

健康医疗大数据视角下面向应用型人才培养的大数据导论课程教学改革实践

王晓东,王超,左风云,赵慧茹,吴雅琴

(内蒙古医科大学计算机信息学院,内蒙古 呼和浩特 010110)

摘要:随着健康医疗大数据应用进程不断深入,在健康医疗临床和科研、健康医疗大数据资源共享开放、健康医疗行业治理等领域,具有医学背景的大数据专业人才将会受到更多用人单位的青睐。在当前健康医疗业务与大数据技术深度融合的背景下,本文针对大数据导论课程特点,提出构建以培养学生理论知识能力、实践操作能力、自主学习能力、综合应用能力的递进模式,开展基于领域知识的大数据导论课程教学改革探索,旨在为应用型人才培养奠定坚实的理论与实践基础,也为其他课程的教学改革提供参考。

关键词:大数据导论;健康医疗大数据;应用型人才;教学改革

中图分类号:G642

文献标识码:B

DOI:10.3969/j.issn.1006-1959.2021.04.007

文章编号:1006-1959(2021)04-0025-04

The Practice of Teaching Reform of the Big Data Introduction Course for the Cultivation of Applied Talents from the Perspective of Healthcare Big Data

WANG Xiao-dong, WANG Chao, ZUO Feng-yun, ZHAO Hui-ru, WU Ya-qin

(School of Computer Information, Inner Mongolia Medical University, Hohhot 010110, Inner Mongolia, China)

Abstract: With the deepening of the application of big data in healthcare, in the fields of healthcare clinical and scientific research, healthcare big data resource sharing and openness, healthcare industry governance, and other fields, big data professionals with medical backgrounds will be favored by more employers. In the context of the deep integration of the current health care business and big data technology, this article proposes to build a progressive model based on the characteristics of the big data introduction course to cultivate students' theoretical knowledge ability, practical operation ability, independent learning ability, and comprehensive application ability. Exploration of the teaching reform of the big data introduction course of domain knowledge, it aims to lay a solid theoretical and practical foundation for the cultivation of applied talents, and also provide a reference for the reform of other courses.

Key words: Introduction to big data; Health and medical big data; Applied talents; Teaching reform

2016年国务院办公厅印发的《关于促进和规范健康医疗大数据应用发展的指导意见》(国办发〔2016〕47号)明确提出,健康医疗大数据是国家重要的基础性战略资源。健康医疗大数据应用发展将带来健康医疗模式的深刻变化,将满足人民群众多层次、多样化的健康需求,同时也将对本领域内专业技能人才提出更高要求。如何培养高质量的技术、经济、管理、社会等领域知识能力并重且能满足多领域需求的大数据类人才是高校面临的巨大挑战^[1]。目前,以医疗和大健康行业为依托与背景进行大数据人才培养的高校少之甚少,既具备必需的医学基础知识,又精通医疗健康大数据知识与技能的专门人才较为缺乏^[2]。2017年内蒙古医科大学(以下简称学校)申报数据科学与大数据技术专业(以下简称大数据专业),并于2018年获批,同年9月开始招收第一届本科生。该专业定位为立足内蒙古自治区,面向全国,培养健康医疗大数据领域复合应用型人才。依托学校蒙医药学、临床医学、中医学、药理学、护理学

等优势专业,为开办以健康医疗大数据为特色的大数据专业提供了基础条件,同时也为师生在医疗卫生领域开展大数据研究提供了平台与保障。在充分吸收借鉴了学校信息管理与信息系统专业(医学方向)与生物医学工程专业的丰富办学经验基础上,以及对行业发展动态、人才市场需求等方面进行广泛深入调研,构建了具备自身鲜明专业特色的大数据专业的培养方案与课程体系^[3]。本文针对数据科学与大数据技术专业的大数据导论课程进行教学改革,从知识体系与能力体系等方面分析了培养的要求与目标定位,在知识点的串接、动手能力意识的培养与实际操作方面提出了可供参考的方案。

1 大数据导论课程的教学定位与目标

大数据导论是针对大数据专业本科生设置的一门先导基础课程,在众多的专业课程中具有举足轻重的地位。大数据导论主要探讨大数据时代及其背后的科学问题--数据科学的理念、理论、方法、技术、工具、应用及最佳实践,为学生更好地掌握专业知识以及提升自己的核心竞争力打下坚实基础^[4]。通过大量文献查阅发现,大数据导论课程一般是在学生第一学年基础课程教学后的第三或第四学期开设,旨在为学生普及相关专业基础知识,引导学生进行下一步的专业学习。大数据导论课程在学科基础课和专业课之间起到了承上启下的作用。学校大数据专业的大数据导论课程与其他核心课程之间的关系见图1。

基金项目:1.2020年全国高等院校计算机基础教育研究会计算机基础教育教学研究项目(编号:2020-AFCEC-283);2.2018年度内蒙古医科大学本科教学质量工程项目(编号:NYJTX201803)

作者简介:王晓东(1978.9-),男,内蒙古锡盟苏尼特右旗人,硕士,副教授,主要从事医学信息辅助决策、健康医疗大数据处理分析工作
通讯作者:吴雅琴(1978.11-),女,内蒙古乌兰察布人,硕士,副教授,主要从事健康医疗大数据处理分析工作

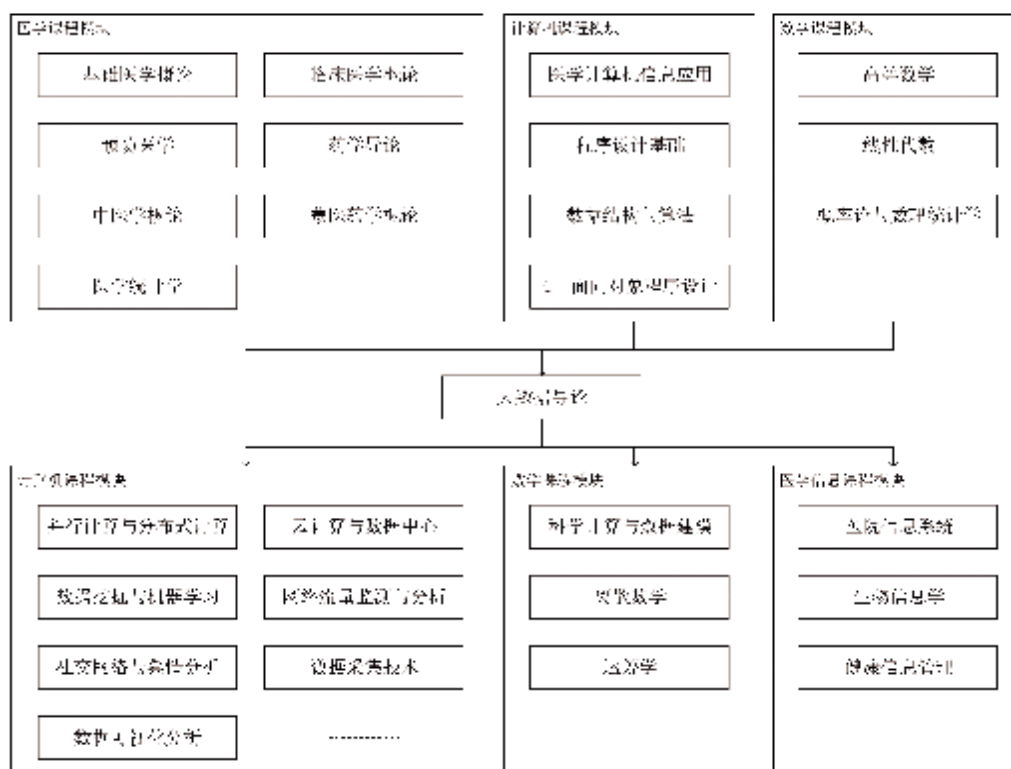


图1 大数据导论课程与其他核心课程之间的关系

2 大数据导论课程教学中存在的问题及解决对策

2.1 专业教学与健康医疗领域知识融合不够 学校大数据专业定位是培养健康医疗领域应用型大数据专业人才,专业特色必须通过课程体系作为载体并且予以体现。医学与医学信息类课程数量在课程体系中占有较大比例,旨在要求学生掌握中医、蒙医、临床、预防、药学、医学信息学等基本理论知识,实现与后续开设的专业课交叉融合。目前大数据导论课程教学以基本概念和基本理论讲授为主,辅以教学案例帮助学生理解掌握,但是在教学过程中存在领域知识与专业内容相互独立,缺乏渗透融合的问题。当前健康医疗大数据涌现出丰富多样的新业态与产业链,只有在健康医疗领域知识应用的场景下,循序渐进开展理论与实验教学,学生才能深入了解行业动态,把握正确努力方向。

2.2 学生现有的专业知识结构水平不能满足案例教学实践需求 在大数据导论课程开课阶段,学生掌握的专业知识有限,独立解决部分教学案例的实际问题有较大难度。在实验教学环节,可以布置验证性实验,即教师分析案例的主要解决问题思路,将核心代码交于学生,由其在实际环境中测试运行程序,查看最终结果是否与预期结果相同,达到让学生深入理解相关大数据理论的内涵以及了解大数据处理分析技术的使用场景的教学目标。在此基础上,课后可以布置模块作业,进一步激励学生通过 SPOC、MOOC 平台或其他教学资源,学习专业知识解决遇到的实

际问题,如此,可以促进其自主学习能力的培养。

2.3 学习的专业知识碎片化,缺乏有机统一 修读完不同课程,学生掌握了不同学科的专业知识,很容易出现知识“碎片化”的现象,即知识之间没有交集,各自为营,缺乏有机统一。设计大数据导论课程的教学案例以及综合实训,将所学内容进行整合,真正实现“学以致用”,同时教学案例以及综合实训的解决问题思路也可以指导学生后续的专业学习。

3 大数据导论课程教学改革

按照大数据专业的培养健康医疗大数据领域应用型人才的要求,为了实现学科深度融合,体现专业特色,在专业课程体系中,增加了医学与医学信息课程模块。在专业总学分固定的限制下,各门课程在满足教学基本需求的基础上对学时进行了压缩。大数据导论课程在大二上学期(第3学期)开课,课时为32学时。到目前为止,已经完成该课程的第一轮教学,通过学生的反馈,教学效果总体较好。但由于以上提及的原因,该课程在教学内容的广度与深度以及教学方式上,还有较大的提升空间。

3.1 大数据导论课程教学改革目标 知识与能力的培养离不开课程学习、训练与实习实践^[9]。课程目标是实现专业人才培养目标的基础^[6]。根据大数据导论课程的现状以及教学中存在的问题,在健康医疗大数据视角下,对大数据导论课程的理论与实验教学内容以及教学方式的教学改革,构建基于领域知识大数据导论课程的能力培养递进模式,旨在为

应用型人才的培养奠定坚实基础。通过基于数据科学基本流程的理论教学,培养学生理论知识能力;通过基于领域知识的真实案例教学,培养学生实践能力;通过基于 SPOC 的拓展学习,培养学生自主学习能力;通过基于解决领域问题的综合实训,培养学生综合应用能力。基于领域知识大数据导论课程的能力培养递进模式见图 2。

3.2 大数据导论课程教学内容改革 这门课程在设计时,首先需要回答一个问题,即如何处理和后续课程的关系,亦即大数据导论课程在广度和深度方面应该如何设计,才能既避免和后续课程的简单重复,同时又能够起到入门和引领的作用^[7]。

3.2.1 理论教学内容设计 针对大数据导论课程的定

位,在原有的七章教学内容的基础上进行整合,去粗取精,调整为六个章节。大数据导论课程理论教学内容设计见表 1。总之,大数据导论课程教学涵盖数据科学各个环节,包括基础理论、数据采集、预处理、存储与管理、处理与分析、可视化等理论内容,以及各流程的处理方法和技术。

3.2.2 实验教学案例设计 课程实验是学生实践教学的重要组成部分^[8]。大数据导论课程的实验教学案例均来源于医院内去除隐私信息的真实病历数据,课程实验时学生可以尽早熟悉健康医疗领域知识与专业融合的应用场景,达到潜移默化的效果。大数据导论课程实验教学案例设计见表 2。

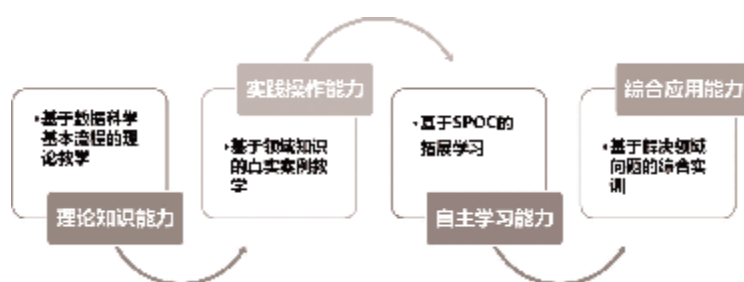


图 2 基于领域知识大数据导论课程的能力培养递进模式

表 1 大数据导论课程理论教学内容设计

| 章节名称 | 学习内容 |
|--------|--|
| 概论 | 数据科学与大数据技术专业的发展历史、知识结构、培养目标与要求、本专业课程设置及特点 |
| 基础理论 | 数据科学中的基础理论,核心术语、研究目的、理论体系与基本原则,数据科学家的主要职责与能力要求,数据科学的发展简史 |
| 理论基础 | 数据科学的学科地位,数据科学的理论基础——机器学习、统计学、可视化理论的知识体系及代表性方法 |
| 流程与技术 | 基本流程、数据加工、数据审计、数据分析、数据可视化、数据故事化以及项目管理 |
| 技术与工具 | 数据科学的技术体系、MapReduce、Hadoop、Spark、NoSQL 与 NewSQL 等 |
| 数据产品开发 | 主要特征、关键活动、数据采录、数据能力、数据战略以及数据治理 |

表 2 大数据导论课程实验教学案例设计

| 案例名称 | 案例内容 |
|---------------------------|--|
| 结构化与非结构化健康医疗大数据存储方式、结构与特点 | 深入了解《全过程医院信息管理平台》内医院信息系统(HIS)、电子病历(EMR)、检验科信息系统(LIS)、影像归档和通信系统(PACS)等系统内结构化与非结构化健康医疗大数据的存储方式、结构与特点 |
| 实验环境与相关软件的下载、安装与使用 | Python、Anaconda 的下载与安装、使用与测试、运行程序 |
| 健康医疗大数据采集 | 从原始数据库 Oracle Database 11g Enterprise Edition Release 11.2.0.3.0 - 64bit Production,对高血压大数据进行数据采集 |
| 健康医疗大数据清洗和预处理 | 在 Anaconda 环境中使用 Python 对高血压大数据进行数据清洗和数据预处理 |
| 电子病历非结构化文本数据挖掘 | 对高血压电子病历大数据中的字段-“诊断结果”非结构化文本进行分词,对文本中出现的不规范表述使用自定义词典和正则表达式进行处理,并统计关键词词频 |
| 健康医疗大数据可视化 | 根据高血压电子病历大数据中的性别、患病严重程度、各年龄段患病人数、血压、主诉关键词词频、诊断结果生成折线图、散点图、柱状图、词云图等,进行可视化分析 |
| 健康医疗大数据机器学习算法应用 | 对高血压电子病历大数据中的收缩压、舒张压、年龄三个字段值作为自变量使用 KNN 算法、朴素贝叶斯算法、随机森林等机器学习算法进行训练并对高血压患病程度进行预测,比较结果选择最优预测模型 |

(下转第 31 页)

(上接第 27 页)

3.3 大数据导论课程教学方式改革

3.3.1 以案例分析为主、理论讲授为辅,侧重学生实践的课堂教学 根据上述理论教学内容可以看出,大数据导论课程中的概念、定义、知识体系等内容较多,抽象而且晦涩难懂。如果教师在课堂上单纯进行理论讲授,很难调动学生的学习兴趣与积极性。为了加强学生应用实践能力的培养,教师在讲解重点和难点内容之后,对案例进行分析,然后由学生动手完成实验。案例来源于健康医疗领域真实数据,把大数据理论知识、领域知识以及相关技术融合,有助于学生加深对理论的理解,同时也可以提高其实践操作能力。

3.3.2 鼓励借助于 SPOC 进行拓展学习的课外教学 为了弥补在有限的学时内,课堂内容不能完全覆盖所有知识面的弊端,可以借助于校企合作建设的 SPOC 帮助学生进行拓展学习。SPOC 平台涵盖大数据开发技术(Hadoop)、分布式数据仓库(Hive)、数据库技术(NoSQL)、消息发布订阅系统(Kafka)、Scala 编程、大数据处理技术 Spark(基础、中级、高级)、Python 编程、数据分析与挖掘、大数据可视化技术、机器学习(基础、中级、高级)等模块。期末按照平台内学生在线时长、作业完成情况等折算一定分数计入本门课程的最终成绩,以此鼓励学生积极开展课外自学与实践,达到培养自主学习能力的目的。

3.3.3 设计以数据科学基本流程解决现实问题的综合实训 课程伊始,教师提前布置综合实训。要求 2~3 人组成一个团队,按照数据化、数据加工、数据规范化、探索性分析、数据分析与洞见、结果展现以及数据产品提供的数据科学基本流程,运行环境和开发工具不限,团队自行设计主题,在期末前完成 1 个

解决医学领域现实问题的综合案例。期末教师安排课外时间,组织各团队演示答辩,教师和其他组同学按不同权重,共同评分折算分数,计入本门课程的最终成绩。综合实训不仅可以对学生课程学习效果进行检验,而且也可以培养学生的团队协作能力和综合应用能力。

4 总结

开设大数据专业的不同高校立足学科融合,大数据导论的教学内容也不尽相同,各有侧重。大数据导论课程改革只有在明确课程目标的前提下,提高课程建设规划性、系统性,避免随意化、碎片化,才能达到满意的教学效果,从而为应用型人才培养提供帮助。

参考文献:

- [1]李辉,张标.涉农高校数据科学与大数据技术专业人才培养思考[J].高等工程教育研究,2019(5):16-22.
- [2]刘巧红,孙丽萍.医学院校医疗健康大数据人才培养的思考与探索[J].教育教学论坛,2020(12):224-225.
- [3]张祖平.新工科背景下大数据专业导论课程的改革与探索[J].大数据,2018,4(6):38-45.
- [4]朝乐门.数据科学理论与实践[M].北京:清华大学出版社,2019.
- [5]贾蓓.学科融合视角下的数据科学与大数据技术专业人才培养思考[J].大学教育,2020(1):166-169.
- [6]牟延林.普通本科高校转型进程中课程改革的思考[J].中国高教研究,2014(9):84-91.
- [7]覃雄派,陈跃国,李翠平,等.“数据科学”课程群与“数据科学导论”课程建设初探[J].大数据,2018,4(6):19-28.
- [8]李传锋,舒云星,李蒙.地方本科院校转型发展下单片机课程教学改革与实践[J].计算机教育,2017(6):157-160.

收稿日期:2020-09-26;修回日期:2020-10-12

编辑/成森