

# 关于数据挖掘技术在软件工程中的应用分析

王玉贞

国网河北省电力有限公司信息通信分公司 河北 石家庄 050000

**摘要:** 由于现代社会经济的迅速发展, 计算机软件对现代社会生产活动起到了十分巨大的促进作用。通过在软件工程领域使用数据挖掘方法, 不但可以提升软件的研究效率, 而且可以高效收集软件研究过程中可能面临的困难并积极加以分析, 由此推动对软件的研究能力的提升。

**关键词:** 数据挖掘; 软件工程; 应用

## 1 数据挖掘

数据挖掘技术是一种有效、方便、快速对信息进行收集、筛选与保存的方法, 其功能在于与现代数据技术相结合, 以便有效的在各类软件中使用。由于科技发展程度的限制, 中国目前很多公司使用的都是传统的计算机技术, 传统信息技术在工作的环境中需要大量的人才, 同时对人才需求也相对较多, 使用范围可以收到较大的效果<sup>[1]</sup>。数据挖掘技术是把数字信息转化为人类可以掌握的商业用语的技术, 通过对各类信息的筛选、分类、判断, 可以将数据转为有商业价值的信息。企业管理软件中运用数据挖掘技术, 可以较好的提高软件系统中信息的处理速度, 提高信息处理的准确性, 提高了信息处理水平与效率, 为公司决策层提供更为有效的资讯, 以便进行更为正确的投资决策。在应用的数据挖掘技术中, 首先就必须针对要研究和解决的问题, 收集大量的信息, 然后利用科学技术的方法对信息进行合理的筛选甄别后对信息进行模型处理, 然后将所分析和处理的问题与建模的成果加以对比, 由此获得科学的信息, 帮助研究和解决难题。数字模型由于不受其他的条线的影响, 并且完全建立在自然科学的原理基础之上, 从而能够较好的提高得出结果的实际准确性。

## 2 数据挖掘技术在软件工程中应用的意义

### 2.1 充分利用数据信息

数据挖掘技术在进行了大量数据采集之后, 就可以根据企业不同的要求, 对各种数据信息进行科学的分类, 并根据数据采集内容丰富、覆盖范围广等特点, 可以针对不同的企业数据有针对性的进行分析, 并通过客观科学的分析结果建立企业内部更加科学和完善的信息管理平台, 为企业内部工作人员查询数据提供了更多便利, 帮助企业管理者更为全面快捷的掌握公司实际管理与运作状况<sup>[2]</sup>。对于应用软件的发展, 数据挖掘能够有效的将各种分散的信息加以集中化管理, 使得公司可以在

各个方面对信息掌握比较充分。数据挖掘主要是把企业的不同数据, 按照特征不同加以合理的划分和有针对性的进行管理, 以达到企业对不同数据的充分利用。

### 2.2 有助于信息的获取

数据挖掘技术还涵盖了数据处理中的所有关键内容。数据采集后, 可以针对数据采集的情况进行分类管理。在这种信息处理机制下, 数据挖掘技术所涉及的业务范围也将进一步扩大, 从而可以结合数据分析的具体特点, 在海量数据资源中建立起功能更加完备的信息管理系统, 并能够协助我们迅速查找必要的信息来源。数据挖掘方法在软件工程中的广泛运用还能够整合大量的信息, 有助于公司人员迅速了解数据内容的各个方面。

### 2.3 提高信息的准确度

数据挖掘系统利用了科学的数字计算理论并拥有超级庞大的计算功能, 不过相对普通数据处理系统而言必须耗费巨大的时间与人工成本, 同时在人工实现计算机操作的过程中, 还将会由于各种问题而出现相应的故障<sup>[3]</sup>。通过应用数据挖掘方法, 该软件能够做到在短时间内迅速的实现了对数据信息的分析计算速度, 从而减少了数据信息或是错误等风险, 从而保证了信息的科学性完整性。同时在大量的资料信息中, 也有许多不合乎学术规律的资料信息, 如不及时加以解决一定的上将导致, 对于软件系统工作的有效性, 通过数据挖掘方法可以首先对无用的信息加以过滤, 消除软件的工作流程中的垃圾, 为其正常工作创造更大的工作空间, 进而提高其工作的效率与质量, 保证数据处理的效率。

### 2.4 提高实效性

软件工程企业通过对大量真实信息的发现与研究, 以便进行数据的合理转换, 对所有看起来没有规律的信息加以梳理, 同时采取合理的管理方法, 有效利用各种数据信息, 以此增强其内容的真实性。在数据挖掘中, 企业能够针对信息的差异采取不同的数据分析方法, 对

各种散乱无规律的信息，做出清晰的梳理，使信息能够发挥出其功能<sup>[4]</sup>。通过已经从资料中获取的信息，从不同方面，对其真实性、准确度加以判断，对比较现实可靠的信息采用建模的方法加以处理，以便将其信息转变成实用的商业信息加以传播，并以此为基础，有效的提升软件开发的质量。因为数据挖掘一个比较繁琐的过程，对各个分析过程的都必须经过不断的运算与检验，直至到达一个比较理想化的阶段，这时候把数据挖掘方法运用于软件的数据分析过程时，就能较有效的缩短数据和分析过程的周期，从而增加了数据分析执行中的稳定性，进而增强了数据分析的有效性。

### 3 数据挖掘技术在软件工程中的具体应用

#### 3.1 开源软件开发中的应用分析

所谓的开源软件，也就是对源代码处于完全开放状态下的软件，一般来说，开源软件普遍对客户免费开放，也正由于开源软件的这一特性，使得对开源环境的监管与研究显得比较麻烦。在这种大环境背景下，在开源软件开发中导入数据挖掘方法，将可以有效提升开源软件产品的研发效率。以日本大阪大学开发的分布式挖掘平台为例，该系统将能够在进行大系统挖掘的同时，实现对各种开源应用的挖掘；最后以由牛津大学所开发的数据挖掘平台为例，该系统实现了对系统开发人员和用户的持续追踪监控，从而大幅的提升了对开源软件的利用率<sup>[5]</sup>。

#### 3.2 在软件项目管理中应用

软件项目具有相对复杂化和系统化的特征，在软件系统中应用数据挖掘技术对于提升软件项目的整体管理水平将起到非常重要作用。首先，合理的建立企业组织关系。随着市场经济的发展，公司规模要求也在进一步的增加，各个部门之前有了比较细致的规划，而不少的管理人员并没有很好的理顺与企业的组织关系。如果在企业开展项目的同时，没有很好的管理企业员工间的人际关系，很有可能会产生许多无谓的纠纷，利用数据挖掘方法能够较好的对企业的组织关系有较为清楚的认识，以便于对企业现场做出更有效的分区管理，从而实现了企业活动的顺利开展，并提高了公司的办事效率。其次，提高软件的稳定性。许多公司在日常的办公中都采用了办公软件，办公软件中收录了许多企业内部的重要信息以及公司的个人隐私，所以提高软件稳定性是十分关键的<sup>[1]</sup>。在管理软件中使用数据挖掘技术能够较好的提高管理软件的稳定性，在软件受到外界系统侵扰的时候，数据挖掘技术能够及时发现，并给出一定的预警，帮助客户迅速找到问题，然后及时地对企业系统软件进

行处理，并修复缺陷，这就是数据挖掘技术会进行对系统修复情况的长期监测，以保证企业系统漏洞全部修补，从而提高了企业系统和人员对数据的稳定性，减少了不想要的经济损失。数据挖掘方法还能够实现对应用软件版本信息进行更有效的挖掘，用户可以比较了解所使用的版本信息，提高对应用软件的认识，从而达到对应用软件的有效使用，从而实现软件的正常运行，提高应用软件水平。

#### 3.3 软件漏洞检测

在软件检查中，漏洞检查成为工作中不能缺少的内容，有利于提高软件开发的效率和系统软件的稳定使用，及时发现并消除隐患。在数据挖掘技术的使用中，缺陷测试不但要对程序进行全面检查内容，还要进行相关检查，对与程序缺陷有关的部分信息也要进行全面检查，以适应软件开发的具体要求<sup>[2]</sup>。数据挖掘方法能够根据记录的资料对数据进行分类，提出数据分析结论，从而选择使用哪一种方法来进行软件测试的，从而使得方法的选取更加适合软件的具体要求。在脆弱性数据处理方面，数据挖掘技术需要首先对数据进行清理，然后再对相关数据进行分析提取，去除没有实际应用价值的信息，然后再补充内容中缺失的部分，以数据的方式展示出来<sup>[2]</sup>。各操作系统之间也可以通过对系统的数据进行对比，选取最恰当的数据。在漏洞检查和分类系统的漏洞和描述过程中，通过更新系统数据库中的软件漏洞信息，使各种其他的程序或操作系统可以在待获取的系统信息中有效处理相关情况，以使用户使用，缩短漏检查所需时间的同时又为漏检查和处理提供了参考数据，从而在发布漏洞信息前可保证数据信息的正确性<sup>[3]</sup>。

#### 3.4 软件故障检测中的数据挖掘

在软件故障检查项目中，通过一定的数据挖掘方法，能够对软件问题做出有效检查，并根据问题出现情况、常见问题，进行维修操作、保养作业。

##### 3.4.1 故障检测

在软件工程中，数据挖掘技术不仅可以提升挖掘任务的完成效率、挖掘程序执行记录及具体说明，同时还可以完成故障检测工作。在项目发掘过程中，首先，要确立工作重点任务，对信息做好跟踪分析，提高程序掌握率，为后期项目的实施提供保障；其次，对于挖掘手段的适当采用，比如采用基于规则的自动化等挖掘工具，也能够使得抽象分析、抽象挖掘任务能够高效实现。

当前，软件由于各种因素的作用，也容易产生故障现象，而一般的方法虽然能够对其加以解决，但它仍存在一些繁琐性、低效性，其的准确性还会出现一些困难<sup>[4]</sup>。如

果利用数据挖掘技术,利用先进的定位方法,如利用程序谱抽象定位问题的位置,利用对比方法对问题进行有效排查,则能够使问题根源能够快速找到,使问题排除速度、分析精度得到提高。

#### 3.4.2 软件维护

在数据挖掘技术下,通过识别和评估软件缺陷程度,能够对软件修复环节漏洞加以分析,对修补的次序进行评估,使各个漏洞的修补途径得到明确。为提高修复质量,应完善语言与操作的故障报告功能,在故障报表中整合了自然语言、执行的相似性数据,以保证依照所定义的向量维度即可进行描述软件缺陷的功能。然后,通过加权和计算相似点,来进行缺陷恢复操作。

#### 3.5 数据挖掘执行记录

数据挖掘的记录是利用研究应用程序的行为,揭示计算机程序内部可能存在的相互联系。实质上,挖掘过程记录就是利用对相关过程行为的研究,实现逆向建模,以此实现特定的目的。挖掘过程记录重点就是可以检验、管理和分析过程<sup>[5]</sup>。

#### 3.6 检测版本信息

在版本信息中使用数据挖掘系统时,必须注意项目成员的数据资料应该是一致的,如此一来,不但给信息数据挖掘带来了便利,而且系统使用起来相当简单。利用版本信息对开发软件以及系统管理进行控制在软件工程中是一件十分寻常的工作,因为通过对不同版本信息的利用,可以获取到大量数据信息和资源,并能对产品存在的缺陷和错误加以检测。由于数据挖掘技术在软件领域的运用日益深入,软件人员对系统的把握也随之提高,促进了应用软件的良好发展。

### 4 在软件工程中数据挖掘技术的应用发展方向

数据挖掘技术在软件工程中应用的根本目的,是为了最大限度地提高软件开发的质量和效率,为后续的管理工作打下良好的技术基础。所以,数据挖掘技术和软件中的任何一个方面都应该随着软件工程的进展而不断完善与创新,从而推动软件工程和数据挖掘技术更加密切的融合,充分发挥其效果。所以,数据挖掘技术在未来的开发与使用过程中,应向着更先进的挖掘研究目标

奋力前进<sup>[1]</sup>。也就是说,根据软件工程和数据挖掘领域面临的具体情况,必须研究一个稳定性较好的数据挖掘方法<sup>[4]</sup>。值得注意的是,目前的软件领域,尽管大多数测试工具都可以检查出偏离给定过程的动态或静态规范而导致的问题,而如果检查出容易产生问题的地方,可通过程序进行,但还是存在较大问题,至于其他更复杂的问题,则很难精确检测。所以,研究对象必须应该是最广泛的复杂问题检测工具。另外,考虑到近年来对我国软件需求的逐渐增加,以及国际软件需求量的逐渐扩大,不管是细化了软件开发程序,还是对软件开发者的参与层面都有了更多的要求,因此软件研发时间也将继续被拉长。面对这一情况,尤其面对大量的数据资料,数据挖掘技术必须围绕着信息降维、特征提取的开放目标发展,进而研究出更加有效的信息预处理技术。另外,我们还需要充分融合人工智能技术、机器人技术等前沿领域的科技,并完善了数据挖掘技术本身,才能更好的服务于软件工程的各个领域<sup>[2]</sup>。

#### 结语

综上所述,数据挖掘方法在软件中的运用能够比较有效提高软件工作效能,从而降低了软件的实际应用与管理过程中的风险,从而提高了软件工作的稳定性,让软件更好的为企业服务。要做好数据挖掘技术在大数据分析、系统漏洞分析、应用数据管理平台、开源软件工程等领域的运用,有效的促进软件智能化发展,以便进一步推动中国信息技术的高速成长。

#### 参考文献

- [1]刘知云.关于数据挖掘技术在软件工程中的应用分析[J].电脑知识与技术,2021,17(24):27-28.
- [2]王春明.在软件工程中数据挖掘技术的应用探索[J].科技资讯,2021,19(12):42-44.
- [3]张建新.软件工程数据挖掘技术应用分析[J].信息记录材料,2021,22(03):163-164.
- [4]张海霞.数据挖掘在软件工程领域中的应用探讨[J].现代信息科技,2020,4(21):13-16.
- [5]张欲丰,储小勤.探析数据挖掘技术在软件工程中的应用[J].信息记录材料,2020,21(03):189-190.