

数据挖掘技术在经济统计中的应用

曹 晓

齐鲁师范学院, 中国·山东 济南 253200

【摘 要】随着经济社会的不断进步,经济统计的数据量开始逐渐加大,传统的数据处理方法已经远远满足不了当今社会经济统计工作的需求,导致在进行经济统计工作带来了阻碍。而数据挖掘技术能更好的解决这一问题。本文就数据挖掘技术的特点,对如何在经济统计工作中对其进行科学运用作了相关方面的分析,希望为经济统计工作提供一定的参考价值。

【关键词】数据挖掘; 经济; 统计; 应用

引言

现今社会,计算机、网络技术不断发展,获得有关资料已经非常简单易行。但是经济的快速发展已经开始超过了人类大脑可以理解的范围之内了。在完成这类庞大、涉及面广的数据分析时,不能仅仅只依靠以往那种简单汇总、按指定模式去分析的统计方法是。因此我们需要对当代先进的统计技术加以有效应用,来对经济发展的趋势进行分析。在我国最近几年的统计学发展过程中,在经济统计活动当中,数据挖掘技术逐步开始应用到经济统计和数据分析的工作当中,更是现今经济工作中的重要统计技术。

1 简述数据挖掘技术

数据挖掘技术简单来说指的就是一种可以从海量信息数据中,快速、精准找到所需要的有用信息数据的技术。用传统方法处理数据时,虽然信息数据量较为庞大,但其中有许多信息数据并不完整,甚至存在诸多错误信息数据。在经济工作中对于选出来的信息数据利用数据挖掘技术进行加工处理,就可以将其中新颖的、有价值的信息数据准确获取出来,为信息统计工作提供真实可靠的数据价值。在经济统计当中运用数据挖掘技术,可以帮助统计人员快速从未加工的经济数据中,筛选出有价值的统计数据,并对其进行深入处理与再加工,为相关工作人员分析经济形势、判断经济走向、制定经济决策决议等提供重要数据参考。此外,由于数据挖掘技术并不只是一种简单的数据分析工具,其可以深入结合信息使用者的实际需求,特别是在我国经济不断发展下,经济数据信息量加大,对经济统计的要求也越来越高。因此将数据挖掘技术运用在经济统计当中,能够有效保障信息评估的可靠性,实现数据处理成效的进一步优化。在运用数据挖掘技术完成对经济数据的深入挖掘与分析处理下,建立起相应的数据库,可以在更好地融合各种数据信息的基础上,不仅减轻统计人员的工作量还减少了重复性的统计工作,给经济管理带来便捷及高效率的服务需求。

2 在经济统计中数据挖掘技术的优点

2.1 超强适用性

在实际工作中,在各行各业的数据统计工作中均会用到数据挖掘技术。由于各行业的经济部门不同,对于数据统计的需求也是存在区别。在一定意义上讲,数据挖掘技术是一项具有整合功能的统计方法,其适应性比较强,通过数据挖掘技术对经济数据进行精准统计处理,将信息充分融合,满足不同经济部门对数据统计的要求,也弥补了原有传统统计方法的缺陷。

2.2 具有良好的效果

在经济数据统计中,采用数据挖掘技术,还可以对大量的数据进行统计、加工和整合,从而将那些大量没有整理的

信息进行有效、科学的信息化处理,实现数据统计的高效性。数据挖掘技术把现有的数据进行精准核算,更能保证经济统计数据的真实性和有效性。

2.3 宏观型的数据库

现阶段,在很多企业的经济部门大多还是采用传统的数据统计方法,在对数据进行处理时,出现了很多问题,收集的统计数据信息不够全面,不够精准。因此,统计工作需要提供更好的新技术。构建经济统计数据库,可以为数据挖掘技术的发展提供了更大的平台。在对经济统计信息管理时,数据管理系统可以做到正确无误,然后经过数据挖掘技术的整合处理,就能得到更丰富的数据资源。这就是所谓建立更为庞大的数据库的意义所在。

2.4 方便进一步调查

在使用数据挖掘技术进行某地区经济统计时,经济统计人员在立足本地区企业历年上报数据的基础上,通过充分结合企业规模变化率以及是否有重大经营事件曾经发生于企业内部,判断此时是否需要继续进行先一步调查分析。具体来说,如果根据数据挖掘结果显示,企业规模变化差异率较大,但并未超过合理范围,则无需进行下一步调查从而获取更加可靠的有效经济信息。如果企业规模变化差异率超过规定值,或是企业规模变化差异率极小,且企业确实曾经发生过重大经营事件,此时需要经济统计人员展开下一步调查分析,从而更加深入地了解企业及本地区存在的经济问题,以便可以制定出更加具有针对性和有效性的经济决策决议,在确保经济统计结果的完整性与可靠性的同时,对企业规模变化差异率进行合理调整,达到帮助企业提高自身核心竞争力,扩大经济效益规模,最终实现推动本地区经济长效发展的目的。

3 数据挖掘技术在经济统计中的运用

数据统计得到的数据信息要具有完整性和真实性,数据挖掘技术正好满足了经济统计工作的需要。基于此,本文总结了以下几种方法。

3.1 预处理方法

在将数据挖掘技术运用在经济统计当中时,首先需要对现有的经济数据进行预处理,这主要是由于在经济统计中搜集得到的各种经济数据,属于初始数据,其中混杂着部分错误和虚假数据,以及不完整数据。对经济数据进行预处理,其根本目的在于快速筛出存在于经济数据集当中,与经济统计工作无关的信息数据,只精准提取出经济统计分析所需的重要数据。例如在某地区的经济统计当中,在对该地区历年企业上报的基础经济数据进行集中整理后,可以采用均值法进行经济数据的预处理。即企业上报的基础经济数据中如有空值或是噪声数据,则要对此类数据加以处理,使用数据库当中该属性全部已知属性均值进行空缺添

补,可以有效保障后续经济基础数据挖掘以及统计分析工作的顺利进行,并由此获得具有高精准性的经济统计分析数据。

3.2 决策树方法

在完成经济数据的预处理之后,需要建立相应的决策树。具体来说,在经济统计中运用数据挖掘技术,建立决策树时,工作人员首先需要利用训练集完成决策树的初步建立,而后借助专门的决策树算法对建立起的决策树进行简化。此时所建立的模型即为数据输出分析模型,随后经济统计工作人员需要对建立起的决策树进行充分利用,对相关经济数据进行准确分类与深入分析。按照从决策树根本开始逐渐延伸至决策树各枝干部分的顺序,直至数据与既定条件相符合后才可终止分割。值得注意的是,如果同时进行两个分割,且分割点上的数据完全相同,此时将会自动终止分割,即顺利完成决策。如果数据本身无分类属性,但可以继续分割输入的经济数据,则此时需要停止分割。在经济统计当中,可以根据当地企业上报的数据,采用数据挖掘技术进行预测值计算,将企业预测值与企业历年上报的实际经济数据进行对比,由此获得差别率。此后统计人员可以对获得的差别率进行一次分类,如按照差别率的占比大小,依次将其分成A、B、c三类,其中A类差别率在20%以上,B类数据中差别率在10%至20%之间,c类数据中的差别率则不足10%。

3.3 数据处理

在统计工作中,数据处理的主要内容就是在不予许删除的情况下,要对数据进行填补,准确处理真正需要分析的数据,然后解决当下某个经济问题。建议可以通过这几个方法来处理。

3.3.1 加权平均法

此类的方法可以运用到常见的占比问题的均值计算中来,比如学校的各项优秀评选活动,是根据每个学生的加权成绩作为衡量标准的。从学生不同方面来加权平均出最终的成绩。

3.3.2 预测法

利用教学手段对经济进行谈论的一种方法。一般就是通过教

学的模式通过曲线之间的契合度来判断经济现象的相关性,比如根据在房屋测量数据预测方面,通过测量预测,测量数据,建立回归方程,就可以得出有关的数据信息了。

3.3.3 集成化处理方法

实际数据处理过程中,会出现多个数据源进行重叠统计的现象,这就要求对数据进行挖掘之前做一个统一整理的过程,即数据的集成化处理。这种方法在统计中被广泛的应用。经过数据集成处理的统计信息真实性高,更加完整,因此,将其可作为挖掘数据的一种处理方法。

4 结束语

综上所述,在经济统计中运用数据挖掘技术,不仅有助于实现经济数据资源的有效扩展,同时也可以为判断经济局势与经济发展走向等提供真实可靠的数据支持。因此在实际运用数据挖掘技术进行经济统计分析的过程中,相关工作人员需要有意识地对获取的经济数据进行预处理,并建立起相应的决策树,在完成经济数据科学分类并确定具体调查对象的基础上,对挖掘提取出的有价值的经济数据进行深入统计分析,从而更好地为经济管理活动提供所需服务。

参考文献:

- [1]赵彬.数据挖掘技术在经济统计中的应用探索[J].南方企业家,2018(02):234.
- [2]赵舰波.数据挖掘技术在经济统计中的应用探索[J].经济研究导刊,2018, No. 362(12):193-194.
- [3]杨留华.数据挖掘技术在经济统计调查中的应用[J].全国商情·理论研究,2019,000(026):55-57.
- [4]谭国梁.数据挖掘技术在湖南省农业统计数据分析中的应用[D].2019.
- [5]张丽丽.试析数据挖掘技术在经济统计中的应用[J].中国商论,2018,000(001):13-14.