

新形势下数据挖掘课程的教学改革的研究

张秀琴

(武警工程大学信息工程学院 陕西 西安 710086)

[摘要]数据挖掘课程建设成为数据科学人才培养和学科建设的重要组成部分。传统数据挖掘课程的教学中存在教学内容难懂、滞后,课程教学模式单一,课程考核简单、不科学等问题,本文从合理化教学内容、多样化教学方法、科学化课程考核方式、强化教学保障四方面突破传统数据挖掘课程教学过程中的阻碍。创新求变,激发学生的学习兴趣。

[关键词]数据挖掘课程; 教学内容; 教学模式; 考核方式; 教学保障

[DOI] 10.12252/j.issn.2096-6288.2021.06.2699

一、引言

数据已经成为国家基础性战略资源。2012年联合国发布BigDataforDevelopment:Challenges&Opportunities^[1],详细阐明了数据生态系统,并分析了各国在运用大数据时所面临的机遇和挑战。为迎接大数据时代带来的重大机遇与挑战,我国制定了一系列的发展规划。2015年国务院印发《促进大数据发展行动纲要》,明确指出全面推进我国大数据发展和应用,并提出加快建设数据强国。2016年发布的《中华人民共和国国民经济和社会发展第十三个五年规划纲要》中提出:“把大数据作为基础性战略资源,全面实施促进大数据发展行动,加快推动数据资源共享开放和开发应用,助力产业转型升级和社会治理创新”。工业和信息化部也制定了《大数据产业发展规划(2016-2020年)》,明确指出要加快建设数据强国,为实现制造强国和网络强国提供强大的产业支撑。近年来,国内许多高校和科研院所也在积极响应国家的大数据战略,加大对大数据相关人才的培养力度,开设大数据相关专业^[2]。

数据挖掘成为大数据时代数据科学领域的核心关键技术。开展数据挖掘课程建设成为数据科学人才培养和学科建设的重要组成部分^[3]。在传统教学模式中下,数据挖掘课程普遍采用知识点覆盖的教学思路,侧重知识的传授而轻能力培养。课程从教学内容、方法、手段以及考核方式都在教师设定好的框架中实施完成,这种模式的“套路”化阻碍学生提升解决问题的团队协作能力,不利于培养学生大数据思维。

二、数据挖掘课程教学现状

2.1 教学内容难懂滞后

数据挖掘课程是一门综合性很强、难度系数最高的本科课程之一,高校课堂中,教学依托的教材知识远远滞后于目前的技术前沿发展,尽管教学内容是实际应用的基础,掌握基础内容是进一步发展的前提,但相对于商业应用中的知识教材知识相对滞后,缺乏实用性。作为大数据时代和人工智能领域的新兴课程,数据挖掘课程与计算机其他传统课程相比,缺乏国际最新和最前沿的教学资源,纸质教材更新严重滞后。经典的数据挖掘教材内容涵盖了非常宽广的数据挖掘知识,各算法均与微积分、线性代数、概率等数学、统计理论紧密相关,理论部分是大量的定理和公式推导过程。要求学生具备扎实的数学基础的同时要求由较强的编程与逻辑思维能力艰涩、复杂的课程内容极易滋生学生畏难和厌学情绪,造成学习动力不足,最终影响教学质量^[4]。

大多数数据挖掘课程实验数据资源都是公共数据资源,这些数据比较陈旧,数据量较小,且很多数据集已经过预处理。然而在实际的数据挖掘项目中,数据收集和数据处理的部分要占全部工作量的80%以上。因此,针对这些数据集进行挖掘,很难让学生掌握数据预处理的基本技能。实验教学案例主要是针对单个经典算法的实现,脱离生活实际,也会较为枯燥和晦涩,学生难以理解,缺乏解决实际相关的问题^[5]。单纯的上机实现单个的经典算法,已经很难满足对学生数据分析能力的培养,难以提高学生的综合解决问题和运用知识的能力。此外没有统一的实验平台和实验资源。如果授课教师没有数据挖掘知识的储备,很容易导致学生觉得数据挖掘课程学不到实质的内容^[6]。

2.2 课程教学模式单一

数据挖掘属于多学科交叉课程,涵盖知识面广,课程内容(如图1)复杂且各个章节之间独立性较强,在有限的授课时间,教师主要依据教学计划选择性讲授部分知识点,在这种“以教师为中心”的单向“灌输”且与需求相分离的教学方式,强调了教师在教与学课堂中的主导地位,却忽视了学生才是课堂的真正主体^[7]。在这种“自顶向下”的课程设计之下,学生按照教师既定的教学计划推进学习进度,不仅容易让学生产生学习倦怠,并对课程价值产生质疑,而且缺乏探究性思考和评判性思维训练。此外,采用课后作业和课后实验对教学内容进行补充,但是大多学生未独立思考通过各种途径获取课后作业的答案和实验内容的源代码,教师无法通过课后作业和课后实验检验学生的学习效果^[8]。

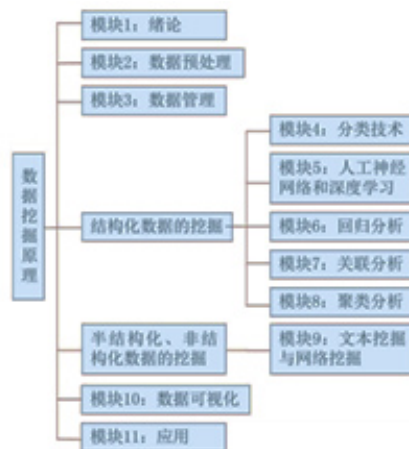


图1. 数据挖掘原理知识点的结构化框架

2.3 课程考核不科学

通常采用平时成绩、期末卷面成绩和期末大作业综合评定的考核方式。这种考核方式刻板单一,重理论轻实践,没有考虑学科之间的差异和特点,没有专门开发基于工程教育认证的课程评价指标体系,也就无法反映出课程教学方法改革的正确导向^[9]。以期末大作业的形式对学生进行评价,则缺乏对学生的过程性监督。由此可见,现有课程在教学内容、实验数据资源及课程评价等方面不尽合理。

三、教学模式改革措施

3.1 合理化教学内容

3.1.1 增添课程思政

在教学中自觉引入思政教育,结合学校特色,课程与案例中有意识地融入了传递精神、不成功决不罢休精神,通过英雄的事迹,传递社会主义核心价值观、社会正能量和正义感等内容,培养具有吃苦耐劳、耐得住寂寞、精益求精的精神。

3.1.2 扩充理论教学案例

数据挖掘课程理论教学的主要目标是使学生掌握数据挖掘的基本概念,了解数据挖掘的定义和功能以及实现数据挖掘的主要步骤和具体方法,掌握数据挖掘的算法。实验主导教学内容,加入的实际案例,并注重与应用单位联系,将相应知识融入到实际案例中,案例将课程中各个章节的内容有机结合起

来,形成完整的知识体系,以实际案例为中心组织教学内容,以问题为核心引出知识点并串联各章节内容。案例的设计还应该紧贴实际需求和行业先进技术,根据教学计划紧扣各章节的重点和难点内容,并通过不同案例的讲解将各个独立章节的内容关联起来。案例设计从竞赛中选取贴近实际问题的案例数据等。从而帮助学生掌握数据挖掘的基本理论和方法,为学习后续课程、开展科学研、进行生产实践打好基础。案例质量的好坏就显得相当重要。根据学生对每个教学组织环节不间断的信息反馈。以及教师对学生掌握知识点的情况评价调整教学资源。

3.1.3 深化实验内容

实验教学的主要目标是使学生能够实现简单的数据挖掘算法编程,了解实现数据挖掘的具体操作。首先,用流行的数据挖掘软件或软件包自己编程逐步实现基础的算法,加深对算法步骤的理解,并对算法的缺陷有一点的了解,尽可能地提出改进的思路。其次,利用理论内容中预先给定的具体案例(案例的数据集是未处理过的),通过互联网资源、开源软件和公开数据集等多种资源开发并完成实验任务。使学生能够实现数据挖掘算法编程从而掌握数据分析和处理的能力。具体的实验过程由合作小组组队一起完成。此外鼓励参与各级各类数据挖掘竞赛,例如国外的Kaggle竞赛网站,国内的数据城堡竞赛平台、天池大数据竞赛平台等。在这些平台上,有很多从企业实际问题出发,延伸而来的数据挖掘赛事。这些赛事目的明确,数据集规范,同时采取学生自主参加、教师辅导的方式,提高学生解决实际问题的能力。

3.2 多样化教学方法

3.2.1 案例驱动激发学习兴趣

以案例为中心,学生为主体,转变传统填鸭式教学为学生自主式、体验式学习的教学模式,引导学生在创设的项目情境中实现“做中学,学中做”。这种教学方法能够激发学生的学习兴趣,并全面提升学生解决实际问题的能力。整个课程内容以多个案例驱动,涉及到案例完成所需的各部分技术,分模块进行讲解、实际操作,将案例融入课程教学的全程。具体包括以下步骤:首先,应该通过案例背景介绍需要实际解决的问题,让学生带着思考进入具体模型和算法的学习中;其次,对问题进行分析并引出解决问题的知识点,让学生对知识点的应用产生初步的认识,激发学生应用知识点解决问题的兴趣;最后,对如何应用知识点解决问题进行讲解,帮助学生掌握课程的重点与难点。通过这种任务驱动的授课方式,学生能够在学学习全程精力集中,抓住学习重点。

3.2.2 分组讨论加强教学互动

教学中对学生分组打破整体教学班级的限制,以3~5人为一组,每个小组推选1名组长,开展互帮互助的合作学习。优势如下,(1)数据挖掘课程中的实际案例整体较为复杂,小组合作有利于培养学生的团队合作能力,促进学生之间相互协调和共同解决问题;(2)教学过程中可以引入对小组的考核,如教师通过随机选择小组的方式参加模型或算法的讨论,对整个小组进行打分,将其作为该小组每个成员的平时成绩,推动小组的整体进步;(3)更有利于对教学内容进行过程性评价,根据整个小组在各个阶段的表现进行学习动机、学习方法、学习目标和学习效果的多重评价,充分展现学生的学习效果和真实水平。为完成项目目标,小组成员相互分工协作,一起研究问题、学习探讨、制定计划,依据能力专长的差异,赋予每个成员不同的角色和职责。有助于培养学生自主终身学习、沟通协调能力和工程职业素养。

3.2.3 混合式教学升华学习内容

课前在对学习目标和学习者特征分析的基础上,利用网络平台为学生提供相关视频资料、知识点动画和其他辅助资料。学生通过网络平台按时完成知识点预习及测验等任务,可查看成绩及推荐答案。也可提出新问题发布到网络平台,师生可参与所有讨论。教师根据测验成绩和议题讨论,可掌握学生预习

完成度、参与度、学习积极性情况和对知识点理解程度等,为不同预习完成度的学生制定不同的课中讨论议题,以期通过个性化教学完成知识点学习。同时根据实际情况不断改进资源和任务。课后结合课前、课中活动学生的表现发布合适的拓展任务和议题至网络平台,完成课后活动,升华学习内容。要求学生对本次课程组织形式等给予反馈。

3.3 科学化课程考核方式

数据挖掘课程教学过程中新设了小组成绩模块,采用小组成绩和个人成绩对学习过程、课程组织等进行有效评价。其中个人成绩包括考勤(占5%)、作业考核(占10%)、笔试考核(50%)。小组成绩包括成果成绩(25%)、个人在组内中的贡献大小(10%)。成果鉴定通过对进度控制、完成质量、汇报表现和团队合作精神等获得。在考核过程中,每个小组成员都被要求至少有一次上台展示、汇报、答辩的机会。这种考核方式重视实践能力、科研能力、表达能力和自我约束能力,有助于培养学生动手实践和独立开展科学研究的能力。使学生能够系统地掌握项目管理的过程、方法和工具,为学生将来从事统计项目研究和管理奠定基础。

此外,可附加竞赛作品的的成绩。借力大学生创新创业计划的激励,促进课程建设,并丰富建设内涵。

3.4 强化教学保障

3.4.1 选优配强师资

教学条件保障课程实施全过程,师资方面教学团队“数据挖掘技术”课程团队教师具有丰富的教学经验,从职称上配备有教授、副教授、讲师和助教,从学历上有博士、研究生,从年龄上采用老、中、青传帮带模式,且以青年教师为主。因青年团队成员在课程建设方面思维超前、锐意创新、不拘一格,教学与实践效果显著。

3.4.2 突出特色著资料

本课程建设突出特色,教学团队编制了数据挖掘教材讲义、辅助阅读等资料。由于国内没有与大纲完全吻合的教材选用,课程团队教师针对大纲内容自编了讲义和实验指导文档,分层次的课程学习模块、逐步拓展深入的实训项目,精心设计多媒体教案、PPT内容,生动醒目的课程演示课件、素材、等保证了教学目标的实现。此外指导学生进行课外资料的选择和阅读,达到了扩充知识面和提高研究能力的目的。

3.4.3 因地创平台

创建了方便的实训环境,配有学习视频,编程语言python学习教程、java自学教程等,平台强化理论学习,实践可以让学生更加深入理解大数据,并且可以提高他们的创新能力与综合素质。提供大数据实训平台作为训练学生处理数据能力的专用平台,使学生能尽早接触大数据分析的实际问题,在学习过程中融会贯通,获得扎实的基础知识

四、总结

数据挖掘课程教学中,教师更新思路,创新求变,激发学生的学习兴趣,带动学生的参与感,提升学生的运用能力。让学生不再是被动接受知识的载体,而是为自己建构知识体系的参与者,不仅提高了学生的学习兴趣 and 积极性,也培养了学生分析问题、独立思考和解决问题的能力及团队合作与创新意识。注重能力培养的同时不忘思政教育和社会正能量的传递。任课教师的角色也从知识的传递者转变为学生学习的促进者。不仅可以提升课程的过程监督,还可以检验人才培养的标准。

参考文献

[1]Big Data for Development: Challenges&Opportunities. <http://www.unglobalpulse.org/sites/default/files/BigDataforDevelopmentUNG10-balPulseJune2012.pdf>, 2012.

[2]全国逾600所高校开设大数据专业北大、中科大、复旦专业实力排名前三. http://news.southcn.com/nfdsb/content/2020-07/29/content_191236851.htm.

[3]陈朝焰,许洪云,刘攀.本科数据挖掘课程教学内容的结构化设计[J].现代计算机,2020(36):77-81+94.

(下转第3075页)

c. 可通过流程中引起批量重工和维修的风险站位和关键参数设定来作为控制点：工夹具的管理及使用、标签重打制程、测试烧码错误、出错货等问题。

d. 生产过程中的非在线操作的流程因容易疏忽，需要列为管控重点：外观不良的机构件维修、维修站位、工单完结的产品，而且所有信息必须要进入生产管理系统进行追溯。

e. 随着工业自动化的导入，人工操作渐渐被自动化设备所取代，但机械自动化因设备变异磨损等原因，加上风险评估因人而异，一个小小的设定错误，会导致整批次的品质异常，故品质抽象和在线品质稽核尤为重要。

2.2 首件的确认

在生产过程中，首件的确认是保证批量生产前最为重要的环节，也是每个制造业必做的流程，但其中有几个关键的环节如下：

2.2.1 首件必须是全流程的检查和确认，不只是产品本身的规格确认，在生产流程中关键流程和参数（CTQ）必须要进行核对，必须要按照制造流程图来进行逐站来确认。

2.2.2 首件最为重要的是非可视化的标准确认，如软件版本、烧录的产品信息、产品的原始密码等等，这些是在现场无法可以看得见的信息，需要仔细核对。

2.2.3 首件的确认时机也是非常重要的，在机种生产转换时，在生产班别转换时，和一切变工发生时，人们往往会忘记在变更时的首件而导致品质问题。

2.2.4 首件参考的文件准确性也非常重要，不只是参考工程师或者是某个单位的信息，需要了解什么是最准确的信息，甚至要看客户原始的要求文档。

3 出货成品检验

在成品检验时，我们对成品的检验也需要做首件，以确保在线确认的首件是100%符合产品规格，对客户定义的要求和

规格做出货前的确认，在成品出货检验时，我们必须要知道产品实际的使用功能，以及在客户端的使用环境，而我们成品的检验也必须最大限度地模拟客户的环境，对产品的所有功能进行检验，很多时候这个关键环节被制造厂给遗忘。

同时，成品检验也要注意工程变工的部分，以防流程和工程变更遗漏和错误导入，成品检验为所有制程的最后确认站位，稍有不仔细，会导致重大客诉和退货，其本身的作业流程和人员管控也是一个重要环节，因为品质人员作业疏忽导致产品缺少附件和功能模式问题也经常发生，所以品质人员对出货品质检验的流程也需要进行失效分析，将所有风险控制工厂内部。

4 结束语

每家的产品和客户有所不同，其关键流程定义也有所不同，在此就不一一介绍，但品质管控的手法大同小异，完全靠劳动密集型的检验方式也已经不适用现代制造业，品质人员应该具备制定关键流程控制点的能力，和改变传统检验方式的大胆创新的思维，要把用最少的人做最好的品质为目标，因地制宜地制定关键管控流程和参数，减少不必要的品质损失，并建立有效的品质控制体系，为企业赢得信赖和永续经营提供有效的保障。

参考文献

[1] 田贻民. SPC在现代制造业品质控制中的应用[J]. 机械制造, 1999(06): 42-43.

[2] 林云卫. 全面的品质管理[J]. 印制电路信息, 2008(10): 48-49.

作者简介:

黄志远(1982 -), 男, 上海松江区, 华南理工大学, 研究生, 主要研究方向工厂品质统计管理方向。

(上接第3043页)

[4] 朱毅, 李云, 强继朋, 袁运浩. 数据挖掘课程教学模式改革与探索[J]. 科教文汇(下旬刊), 2021(06): 108-109.

[5] 贾媚媚, 刘泉, 马晓普, 兰义华, 李贺, 刘金江. 面向新工科的数据挖掘实验课程教学案例设计[J]. 现代计算机, 2021(03): 80-83.

[6] 唐培培, 吴明晖. 基于Kaggle竞赛数据的“数据挖掘技术”课程建设探索与实践研究[J]. 工业和信息化教育, 2021(03): 85-88.

[7] 李会民, 王延仓, 马桂英. “数据挖掘技术”课程教学改革与探索——以北华航天工业学院电子与通信工程专业研究生“数据挖掘技术”课程教学为例[J]. 濮阳职业技术学院学报, 2021, 34(04): 29-32.

[8] 王辉, 张强, 潘俊辉, 王浩畅. 认证导向下基于0BE的数据挖掘课程混合教学模式探索与实践[J]. 微型电脑应用, 2021, 37(04): 18-20.

[9] 林华灵, 丁杰, 卓琳, 陈慧. 关于应用型本科“数据挖掘”课程改革的探索与思考[J]. 福建金融管理干部学院学报, 2020(04): 58-62.

[10] 徐建军, 张国华. 基于Apriori数据挖掘算法的应用与实践[J]. 计算机技术与发展, 2020, 30(04): 206-210.

作者简介:

张秀琴(1988.04)女, 山西省吕梁市, 武警工程大学讲师, 硕士, 大数据, 数据挖掘。

(上接第3054页)

限和整改效果，并将整改方案和措施上报上级机关。

七是执行整改落实。整改方案和措施一经上级机关批准，参评单位应严格执行落实，并确保整改效果，整改完成后参评单位可根据评估指标自我评价，检验整改质效是否执行。

结语

新体制下，教学条件建设绩效评估工作任务重道远，我们要在科学的理论基础上，做好顶层设计、精心筹划、合理安排，完善评估指标，才能不断提高教学条件建设对院校人才培养的贡献率。

参考文献

[1] 康悦, 陈鹏. 军队工程项目绩效评估研究[J]. 基建财务, 2016, 10: 44-46.

[2] 凡. 孙开磊, 曹磊. 对扎实推进教学保障实战化训练落实的思考[J]. 火箭军院校教育, 2017, 79(2): 61-62.

[3] 张文桥, 吴迪, 蒋建明. 军队工程项目绩效评估指标分析[J]. 空军勤务学院学报, 2018, 29(2): 13-16.

[4] 应望江, 李泉英. 高校绩效评价指标体系设计及应用研究[J]. 国家教育行政学院学报, 2010, 02.

作者简介:

张伟涛, 海军勤务学院学员一大队。