

# 基于“Fuzzy-GRNN”的硕士研究生学业绩效评价模型研究

康天姝

北京邮电大学, 北京

收稿日期: 2022年9月9日; 录用日期: 2022年10月6日; 发布日期: 2022年10月13日

## 摘要

硕士研究生学业绩效评价对于客观衡量研究生学术水平, 以此为参考建立科学合理的人才培养机制具有重要的现实意义。本文使用Fuzzy-GRNN方法, 基于教育部对硕士研究生提出的四项基本学业绩效构建评价模型。以北京B大学“特殊人才”为例进行了实证研究, 并将评价结果与德尔菲法、BP神经网络进行对比分析, 证明了本文所提出模型的准确性和合理性, 为研究生学业绩效评价提供依据。

## 关键词

硕士研究生评价, 学术能力, 特殊人才, Fuzzy-GRNN

# A Model Study of Academic Competence Evaluation for Graduate Students Based on Fuzzy-GRNN

Tianshu Kang

Beijing University of Posts and Telecommunications, Beijing

Received: Sep. 9<sup>th</sup>, 2022; accepted: Oct. 6<sup>th</sup>, 2022; published: Oct. 13<sup>th</sup>, 2022

## Abstract

The evaluation of academic ability of postgraduates is of great practical significance for objectively measuring the academic level of postgraduates and establishing a scientific and reasonable talent training mechanism. This paper uses the fuzzy GRNN method to construct the evaluation model based on the four basic academic abilities proposed by the Ministry of education. Taking the

“special talents” of Beijing B University as an example, the evaluation results are compared with Delphi method and BP neural network to prove the accuracy and rationality of the model proposed in this paper.

## Keywords

Master's Evaluation, Academic Ability, Special Talents, Fuzzy-GRNN

Copyright © 2022 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

研究生教育对国家创新驱动发展战略的实施起到重要作用[1]。作为研究生培养的最核心部分，学业绩效是衡量研究生教育成果的关键指标。那么，如何评价硕士研究生的学业绩效？目前我国大部分高校对于研究生学术能力的考评局限于论文发表和毕业设计，门槛较低并伴随着次生性问题。这种狭隘的滞后指标非但不利于激发研究生的学习热情，反而会误导学习方式、降低学习成效。

针对硕士研究生学业绩效评价这一研究议题，国内学者开展了积极探讨。诸多学者提出了从学生个体入手提出了以研究生为测度对象的评价模型，都为我们提供了有力的思考与借鉴。表 1 列出了部分文献在进行研究生评价时选取的模型和被解释变量。

**Table 1.** Models and explained variables selected for postgraduate evaluation of existing literature

**表 1.** 现有文献研究生评价所选取的模型和被解释变量

年份	作者	模型基础	被解释变量
2008	刘凌冰, 张先治[2]	平衡记分卡	综合评价
2015	赵仁铃[3]	结构方程模型	综合素质
2016	汪雅霜, 康敏[4]	IEO 模型	学习成果
2016	李圣, 李勇, 王海燕[5]	粒子群优化层次分析	教育质量
2018	张万朋, 柯乐乐[6]	德尔菲法、层次分析	学习成果
2019	宋陈澄, 夏宗洋, 李德根[7]	层次分析模糊评价	学术道德
2019	薛黎明, 王豪杰, 李长明等[8]	集对云模型	科研能力
2021	齐乾坤, 王文龙[9]	数据生命周期	数据素养
2022	马燕, 郭惠芬, 范文翔等[1]	Kirkpatrick 模型	科研绩效

资料来源：作者整理。

通过梳理文献发现，现有关于学生学业评价的研究提出的评价指标存在可操作性及实用性的问题，获取全部数据需要高昂的时间成本及金钱成本，在理论上成立，但缺乏对于实践的指导意义。在研究方法方面，已有文献多以定量测度为主、定性分析为辅，其中定量方法多采用综合模型法，包括因子分析、层次分析等。由于因子分析容易受到旋转方式的干扰，而层次分析法的赋权受主观因素影响较大，我们

认为,在不同的目标上寻找和校准合适的权重是一件费时费力且非常琐碎的事情,并且缺乏理论来指导如何在硕士研究生培养环境下选择这些权重。加之近年来研究生扩招、新学科建设等一系列政策的出台,教育环境不断变化,原有评估方法已不能满足需求。

而与此同时,我们发现“Fuzzy-GRNN”模型对于模糊问题的评价具有较强的适用性,在我国最早由佟泽华[10]提出,用于企业知识共享能力评价。其后,韩春花等[11]又将该网络与BP神经网络、系统聚类分析评价结果进行了比对,证明了“基于Fuzzy + GRNN的评价模型”不仅具有一般人工神经网络的自学习、并行处理等特点,并且收敛速度快、小样本适应性强、容错性及鲁棒性高。GRNN网络与Fuzzy理论的有效结合使得该模型具备了解决模糊问题的能力。

综上,本文试图提出一种行之有效并且便于操作的硕士研究生评价模型,提出将“Fuzzy-GRNN”模型应用于研究生评价,选取了仅需设定SPREAD值一个参数值的GRNN(广义回归神经网络),结合Fuzzy理论,以期获得更加客观的评价结果。

## 2. 数据说明及模型构建

### 2.1. 计量模型

#### 2.1.1. Fuzzy 理论

Fuzzy理论即模糊理论(Fuzzy Theory),是指使用模糊集或连续隶属函数的理论,由美国工程科学院院士扎德(L.A.zadeh)提出,用以描述生活中的不确定性问题。模糊理论凭借其可以对许多复杂的事物或系统进行较好的模糊测量的属性,现如今已经大范围推广并在许多领域进行了广泛的实际应用[12]。

结合研究生评价的特点,本文设定评价集 $Y\{\text{差,较差,一般,良好,优秀}\}$ ,并且依据最大隶属度”原则来确定评价结果。此处应用专家打分来实现数据的模糊化处理划定评价集,步骤如下:

- 1) 选择十位具有多年研究生教学经验的专家;
- 2) 确定评价目标的影响因素,并将各项内容予以陈列及解释说明;
- 3) 提供相关资料,并匿名征求专家意见;
- 4) 收集反馈数据,进行效度分析;
- 5) 进行意见反馈,得到最终数据;

#### 2.1.2. GRNN 网络

GRNN即广义回归神经网络(General Regression Neural Network),是一种基于非线性回归理论的前馈式神经网络模型,由于学习速度快、能够有效处理稀疏数据等优点,被用于系统识别和预测控制,尤其是在复杂的经济社会中进行系统模拟和预测,现在已被学者们广泛使用[13]。

GRNN通过不断激活神经元的方法,得到近似函数,由输入层、模式层、求和层和输出层4层构成[14]:

- 1) 输入层:输入测试样本,节点个数等于样本的特征维度
- 2) 模式层:其神经元数目等于输入样本数,神经元传递函数为:

$$p_i = \exp\left[-(X - X_i)^T (X - X_i) / 2\sigma^2\right], i = 1, 2, 3, \dots, n$$

(其中 $X$ 为网络输入变量、 $X_i$ 为第 $i$ 个神经元对应的学习样本、 $\sigma$ 为光滑因子)

- 3) 求和层:求和层中使用两种类型的神经元分别进行求和一类计算公式为:

$$\sum \exp\left[-(X - X_i)^T (X - X_i) / 2\sigma^2\right], i = 1, 2, 3, \dots, n$$

对所有模式层神经元的输出进行算术求和，模式层与各神经元的连接权值为 1，传递函数为：

$$S_D = \sum p_i$$

另一类为：

$$\sum Y_i \exp\left[-(X - X_i)^T (X - X_i) / 2\sigma^2\right], i = 1, 2, 3, \dots, n$$

对所有模式层神经元的输出进行加权求和，模式层中第  $i$  个神经元与求和层中第  $j$  个分子求和神经元之间的连接权值为第  $i$  个输出样本  $Y_i$  中的第  $j$  个元素，传递函数为：

$$S_{N_{ij}} = \sum y_{ij} P_i, j = 1, 2, 3, \dots, m$$

4) 输出层：输出层节点个数等于标签向量的维度，表达式为：

$$y_j = \frac{S_{N_{ij}}}{S_D}, j = 1, 2, 3, \dots, m$$

Fuzzy-GRNN 模型的“Fuzzy-GRNN”模型可理解为 Fuzzy 理论与 GRNN (广义回归神经网络)的结合，其网络结构如图 1 所示。

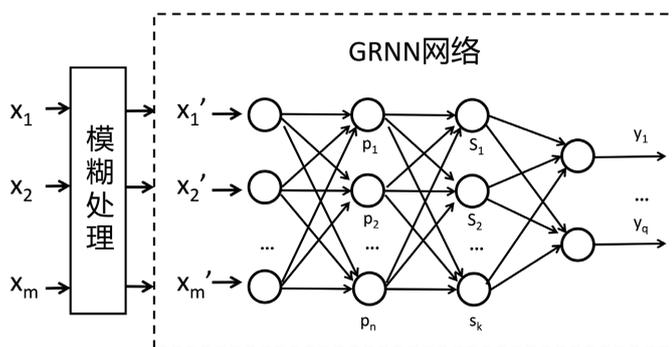


Figure 1. Schematic diagram of Fuzzy GRNN network structure  
图 1. Fuzzy-GRNN 网络结构示意图[9]

## 2.2. 指标的构建

### 2.2.1. 构建原则

本文的指标选取“集中在那些真正起到战略性影响的、至关重要的少数指标上” [15]，致力于寻找研究生评价中的关键绩效指标(Key Performance Indicators, KPI)。关键绩效指标是用来衡量某一组织或人员绩效表现的具体量化指标，是对工作完成效果的最直接衡量方式[2]。

本文的关键技术指标设计遵守 SMART 原则，该原则由彼得·德鲁克提出，该原则要求指评价目标必须是具体的(special)、可量化(measurable)、可实现(achievable)；与战略目标一致(relevant)以及在某一时间区间内能够完成(timely)。

### 2.2.2. 指标体系

基于上述原则，本文构建了包含目标层、准则层和指标层在内的三层指标体系。

目标层即为本文的评价目标——硕士研究生的学业绩效。在准则层的设计上，我们发现，关于硕士研究生的学术能力，国务院学位委员会及教育部从学位授予的角度对获取硕士学位应具备的基本学术能力提出了要求[16]，明确每项素质学科的职责，确保研究生教育的基础质量，创新机制，激发学位授予单位追求卓越的热情和创造力，不断提高人才培养水平。其中明确提出如下四项：获取知识的能力、科学

研究能力、实践能力及学术交流能力。因此，将上述四项能力作为本文的准则层。更进一步的，结合目前我国高等教育的考评体系，我们提炼出能够反映上述能力的各项学业成果，在指标层的设计上，包括英语水平( $X_1$ )、软件运用情况( $X_2$ )、论文发表情况( $X_3$ )、知识产权数量( $X_4$ )；参与项目情况( $X_5$ )、获奖情况( $X_6$ )、参会情况( $X_7$ )。如表 2 所示。

**Table 2.** Evaluation index system of academic performance level of postgraduates

**表 2.** 硕士研究生学业绩效水平评价指标体系

目标层	准则层	指标层
学业绩效	获取知识	英语水平( $X_1$ )
		学习软件运用情况( $X_2$ )
	科学研究	学术论文发表情况( $X_3$ )
		知识产权情况( $X_4$ )
	实践能力	课题项目参与情况( $X_5$ )
		竞赛及评比获奖情况( $X_6$ )
	学术交流	学术会议参与情况( $X_7$ )

### 3. 实证分析

本文选取了北京 B 大学通过“特殊人才保研政策”保送研究生的同学作为研究对象。原因如下：2006 年教育部印发《本科毕业生免试攻读硕士学位研究生工作管理办法》，明确提出，具有突出的学术专长或培养潜质的应届本科毕业生可以不受综合排名限制[9]。至此，“特殊人才保研”成为我国本科生保研的一个重要通道。具有推免资质的高校纷纷开始“特殊人才保研”的实践探索，在校级层面制定并出台相关政策进行人才选拔及后续推免工作。但普遍表现出以选拔出学生作为政策的终点，而缺乏后续跟踪及对于政策方案的反馈与完善。该政策的实施效果如何？在选拔时因突出的创新能力而被寄予厚望的“特殊人才”们是否满足了政策期待？本文试图通过对于北京 B 大学“特殊人才”的追踪予以回答。北京 B 大学是由教育部直属、工信部共建的全国重点大学，位列国家“211 工程”、“世界一流学科建设高校”，在部委直属高校中具有一定的代表性。

具体操作如下：

- 1) 模糊处理：对各样本的学业绩效进行专家打分，得到模糊数据矩阵；
- 2) 模型训练：将数据集分为训练样本和测试样本，利用训练样本进行模型训练；
- 3) 模型测试：将测试样本代入训练好的 GRNN 网络，进行测试并输出结果。

#### 3.1. 数据收集及处理

据北京 B 大学信息门户显示，该校的“特殊人才保研”政策可追溯到 2012 年，首届针对 2009 级本科生。根据北京 B 高校教务处发布得相关文件，“对有特殊学术专长或创新实践能力强、具有突出培养潜质的学生，可不受综合排名限制”，申请获得推免生资格。由于本文选取的评价区间为完整的研究生阶段(即已经完成硕士学位)，共 25 人符合条件，受信息可及性限制，笔者共接触到 21 位研究对象，对其逐一进行了深入访谈，获取第一手资料并进行归纳整理。

接下来邀请 10 位资深教育专家，对每位同学的学业绩效给出“[差，较差，一般，良好，优秀]”的评价结果。如表 3 所示，以样本 A 为例，其  $X_1$  的评价集为[0, 0, 0.9, 0.1, 0]即表示对于样本 A 的  $X_1$  项，

有 9 位专家认为“一般”、1 位认为“良好”。期望输出为[0, 0, 0, 0.1, 0.9]，根据最大隶属度原则，其综合评价结果为“优秀”。

**Table 3.** Evaluation index data of postgraduates exempted from examination recommended by Beijing B University for “Special Talents”

**表 3.** 北京 B 大学“特殊人才”推荐免试研究生评价指标数据

样本	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>	期望输出	
训练集	A	[0, 0, 0.9, 0.1, 0]	[0, 0, 0.2, 0.8, 0]	[0, 0, 0, 0.2, 0.8]	[0, 0, 0, 0.3, 0.7]	[0, 0, 0.1, 0.9, 0]	[0, 0, 0, 0, 1]	[0, 0, 0, 0.6, 0.4]	[0, 0, 0, 0.1, 0.9] 优秀
	B	[0, 0, 0.1, 0.8, 0.1]	[0, 0, 0.1, 0.8, 0.1]	[0, 0.1, 0.9, 0, 0]	[0, 0, 1, 0, 0]	[0, 0, 0.4, 0.6, 0]	[0, 0.1, 0.9, 0, 0]	[0, 0, 1, 0, 0]	[0, 0, 0.9, 0.1, 0] 一般
	C	[0, 0, 0.3, 0.7, 0]	[0, 0, 0.1, 0.8, 0.1, 0]	[0, 0, 0, 0.3, 0.7]	[0, 0, 0, 0.2, 0.8]	[0, 0, 0.3, 0.7, 0]	[0, 0, 0.1, 0, 0.9]	[0, 0, 0.1, 0.1, 0.8]	[0, 0, 0, 0.4, 0.6] 优秀
	D	[0, 0, 0.1, 0.9, 0]	[0, 0, 0.1, 0.8, 0.1]	[0, 0, 0.4, 0.4, 0.2]	[0, 0, 1, 0, 0]	[0, 0, 0.4, 0.6, 0]	[0, 0.1, 0.9, 0, 0]	[0, 0, 1, 0, 0]	[0, 0, 0.8, 0.2, 0] 一般
	E	[0, 0, 0, 0.9, 0.1]	[0, 0, 0, 0.8, 0.2]	[0, 0, 0.4, 0.5, 0.1]	[0, 0, 1, 0, 0]	[0, 0, 0.2, 0.8, 0]	[0, 0, 0.1, 0.9, 0]	[0, 0, 0.1, 0.9, 0]	[0, 0, 0.1, 0.9, 0] 良好
	F	[0, 0, 0.9, 0.1, 0]	[0, 0, 0.6, 0.4, 0]	[0, 0, 0, 0, 1]	[0, 0, 0, 0.1, 0.9]	[0, 0, 0, 0.2, 0.8]	[0, 0, 0, 0.2, 0.8]	[0, 0, 0, 0, 1]	[0, 0, 0, 0, 1] 优秀
	G	[0, 0, 0.9, 0.1, 0]	[0, 0, 0, 0.7, 0.3]	[0, 0.1, 0.5, 0.4, 0]	[0, 0, 1, 0, 0]	[0, 0, 0.2, 0.8, 0]	[0, 0.1, 0.9, 0, 0]	[0, 0.1, 0.9, 0, 0]	[0, 0, 0.9, 0.1, 0] 一般
	H	[0, 0, 0, 0, 1]	[0, 0, 0.1, 0.9, 0]	[0, 0, 0.1, 0.4, 0.5]	[0, 0, 0.1, 0.4, 0.5]	[0, 0, 0.1, 0.9, 0]	[0, 0.1, 0.9, 0, 0]	[0, 0, 0, 0.4, 0.6]	[0, 0, 0.1, 0.7, 0.2] 良好
	I	[0, 0, 0.1, 0.8, 0.1]	[0, 0, 0.2, 0.8, 0]	[0, 0, 0, 0.8, 0.2]	[0, 0, 0, 0.9, 0.1]	[0, 0, 0.2, 0.8, 0]	[0, 0, 0, 0.9, 0.1]	[0, 0, 1, 0, 0]	[0, 0, 0.1, 0.9, 0] 良好
	J	[0, 0, 0, 0.9, 0.1]	[0, 0, 0, 0.9, 0.1]	[0, 0.1, 0.9, 0, 0]	[0, 0, 1, 0, 0]	[0, 0, 0.1, 0.4, 0.5]	[0, 0, 0, 0.2, 0.8]	[0, 0, 1, 0, 0]	[0, 0, 0, 0.9, 0.1] 良好
	K	[0, 0, 0.9, 0.1, 0]	[0, 0, 0, 1, 0]	[0, 0, 0.1, 0.4, 0.5]	[0, 0, 1, 0, 0]	[0, 0, 0.2, 0.8, 0]	[0, 0, 1, 0, 0]	[0, 0, 0.1, 0.9, 0]	[0, 0, 0.2, 0.8, 0] 良好
	L	[0.1, 0.7, 0.2, 0, 0]	[0, 0, 0, 0.4, 0.6]	[0, 0, 0.4, 0.6, 0]	[0, 0, 1, 0, 0]	[0, 0, 0, 0.6, 0.4]	[0, 0.1, 0.8, 0.1, 0]	[0, 0, 1, 0, 0]	[0, 0, 0.7, 0.3, 0] 一般
	M	[0, 0.1, 0.8, 0.1, 0]	[0, 0, 0, 0.8, 0.2]	[0, 0, 0, 0.7, 0.3]	[0, 0, 0, 0.4, 0.6]	[0, 0, 0.1, 0.9, 0]	[0, 0, 0, 0.2, 0.8]	[0, 0, 1, 0, 0]	[0, 0, 0, 0.9, 0.1] 良好
测试集	N	[0, 0, 0, 0, 1]	[0, 0, 0.1, 0.8, 0.1]	[0, 0, 0.1, 0.4, 0.5]	[0, 0, 0.1, 0.3, 0.6]	[0, 0, 0.1, 0.9, 0]	[0, 0.1, 0.9, 0, 0]	[0, 0, 0, 0.4, 0.6]	
	O	[0, 0, 0.3, 0.7, 0]	[0, 0, 0.1, 0.8, 0.1]	[0, 0, 0, 0.3, 0.7]	[0, 0, 0, 0.2, 0.8]	[0, 0, 0.3, 0.7, 0]	[0, 0, 0.1, 0.9, 0]	[0, 0, 0.1, 0.1, 0.8]	
	P	[0, 0, 0.9, 0.1, 0]	[0, 0, 0, 1, 0]	[0, 0, 0.1, 0.4, 0.5]	[0, 0, 0.1, 0.9, 0]	[0, 0, 0.2, 0.8, 0]	[0, 0, 1, 0, 0]	[0, 0, 0.1, 0.9, 0]	
	Q	[0, 0, 0.1, 0.9, 0]	[0, 0, 0.1, 0.8, 0.1]	[0, 0, 0.5, 0.3, 0.2]	[0, 0, 1, 0, 0]	[0, 0, 0.4, 0.6, 0]	[0, 0.1, 0.9, 0, 0]	[0, 0, 1, 0, 0]	
	R	[0.1, 0.7, 0.2, 0, 0]	[0, 0, 0, 0.4, 0.6]	[0, 0, 0.4, 0.6, 0]	[0, 0.1, 0.8, 0.1, 0]	[0, 0, 0, 0.7, 0.3]	[0, 0.1, 0.8, 0.1, 0]	[0, 0, 1, 0, 0]	
	S	[0, 0, 0.1, 0.8, 0.1]	[0, 0, 0.1, 0.9, 0]	[0, 0.1, 0.9, 0, 0]	[0, 0, 1, 0, 0]	[0, 0, 0.4, 0.6, 0]	[0, 0.1, 0.9, 0, 0]	[0, 0, 0.9, 0.1, 0]	
	T	[0, 0, 0, 0.8, 0.2]	[0, 0, 0, 0.8, 0.2]	[0, 0, 0.4, 0.5, 0.1]	[0, 0, 0.8, 0.2, 0]	[0, 0, 0.2, 0.8, 0]	[0, 0, 0.1, 0.9, 0]	[0, 0, 0.2, 0.8, 0]	
	U	[0, 0, 0, 0.9, 0.1]	[0, 0, 0, 0.8, 0.2]	[0, 0, 0.5, 0.4, 0.1]	[0, 0.1, 0.9, 0, 0]	[0, 0, 0.1, 0.9, 0]	[0, 0, 0.2, 0.8, 0]	[0, 0, 0.1, 0.9, 0]	

## 3.2. 网络训练及输出

### 3.2.1. 仿真训练

在 MATLAB 软件中使用前 13 组样本数据数据为训练样本进行仿真。使用 `newgrnn` 函数编写 GRNN 程序，将处理后的模糊数据矩阵作为输入神经元进行网络训练。输入层共包含 35 个神经元，其中存在有 7 个指标，由于进行了模糊化处理，每个指标内包含 5 个元素，所以输入层神经元总数量为 35 (7 \* 5) 个，隐含层神经元数量与输入层一致也是 35 个，输出层神经元数量  $1 \times 5 = 5$  个。网络结构示意图如图 2 所示：

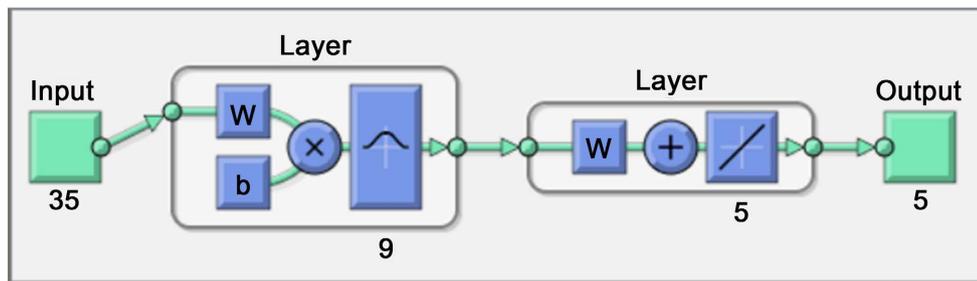


Figure 2. Schematic diagram of network structure  
图 2. 本文网络结构示意图

### 3.2.2. SPREAD 值的选取

在 GRNN 网络中，spread 值的选取尤为重要。由于样本量较少，本文采取交叉验证方法进行训练，并通过循环得到最佳 spread 值。

本程序设置 spread 值从 0.1 到 10 每 0.1 递增循环验证，每个 spread 值都会输出一组评价结果，取评价结果误差最小的那一组数据对应的 spread 值为最佳 spread 值。执行命令后得到图 3，当光滑因子 SPREAD 取值为 1 时，验证集均方误差最低，网络达到最佳评价效果。因而本文选取  $\text{spread} = 1$  时的输出数据作为评价结果。

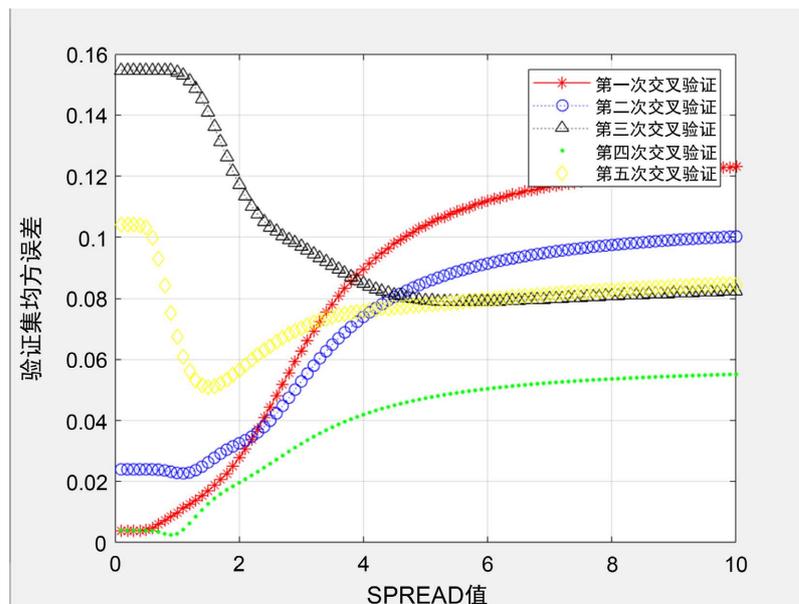


Figure 3. Mean square error under different SPREAD values  
图 3. 不同 SPREAD 值下的均方误差

### 3.2.3. 网络输出结果

最终输出数据表 4 所示(只显示部分), 并依据最大隶属度原则确定评价结果。

**Table 4.** Fuzzy GRNN network evaluation results

**表 4.** Fuzzy-GRNN 网络评价结果

样本	评价集	评价结果
N	[0, 0, 0.1, <u>0.7</u> , 0.2]	良好
O	[0, 0, 0.0977, <u>0.8822</u> , 0.0201]	良好
P	[0, 0, 0.0029, 0.4251, <u>0.5702</u> ]	优秀
Q	[0, 0, <u>0.9</u> , 0.1, 0]	一般
R	[0, 0, <u>0.7</u> , 0.3, e]	一般
S	[0, 0, <u>0.8998</u> , 0.1002, e]	一般
T	[0, 0, 0.1, <u>0.9</u> , e]	良好
U	[0, 0, 0.1, <u>0.9</u> , e]	良好

### 3.3. 与德尔菲法、BP 神经网络评价结果的比较分析

#### 3.3.1. 与德尔菲法评价结果的比较分析

为了证明该模型的准确性, 本文与研究生评价使用的传统方法德尔菲法的评价结果进行比较分析。德尔菲法是指通过与有关专家协商, 对专家意见进行统计、归纳和分析, 根据专家经验形成主观判断, 并对难以通过软件分析等方法量化的相关因素进行适切的估计。我们邀请十位专家对测试集的八份样本进行评价, 给出[差, 较差, 一般, 良好, 优秀]五种评价, 最终依据“最大隶属度”原则确定评价结果, 见表 5。

**Table 5.** Evaluation results of delphi method and “Fuzzy GRNN” method

**表 5.** 德尔菲法与“Fuzzy-GRNN”方法评价结果

样本	德尔菲法评价结果	Fuzzy-GRNN 评价结果	
N	良好	[0, 0, 0.1, <u>0.7</u> , 0.2]	良好
O	良好	[0, 0, 0.097 7, 0.882 2, 0.020 1]	良好
P	优秀	[0, 0, 0.002 9, 0.425 1, 0.570 2]	优秀
Q	一般	[0, 0, 0.9, 0.1, 0]	一般
R	一般	[0, 0, 0.7, 0.3, e]	一般
S	一般	[0, 0, 0.899 8, 0.100 2, e]	一般
T	良好	[0, 0, 0.1, <u>0.9</u> , e]	良好
U	良好	[0, 0, 0.1, <u>0.9</u> , e]	良好

由表 5 可得, “基于 Fuzzy-GRNN 的研究生学术能力评价模型”的评价结果与德尔菲法的结果是完全一致的, 证明了该模型的准确性。但在实际操作中, Fuzzy-GRNN 明显具有更强的可操作性和更高的效率。

### 3.3.2. 与 BP 神经网络评价结果的比较分析

为了进一步说明本文模型的合理性，本文又与 BP 网络进行了比较分析。在参数的设置上，设定训练目标 goal 为  $1e-006$ ，最大训练次数 epochs 为 10,000 步，学习率 lr 应设置为较少值，过大虽然会在开始加快收敛速度，但临近最佳点时会产生动荡，而致使无法收敛，因而本文设定 lr 为 0.01。接下来，使用不同的学习算法进行仿真，如表 6 所示。

**Table 6.** Evaluation results of BP neural network

**表 6.** BP 神经网络评价结果

样本	最速下降法 traingd	弹性算法 trainrp	动量算法 traingdm	Fletcher-Reeves 修正算法
N	[0,0,0.286,0.668,0.651]	[0,0,0.093,0.376,0.004]	[0,0,0.149,0.323,0.649]	[0,0,0.001,0.823,0.903]
O	[0,0,0.016,0.496,0.369]	[0,0,0.012,0.730,0.428]	[0,0,0.003,0.888,0.016]	[0,0,0.002,0.900,0.185]
P	[0,0,0.118,0.782,0.020]	[0,0,0.670,0.269,0.845]	[0,0,0.158,0.867,0.026]	[0,0,0.007,0.900,0.245]
Q	[0,0,0.027,0.054,0.081]	[0,0,0.539,0.004,e]	[0,0,0.007,0.893,e]	[0,0,0.877,0.892,0.000]
R	[0,0,0.650,0.896,0.001]	[0,0,0.873,0.334,0.002]	[0,0,0.644,0.383,0.060]	[0,0,0.800,0.598,0.002]
S	[0,0,0.159,0.828,0.085]	[0,0,0.898,0.208,0.006]	[0,0,0.835,0.153,0.200]	[0,0,0.633,0.192,0.006]
T	[0,0,0.094,0.684,0.566]	[0,0,0.030,0.859,0.002]	[0,0,0.165,0.727,0.044]	[0,0,0.002,0.900,0.565]
U	[0,0,0.100,0.693,0.256]	[0,0,0.028,0.847,0.002]	[0,0,0.189,0.719,0.091]	[0,0,0.008,0.900,0.422]
训练步数	348 步收敛	141 步收敛	973 步收敛	26 步收敛
评价结果	部分不准确，不稳定	接近准确，不太稳定	部分不准确，不稳定	稳定但不准确，偏差较大

由仿真可得：1) 采用最速下降法得到的评价结果存在偏差，如 Q 样本[0, 0, 0.0270, 0.0545, 0.0813]输出数据均较小，难以确定评价结果，且与德尔菲法评价结果相比偏差明显。2) 采用弹性算法得到的评价结果与德尔菲法和 Fuzzy-GRNN 方法一致，但输出数据不太稳定，且效率不如 Fuzzy-GRNN。3) 采用动量算法经过 973 步才收敛，效率明显降低，而且评价结果偏差较大。4) 采用变梯度的 Fletcher-Reeves 修正算法收敛很快，效率较高且数据稳定，但评价结果不准确。

由此可见，在本文语境下，当样本量较小时，BP 神经网络的训练结果不稳定并且偏差较大。相对而言 Fuzzy-GRNN 的效率更高，具有更好的鲁棒性和容错性，且训练结果更加准确，进一步验证了本文所选模型的合理性。

### 3.4. 评价结果与可能的解释

全部 21 份样本的评价结果为：“优秀”、“良好”、“一般”、“较差”、“差”分别占比 19.05%、47.62%、33.33%、0、0。

接下来，我们对上述发现做一些讨论和解释。已知 21 位同学均顺利取得硕士学位，满足学位授予的各项要求，不存在“差”、“较差”两项评价，且其中有 14 人获得良好及以上评价，这意味着 66.7% 的特殊人才保研的同学在硕士研究生阶段取得的学业绩效高于平均水平，且有部分同学创造了突出的学业成果，这可能是因为：

1) 项目经验丰富。在如今的培养模式下，本科生科研成果的孵化过程表现为以竞赛为牵引、以项目为依托，“特殊人才”大多在本科阶段就在主流学科竞赛或双创竞赛中取得优异成绩，积累了丰富的项目经验，这种培养方式与研究生培养一致，使得他们能更快适应研究生教育模式从而更早产生学业成果。

2) 学术投入高。“特殊人才”既然能在本科阶段取得突出学术成果,证明其在学术上进行了大量投入,而这种投入与本科通识教育属性有所不同。“具有创新能力的人们更倾向于创造出创新成果,而未来创新成果的最佳预测指标是过去的创新行为。”[17]本科阶段的项目经历有助于更早地确定研究方向,其研究生阶段的科研工作可基于本科阶段取得的成果之上而进行深入研究。

3) 导学互动紧密。学生在选择导师时倾向于选择了解并认可的老师。“特殊人才”由于前期项目经历丰富,在本科阶段与更多老师有更密切的接触,为其研究生导师的选择提供依据并提前建立默契。我们在调研中发现,大部分“特殊人才”的研究生导师在其本科阶段曾参与的竞赛或课题中予以指导,而研究表明,师生间交流的密切程度和学生的创新能力之间呈正相关关系且十分显著[18]。

## 4. 结论与讨论

结合上述分析,可以得出如下结论:

1) 通过特殊人才保研的学生大部分具备突出培养潜质。特殊人才保研政策确实选拔出了部分综合素质高、培养潜质强的人才,给予特殊人才以深造机会、不使其埋没,有助于推进研究生教育深入实施,进而推动高水平大学建设。但政策覆盖面较小,带动性不强,政策运行仍有进一步提升的空间。

2) “Fuzzy-GRNN”方法可以应用于教育领域。不同专业的指标数值存在较大差异,该模型基于模糊理论,在分析这类数据集方面具有突出优势。并且可通过交叉循环得到 SPREAD 值,回避了以往研究方法中人为确定权重的问题,精度较高且不易受极端值干扰。因此,本研究一定程度上兼具理论意义及实践意义。

3) 基于“Fuzzy-GRNN”的硕士研究生评价模型所得到的评价结果是准确且合理的。一方面,其评价结果与传统的学生评价方法德尔菲法的评价结果完全一致,证明了该模型的准确性,同时表现出更高的效率和可操作性。另一方面,在评价本文语境下的小样本事件时,其评价结果比传统的神经网络 BP 网络更加准确和稳定,证明了选择该模型的合理性。

本文研究尚存在两个方面的不足,这也是后续的研究重点和改进方向:

1) 虽然 Fuzzy-GRNN 模型逻辑清晰并且具有统计理论支撑,其估计结果是可靠的,但本文采用的关键绩效指标关注的是对于硕士研究生学业过程中重点学习活动的衡量,并非整个学业过程的衡量,更加强调“学业成果导向”的质量观。

2) 本文缺乏大样本、不同类型高校之间的比较,样本局限于一所学校且受数据可及性影响而未实现全样本覆盖,虽然 GRNN 可通过交叉训练一定程度上缓解上述问题,但所得结论仍有局限,有望在后续的研究中深入研究并加以解决。

## 基金项目

北京邮电大学研究生教育教学改革与研究项目“新时期研究生‘课程思政’的学术接口及其实现路径研究”项目编号 2021Y009。

## 参考文献

- [1] 马燕,郭惠芬,范文翔,胡慧丽.基于 AHP-熵值法的硕士研究生科研绩效评价研究[J].黑龙江高教研究,2022,40(8):97-103. <https://doi.org/10.19903/j.cnki.cn23-1074/g.2022.08.016>
- [2] 刘凌冰,张先治.基于平衡记分卡的研究生综合评价模型研究[J].中国高教研究,2008(12):34-38.
- [3] 赵仁铃.基于结构方程模型的研究生考生综合素质评价模型研究[J].图书与情报,2015(5):92-98.
- [4] 汪雅霜,康敏.硕士研究生学习成果评价的实证研究——基于 IEO 模型的分析结果[J].国家教育行政学院学报,2016(3):76-82.

- 
- [5] 李圣, 李勇, 王海燕. 研究生教育质量评价指标体系模型构建与研究——基于 PSO-AHP 分析方法的应用与实证[J]. 研究生教育研究, 2016(5): 53-59.
- [6] 张万朋, 柯乐乐. 基于德尔菲法和层次分析法的研究生学习成果评价研究——以教育经济与管理专业为例[J]. 现代大学教育, 2018(1): 93-99.
- [7] 宋陈澄, 夏宗洋, 李德根. 研究生学术道德的 AHP-模糊综合评价模型及应用[J]. 黑龙江畜牧兽医(下半月), 2019(4): 163-167.
- [8] 薛黎明, 王豪杰, 李长明, 朱琳龙, 郑志学, 杜玉琳. 基于集对云模型的矿业高校研究生科研能力评价[J]. 中国矿业, 2019, 28(2): 175-180.
- [9] 齐乾坤, 王文龙. 基于数据生命周期的高校研究生数据素养评价研究[J]. 情报科学, 2021, 39(9): 125-130+145. <https://doi.org/10.13833/j.issn.1007-7634.2021.09.017>
- [10] 佟泽华. 基于 Fussy-GRNN 的企业知识共享能力评价研究[J]. 情报理论与实践, 2012, 35(1): 75-79.
- [11] 韩春花, 佟泽华. 基于 Fussy-GRNN 网络的区域创新能力评价模型研究[J]. 科技管理研究, 2016, 36(14): 55-60.
- [12] 胡宝清. 模糊理论基础[M]. 武汉: 武汉大学出版社, 2004.
- [13] 叶焯, 李贵才, 李莉, 王乾, 张华. 国家级新区功能定位及发展建议——基于 GRNN 潜力评价方法[J]. 经济地理, 2015, 35(2): 92-99.
- [14] 史峰. MATLAB 神经网络 30 个案例分析[M]. 北京: 北京航空航天大学出版社, 2010.
- [15] Schneiderman, A.M. (2001) The First Balanced Scorecard: Analog Devices, 1986-1988. *Journal of Cost Management*, No. 5, 13-18.
- [16] 教育部. 国务院学位委员会、教育部关于加强学位与研究生教育质量保证和监督体系建设的意见(学位[2014]3号) [EB/OL]. [http://www.moe.gov.cn/srcsite/A22/s7065/201402/t20140212\\_165554.html](http://www.moe.gov.cn/srcsite/A22/s7065/201402/t20140212_165554.html), 2021-05-12.
- [17] 罗伯特·J·斯滕博格. 创造力手册[M]. 施建农, 等, 译. 北京: 北京理工大学出版社, 2007.
- [18] 王晓辉. 师生关系对研究生创新能力的研究[D]: [硕士学位论文]. 武汉: 华中师范大学, 2010.