

# 新工科背景下大数据技术通识课程 教学改革与实践\*

董轶群 王芳\*\* 付秀丽 王淑鸿

北京石油化工学院信息工程学院, 北京 102617

**摘要** 针对当前高校大数据通识课程存在的问题,以新工科理念为导向,以大数据思维能力培养为核心,基于OBE理念开展反向的课程设计与教学改革。结合工程教育认证提出大数据技术基础与应用的课程目标,秉持通识理念重构课程内容体系,以大数据处理流程为主线组织设计实践教学,面向企业实际生产环境开展教学评价,阐述具体实施过程,说明课程教学效果。

**关键字** 新工科, 大数据技术, 通识教育, OBE, 工程教育认证

## Teaching Reform and Practice for Big Data Technology General Course under the Background of Emerging Engineering Education

Dong Yiqun Wang Fang Fu Xiuli Wang Shuhong

School of Information Engineering  
Beijing Institute of Petrochemical Technology  
Beijing 102617, China  
{dongyq,0020170026,fuxiuli,wangshuhong}@bipt.edu.cn

**Abstract**—Addressing the prevalent issues in the current big data general education courses in universities, guided by the emerging engineering education, with the cultivation of big data thinking ability as the core, the course design and teaching reform is carried out based on the Outcome-Based Education. Firstly, in alignment with engineering education certification, the course objectives for the fundamentals and applications of big data technology are introduced. Subsequently, the course content system is restructured based on the concept of general education. Practical teaching is organized around the big data processing workflow, and teaching evaluations are conducted in the actual production environments of enterprises. Finally, the specific implementation process and the effectiveness of the course teaching are elaborated.

**Keywords**—emerging engineering education, big data technology, general education, OBE, engineering education certification

## 1 引言

在大数据时代,数据已成为国家基础性战略资源<sup>[1,2]</sup>。习近平总书记做出了“大数据是信息化发展的新阶段”这一重要论断,要求“加快建设数字中国”。党的二十大报告指出数字技术与实体经济深度融合具有重大意义。在新的时代背景下,2017年教育部提出了加快建设“新工科”的号召,将大数据人才培养作“新工科”建设的重要内容<sup>[3]</sup>。

在大数据时代和新工科建设背景下,构建各专业人才的大数据知识结构,掌握在各学科中运用数据的

\***基金资助**: 本文得到教育部产学合作协同育人项目“面向数据科学与大数据技术专业的教师培训”(22060739282934),北京市教育科学“十四五”规划2023年度一般课题“新工科背景下大数据通识课程建设”(CDDB23221),北京石油化工学院校级重点项目“计算机程序设计基础(C语言)”(22032005005-2/015)、“大数据技术与应用”(23033981004/002)资助。

\*\***通讯作者**: 王芳, 0020170026@bipt.edu.cn

基本能力素质,是当前本科通识教育的一项重要任务<sup>[4,5]</sup>。然而目前高校大数据人才培养开展时间短<sup>[6,7]</sup>,且多面向专业教育,尚缺乏成熟的通识课程体系与人才培养模式。大数据通识课程普遍存在:重知识传授,轻工程实践能力培养;课程育人功能发挥不充分<sup>[8]</sup>;教学内容照搬专业课程知识体系,专业性强、普适性差;大数据通识课程教材、慕课、工程案例等教学资源匮乏;教学内容难以反映大数据领域最新技术成果;实践教学环节程序设计门槛高、与企业实际生产环境差距较大。此外,大数据通识课程如何支撑各专业人才培养目标,如何融入计算机通识课程体系等一系列问题仍有待于深入探究。

围绕上述问题,北京石油化工学院以“数据科学与大数据技术”专业申报、建设为契机,同步启动大数据通识课程建设工作。大数据技术基础与应用是面

向各专业本科学生的一门通识选修课程，2 学分，共 32 学时（理论 16 学时+实验 16 学时）。课程以新工科理念为导向，围绕大数据思维能力培养，以大数据处理流程为主线，以实际工程案例为载体，聚焦主流大数据技术，在大数据通识课程的课程目标、知识体系、课程思政、工程实践、教学评价等方面进行了一系列教学探索，经过多轮次的教学实践与持续改进，取得了良好教学效果。

## 2 基于工程教育认证的多维度课程目标设计

### 2.1 4 个维度

在新工科背景下，大数据技术基础与应用贯彻落实“立德树人”根本任务，立足北京石油化工学院高素质应用型人才培养定位，本着“以学生为中心”的教学理念，着重从知识、能力、素质、思政 4 个维度对课程目标进行设计。

**知识维度：**学生能够了解大数据的发展与应用现状，熟悉大数据生态系统基本架构，掌握大数据主流技术组件的功能、基本概念、核心原理与初步的应用方法。

**能力维度：**着重培养学生的大数据思维能力，具体包括选型并部署大数据技术组件的系统架构能力、基于大数据知识的沟通与协作能力、运用大数据技术解决本专业问题的融合创新能力与初步的工程实践能力。

**素质维度：**拓展学生大数据视野，增强学生的数据赋能<sup>[9]</sup>意识，培养团队精神与融合创新意识，提高学生的职业道德水平。

**思政维度：**厚植家国情怀，激发使命担当，帮助学生树立正确的学习目标与理想信念；培养学生勇于创新、理论联系实际的科学精神，认真严谨、精益求精的大国工匠精神；强化学生的大数据伦理意识。

### 2.2 课程目标

秉持通识教育理念，课程对标工程教育认证提出的毕业要求，结合大数据企业前沿核心技术需求，将 4 个维度的内容有机融合，提出大数据技术基础与应用课程的以下 5 个具体课程目标，明确了课程目标与毕业要求之间的支撑关系，如图 1 所示。

**课程目标 1：**了解大数据的关键问题、典型应用与热点趋势，理解大数据特征与大数据思维内涵，建立大数据知识体系基本架构，开阔大数据视野、增强数据赋能与大数据伦理意识。

**课程目标 2：**了解大数据处理一般流程、典型大数据计算模式与软件系统开发生命周期，理解 HDFS、YARN、Spark 等主流大数据技术组件的核心概念、基本原理，掌握 Hadoop+Spark 集群部署方法。

**课程目标 3：**能够使用 VMWare 虚拟机软件在 Linux 操作系统上搭建集群仿真环境，能够结合实际需求合理选择大数据技术组件设计和部署大数据处理系统基础架构，能够运用 Spark 程序设计模式设计实现大数据处理系统，具备初步的大数据工程实践能力与理论联系实际的科学精神。

**课程目标 4：**具有团队意识，能够基于大数据和多学科背景知识与团队中成员开展合作式学习，相互协作完成系统设计与实现。

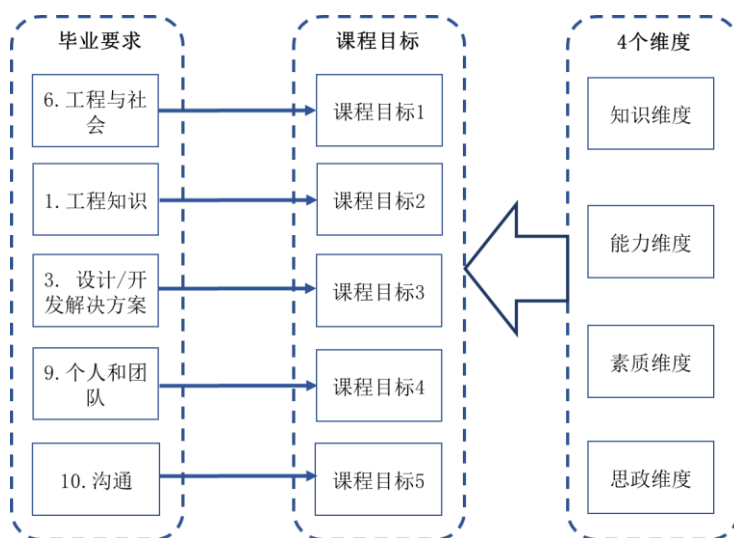


图 1 大数据技术基础与应用课程目标设计思路图

课程目标 5: 能够准确使用大数据相关术语就大数据处理问题进行书面表达与沟通交流, 能够运用多媒体软件以多种形式综合呈现课程设计核心工作。

### 3 秉持通识教育理念的课程内容体系设计

围绕着上述课程目标的达成, 课程本着弱化理论、突出实践、融合思政、呈现前沿的原则设计课程教学内容。首先结合大数据企业大数据核心技术需求, 将前沿技术 Hadoop、Spark 作为课程的教学重点和实践

平台工具。在广度上, 以大数据处理流程(数据获取、数据预处理、数据存储与管理、数据分析、数据可视化与应用)为主线组织内容, 帮助学生开阔大数据视野, 构建大数据基础知识体系; 在深度上, 围绕大数据存储、分析两个关键问题, 重点讲解 HDFS、Spark 大数据技术组件的核心概念、基本原理与应用; 同时深入挖掘课程内容中国家情怀、工程伦理、工匠精神、科学精神等方面的思政元素, 并将其有机融入教学过程中, 增强学生大数据伦理意识、树立正确的数据观和价值观。课程内容体系如图 2 所示。

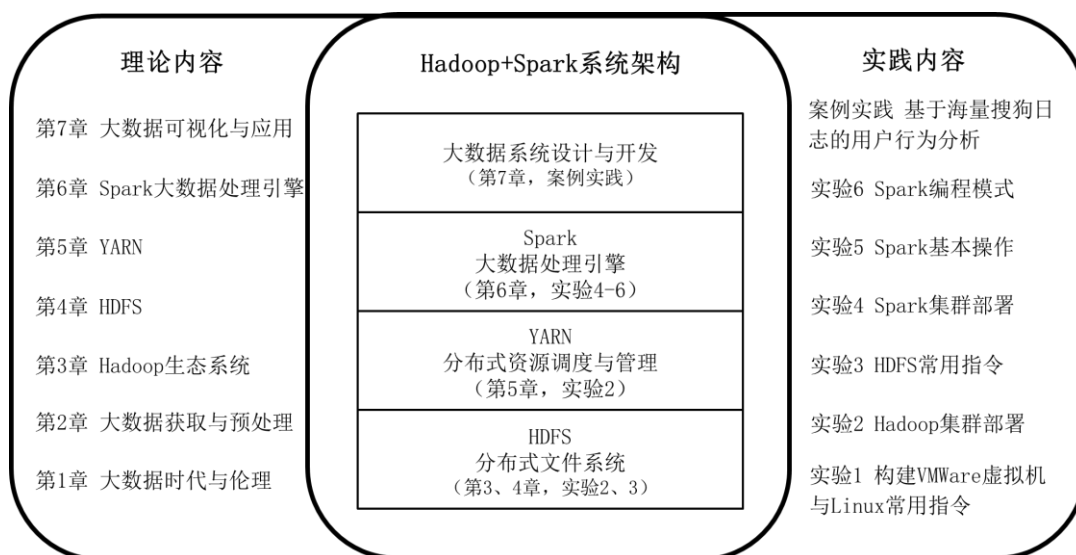


图 2 大数据技术基础与应用通识课程内容体系

## 4 以大数据处理流程为主线的实践教学设计

实践性强是大数据相关课程的一个显著特征。在通识课程中, 各专业学生普遍存在计算机基础薄弱、缺乏工程实践能力等情况, 进一步增加了课程实践教学的难度。为此本课程在实践平台、案例资源、教学模式等方面采取了一系列行之有效的措施, 全面确保实践教学质量。

### 4.1 虚实结合的实践教学平台

现有大数据相关实验多存在实验环境要求高、缺乏整体设计、与实际生产环境脱节等问题。为此课程实践环节首先从基于 VMware 创建虚拟机入手, 帮助学生利用有限资源构建个性化的大数据实验环境。区别传统教学中的伪分布集群部署模式, 本课程面向企业实际生产环境, 指导学生基于虚拟机部署完全分布模式下的 Hadoop+Spark 大数据集群。

### 4.2 分层递进的实践教学内容

以部署的 Hadoop+Spark 大数据集群为实践平台, 课程将实践教学内容划分为验证性、设计性和探究性三类, 并依次组织实践教学, 每一类实践内容又细分为基础和进阶内容, 循序渐进、由浅入深地指导学生开展工程实践, 如图 3 所示。

课程重点围绕新冠疫情防控、电信用户分析、地震预测等时事热点建设了一系列趣味性强、易上手、突显学科交叉和数据赋能的工程案例资源。在教学过程中, 以大数据系统开发生命周期为主线, 着重从构思(C)、设计(D)、实施(I)、运行(O)4个层面组织设计实践内容, 高度还原企业大数据集群实际部署场景与开发流程, 综合培养学生的工程实践能力, 实现学校与企业的无缝衔接。

### 4.3 模式化的 Spark 程序设计

为方便不同专业、不同程序设计基础的学生开展大数据工程实践, 课程归纳总结不同应用场景中

Spark 程序设计的思路与方法。针对流式计算、图计算、交互式查询和批处理四类常见大数据计算模式，设计出相应的 Spark 程序设计模式。这些模式具有逻辑清晰、代码简洁、通用性强等特点，能够有效降低

课程程序设计门槛，有助于不同领域间的拓展迁移，为学生基于大数据开展融合创新提供便捷工具。例如图 4 为基于 Spark 的批处理程序设计模式。

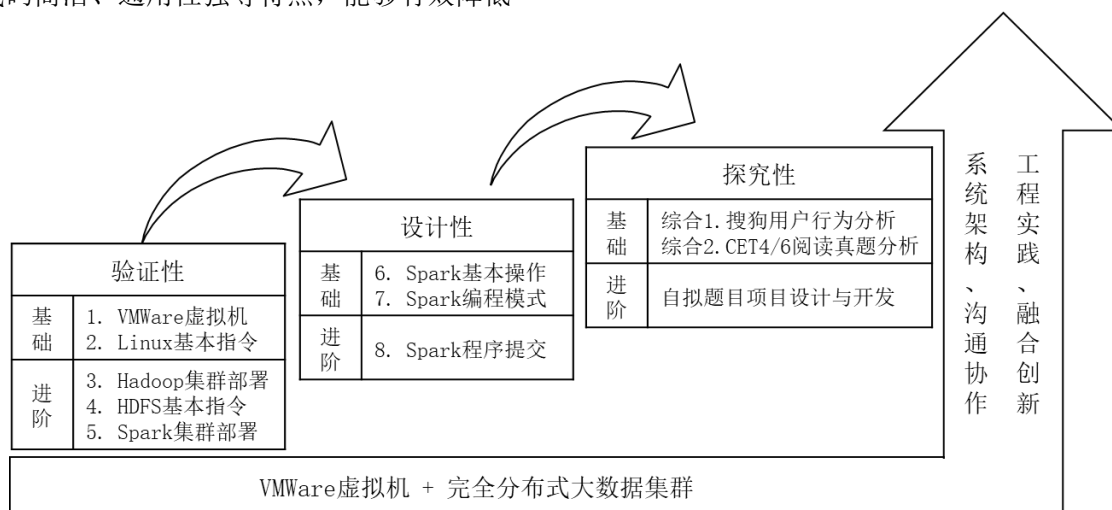


图 3 分层递进的实践教学内容

## 5 面向实际生产环境的教学评价体系

为有效评价教学目标的达成情况，课程构建了量化、多元、过程性的评价体系。参照企业实际生产环境中系统开发场景，围绕大数据岗位核心技术与能力

需求，课程有针对性地设计考核内容与方法，主要涉及过程性评价（40分）、项目验收（20分）、项目汇报（20分）、课程设计报告（20分），其中项目汇报与课程设计报告以团队形式完成。

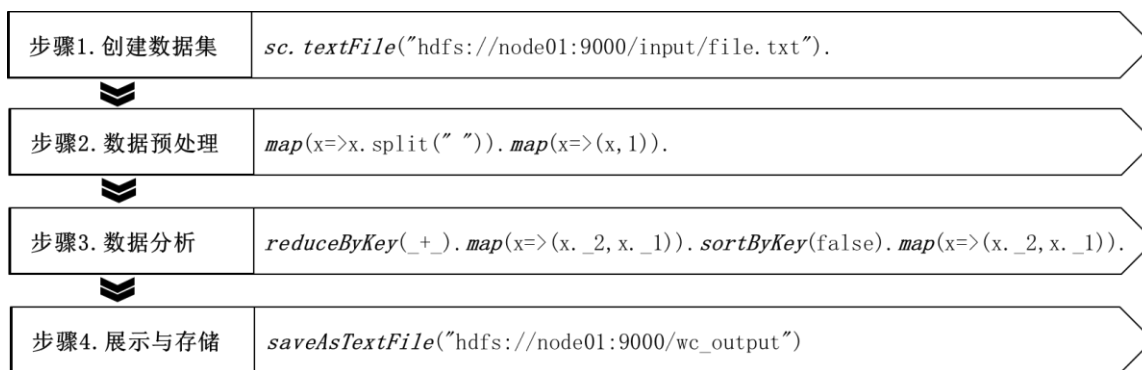


图 4 Spark批处理程序设计模式示例

### 5.1 过程性评价

过程性评价主要由随堂测验、实验报告和课堂表现构成。随堂测验是过程性评价的重点，主要基于云班课测验活动展开，围绕课程具体教学目标，侧重于每节课中的基本概念、教学重点与难点，时间不超过 15 分钟。随堂测验能够帮助学生理解巩固教学内容，实时反馈的测验数据方便教师及时掌握学生的学习情况，发现教学中的问题与难点，指导教师有针对性调整教学设计。为提升学生课堂参与度，课程为随堂测

验赋予较高权重，将对应的经验值转换后计入评价体系。

### 5.2 项目验收

项目验收分为两部分：大数据系统部署验收与应用程序运行结果验收。大数据系统部署验收主要考核学生基于 VMWare、Linux、Hadoop、HDFS、Spark 等组件构建分布式集群的能力；应用程序运行结果验收主要考核学生基于 Spark 的程序设计与实现能力。

### 5.3 项目汇报

项目汇报环节以企业项目汇报要求和场景为背景, 首先要求学生制作 5-8 分钟的项目汇报 PPT, 内容以大数据系统开发生命周期为主线, 反映大数据处理一般流程, 突出课程设计主要工作。PPT 设计方面要求生动、简洁。在项目验收与项目汇报基础上, 教师现场提出 3-5 个课程设计相关问题, 要求学生在规定时间内回答。通过项目汇报着重考核学生团队协作能力、基于应用领域背景知识和大数据技术语言的沟通与交流能力。

### 5.4 课程设计报告

在期末考核环节, 学生可在给定范围内选题完成课程设计, 同时鼓励学生融合创新, 使用本专业相关数据自拟课程设计题目, 旨在促进学生的深度学习, 利用大数据技术发现并解决自身专业相关领域中的问题。

课程要求学生以大数据系统开发生命周期为主线组织课程设计报告内容, 要求能够反映出设计与开发过程中的核心问题、大数据处理的一般流程、设计思路与主要工作。课程设计报告不仅是巩固深化课程教学内容的重要方法, 也是综合评价学生大数据思维能力、工程实践能力有效工具。

## 6 教学成效

大数据技术基础与应用作为北京石油化工学院通识选修课程已累计授课本科生 200 余人, 涉及计算机科学与技术、数据科学与大数据技术等 5 个电子信息类专业, 以及环境工程、物流管理、生物制药等 19 个非电子信息类专业, 受到学生广泛好评。每轮教学实践后围绕着课程目标的达成情况进行定量评估与分析, 近三年课程目标达成度总体持续提升, 如图 5 所示。

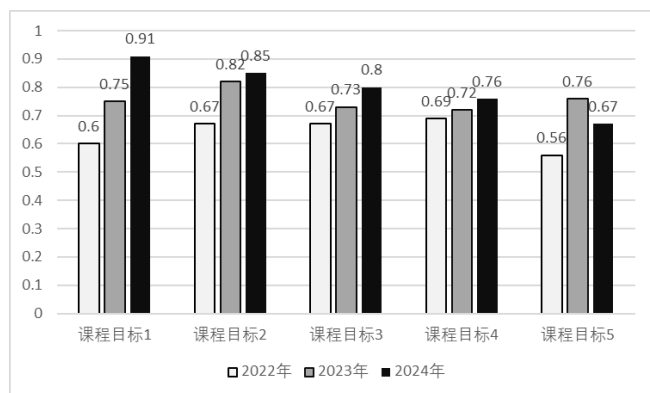


图 5 2022-2024 学年课程各目标达成情况

教学过程中涌现出“十四五规划词频热点统计”、“疫情背景下北京周边游评价的数据变化分析”、“近三年北京空气颗粒物含量统计分析”等一系列大数据与各领域交叉融合的学生自主研发课程设计项目, 其中“基于 Spark 的电信离网用户分析”获得“互联网+”大赛本科生创意项目组校内二等奖。基于本课程, 学生申报市级大学生创新训练项目 (URT) 3 项。

依托本课程编写的教材《大数据技术基础》于 2021 年在蓝墨云平台公开出版, 是该平台首部大数据相关云教材, 已超 500 人下载使用, 为 17 所高校的大数据相关课程建设提供了有益借鉴, 获得师生们广泛好评。2023 年, 大数据技术基础与应用获评北京高等学校优质本科课程。

## 7 结束语

目前大数据通识教育尚缺乏成熟的人才培养模式, 在课程建设理念、课程目标、知识体系、课程思政、实践模式等方面存在诸多挑战。基于 OBE 理念, 通识课程大数据技术基础与应用对标工程教育认证毕业要求提出了多维度的课程目标; 围绕着课程目标的达成, 本着弱化理论、突出实践、融合思政、呈现前沿的原则提出了层次化的课程内容体系; 针对以往大数据实践教学存在的问题, 探索并实践了虚实结合、分层递进和模式化的实践教学模式; 为了有效评价课程目标的达成情况, 设计了面向企业实际生产环境的教学评价体系。课程在实践过程中取得了显著成效, 能够为新工科背景下大数据通识课程建设提供思路, 为各专业实践型、创新型和复合型人才培养提供有力支撑。

## 参考文献

- [1] 中华人民共和国国务院. 促进大数据发展行动纲要. 成组技术与生产现代化[J], 2015(3): 51-58.
- [2] 梅宏. 大数据导论 [M]. 北京: 高等教育出版社, 2018.
- [3] 钟登华. 新工科建设的内涵与行动[J]. 高等工程教育研究, 2017, (3): 1-6.
- [4] 朱涛, 夏玲玲, 袁明. 大数据通识课程教学内容探索[J]. 高教学刊, 2023, 9(02): 110-113. DOI:10.19980/j.cn23-1593/G4.2023.02.027.
- [5] 杨先娣, 熊素萍, 张华, 彭智勇[J]. 计算机技术与教育学报, 2023, 11(2): 101-105.
- [6] 朝乐门, 杨灿军, 王盛杰等. 全球数据科学课程建设现状的实证分析[J]. 数据分析与知识发现, 2017, 1(06): 12-21.
- [7] 徐昊, 秦玥, 黄岚. 面向通识教育的数据科学课程建设[J]. 计算机教育, 2016, 260(08): 158-162.
- [8] 王丽丽, 方贤文. 新工科背景下的大数据专业课程思政建设[J]. 安徽理工大学学报(社会科学版), 2022, 24(03): 104-108.
- [9] 周傲英, 周烜. 数据专业人才培养方案与核心课程体系建设[J]. 中国大学教学, 2020, 358(06): 15-21+33.