

太阳风暴中的“意外”:质子通量反峰与AI预测

罗海艺 唐智宸

(重庆育才中学校 400050)

一、开篇:从一次意外的科学发现说起

2012年秋季的一个傍晚,北半球的夜空突然被耀眼的绿紫色光幕吞没——一场罕见的太阳风暴正以电光火石之势席卷地球。人们在惊叹极光奇观的同时,地面上的电网传感器却连续发出警报。科学家们惊讶地发现:质子通量(指单位时间穿过单位面积的质子数量)并未如预期般大幅上升,反而出现一次显著的“反峰”——数据曲线在高点处骤然下滑,随后又迅速恢复正常。

根据NASA ACE卫星EPAM探测器与多座地面观测站的同步数据,这段异常波形被迅速拼接成一幅复杂的反常波动图,揭示出太阳风暴与质子通量之间并非简单的线性对应关系。质子是太阳喷发的主要粒子;一旦其通量飙升,卫星电子设备将受到剧烈冲击,导航系统也可能陷入瘫痪。2015年,一架瑞典航空波德兰—斯德哥尔摩航班就在太阳风暴突袭中被迫返航,上百名乘客滞留数小时才得以安全降落——无法准确预警的质子风暴正悄然威胁着我们的高科技生活与安全出行。

当极光的帷幕褪去,夜空恢复宁静,科学家们的疑问却更加深邃:太阳的暴怒究竟如何在抵达地球前被“化解”?隐藏在“反峰”背后的秘密是什么?

二、科学背景:太空中的“攻防战”

太阳活动机制

太阳这颗孕育地球生命的恒星,其表面不断爆发的耀斑与日冕物质抛射(CME)是引发地球空间

风暴的根源。耀斑源于日冕磁场的重联,瞬间释放出高能粒子并以光速扩散;CME则携带庞大等离子体与磁场结构,以百万公里每小时的速度掷向太空。在2012年那次事件中,太阳表面的巨大活跃区连续数十次爆发耀斑,随之而来的CME体量之大,仿佛在太空中点燃了一把看不见的火炬。

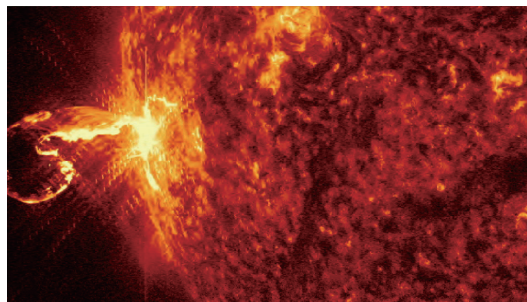


图1 太阳活跃区的示意图

地球防御体系

地球自带一层无形却坚固的“盾牌”——磁层。它由地磁场与太阳风相互作用形成,磁层顶在朝向太阳一侧距地球约十万公里,背向太阳一侧的磁尾可延伸数百万公里。高能质子在此过程中或偏转向太空,或沿磁力线坠入两极,撞击大气原子,激发

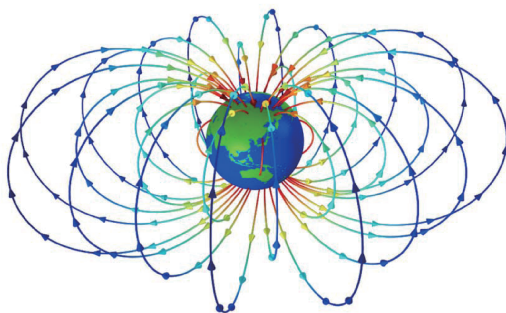


图2 地磁场示意图

绚烂的极光灯光秀。

一个关键物理量是地磁刚度,表示带电粒子要想穿透地球磁场所需的最小动量。能量较低(数百至数千 eV)的太阳风质子大多被磁层阻挡;而来自深空的高能宇宙线质子则可能直接穿越磁层。此外,强烈太阳活动(如大型 CME)期间抛出的低能质子云,在太阳周围形成暂时的屏蔽“外壳”,进一步加强了对外源高能粒子的屏蔽效应。这一复杂交互正是部分强烈活动期出现质子通量“反峰”现象的关键。

三、质子通量预测

对地球周围质子通量的短期估算已可借助 AP9、CREME2009 等经验模型给出可靠范围,广泛用于航天器设计与太空环境评估。然而,要“提前知道”明天是否会出现质子爆发,以及通量具体值,现有理论仍显不足。

主要难点在于:① 太阳高能粒子事件具有随机性;② 质子在传播过程中受太阳风、行星际磁场等多重因素影响,使其行为更加复杂。换言之,虽然我们已能像看天气预报一样判断质子环境的大体趋势,要实现长时间尺度精确的预测,仍需新的突破。

近年来,研究人员尝试引入 xLSTM、Transformer 等深度学习架构,以历史观测序列训练模型,自动提取非线性时序特征,在数小时至数天尺

度上提升预测精度。若能进一步获得更多维度、分辨率更高的数据,并与物理约束模型耦合,这类方法有望成为未来空间天气预报的重要支撑技术。

四、用于质子通量预测的 Transformer+xLSTM 混合模型

技术让计算机能够像人一样,从大量历史数据中学习规律和模式。我们采用的创新方法结合了两种先进的 AI 模型:

1. Transformer 模型:就像一个注意力非凡的分析师,能够同时关注大量数据中的关键信息,发现不同时间点之间的关联。

2. xLSTM 模型:这是传统记忆模型的升级版,特别擅长捕捉时间序列中的长期规律和突然变化。

强强联合的预测系统。我们将这两个模型有机结合,创建了一个智能预测系统。这个系统的工作原理可以分为三个步骤:

学习阶段:系统首先“学习”过去十年的质子通量数据,就像学生复习历史知识一样。

分析阶段:模型会识别出数据中的复杂模式,包括那些人类难以察觉的细微规律。

预测阶段:基于学到的知识,系统能够预测未来一段时间内质子通量的变化趋势。

为什么这个方法更有效呢?传统的预测方法就像只用一种工具解决所有问题,而我们的方法更

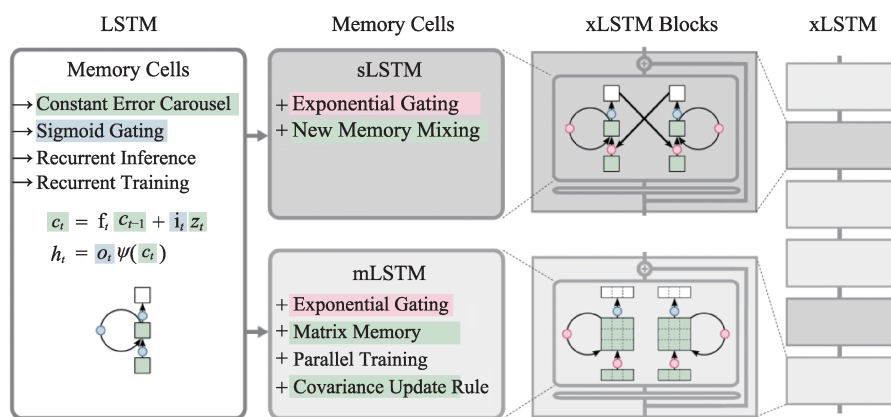


图3 xLSTM 内部机制图^[4]

像是为太空天气预报量身定制的“智能工具箱”。Transformer 模型负责把握大局, xLSTM 模型则专注于理解时间变化规律, 两者配合相得益彰。

模型总体架构

基于上述记忆机制, 本研究构建了如图 4 所示的融合式 Transformer-xLSTM 时序建模框架。整个预测系统的工作流程可以类比为一个专业分析团

队的合作过程, 其运行机制包含以下几个关键阶段: 输入序列首先通过嵌入层映射至高维空间, 这相当于将原始的质子通量数据“翻译”成计算机更容易处理的数字化语言。随后, 数据被送入多头自注意力机制(Multi-Head Attention), 这个过程就像是一组专家同时从不同角度分析数据, 各自关注时间序列中的不同特征模式, 共同捕捉全局依赖关系。

注意力机制的输出结果随后被传递给 xLSTM

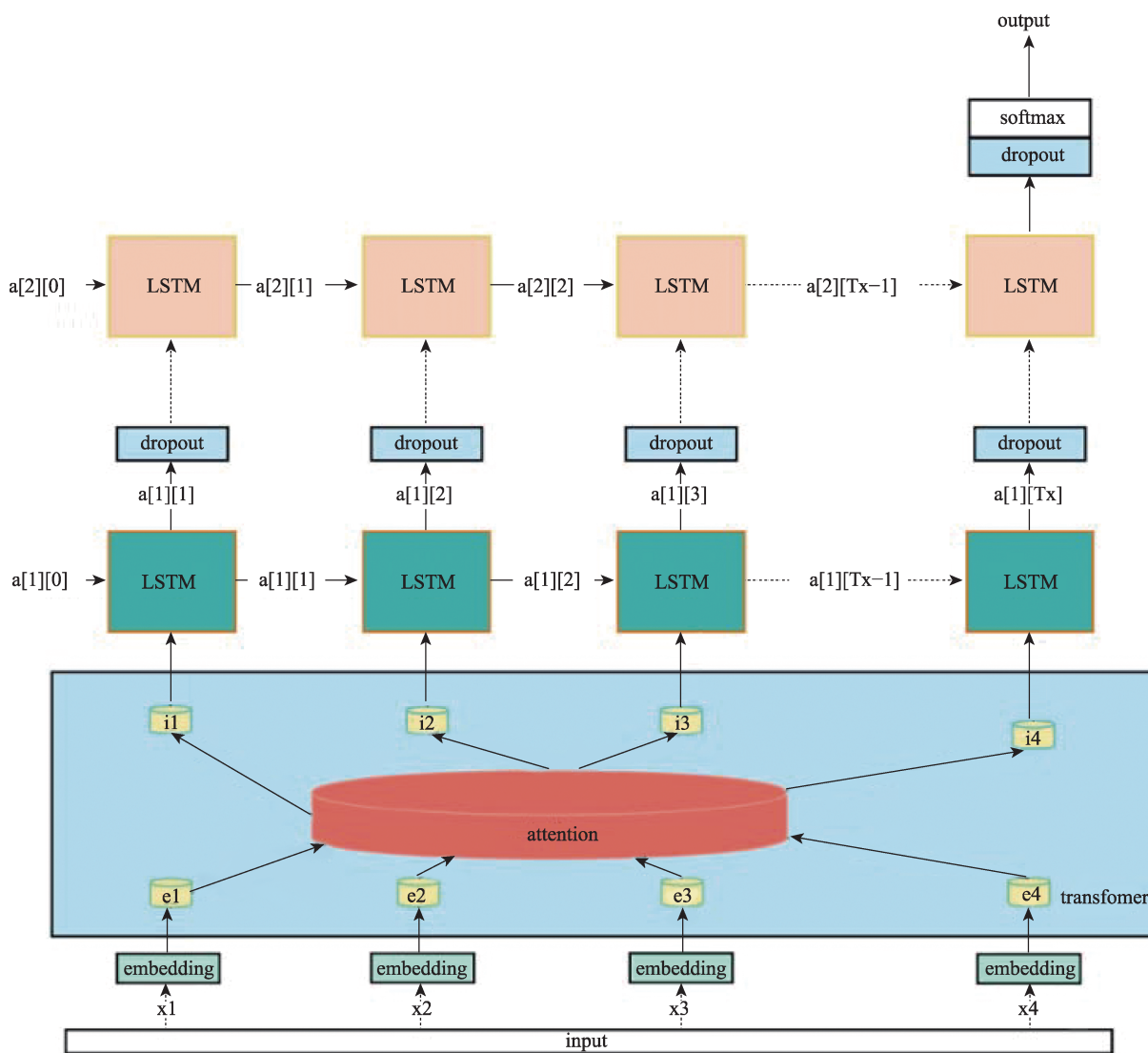


图4 Transformer + xLSTM 预测架构图。模型接收输入序列(x_1, x_2, x_3, x_4), 首先通过嵌入层将其转换为高维向量表示(e_1, e_2, e_3, e_4)。随后, 这些向量经过多头注意力机制处理, 生成蕴含全局不同时间步数下数据关联性信息的加权输出(i_1, i_2, i_3, i_4)。最终, 这些信息被送入双向 LSTM 网络进行深度时序建模, 其各层、各时间步的隐藏状态输出记为 $a[\text{层数}][\text{时间步}]$ (如 $a[1][1]$)。该状态封装了网络至当前时刻的综合记忆与理解。整个架构通过层层递进的特征提取与融合, 实现对复杂时序规律的精准建模

模块,该模块能够结合之前的状态信息实现动态更新,相当于为系统赋予了“记忆能力”,使其能够记住重要的历史信息。在此基础上,模型还通过堆叠的双向LSTM层进一步整合不同时间步数下数据关联性信息,这就像是在分析当前情况时,同时参考过去与当下的相关背景。为了防止模型过度依赖训练数据中的特定模式(过拟合),系统采用了Dropout正则化技术,这相当于偶尔让分析员忽略一些信息,以保持判断的泛化能力。在解码阶段,系统采用对称的xLSTM模块对输出序列进行建模,最终通过线性变换与softmax层输出未来时间步的预测结果。整体而言,这个架构自底向上实现了从局部特征提取、全局不同时间步数下数据关联性融合到未来趋势建模的完整过程。它巧妙地结合了Transformer架构强大的特征表达能力与xLSTM优异的时序记忆机制,在质子通量预测这一复杂任务中展现出了良好的建模深度与预测性能。这种设计使得系统既能够把握大数据中的宏观规律,又能捕捉细微的时序变化特征,从而为太空天气预测提供更加准确可靠的技术支持。

本研究使用的数据集来源于2011年6月至2019年12月AMS-02探测器记录的每日质子通量数据,以及同期卫星和太阳观测站获取的太阳活动数据(包括太阳耀斑强度、日冕物质抛射速度、太阳黑子数和太阳风参数等)。数据经过缺失值处理、异常值剔除和标准化等预处理后,按70%:10%:20%的比例划分为训练集、验证集和测试集,以确保模型的可靠性。

模块结构细节

在模型的核心设计方面,我们融合了多项先进技术,使其像一个高效协作的专家团队,各司其职又密切配合。整个系统的智能核心主要由四个关键部分组成:首先是我们称之为“多专家会诊”机制(多头注意力机制)。这就像同时聘请多位各有所长的分析师,他们分别从不同角度深入研究数据,有人专注识别突然的峰值波动,有人擅长发现周期性

的规律,还有人专门捕捉局部异常。这种并行工作的方式极大地提升了对复杂特征的发掘能力。

第二个核心是增强型记忆中枢(xLSTM记忆模块)。这个组件为系统提供了强大的记忆能力,不仅能够记住近期的重要信息,还能保持对长期规律的深刻理解。就像一位经验丰富的预报员,既能记得最近几天的天气变化,又不会忘记历年同期的气候特征。第三个重要设计是稳定训练的双重保障(残差连接与层归一化)。在深度学习模型训练过程中,容易出现“学不下去”或者“学偏了”的情况。我们引入的这两个技术就像给学习过程加上了安全绳和导航仪,确保训练过程更加平稳可靠,提高了模型的整体性能。最后是深度理解网络。这个组件分布在各个处理层中,负责对特征进行深度加工和转化。它就像团队的智囊团,能够发现数据中隐藏的复杂非线性关系,从而做出更加精准的判断。这些组件并非孤立工作,而是形成了一个有机的整体。“多专家会诊”机制负责发掘数据中的丰富特征,然后将这些特征传递给“增强型记忆中枢”进行时序整合,期间“深度理解网络”不断加深对特征的理解,而“稳定训练的双重保障”则确保整个学习过程高效可靠。这种精心设计的架构使得我们的模型能够充分捕捉质子通量变化中的复杂规律,既把握宏观趋势,又不忽略细微变化,从而实现了预测性能的显著提升。通过这种仿生学的设计理念,我们让人工智能模型模仿了人类专家分析问题的思维过程,既有分工又有合作,既有专注又有整合,这正是我们能够在复杂太空环境预测中取得突破的关键所在。

模型超参数设置

为了让预测模型能够更好地学习和预测质子通量的变化规律,我们对这个“AI大脑”进行了一系列精心设置。首先,我们让模型每次以步长为一天,观察过去连续10天的数据(相当于回顾一段完整的变化过程),基于这些历史信息来预测下一个时间点会发生什么变化。这就像是让气象预报员

表 1 模型的核心配置

参数类别	参数名称	设置值	说明
输入/输出设置	输入时间步长(look_back)	10	使用过去 10 个时间点的数据作为输入
	输出预测步长(T)	1	预测未来 1 个时间点的质子通量值
	编码器/解码器层数	4	Transformer 编码器和解码器各由 4 层结构堆叠而成
模型结构配置	注意力头数(Heads)	4	多头自注意力机制中并行工作的注意力头数量
	嵌入维度(d_model)	32	将输入数据映射到的高维特征空间的维度
	前馈网络隐藏层维度(d_ff)	32	编码器和解码器中前馈神经网络的隐藏层节点数
	损失函数	均方误差 (MSE)	用于衡量模型预测值与真实值差异的函数
训练策略	优化器	Adam	默认使用 Adam 优化器,其学习率见下项
	学习率 (Learning Rate)	0.001	控制模型参数更新步长的关键超参数
	批量大小 (Batch Size)	64	一次训练所选取的样本数
正则化策略	Dropout 比率	0.1	随机丢弃神经元的比例,以防止模型过拟合
模型参数量	总参数量	约 15,000	估算的模型可训练参数总数,直接影响模型复杂度与训练计算成本

分析最近几小时的天气数据,来预测下一小时会不会下雨。

其次,我们的模型采用了对称的 4 层编码器和 4 层解码器结构,这相当于组建了 4 个各司其职的专业分析团队,每个团队都配备了 4 位专项分析师(4 头注意力机制),他们分别从不同角度深入研究数据中的各种特征模式。

第三,在数据处理维度上,我们设置了 32 维的特征空间,这个维度既能够保留足够的细节信息,又不会让模型过于复杂。这好比为分析团队提供了 32 种不同的分析工具,确保他们能够多角度全方位地解析数据。

另外,为了避免模型只是单纯记忆训练数据而缺乏真正的预测能力(这种现象称为“过拟合”),我们引入了轻度 Dropout 机制(比率为 0.1)。这相当于偶尔让分析师忽略极少数信息,促使他们掌握规律而非记住具体数字,从而提升模型的泛化能力。

最后,在训练过程中,我们使用均方误差作为评估预测准确度的标准。我们将学习率设置为 0.001,确保模型能够稳定而扎实地进步,避免学得太快而忽略重要细节。每次训练使用 64 组数据作为一个批次,这样的设置既保证了学习效率,又确保了训练过程的稳定性。通过这些精心设计的参数配置,我们的模型在学习速度与预测精度之间找

到了最佳平衡点,使其能够高效地从历史数据中学习规律,并对未来的质子通量变化做出准确预测。这种配置方案既考虑了模型的表达能力,又兼顾了训练效率,为可靠的太空天气预测奠定了坚实基础。

模型评估

在完成模型构建与训练之后,为全面验证所提模型在质子通量预测任务中的有效性与泛化能力,本文进一步开展了系统的准确性评估分析。通过多种性能指标与可视化手段对模型的预测效果进行量化与对比,评估重点包括趋势拟合能力、误差分布特性以及模型在不同通量水平下的响应表现。图形化结果清晰地展示了模型在捕捉质子通量时间序列主趋势与局部波动方面的表现,为后续对预测精度、稳定性及实际应用潜力的综合判断提供了实证依据。

图 5 比较了模型预测的年度平均质子通量(蓝色圆点)与实际观测值(橙色折线)的变化趋势。从整体趋势上看,模型较好地拟合了质子通量的长期演化规律,包括上升段、平台期以及下行趋势。大部分年份中,预测值与真实值保持较小偏差,且走势一致,表明模型具备一定的全局趋势学习能力。

但在局部如年份索引约为 70~110 之间,预测值波动较小,略显“平滑”,提示模型在应对强烈非线性

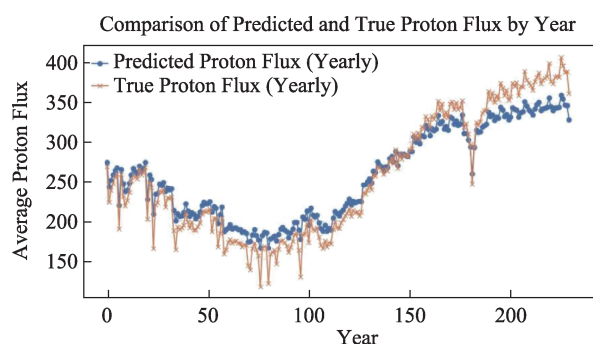


图5 年度质子通量预测值/真实值对比图。横坐标起点为2011年6月,纵坐标单位为 $/(cm^2 \cdot sr \cdot s)$,展示了每平方米每球面度每秒的粒子数

性跳跃(如爆发性通量增加)时存在一定欠拟合。此类特征往往与太阳高活动期相关,说明未来可通过引入辅助物理特征如太阳风速度、F10.7通量等增强模型感知能力。

图6显示了模型预测值相对于真实值的分布,红色虚线表示理想预测线(即预测值等于真实值)。从图中可观测到大多数点靠近该线,尤其在低至中等质子通量区域(0~1000)密集分布在理想预测线周围,说明模型在主流量水平下预测较为准确。图中给出的决定系数 $R^2=0.78$,表示模型对观测值的方差解释能力为78%,为非线性时间序列预测中的良好水平。

然而,在高通量区域(>1200)存在部分预测值聚集在较低位置,出现“压缩效应”或“高值低估”问题。这表明模型在训练集中较少接触高通量样本,可能导致对极端事件响应不足。此问题可通过平

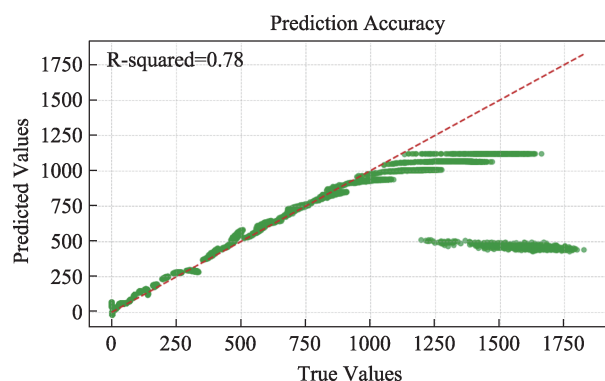


图6 预测精度散点图

衡样本分布或使用分段损失函数予以缓解。

图7展示了模型在10个训练周期(Epoch)中的训练损失与验证损失变化。总体趋势清晰:两者都呈下降趋势,且在第5轮之后趋于平稳,说明模型已接近最优状态,无明显过拟合或欠拟合迹象。

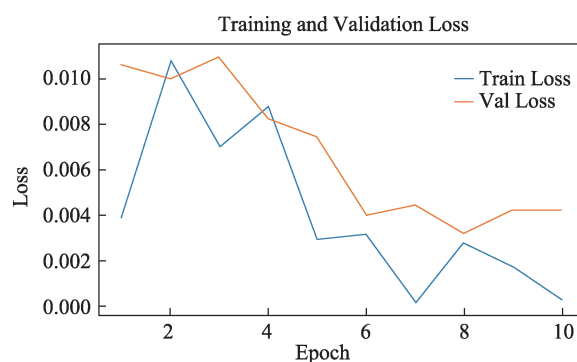


图7 训练与验证损失

值得注意的是,在第3轮左右验证损失短暂上升,可能源于权重更新波动或数据集某一批次的随机性,但并未导致整体性能退化,表明训练过程是稳定且可控的。

模型总结

综合预测结果与性能可视化分析,本研究所提出的Transformer-xLSTM模型在质子通量预测任务中展现出较为优异的整体表现。首先,从趋势拟合角度来看,模型能够有效重建年度质子通量的演化轨迹,准确捕捉周期性波动与趋势性变化,体现出较强的全局建模能力。其次,在主流量数值区间内,模型预测精度较高,决定系数 R^2 达0.78,表明其对常规通量水平具备良好的拟合能力。然而,在高通量极值区域,模型存在一定程度的系统性低估,表明其对突发性或异常增强事件的响应能力尚有提升空间。此外,从误差分布观察可见,预测误差在样本后段呈现缓慢增加趋势,暗示模型在边界样本的预测性能相对较弱,可能与序列不同时间步数下数据关联性信息减少或边界特征异质性有关。最后,从训练过程来看,模型在有限轮次内即实现稳定收敛,训练与验证损失同步下降,未见明显过拟

合,验证了其良好的训练稳定性与泛化性能。综上所述,该模型在质子通量时序预测任务中具有较高的实际应用潜力,为空间天气变化趋势的智能建模与预测提供了有效技术支撑。

参考文献

- [1] Beck, Maximilian & Pöppel, Korbinian & Spanring, Markus & Auer, Andreas & Prudnikova, Oleksandra & Kopp, Michael & Klam-bauer, Günter & Brandstetter, Johannes & Hochreiter, Sepp. (2024). xLSTM: Extended Long Short-Term Memory.
- [2] Hochreiter S, Schmidhuber J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735 – 1780.
- [3] Beck M, Auer A, Pöppel K, et al. xLSTM: Extended Long Short-Term Memory[EB/OL]. (2023-05)[2024-06-04]. <https://arxiv.org/abs/2405.04517>.
- [4] Prudnikova O, et al. Advanced hybrid LSTM-transformer architecture for real-time multi-task prediction in engineering systems[J]. Scientific Reports, 2024, 14: 4890.
- [5] Sak H, Senior A, Beaufays F. Long short-term memory recurrent neural network architectures for large scale acoustic modeling[C]// Proceedings of the 15th Annual Conference of the International Speech Communication Association. 2014.
- [6] Li X, Wu X. Constructing Long Short-Term Memory based Deep Recurrent Neural Networks for Large Vocabulary Speech Recognition[EB/OL]. (2014-10-15)[2024-06-04]. <https://arxiv.org/abs/14-10.4281>.
- [7] Wu Y, Schuster M, Chen Z, et al. Google's neural machine translation system: Bridging the gap between human and machine translation[EB/OL]. (2016-09-26) [2024-06-04]. <https://arxiv.org/abs/16-09.08144>.
- [8] Hochreiter S. Untersuchungen zu dynamischen neuronalen Netzen[D]. München: Technische Universität München, Institut für Informatik, 1991.
- [9] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need [J]. Advances in Neural Information Processing Systems, 2017.

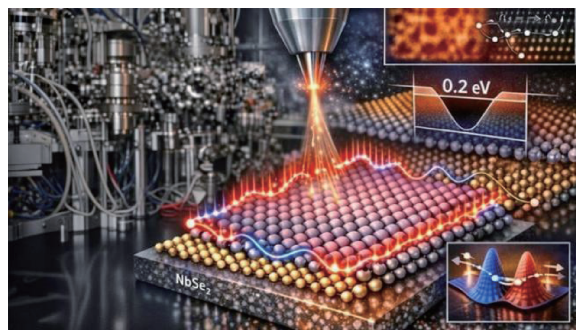
科苑快讯

物理学家终于在实验室中实现了长期预测的二维拓扑晶体

芬兰于韦斯屈莱大学(University of Jyväskylä)和阿尔托大学(Aalto University)的物理学家在实验室中成功地制造了一个二维拓扑晶体绝缘体。这种不寻常的量子材料早在10多年前就被理论预测到,但由于准备合适的材料面临挑战,实验证明很难生产出来。

为了制造这种材料,研究人员在二硒化铌(NbSe_2)衬底上沉积了一层由两层组成的碲化锡(SnTe)原子薄膜。然后,科学家们使用分子束外延和低温扫描隧道显微镜检查了材料的电子行为。这些技术使他们能够以原子级别的精度分析这种结构。

在这个二维系统中,他们检测到沿着材料边缘出现的导电态对。这些边缘状态是拓扑晶体绝缘体的定义特征,并保持稳定,因为它们受到晶格对称性的保护。



这个新实现的系统为研究人员提供了一个强大的实验平台,用于研究可以通过应变调谐的二维拓扑状态。它也可能有助于未来基于自选的电子学和其他纳米技术的发展。

(高凌云编译自2026年3月7日SciTechDaily网站)