

文章编号:1007-757X(2018)06-0050-03

数据挖掘技术在高校思想政治教育中的运用

关翠玲

(陕西财经职业技术学院, 咸阳 712000)

摘要: 主要研究数据挖掘中的聚类技术在高校思想政治教育管理中的应用。通过对辅导员“工作考核量表”的数据分析,按照数据挖掘的聚类分析过程,对各种数据进行预处理,使用划分方法中的 k 均值算法,实现了数据的聚类分析,最终得到具有指导价值的结果。

关键词: 数据挖掘; 聚类算法; 思政教育; 应用研究

中图分类号: G641

文献标志码: A

Application of Data Mining Technology in College Ideological and Political Education

Guan Cuiling

(Shaanxi Vocational College of Finance and Economics, Xianyang 712000)

Abstract: In this paper, we study clustering technology of data mining, and apply to ideological and political education management. Through data analysis of "work assessment quantitative table" according to the clustering analysis of data mining process, sorts of data are processed. By using the *k*-means algorithm, the clustering analysis of data is completed. Finally, the results which have guiding value are obtained.

Key words: data mining; clustering algorithm; ideological and political education; application research

0 引言

随着我国信息化建设进程的不断推进,许多高校都已经建立起各类基于业务的数据库用于日常管理,作为应用广泛的新兴学科,数据挖掘技术在高校教育信息化中的应用前景较好,为高校的管理、建设、服务过程的绝学提供了全新而科学的分析途径。在新形势下,高校学生思政管理工作面临着巨大挑战,所以适时不断调整思想工作的途径,加强先进经验的交流,可以有效的提高高校思政工作的效果,对此,本文借助数据挖掘技术进行尝试,通过聚类结果分析,所挖掘到的信息对学生工作具有一定的参考价值。

1 数据挖掘技术在思想政治教育中的实际应用

1.1 思想政治教育管理

随着高等教育的不断发展与普及,给高校思想政治教育带来一定挑战,在通常情况下,学校相关部门会对教育管理工作进行数据收集,但是目前对这些数据的处理还处于底层的查找与简单分析阶段,不能够挖掘出其中的价值。为了更加具体的了解思政教育工作者的工作情况,学校每学期会组织学生辅导员的工作进行评议,填写辅导员“工作考核量表”如何从中提取有价值的信息,对高校思想政治教育有

非常重要的意义^[1]。

1.2 解决方案

数据挖掘属于一个方案得到肯定的过程,是数据分析研究的深层系手段,将数据挖掘技术运用到辅导员工作考核中具有特别意义。例如:通过数据挖掘技术手段分析“辅导员工作考核量表”中的数据,可以了解“某所高校思政管理整体水平”,在管理中“哪些方面做得好,哪些方面做得不到位”等相关问题。通过这些结论进一步完善高校思政教育管理。本文提出运用聚类分析的数据挖掘技术对辅导员的工作成效数据进行分析,将大批的数据转换为聚类结果,从而更好的对数据加以利用。数据挖掘过程,如图 1 所示。

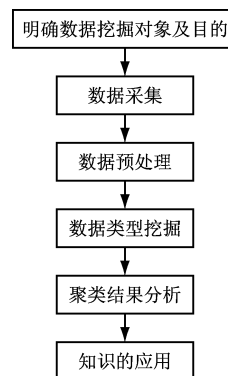


图 1 数据挖掘过程

作者简介:关翠玲(1975-),女,黑龙江双城人,满族,副教授。博士研究生。研究方向:思想政治教育,马克思主义理论。

步骤 1:明确数据挖掘的对象和主要目的,通过数据挖掘虽然不能预测最终结果,但是可以对所研究的问题进行预测,所以挖掘目标的确定是数据挖掘的关键步骤^[2]。

步骤 2:数据采集,该过程的任务比较繁重,并且需要时间比较多。在品势的教育管理中,要认真的收集数据信息,一部分数据是直接可以拿到的,一部分数据则需要通过调研才能获得。

步骤 3:数据预处理,将收集到的数据转变成可分析的数据模型,该模型是根据算法来准备的,不同的算法对数据模型的要求是不一样的。

步骤 4:数据类聚挖掘,通过类聚挖掘能够将数据模型划分为相似的多个组,该过程主要为数据模型的输入过程以及聚类算法的选择进行实现。

步骤 5:聚类结果分析,该过程主要分析研究聚类数据挖掘之后得到的多个组属性。

步骤 6:知识应用,将研究所得的信息集成到辅导员的管理教育环节中,思政工作者通过该结论促进教学管理,形成良好的管理方针^[3]。

2 数据挖掘技术在思政教育工作中具体方案实施

2.1 确定数据挖掘对象

收集并整理某大学 2017 年“辅导员工作考核量化表”,整理其中关于辅导员教育管理的 120 张考核量化表,尝试解答高校思政教育中存在的问题,经过对有价值数据的挖掘,得出结论为教学管理带来有效的指导价值。

2.2 数据采集

从学校学生工作处,搜集 2017 年度“辅导员工作考核量化表”。

2.3 数据预处理

“辅导员工作考核量化表”要求辅导员在“坚持标准,奖惩分明,客观公正的对待每一位学生。”“认真做好勤工助学活动。”“正确分析学生的思想动态”等几个指标项目中,根据辅导员的实际工作表现,划分为“优秀、良好、合格、较差、差”五等类型等级。最终获得比较完整的考核记录工作考核量化表 117 张。

2.4 数据转换

在工作考核量化表中考核等级的项目共 15 项,如何将数据合成到一个聚类分析的模式中非常关键,按照“管理态度”“管理能力”“管理方法”“管理效果”四方面属性来对工作考核量化表中的数据重新组合:其中

“管理态度”=(坚持标准+与同学之间感情融洽+言谈得体+办事客观)/4

“管理能力”=(准确掌握贫困生情况+准确掌握特殊群体+严格教育与查出违纪学生+胜任工作+组织学生做好评优工作)/5

“管理方法”=(每周 3 次以上探入班级宿舍+积极参加检查学生早操+学生奖学金发放到位+有准备的与学生谈

话+检查宿舍卫生)/5

“管理效果”=(积极参加团活班会+课下了解学生思想状况+评论与建议)/3

通过以上处理,可以将工作考核量化表关系到的十五个考评等级统一演化到四个属性中。然后针对 117 份数据样本信息的 4 个属性采取聚类挖掘的方法进行研究。通过样本预处理得到数据样本,如表 1 所示。

表 1 聚类的数据样本

管理态度	管理能力	管理方法	管理效率
0.65	0.6	0.56	0.58
0.65	0.6	0.56	0.58
0.35	0.35	0.31	0.33
0.76	0.76	0.68	0.75
0.8	0.8	0.81	0.83
...

2.5 数据聚类挖掘

数据的聚类挖掘采用划分方法中的经典算法 K 均值以及 K 中心点算法,其中 K 代表类别个数($K=3$),主要挖掘思路为:将 n 个对象划分为 K 个簇,使同一簇中的对象具有较高的相似度, K 均值算法主要是使用簇中对象的平均值作为参考值。 K 均值算法的复杂度可以通过进一步计算得出 $O(nkt)$, n 代表簇的数量, t 代表反复迭代的次数,在一般情况下, k 与 t 都会远小于 n 。针对所要分析的数据样本,四类属性都是通过数据转换而得到的,所要的数据都是算术平均值,所以产生孤立点的可能性非常小,最终选用 K 均值的算法来运用于本研究的数据聚类中。

一般情况下, K 均值算法当局部取得最优解时会终止,所以一定要对数据样本进行改进,考察数据样本信息的综合比例分布情况,采取进一步措施对 K 均值算法进行改进得到三个等级样本,代表三个等级样本数据如表 2 所示。

表 2 三个样本代表数据

	管理态度	管理能力	管理方法	管理效果
四类属性值中表示较好情况	0.75	0.75	0.75	0.75
四类属性值中表示中等情况	0.5	0.5	0.5	0.5
四类属性值中表示较差情况	0.25	0.25	0.25	0.25

3 数据挖掘算法流程

3.1 算法实现的流程

算法实现流程,如图 2 所示。

在 K 均值算法中,函数 LoadPatterns 的作用主要是将数据信息装载到程序中,目的是为了从数据库文件中读取相关信息,并且将文件中的数据转换成样本数组。函数 RunK-Means()的作用算法的主程序,将所有对象同簇中心距离进行对比,然后将对象划分到最近的簇中。函数 Show-

Centers()代表算法所描述的聚类中心。函数 ShowClusters()表示样本的标识符号^[4]。

3.2 主控程序 RunKMeans()的调用

主控程序调用结构,如图 3 所示。

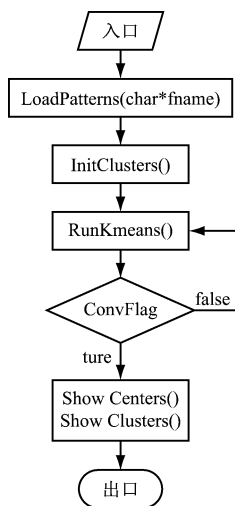


图 2 数据挖掘算法流程图

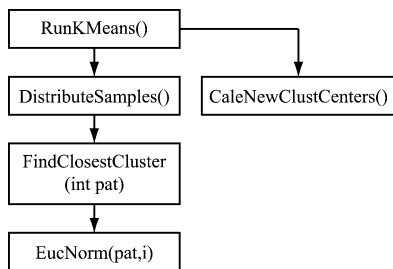


图 3 主控程序调用结构图

图 3 中函数 EucNom(pat,1)目的是算出样本和质心的平方差,从而找到最短距离的簇,然后运用 DistributeSamples()将所有对象划分到最近的簇当中,算出所有簇中对象的平均值,作为新的质心,如果所有新的质心不发生改变,则聚类结束。

4 聚类结果分析

本文运用 K 均值算法对 120 个数据通过数据转换得到的样本数据进行分析,对管理态度、管理能力、管理方法、管理效果 4 个属性进行数据挖掘聚类,设置初始 k 值为 3,最终挖掘到的结果,如表 3 所示。

表 3 簇样本数据

	管理态度	管理能力	管理方法	管理效果	含样本个数
簇 1(较好)	0.77	0.77	0.74	0.79	36
簇 2(中等)	0.61	0.57	0.54	0.56	74
簇 3(较差)	0.31	0.31	0.28	0.30	10

根据以上结果,每个簇所包括的数据样本最后的比例分布范围如下:

簇 1(较好)共计 36 个样本,删除定义样本,剩余 35 个数据样本,占 35/117=30%。

簇 2(中等)共计 74 个样本,删除一个标准样本,剩余 73 个数据样本,占 73/117=62%。

簇 3(较差)共计 10 个样本,删除一个标准样本,剩余 9 个数据样本,占 9/117=8%

“管理态度”=0.77*30%+0.61*62%+0.31*8%=0.634

“管理能力”=0.77*30%+0.57*62%+0.31*8%=0.6092

“管理方法”=0.74*30%+0.54*62%+0.28*8%=0.5792

“管理效果”=0.79*30%+0.56*62%+0.30*8%=0.6082

从总体得分由高到低排序为:管理态度、管理能力、管理效果、管理方法。总体上证明该校的思政管理水平属于中等偏上的。

4 总结

数据挖掘,主要是通过对原始数据的分析、提炼,找到最优价值的信息的过程,属于一类深层次的数据分析方法。将数据挖掘技术运用在高校思想政治教育中,有利于对思政教育工作者的多项工作指标进行分析,对其综合能力进行评定,为高校进一步完善思想政治教育管理决策,准确定位人才培养目标,加强教育团队建设提供有效的数据依据。

参考文献

- [1] 刘强璐,丁养斌.基于数据挖掘技术的高校思政教育管理研究[J].电子测试,2015(1):101-103.
- [2] 范震西,韩松洋.思想政治教育在高校内涵式发展中的重新定位[J].中共珠海市委党校珠海市行政学院学报,2015(4):50-54.
- [3] 吴小龙,张丽丽.大数据视角下高校思想政治理论教育创新[J].江西理工大学学报,2017(8):20-23.
- [4] 李平荣.大数据时代的数据挖掘技术与应用[J].重庆三峡学院学报,2014(5):159.
- [5] 舒正渝.浅谈数据挖掘技术及其应用[J].中国西部科技,2010(2):148-150.

(收稿日期:2018.02.02)