



基于模糊系统的第三代人工智能

陈德旺, 王蕊, 孔令坤, 韩泽明

引用本文:

陈德旺, 王蕊, 孔令坤, 等. 基于模糊系统的第三代人工智能[J]. *智能系统学报*, 2025, 20(5): 1071–1081.

CHEN Dewang, WANG Rui, KONG Lingkun, et al. Third-generation artificial intelligence based on fuzzy systems[J]. *CAAI Transactions on Intelligent Systems*, 2025, 20(5): 1071–1081.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202407011>

您可能感兴趣的其他文章

人工智能范式的革命与通用智能理论的创生

Paradigm revolution in artificial intelligence and the birth of general theory of intelligence

智能系统学报. 2021, 16(4): 792–800 <https://dx.doi.org/10.11992/tis.202103042>

当前人工智能技术创新特征和演进趋势

Main features and development trend in current artificial intelligence technology innovation

智能系统学报. 2020, 15(2): 409–412 <https://dx.doi.org/10.11992/tis.202001030>

因素表示的信息空间与广义概率逻辑

Factorial information space and generalized probability logic

智能系统学报. 2019, 14(5): 843–852 <https://dx.doi.org/10.11992/tis.201810021>

集对分析在人工智能中的应用与进展

Application and development of set pair analysis in artificial intelligence: a survey

智能系统学报. 2019, 14(1): 28–43 <https://dx.doi.org/10.11992/tis.201803030>

泛逻辑学理论——机制主义人工智能理论的逻辑基础

Universal logic theory: logical foundation of mechanism-based artificial intelligence theory

智能系统学报. 2018, 13(1): 19–36 <https://dx.doi.org/10.11992/tis.201711033>

机制主义人工智能理论——一种通用的人工智能理论

Mechanism-based artificial intelligence theory: a universal theory of artificial intelligence

智能系统学报. 2018, 13(1): 2–18 <https://dx.doi.org/10.11992/tis.201711032>

DOI: 10.11992/tis.202407011

网络出版地址: <https://link.cnki.net/urlid/23.1538.tp.20250123.1123.004>

基于模糊系统的第三代人工智能

陈德旺¹, 王蕊¹, 孔令坤², 韩泽明¹

(1. 福建理工大学 交通运输学院, 福建 福州 350118; 2. 福州大学 电气工程与自动化学院, 福建 福州 350108)

摘要: 人工智能经历了第一代人工智能和第二代人工智能 2 个发展阶段, 2 个阶段的人工智能分别运用以控制逻辑为核心的知识驱动和以数据学习为核心的数据驱动建构算法系统, 以模拟人类的生物智能。2 种路径各有优势, 但存在算力有限和可解释性缺陷等缺点, 第三代人工智能理论与方法正致力于发展抗噪、鲁棒且可解释的人工智能。为实现这一目标, 详细讨论了基于知识驱动和数据驱动相融合的第三代人工智能的建模方法, 并在此第三代人工智能的基础上, 探讨将模糊系统与第三代人工智能相结合, 充分利用模糊系统鲁棒性与可解释性强的优势, 推动第三代人工智能的发展, 希望对未来第三代人工智能发展具有一定借鉴意义。

关键词: 第三代人工智能; 模糊系统; 三空间融合模型; 知识驱动; 数据驱动; 通用逼近性; 模糊专家系统; 数据驱动的模糊系统辨识

中图分类号: TP39 文献标志码: A 文章编号: 1673-4785(2025)05-1071-11

中文引用格式: 陈德旺, 王蕊, 孔令坤, 等. 基于模糊系统的第三代人工智能 [J]. 智能系统学报, 2025, 20(5): 1071-1081.

英文引用格式: CHEN Dewang, WANG Rui, KONG Lingkun, et al. Third-generation artificial intelligence based on fuzzy systems[J]. CAAI transactions on intelligent systems, 2025, 20(5): 1071-1081.

Third-generation artificial intelligence based on fuzzy systems

CHEN Dewang¹, WANG Rui¹, KONG Lingkun², HAN Zeming¹

(1. School of Transportation, Fujian University of Technology, Fuzhou 350118, China; 2. College of Electrical Engineering and Automation, Fuzhou University, Fuzhou 350108, China)

Abstract: Artificial Intelligence (AI) has undergone two stages of development. The first-generation AI relied on knowledge-driven algorithmic systems focused on control logic, while the second-generation AI shifted toward data-driven systems emphasizing data learning to simulate the biological intelligence of human beings. Both approaches have their strengths. However, they face limitations such as limited computational power and issues with interpretability. To address these shortcomings, third-generation AI theories and methodologies aim to develop systems that are noise-resistant, robust, and interpretable. This paper discusses the modeling methods for third-generation AI, which are based on integrating knowledge-driven and data-driven approaches. It also examines the potential of combining third-generation AI with fuzzy systems. By leveraging the robust and interpretable characteristics of fuzzy systems, the advancement of third-generation AI can be further accelerated. The insights presented offer a valuable reference for shaping the future development of third-generation AI.

Keywords: third-generation artificial intelligence; fuzzy systems; three-space fusion model; knowledge-driven; data-driven; universal approximation; fuzzy expert systems; data-driven fuzzy system identification

自 20 世纪中叶以来, 人工智能经历的 2 个主要阶段中, 第一代人工智能产业受限于知识驱动及有限的算力并未发展起来。相比之下, 第二代人工智能产业中人工智能技术应用广泛, 但仍存

在规模有限、数据安全、算法安全等问题^[1]。这两代人工智能技术仅从不同角度模拟人类智能, 却未能创造出完全透明、可解释的系统, 也无法完全达到人类智能的深度。因此, 有必要开发新的、可解释且强健的人工智能理论和方法, 推动安全、可信、可靠且具有扩展性的人工智能技术的发展。为实现这个目标, 人工智能领域已经进入了探索第三代人工智能的时代, 第三代人工智

收稿日期: 2024-07-08. 网络出版日期: 2025-01-23.

基金项目: 福建省第三批创新之星人才计划项目 (003002); 福建省财政厅教育科研专项资金项目 (GY-Z21001); 福建理工大学科研基金项目 (GY-Z22071).

通信作者: 陈德旺. E-mail: dwchen@fjut.edu.cn.

能将会实现高系统透明度、高鲁棒性、强可解释性、低参数总数的系统框架, 强调将模糊系统作为第三代人工智能主体框架以实现可解释性人工智能系统的重要性。

第三代人工智能有很多实现方法, 本文探讨了一种将模糊系统改进为第三代人工智能的方法。基于知识驱动的模糊系统具有较好的可解释性和鲁棒性, 基于数据驱动的模糊系统可以应对复杂的任务, 因此, 两者融合之后的新系统更符合第三代人工智能的发展要求。第三代人工智能被广泛认为是人工智能领域的下一个重要发展阶段, 其具体实现方法和方向引起了广泛关注。本文以基于模糊系统的第三代人工智能为研究对象, 分析了数据驱动和知识驱动的模糊系统的数据处理能力, 最后探讨将模糊系统与第三代人工智能的时代需求相结合, 提出了基于模糊系统的第三代人工智能的具体实现框架和思路, 为发展第三代人工智能提出一个可解释性强的解决方案。本论文希望为第三代人工智能的发展提供有益启发, 鼓励更多的研究和创新, 以推动人工智能领域的进一步发展。

1 模糊系统是一种具有强可解释性的人工智能系统

自从 Zadeh 教授发表了开创性论文《Fuzzy Sets》^[2], 奠定了模糊集理论的基础, 模糊数学领域便迎来了迅速的成长。尽管模糊系统在控制学中得到了广泛的应用, 但由于其理论分析得不够严谨, 其潜力在多个领域的发挥受到了限制。为了解决这一问题, 在 1992 年, Wang^[3] 和 Kosko^[4] 提出可将模糊系统作为通用逼近器。此后, 众多学者对模糊系统的一致逼近性质进行了深入和广泛的研究, 证实了模糊系统作为一种通用的逼近工具, 能够在紧致集上以任意精确度逼近任意的非线性系统^[5-9]。

模糊系统作为一种具有高度可解释性的人工智能系统, 已在众多领域中得到了推广和应用^[10-12]。这种系统的强大之处在于其易于理解和解释的模糊规则, 允许专家或操作者以接近自然语言的方式来表述知识, 同时它也能有效地处理那些含糊不清和不确定性的问题。

1.1 模糊系统通用逼近性的研究

在人工智能领域, 模糊系统作为一种强大的工具, 近年来已被广泛证明是一种万能的函数逼近器。模糊系统的基本思想是模仿人类的知识推

理能力, 将复杂的模糊问题转化为函数问题, 并通过组合规则库中的 IF-THEN 规则, 生成输入与输出的映射关系。这种方法不仅能够逼近任意非线性函数, 还能充分利用各种有效信息, 建立误差尽可能小的非线性函数模型。

模糊系统的通用逼近理论是其核心, 这一理论认为模糊系统能够以任意精度逼近紧致集上的任意连续函数。模糊系统的通用逼近性则涉及一个问题: 模糊系统是否能够以任意所需的精度逼近紧致集上的任意连续函数。这意味着模糊系统可以作为一个强有力的工具, 用于模拟和理解复杂的系统和过程。

模糊系统的理论研究和实际应用都以其通用逼近性为基础^[13-14]。在数学上, 模糊系统实现了输入论域到输出论域的映射。它在系统建模和辨识方面的应用, 以及在控制方面的能力, 都取决于其通用逼近性。因此, 对模糊系统通用逼近理论的研究至关重要, 特别是在其非线性逼近能力、可解释性和鲁棒性方面。模糊系统的特点主要有非线性逼近能力强、可解释性好、鲁棒性强等, 其特点如图 1 所示。

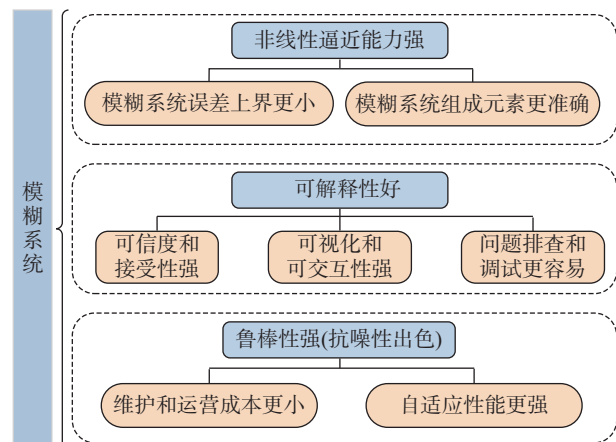


图 1 模糊系统的特点

Fig. 1 Features of the fuzzy systems

为了更好地理解模糊系统的通用逼近性, 研究者们已经进行了大量的工作^[13, 15-16]。这些研究不仅涉及模糊系统作为通用逼近器的充分和必要条件, 还包括对模糊系统逼近精度的深入分析。目前, 关于模糊系统作为通用逼近器的研究主要分为定性研究和定量研究两个方面。定性研究侧重于分析具有通用逼近性的模糊系统及其内在机制, 通常依赖于知识驱动的方法。定量研究则关注于确定模糊系统逼近误差的上界和分析逼近精度, 通常采用数据驱动的方法。这 2 种研究方法共同推动了模糊系统理论和应用的发展。这些研究证明了模糊系统具有通用逼近性的存在性, 为

模糊系统的设计和应用提供了理论基础。因此,对模糊系统及其通用逼近理论的研究在理论和实际应用上都具有重要意义。

1.2 模糊系统的逼近精度分析

为了确保模糊系统的性能,逼近精度是一个关键指标,它决定了模糊系统的误差上界,从而评估系统的优劣。在设计过程中,理解模糊系统如何逼近特定的控制或决策机制至关重要。近年来,模糊系统的发展迅速,其核心优势在于结合逼近函数和语言信息进行设计。但这一优势也带来了挑战,尤其是如何保证模糊系统达到预期的逼近精度^[17]。

Zeng 等^[18-19]通过合理设计隶属度函数来提高模糊系统的逼近精度。他们发现,当相邻隶属度函数的交点靠近中心点时,系统的逼近精度会提高;反之,精度则下降。Wang 等^[20]提出的设计模糊系统的 2 种方法包括聚类法和查表法。聚类法主要关注通用逼近问题,涉及模糊规则的设计;而查表法更适合特定应用,与通用逼近问题的关联较小。其还强调了设计方法和被逼近函数的光滑程度是影响逼近精度的 2 个关键因素。

目前,模糊系统逼近精度的研究尽管在理论、分析方法和实际应用方面已取得一定进展^[21-23],但仍然存在理论上限探索不足、收敛性分析与误差分析不够深入、对实际应用需求考虑不充分以及研究方法过于单一等关键性不足,这些缺陷限制了模糊系统性能的进一步提升和广泛应用。通常情况下,追求更高的逼近精度会导致设计的模糊系统变得更加复杂,这可能会对模糊系统的实际应用造成影响。因此,深入研究模糊系统的逼近精度问题对于提高模糊系统的性能至关重要。

2 第三代人工智能: 数据驱动和知识驱动融合

2.1 数据驱动和知识驱动的方法概述

数据驱动方法和知识驱动方法均基于人类知识的总结和扩展,并以数学理论为支撑。尽管两者均以数学为基础,它们在应用上存在显著差异。数据驱动方法主要依赖样本数据来构建经验模型,通过统计分析或人工智能技术从数据中挖掘特征。相反,知识驱动方法通常基于具体的机理模型或规则描述,其模型形式由功能和需求特点决定,更侧重于问题本质的探索和新理论的开发^[24]。

1) 数据驱动。数据驱动方法主要依赖于大量

的样本数据来构建模型,旨在从数据中自动挖掘规律和模式。其核心是通过统计分析、机器学习或人工智能技术,对大量历史数据进行细致整理和提炼,从中提取特征,从而构建经验性模型并做出决策。数据驱动方法强调实证性,基于数据的实际情况进行预测和决策。数据驱动的策略涉及广泛的数据采集和分析,旨在基于实证结果作出决策。为确保模型的准确性和适应性,需要不断地将模型输出与现实世界的数据进行对比,并利用实时数据进行持续调整和自我完善,数据驱动的流程如图 2 所示。

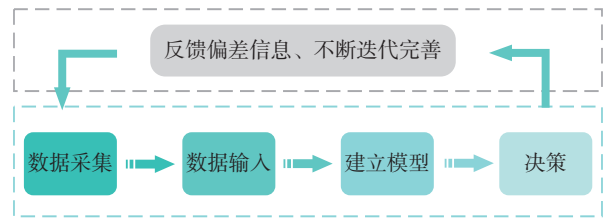


图 2 数据驱动流程

Fig. 2 Data-driven process

2) 知识驱动。知识驱动方法主要依赖于现有的领域知识和专家经验,通过对研究对象的运作机制和基本原理进行人工加工与提炼,来构建数学模型。知识驱动方法利用技术手段对原始数据进行人工提炼,包括分类、归纳和特征提取等过程,从而形成以模型和经验为基础的抽象知识。其过程通常包括知识的获取、表示、整合及应用,往往需要人的参与和判断^[25]。知识驱动方法不仅通过专家知识和制定逻辑规则来增强鲁棒性和可解释性,减少所需训练样本的规模,当面对复杂的系统,它能有效应对数据不足的问题,而且还通过知识的共享和转移,提升资源调度的决策速度^[26]。

2.2 数据和知识混合驱动的优势

知识与数据混合驱动是指,智能体的策略结合了模糊规则和强化学习输出,训练前期侧重模糊规则,训练后期侧重强化学习。在训练初期,利用专家经验确定模糊规则前件权重,以减小强化学习在训练初期由于从零试错而带来的低质量决策影响;随着训练的进行,逐步提高强化学习策略的权重,并优化规则前件的权重,不断改进专家经验。经过大规模训练后,智能体能够兼具专家经验的模糊策略和人工智能优化下的精细决策能力,使其能够灵活适应环境、高效应对挑战,并建立决策上的优势。

目前,数据驱动的人工智能算法在利用大量数据进行自监督或无监督学习时,面临的一项挑战是大多数参数拟合过程像黑箱一样不透明,这使得逻辑、规则和约束等知识难以有效融入。由

于学习过程主要依赖数据驱动, 缺乏对更广泛变化的数据的获取, 导致学习结果的鲁棒性和泛化

能力不足, 这使得算法难以适应更广泛的任务环境^[27]。知识驱动和数据驱动的优缺点如表 1 所示。

表 1 数据驱动和知识驱动的优缺点

Table 1 Advantages and disadvantages of being data-driven and knowledge-driven

驱动	优点	缺点
知识驱动	可解释性好; 自上而下; 理论支撑完备; 具有全局性	无法应对复杂模式任务; 知识凝练代价大, 周期长; 无法持续进化学习
数据驱动	在高层次模式识别下表现好; 可拓展性较强; 持续学习改进; 算法通用性较强	自下而上; 可解释性差; 对数据的质量和数量要求较高

单纯依赖现有的数据驱动人工智能技术, 虽然能够从数据中挖掘出潜在的规律, 但仍然存在一些问题。引入知识则可以有效解决这些问题: 首先它能够大幅减少模型训练所需要的数据量; 其次它能够提升模型的可解释性, 使得机器学习模型的输入、处理过程和输出对于人类来说更加易于理解; 最后, 知识的引入也有助于提高机器学习模型的可靠性和鲁棒性^[27]。

数据驱动方法和知识驱动方法各有其长短, 但将两者结合的研究途径开辟了新的方向, 为单一方法难以应对的学科领域带来了创新的解决方案。结合知识与数据驱动的方法是推动人工智能更真实地模仿人类能力的关键。人工智能的未来应该是知识驱动和数据驱动相结合的人工智能, 未来的人工智能应该是知识驱动和数据驱动相结合的。知识驱动方法依赖于领域专家的知识 and 经验, 因此具有较好的可解释性, 而数据驱动方法在识别和理解复杂、抽象模式和趋势方面具有优势^[28]。因此, 不断探索如何将数据驱动和知识驱动的优势相结合, 对于更好地服务于未来社会至关重要。

2.3 知识驱动与数据驱动的融合

数据驱动和知识驱动方法各自存在缺陷, 但它们的结合使用可以相互取长补短, 从而提升整体方法的性能。数据驱动方法以其强大的非线性表示能力和能够在离线及在线环境下学习的能力而著称, 能够在一定程度上克服知识驱动方法在处理复杂模型、缺乏精确建模以及应对环境不确定性方面的局限。而基于知识的方法可以对复杂问题进行分解降维, 或者对数据驱动方法的初始参数和学习结构进行优化, 有助于促进数据驱动方法的收敛性。

当前, 第一代人工智能以知识驱动为核心, 旨在通过算法和算力, 让机器能够模拟人类的思考过程。这一阶段的人工智能严重依赖于人类专家提供的知识, 构建规则和逻辑来处理信息; 随后, 第二代人工智能在第一代的基础上发展起来, 进

入了数据驱动的时代。这一阶段的人工智能主要通过模拟人类大脑的神经网络来工作, 利用大量的数据来训练模型, 从而实现对复杂数据的分析和预测, 但它仍面临着如何提高模型可解释性的挑战。所以仅使用现有的数据驱动型人工智能, 只能从数据中发现隐藏模式。知识的引入能够有效缓解上述数据驱动模型所面临的问题, 可以减少模型训练所需数据, 例如医学图像领域中, 数据集稀缺一直是限制该领域深度学习模型发展的因素, 此时引入医生知识可以极大程度地提高了这些领域深度学习模型的效果^[29]; 或者具有优异的控制鲁棒性和适应性, 例如可以控制融入电压/无功调控知识提升对电压越限的控制能力, 通过数据学习的方式增强优化性能^[30]。

随着人工智能技术的不断进步, 逐渐认识到仅仅依靠早期的人工智能技术, 即第一代和第二代人工智能, 已经无法满足现实世界日益增长的复杂需求。鉴于前两代人工智能各自的局限性, 张钹院士^[31]提出了第三代人工智能的构想。第三代人工智能的目标是结合第一代的知识驱动和第二代的数据驱动, 同时利用知识、数据、算法和算力这 4 个要素, 构建一个更全面、稳健且可持续的智能系统。第三代人工智能将更加注重建立可靠且可解释的理论和方法, 推动安全、可信和可扩展的人工智能技术的发展, 以期实现一个更加接近人类智能的系统。

未来推进第三代人工智能的发展, 关键在于融合知识驱动和数据驱动的理念, 通过整合知识、数据、算法和算力这 4 个关键要素, 打造出更加强大和高效的人工智能模型。为了实现这一目标, 张钹院士提出了 2 种策略: 双空间模型和单一空间模型。

1) 双空间模型。双空间模型是一种类脑模型, 旨在模拟大脑的认知与感知行为。人脑的认知与感知行为处理过程是无缝融合在一起的, 若计算机上也能实现这种融合, 或许人工智能可达到与人类相似的智能水平, 从而在根本上解决目

前人工智能系统所面临的不可解释性和鲁棒性不足的挑战。

2) 单一空间模型。单一空间模型依赖于深度学习, 将所有处理置于向量空间中, 这样做是为了充分发挥计算机硬件的高速计算能力, 从而加快处理效率。然而, 这种深度学习模型的学习机制与人类大脑的学习过程存在显著差异, 导致在可解释性和鲁棒性方面, 深度学习模型常常显示出不足。

为了推动人工智能的进一步发展并实现第三代人工智能的愿景, 张钹院士提出了一个创新的三空间融合模型, 即将双空间模型和单一空间模型相融合, 如图 3 所示^[31]。这一模型旨在结合第一代人工智能的知识驱动特性和第二代人工智能的数据驱动优势, 通过融合离散符号空间和连续向量空间, 创建一个既能够处理符号和知识, 又能有效利用大数据和深度学习的新一代人工智能系统。

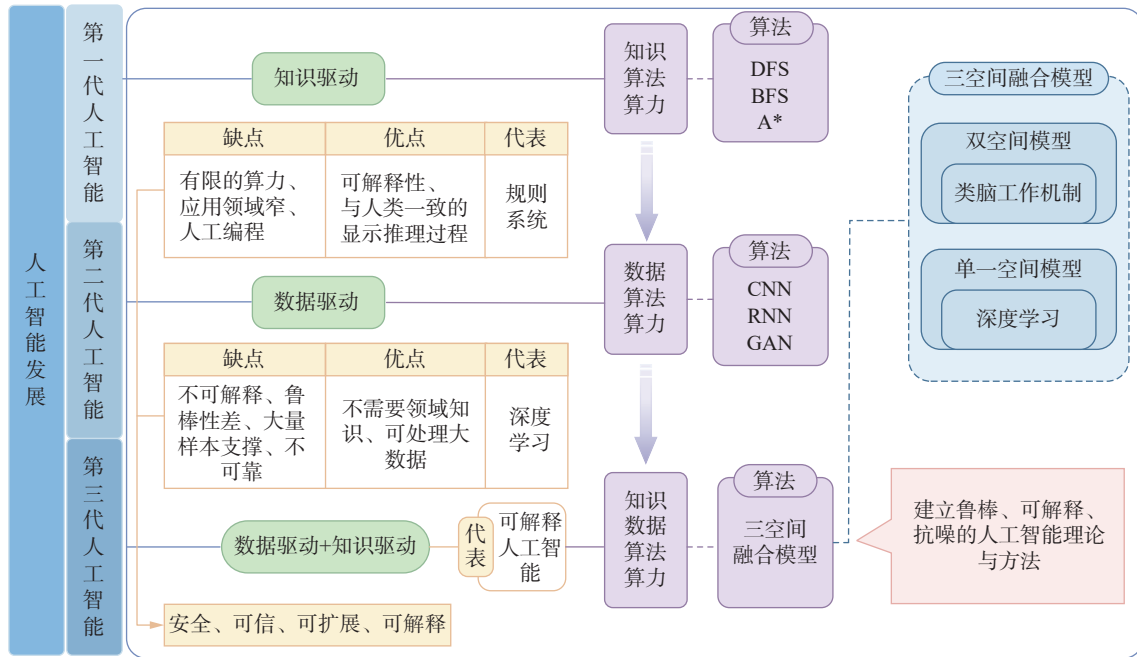


图 3 数据驱动和知识驱动融合的第三代人工智能

Fig. 3 The third generation of AI that blends data-driven and knowledge-driven

双空间模型模拟人脑的工作机制, 旨在使机器展现出类似人脑的行为, 具备更好的可解释性和鲁棒性。然而, 由于目前人们对大脑工作机制的理解仍然有限, 这一研究路径存在一定的不确定性。例如, 机器通过强化学习与环境交互所形成的“内在语义”, 是否与人类通过感知获得的“内在语义”相同, 以及机器是否能够具备意识等问题, 仍然尚无定论。相比之下, 单空间模型主要基于深度学习, 虽然计算效率高, 但也存在深度学习根本性的缺点, 即不可解释和鲁棒性差。如果单空间模型经过改进, 提升其解释性和鲁棒性, 可以朝着第三代人工智能的目标迈进。张钹院士强调, 深度学习模型虽然利用了计算机的强大计算能力, 提高了处理速度, 但其学习机制与大脑存在显著差异, 尤其在可解释性和鲁棒性方面表现不佳。因此, 三空间融合模型的出现, 不仅能够克服前两代人工智能的局限, 还能够结合两者的优点, 实现更加强大和广泛的应用。

为了实现第三代人工智能的宏伟目标, 最佳的途径是并行发展这两种模型, 即实现三空间的深度融合。这样不仅能够利用大数据和深度学习的潜力, 还能够保持对知识和符号的处理能力, 从而构建一个更加全面、稳健且可持续的人工智能系统。

3 基于模糊系统的第三代人工智能

3.1 知识驱动: 基于专家经验的模糊系统

在模糊系统建模研究的初期, 模糊规则主要源自于专家的知识 and 经验。这类模糊系统建模因此被称为基于专家经验的模糊系统建模, 也称为模糊专家系统^[32-34]。模糊专家系统利用模糊集合和模糊逻辑来处理 and 表达知识中的不确定性和不精确性。模糊专家系统的关键在于模糊推理机制, 目前, 模糊专家系统通常采用基于模糊集的合成推理方法, 这使得它们不仅能够进行模糊推理, 还具有较强的逻辑推理能力和良好的解释

性。模糊专家系统所需的知识不要求有精确的结构和量化描述,它能够处理不确定、不完整、不精确和模糊的知识,这极大地便利了领域专家^[35]。一旦将具有不确定性的专家经验以模糊知识的形式录入知识库,就可以通过模糊匹配和推理来实现快速且明确的决策。因此,即使初始信息不完整,知识体系不完备,模糊专家系统也能有效地模拟人类专家的解题思路和方法,进行知识处理,并得出较为合理的最优解。

在基于知识驱动的模糊系统建模中,有几种方法可以处理多个专家提供的模糊规则。一种方法是加权平均法,这为每个专家的规则分配一个权重,反映其可信度或专业性,然后通过对这些规则进行加权平均来生成最终的综合规则,权重的确定可以依赖于专家评估或其他可用信息。另一种方法是投票法,每个专家的规则都有一个投票,最终的规则是根据多数意见来决定的,这在权重难以确定的情况下非常有用。

对于模糊规则可信度的确定可以进行专家评估,即请领域专家根据其经验和知识对规则的准确性和重要性进行评估,从而确定其可信度。还可以采用模型拟合的方法,运用机器学习技术来拟合模型,以确定每条规则的可信度,这可以通过训练数据来估计规则的权重或权值。另外,一致性检验也是一种方法,通过检查多个专家提供的规则是否在某种程度上一致,如果规则之间的一致性较高,则可以为这些规则分配更高的可信度。这些方法可以根据具体情况选择,以确保模糊系统建模过程中规则的可信度得以合理估计。

针对模糊专家系统的冲突规则的处理的方法多种多样。一种方法是通过为模糊规则分配优先级来解决冲突,高优先级的规则将覆盖低优先级的规则,这种优先级可以根据模糊规则的重要性来确定。另一种是通过为规则分配权重来处理冲突,每个模糊规则都被赋予一个权重,然后通过这些权重的线性组合来计算最终的推断结果,这些权重可以基于规则的可信度或重要性来分配。此外,当多个规则冲突时,可以使用模糊逻辑方法(如最大值、平均值、加权平均值等)来合并规则的输出,以生成最终的决策结果。

建模过程中,当多个规则在某些条件下产生相似的推断结果的情况时,为了简化规则库提高系统的性能和效率,有几种方法可以用来合并这些相似规则。规则约简是通过删除相似规则中的一些规则来减少规则库的数量,这可以简化规则库,使其更易于管理,同时仍然保留核心规

则,确保系统性能不受影响。聚类分析可以根据规则之间的相似性度量将相似规则分组,每个组代表一个合并后的规则,这有助于构建更为简洁但功能强大的规则库。另外,特征提取从相似规则中提取共同的特征,并将它们合并为一个更一般性的规则,特征提取可以被看作是对规则间共性的抽象,但需要根据规则之间的相似性来确定哪些特征可以合并,从而创建更简洁但功能强大的规则库。模糊专家系统建模过程中的规则处理方法如图 4 所示。

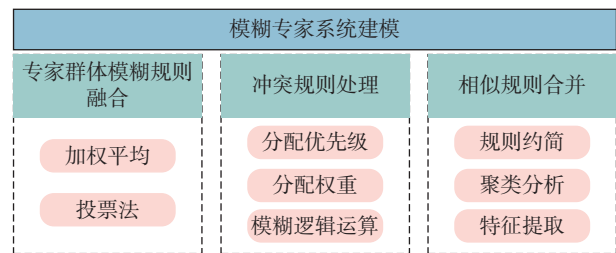


图 4 模糊专家系统建模的规则处理方法

Fig. 4 Rule-based approach to fuzzy expert systems modeling

在模糊专家系统建模过程中,引入专家知识和经验处理模糊规则是至关重要的,这确保了系统能够准确地模拟人类专家在解决问题时的思考和方法,确保系统能够充分吸取专业知识和经验。所采用的模糊规则处理方法共同构成了模糊专家系统建模的核心要素,使系统在具备推理和决策能力的同时,能够保持可解释性。

3.2 数据驱动:从大量数据中获取模糊系统

随着信息量的增长和需求的变换,以及问题的复杂性和专家知识的缺乏,传统的基于知识驱动的模糊系统建模方法可能变得困难或不切实际,因此,研究者们开始转向从海量数据中提取构建模糊系统所需的知识,这种方法被称为基于数据驱动的模糊系统辨识。数据驱动的模糊系统辨识主要依赖于输入/输出数据对样本集,特别是在模糊系统中,这意味着通过提取和优化模糊规则来模拟输入/输出之间的映射关系。模糊模型具有对复杂非线性系统的一致逼近能力和良好的可解释性,因此能够从大量数据中提取知识,深入理解系统的本质。数据驱动的方法能够通过挖掘数据中的模式和联系,协助创建更为精准、更符合现实情况的模糊系统模型。

目前,存在几种成熟的数值优化技术,它们能够从输入/输出数据对中导出模糊规则,例如自适应神经模糊推理系统(adaptive-network-based fuzzy inference system)^[36]、基于遗传算法的模糊规则优化和提取方法^[37]等。此外,聚类方法也被广泛应

用于输入/输出数据的划分和模糊规则前件的提取,例如基于目标函数最优化的 K-means 方法和模糊 C-means 划分方法,后者通过模糊隶属度来描述数据点的聚类程度,表征新的目标函数。

在模糊系统的辨识过程中,一个关键的挑战是确定适当的模糊规则数量,尤其是在处理大量高维数据时。通常情况下,缺乏足够的先验知识时,为了提高模型的准确性,会倾向于使用所有的模糊规则,但这往往导致规则冗余和模型过拟合问题的出现^[38]。随着模糊规则数量的增加,系统的可解释性会相应降低。因此确定最优的模糊规则个数变得至关重要。为了解决这个问题,一些学者采用了正交变换技术,例如基于无偏参数估计的完全最小二乘法^[39-40],来减少模糊规则的数量,特别是在考虑输入输出数据噪声时。

将多种数据驱动方法融合来提取模糊规则可以帮助提高模型的准确性和稳定性,同时可以充分利用不同方法的优势。多种数据驱动方法的融合方式有很多,例如集成方法、模型组合、特征选择和特征加权、模型融合、交叉验证和验证集等。在实际应用中,集成方法和模型组合往往是最常用的融合方法之一。

集成学习方法,包括随机森林、梯度提升树等,它能够多个基础模型的预测结果进行组合,通过投票、平均等方式来决定最终预测结果。这种方法能够减少单个方法的偏差和方差,提高模型的泛化能力。模型组合是将不同的数据驱动方法组合在一起形成一个整体模型,可以通过综合考虑各个方法的结果,得到更准确的模糊规则。例如,可以将聚类算法、关联规则挖掘和模糊聚类等方法的输出整合在一个统一的模型中。这两种方法能够克服单一方法可能存在的缺点,且可以将多种数据驱动方法的输出整合在一起,同时利用不同方法的多样性来提高模型的性能。然而,最终选择哪种方法还需根据具体问题的特点、数据的性质和可行性进行综合考虑。

处理冲突规则和合并相似规则是基于数据驱动的模糊系统辨识中的重要步骤,旨在确保提取出的模糊规则具有一定的一致性和准确性。对于多个冲突的规则,可以根据其可信度进行加权,选择可信度较高的规则作为主要规则。基于数据分析,可以为每个规则设置统计优先级,优先级可以根据它们在历史数据中的性能表现来确定,效果好的规则可能被分配更高的优先级,确保某些规则在冲突情况下具有较高的优先级。

对于相似规则的合并,常用的有3种方法。

第1种是聚类算法,将相似的规则聚类在一起,形成一组具有相似特点的规则,从而减少模型中的规则数量。第2种是规则剪枝,如果多个规则的条件部分高度相似,且某些条件对规则的贡献较小,可以考虑去除或简化这些条件,甚至删除某些规则,以减少模型的复杂度和过拟合的风险。第3种是加权合并,对于相似规则,可进行加权平均或加权融合,以生成一个更一般性的规则,能够适应更多的情况。模糊系统辨识建模过程的规则处理方法如图5所示。

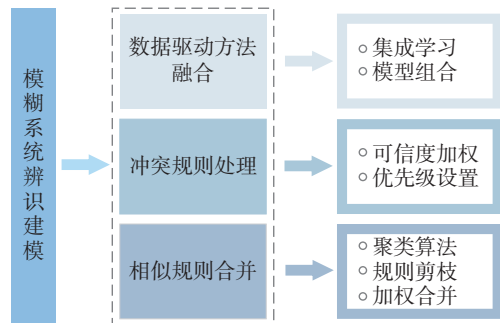


图5 模糊系统辨识建模的规则处理方法

Fig. 5 Rule processing method for fuzzy system identification modeling

基于数据驱动的模糊系统建模是一很活跃的研究内容^[41-42]。目前,基于数据驱动的模糊系统辨识已经成功应用于故障检测、图像分类、决策支持系统以及控制系统的辨识等领域。

3.3 两者融合:形成完整兼容的模糊系统

当前以深度学习为代表的人工智能方法存在鲁棒性差、可解释性差、对数据的依赖性强等基础科学问题,迫切需要研究第三代人工智能方法与技术。因此,可以利用模糊系统的可解释性好和鲁棒性强的优点,将知识驱动的模糊系统与数据驱动的模糊系统相结合,利用各自优势,为解决复杂、不确定和动态问题提供有力工具,使模糊系统更好地适应第三代人工智能的要求,实现更广泛的应用和更高水平的智能化。

当融合基于知识驱动的模糊系统和基于数据驱动的模糊系统时,融合后的模糊系统可以汇集专家经验和大量数据中提取的模糊规则。专家知识提供了人工编制的规则,通常具有较高的可信度,而数据驱动提供了大规模数据中提取的模糊规则。这种融合可以带来多方面的优势,例如,能够克服传统知识驱动模糊系统的局限性,包括专家知识库中可能存在的偏见性和不完备性。通过引入数据驱动,模糊系统可以更好地适应实际数据的变化和复杂模式任务,提高了系统的灵活性和适应性。而且融合后的模糊系统还可以利用

数据驱动的优势,实现自动化的规则生成和优化,从而减轻了专家在规则制定和维护上的负担,降低了系统的人力成本。此外,融合后的模糊系统通常具有可交互性,用户可以根据自己的需求和偏好调整模糊规则或输入数据,这进一步提高了系统的可解释性和用户参与度。

融合过程需要考虑两种模糊系统的规则的合并、消除和确保规则的兼容性,如图 6 所示。规则的合并涉及使用逻辑运算和权重分配,以便将来自不同来源的规则融为一体。这一过程中,人类专家和机器密切合作,共同决定规则的权重分布,以确保规则在整个系统中的一致性和有效性。规则的消除是为了筛选出冗余或不再有效的规则,从而精简规则集合。这个过程包括专家领域的知识与机器学习算法的相互协作,专家提供有关规则有效性的判断依据,同时机器通过分析大量数据支持决策,以帮助识别不再必要的规则。这有助于减少规则集合的复杂性,提高了系统的执行效率。规则的兼容性是确保不同规则能够协同工作的关键环节。这需要规则的转化和使用模糊逻辑,以确保不同规则之间的一致性和互操作性。通过适当的规则转化和模糊逻辑的应用,不同规则可以在统一的架下相互补充,从而提高了系统的综合性能。最后,性能评估由专家和机器共同进行,以评估整体系统的性能。

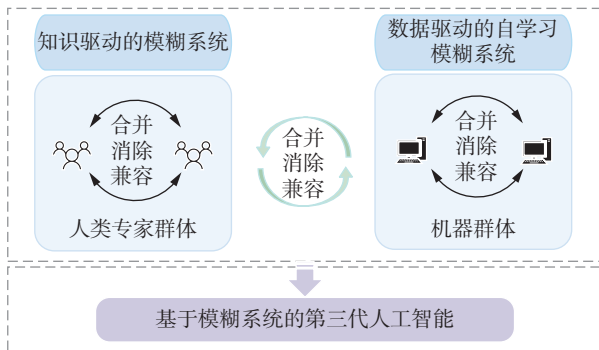


图 6 数据驱动和知识驱动融合的模糊系统

Fig. 6 Fuzzy system integrating data-driven and knowledge-driven

在融合过程中,人类专家的学习经历构成了一个基于先前知识进行逻辑推理的隐性学习过程,从而掌握新的知识。机器学习中的隐性学习过程与人类专家的学习经历有着相似之处,都强调了在没有大量监督和指导下,通过有限的知识和数据进行学习的重要性。这种能力使得机器能够在处理新问题时,不仅仅依赖于大量的数据和重新训练,而是能够利用先前的知识和经验,提高学习效率和决策的准确性。

相比之下,基于模糊系统的第三代人工智能

比知识驱动和数据驱动相融合的第三代人工智能有着明显优势。首先模糊系统在可解释性和鲁棒表现突出,在面对不完全数据或噪声数据时能够有效地处理并作出合理决策。而且即使在数据稀缺或数据质量较差的情况下仍能有效工作,减少了对大规模数据的依赖。模糊系统能够自动生成和优化规则,减轻专家的负担并提高系统的自适应能力。最重要的是,模糊系统能够更自然地整合知识和数据,结合知识驱动和数据驱动的优势,使其既能够处理规则和经验,又能够在数据驱动的环境下进行动态优化的智能系统。

融合基于专家知识和数据驱动的模糊系统,能够实现自适应学习,并在不断变化的环境和任务中改进性能,适应不断变化的环境和任务,也在智能决策制定方面发挥着重要作用。它具备理解复杂情境和问题的能力,提供高效的决策支持,符合第三代人工智能的智能决策制定理念。融合后的模糊系统可以利用专家知识和数据进行知识推理,提供有力的支持,这与第三代人工智能强调的知识推理和问题解决能力相关。在自动化领域,这个模糊系统可用于适应性控制,根据不同情境和需求调整系统行为,与第三代人工智能的智能控制理念一致。这种融合模糊系统为第三代人工智能的发展提供了更强大、更灵活的工具和方法,改善了决策制定、问题解决、自主学习等方面的能力,推动着人工智能领域的进步。

4 结论与展望

本文综合探讨了基于模糊系统的第三代人工智能的实现思路。首先,介绍了模糊系统的基本概念和重要性,以及其在人工智能领域的应用前景。其次,详细讨论了知识驱动和数据驱动两种模糊系统建模方法,并阐述了它们各自的特点和优势。知识驱动的模糊系统依赖于专家知识和经验,而数据驱动的模糊系统则通过大量数据提取知识。随后,分析了两种方法的融合,指出了融合后模糊系统的优势,包括充分利用专家知识和数据、提高系统的灵活性和适应性等。本文总结了模糊系统在第三代人工智能中的重要性,并展望了未来的研究方向和发展趋势。未来,将进一步探索模糊系统在人工智能领域的应用,促进其在智能决策、自动化控制等方面的发展,推动第三代人工智能的进步和应用。

基于模糊系统的第三代人工智能在未来的发展中将面临许多挑战和机遇。首先,随着信息技术的不断进步和应用场景的不断扩展,模糊系统

将更广泛地应用于智能决策、自动化控制、故障检测等领域。随着人工智能技术的不断发展和深入研究,模糊系统的理论基础和应用方法将得到进一步完善和优化,提高系统的性能和效率。另外,随着大数据和人工智能的融合,基于数据驱动的模糊系统将更加强大和智能化,为人类提供更准确、更可靠的智能决策支持。此外,模糊系统与其他人工智能技术的融合也将成为未来研究的重要方向,如深度学习、强化学习等,以实现更高层次的智能化应用。综上所述,基于模糊系统的第三代人工智能将在未来的发展中发挥越来越重要的作用,为人类社会的进步和发展做出更大的贡献。

参考文献:

- [1] 郑南宁. 人工智能新时代[J]. *智能科学与技术学报*, 2019(1): 1-3.
ZHENG Nanning. The new era of artificial intelligence [J]. *Chinese journal of intelligent science and technology*, 2019(1): 1-3.
- [2] ZADEH L A. Fuzzy sets[J]. *Information and control*, 1965, 8(3): 338-353.
- [3] WANG Lixin. Fuzzy systems are universal approximators[C]//IEEE International Conference on Fuzzy Systems. San Diego: IEEE, 1992: 1163-1170.
- [4] KOSKO B. Fuzzy systems as universal approximators[J]. *IEEE transactions on computers*, 1994, 43(11): 1329-1333.
- [5] 刘慧林, 冯汝鹏, 胡瑞栋, 等. 模糊系统作为通用逼近器的10年历程[J]. *控制与决策*, 2004, 19(4): 367-371.
LIU Huilin, FENG Rupeng, HU Ruidong, et al. Decennary development of fuzzy systems as universal approximators[J]. *Control and decision*, 2004, 19(4): 367-371.
- [6] CASTRO J L. Fuzzy logic controllers are universal approximators[J]. *IEEE transactions on systems, man, and cybernetics*, 1995, 25(4): 629-635.
- [7] JANG J R, SUN C T. Functional equivalence between radial basis function networks and fuzzy inference systems [J]. *IEEE transactions on neural networks*, 1993, 4(1): 156-159.
- [8] PARK J, SANDBERG I W. Universal approximation using radial-basis-function networks[J]. *Neural computation*, 1991, 3(2): 246-257.
- [9] ZENG Xiaojun, SINGH M G. Approximation theory of fuzzy systems-MIMO case[J]. *IEEE transactions on fuzzy systems*, 1995, 3(2): 219-235.
- [10] 陈德旺, 蔡际杰, 黄允浒. 面向可解释性人工智能与大数据的模糊系统发展展望[J]. *智能科学与技术学报*, 2019, 1(4): 327-334.
CHEN Dewang, CAI Jijie, HUANG Y (H /X). Development prospect of fuzzy system oriented to interpretable artificial intelligence and big data[J]. *Chinese journal of intelligent science and technology*, 2019, 1(4): 327-334.
- [11] FERNANDEZ A, HERRERA F, CORDON O, et al. Evolutionary fuzzy systems for explainable artificial intelligence: why, when, what for, and where to?[J]. *IEEE computational intelligence magazine*, 2019, 14(1): 69-81.
- [12] CAO Jin, ZHOU Ta, ZHI Shaohua, et al. Fuzzy inference system with interpretable fuzzy rules: advancing explainable artificial intelligence for disease diagnosis: a comprehensive review[J]. *Information sciences*, 2024, 662: 120212.
- [13] 刘福才, 陈超, 邵慧, 等. 模糊系统万能逼近理论研究综述[J]. *智能系统学报*, 2007, 2(1): 25-34.
LIU Fucui, CHEN Chao, SHAO Hui, et al. Researches for universal approximation of fuzzy systems: a survey[J]. *CAAI transactions on intelligent systems*, 2007, 2(1): 25-34.
- [14] SUN Changle, LI Haitao. Construction of universal approximators for multi-input single-output hierarchical fuzzy systems[J]. *IEEE transactions on fuzzy systems*, 2023, 31(12): 4170-4179.
- [15] 游文虎, 王茂, 施佳. PIE+MOM的Mamdani模糊系统通用逼近性充要条件[J]. *哈尔滨工业大学学报*, 2014, 46(11): 1-7.
YOU Wenhui, WANG Mao, SHI Jia. Sufficient and necessary condition of Mamdani fuzzy systems with PIE and MOM as universal approximators[J]. *Journal of Harbin Institute of Technology*, 2014, 46(11): 1-7.
- [16] 王宁, 谭跃, 王丹, 等. 一般齐次T-S模糊系统的逼近性能[J]. *智能系统学报*, 2010, 5(5): 436-442.
WANG Ning, TAN Yue, WANG Dan, et al. Study on approximation capabilities of general homogenous T-S fuzzy systems[J]. *CAAI transactions on intelligent systems*, 2010, 5(5): 436-442.
- [17] SHEN Hanhan, PAN Xiaodong, PENG Xiaoyu, et al. A new type of fuzzy systems in terms of vague partitions[J]. *Journal of intelligent & fuzzy systems*, 2023, 44(6): 9545-9563.
- [18] ZENG Xiaojun, SINGH M G. Approximation accuracy analysis of fuzzy systems as function approximators[J]. *IEEE transactions on fuzzy systems*, 1996, 4(1): 44-63.
- [19] ZENG Xiaojun, SINGH M G. A relationship between membership functions and approximation accuracy in fuzzy systems[J]. *IEEE transactions on systems, man, and cybernetics, part B (cybernetics)*, 1996, 26(1): 176-180.
- [20] WANG Lixin, WEI Chen. Approximation accuracy of

- some neuro-fuzzy approaches[J]. *IEEE transactions on fuzzy systems*, 2000, 8(4): 470–478.
- [21] YAN Wen, ZHAO Tao, GONG Xin. An explicit-time and explicit-accuracy control for a state-constrained nonstrict-feedback uncertain system based on adaptive fuzzy dynamic-approximation[J]. *Journal of the franklin institute*, 2023, 360(9): 6425–6462.
- [22] RAJA P, AGHILI-ASHTIANI A. G-normal fuzzy relational models are universal approximators[J]. *Fuzzy sets and systems*, 2023, 471: 108682.
- [23] TAO Yujie, SUO Chunfeng, WANG Guijun. Approximation factor of the piecewise linear functions in Mamdani fuzzy system and its realization process[J]. *Journal of intelligent & fuzzy systems*, 2021, 41(6): 6859–6873.
- [24] 李峰, 王琦, 胡健雄, 等. 数据与知识联合驱动方法研究进展及其在电力系统中应用展望[J]. *中国电机工程学报*, 2021, 41(13): 4377–4389.
- LI Feng, WANG Qi, HU Jianxiong, et al. Combined data-driven and knowledge-driven methodology research advances and its applied prospect in power systems[J]. *Proceedings of the CSEE*, 2021, 41(13): 4377–4389.
- [25] 王飞跃, 缪青海. 人工智能驱动的科学新范式: 从 AI4S 到智能科学[J]. *中国科学院院刊*, 2023, 38(4): 536–540.
- WANG Feiyue, MIAO Qinghai. *Bulletin of Chinese academy of sciences*, 2023, 38(4): 536–540.
- [26] ZHENG Ying, CHEN Haoyu, DUAN Qingyang, et al. Leveraging domain knowledge for robust deep reinforcement learning in networking[C]//*IEEE Conference on Computer Communications*. Vancouver: IEEE, 2021: 1–10.
- [27] 金哲, 张引, 吴飞, 等. 数据驱动与知识引导下人工智能算法模型[J]. *电子与信息学报*, 2023, 45(7): 2580–2594.
- JIN Zhe, ZHANG Yin, WU Fei, et al. Artificial intelligence algorithms based on data-driven and knowledge-guided models[J]. *Journal of electronics & information technology*, 2023, 45(7): 2580–2594.
- [28] 万超. 知识驱动与数据驱动系统构建综述[J]. *长江信息通信*, 2021(4): 145–147.
- WAN Chao. Summary of knowledge-driven and data-driven system construction[J]. *Changjiang information & communications*, 2021(4): 145–147.
- [29] XIE Xiaozheng, NIU Jianwei, LIU Xuefeng, et al. A survey on incorporating domain knowledge into deep learning for medical image analysis[J]. *Medical image analysis*, 2021, 69: 101985.
- [30] 杨浩, 王佳怡, 易文飞, 等. 知识-数据融合驱动的配电网光伏逆变器 电压/无功优化自适应控制[J/OL]. *中国电机工程学报*. <https://doi.org/10.13334/j.0258-8013.pcsee.241838>.
- YANG Hao, WANG Jiayi, YI Wenfei, et al. Hybrid knowledge-data driven adaptive voltage/var optimization control of photovoltaic inverters in distribution networks [J/OL]. *Proceedings of the CSEE*. <https://doi.org/10.13334/j.0258-8013.pcsee.241838>.
- [31] 张钹, 朱军, 苏航. 迈向第三代人工智能[J]. *中国科学(信息科学)*, 2020, 50(9): 1281–1302.
- ZHANG Bo, ZHU Jun, SU Hang. Toward the third generation of artificial intelligence[J]. *Scientia sinica (informationis)*, 2020, 50(9): 1281–1302.
- [32] NURMINEN J K, KARONEN O, HÄTÖNEN K. What makes expert systems survive over 10 years: empirical evaluation of several engineering applications[J]. *Expert systems with applications*, 2003, 24(2): 199–211.
- [33] KANDEL A. *Fuzzy expert systems*[M]. Boca Raton: CRC press, 1991.
- [34] ROBINSON J A. A machine-oriented logic based on the resolution principle[J]. *Journal of the ACM*, 1965, 12(1): 23–41.
- [35] ESLAMI E, BUCKLEY J J. Inverse approximate reasoning[J]. *Fuzzy sets and systems*, 1997, 87(2): 155–158.
- [36] JANG J R. ANFIS: adaptive-network-based fuzzy inference system[J]. *IEEE transactions on systems, man, and cybernetics*, 1993, 23(3): 665–685.
- [37] CORDON O, HERRERA F, GOMIDE F, et al. Ten years of genetic fuzzy systems: current framework and new trends[C]//*Proceedings Joint 9th IFSA World Congress and 20th NAFIPS International Conference*. Vancouver: IEEE, 2001: 1241–1246.
- [38] 孙富春, 罗敏楠. 模糊系统建模综述[J]. *杭州电子科技大学学报*, 2015, 35(5): 1–6.
- SUN Fuchun, LUO Minnan. A review of fuzzy system modeling[J]. *Journal of Hangzhou Dianzi University*, 2015, 35(5): 1–6.
- [39] JAKUBEK S, HAMETNER C, KEUTH N. Total least squares in fuzzy system identification: an application to an industrial engine[J]. *Engineering applications of artificial intelligence*, 2008, 21(8): 1277–1288.
- [40] YEN J, WANG L. Simplifying fuzzy rule-based models using orthogonal transformation methods[J]. *IEEE transactions on systems, man, and cybernetics Part B, Cybernetics*, 1999, 29(1): 13–24.
- [41] CORDON O, HERRERA F, DEL JESUS M J, et al. A multiobjective genetic algorithm for feature selection and granularity learning in fuzzy-rule based classification systems[C]//*Proceedings Joint 9th IFSA World Congress and 20th NAFIPS International Conference*. Vancouver:

IEEE, 2002: 1253–1258.

- [42] ISHIBUCHI H, YAMAMOTO T. Fuzzy rule selection by multi-objective genetic local search algorithms and rule evaluation measures in data mining[J]. *Fuzzy sets and systems*, 2004, 141(1): 59–88.

作者简介:



陈德旺, 教授, 博士, 电气电子工程师学会 (IEEE) 高级会员、中国自动化学会高级会员, 主要研究方向为人工智能算法、模糊系统和智能交通系统。发表学术论文 200 余篇, 出版学术专著 4 部, 出版科普专著 2 部。
E-mail: dwchen@fjut.edu.cn。



王蕊, 硕士研究生, 主要研究方向为模糊系统、列车运行时刻表优化。
E-mail: wangrui175@163.com。



孔令坤, 博士研究生, 主要研究方向为柔性生产线控制与通信优化。
E-mail: klk@126.com。

第八届中国模式识别与计算机视觉学术会议 (PRCV 2025)

第八届中国模式识别与计算机视觉学术会议 (The 8th Chinese Conference on Pattern Recognition and Computer Vision, PRCV 2025) 将于 2025 年 10 月 15 日—18 日在国家会展中心 (上海) 举办。PRCV 是国内模式识别和计算机视觉领域顶级学术盛会, 也是国际上重要且受到国际学术界认可的会议, 进入《中国计算机学会推荐国际学术会议和期刊目录》(CCF-C)。本次会议由中国图象图形学学会 (CSIG)、中国人工智能学会 (CAAI)、中国计算机学会 (CCF) 和中国自动化学会 (CAA) 联合主办, 上海交通大学承办。

PRCV 2025 旨在汇聚国内外模式识别与计算机视觉领域的学术、工业界研究人员与专业人士, 交流理论研究和发展的最新进展。会议将邀请数十位领域顶尖专家共襄盛会。多位国内外院士、IEEE Fellow 进行主旨报告和特邀演讲, 洞悉领域前沿发展。百余位国家级人才、企业专家参与的十余场专题论坛与讲习班, 涵盖学术前沿、产业应用与技术创新等内容。女科学家论坛、竞赛论坛、博士生论坛等多元舞台交相辉映, 为全球学者搭建展示卓越成果、碰撞思想火花的广阔天地。更有十余家先锋企业共襄会展盛况, 深化“产学研”融合, 助力模式识别与计算机视觉领域乘风破浪, 勇攀高峰。

联系方式

大会官网: <https://www.prcv.cn/>

联系咨询: prcv2025@163.com