

# 多元化高等学校专业建设质量评估体系研究

马梦优, 贾宏恩

太原理工大学数学学院, 山西 太原

收稿日期: 2024年1月7日; 录用日期: 2024年1月31日; 发布日期: 2024年2月7日

## 摘要

根据社会需求、自身办学条件和高校特色, 构建科学、合理、严谨的专业建设质量评价体制, 是专业建设实施过程中不可缺少的重要环节。目前国内现有的专业建设质量评价体系相关研究较少, 本项目基于“高等教育质量监测国家数据平台”数据, 结合国内主流的、认可度较高的高校排行榜数据, 利用因子分析进行数据筛选, 且基于RankNet的神经网络模型和三级模糊综合评价模型确定量化度量模型, 旨在建立一个合理、客观、科学的普通高等学校专业建设质量评价体系。

## 关键词

CIPP评价模式, 三级模糊综合评价模型, RankNet的神经网络模型

## Research on the Quality Evaluation System for the Construction of Diversified Higher Education Majors

Mengyou Ma, Hong'en Jia

College of Mathematics, Taiyuan University of Technology, Taiyuan Shanxi

Received: Jan. 7<sup>th</sup>, 2024; accepted: Jan. 31<sup>st</sup>, 2024; published: Feb. 7<sup>th</sup>, 2024

## Abstract

According to the social needs, its own educational conditions and the characteristics of colleges and universities, building a scientific, reasonable and rigorous professional construction quality evaluation system is an indispensable and important link in the implementation process of professional construction. At present, the existing research on the professional construction quality evaluation system is less, based on the higher education quality monitoring national data platform, combined with the domestic mainstream, high recognition of university ranking data, using the factor analysis, and based on RankNet neural network model and three-level comprehensive evaluation

**model quantitative measurement model, this project aims to establish a reasonable, objective and scientific ordinary institutions of professional construction quality evaluation system.**

## Keywords

**CIPP Evaluation Model, Three-Level Fuzzy Comprehensive Evaluation Model, RankNet Neural Network Model**

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

国内现有的专业建设评价主要以办学资源、科研绩效等量化指标为主,其容易被窄化为专业排名。在“双一流”建设语境[1]下,我们需要引导高校将注意力转向内涵建设,着眼专业长远发展规划,推进体制机制的综合改革。目前我国在高校评价指标体系构建研究上已经取得了很大进展,也有大量的研究团体以及研究人员对本科专业建设质量进行了研究,并且给出了许多排名,但这些评价、排名仍然存在一些不足,包括:没有明确介绍评估目标、评估指标体系和量化度量模型排行榜;各机构的参考指标、评价项目、评价重点往往各不相同;评估方法手段和标准单一;这些研究主要针对全国高校,不考虑地域、高校性质等因素。对于山西省的高校,目前还缺少专门针对此类地区且体系较为健全科学的专业建设质量评估体系。

在此背景下,本文拟以“高等教育质量监测国家数据平台”的数据作为基础,按照一定规则和筛选标准处理数据,并且利用因子分析对所选数据进行校验与筛选。在此基础上,教育部门或高校在确定评估目标和评估作用后,建立适应这种评估目标和评估作用的评估指标体系[2] [3] [4] [5],然后通过基于RankNet神经网络模型来实现各三级指标的权重计算,并使用模糊数学计算模型来实现专业建设质量的排序结果。从而构建本科专业建设质量评价系统,为教育部门、高校提供较科学、合理、可靠的山西普通高等学校专业建设质量模型。

美籍教育批评家斯塔弗尔比姆提出了一种新的课程评估模型即CIPP评估模型[6] [7] [8],它强调“评估是为领导决策提供数据支撑的一种过程”,其基本理念是“评价最重要的目的不在证明,而在改进”。该模型由“背景评估”、“输入评估”、“过程评估”和“结果评估”构成,可以运用于专业建设的总体流程,为建立合理的评价指标体系提供思路。

## 2. 评价指标体系的建设

### 2.1. 研究目标

- 1) 利用山西省各高校相关数据,建立对山西高校全覆盖的专业建设质量评估模型。
- 2) 构建度量量化模型形成完整的专业建设质量评估体系以供山西省各高校不同专业进行实时的质量评价。
- 3) 为各高校的不同专业提供改善专业建设质量可行方向的可靠依据。

### 2.2. 研究方法

研究探索CIPP评价模式和专业周期理论,将其结合专业建设的一般过程进行思考,作为本项目的理论基础。

利用访谈法对行业内专家进行访谈, 访谈专家们对于评价指标体系的建立的因素思考, 结合专家们的建议和文献研究来初步构建专业建设质量评价指标体系。

利用因子分析法对收集所得的数据进行信度与效度检验以此确保数据的可用性, 并利用因子分析筛选影响因子较小的指标将其筛去, 进一步修改评价指标体系。

### 2.3. 研究思路

本文研究思路见图 1:

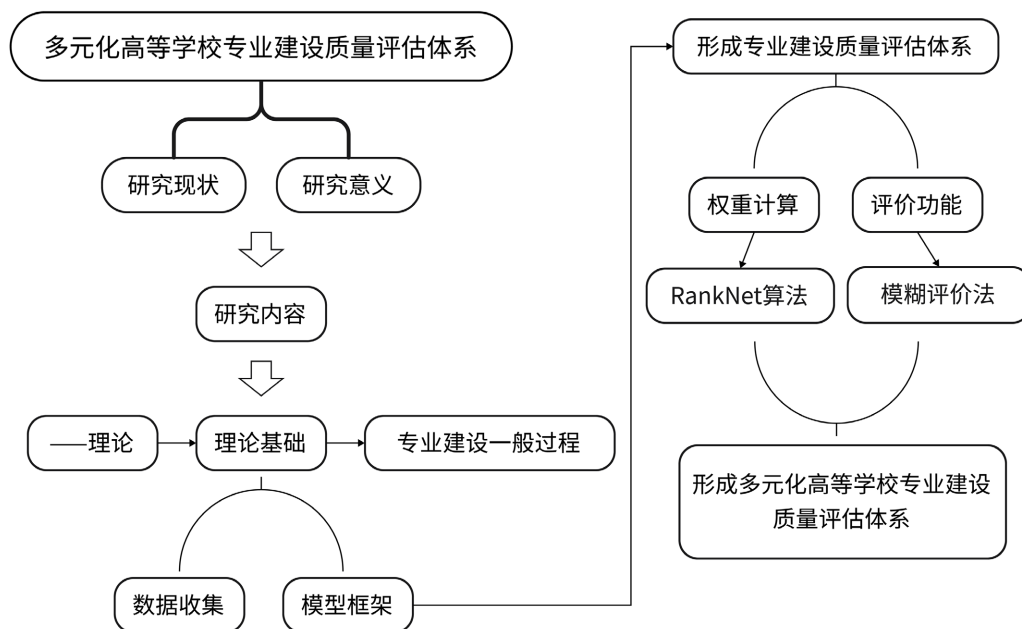


Figure 1. Research ideas and technical diagrams  
图 1. 研究思路技术图

### 2.4. 模型研究

在科学理解高校专业建设基本内涵和参阅相关文献的基础上, 参考教育部颁布的《普通高等本科教育专业教学审核评估指标体系》, 并结合高校的实际情况构建高校专业建设评估指标体系, 见图 2。构建指标体系时, 分为 7 个一级指标、22 个二级指标、49 个三级指标。

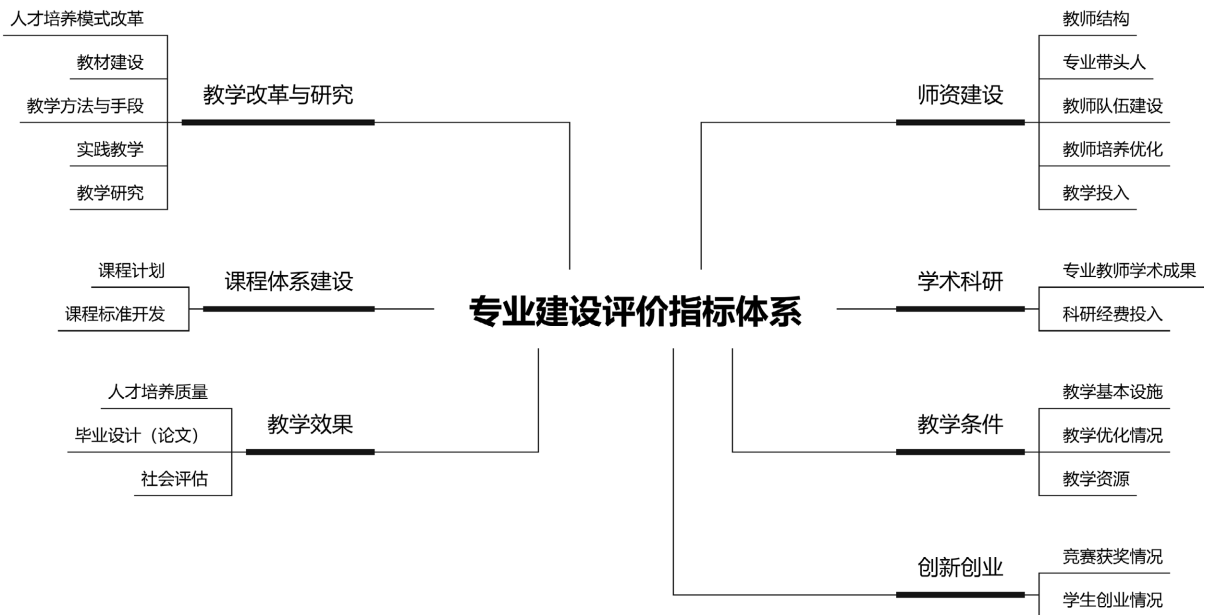
## 3. 权重的确定

### 3.1. 数据收集、整理与检验

通过研讨会及单位走访调研等多种数据采集方式, 以山西各高校的“高等教育质量监测国家数据平台”的数据为主, 结合其他官方网站数据补充来作为研究的数据支撑。通过访谈法对相关领域专家及教师进行深度访谈, 提供指标初步构建的理论依据。由省属教育部门确定本项目的评估目标以及评估作用, 并以此为依据指导评估指标体系的初步确立。依据初步确立的指标体系进行数据选择, 完成数据收集。

首先对数据进行缺失值填补, 这里使用 K-means 算法进行有效填补。然后对填补完整的数据进行无量纲化处理。设参与数据分析的各高校专业数量有  $n$  个, 指标数量有  $m$  个, 将各个指标值转换为标准化指标:

$$X_{ij} = \frac{x_{ij} - x_j}{S_j}, (i = 1, 2, \dots, n; j = 1, 2, \dots, m) \quad (1)$$



**Figure 2.** Professional construction evaluation index system  
**图 2.** 专业建设评价指标体系

为了检验收集与整理所得数据的可靠性与真实性, 这里对数据进行信度与效度分析, 利用 SPSS 软件进行信度与效度分析, 见表 1。由测量结果可以看出, 信度与效度良好。

**Table 1.** Reliability and validity testing  
**表 1.** 信度与效度检验

可靠性统计	
克隆巴赫 Alpha	项数
0.851	49

接着进行 KMO 值和 Bartlett 球形检验, 见表 2, 这可以反映数据之间是否具有较强的相关性, 由结果知因子之间具有较强相关性, 且适合做因子分析。

**Table 2.** KMO value and Bartlett sphericity test value  
**表 2.** KMO 值和 Bartlett 球形检验值

KMO 和巴特利特检验 a		
KMO 取样适切性量数		0.720
	近似卡方	151.606
巴特利特球形度检验	自由度	43
	显著性	0.000

### 3.2. 基于 RankNet 算法确定权重

本文采用 RankNet 的神经网络模型对专业建设指标的权重进行计算。RankNet 算法是一种从概率的角度来解决排序问题的方法, 其核心思想是构造概率损失函数来学习 Ranking Function, 并以此为工具对文档进行排序。RankNet 不依赖于特定的机器学习模型, 而是基于神经网络实现, 使用误差反向传播训练模型。

基于 RankNet 神经网络模型的指标权重确定过程:

输入数据: 将专业建设三级指标相关数据构成的特征向量组输入进 RankNet 神经网络中;

成对数据生成: 随机选中一组指标数据进行计算;

模型建立: RankNet 神经网络中, 将每一对指标的数据作为输入, 并将它们的权重值作为输出, 使用交叉熵损失函数来训练模型;

训练模型: 使用随机梯度下降算法来最小化交叉熵损失函数, 更新模型参数, 使得模型计算的权重值能够尽可能地接近真实的结果;

模型迭代: 重复步骤 2~4, 直到达到指定的迭代次数或者损失函数收敛。

在本文中 RankNet 网络选择特征向量 49 个, 即输入层神经元个数为 49, 隐藏层设置为 1 层, 将样本输入到 RankNet 神经网络进行训练后, 得出专业建设指标权重, 见表 3:

**Table 3.** Weights of tertiary indicators for professional construction

**表 3.** 专业建设三级指标权重

一级指标	二级指标	三级指标
师资建设 (25.415%)	教师结构(5.193%)	专任教师师生比(本科生 + 研究生)(1.029%)
		专任教师中职称结构教授与副教授总人数占比(1.381%)
		专任教师的学历结构博士占比(0.561%)
	专业带头人(8.572%)	实践教学师资队伍占比(2.222%)
		学术头衔拥有者比例(8.572%)
	教师队伍建设(3.4%)	近五年专业新引进教师平均每年人数(2.547%)
		近五年专业教师外出进修次数(0.853%)
		近五年专业教师近五年来年均评职称人数(4.135%)
	教师培养优化(7.647%)	近五年专业教师中教学名师和优秀教师荣誉获得人数(2.099%)
	教学投入(0.603%)	教师赴国(境)外交流人数(1.413%)
学术科研 (10.905%)	专业教师学术成果 (9.94%)	主讲本科课程教授占教授总数的比例(0.603%)
		近五年专业教师发表 SCI 论文数量(1.073%)
		近五年专业教师科研项目获得数量(0.806%)
		近五年专业教师参加学术活动数量(1.466%)
	科研经费投入(0.965%)	近五年专业教师出版书籍数量(5.63%)
		近五年专业科研经费申报总数额(万元)(0.965%)

续表

教学条件 (10.16%)	教学基本设施(5.85%)	专业研究生工位覆盖率(0.801%)
		专业教研室数量(1.254%)
	教学优化情况(2.481%)	实验设备与学生总数比例(1.313%)
		实习基地数量(2.482%)
教学条件 (10.16%)	教学资源(1.829%)	近五年学科资源、科研成果转化为教学资源数目(1.026%)
		近五年产业技术发展成果、产学研合作项目转化为教学资源数目(0.803%)
创新创业 (9.061%)	竞赛获奖情况(1.801%)	近五年竞赛或省级及以上奖项数量(1.212%)
		近五年竞赛年均报名人数(0.589%)
	学生创业情况(7.26%)	近五年年均学生创业人数(1.151%)
		近五年学生技术成果转化数(6.109%)
教学改革与 研究 (24.959%)	人才培养模式改革 (1.324%)	近五年本科生专业组织实习或实践次数(1.324%)
	教材建设(12.016%)	专业内教师任职期间教材出版数量(3.485%)
		优秀教材获奖数量(8.531%)
	教学方法与手段 (1.042%)	近五年信息化教学学时占比(1.042%)
	实践教学(1.129%)	近五年实验室利用率(1.129%)
教学研究(9.448%)	近五年主持校级以上项目教改立项数量(4.943%)	
课程体系建 设(8.891%)	课程计划(2.706%)	双语教学课程占比(1.157%)
		教材反映学科最新知识占比(1.549%)
	课程标准开发(6.185%)	近五年校级及以上立项建设课程数量(2.129%)
		线上精品课程数量(4.056%)
教学效果 (11.577%)	人才培养质量(5.973%)	本科生以第一作者/通讯作者在公开发行人期刊发表的论文数和本科生获批国家发明专利数(1.755%)
		体质测试达标率(0.645%)
		近五年在学期间获得国家认可的职业资格证书学生数占在校生数的比例(0.229%)
		近五年年均学生毕业且获得学位比例(1.17%)
		近五年年均应届本科生初次就业 + 升学人数占比(0.691%)
		近五年年均学生对教师授课满意率(1.483%)

续表

教学效果 (11.577%)	毕业设计(论文) (3.603%)	校级优秀论文比例(2.215%)
		近五年选题的性质、难度、分量等全面反映目标要求占比(1.388%)
	社会评估(2.001%)	近五年用人单位满意率(0.985%)
		近五年专业本科毕业生平均月薪(1.016%)

#### 4. 三级模糊综合评价量化度量模型

依据 RankNet 算法计算各个指标权重, 模糊综合评价法以 RankNet 算法所计算的权重为已知数据, 在 RankNet 算法的基础上进行三级模糊综合评价。

##### 4.1. 一级模糊综合评判

设评价集  $V = \{1, 2, 3, \dots, m\}$  代表  $m$  个不同的高校。

一级指标“师资建设”下属的三级指标是定量指标, 有具体的数据, 对这些数据归一化即求出各个高校的该指标与总指标的比重, 得到单因素评价价值。对于其他是量指标的因素, 按照同样的方法得到比重。就可得到  $M$  个高校的单因素评价价值, 其中每一行数据的和为 1, 部分数据示例见表 4。

Table 4. Example of normalized teacher structure data

表 4. 归一化后教师结构数据示例

	$i$ 高校工程力学	$i$ 高校车辆工程	$i$ 高校机器人工程	$i$ 高校材料成型及控制工程	$i$ 高校材料科学与工程	$i$ 高校材料物理
专任教师师生比(本科生 + 研究生)	0.16	0.29	0.25	0.12	0.13	0.19
专任教师中职称结构教授与副教授总人数占比	0.64	0.45	0.49	0.68	0.56	0.56
专任教师的学历结构博士占比	0.75	0.91	0.75	0.81	0.63	0.73
实践教学师资队伍占比	0.0075	0.0091	0.0075	0.0081	0.0063	0.0073

得到“教师结构”的各因素的单因素评价矩阵, 矩阵中的数据即为表中相应位置的数据。记为  $R_{11}$ 。然后利用已有的三级指标权重归一化后对涉及到的几个因子进行权重分配, 记为  $A_{11}$ 。则“教师结构”因素的一级评判为:  $B_{11} = A_{11} \cdot R_{11}$ 。

$$R_{11} = \begin{bmatrix} 0.073 & 0.115 & 0.069 & 0.087 & \dots \\ 0.067 & 0.093 & 0.069 & 0.075 & \dots \\ 0.076 & 0.083 & 0.068 & 0.072 & \dots \\ 0.048 & 0.086 & 0.058 & 0.087 & \dots \end{bmatrix} \quad (2)$$

而权重分配  $A_{11} = [0.20 \quad 0.26 \quad 0.11 \quad 0.43]$ 。

因此  $B_{11} = [0.053 \quad 0.098 \quad 0.029 \quad 0.014 \quad \dots]$ 。

## 4.2. 二级模糊综合评判

对于师资建设、学术科研、教学条件、创新创业、教学改革与研究、课程体系建设以及教学效果下属的单因素指标进行二级评判。专业建设质量情况的二级评判可以得到 7 个矩阵, 对于师资建设, 记其二级评判为  $R_1$ ,  $R_1$  为 5 行  $m$  列的矩阵, 将教师结构的一级评判结果作为二级评判的单因素评价价值, 即学生信息评判矩阵  $R_1$  的第 1 行, 给出师资建设一级指标各因子的权重  $A_1$ , 通过  $B_1 = A_1 \cdot R_1$  计算得到师资建设对专业建设质量的隶属度。同理计算其余一级指标对专业建设质量的隶属度。

## 4.3. 三级模糊综合评判

将二级评判结果  $B_1$ 、 $B_2$ 、 $B_3$ 、 $B_4$ 、 $B_5$ 、 $B_6$ 、 $B_7$  作为行, 组成三级评判的单因素评判矩阵

$$R = \begin{bmatrix} B_1 \\ B_2 \\ B_3 \\ B_4 \\ B_5 \\ B_6 \\ B_7 \end{bmatrix} \quad (3)$$

而各一级指标权重集  $A = [0.25 \ 0.11 \ 0.10 \ 0.09 \ 0.25 \ 0.09 \ 0.12]$ , 根据  $B = A \cdot R$  即可得到一个一行  $m$  列的矩阵, 从矩阵每个元素的数值大小就可以判断出某高校的某专业建设质量的分值, 分值越高, 表示专业建设排名越高。从而得到不同高校的不同专业建设质量排名。

可以依据学科评估等级划分为山西省内高校专业建设质量设定评级, 当参与的各学校各专业足够多时, 专业建设质量的排名也就愈发准确, 评级也愈发真实, 选取学科评级可以让模糊评价法得到的结果更加可视化。进行评分评级后, 可通过与各指标标准值进行距离计算来观察所评估专业建设质量的优势与劣势, 分析原因, 基于此对弱势指标进行大力改进, 保持并优化优势指标, 稳步推进专业建设质量。

## 5. 研究结论

本研究以  $i$  个高校不同专业为例, 通过对较多高校各专业进行实证分析, 以 CIPP 评价模式与专业生命周期理论为理论基础, 以推动发展山西高校专业建设目标为导向, 围绕师资建设、学术科研、教学改革与研究、课程体系建设等对山西省高校专业建设进行了系统性评价。首先, 以教育部颁布的《普通高等本科教育教学审核评估指标体系》作为基础, 并参考专家等给出的建议构建高校专业建设评价指标体系。其次, 以“高等教育质量监测国家数据平台”所提供数据为依据, 通过 RankNet 神经网络模型计算专业建设指标体系中各指标权重。最后通过三级模糊评价模型, 从而可以快速对高校各专业进行打分排名及评级, 然后通过与指标标准数据进行距离计算分析专业建设质量优势劣势; 进而以分析结果与指标距离为依据, 各高校可针对性地对专业建设提出对策与建议并进行改进。

## 致 谢

感谢同事对于本文算法的建设性建议。

## 基金项目

山西省高等学校教学改革项目支持研究。



## 参考文献

- [1] 杨频萍, 汪霞. “双一流”建设背景下我国学科专业评价创新研究[J]. 高校教育管理, 2018, 12(6): 65-73.  
<https://doi.org/10.13316/j.cnki.jhem.20181108.011>
- [2] 李俊龙, 林江辉, 胡锋. 对高校如何开展特色专业建设的认识和思考[J]. 中国大学教学, 2008(4): 59-61.
- [3] 刘洪彬, 刘文合, 王楠, 王洪来, 李竹林, 高爽, 孟思雯. 高等学校本科专业建设质量综合评价指标体系的构建与实践[J]. 高等农业教育, 2021(5): 28-33. <https://doi.org/10.13839/j.cnki.hae.2021.5.005>
- [4] 刘卫东, 熊杨, 张丹平. 基于过程要素模型的新工科专业建设质量分析与评价[J]. 高等工程教育研究, 2019(1): 34-40.
- [5] 吴迪. 云南省高职院校高水平专业群建设绩效评价研究[D]: [硕士学位论文]. 昆明: 云南财经大学, 2022.  
<https://doi.org/10.27455/d.cnki.gycmc.2022.000634>
- [6] 孔庆聪, 李宗山. 基于 CIPP 的“双高”背景下高职院校专业群建设评价体系构建[J]. 长江工程职业技术学院学报, 2023, 40(1): 27-31+36. <https://doi.org/10.14079/j.cnki.cn42-1745/tv.2023.01.007>
- [7] 沈军. 职业院校专业建设 CIPP 评价模式实践研究[D]: [博士学位论文]. 重庆: 西南大学, 2016.
- [8] 吴雪. 基于 CIPP 模型的专业建设研究[J]. 中阿科技论坛(中英阿文), 2019(2): 87-89+117-120.