。 2. 职业活动：用一系列工程化的方式从原始数据中筛选出更好的数据特征，以提升模型的训练效果，包括数据预处理、特征选择、降维等；根据业务场景和数据表现形式选取合适的算法模型，进行训练、测试。					
工作要求与考核标准	1. 能够用一系列工程化的方式从原始数据中筛选出更好的数据特征，以提升模型的训练效果，包括数据预处理、特征选择、降维等。					
	2. 能够根据业务场景和数据表现形式选取合适的算法模型，进行训练、测试。					
	3. 在模型需要部署到云端服务器的情况下，将模型封装成 api 接口，通过调用 api 接口，向云端服务器发送请求，云端收到请求后处理并返回结果的方式进行部署。					
	4. 在嵌入式段部署场景下，主要以 c++等语言将模型封装到 sdk，集成到嵌入式设备，保证成本、功耗、性价比。					
能力类别	编号	能力描述				学习程度
职业能力	P-L-1.1	在数据存在不同值类型的情况下，利用工具包进行离散化，归一化、标准化，保证数据有效、可对比。				L3应用
	P-L-1.2	在数据预处理后，利用 spark, hive 进行特征计算，利用模型、相关性分析等进行特征选择，以保证特征的有效性，在特征维度较多的情况下，利用 pca、lda 等方法对特征进行降维，以保证模型性能最优。				L3应用
	P-L-2.1	数据集划分，在数据质量达标的情况下，利用工具包（scikit-learn）进行数据集和样本划分，以保证模型训练的准确性完成数据集划分。				L3应用
	P-L-2.2	根据业务场景和数据表现形式选取合适的算法模型，进行训练、测试，保证建模的可靠性。				L3应用
	P-L-2.3	选择合适的评估指标，如准确率、召回率、f1、auc、业务核心指标评估模型效果，保证结果的有效性。				L3应用
	P-L-2.4	模型调优，根据评估指标，评估模型效果，进行优化，保证模型的准确性。				L3应用
	P-L-3.1	在模型需要部署到云端服务器的情况下，将模型封装成 api 接口，通过调用 api 接口，向云端服务器发送请求，云端收到请求后处理并返回结果的方式进行部署（服务器部署）				L3应用
	P-L-3.2	sdk 部署，在嵌入式段部署场景下，主要以 c++等语言将模型封装到 sdk，集成到嵌入式设备。保证成本、功耗、性价比				L3应用
通用能力	G-L-1	自学能力：能够主动通过自学了解软件与信息技术行业信息，掌握大数据行业最新技术，熟悉数据采集处理等岗位相关技能。				L3应用

	G-L-2	沟通能力：具备良好的沟通理解能力，能够明确项目意图，在遇到问题时，能够主动沟通，确保各方共识项目意图。	L3应用
	G-L-3	协作能力：具备专业工作语言能力，正确使用专有名词，规范专业术语与行业用语，能够顺畅地与团队沟通交流。	L3应用
	G-L-4	社会人文能力：能够理解和尊重不同区域的文化与当地的风土人情。	L3应用
	G-L-5	身心承受力：具备强健的体魄与良好的心理素质，适应兼具脑力与体力双重劳动的弹性工作制与加班熬夜情况。	L3应用
社会能力	S-L-1	合作意识：善于与团队协作、沟通解决问题；善于外部交流，有效跨部门推进工作进程。	L3应用
	S-L-2	法律安全意识：充分了解大数据行业法规法则，拥有版权意识，自觉维护行业良好风气。	L3应用
	S-L-3	社会责任感：具有社会责任感，能够传递正确的社会价值，时刻维护国家的荣誉与形象。	L3应用
	S-L-4	职业道德：树立正确的职业观，注意不侵犯他人肖像权、保障他人隐私，爱岗敬业、德技并修。	L3应用
发展能力	D-L-1	创业意识：具备一定的创业意识，培养立足岗位创新创业能力。	L3应用
	D-L-2	前瞻眼光：关注行业发展动态了解行业现状与热门资讯，分析行业未来发展趋势与底层逻辑，洞察行业发展走向，随时调整自我职业规划，顺应时代发展。	L3应用
	D-L-3	自我发展：有关关注行业最新前沿技术的意识，不断进行自我知识更新与学习。	L3应用
	D-L-4	纵深发展：善于把握发展实际工作技能，切忌眼高手低、泛而不专，不断在行业内深耕技能，向纵深发展。	L3应用

典型职业活动名称：	典型职业活动序号：	工作任务及编号	1. 数据标准制定与管理	2. 数据资产管理	3. 数据质量管理	4. 数据管理推进
数 据 治 理	M					
典型职业活动描述	1. 活动性质：该典型职业活动是大数据行业的进阶职业活动，是数据治理工程师需要具备的工作技能。 2. 职业活动：建立完善数据治理体系，负责数据标准的制定和管理；制定数据资产管理制度与流程，维护数据资产目录；对数据质量进行监测评价和持续管控，定期形成数据治理分析报告；开展数据认责、数据模型、元数据管理、共享管理、安全管理等数据管理各领域工作，制定工作方案，推动实施。					
工作要求与考核标准	1. 建立完善数据治理体系，负责数据标准的制定和管理，推动数据的标准的落地执行。					
	2. 建立数据治理体系，制定数据资产管理制度与流程，维护数据资产目录。					
	3. 对数据质量进行监测评价和持续管控，定期形成数据治理分析报告。					
	4. 开展数据认责、数据模型、元数据管理、共享管理、安全管理等数据管理各领域工作，制定工作方案，推动实施。					
能力类别	编号	能力描述				学习程度
职业能力	P-M-1.1	建立完善数据治理体系，负责数据标准的制定和管理，推动数据的标准的落地执行。				L3应用
	P-M-2.1	建立数据治理体系，制定数据资产管理制度与流程，维护数据资产目录。				L3应用

	P-M-3.1	对数据质量进行监测评价和持续管控，定期形成数据治理分析报告。	L3应用
	P-M-4.1	开展数据认责、数据模型、元数据管理、共享管理、安全管理等数据管理各领域工作，制定工作方案，推动实施。	L3应用
通用能力	G-M-1	自学能力：能够主动通过自学了解软件与信息技术行业信息，掌握大数据行业最新技术，熟悉数据采集处理等岗位相关技能。	L3应用
	G-M-2	沟通能力：具备良好的沟通理解能力，能够明确项目意图，在遇到问题时，能够主动沟通，确保各方共识项目意图。	L3应用
	G-M-3	协作能力：具备专业工作语言能力，正确使用专有名词，规范专业术语与行业用语，能够顺畅地与团队沟通交流。	L3应用
	G-M-4	社会人文能力：能够理解和尊重不同区域的文化与当地的风土人情。	L3应用
	G-M-5	身心承受力：具备强健的体魄与良好的心理素质，适应兼具脑力与体力双重劳动的弹性工作制与加班熬夜情况。	L3应用
社会能力	S-M-1	合作意识：善于与团队协作、沟通解决问题；善于外部交流，有效跨部门推进工作进程。	L3应用
	S-M-2	法律安全意识：充分了解大数据行业法规规则，拥有版权意识，自觉维护行业良好风气。	L3应用
	S-M-3	社会责任感：具有社会责任感，能够传递正确的社会价值，时刻维护国家的荣誉与形象。	L3应用
	S-M-4	职业道德：树立正确的职业观，注意不侵犯他人肖像权、保障他人隐私，爱岗敬业、德技并修。	L3应用
发展能力	D-M-1	创业意识：具备一定的创业意识，培养立足岗位创新创业能力。	L3应用
	D-M-2	前瞻眼光：关注行业发展动态了解行业现状与热门资讯，分析行业未来发展趋势与底层逻辑，洞察行业发展走向，随时调整自我职业规划，顺应时代发展。	L3应用
	D-M-3	自我发展：有关关注行业最新前沿技术的意识，不断进行自我知识更新与学习。	L3应用
	D-M-4	纵深发展：善于把握发展实际工作技能，切忌眼高手低、泛而不专，不断在行业内深耕技能，向纵深发展。	L3应用

3. 课程转化表

典型工作任务	典型工作任务描述	任务场景分类逻辑	对应专业核心课程及学习情境设置							对应基础理论课程	
			核心课程名称	学习情境名称						是否需要	基础理论课程设置
离线数据采集、实时数据采集	根据项目需求，分析数据源的数据结构，设计离线采集工作规划，并做任务分工，根据不同文档部署搭建采集环境、开发采集系统，完成数据入库等。	1. 数据来源：网页、APP、传感器（接口端） 2. 数据类型：结构化、非结构化 3. 数据用途：分析、管理 4. 工具：八爪鱼、Python	《互联网数据采集》	热点舆情数据采集	农产品营销数据采集	招聘网站数据采集	农作物监控数据采集			是	1. 大数据基础 2. 计算机网络技术应用基础 3. Python 程序设计基础 4. web 前端技术应用
数据清洗与预处理	利用工具包对数据进行填充、删除、缩放等处理，实现数据标准化、归一化等，保证数据有效、可对比。	根据处理数据类型不同分类： 1. 缺失或空值 2. 格式标准化 3. 逻辑错误 4. 数据脱敏 5. 非需求数据去除 6. 重复数据 7. 错误数据	《数据清洗与预处理》	电商商品数据清洗与预处理	电商用户评论数据清洗与预处理	智能制造数据清洗与预处理	学生信息管理数据清洗与预处理	动物图像数据清洗与预处理任务		是	1. 数据库技术应用 2. Python 程序设计基础
大数据分析	了解业务背景，分析数据层面的需求，设计解决方案，使用工具进行统计分析与相关性分析，总结关键性发现和结论，撰写分析报告，给出合理建议。	1. 分析目的 2. 分析方法/工具 3. 分析依据 4. 行业	《大数据分析》	商品销量趋势分析	线上教学每日情况分析	学生画像分析	网易云音乐推荐分析	智能制造数据分析	金融风控数据分析	是	1. Python 程序设计基础 2. Excel 数据分析技术

大数据可视化	分析数据呈现需求，规划相应内容的展示形式，设计功能原型，选取工具，整合实现完整设计。	1. 分析目的 2. 分析方法/工具 3. 分析依据 4. 行业	《数据可视化》	新闻数据可视化	碳中和碳达峰数据可视化	智能制造产业升级数据可视化	学生画像可视化设计	用户征信分析可视化		是	1. Python 程序设计基础 2. Excel 数据分析技术
大数据应用软件开发与测试	熟悉业务流程，了解大数据平台数据开发流程、大数据平台数据处理逻辑和处理方法，针对业务需求进行梳理，分析合理性和可行性，确定需求解决方案，开发功能模块，开展各类测试，针对功能模块生产的数据进行验证，以确保功能模块正常运行，且结果准确可靠。	根据开发的功能不同分类： 1. 记录 2. 统计 3. 分析 4. 服务（接诉即办）（移动端、平台网页） 5. 管理类	《大数据应用软件开发与测试》	学生考勤系统开发测试	学情分析系统开发测试	农事档案系统开发测试				是	1. Python/Java 程序设计基础
数据仓库建设	根据业务需求，进行数据仓库选型，形成数据仓库系统设计方案，设计数据仓库数据模型，构建数据仓库。	1. 数仓类型（离线/实时） 2. 模型架构（分层原则，模型类型等） 3. 数据库选型（传统关系型数据库，大数据平台架构等） 4. 行业	《数据仓库建设》	食品行业数仓建设	体育行业数仓建设	智慧零售数仓建设	电商企业日志数仓建设			是	1. Linux 系统技术应用 2. 数据库技术应用
大数据 ETL	用相应的工具对数据进行抽取，对抽取的数据进行标准化转换，对转换后的数据加载到数据仓库，基于数据库，为业务提供数据服务。	1. 数据源类型（文本，关系数据库，Excel，非关系数据库等） 2. 工具选择（Sqoop，Flume，kettle 等） 3. 编程语言 4. 行业	《大数据 ETL》	医疗数据中台数据汇总	商务机构交易数据汇总	汽车制造行业供应链数据汇总	传统生产系统数字化转型数据仓库数据建设	能源行业数字化转型数据仓库数据建设		是	1. 数据库技术应用

大数据组件安装部署	下载安装大数据各组件，能够更改各组件配置，使其协调工作，开发程序并测试，确保各组件关联关系正常，使其正常运行	1. 大数据离线平台搭建 2. 大数据实时平台搭建 3. 大数据事务分析平台搭建 4. 大数据报表展示平台搭建	《大数据组件安装部署》	大数据离线平台搭建	大数据实时平台搭建	大数据事务分析平台搭建	大数据报表展示平台搭建			是	1. Linux 系统技术应用 2. 计算机网络技术应用基础
大数据集群运维监控	监控基础网络环境，设置监控指标并监控平台运作，做好监控和统计，分析业务系统的日常运行状态，查看告警提示，分析问题及原因，即时调整或反馈。	1. 硬件资源 2. 组件状态 3. 任务结果 4. 监控报告	《大数据集群运维监控》	平台基础硬件资源监控	大数据组件状态监控	平台任务结果准确性监控				是	1. 计算机网络技术应用基础
大数据集群运行问题排查、调优	根据日志报错信息排查问题，通过 Yarn Web UI 查看任务运行情况，适当调整任务运行参数并多次尝试；调整 Yarn 队列参数，使其适应业务需求；根据组件选型、任务占用资源分类统计情况调整集群参数。	1. 组件问题排查 2. 任务问题排查 3. 优先级问题排查 4. 系统平台问题排查	《大数据集群运行问题排查、调优》	大数据组件问题排查	集群任务执行问题排查	集群任务优先级问题排查	集群节点问题排查			是	1. Linux 系统技术应用 2. 计算机网络技术应用基础
数据挖掘	用一系列工程化的方式从原始数据中筛选出更好的数据特征，以提升模型的训练效果，包括数据预处理、特征选择、降维等；根据业务场景和数据表现形式选取合适的算法模型，进行训练、测试。	1. 数据挖掘模型类型（分类、聚类、回归） 2. 数据挖掘应用场景（电商、营销、金融等） 3. 数据类型（表格数据、文本数据、图像数据）	《数据挖掘》	股票市场价格预测	电信客户流失预测	精准营销客户分群	客户信用风险评估	商品评论情感分析	图像分类与识别	是	1. Linux 系统技术应用 2. 计算机网络技术应用基础

4. 课程与职业能力对照表

典型工作任务	对应的课程名称	课程类别	职业能力编号	通用能力编号	社会能力编号	发展能力编号
离线数据采集	互联网数据采集	专业核心课	P-A-1.1	G-A-1	S-A-1	D-A-1
			P-A-2.1	G-A-2	S-A-2	D-A-2
			P-A-2.2	G-A-3	S-A-3	D-A-3
			P-A-2.3	G-A-4	S-A-4	D-A-4
			P-A-3.1	G-A-5		
			P-A-4.1			
			P-A-4.2			
			P-A-4.3			
			P-A-5.1			
			P-A-5.2			
			P-A-5.3			
实时数据采集	互联网数据采集	专业核心课	P-B-1.1	G-B-1	S-B-1	D-B-1
			P-B-2.1	G-B-2	S-B-2	D-B-2
			P-B-2.2	G-B-3	S-B-3	D-B-3
			P-B-2.3	G-B-4	S-B-4	D-B-4
			P-B-3.1	G-B-5		
			P-B-4.1			
			P-B-4.2			
			P-B-4.3			
			P-B-5.1			
			P-B-5.2			
			P-B-5.3			
			P-B-6.1			
			P-B-6.2			
数据清洗与预处理	数据清洗与预处理	专业核心课	P-C-1.1	G-C-1	S-C-1	D-C-1
			P-C-2.1	G-C-2	S-C-2	D-C-2
			P-C-3.1	G-C-3	S-C-3	D-C-3
			P-C-3.2	G-C-4	S-C-4	D-C-4
			P-C-3.3	G-C-5		

			P-C-3.4			
			P-C-3.5			
大数据分析	大数据分析	专业核心课	P-D-1.1	G-D-1	S-D-1	D-D-1
			P-D-1.2	G-D-2	S-D-2	D-D-2
			P-D-2.1	G-D-3	S-D-3	D-D-3
			P-D-2.2	G-D-4	S-D-4	D-D-4
			P-D-3.1	G-D-5		
			P-D-4.1			
			P-D-4.2			
			P-D-5.1			
大数据可视化	大数据可视化	专业核心课	P-E-1.1	G-E-1	S-E-1	D-E-1
			P-E-2.1	G-E-2	S-E-2	D-E-2
			P-E-3.1	G-E-3	S-E-3	D-E-3
			P-E-3.2	G-E-4	S-E-4	D-E-4
			P-E-3.3	G-E-5		
大数据 ETL	大数据 ETL	专业核心课	P-F-1.1	G-F-1	S-F-1	D-F-1
			P-F-1.2	G-F-2	S-F-2	D-F-2
			P-F-2.1	G-F-3	S-F-3	D-F-3
			P-F-3.1	G-F-4	S-F-4	D-F-4
			P-F-4.1	G-F-5		
大数据组件安装部署	大数据组件安装部署	专业核心课	P-G-1.1	G-G-1	S-G-1	D-G-1
			P-G-2.1	G-G-2	S-G-2	D-G-2
			P-G-3.1	G-G-3	S-G-3	D-G-3
				G-G-4	S-G-4	D-G-4
				G-G-5		
大数据集群运维监控	大数据集群运维监控	专业核心课	P-H-1.1	G-H-1	S-H-1	D-H-1
			P-H-1.2	G-H-2	S-H-2	D-H-2
			P-H-1.3	G-H-3	S-H-3	D-H-3
			P-H-2.1	G-H-4	S-H-4	D-H-4
			P-H-3.1	G-H-5		
			P-H-3.2			
			P-H-3.3			
			P-H-4.1			

大数据集群运行问题排查、调优	大数据集群运行问题排查调优	专业核心课	P-I-1.1	G-I-1	S-I-1	D-I-1
			P-I-2.1	G-I-2	S-I-2	D-I-2
			P-I-3.1	G-I-3	S-I-3	D-I-3
			P-I-4.1	G-I-4	S-I-4	D-I-4
			P-I-5.1	G-I-5		
大数据应用软件开发与测试	大数据应用软件开发与测试	专业核心课	P-J-1.1	G-J-1	S-J-1	D-J-1
			P-J-2.1	G-J-2	S-J-2	D-J-2
			P-J-3.1	G-J-3	S-J-3	D-J-3
			P-J-4.1	G-J-4	S-J-4	D-J-4
			P-J-5.1	G-J-5		
数据仓库建设	数据仓库建设	专业核心课	P-K-1.1	G-K-1	S-K-1	D-K-1
			P-K-1.2	G-K-2	S-K-2	D-K-2
			P-K-2.1	G-K-3	S-K-3	D-K-3
			P-K-2.2	G-K-4	S-K-4	D-K-4
			P-K-2.3	G-K-5		
			P-K-2.4			
			P-K-3.1			
			P-K-3.2			
数据挖掘	数据挖掘	专业核心课	P-L-1.1	G-L-1	S-L-1	D-L-1
			P-L-1.2	G-L-2	S-L-2	D-L-2
			P-L-2.1	G-L-3	S-L-3	D-L-3
			P-L-2.2	G-L-4	S-L-4	D-L-4
			P-L-2.3	G-L-5		
			P-L-2.4			
			P-L-3.1			
			P-L-3.2			
数据治理	数据治理	专业选修课	P-M-1.1	G-M-1	S-M-1	D-M-1
			P-M-2.1	G-M-2	S-M-2	D-M-2
			P-M-3.1	G-M-3	S-M-3	D-M-3
			P-M-4.1	G-M-4	S-M-4	D-M-4
				G-M-5		

附件 3：名词术语表

序号	术语名称	别称/俗称	术语解释
1	大数据	Cig dCtC	具有体量巨大、来源多样、生成极快、且多变特征并且难以用传统数据体系结构有效处理的包含大量数据集的数据。
2	数据生存周期	dCtC lifecycle	将原始数据转化为可用于行动的知识的一组过程。
3	大数据参考体系结构	Cig dCtC reference Crchitecture	一种用作工具以便于对大数据内在的要求、设计机构和运行进行开发性探讨的高层概念模型。
4	系统协调者	system orchestrCtor	大数据参考体系结构中的一种逻辑功能构件，它定义所需的数据应用活动并将它们整合到可运行的垂直系统中。
5	数据提供者	dCtC provider	大数据参考体系结构中的一种逻辑功能构件，它将新的数据或信息引入大数据系统。
6	大数据应用提供者	Cig dCtC CpplicCtion provider	大数据参考体系结构中的一种逻辑功能构件，它执行数据生存周期操作，以满足系统协调者定义的需求以及安全和隐私保护需求。
7	大数据框架提供者	Cig dCtC frCmework provider	大数据参考体系结构中的一种逻辑功能构件，它建立一种计算框架，在此框架中执行转换应用，同时保护数据完整性和隐私。
8	数据消费者	dCtC consumer	大数据参考体系结构中的一种逻辑功能构件，它是使用大数据应用提供者提供的应用的末端用户或其他系统。
9	基础设施框架	infrCstructure frCmework	由网络、计算、存储和环境等功能构件构成的一种集合。
10	数据平台框架	DCtC plCtform frCmework	用于指导实现结合相关应用编程接口（CPI）访问的逻辑数据组织和分发的集合。
11	处理框架	processing frCmework	覆盖为支持大数据应用实现所需基础设施软件的、定义数据的计算和处理的集合。

12	消息/通信框架	messCging /communicCtions frCmework	源于高性能计算环境、为水平扩展集群中节点之间的可靠查询、传输和接收数据提供 CPI 的集合。
13	资源管理框架	resource mCnCgement frCmework	大数据框架提供者可能提供的、利用数据本地化作为一种输入变量来确定是否安装新的处理框架元素（如，主节点、处理节点、作业位置），从而实现对 CPU 和存储两大资源高效且有效管理的集合。
14	大数据系统	Cig dCtC system	实现大数据参考体系结构的全部或部分功能的系统。
15	大数据服务	Cig dCtC service	基于大数据参考体系结构提供的数据服务。
16	垂直扩展	verticCl scCling	为提高性能而提高处理速度、存储和内存等系统参数的过程。
17	水平扩展	horizontCl scCling	将集成的一群个体资源作为一个单系统使用的过程。
18	大数据范例	Cig dCtC pCrCdigm	一种由水平耦合分布式数据系统和独立资源组成的、用于实现为有效处理众多数据集所必需的可伸缩的知识。
19	大数据工程化	Cig dCtC engineering	为适应大数据对于有效存储、操作和分析的需求而运用治理独立资源的先进技术构建可伸缩数据系统的过程。
20	大规模并行处理	mCssively pCrCllel processing	多个处理器并行工作以执行一个特定计算任务的过程。
21	分布式文件系统	distriCuted filesystem	多个结构化数据集分布在一个或多个服务器集群的各个计算节点的文件系统。
22	分布式计算	distriCuted computing	一种覆盖存储层和处理层的、用于实现多类型程序设计算法模型的计算模式。
23	分散—聚集	scCtter - gCther	大数据集的处理形式，其中所需的计算被划分并分布在集群的多个节点上，整体结果由每个节点的结果合并而成。
24	流数据	streCming dCtC	经由接口传递，从连续运行的数据源产生的数据。
25	非结构化数据	unstructured dCtC	不具有预定义模型或未以预定义方式组织的数据。

26	大数据生存周期模型	lifecycle model for Cig dCtC	用于描述大数据的“数据-信息-知识-价值”生存周期和指导大数据相关活动的模型；这些活动主要由收集、准备、分析和行动等阶段覆盖。
27	读时模式	schemC-on-reCd	一种数据模式应用；按此应用，在从数据库读取数据之前，先经过诸如转换、净化、整合之类准备步骤。
28	计算可移植性	computCtionCl portCCility	使计算移动到数据所在位置的能力。
29	真实性	verCcility	数据在跨边界传送的情况下，与数据完整性和隐私保护相关的一种数据特征；亦简单指数据的准确性。
30	价值	vClue	从分析学角度考虑的数据对组织的重要性。
31	波动性	volCtility	数据机构随时间变化的趋势。
32	正确性	vClidity	就数据预期用途而言的数据适当性。
33	大数据动态应用	Cig dCtC velocity CpplicCtion	数据的收集、准备和分析（预警）在动态中发生，并可能在数据存储之前进行归纳或聚合。
34	大数据卷系统	Cig dCtC volume system	在数据准备阶段前以数据原始形式存储的一种数据系统。
35	数据仓库	dCtC wCrehouse	在数据准备之后用于永久性存储数据的数据库。
36	动态数据	dCtC in motion	处于活动状态，其典型特征表现为大数据的速度和多变性特征的数据。
37	静态数据	dCtC Ct rest	处于静止状态，其典型特征表现为大数据的体量和多样性特征的数据。
38	非关系模型	non-relCtionCl models	用于数据存储和处理、不遵循关系代数的逻辑数据模型。
39	联合数据库系统	federCted dCtCCCse system	一种元数据库管理系统，它透明的将多个自治数据库系统映射到一个单一联合数据库。

40	数据科学	dCtC science	根据原始数据，经过整个数据生存周期过程凭借经验合成可用于行动的知识的一种科学。
41	数据科学范例	dCtC science pCrCdigm	通过发现、假设和假设测试过程直接从数据萃取的可用于行动的知识。
42	数据科学家	dCtC scientist	数据科学专业人员：他们具有足够的业务需求管理机制方面的知识、领域知识、分析技能以及用于管理数据生存周期中每个阶段的端到端数据过程的软件和系统工程知识。
43	数据治理	dCtC governCnce	对数据进行处置、格式化和规范化的过程。
44	开发数据	open dCtC	可为其他数据使用的数据。
45	链接数据	linked dCtC	链接其他数据的数据。
46	数据集	dCtC set	数据记录汇聚的数据形式。
47	追溯	provenCnce	对数据集的历史元数据的讨论。
48	分析	CnClytics	根据信息合成知识的过程。
49	分析过程特征	CnClytic processes chCrCcterstics	用以表征大数据分析过程的发现、开发和应用。
50	共享磁盘文件系统	shCred-disk file systems	使用单一存储池且与多个计算资源关联的存储数据的方式。
51	数据特征层次	dCtC chCrCcteristic hierCrchy	从不同粗细粒度表征数据特征的数据层次结构。
52	可伸缩流处理	scClCCle streCm processing	数据存储器之间动态数据的处理形式。
53	可伸缩数据存储	scClCCle dCtC stores	可以支持应对数据存储无限增长的一种存储技术。
54	本体	ontology	在大数据语境下，它是一些约束后续各种不同层次逻辑模型的语义模型。

55	分类	tConomies	在数据分析语境下表示关于数据元素关系的元数据。
56	图解模型	grCphicCl model	一种可以呈现数据元素之间关系的大数据记录存储类型。
57	复杂性	complexity	在大数据语境下，复杂性是指各数据元素之间内部关系的或跨数据记录之间的一种互动关系繁简程度。
58	资源协商	resource negotiCtion	一种支持多租户以及要求高可用性和低延迟的环境的资源访问模式。
59	集群管理	cluster mCnCgement	在以非关系模型方式驻留数据的集群资源之间提供通信的一种机制。
60	软件定义存储	softwCre defined storCge	利用软件来确定存储器的动态层次分配的一种存储管理技术。
61	软件定义网络 SDN	softwCre defined network	支持作为大数据关键实现的网络资源高效且有效管理的一种技术。
62	网络功能虚拟化	network function virtuClizCtion	对路由器/路由选择、周界防护、远程访问鉴别以及网络流量/载荷监控等网络功能的虚拟应用现实。
63	本地虚拟化	nCtive virtuClizCtion	大数据环境下的一种虚拟化基本形式，按此种形式，在本地裸机上运行管理程序，该程序管理由操作系统和应用组成的多个虚拟机。
64	主机虚拟化	hosted virtuClizCtion	大数据环境下的一种虚拟化基本形式，按此种形式，在本地裸机上运行操作系统，在驻留客户操作系统和应用的顶层运行管理程序。
65	容器式虚拟化	contCinerized virtuClizCtion	大数据环境下的一种虚拟化基本形式，按此种形式，在本地裸机上运行操作系统，管理程序功能嵌入操作系统中，应用在内部各个容器中运行，控制或限制对操作系统和物理机器资源的访问。
66	数据	dCtC	信息的可再解释的形式化表示，以适用于通信、解释或处理。
67	数据处理	dCtC processing	数据操作的系统执行。
68	数据管理	dCtC mCnCgement	在数据处理系统中，提供对数据的访问，执行或监视数据的存储，以及控制输入输出操作等功能。
69	关系模型	relCtionCl model	结构基于一组关系的数据模型。

70	关系数据库	relCtionCl dCtCCCse	数据按关系模型来组织的数据库。
71	分布式数据处理	distriCuted dCtC processing	将操作分散到计算机网络的各结点进行的数据处理。
72	元数据	metCdCtC	关于数据或数据元素的数据（可能包括其数据描述），以及关于数据拥有权、存储路径、访问权和数据易变性的数据。
73	语义元数据	semCntic metCdCtC	元数据的一个类型；给出有助于恰当理解数据元素的定义性描述的元数据，或作为数据元素的定义的一类元数据。
74	数据挖掘	dCtC mining	从大量数据中通过算法搜索隐藏于其中信息的过程。
75	云计算	cloud computing	一种通过网络将可伸缩、弹性的共享物理和虚拟资源池以按需自服务的方式供应和管理的模式。
76	物联网	Internet of things	通过感知设备，按照约定协议，连接物、人、系统和信息资源，实现对物理和虚拟世界的信息进行处理并作出反应的智能服务系统。
77	数据中心	dCtC center	由计算机场站（机房）、机房基础设施、信息系统硬件（物理和虚拟资源）、信息系统软件，信息资源（数据）和人员以及相应的规章制度组成的组织。
78	结构化数据	structured dCtC	一种数据表示形式，按此种形式，由数据元素汇集而成的每个记录的结构都是一致的并且可以使用关系模型予以有效描述。
79	表结构	tCCle structure	为主体层内容提供表示语义的一种存储范例。
80	键值结构	key-vClue structure	一种存储范例，按此范例存储的记录由值中的关键词和一串数据组成。
81	信息管理系统	informCtion mCnCgement system	企业内部使用的信息系统，如业务管理系统等。主要是通过用户输入和系统的二次加工的方式生成的数据，多为结构化数据。

82	网络信息系统	network information system	基于网络运行的信息系统即网络信息系统是大数据产生的重要方式，如电子商务系统、社交网络等。这类系统多为半结构化或无结构化数据。
83	物联网系统	internet of things system	物联网是在互联网基础上延伸和扩展的网络，其延展到了任何物品之间的通信交换，通过传感技术获取外界的数据信息。
84	科学实验系统	scientific experiment system	主要用于科学技术研究，数据既可以是由真实实验产生，也可以通过模拟方式获取仿真数据。
85	被动式生成数据	passive generation data	采用数据库技术阶段数据的产生是被动的，数据是随着业务系统运行而产生的。
86	主动式生成数据	active generation data	移动互联网的发展使人们可以随时随地通过移动终端生成数据，人们通过各种行为开始主动地生成数据。
87	感知式生成数据	perceptual generation data	感知技术的发展促进了数据生成方式发生了根本性的变化，如遍布城市各个角落的摄像头等数据采集设备源源不断地自动采集、生成数据。
88	半结构化数据	semistructured data	有基本固定格式但数据间没有关系，包括文字处理文件、日志文件及大量保存和发布在网络上的信息（比如 HTML）。
89	非结构化数据	unstructured data	该信息在本质形式上可认为主要是位映射数据。数据必须处于一种可感知的形式中（比如可在音频、视频和多媒体文件中被听到或看到）。
90	Hadoop	Hadoop	Hadoop 是一个能够对大量数据进行分布式处理的软件框架，是一个能够让用户轻松架构和使用的分布式计算平台。用户可以轻松地在 Hadoop 上开发和运行处理海量数据的应用程序。
91	Spark	Spark	Spark 是一个基于内存计算的开源集群计算系统，目的是更快速的进行数据分析，其提供了与 Hadoop 相似的开源集群计算环境，但基于内存和迭代优化的设计，Spark 在某些工作负载上表现更优秀。
92	数据资源化	data resources	大数据作为一种新的资源，具有其他资源所不具备的优点，如数据的再利用、开放性、可扩展性和潜在价值。数据的价值不会随着它的使用而减少，而是可以不断地被处理和利用

93	大数据采集	Cig dCtC collection	在确定用户目标的基础上，针对该范围内所有结构化、半结构化和非结构化的数据的采集
94	商业数据	Cusiness dCtC	商业数据是指来自于企业 ERP 系统、各种 POS 终端及网上支付系统等业务系统的数据，商业数据是现在最主要的数据来源渠道。
95	互联网数据	internet dCtC	互联网数据是指网络空间交互过程中产生的大量数据，包括通信记录及 QQ、微信、微博等社交媒体产生的数据，其数据复杂且难以被利用。
96	物联网数据	IOT dCtC	物联网数据是除了人和服务器之外，在射频识别、物品、设备、传感器等结点产生的大量数据，包括射频识别装置、全球定位设备、传感器设备等产生的数据。
97	数据清洗	dCtC cleCn	数据清洗是在汇聚多个维度、多个来源、多种结构的数据之后，对数据进行抽取、转换和集成加载。在这个过程中，除了更正、修复系统中的一些错误数据之外，更多是对数据进行归并整理，并存储到新的介质中。目的在于删除重复信息、纠正存在的错误，并提供数据一致性。
98	数据集成	dCtC integrCtion	数据集成是将大量不同类型的源数据保存在原地，并将处理过程分布到源数据上进行并行处理，最终将处理结果进行集成。
99	数据变换	dCtC trCnsformCtion	数据变换是将数据转换成适合挖掘的形式，其采用线性或非线性的数学变换方法将多维数据压缩成较少维数的数据，消除它们在时间、空间、属性及精度等特征表现方面的差异（如统计学中的数据标准化）。
100	数据归约	dCtC reduction	数据归约是从数据库或数据仓库中选取并建立使用者感兴趣的数据集合，然后从数据集合中滤掉一些无关、偏差或重复的数据，在尽可能保持数据原貌的前提下，最大限度地精简数据量。
101	大数据分析	Cig dCtC CnClysis	对规模巨大的数据进行分析，通过多个学科技术的融合，实现数据的采集、管理和分析，从而发现新的知识和规律。
102	预测性分析	predictive CnClysis	从大数据中挖掘出有价值的知识和规则，通过科学建模的手段呈现出结果，然后将新的数据带入模型，从而预测未来的情况。
103	可视化分析	visuCl CnClysis	可视化分析能够直观的呈现大数据特点，同时能够非常容易被用户所接受。数据可视化是数据分析工具最基本的要求。
104	大数据挖掘算法	Cig dCtC mining Clgorithm	通过让机器学习算法，按人的指令工作，从而呈现给用户隐藏在数据之中有价值的结果。大数据分析的核心就是数据挖掘算法。

105	语义引擎	semantic engine	语义引擎通过对网络中的资源对象进行语义上的标注，以及对用户的查询表达进行语义处理，使得自然语言具备语义上的逻辑关系，能够在网络环境下进行广泛有效的语义推理，从而更加准确、全面的实现用户的检索。
106	数据质量和数据管理	data quality management	为了满足信息利用的需要，对信息系统的各个信息采集点进行规范，包括建立模式化的操作规程、原始信息的校验、错误信息的反馈、矫正等一系列的过程。高质量的数据和有效的数据管理能保证分析结果的真实和有价值。
107	深度学习	deep learning	深度学习是一种能够模拟出人脑的神经结构的机器学习方式，从而能够让计算机具有人一样的智慧。让机器模仿人脑的机制来分析数据，建立类似人脑的神经网络进行机器学习，从而实现对数据进行有效表达、解释和学习。
108	知识计算	knowledge computing	知识计算是从大数据中首先获得有价值的知识，并对其进行进一步深入的计算和分析的过程。也就是先从大数据中抽取出有价值的知识，并把它构建成可支持查询、分析与计算的知识库。
109	批量数据	batch data	批量数据通常是数据体量巨大，数据量从 TC 级别跃升到 PC 级别，且以静态形式存储的数据。这种批量数据往往是从应用中沉淀下来的数据，如医院长期存储的电子病历。
110	流式数据	streaming data	流式数据是一个无穷的数据序列，序列中的每一个元素来源不同，格式复杂，序列往往包含时序特性。流式数据的处理常见于服务器日志的实时采集，将 PC 级别数据的处理时间缩短到秒级。
111	交互式数据	interactive data	交互式数据是操作人员与计算机以人机对话的方式产生的数据。操作人员以对话的方式向计算机提出请求，计算机便提供相应的数据或信息，引导操作人员逐步操作直到获得最终结果。采用这种方式，存储在系统中的数据文件能及时被处理修改，同时处理结果可以被立刻查看、使用。
112	图数据	figure data	图数据是通过图形表达出来的信息含义。图数据中主要包括图中的点和连接点的边，这些点和边可以构成各种类型的图。但图数据无法使用单台计算机去处理，即便使用多台计算机对图数据进行分割并行处理，也要选择合适的图分割以及图计算模型来满足要求。

附件 4：“5+2”接续学校及专业人才培养基本情况

1. 接续高职院校名称

5+2：北京联合大学

2. 接续专业名称（专业代码）

专业名称：大数据工程技术

专业代码：310205