

刘泽润, 刘超. 可持续建成环境研究的机器学习应用进展与展望[J]. 风景园林, 2023, 30 (7): 51-59.

可持续建成环境研究的机器学习应用进展与展望

刘泽润 刘超*

摘要: 【目的】大数据、物联网和人工智能技术正在经历快速发展阶段, 其中机器学习的应用尤为瞩目, 探索机器学习对可持续建成环境研究的影响具有理论和实践价值。【方法】基于文献综述, 聚焦城市公共健康、能源碳排放、气候环境、生态系统、绿色出行 5 个可持续建成环境重要议题, 详述机器学习的概念、分类、重要算法及关键应用。【结果】提出机器学习应用预测性有余解释力不足的特点, 梳理机器学习发展从预测性到解释性的趋势, 分析机器学习应用对研究的影响。【结论】结果表明: 解释性方法和可读模型增多, 研究目的更加侧重决策解读和规律总结, 但基于实证研究的因果机制探索仍较少。基于此, 比较分析了机器学习在不同议题中的典型应用, 展望未来的发展前景。

关键词: 人工智能; 解释性机器学习; 公共健康; 能源碳排放; 气候环境; 生态系统

基金项目: 国家自然科学基金青年项目“基于人工智能的城市 PM_{2.5} 时空格局”(编号 52108060); 国家重点研发计划“十四五重点研发子课题: 形流匹配的国土空间综合效能评价”(编号 2022YFC3800804); 上海市自然科学基金项目“人工智能评估城市 PM_{2.5} 浓度的时空格局及规划响应——以上海市为例”(编号 21ZR1466500); 上海市科技支撑双碳专项“面向区域碳中和的超大城市规划设计关键技术”(编号 22DZ1207800)

中图分类号: TU984; TU986; TP181

文献标识码: A

文章编号: 1673-1530(2023)07-0051-09

DOI: 10.12409/j.fjyl.202303010108

收稿日期: 2023-03-01

修回日期: 2023-05-21

开放科学 (资源服务)
标识码 (OSID)



1 背景

1.1 大数据和物联网加快智能化进程

大数据和物联网的兴起和发展加快了智能化的进程, 为建成环境可持续发展带来了新的机遇和挑战。城市涌现海量的数据, 对建成环境的认知手段更加智能化, 实时的监测和数据收集也丰富了研究和实践的资源, 物联网也进一步将人流、物流、能量流和信息流编织到城市系统中, 使得建成环境各个维度之间的关系更加紧密^[1]。建立在智能化数据和信息基础上的新方法和分析体系蓄势待发, 以应对可持续建设中日新月异的问题, 响应新的需求和应用场景, 实现智能化的建设、改造和管理, 推动城市结构的转变, 促进建成环境的可持续发展^[2]。

1.2 可持续建成环境研究受到更多关注

当前可持续环境的特点和热点的讨论是历久弥新的问题。自第二次工业革命以来, 人类活动范围和对环境的影响逐步扩大, 随着自然科学体系的蓬勃发展, 基于物理、化学、生态、热力学等学科的综合理论推动了对环境经济和人类活动的认知和思考, 20 世纪 70 年代, 对技术发展和资源分配的反思促进

了可持续概念的诞生, 社会、自然、建成环境的“复杂巨系统”不断扩展了可持续的内涵, 涉及财富和社会公平、未来发展、生态系统保护、公众参与及文化的方方面面^[3]。进入 21 世纪, 建成环境和可持续的内涵被进一步扩展^[4], 建成环境从自然系统发展到人类系统, 以及人类的感知和历史维度, 可持续的体系也建立在文化性、社会性、生态性、政治性、经济性之上, 研究可持续建成环境成为影响人类命运和建设发展的重要问题。

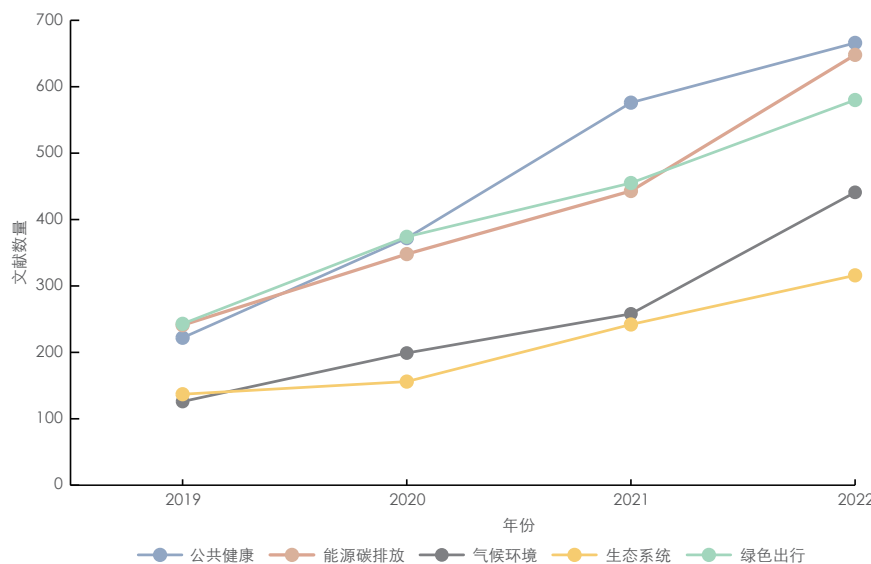
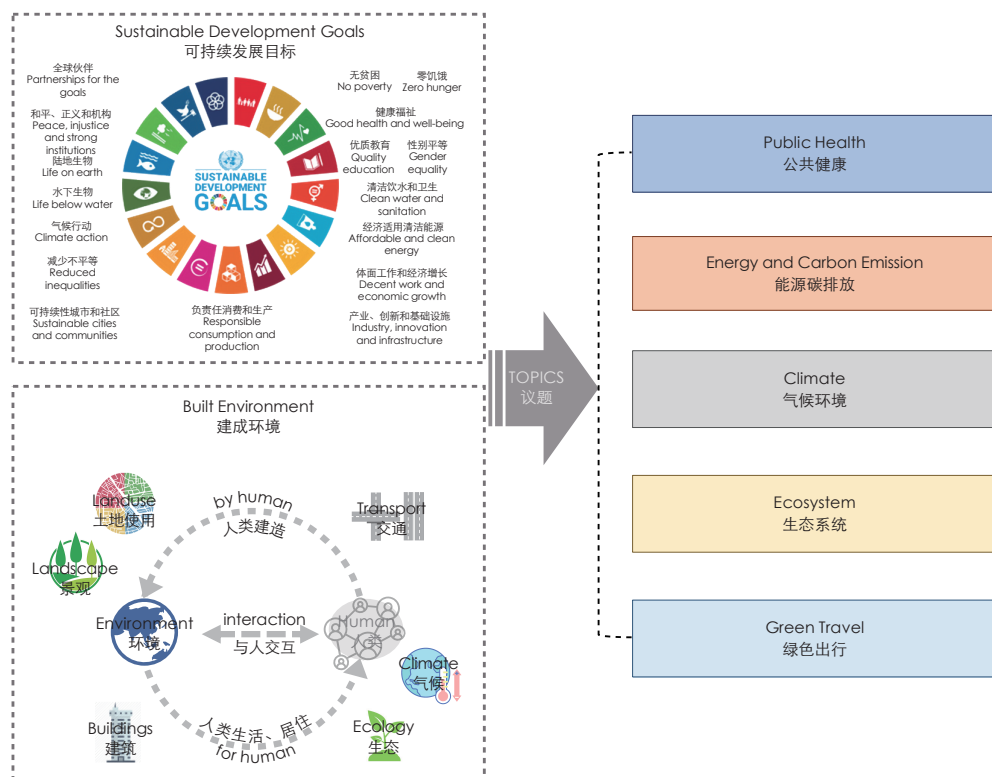
在建筑、城市规划和风景园林领域, 这一问题尤为重要。一方面, 建成环境中物质空间与建筑师、规划师和风景园林师的实践和工作紧密相关, 可持续发展是当下及未来的建设导向。另一方面, 建成环境多维度复杂化的性质对可持续提出了更多要求。不止物质空间, 空间中人的活动、意识形态和交互关系也是这一领域研究的对象, 越来越多的学者投入到相关的工作中, 致力于挖掘新的方向。因此, 探讨建成环境的可持续发展的研究议题, 分析其实践成果并展望趋势, 能引起当下建筑、城市和风景园林研究者的共鸣。

为了遴选关键和前沿的可持续建成环境

议题, 比较联合国可持续发展目标^[5]和建成环境概念, 发现它们都包含促进健康、气候行动、绿色出行、生态性景观的内容, 可以归纳为公共健康、能源碳排放、气候环境、生态系统和绿色出行这 5 个议题 (图 1)。同时, 中国知网和 Web of Science 数据统计显示, 近 5 年来, 大数据和机器学习 (machine learning, ML) 的研究在公共健康、能源碳排放、气候环境、生态系统和绿色出行这 5 个议题中增长较快 (图 2), 有必要对这几个议题开展深入研究。

1.3 机器学习给各研究领域带来新影响

作为人工智能的一个分支, 机器学习在过去几年中快速崛起, 并在不同学科领域展示了强大的能力^[2, 5-6]。数据类型和量级的增长为机器学习奠定了基础, 借由计算机硬件的升级, 越来越多的复杂算法介入不同研究主题的分析、解读和决策过程。环境的高度智能化和庞大复杂的运算量使得计算机辅助人类成为必然趋势, 而机器学习的高效性和可靠性也使得其频繁出现在制造、医学、社交、生物、互联网等各行业。对于可持续建成环境中的议题, 机器学习也依靠其能够处理



1 可持续建成环境概念及重要议题

Concept and key topics of sustainable built environment

2 重要议题 2019—2023 年文献数量变化趋势

Quantity changing trend of literature on important topics during the period from 2019 to 2023

复杂大量信息的优势，迅速生根繁荣，在实证研究和决策实践中都得到了广泛应用。其中，在智能交通、可持续能源系统等机器学习发展的热门话题下，学者对机器学习方法进行了系统的回顾^[7-10]，综述了机器学习的范式和分类，分析了机器学习方法对研究产生的影响，总结了不同方向的发展程度，为未

来研究提供参考；也有学者讨论机器学习的重要算法对研究的贡献和之后的趋势^[11-13]。虽然机器学习愈发重要，但现有的综述多关注到建成环境的某一方面或者机器学习的某类算法，定义机器学习范式、分类和优劣势，缺乏关于可持续建成环境领域的综合性论述，以及机器学习应用于城市研究的解释局限性

分析。

因此，有必要分析机器学习对可持续建成环境领域的研究带来的影响。笔者通过对机器学习相关概念的回顾，分析机器学习在可持续建成环境中的应用重点和发展趋势。进一步聚焦公共健康、能源碳排放、气候环境、生态系统和绿色出行 5 个可持续建成环境领域核心议题，梳理不同议题用到了何种机器学习方法，以及这些方法如何帮助研究者解决研究问题。虽然本研究综述的文献跨度较大，涉及不同地区和学科背景，但是本研究扎根可持续建成环境领域，紧扣机器学习应用这一核心线索，总结机器学习给可持续建成环境研究带来的影响，包括研究范式的变革、研究方向的创新、研究对象的多元和研究结论的提升。比较分析 5 个重要议题的侧重，展现不同议题在目标、方法、对象和发展程度上的差别。本研究的分析内容和结论旨在完善机器学习在可持续建成环境领域研究的应用综述，展望机器学习的方法为研究实践带来的变化，从而促进学科向更科学、更智能、更全面的方向发展。

2 从预测性到解释性：机器学习的发展

2.1 机器学习的概念、分类和方法

在统计学等的基础上，机器学习模拟实现人类的学习过程，从而让计算机能自主从经验数据中学习规律，并用于新的样本和数据中^[6, 14]。机器学习是人工智能领域最为成功的实践之一。早在 1959 年，Arthur Samuel 提出“机器学习是指赋予计算机无须明确编程即可学习的能力”的概念^[14]，1997 年，Tom Mitchell 提出了更加工程化的定义^[14]：“一个计算机程序利用经验 E 来学习任务 T，性能是 P，如果针对任务 T 的性能 P 随着经验 E 不断增长，则称之为机器学习。”从最初的简单算法开始，经历概率理论完善、算法升级、数据驱动和算力飞跃几个阶段，机器学习表现不断提高，在预测准确度和自主学习能力上具有巨大提升。在广泛的实际应用中帮助人们解决复杂问题，提升决策效率，适应新数据环境中的波动，洞察大量数据并学习认识新的规律。

表1 解释性机器学习分类及常见算法总结^[18-20]
Tab. 1 Summary of interpretable ML categories and common algorithms^[18-20]

解释性分类	常见方法		
事前可解释/内在可解释 (intrinsic interpretability)	线性回归、逻辑回归、(浅层)决策树、K近邻算法、广义加性模型、(朴素)贝叶斯和基于规则的方法		
事后可解释 (post-hoc interpretability)	全局可解释(global interpretability): 对完整数据集的整体解释	特征重要性	衡量特征的重要程度
		部分依赖图	特征对结果的边际效应
		累积局部效应图	特征平均如何影响结果贡献
	局部可解释(local interpretability): 对单个实例的本地解释	全局代理模型	简单代理模型
		局部代理模型	单个样本的简单代理模型
		个体条件期望	可看作部分依赖图在每个样本上分解
		Shapley方法/SHAP	SHAP基于Shapley计算得出,反映特征对模型的贡献

根据训练过程中数据集的监督数量和类型的不同,可以简单将机器学习分成有监督学习 (supervised learning, SL)、无监督学习 (unsupervised learning, UL)、半监督学习 (semi-supervised learning SSL) 和强化学习 (reinforcement learning, RL) 4类。有监督学习通过训练有标记的数据集 (labeled datasets) 实现算法,这也是目前机器学习领域研究最为成熟的方法,常见的有监督学习算法包括 K 近邻算法、线性回归、逻辑回归、支持向量机 (support vector machine, SVM)、决策树和随机森林以及神经网络等,这些算法能较为准确地解决回归和分类问题,例如预测公园人流量、提取并分类城市街景元素、进行城市绿视分析、分析社交网络和文本的热点等。无监督学习使用的是未经标记的数据 (unlabeled datasets), 重要的无监督学习算法有聚类算法、K 均值算法、具有噪声的基于密度的聚类方法 (density-based spatial clustering of applications with noise, DBSCAN)、分层聚类分析 (hierarchy clustering analysis, HCA)、异常检测和新颖性监测、单类支持向量机 (one-class support vector machine, OCSVM)、孤立森林、主成分分析、关联规则学习等,无监督学习善于发现数据未被人类认识的潜在规律,主要解决聚类问题,常被用于城市形态的分类和人群画像等。半监督学习可以处理部分标记的数据,大部分的半监督学习算法是无监督和有监督算法的结合,解决分类、回归和聚类问题。强化学习则强调个体与环境的交互,借由策略 (policy)、奖励 (reward)、价值 (value) 和模型 (model)

要素,进行目标导向的学习和决策过程,常见算法有自适应动态规划 (adaptive/approximate dynamic programming, ADP)、时间差分 (temporal difference, TD) 学习、Q 学习 (Q-learning) 和深度强化学习 (deep reinforcement learning, DRL) 等,典型的运用如 DeepMind 的 AlphaGO 学习围棋的制胜策略,未来能应用到城市管理、土地预测等方面。

2.2 “黑盒”模型的应用瓶颈

随着模型复杂程度和数据量的提升,现有的机器学习模型已经可以达到很高的准确度,但是,仅依靠预测的准确率无法满足实际需求的需求。因为实际问题的应用场景往往不是结果导向的,更多关注到决策的过程和背后的机制,所以,对机器学习的模型解释尤为重要,机器学习的研究导向也从提高预测性转变为优化解释性。通过对模型的解释,可以有效提高模型的可信度,降低决策过程中的偏见和偶然性,改进模型的决策机制。同时,解释的过程可以帮助研究者捕捉潜在规律,探索未知领域,解读海量数据和复杂关系,为学习新知识、建立新体系提供支持。然而,计算机语言的难读性、计算过程的复杂性、决策的不透明性使得机器学习成为“黑盒”模型,为解释机器学习提出挑战。

2.3 解释性机器学习的兴起与发展

为了更好地理解“黑盒”模型的内部工作原理,理解模型的决策过程,提高模型的透明性和可读性,机器学习的解释性方法逐步发展起来。机器学习的可解释性可以被看作对抽象概念的解释、对决策过程的解读和

对功能的理解。可解释性的目的在于将复杂的计算机模型映射到人类可理解的领域,从而实现从计算机语言到人类语言的转移,帮助人们读懂模型隐藏的步骤^[15-17]。

根据解释发生时间的不同,可以将解释性机器学习 (interpretable ML) 分为本身可解释的机器学习方法的事前可解释和本身不可解释模型的事后可解释方法^[18-20] (表 1)。前者是相对简单和可读的“白盒”模型,如线性回归、逻辑回归、(浅层)决策树、K 近邻算法、广义加性模型、(朴素)贝叶斯和基于规则的方法等。后者是在完成的“黑盒”模型基础上,进行模型内部关系的解读,也是当下研究的重点。根据解释范围的不同,事后可解释方法可以分为全局可解释和局部可解释^[18,20]。全局可解释方法关注对所有输入数据的整体理解,而局部可解释强调对特定样本的解读,能聚焦个性化的差异。目前,常用的全局可解释方法有特征重要性 (feature importance)、部分依赖图 (partial dependence plot, PDP)、累积局部效应图 (accumulated local effect plot, ALE) 和全局代理模型 (global surrogate model) 等,局部可解释方法主要有局部代理模型 (interpretable model-agnostic explanation, LIME)、个体条件期望 (individual conditional expectation, ICE) 和 SHAP 方法 (Shapley additive explanations)。

目前,各领域专家学者已对可持续建成环境研究做出了突出贡献,建立了相对完整的方法体系,并综合运用传统统计方法和一般性机器学习方法得出了富有启发性的研究成果。然而,随着数据的多样化和精细程度

的提升,一般的机器学习方法已经不能满足复杂机制的模拟需求,如何解释模型背后的原因,推演决策过程成为新的研究难题。传统的统计方法和一般性机器学习的方法无法精确描述多要素之间的复杂关系,难以完成这一复杂框架的评估。为突破方法上的瓶颈,解释性机器学习弥补了传统统计模型精确性不足和一般机器学习解释性不足的缺陷,具有更广泛应用前景。在可持续建成环境领域,亟待引入解释性机器学习的方法,用新技术、新数据来深入挖掘其中的规律和复杂机制,实现模型精度、可信度和可读性的平衡。当前解释性机器学习的探索正处于起步阶段,在健康、节能减排、气候、交通、生态等方面仍需更多实证研究支持和发展。

3 重要议题的机器学习应用

机器学习看似晦涩难懂,但其核心是通过数据对不同议题进行预测和解释。传统的研究过程是通过对理论原理的深度剖析,推导公式或算法,预测事物的运行,解释其背后的原因。然而,机器学习“暴力”地降低了这一过程的难度,通过对数据的推断,机器学习能自己明白其中的方法,将庞大的数据转化为科学知识,数据越多,算法给出的答案也就越准确。这无疑扫清了许多研究上的障碍:一方面,它给了研究者通往未知领域的“钥匙”,借助机器学习,很多过去难以用公式定义的、复杂的、多尺度的、动态的问题得以解决,效率得到极大的提升;另一方面,它打破了知识的壁垒,研究者能够通过机器学习快速入门,甚至跨学科地解决复杂问题。当然,机器学习并不是万能的,数据的体量和质量都极大影响着结果的可靠性和准确度,现阶段,研究者更多地把机器学习作为数据分析的工具,帮助快速识别关键信息,发现其中规律。在可持续建成环境领域的公共健康、能源碳排放、气候环境、生态系统和绿色出行 5 个重要议题下,笔者选取近 5 年内具有代表性的机器学习研究,选择的范围旨在能覆盖更全面的机器学习方法,包括有监督学习、无监督学习和强化学习的各类算法,力图涵盖算法应用的不同目

的,包括是预测导向或是解释导向,以及对数据处理和决策过程的辅助,并按照任务的简单到复杂、单一到多样对不同领域进行整理总结(表 2)。但机器学习的发展十分迅速,国内外文献也与日俱增,筛选难免带有一定的时效性、主观性和随机性,但笔者力求能最大程度遵循客观多样性,概括性地分析不同算法的特点、适用的数据和输出的结果,比较分析不同领域的应用侧重,帮助读者快速理解不同机器学习算法如何运用到研究中,如何支持研究结论,以及在不同议题下的发展进展。

3.1 公共健康议题有大量的解释性需求

分析国内外较为典型的建成环境与公共健康的研究,并总结相应的机器学习算法,发现对公共健康话题的讨论十分多元,涵盖传染病、慢性病、健康行为习惯、心理健康问题等不同方面,尤其是对慢性病、行为习惯、心理问题的探索不断增多,人们愈发意识到环境对居民行为的塑造作用,对身心健康的潜在影响。伴随健康数据和建成环境数据的增多,机器学习模型被广泛应用于分析建成环境对各类健康问题的影响。在大量经验数据的基础上,有监督学习算法主要应用于挖掘关键的建成环境影响要素,识别建成环境与不同健康风险之间的关系。线性回归模型因其高效、简单和实用性,出现在许多健康问题的研究中,帮助研究者快速分析显著影响因子,而逻辑回归模型多用于健康行为决策的分析。决策树和随机森林能够更有效构建多要素复杂模型,同样帮助理解关键要素,同时,通过部分依赖图的解释性机器学习展示了要素之间的复杂关系。无监督学习主要应用于健康人群划分、健康风险分类和样本地区分类等问题,K 均值、主成分分析等聚类算法频繁出现,为健康与建成环境的研究提供分类依据。强化学习主要应用于与健康有关的传感器,通过对传感器具体流程的学习,帮助传感器更好地收集数据并实现交互。总的来说,对于建成环境与健康关系的解释探索仍是当下和未来的热点,该议题的研究将面临数据更多元化和模型更复杂化的形势,现有研究对机器学习的解释度挖

掘有限,可解释机器学习的应用较少,尤其是基于临床实验的因果机制挖掘,随着各类健康问题理论的完善,亟须结合病理的可解释的机器学习算法来帮助人们梳理其中的关系和决策。

3.2 能源碳排放议题更多关注预测与决策过程

建成环境与能源方面的研究主要关注到区域尺度、城市尺度和建筑尺度,包括燃料消耗、能源消费、电网负荷等。而碳排放方面的研究,多是在国家、城市和家庭层面,综合运用多种机器学习方法,对碳排放量进行计算。基于燃料、能源消费、灯光影像、遥感、环境要素和天气等数据,多种有监督学习算法可以用来预测能源消耗量和碳排放计算,同时分析不同要素的影响程度。包括 K 均值聚类算法等在内的无监督学习主要应用于对城市进行分类、对类别划分和对特征类型的提取。此外,也有研究综合有监督学习和无监督学习算法,在集成树模型的基础上,借助 SHAP 的解释性机器学习方法,理解机器学习的决策过程。强化学习在能源和碳排放领域被应用于寻找最优节能减排的各场景中,包括但不限于生成对成本、舒适度、收益、满意度和可持续性的平衡多目标最优策略。考虑到能源和碳排放有科学详细的量化基础,结果与因素往往可通过公式推理,有迹可循,机器学习算法多用于对结果的预测,通过算法和数据的“暴力破解”,节省公式计算的成本和时间。面对更为精细的系统,可解释性方法也能帮助人们捕捉数据背后隐藏的线索。总体来说,该议题之下机器学习的预测性应用更多,解释性应用较少,但解释性方法探索已初见成效。此外,寻找发展和保护的平衡是该议题下机器学习的另一个主要应用方面,基于决策过程并结合各种深度学习算法的强化学习能有效助力这一过程,提高与能源碳排放相关决策的效率和合理性。

3.3 气候环境议题兼顾预测性与解释性

对于建成环境中气候、空气质量和热岛效应等话题的研究主要关注到时空变化和驱动力分析,在数据监测和量化体系逐步完善

表2 可持续建成环境领域典型机器学习研究应用总结^[21-60]
Tab. 2 Summary of typical ML researches and applications in the field of sustainable built environment^[21-60]

主题	作者（年份）	机器学习算法	算法类别	应用要点	文献来源
公共健康	王兰等（2020）	多元线性回归	SL	线性回归分析探索建成环境对居民自评健康的影响	[21]
	Cuesta等（2019）	决策树	SL	决策树帮助理解影响健康效果的关键要素	[22]
	张雨晴等（2022）	随机森林、多因素逻辑回归	SL	随机森林模型识别非吸烟女性肺癌的重要因素，Logistic回归判断重要暴露对非吸烟女性肺癌的风险大小	[23]
	裴泽华等（2022）	多元线性回归、随机森林、部分依赖图	SL	利用线性回归和随机森林探究HDL-C水平空间分布的环境影响因素及其指示作用	[24]
	Wattanapisit等（2022）	聚类算法、多因素逻辑回归	UL、SL	聚类算法识别健康风险类型，逻辑回归评估人口特征和健康风险类型的关系	[25]
	Sinha等（2022）	K均值聚类、主成分分析	UL	聚类分析划分研究区域	[26]
	Rehman等（2022）	强化学习的最优解	RL	强化学习优化实时数据的质量，实现传感器的交互	[27]
	Lin等（2020）	逆熵强化学习	RL	利用强化学习分析不同健康群体中基于传感数据的行为偏好	[28]
能源与碳排放	姚婷婷等（2017）	多元线性回归	SL	建立拟合燃料含碳量的多元线性模型	[29]
	霍娟等（2018）	随机森林、支持向量机	SL	预测电力负荷，分析要素影响程度	[30]
	窦嘉铭等（2022）	神经网络	SL	基于多种神经网络预测建筑能耗	[31]
	郭芳等（2021）	蒙特卡洛方法、K均值聚类	UL	利用聚类算法对样本城市的碳达峰趋势进行分类分析	[32]
	Yu等（2019）	SHAP、提升树、聚类分析	SL、UL	SHAP帮助识别重要特征，聚类提取操作类型	[33]
	Rocchetta等（2019）	强化学习、神经网络	RL	强化学习用于优化，提出电网运行的最佳决策	[34]
气候环境	侯慧等（2023）	深度Q网络	RL	利用深度Q网络算法实现用户电费-碳交易成本和满意度的平衡，生成不确定环境中实时优化管理模型	[35]
	廖了等（2022）	多元线性回归	SL	发现影响热环境的显著正负因子	[36]
	侯赛男等（2022）	支持向量机、随机森林、极度提升树	SL	利用机器学习模型实现对MODIS地表温度产品缺失像元值的最优模拟	[37]
	袁帅等（2022）	梯度提升回归树、部分依赖图	SL	量化建筑形态对地表温度的影响，识别重要因子，展示非线性关系	[38]
	夏晓圣（2022）	随机森林、广义可加模型、神经网络、特征重要性、部分依赖图	SL	多种模型预测PM _{2.5} 浓度，利用特征重要性和PDP解析重要影响要素和相关性，识别非线性响应和因子的交互	[39]
	Araujo等（2022）	神经网络	SL	神经网络模型提高预测准确度，减小预测误差	[40]
	左芝鲤等（2018）	K均值聚类	UL	根据经济、人口、交通、能源等划分长江经济带	[41]
	Vignesh等（2019）	深度强化学习、时间差分学习	RL	深度强化模型提高天气预测能力	[42]
生态系统	黄隆杨等（2019）	二元逻辑回归	SL	二元逻辑回归模型帮助量化生态用地演化驱动力	[43]
	张文强等（2021）	多元线性回归、随机森林	SL	多种模型分析气候变化和具体的人类活动对归一化植被指数的影响	[44]
	Yang等（2018）	神经网络	SL	机器学习估算生物总量，筛选重要因子	[45]
	Moros-ocha等（2022）	神经网络	SL	利用深度神经网络算法预测生物量和生态足迹	[46]
	盛硕等（2022）	K中心聚类	UL	利用聚类算法划分生态修复单元	[47]
	Sunehag等（2019）	强化学习	RL	强化学习模拟生态系统中群体和共生行为	[48]
	Wen等（2022）	随机森林	SL	利用随机森林模型估算交通工具碳排放	[49]
	Vaiyapuri等（2021）	支持向量机、随机森林、K近邻、逻辑回归	SL	比较多种算法预测出行事故严重程度的准确度，分析评估不同指标作用	[50]
	Yan等（2022）	随机森林、部分依赖图	SL	随机森林算法预测交通事故，部分依赖图的解释性方法分析影响要素作用	[51]
	Wet等（2022）	轻度提升树、SHAP	SL	集成算法模型预测风险和事故发生，解释性方法SHAP分析影响要素	[52]
	Das等（2020）	随机森林、支持向量机、极度提升树、LIME	SL	多种机器学习算法分类事故，解释性方法展示案例中关联度贡献	[53]
	饶丹等（2022）	深度神经网络、深度聚类	SL	识别航空交通流并检测出异常轨迹	[54]
	李闻等（2022）	卷积神经网络	SL	分析图像内的信息，识别交通工具	[55]
	Li等（2022）	分层聚类	UL	非监督学习算法被用于轨迹的压缩、评估和聚类分析	[56]
	Shang等（2022）	K均值聚类、自调整聚类	UL	多种聚类算法对交通状态分类	[57]
	曹欢（2022）	静态分配、深度强化学习	RL	针对形式不同场景，实现全局路径动态再分配及更新	[58]
	Yan等（2022）	深度Q学习	RL	基于图合作的深度Q学习算法，实现交通信号等之间有效合作、缓解交通拥堵、减少尾气排放的多目标平衡，生成针对不同交通场景的实时交通信号优化控制策略	[59]
	王金永等（2021）	强化学习	RL	提出了一种形式化时空同步约束的安全强化学习方法，保证系统安全同时提高复杂任务的可解释性	[60]

的当下，多元数据处理、多情景估算和影响要素分析愈发丰富，虽然传统的统计分析、空间分析和空间模型应用已比较成熟，但机器学习仍展现出优势和潜力，兼顾预测问题和解释问题。多元线性模型、支持向量机、随机森林等有监督学习被用于预测估算和要素的识别上，基于梯度提升的树模型和神经网络、实施特征重要性和部分依赖图等可解释性方法来解释因子的非线性关系和交互作用，同时神经网络模型也被用于对图像数据的处理。无监督学习的聚类算法用于评价中的分级分类过程，而深度强化学习协助监测系统完成对数据采集处理的优化，提高预测准确度。该议题下，机器学习应用更多体现在预测和模拟的环节，解释性能力发展相对

较慢：一方面气候环境的复杂性导致数据难以收集，建立解释性机器学习模型的数据基础较薄弱；另一方面，原有的气候模型与机器学习各种算法的结合尚未成熟，仍需更多的探索实践。伴随数据和监测手段的多元化，机器学习方法会越来越地参与到数据处理和预测模拟的过程中，解释应用的场景也将更加多样。

3.4 生态系统议题处于解释性初步探索阶段

实测、模型和遥感技术的发展使人们对生态系统和生物的认知更全面和详细，机器学习参与到生态演化驱动力、物种分布、生态足迹和生物量的分析中，提供预测和解释的能力支持。对于有监督学习，逻辑回归模型多用于生态演化驱动力的量化分析，线性模型、树模型和各种集成算法用于对生物分布的模拟和生物量的估算，借助神经网络也实现了对生态足迹的预测和重要因子的筛选。对于无监督学习，诸如主成分分析等算法用于特征的提取和综合指标体系的建立，K 中心聚类、K 均值聚类等方法用于对生态单元的划分。而强化学习更多用在生物行为模型、生物动力分析等方面，通过对动态数据的挖掘，模拟实时交互过程。目前，预测估算和模拟工作是该议题下的主流，解释性方面，驱动力分析的量化较为简单，对模型的解读也比较初步地停留在重要因子的识别上，仍需更多的探索挖掘。

3.5 绿色出行议题发展成熟且应用前沿

机器学习在有关绿色出行的研究应用上积累了丰硕的成果，在数据处理、预测、模拟、决策、解释方向都有前沿性的探索和突破，涵盖的话题包括但不限于出行行为、出行安全、交通流、出行工具、拥堵风险、道路网络等的研究。对于有监督学习，各类模型应用于对指标影响的分析和结果的预测估算，在集成树模型的基础上，特征重要性、部分依赖图、SHAP、LIME 等全局和局部可解释方法被用来理解要素的贡献度和相互关联作用，并对具体实例做出解读，更为复杂的神经网络模型用于异常的识别和图像信息的提取，帮助理解决策过程、提高预测精度，同时基于神经网络建立的更为透明的算

法被应用于决策解读和因子识别中。而无监督学习算法被用于压缩、评估和分类的多元问题中，兼容对静态和动态数据的挖掘。强化学习的应用场景也十分广泛，利用丰富的动态数据，路径规划、信号灯控制、智能交通仿真、无人驾驶安全的任务可以被高效实时处理，不断趋近交通的可持续性理想目标，而约束法则的架构也提升了复杂任务的解释度，使得决策生成更为透明。不难推测，对可持续交通问题的研究是过去、当下和未来的热点，面向未来的多样场景，机器学习的应用仍有更大的潜力和更多可能的探索方向，机器学习本身已成为可持续智能交通体系不可或缺的一部分，二者是相互依存共同促进的关系，预测与解释将会更加精细和深入，为更为智能的决策提供基础。

4 讨论

总的来说，机器学习的发展和应用对可持续建成环境的研究方法、研究方向、研究对象和研究结论均产生了重要影响。机器学习解释能力的提高显著促进了研究方法的变革，越来越多的研究者实践基于机器学习的研究框架，应用各种机器学习模型，发展具有专业特色的可操作方法路径。机器学习日渐强大的性能极大拓宽了研究方向和研究对象，机器学习擅长处理复杂的、多元的、动态的、大规模的数据，以往被认为难以解决的复杂问题成为机器学习的破译对象。这为解答多地区、多尺度、长时间跨度的动态复合问题提供了新视角和可能。同时，机器学习自身不断向着智能化、透明化、可解释化发展，通过对以往经验和人类思路的学习，机器学习能帮助研究者得出最优决策方案，揭示决策机制，发现潜在规律，研究结论也愈加科学、动态、智能。

对于可持续建成环境的各重要议题，研究者都注意到机器学习在预测和解释上的强大能力，并开发了基于具体场景和问题的解释性方法和可读模型，研究目的也逐渐侧重于决策的理解和规律的挖掘，但不同议题的发展程度不同，同时基于实证研究的因果机制探索仍有大量空白。对于公共健康，解释

建成环境对健康效应的影响是核心论点，但解释性机器学习的应用相对简单初步，解释性方法的实证探索较少，缺乏对病理的解释和临床实验，需要更多理论的完善和方法的探索，以理解其中的复杂因果机制。而机器学习在能源碳排放、气候环境领域的应用通常兼顾预测和解释的问题，也在一定程度上展开对潜在规律的讨论：一方面，通过对数据处理的升级和算法的综合，提高估算预测的准确度；另一方面，结合算法的特点和对原理解读，模型的决策过程也得以解释。对于建成环境与生态系统，机器学习的预测应用更多，解释力相对不足，由于分析过程中涉及的要害庞杂，模型的设计与实施仍需完善。对于绿色出行问题，机器学习应用已有成熟的发展，兼具前沿性和多样性，在预测能力和解释能力发展的同时，不断转向对智能决策的探索，在分析解释的基础上寻找改进优化策略。

5 结论

本研究从现有的研究应用出发，提出机器学习预测性有余而解释力不足的特点，总结各类机器学习算法对研究的影响。有监督学习被用于数据较全的回归和分类问题上：线性回归应用最广泛，适用于简单的预测；集成算法发展迅速，随机森林、集成的树模型适用于复杂的预测任务，并结合多种可解释性方法，揭示内在规律；神经网络也日趋复杂，但目前可读性不高。无监督学习被用于陌生数据或缺失数据的聚类和降维问题上，且常与有监督学习结合应用：K 均值聚类是最常用的聚类算法，能快速实现数据的分组；主成分分析是较为常用的降维方式，能高效凝练关键信息。强化学习被用于优化和动态规划问题上，且往往是经过有监督学习或无监督学习的数据处理后的最后一步，算法通常会结合具体问题而调整，有很强的针对性。机器学习的理论和方法日新月异，本研究仅提及最为主流且应用广泛的方法，为学者了解相关研究和涉足该领域实践提供初步的图景。

未来，随着数据和模型的发展日趋完善，机器学习应用的资源和场景也更丰富多元，

可持续建成环境的研究将更多依赖并受益于机器学习的发展。不同议题的发展程度不同,将来可能有不同的侧重。绿色出行等交通问题的机器学习应用更为前沿,未来将持续探索从预测、解释到智能决策的完整体系,更多关注动态数据,更多开发人机互动场景;对于能源、碳排放和气候环境,解释性的应用将不断增多,方法将会更加多元化和专业化;而生态系统和公共健康的机器学习应用将进一步普及,此外,这2个议题也存在大量的解释性需求,可解释的方法和模型亟须大量实证探索,相关研究将成为下一步的热点。当然,本研究仅在文献综述的基础上关注重点议题,对较为典型的应用进行了总结分析,并未对机器学习方法本身的原理、技术和道德问题做更深入的探索。总之,机器学习对可持续建成环境的研究方法、研究对象和研究结论带来的影响和变革毋庸置疑,有效推动了该领域更定量、更多元、更智能、更科学的发展,可以预见,这一领域的机器学习研究应用正在且将不断快速发展,且具备长久广阔的发展前景。

参考文献 (References):

- [1] 龙瀛.颠覆性技术驱动下的未来人居:来自新城市科学和未来城市等视角[J].建筑学报, 2020 (Z1): 34-40.
LONG Y. Future of Human Habitats Driven by Disruptive Technologies: Perspectives from the New Science of Cities and Future Cities[J]. Architectural Journal, 2020 (Z1): 34-40.
- [2] 鲍毅, 黄舟, 郭庆华, 等.城市建成环境存量的空间计算:进展及展望[J].遥感学报, 2022, 26 (10): 1909-1919.
BAO Y, HUANG Z, GUO Q H, et al. Spatial Calculation of Urban Built Environment Stock: Progress and Prospects[J]. National Remote Sensing Bulletin, 2022, 26 (10): 1909-1919.
- [3] 牛文元.可持续发展理论的内涵认知:纪念联合国里约环发大会 20 周年[J].中国人口·资源与环境, 2012, 22 (5): 9-14.
NIU W Y. The Theoretical Connection of Sustainable Development: the 20th Anniversary of UN Conference on Environment and Development in Rio De Janeiro, Brazil[J]. China Population, Resources and Environment, 2012, 22 (5): 9-14.
- [4] 崔桂彬.可持续发展理念下城市园林景观设计探讨[J].工程建设与设计, 2022 (16): 56-58.
- [5] 陈凯, 朱钰.机器学习及其相关算法综述[J].统计与信息

- 论坛, 2007 (5): 105-112.
- CHEN K, ZHU Y. A Summary of Machine Learning and Related Algorithms[J]. Journal of Statistics and Information, 2007 (5): 105-112.
- [6] 何清, 李宁, 罗文娟, 等.大数据下的机器学习算法综述[J].模式识别与人工智能, 2014, 27 (4): 327-336.
HE Q, LI N, LUO W J, et al. A Survey of Machine Learning Algorithms for Big Data[J]. Pattern Recognition and Artificial Intelligence, 2014, 27 (4): 327-336.
- [7] DE SOUZA J T, DE FRANCISCO A C, PIEKARSKI C M, et al. Data Mining and Machine Learning in the Context of Sustainable Evaluation: A Literature Review[J]. IEEE Latin America Transactions, 2019, 17 (3): 372-382.
- [8] DE LA TORRE R, CORLU C G, FAULIN J, et al. Simulation, Optimization, and Machine Learning in Sustainable Transportation Systems: Models and Applications[J]. Sustainability, 2021, 13 (3): 1551.
- [9] CIOFFI R, TRAVAGLIONI M, PISCITELLI G, et al. Artificial Intelligence and Machine Learning Applications in Smart Production: Progress, Trends, and Directions[J]. Sustainability, 2020, 12 (2): 492.
- [10] AHMAD T, MADONSKI R, ZHANG D, et al. Data-Driven Probabilistic Machine Learning in Sustainable Smart Energy/Smart Energy Systems: Key Developments, Challenges, and Future Research Opportunities in the Context of Smart Grid Paradigm[J]. Renewable and Sustainable Energy Reviews, 2022, 160: 112128.
- [11] WANG J, BILJECKI F. Unsupervised Machine Learning in Urban Studies: A Systematic Review of Applications[J]. Cities, 2022, 129: 103925.
- [12] SHAFIQ M, TIAN Z, BASHIR A K, et al. Data Mining and Machine Learning Methods for Sustainable Smart Cities Traffic Classification: A Survey[J]. Sustainable Cities and Society, 2020, 60: 102177.
- [13] RANGEL-MARTINEZ D, NIGAM K, RICARDEZ-SANDOVAL L A. Machine Learning on Sustainable Energy: A Review and Outlook on Renewable Energy Systems, Catalysis, Smart Grid and Energy Storage[J]. Chemical Engineering Research and Design, 2021, 174: 414-441.
- [14] EL NAQA I, MURPHY M J. What is Machine Learning? [M/OL].(2015-06-19)[2023-03-01]. <https://link.springer.com/book/10.1007/978-3-319-18305-3>.
- [15] 孔祥维, 唐鑫泽, 王子明.人工智能决策可解释性的研究综述[J].系统工程理论与实践, 2021, 41 (2): 524-536.
- KONG X W, TANG X Z, YANG Z M. A Survey of Explainable Artificial Intelligence Decision[J]. Systems Engineering-Theory & Practice, 2021, 41 (2): 524-536.
- [16] 纪守领, 李进锋, 杜天宇, 等.机器学习模型可解释性方法、应用与安全研究综述[J].计算机研究与发展, 2019, 56 (10): 2071-2096.
- JI S L, LI J F, DU T Y, et al. Survey on Techniques, Applications and Security of Machine Learning Interpretability[J]. Journal of Computer Research and Development, 2019, 56 (10): 2071-2096.
- [17] 陈珂锐, 孟小峰.机器学习的可解释性[J].计算机研究与发展, 2020, 57 (9): 1971-1986.
- CHEN K Y, MENG X F. Interpretation and Understanding in Machine Learning[J]. Journal of Computer Research and Development, 2020, 57 (9): 1971-1986.
- [18] MOLNAR C. Interpretable Machine Learning[M/OL].(2020-02-28)[2023-03-01]. <https://dlib.hust.edu.vn/bitstream/HUST/18122/3/OER000000269.pdf>.

- [19] 张旭.基于机器学习的精细化可解释公共盗窃犯罪分布研究[D].广州: 广州大学, 2022.
- ZHANG X. Research on the Distribution of Fine Explainable Public Theft Crimes Based on Machine Learning[D]. Guangzhou: Guangzhou University, 2022.
- [20] 杨丰春, 郑思, 李姣.可解释机器学习方法在疾病预测中的应用: 脓毒血症患者死亡风险研究[J].首都医科大学学报, 2022, 43 (4): 610-617.
- YANG F C, ZHENG S, LI J. Interpretable Machine Learning Methods Applied in Disease Risk Prediction: A Case Study of Sepsis Mortality Risk Prediction[J]. Journal of Capital Medical University, 2022, 43 (4): 610-617.
- [21] 王兰, 张苏榕, 杨秀.建成环境满意度对乡村新型社区居民自评健康的影响分析:以成都市远郊4个社区为例[J].风景园林, 2020, 27 (9): 57-62.
- WANG L, ZHANG S R, YANG X. Impacts of Built Environment Satisfaction on Self-rated Health Outcomes in New Types of Village Communities: A Case Study of Four Communities in Chengdu Outskirts[J]. Landscape Architecture, 2020, 27 (9): 57-62.
- [22] CUESTA H A, COFFMAN D L, BRANAS C, et al. Using Decision Trees to Understand the Influence of Individual- and Neighborhood-Level Factors on Urban Diabetes and Asthma[J]. Health & Place, 2019, 58: 102119.
- [23] 张雨晴, 许宁, 武云云, 等.应用随机森林分析非吸烟女性肺癌风险因素[J].环境卫生学杂志, 2022, 12 (2): 80-86.
- ZHANG Y Q, XU N, WU Y Y, et al. Using Random Forest Method to Analyze Risk Factors for Lung Cancer among Non-smoking Women[J]. Journal of Environmental Hygiene, 2022, 12 (2): 80-86.
- [24] 裴泽华, 葛蕊, 李浩, 等.基于随机森林模型的中国中老年人HDL-C环境影响因素研究[J].地球信息科学学报, 2022, 24 (7): 1286-1300.
- PEI Z H, GE M, LI H, et al. Environmental Factors Influencing HDL-C in Middle-aged and Elderly Chinese Population based on Random Forest Model[J]. Journal of Geo-information Science, 2022, 24 (7): 1286-1300.
- [25] WATTANAPISIT A, ABDUL RAHMAN H, CAR J, et al. The Clusters of Health-Risk Behaviours and Mental Wellbeing and Their Sociodemographic Correlates: A Study of 15,366 Asean University Students[J]. BMC public health, 2022, 22 (1): 1-12.
- [26] SINHA K, DWIVEDI J, SINGH P, et al. Spatio-Temporal Dynamics of Water Quality in River Sources of Drinking Water in Uttarakhand with Reference to Human Health[J]. Environmental Science and Pollution Research, 2022, 29 (43): 64756-64774.
- [27] REHMAN A, SABA T, HASEEB K, et al. Sustainability Model for the Internet of Health Things (IoHT) Using Reinforcement Learning with Mobile Edge Secured Services[J]. Sustainability, 2022, 14 (19): 12185.
- [28] LIN B, COOK D J. Analyzing Sensor-based Individual and Population Behavior Patterns Via Inverse Reinforcement Learning[J]. Sensors, 2020, 20 (18): 5207.
- [29] 姚婷婷, 马晓茜, 王梓恒.基于多元线性回归模型和碳平衡的CO₂排放量简便算法[J].环境污染与防治, 2017, 39 (11): 1264-1267.
- YAO T T, MA X Q, WANG Z H. A Simple Algorithm for CO₂ Emission based on Multiple Linear Regression Model and Carbon Balance[J]. Environmental Pollution & Control, 2017, 39 (11): 1264-1267.
- [30] 霍娟, 孙晓伟, 张明杰.电力负荷预测算法比较-随机森

林与支持向量机[J].电力系统及其自动化学报, 2019, 31 (7): 129-134.

HUO J, SUN X W, ZHANG M J. Comparison between Power Load Forecasting Algorithms based on Random Forest and Support Vector Machine[J]. Proceedings of the CSU-EPSC, 2019, 31 (7): 129-134.

[31] 冀嘉铭, 马鸿雁.基于神经网络组合的建筑能耗预测研究[J].计算机仿真, 2022, 39 (5): 438-443.

DOU J M, MA H Y. Prediction of Building Energy Consumption Based on Ensemble Artificial Neural Networks[J]. Computer Simulation, 2022, 39 (5): 438-443.

[32] 郭芳, 王灿, 张诗卉.中国城市碳达峰趋势的聚类分析[J].中国环境管理, 2021, 13 (1): 40-48.

GUO F, WANG C, ZHANG S H. Cluster Analysis of Carbon Emissions Peaking Trends in Chinese Cities[J]. Chinese Journal of Environmental Management, 2021, 13 (1): 40-48.

[33] YU X, ERGAN S, DEDEMEN G. A Data-Driven Approach to Extract Operational Signatures of Hvac Systems and Analyze Impact on Electricity Consumption[J]. Applied Energy, 2019, 253: 113497.

[34] ROCCHETTA R, BELLANI L, COMPARE M, et al. A Reinforcement Learning Framework for Optimal Operation and Maintenance of Power Grids[J]. Applied Energy, 2019, 241: 291-301.

[35] 侯慧, 陈跃, 吴细秀, 等.非预测机制下计及碳交易的家庭能量低碳优化实时管理[J].电网技术, 2023, 47 (3): 1-13.

HOU H, CHEN Y, WU X X, et al. Low-carbon Optimal Real-Time Management Strategy for Home Energy Considering Carbon Trading Under Non-prediction Mechanisms[J]. Power System Technology, 2023, 47 (3): 1-13.

[36] 廖了, 魏汝焱, 卢军, 等.山地城市绿地和水体对热环境的影响[J].重庆大学学报, 2022, 45 (S1): 125-129.

LIAO L, WEI R Y, LU J, et al. The Influence of Green Space and Water on Thermal Environment in Mountainous Cities[J]. Journal of Chongqing University, 2022, 45 (S1): 125-129.

[37] 侯赛男, 党国锋.机器学习模拟热环境及热岛时空变化特征研究[J].测绘科学, 2022, 47 (9): 200-207.

HOU S N, DANG G F. Simulation of Urban Thermal Environment Based on Machine Learning and Research on the Temporal and Spatial Variation Characteristics of Urban Heat Island[J]. Science of Surveying and Mapping, 2022, 47 (9): 200-207.

[38] 袁帅, 张金梦, 徐志刚, 等.城市热岛效应对建筑形态的响应及其昼夜变化研究[J].地理与地理信息科学, 2022, 38 (5): 40-46.

YUAN S, ZHANG J M, XU Z G, et al. Response of Urban Heat Island Effect to Morphology Characteristics of Buildings and its Diurnal Variation[J]. Geography and Geo-Information Science, 2022, 38 (5): 40-46.

[39] 夏晓圣, 陈菁菁, 王佳佳, 等.基于随机森林模型的中国 PM_{2.5} 浓度影响因素分析[J].环境科学, 2020, 41 (5): 2057-2065.

XIA X S, CHEN J J, WANG J J, et al. PM_{2.5} Concentration Influencing Factors in China based on the Random Forest Model[J]. Environmental Science, 2020, 41 (5): 2057-2065.

[40] ARAUJO G, ANDRADE F A. Post-Processing Air Temperature Weather Forecast Using Artificial Neural Networks with Measurements from Meteorological Stations[J]. Applied Sciences, 2022, 12 (14): 7131.

[41] 左芝鲤, 郭海湘, 成金华.长江经济带空气质量影响因素研究[J].环境经济研究, 2018, 3 (4): 150-167.

ZUO Z L, GUO H X, CHENG J H. The Study on the Influential Factors of Air Quality in the Yangtze River Economic Belt[J]. Journal of Environmental Economics, 2018, 3 (4): 150-167.

[42] VIGNESH R S, SIVAKUMAR A, SHYAM M, et al. Deep Reinforcement Learning Based Weather Monitoring System using Arduino for Smart Environment[J]. International Journal of Recent Technology and Engineering, 2019, 8 (4): 2277-3878.

[43] 黄隆杨, 刘胜华, 李健.城市生态用地时空动态及其相关驱动力: 以武汉市为例[J].长江流域资源与环境, 2019, 28 (5): 1059-1069.

HUANG L Y, LIU S H, LI J. Spatial and Temporal Dynamics of Urban Ecological Land Use and Its Related Driving Forces: A Case Study of Wuhan City[J]. Resources and Environment in the Yangtze Basin, 2019, 28 (5): 1059-1069.

[44] 张文强, 罗格平, 郑宏伟, 等.基于随机森林模型的内陆干旱区植被指数变化与驱动力分析: 以北天山北坡中段为例[J].植物生态学报, 2020, 44 (11): 1113-1126.

ZHANG W Q, LUO G P, ZHENG H W, et al. Analysis of Vegetation Index Changes and Driving Forces in Inland Arid Areas Based on Random Forest Models: A Case Study of the Middle Part of Northern Slope of the North Tianshan Mountains[J]. Chinese Journal of Plant Ecology, 2020, 44 (11): 1113-1126.

[45] YANG S, FENG Q, LIANG T, et al. Modeling Grassland Above-Ground Biomass Based on Artificial Neural Network and Remote Sensing in the Three-River Headwaters Region[J]. Remote Sensing of Environment, 2018, 204: 448-455.

[46] MOROS-OCHOA M A, CASTRO-NIETO G Y, QUINTERO-ESPAÑOL A, et al. Forecasting Biocapacity and Ecological Footprint at a Worldwide Level to 2030 Using Neural Networks[J]. Sustainability, 2022, 14 (17): 10691.

[47] 盛硕, 王云才.基于生态特征与问题聚类识别的平原农业区乡村生态修复策略: 以辽宁省黑山县为例[J].中国园林, 2022, 38 (12): 20-25.

SHENG S, WANG Y C. Rural Ecological Restoration Strategies in Plain Agricultural Areas Based on the Identification of Ecological Characteristics and Issues Through Clustering: A Case Study of Heishan County, Liaoning Province[J]. Chinese Landscape Architecture, 2022, 38 (12): 20-25.

[48] SUNEHAG P, LEVER G, LIU S, et al. Reinforcement Learning Agents Acquire Flocking and Symbiotic Behaviour in Simulated Ecosystems; Proceedings of the ALIFE 2019: The 2019 Conference on Artificial Life, F, 2019[C]. Cambridge, MA: MIT Press, 2019: 103-110.

[49] WEN Y, WU R, ZHOU Z, et al. A Data-Driven Method of Traffic Emissions Mapping with Land Use Random Forest Models[J]. Applied Energy, 2022, 305: 117916.

[50] VAIYAPURI T, GUPTA M. Traffic Accident Severity Prediction and Cognitive Analysis Using Deep Learning [J/OL]. Soft Computing, 2021(2021-11-23)[2023-03-01]. <https://link.springer.com/article/10.1007/s00500-021-06515-5>.

[51] YAN M, SHEN Y. Traffic Accident Severity Prediction Based on Random Forest[J]. Sustainability, 2022, 14 (3): 1729.

[52] WEI T, ZHU T, LIU H. Risk Prediction and Factor Analysis of Rider's Injury Severity in Passenger Car and E-Bike Accidents Based on Interpretable Machine Learning [J/OL]. Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering, 2022(2022-09-07)[2023-03-01]. <https://journals.sagepub.com/doi/abs/10.1177/09544070221121864>.

[53] DAS S, DUTTA A, DEY K, et al. Vehicle Involvements in Hydroplaning Crashes: Applying Interpretable Machine Learning[J]. Transportation Research Interdisciplinary Perspectives, 2020, 6: 100176.

[54] 饶丹, 时宏伟.基于深度聚类的航空交通流识别与异常检测研究[J].计算机科学, 2023, 50: 1-11.

RAO D, SHI H W. Study on Air Traffic Flow Recognition and Anomaly Detection Based on Deep Clustering[J]. Computer Science, 2023, 50: 1-11.

[55] 李闻, 常秋实.基于卷积神经网络的交通工具分类系统[J].信息与电脑 (理论版), 2022, 34 (20): 171-174.

LI W, CHANG Q S. Vehicle Classification System based on Convolutional Neural Network[J]. Information & Computer, 2022, 34 (20): 171-174.

[56] LI H, LAM J S L, YANG Z, et al. Unsupervised Hierarchical Methodology of Maritime Traffic Pattern Extraction for Knowledge Discovery[J]. Transportation Research Part C: Emerging Technologies, 2022, 143: 103856.

[57] SHANG Q, YU Y, XIE T. A Hybrid Method for Traffic State Classification Using K-Medoids Clustering and Self-Tuning Spectral Clustering[J]. Sustainability, 2022, 14 (17): 11068.

[58] 曹欢.交通场景下基于深度强化学习的感知型路径分配算法[J].信息技术与网络安全, 2022, 41 (6): 43-49.

CAO H. Perceptual Path Allocation Algorithm based on Deep Reinforcement Learning in Traffic Scene[J]. Cyber Security And Data Governance, 2022, 41 (6): 43-49.

[59] YAN L, ZHU L, SONG K, et al. Graph Cooperation Deep Reinforcement Learning for Ecological Urban Traffic Signal Control[J]. Applied Intelligence, 2023, 53 (6): 6248-6265.

[60] 王金永, 黄志球, 杨德艳, 等.面向无人驾驶时空同步约束制导的安全强化学习[J].计算机研究与发展, 2021, 58 (12): 2585-2603.

WANG J Y, HUANG Z Q, YANG D Y, et al. Spatio-Clock Synchronous Constraint Guided Safe Reinforcement Learning for Autonomous Driving[J]. Journal of Computer Research and Development, 2021, 58 (12): 2585-2603.

图表来源(Sources of Figures and Tables):

图 1 由作者绘制, 其中可持续发展目标部分来自参考文献 [3]; 图 2 由作者绘制; 表 1 根据参考文献 [18]~[20] 整理; 表 2 根据参考文献 [21]~[60] 整理。

(编辑 / 王一兰)

作者简介:

刘泽润 / 女 / 同济大学建筑与规划学院在读硕士研究生 / 研究方向为城市可持续发展

刘超 / 女 / 同济大学建筑与规划学院助理教授 / 智慧社区试验中心负责人 / 研究方向为智能可持续城市规划
通信作者邮箱: liuchao1020@tongji.edu.cn

LIU Z R, LIU C. Application of Machine Learning in the Research on Sustainable Built Environment: Progress and Prospect[J]. Landscape Architecture, 2023, 30(7): 51-59. DOI: 10.12409/j.fjyl.202303010108.

Application of Machine Learning in the Research on Sustainable Built Environment: Progress and Prospect

LIU Zerun, LIU Chao*

Abstract:

[Objectives] The development of big data and the internet of things (IoT) has accelerated the process of artificial intelligence, with machine learning (ML) becoming particularly noticeable. More people are realizing the influence of advanced technology on reframing the researches in the field of sustainable built environment. Thus, this research aims to identify the role of machine learning in the research on sustainable built environment.

[Methods] By reviewing the conceptions of sustainability and built environment, the research summarizes several important topics in the wide-ranging field of sustainable built environment. Based on literature from Web of Science (WoS) and China National Knowledge Infrastructure (CNKI) over the last five years, the research finds that the increasing trends of ML-related research are significant for topics such as urban public health, energy consumption and carbon emission, climate and environment, ecosystem, and green travel, which may serve as a useful clue for structure organization and conclusion drawing of this research. Initially, the research elaborates on the concept, classification and algorithms of ML, and presents the application characteristics of strong prediction capability but poor interpretation capability. Then the research analyzes the trend change of ML from predictability to interpretability, followed by the descriptions of scope, taxonomy and core algorithms of interpretable ML. Moreover, the research establishes a thorough table to list multiple ML algorithms and their roles in the research. Far from being all-inclusive, the research only tries to incorporate as many methods and objectives as possible across the world. The research particularly focuses on the predictability and interpretability of the results of ML algorithms, and the specific roles of such algorithms in the research direction, methodology and conclusion of sustainable built environment.

[Results] The results show that ML's interpretability capability substantially aids in the transformation of research techniques. More and more researchers are practicing ML-based research frameworks and building practical methodological routes, accompanied by increasingly diversified research directions and objects. ML has excellent performance in dealing with complex, heterogeneous, dynamic and large-scale data. Complicated issues previously considered difficult to solve have become the deciphering objects of ML, providing new insights and possibilities for solving composite puzzles. Meanwhile, ML itself is progressing toward intelligence, transparency and interpretability. By integrating experience and human wisdom, ML can help researchers formulate optimal decision strategies, reveal underlying mechanisms, and uncover universal laws. Furthermore, the research findings are more scientific and sophisticated. The majority of researches in the five key topics recognize the great advantages of ML in prediction and interpretation and start to develop interpretable and readable methods based on specific scenarios and problems. It is clear that more researches are focusing on

understanding decision-making processes and mining patterns, although the levels of progress thereof vary depending on specific issues. As for public health, several articles discuss the effects of the built environment on health concerns. The applications of interpretable ML are relatively simple and preliminary, with less empirical investigation of transparent models. More medical and clinical experiments are needed to support theoretical analyses. As for energy consumption and carbon emission, as well as climate and environment, both predicting and interpreting tasks are important. By upgrading data processing, the accuracy of estimation and prediction can be improved. Additionally, models' inner processes can be better understood by optimizing algorithms. As for ecosystems, there are more researches on prediction than on interpretation. The complexity of ecological elements might be the primary cause for this situation, which may be changed when more powerful models are created. As for green travel, ML applications have reached a mature stage. Lots of cutting-edge technologies have been developed and widely implemented. Except for the predictable and interpretable tasks, some researches have further stepped into the next level of intelligent decisions and optimal strategies, providing a vibrant foundation for human-machine interaction.

[Conclusions] In summary, there is no doubt that ML plays an important role in transforming research methods, objects and conclusions in the field of sustainable built environment, and can effectively promote the development of researches in this field in a more quantitative, diverse, intelligent and scientific manner. It can be foreseen that the researches and applications of ML in sustainable built environment will continue to advance swiftly, with long-term and broad prospects. Besides, the shift from predictability to interpretability is obvious. On the one hand, there are more interpretable methods and readable ML models, and research objectives tend to emphasize understanding the decision-making process and revealing the underlying rules; on the other hand, the exploration of casual mechanisms based on clinical experiments is still insufficient. In terms of the five key topics, the development of ML research and application is the fastest in transportation issues such as green travel, forming a mature system of prediction, interpretation, and decision-making, while continuously promoting the exploration of more innovative applications and methods of human-machine interaction. There are also quite full prediction and interpretation tools for researches on energy consumption and carbon emission, and climate and environment. However, both the variety and professionalism of research methods remain insufficient. Though the ML applications in ecosystems and public health are still in the early stage, many researchers have realized the great potential of ML in these areas, implying that additional chances and breakthroughs in related fields are likely to emerge in the future.

Keywords: artificial intelligence; interpretable machine learning; public health; energy consumption and carbon emission; climate and environment; ecosystem

Authors:

LIU Zerun is a master student in the College of Architecture and Urban Planning (CAUP), Tongji University. Her research focuses on urban sustainable development.

LIU Chao is an assistant professor in the College of Architecture and Urban Planning (CAUP), Tongji University, and leader of Smart Community Lab. Her research focuses on intelligent sustainable urban planning.

Corresponding author Email: liuchao1020@tongji.edu.cn