

数据挖掘技术在高校教学质量评价中的应用研究

许瑞斌

泉州信息工程学院

DOI:10.12238/er.v4i5.3927

[摘要] 教学质量评价是每一个高校在教学管理过程中都比较重视的环节,因为它对提高整体教学水平起着至关重要的作用。本文利用数据挖掘技术对收集到的评价数据源进行清理和转换,采用关联规则 Apriori 算法对处理后的数据建模,从而获得师生信息与教学质量评价结果之间蕴含的有实用价值的关联,为高校更好地进行教学管理提供决策服务。

[关键词] 数据挖掘技术; 教学质量评价; Apriori 算法; 关联规则

中图分类号: G642.0 **文献标识码:** A

Application of Data Mining Technology in College Teaching Quality Evaluation

Ruibin Xu

Quanzhou Institute of Information Engineering

[Abstract] Teaching quality evaluation is a link that every college attaches more importance to in the process of teaching management, because it plays a crucial role in improving the overall teaching level. This paper uses data mining technology to clean up and transform the collected evaluation data sources, and model the processed data with the correlation rule Apriori algorithm, so as to obtain the practical correlation between teacher and student information and teaching quality evaluation results, and provide decision-making services for universities to better conduct teaching management.

[Key words] Data mining technology; teaching quality evaluation; Apriori algorithm; and association rules

引言

随着教育改革不断地深入发展,教学质量评价更是人们非常关心的话题。高校教务部门的管理者们都非常重视教学质量评价,基本上都把教学质量评价提升到新的战略高度,从备课充分与否、上课效果如何、课后辅导答疑情况、课程思政等各个教学环节进行不间断地评价。传统的教学质量评价方式多种多样,比如有学生评教、师生互评、督导听课、领导谈话等等,再根据各种评价方式的重要程度不同设置不同的权值,最后进行加权求和,从而得到被评价教师的得分,学校以此评价结果作为对教师进行考核的重要依据。然而在上述的教学质量评价方式中,或多或少存在评价内容单一且机械化、评价指标设置不合理、权值分配不具科学性等诸多问题,如果这些问题不解决,教学评价就有失其本

身应有的价值取向。随着数据挖掘技术的不断发展,探讨如何合理地利用数据挖掘技术对在不同方式的教学质量评价过程中所获取的数据进行重新梳理,并寻找出大量数据之间隐藏的有实际应用价值关联信息,从而提高教学质量提供理论支持,将会是近年来教育教学改革的研究重点和新方向,并赋予新的科学内涵。

本研究先是利用数据挖掘技术从教学质量评价原始数据中采集到的数据源进行清理和转换,接着将这些数据采用数据挖掘技术的 Apriori 算法进行建模,最后得出师生信息与教学质量评价结果之间的隐含联系,为高校更好地进行教学管理提供决策服务。

1 关键技术与算法

1.1 数据挖掘

数据挖掘(Data mining)又称为知

识发现,它是通过分析大量的、模糊的、随机的数据,并从中寻找其事先并未发觉的、具有潜在应用价值的规律。数据挖掘通常要做数据准备、规律寻找和规律表示三个步骤。①数据准备是收集大量的、真实的、含噪声的数据,并从中获取用户需要的相应数据来构造用于数据挖掘的数据集;②规律寻找是用数据挖掘算法将收集到的数据集进行清洗和转换,从而找出数据对象之间隐含的规律;③规律表示是用人们可听可看、可理解、可使用的方式来表示所找到的规律,从而为人们的应用决策提供帮助。

1.2 Apriori 算法及分析

关联规则挖掘目前比较流行,也是数据挖掘领域的重点研究方向,它可以反映数据对象之间的函数依赖关系,比如一个数据对象与另一个数据对象之间存在关联,则一个数据对象的后继发展

表1 部分教学评价信息表

序号	教工号	课程代码	职称	年龄	评教人数	评价结果
01	T04003	C0001	教授	53	75	95.00
02	T07001	C0005	讲师	35	135	91.21
03	T05023	C0004	讲师	31	268	71.25
04	T08211	C0015	助教	25	90	92.15
05	T10245	C0054	副教授	41	20	98.21
06	T14196	C0024	教授	60	35	94.03
07	T09518	C0009	讲师	36	58	86.45

结果可以通过另一个数据对象来进行预测,从而提高两个数据对象的共同发展,这也是国家领导人提到的“共同体”的理念。其中最具代表性就是沃尔玛超市中啤酒与尿布的故事,只要我们找到了对象之间的这种隐含的关联,我们将创造更多的惊喜。本文所应用的算法就是目前最经典的算法之一,Apriori算法。

Apriori算法是用在两个相辅相成的阶段中寻找频繁项集的递推方式来实现的:

1.2.1通过不断地迭代循环,找出所有频繁项集,且它们的支持度一定不能小于用户设定的最小支持度。

1.2.2由找出的满足条件的频繁项集继续去生成置信度,且它们一定不能小于用户设定的最小置信度,从而形成所需的关联规则。

2 关联规则挖掘在教学质量评价中的应用研究

本文将关联规则挖掘技术进行深入研究,探索在高校教学质量评价中的具体应用,从中发现师生信息与教学质量评价结果之间的存在的隐含的关联,为教务部门的日常教学管理提供理论依据,从而进一步提高高校教学质量。

2.1 数据处理与数据集成

为了提高挖掘结果的可靠性,就必须要去掉收集到的海量数据中的噪声和不一致。本研究收集了泉州信息工程学院软件学院2016-2020学年学生的评教结果及近几年的师资汇总表和学生信息表等作为数据源,共有8000多项数据,然后对收集来的数据进行不完整性与冗余性的消除处理及敏感信息做了简化处理,接着采用数据集成方法将上述处理过的数据进行集成,集成后的部分数据如表1所示。

2.2 数据转换

集成后的数据相对还是比较粗糙的,为了提高数据挖掘信息数据的实用性,要对经过集成后的数据进行必要的转换,一般采用数据泛化和数据离散化两种不同方法的混合使用,将原本较为粗糙的集成数据进行转换并映射到指定的特定区间内或将其化分为不同的类别标识。

具体的转化方法:将表1中的学生评教人数泛化为四个区间,即N1[20, 60], N2[61, 120], N3[121, 200], N4[201, 300];教师的年龄泛化为四个区间,即A1[25, 33], A2[34, 41], A3[42, 49], A4[50, 60];教师的职称离散化为四个类型标识,即L1助教, L2讲师, L3副教授, L4教授;评价结果泛化为四个区间,即S1[75, 84], S2[85, 90], S3[91, 94], S4[95, 98]。

2.3 关联规划挖掘分析

将转换后的数据采用Apriori算法进行数据建模,根据用户需求设置最小支持度和最小置信度,寻找符合算法条件的频繁项集,分析数据源表中学生评教人数、教师年龄和教师职称等信息与评价结果之间的内在联系。

表2 学生评教人数和评价结果的频繁项集

输入项	目标项	支持度S(%)	置信度C(%)
N4	S2	11.87	45.00
N1	S3	31.13	54.22

2.3.1学生评教人数与评价结果之间的联系。利用Apriori算法,以学生评教人数作为输入项,评价结果作为目标项,对它们进行关联规则分析,得到学生评教人数和评价结果的支持度和置信度。根据用户需求,我们设定了两个阈值,即最小支持度和最小置信度,它们的值分别为11%和35%,从而得到两者之间的频繁项集,如表2所示。

由表2可知,学生评教人数在200-300人时,评价结果处于85-90分的支持度为11.87%,置信度为45.00%;学生人数在20-60人时,评价结果为91-94分的支持度为31.13%,置信度为54.22%。从数据上来查看,评教人数较少,即学生人数较少的班级,分析得到的支持度和置信度都比人数多的班级来得更高,所以人数少上课效果更好。

2.3.2教师年龄与评价结果之间的关系。利用Apriori算法,以教师年龄作为输入项,评价结果作为目标项,对它们进行关联规则分析,得到教师年龄和评价结果的支持度和置信度。根据用户需求,我们设定了两个阈值,即最小支持度和最小置信度,它们的值分别为12%和42%,从而得到两者之间的频繁项集,如表3所示。

表3 教师年龄和评价结果的频繁项集

输入项	目标项	支持度S(%)	置信度C(%)
A1	S2	16.30	40.33
A3	S4	22.39	53.15

由表3可知,当教师年龄处于25-33岁时,评价结果处于85-90分的支持度为16.30%,置信度为40.33%;当教师年龄处于42-49岁时,评价结果处于95-98分的支持度为22.39%,置信度为53.15%。从数据上来查看,教师年龄较大,分析得到的支持度和置信度都比教师年龄较小来得更高,所以教师年龄较大效果更好。

2.3.3教师职称与评价结果之间的关系。利用Apriori算法,以教师职称作为输入项,评价结果作为目标项,对它们进行关联规则分析,得到教师职称与评价结果的支持度和置信度。根据用户需求,我们设定了两个阈值,即最小支持度和最小置信度,它们的值分别为16%和40%,从而得到两者之间的频繁项集,如表4所示。

表4 教师职称和评价结果的频繁项集

输入项	目标项	支持度S(%)	置信度C(%)
L2	S3	18.17	41.00
L3	S3	20.91	42.00

由表4可知,当教师职称是讲师时,评价结果在91-94分的支持度为18.17%,

置信度为41.00%;教师职称称为副教授时,评价结果在91-94分的支持度分别是20.91%和42.00%。从数据上来查看,教师职称不一样时,分析得到的支持度和置信度都差不多,所以教师职称对教学效果影响不大。

3 研究结论

根据以上数据挖掘关联规则分析,得到以下三个结论:

(1) 学生评教人数与评价结果之间呈现反比关系。从本研究中要可以发现班级学生人数较少时,即学生评教人数不多时,评价结果较好,反之当学生人数达到一定数量的话,则评价结果就较差。显然,学生评教人数对教学质量的提高有着至关重要的作用,学生人数超过一定数量就会影响教师的积极性和学生的学习效率,毕竟学生人数多了,但是时间是固定的,这样肯定会造成相当一部分学生无法得到老师的辅导,从而降低教学效率。由此可见,要想提高整体教学质量,班级的学生人数一定要有所控制,不能太多,就像现在比较流行的翻转课堂,让教师能够有更多的时间进行课前的备课、课中的授课和互动、课后的答疑和辅导,学生也可以在轻松愉快的环境中学到更多的知识,学校应尽量避免因授课人数过多而造成教学质量的不高。

(2) 教师年龄对评价结果的好坏也有较大的关联。从本研究中可以发现年龄较大的教师,评价结果为高分的概率较大,即评价结果较好;年龄较小的教师,评价结果的分数较低。由此可见,学校可以充分发挥校内年龄较大教师的优势,让他们进行传帮带,形成一种积极向上的学习氛围,以期尽快提高青年教师的教学水平和业务能力。研究表明年龄对

于教师而言具有天然优势,他们所授课程被评价为好的置信度比起青年教师来说都普遍较高。不可否认,教师年龄越大,教学水平就越高,教学效果较好,所以学生对年龄较大的教师的教学水平的认可度就相对较高。所以“以老带新”的导师制是完全可行的。高校的领导者们可以以本研究结果加强探讨力度,并制定出有利于青年教师成长的新的培养模式,提出新的教改方案,让青年教师在年龄较大的教师的带领下不断地提高自我,超载自我。学无止境,相信在这样的模式下,教师的教学水平和教学质量都会有较大的提升。

(3) 教师职称与评价结果之间相关性不大。职称为讲师和副教授的老师,他们的所授课程,学生的评价结果的支持度的置信度从表3可以明显看出来,两者之间并无多大的不同,即无论职称如何,都不会影响到教学质量,学生的评价也都差不多。所以建议学校不用过多的关注教师的职称,提高教学质量的本质是教师教得好,学生学得好,学校应将更多的精力放在年青老师的专业素质和能力水平的培养上,而不是片面地追求高职称的教师,以学生为本,一切从学生是否能快乐的学习出发,提高教学质量那就可能是顺理成章的了。

4 结束语

教学质量就像列车的控制系统,关系着高校优质发展,而教学质量评价更是保障该系统永不停歇的重要手段。各个高校的领导及管理者们也是时刻不敢放松,对教学质量评价的把控也是越来越严格,因为教学质量不仅是关乎一所高校办学水平的高低,也是关系到每一个大学生的前途与命运。本研究采用数

据挖掘技术中的关联规则对高校教学质量评价数据进行挖掘、整理,得出的结论行之有效,为高校的教学管理、教师成长等管理工作提供借鉴。随着数据挖掘技术不断地发展,如何将它很好地运用在高校教学质量评价当中去是一项新课题,我们应该不断挖掘内涵、拓展其应用范围,让它在教育改革的浪潮中发挥更大的作用、提供更多有用的决策信息,助力高校教学水平更上一层楼。

【课题项目】

项目1:福建省2018年本科高校教育教学改革研究项目《软件专业“双创型”人才培养模式改革实践》编号:FBJG20180122。

项目2:2018年福建省中青年教育教研项目《数据挖掘技术在构建高校教学质量评价体系的应用研究》编号:JT180707。

【参考文献】

[1]王平,邱腾雄.高职院校专业教学质量评价体系的构建[J].工程技术研究,2018,(06):236-238+241.

[2]赵倩.基于数据挖掘的教学质量评价体系研究[J].电脑知识与技术:术交流,2017,13(28):10-11.

[3]邵峰晶,于忠清,王金龙,等.数据挖掘原理与算法[M].北京:科学出版社,2016:90-95.

[4]祁瑞丽,郭学涛,孟军英.数据挖掘技术在高校教学评价中的应用研究[J].石家庄学院学报,2020,22(03):65-68.

[5]王忠洋.教育大数据的应用与挑战分析[J].无线互联科技,2019,(13)95-96.

作者简介:

许瑞斌(1979--),男,汉族,福建泉州人,硕士,讲师,从事数据库、数据挖掘研究。