

# 郑州智能科技职业学院

## 2025 级专业人才培养方案

专业名称： 大数据技术

专业代码： 510205

学 制： 三年制

层 次： 专科

合作企业： 广州小鹏汽车科技有限公司

撰 写 人： 张成林、张承谦、田瑞敏

审 核 人： 付晓炎

制订时间：2025 年 7 月

# 目录

一、专业名称与代码 .....	1
二、入学要求 .....	1
三、修业年限 .....	1
四、职业面向与岗位能力分析 .....	1
五、培养目标与培养规格 .....	4
六、课程设置与要求 .....	5
七、教学进程总体安排 .....	12
八、实施保障 .....	17
九、毕业要求 .....	21

# 大数据技术专业人才培养方案

## 一、专业名称与代码

专业名称：大数据技术

专业代码：510205

## 二、入学基本要求

中等职业学校毕业、普通高级中学毕业或具备同等学力。

## 三、基本修业年限

基本修业年限为 3 年。

## 四、职业面向与岗位能力分析

### (一) 职业面向

表 1：职业面向表

所属专业大类（代码）	电子与信息大类（51）
所属专业类（代码）	计算机类（5102）
对应行业（代码）	互联网和相关服务（64） 软件和信息技术服务业（65）
主要职业类别（代码）	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)
主要岗位（群）及技术领域	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
职业类证书	计算机技术与软件专业技术资格 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) 职业岗位与能力需求分析

表 2：职业岗位与能力需求分析表

工作岗位	关键能力	典型工作任务	职业能力与素质要求
大数据运维工程师	1. 大数据平台（Hadoop、Spark、Kafka 等）部署、配置与集群管理能力。	1. 负责大数据平台的安装部署、集群扩容与配置管理。 2. 监控平台运行状态。	1. 具备对新知识、新技术的快速学习能力，具备持续自我提升与技术创新、岗位适应与转型的能力。

	<p>2. 系统运行状态监控、日志分析与故障排查能力。</p> <p>3. 运维自动化脚本编写与任务调度能力。</p> <p>4. 数据平台安全管理、性能优化与资源调度能力。</p>	<p>态, 进行系统日志分析和故障处理。</p> <p>3. 编写自动化运维脚本, 实现集群巡检、资源监控与告警推送。</p> <p>4. 进行权限配置、安全加固、系统升级与数据备份恢复, 保障平台高可用性。</p>	<p>能力和素质。</p> <p>2. 拥有良好的职业道德、敬业精神与责任意识, 工作严谨, 遵守行业规范与工作纪律。</p> <p>3. 具备良好的人际沟通能力、公共关系处理能力与团队协作精神, 能够胜任跨部门协作与项目配合。</p>
数据采集工程师	<p>1. 多源数据采集与接入能力, 包括网页爬取、API 调用、日志收集等。</p> <p>2. 数据采集脚本开发与调试能力, 掌握数据采集流程与常用框架(如 Scrapy、Requests)。</p> <p>3. 数据清洗、格式转换与初步预处理能力。</p> <p>4. 数据采集系统的运行监控、异常检测与容错处理能力。</p>	<p>1. 负责业务需求对接, 制定数据采集方案, 确定数据源与接口规则。</p> <p>2. 编写采集脚本实现对网页、接口、日志等结构化与非结构化数据的抓取与整理。</p> <p>3. 对接数据库或数据平台, 完成数据入库、传输与调度配置。</p> <p>4. 实现采集日志记录与异常监控, 定期优化采集效率与系统稳定性。</p>	<p>4. 拥有适应高强度、快节奏工作环境的身体素质与心理素质, 具备较强的抗压能力。</p> <p>5. 具备扎实的逻辑思维能力和问题分析与解决能力, 能够独立处理技术问题并持续优化工作流程。</p> <p>6. 能够胜任数据相关工作对信息敏感度和严谨性的要求, 具备数据安全意识和风险防范能力。</p> <p>7. 遵守职业道德与法律法规, 尊重并保护客户数据和企业信息安全, 确保数据处理和系统运维等活动合法合规、符合伦理标准。</p>
数据挖掘工程师	<p>1. 海量数据分析与建模能力, 掌握聚类、分类、回归、关联规则、异常检测等常见算法。</p> <p>2. 特征工程与模型评估能力, 具备算法调参与结果解释能力。</p> <p>3. 熟练使用 Python、SQL、Spark 等挖掘与建模工具。</p> <p>4. 能够结合业务场景进行数据建模与分析, 为决策提供数据支撑。</p>	<p>1. 根据业务需求进行数据探索与分析, 识别可挖掘的问题点。</p> <p>2. 构建特征体系, 选择合适算法模型并进行训练、调优与评估。</p> <p>3. 分析挖掘结果, 撰写数据分析报告并进行可视化展示。</p> <p>4. 跟踪模型上线效果, 根据反馈进行模型优化与迭代。</p>	
大数据分析工程师	<p>1. 大规模数据提取、整理、分析与建模能力。</p> <p>2. 熟练掌握 SQL、Hive、Python、Spark 等分析工具与平台。</p>	<p>1. 对接业务部门需求, 进行数据采集、整理与加工处理。</p> <p>2. 构建分析模型, 挖掘用户行为、业务流程或运营数据规律。</p>	

	<p>3. 指标体系构建与数据可视化能力，能输出高质量分析报告。</p> <p>4. 基于业务场景进行数据洞察与趋势预测，支持运营与决策。</p>	<p>3. 设计并实现数据可视化图表，撰写分析报告。</p> <p>4. 提出优化建议或策略支持，推动业务数据化决策。</p>	
人工智能训练工程师	<p>1. 具备数据采集、清洗、标注、特征分析及挖掘脚本编写能力。</p> <p>2. 具备模型选择、搭建、训练、测试和评估能力。</p> <p>3. 具备深度学习框架安装、模型训练与推理部署能力。</p> <p>4. 具备基于行业典型场景完成人工智能技术综合应用能力。</p>	<p>1. 依据业务需求完成文字、图像、视频、语音等多模态数据采集与高质量标注。</p> <p>2. 使用深度学习框架进行神经网络模型训练、调参与性能评估。</p> <p>3. 对训练结果进行误差分析、模型优化与迭代更新。</p> <p>4. 输出训练报告、模型文件及部署说明文档。</p>	
数据处理工程师	<p>1. 数据采集、清洗、标注以及脚本编写能力。</p> <p>2. 模型选择、训练与评估能力。</p> <p>3. 深度学习框架安装、推理部署能力。</p> <p>4. 基于行业场景解决业务需求的数据综合应用能力。</p>	<p>1. 完成文字、图像、视频、语音等多模态数据采集与标注。</p> <p>2. 使用标注工具对数据进行分类、统计、审核，生成高质量数据集。</p> <p>3. 运用数据分析与可视化工具进行源数据分析并生成图表展示。</p> <p>4. 根据业务需求对数据进行深度信息挖掘，发现数据间关联关系。</p>	

### (三) 岗位相关职业资格(专业技术)证书

表 3：岗位相关职业资格(专业技术)证书表

职业岗位	职业资格证书名称	等级	发证单位	证书要求
大数据运维工程师	信息技术处理员、信息技术支持工程师、信息系统运行管理员、信息系统管理工程师、信息安全工程师	初级/中级	工业和信息化部	完成对应大数据运维课程学习
大数据分析工程师、数据挖掘工程师	大数据分析师、BDA 数据分析师	初级/中级	工业和信息化部、中国信息协会市场研究业分会	完成对应数据分析课程学习

人工智能训练工程师	人工智能训练师、人工智能深度学习工程应用职业技能等级证书、数据标注职业技能等级证书	初级/中级	人力资源和社会保障部、北京百度网讯科技有限公司、工创集团有限公司	完成对应课程学习
数据处理工程师	人工智能工程技术人员、人工智能数据处理职业技能等级证书、数据应用开发与服务（Python）职业技能等级证书	初级/中级	人力资源和社会保障部、科大讯飞股份有限公司、北京中软国际信息技术有限公司	完成对应数据处理课程学习

## 五、培养目标与培养规格

### （一）培养目标

本专业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，面向互联网和相关服务、软件和信息技术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（群）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）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）掌握支撑本专业学习和可持续发展必备的语文、数学、外语（英语等）、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；

（4）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；

- (5) 掌握数据库基本原理、程序设计、linux 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；
- (6) 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；
- (7) 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；
- (8) 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；
- (9) 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、大数据技术服务等实践能力；
- (10) 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；
- (11) 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；
- (12) 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；
- (13) 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；
- (14) 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。

## 六、课程设置与要求

课程设置包括公共必修课程、公共选修课、专业基础课程、专业核心课、专业拓展课和专业实践课。

### (一) 公共必修课程

根据党和国家有关文件规定和学校特色，将思想道德与法治、毛泽东思想与中国特色社会主义理论体系概论、习近平新时代中国特色社会主义思想概论、形势与政策、军事理论、军事技能训练、心理健康教育、创新创业教育、信息技术基础、大学英语、大学体育、职业发展与就业指导、中华优秀传统文化、国家安全教育、安全教育、劳动教

育等课程列为公共必修课程，将党史国史、中华民族共同体概论、数学等课程列为必修课程或限定性选修课程。

## (二) 公共选修课程

按照上级教育行政部门要求，结合学校特色、学生全面素质教育和个性发展，将口才艺术、中国书法、音乐欣赏、影视鉴赏、信息检索、数学建模、诗文与修养、交响乐欣赏、瑜伽、社交礼仪、大学生疾病与健康等课程列为公共选修课程。

## (三) 专业基础课程

专业基础课程按照专业群进行规划组合，包括相近相关专业互通的专业基础知识课程，它与公共基础课程一起为学生构筑学习专业知识而必须掌握的基础知识和技能。建设完善、规范、科学的知识体系，为学生拓宽专业口径和专业学习奠定宽厚的基础，详见表 4。

表 4：大数据技术专业基础课程一览表

序号	课程名称	主要教学内容及要求	学时数
1	Python 程序设计	<p>1. 主要教学内容：Python 编程环境安装与配置、Python 语言基础—值、变量、数据类型、表达式和运算，程序流程控制—if 语句、循环语句，函数定义与调用、面向对象编程基础、文件 I/O 操作和异常处理等。</p> <p>2. 要求：了解 Python 编程环境的安装与配置能力，掌握 Python 语言的基础知识语法和程序设计方法，培养学生的逻辑思维和编程能力，树立规范编程和创新思维的职业素养。</p>	64
2	Python 应用开发	<p>1. 主要教学内容：Python 程序开发语言基本语法知识、Python 开发应用系统相关技术、Python 第三方库的使用方法、Python 桌面应用系统架构设计、Python 桌面系统测试与部署等。</p> <p>2. 要求：具备 Python 语言程序设计基础能力，具备 Python 桌面应用系统和 Web 应用系统开发能力，树立规范编码和系统设计的职业素养。</p>	64
3	网页设计与制作	<p>1. 主要教学内容：网页开发工具的使用，基础标签、常用样式表使用、盒子模型、浮动、定位、动画。</p> <p>2. 要求：掌握 HTML 基础结构搭建、CSS 样式控制，能独立完成静态网页开发并符合规范、增强网络空间法治观念培养。</p>	48
4	人工智能基础	<p>1. 主要教学内容：人工智能三次浪潮、人工智能的内涵和外延、人工智能产业链、人工智能的基本概念、人工智能的主要技术与工具、常见 AI 工具的基本使用方法、人工智能技术在交通、电商、建筑、工业、农业、医疗等行业的应用。</p> <p>2. 要求：培养使用 AI 解决专业问题的意识，明晰 AI 应用的伦理边界与社会责任。</p>	32
5	计算机操作系统	<p>1. 主要教学内容：Linux 操作系统的安装、使用及配置、命令行操作，掌握文件管理、用户权限、进程控制等核心命令、</p>	48

		<p>文本编辑工具使用,如 vim 的基本操作与高级技巧、系统管理与维护、网络及防火墙配置、常见服务搭建、Shell 脚本基础。</p> <p>2. 要求:理解 Linux 系统的架构及特性、掌握核心服务的工作原理、能独立完成 Linux 系统的安装、初始化配置及日常维护、培养严谨的工作态度,养成规范的操作习惯。</p>	
6	数据库技术	<p>1. 主要教学内容:数据库的设计、使用、检索、管理、数据库基本概念及操作、SQL 实现增、删、改、查的操作、数据查询与统计、索引、存储过程、数据库权限管理。</p> <p>2. 要求:掌握 MySQL 安装配置、数据库与表的创建及管理、熟练运用 SQL 语句,包括查询、插入、更新、删除及复杂查询、理解数据库约束、索引、视图及存储过程的作用与使用方法、培养规范的数据库操作习惯和数据安全意识。</p>	48
7	计算机网络技术	<p>1. 主要教学内容:计算机网络的基础知识、基本原理,以太网技术、基本网络规划、交换机的配置、路由器的基本配置等;操作系统的管理维护、网络服务器的搭建配置和管理。</p> <p>2. 要求:掌握计算机网络的基本概念、协议原理及网络设备的工作机制、了解网络安全的基本原理,熟悉常见安全防护技术的应用场景、养成规范的网络操作习惯,确保网络配置的准确性与安全性。</p>	48

#### (四) 专业核心课程

专业核心课程要提供与学生未来从事的职业、适应未来工作环境所必需的专业知识、技能和相关职业考证所需等课程,根据职业岗位要求和人才成长规律及国家专业教学标准设置专业核心课程,详见表 5。

表 5: 大数据技术专业核心课程一览表

序号	课程名称	主要教学内容及要求	学时数
1	数据采集技术	<p>1. 主要教学内容:数据采集基础知识、数据采集流程与方法、数据采集相关法律法规、网页数据采集与解析、数据库与日志数据采集、采集工具安装与使用、数据采集程序设计与调试、采集数据的初步清洗与存储。</p> <p>2. 要求:掌握数据采集的基本原理与流程,熟悉常见的网页解析与爬取方法,能编写并调试简单的数据采集程序;掌握数据库及日志数据的采集方法;能独立搭建采集环境并配置采集工具;了解数据采集合规要求,具备数据安全与隐私保护意识。</p>	64
2	数据预处理技术	<p>1. 主要教学内容:数据 ETL 基础知识、常用 ETL 工具的安装与配置、数据清洗与质量提升方法、缺失值与异常值处理、重复与不一致数据识别、文本数据与结构化数据的抽取与加载、不同数据源的数据迁移与装载、数据格式转换与多源数据整合优化。</p> <p>2. 要求:掌握数据 ETL 的基本概念与流程,熟悉</p>	48

		常用 ETL 工具的使用方法；能对原始数据进行缺失值、异常值和重复值处理；能实现不同数据源间的数据抽取、转换与加载；具备多源数据整合与优化的能力；培养严谨的数据质量意识与规范化处理习惯。	
3	大数据分析技术应用	<p>1. 主要教学内容：数据分析与计算基础知识、常用数据分析工具的安装与使用、数据结构封装与操作、数据聚合与分组计算、时间序列分析方法、批量与实时数据计算技术、大数据分析平台的应用实践、数据分析结果展示与报告撰写。</p> <p>2. 要求：掌握大数据分析的基本原理与计算方法，熟练使用常见数据分析工具；能独立完成数据聚合、分组及时间序列分析；掌握批量与实时计算任务的实现方法；能运用大数据分析平台开展数据处理与结果展示；具备数据分析报告撰写与结果表达能力。</p>	48
4	数据挖掘应用	<p>1. 主要教学内容：数据特征管理基础知识、监督学习、无监督学习与半监督学习方法、常见数据挖掘算法（回归、分类、关联、聚类）的原理与应用、训练集与测试集划分方法、机器学习典型模型的构建与应用、模型性能评估与结果分析。</p> <p>2. 要求：掌握数据挖掘的基本原理与流程，熟悉常见的机器学习算法及其应用场景；能完成数据集划分与模型训练、验证；了解典型数据挖掘模型的原理与实现过程；掌握模型性能的基本评估方法；具备运用数据挖掘技术解决实际问题的能力。</p>	64
5	大数据平台部署与运维	<p>1. 主要教学内容：Hadoop 平台架构与设计思想、Hadoop 集群的部署与配置、HDFS 文件系统的管理与维护、MapReduce 编程模型与任务实现、HBase 数据库的基本操作与管理、大数据平台运行监控与性能优化。</p> <p>2. 要求：掌握 Hadoop 平台的基本原理与部署流程，熟悉 HDFS 的管理与维护方法；能编写并运行简单的 MapReduce 程序；掌握 HBase 的安装配置与基本使用；具备独立搭建与运维大数据平台的能力，具备系统监控与性能优化意识。</p>	48
6	数据可视化技术与应用	<p>1. 主要教学内容：数据可视化的基本概念、目标与流程、常见可视化图表类型及应用场景、文本与网络可视化方法、主流数据可视化工具的使用、可视化设计原则与实现方法、可视化组件库的开发与应用、可视化结果展示与分析报告编写。</p> <p>2. 要求：掌握数据可视化的基本原理与设计流程，熟悉常见图表类型及可视化工具的使用方法；能根据数据特征选择合适的可视化形式；能使用组件库进行可视化开发与展示；具备撰写数据可视化分析报告和表达分析结论的能力。</p>	48

## (五) 专业拓展课程

专业拓展课程是根据学生发展需求横向拓展和纵向深化的课程，是提升综合职业能力的延展课程，详见表 6。

表 6：大数据技术专业拓展课程一览表

序号	课程名称	主要教学内容及要求	学时数
1	云计算平台技术应用	<p>1. 主要教学内容：云计算的基本概念与核心技术、虚拟化技术（KVM、Docker）的原理与应用、云存储架构与实现方法、OpenStack 平台的部署与管理、公有云与私有云的运维管理、云平台资源配置与性能优化、云环境的安全与监控。</p> <p>2. 要求：掌握云计算平台的基本原理与核心技术，熟悉 KVM、Docker 等虚拟化工具的使用方法；能完成 OpenStack 等云平台的安装部署与组件配置；具备云资源管理与运维能力；了解公有云与私有云的运行机制，能独立完成云环境部署与维护任务。</p>	48
2	NoSQL 数据库技术应用	<p>1. 主要教学内容：NoSQL 数据库的基本概念、特点与应用场景，HBase 的原理与操作方法，Redis 的安装配置与应用，MongoDB 的数据存储与查询操作，Kafka 的消息队列机制与使用方法，NoSQL 数据库的综合应用与性能优化。</p> <p>2. 要求：理解 NoSQL 数据库的特点与适用范围，掌握 HBase、Redis、MongoDB、Kafka 等主流 NoSQL 系统的基本操作与使用方法；能完成数据存储、查询及消息传递等任务；具备在大数据环境中选择并应用合适 NoSQL 数据库的能力。</p>	48
3	Web 前端开发框架技术	<p>1. 主要教学内容：主流前端框架（Vue、React、Angular）的核心概念与应用，组件化开发与状态管理、路由配置与页面导航、Webpack 与 Vite 等前端构建工具的使用、Git 版本控制与项目协作、前后端数据交互实现、单页面应用（SPA）的开发与部署。</p> <p>2. 要求：掌握主流前端框架的基本原理与使用方法，熟悉组件化开发、状态管理及路由配置；能使用 Webpack 或 Vite 完成项目构建；能运用 Git 进行版本控制与团队协作；具备基于框架开发单页面应用并实现前后端数据交互的能力。</p>	48
4	Scala 程序设计	<p>1. 主要教学内容：Scala 语言的核心语法与特性、面向对象与函数式编程范式、集合操作与高阶函数应用、并发编程的 Actor 模型与 Future/Promise 异步处理机制、Play 与 Akka 框架的基本使用方法、基于 Scala 的分布式应用开发与实践。</p> <p>2. 要求：掌握 Scala 语言的语法结构与编程范式，熟悉面向对象与函数式编程的结合使用；能运用集合与并发编程技术实现复杂逻辑；掌握 Actor 模型与异步编程机制；具备使用 Play 或 Akka 框架开发基础分布式应用的能力。</p>	48
5	Spark 应用技术	<p>1. 主要教学内容：Spark 的设计思想与运行原理、Spark 环境的搭建与配置、RDD 的编程模型与操作方法、Spark SQL 的数据查询与分析、Spark Streaming 的实时数据处理、Spa</p>	48

		<p>rk MLlib 的机器学习应用、Spark 在大数据处理中的综合实践。</p> <p>2. 要求：掌握 Spark 平台的基本原理与运行机制，熟悉 Spark 环境的部署与使用；能使用 RDD 进行数据处理与编程操作；掌握 Spark SQL、Spark Streaming 及 Spark MLlib 的基本应用方法；具备利用 Spark 框架实现批处理与实时计算任务的能力。</p>	
6	大数据技术服务	<p>1. 主要教学内容：大数据技术服务的基本概念与应用场景、大数据平台的部署与运维方法、数据采集与预处理流程、大数据分析与可视化应用、数据仓库技术与数据管理方法、数据安全与合规性要求、Hadoop 与 Spark 等工具在数据服务中的综合应用。</p> <p>2. 要求：掌握大数据平台运维与服务的基本方法，熟悉数据采集、分析与可视化的整体流程；能运用 Hadoop、Spark 等工具完成数据处理与分析任务；了解数据仓库构建与数据安全管理要求；具备提供大数据技术支持与服务的综合能力。</p>	48
7	生成式人工智能应用	<p>1. 主要教学内容：生成式人工智能的基本原理与发展概况、AIGC 多模态模型的结构与应用、提示工程设计方法、文本、图像、音频与视频的生成技术、生成式 AI 在数据分析、内容创作、办公与学习场景中的应用、主流生成式 AI 工具的使用实践。</p> <p>2. 要求：理解生成式人工智能与 AIGC 多模态模型的基本原理，掌握提示工程的设计与应用方法；能熟练使用主流生成式 AI 工具进行文案、图像、视频等内容创作；具备利用生成式 AI 开展数据可视化、办公文档与多媒体自动生成的能力；培养创新意识与负责任使用人工智能技术的职业素养。</p>	48

## (六) 专业实践课

专业实践课包括认知实习、岗位实习、专业实训等课程，详见表 7。

表 7：大数据技术专业实践课程一览表

序号	课程名称	主要教学内容及要求	学时数
1	数据采集实训	<p>1. 主要教学内容：外汇、名言、电影、图书、商城、景区等网站的数据采集项目实践，Python 网络访问与数据获取技术、BeautifulSoup 网页解析与数据提取、深度优先与广度优先爬取策略、多线程网页爬取方法、Scrapy 分布式爬取框架的应用、AWS DynamoDB 的 NoSQL 数据存储与管理。</p> <p>2. 要求：熟练掌握 Python 网页爬取与数据解析技术，能独立完成多类型网站的数据采集任务；掌握深度优先、广度优先及多线程爬取策略；能使用 Scrapy 框架实现分布式采集；了解云端 NoSQL 存储的基本应用；具备从采集到存储的完整数据获取与处理能力。</p>	32

2	大数据平台实训	<p>1. 主要教学内容：大数据平台部署与运维的基本技术与方法，Hadoop 集群的安装与配置、Spark 运行环境的搭建、Kafka 消息队列的部署与应用、集群监控与性能优化、故障诊断与恢复、Linux 系统管理与 Shell 脚本编写、集群资源管理与服务高可用技术。</p> <p>2. 要求：掌握大数据平台的部署流程与运维方法，熟悉 Hadoop、Spark、Kafka 等核心组件的安装与配置；能独立完成分布式存储与计算环境的搭建；掌握 Linux 系统操作与 Shell 脚本编写；具备集群监控、性能调优及故障处理能力，能保障平台稳定高效运行。</p>	32
3	数据预处理实训	<p>1. 主要教学内容：数据预处理的核心方法与实操训练，包括销售数据的缺失值填补与异常值处理、客户数据的标准化与归一化、多源数据的集成与去重、SQL 在数据清洗与转换中的应用、数据规约与可视化分析。通过典型案例项目，综合训练数据清洗、集成、转换与规约能力。</p> <p>2. 要求：掌握常见数据质量问题的识别与处理方法，能熟练运用 SQL 及相关工具进行多源数据整合与转换；具备数据标准化、归一化与规约处理能力；能独立完成数据清洗与初步分析任务；培养规范、高效的结构化数据预处理实战技能。</p>	32
4	大数据分析实训	<p>1. 主要教学内容：大数据分析的核心技术与实操训练，包括电商用户行为数据的清洗与特征工程、金融交易数据的存储与查询、基于 Spark MLlib 的客户流失预测建模、分布式数据处理与机器学习算法应用。通过典型行业案例项目，综合训练数据预处理、分布式计算与模型构建能力。</p> <p>2. 要求：掌握大数据分析的基本流程与方法，熟练使用 Spark 及 MLlib 进行数据建模与分析；能完成从数据清洗、特征构造到模型评估的完整分析过程；具备分布式计算与机器学习建模的综合应用能力；培养解决实际业务问题的企业级大数据分析实战技能。</p>	32

5	数据可视化实训	<p>1. 主要教学内容：数据可视化的核心技术与实操训练，包括使用 Tableau 制作零售销售数据动态看板、利用 Power BI 实现金融指标的多维度分析、运用 Python 与 Matplotlib 绘制科研数据图表、基于 Echarts 开发交互式 Web 可视化大屏。通过多场景案例项目，综合训练数据展示与可视化设计能力。</p> <p>2. 要求：掌握主流数据可视化工具的使用方法，能独立完成可视化看板、报表与交互式图表的设计与实现；熟悉 Tableau、Power BI、Matplotlib 与 Echarts 等工具的应用场景；具备从数据分析结果到可视化呈现的完整实现能力；培养数据表达、视觉设计与分析沟通的综合技能。</p>	32
6	数据挖掘实训	<p>1. 主要教学内容：数据挖掘的核心技术与实操训练，包括使用 Python 与 Scikit-learn 实现零售商品关联规则挖掘、利用 Weka 完成客户分群聚类分析、基于 SPSS Modeler 构建电商用户画像。通过典型商业案例项目，综合训练数据准备、模式发现与模型构建能力，强化数据驱动的决策支持实践。</p> <p>2. 要求：掌握常用数据挖掘算法与工具的使用方法，能独立完成关联分析、聚类分析与画像建模任务；熟悉 Scikit-learn、Weka、SPSS Modeler 等平台的基本操作；具备从数据准备到模型评估的完整实现能力；培养面向商业智能与决策支持的数据挖掘实战技能。</p>	32

## 七、教学进程总体安排

### (一) 课程学时、学分结构表

表 8：学时学分结构表

课程性质	公共必修课	公共选修课	专业基础课	专业核心课	专业拓展课	专业实践课	合计
学时数	872	64	352	320	240	880	2728
学分数	45	4	22	20	15	34	140
占总学时比例	31.96%	2.35%	12.90%	11.73%	8.80%	32.26%	100%

注：本专业总学分 140 学分，总学时 2728 学时，其中理论课 920 学时，占比 33.72%；实践性教学 1808 学时，占比 66.28%；选修课 304 学时，占比 11.14%。

## (二) 课程设置及学时安排

表 9: 大数据技术专业教学计划进程表

课程性质	课程编码	课程名称	学分	总学时	学时分配		学期课程安排						考核方式		备注	
					理论	实践	第一学年		第二学年		第三学年		考试	考查		
							1	2	3	4	5	6				
公共必修课	0120011001	思想道德与法治	3	48	32	16	3							√		
	0120011002	毛泽东思想和中国特色社会主义理论体系概论	2	32	32	0		2						√		
	0120011003	习近平新时代中国特色社会主义思想概论	3	48	32	16		3						√		
	0120011004	形势与政策(1)	0.25	8	8	0	1							√		
	0120011005	形势与政策(2)	0.25	8	8	0		1						√		
	0120011006	形势与政策(3)	0.25	8	8	0			1					√		
	0120011007	形势与政策(4)	0.25	8	8	0				1				√		
	0121011004	中华优秀传统文化	2	32	32	0		2						√		
	0121011005	大学英语(1)	3	48	32	16	3							√		
	0121011006	大学英语(2)	3	48	32	16		3						√		
	0101011002	信息技术基础	3	48	16	32	3							√		
	0121011008	就业指导	1	16	8	8				1				√		
	0121011009	大学生职业生涯规划	1	16	8	8	1							√		
	0121011010	创新创业教育	1	16	0	16			1					√		

	0121011013	大学体育 (1)	2	36	4	32	2								√	
	0121011014	大学体育 (2)	2	36	4	32		2							√	
	0121011015	大学体育 (3)	2	36	4	32			2						√	
	0121011016	大学体育 (4)	2	36	4	32				2					√	
	0121011002	军事技能	3	168	0	168	3周								√	军训三周
	0121011001	军事理论	2	32	32	0	2								√	
	0122011001	心理健康教育	2	32	16	16	2								√	
	0121011003	国家安全教育	1	16	16	0	1								√	
	0121011019	劳动教育 (1)	1	16	0	16	1								√	
	0121011020	劳动教育 (2)	1	16	0	16		1							√	
	0121011011	高等数学 (1)	2	32	32	0	2								√	
	0121011012	高等数学 (2)	2	32	32	0		2							√	
	0121011007	安全教育														贯穿全过程
	公共必修课小计		45	872	400	472	21	16	4	4	0	0				
公共选修课		公共选修课	4	64												
	公共选修课小计		4	64												公共选修课由教务科研处统一安排至前四个学期修读完成，其中艺术类课程至少修读 2 学分。
专	0101013021	网页设计与制作	3	48	24	24	3								√	

业 基 础 课	0101013022	Python 程序设计	4	64	32	32	4						√		
	0101013023	Python 应用开发	4	64	32	32		4					√		
	0101013024	人工智能基础	2	32	16	16		2					√		
	0101013025	计算机操作系统	3	48	24	24		3					√		
	0101013026	数据库技术	3	48	24	24			3				√		
	0101013027	计算机网络技术	3	48	24	24		3					√		
	专业基础课小计		22	352	176	176	7	12	3	0	0	0			
专 业 核 心 课	0101014021	数据采集技术	4	64	32	32			4				√		
	0101014022	大数据平台部署与运维	3	48	24	24			3				√		
	0101014023	大数据分析技术应用	3	48	24	24				3			√		
	0101014024	数据挖掘应用	4	64	32	32				4			√		
	0101014025	数据预处理技术	3	48	24	24			3				√		
	0101014026	数据可视化技术与应用	3	48	24	24				3			√		
	专业核心课小计		20	320	160	160	0	0	10	10	0	0			
专 业 拓	0101015021	云计算平台技术应用	3	48	24	24			3				√		
	0101015022	Scala 程序设计	3	48	24	24			3				√		
	0101015023	Web 前端开发框架技术	3	48	24	24			3				√		

展 课	0101015024	NoSQL 数据库技术应用	3	48	24	24				3				√	
	0101015025	Spark 应用技术	3	48	24	24				3				√	
	0101015026	大数据技术服务	3	48	24	24				3				√	
	0101015027	生成式人工智能应用	3	48	24	24				3				√	
	专业拓展课小计		15	240	120	120	0	0	6	9	0	0		最低选修要求	
专 业 实 践 课	0101017021	数据采集实训	2	32	0	32					2			√	
	0101017022	大数据平台实训	2	32	0	32					2			√	
	0101017023	数据预处理实训	2	32	0	32					2			√	
	0101017024	大数据分析实训	2	32	0	32					2			√	
	0101017025	数据可视化实训	2	32	0	32					2			√	
	0101017027	岗位实习	24	720	0	720						30		√	第 5、6 学期完成 6 个月岗位实习
	专业实践课小计		34	880	0	880	0	0	0	0	10	30			
合计			140	2728	920	1808	28	28	23	23	10	30			

## **八、实施保障**

主要包括师资队伍、教学设施、教学资源、教学方法、学习评价、质量管理等方面。

### **(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**

本专业共有专职教师 12 名，兼职教师 4 名，生师比不高于 22 : 1，“双师型”教师占专业教师数比例为 60%，中、高级职称专任教师的比例为 42%，45 岁以下青年教师占比 88%，专任教师队伍在职称、年龄方面组成合理，既有在教育行业工作三四十年的老教师，也有刚走出校园的年轻教师。我院充分利用校内外优质人才资源，选聘了大数据行业知名企事业单位技术人员担任行业导师，组建校企合作、专兼结合的教师团队，建立定期开展专业教研机制。

#### **2. 专业带头人**

学校专业带头人具有副高级职称，从事高等教育与科研工作近 20 年，长期承担大数据技术及相关课程的教学任务，积累了扎实的教学经验和较强的课程建设能力。其同时具备丰富的行业实践背景，先后主持或参与过与小鹏汽车、华为、海康威视等国内知名大数据与人工智能企业的产学研合作项目，熟悉大数据平台建设、数据挖掘与分析、智能应用开发等前沿技术，能够将实际工程案例引入课堂教学。在专业建设方面，该教师具备较强的学术研究与成果转化能力。其长期关注国内外软件和信息技术服务业、互联网产业及大数据产业的发展趋势，能够敏锐把握行业动态和人才需求变化，及时将行业前沿融入课程体系。此外，该教师在校企合作、实训基地建设、课程体系优化等方面经验丰富，善于组织行业专家、企业导师与教师团队开展联合培养，推动了“课程—实训—就业”一体化人才培养模式的落实。在专业改革与发展中发挥了核心引领作用，为培养适应行业需求的高素质技术技能人才提供了有力保障。

#### **3. 专任教师**

本专业专职教师 12 人，全部取得高校教师资格证书，学历背景涵盖数据科学与大数据技术、大数据工程技术、计算机科学与技术等相关专业，硕士及以上学位占比 60%。教师团队整体教学经验丰富，具有较强的专业理论功底和实践能力，能够全面支撑本专业课程体系建设与人才培养需求。教师团队中既有长期从事大数据教学的“双师型”教师，也有具有企业大数据开发、运维、数据分析等实践经验的教师，形成了“教学+科研+实践”相结合的队伍结构。教师们积极参与省市级教育教学改革与科研项目，承担教材编写与课程资源建设任务，能够将行业前沿成果及时引入课堂。专业教师每年至少

1个月在企业或实训基地锻炼，每5年累计不少于6个月的企业实践经历。

#### 4. 兼职教师

本专业兼职教师4人，来源广泛、结构合理，既包括具有副高级职称的高校教师，也有在大数据与人工智能领域具有影响力的校外知名教师，同时还引入了来自知名互联网企业和大数据企业的一线技术骨干。多数具有中级及以上专业技术职务（职称），他们熟悉行业现状与前沿趋势，具备较强的工程实践与项目管理能力，确保兼职教师能够在教学、科研和校企合作中发挥实质性作用。通过兼职教师的引入，进一步加强了专业建设与企业实践的紧密结合，促进了产教融合，提升了学生的职业技能和就业竞争力。

## （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、实训室和实习实训基地。

### 1. 专业教室基本要求

现有14间机房，具备利用信息化手段开展混合式教学的条件，教室均配备黑（白）板、智慧黑板、多媒体计算机、投影设备等，仪器设备总值700余万元，并具有网络安全防护措施。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。

### 2. 校内外实习实训基地基本条件

学校具有稳定的校内外实习基地，与广州小鹏汽车科技有限公司、杭州海康威视数字技术股份有限公司、东软教育科技集团有限公司、软通动力信息技术（集团）股份有限公司、无锡鑫巨宏智能科技股份有限公司等开展校外实习合作，为学生提供大数据平台部署与运维、数据采集、数据预处理、大数据分析、数据挖掘、数据可视化等相关实习实践。校企双方建有稳定的双师型、专兼职结合的师资队伍，对学生实习进行指导和管理。详情见表10—表11。

表 10：校内实训室一览表

序号	名称	主要设备	实训内容
1	大数据技术综合实训室	计算机（或云桌面）、服务器、交换机、无线AP、网络机柜、多媒体中控台、投影仪、无线投屏器、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、Java项目开发软件、数据库开发软件、Python项目开发软件、项目管理软件。	用于网页设计与开发、Linux操作系统、数据库技术、Python程序设计、云计算技术基础、大数据应用开发项目实践等实训教学。

2	大数据平台实训室	配备计算机(或云桌面)、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、无线投屏器、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、数据可视化技术与应用等实训教学。
3	数据采集与分析实训室	配备计算机、服务器、交换机、网络机柜、多媒体中控台、投影仪、无线投屏器、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。
4	大数据可视化实训室	配备计算机、实训系统服务器、交换机、网络机柜、多媒体中控台、投影仪、无线投屏器、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 11：校外实习（实训）基地一览表

序号	实习(实训)基地名称	合作企业	实训内容
1	广州小鹏汽车实习基地	广州小鹏汽车科技有限公司	数据分析实训
2	杭州海康威视实习基地	杭州海康威视数字技术股份有限公司	数据采集实训
3	大连东软教育实习基地	东软教育科技集团有限公司	大数据平台实训
4	北京软通动力实习基地	软通动力信息技术(集团)股份有限公司	数据预处理实训
5	无锡鑫巨宏实习基地	无锡鑫巨宏智能科技股份有限公司	数据可视化实训

### **(三) 教学资源**

#### **1. 教材选用基本要求**

按照国家规定，经过规范程序选用教材，优先选用国家规划教材和国家优秀教材。专业课程教材应体现本行业新技术、新规范、新标准、新形态，并通过数字教材、活页式教材等多种方式进行动态更新。

#### **2. 书文献配备情况**

学校图书馆纸质藏书约 36 万册，电子图书约 20 万册，纸质期刊近 14 种，电子期刊 0.65 万种。图书文献配备丰富，为本专业师生提供了充足的文本信息、数据资料等知识服务，基本能满足人才培养、专业建设、教科研等工作需要。专业类图书文献主要包括：大数据行业政策法规资料，有关大数据岗位的技术、标准、方法、操作规范以及实训案例类图书等。及时配置新经济、新技术、新工艺、新材料、新管理方式、新服务方式等相关的图书文献。

#### **3. 数字教学资源建设情况**

建设、配备与本专业有关的音视频素材、教学课件、数字化教学案例库、虚拟仿真软件等专业教学资源库，种类丰富、形式多样、使用便捷、动态更新、满足教学。

### **(四) 教学方法**

教学方法以“技术为基、能力为核”为导向，融合项目驱动、案例教学、情境模拟与翻转课堂等方法，按课程特性灵活组合。依托真实业务设计任务，通过分组协作、上机实操、岗位模拟强化技能；采用“线上自学+线下研习”混合模式衔接理论与实践，以“做中学”培养学生解决复杂问题能力。动态调整教学策略，推动课证融通，培养兼具数据思维、职业素养与创新精神的高技能人才。

以“实践”为核心，贴近企业场景，打破“理论课 + 实训课”的割裂模式，将“知识点讲解”与“即时实操”同步推进，让学生在“学中做、做中学”，解决“理论听不懂、实操不会用”的痛点。以“完整项目”为载体，将分散的知识点串联成“岗位能力模块”，模拟企业真实开发流程，让学生在完成项目的过程中掌握技术、理解协作，实现“毕业即能对接岗位”。通过“典型案例”拆解抽象知识点，让学生理解“技术在实际场景中的应用方式”，避免“学了语法却不知道用在哪”的困惑，在案例选择上，贴近行业热点并尽可能覆盖核心知识点。

### **(五) 学习评价**

采用过程性评价与终结性评价相结合的方法，终结性评价采用期末一次性目标评价和每个开发任务完成后目标评价相结合的方式进行。考核内容包括知识掌握、技能水平、职业素质、自主学习、团队意识、解决实际问题的能力等；过程性评价注重评价的多样

性，结合出勤、课堂提问、作业、平时测验、技能训练过程及期末考试等综合评价学生成绩。也要根据不同的课程，设定不同的考核要求。

1. 必修考试课成绩评定：总成绩=平时成绩×50%+期末考试成绩×50%

2. 选修、考查课程成绩评定：总成绩=平时成绩×40%+期末考试成绩×60%

3. 实习考核：认知实习的考核由任课教师根据实习表现和实习报告给与成绩；顶岗实习的考核由实习企业和实习指导老师共同完成：企业考核成绩（60%）+指导老师考核（40%）；毕业实习的考核由实习企业和毕业实习指导教师共同完成：企业考核成绩（60%）+毕业实习指导教师考核成绩（40%）；考核合格以上等次的学生获得学分，并纳入学籍档案。实习考核不合格者，不予毕业。考核形式注重学生的学习态度、平时成绩、卷面成绩、课堂表现、技能掌握情况等。

根据课程需要采用多样考核方法，如闭卷考试、开卷考试、实操等。鼓励学生积极参加国家、省各有关部门及学院组织的各项专业技能竞赛。

## （六）质量管理

1. 健全综合质量保障机制：学校与二级院系建立专业人才培养质量保障机制，完善教学质量监控制度。评价体系上，改进结果评价、强化过程评价、探索增值评价，并积极吸纳行业与企业参与。通过及时公开信息、接受教育督导与社会监督，形成综合评价。同时，夯实人才培养方案、课程标准、课堂教学、实验实训、毕业设计等各环节的质量建设，通过“教学实施-过程监控-质量评价-持续改进”的闭环管理，确保人才培养目标的实现。

2. 完善教学运行与管理机制：学校与二级院系加强日常教学组织与管理，定期开展课程建设、日常教学及人才培养质量的诊断与改进工作。建立健全巡课、听课、评教、评学等制度，并建立与企业联动的实践教学督导制度。要严明教学纪律，强化教学组织功能，定期组织公开课、示范课等教研活动，促进教学交流与提升。

3. 强化专业教研组织功能：专业教研组织应建立线上线下相结合的集体备课制度，定期召开教学研讨会。要善于运用各类评价分析结果，精准诊断教学问题，有效改进教学方法，从而持续提高人才培养质量。

4. 建立毕业生跟踪与社会评价机制：学校应建立常态化的毕业生跟踪反馈机制及社会评价机制。通过对生源情况、职业道德、技术技能水平、就业质量等数据的系统分析，定期评估人才培养的整体质量，并检验培养目标的达成度，为专业发展和教学改革提供数据支持。

## 九、毕业要求

根据大数据技术专业培养特色及专业培养目标的要求，通过公共基础课、专业（技

能)课、职业拓展课等的课堂教学、讲座、社会活动、文化活动、各种竞赛、大学生创新实验、实习、辅导、座谈等教学环节，在确保学生德育审核合格的基础上，引导大数据技术专业学生修满规定的140学分，使其能力达到基本要求，且各项考核全部合格，方可毕业。

# 大数据技术专业人才培养方案

## 专家评审意见表

人 才 培 养 方 案 评 审 组 成 员	姓名	单位	职务/职称	签名
	付晓炎	郑州智能科技职业学院	高级工程师	付晓炎
	王国敬	郑州智能科技职业学院	副院长	王国敬
	李胜辉	河南机电职业学院	副教授	李胜辉
	李巧君	河南工业职业技术学院	教授	李巧君
	帖莎娜	华为技术	总监	帖莎娜

评审组意见：

同意该方案通过审核。

评审组组长签字：李巧君

日期：2015年9月21日