

基于食堂消费大数据分析的精准资助对策研究

——以湖南科技大学为例

赵前程^a,王艳^b,梁宏军^b,李兵^b

(湖南科技大学 a.党办校办;b.后勤管理处,湖南 湘潭 411100)

摘要:高校家庭经济困难大学生的精准认定一直是高校资助工作中的难题,大数据分析为解决这一难题提供了方法论支撑。高校食堂的消费数据是大学生消费大数据的重要组成部分,利用大学生食堂消费大数据提升学生的精准化认定和资助,是当前一个热点。以湖南科技大学为例,利用大学生食堂就餐率和平均消费额数据进行统计分析,提出贫困大学生认定和资助基本方案,并以此为基础开展访谈与座谈调研,进行探究学生就餐率与就餐满意度关系的问卷调查,提出更加科学和精准的数据分析指标,建立食堂消费大数据分析的“决策树模型”,提升资助精准度。

关键词:食堂消费;大数据;决策树;精准资助

中图分类号:G47

文献标志码:A

文章编号:1674-5884(2023)04-0015-07

1 引言

1.1 背景介绍

高校对家庭困难学生的精准化认定和资助面临一些挑战和问题。首先,传统的认定方法过于依赖静态的信息,如家庭收入和贫困指标,无法全面了解学生的真实需求。其次,传统的资助方式个性化和差异化程度较低,不能满足不同学生的需求。此外,精准化认定和资助涉及大量的信息收集和处理工作,如果仅仅依靠传统的手工方式进行,会耗费大量的时间和人力成本。

大数据技术的快速发展为高校精准化认定和资助家庭困难学生工作带来了新的机遇。大数据是指庞大而复杂的数据集合,通过应用高效的技术和算法进行分析,可以揭示隐藏在数据背后的有价值的信息。大数据在教育领域的应用日趋明显,主要体现在学生群体分析、学生学业风险预测与干预、根据学生的学习情况和特点进行个性化教育和辅导、教学质量评估等方面。马永健《高校智慧校园建设中大数据技术的应用》探讨了高校智慧校园建设中大数据技术的具体应用^[1]。张琼霞

《基于一卡通数据的校园大数据技术的应用》对一卡通系统数据进行聚类分析并构建数据分析模型,利用聚类算法对学生进行各种分类^[2]。

大数据在教育领域的应用为高校精准化认定和资助政策的改进和优化提供了新的解决方案。高校学生管理和资助精准化可以利用大数据进行管理优化,如通过大数据分析可以更精确地确定资助对象,然后根据学生的真实需求为其提供个性化资助方案,从而强化资助的效果和公平性等。李剑富等《高校贫困生资助政策的历史回顾、演进逻辑及未来走势》对高校贫困生资助政策的历史演进、未来走势和逻辑进行了研究,并通过对动力机制、价值取向、制度规范和话语规则的分析,为未来资助政策的发展提供了建议^[3]。董波《高校学生资助工作存在的问题及对策》聚焦高校学生资助工作存在的问题并提出相应的对策,指出高校学生资助工作中存在的问题包括经济困难生认定材料造假、诚信意识淡薄、心理自卑等,为了解决这些问题,需要完善资助工作机制,实现资助育人模式的转变^[4]。张存禄等《食堂消费大数据

收稿日期:2023-07-01

作者简介:赵前程(1969—),男,安徽肥西人,湖南科技大学党委委员、副校长,教授,博士,主要从事机械测量和大数据分析等方面研究。

可以精准识别贫困生吗?——基于本科生行为数据、行政数据和问卷数据的实证研究》围绕食堂消费大数据对贫困生的精准识别展开研究,通过分析某“双一流”大学本科生的食堂消费数据和家庭基本信息问卷调查,结合统计学习模型,探讨利用食堂消费大数据识别贫困生的问题,研究得出:食堂消费大数据识别贫困生的精准率有待提高,但结合其他数据和模型可以改进识别结果^[5]。

综上所述,大数据在高校的应用研究涉及不同的主题和研究对象,涵盖学习分析、学生行为建模、在线教育、教师支持与培训、教育决策支持等多个方面,这些研究进展为教育改革和教育质量的提升提供了新的思路和方法。特别是在高校家庭经济困难学生资助方面的研究,从资助政策的演进逻辑和问题、资助工作中存在的问题及对策、食堂消费大数据识别等方面提出一系列观点和建议,这些文献为高校家庭经济困难学生资助研究提供了新的视角和方法,并为具体实践提供了有益借鉴。

1.2 研究目的和重要性

高校家庭经济困难学生精准化认定和资助对于学生个体发展和整体教育公平性至关重要。个体发展方面,每个学生的背景和需求不同,精准化认定和资助可以确保他们获取适当的支持和资源,为个体发展提供平等的机会。同时,精准化认定和资助也有助于提升整体教育公平性,为经济困难或特殊困境学生提供帮助,减轻负担,有助于建立更加公平、包容和多样化的高等教育体系。

利用大数据技术提升精准化认定和资助水平的必要性不言而喻。大数据技术可以收集和分析大量的数据,包括学生的学习成绩、家庭背景、社会经济状况等信息。通过对这些数据的分析,可以建立相应的模型和算法,精确评估学生的资助需求,确保资金分配的公正性和准确性。大数据技术还可以帮助高校通过学生在食堂的消费大数据来了解学生的消费行为和生活方式,从而更好地满足他们的实际需求,提供个性化的支持和服务。

本文的研究目标是通过分析高校食堂消费数据来探索提高精准化认定和资助水平的途径。高校食堂消费数据是一种重要的学生行为信息来源,反映了学生的消费偏好、经济状况和生活习惯等。通过对这些数据的挖掘和分析,可以识别学

生群体的消费特征和经济需求。进一步开发基于学生食堂消费数据的算法和模型,以便更准确地确定学生的资助需求,并提供个性化资助方案。

具体研究包括以下几个方面:收集和整理高校食堂消费数据;利用食堂消费数据推断学生的经济状况和资助需求;基于食堂消费数据和其他相关信息,确定合适的资助标准和分配机制。这些研究为探索如何通过分析高校食堂消费数据来提高精准化认定和资助水平提供方法。研究主要通过统计分析方法、机器学习等技术手段分析问题,从而强化学生资助的准确性和公平性。

1.3 研究方法和数据来源

本文主要使用文献调研方法来进行大数据分析中指标选取的定性分析,使用问卷调查方法进行学生满意度的数据收集,在问卷数据分析中主要使用描述性统计分析和因子分析的方法来获取问卷数据中的有用信息。在食堂消费数据的大数据分析中,使用机器学习中常用的特征提出方法并结合文献分析中的定性分析方法来提取特征,然后针对提取的特征,使用决策树法得到贫困学生的精准认定模型。食堂就餐数据主要为湖南科技大学10个食堂2022年9月至12月的刷卡消费数据,数据跨度100天,共计635万余条,包含的字段较多,能够挖掘的信息较丰富。

2 湖南科技大学精准化资助案例研究

2.1 初次精准化资助和就餐数据统计

学校从2023年开始每年拨款100万元给学生发放就餐补助,一年分4次发放,每次25万元。为了让餐补能够发放到真正需要的同学手中,学校试图利用食堂的学生消费大数据,使用大数据分析方法来餐补学生进行精准认定。首先,对食堂消费数据进行了分析。对数据进行简单的描述统计后,发现学生的消费金额区间分布如图1所示。

从图1可以看出,女生每日在食堂的消费额集中在21~30元之间,男生每日在食堂的消费额集中在26~35元之间,由此可以看出,每日的消费金额是可以部分反映学生消费能力的。

对学生的就餐率和日均消费额进行统计分析,得到的分布图如图2所示。

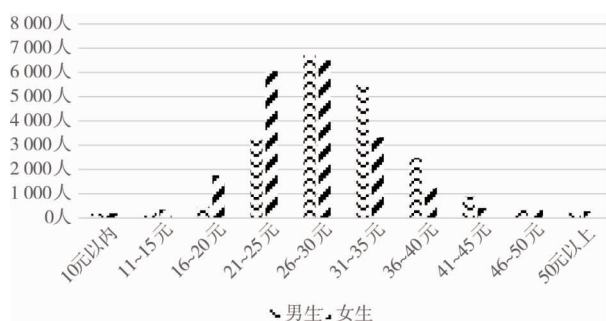


图1 全校学生日均消费区间分布

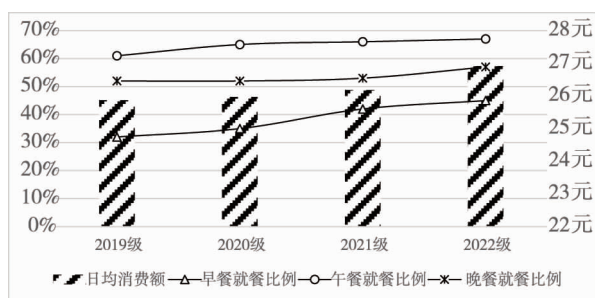


图2 不同年级学生就餐率及日均消费额对比

从图2可以看出,学生的就餐率和日均消费额有一定的关系,就餐率也比较稳定。同时,通过走访发现,学生中每日生活消费额度低的同学更倾向于在学校食堂就餐。

综合这些数据的分析结果,结合走访学生得到的信息,选择“就餐率”和“日均消费额”两个指标来进行需要补助学生的认定。为精准高效地完成餐补发放任务,我们使用数据挖掘方法建立餐补发放判定模型。以学生在食堂的餐均消费和就餐频率作为特征,构建深度为2的决策树模型,模型结构图如图3所示。

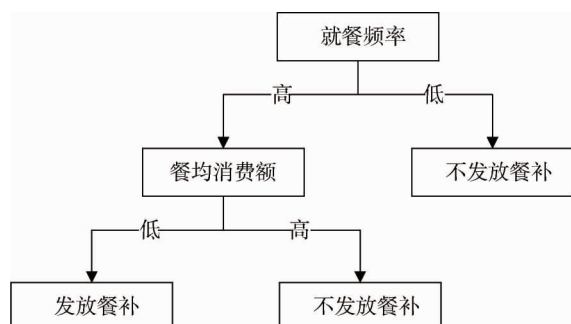


图3 餐补发放判定决策树

2.2 餐补发放工作实际开展情况

根据上述模型初步选定学生后,学校进一步筛选就餐率较高、日均消费额较低的学生进行就餐补助发放。截至目前,学校已经开展两次餐补

发放,具体情况见表1和表2。

表1 2023年第一次餐补发放情况统计

发放标准	女生人数	男生人数	金额合计(元)
100元/人	812	806	161 800
200元/人	269	172	88 200
合计	1 081	978	250 000

表2 2023年第二次餐补发放情况统计

发放标准	女生人数	男生人数	金额合计(元)
100元/人	688	630	131 800
200元/人	331	260	118 200
合计	1 019	890	250 000

从两次餐补发放的统计表可以看出,每次发放人数在2 000人左右,女生人数稍高于男生人数,分别为21 586和21 663,可见,存在男女生发放比例略失衡的情况。

2.3 学生访谈和调研结果分析

学生就餐补助方案第一次实施后,大部分学生对实施方案反响强烈,称学校的举措“有爱又有心”,还表示“这才是大数据的意义”,“一口口吃下去的全是爱”,但小部分学生对方案实施标准的科学性提出质疑。社会上对这一方案的实施赞扬较多,同时也提出了很多建设性建议,如统计范围是否可以更加全面等。结合这些意见和建议,为了更好地保证资助对象选择标准的科学性,学校在学生中进行了多次深入访谈和问卷调查,以探究影响资助对象选择的指标。

在消费数据的分析和数据来源的探究中发现,存在消费数据来源不够全面的问题。学校食堂虽然安装有500余台收款机,但受技术限制,只支持校园卡、企业微信付款码、云闪付等支付方式,不支持微信、支付宝等学生较常用的支付方式,此外,个别食堂和档口存在私自收款行为,这部分私自收款数据学校无法统计,从而导致数据来源不够全面。

在根据指标进行发放学生认定的过程中发现,消费数据统计指标利用率不够高。在大数据统计及补助发放过程中,学校选取食堂就餐率和日均消费额两个指标进行决策树建模识别需要补助的学生,其方法基本上是科学的。但在对识别出的学生进行统计分析的过程中,发现存在性别

略有不平衡的问题,后续实际发放中需要结合学工处提供的学生信息进行调整,如在改进模型中加入性别指标,从而实现家庭经济困难学生的有效识别。此外,结合学生访谈和问卷调研的情况发现,只考虑这两个指标导致对学生家庭经济情况的识别不够精准,因此,需要加入更多维度的有效指标。

2.4 就餐满意度问卷调查和数据分析

2.4.1 设计问卷调查的目的和问卷数据检验

为了解学生就餐满意度和影响满意度的主要因素,特设计了关于湖南科技大学校内 10 个食堂满意度的调查问卷。满意度量表借鉴了通常的餐饮满意度量表,将满意度度量分为八个维度,同时,为了方便学生填写,将量表设计为三级量表。研究将问卷通过网络形式在线发放,结合线下食堂随机截取学生填写,最后收集到问卷 1 180 份。对问卷的信效度检验结果如表 3 所示。

表 3 问卷信度检验

克隆巴赫 Alpha	项数
0.927	90

从表 3 可以看出,克隆巴赫信度系数高于 0.9,表示测验的内部一致性强,即同一测试项在不同条件下得到的结果相似。这说明该问卷信度相当高,即问卷数据可信度高。

KMO 和巴特利特检验结果如表 4 所示。

表 4 KMO 和巴特利特检验

KMO 取样适性量数		0.889
巴特利特球形度检验	近似卡方	8 862.162
	自由度	28.000
	显著性	0.000

对问卷中的量表数据进行效度检验,从检验结果可以看出,巴特利特球形度检验的显著性小于 0.001,KMO 值为 0.889。根据 Kaiser 给出的常用 KMO 度量标准,KMO 值高于 0.8 表示适合进行因子分析。综上,该问卷的量表数据适合做因子分析,且问卷效度高,数据能有效反应所研究的问题。

2.4.2 问卷数据分析结果

对各食堂满意度数据进行因子分析,结果如表 5 所示。

表 5 各食堂满意度因子分析

因子	一粟轩成分得分		三白馆成分得分		五味苑成分得分		六和堂成分得分	
	成分 1	成分 2	成分 1	成分 2	成分 1	成分 2	成分 1	成分 2
饭菜价格	-0.116	0.335	-0.183	0.407	-0.226	0.434	-0.153	0.397
饭菜分量	0.099	0.152	-0.201	0.439	-0.093	0.371	-0.160	0.466
饭菜品质	-0.182	0.470	-0.052	0.330	-0.075	0.324	0.041	0.205
菜品设置	-0.145	0.419	0.005	0.229	-0.124	0.300	-0.006	0.253
卫生情况	0.221	0.044	0.310	-0.088	0.353	-0.116	0.336	-0.091
就餐环境	0.477	-0.237	0.392	-0.191	0.382	-0.189	0.340	-0.118
就餐秩序	0.444	-0.175	0.386	-0.156	0.378	-0.165	0.326	-0.115
服务水平	0.209	-0.007	0.251	-0.012	0.285	-0.042	0.218	0.011

因子	七星阁成分得分		八珍楼成分得分		二至坊成分得分		南园食府成分得分		北庄食府成分得分		昭潭食府成分得分	
	成分 1	成分 2	成分 1	成分 2	成分 1	成分 1	成分 1	成分 1	成分 1	成分 1	成分 1	成分 1
饭菜价格	0.335	-0.149	-0.119	0.311	0.106	0.134	0.135	0.135	0.135	0.135	0.135	0.135
饭菜分量	0.152	-0.291	-0.223	0.555	0.186	0.168	0.180	0.180	0.180	0.180	0.180	0.180
饭菜品质	0.470	-0.005	-0.037	0.343	0.170	0.155	0.185	0.185	0.185	0.185	0.185	0.185
菜品设置	0.419	0.072	0.032	0.285	0.180	0.167	0.141	0.141	0.141	0.141	0.141	0.141
卫生情况	0.044	0.284	0.310	-0.094	0.164	0.159	0.133	0.133	0.133	0.133	0.133	0.133
就餐环境	-0.237	0.393	0.348	-0.179	0.150	0.138	0.132	0.132	0.132	0.132	0.132	0.132
就餐秩序	-0.175	0.367	0.358	-0.155	0.187	0.135	0.136	0.136	0.136	0.136	0.136	0.136
服务水平	-0.007	0.232	0.261	-0.021	0.178	0.141	0.141	0.141	0.141	0.141	0.141	0.141

从表 5 可以观察到,一粟轩、三白馆、五味苑、六和堂、七星阁、八珍楼、学术交流中心的因子被归纳成两个成分,而综合来看这两个成分又大致

代表着学生对环境的满意度和对饭菜的满意度,二者得分差异越大表明该食堂的环境和所提供的饭菜越不相称。二至坊、南园食府、北庄食府和昭

潭食府的因子被归纳成一个成分,表明学生对每个因子的评分比较相似。

针对提炼出的影响满意度的两个因子,可得湖南科技大学全食堂成分得分系数矩阵,如表6所示。

表6 湖南科技大学全食堂成分得分系数矩阵

因子	成分得分	
	成分1	成分2
饭菜价格	-0.215	0.435
饭菜分量	-0.197	0.460
饭菜品质	-0.027	0.289
菜品设置	0.000	0.259
卫生情况	0.305	-0.086
就餐环境	0.400	-0.206
就餐秩序	0.360	-0.144
服务水平	0.264	-0.045

从表6可以看出,学生对食堂的评价因子降维成两个成分:“成分1”受就餐环境、就餐秩序、卫生情况和服务水平影响较大,而不在意与饭菜相关因子的表现;“成分2”受饭菜分量、饭菜价格、饭菜品质和菜品设置的影响较大,却不在意就餐环境相关因子的表现。

再探究各食堂因子得分和餐位重复使用次数间的关系,可得数据如表7所示。

表7 湖南科技大学各食堂得分与就餐比例

食堂	成分1 平均得分	成分2 平均得分	餐位重复使用次数
一粟轩	1.930	1.264	2.4
二至坊	2.089	1.428	1.5
三白馆	2.027	1.547	3.0
五味苑	2.002	1.232	1.2
六和堂	2.085	1.434	2.9
七星阁	2.020	1.822	2.9
八珍楼	2.238	1.309	1.2
南园食府	2.081	1.879	1.4
北庄食府	2.044	1.823	2.0
昭潭食府	1.454	1.179	0.5

“成分1”得分越高表示就餐环境越舒适;“成分2”得分越高表示饭菜越符合学生心意。根据表中数据探讨各个成分对就餐比的影响,对成分得分和就餐比进行相关分析,得到结果:“成分1 平均得分”和“餐位重复使用次数”的相关系数为-0.715 25,说明就餐环境越好,餐位重复使用次数反而越小;“成分2 平均得分”和“餐位重复使用次数”的相关系数为0.200 39,说明饭菜越符合学生心意,餐位重复使用次数越大;“成分1 平均得分”和“成分2 平均得分”的相关系数为

0.391 787,说明“成分2 平均得分”对“餐位重复使用次数”的正向影响被“成分1 平均得分”对“餐位重复使用次数”的负向影响抵消了。

综合上述对满意度影响因素的分析及对餐位重复使用次数与满意度之间关联关系的分析可以看出,要想弥补只使用“就餐率”和“日均消费额”两个指标构建决策树模型来认定需要补助学生导致的性别不均衡缺陷,还需要在利用食堂消费大数据进行认定时增加相关的消费数据指标。

3 餐补发放判定模型的改进

对餐补发放对象的消费数据进行数据挖掘分析时发现,初步建立的决策树模型识别发放对象时出现了男女比例失衡的问题,女生人数稍多于男生。因此,需要对已有模型进行改进,充分考虑不同性别学生在用餐消费上的差异,以及不同性别学生消费价格标准的不同。

根据伍智鑫《基于一卡通数据挖掘的高校贫困生精准扶贫应用研究》中的研究并结合实际工作经验可知,就餐消费的稳定性也是发放餐补的重要依据^[6]。因此,餐补发放判定模型中需将消费的稳定性纳入考虑范畴。统计学中常用标准差来衡量数据的离散程度。考虑到早餐与中晚餐消费金额的明显差距会极大干扰消费金额稳定性的衡量,我们通过学生连续三个月中、晚餐消费有效数据的标准差来衡量学生用餐金额的稳定性。

结合初始模型,增加性别和就餐稳定性两个特征,建立深度为4的决策树模型如图4所示。

为了验证改进后模型的效果,同样利用635万条食堂消费数据进行分析,并对日就餐频数分布和用餐稳定性指标频数分布进行对比,如图5所示。

以消费金额的标准差为衡量就餐稳定性的标准,标准差越低表示就餐金额越稳定,即该学生每餐花费的金额相差不大。左图为原始模型筛选出来的学生消费稳定性直方图,从中可以看出改进前部分同学的消费波动很大,对于这类人群即使平均消费较低也不能认为他们需要补助。原始模型没有考虑到消费稳定性的影响,而改进后模型解决了这个问题。右图为改进后模型筛选出来的学生消费稳定性直方图,从中可以看到筛选出来的群体中不存在消费金额差别很大的个体,这说明改进后的模型可以更精准地定位需要补助的群体。

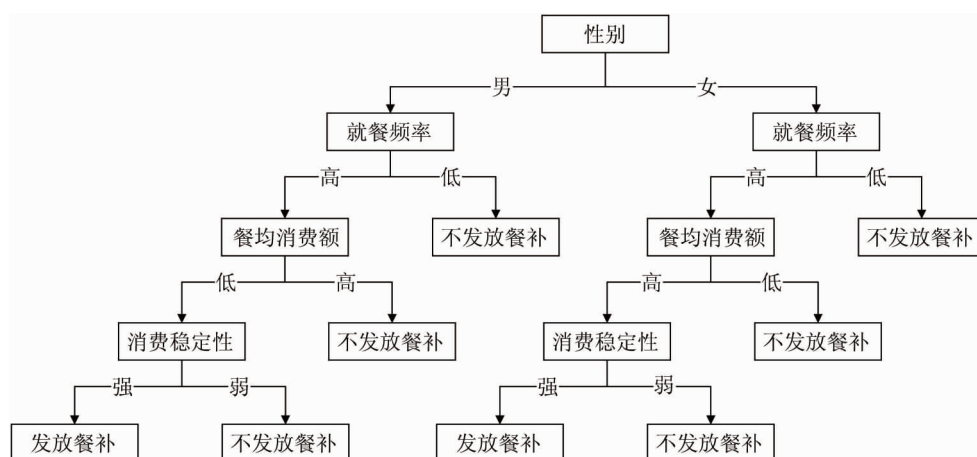


图4 改进后的餐补发放判定决策树

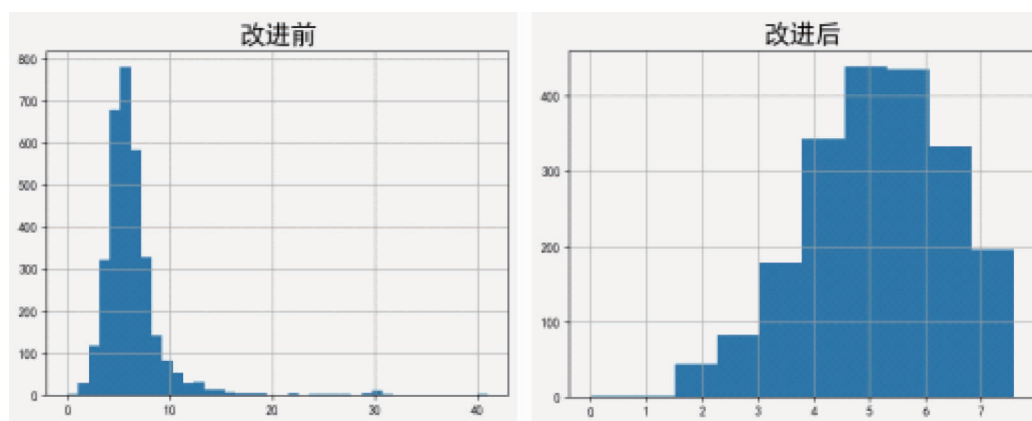


图5 模型改进前后消费稳定性对比

再次利用 635 万条数据进行分析,改进后模型的判定结果中,被判断需要补助的群体中男女生比例趋于平衡。改进前后判别结果中的性别比例对比分析如表 8 所示。

表 8 模型改进前后性别比例的卡方检验

改进状态	统计量	值	自由度	渐进显著性(双侧)
改进前	皮尔逊卡方	199.655	1	0.000
	连续性修正	199.149	1	0.000
	似然比	201.276	1	0.000
改进后	皮尔逊卡方	0.087	1	0.512
	连续性修正	0.075	1	0.526
	似然比	0.087	1	0.512

结果显示:改进前的渐进显著性 <0.001 ,表明需要补助的学生群体性别构成与全体学生性别构成有明显不同;改进后的渐进显著性 >0.05 ,表明需要补助的学生群体性别比例不存在失衡问题。为进一步验证性别对贫困生判定的影响,接下来进行相关性检验。

表 9 模型改进前后性别和贫困指标的皮尔逊相关性

统计量	性别 & 贫困指标	性别 & 改进的指标
皮尔逊相关性	-0.073	0.003
Sig.(双尾)	0.000	0.512
个案数	37198	37198

性别和改进前贫困指标的皮尔逊相关系数为 -0.073 ,渐进显著性 <0.001 ,表明性别和贫困指标之间显著相关;性别和改进后贫困指标的皮尔逊相关系数仅有 0.003 ,渐进显著性 >0.05 ,不能认为性别和贫困指标之间显著相关。

综合上述分析可以看出,改进后模型的识别结果更科学也更符合实际情况。将识别的学生情况与学工和教务部门的数据进行对比,发现此模型的识别准确率达到 90% 以上,相比原模型节约了大量的后期核实工作。

4 精准资助工作的建议

通过本次调查研究发现,对家庭经济困难学生的精准认定是一项长期而艰巨的工作。充分利

用食堂消费大数据进行统计分析并实现精准资助,从而进一步促进高等教育公平公正,应从以下几个方面着手。

一是提升学生在食堂就餐的满意度。学校可通过完善食堂监管制度、控制饭菜价格、畅通反馈机制、加强从业人员队伍建设等进一步提升学生对食堂的满意度,提升学生的食堂就餐率,从而保证学生就餐数据的可靠性。

二是拓宽统计数据来源渠道。通过加强监管或升级收款设备,进一步规范支付方式,最大限度地将在食堂的所有支付方式都纳入学校统一管理,确保统计数据的完整性。

三是拓展数据统计指标的深度和维度。进一步对现有数据进行挖掘,增加就餐稳定性等统计指标,在条件成熟的前提下还可以增加家庭成员、家庭收支等信息,进一步拓展数据挖掘模型统计指标的维度,从而获取更加充分的信息进行认定。

四是加强对困难学生消费行为的监管。通过监控学生的消费情况,可以有效监管学生的消费行为,必要时可以取消具有不良消费行为学生的资助资格。

五是拓宽资助渠道。餐补是学校资助工作的

重要补充。通过消费大数据统计分析形成的精准识别结论,可作为学校评判家庭经济困难学生困难等级的重要依据,为学校学生资助部门等开展助学金发放、助学岗位设定等工作提供参考。其次,结合大数据精准认定结论,动员教职工并走访校友和社会爱心人士,在他们与特别困难学生之间建立起“一对一”精准资助连接线,确保每一个家庭经济困难学生都能得到有效资助。

参考文献:

- [1] 马永健.高校智慧校园建设中大数据技术的应用[J].中国新通信,2020(21):117-118.
- [2] 张琼霞.基于一卡通数据的校园大数据技术的应用[J].莆田学院学报,2022(5):66-71.
- [3] 李剑富,韩岳丰.高校贫困生资助政策的历史回顾、演进逻辑及未来走势[J].黑龙江高教研究,2019(10):79-84.
- [4] 董波.高校家庭经济困难学生资助工作存在的问题及对策[J].黑龙江科学,2021(13):82-83.
- [5] 张存禄,马莉萍,陈晓宇.食堂消费大数据可以精准识别贫困生吗?——基于本科生行为数据、行政数据和问卷数据的实证研究[J].教育经济评论,2022(4):41-56.
- [6] 伍智鑫.基于一卡通数据挖掘的高校贫困生精准扶贫应用研究[D].合肥:安徽农业大学,2019.

A Strategy Study of Targeted Subsidization Based on Big Data of Canteen Consumption: A Case Study of Hunan University of Science and Technology

ZHAO Qiancheng^a, WANG Yan^b, LIANG Hongjun^b, LI Bing^b

(a. Party Office and President's Office; b. Logistics Department, Hunan University of Science and Technology, Xiangtan 411201, China)

Abstract: It is a tricky problem to accurately identify students with financial difficulties in colleges, but the big data has provided a methodological support for addressing this challenge. The data of the college students' canteen consumption plays a necessary role in their consumption data. Currently, it is a heated topic to improve the accuracy of identifying students with financial difficulties and financial aid by using the big data of canteen consumption. Therefore, based on the identification and subsidization of impoverished college students reflected by the data of dining rate and average consumption amount, the paper takes Hunan University of Science and Technology as an example to carry out interviews and talks, as well as questionnaires on the relationship between the dining rate and satisfaction rate. In addition, the paper puts forward data analysis indicators and establishes "decision-making tree" model of big data of canteen consumption by analyzing the consumption data, financial aid data and survey data.

Key words: canteen consumption; big data; decision-making tree; targeted subsidization

(责任校对 王小飞)