DOI: 10.12361/2705-0866-05-02-116822

基于计算机软件工程中的数据挖掘技术实践分析

蔡俊伊 曹毅强 高紫怡

武汉东湖学院, 中国·湖北 武汉 430212

【摘要】随着大数据时代的到来,在计算机软件工程中,通过数据挖掘技术获得更大的数据价值,已经成为软件工程开发人员的重要工作目标。对数据挖掘技术进行全面、深入的研究,是现代软件工程领域发展的重要趋势。数据挖掘技术优势非常明显,在不同的挖掘方法支持下,可以确保软件工程开发和实践中,对各种数据的科学有效掌握。通过源代码开发、软件管理信息、程序代码以及软件故障检测中的数据挖掘技术应用,可以进行高效的软件数据管理。基于此,对计算机软件工程中的数据挖掘技术实践进行研究,以供参考。

【关键词】计算机软件工程; 数据挖掘技术; 实践分析

Practical Analysis of Data Mining Technology Based on Computer Software Engineering

Junyi Cai, Yiqiang Cao, Ziyi Gao

Wuhan East Lake University, Wuhan, Hubei, China 430212

[Abstract] With the arrival of the era of big data, in computer software engineering, obtaining greater data value through data mining technology has become an important goal of software engineering developers. It is an important trend in the development of modern software engineering to conduct comprehensive and in-depth research on data mining technology. The advantages of data mining technology are very obvious. Supported by different mining methods, it can ensure the scientific and effective grasp of various data in software engineering development and practice. Through the application of data mining technology in source code development, software management information, program code and software fault detection, efficient software data management can be carried out. Based on this, the data mining technology practice in computer software engineering is studied for reference.

[Keywords] Computer software engineering; Data mining technology; Practical analysis

引言

“数据挖掘”这一技术已经被广泛地运用到生活的方方面面。不管是在医疗卫生方面或者是APP开发方面,还是金融、教育等领域都可以到处看见数据挖掘技术留下的踪迹,各行各业使用数据挖掘技术显示出大数据在社会中的重要地位。比如说在金融方面,大部分银行都为客户提供了丰富多样的储蓄服务、保险服务或者投资服务等;在销售方面,数据挖掘技术在积累销售顾客、划分销售群体、统计销售数据等方面发挥了重要的作用;在科学研究方面,数据挖掘的应用如观测卫星、遥感器、DNA分子技术等都有很重要的体现。

1 数据挖掘技术

数据挖掘技术,是通过软件工程中的程序编辑和模型建立,对大量信息数据进行分析和提取,找到符合运行目标的数据。数据挖掘技术的实施,具有较强的流程性和规则性。数据挖掘的对象目标,往往是整个数据库的数据,在数据库建立时,就要考虑到数据挖掘的需要,从数据格式、数据分类等方面做好铺垫。数据挖掘过程

按照既定步骤,首先进行数据的探查,数据探查是对数据的信息的全面掌握,包括数据源、数据链接、数据格式等,在数据探查中可以对数据进行必要的统计分类,这样可以为后续的提取和挖掘做好准备。其次进行数据提取,数据提取必须有相应的信息标准,也就是计算机程序必须明确,提取的数据必须具备哪些信息特征,这样可以确保数据提取的针对性和有效性。最后是数据模型建立,数据模型是根据计算机软件工程要求,根据数据规律的掌握和挖掘目标情况,在具体数据中找到一个统一的框架,这个框架就是一种数据模型,当模型建立后,只需要更换具体数据,就可以自动快速地获得相关结果,使软件工程可以完成既定的任务目标。

2 计算机软件工程中的数据挖掘技术应用

2.1 源代码开发中的数据挖掘应用

计算机软件工程中,需要基于软件开发目标和要求,做好数据挖掘工作。软件开发是基于源代码编辑实现的,软件工程人员需要根据需要,分解开发项目,并通过对软件应用范围、标准和执行路径等的理解和掌握,在源代码开发中利用数据挖掘,提高软件开发

能力。软件开发过程中,由于各种源代码来源非常广泛,在利用这种源代码重新构建新的软件模块时,需要依据开发对象的特征,找到数据挖掘的标准和原则。数据挖掘技术应用中,利用不同的数据挖掘方法,可以构建一个专门用于源代码分析、提取和建模的数据系统。该数据系统可以完成规模性的开发任务。如软件工程开发中,存在一个软件系统多项具体应用功能时,数据挖掘技术就可以显示较强的应用价值。在数据挖掘时,基于源代码的功能、目标等分类,可以提取出符合开发对象需要的一系列数据,这些数据可以重新构成一个数据库,以供软件开发人员进行借鉴与使用。

2.2 软件管理信息中的数据挖掘应用

软件工程发展中,需要考虑到软件后续维护方面的需要。利用数据挖掘技术,可以对软件项目版本信息更好地掌握,从而增强维护能力。对于软件而言,随着功能的完善和漏洞弥补等,对软件会进行必要的更新和修复,这个过程必然会涉及到版本信息的变化。数据挖掘技术运用于版本管理方面,可以在计算机软件系统安装和更新时,快速准确查找到版本的一系列信息情况,包括版本号、版本更新时间和更新内容,以及版本修复记录等。数据挖掘技术,将这些挖掘到的软件信息予以归类 and 呈现,可以让软件开发和使用者,对软件系统有更全面地掌握,从而在具体工作中可以规避一些异常使用情况。

2.3 程序代码中的数据挖掘应用

代码构件挖掘,程序代码数据挖掘时,可以实现代码构件的提取和掌握。代码构件是一种对程序代码的结构检索模式,由于计算机程序编辑中,程序代码主要为字符格式,代码构件就是基于这个特点,建立代码关键词检索功能。代码检索对于数据挖掘具有相关性,在数据挖掘技术应用中,能够基于检索到的代码关键字,了解到具体代码程序的编辑逻辑结构,使用的计算机程序语言,以及软件工程师对代码的描述等,这些挖掘到的数据具有内在相关性,也就是都与检索关键字有一定联系。根据这样的数据挖掘规则和方法,软件工程领域开发出了诸如百度、谷歌检索网站,用户能够在仅仅掌握一些关键字,或者对检索目标有一定了解的情况下,通过数据挖掘获取更多相关软件内容。

2.4 软件故障检测中的数据挖掘应用

数据程序切片,随着软件工程领域的快速发展,软件工程规模更大,程序语言更多,给软件故障检测带来一定的困难。在数据挖掘技术实践中,可以针对软件故障程序,进行数据程序切片,将复杂的数据程序,分解为多个独立的程序单元,这样可以将故障更好地呈现出来。数据程序切片,是利用数据挖掘技术,对软件故障程序进行综合分析和分类。诸如采用关联挖掘技术,根据科学合理的数据关联原则,对程序进行切片,这样的程序切片可以保证软件程序在功能、逻辑等方面保持一定的完整性,不会将一个独立运行的程序分割为无法获知详细编辑目标和执行过程的碎片程序。数据程序切片采用的挖掘技术,要进行程序故障的定位,也就是故障程序应当包括在数据切片中,即便是故障程序处于不同切片,也可以基

于数据挖掘技术,对故障程序运行情况,构建数据轨迹,这些数据轨迹也是数据挖掘的一部分,由此进行程序校对和检测,可以有效提升程序故障检测水平。如果一个数据程序切片检测完成后,可以依据数据挖掘的规则,将相关数据程序切片进行定位,这样能够较为全面地兼顾到程序故障的影响范围,强化程序故障检测质量。

3 数据挖掘技术与计算机网络病毒防御

在计算机网络病毒防御系统中应用数据挖掘技术,可以借此技术对数据信息进行筛选与处理,以带入关键信息的方法将与病毒有关的代码筛选出来,对结果进行分类整理,并结合此项技术转换数据信息的格式。在一般情况下,网络病毒并不会以文字或数字的形式出现,而是通过代码的形式攻击计算机网络系统,而代码的形式需要经过有效的转换,才能够被展示出来,但是只有将病毒代码转换为计算机可以进行识别的形式后,才能够破解病毒的类型。数据挖掘技术中的处理技术主要功能便是转换病毒代码的形式。在实际处理病毒代码的过程中,需要对数据源信息、源IP的具体位置及数据源的位置等内容进行深入的挖掘,进而实现对网络病毒传播途径的有效阻止。

结束语

随着人们对信息技术需求不断提升,随之而来的是计算机性能的不断优化,再加上各学者对数据挖掘技术研究力度不断加深,数据挖掘在各行各业都已经占据着越来越重要的地位。可以说,在这种背景下,种种迹象都表明数据挖掘技术代表未来面对大量数据处理技术的一个大体趋向,其所涉及的数据采集、数据预处理、聚类分析和数据变换等各种核心技术和算法都将会成为未来各行业处理数据不可或缺的一部分,有了这些技术,无论是学者的研究还是各行业的运营,都将会达到事半功倍的效果。

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] 陈翠娟. 基于关联分析的计算机软件数据挖掘技术[J]. 安阳师范学院学报, 2021(02): 28-31.
- [3] 陈丽. 数据挖掘技术在计算机网络病毒防范中的应用[J]. 电子技术与软件工程, 2021(08): 247-248.
- [4] 田琴琴. 数据挖掘技术在软件工程中的应用[J]. 信息通信, 2020(08): 157-158.
- [5] 龙艳. 数据挖掘技术在软件工程中的应用[J]. 信息与电脑(理论版), 2019, 31(23): 122-123+126.
- [6] 彭湘华. 数据挖掘技术在软件工程中的应用[J]. 电子技术与软件工程, 2018(16): 156.
- [7] 蒯天宇. 数据挖掘技术在软件工程中的应用[J]. 数字技术与应用, 2018, 36(05): 112-113.
- [8] 王蕊. 数据挖掘技术在软件工程中的应用分析[J]. 信息与电脑(理论版), 2018(07): 131-132+135.