

基于熵值模型的高校贫困生认定系统研究^①

吉朝明¹, 王 华¹, 宋铁成²

1. 四川交通职业技术学院 信息工程系, 成都 611130; 2. 重庆邮电大学 通信与信息工程学院, 重庆 400065

摘要: 为客观量化认定高校贫困生的贫困状态, 本文利用校园卡消费数据和学生基本信息, 构建了学生消费恩格尔系数、月均食堂消费水平、早餐月均消费额校均比、中餐月均消费额校均比、晚餐月均消费额校均比、食堂就餐率、助学贷款总额、勤工助学次数等 8 项贫困生评定指标, 根据熵值法客观确定指标权重并计算出学生贫困指数。最后, 结合高校贫困生认定的实际情况, 采用熵值模型分析与人工审核相结合的方式系统设计。测试表明, 该系统能帮助高校有效解决长期以来贫困生认定工作难量化、不客观的问题。

关键词: 贫困生认定; 指标体系; 熵值模型; 认定系统; 研究与设计

中图分类号: TP311

文献标志码: A

文章编号: 1000-5471(2021)04-0122-08

随着我国经济社会和高等教育的发展, 国家高度重视对贫困生的资助, 但如何提升资助育人成效, 对贫困生进行精准识别和对贫困度进行客观计算, 成为长期以来高校贫困生认定工作中的热点和难点问题, 较大程度地影响了贫困生资助的公平性和公正性^[1]。

随着校园卡建设向服务深化、技术融合、平台互连的方向发展, 校园卡承载了金融消费和身份识别两大主要功能, 被广泛应用于食堂、校园超市、图书借阅、考勤等校园业务场景中, 方便了学生的日常生活; 校园卡的数据也成为了获取学生生活和经济情况的有效途径。目前利用校园卡数据进行分析 and 数据挖掘处理, 客观评定学生的经济状况, 已经有一些初步研究。比如, 饶亮^[2]以校园一卡通数据为研究对象, 运用改进了的 Apriori 方法挖掘困难学生群体一卡通中各项数据的关联规则, 发现其中的相关性和潜在规律。文献[3]基于学生基本信息、校园卡消费信息等数据, 运用随机森林算法和决策树进行贫困生的判别。柴政等人^[4]通过分析学生校园一卡通的消费记录, 然后利用基于神经网络的数据挖掘方法进行贫困生的识别。陆桂明等人^[5]基于学生的校园卡消费数据, 结合学生的消费和生活规律, 运用机器学习中 XGBoost(eXtreme Gradient Boosting)模型和主成分分析法、过采样算法, 对高校贫困生进行预测。程茜宇^[6]利用校园一卡通数据, 以深度学习方法为技术依托, 构建精准识别模型进行高校贫困生的识别。何功炳^[7]基于改进的主成分分析法对一卡通消费数据进行简约处理, 再利用 GBDT(Gradient Boosting Decision Tree)算法实现贫困判定模型的训练以及贫困生的预测。

这些研究主要集中在理论研究上, 能运用于实际工作中的方法和系统较少。鉴于此, 本文在已有研究成果的基础上, 利用校园一卡通消费数据和学生基本信息, 构建贫困生认定指标体系, 建立基于熵值法的客观数据模型, 计算学生的贫困指数, 客观地识别出贫困生, 并通过 Java EE 技术实现相应的管理平台并运用于实际工作中。

1 关键技术研究

熵值法通过信息熵来计算指标权重, 被广泛应用于分析评价领域中, 能较科学地考虑数据之间的相关

① 收稿日期: 2020-05-06

基金项目: 国家自然科学基金资助项目(61702065)、四川省教育信息化应用与发展研究中心项目(JYXX18-030)、四川交通职业技术学院科研基金项目(2019-120-07)。

作者简介: 吉朝明, 硕士, 讲师、工程师, 主要从事软件技术、网络安全等方面的研究。

关系和变异程度,从而弥补了不精准的主观赋权模式的不足. 基于熵值模型的贫困生认定系统的研究主要包含 2 个关键问题,一是如何利用校园卡的消费数据和学生的基本信息来科学地构建贫困生认定指标体系;二是如何确定所构建的指标体系的权重.

1.1 贫困生认定指标体系的构建

指标体系的构建是贫困生认定模型中极为重要的一步,而指标设计是否合理将对认定结果起决定性作用. 一般情况下,通过人工经验进行贫困生认定指标体系的构建. 本文采用人工经验与文献分析相结合的方式构建指标体系,根据校园卡消费数据和学生基本信息,构建出学生消费恩格尔系数、月均食堂消费水平、早餐月均消费额校均比、中餐月均消费额校均比、晚餐月均消费额校均比、食堂就餐率、助学贷款总额、勤工助学次数等 8 项贫困生评定的指标,具体指标及其说明如表 1 所示. 其中,恩格尔系数表示学生食堂消费占个人消费支出总额的比重. 依据恩格尔定律公式^[6],本文恩格尔系数可表示为

恩格尔系数(X_1)=食堂消费支出/(食堂消费支出+校园购物消费支出)

其中,食堂消费支出是指学生利用校园卡在学校各食堂进行餐饮消费的支出;校园购物消费是指学生利用校园卡在学校商贸超市等进行购物的消费支出.

表 1 贫困生认定指标体系

指标体系	指标说明	权 重
恩格尔系数(X_1)	三餐消费在整体消费中的占比(%)	W_1
月均食堂消费水平(X_2)	月均食堂消费额/全校学生月食堂均消费额	W_2
早餐月均消费额校均比(X_3)	(总早餐消费金额/早餐消费次数)/(全校平均早餐消费额)	W_3
中餐月均消费额校均比(X_4)	(总中餐消费金额/中餐消费次数)/(全校平均中餐消费额)	W_4
晚餐月均消费额校均比(X_5)	(总晚餐消费金额/晚餐消费次数)/(全校平均晚餐消费额)	W_5
食堂就餐率(X_6)	该生食堂就餐率	W_6
助学贷款总额(X_7)	年均助学贷款金额	W_7
勤工助学次数(X_8)	勤工助学累计次数	W_8

1.2 确定指标权重

传统的指标赋权方法如专家评判法、层次分析法等具有很大的主观性和盲目性,而熵值法是一种客观的赋权方法^[8],能够克服主观赋权的不足. 其原理是利用信息熵的特性,依据指标变异性的大小来确定指标的客观权重^[9]. 一般情况下,若某个指标的信息熵 E_j 越小,表明该指标的变异性越大,其涵盖的信息量就越多,在综合评价中起到的作用就越大,因此其权重就越大;反之,如果某个指标的信息熵 E_j 越大,表明该指标的变异性越小,其涵盖的信息量就越少,在综合评价中起到的作用就越小,因此其权重就越小^[10]. 计算步骤为:

1) 构建矩阵. 根据设计的指标,取各指标的数据值,构建矩阵 R

$$R = (r_{ij})_{m \times n} (i = 1, 2, \cdots, m; j = 1, 2, \cdots, n)$$
 (1)

式(1)中, m 是指标个数, n 是对象个数, r_{ij} 为指标 i 在区域 j 上的值^[11]. 本文依据设计出的指标,将校园卡消费的原始数据经过预处理后得到贫困生指标数据,从而构建出以上矩阵.

2) 数据标准化处理. 由于各指标之间计量单位等属性差异很大,因此需要对数据采取标准化处理,用 min-max 标准化法对得到的指标数据进行线性变换^[12],得到标准化后的数据 C_{ij} ,其取值范围为 $[0, 1]$,计算公式为

$$C_{ij} = \frac{r_{ij} - \min(r_{ij})}{\max(r_{ij}) - \min(r_{ij})}$$
 (2)

$$C_{ij} = \frac{\max(r_{ij}) - r_{ij}}{\max(r_{ij}) - \min(r_{ij})}$$
 (3)

式(2)、(3)中, $\max(r_{ij})$ 、 $\min(r_{ij})$ 分别为某指标所在行中的最大、最小值. 式(2)适用于正向指标,式(3)适用于逆向指标.

3) 计算贡献度. 将数据矩阵 R 经过标准化处理后,建立新数据矩阵 P ,然后确定各指标的贡献度

$$P_{ij}^{[13]}, P_{ij} = \frac{P_{ij}}{\sum_{j=1}^n P_{ij}}.$$

4) 计算各指标的信息熵. 根据公式(4) 计算各个指标的信息熵 E_j ,

$$E_j = -(\ln n)^{-1} \sum_{j=1}^n (P_{ij} \cdot \ln P_{ij}) = -K \cdot \sum_{j=1}^n (P_{ij} \cdot \ln P_{ij}) \quad (4)$$

式(4) 中, $K = \frac{1}{\ln n}$, P_{ij} 为贡献度.

5) 计算指标权重. 在上述计算中, 若 $P_{ij} = 0$, $\lim_{P_{ij} \rightarrow 0} P_{ij} = 0$, 在确定各指标的信息熵值后, 根据下述公式来计算指标的权重:

$$W_i = \frac{1 - E_i}{m - \sum_{i=1}^n E_i} \quad (5)$$

其中, $1 - E_i$ 表示第 i 项指标的差异系数, 差异系数越大, 指标就越重要.

1.3 贫困指数的计算

贫困指数是用以衡量一个学生贫困程度的指数, 用 PM 表示, 其取值范围在 $0 \sim 1$ 之间, 数值越大贫困程度越高. 依据本文构建的贫困因子及其权重, 学生贫困指数的计算逻辑表示为:

$$PM = \frac{X_1 * W_1 + X_2 * W_2 + X_3 * W_3 + \dots + X_6 * W_6 + Y_7 + Y_8}{W_1 + W_2 + W_3 + \dots + W_6 + W_7 + W_8} \quad (6)$$

其中, 当 $0 \leq X_7 \leq 5000$ 时, $Y_7 = (X_7 * W_7)/5000$; 当 $X_7 > 5000$ 时, $Y_7 = W_7$; 当 $0 \leq X_8 \leq 3$ 时, $Y_8 = (X_8 * W_8)/3$; 当 $X_8 > 3$ 时, $Y_8 = W_8$. PM 代表贫困生的贫困指数, 将贫困指数排序, 根据从高到低的原则可以得出学生的贫困程度.

2 系统设计

2.1 系统架构设计

为了达到系统高扩展性、高效性和高度安全的目标, 本文设计的系统由系统前端、服务器层、数据库 3 层结构组成, 系统结构图如图 1 所示. 系统前端分为智能终端和 PC 浏览器端; 服务器层采用 Web 层、服务层和持久层 3 层结构进行设计, 持久层实现与数据库的交互操作, 服务层通过调用持久层的接口实现业务逻辑, Web 层通过调用服务层的接口实现客户端的请求处理; 数据库层采用 Oracle 数据库存储系统数据, 同时采用 Redis 作为缓存服务器以存储规则数据、系统数据等热点数据^[14].

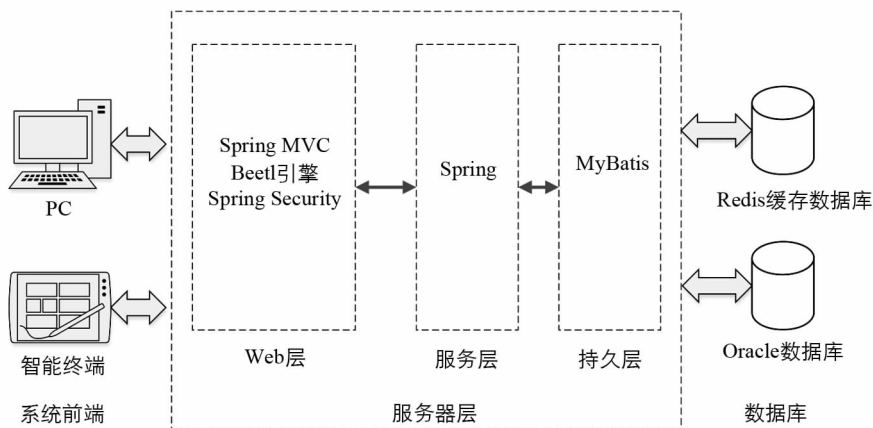


图 1 系统结构图

2.2 系统功能设计

本文设计的系统通过选取贫困指标建立数学模型, 对贫困生的贫困指数进行量化计算, 根据定义的贫困生规则帮助学校找出真正的贫困生, 辅助学校进行贫困生的精准资助. 本系统包含 8 个子模块: 数据预

处理模块、贫困生分析模块、贫困生查询模块、统计分析模块、贫困生规则定义、数据更新配置、审批处理模块和系统基础模块。系统功能结构图如图 2 所示。

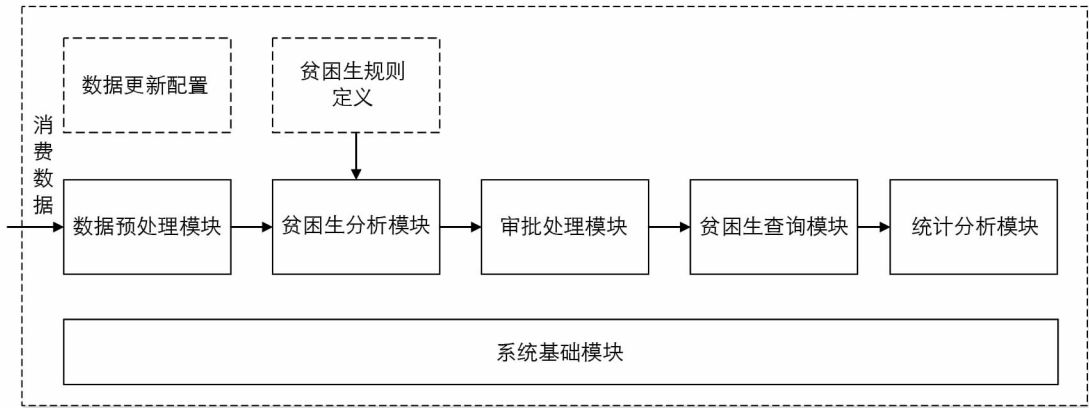


图 2 系统功能结构图

- 1) 数据预处理模块: 本文对校园卡消费数据首先进行预处理, 预处理的内容主要包含:
- ①对学生每月食堂消费的次数、总金额进行统计处理. 考虑到食堂消费的时段问题, 定义一日三餐的时间, 早餐时间段为 06:00—9:59, 午餐时间段为 11:00—14:59, 晚餐时间段为 16:30—21:30; 根据这 3 个时间段统计得到学生每月三餐消费的次数和金额. 同时, 根据学生每月食堂消费的次数, 并结合每月上学的天数计算学生的食堂用餐率.
 - ②消费异常处理. 在同一个时间段内, 如果学生的消费次数 >1 次, 则视为消费异常. 为排除该情况对模型的干扰, 在消费次数 >1 的情况中, 将单次消费金额 $\leq$ 阈值 β (本文 β 取值为 10)的情况认为是正常消费, 将此时间段内的消费金额进行累加, 得到学生该餐的消费总金额; 若消费金额 $>$ 阈值 β , 则将此时间段内的消费均值作为学生该餐的消费金额. 另外, 对于消费金额为 0 的记录同样视为消费异常, 将此类数据进行剔除处理.
 - ③个体异常处理. 由于高校的管理模式, 有些学生使用校园卡的次数较少, 本文对每学期校园卡消费总次数 ≤ 50 次的学生视为异常个体, 将其数据记录剔除不予考虑. 另外, 对于学生某一餐缺失的数据, 其该餐的消费金额记为 0, 消费次数不计, 此类处理不会影响该生整体的消费水平和模型分析的准确性.
 - ④超市消费处理. 对学校超市的消费数据进行处理, 得到学生每月超市消费的次数与金额.
- 2) 贫困生分析模块: 该模块为系统的核心模块, 根据预处理之后的数据得到学生消费特征数据, 结合系统定义的贫困生规则, 按照熵值法模型计算各指标的权重, 进而计算学生的贫困指数, 并进行排序; 依据排序和规则划定贫困生等级, 得到最新的贫困生名单, 并进入审批环节.
- 3) 贫困生查询模块: 该模块主要用于查询在校贫困生的情况, 主要包含两类, 一类是历史贫困生名单, 另一类是新增的贫困生名单. 历史贫困生名单是上一期根据贫困生规则定义生成并经过审批后, 实际执行的贫困生名单; 新增贫困生名单是在本次评定周期内, 根据贫困生规则定义进行分析, 并结合历史贫困生数据, 通过审批后生成的最新贫困生名单.
- 4) 统计分析模块: 该模块结合图、表等形式呈现在校贫困生的整体信息, 如全校贫困生的历史人数对比、院系人数分布、籍贯和政治面貌分布等; 同时, 也对全校贫困生的消费情况进行分析, 便于学校全面了解在校贫困生的整体情况.
- 5) 贫困生规则定义: 该模块主要是对学校的贫困生评定规则进行定义, 包括贫困因子权重配置、贫困等级评定配置. 贫困因子权重配置就是学校可根据实际情况对系统默认的因子权重进行一定的调整, 贫困等级评定配置是学校根据系统分析出的学生贫困指数量化分值排名来配置学校贫困生的贫困等级. 比如, 将贫困指数量化分值排名前 5% 的学生认定为特别贫困; 将贫困指数量化分值排名前 5%~15% 的学生认定为困难生; 将贫困指数量化分值排名前 15%~30% 的学生认定为一般贫困生.
- 6) 数据更新配置: 考虑到学生在校学习的时间情况, 该模块用于对原始消费数据的抽取和分析处理的时间周期进行配置. 系统每学期进行一次数据抽取和分析, 抽取周期默认为 3 个月的消费数据. 比如, 每

年上学期的数据抽取 3 月、4 月和 5 月份的数据进行分析处理,数据更新时间设置为 9 月 1 日;每年下学期的数据抽取 10 月、11 月和 12 月份的数据进行分析处理,数据更新时间设置为 3 月 1 日。

7) 审批处理模块:本文设计的系统通过自动化分析与人工评定相结合的方式得出最终的贫困生名单,因此,该模块主要是实现对自动化分析评定结果的审批处理.系统默认设计四级结果审核机制,分别是辅导员审核、院系审核、学工部审核、学校审核,学校可根据实际自定义审批流程。

8) 系统基础模块:该模块是系统的基础模块,包含用户管理、角色管理、菜单管理、权限管理以及参数配置等.用户管理用于对学生及用户的基本信息进行新增或者同步管理;角色管理用于配置和分配用户角色;菜单管理用于配置系统功能菜单;权限管理用于对用户角色进行权限分配;参数配置模块用于配置各类系统参数。

2.3 数据库设计

本系统除了系统基础模块涉及的用户表、角色表、权限表、菜单表、系统参数表等基础数据表外,涉及的主要业务数据表有学生基本信息表、一卡通消费原始记录表、消费数据预处理记录表、学生消费特征数据表、贫困指数表、贫困生规则定义表、贫困生等级定义表、审批结果表等,以上几个数据表设计如图 3 所示。

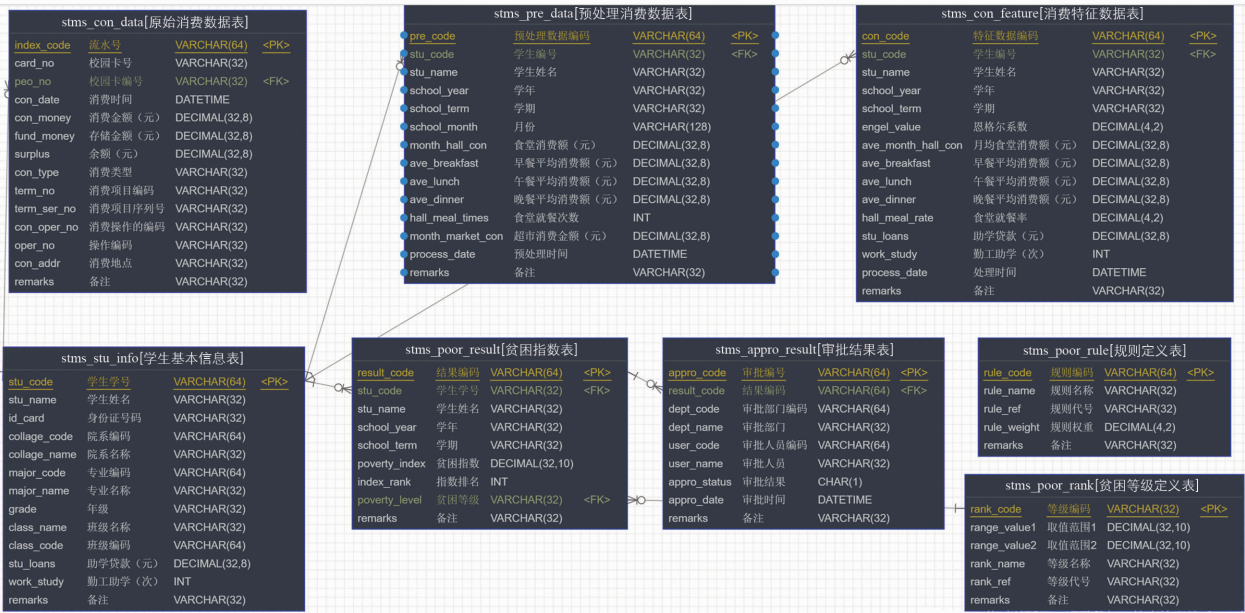


图 3 系统主要业务表设计图

3 系统实现与结果分析

3.1 系统实现

本系统的实现以 Spring Boot 2 为基础,通过 Maven 构建多模块项目,模块之间松耦合设计,方便模块的升级与增删.前端基于 Bootstrap 4 的 HTML 页面和语法直观、性能超高的 Beetl 模板引擎进行开发.后端服务器基于 Spring Boot 2.0+Spring Security+Spring MVC+MyBatis 的技术体系进行开发,从而去除了大量的配置操作,简化了开发流程^[15].将 Tomcat 作为后端服务器;采用 Spring Security 作为安全访问框架,实现对用户身份的认证和资源访问权限的控制;使用 MyBatis 作为持久层框架,同时引入 Mapper 和 Page Helper 插件,可以快速对表进行增删改.通过这几种经典技术的组合,使系统安全稳定并具有可扩展性.在系统实现过程中,利用熵值法求指标权重的代码,原始消费数据预处理实现代码,利用预处理数据计算消费特征数据的实现代码,以及计算贫困指数的实现代码为该系统实现的主要核心代码.其中,利用熵值法求指标权重的 Java 实现代码如下:

```
public static List<Double> getWeight(List<List<Double>>list){
//用于存放每种指标下所有记录归一化后的和
List<Double> listSum=new ArrayList<Double>();
```

```

//用于存放每种指标的差异系数
List<Double> gList=new ArrayList<Double>();
//用于存放每种指标的最终权重系数
List<Double> wList=new ArrayList<Double>();
double sumLast=0;
//计算  $k$  值  $k=1/\ln(n)$ 
double  $k=1 / \text{Math.log}(\text{list.get}(0).\text{size}());$ 
//数据归一化处理  $(i-\min)/(\max-\min)$ 
for (int  $i=0$ ;  $i<\text{list.size}()$ ;  $i++$ ) {
    double  $sum=0$ ;
    double  $max=\text{Collections.max}(\text{list.get}(i));$ 
    double  $min=\text{Collections.min}(\text{list.get}(i));$ 
    for (int  $j=0$ ;  $j<\text{list.get}(i).\text{size}()$ ;  $j++$ ) {
        double  $temp=(\text{list.get}(i).\text{get}(j) - \min) / (\max - \min);$ 
         $sum += temp;$ 
         $\text{list.get}(i).\text{set}(j, temp);$ 
    }
     $\text{listSum.add}(sum);$ 
}
//计算每项指标下每个记录所占比重
for (int  $i=0$ ;  $i<\text{list.size}()$ ;  $i++$ ) {
    //每种指标下所有记录权重和
    double  $sum=0$ ;
    for (int  $j=0$ ;  $j<\text{list.get}(i).\text{size}()$ ;  $j++$ )
        if( $\text{list.get}(i).\text{get}(j) / \text{listSum.get}(i) == 0$ ) {
             $sum += 0;$ 
        }
    else {
         $sum += (\text{list.get}(i).\text{get}(j) / \text{listSum.get}(i)) * \text{Math.log}(\text{list.get}(i).\text{get}(j) / \text{listSum.get}(i));$ 
    }
}
//计算第  $i$  项指标的熵值
double  $e=-k * sum;$ 
//计算第  $j$  项指标的差异系数
double  $g=1-e;$ 
 $sumLast+=g;$ 
 $gList.add(g);$ 
}
//计算每项指标的权重
for (int  $i=0$ ;  $i<gList.size()$ ;  $i++$ ) {
     $wList.add(gList.get(i) / sumLast);$ 
}
return wList;
}

```

经过编程实现,利用熵值法求出本文设计的 8 个指标权重分别为: $W_1=0.182\ 1$, $W_2=0.153\ 3$, $W_3=0.102\ 1$, $W_4=0.124\ 2$, $W_5=0.105\ 6$, $W_6=0.087\ 2$, $W_7=0.128\ 3$, $W_8=0.117\ 2$. 经过测试,本文设计与开发的系统所分析的结果示例图如图 4 所示. 按照系统审批流程,系统分析结果经过相应的审核流程后,可最终确定贫困生资助名单和资助等级.

序号	学号	姓名	院系	◆ 年级	专业	◆ 性...	◆ 贫困...	◆ 贫困级别	◆ 资助金额...	▼ 名单来源	查看
1	201501011310002	李思思	哲学与社会发...	2015	哲学	女	0.65	特殊困难	500	新增贫困生名单	Q
2	201501011310001	张明	哲学与社会发...	2015	哲学	男	0.64	困难	300	新增贫困生名单	Q
3	201501011310003	王小红	哲学与社会发...	2015	宗教学	女	0.54	特殊困难	500	新增贫困生名单	Q
4	201501011310007	陈小明	哲学与社会发...	2015	社会学	男	0.06	困难	300	新增贫困生名单	Q
5	201501011310008	林小燕	哲学与社会发...	2015	社会学	女	0.52	特殊困难	500	新增贫困生名单	Q
6	201501011310012	黄小强	哲学与社会发...	2015	社会工作	男	0.79	特殊困难	500	新增贫困生名单	Q
7	201501011310004	赵小华	哲学与社会发...	2015	人类学	男	0.08	困难	300	新增贫困生名单	Q
8	201501011310013	周小波	哲学与社会发...	2015	人类学	男	0.01	困难	300	新增贫困生名单	Q
9	201501011310009	孙小丽	哲学与社会发...	2015	人类学	女	0.4	困难	300	新增贫困生名单	Q
10	201501011310015	陈小明	经济学院	2015	经济学	女	0.09	困难	300	新增贫困生名单	Q

图 4 贫困生评定分析结果示例图

3.2 结果分析

高校贫困生认定系统的研究重点包括两方面,一是识别出哪些学生更符合学校的补助标准;二是根据数据分析出学生的补助等级. 因此,为了验证系统分析结果的准确性,本文利用某高职院校 2016 级 4 218 名学生在 2017 年的一卡通消费数据和基本信息数据进行测试分析. 结合该校的实际情况,系统将贫困指数量化分值排名前 2% 的学生认定为特别贫困;排名前 3%~5% 的学生认定为普通困难;排名前 6%~10% 的学生认定为一般贫困. 根据测试结果,验证系统分析出的贫困生名单与学校当年实际评定名单的符合度.

表 2 测试结果分析表

验证指标	实际名单人数	与实际名单相符的人数	准确率
贫困生总数	422	367	86.97%
特别贫困	84	69	82.14%
普通贫困	127	107	84.25%
一般贫困	211	191	90.52%

按照设定的分析比例,系统分析出的 422 个学生总名单中有 367 个学生与实际名单和贫困等级相吻合,精准度达到 86.97%. 其中,特别贫困生的准确率为 82.14%,普通贫困生的准确率为 84.25%,一般贫困生的准确率为 90.52%. 通过测试结果显示,该系统对于贫困度相对较低的学生的识别准确率相对高一些,能较为准确地筛选出学生所属的贫困等级范围,学校再结合人工审核的方式进行最终评定,将提高学校贫困生认定工作的效率.

4 结 论

为实现高校对贫困生的精准资助,提高资助育人成效,解决贫困生认定工作中难量化、不客观的问题,本文利用校园卡消费数据、学生基本信息数据,结合恩格尔定律,设计了学生消费恩格尔系数、月均食堂消费水平、早餐月均消费额校均比、中餐月均消费额校均比、晚餐月均消费额校均比、食堂就餐率、助学贷款总额、勤工助学次数等 8 项贫困生认定指标,并运用熵值法计算各指标的权重,客观计算出学生的贫困指数,确定贫困等级. 最后,通过 Java EE 技术实现了相应的系统,经过测试系统可以很好地运用于实际工作中,能帮助学校解决贫困生评定过程中存在的难量化、不客观等问题. 未来,将研究利用学生的移动支付和网络支付等途径的消费数据,使数据更加完整,涵盖的范围更加广泛,系统识别的准确性也将进一步提高.

参考文献:

[1] 肖凤翔, 俸晓玲. 高校贫困生精准资助的宗旨、原则和策略 [J]. 现代教育管理, 2020(3): 117-122.

[2] 饶 亮. 改进的 Apriori 算法在贫困生助学系统中的应用 [D]. 哈尔滨: 哈尔滨工程大学, 2011.

[3] 王泽原, 赵 丽, 胡 俊. 大数据环境下利用随机森林算法和决策树的贫困生认定方法 [J]. 湘潭大学学报(自然科学版), 2018, 40(6): 115-120.

[4] 柴 政, 屈莉莉, 彭贵宾. 高校贫困生精准资助的神经网络模型 [J]. 数学的实践与认识, 2018, 48(16): 85-91.

[5] 陆桂明, 张 源, 周志敏. 基于机器学习的贫困生分类预测研究 [J]. 计算机应用与软件, 2019, 36(1): 316-319.

[6] 程茜宇. 基于深度神经网络的高校贫困生精准识别研究 [D]. 南昌: 江西财经大学, 2019.

[7] 何功炳. 基于校园卡消费数据挖掘的高校贫困生认定方法研究 [D]. 合肥: 安徽大学, 2019.

[8] 柴 政, 屈莉莉, 彭贵宾. 借鉴恩格尔定律测度高校贫困生精准资助指数 [J]. 教育现代化, 2018, 5(6): 186-188, 199.

[9] 廖文婷, 何多兴, 唐 傲, 等. 基于改进熵值法的土地储备融资风险评价——以重庆市江北区为例 [J]. 西南师范大学学报(自然科学版), 2016, 41(8): 95-100.

[10] YIN L K, DENG Y. Toward Uncertainty of Weighted Networks: an Entropy-Based Model [J]. Physica A: Statistical Mechanics and its Applications, 2018, 508: 176-186.

[11] 张茜茜, 廖和平, 杨 伟, 等. 基于熵权 TOPSIS 模型的乡村土地利用转型评价研究——以重庆市渝北区为例 [J]. 西南大学学报(自然科学版), 2018, 40(10): 135-144.

[12] ONAL E, SEKER S. Entropy Model for Optimal Coordination in High-Voltage Dielectric Systems [J]. IET Generation, Transmission & Distribution, 2018, 12(13): 3130-3135.

[13] WU D M, LI Z L. Work Safety Success Theory Based on Dynamic Safety Entropy Model [J]. Safety Science, 2019, 113: 438-444.

[14] 李法平, 王成良. Redis 可扩展与高效复制方案的设计与实现 [J]. 西南师范大学学报(自然科学版), 2018, 43(3): 83-90.

[15] 胡 鑫, 姚 宇, 徐英杰. 基于 ElasticSearch 的 TEE 病例库检索系统设计与实现 [J]. 计算机应用, 2018, 38(S1): 91-94.

Research of an Identification System for Poverty-Stricken College Students Based on the Entropy Model

JI Chao-ming¹, WANG Hua¹, SONG Tie-cheng²

1. Department of Information Engineering, Sichuan Vocational and Technical College of Communications, Chengdu 611130, China;
2. Department of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

Abstract: To objectively quantify the poverty status of poverty-stricken college students, this paper constructs an index system containing 8 indicators to assess poverty-stricken students based on the campus card consumption data and basic student information. These indicators include the Engel coefficient of student consumption, the average monthly canteen consumption level, the average monthly breakfast/lunch/dinner consumption ratio between one person and all students, dining room rate, total student loan, and work-study times. The entropy value model is used to determine the weights of the indicators and calculate the student poverty index. Finally, considering the reality of identifying poverty-stricken college students, a systematic research and design is carried out by combining the entropy model analysis and manual audits. The results of a test show that our system can efficiently address the difficulties of identifying poverty-stricken students which have long plagued the colleges.

Key words: identification of poverty-stricken college students; index system; entropy model; identification system; research and design