

特色专题

面向时间序列和时空数据分析的大模型研究进展

罗远波¹, 孙嘉¹, 陶俐芝^{2*}

摘要 近年来,大模型凭借其强大的序列建模能力、优异的特征学习潜力和灵活的预训练-微调范式,在自然语言处理和计算机视觉领域取得了突破性进展,也为时间序列和时空数据的挖掘分析带来新的发展机遇。综述了大模型在时间序列与时空数据分析中的研究进展,重点涵盖大语言模型赋能与专用基础模型构建2条路径。前者通过提示工程、标记化设计和参数微调等手段,使大模型能够适配时序与时空任务;后者则依托跨域大规模预训练,形成统一的动态表征能力。大语言模型赋能路径具有开发成本低、零样本/少样本迁移灵活等优势,而专用基础模型在跨域泛化方面表现更优。与此同时,这2类方法仍存在可解释性不足、多模态语义对齐困难等问题。未来研究亟需在可解释性增强、多模态联合建模以及模型架构创新等方面取得突破。

关键词 时间序列;时空数据;大语言模型;预训练基础模型

时间数据主要包含时间序列和时空数据,它们蕴含着丰富的动态信息,广泛应用于现实世界中^[1-2]。它们捕获动态系统的测量值,并由物理和虚拟传感器大量产生。分析这些数据对于理解现实世界系统的复杂性和演化规律至关重要。

近年来,人工智能领域见证了基础模型(foundation models, FMs)的崛起与蓬勃发展。这些深度神经网络模型通过自监督学习在大规模数据集上进行预训练,能够学习到通用知识表征,并通过迁移学习广泛适用于各种下游任务,取得了变革性成功。基础模型从早期预训练语言模型发展到现在的大语言模型(large language models, LLMs)和视觉语言模型(vision-

language models, VLMs),在多模态理解、推理和跨域泛化等方面展现出卓越性能。随着多模态 LLMs 的不断进步,LLMs 的下游任务已远超传统自然语言的范畴,渗透到更广泛的数据模态和应用场景中^[3]。

值得注意的是,时间序列和时空数据与自然语言序列在序列结构上具有本质的相似性。时间序列是按时间顺序排列的单变量或多变量数值序列,主要反映系统随时间演化的动态过程;而时空数据则在此基础上引入空间维度,用于描述特定空间位置上随时间变化的状态。两者虽形式不同,但本质上都是时间数据,均依赖上下文信息进行建模。通过合适的离散化方式,这些数据可以转化为类似自然语言词元(token)的离散符号序列,与 LLMs 处理文本的方式在结构上具有相似性。这种结构上的相似性自然地引出一个关键问题:我们能否以及如何有效利用现有强大的 LLMs,赋能通用的时间序列与时空数据的分析任

1. 南方海洋科学与工程广东省实验室(广州),广州 511458

2. 江西师范大学鄱阳湖湿地与流域研究教育部重点实验室,南昌 330022

收稿日期: 2025-05-08; 修回日期: 2025-07-03

基金项目: 国家自然科学基金青年科学基金项目(42301041); 中国科学院基础与交叉前沿科研先导专项(XDB0740200); 广东省科技计划项目(2024B1212080002)

作者简介: 罗远波,博士研究生,研究方向为时间序列建模与地理时空数据智能分析,电子信箱: luo_yuanbo@gmlab.ac.cn; 陶俐芝(通信作者),助理研究员,研究方向为

水文气象多尺度问题建模与分析,电子信箱: taoliz@jxnu.edu.cn

引用格式: 罗远波,孙嘉,陶俐芝. 面向时间序列和时空数据分析的大模型研究进展[J]. 科技导报, 2025, 43(18): 48-56; doi:10.3981/j.issn.1000-7857.2025.05.00037

务? 这可能涉及将时间序列/时空数据编码为 LLMs 能够理解的语言序列, 或利用 LLMs 强大的上下文理解和推理能力进行零样本或少样本预测、异常检测等。或者说, 是否需要借鉴 LLMs 的成功范式, 构建专门针对时间序列与时空数据特性的基础模型? 这一范式强调在海量、跨域的时序/时空数据集上进行大规模预训练, 学习时间和空间维度上的通用表征和动态规律, 然后通过迁移学习技术适配到各种具体的下游任务, 最终实现数据通用和任务通用的时间序列和时空数据基础模型。

探索上述问题, 将大模型的技术优势引入时间序列与时空数据的挖掘分析中已成为当下的学术前沿热点^[4]。本文旨在全面梳理这一新兴交叉领域的研究现状, 回顾利用大模型进行时间序列与时空数据分析的最新进展, 探讨关键技术、代表性模型、主要应用场景, 分析当前面临的挑战, 并展望未来的发展趋势, 以期对相关研究和引用提供有价值的参考。

1 基础理论与关键技术

1.1 时间序列与时空数据分析

时间序列(time series)定义为按时间排序的随机变量的集合 $\{X_t | t \in T\}$, 时间序列的观测值可以认为是随机变量的实现。时间序列可以分为单变量时间序列和多变量时间序列。对于长度为 T 的单变量时间序列样本 $\mathbf{X} = \{x_1, x_2, \dots, x_T\} \in \mathbf{R}^T$, $x_t \in \mathbf{R}$ 是时刻 t 对应的观测值; 具有 $D(D > 1)$ 个通道的多变量时间序列样本 $\mathbf{X} = \{x_1, x_2, \dots, x_T\} \in \mathbf{R}^{T \times D}$ 由一组有序的 D 维向量组成, 其中 $x_t \in \mathbf{R}^D$ 是时刻 t 对应的观测值。一般来说, 主流的时间序列分析任务包括预测、插补、异常检测和分类。

时空数据(spatio-temporal data)是指同时包含时间维度和空间维度的数据集, 用于记录和分析时空动态变化规律及其内在关联。它可以被理解为一组时间序列, 每个序列都与特定的空间位置相关联。在实际应用中, 时空数据常见的数据结构包括位置(location)、轨迹(trajec-tory)、事件(event)、时空栅格(spatio-temporal raster)和时空图(spatio-temporal graph)。时空数据分析包含与时间序列分析类似的任务, 但更注重针对特定应用场景的优化与扩展。

1.2 LLMs

LLMs 通常是具有大规模参数和计算能力的自然语言处理模型。通过规模扩展(例如增加模型参数规模或数据规模)通常会带来下游任务的模型性能提升, 这些大规模的预训练语言模型在解决复杂任务时表现出与小型预训练语言模型不同的行为^[5]。这种大模型具有但小模型不具有的能力通常被称为“涌现能力”(emergent abilities), 为序列建模带来了前所未有的潜力, 并重塑了人类使用人工智能的方式。

1.3 预训练基础模型

预训练基础模型(pre-trained foundation models, PFM)是指在大规模语料库上进行预训练的通用人工智能模型。在自然语言处理领域, BERT^[6]和 GPT-3^[7]等基础模型彻底改变了文本理解和生成任务。在计算机视觉领域, 文本提示模型 CLIP^[8]和视觉提示分割模型(segment anything model, SAM)^[9]等模型推动了图像识别、目标检测等领域的进步。随着 PFM 的巨大成功, 针对时间数据设计和预训练的基础模型也在不断涌现, 但仍处于早期发展的阶段。

1.4 骨干网络

1) Transformer 架构。Transformer 网络结构作为自然语言处理领域的一个革命性进展, 其核心创新在于对注意力机制的利用, 该机制允许模型动态地关注输入数据的不同部分。根据这种网络架构的变形, 目前主流的框架可分为 Encoder-decoder(编码器-解码器)、Encoder-only(仅编码器)、Decoder-only(仅解码器)3 种^[10], 如图 1 所示。

2) 扩散(Diffusion)模型。扩散模型代表了一类概率生成模型, 这些模型通过在训练样本中注入和去除噪声的两步过程进行优化和生成^[11]。具体而言, 其通过一个迭代的正向扩散过程系统地、缓慢地破坏数据分布中的结构, 然后学习一个反向去噪过程来恢复数据中的结构, 从而得到一个高度灵活且易于处理的生成模型(图 2)。

3) 其他模型。除 Transformer 和扩散模型这类近年来备受关注的架构外, 一些经典的深度学习模型, 包括多层感知机(multi-layer perceptron, MLP)、卷积神经网络(convolutional neural network, CNN)、循环神经网络(recurrent neural network, RNN)和图神经网络(graph neural network, GNN)仍然是时间序列与时

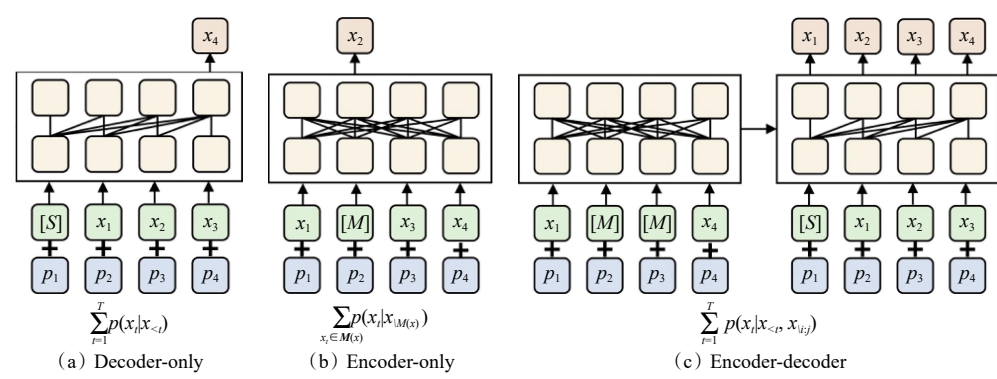


图1 3种主流模型训练框架

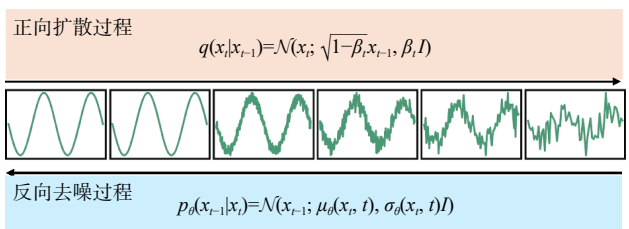


图2 扩散模型用于时间序列和时空数据分析

空数据分析领域的重要基石。

2 LLMs 赋能时间序列与时空数据分析

自然语言处理领域 LLMs 的快速发展揭示了其在序列建模和模式识别方面前所未有的能力,这为时间序列与时空数据分析提供新的研究范式。本节基于 LLMs 的时序与时空数据分析方法,并依据模型架构设计将其划分为 3 类:直接推理方法以零样本或少样本方式直接利用预训练 LLMs 进行挖掘分析;定制微调方法通过领域适配的微调策略使 LLMs 适应特定任务,具体包括标记化设计、提示设计和微调策略 3 个关键环节(图 3);特征增强方法将 LLMs 作为辅助模块,与传统模型协同提升性能。

2.1 直接查询 LLMs

PromptCast^[12]引入了一种新的基于提示(Prompt)的学习范式,通过设计特定的提示词模板将数值序列转化为自然语言提示,利用 LLMs 以句子对句子的方式进行序列预测。研究表明,在特定领域场景中,尤其是当结合领域相关的上下文提示和先进的预训练 LLMs 时,直接查询 LLMs 可能带来显著优势。

2.2 时间数据的标记化设计

LLMs 被设计用来处理离散的 token,对于时间序

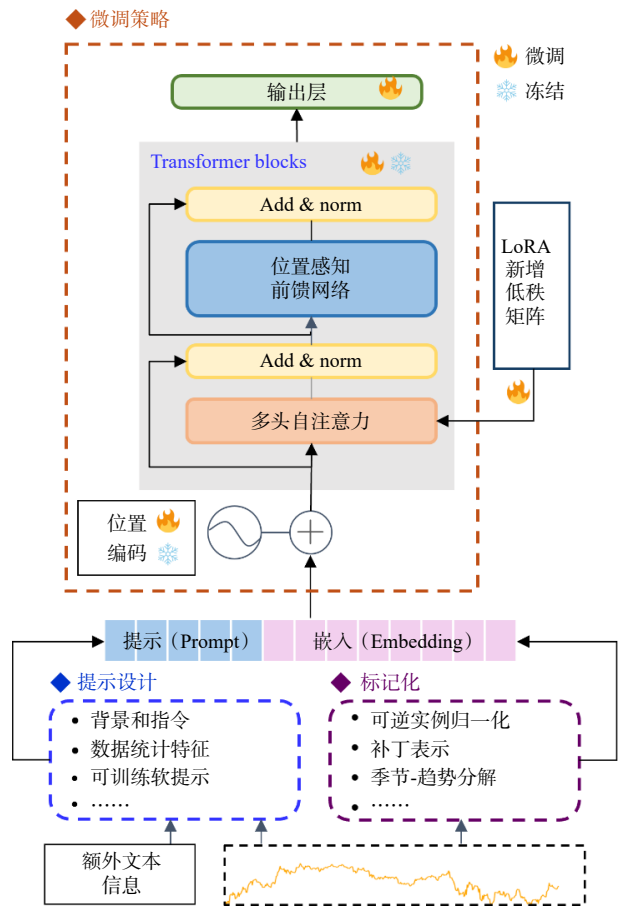


图3 微调 LLMs 组件设计分类

列与时空数据的数值序列输入,将其转化为字符串 token 是关键步骤。Nie 等^[13]提出了时间序列的补丁表示方法,对长度为 L 的单变量时间序列 $\mathbf{X}_{1D} \in \mathbf{R}^L$,补丁操作首先在原始序列中重复最后的数值 S 次。然后,通过具有补丁长度 P 和步长 S 的滑动窗口展开输入单变量时间序列。通过补丁操作,单变量时间序列被转换为二维表示: $\mathbf{X}_p \in \mathbf{R}^{P \times N}$,其中 $N = \left\lfloor \frac{L-P}{S} \right\rfloor + 2$,是补丁的数目。补丁表示方法保留了数据的原始相对

顺序,并将局部信息聚合到每个补丁中。OFA^[14]、LLM4TS^[15]、TEST^[16]、TEMPO^[17]和 Time-LLM^[18]等工作主要采用这种补丁表示方法对时间序列数据进行标记化。为了协调数值数据和自然语言的模态,TEST 引入了一个辅助损失,以增强时间序列标记的嵌入与选定文本原型之间的余弦相似性,并确保相似时间序列实例在文本原型空间中的邻近性。出于类似的目的,Time-LLM 提出使用多头注意力机制来将时间序列补丁表示与通过线性探测获得的预训练文本原型嵌入对齐。此外,基于提示选择的生成 Transformer 模型 TEMPO 使用加性 STL(seasonal and trend decomposition using loess)分解来提取趋势、季节性和残差成分,以更好地标记输入时间序列。

2.3 提示设计

PromptCast 为利用 LLMs 进行时间序列预测开发了基于模板的提示,而一些方法通过整合 LLMs 生成的或收集的背景信息来丰富提示设计。除了背景和指令提示之外,Time-LLM 还添加了时间序列数据的关键统计信息,包括最小值、最大值、中位数以及整体趋势等信息。与固定且不可训练的提示相比,软提示和可训练的提示使 LLMs 更容易理解和与输入对齐。软提示将文本提示转化为可训练的连续嵌入向量引入到模型的输入层,并在训练过程中根据特定任务进行调整。TEST 训练了针对时序数据的软提示,这些软提示针对特定任务嵌入,可以从均匀分布中随机初始化,或从下游任务标签的文本嵌入中获取初始值、从词汇表中最常见的词汇中获取初始值等。为应对分布偏移的挑战,TEMPO 引入一个共享的提示池,以不同的键值对的形式存储,然后通过相似性分数匹配机制为微调选择最具代表性的软提示,从而让模型更好地利用已有的经验知识。

2.4 微调策略

微调作为一种强大的工具,能够使大型预训练模型适应于特定的任务和应用场景。已有的工作通过直接微调预训练的 LLMs 以进行时间序列分析,区别在于微调过程中模块参数的更新方式。作为一项标准做法,OFA 微调了位置嵌入和归一化层,并冻结了自注意力层和前馈神经网络。TEMPO 进一步使用低秩自适应(LoRA)微调自注意力模块,以适应不同的数据分布。LLM4TS 提出了一种 2 阶段微调方法,首

先通过监督微调将模型与时间序列的特性进行对齐,引导 LLMs 适应时间序列,接下来以下游预测任务为导向进一步对模型进行微调,从而保障在不破坏语言模型固有特性的基础上使得模型能够更好地适配各类不同域的数据及不同的下游任务。

2.5 集成 LLMs 作为模型组件

相较于依赖直接查询或微调 LLMs 来生成输出,一些研究选择将冻结的 LLMs 作为模块嵌入,以扩展和增强输入序列的特征表示空间。冻结的 LLMs 可以作为多阶段建模中的一个功能强大的组件,它提供对数据的中间处理或前一个组件的输出,并将它输入到随后的神经网络或回归分析^[19]。

3 构建专用的时间序列与时空数据基础模型

3.1 时间序列基础模型研究进展

自然语言处理和计算机视觉领域的基础模型近年来呈现指数级增长,而时间序列领域的基础模型研究仍处于早期阶段。这一差距主要源于可用的时间序列数据集规模较小,以及时间序列数据本身的高噪声、非平稳性和跨领域差异性等固有挑战。尽管如此,近期已有一些开创性的工作开始填补这一空白(表 1),这些模型借鉴了 LLMs 的成功经验,探索了时间序列预训练的范式,为该领域的基础模型发展奠定了重要基础。

TimeGPT-1 是由 Nixtla 提出的第一个时间序列大模型,无需额外训练就可以在金融、电力、气象等各种领域生成准确的预测。其通用的预训练模型构成了一项突破性的创新,为预测实践开辟了新范式。在 TimeGPT-1 发布后,时间序列基础模型的发展进入加速阶段,包括 Lag-Llama、TimesFM、Chronos、Moirai 和 Time-MoE 等代表性工作。作为首个开源大型预训练时间序列基础模型,MOMENT 将大量的公共时间序列数据整合成时间序列堆栈,然后通过对时间序列的掩码预测任务进行预训练,其流程可概括为:分段(Patching)→编码(Encoding)→掩码(Masking)→重建(Reconstruction)。Chronos 通过尺度缩放(scaling)和量化(quantization)技术,将连续时间序列值离散化为固定词汇表中的标记,并直接采用与语言模型相同的交叉熵损失函数进行训练。这一方法

表 1 代表性时间序列基础模型汇总

模型名称 (method)	骨干网络 (architecture)	主研机构 (main institute)	预训练数据集 (pre-training scale)	下游任务 (downstream tasks)	年份 (year)
TimeDiT ^[20]	Diffusion	USC	—	预测、插补	2024
Time-MoE ^[21]	Decoder-only	PU	309 B	预测	2024
Moirai-MoE ^[22]	Decoder-only	Salesforce	—	预测	2024
TTMs ^[23]	MLP	IBM	—	预测	2024
Timer ^[24]	Decoder-only	Tsinghua	28 B	预测、插补、异常检测	2024
UniTS ^[25]	Encoder-Decoder	Harvard & MIT	—	预测、插补、异常检测、分类	2024
Moirai ^[26]	Encoder-only	Salesforce	27.65 B	预测	2024
MOMENT ^[27]	Encoder-only	CMU	1.13 B	预测、分类、异常检测	2024
Chronos ^[28]	Encoder-Decoder	Amazon	84 B	预测	2024
ForecastPFN ^[29]	Encoder-only	Abacus.AI	—	预测	2023
Lag-Llama ^[30]	Decoder-only	—	0.36 B	预测	2023
TimesFM ^[31]	Decoder-only	Google	100 B	预测	2023
TimeGPT-1 ^[32]	Decoder-only	Nixtla	100 B	预测、异常检测	2023

巧妙地将时间序列预测问题转化为类似语言建模的任务,为时间序列基础模型的发展提供了新思路。与大多数采用补丁标记化(patch tokenization)的模型不同,Time-MoE 采用逐点标记化(point-wise tokenization)方法进行时间序列嵌入以完整保留时序信息。

作为时序预测领域的新兴技术,时间序列基础模型各具优势,适用于不同的应用场景,但也伴随着一定的局限性。Lag-Llama 和 TimesFM 均基于 Transformer 架构,前者在概率预测方面表现稳定,后者则在零样本泛化方面表现突出。Chronos 和 TSFM 擅长处理金融时序和多频率数据,而 Moirai 和 MOMENT 凭借跨领域适应性和灵活定制能力受到关注。然而,这些模型在复杂场景中的表现仍存在局限,特别是在处理高频数据、长跨度预测任务时,其效果往往不及传统统计方法。当面对上下文信息不足的情况时,模型的泛化能力也会明显下降。此外,由于需要针对不同垂直领域独立开发,这些模型的训练过程通常需要消耗大量计算资源。虽然通过融合多模态信息和改进预训练策略有望提升模型性能,但在当前发展阶段,传统方法在大多数实际应用中仍保持着可靠性和效率方面的优势。

3.2 时空数据基础模型研究进展

相较于单一时间序列数据,时空数据普遍呈现出 2 个关键的基本特性:异质性与自相关性。异质性表

现为数据分布的高度不均匀性和复杂性,体现在不同地理位置或时间节点上观测变量的特征及其相互关系存在显著差异。自相关性则揭示了时空数据中邻近观测值之间的统计依赖性,即空间或时间上相近的测量值往往具有更强的相似性。这些固有特性不仅显著增加了时空数据建模的复杂度,也对基础模型提出了更高的泛化能力要求。当前时空数据基础模型的发展仍处于相对分散的状态,应用场景也呈现出明显的碎片化特征(表 2)。

3.2.1 城市时空动态建模

针对城市时空数据的多模态特性,UrbanVLP 提出一种视觉-语言-时空联合预训练框架,通过融合卫星影像、城市传感器数据及文本语义信息,构建统一的多模态表征空间。其核心创新在于设计时空对比学习策略,将不同模态的时空模式对齐到共享的潜在空间。类似的,CityFM 专注于城市多任务泛化,采用层次化时空注意力机制,对交通流量、人群移动、空气质量等异构数据进行联合建模,并通过任务解耦的微调策略实现跨场景迁移。OpenCity 进一步引入地理编码先验知识,将城市区域的空间拓扑结构嵌入模型架构,增强了模型对空间异质性的建模能力。UniST 作为第一个无需自然语言的纯时空通用模型,实现了对城市多样化时空动态性的统一建模和精准预测,尤其是在跨场景零样本预测中,其性能超过了

表 2 代表性时空基础模型汇总

模型名称 (method)	骨干网络 (architecture)	应用 (applications)	数据结构 (data types)	年份 (year)
UrbanVLP ^[33]	Encoder-only	城市计算	位置	2025
CityFM ^[34]	Encoder-only	城市计算	位置	2024
UniTraj ^[35]	Encoder-Decoder	城市计算	轨迹	2024
TrajFM ^[36]	Encoder-only	城市计算	轨迹	2024
PTrajM ^[37]	Mamba	城市计算	轨迹	2024
OpenCity ^[38]	Encoder-only	城市计算	时空图	2024
GraphCast ^[39]	GNN	气候	时空图	2023
UniST ^[40]	Encoder-Decoder	城市计算	时空栅格	2024
UrbanDiT ^[41]	Diffusion	城市计算	时空栅格	2024
ClimaX ^[42]	Encoder-only	气候	时空栅格	2023
FengWu ^[43]	Encoder-only	气候	时空栅格	2023
Pangu ^[44]	Encoder-Decoder	气候	时空栅格	2023

少样本基线方法,展现出强大的通用性和泛化能力。

对于轨迹数据的 PFMs,旨在从输入数据中学习通用的顺序模式。TrajFM 是这一方向上的先驱,它引入了一种能够支持区域和任务转移的轨迹基础模型,采用轨迹屏蔽和自回归恢复机制来增强其学习能力。UniTraj 轨迹基础大模型通过收集学习全球十亿级轨迹点的高质量轨迹数据,并集成多种重采样和掩码策略,实现跨任务、跨区域的泛化能力,并在不同数据质量下保持鲁棒性。PTrajM 则是引入 Trajectory-Mamba 作为可学习模型,通过整合运动行为参数化和轨迹状态空间模型,实现对连续运动行为的有效提取。

3.2.2 气象与地球系统建模

气象预测是时空数据基础模型的典型应用场景。谷歌的 DeepMind 研究所开发的天气预报模型 GraphCast 基于图神经网络和 Novel 的高分辨率多尺度网格表示,构建了全球大气状态的高分辨率预测模型,其通过物理约束的图结构建模大气变量间的复杂相互作用,显著提升了中长期气象预测的精度。微软团队发布的 ClimaX 使用 Novel 的编码和聚合块扩展了 Transformer 架构,通过微调技术来解决广泛的气候和天气任务。上海人工智能实验室发布的 FengWu 大模型则采用多模态神经网络和多任务自动均衡权重解决多种大气变量表征和相互影响的问题。此外,华为提出的 Pangu-Weather 通过引入地球坐标的 3D Transformer 和层次化时域聚合技术提前一周预测全球天气。

4 挑战与机遇

面向时间序列和时空数据分析的大模型研究正处于快速发展阶段,但仍面临众多挑战,同时也蕴藏着新的机遇。

4.1 增强模型的可解释性

当前,尽管 LLMs 在时间序列和时空数据分析任务中展现出一定的泛化能力,但其内部决策机制的“黑箱”特性仍构成关键研究瓶颈^[45]。构建系统性理论框架以解析 LLMs 在时间序列和时空数据分析中的学习模式与应用逻辑,成为亟待突破的关键问题。此外,探索 LLMs 在时间推理与因果关系推断中的能力增强路径同样意义重大,需研发适用于时间序列和时空数据的因果关系识别方法。

4.2 发展多模态模型

在现实应用场景中,许多时间序列和时空数据通常伴随文本描述、图片等补充信息。在此背景下,LLMs 可通过适配学习构建多模态联合表示—既捕捉时序数据的序列依赖性特征,又保留其他模态的独特信息特性^[46]。LLMs 能够有效处理多模态时序数据在时间分辨率上的异质性,通过动态对齐或尺度转换等机制,实现跨时间分辨率的多模态信息融合,从而充分挖掘不同时间尺度下的互补信息,进一步提升任务表现。

4.3 适应数据分布偏移

基础模型通过大规模多源数据训练,理论上具备

泛化多样化应用的潜力。但在实际推理过程中,模型性能仍高度依赖训练数据的统计分布特性。如果测试数据与训练分布存在显著偏移时,模型的表现可能会急剧下降。更关键的是,时空数据固有的高维度、高耦合特性进一步增加了实际应用中面临未见过模式或分布外(out-of-distribution, OOD)场景的概率。基础模型的适应性技术^[47]为应对这一挑战提供了极具潜力的解决路径。最新的研究聚焦于领域对抗训练等技术,通过对齐不同领域的特征分布,推动模型学习跨域不变性特征;此外,元学习技术^[48]通过快速学习新分布的适配规则,仅需少量新分布数据即可实现模型对新分布的快速迁移,进一步强化了基础模型应对分布偏移的能力。

4.4 创新模型架构

当前研究主要聚焦于 Transformer 架构,然而传统的 Transformer 架构和基于它们的 LLMs 由于注意力机制的特性,具有受限的上下文依赖记忆,特别是在处理长序列时面临挑战。为了解决这些挑战,需要创新性的方法,如稀疏注意力机制、高效的基于图的表示以及将 Transformer 与其他架构结合的混合模型。最近,谷歌提出的 Infini-Transformer^[49]通过引入“无限注意力”机制,使得 LLMs 能够在有限的内存条件下处理极长的输入;Meta 等开发新架构 Megalodon^[50]通过突破性的设计实现了无限上下文处理能力,在特定训练任务中展现出非凡的效率。这些前沿研究为提升 LLMs 的序列处理能力开辟了新的技术路径,对时间序列和时空数据的分析具有借鉴意义。

5 结论

最新研究表明,大模型在时间序列和时空数据分析中展现出应用潜力。一方面,基于 LLMs 的提示工程、标记化和参数微调等技术,为自然语言与时序/时空数据的模态对齐提供了可行路径。另一方面,面向时序与时空特性的专用基础模型通过跨域预训练,逐步形成统一的时空动态表征能力。从范式上看,LLMs 赋能路径开发成本低、迁移灵活、适用于少样本场景,但在面临长序列输入与复杂时空结构上存在一定局限。从零训练的基础模型路径则更强调对时间与空间结构的建模能力及特定任务的泛化表现,适

用于高频与长序列预测等复杂场景,但也面临预训练资源需求大、迁移门槛高等现实挑战。

作为一个新兴且快速发展的领域,国内外关于时间序列和时空数据大模型的研究尚处于初步阶段。当前研究仍面临若干挑战:(1) 时空数据的异质性和自相关性导致模型泛化能力受限,跨领域迁移时易受分布偏移影响;(2) 多模态对齐存在语义鸿沟,数值型时空数据与文本/视觉等模态的联合表征仍需突破;(3) 模型可解释性不足制约其在关键领域的可信应用。

未来,面向时间序列与时空数据分析的大模型需聚焦以下方向:(1) 模型架构创新,融合物理机理与数据驱动优势,提升对非平稳动态的建模能力;(2) 多模态协同学习技术突破,建立时空-语言-视觉的统一表征,实现跨模态知识迁移;(3) 可信赖性增强机制构建,发展因果推理、不确定性量化等技术提升模型可解释性。随着这些技术瓶颈的突破,时间序列和时空数据大模型有望成为理解复杂现实系统、支撑科学决策的通用基础设施,为智慧城市、气候预测、精准医疗等重大领域带来颠覆性创新。

致谢:感谢“坤元”地理科学大模型团队在文献分析中提供的技术支持。

参考文献(References)

- [1] 韩敏,任伟杰,李柏松,等.混沌时间序列分析与预测研究综述[J].信息与控制,2020,49(1):24-35.
- [2] 王家耀,武芳,郭建忠,等.时空大数据面临的挑战与机遇[J].测绘科学,2017,42(7):1-7.
- [3] 李耕,王梓烁,何相腾,等.从ChatGPT到多模态大模型:现状与未来[J].中国科学基金,2023,37(5):724-734.
- [4] 邵泽志,余澄庆,李雨杰,等.决策智能中的时间序列预测大模型[J].指挥与控制学报,2025,11(2):146-157.
- [5] 舒文韬,李睿潇,孙天祥,等.大型语言模型:原理、实现与发展[J].计算机研究与发展,2024,61(2):351-361.
- [6] Devlin J, Chang M W, Lee K, et al. BERT: Pre-training of deep bidirectional transformers for language understanding[J/OL]. (2019-05-24) [2023-10-12]. <http://arxiv.org/abs/1810.04805>.
- [7] Brown T, Mann B, Ryder N, et al. Language models are few-shot learners[J]. Proceedings of the 34th International Conference on Neural Information Processing Systems, 2020, 33: 1877-1901.
- [8] Linear-probe A. Learning transferable visual models from natural language supervision[J/OL]. (2021-02-26) [2025-04-

- 27]. <http://arxiv.org/abs/2103.00020>.
- [9] Kirillov A, Mintun E, Ravi N, et al. Segment anything[J/OL]. (2023-04-05) [2025-04-27]. <http://arxiv.org/abs/2304.02643>.
- [10] 王耀祖, 李擎, 戴张杰, 等. 大语言模型研究现状与趋势[J]. 工程科学学报, 2024, 46(8): 1411-1425.
- [11] 闫志浩, 周长兵, 李小翠. 生成扩散模型研究综述[J]. 计算机科学, 2024, 51(1): 273-283.
- [12] Xue H, Salim F D. PromptCast: A new prompt-based learning paradigm for time series forecasting[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2024, 36(11): 6851-6864.
- [13] Nie Y Q, Nguyen N H, Sinthong P, et al. A time series is worth 64 words: Long-term forecasting with transformers [J/OL]. (2023-03-05) [2023-11-13]. <http://arxiv.org/abs/2211.14730>.
- [14] Zhou T, Niu P, Wang X, et al. One fits all: Power general time series analysis by pretrained LM[J]. *Advances in Neural Information Processing Systems*, 2023, 36: 43322-43355.
- [15] Chang C, Peng W J, Chen T F. LLM4TS: Aligning pre-trained LLMs as data-efficient time-series forecasters [J/OL]. (2024-01-18) [2024-12-12]. <http://arxiv.org/abs/2308.08469>.
- [16] Sun C X, Li Y L, Li H Y, et al. TEST: Text prototype aligned embedding to activate LLM's ability for time series[J/OL]. (2024-02-22) [2024-12-12]. <http://arxiv.org/abs/2308.08241>.
- [17] Cao D F, Jia F R, Arik S Ö, et al. TEMPO: Prompt-based generative pre-trained transformer for time series forecasting[J/OL]. (2024-04-02) [2024-12-12]. <http://arxiv.org/abs/2310.04948>.
- [18] Jin M, Wang S Y, Ma L T, et al. Time-LLM: Time series forecasting by reprogramming large language models[J/OL]. (2024-01-29) [2024-03-19]. <http://arxiv.org/abs/2310.01728>.
- [19] Shi X, Xue S, Wang K, et al. Language models can improve event prediction by few-shot abductive reasoning[J]. *Advances in Neural Information Processing Systems*, 2023, 36: 29532-29557.
- [20] Cao D F, Ye W, Zhang Y Z, et al. TimeDiT: General-purpose diffusion transformers for time series foundation mode[J/OL]. (2025-02-11) [2025-04-27]. <http://arxiv.org/abs/2409.02322>.
- [21] Shi X, Wang S Y, Nie Y Q, et al. Time-MoE: Billion-scale time series foundation models with mixture of experts [J/OL]. (2024-10-02) [2024-12-13]. <http://arxiv.org/abs/2409.16040>.
- [22] Liu X, Liu J C, Woo G, et al. Moirai-MoE: Empowering time series foundation models with sparse mixture of expertss[J/OL]. (2024-10-14) [2024-12-13]. <http://arxiv.org/abs/2410.10469>.
- [23] Ekambaram V, Jati A, Nguyen N H, et al. Tiny time mixers (TTMs): Fast pre-trained models for enhanced zero/few-shot forecasting of multivariate time serie[J/OL]. (2024-11-07) [2024-12-13]. <http://arxiv.org/abs/2401.03955>.
- [24] Liu Y, Zhang H R, Li C Y, et al. Timer: Generative pre-trained transformers are large time series models[J/OL]. (2024-10-18) [2024-12-13]. <http://arxiv.org/abs/2402.02368>.
- [25] Gao S, Koker T, Queen O, et al. UniTS: A unified multi-task time series model[J/OL]. (2024-11-25) [2024-12-12]. <http://arxiv.org/abs/2403.00131>.
- [26] Woo G, Liu C, Kumar A, et al. Unified training of universal time series forecasting transformers[J/OL]. (2024-05-22) [2024-12-12]. <http://arxiv.org/abs/2402.02592>.
- [27] Goswami M, Szafer K, Choudhry A, et al. MOMENT: A family of open time-series foundation models[J/OL]. (2024-10-10) [2024-12-12]. <http://arxiv.org/abs/2402.03885>.
- [28] Ansari A F, Stella L, Turkmen C, et al. Chronos: Learning the language of time series[J/OL]. (2024-11-04) [2024-12-12]. <http://arxiv.org/abs/2403.07815>.
- [29] Dooley S, Khurana G S, Mohapatra C, et al. ForecastPFN: Synthetically-trained zero-shot forecasting[J/OL]. (2023-11-03) [2024-12-14]. <http://arxiv.org/abs/2311.01933>.
- [30] Rasul K, Ashok A, Williams A R, et al. Lag-Llama: Towards foundation models for time series forecasting [C]//R0-FoMo: Robustness of Few-shot and Zero-shot Learning in Large Foundation Models, 2023.
- [31] Das A, Kong W, Sen R, et al. A decoder-only foundation model for time-series forecasting[J/OL]. (2024-04-17) [2024-12-12]. <http://arxiv.org/abs/2310.10688>.
- [32] Garza A, Challu C, Mergenthaler-Canseco M. TimeGPT-1 [J/OL]. (2024-05-27) [2024-12-12]. <http://arxiv.org/abs/2310.03589>.
- [33] Hao X X, Chen W, Yan Y B, et al. UrbanVLP: Multi-granularity vision-language pretraining for urban socioeconomic indicator prediction[J/OL]. (2025-01-22) [2025-04-27]. <http://arxiv.org/abs/2403.16831>.
- [34] Balsebre P, Huang W M, Cong G, et al. City foundation models for learning general purpose representations from OpenStreetMap[C]//Proceedings of the 33rd ACM International Conference on Information and Knowledge Management. New York: ACM, 2024: 87-97.
- [35] Zhu Y, Yu J J, Zhao X, et al. UniTraj: Learning a universal trajectory foundation model from billion-scale worldwide traces[J/OL]. (2024-11-16) [2025-04-27]. <http://arxiv.org/abs/2411.03859>.
- [36] Lin Y, Wei T L, Zhou Z Y, et al. TrajFM: A vehicle trajectory foundation model for region and task transferability [J/OL]. (2024-08-09) [2025-04-27]. <http://arxiv.org/abs/2408.15251>.
- [37] Lin Y, Liu Y C, Zhou Z Y, et al. PTrajM: Efficient and semantic-rich trajectory learning with pretrained trajectory-mamba[J/OL]. (2024-08-09) [2025-04-27]. <http://arxiv.org/abs/2408.15251>.

- org/abs/2408.04916.
- [38] Li Z H, Xia L, Shi L, et al. OpenCity: Open spatio-temporal foundation models for traffic prediction[J/OL]. (2024-08-16) [2025-04-27]. <http://arxiv.org/abs/2408.10269>.
- [39] Lam R, Sanchez-Gonzalez A, Willson M, et al. Learning skillful medium-range global weather forecasting[J]. *Science*, 2023, 382(6677): 1416–1421.
- [40] Yuan Y, Ding J T, Feng J, et al. UniST: A prompt-empowered universal model for urban spatio-temporal prediction [C]//Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. New York: ACM, 2024: 4095–4106.
- [41] Yuan Y, Han C H, Ding J T, et al. UrbanDiT: A foundation model for open-world urban spatio-temporal learning [J/OL]. (2024-11-19) [2025-04-27]. <http://arxiv.org/abs/2411.12164>.
- [42] Nguyen T, Brandstetter J, Kapoor A, et al. ClimaX: A foundation model for weather and climate[J/OL]. (2023-12-18) [2025-04-27]. <http://arxiv.org/abs/2301.10343>.
- [43] Chen K, Han T, Gong J, et al. FengWu: Pushing the skillful global medium-range weather forecast beyond 10 days lead[J/OL]. (2023-04-06) [2025-02-08]. <http://arxiv.org/abs/2304.02948>.
- [44] Bi K F, Xie L X, Zhang H H, et al. Accurate medium-range global weather forecasting with 3D neural networks[J]. *Nature*, 2023, 619(7970): 533–538.
- [45] 赵延玉, 赵晓永, 王磊, 等. 可解释人工智能研究综述[J]. *计算机工程与应用*, 2023, 59(14): 1–14.
- [46] 王金桥, 杨蓓莹. 多模态大模型的发展综述及思考[J/OL]. *无线电工程*, 1–21. [2025-09-04]. <https://link.cnki.net/urlid/13.1097.TN.20250812.0837.002>.
- [47] Liu F, Zhang T S, Dai W W, et al. Few-shot adaptation of multi-modal foundation models: A survey[J]. *Artificial Intelligence Review*, 2024, 57(10): 268.
- [48] Wharrie S, Eick L, Makinen L, et al. Bayesian meta-learning for improving generalizability of health prediction models with similar causal mechanisms[J/OL]. (2024-12-30) [2025-04-30]. <http://arxiv.org/abs/2310.12595>.
- [49] Munkhdalai T, Faruqi M, Gopal S. Leave No context behind: Efficient infinite context transformers with infinite attention[J/OL]. (2024-08-09) [2025-05-06]. <http://arxiv.org/abs/2404.07143>.
- [50] Ma X Z, Yang X M, Xiong W H, et al. *Megalodon*: Efficient LLM pretraining and inference with unlimited context length[J/OL]. (2024-04-16) [2025-05-06]. <http://arxiv.org/abs/2404.08801>.

Research progress of large models for time series and spatio-temporal data analysis

LUO Yuanbo¹, SUN Jia¹, TAO Lizhi^{2*}

1. Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), Guangzhou 511458, China

2. Key Laboratory of Poyang Lake Wetland and Watershed Research of Ministry of Education, Jiangxi Normal University, Nanchang 330022, China

Abstract In recent years, large models have made breakthroughs in natural language processing and computer vision by virtue of their powerful sequence modeling capability, excellent representation learning potential, and flexible pre-training–fine-tuning paradigm, which also bring new development opportunities for time-series and spatio-temporal data intelligent analysis and are expected to revolutionize the analysis paradigm. This paper provides a systematic review of research progress on large models for time series and spatio-temporal data analysis, focusing on two major directions: empowering large language models and building dedicated foundation models. The former leverages prompt engineering, tokenization, and parameter-efficient fine-tuning to adapt large models to time series and spatio-temporal tasks, while the latter employs large-scale cross-domain pre-training to establish unified dynamic representations. The path of empowering large language models offers advantages such as low development costs and flexible zero-shot/few-shot transfer learning, while specialized foundational models demonstrate superior cross-domain generalization capabilities. At the same time, both approaches still face challenges including insufficient interpretability and difficulties in multimodal semantic alignment. Future research urgently requires breakthroughs in enhancing interpretability, advancing multimodal joint modeling, and innovating model architectures.

Keywords time series; spatio-temporal data; large language models; pre-trained foundation models ●



(责任编辑 赵庆圆)