

情感分析在学生评教中的应用

陈国心

西安航空学院, 陕西 西安
Email: 201803013@xaau.edu.cn

收稿日期: 2021年7月21日; 录用日期: 2021年8月23日; 发布日期: 2021年8月31日

摘要

为适应教育信息化的发展趋势, 各高校将部分传统教学环节转移到线上, 因此每年会产生大量的教学管理及评学评教的数据, 其中, 分析学生评教数据是教师了解学生需求, 提高教学质量的有效途径。论文利用某高校真实的学生评教数据, 探索学生对各教学环节的关注与情感倾向。论文首先对学生评教的历史与现状进行说明; 再结合SnowNLP情感分析的特点, 利用情感词典生成情感分析模型; 将数据预处理后, 通过模型获得评教数据的正负性; 再通过统计分析的方法, 将结果可视化呈现。分析发现, 学生对备课最关注且积极评价比例最高, 而对作业评讲环节关注度最低, 课堂管理与课堂互动环节消极评价比例最高。分析结果可为教师改善教学过程提供参考依据。

关键词

评教, 自然语言处理, 情感分析, SnowNLP

Application of Sentiment Analysis in Student Assessment

Guoxin Chen

Xi'an Aeronautical University, Xi'an Shaanxi
Email: 201803013@xaau.edu.cn

Received: Jul. 21st, 2021; accepted: Aug. 23rd, 2021; published: Aug. 31st, 2021

Abstract

In order to meet the development trend of education informatization, universities have transferred some traditional teaching sessions online, so a large amount of data on teaching management and evaluation of learning and teaching will be generated every year, among which student evaluation

data is an effective way for teachers to understand students' needs and improve teaching quality. The paper uses real student evaluation data from a university to explore students' concerns and emotional tendencies towards each teaching session. The paper first explains the history and current situation of students' evaluation of teaching; then combines the features of SnowNLP sentiment analysis and uses sentiment dictionary to generate sentiment analysis model; after pre-processing the data, the positive and negative aspects of the evaluation data are obtained through the model; and then the results are presented visually through statistical analysis. The analysis found that students were most concerned about lesson preparation and had the highest proportion of positive evaluations, while they were least concerned about the homework assessment session and had the highest proportion of negative evaluations in classroom management and classroom interaction. The analysis results can provide a reference basis for teachers to improve the teaching process.

Keywords

Evaluation of Teaching, Natural Language Processing, Emotional Analysis, SnowNLP

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

教学评价是依据教学目标，以保证教学效果、提高教学质量为目的，根据一定的标准，运用科学有效的方法和手段，对教学活动及其结果进行判定，并为教学决策服务的活动[1]。其中，以学生评教最能反映教师在授课过程中的优缺点。

20世纪80年代中期，我国部分高校开始引入学生参与评教，后逐步获得多数高校的认可，并逐渐纳入到高校考核中[1]。但由于早期受到技术限制，各高校主要以调查问卷或填涂信息卡的方式进行评教数据的收集与统计。

当下，信息技术发展一日千里，各高校为不断适应现代教育环境与方法，也将部分教学活动转移到了线上[2]，学生评教也由早期的调查问卷改变为线上的评分与评论相结合的方式，教育方式也由曾经的以教师为主体，逐步过渡到以学生为主体。

学生作为课堂教育的主体，直接参与到了教师授课过程中，对于教师的备课、语言表述、课堂互动、课堂管理、授课技巧等各个授课环节的感受，通过学生评教能够最直接地表达出来，因此，分析课堂评教内容，可以及时掌握学生对本课程教师授课活动的建议，使教师及时发现自身在授课过程中存在的优缺点，促进教师提升教学技巧、提高教学质量；也有助于学校建立行之有效的量化机制，了解教师教学水平，提高教学量化考核的质量。

无论是线上还是线下教育，普遍采用的是一对多的模式[3]，即一个老师对应多个学生。因此，在线上的学生评教活动中，会产生大量的数据，一部分是由评分方式产生的数值型数据，另一部分是由学生自然通过自然语言描述产生的文本类数据。对数值型数据的分析相对简单，但分析结果无法有效反映教师在教学活动中具体的优缺点，因此，文本类数据是了解学生状况、关注学生差异、获得学生反馈信息、提升教学效果的重要途径。

但由于学生基数大，且水平参差不齐[4]，因此学生评教内容存在主观性和随意性，导致对评教结果

的判定多以统计评分为主，而忽略了学生评教的文本内容，使教师无法更加详细客观地了解学生的评教意见，从而无法及时改进教学方式，提高教学质量。

通常评教指标包括教学态度、教学内容、教学方法及教学效果，教学效果通常以分数形式呈现，故本文将某高校学生评教文本数据在一级指标[5]的基础上细分为5个维度，包括：备课、课堂管理、作业评讲、语言表达、课堂互动，见表1；再通过自然语言情感分析对文本内容进行积极和消极判定，最后依据结果的可视化分析，讨论教师在教学活动中改进的方向。

Table 1. Breakdown of evaluation dimensions

表 1. 评教维度细分

一级指标	评教描述	归一化标签
教学态度	备课充分度、板书、PPT	备课
	课堂管理类描述	课堂管理
教学内容	讲评类描述	作业评讲
教学方法	语言表达类描述	语言表达
	课堂气氛类描述	课堂互动

2. 自然语言处理

自然语言处理是计算机科学和人工智能的研究领域之一，通常将自然语言转换成计算机能够用于理解这个世界的的数据或用于生成能够体现这种理解的自然语言文本[6]。

2.1. SnowNLP

SnowNLP 是一个 GitHub 开源项目，是针对简体中文文本中的主观信息(如态度、情感)进行提取、分析、处理、归纳的 Python 库[7]。

SnowNLP 功能强大，能够实现中文分词、情感分析、文本分类、繁体字转简体字、文本相似性分析等功能[8]。其中情感分析是基于 SnowNLP 自带的两个积极与消极语料库(neg.txt、pos.txt)实现情感数值计算。情感数值取值范围从 0 到 1 表示情感由消极到积极的变化，即越接近 0 表示负面情绪越高，越进阶 1 表示正面情绪越高。例如，对“缺勤人数多”进行情感评价，获得情感值为 0.27841740057，表示该语句偏向表达负面情绪；而语句“重难点讲解透彻”的情感值为 0.9981557064，表示该语句偏向表达正面情绪。

但 SnowNLP 自带语料库是关于酒店、书本等电商留言评论内容[9]，部分内容较片面，不能够满足学生评教数据情感分析，因此需要更换语料库进行训练。

2.2. 语料库

语料库是通过自然语言运用的随机抽样，以一定大小的语言样本来代表某一研究中所确定的语言运用的总体，主要应用于语言定量分析、语法研究、法律语言研究、自然语言理解、机器翻译等方面[10]。

目前网上开源的中文语料库较多[11]，例如酒店评论语料、商品评论语料、微博情绪语料、影评数据集等，但针对教学评论的语料库内容极少，本文使用通用情感词典生成新的积极与消极情感语料库，词典包括：大连理工大学中文情感词汇库、台湾大学中文情感极性词典、清华大学中文褒贬义词典、玻森中文语义词典、Synonyms 中文近义词库。

3. 数据预处理

本文对某高校某门课程的 265 条学生评教文本数据进行情感倾向分析, 该评教数据是针对教师教学过程中的优点及不足两方面进行评价, 如图 1 所示。

263	您认为该教师教学的长处或精彩之处有哪些:上课认真, 结合实际, 采用随堂测验的方式了解学生的情况; 您认为该教师的教学有哪些需要改进的地方:设备的介绍应重点突出, 多和学生沟通, 调动积极性;
264	您认为该教师教学的长处或精彩之处有哪些:教学准备充分; 讲授重点突出; 课堂气氛活跃; 教学效果较好。; 您认为该教师的教学有哪些需要改进的地方:注意进一步调动学生思维; 适当提高音量, 降低语速。;
265	您认为该教师教学的长处或精彩之处有哪些:课程准备充分, 讲稿手写认真, 内容丰富。讲解细致, 语速适度。 交流: 该课共8周32学时, 教材内容庞杂丰富, 在短学时内讲解全部内容则时间紧张。故, 讲解时, 具体编程的每条指令不必过渡细致讲解, 给出模型及编程和算法思路以及编程流程图即可; 您认为该教师的教学有哪些需要改进的地方:1.板书书写不宜太密集; 2.对内容庞杂, 丰富的短学时课程, 注意知识点的衔接, 脉络联系, 强调出重点, 难点即可; 且授课前应提前进行讲解思路和讲解方法的设计。;

Figure 1. Student review data

图 1. 学生评论数据

3.1. 评教数据预处理

数据预处理之前要进行数据质量分析[12], 数据质量分析的主要目的就是检查原始数据中是否存在“脏数据”。脏数据包括缺失值、异常值、不一致的值和重复数据。

首先将评教数据进行质量分析, 对文本类数据检测发现存在 19 条无评教内容, 及存在“缺失值数据”, 最终获得有效评教数 246 条; 再将评教数据以句为单位进行拆分, 得到 764 条评教数据; 最后依据评教指标的 5 个维度进行人工识别和分类, 若存在某一句中的评教内容包含 2 个以上评教维度, 则将该条评教进行人工分离成对应的语句数, 并添加对应维度。最终获取到具有有效评教维度的数据 815 条, 如图 2 所示。

804	加强与学生互动 → 互动 CRIF
805	作业讲评认真 → 讲评 CRIF
806	该教师教学态度认真 → 语言 CRIF
807	教学准备充分 → 备课 CRIF
808	对教学内容熟悉 → 备课 CRIF
809	讲课语速适中清楚 → 语言 CRIF
810	PPT图文并茂图片丰富设计工整清晰 → 备课 CRIF
811	讲解中能够适当加入实例对内容进行阐释 → 语言 CRIF
812	也吸引了学生注意力 → 互动 CRIF
813	要在调动学生听课兴趣方面做努力 → 互动 CRIF
814	ppt内容偏多 → 备课 CRIF
815	ppt字号偏小看不清楚 → 备课

Figure 2. Data pre-processing

图 2. 数据预处理

3.2. 情感词典预处理

提取通用情感词典中的积极词汇与消极词汇分别生成积极情感倾向语料库与消极情感倾向语料库, 并去掉重复词后, 得到积极词数 32550 条, 消极词数 25407 条, 见表 2。其中玻森中文语义词典中存在大量非中文词语[13], 且词表中的情感倾向性权重取值范围为[-6.704000126, 6.37503943135]之间, 因此先去掉所有非中文词语, 再将权重小于 -1 的中文词语划入消极词表, 权重大于 1 的中文词语划入积极词表, 而权重在 -1 到 1 之前的词语偏向中性, 本论文中暂不讨论。

情感词典集合的生成基本实现了通用性[14],但在评教数据分析的过程中发现,部分词语的积极与消极性与评教表达的含义不符,例如:建议、需要、加强等,在评教中表示教师需要改进之处,应替换到消极词表中,因此需要利用 Synonyms 中文近义词库进一步核准情感词典集合中的积极性与消极性。

Table 2. Emotional dictionary collection
表 2. 情感词典集合

词典名称	积极词数	消极词数
大连理工情感词库	11,229	10,783
台湾大学情感词典	2811	8277
清华大学褒贬义词典	5567	4469
玻森中文语义词典	26,894	9902
总计	46,501	33,431
去掉重复词	32,550	25,407

4. 模型生成与结果分析

4.1. 模型生成

在情感分析前,需要对情感词典集合进行训练[10],代替 SnowNLP 包中原有模型,训练方法如下:

首先找到 SnowNLP 的安装目录,在目录下的 sentiment 文件夹保存着 SnowNLP 进行情感分析的默认语料文件(pos.txt, neg.txt)与通过训练该语料得到的模型,如图 3 所示。其中 sentiment.marshal.3 是 Python3 版本的序列化模型文件, sentiment.marshal 是 Python3 版本之前的序列化模型文件。

名称	修改日期	类型	大小
init.py	2014-4-18 23:59	Python File	2 KB
neg.txt	2013-12-7 22:21	文本文档	3,360 KB
pos.txt	2013-12-7 22:21	文本文档	3,888 KB
sentiment.marshal	2014-4-19 0:21	MARSHAL 文件	295 KB
sentiment.marshal.3	2014-4-19 0:42	3 文件	307 KB

Figure 3. Sentiment catalog

图 3. Sentiment 目录

再指定生成的情感词典集合中的积极词表(train.pos.txt)和消极词表(train.neg.txt),通过 SnowNLP 提供的模型训练与保存方法训练和保存模型,并替换原 sentiment 目录下的模型。

1) #训练模型

```
sentiment.train ('train.neg.txt', 'train.pos.txt')。
```

2) #保存模型

```
Sentiment.save ('sentiment.marshal')。
```

4.2. 结果分析

调用训练好的模型对学生评教数据进行情感倾向分析,获得情感分值如图 4 所示。通过观察数据结

果可以发现，积极评论的情感分值基本大于 0.6，消极评论的情感分值与之相反。由此，将评教数据中情感分值小于 0.6 的统一用 0 表示，情感分值大于 0.6 的用 1 表示，实现积极与消极情感分值的归一化，如图 5 所示。

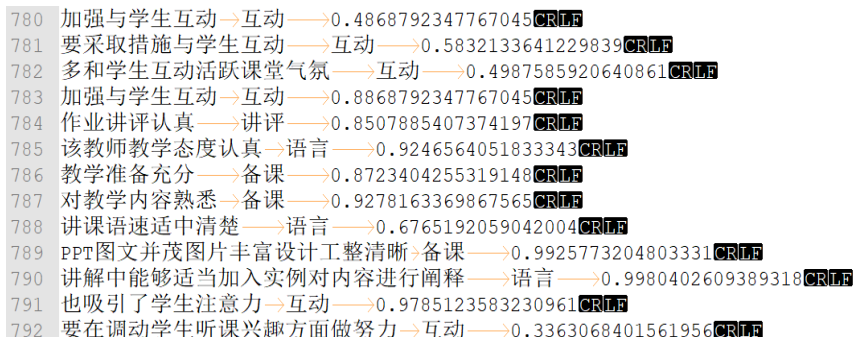


Figure 4. Emotional score

图 4. 情感分值

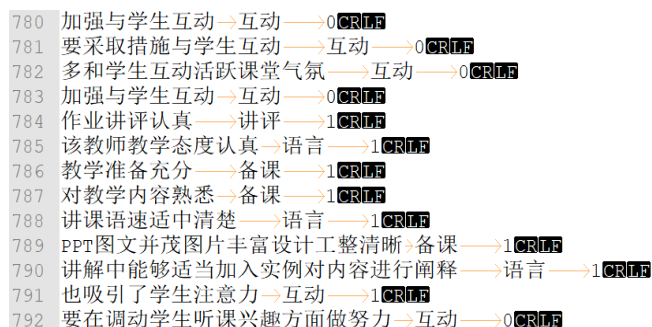


Figure 5. Emotional scores normalization

图 5. 情感分值归一化

对情感倾向分析结果进行分类统计，获得 5 个维度的积极与消极评教数，见表 3。将结果生成饼状图与叠加柱状图，如图 6，图 7。其中备课环节评教数最多，占总评教数的 46.9%，且该环节的积极评教数与消极评教数之比约为 4:1；作业讲评环节评教数最少，只占总评教数的 4.7%；课堂管理与课堂互动环节的积极评教数与消极评教数之比约为 1:3 与 1:2.5；语言表达环节的积极评教数与消极评教数之比约为 1:2.5。

Table 3. Statistics of teaching evaluation

表 3. 评教数统计

评教维度	积极评教数	消极评教数	总数
备课	291	91	382
课堂管理	42	123	165
作业讲评	27	11	38
语言表达	93	36	129
课堂互动	28	72	101

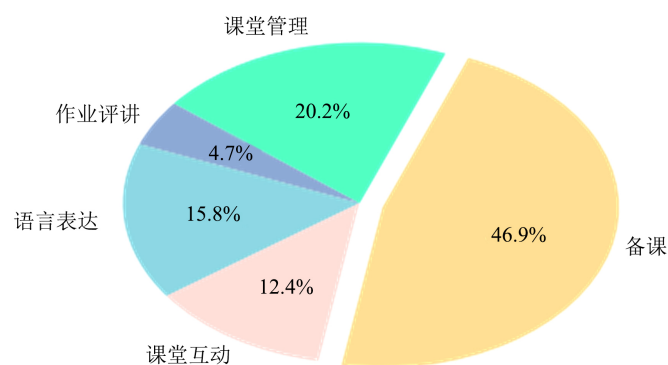


Figure 6. Pie chart

图 6. 饼状图

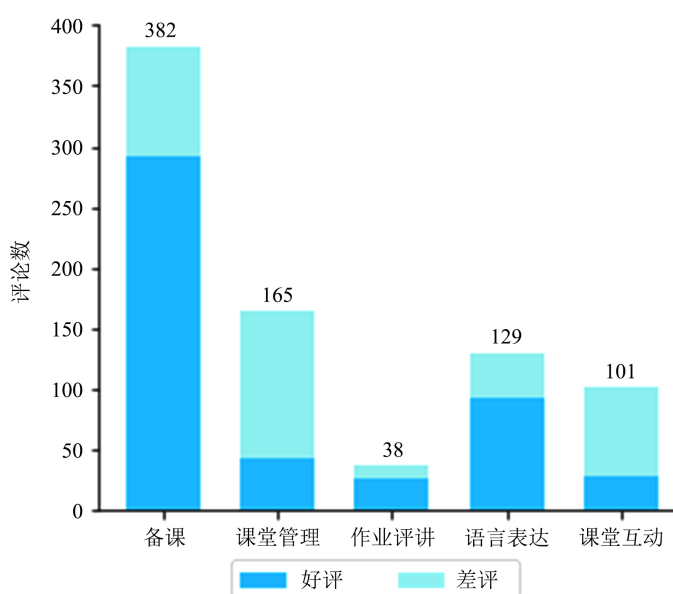


Figure 7. Superimposed bar chart

图 7. 叠加柱状图

综上所述，可以发现：

1) 学生对本课程授课过程中的教师备课环节最为关注，且该环节积极评论数远高于消极评论数，证明教师在该环节准备充分，应继续保持。

2) 学生对作业评讲环节关注度最低，证明该环节存在感较低，学生对该环节的重视度不够。教师应在授课过程中适当提高讲评环节的时间占比，从而增强学生对作业讲评的重视程度。

3) 课堂管理与课堂互动环节消极评论数多于积极评论数，证明教师在授课过程中对课堂纪律要求不够严格，且在调动学生积极性、活跃课堂气氛方面有所欠缺，需要在后期教学过程中，加强课堂管理，并努力提高授课技巧。

5. 结束语

本文在教育现代化背景下，利用 SnowNLP 对学生评教文本数据进行情感倾向正负性分析，再经过统计与可视化展示，判断教师在授课过程中存在的优缺点，为教师改善教学过程提供了理论依据。

参考文献

- [1] 李杨, 尹天光. 高校学生评教异化原因分析及解决方法研究[J]. 教育教学论坛, 2021, 4(21): 21-24.
- [2] 中华人民共和国教育部. 《教育信息化十年发展规划(2011-2020年)》[R]. 2012-03.
- [3] 杨现民, 田雪松. 互联网+教育: 中国基础教育大数据[M]. 北京: 电子工业出版社, 2016.
- [4] Huang, J.T., Zeng, G.H., Huang, B., Gao, Y.B., Liu, J. and Shi, Z.C. (2021) Knowledge Enhanced Pre-Training Model for Vision-Language-Navigation Task. *Wuhan University Journal of Natural Sciences*, **26**, 147-155.
- [5] 高祥. 基于扩展情感词典的短文本情感分析技术研究[D]: [硕士学位论文]. 青岛: 青岛理工大学, 2020.
- [6] (美) Hobson Lane, 等, 著. 自然语言处理实战[M]. 史亮, 等, 译. 北京: 人民邮电出版社, 2020.
- [7] 余同瑞, 金冉, 韩晓臻, 李家辉, 郁婷. 自然语言处理预训练模型的研究综述[J]. 计算机工程与应用, 2020, 56(23): 12-22.
- [8] 郭荣荣, 闵素芹. “线上学习”舆情分析与在线教学提升策略[J]. 中国传媒大学学报(自然科学版), 2020, 27(6): 48-54+78.
- [9] 钟锦燕. 基于深度学习的文本分类研究[D]: [硕士学位论文]. 成都: 电子科技大学, 2020.
- [10] 曾小芹, 余宏. 基于 Python 的商品评论文本情感分析[J]. 电脑知识与技术, 2020, 16(8): 181-183.
- [11] 赵志升, 靳晓松, 温童童, 梁俊花. 基于 Python-Snownlp 的新闻评论数据分析[J]. 科技传播, 2018, 10(18): 104-105.
- [12] 夏玉芹, 单雪微. 基于 Python 的简单文本情感分析[J]. 阴山学刊(自然科学版), 2018, 32(4): 58-62.
- [13] 王恩慧. 基于情感倾向点互信息算法的情感分析方法研究[J]. 科学技术创新, 2021, 4(12): 89-90.
- [14] 杜彦熳, 张军, 郑棋方, 赵贞箏, 王寄權. 基于 SnowNLP 技术的微博评论分析研究——以“李子柒”微博为例[J]. 科技经济导刊, 2020, 28(18): 37.