

# 融合强化学习与大数据分析的动态决策支持系统优化算法

张瀚文

山西大学, 山西 太原 030006

**摘要:** 随着全球数据量持续快速增长, 动态决策支持系统在应对复杂环境下的实时决策问题中扮演着越来越重要的角色。然而, 高维度、异构性数据的激增, 以及动态环境中的不确定性, 为现有技术带来了严峻考验。在此背景下, 本文提出了一种融合强化学习与大数据分析的优化算法, 该方法有效缓解了传统强化学习在高维数据场景中的性能限制, 同时增强了决策过程的透明度和可解释性。这一改进为金融、医疗、交通等领域的智能化发展提供了重要支持, 展现出广阔的应用前景。

**关键词:** 融合强化学习; 大数据分析; 动态决策

## Optimization Algorithm of Dynamic Decision Support System Integrating Reinforcement Learning and Big Data Analysis

Zhang, Hanwen

Shanxi University, Taiyuan, Shanxi, 030006, China

**Abstract:** With the continuous and rapid growth of global data, dynamic decision support system plays an increasingly important role in dealing with real-time decision-making problems in complex environment. However, the proliferation of high-dimensional and heterogeneous data and the uncertainty in the dynamic environment have brought severe tests to the existing technology. In this context, this paper proposes an optimization algorithm that combines reinforcement learning with big data analysis. This method effectively alleviates the performance limitations of traditional reinforcement learning in high-dimensional data scenarios, and at the same time enhances the transparency and interpretability of the decision-making process. This improvement has provided important support for the intelligent development of finance, medical care, transportation and other fields, showing broad application prospects.

**Keywords:** Fusion reinforcement learning; Big data analysis; Dynamic decision-making

DOI: 10.62639/sspsstr22.20250202

随着全球数据量的爆炸式增长, 动态决策支持系统在各个行业的应用需求正在日益凸显。据统计, 2022 年全球每天产生的数据量已超过 97 泽字节 (zettabytes), 而这一数字预计将在未来几年持续攀升。在此背景下, 如何从海量数据中挖掘出关键信息, 并实时转化为优化决策的策略, 成为一个亟待解决的问题。另一方面, 强化学习技术的迅速发展为解决动态决策问题提供了新的技术支持。以深度强化学习为代表的方法在复杂环境下的表现尤为突出, 诸如 AlphaGo 等成果已经展现了这一技术的巨大潜力。然而, 单一的强化学习在面对高维度数据、实时响应和透明性需求时, 仍存在一定的局限性<sup>[1]</sup>。因此, 将强化学习与大数据分析技术相结合, 构建智能化、动态化的决策支持系统, 不仅符合当下技术发展趋势, 也契合了产业对更高效、更可靠决策工具的迫切需求。

### 一、动态决策支持系统的优化框架

#### (一) 系统架构

##### 1. 数据层

数据层是系统的基础, 主要负责收集、处理和管理多源异构数据。这一层的作用在于为后续

的决策和反馈提供可靠、结构化的信息输入。首先, 数据采集是数据层的起点, 需要整合来自不同来源的数据, 例如传感器网络、历史数据库以及实时流数据。由于数据来源的多样性, 这一过程常常面临数据格式不一致和噪声较多的问题。在这一背景下, 智能化的数据清洗机制显得尤为重要。通过对噪声值的识别和异常数据的剔除, 数据清洗能够有效提高输入数据的可靠性。接着, 预处理任务进一步对数据进行标准化和归一化等操作, 使其适应后续算法的计算需求<sup>[2]</sup>。此外, 高维数据中常常存在冗余和无效特征, 特征提取作为数据层的关键步骤, 利用降维和特征筛选技术, 将数据压缩为更加紧凑且信息密度更高的形式, 为决策层提供更高质量的输入。

##### 2. 决策层

决策层是系统的核心, 其主要功能是基于数据层提供的信息进行策略制定, 并在动态环境中不断优化决策结果。强化学习算法在这一层中发挥了重要作用, 能够在交互中迭代学习最佳策略。具体而言, 系统首先对环境状态进行建模, 利用强化学习算法生成初始策略。此时, 模型不仅考虑当前的即时回报, 还会关注长期利益, 从而制定更加全局化的行动计划。接下来的关键在于优化, 决策层依赖环境反馈对策略进行调整。例如,

(稿件编号: SSTR-25-2-1028)

**作者简介:** 张瀚文 (2002-), 男, 汉族, 籍贯: 上海市, 本科, 研究方向: 数据科学与大数据技术、人工智能及算法。

当策略表现未达到预期时,模型会通过更新参数以适应新的环境状态<sup>[3]</sup>。在多变的动态环境中,强化学习的自适应特性为系统提供了重要优势。此外,现代决策层往往结合多种学习方法,例如监督学习的预训练技术,可以缩短强化学习的探索时间,提升模型的学习效率。整个决策过程既体现了系统对环境的反应能力,也反映了其长效优化的能力。

### 3. 反馈层

反馈层是系统实现闭环优化的关键部分,其核心任务是评估决策效果并向决策层提供调整依据。与传统静态反馈不同,动态决策支持系统中的反馈层具备实时性和适应性特点。首先,反馈层监控系统输出结果,评估其与期望目标之间的差异。如果发现偏差,反馈层会识别偏差来源并将相关信息传递给决策层,用于后续的策略修正。反馈过程不仅限于单次决策的矫正,更会在多次迭代中累计经验,从而提升整体系统性能。例如,反馈层通过分析多次决策效果,能够总结出环境变化的趋势信息,并以此优化未来的决策方向<sup>[4]</sup>。此外,反馈层还可通过可视化手段直观呈现系统运行状态和决策逻辑,这一功能为系统的可信度和可解释性提供了支持。整体而言,反馈层通过动态监控和信息回馈,成为系统实现自适应优化的强大动力。

## (二) 优化算法设计

### 1. 特征提取与降维

在特征提取与降维方面,如何从海量高维数据中提炼出对决策至关重要的信息是首要问题。深度学习模型以其强大的非线性映射能力,成为了这一任务的优选工具。尤其是卷积神经网络(CNN)在图像数据的处理、循环神经网络(RNN)在时间序列数据的挖掘方面表现出色。模型的多层结构能够逐步从原始数据中提炼高层次语义信息,而这些信息通常更直接与决策变量相关。以降维为例,主成分分析(PCA)与自动编码器等技术可以压缩数据维度,降低计算复杂度,同时保留关键信息。这不仅减轻了后续算法的计算负担,还能避免数据维度过高导致的“维度灾难”问题。

数据的高维性和冗余性需要从源头上进行处理。数据预处理阶段可结合数据分布特征,将异常值过滤掉,并对数据进行标准化处理,为模型提供一致性输入。在设计优化算法时,必须明确如何评估提取特征的重要性。采用如互信息、熵增益等指标,可以量化特征与目标之间的关联度,从而更科学地选择适配的特征提取方法。这种“用什么,弃什么”的取舍逻辑直接决定了模型的输入质量,对最终性能影响深远。

### 2. 强化学习与监督学习结合

强化学习的优势在于它能够动态适应环境变化,但由于探索与利用的平衡问题,其初期收敛过程往往较为缓慢。此时引入监督学习预训练模型策略,是减少探索成本的可行途径。监督学习在已有数据的基础上,通过学习数据中的输入输

出映射关系,为强化学习的初始策略提供了一个相对优良的起点。例如,通过分类器或回归模型的预测结果,强化学习可以减少随机探索的频率,集中于那些更有可能带来高回报的策略。

在实际应用中,预训练策略的质量取决于训练数据的代表性。如果数据无法覆盖环境的关键特征,可能导致策略的局限性。在设计过程中,应充分考虑如何挑选训练数据。比如采用数据增强技术生成多样化样本,或者通过仿真环境模拟更多场景,从而提升预训练的泛化能力。此外,监督学习与强化学习的结合不应仅限于初期阶段。交替训练策略,例如先使用监督学习调整模型,再以强化学习微调,可在整个学习过程中提供持续的优化效果。通过这种“承上启下”的结合方式,算法能够兼顾效率与效果,为决策提供更高层次的保障。

### 3. 实时动态调整

在复杂动态环境中,优化算法能否实时调整策略是衡量其适应能力的重要标准。环境的变化常常是难以预测的,而实时调整则要求算法能够快速感知并响应这些变化。一个关键问题在于,如何将环境状态转化为可用信息。状态编码机制可以发挥重要作用,将复杂的环境信息映射为结构化的特征向量,以便强化学习模型直接调用。同时,算法还应具备在线学习的能力,这意味着它在面对新数据或新环境时,可以动态更新策略,而非依赖固定模型参数。

设计实时动态调整算法时,自适应成为核心考量。通过引入动态目标函数,算法能够根据实时反馈调整优化方向。例如,当环境中的某些约束条件发生变化时,目标函数权重也应随之重新分配。此外,为提高调整过程的稳健性,可以采用带有记忆机制的强化学习算法(如深度Q网络与长短期记忆网络结合),使模型能够根据历史信息更好地判断当前状态的特性。这种“前瞻性”调整使得算法不仅能够应对突发变化,还能在长期趋势中实现持续优化。

## 二、当前存在的主要问题

### (一) 数据质量差且维度过高

动态决策支持系统对数据的依赖程度很高,但现实中数据可能存在噪声、缺失值以及冗余特征。这不仅降低了算法输入的质量,还导致模型难以从高维数据中有效提取信息,增加了计算难度。

### (二) 算法响应慢且资源消耗高

实时动态调整要求算法能够快速响应环境变化,但现有算法常因复杂的训练过程而收敛速度较慢,特别是在处理高维数据时,计算资源需求极高。这在需要实时决策的场景中表现尤为明显,可能造成系统延迟。

### (三) 决策过程不透明,用户难以信任

强化学习与大数据分析的结合虽然提升了系统性能,但其决策过程往往呈现“黑箱”性质,缺乏可解释性。这会导致用户对决策结果的信任

感下降,尤其在医疗、金融等对安全性要求较高的领域可能阻碍实际应用<sup>[5]</sup>。

### 三、改进方向

#### (一) 优化数据质量与特征表达

在动态决策支持系统中,数据质量直接决定了模型的性能与稳定性,因此优化数据处理机制是提升整体效果的基础。传统的数据清洗方法往往采用规则驱动的模式,难以应对复杂且动态的数据环境。智能化的数据清洗机制可以弥补这一不足。例如,利用自监督学习技术,让系统在缺乏标签的情况下自我学习数据的内在特征,从而更精准地识别噪声和异常值。此外,对于缺失值的填补,基于上下文感知的生成模型(如变分自编码器)可以在多维数据中捕捉变量间的潜在关系,以更真实的方式补充缺失数据。冗余数据则需要采用维度的相关性分析工具进行筛选,通过衡量特征与目标变量的统计相关性,剔除对决策无关或贡献较低的特征,最终形成更加精炼和高效的数据集。

多模态数据融合技术是优化特征表达的重要方向。现代决策支持系统通常需要处理来自传感器、图像、文本和时间序列的多源异构数据,而简单的数据叠加会导致特征表达失真或冗余信息的增多。针对这一问题,可以构建深度学习驱动的特征融合框架,将不同模态的数据映射到一个共享表示空间中。这种方式不仅能够保留各数据源的关键特征,还能挖掘模态之间的协同信息。例如,图像数据和文本数据可以通过联合注意力机制提取其关联特征,为决策模型提供更加丰富的语义表达。此外,多层特征交互网络能够进一步优化这些融合特征,使其具备更强的鲁棒性和表达力,从而增强对复杂环境的适应能力。

#### (二) 提升算法效率与响应能力

算法效率一直是动态决策支持系统发展的瓶颈,特别是在实时性要求较高的场景中,计算复杂度的增加往往与资源消耗直接挂钩。轻量化模型的设计是解决这一问题的关键路径。剪枝技术可以剔除冗余神经元与权重连接,从而减小模型规模,而量化技术则通过将浮点权重映射为低比特值来降低存储需求。这两种技术结合使用,能够在保持模型性能的同时大幅缩减计算开销。此外,知识蒸馏技术也被证明是有效的优化手段之一,通过让较小的学生模型学习较大教师模型的预测分布,达到在模型规模和性能间取得平衡的目标。

实时决策需要的不仅是快速处理能力,还要求系统能够稳定应对高并发与复杂计算场景。分布式计算架构的引入可以缓解单节点资源限制的问题,让任务在多节点间并行处理,从而显著提升吞吐量。另一方面,边缘计算架构则能将部分计算任务下放至靠近数据源的边缘设备,从而降低中心服务器的计算压力并减少数据传输延迟。比如,在智能交通系统中,可以将车辆路径优化任务直接分配至路侧单元进行计算,避免因中心

节点过载而导致的响应延迟。算法设计中还应注意引入动态负载均衡机制,自动调整计算任务在资源中的分配比例,进一步保障实时响应的稳定性。

#### (三) 增强决策透明性与用户信任

增强系统的透明性不仅关乎技术性能,也直接影响用户对系统的接受程度和信任度。传统的强化学习模型由于复杂的非线性和高维性,往往被视为“黑箱”,用户难以理解其决策过程背后的逻辑。因果推断方法为这一问题提供了新的解思路。通过构建因果图模型,系统可以明确描述变量之间的因果关系,并在决策过程中解释每一步操作的因果依据。例如,在医疗决策支持系统中,因果推断可以帮助医生理解为什么推荐某种治疗方案,而非简单地输出预测结果。这种可解释性不仅提升了用户体验,也有助于模型性能的诊断与优化。

决策可视化工具是进一步加强透明性的有效手段。采用交互式图表和多维度展示,用户能够直观地了解模型在不同情境下的预测表现与行为规律。例如,在金融投资场景中,系统可以通过风险-收益曲线和模拟交易策略图,展示每项决策的潜在结果及其背后的逻辑路径。此外,设计基于规则的后验分析模块能够帮助用户回溯历史决策,将输出结果与输入特征间的关系进行解构分析。这种分析方式尤其适合需要审计和追责的场景,如法律或公共政策领域,让用户更加信赖系统的科学性与公平性。

### 四、结语

总的来看,融合强化学习与大数据分析的动态决策支持系统展现了广阔的潜力,在复杂环境下的实时决策中发挥了重要作用。从数据处理到策略优化,再到反馈机制的完善,这一系统提供了全方位的支持。然而,实际应用中依然存在一些难题,如数据质量不稳定、模型对计算资源的要求较高,决策过程不够透明。未来的研究中,需要更细致地考虑算法的性能改进,同时更贴近用户的实际需求,让系统更加直观和易用。只有在技术能力提升的同时,关注实践中的真实需求,这样的系统才能在更多领域发挥价值,助力智能化发展迈向新高度。

#### 参考文献:

- [1] 姜北. 基于大数据分析的ERP系统优化与决策支持研究[J]. 自动化应用, 2024, 65 (07): 54-56.
- [2] 王雪文, 陈艳艳, 王海君. 江苏省工业用地动态调查和智能分析技术研究[J]. 江苏科技信息, 2023, 40 (27): 77-80.
- [3] 马骋, 宋焘. 智能精准开采大数据分析决策系统关键技术[J]. 陕西煤炭, 2021, 40 (05): 57-61+84.
- [4] 赵志宏. 数字化时代商业银行零售客户动态分层——以“i世代”群体为例(下)[J]. 当代金融家, 2020, (10): 70-72.
- [5] 徐选华, 马志鹏, 陈晓红. 基于公众偏好大数据分析的大群体应急决策质量动态演化研究[J]. 中国管理科学, 2022, 30 (07): 140-149.