

大数据环境下高校贫困生精准资助模式初探

吴朝文¹, 代 劲², 孙延楠²

(1. 电子科技大学 马克思主义教育学院 四川 成都 611731; 2. 重庆邮电大学 软件工程学院 重庆 400065)

摘要: 在中央“精准扶贫”的总体部署下,实现高校贫困生的精准资助具有重要意义。现行贫困生认定政策采取定性和定量相结合的方式,但定性环节的民主评议具有较多不确定性,同时定量标准也缺乏客观依据。基于高校“智慧校园”的建设发展,为运用大数据技术开展贫困生资助工作创造基础条件。随着学生日常生活的数字化,大数据客观真实地反映学生的生活轨迹和行为特征,在现有贫困生认定情况的基础上,通过对学生消费行为的特征分析,实现生活状况的评价。配合学生资助管理部门,将大数据分析的结果应用于对贫困生资助体系的验证性评估和特殊困难学生群体的预警,实现对现行贫困生认定有效的补充,从而实现高校贫困生的精准资助。

关键词: 智慧校园; 精准扶贫; 大数据; 贫困生资助

中图分类号: G645

文献标志码: A

文章编号: 1003-2614(2016)12-0041-04

扶贫是加速社会建设、构建和谐社会,体现中国共产党领导和中国特色社会主义制度优越性的重要工作。2015 年 11 月,中央扶贫工作会议召开,习近平同志指出,脱贫攻坚已经到了啃硬骨头、攻坚拔寨的冲刺阶段,要坚持精准扶贫、精准脱贫,重在提高脱贫攻坚成效。扶贫工作贵在精准,重在精准,成败之举在于精准。精准扶贫是全面建成小康社会、实现中华民族伟大复兴“中国梦”的重要保障。教育公平是社会公平的重要基础,高校家庭经济困难学生资助是扶贫工作的重要组成部分。

一、现行高校学生资助政策体系

国家高度重视高校家庭经济困难学生资助问题,将建立健全国家学生资助政策体系,保障所有家庭经济困难学生都有平等接受教育的机会,作为促进教育公平的重要举措和途径。目前,国家在高校设立国家奖学金、国家励志奖学金、国家助学金、国家助学贷款(包含校园地国家助学贷款和生源地信用助学贷款)、勤工助学、学费减免等多种形式有机结合的高校家庭困难学生资助政策体系。此外,国家还积极引导和鼓励社会团体、企业和个人面向高校设立奖学金、助学金,共同帮助高校家庭经济困难学生顺利入学并完成学业。

在政策体系中,将家庭经济困难学生界定为“学生本人及其家庭所能筹集到的资金,难以支付其在校期间学习和生活基本费用的学生”。对于认定程序规定,学生首先向学校申报家庭经济困难,并提供家庭所在地的乡(镇)或街道民政部门加盖公章予以确认的《高等学校学生及家庭情况调

查表》。然后学校根据有关部门设置的标准和规定的程序,组织学校、学院(系)、班级分层分级以民主评议的方式认定。

现行高校贫困生资助政策体系结构完善,资助力度大,对于保障所有家庭困难学生都有平等接受教育的机会,促进教育公平,传递党和政府的关怀、社会主义大家庭的温暖发挥重要的作用。但是,在贫困生认定中还存在贫困状况核实不够客观、评价标准不够统一、主观评价不够准确等现实问题。高校贫困生认定工作需要建立对学生贫困状况客观评价体系,对现行认定方式做有益的补充,从而更加精准地评定资助对象,更加精准地安排资助项目,使国家资助项目发挥最大的效能。

二、“智慧校园”为“精准扶贫”提供大数据环境

1994 年,我国正式接入国际互联网,随着中国教育和科研计算机网(CERNET)的建设发展,20 世纪 90 年代高校进入校园网建设的热潮。在进入 21 世纪后,随着基于校园网的应用系统建设成为高校信息化建设新主题而进入数字校园建设阶段。“十一五”期间,以统一身份认证、统一门户和公共数据平台为特征的数字校园建设成为各高校的建设热点,从整体上提升高校信息化水平。到“十二五”中期,随着云计算、物联网、移动互联等信息技术的普及化,传统意义上的数字校园建设出现新的提升,建设“智慧校园”不再是理念口号,已经成为许多高校新的建设目标和实际行动。可以说,当前的高校信息化建设已经进入以“智慧校园”建设为主题的新时期。随着社会不断进步,信息技术不断发展,新技

收稿日期:2016-11-02

基金项目:国家社科基金西部项目“移动互联网时代高校思想政治教育路径创新研究”(编号:14XKS038)。

作者简介:吴朝文,电子科技大学马克思主义教育学院,博士研究生,研究方向:马克思主义大众化;代 劲,重庆邮电大学软件工程学院副教授,博士,研究方向:数据挖掘、大数据分析;孙延楠,重庆邮电大学软件工程学院,研究方向:数据挖掘。

术不断完善,以及上层应用系统不断增加,校园信息化环境中积累的数据也在迅速膨胀,已经初步形成一个典型的校园大数据环境。

在“智慧校园”中,大学生的学习生活处在信息化、网络化的环境中,其个人信息、行为轨迹、习惯爱好、表现结果等都被应用系统和设备记录存储,实现现实生活的数字化。数据是大学生的自在状态下自我行为的对象性反映,展现一个客观存在但又一直不能被大家认识的现实世界,赋予人们一个认识的崭新维度,深入探索学生成长成才规律。

对于家庭经济困难学生学习生活在共同的“智慧校园”中,其家庭基本信息、欠费贷款情况、日常生活习惯、校园消费明细、学习获奖成绩、奖助学评定情况等信息都被各种系统记录并数字化。这些数据真实、客观、全面地反映学生在校期间的行为表现,通过数据关联分析得到家庭经济困难学生行为的基本特征,可以为高校贫困生认定工作提供有力的支撑。

三、基于校园大数据环境下贫困生认定

通过对高校“精准扶贫”意义的认识、高校学生资助政策和“智慧校园”大数据的分析,探索开展应用大数据的方法认识贫困生状况。贫困生的界定基于以下思路:一是根据已经认定的贫困生历史数据,取三分之一作为训练样本集 T_r ,三分之二作为测试样本集 T_e ;二是采集全校在一定时间段(比如一个季度)的全体学生的食堂消费记录情况,根据训练样本集 T_r 得到贫困标准线;三是依次计算全体学生与贫困标准线之间的距离,经过归一化处理后得到贫困指数;四是采用测试样本集 T_e 进行精度评估。

整个界定具体步骤如下:研究选取重庆邮电大学作为样本,学校为全日制普通本科高校,是教育信息化示范单位,现有本科生两万余人,研究生三千余人。样本具有普遍性和代表性,具备较好的基础数据条件和一定的数据规模。

1. 数据采集和清洗。学校现有二十余个应用系统在线运行,“校园一卡通”普及使用(见表1),丰富的信息化应用建立学生完整的数据信息(见图1),为大数据的应用提供丰富的资源。

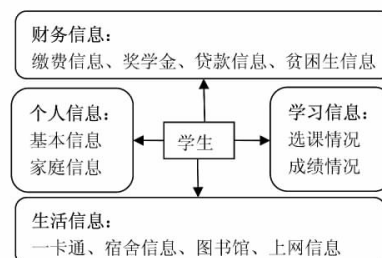


图1 “智慧校园”学生信息的组成

针对贫困生认定工作,将以上信息分为三个类别:一是基础信息,包括学生和家庭的的信息;二是轨迹信息,包括生活中一卡通消费信息等;三是结果信息,包括缴费信息、奖贷勤补信息和贫困生认定信息。在现行贫困生认定程序中,充分考虑基础信息和结果信息,同时通过民主评议的方式反映轨迹信息,最后综合得到贫困生认定结果。现行贫困生认定程序中主观性最强的不确定信息是轨迹信息,这是大数据方法需要解决的主要问题。

基于在校学生校园全部数据的分析,重点采集2015年秋季贫困生认定名单,共有贫困生5593人,其中特困生994人;采集同期2015年9-11月食堂一卡通消费信息360万余条。

表1 一卡通数据示例

一卡通号	学生姓名	刷卡时间	刷卡类型	刷卡金额	刷卡地点
16 * * * 34	邹 * *	2015/10/15 07:08	消费	5.5	食堂一
16 * * * 34	邹 * *	2015/10/15 11:32	消费	11.2	食堂二
16 * * * 34	邹 * *	2015/10/15 17:29	消费	12.0	食堂三

一卡通信息完整客观地记录学生在校期间消费的轨迹信息,其中的用餐信息是最刚需的信息,消费情况最能反映生活水平。在分析前对数据进行清洗,分早、中、晚等三个时段对数据进行合并、去噪等处理。

2. 贫困生消费特征分析。根据贫困生认定数据情况首先将学生分为三个类别:特困生、贫困生和非贫困生。同时,按照一卡通消费数据情况将学生食堂用餐分为三个观测点:用餐次数、平均用餐金额和用餐金额波动情况。在分类的基础上对三个类别学生的消费情况进行分析,从数据分析结果(见图2)可以看出,在没有预设结果的情况下,从客观数据真实反映出贫困生在食堂用餐中的趋势。一是贫困生在食堂的用餐次数显著高于非贫困生,其中,特困生较非贫困生

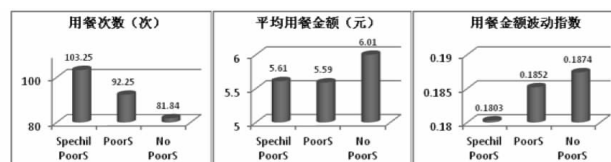


图2 学生食堂用餐情况

用餐次数高出26%,学校食堂仍然是贫困生首选的用餐场所。二是贫困生在食堂的平均用餐金额略低于非贫困生。三是贫困生每次用餐金额较非贫困生更为稳定。

由此,可以描述出贫困生食堂用餐的特征。贫困生大部分时间仍然在学校食堂用餐,并且平均用餐金额低于非贫困生,同时每次用餐金额更加稳定。大数据得到的结论和日常

思维的判断是基本一致的,但最大的不同是,大数据将一切概念思维进行量化,而量化为进一步分析奠定基础。

3. 贫困指数的确定。从数据分析的结果中,我们选择特困生和非贫困生两个类别学生的用餐次数、平均用餐金额和用餐金额波动等三个观测点,共六个指标进行分析(见图3)。通过已经认定的贫困生得到相应的贫困标准线,比较每名与贫困标准线的偏离程度。偏离程度越小,则数据更接近特困生,那么该名学生与贫困生的相似度更大;反之,则与非贫困生的相似度更大。

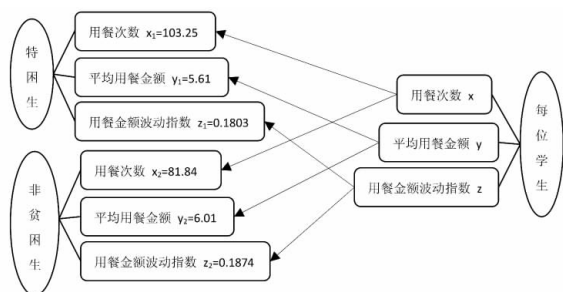


图3 学生贫困认定方法

通过对数据的归一化处理 and 欧氏距离方法得到每位学生的一个相似度数据 L , L 值越低, 该名学生与贫困生的相似度越高, 这为用餐数据评估学生生活状态提供量化指标。因为 L 值范围太大, 大小的概念是相对的, 使用不够直观, 所以引入贫困指数。

$$\text{贫困指数公式: } P_i = 1 - \frac{L_i - \min L}{\max L - \min L}$$

贫困指数反映学生与贫困标准线的偏离程度, 可以更加直观地体现学生贫困情况。在贫困指数 P_i 中, L_i 为学生与贫困标准线的距离, $\min L$ 、 $\max L$ 分别为与贫困标准线的最近距离与最远距离, 取值范围在 0-1 之间。该指数越高, 该名学生贫困的概率越大; 反之, 指数越低贫困的概念越低。学生生活状态的定性评价通过大数据的方法转化为定量评价, 为评估贫困生资助体系运行状况和贫困生预警实现精准扶贫提供重要依据。

四、贫困生大数据结果的应用

基于消费数据的分析, 获得每个学生消费状态的量化指标, 并且经过标准化处理得到每个学生的贫困指数。所有数据来源于学生生活的现实活动, 通过大数据的方式, 帮助我们认识到客观存在但是传统方法不能认识的规律, 最终认识必须作用于实践才彰显出其价值。以下从宏观评价和微观作用两个方面探索大数据结果的应用。

1. 评估贫困生资助体系运行状况。将全体学生的贫困指数由高到低排列, 得到反映贫困指数分布情况的贫困指标图(见图4)。图中横坐标为学生人数, 纵坐标为学生贫困指数。在指标图的光滑曲线中, 根据拐点斜率设定三个分割点, 将整个曲线划分为四段, 并且定义第一段为贫困预警区间, 第二段为贫困区间, 第三段为正常区间, 第四段为非贫困区间。

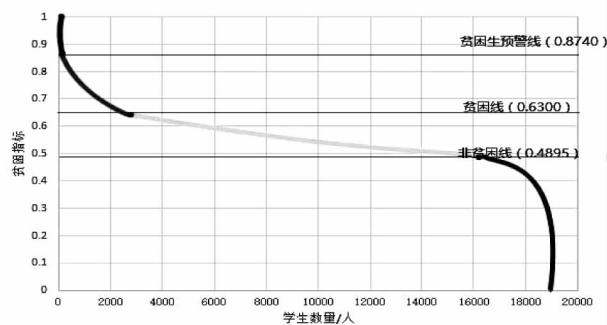


图4 所有本科生贫困指标图

同时, 将贫困指数由高到低的名单与贫困生认定的名单进行对比, 得到贫困生指标与贫困生名单符合度表格(见表2)。

表2 贫困生指数与贫困生名单符合度

贫困指标 靠前人数	其中的贫 困生人数	其中的特 困生人数	贫困生 符合度
100	100	17	100.00%
500	494	78	98.80%
1000	971	165	97.10%
2000	1830	312	91.50%

通过以上图表, 结合日常工作实际情况, 对学校贫困生资助体系状况做出如下评估: 一是现行贫困生资助体系运行良好, 发挥很好的作用。70%的学生在食堂用餐消费方面处于平均值附近, 没有明显差异; 15%的学生生活条件较好, 用餐消费较高; 15%的学生用餐消费水平略低于平均水平。二是贫困生认定组织规范有效, 认定结果与学生实际生活状态具有很高的符合度。前100名贫困指标较高的学生全部是认定的贫困生, 在前2000名贫困指标较高的学生中, 91.5%是认定的贫困生。三是在现行资助体系中, 特困生资助力度大, 半数以上特困生在用餐消费水平方面和全体学生平均值接近, 仅有2%的特困生进入预警区域, 表明特困生用餐问题基本解决。

2. 贫困生预警与精准资助。依据贫困生指标图(图4)中的贫困生预警线进行数据挖掘, 线上共有学生121名。将这部分贫困预警学生名单与贫困生认定名单进行对比。121名学生均是认定的贫困生, 其中, 特困生18名。比对贫困生认定过程中填写的《高等学校学生及家庭情况调查表》信息得到如下统计(见图5)。

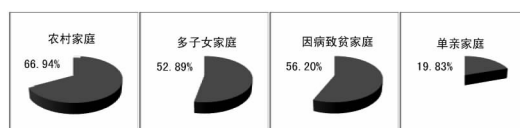


图5 贫困产生原因分析

选取其中典型学生, 分析行为轨迹, 对处于贫困生预警范围的学生进行画像, 了解学生生活状态。将学生的行为分为时间、地点、用餐经费等三个指标, 同时对照食堂价格表估计用餐情况, 得到表3。

表3 贫困生消费情况表

一日三餐	通常用餐时间	通常用餐地点	用餐平均金额	推测用餐情况
早餐	06:20	食堂一	1.48 元	两个馒头
午餐	11:40	食堂二	5.41 元	一荤一素
晚餐	17:30	食堂三	5.13 元	一荤一素

通过以上对贫困预警线范围内学生的整体情况分析和典型学生画像,我们得到以下结果:一是在坚持资助特困生的同时,需要特别关注资助等级低,并且未获奖学金的困难学生。在目前用餐消费水平偏低的学生群体中,这部分学生占主要部分;二是在贫困生资助工作中需要更加关注农村家庭、多子女家庭和因病致贫家庭的学生;三是用餐消费水平偏低的学生群体基本温饱问题已经得到解决,但生活质量水平还比较低,需要进一步加大资助力度。

校园大数据的应用需要建立学校相关部门的联动机制。数据来源需要学生工作系统、后勤保障系统和财务管理系统的协调支持,分析过程需要技术和管理的结合,结果处理也需要各个系统共同作用。大数据分析得到的“精准资助”名单提供学生工作系统,通过安排专项勤工俭学项目的方式,在“征求学生本人意见”的前提下增加资助项目。

在高校贫困生资助工作中,采取大数据技术以最客观、最全面的方式还原学生现实生活状态,为贫困生认定提供量化的指标,是对现行资助体系的有效补充。通过对学生消费情况的大数据分析,既能够对高校贫困生资助体系进行验证的评估,也可以挖掘消费水平偏低的学生群体,从而有针对性地采取资助措施,实现高校贫困生资助的“精准扶贫”。贫

困生资助是一项面向人的工作,大数据也是一种面向人的现实活动的认识方法。大数据能够为贫困生认定等提供支撑和有效的补充,但是做好贫困生资助工作从根本上还需要相关学生资助工作系统的教师深入关心学生、了解学生、充分认识工作对象,才能最大限度地形成合力,实现贫困生资助工作的良性发展。

大数据的本质是人的自由活动的数字化,为认识人的现实活动规律提供崭新的维度,开发利用大数据将获得重要价值的源泉。在高校“智慧校园”建设发展的背景下,校园大数据发展已具备良好的条件。在优化数据构架、整合数据资源、规范数据标准、开放数据共享的基础上,建立跨工作系统、跨业务流程的工作机构,通过业务和技术的结合,围绕人才培养的中心工作,开发利用大数据宝贵资源,将为提高高校管理水平、服务师生发展发挥重要的作用。

参考文献:

- [1] [英]维克托·迈尔-舍恩伯格,肯尼斯·库克耶. 大数据时代[M]. 盛杨燕,周涛,译. 杭州:浙江人民出版社,2013.
- [2] 胡树祥,谢玉进. 大数据时代的网络思想政治教育[J]. 思想教育研究,2013(6).
- [3] 黄欣荣. 大数据哲学研究的背景、现状与路径[J]. 哲学动态,2015(7).
- [4] 蒋东兴,等. 大数据背景下的高校智慧校园建设探讨[J]. 华东师范大学学报:自然科学版,2015(S1).

Research on the Targeted Poverty Reduction Model of the Needy Undergraduates in the Big Data Environment

WU Chao-wen¹, DAI Jin², SUN Yan-nan²

(1. Marxist Education Institute, The University of Electronic Science and Technology of China, Chengdu 611731, China;

2. School of Software Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

Abstract: Targeted funding for needy undergraduates plays an important role in the central government's overall scheme of precision-targeted poverty alleviation. The current approach to identify poverty students is evaluating the situation of the students qualitatively and quantitatively. However, the voting session of the qualitative process is full of uncertainty and the qualitative standard is lack of objective basis. The Smart Campus program offers a basis data for the evaluation of the financial situation of a student by the big data technology. The big data relevant to the daily life of a student reflect his habitual behaviors and life traces, by which the consuming behavior of the student could be analyzed to evaluate the financial situation. It could also be used as an early warning of extremely impoverished students. This will be an effective supplement to the affirmation of the poor students in the work of a precision-targeted subsidization.

Key words: smart campus; targeted poverty reduction; big data; penury students' financial assistance

[责任编辑: 郝文斌]