

数据挖掘技术在软件工程中的有效应用研讨

杨 阳, 王雅倩

(新疆天山职业技术学院, 新疆乌鲁木齐 830000)

【摘要】本文以软件工程中数据挖掘技术的有效应用为探讨主题,对数据挖掘技术的基本概念进行详细分析,从实现多元化数据信息的高效整合、提高数据传输准确性与缩短处理时间三方面阐述数据挖掘技术在软件工程中的作用,总结数据挖掘在系统结构、软件管理与开发、聚类以及故障检测中的实际运用,以期为软件工程在数据挖掘等人工智能技术助力下的高速发展提供合理建议。

【关键词】数据挖掘技术; 软件工程; 应用

Research on the Effective Application of Data Mining Technology in Software Engineering

【Abstract】This article takes the effective application of data mining technology in software engineering as the topic of discussion, and analyzes the basic concepts of data mining technology in detail, from the realization of efficient integration of diversified data information, improvement of data transmission accuracy and shortening of processing time. Explain the role of data mining technology in software engineering, summarize the practical application of data mining in system structure, software management and development, clustering and fault detection, in order to provide software engineering with the rapid development of artificial intelligence technology such as data mining. Reasonable suggestion.

【Keywords】data mining technology; software engineering; application

【中图分类号】TP311.13;TP311.5

引言:传输与存储信息数据是软件工程的核心业务要求,在开发、优化软件的过程中,技术人员需要以满足多样化的使用需求为基准,将其适用层场景不断丰富,促使用户获得更好的应用体验。数据挖掘是人工智能技术的重要组成部分,其具有高效提取与整合数据信息的能力,在软件工程中对该项技术予以有效运用,不仅有助于软件开发机制的完善化构建,还能够加强软件工程针对性特征。

一、数据挖掘技术概述

要实现数据挖掘在软件工程中的深入渗透与灵活运用,关键在于技术人员在思想理念与思维模式层面上的转变,即需要对数据挖掘技术具有正确认知,抓

住应用技术的要点,以此为基础准确地梳理后续的技术应用流程,能够创新并针对性地改进使用方法。数据库的飞速发展存储各类型数据信息提供了十分便利的技术条件,无论是图像音频,还是视频资料,存储大体量数据的难度已经逐渐降低。为了进一步提高提取数据、传输信息的交互效率,确保用户的多元化需求得到满足,选择将数据挖掘技术应用到软件工程这一有效途径,以达到快速获取数据、实时交互有效数据的目的。与此同时,在智能技术手段的支持下,还可以对相关活动进行预测^[1]。整合信息是数据挖掘的核心技术手段,现阶段其已发展到较为成熟的时期,借助于完整的算法程序,其具有统计与分析大量冗余数据的功能,还能够实现精准的检索与处理,最

后达到高效整合有效数据信息的目的。在处理数据源的过程中，其所属类型不会对数据挖掘技术的应用产生限制，因而在管理信息与查询信息等方面，其能够顺利地开展工作。除此以外，还能够提供科学合理的决策支持与数据预测。对于软件工程来说，在开发新软件、优化软件功能方面，数据挖掘技术带来的支撑力量是十分显著的，从有效性与针对性两方面促进软件开发质量的大幅提升，切实推动了以数据挖掘为核心的软件在交通与金融等行业领域的广泛应用。

借助于对数据挖掘技术的合理运用，优化软件工程计划方案合理性的实践成效也大大提升，在软件工程的设计与规划阶段，涉及的工作环节较为复杂，需要调配技术应用、人力资源与资金等，应用数据挖掘能够构建与之对应的改进机制，对存在的问题与误差等进行纠错与纠偏，以便在第一时间对软件的整个运行过程进行监察，进而快速判定其是否出现漏洞，获取的数据资料作为日后优化处理的参考依据，实现对软件工程的针对性调整与管理。对数据挖掘技术的充分运用带来的是更为便利的作业条件，软件检测的固定周期也大幅缩短，大幅节约了在运行检测方面的投入成本，管控整体软件工程的实践难度显著降低（如图1）。

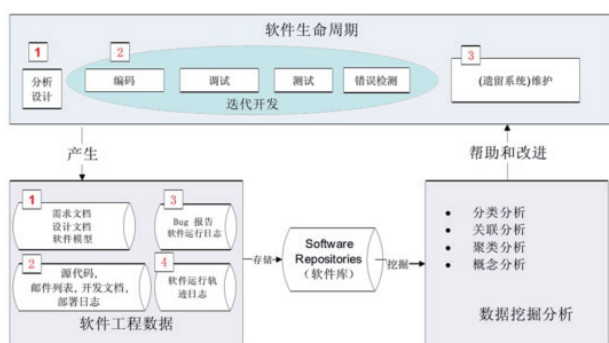


图1：软件工程中数据挖掘技术的应用

二、软件工程中应用数据挖掘技术的重要作用

（一）实现多元化数据信息的高效整合

数据挖掘技术的结构系统较为复杂，涵盖较多的应用功能，例如得到广泛运用的数据处理，或是指令性收集并整理分类多类型的数据信息等。在存储与归

纳指定的信息时，也可以依照一定的类别规则。由此可见，在面对具有多样性与复杂性特征的数据信息，数据挖掘技术依然可以实现规范化、系统化管理，便于人们搜集、查阅自身所需的信息资源。将数据挖掘与软件工程相结合，关键在于达到高效性、便捷化整合数据信息的目的，以便于企业的工作人员对整体资源架构具有全方位的了解，准确掌握数据与信息的管理现状^[2]。

（二）提高数据信息传输准确性

运算数据也是数据挖掘技术的突出性功能之一，通过观察与分析传统数据信息系统的实际使用效能可以发现，在数据运算的过程中，其往往需要投入较多的资金成本，并耗费一定的运行时间。若需要运算规模体量较大的数据信息，运行系统很可能会出现卡壳、瘫痪等严重问题，甚至破坏重要的数据信息。在研究解决此类问题的过程中，可以发现对数据挖掘技术的合理运用是一种有效途径，通过优化与改进系统整体架构，缩短其数据处理的运行时间，避免因故障问题而丢失数据信息，加强落实实效性为先的数据处理原则。若需要运算的信息数据体量较大，难以充分利用所有数据，可以借助于数据挖掘技术开展精准性运算，剔除长期留存的没有实用价值的信息数据，避免其对整个系统的运行效率产生制约或负面影响。只需将有价值的信息数据予以留用即可，进一步提高软件系统运行效率，以及数据处理质量^[3]。

（三）数据信息处理时间的缩短

将数据挖掘技术应用到软件工程中有利于缩短处理信息数据耗费的时间，将杂乱无章的资源内容进行科学化分类处理，以达到快速转换、调用所需数据的目的。与此同时，在深入挖掘、分类处理数据的过程中，数据挖掘技术的实用效率更高，还具有第一时间清理模糊数据的功能，优化软件系统内环境，确保现存数据的使用价值进一步提升。在搜集并获取到自身所需的数据信息后，以往人们普遍需要予以反复核对，确认自身得到的是真实的数据内容，应用数据挖掘可以省去人为检查这一环节，

其智能化核对数据的运行效率更高，有助于节省自主核验时间。

三、软件工程中数据挖掘技术的有效应用

(一) 系统结构

在软件工程中应用数据挖掘技术时，可以遵循含有三个主要阶段的规范流程有序开展：一是预处理数据、二是数据挖掘、三是评估模式和表示知识（如图2）。在第一阶段内，要想有效提升获取原始数据的作业效率，关键在于对任务处理对象的精准划分与明确，确保获取的数据信息能够满足软件工程的建构要求。与此同时，还需借助于其数据清洗功能，面对参

差不齐的原始数据，对其存在的漏洞与缺陷等进行针对性弥补，争取获得更为完整的数据内容。在从数据库中抽取数据时，需要注意信息内容是否与软件工程的建构任务相符。若需要转换数据，应注重强调准确转化数据原有的格式，确保数据信息的适用性满足软件要求。挖掘数据是软件工程的核心作业环节，这需要技术人员预先完成挖掘任务与目标的制定。在划分数据类别、开展评价总结时，需要选择适宜的运算方法，合理推敲获取的数据结果。评估模式、知识表示是表达挖掘成功的一种有效途径，其所遵循的衡量标准是兴趣度的高低，确保识别数据表达的功能水平有效提升^[4]。

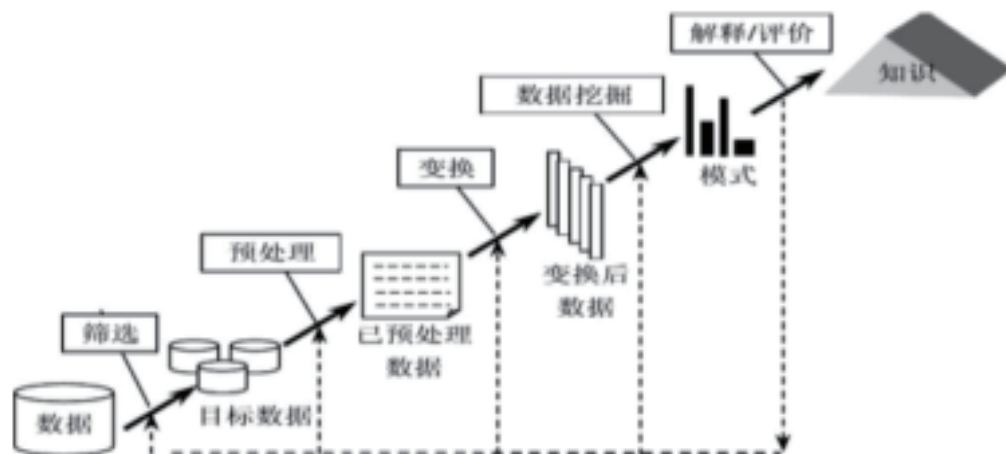


图2：数据挖掘技术模型

若要实现数据挖掘技术在软件工程中的进一步优化运用，可以以系统结构为着手点，一是对软件工程中是否存在克隆代码进行检测，在现行软件编制标准下，克隆代码主要是复制某一串代码，达到转移代码的目的。这样开发软件的作业周期会大幅缩短。但站在软件运行的角度来看，在实际克隆时，往往需要略微修改部分代码，确保运行软件程序的流畅度更高。因而，灵活运用数据挖掘技术可以准确、高效检测软件代码。二是数据信息挖掘法，其主要原理是对比并分析软件系统内存在的各个关键节点，依托于重建候选集，达到完善处理软件横切点的最终效果，促使编写与改进原有代码的完成效率更高，为开发软件工程提供便利，在编写软件代码方面缩减成本费用^[5]。

(二) 软件管理

难度大是软件工程项目管理的主要特点，若将数据技术应用到软件工程的管理工作中，可以从两方面分析并总结其具体表现，一是挖掘组织关系，二是挖掘版本控制信息。挖掘组织关系的核心与关键点在于对人力资源的合理调度，尤其是在工程建构的进行过程中，人与人之间存在的信息交互规模是十分庞大的。为了有效避免其出现严重不可控的紊乱现象，可以灵活运用数据挖掘技术，重新规整并梳理其内在的组织关系，为进一步展开软件管理工作奠定良好基础。

挖掘版本控制信息是软件项目管理的重要环节，以文件内容为基础，将其在某一阶段内的变化趋势完整地呈现出来，便于工作人员查看文件各个版本的具

体修订情况。除此以外，引入数据挖掘技术的根本目的是为软件项目管理提供有力的技术支撑，以便在后期管理阶段，实现对维修成本的合理管控，最大程度上降低成本支出。

现阶段，系统程序与现实应用的协同发展是软件工程的重要发展趋势，在开发与研制新型的软件产品时，一方面需要严格遵循各项参数指标，另一方面还需达到各类功能要求的规范标准，这大幅增加了开发软件的实践难度。对数据挖掘技术的灵活运用能够紧密结合数据库系统，将其内在信息的应用价值最大化发挥出来，为软件工程师开发的进步奠定基础。在深层次的挖掘信息数据时，数据挖掘技术还具有实时更新数据的作用，在质量层面上开发出的软件价值更高、实用性更好，原有的操作流程得到针对性优化。

（三）聚类

分析数据挖掘各个流程涉及的数据内容是聚类的本质含义，与此同时，其还能够理解与把握软件工程各项建构要求的基础上，细化数据类型，保障原始数据的固有类型不会受到外部因素影响。依托于数据挖掘技术的聚类功能，在面对属于同种类型的数据信息时，可以快速地找到其存在的相似性特点，并突出显示出个体之间的差异性，达到科学划分聚类的目的。在这种情况下，无法预测面对的具体对象，也可以实现对数据信息的有效挖掘，适用范围更为广泛，独特的数据分析方式带来的是真实性更高、有效性更佳检测结果。

（四）故障检测

传统的软件开发模式对人工操作具有一定的依赖性，大部分程序的执行过程均需要被人为地记录下来，应用数据挖掘技术不仅能够便捷、快速地获取程序执行记录，还能够对实际的挖掘方法做出具体说明。因而将数据挖掘应用到软件故障检测中能够实现对运行问题的精准测定，以及对潜在故障的准确判断（如图3）。

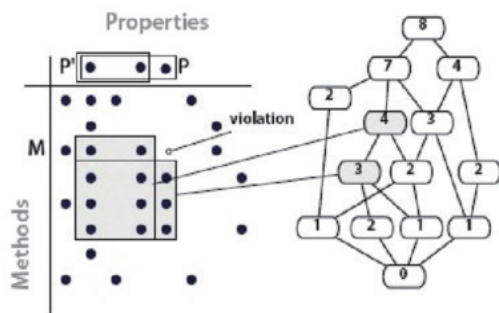


图3：故障检测流程图

一般情况下，引发软件故障的原因比较多，既可能来源于内部环境，也可能是外部因素，因而借助于数据挖掘抽象化定位软件程序的运行轨迹，便于工作人员在短时间内迅速确定故障源头，再寻求科学有效的解决手段。

四、结束语

随着软件工程在各行业领域内的深入渗透与广泛应用，其显露出的对技术支持力量的渴求是不容忽视的，要实现数据挖掘技术在软件开发、故障检测等环节中的高效运用，还需技术人员做出进一步的研究与探索。

参考文献：

- [1] 金巨波，葛雷，徐秀丽．计算机数据挖掘技术的开发及其应用[J]．浙江水利水电学院学报，2019，31(06)：68-72.
- [2] 贾淑滢．试论数据挖掘技术在软件工程中的应用[J]．山西广播电视大学学报，2019，24(04)：79-82.
- [3] 李高杰．软件工程中数据挖掘技术的应用与研究[J]．无线互联科技，2019，16(21)：141-142.
- [4] 魏凡其．数据挖掘技术在软件工程中的应用[J]．计算机产品与流通，2019(10)：20.
- [5] 郭竑晖，钟智慧．电信客户关系管理中应用数据挖掘技术的研究[J]．电脑知识与技术，2019，15(22)：5-7.