

# 校园精准资助论文大纲

## 一、自拟题目描述

问题背景 (精准资助的政策需求、高效贫困生资助现状与痛点、传统认定方法的局限性、大数据技术在教育公平中的应用趋势)、问题提出两个板块。需对广财 2022 年数据分析大赛赛题“1.2 校园精准资助”进行改编。本次赛制要求必须设置 3-4 个子问题进行解答,子问题设置需要由易到难,或者从对特殊问题的研究拓展到对一般性问题研究。改编题目参考附后。

## 二、问题分析

问题拆解、关键思路、方法选择、模型可能成效

## 三、模型假设

## 四、符号说明

## 五、模型的建立与求解

### 5.1 问题一模型的建立与求解

**问题:** (1) 分析学生不同消费场所的活跃用户规模、消费频次分布、人均消费水平等统计特征,并识别单次消费金额显著低于该场景常态水平的异常消费行为模式。

#### 5.1.1 数据预处理

(1) 剔除异常或相关性较小的数据。将消费地点分为“核心场景”和“辅助场景”,辅助场景即使消费金额低也保留,可用于交叉验证贫困生的学习相关消费特征。

(2) 合并同一地点、同一消费类型、10 分钟内连续消费。

#### 5.1.2 统计分析,图表呈现分布情况,加以文字描述

#### 5.1.3 异常消费识别 (Z-score)

计算各场景消费金额的百分位数 (或者正态分布的  $3\sigma$  法则、箱型图法),将单次消费金额低于某个百分位数的行为定义为“单次异常低消费”。综合起来总结“高频异常低消费”情况。

#### 5.1.4 多维度统计

统计各场景不同时段 (如工作日/周末、早中晚) 的消费占比,贫困生可能在早餐时段消费占比更高。衡量稳定性,贫困生该系数应较低。

#### 5.1.5 沿用画像,类似做个小总结

## 5.2 问题二模型的建立与求解

问题：（2）在（1）的分析结果中建立模型，从研究样本中识别并筛选出经济困难程度最高的前 50 名学生作为最需要资助的候选对象。

### 5.2.1 特征指标构建

### 5.2.2 综合评价模型（Topsis 法+熵权法）

### 5.2.3 模型检验 or 模型优化

主要内容是反向验证与调整，确保无遗漏和误判，给出排除机制和补选机制

## 5.3 问题三模型的建立与求解

问题：（3）划分最需要资助的 50 名学生的资助金额等级，测算科学资助的金额区间。

### 5.3.1 K-means 聚类

将 50 名学生分为 K 个聚类，每个聚类代表一个资助等级

### 5.3.2 金额区间测算

基于平均值或者某个合理百分位数确定能保证学生日常消费的金额

## 六、模型的评价及优化

### 6.1 模型检验

### 6.2 模型的优点

可能实现的创新点：“消费-学习-健康”多维贫困评估框架、建立资助效果闭环评估系统

### 6.3 模型的缺点

### 6.4 模型的推广

## 七、参考文献

## 《校园卡里的“精打细算”密码：多场景消费如何解码“经济隐形人”？》

### ——基于全维度消费行为的高校隐形资助模型研究

随着社会经济发展，人民生活水平整体提升，但在高校群体中，仍存在部分学生因家庭经济困难，其生活质量与教育支出承受能力显著低于校园平均水平。这些学生，特别是来自低收入家庭的大学生，常需在学业之外承受巨大的经济压力，为减轻家庭负担而不得不大幅压缩日常开支，其消费行为模式呈现出明显区别于普通学生的特征。当前高校贫困生资助工作面临诸多挑战：传统认定方式存在认定精度有限、易受主观因素影响、涉及隐私暴露、易引发心理负担等痛点；同时，资助资源有限性与需求广泛性之间的矛盾，亟需精准化、科学化、动态化的资助政策来提升资源配置效率与公平性。在此背景下，大数据技术在促进教育公平、实现“精准资助”方面展现出巨大潜力。我们将借助某高校 300 余名学生产生的逾 2 万条校园卡消费流水数据，运用大数据分析技术与数学建模方法，探索构建一种既能有效识别经济困难学生，又能充分保障其个人尊严与隐私的“隐形资助”机制。基于上述背景，我们拟定研究以下问题：

(1) 分析学生不同消费场所的活跃用户规模、消费频次分布、人均消费水平等统计特征，并识别单次消费金额显著低于该场景常态水平的异常消费行为模式。

(2) 在 (1) 的分析结果中建立模型，从研究样本中识别并筛选出经济困难程度最高的前 50 名学生作为最需要资助的候选对象。

(3) 划分最需要资助的 50 名学生的资助金额等级，测算科学资助的金额区间。