

经济发展的统计应用方法概述

李旭兴

中国人民大学统计学, 中国·北京 100872

【摘要】近年来, 市场经济的高速发展, 为我国各领域创新改革带来重大机遇, 特别是对于经济统计行业而言更是如此, 在引入先进的信息技术以后, 数据调查中自身的复杂性和统计中的繁琐性可以得到有效的改善。数据挖掘技术作为互联网成熟发展的产物, 应用到经济统计中, 可以利用自身更大的优势, 从海量数据中挖掘出更多有用的信息, 利用算法和分析软件进行处理整合, 从而解决实际遇到的问题。鉴于这种现象, 此文着重分析数据挖掘技术在经济发展中的应用, 了解应用的优势和发展方向, 并提出具体优化策略, 旨在更好的提高数据挖掘技术在经济统计中的应用效果, 完成多学科的综合, 提高社会经济发展分析水平。

【关键词】 大数据挖掘; 经济发展分析; 统计技术

前言

一般来讲, 数据挖掘基础是对于相关数据信息进行更加深层次的剖析, 而这一技术原本是相对复杂的, 并且有着多种变化, 但可以对数据库进行简化, 对经过简化后的数据库进行深度的分析, 能够找到其中可以被利用的数据, 通过多种数据的整理分析, 能够提升其数据信息应用的效果, 而目前我国现阶段的经济统计方面所涉及到的调查数据较多, 但是效果却不是十分理想, 采用数据挖掘基础却能够有效缓解这些问题, 通过软件程序算法, 能够对原本并非十分完整的数据进行综合性的分析和整理, 也能够更加直接地对数据进行提取和使用^[1]。

1 数据挖掘技术相关概述

1.1 概述

数据挖掘技术作为信息化时代下成熟发展的产物, 随着网络发展以及信息爆炸式增长。特别是在网络成为人们生活中重要组成部分以后, 人们对于网络的依赖和需求明显增加, 通过网络可以打破时间、空间的限制, 为人们的生活带来巨大的便利, 但是不可否认的是, 所产生的相关数据也呈现出爆炸式的增长, 也无形中阻碍了人们对于有价值信息的获取, 海量的数据加剧了鉴别的难度, 更能够帮助人们在海量信息中通过特定筛选的方式来获取更加有价值的信息, 促使信息可以真正的服务于人们的生产和生活。在海量信息中, 如何进行更加精确的筛选和甄别仍然是极需解决的重要问题, 在这一需求和趋势下, 数据挖掘技术应运而生, 凭借着数据系统处理到人工智能分析, 再到后来的机器学习技术, 功能逐渐强大, 效率也随之提升, 经过多年来的发展和实践, 数据挖掘技术的应用更加广泛, 在很多行业中发挥着不可替代的作用。因为我国经济发展和时代进步创设了更多的力量, 而数据挖掘和其他学科之间彼此融入的深入, 特别是以经济统计学的融合, 也能够为统计工作的效率和准确性, 提供更加有力的保障^[2]。这样一来, 统计工作的效果也能够达到预期的状态, 提升了整体工作质量和效率, 许多数据复杂问题也迎刃而解^[3]。

1.2 主要流程

第一点是定义目标, 在使用数据挖掘技术投入到经济统计中, 需要由相关领域专家、学者对于所涉及到的挖掘数据进行目标定义, 而目标定义效果会直接影响, 甚至决定着后续的数据

挖掘所产生的效果, 因此数据挖掘初始定义也有着十分重要的作用, 一边定义一边进行算法的演算, 更能够确保定义的效果更加清晰明确, 也能够提高其效率, 但是第一要求需要与实际工作需求相对应, 第二点是数据准备作为整个数据挖掘技术中最为核心且消耗最长的环节, 数据准备阶段一般会分为三个步骤, 分别为数据的挑选、数据预处理、变化分析^[4]。其中数据的挑选作为整个宏观数据库中所涉及到所有数据的排列和筛选, 并且定义到目标区域数据, 紧接着利用数据处理能力, 对于所排列的所有数据进行初步的筛选和处理, 找到其中符合要求的数据, 最后利用数据变化的功能, 对于相关数据进行简化处理, 也能够找到最符合目标定义的数据信息。第三点是数据挖掘作为首要步骤, 也能够对算法进行预订。需要设计出合适算法来变换数据挖掘关键点所在。第四点是评估与显示, 在经过上述一系列步骤以后会得到最终的数据结果, 但是数据使用者仍然需要对其结果进行评估, 并展开全方位的显示, 自动生成数据价值评估后, 筛选有利的信息, 将一些无意义的信息和数据及时删除^[5]。

2 数据挖掘技术在经济统计中的应用优势及发展方向

2.1 应用优势

第一点是有着较强的综合性应用能力, 目前许多数据挖掘技术自兴起以来得到了广泛应用, 特别是在对数据进行分析、归纳、处理的过程中有着举足轻重的作用, 无论是经济统计、物理分析或数学统计等方面, 采用数据挖掘技术, 均能够凭借其较大的统计能力和数据挖掘能力以及后续的数据自动分析统计功能, 协助使用者对数据进行完整的处理。因此, 数据挖掘技术在经济统计中会起到十分重要的作用^[6]。第二点是有效性较强。目前我国的数据挖掘技术方面钻研的应用时间相对有限, 基于技术的角度与发达国家相比仍然存在较大的差距。但是根据目前数据挖掘技术而言, 运用到经济统计中仍然有着十分理想的效果, 促使统计工作的性能可以得到进一步增强。并且我国在经济数据领域分析和统计与整理能力中也进行了强化, 利用数据挖掘技术对于更多有价值有意义的数据进行深度挖掘, 也真正提升了其数据高效性和准确性^[7]。第三点是宏观性的数据应用。对于如今宏观性的数据库, 数据挖掘技术应用其中能够取得理想的效果。目前我国经济数据统计分析中所应用到的

更多方式仍然以传统经济统计为主,对于信息收集、处理、整合的方法仍然存在滞后性,无法与所涉及到的信息匹配成一个完整的整体^[8]。很多数据使用者在对数据进行全方面管理,分析师仍然会面临较多的问题,无法达到预计的统计效果。针对这一情况,也需要引进全新的方式来提高经济数据处理的效率,宏观经济统计数据库也能够为数据挖掘技术创设出一个能够完全展示其功能的平台^[8]。

2.2 发展方向

第一点,可以结合具体的使用情况进行针对性的数据挖掘。由于数据的使用者所涉及到的数据会和目标有着很大的差异性,因此,采用数据挖掘基础所对应的系统也需要适应更多的变化,而目前数据挖掘技术的发展并没有十分成熟且彻底,仍然会存在许多疑难问题。主要体现在数学数据难题无法匹配相对应的功能性数据挖掘技术,也正是因为如此,许多设计者所面临的任务仍然是设计出更多的数据挖掘系统来解决层出不穷的数据处理和统计的难题。

第二点是利用数据挖掘技术来提升结果的有效性。数据挖掘技术应用到经济统计中仍然会存在精确性不够理想这一问题。在宏观数据库中,进行数据的筛选和归纳,有着较强的可表达性有效性和确定性。而利用自动化原理对于所表达的数据进行深度挖掘时,也能够充分显示出数据库中的主要内容,并且将相应的算法应用到数据统计处理中。采用数据挖掘技术也可以对一些有着明显缺陷的数据进行补充和处理,对于一些曲线不够明显的数据也能够按照相似数据的规律展现出来。

第三点是数据挖掘让结果更加简化。在未来数据挖掘技术应用到经济统计中需要进一步减少数据挖掘的频率难度和时间,能够让结果变得更加精确且简单。即便是一些非专业学者也能够采用数据挖掘技术对于经济统计中的相关数据进行简化。

第四点是交互性的数据挖掘技术发展。数据挖掘技术资源的技术交互性是其特有的优势之一,利用交互式的数据挖掘,能够对于所需要处理分析整理的数据进行精确的挖掘,利用数据焦点的动态变化等特点,也能够从多个角度对数据进行分析处理,采用抽象层次变化属性对归纳的数据进行挖掘,也能够实现数据交互。

第五点是数据挖掘的安全性。由于经济数据具有一定的保密性和很高的安全性需求,也会随着科技进步、时代发展在安全保密角度有着更大的进步。

3 数据挖掘技术在经济统计中的应用分析

3.1 经济统计数据调查中的预处理

在经济统计学中数据挖掘技术已经兴起和应用并取得了十分理想的效果,并发挥着重要的作用。主要是因为数据和字段之间的关系是很复杂有函数关系、关键词关系,甚至是数据引申含义关系。而无论是哪一种,若想要通过统计的方法来明确其运营的规律,都需要对数据进行进一步筛选,这也是需要根据统计学的原理对体量庞大的数据进行有效的分析,从而将更多有价值的信息从海量数据中提取出来,再对应统计学的方法而对比。十分常见的仍然是以神经网络策略为主。该策略可以通过计算机来模拟人类的大脑对于数据进行深度加工和处理,对

数据进行分析和处理的效率也会十分理想。针对不同类型的数 据也能够提供更加成熟完善的处理机制,一方面可以提高统计工作准确性,另一方面也能够减少数据挖掘中存在的海量数据挖掘难的问题。利用神经网络也能够使统计的结果变得更加实用化和形象化,让人们能够更加方便地使用数据。这种处理方法的 优势很多,其中最为明显的优势之一是初级理论策略的应用不仅操作简单,无需给出额外信息,更能节省空间,提升数据信息的安全性和可靠性。特别是对于一些有着隐藏安全隐患的经济统计而言,数据挖掘可以发现其潜在的风险,由相关人员第一时间做出应对。数据预处理是数据统计活动中十分常见的一种方法,因为需要对数据进行预处理,也因为挖掘技术自身可能会受到多种经济条件的限制,因此,无法更好的取代经济系统中原有的数据收集。但仅仅是对于技术信息进行智能化分析,以及在统计调查中得到数据,复杂性处理的内容也会涉及到较多的种类,例如处理不准确的数据或是处理错误,以及处理差距较大的相关信息数据等等,这些内容实质上也会被称之为数据清洗。

3.2 调查数据分析中的应用

数据挖掘技术应用到经济统计中的具体应用会分为两个大致方向,分别为描述方向和预测方向,其中前者会包括聚类以及描述变量和数据值。之间所涉及到各种关系的算法,而聚类也会将相关关系中较大的变量或个体均为一类,称之为R型聚类或是Q型聚类及实质上会起到降温的效果,而后者描述变量或数据之间的相关关系,也有着经典的相关分析,通过研究典型的相关系数,可以揭示两组变量之间的对应关系,研究两组分类变量或是分类变量,内部各个状态预测方向主要也会包括分类和回归,其中前者的主要工作是用于贴标签,具体算法有决策树、神经网络遗传算法、SVM等等。而后者回归模型会包括线性回归等等。其中,聚类在经济统计调查中也最为常见,工作人员为了进一步提高其调查的精度,会在抽样调查中先将对象进行聚类,再从不同的类别中分别抽取样品,在抽样调查中的分层抽样和整群抽样都可能用到聚类,例如在对居民收入情况进行抽样调查时,可以先根据居民的消费数据,对齐聚类进行分析,一类分为一层,进而对于居民进行分层抽样,从而提升最终的调查精度,而降维方法也会广泛地应用到经济统计调查中,由于统计调查进行一次需要花费大量的人力、物力,因此,在进行统计调查过程中所涉及到的指标十分复杂,且数据量较多,这也需要数据挖掘中采用降维技术将相关性大的变量合并在一起,促使数据的分析变得更加简单,而常见的降维方法有因子分析法、主成分法等等。

4 结论

总而言之,在经济新常态的背景下,我国市场经济高速发展离不开成熟完善的经济统计和分析,而在海量数据中,如何更好地提取相关数据并提升数据统计分析的正确性和有效性仍然是亟待解决的重要问题,数据挖掘技术作为信息技术成熟发展的产物,在投入到经济统计中,可以从海量数据中挖掘出更多具有价值的信息进行聚类分析,也能够提升其统计效率和效果,而经济统计的效果也会直接影响着经济长远发展前景,数

据挖掘技术应用在经济统计领域大范围的应用也能够使我国目前的数据统计分析领域有着更加优异的表现。因此,数据挖掘技术会得到更加广泛的应用。也会使得统计工作的结果变得更加真实且可靠,更能够为相关部门的经济统计工作提供明确的理论依据和保障。

参考文献:

- [1] 刘西清. 数据挖掘技术应用于经济统计中的研究[J]. 人力资源管理, 2016, 000(010): 293-294.
- [2] 辛金国, 柯芳, 李绍君, 等. 数据挖掘技术在浙江省经济统计中的应用[J]. 统计科学与实践, 2009, 000(004): 31-33.
- [3] 杨梅冰, 梁思思. 关于数据挖掘技术在经济统计中的应用[J]. 时代金融, 2015, 000(008): 173-175.

[4] 张文彤, 姜庆五, 赵耐青, 等. 数据挖掘技术在生物信息学基因变异规律研究中的应用[J]. 中国卫生统计, 2005.

[5] 董世彪. 数据挖掘技术在少年健康体育行为应用中的研究[J]. 现代电子技术, 2017(09): 120-122+128.

[6] 王海荣. 数据挖掘技术在教学管理中的应用研究[J]. 电子设计工程, 2013(03): 43-46.

[7] 周光华, 李岳峰. 数据挖掘技术在卫生统计信息工作中的应用研究[J]. 中国卫生信息管理杂志, 2012, 9(006): 82-86.

[8] 朱世武, 崔巍, 张尧庭, 等. 数据挖掘与其他技术的比较[J]. 统计研究, 2003, 07(7): 58-58.

作者简介:

李旭兴(1990.10.05-)男,汉族,籍贯:山西省运城市,本科,职称:无,研究方向:统计方法研究。