

# 大数据背景下的数据挖掘技术与应用

大数据技术能够对数据实施集优化处理与数据管理，但实际数据操作中还存在数据搜索不精准等问题，影响用户对数据的使用情况。数据挖掘技术的出现能够有效解决该问题，保证数据搜索的精准性，同时还能对数据进行必要的编辑处理。在大数据时代下，数据挖掘技术已经在各领域中有广泛应用，为推动社会进步发展做出了巨大贡献。基于此，加强对大数据时代数据挖掘技术与应用的研究具有十分现实的意义。

## 1 数据挖掘技术概述

数据挖掘主要是对随意性强、模糊性质数据的处理，能够实现对不是十分精确数据的深度挖掘。数据挖掘的过程相对复杂，且存在多种数据挖掘方法，不同挖掘方法在数据处理中相关步骤基本相同：第一，对待挖掘数据特点、形式等进行前提判断，并分析这些数据是否具有挖掘的意义；第二，确定数据衡量标准，选择恰当标准，将残余数据进行清理；第三，实施深度挖掘，最终得到数据挖掘的结果。

目前，在各领域中数据挖掘都发挥了重要作用，能够实现对行业所需数据实施精准定位，将其中实用数据进行挖掘。同时，能够对数据信息进行深度挖掘，根据数据实施市场预测，做出更加合理的判断。基于数据挖掘技术，从大数据库中提取出有用的信息，其具有一定判定功能，通过对数据中变量动态分析，客观判定数据，对根据数据情况将挖掘对象的特征描绘出来。

## 2 数据挖掘技术方法

现阶段，经过科学技术发展，数据挖掘技术方法呈现出不同类型，以便于适应更多行业发展需求。大数据作为一个海量数据资源库，利用数据挖掘技术想要实现精准分析，必须根据实际情况选择针对性的数据挖掘技术。

### 2.1 聚类分析法

聚类分析法指的是将不同类型数据进行有效的归类处理，整理零碎的数据信息，提升数据信息的条理性。聚类和分类有所区别，不是简单的将数据进行分类，在数据对象不清晰情况下，寻找数据组中具有价值意义的信息。聚类分析法缺点也十分明显，即无法准

确进行数据分类，因此目前聚类分析法仅运用于数据识别、心理学等领域中。

## **2.2 关联分析法**

根据相关理论发现任何事物之间都存在关联性，这也是数据最为基本特征，为了在海量数据库中提取有用的信息，必须能够探索出数据和数据之间的规律。而关联分析法即在这个背景下产生，利用数据之间关联性进行数据挖掘，利用诸多具有关联的数据最终找到想要的的数据，并对数据进行辨别处理，最后将其运用到行业分析中。

## **2.3 特征分析法**

在我国产业发展中，数据类型不断增多，且数据量呈现持续增多的趋势。面对这些数据量庞大、结构复杂、种类繁多的数据，需要针对其特征加以分类。这个过程中需要借助于先进的计算机设备，对数据加以虚拟分类，之后基于所需要数据的特征，对已经分类数据深度挖掘，最终得到想要的的结果。

随着科学技术发展，数据挖掘方法会越来越丰富，也将发挥更大的作用。

# **3 大数据背景下数据挖掘技术的具体应用**

现如今，数据挖掘技术已经广泛应用于不同领域中，是由于其能适应各行业，能够对市场进行精准预测，对促进企业发展，提高其生产效益具有重要意义。

## **3.1 科研领域**

在科学研究过程中，必须以大量数据为基础。可以说，数据对科研领域具有重要的作用，包括原始资料数据、实验数据等，都关系到科研工作的成败，不仅需要分析数据之间关系，同时需要对数据进行统计学分析。因此，在科研领域中可以利用数据挖掘技术，根据科研项目提供数据为基础，对所需要数据进行深度挖掘，快速罗列出所需数据，能够为科研工作人员提供必要参考。

## **3.2 电信领域**

在计算机互联网发展过程中，电信产业逐渐兴起并发展起来，成为人们日常生活中不可缺少的一部分，而电信企业发展中必须应用海量数据，才能为用户提供更加优质的服务。电信数据往往具有较强的综合性特点，一旦受到限制，则严重影响服务质量。利用数据挖掘技术，对相关数据条理化处理，更加全面的了解和记录用户信息，对服务过程中出现的各种问题。利用数据挖掘为电信企业提供精准的用户信息，对电信企业优化、高效运

行提供便利。

### **3.3 教育领域**

新时期，现代化教育工作开展中，对于学生个人发展情况、学习情况较为关注，而只有相关数据采用准确反映学生个人情况。在学校中，学生成绩、各科学学习都会形成数据，而在教育领域中使用数据挖掘技术，能更准确的分析学生各项素质，最终得到的分析结果能为教师教学、学校管理工作提供参考，对促进现代化教育教学工作发展具有重要意义。

### **3.4 在制造业领域**

经济发展下，我国人民生活水平有了质的飞跃，对产品的质量、功能都提出更高要求。在制造业领域中，产品数据、消费者需求数据等是制造业关键，通过数据挖掘技术对产品数据分类整理，并进行统计学分析，能够对比出产品的优点与缺点，根据数据结果对产品进行改良，不仅有利于提升产品生产的效率，同时也能保证产品更加符合消费者需求，为制造商创造更大的经济效益，同时也制造出更多优质的产品。可以说，制造业领域中应用数据挖掘技术，对其自身的推动作用巨大的。

### **3.5 医疗领域**

新时期，医学逐渐发达起来，但要想精准的进行疾病诊断依旧是一项难度较大的工作。通过数据挖掘技术，能在疾病诊断方面提供重要帮助，提升相关检测精准度，还能降低医务工作者的劳动量，提升疾病诊治的效率，从疾病中拯救更多患者。在长期积累下，医院数据量十分庞大，且具有处理方式繁杂、类型复杂等，传统数据处理方式并无法发挥这些数据价值。应用数据挖掘正好能解决该问题，对医生进行疾病诊断提供重要帮助。

### **3.6 在计算机安全领域**

通过数据挖掘技术，能够帮助用户有效筛选、甄别网络中存在的威胁计算机安全的软件、病毒等。一些不法分子会利用伪装，将一些病毒隐藏在软件中，而一般杀毒软件无法顺利检测出来。利用数据挖掘技术能够轻松检测出恶意软件行为模式，并对其益害进行判断，对净化网络环境，维护计算机网络用户安全具有重要帮助。

## **4 大数据时代下数据挖掘技术的发展**

发展至今，各行业都意识到数据挖掘技术在其发展中扮演的角色，也越来越重视数

据挖掘开发。基于大数据时代背景，数据挖掘技术必须有进一步发展，才能更好迎合时代发展需求。

#### **4.1 多媒体数据发展**

多媒体指的是包括文字、图片、音频、视频为一体的集合媒体形式，在计算机网络发展中呈现出更多元化形式，多媒体数据类型更多、复杂程度更甚。为了能够对这些更加多元复杂数据分析，传统数据收集、分析技术已经无法满足，需要采用新型的数据挖掘技术，对多媒体数据实施即时精准处理，探求到掩藏在数据下的真实信息。

#### **4.2 改进数据挖掘算法**

算法是数据挖掘技术中关键环节，大数据让数据搜集更加便利，但同时也标志着数据量呈现出几何倍数增加。为了能够快速处理海量数据信息，必须更精确的数据挖掘技术，必须对算法加以创新、改进，提升数据挖掘技术效率，促进其在更多领域能够发挥更大作用。

#### **4.3 与其他系统的集成**

数据挖掘实施过程中，并非只需要一种方法，很多时候需要多种计算模式共用，才能更加及时获取有用信息。在应用数据挖掘技术中，应该保持更加灵活的原则，与各系统集成发展，保证数据挖掘发挥更大的价值。