

# Application of computer artificial intelligence identification technology in the big data environment

Ya Yang

Yunnan Light and Textile Industry Vocational College, Kunming, Yunnan, 650300, China

## Abstract

In the era of big data, the application of computer artificial intelligence identification technology is rapidly changing the operation mode of all walks of life. The rapid increase in information makes the efficient processing and analysis of massive data the key to improving decision-making efficiency and innovation ability. With its powerful data processing and pattern recognition capabilities, AI recognition technology has been widely used in many fields, which has greatly promoted the intelligent transformation. Through deep learning, machine learning and other methods, artificial intelligence can extract valuable information from complex data, and can also conduct prediction and analysis based on historical data, so as to provide accurate decision support for various industries.

## Keywords

big data environment; computer; artificial intelligence identification technology; application

# 大数据环境下计算机人工智能识别技术的应用

杨亚

云南轻纺职业学院, 中国·云南昆明 650300

## 摘要

在大数据时代, 计算机人工智能识别技术的应用正迅速改变各行各业的运作方式。信息量的急剧增加, 让高效处理和解析海量数据成为提升决策效率和创新能力的关键。人工智能识别技术凭借其强大的数据处理和模式识别能力, 已经广泛应用于多个领域, 极大地推动了智能化转型。通过深度学习、机器学习等方法, 人工智能就能从复杂的数据中提取有价值的信息, 还能根据历史数据进行预测分析, 从而为各行业提供精准的决策支持。

## 关键词

大数据环境; 计算机; 人工智能识别技术; 应用

## 1 引言

在大数据时代, 信息量的爆炸式增长对各行业提出了更高的要求, 尤其是在数据处理和分析领域。计算机人工智能(AI)识别技术作为一种工具正在逐步渗透到各个应用场景中, 利用机器学习、深度学习的方法, 人工智能可以从海量数据中提取有价值的信息, 帮助各行各业做到自动化、智能化的决策支持。

## 2 大数据环境下计算机人工智能识别技术的应用优势

### 2.1 高效的数据处理能力

大数据环境下的数据量极为庞大且复杂, 数据处理方法无法在短时间内完成对数据的分析, 而计算机人工智能识别技术能结合自动化的方式对这些海量数据进行高效处理,

AI算法在大数据集上执行高速的数据挖掘与分析, 自动识别其中的模式和趋势, 极大地提高了数据分析的效率。以深度学习为例, 它通过多层神经网络对输入的数据进行分层抽象, 逐步提取数据的特征, 使机器自动识别并分类各类数据而无需人为干预, 这种高效性会让AI技术在大数据环境中具有无可替代的优势<sup>[1]</sup>。

### 2.2 精准的模式识别与预测能力

人工智能识别技术通过大规模数据训练模型, 在数据中发现潜在的关联关系与规律, 机器学习算法特别擅长处理高维度和非结构化数据如图像、语音和文本数据, 在大数据的支持下, AI能从复杂、噪声较大的数据中提取出精准的模式并对未来趋势做出科学的预测, 例如在金融领域, AI可以分析历史交易数据, 预测市场的波动趋势, 这种精准的预测能力提高了决策的科学性, 为各行业提供更可靠的支持。

### 2.3 自动化与智能化的决策支持

大数据环境下人工智能识别技术能做好高度的自动化, 帮助企业 and 组织优化决策流程, 决策过程依赖于专家经验和

【作者简介】杨亚(1979-), 男, 中国云南昆明人, 本科, 中级实验师, 从事计算机、人工智能及大数据研究。

人工分析,既耗时又容易受到人为因素的影响,AI技术通过学习优化,根据实时数据自动生成决策方案,比如在供应链管理中,AI根据市场需求、库存状况、运输情况的数据自动调整采购和库存策略,从而做到高效的资源配置和成本控制,AI还可以在制造业中结合实时监控生产线上的数据自动调整生产计划,降低资源浪费和生产停滞的风险,这种自动化和智能化的决策支持系统节省了大量的人工成本。

### 3 大数据环境下计算机人工智能识别技术的应用措施

#### 3.1 加强数据采集与管理

为了应对大数据环境下日益复杂的应用需求,需要从多个渠道和多种类型的数据源采集,通过整合来自不同领域、不同形式的数据获得更为全面的信息,这包括结构化数据,还应涵盖半结构化和非结构化数据。以医疗行业为例,除了传统的病历数据外,医生的诊断记录、影像资料、基因信息的数据都应当纳入数据采集的范围,利用这些多维度的数据来源为人工智能模型提供更多的背景信息,提升其对复杂情况的识别能力,在数据采集过程中要利用传感器、物联网设备、社交平台的数据采集工具,确保涵盖更广泛的应用场景和数据维度<sup>[2]</sup>。同时为保障采集的频率和实时性,特别是在那些要求即时决策的领域,数据的采集必须具有极高的实时性,保证能够及时获取新的数据,从而支撑人工智能技术进行实时识别与预测。因为大数据环境中的数据来源复杂,数据的准确性和一致性无法得到保证,怎样在采集阶段尽量减少数据的噪声和误差是一个非常重要的问题,为了保证数据的质量,可以采用多种方法如校验规则、数据清洗和去重技术。在数据采集的过程中要为每个数据点设计严格的验证机制,通过自动化工具和人工审核相结合的方式,尽量避免无效数据的进入,例如在电商行业中,采集用户行为数据时,只要未对数据进行有效校验,就容易出现由于网络问题、设备故障或用户操作错误导致的无效数据,这些数据未能被及时剔除,将直接影响后续模型的训练结果,因此数据采集时必须具备相应的清洗机制,让输入的数据具备较高的准确度。除了保证数据采集的全面性和质量外,在大数据环境下还必须建立高效的数据存储系统,要求在面对海量数据时,可以采用分布式存储和云计算技术将数据分布在不同的服务器或数据中心,以此提高数据存储的扩展性,数据存储系统还需要具备高可用性和灾备能力,这样就能应对自然灾害或系统故障带来的数据丢失风险,结合使用云存储和容灾备份技术来保证数据在任何情况下都不会丢失,同时还能提供更强的数据访问能力。为了提高存储效率,可以做好数据压缩或是去重,减少数据冗余和存储成本,保证存储系统的可扩展性<sup>[3]</sup>。

#### 3.2 优化人工智能模型与算法

大数据环境下,数据的复杂性让模型的训练过程面临

更高的要求。为了提升模型的性能,必须通过合理的特征选择和特征工程来提升模型的学习能力,特征工程通过从原始数据中提取有效特征,以此提高模型的预测准确性,在面对高维数据时,可以使用降维技术如主成分分析(PCA)、线性判别分析(LDA)来帮助减少输入数据的复杂度,从而提高算法的计算效率,针对不同的数据类型和应用场景,应该利用数据增强技术进一步扩展训练数据集,提升模型的泛化能力。数据增强技术通过人工合成新样本克服数据量不足或数据不均衡的问题,保证模型在真实场景中的表现更加稳定。值得注意的是,优化人工智能算法的一个重要途径是改进模型结构与算法设计,所以新的算法架构要被引入以解决大数据的挑战。深度学习作为一种强大的模型架构,近年来在大数据处理中表现出色,例如,通过剪枝技术去除网络中冗余的神经元或连接,从而简化模型,降低计算资源的消耗。与此同时,使用模型蒸馏技术,可以将一个大型深度学习模型的知识传递给一个较小的模型,从而保持其预测能力的同时,减小模型的体积和计算需求。对于某些需要实时处理的应用场景,还可以采用边缘计算技术将部分计算任务从中心服务器转移到边缘设备上,减少延迟,提升响应速度。在大数据环境下,训练神经网络这类型的复杂模型要大量的计算资源和长时间的训练周期,因此如何加速训练过程并在有限的资源下达到更好的效果,成为优化的重点,并行计算和分布式训练技术的引入能解决这一问题,将训练过程分配到多个计算节点上,可以加速训练过程,缩短模型训练时间,同时分布式计算支持处理更大规模的数据集,解决单一机器无法承载海量数据的困境。

#### 3.3 推动行业数据共享与协作

推动行业数据共享与协作的前提是政策法规的落实,在数据共享过程中,尤其是跨行业的数据共享,应该遵循数据隐私保护与信息安全的法律法规,所以要求制定完善的数据隐私保护政策以及跨行业的数据共享标准,政府和相关部门需要出台相关法律框架,对数据共享的范围、方式、数据所有权以及数据处理责任进行详细规定,保障各方在共享过程中的权益。以欧盟的《通用数据保护条例》(GDPR)为例,它为个人数据的处理与共享设立了严格的规定,特别是在数据存储、传输和访问控制方面提供了法律框架,以此提升数据共享的信任度和合法性,各个行业的监管机构也要制定针对行业特点的具体数据共享政策,避免在数据共享过程中出现信息滥用、泄露的问题,通过规范的政策和法规,为各行业的数据共享建立法律保障,降低数据共享中的风险,确保各方在共享数据时不受法律和伦理问题的困扰。因为在数据共享过程中,不同的数据来自不同的行业、企业甚至国家,为了实现数据共享,必须构建一个统一的技术架构来让不同类型、格式、标准的数据能方便地在平台之间流动,一种有效的解决方案是采用标准化的数据接口和协议,让数据源能够在保证安全的前提下快速接入共享平台,其间云计

算平台能为跨行业数据共享提供强大的计算和存储能力，尤其是在大数据量的传输和处理过程中，云计算能提供灵活、高效的资源调度，使数据共享不受地域限制，再结合大数据技术帮助处理异构数据进行数据清洗和标准化，让共享数据的质量得到提升。

### 3.4 加强计算资源和基础设施建设

在大数据时代，计算任务的规模呈现出指数级增长，因此传统的计算设备已经无法满足需求，这就要求及时投资和引入更先进的硬件设施，提升计算能力，其中高性能计算（HPC）集群的建设是一个非常重要的方向，通过部署大量的高性能计算服务器搭建大规模的计算集群，可以提升数据处理和分析的速度，满足大数据环境下对计算能力的高要求。近年来图形处理单元（GPU）在深度学习领域的应用已成为一种趋势，GPU的并行计算能力让它在处理大规模数据时的效率远远高于传统的中央处理单元（CPU），因此针对人工智能任务，尤其是深度学习模型的训练，可以结合GPU集群来加速计算过程，减少训练时间。在未来，计算需求的进一步增加，采用专用的硬件加速器如TPU（张量处理单元）等也可以进一步提升计算性能，推动人工智能识别技术的快速发展。除了计算能力的提升，还要注意做好存储资源的优化，在大数据环境下，海量数据的存储和高效管理是推动人工智能技术发展的关键，分布式存储技术的引入能解决这一问题，通过分布式存储，让数据可以在多个节点上进行存储和备份，进而提高存储系统的扩展性，同时还应该结合分布式存储提供高吞吐量和低延迟的存取速度，满足大数据环境下对存储效率的严格要求，例如Hadoop分布式文件系统（HDFS）就广泛应用于大数据处理场景中，其能将数据分散存储在多个物理节点上，并通过并行计算技术提供高速的数据读取和处理能力。在此基础上，云存储技术也逐渐成为主流选择，注意将存储资源放在云端，以此简化企业和组织在存储资源上的投入，同时能动态调整存储容量，根据数据量的变化进行弹性扩展，以便满足不同规模和类型的数据存储需求。

### 3.5 实施智能化决策支持系统

在大数据环境下，数据源的多样性要求决策支持系统必须具备强大的数据处理能力，为了实现智能化决策，必须从数据源的整合开始让系统获取到各类重要的数据，这些数据来自企业内部的生产、销售、客户反馈环节，也可以来自

外部的市场动态、竞争态势以及社会环境的情报。在收集数据时要采用大数据采集技术如传感器、物联网设备等获取实时的、精准的数据，这些数据是非结构化的。在数据采集的基础上，分布式存储能将数据切分并分布在多个节点上，通过并行计算来提高存储效率和读取速度，让海量数据快速访问，云计算平台提供了弹性存储和计算资源的调度能力，利用该技术来保障在数据量快速增长的情况下灵活扩展存储空间，保障数据的长期存储和随时调用，通过数据存储与管理来保障决策支持系统具备快速响应的能力，可以在数据量激增的情况下依然保持系统的稳定性。智能化决策支持系统的实现还需要借助人工智能算法，机器学习会对历史数据的训练，提取出数据中的潜在模式，并将这些模式应用于新的数据集上，帮助决策者做出更加精准的预测判断。为了提高决策支持系统的智能性，可以采用深度学习技术，尤其是在面对复杂的非结构化数据时如图像、文本和语音数据，深度学习自动提取数据中的特征并在此基础上做出高效的决策，例如基于自然语言处理技术，深度学习模型能分析客户反馈或社交媒体中的情感趋势，为企业的营销策略提供实时调整依据。注意优化算法和模型，智能化决策支持系统能根据新的数据和情况快速调整和更新决策策略，形成闭环反馈机制，以此保持决策过程的精准性。

## 4 结语

综上所述，大数据和计算机人工智能识别技术的发展，让很多行业经历深刻变革，这些技术提升了数据处理的效率，还为决策提供支持，从而推动智能化、自动化的进程。无论是从企业运营、公共管理，还是从个体生活的方方面面，人工智能的广泛应用都在创造新的可能。在加速技术创新的同时，如何规范应用和加强数据安全依然是必须面对的课题。将来技术的成熟与完善，会让计算机人工智能识别技术在更多领域发挥作用，为社会带来更多的智能化变革。

### 参考文献

- [1] 胡博,钱鑫.基于人工智能的计算机图像自动识别[J].电子设计工程,2024,32(22):191-195.
- [2] 陈远祝,禹谢华.基于人工智能的计算机图像识别技术发展研究[J].昭通学院学报,2024,46(05):76-82.
- [3] 杨昌尧,颜源.大数据环境下人工智能技术在计算机网络中的应用研究[J].华东科技,2022,(09):86-88.