



开放科学
(资源服务)
标识码
(OSID)

基于机器学习的科技人才综合学术水平评价指标与模型研究

李勃慧 王运红 杨代庆 郑楚华 陈国娇

中国科学技术信息研究所 北京 100038

摘要: [目的/意义] 从学术诚信、科研产出、科研活动和学术声誉 4 个维度,打破单一指标局限性,为人才评价中“破四唯”“立新标”提供新的思路和参考。[方法/过程] 针对科技人才的特征,构建了 98 项定量指标与定性特征指标。选取研究样本,进行数据集构建和数据预处理,形成最终识别预测模型的输入数据;通过机器学习获得样本特征指标与评价结果之间的隐含关系,比较 9 种模型算法下的人才学术能力表现,综合得出最优模型进行后续训练和调优,通过对比模型输出与实际结果进行模型评价。[局限] 模型仅基于实验样本数据量和特征,推广应用还需要更多的样本数据进行训练调优。[结果/结论] 实证研究表明,模型对高水平科技人才评价具备较好的适用性,为科技人才评价提供新的视角和方法。

关键词: 科技人才; 人才评价; 人才识别; 评价模型

中图分类号: G350

Research on Comprehensive Academic Level Evaluation Indicators and Models for Scientific and Technological Talents Based on Machine Learning

LI Bohui WANG Yunhong YANG Daiqing ZHENG Chuhua CHEN Guojiao

Institute of Scientific and Technical Information of China, Beijing 100038, China

Abstract: [Objective/Significance] This study breaks the limitations of a single indicator from four dimensions: academic integrity, research output, research activities, and academic reputation, providing new ideas and references for breaking the “4 Single dimension” and “setting new standards” in talent evaluation. [Methods/Processes] For the characteristics of scientific and technological talents, 98 quantitative indicators and qualitative characteristic indicators were constructed. This study selected research samples for dataset construction and data preprocessing to form the input data for the final recognition and prediction

基金项目 国家社会科学基金项目“大数据环境下同行评议方法模式研究”(19BTQ082)。

作者简介 李勃慧(1994-),女,硕士,助理研究员,研究方向为科技评价、数据挖掘, E-mail: libh@istic.ac.cn; 王运红(1971-),女,硕士,研究员,研究方向为科学计量、科技评价; 杨代庆(1975-),男,硕士,研究员,研究方向为科学计量、科技评价。郑楚华(1991-),女,博士,助理研究员,研究方向为科技人才评价; 陈国娇(1992-),女,博士,助理研究员,研究方向为科技人才评价。

引用格式 李勃慧,王运红,杨代庆,等. 基于机器学习的科技人才综合学术水平评价指标与模型研究[J]. 情报工程, 2024, 10(5): 115-127.

model; Obtain the implicit relationship between sample feature indicators and evaluation results through machine learning, compare the academic performance of talents under 9 model algorithms, comprehensively determine the optimal model for subsequent training and optimization, and evaluate the model by comparing the model output with actual results. [Limitations] This research model is only based on experimental sample data size and features, and further training and optimization of sample data are needed for its widespread application. [Results/Conclusions] Empirical research has shown that the model has good applicability for the evaluation of scientific and technological talents, providing new perspectives and methods for the evaluation of scientific and technological talents.

Keywords: Scientific and Technological Talents; Talent Evaluation; Talent Identification; Evaluation Models

引言

科学、合理的科技人才评价体系是人才工作的“指挥棒”，不仅影响着科技人才的质量和水平，同时影响科技人才引进、管理、培养和使用等。近年来，中共中央办公厅、国务院办公厅、科技部、教育部等部门针对科技人才评价连续发文；2022年，科技部会同教育部等7部门联合召开科技人才评价改革试点启动会，部署推进《关于开展科技人才评价改革试点的工作方案》，推动人才评价改革落地见效，着力克服“唯论文、唯职称、唯学历、唯奖项”倾向。由此可见，对科技人才的评价问题不仅仅是学术界研究和探讨的热点，更是影响国家人才培养、使用的重要命题。

科技人才评价改革提出“破四唯”，但是以“立新标”为突破口的研究还在探索中。在“破四唯”的要求下，科技人才能力的评价必须打破原来的标准，对科技人才的科研综合能力能够进行全方位、客观、合理和可操作的评价。当前科技人才评价工作还存在一些问题，如缺乏科学合理、各有侧重的人才评价标准，评价标准动态更新调整机制不够及时有效；在采用成果影响力评价科研能力时，将SCI和国内核

心期刊论文发表数量、引用榜单和影响因子排名等作为评价参考的重点，忽视了标志性成果的质量、项目能力、学术活跃度、学术声誉贡献等综合指标；评价指标的主观性较强，不利于操作性和公平性。

本研究针对科技人才的个体特征和学术特性，构建了4个维度，包含98项定量指标与定性指标。同时，选取了531名工程院院士候选人作为研究对象，进行数据集构建和数据预处理，形成最终识别预测模型的输入数据；通过机器学习获得样本特征指标与评价结果之间的隐含关系，比较9种模型算法下的人才学术能力表现，综合得出最优模型进行后续训练和调优。实证结果表明，该方法可应用于科技人才的评价和识别工作实践中，为多维度评价指标下，应用机器学习模型进行科技人才的评价识别提供参考。

1 国内外研究现状

1.1 科技人才评价方法

学术界呼吁当前亟须创新人才评价机制，建立健全以创新能力、质量、贡献为导向的科

科技人才评价识别体系，形成并实施有利于科技人才潜心研究和创新的评价制度。科技人才评价则是对个体能力、表现和潜力进行系统性分析和评估，在科技人才评价方法研究上，国内学者开展的主要研究如下。

一是同行评议、专家访谈、调查问卷等定性研究方法。这些定性方法在揭示科技人才评价的隐性因素和复杂关系等方面具有明显优势，但存在过度依赖主观分析和判断影响、获取和处理专家意见带来的人力成本、不同环境或条件下的适用性和可迁移性差等问题。

二是基于科学计量学的定量研究方法。通常采用“构建评价指标体系—设置各项参评指标权重或赋分规则—计算得分”路径，研究重点和创新点多集中在新型评价指标的选取和指标权重确定方法两方面。在评价指标选取方面，刘璇等^[1]通过合著网络的中心度指标发现能够带领整个领域发展方向的学术带头人，宋培彦等^[2]提出颠覆性指数，基于文献引用的聚集效应评估了文章的创新程度，常鏐鏐等^[3]使用调和H指数综合合著者数量和合著者顺序影响，实现了合著贡献程度的量化。在指标权重设置方面，常见的权重确定方法包括利用数字的相对大小信息进行权重计算，通常与专家打分或问卷调研等方法结合使用，常见方法包括层次分析法和优序图法等；根据各项指标值的变异程度进行权重计算，常见方法包括熵权法^[4]、CRITIC等；基于变量之间的相关性进行权重计算，常见方法包括主成分分析法、因子分析法^[5]等。此类科学计量学方法通过构建指标体系和计算得分，提供了一种相对客观和标准化的评价过程，减少了主观判断的干扰；强调评价指

标体系的系统性构建，能够从多个维度全面反映被评价对象的综合素质和能力。然而，依赖线性模型的方法忽视了非线性的复杂关系，丢失信息的丰富性，可能导致评价结果的简化和失真。

三是基于数据挖掘及机器学习模型构建的研究方法。该方法能够自动从大规模数据集中识别复杂非线性关系，降低了人工判断的主观性以及传统计量分析的约束性。例如基于神经网络^[6]的动态社交分析高潜力人才早期识别，基于支持向量机^[7]的高、低层次人才二分类研究，基于决策树C4.5分类器^[8]的人才绩效模式预测，基于动态神经网络^[9]的人才流动和工作绩效的建模，基于递归神经网络^[10]的人才招聘场景下求职者经验与匹配度评价，基于决策树^[11]的人才类型与人才结构划分等。尽管上述基于机器学习的人才评价方法在多个数据集和应用场景中较为有效，但此类方法在指标维度和数据集的构建方面具有高度依赖性，在科技人才评价场景下的适应性及效能尚未得到验证。

1.2 科技人才综合学术水平评价维度

单一的评价维度已不能反映人才的综合学术能力和水平，不再是学者研究的重点。随着人才评价体系研究的不断完善和深化，对科技人才综合学术水平的评价呈现出以下三方面的特征。

一是评价维度更加全面多元，突破了传统维度中对于评价学术历程和经验的偏向性，考虑创新能力、团队领导力、产学研结合能力等。例如，田军等^[12]针对陕西省科技人才评价的实

际需要,提出创新知识、道德素质、创新动机、影响力、创新能力、产出绩效6维度评价模型;丁宁等^[13]基于医学人才的学科特点,提出临床、科研、教学相结合的指标体系;张熠等^[14]从创新驱动对人才的需求出发,建立了以基本素质为基础、以创新能力为核心、以创新成果为导向的评价体系。余波等^[15]在传统高校人才评价指标基础上提出了社会影响力和可持续发展、国际化和国际交流等指标。王运红等^[16]在对科技人才科研综合能力分析的基础上,构建科技人才科研综合能力评价模型,设计基于科技信息大数据的评价指标体系,包括基本素养、科研产出影响力、科研管理能力、学术潜力和学术地位5个维度,并研究其在人才引进工作中的应用,其评价指标维度的设计重点关注科研成果的数量,如论文和专利数量;关注成果的质量,如高影响期刊发表、高被引论文;科研成果的实际应用和转化,如技术许可和商业化成功案例。

二是评价维度更注重标准化、规范化与连续性。科睿唯安引文桂冠奖^[17]通过统计过去30年里4个学科领域内各位高被引作者的总引用量、单篇引用量、高被引文章总数,以及领域内论文的篇均被引次数,或每位作者的篇均被引次数等多个指标,以衡量这些作者是否是该领域的开创者、是否曾经因该项工作多次获奖等,最终对引文桂冠奖获得者做出判定。

三是新兴评价维度基于国家需求和政策导向,更加重视创新能力、贡献、科学家精神等层面。例如,井润田^[18]采用多案例研究方法归纳总结出科技人才能力特征的三个较为特殊的维度,包括学术热情和学科敏感性、战略规划

及其实现路径、团队氛围构建。冯燊等^[19]基于社会背景、教育背景、工作经历、科技贡献、科学精神等五个维度归纳科技人才的规律性特征,其中科技贡献引入了对科研影响力的更广泛评估,如社会影响、政策制定、科学传播等,科学精神则集中体现科学的理想、信念、态度、思维和视野。

国内外学者构建了全面多元、标准化规范化、与国家需求和政策导向紧密结合的科技人才综合学术水平评价维度,为人才识别和评价研究提供了更为全面和深入的认知。本研究基于前述工作,补充战略性科技人才特征指标,分析不同学科特性,评估指标的表征性和可获取性,结合定量与定性指标,提出科技人才综合学术水平评价维度。

综上所述,同行评议方法具有依赖主观判断的问题,基于科学计量学的人才评价方法过于依赖科学成果产出维度的指标。同时,在科研数据呈爆炸式增长的情况下,简单的赋权重定量评价方法在辨识复杂模式、泛化能力以及处理高维数据方面具有很大局限性。因此,科技人才的遴选评价体系需要融合更多的数据挖掘和机器学习方法,以提升评价的准确性和适应性。

2 综合学术水平评价指标与模型构建

2.1 评价指标与模型构建流程

本研究运用机器学习方法进行科技人才评价的模型构建,具体流程如图1所示。

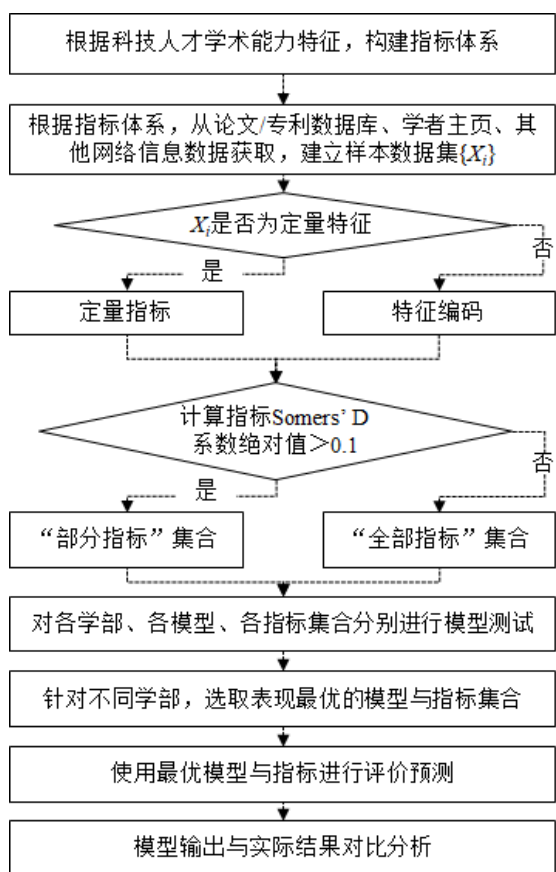


图1 科技人才评价模型构建流程

第一步，评价指标体系构建。科技人才评价与识别指标体系设计的原则一方面要着重考虑指标的有效性，另一方面也要兼顾指标的可获得性，确保指标体系的建立简单易行。在指标的有效性方面，根据科技人才学术能力的特征研究构建指标体系；在指标的可获得性方面，考虑从论文/专利数据库、学者主页、其他公开网络渠道获取数据。

第二步，数据集构建和数据预处理。首先根据指标体系从多来源公开渠道收集并清洗、整合数据。在数据预处理阶段，根据数据类型区别，针对分类变量，采用独热编码方式，将每个分类变量转换为独立二进制特征；针对数值变量，使用 Somers’D 系数衡量各项自变量 X_i

与因变量 Y 之间的关联程度，筛选出系数值介于 -1 和 1 之间的变量以构建“部分指标”集合，并与“完全指标”集合共同构成形成最终科技人才识别预测模型的输入数据。

第三步，构建机器学习模型并进行模型评价。考虑不同类型的人才指标集的特征差异，对各学部、各机器学习模型、各指标集合分别进行模型测试，选取表现最优的模型与指标集合进行后续训练和调优。预测完成后，通过对比模型输出与实际结果进行模型评价。

2.2 综合学术水平评价维度

推动人才评价改革落地的重要目标就是着力克服“唯论文、唯职称、唯学历、唯奖项”倾向，本研究基于科技人才具有较高学术能力的特征，对其综合能力设计多维度的多元评价指标体系，破除“四唯”带来的片面性。

除科研人员具备的综合学术能力外，科技人才还要具备科技创新能力强、科研任务组织领导能力强等特征，其职责和目标与普通科研人员相比有更高层面的要求，需要根据科技人才的特征来设计其评价维度。本研究选择学术诚信、科研产出、科研活动、学术声誉 4 个维度来衡量，如表 1 所示。

表1 科技人才综合学术水平评价维度

评价维度	说明
学术诚信	道德诚信，含成果真实性和履历真实性
科研产出	包括学者各类科研产出数量、质量、影响力等
科研活动	包括学者参与各类科研项目，参加学术会议情况
学术声誉	学术界对于学者的认可，包含科研奖励、学术荣誉等

学术诚信主要从人才的履历真实性和产出成果真实性两个方面考察。履历真实性考察教育背景和工作经历,从学历、工作任职等个人基本信息判断是否有造假和瞒报;产出成果真实性从学术道德上进行判断,如论文存在抄袭、造假而被撤稿,专利有侵权而被追诉,专著或者标准等任何一项产出成果存在的抄袭、造假等问题。

科研产出直观反映人才的研究成果影响力,体现了人才在特定时间内的学术贡献。该维度可以具体量化人才的研究能力和创新程度,同时反映了其研究的深度和广度。常见指标包括在高质量期刊上发表的论文数量与质量、独立研究成果、学术专著的出版、专利的申请和授权情况等。

科研活动展示了人才参与学术交流和合作研究的情况,反映了其在学术界的活跃度和影

响力,它不仅体现了人才的学术实践能力,还显示了其领导与合作的能力。常见指标包括参与国家级或国际合作的科研项目、与其他领域专家和知名学者的学术合作,以及组织学术会议和进行学术演讲等。

学术声誉是学术界和社会上对人才学术成就及其贡献高度认可的体现,体现了学术界对个体学者整体贡献的认可和尊重。科技人才获得的科研奖励、学术荣誉等反映了学者在学术界的声望。

2.3 综合学术水平评价指标体系

基于学术诚信、科研产出、科研活动、学术声誉4个维度来衡量的科技人才评价指标体系,将定性维度细化为可测度的指标,具体如表2所示。

表2 综合能力定量指标列表

维度	指标名称	说明
学术诚信	成果真实性、履历真实性	学术诚信(2项),若任何一项发现有问題,可“一票否决”
	论文总数、核心期刊论文数、第一作者论文数、总被引频次、篇均被引频次、单篇最高被引频次、被引论文章量、被引率、h指数	中文论文(9项)
科研产出	论文总数、通讯作者论文数、第一作者论文数、高被引论文数、总被引频次、篇均被引频次、单篇最高被引频次、被引论文章量、被引率、h指数	英文论文(10项)
	论文总数、第一作者论文数、通讯作者论文数	SCI论文(15项),分别统计Q1、Q2、Q3、Q4分区情况及SCI总体情况
	论文总数、第一作者论文数、总被引频次、单篇最高被引频次、被引论文章量、h指数	CPCI论文(6项)
	论文总数、第一作者论文数、通讯作者论文数	EI论文(3项)
科研活动	专利总量、第一发明人专利数、发明专利数量、实用新型专利数量、外观设计专利数量、专利有效量、质押数、许可数、转让数、实质审查数、公开数、授权数、权利终止数、撤回数、驳回数、放弃数、专利合享价值度、被引证次数、h指数	专利(19项)
	863计划、973计划、自然科学基金、国家重点研发计划、国际合作项目、科技成果转化计划、科技基础性专项工作、科研院所研究开发专项、星火计划、火炬计划、支撑计划、重大仪器专项、重大专项、富民强县、公益性行业科研专项计划、创新方法专项、项目总计	项目(17项)仅统计项目负责人
	国家自然科学奖一等奖、国家自然科学奖二等奖、国家技术发明奖一等奖、国家技术发明奖二等奖、国家科学技术进步奖创新团队、国家科学技术进步奖特等奖、国家科学技术进步奖一等奖、国家科学技术进步奖二等奖、奖项总计	奖励(9项)统计全部获奖人
学术声誉	杰出青年、长江学者	荣誉称号(2项)

本研究还补充了定性指标 8 项，补充定性指标主要从人才个人特征和背景信息，包括性别、年龄、所在机构等作为评价的背景变量指标，定性指标经特征编码后可以形成虚拟变量。因本研究对象为工程院院士候选人，提名渠道也反映肯定其科研能力和贡献度的重要指标，如表 3 所示。

表 3 补充定性指标

维度	指标名称	说明
基本信息	性别	男性 =1, 女性 =0
	年龄	
	机构类型是否为军工单位	是 =1, 否 =0
	机构类型是否为高校	是 =1, 否 =0
	机构类型是否为科研院所	是 =1, 否 =0
	机构类型是否为企业	是 =1, 否 =0
	提名渠道是否为院士提名	是 =1, 否 =0
提名渠道是否为科协提名	是 =1, 否 =0	

以上两个表中，反映学术诚信、科研产出、科研活动和学术声誉的可量化评价指标共计 92 项，学术诚信的 2 个指标中，若任何一项发现有造假或者瞒报等问题，可“一票否决”，不参与学术综合水平的定量计算，参与模型计算的定量指标与描述性指标共计 98 项。

3 指标分析与模型结果验证

3.1 研究样本与数据来源

本研究以中国工程院 9 个学部，2019 年 531 名院士候选人为研究样本。对 531 名院士候选人的综合学术水平进行评价，以是否入选作为命中率判定模型效果。

中文论文数据来源于北京万方数据股份

有限公司的万方数据；英文论文数据来源于科睿唯安 Web of Science 核心合集（含 SCI 及 CPCI）、爱思唯尔 EI 数据库；专利数据来源于 IncoPat 数据库；基金项目、科技奖励、荣誉称号等信息均来源于互联网公开信息；基本信息包括性别、年龄、所在机构、提名渠道等，均来源于互联网公开信息。

3.2 特征指标分析

定量指标与描述性指标等各项特征指标共计 98 项，记为自变量 X_i ；因变量 Y 为候选人实际通过第二轮评选的情况。为避免过拟合，需去除冗余特征，对 98 项指标 (X_i) 进行降维。由于工程院的 9 学部的学科特点不同，各项学术成果产出或学术活动对于能否当选院士的影响程度也不同，故对于每个学部分别进行特征选择。

使用 Somers' D 系数计算各项特征 X_i 对因变量 Y 的关联程度，Somers' D 的绝对值越大，表明该 X_i 与 Y 之间的序数关联越强，意味着该特征对因变量 Y 具有较大的影响。针对 9 个学部，分别计算并筛选 Somers' D 绝对值大于 0.1 的 X_i ，得到指标数量在 24 项至 67 项之间，形成“部分指标”集合。每个学部选择指标具体情况如表 4 所示。

由表 4 可以看出，不同学部中，各项特征 X_i 与因变量 Y 的关联程度差异较大。除了工程管理学部和信息与电子工程学部外，提名方式对于能否当选院士均存在较大关联；专利类指标在化工、冶金与材料工程学部，机械与运载工程学部，信息与电子工程学部关联性较为明显；在医药卫生学部，中文论文类指标存在显著的负面影响。

表4 不同学部指标选择情况

学部	筛选指标数量	筛选指标平均 Somers' D 绝对值	指标关联度 TOP5
工程管理学部	39	0.17	Q1 论文总数 (0.3644***) Q1 通讯论文数 (0.3401***) 英文通讯论文数 (0.3178***) SCI 论文总数 (0.3158***) Q4 论文总数 (0.3097***)
化工、冶金与材料工程学部	46	0.19	院士提名 (0.4390***) 科协提名 (-0.3659***) 专利权利终止数 (0.3608***) 实用新型专利数 (0.3100***) 第一发明人专利数 (0.2998***)
环境与轻纺工程学部	49	0.20	奖项总计 (0.4023***) 院士提名 (0.3530***) 科协提名 (-0.3530***) Q3 一作论文数 (0.3465***) Q3 论文总数 (0.3350***)
机械与运载工程学部	67	0.25	院士提名 (0.4366***) 科协提名 (-0.4366***) 第一发明人专利数 (0.3918***) SCI 一作论文数 (0.3860***) 专利有效量 (0.3782***)
能源与矿业工程学部	51	0.24	院士提名 (0.5122***) 科协提名 (-0.5122***) SCI 通讯论文数 (0.4129***) SCI 论文总数 (0.4077***) EI 论文总数 (0.3782***)
农业学部	57	0.26	院士提名 (0.5153***) 科协提名 (-0.4565***) Q2 论文总数 (0.4437***) Q2 通讯论文数 (0.4309***) 奖项总计 (0.4207***)
土木、水利与建筑工程学部	60	0.23	项目总计 (0.4335***) CPCI 论文总计 (0.3780***) 奖项总计 (0.3687***) 自然科学基金项目数 (0.3651***) 院士提名 (0.3370***)
信息与电子工程学部	24	0.16	实用新型专利数 (-0.2831***) 国家科学技术进步奖二等奖 (0.2308***) SCI 一作论文数 (0.2099***) CPCI 论文总计 (0.1995***) 项目总计 (0.1864***)
医药卫生学部	26	0.22	中文论文篇均被引频次 (-0.3937***) 中文论文 h 指数 (-3805***) 院士提名 (0.3714***) 中文论文总被引频次 (-0.3714***) 中文论文单篇最高被引频次 (-0.3490***)

3.3 模型选择与训练

(1) 模型选择

本研究运用机器学习方法进行模型构建，机器学习算法能够有效处理具有大量特征的高维数据，擅长揭示非线性关系，在多学科整合应用中具有较强灵活性和泛化能力。

变量 Y 代表候选人实际是否通过第二轮评选，取值只有“1”（通过）或“0”（未通过），故本模型处理的是一项二分类问题，根据数据分布情况选取合适的机器学习模型进行训练。鉴于不同机器学习模型对数据集的适应性各有差异，本研究中的样本数据集的特征概括如下：①数据集规模较小（样本数量 531 项），无需考虑计算资源因素；②数据集维度较高（全部指标共计 98 个），需重点考虑模型复杂度和过拟合问题；③数据缺失值较少，各项指标数据

均通过公开渠道收集获得，无需考虑机器学习模型对缺失值的敏感度；④指标特征既包括分类变量，又包括数值变量，需要模型支持混合型的数据输入。

为了找到最佳拟合本研究数据集的模型，选择多种类型的机器学习方法进行测试，包括：传统统计模型，如逻辑回归（LR）；基于边界的模型，如支持向量机（SVM）；基于树的方法，如随机森林（Random Forest）、梯度提升树（GBDT）、极端梯度提升（XGBoost）和轻型梯度提升（lightGBM）；神经网络方法，如多层感知机（MLP）；基于实例的方法，如 K 近邻（KNN）；贝叶斯方法，如朴素贝叶斯（Naive Bayes）。

在深入分析数据分布特性的基础上，遴选与数据内在结构相匹配的模型，以确保学习过程的最优性能，具体模型选择说明如表 5 所示。

表 5 模型选择说明

序号	模型名称	选择原因
1	逻辑回归（LR）	最常见的二分类模型；计算效率高；易于解释和理解
2	支持向量机（SVM）	对于数据集的特征维度和样本量都有较好的适应性；在复杂非线性问题上表现良好
3	多层感知机（MLP）	常见深度学习模型；具有很强的表达能力和学习能力，适用于各种类型的数据集
4	朴素贝叶斯（Naive Bayes）	计算速度快；对小规模数据集表现良好
5	K 近邻算法（KNN）	适用于本项目小样本量数据集
6	随机森林（Random Forest）	能够处理高维数据和大规模数据集；对于特征的重要性排序具有较好的解释能力
7	梯度提升树（GBDT）	具有较强的拟合能力和预测精度
8	极端梯度提升（XGBoost）	在零膨胀数据集中表现出色
9	轻型梯度提升（lightGBM）	适用于高维特征数据集

(2) 模型训练

将每个学部的部分指标和全部指标作为特征参数输入模型，通过模型训练找到 $Y=f(X_i)$

的隐含关系，从而根据 X_i 值计算出候选人当选院士的概率，可视为院士候选人评价的综合得分。对 9 个学部分别构建模型，每个学部均采

用9种机器学习模型评价预测,每个模型分别构建 $9 \times 9 \times 2 = 162$ 个机器学习模型,预测准确使用全部指标和部分指标作为数据输入,即共率如表6所示。

表6 模型预测准确率(单位:%)

各模型 预测准确率		LR	SVM	MLP	Naive Bayes	KNN	Random Forest	GBDT	XGBoost	Light GBM	综合 预测 准确率
工程管理学 部	部分指标	60.6	59.2	57.6	61.8	55.7	58.2	54.4	57.7	57.8	58.1
	全部指标	55.8	53.5	52.4	44.3	50.6	54.9	52.5	49.5	57.8	52.4
化工、冶金 与材料工程 学部	部分指标	66.8	59.4	60.2	66.8	64.0	61.0	56.0	58.1	60.8	61.5
	全部指标	61.3	60.7	58.2	64.8	63.5	60.3	57.9	58.4	59.9	60.6
环境与轻纺 工程学部	部分指标	66.6	64.7	58.2	57.2	60.5	57.7	53.2	54.9	54.8	58.6
	全部指标	59.3	59.9	52.1	40.4	49.4	58.2	50.1	49.8	53.3	52.5
机械与运载 工程学部	部分指标	66.8	56.5	58.4	65.2	56.3	61.1	62.3	59.8	60.7	60.8
	全部指标	66.0	56.0	57.9	55.8	58.3	60.7	60.5	58.7	58.7	59.2
能源与矿业 工程学部	部分指标	67.9	73.0	65.7	69.6	70.1	70.9	68.6	68.8	65.1	68.9
	全部指标	62.9	66.0	63.4	63.9	63.9	66.4	66.6	68.3	62.3	64.9
农业学部	部分指标	71.2	66.0	65.2	65.4	62.9	61.0	63.3	63.1	58.8	64.1
	全部指标	68.2	62.7	62.9	57.6	62.3	61.1	65.6	63.9	58.4	62.5
土木、水利 与建筑工程 学部	部分指标	57.1	62.5	58.8	61.4	56.0	61.1	63.3	60.8	62.3	60.4
	全部指标	57.7	63.2	59.5	60.5	60.8	62.1	63.6	62.2	61.9	61.3
信息与电子 工程学部	部分指标	56.8	59.8	53.5	57.0	58.4	56.7	56.1	54.7	58.8	56.9
	全部指标	52.8	55.7	54.7	60.0	52.6	52.0	49.2	47.3	58.4	53.6
医药卫生学 部	部分指标	70.1	66.9	65.5	68.9	64.4	61.2	63.1	63.1	64.9	65.3
	全部指标	63.5	61.1	57.5	61.1	52.9	58.3	56.9	56.2	61.8	58.8

针对每个学部,根据测试结果选取表现最好的模型及指标集合用于后续模型预测,具体选择模型以及指标情况如表7所示。

表7 各学部模型及参数选择

学部名称	选择预测模型	选择指标数量	模型准确率
工程管理学部	Naive Bayes	39	61.80%
化工、冶金与材料工程学部	LR	46	66.80%
环境与轻纺工程学部	LR	49	66.60%
机械与运载工程学部	LR	67	66.80%
能源与矿业工程学部	SVM	51	73.00%
农业学部	LR	57	71.20%
土木、水利与建筑工程学部	GBDT	97	63.60%
信息与电子工程学部	Naive Bayes	97	60.00%
医药卫生学部	LR	26	70.10%

3.4 模型评价效果分析

本研究借鉴推荐系统思想，使用 hits 指标评价模型效果，hits@n 即模型评价综合分值的前 n 个候选人在不同年份增选时累计命中的数量。针对样本群体，使用表 7 中的模型及参数，

分别对不同学部候选人通过评审的概率进行计算，去除实际未进入第二轮评选的人员后，输出每个学部入选概率最高的人员名单。将模型输出与 2019、2021、2023 年工程院院士增选实际结果对比，如表 8 所示。

表 8 模型评价结果在不同年份的累计命中情况

hits@n	年份	工程管理学部	化工、冶金与材料工程学部	环境与轻纺工程学部	机械与运载工程学部	能源与矿业工程学部	农业学部	土木、水利与建筑工程学部	信息与电子工程学部	医药卫生学部
hits@3	2019	2	1	0	1	0	1	2	1	0
	2021	2	2	1	2	1	1	3	1	2
	2023	2	3	1	3	2	1	3	2	2
hits@5	2019	2	2	0	1	2	2	4	1	0
	2021	2	4	1	3	3	2	5	1	4
	2023	2	5	1	4	4	2	5	2	4
hits@10	2019	3	4	1	4	3	4	5	4	3
	2021	4	6	3	8	5	5	7	4	7
	2023	4	7	4	9	6	6	9	5	8
hits@15	2019	4	5	4	5	5	5	7	6	6
	2021	7	8	6	11	7	6	10	9	11
	2023	7	11	8	12	8	8	13	10	12

通过对比模型输出与 2019、2021、2023 年院士增选的实际情况，结果如下：①模型在所有学部中评价结果较为准确。在全部学部中 hits@3 指标均大于 1，表示模型在各学部推荐的前 3 名候选人，截至 2023 年均有至少 1 人增选成为工程院院士。②不同学部的模型命中率存在差异，整体命中率为 46.67%~86.67%。其中，表现最好学部为土木、水利与建筑工程学部，模型评价并推荐的 15 位学者，有 13 人在未来 3 届内成功当选为工程院院士。③三年新增选结果对比发现，模型的评价结果与实际增选结果的一致性逐年提升，说明模型在院士增选评价预测场景下具有一定稳定性和持续性。

4 研究结论与展望

本研究基于科技人才的定量和定性指标数据，以中国工程院院士候选人作为研究对象，采用机器学习方法，训练并构建形成适用于不同学科领域的人才评价模型，进行实证研究。

研究表明，模型在中国工程院院士评价预测场景中具有较好适用性，在土木、水利与建筑工程领域，模型输出的 15 人名单中，高达 86.67% 的学者在未来 3 届内成功当选为工程院院士。机器学习模型输出的评价结果基于客观数据，不涉及专家打分或指标赋权，具有客观性；模型能够适应不同学科领域的特点，快

速地对大量数据进行分析和处理,为科技人才评价工作提供了客观、准确、高效的方法。总体而言,本模型在识别出具有院士科研能力和潜质的高层次科技人才上具有很高的适用性,能够在较大范围内快速协助识别潜在人才,或为同行评议提供参考和支持,在人才规划、人才储备、人才引进等工作中具有现实意义。此外,本研究所采用的机器学习方法具备较强的泛化能力,通过挖掘样本特征与评价结果之间的隐含关系,能够灵活应用于不同领域和层次的人才评价场景,尽管目前的研究样本主要集中在工程院院士及候选人,但通过相同的方法和流程,可以将这些模型和指标推广应用于更广泛科技人才评价中。

本研究尚存在一定的局限性和不足之处,一方面,限于时间和数据采集成本等因素,模型的训练数据时间跨度较小,若能补充更长时间范围的数据进行模型训练,则能够识别到更多历史趋势和变化特征;另一方面,模型选择的研究对象为院士候选人,输出结果仅在高层次人才层面展现出较好的适用性,为确保其在更广泛的人才群体中保持高效性和准确性,对于某些特定领域、不同类型的人才,需要补充采集相应数据进行训练和调优。

为提升当前模型在不同学科领域中的科技人才评价准确率,未来计划在以下两方面进一步对模型进行优化和扩展。一是增加更多反映科技人才科研产出和学术活动的指标,使评价的维度更加全面、多元,提升模型成功拟合实际影响因素的概率;二是扩充样本数据量,增强模型的训练基础,从而提高其评价结果的精度和稳定性。此外,本研究所采纳的指标体系

与模型方法具备良好的扩展性,能够适用于不同的数据集进行分析,未来可以扩大至其他科技奖励或荣誉的潜在人才识别应用。

参考文献

- [1] 刘璇,段宇锋,朱庆华.基于合著网络的学术人才评价方法研究[J].情报杂志,2014,33(12):77-82.
- [2] 宋培彦,冯超慧,龙晨翔,等.基于颠覆性指数优化的细分领域优秀科技人才发现研究[J].情报杂志,2022,41(5):61-65.
- [3] 常鏐鏐,高继平,师丽娟.调和H指数视角下的技术创新人才识别方法与实证研究[J].中国科技论坛,2021(11):104-112.
- [4] 孙峰.“双一流”建设背景下我国高校教学科研人才多维度评价体系构建探析:以华南理工大学为例[J].科技管理研究,2022,42(14):79-84.
- [5] 王喆,杨国栋.高校博士后科技人才培养绩效评价指标体系构建研究——基于平衡计分卡方法[J].科技管理研究,2021,41(15):127-133.
- [6] YE Y, ZHU H, XU T, et al. Identifying high potential talent: A neural network based dynamic social profiling approach[C]//2019 IEEE International Conference on Data Mining. Beijing: IEEE, 2019: 718-727.
- [7] YE J, HU H, CHAI C. A talent classification method based on SVM[C]//2009 International Symposium on Intelligent Ubiquitous Computing and Education. Chengdu: IEEE, 2009: 160-163.
- [8] JANTAN H, HAMDAN A R, OTHMAN Z A. Human talent prediction in HRM using C4.5 classification algorithm[J]. International Journal on Computer Science and Engineering, 2010, 2(8): 2526-2534.
- [9] SUN Y, ZHUANG F Z, ZHU H S, et al. The impact of person-organization fit on talent management: a structure-aware convolutional neural network approach[C]// In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York: Association for

- Computing Machinery, 1625-1633.
- [10] QIN C, ZHU H, XU T, et al. An enhanced neural network approach to person-job fit in talent recruitment[J]. ACM Transactions on Information Systems, 2020, 38(2): 1-33.
- [11] NIJS S, DRIES N, VAN V, et al. Reframing talent identification as a status-organising process: Examining talent hierarchies through data mining[J]. Human Resource Management Journal, 2021, 32(1): 169-193.
- [12] 田军, 刘阳, 周琨, 等. 陕西省科技人才评价指标体系与评价方法构建 [J]. 科技管理研究, 2022, 42(4): 89-96.
- [13] 丁宁, 苏颖, 胡豫, 等. “破四唯”背景下我国高校附属公立医院人才评价体系分析 [J]. 中华医院管理杂志, 2021, 37(12): 953-957.
- [14] 张熠, 倪集慧. 科技创新人才评价指标体系构建 [J]. 统计与决策, 2022(16): 172-175.
- [15] 余波, 陈仕吉, 赵嘉懿. 中国高校科技人才评价的影响因素及指标体系构建研究 [J]. 农业图书情报学报, 2023, 35(7): 63-74.
- [16] 王运红, 潘云涛, 赵筱媛. 基于科技信息大数据的科技人才科研综合能力评价及应用研究 [J]. 医学信息学杂志, 2017(38): 7-14.
- [17] DAVID PENDLEBURY. Researchers of Nobel class: Citation Laureates 2023 [EB/OL]. (2023-09-19) [2024-01-28]. <https://clarivate.com/blog/researchers-of-nobel-class-citation-laureates-2023/>.
- [18] 井润田. 高校科研团队管理与战略科学家能力建设 [J]. 上海交通大学学报 (哲学社会科学版), 2022, 30(4): 43-56.
- [19] 冯黎, 童杨, 闫金定. 关于培养使用科技人才的思考——基于中外 100 位科技人才的履历分析 [J]. 科技导报, 2022, 40(16): 38-45.