

# 数据挖掘与分析课程教学改革与探索

肖 薇\*, 任菊慧, 林良钊

上海电机学院电子信息学院, 上海

收稿日期: 2022年9月19日; 录用日期: 2022年10月19日; 发布日期: 2022年10月26日

## 摘 要

为了使研究生培养契合大数据时代人才发展的需求, 以“数据挖掘与分析”课程为例, 以充分激发和培养研究生的独立思考与实践能力和指导思想, 分析该课程目前教学环节中存在的问题, 提出“结合行业背景、注重案例分析、提高实践能力、改善考核方式”的教学改革思路。该课程教学改革与探索研究有助于提升研究生的学习兴趣与实践能力, 也为同类院校开展研究生新工科建设教学改革提供了一定的参考。

## 关键词

数据挖掘与分析, 教学改革, 研究生教育

# Teaching Reform and Exploration of the Course of Data Mining and Analysis

Wei Xiao\*, Juhui Ren, Liangzhao Lin

School of Electronic and Information, Shanghai Dianji University, Shanghai

Received: Sep. 19<sup>th</sup>, 2022; accepted: Oct. 19<sup>th</sup>, 2022; published: Oct. 26<sup>th</sup>, 2022

## Abstract

To make the postgraduates cultivation meet the demand for the talent development in the era of big data, this paper takes the course of “data mining and analysis” as an example, and fully stimulates and cultivates the independent thinking and the practical ability of postgraduates as the guiding ideology. By analyzing the existing problem in the current teaching of the course, the teaching reform idea of the “combining the industry background, focusing on the case analysis, improving the practical ability, and improving the assessment methods” is proposed. The research of the teaching reform and the exploration of this course are helpful to improve the learning in-

\*通讯作者。

terest and the practical ability of postgraduates, and also provide a certain reference for similar colleges to carry out the teaching reform of new engineering construction for postgraduates.

## Keywords

Data Mining and Analysis, Teaching Reform, Graduate Education

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

随着研究生综合教育改革的深入推进,研究生招生比例逐年上升[1],培养规模逐渐扩大,成为培养高层次应用型人才的主要途径。研究生综合教育改革紧密围绕“服务需求、提高质量”的主线,不断深化综合改革,努力构建符合专业学位研究生教育规律的培养模式和评价机制,大力提升专业学位研究生教育质量,为高质量研究生教育的建设提供有力保障。在研究生培养的整个过程中,课程学习是研究生培养的重要组成部分,是研究生学习基础理论和系统专业知识的重要途径,它体现并影响着学校研究生教育质量,课程考试是课程学习质量的重要抓手。2015年8月,《国务院关于印发促进大数据发展行动纲要的通知》明确指出:创新人才培养模式,建立大数据人才培养体系;鼓励高校设立数据科学等专业,重点培养大数据专业人才;鼓励跨校、跨学科联合,大力培养跨界复合型大数据人才;鼓励高校、职校和企业合作,积极培育大数据技术和应用创新型人才;依托社会化教育资源,开展大数据知识普及和教育培训[2]。国内许多应用型本科院校纷纷启动大数据背景下相关专业的课程体系的重构、教学方法与方式的改革、数据科学或大数据专业培养计划的制定等措施,以期培养应用型大数据人才和输出能力,解决大数据人才的需求瓶颈[3]。从目前全国各高校的信息化规划来看,都将大数据列为重点工作内容。高校中的数据有很高挖掘分析价值,近几年来高校数字化校园的建设丰富了学校教学、科研、行政管理、生活服务等方面的数据。贵州省、浙江省、广东省、上海市、重庆市、北京市、新疆维吾尔自治区等省、市、自治区陆续制定大数据发展政策,并强调加强大数据人才培育引进等队伍建设,建立大数据人才研发和创新体系,完善大数据人才培养机制。我国大数据人才培养的现实需求旺盛,各级政府通过制定政策,积极为高校人才培养、企业和政府人才引进等提供良好的制度保障,促进大数据人才培养工作取得实质性成效。因此,在强大的社会需求下,培养出优秀的数据挖掘分析师,是高等院校研究生培养过程中的重要组成部分[4]。

数据挖掘是指对大量数据进行分析与处理,从而揭示其中隐含的、先前未知的并具有潜在价值信息的过程。而“数据挖掘与分析”课程是一门相对较新的交叉学科,涵盖了概率统计、机器学习、数据库等学科的知识内容,难度较大,因此该课程陆续在各高校研究生培养阶段开设。不同学校将其归入不同的专业中,如计算机、信息管理、统计学、医学等,以促进学生掌握数据挖掘的基本原理和各种数据挖掘算法,掌握数据分析与处理、高级数据库编程等技能,进而达到数据聚类、分析、关联分析的目的。在此基础上,“数据挖掘与分析”课程教学改革与探索在课程教学中以数据挖掘、数据分析为抓手,融入机器学习等当前研究领域中的技术热点和实际应用场景,以打破传统“偏理论,轻应用”的研究生教学模式,通过深入研究与探索数据挖掘与分析课程体系建设的创新型教学方法,构建出面向大数据应用技术场景的硕士研究生课程建设体系,以培养符合大数据技术发展领域的技术人才。

## 2. 传统课程教学方法存在的问题

### 1) 教学模式陈旧，学习氛围弱

目前较多高校在对研究生传授“数据挖掘与分析”课程时仍采用本科化的教学模式和方法[5]。授课教师参照数据的基本挖掘方法与分析工具等内容进行传统理论讲授，而缺乏对研究生理论知识应用及动手实践的培养。学生坐在教室以听为主，而缺乏独立思考，且缺少与教师沟通。因此，采用这种培养模式没有对研究生与本科生进行差异化培养，同时无法提高研究生掌握大数据基础理论知识与实践应用能力的自主能动性和科研创新能力。

### 2) 脱离行业背景，教学效果欠佳

数据挖掘与分析课程中涵盖了概率统计、机器学习、数据库等学科知识，这类知识较多依赖于算法推导与数学建模，而大量具有难度的数据挖掘理论与算法的学习使得研究生们丧失了学习兴趣，学完即忘，不知所用。而紧密联系研究生生活实际的案例不仅能够调动学生的学习积极性，同时还能够挖掘数据隐藏信息的深层次信息，帮助理解将基础理论知识运用于实践。

### 3) 缺少案例演练，动手实践能力缺乏

研究生培养需要在强化数据的基本挖掘方法与分析工具等理论知识的基础上，着重培养其运用大数据理论知识及将其转化为应用实践的能力，特别是结合自身研究方向的问题解决海量数据存储与运算的能力。而传统授课过程缺少案例演练或直接引入传统案例[6]，无法紧跟大数据行业发展前沿技术或问题，从而导致研究生缺乏动手实践能力，或在课堂中掌握的数据挖掘与分析方法难以与其实际科研动态吻合。

### 4) 课程过程考核不突出

研究生课程需要强调将大数据挖掘与分析的基础理论知识运用于解决实际应用的能力，而对实际应用能力的考核无法单独通过基础理论知识的考核来衡量。因此，打破传统课程考核方式，在教学内容和课程体系改革中课程考核突出实践教学，强调对研究生进行数据挖掘与分析过程的考核，才能真正提升学生的创新性和系统考虑周密性等[7]。

## 3. 教学改革与探索

针对当前“数据挖掘与分析”课程实际开展情况，拟从以下方面来深化课程建设：

1) 深入挖掘与研究课程基本理论知识及其在大数据技术场景中的应用，加强理论与实际应用场景的结合。

传统课程教学中采用理论教学和纯粹验证教学。“数据挖掘与分析”课程主要针对行业背景是智能制造领域等，在研究生课程改革中，教学方法上拟定由验证型、提高型向综合型、设计开发型和创新型转变。课程教学内容划分上按教学规律分阶段递进展开，按照由认识到理解、由理解到掌握、由具体到抽象、由微观到宏观的学习规律递进，以渐进的形式达到对学生动手能力和创新能力的培养。目标由以往单纯的教学型向产学研相结合的方向转变。采用案例化教学(比如围绕新能源生产的数据预处理、关键特征提取与处理、特征分析等)，以项目驱动法和案例为新元素，广泛收集大数据应用开发的最新案例。通过以实际案例为导向组织教学内容，从而强化各知识点的运用，让研究生及早了解行业现状，对毕业后即将接触到的问题有初步认识，对企业当前的研发水平有较清醒的认识。

2) 设计贴近行业的研究生教学案例，结合大数据技术在特定行业中的实际应用，提升实践环节的的前沿性。

数据挖掘与分析课程的行业背景是智能制造领域等，主要就业区域和行业领域分别是长三角地区和制造业。因此，在强化理论知识的基础上，通过充分利用学校长期积累的广泛社会基础和行业企业或者

校企合作平台等途径,调研这些区域和行业领域对应用型大数据技术人才的职业岗位要求,积极了解企业的前沿方向与技术,再依据职业岗位解读其蕴含的知识、能力和素质要求,从而为硕士研究生培养方案和课程设置提供基本依据,进一步丰富与优化本课程中实际案例的导入,为有针对性地培养“适销对路”的人才奠定基础。此外,从理论与实践两个方面入手,实现行业倾向性。在传统授课过程中,围绕智能制造企业的数据与案例(比如围绕起重机数据的采集、清洗、去噪、处理、可视化与故障预警),进行知识点的讲解。同时,还可以增加以小组为单位的创新大作业。学生可以结合周边企业的大数据处理技术与方案,实现与大数据挖掘与分析相关的课题,这样可以进一步提高实践动手能力和团队合作能力。通过课题的不断改进和完善给研究生带来成就感,从心理上对专业前景充满信心,提高研究生专业技能和就业的核心竞争力,帮助研究生提前熟悉毕业设计制作过程,培养研究生的职业能力。在此基础上,在条件允许时,可邀请企业技术总监或校友进课堂,能够让学生及早了解行业现状,对毕业后即将接触到的问题有初步认识,对企业当前的研发水平有较清醒的认识。

3) 熟悉行业的大数据处理操作,提升动手实践能力,挖掘海量数据中有价值的信息。

硕士研究生的培养要求需要同时加强专业知识、技能和素养的学习。按照硕士研究生专业能力达成的内在逻辑关系与毕业要求的映射,以及与课程体系设计的逻辑关系的匹配程度,设计及持续优化合理的课程评价机制,从而为建立完整且符合时代发展特性的课程体系提供可靠依据。数据内在信息挖掘与分析操作需要在课程建设过程中提升系统理论深度,着重培养研究生的系统性与逻辑性;同时,加强数据收集与预处理方法的讲解,为将来系统规划与算法选型夯实基础。目前,学校平台较多采用实验模拟或使用公共数据源或软件自带数据源,不仅数据量小,且大多为处理之后的数据。而通过实际爬虫或真实生产和实验获得数据,通常数据量较大,且存在较多冗余或不完整记录。因此,为了让研究生了解数据本源和特性,增加对数据价值的深层次理解,需要加强对大数据收集与预处理环节的重视,从而将理论知识学习与企业生产实际接轨。数据预处理主要是检查数据质量,包括准确性、完整性、一致性、及时性、可信度、可解释性等内容,是数据挖掘项目中最费时费力的环节,并对后期的数据分析工作产生严重影响。因此,可以借助产教融合等机制,让研究生深入企业生产环境实时采集数据或利用企业采集的离线数据进行处理。同时,根据企业实际生产需求,设计合理可行的数据处理步骤,并借助不同质量数据的不同的处理算法来去除数据集中的噪声,从而有助于帮助研究生们了解数据集的分布状况与重要特征。

4) 改善考核方式,研究大数据基础理论知识与实际应用案例结合的教学方法,以学术论文形式总结与提炼实践教学成果。

在课程建设中新课程考核突出实践教学,注重研究生在实践过程中应用能力和过程的考核。突破传统以理论为主的闭卷考试,面向数据规划、设计、分析与应用能力,围绕智能制造场景设计综合考核体系,强调过程考核,注重研究生实践过程中发现问题与解决问题的能力,考察研究生知识的创新运用与团队协作能力。在成绩综合评定上,减少期末考试所占比例,增强课程过程考核的比例,也即增大研究生的实践动手过程的考核。拟通过实践创新项目大作业,进一步锻炼学生运用基本理论与技术解决实际工程问题的能力,尤其考查实验过程中学生的创新性和系统考虑周密性等。同时,依托专业教学平台,在课程体系的基础上,通过直播、借鉴示范课程等途径,为研究生提供全天候学习环境。

## 4. 结论

随着我国大数据产业快速发展,大数据技术与应用的创新突破,大数据产业在促进经济发展与保障社会运行等多方面起到积极的促进作用。在大数据快速发展的契机下,将大数据与研究生人才培养深度融合,有助于形成一种新思路、新模式、新理念的人才发展路线。本文对研究生“数据挖掘与分析”课

程教学改革进行了探索,针对“教学模式陈旧、脱离行业背景、缺少案例演练、课程过程考核不突出”的问题,提出了“加强理论在实际场景中的应用、设计贴近行业的研究生教学案例、提升动手实践能力、增强课程过程考核”的改革思路,旨在丰富研究生理论知识及提升其实践动手能力的基础上,重点充分激发和培养研究生的独立思考与实践能力,从而使研究生掌握大数据时代的数据分析与处理能力,成为顺应大数据时代发展需求的应用型人才。

## 参考文献

- [1] 熊丙奇. 研究生再扩招, 如何避免“身价”贬值[J]. 年轻人: (A 版), 2020, 40(3): 56.
- [2] [http://www.gov.cn/zhengce/content/2015-09/05/content\\_10137.htm?url\\_type=39&object\\_type=webpage&pos=1](http://www.gov.cn/zhengce/content/2015-09/05/content_10137.htm?url_type=39&object_type=webpage&pos=1)
- [3] 常浩娟, 余小平, 吴琼. 项目导入和分层式任务驱动在“数据挖掘”教学中的应用[J]. 科教导刊, 2017(10): 114-115.
- [4] 姜明洋, 白青海. 数据科学与大数据技术专业建设研究——以数据挖掘课程为例[J]. 内蒙古民族大学学报: 自然科学版, 2020, 35(4): 350-352.
- [5] 高建瓴, 潘成成. 以“问题驱动”为基础的“数据挖掘及应用”课程教学实践[J]. 新型工业化, 2020, 10(9): 187-189.
- [6] 孟凡奇, 孙昊晨, 王敬东. 新工科研究生“数据挖掘”课教学改革探索[J]. 黑龙江教育(高教研究与评估), 2021(4): 54-55.
- [7] 叶志伟, 高榕, 林姗, 邓娜. 面向研究生思辨能力培养的数据挖掘课程教学实践[J]. 软件导刊, 2022, 21(7): 240-243.