

高校社会声誉评价指标体系研究

刘盛博,刘佳新,刘苗苗

(大连理工大学 高等教育研究院,辽宁 大连 116024)

摘要:在“互联网+”时代,声誉越来越成为高校之间竞争与高校评价的关键因素。已有的高校社会声誉评价侧重主观的定性评价。因此,本研究尝试采用定量评价方法,以高校社会声誉内涵为基点,从高校社会声誉生成机制出发,构建一套量化测度高校社会声誉的指标体系,并以我国104所公办本科院校为实例,验证了评价指标体系的有效性。研究结果可为完善我国高校评价体系提供借鉴。

关键词:高校社会声誉;指标体系;声誉评价

Research on the Evaluation Index System for the Social Reputation of Higher Education Institutions

LIU Shengbo, LIU Jiixin, LIU Miaomiao

(Institute of Higher Education, Dalian University of Technology, Dalian 116024, Liaoning, China)

Abstract: In the "Internet+" era, reputation has increasingly become a key factor in the competition and evaluation of higher education institutions. The existing social reputation evaluation of higher education institutions focuses on subjective qualitative evaluation. Based on the connotation of the social reputation of higher education institutions and starting from its generation mechanism, this research attempts to construct a set of index system for quantitatively measuring the social reputation of higher education institutions by adopting a quantitative evaluation method. Then 104 public undergraduate institutions are chosen as examples to verify the effectiveness of the evaluation index system. The research results could serve as valuable references for improving the evaluation system of higher education institutions in our country.

Keywords: social reputation of higher education institutions; index system; reputation evaluation

一、引言

在“双一流”建设的背景下,高校之间的竞争越发白热化,政府投资、学生、师资等资源成为各个高校奋力争夺的对象。为了在竞争中取得相对优势,高校声誉及其培育越来越受到重视。高校声誉是大学给社会公众留下的综合印象^[1]。高校声誉作为企业声誉的延伸,是指大学作为一类特殊的社会组织形式所获得的社会大众的赞美和支持的程度,是大学整体实力和社会形象的外在表现形式^[2]。高校声誉作为一种宝贵的无形资产,越来越受到高校及各个利益相关群体的关注^[3]。袁浪华^[4]针对民办本科高校声誉困境,提出民办本科高校声誉的提升不仅需要第三方机构客观公

正的参与,还需要政府管理部门的政策支持。段婕^[5]从利益相关者角度构建了高校声誉管理动态模型,指出高校声誉管理不仅需要对内管理,还需要对外管理。当前国内外很多具有影响力的大学排行榜都将声誉指标列为其大学评价指标的重要构成要素之一,表1列出了一些国内外知名大学排行榜最新的声誉指标及其权重。

在高校社会声誉的评价方法方面,国外大学排名中多采用声誉调查法作为间接的质量评估方法。声誉调查法和同行评价在国外的大学排名中扮演着重要角色,是最为常见的排名方法^[6]。有研究认为,这种声誉调查评价方法存在着各种各样的潜在问题,如声誉调查的评价者对高校存在偏见问题^[7]。通常来说,

表1 国内外知名大学排行榜声誉指标及其权重

排行榜名称	排行机构	声誉指标详情	声誉指标权重
QS世界大学排名	英国QS	学术领域同行评价, 40%; 全球雇主评价, 10%; 教学声誉调查, 15%; 研究声誉调查, 18%	50%
THE泰晤士世界大学排名	《泰晤士报高等教育增刊》	区域研究声誉, 12.5%; 全球研究声誉, 12.5%	33%
U. S. News世界大学排名	《美国新闻与世界报道》	声誉调查, 20%	25%
加拿大麦凯琳大学排行 校友会中国大学排名	加拿大《麦凯琳》杂志 中国校友会	社会影响, 13.76%	20% 13.76%

评价者无法得知有关待评项目的全部信息^[8],他们最熟悉、最了解的往往是自己工作或者学习过的地方,而对其他高校的情况往往只是通过表面文字的简单了解^[9]。1993年,美国国家研究委员会(National Research Council, NRC)关于研究生院的声誉调查中,有一半以上的评分者对2/3的待评项目都不了解^[10]。2009年,《美国新闻与世界报道》声誉调查的问卷回收结果显示,佛蒙特大学校长对262所参评高校中近乎一半的高校都不了解,选择了“不知道”选项^[11]。通常来说,非学术界的评价者所知更少^[12]。在声誉调查对象的选择问题和声誉指标的权重问题上,均存在不同程度的争议。高校社会声誉往往表现为一种主观评价,却真实地反映出高校的社会地位及影响,并在很大程度上左右着大学的未来发展^[3]。现有对高校声誉的研究多集中在对高校学术声誉的研究,对高校社会声誉的专门研究较少;并且在评价方法上,多采用定性评价方式,这种方法存在着主观性强的弊端。互联网技术和自媒体信息的发展,为我们开展高校社会声誉研究提供了机遇。本研究在借鉴已有研究的基础上,基于自媒体和网络多源数据,旨在为高校社会声誉提供一种定量测度的方法,并建构一个具有一定科学性和可操作性的评价指标体系。

二、高校社会声誉指标体系构建

高校社会声誉是高校间竞争的重要因素,对高校社会声誉进行评价是互联网时代高校评价的重要环节。本文将在高校社会声誉的概念及内涵基础上,探析高校社会声誉的形成机理,构建高校社会声誉评价的指标体系,并用层次分析法对各指标进行权重赋予。

(一) 高校社会声誉的概念及形成机理

高校声誉是高校长期累积的结果,可以分为学术声誉和社会声誉。学术声誉一般指学术界对一所学校学术水平的评价。学术水平往往较为集中地体现在学术论文及专著的发表和引用、科研项目的承担或

重要成果的拥有等方面^[13]。高校社会声誉往往指一所高校的精神、行为、办学条件、社会贡献等大学身份识别要素在社会公众心目中产生的认知结果和情感反应,能从根本上较为准确地反映一所大学的地位和影响^[14]。

高校社会声誉在学术界一直没有十分清晰的界定,学者大多从高校社会声誉作用的角度认为高校社会声誉是指高校在长期的办学实践过程中形成的以综合实力为基础、以培养人才的质量为核心、以服务社会产生的经济及社会效益为指标的多方面社会影响力^[15],是大学整体价值和竞争力的外显形式,获得良好的社会声誉是高校在竞争群体中脱颖而出的关键^[16]。

声誉理论源于经济学领域的企业声誉。从本质上来讲,企业声誉来自企业对于公众的情感吸引,良好的企业声誉有助于增进消费者及其他利益相关者对企业本身的信任。根据声誉理论,随着高校的社会价值不断显现,高校在社会公众的心目中会形成一个良好的认知,这种良好的认知也正是高校社会声誉的形成过程。高校社会声誉建立后,在高校与社会公众后续的互动过程中,高校会进一步强化其行动。在高校的价值不断显现的过程中,高校也能够获得超额的回报,具体表现在获得更加优质的学生和教师资源,这一系列互动的活动最终会促进整个高校的发展。参考源于经济界的声誉理论,本研究将高校声誉理解为社会公众对高校及其所产生的社会价值的一种综合评价,在当今时代,这种综合评价主要通过网络等渠道表现出来。

社会公众对大学产生形象认知的过程是一个传播和反馈的过程,大学、媒体和社会公众是这一传播过程的三个基本要素,“大学形塑—媒介传输—公众认知”构成了高校声誉的基本传播模式(如图1所示)。社会公众是大学形象的认知者和评价者。一方面,大学的理念和行为通过媒介输出信息成为社会公众形象认知的根据;另一方面,社会公众通过媒介接收大学信息,结合主观知识,形成对大学的感知和判断,经过汇聚整合构成公众群体的整体意识,最终形成社会公众对大学形象的认知。公众通过自身认知对高校给予反馈,社会公众在高校社会声誉的形成与发展过程中发挥着重要作用。总之,高校社会声誉是高校在办学过程中不断形成的高校与其利益相关者认知两方面相互作用的结果,并最终呈现为对于这所高校的综合性评价^[17]。

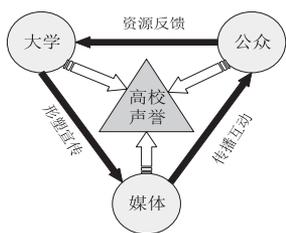


图1 高校社会声誉生成机制示意图

大学形塑是高校社会声誉产生的基础。只有通过良好的大学形塑,大学的形象才会出现在媒体中,才会被公众所认知。大学形塑包括科研能力、人才培养能力、师资队伍水平、服务社会能力等诸多方面。大学形塑完成后,需要运用传播手段使其被广大社会公众所熟知。媒体传播是形成高校社会声誉的重要途径。媒体传播既包括高校自身的宣传,又包括外部媒体对高校的宣传。前者主要通过高校网站、微博、微信公众号等媒体展开;后者主要包括学术网站中的宣传影响要素(如科学网)和社会公共网站中的宣传影响要素(如新浪新闻、百度新闻等)。媒体传播完成之后,接下来是公众认知和反馈的过程。借助传播学中的涵化分析理论,公众在媒体传播过程中得到了对于高校的认知,或者通过搜索引擎、官方网站的途径对高校进行了主动求知。公众通过从外界得到的信息,再结合自己的认知,形成对高校的整体认识。在这个过程中,公众也对高校形成反馈。不同公众群体的反馈途径有所不同,如:学生群体通常通过百度贴吧、微博、知乎等途径反馈,雇主群体通过招聘信息或薪酬等反馈,其他公众通过网络检索及网站访问等方式反馈。在本研究中,借助声誉理论和涵化分析理论,将上述声誉形成过程中涉及的要素及其相互关系抽取出来,为构建高校社会声誉评价指标提供理论支撑。

(二)高校社会声誉指标体系的初步确立

根据涵化分析理论,传播媒介在潜移默化中影响受众的看法与观点。高校在利用传播媒介进行传播活动时,同样会在潜移默化中影响公众对高校的看法,进而影响高校的社会声誉。在互联网飞速发展的时代,高校进行信息传播的方式也更加多样。微信公众号为媒体和个人提供了一种新的信息传播方式。微博是一种基于用户关系信息分享、传播以及获取的广播式社交媒体,具有极强的便捷性、传播性和原创性。百度是我国应用最为广泛的搜索引擎,百度指数是百度推出的以百度海量网民为基础的数据分享平台,其海量的数据存储在近年来被许多研究者应用到了

旅游、金融、社会热点等研究领域中。知乎是一个真实的网络问答社区,用户可以在该平台上自由分享彼此的知识、经验与见解。知乎用户遍布各行各业,且用户群体庞大,越来越被更多的用户所接受。微信、微博、官方网站等已经成为高校进行信息传播与形象展示的窗口,知乎与网络搜索引擎为广大社会公众提供了了解高校的途径。

社会声誉是社会公众对高校的认知结果和情感反应,是对高校的一种综合评价。公众对高校的认知分为三类:第一类公众对高校关注较少,主要是被动认知高校,通过媒体宣传对高校形成初步认知;第二类公众对特定高校关注较多,并且主动关注相关信息,如通过搜索引擎查找高校信息、对高校微博进行关注等;第三类公众对高校较为熟悉,并且可以为其他公众提供高校在某些方面表现的精准信息,如高校在校生和毕业生可为公众提供有关高校师资力量和教学水平的精准信息,企业雇主可为公众提供有关高校人才培养能力的信息等。各类大学排行榜中的高校社会声誉主要将雇主调查反馈作为评价指标。本研究根据不同公众群体对高校的认知差异,将高校社会声誉评价指标分为媒体传播、公众认知和公众反馈3个一级指标,然后基于数据可获取性、完整性和准确性原则确定7个二级指标,包括微信传播指数、网页反链数、百度指数-媒体指数、微博粉丝数、百度指数-搜索指数、知乎评论和雇主评价-毕业生平均薪酬,具体如下表2所示。

表2 高校社会声誉评价指标

指标体系	一级指标	二级指标
高校社会声誉	媒体传播	微信传播指数
		网页反链数
		百度指数-媒体指数
	公众认知	微博粉丝数
		百度指数-搜索指数
	公众反馈	知乎评论
		雇主评价-毕业生平均薪酬

为避免主观评价的影响,本研究在指标数据获取过程中尽量避免直接的学者调查或学生调查,所有指标均采用客观数据进行分析,进而增强指标体系的科学性与可操作性。

(三)指标权重的确定

本研究在确定高校社会声誉评价指标的权重时采用层次分析法。基本思路和步骤如下:

- (a)建立层次结构模型;
- (b)根据层次结构模型,建立两两比较的判断矩阵;
- (c)由判断矩阵所获得的数据计算特征向量;

(d)对判断结果的一致性进行检验,得出目标权重。

1. 构建高校社会声誉评价层次结构模型

依据指标体系构建高校社会声誉评价的层次结构模型,并把结构模型分为目标层A、准则层B和指标层C。目标层表示决策的最终目的或者要实现的总体目标,在本研究中,目标层为高校社会声誉。其他各层的具体指标和指标间的相互关系如图2所示。

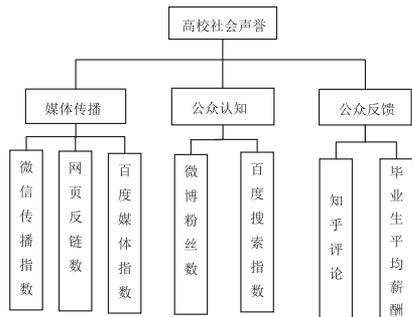


图2 高校社会声誉评价的层次结构模型

2. 建立判断矩阵

通过两两比较的方法判断指标重要性在解决复杂问题时尤为重要。为避免研究者主观价值偏好的臆断,本研究通过构造判断矩阵,用专家打分的方法确定各指标的重要程度,对各指标进行两两比较赋值,采取1-9的标度赋值的方式两两比较各指标的重要性(见表3)^[18]。

表3 各个指标的重要性等级及其赋值

标度	内涵
1	表示 <i>i</i> 因素与 <i>j</i> 因素的重要性相同
3	表示 <i>i</i> 因素比 <i>j</i> 因素的重要性稍强
5	表示 <i>i</i> 因素比 <i>j</i> 因素的重要性强
7	表示 <i>i</i> 因素比 <i>j</i> 因素的重要性明显强
9	表示 <i>i</i> 因素比 <i>j</i> 因素的重要性非常强
2, 4, 6, 8	表示两因素的重要性为上述相邻两级判断的中间值
1, 1/2, 1/3, …, 1/9	表示两因素的重要性之比和上述意义相反

根据层次分析法的重要性等级并结合层次分析结构模型,建立各个指标两两判断的矩阵。设计指标两两比较的调查问卷,并采用纸质问卷的发放形式。共咨询了5位高等教育评价领域的专家,他们都具有多年高校工作及学习经验。

3. 单层权重的计算

在层次分析过程中,本研究选择的是判断矩阵集结的方式,即首先分别计算各专家判断矩阵,再对各专家判断矩阵中各要素求均值,得到均值判断矩阵后,再计算排序权重。具体过程为:

(1)计算集结矩阵,将各专家判断矩阵中各要素求均值

$$\text{集结判断矩阵 } X = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1n} \\ X_{21} & X_{22} & \cdots & X_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ X_{n1} & X_{n2} & \cdots & X_{nm} \end{bmatrix}$$

$$\text{集结判断矩阵中各元素 } X_{ij} = \frac{1}{m} \sum_{j=1}^m x_{ij}$$

其中 X_{ij} 为各专家判断矩阵中对应元素值, m 为专家数量,在本研究中 $m=5$ 。

(2)计算单层权重

集结判断矩阵求得后,需计算各矩阵单层权重,即集结判断矩阵的特征向量求解。表4为一级指标集结判断矩阵及对应指标权重。

表4 高校社会声誉评价一级指标判断矩阵及权重

一级指标	媒体传播	公众认知	公众反馈	R_i
媒体传播	1.0000	2.4000	0.6000	0.3390
公众认知	0.4333	1.0000	0.4333	0.1719
公众反馈	1.8000	2.4000	1.0000	0.4890

(3)判断矩阵的一致性检验

各指标权重得出后,指标权重是否合理,需要进一步检验。具体操作步骤为:首先计算判断矩阵的最大特征根 λ_{max} ,将最大特征根的特征向量作为权向量。其次计算判断矩阵一致性检验指标 CI 的值,计算公式如下:

$$CI = \frac{\lambda_{max} - n}{n - 1}$$

然后查询随机一致性指标表(表5),将对应的随机一致性指标 RI 值代入公式 $CR=CI/RI$ 中,得出最终指标一致性的 CR 值。如果 $CR < 0.1$,则认为 X 的不一致程度在容许范围内,通过一致性检验。如果 $CR \geq 0.1$,则说明未通过一致性检验,需要重新校对判断矩阵 X ,对 X_{ij} 进行调整。

表5 随机一致性指标RI数值表

n	1	2	3	4	5	6	7	8	9	10
RI	0.00	0.00	0.58	0.90	0.12	1.24	1.32	1.41	1.45	1.49

表6为各集结判断矩阵一致性检验结果,各判断矩阵的 CR 值均小于0.1,表明各判断矩阵所计算出的权重合理,可以确定为最终权重。

表6 各集结判断矩阵的结果

矩阵	λ_{max}	CR
一级指标	3.0872	0.0752
二级指标(一)	3.0656	0.0566
二级指标(二)	2	0
二级指标(三)	2	0

(4)计算各层指标的组合权重

组合权重又称目标权重,是通过层次分析法得到的各指标的最终权重值。上文所得到的权重系数为每一层中各级指标的单层权重,形成最终的指标体系还需计算各层指标的组合权重。组合权重的计算方式为上一层的对应指标权重系数乘以本层各指标对应的单层权重系数,各指标单层及组合权重计算结果如表7所示。

表7 各指标对应的单层权重及目标权重

指标体系	一级指标	单层权重	二级指标	单层权重	目标权重
高校社会声誉	媒体传播	0.3385	微信传播指数	0.4554	0.1541
			网页反链数	0.1783	0.0603
			百度指数-媒体指数	0.3663	0.1240
	公众认知	0.1719	微博粉丝数	0.4365	0.0750
			百度指数-搜索指数	0.5635	0.0969
	公众反馈	0.4896	知乎评论数据	0.4699	0.2301
		雇主评价-毕业生平均薪酬	0.5301	0.2595	

从各对应指标的最终权重来看,一级指标中,权重最大的是公众反馈,占比近50%,其次是媒体传播,权重最小的是公众认知。在所有二级指标中,权重最大的是雇主评价-毕业生平均薪酬,该指标所占权重约为26%;权重最小的是网页反链数,占比约为6%。

三、多源数据获取与处理

自媒体时代为高校社会声誉评价提供了巨大的便利。网络是一个巨大的资源宝库,蕴含着大量的数据。本章将对本研究从网络上获取的各类多源数据的内涵以及获取方式进行介绍,并对数据处理方法进行简要阐述。

(一)多源数据获取

1.微信传播指数

在近些年互联网的发展进程中,微信的出现和兴盛是通信技术发展的典型代表。微信公众平台聚合了数字阅读共享、移动、即时等诸多典型特征,表现出一种大众传播的交流方式,同时成为众多新媒体宣传自己的重要场地。因此,微信公众平台迅速占据了大部分注意力资源且影响甚广,是大数据时代的当下进行研究的重要切口^[19]。微信已经得到各行各业的广泛认可,并且与广大社会公众的生活实现了高度融合^[20]。高校也大多开通了自己的微信公众号,用于宣传文化、展示成果等,是高校对外宣传的一个重要窗口。

微信传播指数是清博大数据公司开发的用于测度微信公众号传播力的一项指标。其主要包括整体传播力、篇均传播力、头条传播力、峰值传播力四项一级指标,指标权重由专家主观赋权法和客观因子分析法获得,具体指标权重及指标计算方式如表8所示。

表8 微信传播指数WCI计算方式^[21]

一级指标	二级指标	指标权重	标准化方法
整体传播力 O (40%)	日均阅读数 R/d	77%	$O=0.77\ln(R/d+1)+0.17\ln(Z/d*10+1)+0.06\ln(L/d*10+1)$
	日均在看数 Z/d	17%	
	日均点赞数 L/d	6%	
篇均传播力 A (10%)	篇均阅读数 R/n	77%	$A=0.77\ln(R/n+1)+0.17\ln(Z/n*10+1)+0.06\ln(L/n*10+1)$
	篇均在看数 Z/n	17%	
	篇均点赞数 L/n	6%	
头条传播力 H (20%)	头条(篇均)阅读数 Rt/nt	77%	$H=0.77\ln(Rt/nt+1)+0.17\ln(Zt/nt*10+1)+0.06\ln(Lt/nt*10+1)$
	头条(篇均)在看数 Zt/nt	17%	
	头条(篇均)点赞数 Lt/nt	6%	
峰值传播力 P (30%)	最高阅读数 Rmax	70%	$P=0.7\ln(Rmax+1)+0.2\ln(Zmax*10+1)+0.1\ln(Lmax*10+1)$
	最高在看数 Zmax	20%	
	最高点赞数 Lmax	10%	

$$WCI=[0.4*[0.77\ln(R/d+1)+0.17\ln(Z/d*10+1)+0.06\ln(L/d*10+1)]+0.1*[0.77\ln(R/n+1)+0.17\ln(Z/n*10+1)+0.06\ln(L/n*10+1)]+0.2*[0.77\ln(Rt/nt+1)+0.17\ln(Zt/nt*10+1)+0.06\ln(Lt/nt*10+1)]+0.3*[0.7\ln(Rmax+1)+0.2\ln(Zmax*10+1)+0.1\ln(Lmax*10+1)]]^2*1.2*10$$

2.微博粉丝数

近年来,社交媒体广泛流行, Twitter和新浪微博等已经成为人们了解新闻信息、探讨社会问题的主要平台。微博信息的高度动态性、交互性与海量性,给民众带来了极大的便利。微博也是高校进行自我宣传、自我展示以及危机公关的重要途径。微博粉丝数能较好地反映该微博账号的受关注情况,侧面反映微博账号的运营质量。

3.官网反链数

高校官方网站是高校进行通知发布、宣传展示、成果分享的重要平台。官方网站的反链数是指指向该网站的链接的数量。举例来说,如果a网站指向b网站,那么对于b网站来说,a指向的这个链接就是b网站的反向链接。网站反链数能从某种程度上反映该网站的经营质量。

4.百度搜索指数和媒体指数

本研究中所涉及的百度指数数据主要包括搜索指数和媒体指数两个模块。搜索指数是指以网民在百度的搜索量为数据基础,以关键词为统计对象,经过科学分析并计算出的各个关键词的搜索频次的加权和。媒体指数是指各大互联网报道的新闻中,与关键词相关的,被百度新闻频道收录的数量,采用的是新闻标题包含关键词的统计标准。本研究主要以各高校名称作为关键词,利用网络爬虫技术获取百度搜索指数和媒体指数两个模块的网络原始数据,然后对数据进行清洗、统计与分析。

5.知乎评论数据

问答社区是当前互联网用户获取相关信息与知识的重要途径。知乎是一种被广大社会公众所接受的网络问答平台,在知乎用户群体中学生占据大多数^[22]。

本研究采取网络爬虫的方式,爬取公众对高校进

行评价的评论数据。具体方法为:在知乎搜索框中输入“***大学怎么样?”,点击搜索,列表中会显示与这个关键词有关的提问及回答;剔除无关数据,对有效数据进行爬取,爬取字段为标题、答主名称、发布时间、赞同数、评论数和回答正文。

6. 毕业生平均薪酬

雇主是高校培养人才的直接“消费者”,雇主评价是社会公众对高校形象和高校声誉的重要反馈,大学生就业状况与雇主评价直接影响着高校社会声誉。毕业生薪酬是雇主评价的外在数据表现。因此,用高校本科毕业生平均薪酬排行来测度雇主评价,具有一定的说服力。中国薪酬网调查并公布了2020年中国大学本科毕业生薪酬排行榜,本研究将借用中国薪酬网的调查成果,用各大学毕业生平均薪酬来代表雇主评价。对于部分2020年数据缺失的高校,借助校友会大学排名,采用滑动平均窗口法对其进行插值处理。

(二)数据的处理与分析

1. 数值型数据的处理

在高校社会声誉评价指标体系所有的二级指标中,除知乎评论数据为文本型数据,其余均为数值型数据。对于数值型数据,本研究中的处理方式如表9所示。

表9 各类数值型数据的处理方式

数据名称	处理方式
微信传播指数	组内取平均值后,进行0-1标准化处理
网页反链数	直接进行0-1标准化处理
百度指数-媒体指数	组内取平均值后,进行0-1标准化处理
微博粉丝数	直接进行0-1标准化处理
百度指数-搜索指数	组内取平均值后,进行0-1标准化处理
雇主评价-毕业生平均薪酬	直接进行0-1标准化处理

2. 文本型数据来源处理

基于情感词典的自然语言处理技术是对文本进行情感分析的方法之一。自然语言处理技术在对文本进行情感分析时应用较为广泛,这种方法能够对公众情感的表达、情绪和用词进行准确分析。在中文领域,当前研究中常用的情感词典多为通用情感词典,如HowNet情感分析用词语集、台湾大学自然语言处理实验室构建的情感词典“NTUSD”、大连理工大学信息检索研究室发布的情感词汇本体库等。这些词典具有比较好的通用性,但是也存在领域适应性差的缺点^[23]。

本研究采用大连理工大学信息检索研究室发布的情感词汇本体库,该资源包括了27466个中文词汇或者短语,词语涵盖广泛,通用性较强,但是也存在领域适应性不足的缺点。现有的情感分析词典中,鲜有

对高等教育领域的专有名词进行标注,如985、211等。因此需要对现有词典进行扩充,补充高等教育领域部分专有词汇,提高网络舆情分析的准确性。

(1)情感词典的扩充

利用网络爬虫的方式,共在知乎平台获取217362条高校评论相关数据,去除噪音数据后得到199998条数据。情感词汇本体库的扩充主要分为以下几个步骤:

第一步:利用jieba分词技术对这些数据进行分词处理,并去除停用词。

在中文文本的情感表达中,词语是表达感情的最基本单元,所以对中文文本进行情感分析,首先需要进行分词处理。jieba分词是当前较为常用的分词技术,可通过开源编码实现。如:“大概就是一个普通985”,利用jieba分词并去除停用词后,分解为“大概”“普通”“985”。

第二步:标注词性。

标注词性即根据词的特点对词语进行分类。词性是进行情感判断的基础,在对文本的情感分析中,多数情感极性和情感强度的判断都是基于词性的。如:表示高校名称的词都是名词;形容词和副词是重要的情感词,是判断一句话情感表达倾向与强度的关键。

第三步:筛选高频词。

对分词后的词语进行词频统计,筛选出现频率在100次以上的词。除去大连理工大学情感词汇本体库中已经包含的词语,共获得688个高等教育领域的专有名词、形容词和副词,这些词语经过情感标注后将扩充到情感词典中。

第四步:标注极性和情感强度。

极性和情感强度是词语的“价格标签”,是词典的重要组成部分。参考大连理工大学信息检索研究室情感词语标注方法,对从语料库中筛选出的688个词的词语极性和强度进行标注。中文情感词汇本体库中将词语的情感极性分为了四类:0代表中性,1代表褒义,2代表贬义,3代表兼具褒贬两性。在本研究中,为了方便计算和表达,对原词库中的极性分类做了如下修改:1代表褒义,-1代表贬义,0代表中性和兼具褒贬两性。中文情感词汇本体库中将情感强度分为1、3、5、7、9五档,情感强度依次增强。在情感分类方面,中文情感词汇本体库中将情感分为“乐、好、怒、哀、惧、恶、惊”七大类,其下又细分为21小类。在本研究中,由三位硕士研究生对新增688个词的情感极性

和强度进行标注,最后对这三位硕士生的标注进行整理,取平均值,确定词语最后的极性和情感强度。

(2)计算知乎评论数据的情感得分

本研究所构建的高等教育领域情感分析词典共计28154个词。建立情感词典的目的就是对文本评论的情感倾向和强度进行计算打分,将文本评论进行量化。本研究在新编制完成的高等教育领域情感分析词典的基础上,计算每一条知乎评论数据的情感得分。计算原理如下:

首先,假设评论语句为 $M=\{m_1, m_2, m_3 \dots m_n\}$, M 为爬虫后得到的评论语句, $\{m_1, m_2, m_3 \dots m_n\}$ 为 jieba 分词后得到的单个词语。

其次,在高等教育领域情感分析词典中查找 $\{m_1, m_2, m_3 \dots m_n\}$ 对应的情感极性为 $\{j_1, j_2, j_3 \dots j_n\}$, 情感强度为 $\{q_1, q_2, q_3 \dots q_n\}$ 。

最后,情感得分计算结果为: $D = \sum_{i=1}^n j_n * q_n$ 。D 为

正表示该条评论具有正面积的情感倾向,为负则表示该条评论具有负面的情感倾向;D 的绝对值越大,则表示其情感强度越强。

四、我国公办本科高校社会声誉评价实例

截至2020年6月,全国共有高等院校3005所,其中包括本科院校1272所。本研究以省或市为单位,采取分层抽样的方式,随机选取样本高校104所。这些高校中有整体发展较好的“985”或“211”高校,也有发展程度一般的普通公办本科院校,高校类型、层次涵盖全面。用这些高校作为样本,能够在一定程度上代表我国公办本科高校社会声誉的整体状况。

根据表7所得出的各指标的目标权重,计算各样本高校社会声誉得分。表10呈现的是社会声誉总得分排在前25名的高校。清华大学、浙江大学和南京大学的高校社会声誉总得分排名位于所有样本高校的前三位。

从表10中可以看出,高校社会声誉与高校层次密切相关,排名靠前的高校中多数为“985”高校。除表中所列高校外,部分高校存在媒体传播指数高而公众反馈不高的情况,如内蒙古工业大学、西昌学院等。虽然这些高校重视到了媒体传播与自身宣传,公众对这些高校具有较强的认知欲,但是公众反馈的认可度不足,学生和雇主对于这些高校的体验感和评价不高。由此也说明,高校仅仅依靠媒体传播与自身宣传并不能从根本上提高其社会声誉。

表10 各高校社会声誉具体计算结果(前25名)

排名	学校名称	微信传播指数	网站反链数	媒体指数	微博粉丝数	搜索指数	知乎评论得分	薪酬排名	得分
1	清华大学	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2	浙江大学	0.552	0.337	0.226	0.336	0.832	0.358	0.753	0.517
3	南京大学	0.304	0.195	0.174	0.102	0.731	0.418	0.719	0.442
4	天津大学	0.921	0.142	0.103	0.217	0.791	0.112	0.597	0.437
5	中南大学	0.713	0.111	0.227	0.076	0.830	0.080	0.528	0.386
6	同济大学	0.665	0.178	0.122	0.127	0.236	0.138	0.733	0.383
7	上海交通大学	0.224	0.290	0.046	0.528	0.261	0.157	0.787	0.363
8	东南大学	0.435	0.106	0.017	0.077	0.476	0.227	0.694	0.360
9	厦门大学	0.736	0.136	0.093	0.209	0.109	0.080	0.660	0.349
9	西安交通大学	0.579	0.171	0.255	0.363	0.061	0.049	0.667	0.349
11	吉林大学	0.674	0.187	0.303	0.037	0.271	0.300	0.357	0.343
12	武汉大学	0.770	0.233	0.407	0.448	0.019	0.185	0.284	0.335
13	哈尔滨工业大学	0.483	0.131	0.061	0.171	0.235	0.175	0.642	0.332
14	四川大学	0.604	0.164	0.414	0.131	0.215	0.207	0.357	0.325
15	中国海洋大学	0.780	0.072	0.306	0.150	0.417	0.054	0.357	0.319
16	大连理工大学	0.529	0.093	0.214	0.070	0.136	0.172	0.560	0.317
17	华南理工大学	0.288	0.166	0.045	0.018	0.485	0.122	0.684	0.314
18	北京理工大学	0.320	0.152	0.166	0.011	0.371	0.143	0.618	0.309
19	重庆大学	0.521	0.121	0.217	0.130	0.585	0.136	0.357	0.305
20	湖南大学	0.678	0.083	0.044	0.092	0.219	0.258	0.392	0.304
20	南开大学	0.405	0.150	0.024	0.213	0.043	0.290	0.549	0.304
22	华中科技大学	0.652	0.179	0.080	0.247	0.080	0.270	0.357	0.302
23	中国科学技术大学	0.489	0.132	0.023	0.085	0.186	0.107	0.640	0.301
24	郑州大学	0.669	0.086	0.111	0.118	0.672	0.135	0.277	0.299
25	南昌大学	0.778	0.083	0.255	0.034	0.025	0.104	0.325	0.270

将本研究的评价结果与中国校友会大学排名、QS大学排名和软科大学排名进行比较发现,在大学排名中位居前列的高校普遍具有较高的社会声誉排名,但也存在一些大学排名较高的高校,其社会声誉排名并未名列前茅,如样本中的山东大学并未排进前25名。样本高校中存在一些行业特色较为明显的高校,这些高校取得了较高的社会声誉,如中国海洋大学、北京化工大学,说明高校类型对高校社会声誉存在一定影响。

五、总结

本研究从社会公众的视角出发,结合高校声誉形成机制,构建了一套量化测度高校社会声誉的指标体系,并通过实例验证了该指标体系应用的可行性。通过对评价结果的分析发现,高校社会声誉评价结果与国际上的QS大学排名、软科大学排名存在一定差异。虽然大学排名高的高校普遍具有较好的社会声誉,但社会声誉排名与高校类型、所在地域等要素存在紧密关联。社会声誉评价指标的意义在于,一方面可以直接体现出高校的社会资本,表征其吸引社会资源的能

力;另一方面,此项指标是大学评价不可或缺的评价要素,研究结果对丰富我国大学评价体系具有一定现实意义。

本文虽然构建了一套高校社会声誉量化评价指标体系,但仍然存在一些不足之处。一方面,本研究对数值型数据采用标准化处理方式,差异性较大的数据样本可能会对社会声誉评价结果产生一定影响。另一方面,指标体系构建时,限于数据可获取性和准确性,并未将所有网络媒体和平台指标纳入指标体系,如360搜索指数、微博指数、微博平台等数据指标。随着网络媒体技术的发展,将会涌现出更多的可操作指标用于高校声誉评价,也将会促使高校社会声誉定量评价得到进一步完善和推广。

参考文献

- [1]陈卓.大学声誉制度研究——基于制度主义的视角[J].中国高教研究,2016(3):73-79,95.
- [2]舒颖岗.大学声誉培育与高水平大学建设[J].国家教育行政学院学报,2011(12):21-25.
- [3]蒋丽霞.高校无形资产转化研究[D].广州:华南理工大学,2017.
- [4]袁浪华.我国民办本科高校的声誉困境及其提升策略[J].黑龙江高教研究,2020,38(1):1-7.
- [5]段婕.我国高等院校声誉管理的动态模型研究[J].高教探索,2014(1):36-39,68.
- [6]WEBSTER D S. Academic Quality Rankings of American Colleges and Universities[M]. Springfield: Charles C Thomas Publisher, 1986: 237-245.
- [7]王林春,江育恒.声誉调查在国外大学排名中的应用及反思[J].世界教育信息,2019,32(20):40-47.
- [8]CLARKE M. Some Guidelines for Academic Quality Rankings[J]. Higher Education in Europe, 2002, 27(4): 443-459.
- [9]STUART D L. Reputational Rankings: Background and Development [J]. New Directions for Institutional Research, 1995(88): 13-20.
- [10]DOMETRIUS N C, HOOD M V, SHIRKEYK A, et al. Bugs in the NRC's Doctoral Program Evaluation Data: From Mites to Hissing Cockroaches[J]. PS: Political Science and Politics, 1998, 31(4): 829-835.
- [11]STEPHANIE L. Reputation Without Rigor[EB/OL]. (2009-08-19) [2021-05-05]. <https://www.insidehighered.com/news/2009/08/19/reputation-without-rigor>.
- [12]ESPELAND W N, SAUDER M. Rankings and Reactivity: How Public Measures Recreate Social Worlds[J]. American Journal of Sociology, 2007, 113(1): 1-40.
- [13]何亚明.高校声誉管理:内涵、意义与策略[J].黑龙江教育(高教研究与评估),2007(Z2):140-142.
- [14]袁立莉.论社会声誉对现代大学发展的影响[J].黑龙江教育(高教研究与评估),2008(Z1):6-7.
- [15]叶琳娜,陈纲.社会声誉与高校筹资工作:关系、作用与机制[J].理工高教研究,2008(5):50-52.
- [16]季小天,江育恒,赵文华.大学社会声誉的形成机理初探:基于“身份—形象—声誉”分析框架[J].江苏高教,2019(8):21-30.
- [17]赵振杰.新建本科院校的声誉管理问题及其应对策略[J].教书育人(高教论坛),2017(24):9-11.
- [18]马焕灵,黄丽静.研究生导师立德树人职责履行评价指标体系的构建[J].现代教育管理,2020(8):84-92.
- [19]刘果,汪小伢.标题特征对数字媒介内容传播效果的影响——基于新闻评论类微信公众号标题的实证研究[J].新闻与传播评论,2020,73(6):29-39.
- [20]李军海,卢立飞.基于微信公众平台的大学生党员教育研究[J].高教学刊,2020(32):189-192.
- [21]清博指数.微信传播指数 WCI(V14.2)[EB/OL]. [2021-02-26]. <http://www.gsdata.cn/site/usage>.
- [22]吃语.大数据报告:知乎百万用户分析[EB/OL]. (2017-07-23) [2020-09-12]. <https://www.jianshu.com/p/6815e9553aba>.
- [23]李长荣,纪雪梅.面向突发公共事件网络舆情分析的领域情感词典构建研究[J].数字图书馆论坛,2020(9):32-40.

收稿日期:2021-07-31

基金项目:国家自然科学基金面上项目“基于多源数据的高校声誉评价研究”(72074039);教育部人文社会科学研究青年基金项目“基于多源数据的‘双一流’政策对高校的影响机制及影响效果评价研究”(20YJC880057)。

作者简介:刘盛博,大连理工大学高等教育研究院副教授;刘佳新,大连理工大学高等教育研究院硕士研究生;刘苗苗,大连理工大学高等教育研究院博士研究生,本文通讯作者。