

AHP层次分析与K-means聚类相结合的 博士学位论文评价指标权重确定方法

程仕平, 陈明, 殷悦

(中南大学研究生院, 湖南长沙, 410083)

[摘要] 博士学位论文评价指标体系中各项指标权重系数的确定, 关系到博士学位论文最终的得分与评价, 对整个博士学位论文质量评价体系有着不可忽视的作用。以确定博士学位论文质量评价系统中各项评价指标权重为研究目标, 通过采集国内某高校2019—2020年博士学位论文双盲评审结果数据, 建立以AHP层次分析法与K-means聚类方法为基础的层次数据驱动权重确定模型, 得出博士学位论文评价体系中选题、创新性、科研能力、论文规范性等评价指标的权重系数, 并对案例的结果进行分析, 提出创造性建议。

[关键词] 数据驱动; 博士学位论文; 指标权重确定; 评价指标; 质量; 方法

[中图分类号] G647 **[文献标识码]** A **[文章编号]** 1674-893X(2021)05-0072-05

人才培养质量是高等教育的生命线, 而博士学位论文是博士生自我价值最为重要的体现, 集中反映其科研能力、创新精神, 以及独立运用基础理论和专门知识解决实际问题的能力, 同时也在一定程度上也反映了其导师的学术水平和职业精神。在探索提高博士学位论文质量方法的过程中, 众多学者的研究重点在于如何建立合理的博士学位论文质量评价体系。相较于国外, 我国研究生教育存在起步晚、研究生基数大等问题, 致使我们不得不建立一套符合我国教育国情的研究生学位论文评价指标体系, 这无疑是一种以问题为导向的解决方式。科学的学位论文质量评价, 不仅可以强化导师立德树人的使命, 而且可以促进研究生教育水平的提高^[1]。就目前而言, 博士学位论文质量评价体系存在不够健全, 指标权重确定方法不尽科学等问题, 尤其是学位论文质量评价体系缺乏基于数据的指标权重确定方法, 如何有效地确定博士学位论文质量评价体系中各指标的权重已刻不容缓。

本研究基于博士学位论文评价体系中各指

标确定的情况下, 以数据驱动的方式, 充分利用历史数据, 结合AHP层次分析法与K-means聚类方法对博士学位论文评价指标权重进行确定, 促进博士学位论文质量评价体系的形成。

一、文献回顾

通常情况下, 博士学位论文评价指标权重是在评价体系构建的过程中随之形成的。黄纯美等运用专家访谈法和层次分析法(AHP)构建了一套中医学专业的博士学位论文评价指标体系, 其运用专家调查法(德尔菲法)明确评价指标, 运用层次分析法确定评价体系的指标权重^[2]。王玉环等以评阅小组的关注点为突破口, 以论文是否具有创新性为中心, 设计了一种新颖的评价指标体系, 在关注程度和主观想法的共同作用下完成指标权重的确定^[3]。关志民等采用模糊综合评判的方法建立了学位论文质量主要影响因素的评价模型, 在权重确定环节使用德尔菲法和层次分析法^[4]。马莉萍运用统计学方法和层次分析法, 结合评审专家与评价管理部门的工作, 对现行学位论文质量评价提出了改进意见, 使用德尔菲和层

[收稿日期] 2021-07-05; **[修回日期]** 2021-10-15

[基金项目] 中南大学学位与研究生教育教学改革研究项目“博士生学位论文质量评价指标体系及权重研究”(2020JGB086)

[作者简介] 程仕平, 湖南益阳人, 博士, 中南大学研究生院副研究员, 主要研究方向: 学位与研究生教育管理, 联系邮箱: spcheng@csu.edu.cn

次分析法确定权重^[5]。王琳娜利用专家访谈法构建了军队医学院校的博士学位论文质量评价体系模型, 通过评价指标的聚类进行权重的确定^[6]。刘宁利用 SPSS 统计分析软件, 通过层次分析法完善了现行评价指标体系, 同样在指标权重确定过程中引入层次分析法^[7]。

虽然现有的研究成果已在学位论文评价指标体系构建中取得了较好的成效, 但其评价指标权重的确定均是基于专家调查法(德尔菲法)和层次分析法, 使指标权重对专家组平均意见过于依赖, 进而缺失了一定程度的自由度。

二、研究方法

本文在专家调查法(德尔菲法)和层次分析法的基础之上, 通过分析各专家的意见得出各评价指标的可行范围, 并结合已有的博士学位论文评分数据, 通过 K-means 聚类方法以及遗传算法提出了一种新颖的数据驱动权重确定方法。

(一) 研究方法概述

聚类问题作为无监督机器学习中最重要研究领域, 正吸引着越来越多研究人员的注意。K-means 聚类是聚类方法中最为经典的方法, 也称为 K 均值聚类方法。简单来说, K-means 方法是通过计算各数据样本点之间的欧式距离构造一个刻画数据空间的度量矩阵, 接着根据靠近便相似的思想, 将彼此靠近的数据点视为同一类, 彼此疏远的数据点视为不同类, 进而将样本聚类成 K 个不同类, 即认为两个数据点之间越接近, 那么他们之间的相似性越高, 越有可能分到同一个类别当中。K-means 算法描述如下: 第一步, 从样本集合中随机挑选 K 个样本作为初始聚类中心, $C = \{c_1, c_2, \dots, c_k\}$; 第二步, 计算样本集中每个样本到 K 个聚类中心的欧式距离, 将样本分给距离最小的聚类中心所属的类; 第三步, 针对 K 个类别, 重新计算它的聚类中心, $c_i = (\sum_{x \in c_i} x) / |c_i|$ (即属于该类的质心); 第四步, 重复第二、三步直到聚类中心不再变化或其他的终止条件(例如迭代次数限制)。

遗传算法(Genetic Algorithm, GA)作为群智能算法与启发式算法中最为经典的方法, 其寻找目标函数最优解的优异能力使遗传算法得到了广泛的应用。简单来说, 遗传算法是通过模拟达尔文生物进化论中的自然选择和遗传学机理来

求解给定优化模型的最优解的方法。算法描述如下: 第一步, 随机产生种群(初始解); 第二步, 个体评价, 根据求解问题的目标, 给出每个个体的适应度值; 第三步, 选择运算, 对适应度值依据轮盘赌等方法选择个体, 即适应度值高的个体被选择的概率更高, 适应度值低的个体将被淘汰; 第四步, 交叉运算与变异运算, 按照一定的交叉概率和交叉方法、变异概率和变异方法, 生成新的个体, 产生新一代的种群; 第五步, 终止条件判断, 若满足终止条件, 则将具有最大适应度值的个体作为最优解输出, 否则返回第二步。数据驱动是指通过各种手段收集数据, 随后将数据进行汇总与整理, 并抽取数据中的相关信息, 在数据的基础上做出决策。简单而言, 数据驱动就是以数据为核心依据进行决策。

(二) 层次数据驱动权重确定方法

层次数据驱动权重确定方法是基于层次分析法和 K-means 聚类方法的博士学位论文评价指标权重确定方法, 其建立层次数据驱动权重模型, 而后通过数据驱动确定权重, 具体分为两个步骤。

1. 运用层次分析法确定权重的区间

以国内某高校的博士学位论文为例, 介绍层次数据驱动权重确定模型中权重区间的构造方法。该校博士学位论文评审主要包括四项评议项目: 论文选题、创新性及论文价值、基础知识及科研能力、论文规范性。为了更好地了解各评议项目之间的权重关系, 首先基于专家调查法(德尔菲法)的思想征求各学科各研究方向的各专家意见, 接着对所得到的意见采用层次分析法确定四项评议项目的初始权重(该步骤要对所收集到的专家意见判断矩阵—各指标之间关系矩阵, 进行方差分析以及一致性检验, 剔除异常的意见矩阵, 以保证意见的可靠性), 然后根据专家对各评议指标意见的分歧程度得到各评议指标的变化范围, 进而得到各评议指标的取值范围。

2. 运用聚类与遗传算法求解

博士学位论文评审结果一般分为优秀、良好、中等、一般、不及格 5 种类型。因此, 首先运用 K-means 聚类方法对所收集到的原始数据进行聚类, 使同一类数据间的相似度尽可能高, 不同类数据间的相似度尽可能低, 并尽量保证各类

型人数符合标准正态分布。然后利用数据特点,以及实际情况中各类型的分数阈值,设置每类人群的代表分数。最后,结合步骤1中所得到的各评议指标权重范围,以及最小二乘法的思想,建立回归模型,并通过遗传算法,求解得到对应每一个评价指标的权重。

三、实例验证

以国内某高校2019—2020年通过教育部学位中心平台盲审的博士学位论文评审结果数据为例,研究各评议项目合适的权重。由于各评议项目之间存在相互影响的关系,因此通过构造成对比较矩阵来计算各评议项目之间的权重。

以层次分析法与K-means聚类方法为基础的层次数据驱动权重确定模型,确定四项评议项目的权重,具体步骤分为以下四步。

(一) 确定初始权重范围

设某一层有 n 个因素, $X=\{x_1, x_2, \dots, x_n\}$,比较该层每个因素对上一层某个因素的影响程度,确定在该层中相对于某一准则所占比重。

假设上一层有 m 个因素,该层有 n 个因素,那么对于该层我们需要构建 m 个 $n \times n$ 的成对比较矩阵。用 a_{ij} 表示第 i 个因素相对于第 j 个因素的比较结果,比较时取1~5尺度。记 A 为所需的成对比较矩阵:

$$A = [a_{ij}]_{n \times n} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$$

其中, $a_{ij} = \frac{1}{a_{ji}}$ 。

在本节中,运用专家经验法对四个指标(选题、创新性、科研能力、论文规范性)构造成对比较矩阵(如表1所示)。

表1 四个指标的对比矩阵

	选题	创新性	科研能力	论文规范性
选题	1.00	0.20	0.33	1.00
创新性	5.00	1.00	2.00	5.00
科研能力	3.00	0.50	1.00	3.00
论文规范性	1.00	0.20	0.33	1.00

如果成对比较矩阵非一致阵,Saaty等人建议用其最大特征根对应的归一化特征向量作为权重向量 ω 。根据检验系数 CR 的计算公式:

$$CR = \frac{CI}{RI} = \frac{\lambda - n}{n - 1}$$

$$CR = \frac{CI}{RI} = \frac{0.001851}{0.91} = 0.002 < 0.1$$

当 $CR < 0.1$ 时,认为成对比较矩阵 A 的不一致程度在容许范围之内,可用其归一化特征向量作为权重向量,否则要重构成对比较矩阵,对成对比较矩阵 A 加以调整。由MATLAB计算出 $CR = 0.002 < 0.1$,通过一致性检验。因此我们取其最大特征根对应的归一化特征向量作为初始权重向量(0.0987, 0.5186, 0.2840, 0.0987)。接着,我们根据专家们对各评议指标的分歧程度赋予各评议指标一定的变化范围,最终得到四项评价指标的权重约束范围为(5%—15%, 46%—56%, 24%—34%, 5%—15%)。

(二) 确定聚类与类别分数

对所收集到的国内某高校2019—2020学年博士学位论文评审结果数据进行标准化处理,使得每一项评审项目(选题、创新性及论文价值、基础知识及科研能力、论文规范性)的分数映射到0~100分之间,进而得到全部博士学位论文所对应的分数矩阵 $X = [x_1, x_2, \dots, x_n]$,其中 $x_i = (x_{i1}, x_{i2}, x_{i3}, x_{i4})$, $x_{ij} \in \{0, \dots, 100\}$ 代表第 i 篇博士学位论文在选题、创新性、科研能力、论文规范性这四项评价意见上的分数。接着通过K-means聚类方法,以四种评议项目为特征,将所有样本分为5类,以对应5种不同的评价级别:优秀、良好、中等、一般、不及格。然后,将各类中心的总分作为对应评价级别的分数阈值,进而得到全部博士学位论文所对应的评价级别,记为 $y = (y_1, \dots, y_n)$ 。

(三) 建立模型

根据上一小节所得到的全部博士学位论文分数矩阵 $X \in R^{n \times 4}$ 以及对应的评价级别向量 $y \in R^{n \times 1}$,结合最小二乘法的思想,建立如下回归模型,以得到满足专家意见的评议项目的最优权重,具体模型如下:

$$\min_{\omega} \|y - X\omega\|_2^2$$

$$\text{s.t.} \begin{cases} \omega_1 + \omega_2 + \omega_3 + \omega_4 = 1 \\ 5\% \leq \omega_1 \leq 15\% \\ 46\% \leq \omega_2 \leq 56\% \\ 24\% \leq \omega_3 \leq 34\% \\ 5\% \leq \omega_4 \leq 15\% \end{cases} \quad (1)$$

其中 $\omega = [\omega_1, \omega_2, \omega_3, \omega_4]^T$ 表示 4×1 的权重向量, 分别代表选题、创新性及论文价值、基础知识及科研能力、论文规范性四项评价意见的权重。

(四) 求解模型

由于上述优化模型(1)的目标函数关于决策变量 ω 是非凸的, 且带有等式约束与不等式约束, 这使得传统的优化算法很难高效地求解得到局部极小值点, 为此, 我们采用遗传算法这种启发式算法来求解该优化模型。

遗传算法作为成熟的启发式算法, 因为其强大的适应性, 已经被广泛地应用于各种复杂的优化模型的求解当中。在本文中, 我们针对上述优化模型(1), 使用遗传算法进行求解, 并取 20 次实验结果的平均值作为最终结果, 以保证结果的可靠性。

通过上述博士学位评审实例计算, 最终得到论文选题、创新性及论文价值、基础知识及科研能力、论文规范性的权重结果为[0.11, 0.46, 0.30, 0.13]。

相较于 AHP 方法下的初始权重[0.10, 0.51, 0.29, 0.10], 其创新性所占权重为 0.46, 在创新性权重范围(0.46-0.50)的左边界上, 表现出从计算数据中所体现出来的创新性远远不如专家所认可的创新性程度, 说明博士论文的创新性水平还有待提高与加强。

四、方法总结

本文所提出的适用于博士学位论文质量评价指标体系的权重确定方法即层次数据驱动权重确定法, 充分考虑专家意见, 并且利用现已有的实际数据来驱动评价模型中权重的确定, 不仅克服了仅由层次分析法确定权重时主观性太强的缺点, 也克服了标准离差、熵值法等传统权重确定方法偏离实际情况的缺点, 做到结合专家主观想法的同时, 充分利用历史数据中所包含的有价值信息。

无论是博士学位论文质量评价体系, 还是各类其他评价体系, 评价指标的权重确定都至关重要, 本文所提出的方法都能够适应于各类体系中的权重确定。但是该方法还是存在着一些缺陷: ①由于该方法是在专家调查法(德尔菲法)和层次分析法的基础上建立的, 所以其所得到指标权重

很大程度上依赖于前期的指标体系建立的质量; ②虽然遗传算法与 K-means 算法已被广泛地应用于各种各样的实际问题, 并得到了不错的数值表现, 但由于本身的计算机理的限制, 导致遗传算法与 K-means 算法所得到的近似解往往不具有唯一性, 使得在使用上述方法时, 需要多次计算, 取平均值来作为最终的结果, 这大大影响了求解问题的效率。

五、结论与展望

本文选取教育部学位中心盲审学位论文评审大数据结果, 以国内某高校各指标确定情况下的博士学位论文评审结果为实例, 通过层次数据驱动权重确定法对博士学位论文评价指标及权重进行评估, 在充分考虑专家意见的同时, 还对历史数据中所包含的有价值信息进行挖掘, 克服了仅由层次分析法确定权重时主观性太强, 传统权重确定方法标准离差、熵值法等偏离实际情况的缺点, 得出了一个既有理论支撑又符合实际的指标权重, 验证了层次数据驱动权重确定方法的可行性, 为构建科学的博士学位论文质量评价指标体系提供了理论依据。

全面评估博士研究生学位论文质量, 需要进一步完善博士学位论文评价指标体系, 不仅仅是在指标确定的基础上进行指标权重的确定, 而是将博士学位论文评价体系中指标的确定与指标权重的确定同时进行, 做到全局优化, 适时监控各指标体系权重, 这也是层次数据驱动确定权重的方法能否得到广泛推广的关键。

从目前评审结果所反映的学位论文质量来看, 创新性仍有待提高。因此, 高校应着力在“治标”(博士学位论文创新性提高)和“治本”(博士研究生创新能力提高)两方面同时发力, 做到标本兼治, 全面构建“学生—导师—学科—学院—学校”五位一体的学位与研究生教育质量保障体系, 通过学术训练、学术交流, 不断增强创新意识的培育、创新思维的拓展、创新能力的喷发; 明确导师“第一责任人”职责, 坚持立德树人根本任务, 全面加强师德师风建设, 充分发挥导师言传身教作用, 用科学的世界观引领研究生开展创新性研究; 坚持内涵式发展, 优化学科布局, 切实优化人才培养类型结构, 建立多类别的创新

拔尖人才选拔机制,以国家需求为导向,致力攻坚核心技术,造就一大批复合型科技领军人才和高水平创新团队,为国家做出创新性贡献。

参考文献:

- [1] 俞婉琦. 研究生学位论文质量评价指标体系研究[D]. 杭州: 浙江大学, 2017.
YU Wanqi. Research on the quality evaluation index system of graduate dissertations[D]. Hangzhou: Zhejiang University, 2017.
- [2] 黄纯美, 许仕杰, 方熙茹, 等. 中医学博士学位论文质量评价指标体系的构建与实践[J]. 中国高等医学教育, 2013(9): 118-119.
HUANG Chunmei, XU Shijie, FANG Xiru, et al. Construction and practice of quality evaluation index system for doctoral dissertations in traditional Chinese medicine[J]. China Higher Medical Education, 2013(9): 118-119.
- [3] 王玉环, 程杰贤, 任健华. 基于专家关注度的博士学位论文评价指标体系分析[J]. 学位与研究生教育, 2012(9): 49-53.
WANG Yuhuan, CHENG Jiexian, REN Jianhua. Analysis of doctoral dissertation evaluation index system based on expert attention[J]. Degree and Graduate Education, 2012(9): 49-53.
- [4] 关志民, 束军意, 马钦海. 学位论文质量的多层次模糊综合评价模型及其应用[J]. 科研管理, 2005, 26(3): 152-157.
GUAN Zhimin, SHU Junyi, MA Qin Hai. Multi level fuzzy comprehensive evaluation model of dissertation quality and its application[J]. Scientific Research Management, 2005, 26(3): 152-157.
- [5] 马莉萍. 研究生学位论文质量评价体系的构建[J]. 天津理工大学学报, 2005, 21(6): 85-88.
MA Liping. Construction of quality evaluation system for graduate dissertations[J]. Journal of Tianjin University of Technology, 2005, 21(6): 85-88.
- [6] 王林娜. 军队医学院校博士学位论文质量评价指标体系的构建[D]. 重庆: 第三军医大学, 2005.
WANG Lina. Construction of quality evaluation index system for doctoral dissertations in military medical colleges[D]. Chongqing: Third Military Medical University, 2005.
- [7] 刘宁. 天津大学博士学位论文质量评估体系研究[D]. 天津: 天津大学, 2009.
LIU Ning. Research on quality evaluation system of doctoral dissertation of Tianjin University[D]. Tianjin: Tianjin University, 2009.

Method for determining the weights of doctoral dissertation evaluation index based on AHP and K-means

CHENG Shiping, CHEN Ming, YIN Yue

(School of Graduate Central South University, Changsha 410083, China)

Abstract: The determination of the weight coefficient of each index in the evaluation index system of doctoral dissertation is related to the final score and evaluation, and it plays an important role in the quality evaluation system of the entire doctoral dissertation. Aiming at determining the weight of each evaluation index in the doctoral dissertation quality evaluation system, and by collecting the data of the double-blind review of the doctoral dissertation of our school from 2019 to 2020, this paper establishes a hierarchical data-driven weight determination model based on AHP and K-means, and obtains the weight coefficients of evaluation indicators such as topic selection, innovation, scientific research ability, and thesis standardization in the doctoral dissertation evaluation system, analyzes the results of the case, and puts forward creative suggestions.

Key Words: data-driven; doctoral dissertation; index weight determination; evaluation index; quality; method

[编辑: 何彩章]