

数据挖掘技术在科研项目全流程质量统计评价中的应用研究

张增虎

广西民族师范学院, 广西 崇左 532200

DOI:10.61369/EIR.2026070019

摘 要： 随着当前科研工作越来越复杂化并且产生的数据越来越多，在这种大环境下传统的科研项目的质量评价方式已经无法做到准确高效的完成任务。而作为一项强有力的数据分析手段——数据挖掘技术给整个科研工作的各个阶段的质量统计评价带来了一线生机，本篇文章正是为研究这一项新技术如何应用于其中，并利用来实现对科研工作中的一些重要环节的把握如及时掌握科研项目的发展情况及变化状况等，同时还能对未来可能存在的问题做出相应的预判并对其结果作出评估从而达到合理分配资源的目的并使得相关制度更加完备。

关 键 词： 数据挖掘技术；科研项目；全流程；质量统计评价；应用

Research on the Application of Data Mining Technology in Quality Statistical Evaluation throughout the Entire Process of Scientific Research Projects

Zhang Zenghu

Guangxi Minzu Normal University, Chongzuo, Guangxi 532200

Abstract： With the increasing complexity of current scientific research and the growing volume of data generated, traditional quality evaluation methods for scientific research projects are no longer able to accurately and efficiently complete tasks in this environment. As a powerful data analysis tool, data mining technology brings a ray of hope to the quality statistical evaluation at various stages of scientific research. This article aims to explore how this new technology can be applied to grasp important aspects of scientific research, such as timely understanding of the development and changes in scientific research projects. It also enables the prediction of potential future issues and evaluation of their outcomes, thereby achieving the goal of rational resource allocation and enhancing the completeness of relevant systems.

Keywords： data mining technology; scientific research projects; entire process; quality statistical evaluation; application

引言

科研项目质量评价作为科研管理工作的重要一环，在很大程度上决定科研资源是否合理利用及科研成果能否有效转化等重要问题，并且也直接影响着整个科研工作的整体发展情况。传统的评价方式主要是通过专家评议或采用文献计量的方法来实现，虽然可以大体反映出一个项目的价值所在，但也有许多不足之处。开展数据挖掘技术在科研项目全流程质量统计评价中的应用研究，希望实现科研项目全流程质量统计评价水平的进一步提升。

一、数据挖掘技术在科研项目全流程质量统计评价中的应用目标

（一）精准识别科研项目关键要素

对于科研项目的质量管理来说，在开展相关工作时，需要对

其所涉及的关键性因素进行全面了解与掌握，并以此为基础来实现其质量的有效评估。而借助于数据挖掘技术，则可以基于其中的相关方法（例如关联规则分析以及分类算法）将大量有关科研的数据信息转化为一些重要的元素，比如相关的研究方向、人员组成或资金安排等等。

基金项目： 本文系广西民族师范学院2024年度校级科研项目《科研项目数据库数智化建设》（项目编号：2024FW013）研究成果。

个人简介： 张增虎（1978.09-），男，壮族，广西天等人；大学本科学历，专业：英语；研究方向：翻译理论与实践，数智库，语料库；职称：副教授；广西民族师范学院就职，主讲大学英语课程；主持各级各类研究项目10项，其中省部级研究项目1项，厅级研究项目3项，校级研究项目6项；出版学术著作1部，主编学术著作1部，主编教材1部；公开发表学术论文26篇；获得计算机软件著作权1项。

例如，对于高校中的科研项目来说，在运用 AprioriTid 算法来处理省部级以及自然学科类型的项目时，会发现在总数量中有 1/3 是花费 ¥10000 至 ¥40000 之间，并且从整体来看科研结果往往会被专家们评为优秀等级。除此之外，通过数据分析还可以将过去的信息相结合，进而找出各方面的关联性，从而为项目策划及资源安排提供一定的参考意见。而经过一系列重要因素的提炼并加以分析之后，则使得数据挖掘技术为基础科研项目的质量评定工作打下良好的基础。

（二）实时监测项目进展与质量波动

在科研项目的整个过程中实施动态监督也是保证其顺利完成的一个重要方面，在传统的评价方法中一般通过阶段性的汇报或者人去查看的方式进行评价，并不能够做到随时对其进行观察了解。但是利用数据挖掘的方法可以不断的对其过程的数据信息进行分析研究，以便于及时发现问题存在的问题以及错误之处，进而做出相应的评价。

例如，采用大数据技术建立起来的企业科研项目管理信息系统可以随时了解每个年度内研究人员所开展的研究工作的情况，并且将这些实际情况与预定的目标进行比较，发现其中存在的问题。另外在大学中使用的科研系统中存在的计算机数据分析功能也可以对大量的数据信息进行深入的处理，使管理人员能够在第一时间就发现问题并解决问题以提高工作效率。这样的方式使得对于整个工作的监控更加细致化了并且有利于保证科研的质量。

（三）预测项目潜在风险与成果价值

在科研项目的实施过程中，会遇到很多不确定的因素，这就需要我们对其可能出现的问题以及取得的结果做出相应的估计。而通过对这些方面作出相应判断的过程就是数据分析的一个重要组成部分——利用数据挖掘技术建立一个预测模型，根据已有的数据、信息对未来出现的情况做一个预估。如在教改课题及科研项目管理中运用了数据挖掘技术来分析以前的大数据，从中找出项目的研究情况中存在的规律并对其进行预测，从而得出一些问题供决策人员参考等。

此外，通过分析项目的产出结果来判断一个项目是否具有很高的科学价值或者很大的发展前景是利用数据挖掘技术可以做到的事情。比如采用大数据的技术来进行科研方面的评价就可以发现一些隐藏在深处以及比较滞后的科学研究上的巨大价值，这样就能够使得对于项目的研究成果做出更加准确地评价。这样的预测功能不仅可以减少出现项目研究不成功的情况的发生，而且也能够一定程度上给相关负责人进行资源的分配与设置优先权方面带来一定的指导意义。

（四）优化科研资源配置

项目的成功开展离不开良好的科研环境，在这个过程中需要投入大量的资金、人力等资源，并不是每个项目都能够得到充足的经费支持与人员配备，而通过对项目的数据进行研究可以解决这一问题提供参考意见。例如，对于高校而言，在科研管理方面运用到数据挖掘技术时，可以根据以往的表现情况以及自身实力还有研究的方向来计算出所需要的资源数量，然后将这些资源合理的分发出去。除此之外，通过对国有企业的相关研究发现，利

用大数据的技术手段来进行数据分析之后，管理人员就可以更好的了解所使用的资源是否有效率，进而根据此做出相应的改变与调整，使整个过程中的资源得到最大程度上的利用。这样一来，就使得整体的科研工作得到了有效的安排，同时也可以保证一些质量较高的科研工作的顺利开展。

（五）完善科研项目评价体系

评价体系是否合理、全面将影响到其最终的结果是否公平、真实可靠。而传统的方式一般只凭个人感觉或者一个标准来评定，很难做到完全客观。利用数据挖掘的方法从多个方面考虑问题并加以权衡可以大大提高评价体系的合理性与完整性。

例如，在高校科研评价过程中运用大数据来收集一些不规则或者半规则的数据信息，并且以此为基础建立一个比较完整的评价标准，包括关于专利以及科研的方向等等各个方面。除此之外，在研究如何找到科技人才评价机制时还发现，可以采用数据挖掘的方法从不同角度、不同等级上对科技人才进行评价，这样就可以给项目评价带来更多的参考意见。通过对各项指标以及权重不断的修改和完善，使得数据挖掘能够更好地帮助人们建立起更完善、更科学的科研项目的评价方式，也促进了整个科研水平得到更大的提高。

二、数据挖掘技术在科研项目全流程质量统计评价中的应用方法

（一）关联规则挖掘分析项目要素关联

关联规则挖掘是有效的一种数据分析方法，可以从大量的数据中找到一些隐藏的关系和规律。在科研项目的整个流程的质量统计评价过程中应用了这一手段来找出各个因素之间存在的关系。如利用 Apriori Tid 算法来分析高校科研项目的数量化数据时发现了研究领域的大小以及经费多少同研究成果的数量存在一定的相关性等。

具体而言，对于省部级或自然科类、经费金额介于 ¥10000 至 ¥40000 之间的科研项目来说，其获得较高成果评价价值的可能性较大，此类项目数量约占总项目的 1/3，在整体上大部分被评定为优秀的水平。除此之外，在教改课题及科研项目管理方面，通过运用关联规则挖掘的方法可以发现项目申报要求与其所涉及的研究内容之间存在的相应性问题，从而为改善项目的设计方案提出建议意见。由此可见，利用关联规则挖掘不仅可以找出科研项目中的各个元素间可能存在的相互作用，还可以为其做出的质量评估提供有力的支持。

（二）分类算法评估项目质量等级

分类算法是一种重要的数据分析方法，在数据挖掘过程中主要用来把数据按一定的标准划分到相应的类别的过程。在科研项目的整个流程的质量统计评价工作中，运用了这种分类算法来实现对于科研项目的迅速分级，以达到提升工作效率的目的。如通过采用决策树、支持向量机等分类算法，可将科研项目区分为“优秀”，“良好”及“一般”的质量水平等等。通过以上步骤可以以以往的数据作为基础来进行培训，并且形成一个分类模型，

在之后就可以根据这个模型来实现对于新项目的自动评估。同时也可以借助于分类算法与一些前段图形化的显示方法相结合的方式给管理人员呈现出有的项目质量的情况,这样就能够使他们更加清楚地了解到整个项目的总体情况。在实际应用中发现采用这种分类算法能够提高准确度并且大大减少工作人员所付出的努力程度,这对于科研项目管理来说是十分有利的。

(三) 聚类分析识别项目群体特征

聚类分析是通过寻找数据集中相同或相近的模式来实现的一种无监督的学习过程。其目的是把有某些共性的对象划分到同一个组里去,以发现其中所蕴含的一些规律性的东西。在科研项目的全过程质量管理统计评价过程中经常用它来进行不同群体的研究成果的质量特性及趋势分析。比如,在研究若干个方面的项目情况(包括研究的方向、人员组成以及资金多少等等)之后,利用聚类的方法把这些性质类似的情况放在一个大组里面,然后找出每个小组的特点并对其质量情况进行评价。通过该方法不仅可以找出不同项目的区别所在,并且可以作为合理的安排科研资源的一个依据,在国有企业的科研项目管理上尤其如此,通过对这些企业所拥有的科研项目进行分类之后就可以找到一些比较有前途的研究项目,然后将更多的研究资金投入其中去保证其发展与完善。所以,运用聚类法对于科研项目来说是10分重要的。

(四) 时序模式挖掘跟踪项目进展规律

时序模式挖掘是针对时间序列数据的一种数据分析方法,在这种数据挖掘方式下可以从过去的事件中找出随着时间推移而形成的有规律性的模式。对于科研项目的整个流程的质量统计数据来进行评价的时候,通过这种方式来追踪项目的发展情况,并且能够从中得到一些有用的结论供我们参考、预测等,在研究的过程中,会根据所获取的时间顺序的资料进行分析,这样就能得出一个比较重要的现象,比如说经费使用的高峰时期是在什么时候或者是一些研究成果产生的阶段有什么样的起伏等等。

此外,利用时序模式挖掘可以将大数据运用到实际中去,在一定程度上能够实现对项目进展情况的跟踪,并且将其与事先设定的目标相比较、分析,以便于发现一些不正常的情况以及相应的解决办法。这不仅可以使项目管理水平得到提升,而且对于以后相关工作的开展也具有一定的借鉴意义。

(五) 多种数据挖掘技术融合应用

在发挥各种数据挖掘方法优势的基础上,以保证科研项目质

量统计评价结果准确可靠为目的,一般要综合应用多种数据挖掘技术来实现。如在研究科研项目的全过程信息中,先用关联规则发现各因素间的关系,再运用分类算法确定其等级,并且最终通过聚类找出各个项目组的特点和发展趋势等。

此外,利用时序模式挖掘来了解项目的进展情况,并将这些结果作为其它方法的一个附加来源。通过结合各种技术的方法不仅可以克服单独使用的缺点而且可以从不同侧面全面地评估一个项目。研究发现采用几种数据挖掘技术相结合的方式,在提高评价质量和做出正确决策等方面具有明显的优势,有力地保证了整个科研项目的质量统计评价工作。

利用数据挖掘技术来开展科研项目的全过程的质量统计评价工作,给科研管理工作提供了一种新的思路和方法,在实际工作中通过对相关联规则、分类法、聚类分析及时序模式挖掘等多种技术方式的应用,达到了对科研项目的关键因素有效把握、质量问题的有效监控、可能存在的隐患提前预警、合理调配资源的目的,并取得了很好的效果。

由于数据挖掘需要处理大量高质量的数据,目前科研中的数据存在较为严重的碎片化以及不标准的现象,在某种程度上阻碍着其有效运用。各个领域的研究项目千差万别,想要建立一个适用范围广并且能应对各种情况下的数据挖掘模型还是一大挑战。另外对于数据的安全性和私密性的考虑也是不可小觑的,尤其是涉及到一些比较重要的信息时更是如此,如何做到既能充分利用又能加以保护是值得思考的一个方面。

三、结束语

展望未来,数据挖掘技术在科研项目全过程的质量统计评价中还有很大的发展空间,在未来的人工智能以及大数据等先进技术的支持下,将会出现更多更先进的数据分析手段被运用到科研项目的管理工作中来,这样会使研究结果更为准确高效地完成。另外,各个领域的专家共同参与也会是促进其发展的一大助力,综合了计算机方面的知识、统计学的知识以及其他专业知识,会使得评价模型更为合理有效。最后,制定相关的法律制度及规范,并且建立相应的管理体系也是保证其更好地运用于科研工作的前提条件之一。

参考文献

- [1] 刘思言. 科研管理中数据挖掘技术的应用之研究 [J]. 科技风, 2018, (30): 108-108.
- [2] 刘在洲. 大数据应用于高校科研评价的价值意蕴与运用构想 [J]. 科技管理研究, 2021, 41(4): 109-116.
- [3] 陈小健; 宋承继. 数据挖掘技术在科研项目量化与评价中的应用 [J]. 电子设计工程, 2015, 23(12): 37-40.
- [4] 肖鹏飞; 谢英; 刘秀兰. 数据挖掘技术在教改科研项目中的应用研究 [J]. 互联网周刊, 2023, (17): 68-70.
- [5] 聂迪. 数据挖掘技术在科研管理方面的应用 [J]. 黑龙江科学, 2020, 11(20): 120-121.