

实时双向教学评价体系研究

孙 庚^{1,2}, 冯艳红^{1,2}, 于 红^{1,2}, 史鹏辉^{1,2}

(1.大连海洋大学信息工程学院, 辽宁 大连 116023;

2.辽宁省海洋信息技术重点实验室, 辽宁 大连 116023)

摘 要:针对目前教学评价体系中存在的评价内容和评价方式单一、评价行为严重滞后等问题,提出了一套实时双向教学评价体系。为保证实时性,划分了详尽的评价阶段,为学生和教师的不同评价阶段设置了不同的评价内容,并给出获取评价数据的平台和软件架构,最终给出定量评价方法和定性评价方法。对于定量评价方法,给出了基于相关系数的抽象算法模型,对于定性评价方法,给出了基于自然语言处理技术的软件架构。利用本文给出的抽象模型和思路,可实现一套实时的、双向的、科学的、客观的评价体系,进而提升教学效果。

关键词:教学评价体系;评价内容;评价阶段;软件架构;自然语言处理

中图分类号: TP311 **文献标识码:** A

Research on the Real-time Two-Way Teaching Evaluation System

SUN Geng^{1,2}, FENG Yanhong^{1,2}, YU Hong^{1,2}, SHI Penghui^{1,2}

(1.College of Information Engineering, Dalian Ocean University, Dalian 116023, China;

2.Key Laboratory of Marine Information Technology of Liaoning Province, Dalian 116023, China)

Abstract: Aiming at the problems existing in the current teaching evaluation system, such as the simplicity of content and limited mode of evaluation and the serious lag of evaluation behavior, the paper proposes a real-time two-way teaching evaluation system. In order to ensure real-time performance, detailed evaluation stages are divided, different evaluation contents are set up for different evaluation stages for students and teachers, the platform and software architecture for obtaining evaluation data are provided, and finally quantitative and qualitative evaluation methods are produced. For the quantitative evaluation method, the abstract algorithm model based on correlation coefficient is introduced, and for the qualitative evaluation method, the software architecture based on natural language processing technology is proposed. Using the abstract model and ideas given in this paper, a real-time, two-way, scientific and objective evaluation system can be realized, and the teaching effect can be improved.

Keywords: teaching evaluation system; evaluation contents; evaluation stages; software architecture; natural language processing

1 引言(Introduction)

目前的教学评价一般采用了简单的双向教学评价^[1,2]。申瑶等人提出在医学教学中采用双向评价法,促进了医学教学效果的提升^[3]。江丽丽利用数据挖掘技术研究了双向教学评价问题,给出解决方案^[4]。徐蕾在博士论文中探讨了评价对教学质量保证的作用^[5]。刘红以英语课程为例进行双向教学评价方法的实践^[6]。在这些双向教学评价方法中,对于教师对学生的评价,一般都是采用平时的形成性考核和期末终结性考核一次性评价。学生对教师的评价,通常是学生在期末对教师进行评教,导致无法准确而合理的评价整个教学过程。所以,传统的评价体系主要存在两个问题。

第一,评价内容和评价方式单一,评价方法不够合理。传统的评价内容一般由试卷体现,而合理的评价内容不但包括试卷上所体现出来的知识点的掌握程度,还应包括体现学生其他素质的评价。目前的评价方式都是以考试分数定论结果,掩盖了学生的学习效果和教师讲授效果的目的,扭曲了评价功能,违背了教育的初衷。

第二,评价行为严重滞后,失去了评价的意义。传统的评价行为通常都是在学期结束时进行。王志波提出实时评价的思路^[7],但没有考虑双向的问题。对于一门课程而言,教学过程贯穿整个学年;对于一名学生而言,教学过程贯穿整个大学四年;对于一名教师而言,教学过程贯穿整个职业

生涯。而传统的评价行为的评价时间点过少。讲授的效果和学习的效果在教学过程中已经得以体现，而教学评价却是在教学过程结束后才进行，所以这种滞后的评价失去了评价的意义。对于学生来说，错过了最佳的学习机会，对于教师来说，错过了最佳的更正讲授方法和内容的机会。

所以针对目前的教学评价体系中存在的两个主要问题，本文提出一种实时的、双向的教学评价体系，并给出评价体系的软件架构。对双向教学评价的评价阶段、评价内容、评价方法三个内容进行改革。

2 实时双向教学评价体系软件架构(Software architecture of the real-time two-way teaching evaluation system)

本文提出的实时双向教学评价体系的软件架构如图1所示。

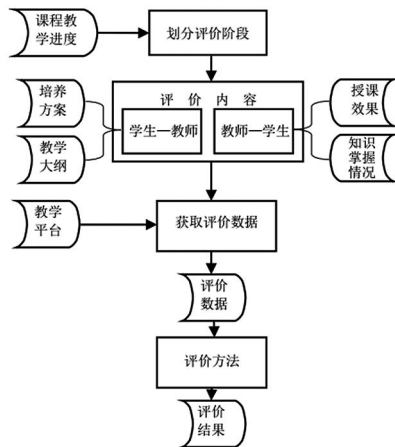


图1 实时双向教学评价体系软件架构

Fig.1 Software architecture of the real-time two-way teaching evaluation system

为解决评价滞后的问题，本文根据课程教学进度划分评价阶段，在各个教学环节设置评价结点，及时评价。为解决评价粒度问题，根据培养方案的要求和课程教学大纲的目标，划分课程的知识点，根据不同知识点、对学生各种能力培养等方面，制定教师对学生的评价内容。根据授课效果、学生的知识点掌握情况等，制定学生对教师的评价内容。根据评价内容，确定需要实时获取和存储的评价数据，本文利用已有的教学平台获取和存储评价数据。最后，依据评价数据，采用机器学习和自然语言处理技术对数据进行分析 and 理解，给出科学的评价结果。

3 评价阶段的划分(Division of evaluation stages)

为保证实时性，教学评价分多个阶段进行，以获得最优的评价效果。主要研究在哪些次课、哪些周等时间点设置评价过程才能提升教学评价的效果。评价时间点设置不宜过少，否则难以保证实时性，且不宜过多，以免增加学生的负担和教师的工作量。每次课、每周、每月、每学期都设置评价时间节点，让一门课程的评价贯穿始终。每次课结束，老师和学生互相反馈结果，提高评价的及时性；每一个时间节点结束，也进行相应的评价。本文以《Java语言程序设计》课程为例，阐述评价阶段划分的思路，阶段划分如图2所示。



图2 评价阶段划分

Fig.2 Division of evaluation stages

《Java语言程序设计》共计64学时，根据《课程教学进度表》，每周2次课，每次课两学时。图2中1t中的t(time)表示次评，1t表示第一次次评，图2中|表示次评，共计有16次次评，在每次下课时对每次课程内容进行评价，这类评价的特点是评价内容粒度细，评价所消耗的时间短，学生和教师通过移动教学平台在课间就可完成。图2中1w的w(week)表示周评，1w表示第一次周评，图2中|表示周评，共计12次周评，在每周的第二次课进行评价，主要针对本周内容进行评价，这类评价粒度稍高于次评，消耗的时间也略高一些。图2中1M(Month)表示月评，1M表示第一次月评，图2中|表示月评，共计3次月评，在每四周结束时进行，主要针对本月内容进行评价，这类评价的粒度高于周评，消耗的时间也更多一些。图2中1e(end)的e表示末评，图2中|表示末评，末评只有一次，针对大纲所有内容进行评价，这种阶段的评价就是期末考试，对学生的评价时间为100分钟，即考试时间，对老师的评价为期末评教过程。

4 评价内容选取(Selection of evaluation contents)

不同的评价阶段需要选取不同的评价内容，才能更合理的给出评价结果，而且老师对学生和学生对老师的评价内容完全不同，本文以《Java语言程序设计》为例，评价内容如表1所列，读者在实践中使用本文评价体系时，评价内容不限于表1所列。

表1 不同评价阶段的评价内容

Tab.1 Evaluation contents in different evaluation stages

阶段	次评	周评	月评	末评	次评	周评	月评	末评
对象	学生—老师				老师—学生			
数据类型	分类、不同数据类型的表达能力和范围、数值类型的运算、不同数据类型间的转换	本周两次课的知识	本周四次课的知识		学生当次课提出的疑问的解答、对当次课知识点掌握程度的评价和统计	分组考查	月考	
评价内容	教学态度	是否改进态度	期末评教		学生的学习习惯	是否改进习惯	知识运用能力	期末考试
	教学方法	是否改进方法	月末评教		学生的学习态度	是否改进态度	学习能力	
	教学风格	是否改进风格			学生提出的教学态度、方法和风格的反馈	对学生的反馈	综合素质	

学生对老师的次评的实时评价包括对本次课教案要求的知识点是否听懂、教师的教学态度是否认真、教学方法是否适合本课知识点、教学风格是否接受等，这类评价的评价结果为布尔类型的数据。以“基本数据类型”一节为例，教案要求的知识点，以及老师对学生的次评的评价详见表1所列。

学生对老师的周评主要是本周的知识点是否掌握、次评时对教师的态度、方法和风格是否改进的评价；月评主要是对本月知识点是否掌握及通过教务系统进行的月末评价；末评主要是学生通过教务系统进行的期末评教。老师对学生的周评主要是分组考查、反馈给学生的学习习惯不好的问题是

否改进、学习态度是否转变、对学生提出的教学态度、方法和风格的反馈;月评的评价主要包括月考及学生的知识运用能力、学习能力和综合素质;末评主要是传统的期末考试。

5 评价方法(Evaluation methods)

不同阶段的评价内容设置好之后,通过多种教学平台和软件便可获取评价数据,评价数据主要有两种类型,一种是数值型,一种是文本型,本文针对两类评价数据研究两类评价方法,针对数值型数据采用定量评价方法,文本型的数据采用定性评价方法,全面地评价教师的整个教学过程,以及学生的整个学习过程。

5.1 定量评价算法模型

首先确定定量评价指标,主要包括成绩、布尔类型数据等,用 $X=(x_1, x_2, \dots, x_n)$ 表示, x_i 表示第 i 个指标($i=1, 2, \dots, n$),定量评价方法如公式(1)所示。

$$Y = F(X) \quad (1)$$

其中, Y 为评价结果, F 为评价函数。传统评价方法的评价函数采用加权平均的方式,如公式(2)所示。

$$Y = F(X) = w_1 x_1 + w_2 x_2 + \dots + w_n x_n \quad (2)$$

其中, w_i 为第 i 个指标的权重,通常取课程的学分作为权重。

这种传统的评价方法没有考虑课程之间的相关性,例如数学成绩好的学生通常物理成绩也会很好,传统的定量评价方法将数学和物理成绩加权了两次,这种评价方法会由于某一方面能力的突出导致学生的综合素质突出,不能客观的评价学生的能力。相关分析方法得到了广泛应用^[8,9]。本文提出一种利用相关性分析的评价方法,将多个相关性很强的指标作为一个指标进行评价,相关系数如公式(3)所示。

$$\text{Corr}(x_i, x_j) = \frac{\text{Cov}(x_i, x_j)}{\sigma_{x_i} \sigma_{x_j}} \quad (3)$$

$\text{Cov}(x_i, x_j)$ 为 x_i 和 x_j 的协方差,如公式(4)所示。

$$\text{Cov}(x_i, x_j) = E(x_i - \mu_{x_i})(x_j - \mu_{x_j}) \quad (4)$$

其中, x_i 和 x_j 为任意两个指标向量, $j=1, 2, \dots, n$, μ_{x_i} 为指标 x_i 的均值, μ_{x_j} 为指标 x_j 的均值, σ_{x_i} 为指标 x_i 的方差, σ_{x_j} 为指标 x_j 的方差, E 为指标的期望。

相关系数 $\text{Corr}(x_i, x_j) > T$ ($0 < T < 1$)时,去掉指标 x_i ,通过令公式(2)中 $w_i = 0$ 实现该方法,进而计算评价结果。

5.2 定性评价

根据教师和学生的教学体验可知,文字形式的评价较定量打分形式的评价更能触动教师和学生,所以对文本型的数据的评价也非常有价值。本文给出基于自然语义处理的定性评价方法的软件架构,如图3所示。

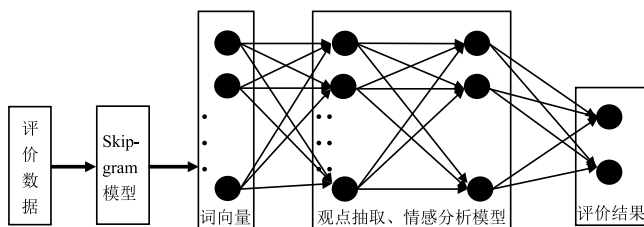


图3 定性评价方法的软件架构

Fig.3 Software structure of the qualitative evaluation method

基于深度学习的自然语言处理领域的评论观点抽取和情感分析技术较为成熟。图3中将本文学生对老师文本评价数据作为输入,通过深度学习模型Skip-gram得到词向量,将词向量作为观点抽取、情感分析模型的输入,通过标注好的评价结果可以进行模型的训练。当使用本软件时,评价数据输入系统,评价结果就是老师看到的学生对老师的客观的评价结果,包括评价是正面还是负面、学生评论监控等。

6 结论(Conclusion)

本文研究了评价体系中的评价阶段、评价内容和评价方法,给出了一套实时的双向的评价体系的软件架构,完善了教学评价体系,丰富了教学评价的内容,调动学生学习积极性,提高教学质量和教学效果。利用机器学习技术中的相关性分析方法对定量评价方法进行研究,并给出了切实可行的算法模型,利用自然语言理解技术对定性评价方法进行研究,给出了思路。本文提出的方法需要教师在课堂上和课堂下与学生进行实时的沟通、互动和评价,这需要教师进行规划并付出一定的精力;数据需要通过网络化的软件平台获取,给出相应的软硬件平台的软件架构;所以本文下一步工作就是逐一进行这些软件的分析、设计和实现。

参考文献(References)

- [1] 王小群,赵道亮.基于集对分析的双向综合教学评价模型应用[J].应用技术学报,2012,12(02):122-125.
- [2] 张亚媛,黄春兰.基于学生满意度的翻转课堂教学评价[J].重庆电力高等专科学校学报,2018,23(06):1-4.
- [3] 申瑶,田利杰.教学双向评价法在消毒供应中心带教管理中的应用[J].中华护理教育,2011,8(1):34-35.
- [4] 江丽丽.浅析教学双向评价系统中的数据挖掘技术[J].山东工业技术,2013(10):175.
- [5] 徐蕾.我国应用技术型大学质量保障研究[D].武汉:武汉大学,2016:83-100.
- [6] 刘红.大学英语双向教学评价体系构建[J].武汉工程大学学报,2010,32(8):89-92.
- [7] 王志波,胡佳慧,王骞.基于众包思想的实时教学评价模式探索[J].计算机教育,2018(01):36-38.
- [8] 梁吉业,冯晨娇,宋鹏.大数据相关分析综述[J].计算机学报,2016,39(01):1-18.
- [9] Jiangqing Fan, Han Liu. Statistical analysis of big data on pharmacogenomics[J]. Advanced Drug Delivery Reviews, 2013, 65(07): 987-1000.

作者简介:

孙 庚(1979-),男,硕士,副教授.研究领域:嵌入式软件开发.

冯艳红(1980-),女,硕士,副教授.研究领域:自然语言处理,机器学习.本文通讯作者.

于 红(1968-),女,博士,教授.研究领域:数据挖掘,信息检索.

史鹏辉(1975-),男,硕士,副教授.研究领域:网络优化.