

专题: 大数据支撑的智能应用

Big Data-Enabled Intelligent Applications

客座编辑



周斌 (1971-), 男, 博士, 国防科技大学计算机学院研究员、博士生导师, 国产基础软件工程研究中心副主任, 中国计算机学会大数据专家委员会委员, 《大数据》期刊编委。长期从事数据挖掘、分布计算、信息检索等研究工作。目前主要从事面向互联网的文本内容分析与挖掘技术、网络舆情分析与挖掘技术、大规模数据处理技术的研究工作。作为课题负责人和技术骨干主持或参与国家重点研发计划、国家自然科学基金项目等国家级项目十多项, 获得国家科学技术进步奖二等奖4项, 发表学术论文200多篇。



秦永彬 (1980-), 男, 博士, 贵州大学计算机科学与技术学院教授、院长、博士生导师, 公共大数据国家重点实验室副主任。中国计算机学会高级会员, 中国计算机学会大数据专家委员会委员, 《大数据》期刊编委。ACM中国理事会重庆分会副主席, 贵州省计算机学会副理事长兼秘书长。主要研究方向为文本计算与认知智能、大数据治理与应用。近年来聚焦政法大数据和人工智能的应用研究。目前, 主持国家自然科学基金项目1项, 贵州省自然科学基金重点项目1项。

导读

信息技术的发展使得数据采集、存储、管理等成本下降,同时也给机器学习等人工智能方法提供了足够的训练样本,使大数据成为人工智能发展的三大重要基础(数据、算法和算力)。在近年来大数据技术发展的基础上,人工智能在技术发展和落地应用等方面都获得了诸多突破。在当前大数据产业链日趋成熟的背景下,大数据与人工智能的结合也在向更全面的方向发展,跨越了棋类游戏、网页搜索、语音合成、人脸识别等单点应用,逐步发展到无人驾驶、智慧交通、智能司法、工控机器人、智能网络应用等更复杂的应用形态。本专题选取了材料科学、智能司法、内容分析等多个领域中的一些成功案例,探讨了大数据技术如何支撑人工智能中的知识表示、推理预测、智能算法、认知计算等方面和环节,达到支撑整个领域智能应用的目标。

任帅等人在《基于材料数值计算大数据的材料辐照机理发现》中,研究了材料辐照效应的数值模拟问题,挖掘分析了辐照损伤机理和性能演化规律,达到认识核材料服役性能的目的。在综述了近年来该领域相关工作的基础上,文章重点介绍了自主研发的材料辐照效应分子动力学软件和随机团簇动力学软件,以及其在国产超级计算机上的实际算例。

近年来,大数据、人工智能与司法实践的不断融合推动了智慧法院的建设发展。本期专题选取了贵州大学秦永彬老师团队的两个相关工作。黄辉等人在《基于BERT阅读理解框架的司法要素抽取方法》中,针对司法要素抽取的现状问题,提出基于BERT阅读理解框架的要素抽取方法,引入标签信息和法律先验知识构造辅助问句,利用BERT机器阅读理

解模型建立辅助问句和裁判文书之间的语义联系,并在婚姻家庭、劳动争议、借款合同3类案件上进行了验证。孙倩等人在《结合案件要素序列的罪名预测方法》中,提出了一种结合案件要素序列的罪名预测方法。该方法将案情事实过程表示为一系列以“行为”为核心、具有时序关系的案件要素序列,利用图卷积神经网络进行要素表示,融合文本语义特征来预测案件罪名。

张宝华等人在《基于多输入模型及句法结构的中文评论情感分析方法》中,针对基于规则的方法中句法规则特征的作用,提出了一种将句法规则特征融入神经网络的新模型。该模型结合MCNN、LSTM和全连接神经网络,同时将文本向量、情感词向量和语法规则向量输入独立的神经网络模型中,对模型提出的特征进行拼接,得到更加全面的文本特征。

杜雪涛在《大数据认知计算在内容安全管控中的应用》中,从通信网络中的海量垃圾和不良信息特征提取及拦截封堵需求出发,研究了如何使用基于大数据的认知计算技术代替人工进行海量数据分析和理解,帮助人们制订内容安全管控策略。针对电信诈骗治理、不良消息治理、变体消息治理和不良网站治理4个方面遇到的实际问题,分别提出了大数据认知计算的解决方案。

限于篇幅,本专题难以覆盖大数据对人工智能、智能应用的支撑作用的其他方面,也无法深入探讨各类技术方法的核心细节。希望本专题能在大数据与人工智能、智能应用的联系方面起到抛砖引玉的作用,为学术界和产业界带来一些样例和启示。