

项目带动法在医学数据挖掘教学中的应用*

马 瑾 刘尚辉 娄 岩

(中国医科大学 沈阳 110122)

〔摘要〕 在医学数据挖掘教学过程中实施项目带动法,分析项目带动法的意义,从数据收集、清洗和挖掘几方面具体介绍教学过程,实现理论教学和实践训练有机结合,创造以学生为主体、以项目为驱动、以实践成果为考核目标的新型教学方式。

〔关键词〕 项目带动法;数据挖掘;计算机教学

〔中图分类号〕 R-056 〔文献标识码〕 A 〔DOI〕 10.3969/j.issn.1673-6036.2016.02.020

Application of Project-driven Method in the Medical Data Mining Teaching MA Jin, LIU Shang-hui, LOU Yan, China Medical University, Shenyang 110122, China

〔Abstract〕 Project-driven method is applied in the medical data mining teaching, the paper analyzes the significance of project-driven method, specifically introduces the teaching process from the aspects of data collection, cleansing and mining, to achieve the organic combination of theoretical teaching and practical training, create a new teaching method that takes students as theme, takes project as driving force and takes practical results as the examining goal.

〔Keywords〕 Project-driven method; Data mining; Computer teaching

1 引言

在以往的医学数据挖掘教学中,理论与实践的关系仅仅是相互验证,学生在实践课上操作老师理论课上给出的案例,得到与理论教学一致的结论即可,这样大大束缚了学生的思考能力,更谈不上锻炼拓展性思维的能力。而医学数据挖掘的教学重点

应该侧重于数据挖掘技术在医学领域的应用,其教学目的是使学生通过数据的获取、存储、清洗及分析等计算机技术来研究和解释医学中所发生的现象和遇到的问题并解决这些问题^[1]。教师在教学中应该培养学生创新与自主学习的能力,新的教育模式将项目与团队合作的思想引入课程教学体系,实现人才培养目标^[2]。以下结合中国医科大学教研室从事医学数据挖掘教学的经验,以医学生研究妇科炎症的致病相关因素调查项目为例,介绍将项目带动法运用于这门课程教学的实践。

2 应用项目带动法的意义

项目带动法教学模式来源于建构主义学习理论,即教师与学生共同通过讨论研究并实施完成一个具体的项目工作而进行的教学活动,是一种“以

〔修回日期〕 2015-11-11

〔作者简介〕 马瑾,副教授,发表论文3篇,参编教材6部;通讯作者:娄岩,教授。

〔基金项目〕 中国医科大学“十二五”第二批医学教育科学研究立项课题(项目编号:YDJK2012029);辽宁省教育厅教改课题(项目编号:YDJK2013006)。

学生为主体”的教学方式。项目带动法融合“教、学、做”3 方面,寓教于做,尊重学生,充分发掘学生的自身潜能,培养和提高学生的实践操作能力、分析解决问题的能力,有效促进学生对知识的理解与深化^[3]。在数据挖掘课程教学中实施的项目带动法,是指对实践性与操作性比较强的内容,由任课教师组织学生以团队形式参与,有针对性地选择可以贯穿课程始终、与社会现实密切相关的实践项目,指导学生分组随同理论教学的进程,同步展开实践,在课程结束的同时,取得实践项目成果,以项目成果为主来考核学生学习效果的一种教学模式^[4]。以项目带动法构建医学数据挖掘课程的意义在于^[5]:(1) 以与医学结合的实际项目为驱动,促进医学生的人才培养从面向理论转为面向应用。(2) 以培养医学生的实践能力为主线,引导学生组建团队,有利于学生之间的沟通能力及团队合作能力的培养。(3) 以学有所用为前提,有利于学生尽早把课堂知识转化为实际能力,把学到的医学基础知识转化为实践动力,避免学生丧失学习兴趣。充分调动学生学习的积极性与主动性。

3 教学过程

3.1 概述

数据挖掘课程第 1 章概论结束后,教师就给学生布置将要进行的调查项目,让学生带着项目学习。围绕这一项目,教师指导学生分组讨论如何完成。可以由任课教师提出几个实践题目的方向,让学生们自选题目或提出新的题目,按学生自己选定的感兴趣的内容将每 4~5 人形成项目组,每组选出一名组长。以调查妇科门诊阴道炎致病相关因素为例,在课程中以此项目为主线,将需要掌握的知识点贯穿于项目的各个模块中,如数据的收集、清洗,分析数据应该用何种模型等。教师按学时数规定好各个模块完成的时间,制成项目进展情况一览表,表结构一般包含项目名称、项目成员、成员主要工作内容、项目进度以及项目完成情况等内容。教学过程中各个项目组组长按规定的时间在表格中记录项目进展情况及组员的完成情况,督促每个组

员在规定的学时数内完成项目内容。这样保证每个学生都积极参与到项目的实践过程中,从而对各个知识要点做到从熟悉到掌握。应用项目带动法,将数据挖掘课程的内容划分为数据收集、数据清洗及数据挖掘 3 大部分。

3.2 数据收集

数据收集是根据确定的数据分析对象抽象出在数据分析中所需要的特征信息,然后选择合适的信息收集方法,将收集到的信息存入数据库。在此学习阶段,教师引导学生确定调查内容,以什么方式取得需要的详实的调查资料,以及如何拟定合理的调查表,其中的重点是通过不同方案的对比,设计出可行的调查问卷,收集可靠的数据资料^[7]。有的团队充分利用互联网的资源进行数据收集,有的团队对身边的同学进行访问调查,还有的团队深入医院科室对患者或患者家属进行问卷调查。在调查妇科门诊阴道炎致病相关因素项目中,该小组的成员多次前往医院妇科门诊化验室,与工作人员进行沟通,收集纸质数据,再转化为电子表格。学生在此阶段已经对数据挖掘这门课程产生了极大的兴趣,所有的资料都是通过努力获得的真实数据,学习的积极性被充分调动起来。在此过程中,还要让学生理解硬件技术不可缺少,对于海量数据,选择一个合适的数据仓库是至关重要的。

3.3 数据清洗

数据清洗的目的是为挖掘提供准确而有效的数据,提高挖掘效率。收集到数据库中的数据有一些是不完整的、含噪声的,并且是不一致的,因此需要进行数据清洗,将完整、正确、一致的数据信息存入数据仓库中。否则挖掘的结果会差强人意。简单来说,数据清洗就是把不符合要求的数据按照一定的规则过滤掉^[8]。各个项目组得到大量的原始资料后,由教师指导学生利用计算机软件对数据进行清洗,比较简单的方法是采用 Excel 对收集的资料分门别类进行汇总,然后对不同类型的数据进行统一的表示,如对标志型的数据可以分别以 0 或 1 来表示,对连续型的数据可以划分成集合表示。而错

误数据有可能是学生在录入的时候造成的,如数值数据输成全角数字、字符或日期格式不正确、日期越界等。这一类数据可以通过写 SQL 语句的方式找出来,然后要求学生统一修改。在此阶段工作量较大,学生对大量数据进行分组处理,分析问题、解决问题的能力得到了充分的锻炼,而且增强了团队合作意识。

3.4 数据挖掘

数据挖掘一般是指从大量的数据中自动搜索隐藏在其中的有着特殊关系信息的过程,数据挖掘的任务有关联分析、聚类分析、分类分析、异常分析、特异群组分析和演变分析等等^[9]。在此阶段教师指导学生应用数据挖掘软件对清洗后的数据进行数据挖掘操作,根据数据内容选取操作任务,对不同类型的数据选择不同的算法,再根据算法的不同对数据源、字段及记录进行有效性设置,根据每个项目组的数据内容构建模型。以聚类分析为例,它是一种探索性的分析,在分类的过程中,人们不必事先给出一个分类的标准,聚类分析能够从样本数据出发,自动进行分类^[10]。例如在调查妇科门诊阴道炎致病相关因素项目中,学生可以根据众多相关因素聚类的结果构建出决策树,进行不同年龄段妇科常见阴道炎发病情况的研究,探讨使用数据挖掘技术分析阴道炎发病情况,为临床提供参考依据。构建决策树是数据挖掘技术中的一种,部分学生使用 C5.0 算法构建决策树模型,C5.0 算法是以信息熵的下降速度为选取测试属性的标准,直到生成的决策树能完美分类训练样例。构建出的决策树可以用于分析数据,同样也可以用来做预测。在 C5.0 算法中,通过类型节点设置年龄变量的方向为输出,其他致病因素变量的方向为输入得到一些结论。如 21-30 年龄组念珠菌性阴道炎和细菌性阴道炎的混合性感染比例较高;31-40 年龄组感染细菌性阴道炎程度较重;近几年滴虫性阴道炎发病较少;年岁较大的患者更容易患细菌性阴道炎等。学生再根据数据挖掘的结果查阅相关文献对挖掘出的结果进行验证,21-30 年龄组易患混合性感染是由于性行为越来越开放,这一时期的女性激素水平

和阴道内环境容易导致细菌感染,并且此时期妇女处于性活跃期,容易因为直接传染而导致细菌性阴道炎的发生。在数据挖掘的过程中已经将课程的内容融入到所有的操作中,同时按规定的进度进行学习,使所有项目组都完成操作。

4 教学考核

传统教学的考核只是将理论考试成绩作为评价标准,而项目带动教学法不同于传统教学方法,单一的期末理论试卷不能全面反映学生对这门课程的学习与掌握情况。改革考核方式,将学生的学习态度、合作性、发展性、探究性、课程设计结果等放入形成性评价中,而不只是课程结束时的笔试考核。项目带动法的考核过程与项目的完成情况同步,在每一阶段都有相应的学生自我评价和教师评价记录。考核时还要注意不仅要考虑团队项目的完成情况,还要评价学生个体的表现。最后,教师对每个项目组提交的结论进行验收与评价,学生提交项目开发相关报告。教师按照评分标准给出成绩,对优秀作品进行展示与点评^[11]。

5 结语

项目带动法最重要的是以学生为主,引导学生参与项目进展的全过程。在实施项目带动法进行医学数据挖掘教学的过程中要突出与医学数据密切结合的原则,将以往教师枯燥的理论讲解和复杂的公式推导改为有针对性地介绍实际所用的知识和理论,使学生在有限的课时里学到与医学实践密切相关的知识。此外教师也在此过程中积累了丰富的医学教学案例,加深了教师自身对这门课程的理解,而且能从学生的作品中获得更多的医学知识,将学生作品中优秀的项目作为下次教学中的案例。

教学实践表明,项目带动法充分调动了医学生的学习兴趣,提高了实践技能与自学创新能力。在调查妇科门诊阴道炎致病相关因素项目中,学生对妇科阴道炎的致病相关因素进行了主动调查学习,了解到致病因素与年龄、季节、饮食等多方面有密

切相关性，用自己收集的数据作为基础进行探讨研究，甚至开始在本科阶段尝试撰写流行病学方面的论文。然而把项目带动法应用于教学中仍存在一些不足，如由于过分追求项目的完整性和实用性，容易疏漏理论基础，忽视学生基础知识的掌握情况；实际操作过程中，经常根据项目的完成情况对整个项目组进行评价，但对学生的个人能力缺乏有效的过程监控。在今后的教学中教师要不断积累经验，科学完善考核评价标准，使项目带动法在教学中发挥更大的作用。

参考文献

- 1 吴辉群, 蒋葵, 姚蓓, 等. 基于项目的学习模式在医学信息学专业创新实践教学中的思考 [J]. 医学信息学杂志, 2011, (11): 83-86.
- 2 王庭之. “项目带动式”课程体系在卫生信息管理专业中的实践——来自盐城卫生职业技术学院的课程改革 [J]. 职教论坛, 2010, (33): 17-21

- 3 徐琦. 医学信息管理专业学生计算思维能力的培养 [J]. 医学信息学杂志, 2011, (11): 90-93.
- 4 李倩. “项目带动教学”模式在《静态网页设计》课程中的应用 [J]. 科技风, 2011, (7): 193-194
- 5 蒋勇清. 论数据清洗对信息检索质量的影响及清洗方法 [J]. 中国索引, 2012, (1): 35-39.
- 6 王众托. 教育数据挖掘与知识发现 [J]. 基础教育课程, 2013, (1): 23-24.
- 7 陈少晖. 项目带动课程教学模式中的课程交叉渗透教学 [J]. 统计教育, 2005, (3): 48-49.
- 8 黄云玲. “项目带动, 工学结合”人才培养模式的探索与实践 [J]. 当代职业教育, 2012, (11): 34-35.
- 9 钟小敏. 项目带动课程教学模式的实践探讨 [J]. 广西教育, 2011, (9): 30-31.
- 10 张丽. 高校“项目带动式”实践教学模式探讨 [J]. 内江科技, 2010, (8): 20-21.
- 11 罗凌. 基于项目驱动的“ASP.NET 程序设计”课程创新教学法研究 [J]. 计算机教育, 2009, (13): 100-102.

(上接第 73 页)

- 2 董家鸿. 迈向“精准医学”时代 [N]. 光明日报, 2015-07-04 (010).
- 3 张莉萍. 学科馆员与科研课题跟踪服务的若干思考 [J]. 现代情报, 2006, (7): 32-34.
- 4 戚敏. 面向高校用户的嵌入式学科化服务模式探究——以广西科技大学为例 [J]. 图书馆界, 2015 (2): 13-16.
- 5 医脉通. 《2015 年 ADA 糖尿病医学诊疗标准》更新要点 [EB/OL]. [2014-12-25]. [http://](http://guide.medlive.cn/guidelinesub/2174)

guide.medlive.cn/guidelinesub/2174.

- 6 吴若琪. 肿瘤精准医疗联盟成立 [N]. 中国医药报, 2015-07-01 (006).
- 7 黄孝群. 论高校图书馆的个性化信息服务 [J]. 山西财经大学学报, 2011, (4): 184.
- 8 邱坚, 陈怡, 俞国雄. 医院图书馆开展学科化服务中需要思考的问题 [J]. 中华医学图书情报杂志, 2012, (21) 8: 9-11.

欢迎订阅 欢迎赐稿