

PsyINN: 面向教育数据挖掘的心理信息神经网络

孙建文^{1),2),3)} 陈麒戎^{2),3)} 李伏清⁴⁾ 罗亚威⁵⁾ 李 卿^{1),2),3)}
沈筱譔^{1),2),3)} 刘三女牙^{1),2),3)}

¹⁾(华中师范大学人工智能与教育新形态实验室 武汉 430079)

²⁾(华中师范大学教育大数据应用技术国家工程研究中心 武汉 430079)

³⁾(华中师范大学人工智能教育学部 武汉 430079)

⁴⁾(澳门大学协同创新研究院 澳门 999078)

⁵⁾(浙江大学软件学院 浙江 宁波 315048)

摘 要 随着大数据、人工智能等技术在教育领域的深化应用,教育数据挖掘在智能教育系统构建中扮演着越来越重要的角色,通过对学习者行为数据的挖掘与建模,研究者得以结合认知机制,实现对学习过程中知识掌握状态与认知演化趋势的有效追踪与预测。当前主流方法多依赖数据驱动的神经网络,虽具备较强预测能力,但普遍存在可解释性不足、噪声敏感和泛化能力有限等问题。以物理信息神经网络为代表的知识驱动型神经网络模型展现出良好的泛化性与解释力,有望成为教育数据挖掘的新框架。本文探索了一个融合教育心理学知识与学习者行为数据驱动建模的神经网络框架 PsyINN(Psychology-Informed Neural Network),针对心理学理论模型存在的参数不可观测与理论形式不统一等问题,引入符号回归方法以表达和演化结构先验作为神经网络的软约束,促进神经网络学习符合教育心理学理论的内在表征,并通过动态对齐,实现理论模型与神经网络模型的协同优化。具体而言,PsyINN 采用“双阶段建模策略”,首先,对记忆、练习等教育心理学理论模型进行参数学习,并作为符号回归的初始化方程;其次,基于历史学习行为数据,同时优化符号回归模块与神经网络模型,构建知识对齐函数,实现双向知识蒸馏与联合优化。为适应不同模型能力的差异,PsyINN 设计了动态知识对齐机制,可根据模型当前预测能力调整知识蒸馏强度,从而提升协同效率。本文在数学练习、外语词汇记忆和认知技能习得三个典型学习场景中开展验证,结果表明 PsyINN 可作为一个灵活的插件,显著提升现有主流神经网络模型的预测精度。其中,AUC 平均提升 2.14%,G-Means 平均提升 9.00%,Precision 提升 2.22%。此外,PsyINN 可基于经典理论方程演化出更加契合学习场景的方程,具备良好的结构可解释性与理论一致性。本研究为人类复杂学习行为的机制建模提供了一条可扩展的新路径,有助于推动教育建模从“黑箱预测”迈向“结构表达与机制解释”的融合范式,也有望为“人文社会科学智能”中“数据与机理双驱动”提供通用的建模框架与方法论支持。

关键词 教育数据挖掘;心理信息神经网络;知识蒸馏;深度学习;符号回归

中图法分类号 TP391

DOI 号 10.11897/SP.J.1016.2026.01027

PsyINN: A Psychology-Informed Neural Network for Educational Data Mining

SUN Jian-Wen^{1),2),3)} CHEN Qi-Rong^{2),3)} LI Fu-Qing⁴⁾ LUO Ya-Wei⁵⁾ LI Qing^{1),2),3)}
SHEN Xiao-Xuan^{1),2),3)} LIU San-Nyu-Ya^{1),2),3)}

¹⁾(Laboratory for Artificial Intelligence and New Forms of Education, Central China Normal University, Wuhan 430079)

²⁾(National Engineering Research Center for Educational Big Data, Central China Normal University, Wuhan 430079)

收稿日期:2025-08-13;在线发布日期:2026-02-03。本课题得到新一代人工智能国家科技重大专项项目(No. 2022ZD0117103)、国家自然科学基金重大项目(No. 62293554)、国家自然科学基金面上项目(No. 62577028)、中国自动化学会青少年人工智能创新后备人才培养工程 2025 年度教科研重大项目(No. HBRC-JKYZD-2024-310)、湖北省自然科学基金创新群体项目(No. 2023AFA020)资助。孙建文,博士,教授,中国计算机学会(CCF)会员,主要研究领域为教育数据挖掘、计算学习科学。E-mail:sunjw@ccnu.edu.cn。陈麒戎,硕士研究生,主要研究领域为教育数据挖掘。李伏清,硕士研究生,主要研究领域为智能教育、学习分析。罗亚威,博士,研究员,中国计算机学会(CCF)会员,主要研究领域为计算机视觉、智能教育。李 卿,博士,研究员,主要研究领域为智能教育、数据挖掘。沈筱譔(通信作者),博士,副教授,主要研究领域为深度神经网络、智能教育。E-mail:shenxiaoxuan@ccnu.edu.cn。刘三女牙(通信作者),博士,教授,长江学者,中国计算机学会(CCF)会员,主要研究领域为智能教育、教育大数据。E-mail:liusy027@ccnu.edu.cn。

³⁾(Faculty of Artificial Intelligence in Education, Central China Normal University, Wuhan 430079)

⁴⁾(Institute of Collaborative Innovation, University of Macau, Macao 999078)

⁵⁾(School of Software Technology, Zhejiang University, Ningbo, Zhejiang 315048)

Abstract With the pervasive integration of big data and artificial intelligence technologies within the educational sector, Educational Data Mining (EDM) has assumed a pivotal role in the architecture of next-generation intelligent tutoring systems. By rigorously mining and modeling learner behavioral data, researchers aim to incorporate cognitive mechanisms to effectively track knowledge mastery states and predict cognitive evolution trends throughout the complex learning process. However, the majority of contemporary approaches rely heavily on data-driven neural networks. While these models demonstrate robust predictive capabilities, they frequently suffer from inherent limitations, including a “black-box” nature that lacks interpretability, high sensitivity to sparse or noisy educational data, and constrained generalization across diverse learning contexts. In contrast, knowledge-driven paradigms—exemplified by Physics-Informed Neural Networks (PINNs)—have exhibited superior generalization and explanatory power by embedding physical laws into deep learning architectures, suggesting a promising new direction for EDM. Drawing inspiration from this paradigm, this paper proposes PsyINN (Psychology-Informed Neural Network), a novel framework that synergistically integrates educational psychology knowledge with behavior-driven modeling. Addressing the unique challenges of cognitive psychology—specifically the issues of unobservable latent parameters and inconsistent theoretical formulations—PsyINN introduces symbolic regression to represent and evolve structural priors. These priors function as “soft constraints” within the loss function, guiding the neural network to learn internal representations that are mathematically aligned with established psychological theories. Furthermore, the framework achieves a collaborative optimization between theoretical constructs and neural computations via a dynamic alignment strategy. Methodologically, PsyINN adopts a comprehensive “two-stage modeling strategy”. In the first stage, theoretical educational psychology models governing memory and practice are parameterized. These parameterized models are then utilized as initialization equations for symbolic regression. In the second stage, utilizing historical learning data, the symbolic regression module and the neural network undergo joint optimization. This is facilitated by a sophisticated knowledge alignment function that supports bidirectional knowledge distillation, ensuring that the data-driven and theory-driven components mutually enhance one another. To accommodate the fluctuating capabilities of different models during training, PsyINN incorporates a dynamic knowledge alignment mechanism. This mechanism adaptively adjusts the intensity of knowledge distillation based on the model’s real-time predictive performance, thereby maximizing collaborative efficiency and preventing negative transfer. The framework was extensively validated across three representative learning scenarios: mathematical problem solving, foreign language vocabulary memorization, and general cognitive skill acquisition. Experimental results indicate that PsyINN functions as a highly flexible plug-in module that significantly bolsters the predictive precision of mainstream neural network models. Specifically, the integration of PsyINN yielded an average improvement of 2.14% in AUC, 9.00% in G-Means, and 2.22% in Precision. Beyond quantitative metrics, PsyINN demonstrates the unique ability to evolve classical theoretical equations into refined forms that are better suited to specific empirical learning scenarios, offering profound structural interpretability and theoretical consistency. In conclusion, this research provides a scalable and robust pathway for the mechanism-oriented modeling of complex learning behaviors. It promotes a fundamental paradigm shift in educational modeling, moving from opaque “black-box prediction” toward a transparent fusion of

“structural expression and mechanism interpretation”. Ultimately, this work offers a generalizable modeling framework and methodological support for the broader vision of “AI for Social Sciences and Humanities”, successfully bridging the divide between data-driven intelligence and mechanistic reasoning.

Keywords educational data mining; psychology-informed neural networks; knowledge distillation; deep learning; symbolic regression

1 引言

在当前教育数字化转型和智能化升级的时代潮流下,教育科学研究正迈向数智驱动的新范式^[1]。教育数据挖掘(Educational Data Mining, EDM)作为促进教育研究范式跃迁的关键技术,旨在从学习行为数据中提取潜在规律,以辅助理解学习机制、优化教学策略设计、提升个性化精准化教育体验,其典型任务包括知识追踪、成绩预测、路径规划等,广泛应用于各类智能教育系统。

近年来,深度神经网络技术在教育数据挖掘中的应用不断深化,以其强大的非线性建模与多层次特征表达能力,在多个教育数据挖掘任务中都取得了显著的预测精度提升。然而,这类方法高度依赖训练数据质量,同时,模型结构复杂难以解释预测结果背后的因果机制,存在典型的“黑箱”问题。

为克服传统数据驱动建模在泛化性、鲁棒性与可解释性方面的固有局限,研究者提出了“知识驱动”建模的新范式。在自然科学领域,这一思路已通过物理信息神经网络(Physics-Informed Neural Networks, PINN)^[2]获得广泛成功。PINN 通过将微分方程等物理定律直接嵌入神经网络训练过程,即使在仅有少量观测数据的情况下,也能在保持高预测精度的同时,实现对问题结构的物理一致性建模。相比传统神经网络,PINN 在数据稀疏、边界不明确或目标函数复杂的情境下展现出更强的泛化能力、鲁棒性与可解释性。更重要的是,这种结构化知识不仅提升了模型性能,还能作为解释模型行为的“代理机制”,赋予网络以符号层面的推理能力。在此启发下,不禁思考是否能够将类似的“结构性先验知识”引入教育数据挖掘任务中?学习过程本质上由一系列心理与认知机制驱动,如记忆的保持与遗忘、知识的迁移现象、技能的习得与整合等。如果能够将这些学习过程背后的理论知识系统性地引入建模,或

许能够同时兼顾数据驱动方法的灵活性与理论建模的解释力,从而实现“学习机制表达”的自动建构。但如何从海量学习者行为数据中提取具备教育心理学含义、可泛化的建模结构,将成为教育数据挖掘领域的关键挑战。这与当前“人文社会科学智能(AI for Social Sciences and Humanities, AI4SSH)”倡导的研究范式高度契合,即通过引入领域结构知识,引导机器学习模型在解释力与泛化力之间取得平衡,实现从“纯粹数据驱动”向“数据与机理双驱动”的转型^[3-5]。其中,AI for Educational Science(AI4ES)^[6]和 AI for Learning Sciences(AI4LS)^[7]作为教育学与学习科学研究的最新范式,主张引入认知理论与心理机制指导建模过程,从人类学习行为数据中揭示潜在的认知规律,并将这些规律知识反馈用于模型优化与个性化教学干预,以提升模型的解释力、泛化性和教育应用价值。刘三女牙等人^[8]提出了一种结合神经网络与符号回归的两阶段建模方法,从大规模自然发生的认知训练日志数据中自动提取技能习得的符号化规律方程;孙建文等人^[9]则构建了 SP-syINN 模型,将经典记忆理论与神经网络模型有机融合,并通过动态异步优化实现知识对齐与协同演化。这些研究在模型结构表达、心理机制嵌入和预测性能提升等方面取得了良好效果,为探索“结构性先验知识”在教育数据挖掘任务中的初步尝试提供了有益参考。

尽管 PINN 在自然科学领域取得广泛和显著的成功,但将其机制直接迁移到教育数据挖掘建模中仍面临诸多挑战。如图 1 所示,与物理规律(如单摆运动)高度形式化、参数易观测的特性相比,教育心理学理论存在显著不同:首先,心理学理论形式化程度较低,大多以定性描述为主,缺乏统一、严谨的数学表达框架,难以直接构造类似物理定律的显式约束条件。其次,即便尝试将部分认知理论(如遗忘规律)结构化,也会面临关键变量难以观测的问题。这些心理参数具有高度的个体差异性和动态性,难以稳定量化并嵌入模型。更为关键的是,理论的生态

效度较低。已有心理学理论多数建立于受控实验环境,往往未能充分覆盖真实学习环境中的多样性与复杂性。例如,词汇记忆、数学练习、技能训练等任务涉及的认知过程和学习机制本质上不同,难以通过单一理论统一建模。即便构建了多个理论约束,如何组合、适配与动态调整这些理论以匹配具体教育任务和学习情境,本身也是一项巨大的建模挑战。因此,在教育数据挖掘任务中,难以像 PINN 那样将理论知识作为“完美约束”直接硬编码进模型。相反,应当将理论视为一种结构性先验或软约束,并与神经网络的表示能力协同优化,以实现数据驱动与理论引导的融合建模。这种柔性融合方式不仅更契合教育场景的异质性,也有助于提升模型的通用性与可解释性^[10]。与此同时,教育数据挖掘建模中理论知识通常更适合以可演化的结构性先验知识参与建模过程,与神经网络形成动态互补。随着训练深入,理论知识与数据驱动的模式表征能力将呈现主导权的交替变化,因此,有必要设计双模动态协同机制以实现持续互导与联动演化。

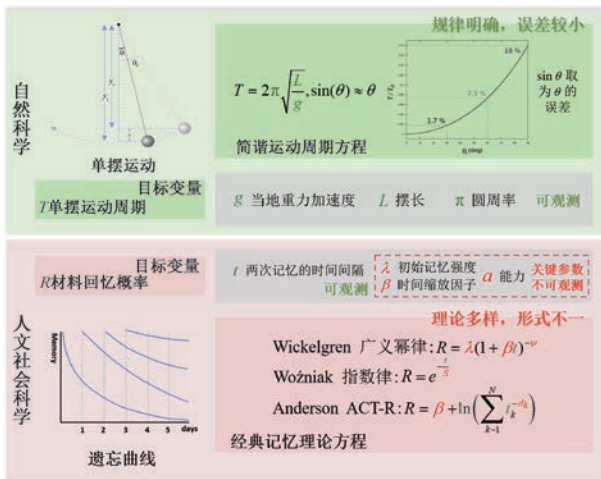


图1 规律方程对比

为融合教育心理学理论的结构性先验知识与神经网络的数据学习能力,本文提出一个知识引导与数据驱动协同的统一框架——心理信息神经网络(Psychology-Informed Neural Network, PsyINN)。该框架构建了符号回归与神经网络双模块,并在训练过程中引入基于模型性能动态演化的协同优化机制。其目标在于兼顾教育心理学理论的可解释性与神经网络的强大表征学习能力,从而提升模型在复杂真实学习场景下的预测性能与理论知识一致性。由于教育心理学理论普遍存在形式化程度低、参数难观测、适用性受限等问题,难以像物理定律那样直

接嵌入模型作为硬约束,本文引入符号回归作为理论表达的软约束方式。符号回归不仅能以数学结构形式表达经典学习规律,还具备演化方程的结构调整能力,使其能够在建模过程中自适应优化,兼顾理论引导与场景适应性。PsyINN 采用“双阶段建模策略”:第一阶段,基于已有心理学理论(如记忆规律、练习规律),构建初始的符号回归结构,在模型训练早期提供结构化先验知识引导;第二阶段,在符号模型的理论引导基础上,引入神经网络等数据驱动建模方法,通过联动优化实现表达式结构与深度表征学习之间的协同演化,从而增强模型的任务适应性与结构灵活性。进一步地,考虑到符号模型与神经网络在不同优化阶段的学习能力存在差异,本文设计了动态知识对齐损失函数(Dynamic Knowledge Alignment Loss, DKAL)。该机制根据神经网络当前预测能力,动态调整符号模型对其的约束强度:在训练初期强调理论主导,强化先验知识;而在模型性能提升后逐步降低约束强度,增强模型的表征灵活性。本质上,DKAL 实现了知识引导与特征学习之间的柔性协同与双向知识蒸馏,不仅有效提升了模型性能,也增强了其对心理学理论的契合度与可解释性。具体而言,本研究主要贡献如下:

(1) 提出融合教育心理学理论与神经网络的通用框架 PsyINN,其“双阶段建模策略”实现了先验知识引导与行为数据驱动建模的动态协同;

(2) 设计了动态知识对齐损失函数(DKAL),通过双向知识蒸馏机制实现符号模型与神经网络的有效融合与联合优化;

(3) 在三种典型学习场景上的实验验证了 PsyINN 在提升模型预测性能(AUC, G-Means, Precision)和可解释性(演化表达式契合理论)方面的有效性。

PsyINN 在整体建模过程中采用“先验知识引导+行为数据驱动”的建模思想,实现了模型结构的自适应演化。相比于传统的 PINN 方法,它无需显式嵌入具体函数表达,便可自动挖掘与任务场景相契合的心理机制结构,从而为教育数据挖掘提供了一种能够结合心理学理论并提升模型可解释性的建模路径。

2 相关工作

2.1 学习规律

人类学习过程由多个认知机制共同驱动,包括

记忆与遗忘、巩固练习与技能习得以及知识的迁移现象等。长期以来,心理学家围绕这些核心认知过程进行了深入研究,并通过数学方程对其进行建模与描述,为学习规律建模提供了理论基础。

2.1.1 记忆与遗忘

记忆与遗忘是学习过程中最基础的认知机制之一,指个体在信息接收后,随时间推移对其保持或消退的过程。早期研究中,艾宾浩斯(Ebbinghaus)于1885年首次系统研究遗忘现象,提出著名的“遗忘曲线”理论^[11],认为“学习后记忆衰退具有快速初期和缓慢后期的特征”。这一规律最初被用指数函数或幂函数进行建模,奠定了后续关于记忆保持与衰退的理论基础。例如,Wickelgren^[12]在1974年提出广义幂律模型($R = \lambda(1 + \beta t)^{-\psi}$),强调记忆强度与时间衰退之间的非线性关系;Wozniak^[13]在1995年引入双成分指数模型($R = e^{-t/S}$),以描述长期记忆中的遗忘规律;而Anderson^[14]则在2004年提出基于认知架构的ACT-R模型,将回忆行为建模为多个记忆痕迹的联合激活($R = \beta + \ln(\sum_k t_k^{-d_k})$),更接近对复杂人类认知过程的刻画。Pashler^[15]提出的MCM模型通过加权多个记忆轨迹模拟重复学习的效果;Lindsey^[16]提出的DASH模型引入认知状态、题目难度与练习历史的多因素建模;2016年,Half-Life Regression^[17](HLR)则以“记忆半衰期”为核心变量,量化知识点保持时间,推动记忆预测精度的发展。

2.1.2 练习规律

早在20世纪初,Thorndike^[18]提出练习律(Law of Exercise),指出“练习越多,联系越强”,揭示了练习在促进记忆稳固与技能内化过程中的作用。这一观点成为后续技能学习建模的重要理论基础。在练习律的理论框架下,巩固练习与技能习得被视为通过重复性行为训练以优化行为表现的关键过程。其成效具体量化为:执行时效的显著缩短、操作精准度的提升,以及认知与动作自动化水平的增强。Fitts与Posner提出了经典的三阶段模型^[19],将技能形成过程划分为认知阶段(cognitive phase)、联结阶段(associative phase)和自动化阶段(autonomous phase)。随着研究的深入,这一阶段性认知转化过程被逐步量化建模,用以描述练习对执行速度与正确率的影响,其中,幂律方程(Power Law of Practice)成为最具代表性的理论框架之一。具体而言,在技能习得研究中,指数模型(Exponential Mod-

el)、双曲模型(Hyperbolic Model)与幂律模型(Power Law Model)被广泛用于刻画练习次数与执行表现之间的变化关系。虽然三者 in 数学形式上有所不同,但均能有效反映练习促进技能发展的基本趋势。指数模型的基本表达式为 $T = A + Be^{-\alpha N}$,其中, T 表示完成任务所需时间, A 是随着练习次数趋于无穷大时的理论最低执行时间, B 代表初始表现与极限水平之间的差距, α 为学习速率,决定了时间下降的快慢。该模型特点是初期执行时间下降迅速,随后逐渐趋于稳定。双曲模型则表示为 $T = A + B/(N + E)$,其中, E 为偏移参数,用以调整起始练习次数,模型描绘了执行时间以双曲速率下降的过程,下降趋势较为平缓,强调了初期进步显著、后期改进减缓的特征。幂律模型则采用表达式 $T = A + B(N + E)^{-\alpha}$,结合了双曲特性与灵活的下降速率调节,其中 α 同样反映学习速度, E 用于平滑起点,揭示了执行时间按照幂律模型持续下降,体现出长期缓慢而稳定的技能改进趋势。总体而言,虽然这三种模型在实验室条件下均能较好地拟合技能习得的学习曲线,但在实际应用中,模型的选择需依据具体任务情境而定,且目前尚无统一、明确的选用标准。此外,由于这些模型大多基于受控实验环境所得,受限于外部变量干扰,其可重复性与对复杂现实环境的外推性仍面临一定挑战。

2.1.3 学习迁移现象

学习迁移现象是指个体将已有的知识、技能或策略应用于新任务或情境的过程,是学习效果的重要体现。早期的“共同要素理论”认为,任务间的相似性决定了迁移效果^[20];而后续理论则强调迁移依赖于对任务结构的理解与抽象表征的建立^[21]。在建模方面,尽管学习迁移现象尚未形成如遗忘曲线或练习律那样统一的函数表达,但已有研究尝试对迁移现象进行形式化描述。例如,结合行为与神经数据探讨迁移机制^[22]和从系统理论角度提出迁移学习形式化框架^[23]。

尽管这些模型在解释学习迁移现象方面提供了重要启示,但它们仍面临三大挑战:在理论表达方面,理论模型多基于简化假设,难以涵盖真实学习场景中的多元复杂因素;在模型实现方面,许多变量(如记忆强度、偏移参数)难以量化或观测,限制了模型的实际适用性;在大规模应用方面,传统模型在预测学习表现方面的精度与泛化能力有限,难以适用大规模在线学习场景。因此,如何在保持理论深度的同时引入数据驱动的方法,整合心理学机制与计

算建模,成为提升学习模型解释力与预测力的关键课题,也是本文试图回应的核心挑战。

2.2 行为数据驱动的教育数据挖掘

随着在线教育平台和智能学习应用的普及,学习者在交互过程中的行为数据(如答题记录、学习路径、点击行为、复习频率等)被大规模收集,为深入挖掘学习规律提供了前所未有的机会。在这一背景下,数据驱动的方法逐渐成为教育数据挖掘的主流方向。研究者通过机器学习、深度学习和统计建模技术,从大量真实学习数据中提取影响学习成效的关键因素,构建能预测学习表现、优化学习策略的智能模型。

知识追踪(Knowledge Tracing, KT)是当前最具代表性的任务之一^[24],其目标是通过建模学习者在不同时间点对知识点的掌握程度,进而预测其未来的学习表现。早期的 KT 模型多基于贝叶斯方法(如 BKT^[25]),假设学习状态在题目之间呈隐马尔可夫过程,但存在简化假设过多、特征建模能力有限的问题。进入深度学习时代后,深度知识追踪(Deep Knowledge Tracing, DKT)^[26]的提出成为突破口,其首次将循环神经网络(RNN)用于建模学习过程中的时序依赖性,显著提升了对学习状态的预测精度。随后,长短期记忆网络(LSTM)和 Transformer 等模型逐渐取代传统 RNN 结构,表现出更强的建模能力^[27-29]。为了进一步提升对学习规律的挖掘能力,研究者在 KT 模型中引入了多种语义特征,如知识难度水平^[30]、复习策略与间隔时间^[31]、知识点之间的语义关联与层级结构^[32]等,从而推动 KT 模型逐步向“认知建模”与“策略优化”方向发展。此外,还有部分研究尝试将 KT 与个性化推荐系统结合,实现在适应性场景下的学习路径推荐与干预。

尽管深度学习技术极大提升了模型的拟合能力和预测精度,但其“黑箱”特性仍是当前难以忽视的问题。多数模型难以提供明确的“为什么”与“如何学习更好”的理论解释,限制了其在教育教学中的可解释性与实用性。同时,这类模型对大量高质量行为数据具有高度依赖性,而现实中因数据噪声、标注缺失、学习路径多样化等问题,模型泛化能力仍有待提升。因此,在推动数据驱动建模的同时,也亟需探索结合理论知识的混合建模范式,以提升模型的理论解释力与应用可靠性。

2.3 物理信息神经网络

近年来,物理信息神经网络(Physics-Informed

Neural Networks, PINN)作为一种融合先验知识与数据驱动能力的混合范式,在多个领域取得了显著进展,尤其是在动力学建模^[33]、流体力学^[34]和微分方程求解^[35]等方面。PINN 的核心思想是将已知的物理规律(如守恒定律、边界条件)直接融入神经网络的损失函数中,从而在训练过程中使模型在拟合数据的同时,满足物理约束条件。这一方法显著增强了模型的可解释性和泛化能力^[36]。例如,研究者通过结合 Navier-Stokes 方程,成功地模拟了流场特性,并提高了水轮机在复杂流场中的能量转换效率,优化了结构设计^[37]。

受此启发,将学习规律嵌入神经网络模型,以刻画学习过程中的行为演化与知识掌握变化,正成为教育智能领域的前沿课题。不同于物理系统的可精确建模,学习过程涉及认知转化、练习效应、动机波动等复杂因素,尚缺乏统一的数学表达形式。尽管已有研究提出记忆遗忘曲线、技能习得规律等形式化描述,但它们多基于理想化假设,难以涵盖真实教学场景中高维、动态、多源的学习行为特征。

此外,传统学习理论中广泛存在的抽象变量(如认知负荷、掌握程度、练习收益)难以直接量化,也缺乏明确的函数表达式,使得其在深度学习框架中的融合与解释面临巨大挑战。因此,如何借鉴 PINN 的范式,将认知科学中的学习规律(如练习对技能表现的边际收益规律、不同任务间的迁移依赖关系)显式地约束神经网络的推理过程,成为学习规律挖掘的一项关键突破方向。但与物理方程相比,心理机制变量多为隐性、动态的,不具备显式硬编码约束的可行性,因此更适合以可调结构性表达参与建模,探索以符号建模为桥梁的柔性引导方法。

2.4 符号回归

符号回归(Symbolic Regression, SR)作为一种典型的可解释建模方法,长期用于从数据中自动发现解析表达式。早期方法大多依赖遗传编程(Genetic Programming, GP),如 Koza^[38]提出的进化框架以及 Schmidt 与 Lipson 的基于物理启发的符号回归方法^[39]。孙浩等人提出的以蒙特卡洛树搜索等启发方法辅助符号物理学习的方法是将符号表达式搜索与现代搜索策略结合的典型代表^[40]。该类工作在复杂物理系统的公式发现中展现了较好的效率与可解释性,增强了符号回归在实际科学场景下的可用性^[41]。这类方法在表达力上表现突出,但在大规模数据和高维特征条件下计算开销较大,难以直接应用于复杂教育行为数据。

近年来,PySR^[42]和 AI Feynman^[43]等工具高效地实现了结合启发式搜索与深度学习辅助策略,能够在更大规模数据中快速挖掘简洁且具有解释性的方程。然而在教育数据挖掘场景下,符号回归直接应用仍存在局限:一方面,学习行为数据噪声较多,纯粹依靠方程搜索容易陷入过拟合;另一方面,缺少与神经网络的协同使得符号回归的预测性能难以保持稳定。基于此,本文希望利用符号回归的结构先验优势,同时克服其在复杂数据下性能不足的问题,因此本文设计与神经网络结合的双模协同机制,让 SR 不仅提供可解释结构,还能在动态蒸馏框架下与神经网络相互促进。

2.5 知识蒸馏

知识蒸馏(Knowledge Distillation, KD)的核心是以软目标或中间表征将教师模型的知识迁移给学生模型,从而在不显著牺牲性能的前提下降低模型复杂度、提升泛化能力^[44]。此后, KD 在模型压缩、半监督、跨模态与大模型生态中得到广泛发展,并在训练范式与蒸馏源(logits、特征、注意力、提示等)方面形成了系统化方法论^[45]。

然而在教育数据挖掘场景中,传统单向“教师→学生”的 KD 可能受限:一是数据长尾与稀疏导致蒸馏信号不稳;二是单向模式难以捕获多源知识的互补。为此,本文提出一套动态知识蒸馏机制:令符号回归与神经网络在训练过程中根据性能动态切换教师-学生角色,形成双向、持续的知识交互。在该机制下,符号回归的结构先验与神经网络的表征能力能够互补并自适应权衡;再结合上节的结构演化,整体实现解释性-预测性的协同提升。

3 方法

3.1 问题定义

教育数据挖掘旨在理解学习者知识状态的演化与其背后的认知机制。然而,认知理论常涉及抽象变量、缺乏统一函数表达,使其难以直接嵌入神经网络。为此,本文希望借助大规模真实学习数据,设计一种能融合教育心理学理论与学习行为数据双驱动的建模框架。这些数据来源于多个实际应用程序中记录的用户交互日志,涵盖数学练习、词汇记忆、技能习得等典型学习任务。用户通过软件进行题目或词汇的反复练习(如图 2 所示),以加深理解、强化记忆,从而逐步建构知识状态。通过分析学习者在不同时间点的答题情况,可以揭示其知识状态的演化

与内部机制的核心规律。本文的目标是基于学习者在题目测试、词汇测试等的历史交互数据,构建心理学理论嵌入的计算模型,估计他们对每个题目、知识点和词汇的知识状态,并准确地预测其在未来特定题目或词汇的测试表现。

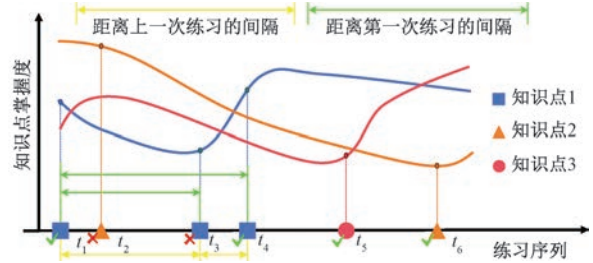


图 2 问题场景示例

为了形式化定义问题,设所有用户的集合表示为 $U = \{u_1, u_2, \dots, u_n\}$,所有题目、词汇或技能的集合表示为 Q 。该数据集包含了所有用户的题目测试行为,表示为 $D = \{D_{u_1}, D_{u_2}, \dots, D_{u_n}\}$,其中每一个用户的行为 D_u 定义如下:

$$D_u = \{(q_1, y_1, t_1), (q_2, y_2, t_2), \dots, (q_m, y_m, t_m)\} \quad (1)$$

其中, D_u 表示用户 u 的学习日志数据,由一系列三元组组成,分别表示题目 q 、标签 y 和时间戳 t 。

对于某一学习行为日志 (q, y, t) ,本文用 $x_u^{t_i}$ 表示用户 u 在时间 t_i 的历史学习行为特征,这些特征是从其所有先前的行为记录中提取出来的。这些特征包含 8 个主要变量,其定义与计算方法如图 3。相应地, $y_u^{t_i}$ 表示用户 u 在时间 t_i 对题目 q 的测试表现。本文的目标是构建一个模型 f ,使得 $y_u^{t_i} = f(x_u^{t_i})$ 。

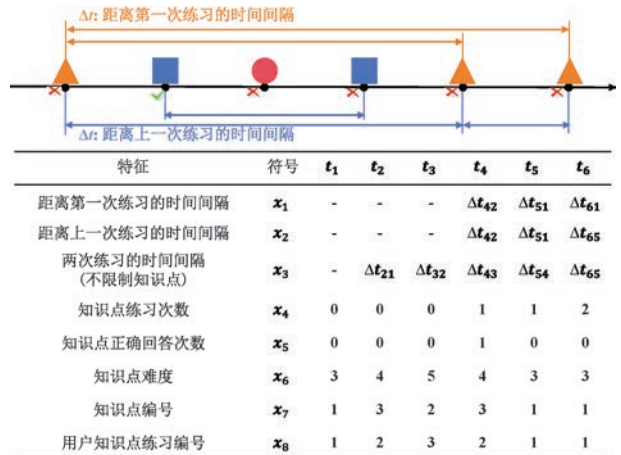


图 3 数据特征定义

3.2 PsyINN 框架

本文提出了一种心理信息神经网络(Psychol-

ogy-Informed Neural Network, PsyINN) 框架, 该框架主要由两个核心模块组成: 深度神经网络模块与符号回归模块(如图 4 所示), 两种模型在结构表达与泛化能力方面具有互补性。考虑到它们在训练不同阶段的建模能力存在差异, 本文进一步引入了动态协同机制, 使模型结构和数据表示能力能够互导

演化。深度神经网络致力于从大规模行为数据中建模学习者复杂的学习状态, 以实现高精度的行为预测; 而符号回归模块则以经典的学习规律方程为基础, 从数据中挖掘出更符合学习规律的数学表达式, 同时将其知识蒸馏至神经网络中。本文将神经网络与符号回归的预测输出统称为 $\hat{y}_{NN}^{u_i}$ 和 $\hat{y}_{SR}^{u_i}$ 。

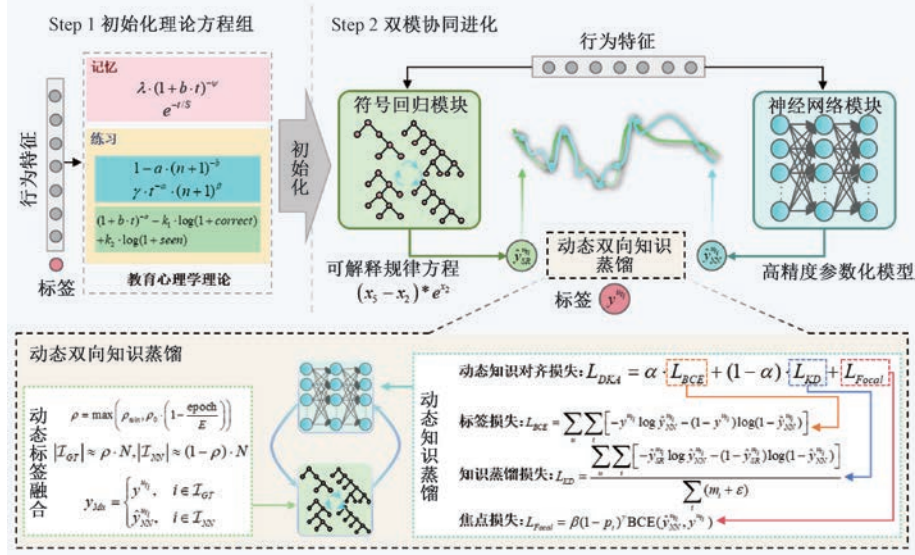


图 4 PsyINN 框架图

3.2.1 神经网络模块

神经网络模块负责从高维、动态的行为序列中建模个体知识状态的时序变化。本文并不对神经网络的具体架构做限定, 它可以采用任意时间序列模型, 如 LSTM^[46]、Transformer^[47]、Mamba^[48], 或其他自定义设计的网络结构。

神经网络模块以输入序列 $x_u^{1:t}$ 为条件, 预测个体在下一时刻的学习表现:

$$\hat{y}_{NN}^{u_i} = f_{NN}(x_u^{u_i}, \theta) \quad (2)$$

神经网络通过反向传播进行训练, 主要优化其与真实标签之间的交叉熵损失, 具体计算方式如下:

$$L_{BCE} = -\frac{1}{|D|} \sum_{(u, t_i) \in D} [y^{u_i} \log \hat{y}_{NN}^{u_i} + (1 - y^{u_i}) \log(1 - \hat{y}_{NN}^{u_i})] \quad (3)$$

3.2.2 符号回归模块

符号回归模块(Symbolic Regression, SR)的核心目标是以可解释数学表达式建模潜在学习机制。本文采用的是一类支持预设表达式初始化的符号回归框架, 其方法基础是基于遗传算法的进化搜索, 通过结构初始引导和表达式演化, 实现理论约束与数据拟合能力的权衡。该框架通过进化过程生成结构化表达式, 包括初始化种群、适应度评估、选择、

变异与交叉等操作, 并具备高度灵活性, 可替换为其他支持预设的符号回归算法。SR 模块的预测形式如下:

$$\hat{y}_{SR}^{u_i} = f_{SR}(x^{u_i}, \Phi, \tau) \quad (4)$$

其中, $\hat{y}_{SR}^{u_i}$ 表示原始数据上的函数值, $f_{SR}(\cdot)$ 是从教育心理学理论方程中优化得到的函数, Φ 表示符号操作集合, 包括加、减、乘、除等基础算子, 符合教育心理学理论中使用的运算符, τ 表示当前的符号表达式。

SR 模型的预测能力通过如下适应度函数进行评估:

$$L_{\hat{S}} = \frac{1}{|\mathcal{D}|} \sum_{u \in \mathcal{U}} \sum_{i=1}^m (\hat{y}_{SR}^{u_i} - y^{u_i})^2 \quad (5)$$

该函数用于衡量所搜索方程对训练数据的拟合程度, 从而确保生成的方程能够贴合数据特征。

3.2.3 双模协同进化

为实现理论知识与数据驱动能力的深度融合, PsyINN 框架旨在构建一个同时具备结构可解释性与表达泛化能力的协同优化体系。符号回归(SR)若仅从数据中搜索符号表达, 往往存在规律不可解释或结构不稳定的问题, 因此有必要为其提供良好的初始化基础以引导演化方向。同时, SR 与神经网络

络(NN)在训练过程中拥有独立的优化机制和参数空间:前者基于多线程进化搜索,后者则依赖梯度反向传播,两者在更新策略和收敛速度上存在明显差异。这种拟合能力的不对称性可能导致训练阶段的协同效率下降,甚至产生优化步调失衡的问题,影响整体模型的鲁棒性与泛化性能。为缓解上述矛盾,有必要设计相应机制以促进双模在训练过程中的信息同步、结构对齐与动态协同。

如图4所示,本文构建了双模协同进化机制,并将其动态优化过程划分为两个阶段:预设方程组的初步进化与动态双向知识蒸馏。第一阶段中,引入教育心理学中的典型认知方程作为SR模块的初始化表达,构建结构良好、理论一致的起始模型,并在训练初期通过小规模标注样本拟合建立基础表达能力。此阶段中,SR作为“教师”向初始状态下尚未收敛的NN模块传递符号知识,引导其优化方向。第二阶段中,模型进入SR与NN的深度协同演化过程,即动态双向知识蒸馏阶段。该阶段由两部分机制组成:①动态知识蒸馏,根据两个模块在不同阶段的性能表现,自适应调整蒸馏方向与强度,实现“谁优谁导”的角色动态转化;②动态标签融合,构造混合标签以引导SR吸收NN的分布表达能力,逐步增强其泛化水平。为进一步增强模型交互频率与优化一致性,训练过程中引入了交替训练策略,使SR与NN在每轮更新中轮流利用对方的最新输出进行重构与拟合,从而提升双模协同效率,促进结构与表达层的双向对齐。

(1) 预设方程组的初步进化

为应对SR随机起步易陷入解释性不足与不稳的问题,并向尚未收敛的NN提供早期、可解释的蒸馏源,本阶段引入教育心理学三类典型方程作为先验并随数据多线程进化,分别覆盖记忆规律、练习规律。这些理论方程作为符号回归(SR)模块的先验候选,构成结构良好的初始表达种群,为建模过程提供了明确的认知机制引导,并以软约束的形式嵌入模型,使其在训练初期不再完全依赖数据拟合。初始训练轮次(Epoch 0)中,系统将从完整训练数据中随机抽取一小部分标注样本,对这些理论方程进行参数拟合和结构筛选,进而获得具备基础表达能力的SR模型。需要强调的是,这些方程并非静态不变,而是通过多线程进化搜索策略在训练过程中不断演化、动态调整,以适应数据中潜在的学习行为规律。该阶段的核心目标是:在保持理论一致性的前提下,为SR提供解释性良好的结构起点,并为NN

提供符号知识的早期蒸馏源,从而实现结构初始化、知识注入与训练稳定性的三重保障。

(2) 动态双向知识蒸馏

为缓解SR与NN的优化机理与速率不一致与阶段性主导权更迭带来的对齐难题,本阶段通过动态知识蒸馏与动态标签融合实现“谁优谁导”的自适应协同。本文设计了一种动态知识对齐(Dynamic Knowledge Alignment)机制,实现在训练过程中符号回归(SR)与神经网络(NN)之间的深度交互。该机制包含动态知识蒸馏与动态标签融合两个部分,共同实现双向且自适应的知识蒸馏。该机制的核心思想在于:随着训练推进,SR和NN的建模能力不断变化,其“教师—学生”的角色也将发生动态转化,因此需要设计一个能适应性能主导权变化的交互优化过程。

①动态知识蒸馏。本文设计了动态调节的蒸馏损失系数,用于在不同阶段调整蒸馏强度,使得NN能在不同学习阶段合理接收来自SR的引导。具体而言,从Epoch 1起,SR与NN开始进入协同进化阶段。每个训练batch中,输入序列先通过NN模型获得预测概率 $\hat{y}_{NN}^{u_i}$,同时由SR模块对当前特征的方程输出 $\hat{y}_{SR}^{u_i}$ 。训练过程中模型的损失函数由衡量神经网络拟合效果的 L_{BCE} 、衡量NN输出与SR输出之间距离的知识蒸馏损失 L_{KD} 和抑制模型对置信度极高但错误预测的焦点损失 L_{Focal} 构成。已有研究表明,标签知识蒸馏方法具有实现简单、适用性强且效果显著等优点。本文使用交叉熵损失计算NN模块与SR模块之间的标签知识损失,知识蒸馏损失计算方式如下:

$$L_{KD} = \frac{\sum_u \sum_{i=1}^N m_i \cdot [-\hat{y}_{SR}^{u_i} \log \hat{y}_{NN}^{u_i} - (1 - \hat{y}_{SR}^{u_i}) \log(1 - \hat{y}_{NN}^{u_i})]}{\sum_{i=1}^N (m_i + \epsilon)} \quad (6)$$

其中,权重项 m_i 的定义如下:

$$m_i = \begin{cases} 1, & \text{if } \hat{y}_{SR}^{u_i} < 1 - \vartheta \text{ or } \hat{y}_{SR}^{u_i} > \vartheta, \vartheta = 0.55 \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

焦点损失 L_{Focal} 定义为

$$L_{Focal} = \beta(1 - p_i)^\gamma \text{BCE}(\hat{y}_{NN}^{u_i}, y^{u_i}) \quad (8)$$

其中, β 是正负样本平衡因子; p_i 预测的概率,针对

当前样本的真实类别; γ 是调节“难易样本”的参数。

最终用于训练神经网络的总损失函数定义为

$$L_{DKA} = \alpha \cdot L_{BCE} + (1 - \alpha) \cdot L_{KD} + L_{Focal} \quad (9)$$

为了动态估计两个模块的性能并指导蒸馏权重系数 α 的调整,权重系数计算方式如下:

$$\alpha = \frac{L_{BCE}^{SR}}{L_{BCE}^{NN} + L_{BCE}^{SR}} \quad (10)$$

其中, L_{BCE}^{NN} 和 L_{BCE}^{SR} 分别为 NN 模块和 SR 模块的交叉熵损失。该机制使模型能够根据当前神经网络与符号回归模块的拟合效果,灵活分配各自的学习权重。

②动态标签融合。为了提升 SR 的泛化能力,本文构造了基于 NN 输出与真实标签的混合标签,并设计了随训练轮次自适应调整伪标签比例的机制。在训练初期以真实标签为主,保证表达式稳定性;训练后期逐步提升 NN 预测结果在混合标签中的权重,引导 SR 逐渐吸收 NN 的泛化表达能力。具体而言,符号回归同样从神经网络中进行动态学习,每个 epoch 结束后,SR 利用 NN 输出结果对方程进行重新拟合,并根据真实标签与预测值的混合构造目标变量,逐步演化得到更具泛化能力的行为方程。随着迭代轮次不断增加,神经网络预测的伪标签在混合标签中的占比便会不断提高,混合标签的计算方式如下:

$$\rho = \max\left(\rho_{\min}, \rho_0 \cdot \left(1 - \frac{\text{epoch}}{E}\right)\right) \quad (11)$$

$$|\mathcal{I}_{GT}| \approx \rho \cdot N, |\mathcal{I}_{NN}| \approx (1 - \rho) \cdot N \quad (12)$$

$$y_{Mix} = \begin{cases} y^{u_i}, & i \in \mathcal{I}_{GT} \\ \hat{y}_{NN}^{u_i}, & i \in \mathcal{I}_{NN} \end{cases} \quad (13)$$

其中, ρ 表示真实标签占比, \mathcal{I}_{GT} 表示真实标签子集, \mathcal{I}_{NN} 表示神经网络预测值的子集, y_{Mix} 表示混合标签后的标签数据。

上述两阶段分别解决起始解释性、稳定性与训练过程对齐与自适应优化的问题,共同形成理论—数据双驱动的协同优化闭环。

4 实 验

4.1 数据集与对比方法

4.1.1 数据集

为验证所提出的 PsyINN 框架在不同类型认知任务中的建模能力,本文在三个具有代表性的真实数据集上开展了广泛实验,涵盖词汇记忆(MaiMemo)、

知识追踪(EdNet)以及技能习得(Lumosity)三类任务,如表 1 所示。三个数据集分别从原始平台中随机选取部分活跃用户样本,包含 10000 名 EdNet 用户、100000 名 MaiMemo 用户和 1000 名 Lumosity 用户,交互总量均超过千万级;为更直观地衡量不同数据集的交互稠密程度,本文定义了“密集度”指标,用于表示平均每位用户在每个题目上的交互次数,数值越高意味着数据集在用户-题目维度上的交互更为丰富、学习行为更为密集。所有数据统一采用滑动窗口机制构造长度为 15 的序列样本,并设计了包括时间间隔、练习频次、历史表现、题目难度等在内的 9 维特征向量,最终按照 7 : 1 : 2 的比例划分训练集、验证集与测试集,确保实验的可比性与一致性。

表 1 数据集基本情况

| 数据集 | EdNet | MaiMemo | Lumosity |
|------|--------------------|--------------------|--------------------|
| 用户数 | 10000 | 100000 | 1000 |
| 题目数 | 1495 | 1297 | 768 |
| 交互次数 | 1.05×10^7 | 2.58×10^7 | 2.13×10^7 |
| 正确率 | 0.6619 | 0.5385 | 0.4976 |
| 密集度 | 1.27 | 1.18 | 31.61 |

4.1.2 对比方法

为了验证本文提出的方法的有效性,将本文提出的 PsyINN 框架应用于以下六个基准模型进行对比:

(1) DKT(Deep Knowledge Tracing)^[26]:这是首个将循环神经网络(RNN)引入知识追踪任务的模型,能够通过建模学生的答题序列来预测其未来的答题表现。

(2) DKT-F^[49]:这是对原始 DKT 模型的扩展,加入了“遗忘机制”以提高用户表现的预测能力。作者引入了三个与时间相关的特征来增强模型表现:重复间隔、序列间隔以及过往尝试次数。

(3) FIFAKT^[27]:该模型通过注意力机制动态整合与遗忘、题目形式以及词语语义相似度相关的关键信息,使模型更全面地捕捉影响学生表现的因素,从而提升预测精度。

(4) SimpleKT^[29]:该模型通过显式建模题目层级的差异性,并使用标准点积注意力机制,有效捕捉题目特性与时间行为信息,能够更准确地反映学生的学习动态与知识状态演化。

(5) QIKT^[32]:该模型结合题目敏感的认知表示和项目反应理论(IRT),增强了模型对学生知识状态的建模能力,强调题目属性对学习过程的影响,从而提升模型的解释性。

(6) MIKT^[28]:该模型同时追踪学生的领域知识状态(粗粒度)与概念知识状态(细粒度),并融合 Rasch 模型与 IRT 模块,实现多层次的知识建模,提升了模型的预测性能与可解释性。

4.2 实施细节和评估指标

为了训练和验证模型,本文使用了 80% 的学生序列数据进行训练,其余 20% 用于评估。所有模型均采用 Adam 优化器进行优化,并在训练过程中引入早停策略以防止过拟合。所有实验结果均为在相同设置下独立运行 5 次后所得指标的均值。

本文使用的 SR 模型是 PySR 工具^[42],PySR 支持用户自定义初始化种群,以进一步进行方程演化,本文使用的算子集为 $\Phi \in \{+, -, \times, \div, \text{pow}, \text{exp}, \text{ln}\}$,以保证表达能力的同时维持可解释性,该算子集包含了经典教育心理学认知规律方程中的基础算子。迭代次数设置为 50,每次迭代包含 40 个种群,种群规模为 40,每个迭代循环执行 100 次。表达式最大深度限定为 3,最大复杂度为 25,以避免深度嵌套和过度复杂的结构。

所有模型均基于 PyTorch 实现,并在搭载 NVIDIA GeForce GTX 4090 GPU 的 Linux 服务器集群上进行训练。本文采用的 AUC(ROC 曲线下面积)指标是与分类阈值无关的全局性指标,能更鲁棒地衡量模型的泛化和排序能力;Precision(精确率)来评估模型在预测结果中正类识别的准确性,主要反映其在实际应用中的判别能力。同时引入 G-Means 作为补充指标,用于衡量模型在标签分布不均情况下

的整体泛化能力,尤其关注对少数类的识别效果。各项指标的计算方法如下所示:

$$\text{AUC} = \int_0^1 \frac{\text{TP}}{\text{TP} + \text{FN}} d\left(\frac{\text{FP}}{\text{FP} + \text{TN}}\right) \quad (14)$$

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \quad (15)$$

$$\text{G-Means} = \sqrt{\text{Precision}_0 \times \text{Precision}_1} \quad (16)$$

4.3 对比实验

为全面验证所提出 PsyINN 框架在多种知识建模任务中的适应性与有效性,本文在三个真实数据集上开展了系统性实验。所使用的数据集包括 EdNet、MaiMemo 和 Lumosity。本文将 PsyINN 框架分别集成至 DKT、DKT-Forget、FIFAKT、SimpleKT、QIKT 与 MIKT 共六种主流基准模型中,对比实验结果如表 2 所示。从 EdNet 数据集结果来看,PsyINN 框架在多个模型上显著提升了 G-Means 指标,体现出模型在样本类别区分和鲁棒性方面的增强。例如,DKT-Forget 的 G-Means 从 0.1626 提升至 0.3091 ($\uparrow 90.10\%$),QIKT 从 0.4536 提升至 0.5606 ($\uparrow 23.59\%$)。六个模型在 AUC, G-Means 和 Precision 上的平均提升幅度分别为 2.14%, 17.20% 和 1.23%。这些结果表明,PsyINN 引入的心理先验与双重建模机制,有助于增强模型对复杂学习状态的感知能力,提升其对知识掌握变化的响应速度与稳定性,尤其在结构较深或具备记忆机制的模型(如 DKT-Forget 和 QIKT)中表现更为显著。

表 2 三种大规模数据集上应用 PsyINN 框架的对比实验结果(* $p < 0.05$)

| 模型 | EdNet | | | MaiMemo | | | Lumosity | | |
|----------------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| | AUC | Precision | G-Means | AUC | Precision | G-Means | AUC | Precision | G-Means |
| DKT ^[26] | 0.6211 | 0.6795 | 0.3282 | 0.6937 | 0.7390 | 0.5204 | 0.8693 | 0.7858 | 0.7888 |
| + PsyINN | 0.6233* | 0.6828* | 0.3610* | 0.7000* | 0.7490* | 0.5612* | 0.8788* | 0.8246* | 0.7839 |
| DKT-Forget ^[49] | 0.5862 | 0.6658 | 0.1626 | 0.6941 | 0.7396 | 0.5192 | 0.8261 | 0.7534 | 0.7586 |
| + PsyINN | 0.5997* | 0.6746* | 0.3091* | 0.7029* | 0.7517* | 0.5685* | 0.8401* | 0.7727* | 0.7626* |
| FIFAKT ^[27] | 0.6010 | 0.6775* | 0.3137* | 0.6972 | 0.7364 | 0.5088 | 0.8619 | 0.7912 | 0.7947* |
| + PsyINN | 0.6145* | 0.6706 | 0.2308 | 0.6993* | 0.7512* | 0.5732* | 0.8809* | 0.8316* | 0.7944 |
| SimpleKT ^[29] | 0.5925 | 0.6807 | 0.3369 | 0.6356 | 0.7398 | 0.5251 | 0.8270 | 0.7404 | 0.7643 |
| + PsyINN | 0.6262* | 0.6818* | 0.3473* | 0.6926* | 0.7570* | 0.5865* | 0.8450* | 0.7743* | 0.7656* |
| QIKT ^[32] | 0.6157 | 0.6953 | 0.4536 | 0.6937 | 0.7390 | 0.5204 | 0.8711 | 0.7833 | 0.7852* |
| + PsyINN | 0.6232* | 0.7388* | 0.5606* | 0.6990* | 0.7505* | 0.5670* | 0.8881* | 0.8141* | 0.7828 |
| MIKT ^[28] | 0.6193 | 0.6831 | 0.3607 | 0.6887 | 0.7396 | 0.5192 | 0.8402 | 0.8038 | 0.7889* |
| + PsyINN | 0.6255* | 0.6843* | 0.3711* | 0.6947* | 0.7501* | 0.5653* | 0.8693* | 0.8134* | 0.7845 |

在 MaiMemo 数据集中,PsyINN 框架的增强效果更为稳定。所有模型的 G-Means 均获得正向提升,平均提升约为 9.92%,其中 SimpleKT 和 FIFAKT 分别提升 11.70% 和 12.66%;AUC 和 Pre-

cision 平均增长为 2.18% 和 1.72%。考虑到该数据集聚焦于间隔练习与记忆建模,实验结果进一步验证了 PsyINN 在节奏性记忆行为、强化与遗忘机制建模方面的表达优势,能够提升模型对学习过程

中文本回忆强度与效果变化的刻画能力。

在 Lumosity 数据集中,所有模型的 AUC 和 Precision 均获得提升,平均增长约为 2.10% 和 3.72%,其中 DKT 和 FIFAKT 在 Precision 的增量为 4.94% 和 5.11%。与此同时,少数基线的 G-Means 变化有限或出现轻微波动(如 DKT、QIKT、MIKT 小幅回落;DKT-Forget、SimpleKT 小幅上升)。该现象并非方法失效,主要来自指标定义的差异与评估时采用的固定阈值:AUC 衡量阈值无关的排序质量,Precision 关注固定阈值下的正类准确性;而 G-Means 需要在同一阈值下同时兼顾正类召回与负类特异性,因而对阈值更敏感。在高频且近均衡的 Lumosity 场景中,收益通常先体现为排序和置信度整体改善(AUC 和 Precision 上升)。

综上,PsyINN 框架在多种任务类型与模型结构下均实现了稳定的性能提升,验证了其在泛化建

模与心理认知建构方面的有效性。为构建具备可解释性与适应性的智能教育模型提供了新的研究思路与技术路径。

4.4 消融实验

为验证所提出方法中心理知识引导的初始符号回归(Init SR)与动态蒸馏损失权重调整机制(Dynamic)的有效性,本文在多个典型知识追踪模型(如 DKT、FIFAKT、QIKT 等)上进行了消融实验。如图 5 所示,去除任一组件均会不同程度地削弱模型性能,表明两者在提升模型预测能力方面均发挥了重要作用。在去除“心理知识初始化”模块(Init SR)后,各类模型在多个数据集和指标上的表现均出现不同程度的下降,尤其是在衡量模型分类平衡能力的指标上表现更为明显。这说明该模块在建模初期为模型提供了结构性先验,帮助其更好地识别低频样本与潜在认知结构,提升了模型在复杂任务下的表达能力与泛化能力。

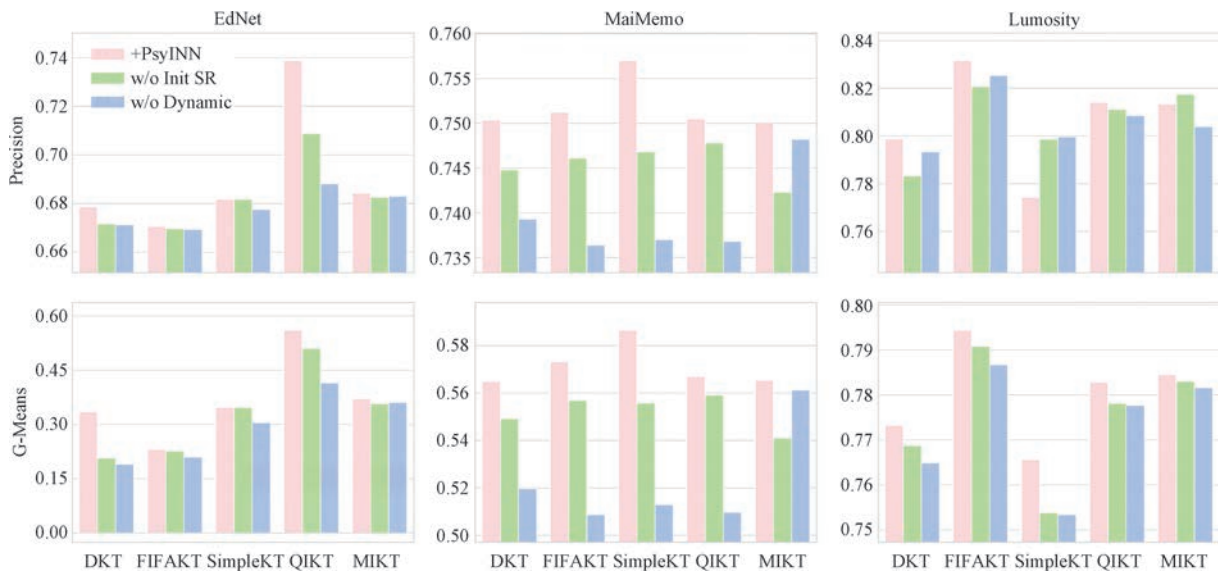


图 5 三个数据集上不同基模型的消融实验结果

进一步地,当移除“动态蒸馏”机制后,模型对难分类样本的辨别能力减弱,多项指标均出现性能退化,特别是在类间分布不均的任务中更为突出。这表明该机制有助于引导模型更有效地吸收教师模型中的判别性知识,优化学习过程中的样本利用策略,从而缓解类别不平衡带来的影响。

总体来看,心理知识初始化和动态蒸馏机制分别从结构建构与学习动态两个层面对模型提供支持。前者为模型奠定了理论驱动的建模基础,后者则在训练过程中强化了模型的适应性与稳定性。两者协同作用,显著提升了 PsyINN 框架在多种模型与任务场景下的表现,验证了其设计的合理性与通用性。

4.5 协同进化案例分析

如图 6 所示,蓝色曲线展示了动态权重 α 在训练过程中的变化趋势。在前期(Epoch 1-15), α 保持在较低水平,表明此阶段以 SR 为主导,SR 的先验结构对 NN 学习起到了主要引导作用。随着训练的推进(Epoch 15 之后), α 值逐渐上升并趋于稳定,说明 NN 的预测能力逐渐增强,并在后期成为主要“教师”。与此同时,NN 的 AUC 与 G-Mean 指标整体呈上升趋势,进一步验证了动态权重调节对 NN 性能提升的有效性。综合来看,动态权重的引入保证了 SR 与 NN 在不同训练阶段的合理分工,实现了从结构先验引导到数据驱动优化的平滑过渡,从

而体现了动态知识蒸馏机制的有效性。

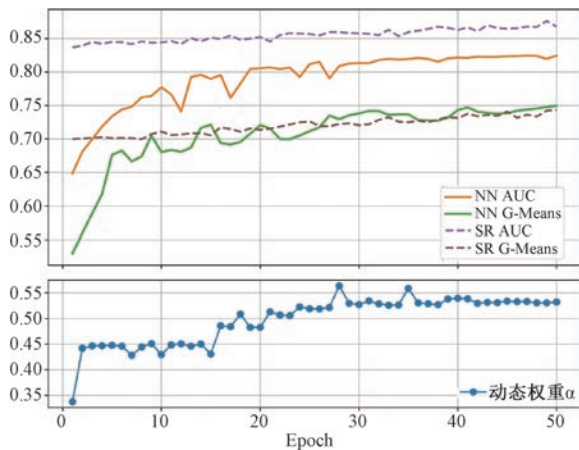


图6 Lumosity数据集上DKT-Forget+PsyINN的模型性能与动态蒸馏权重 α 随训练轮次的演化

4.6 适用性边界与模型适配性分析

在前述对比实验与消融实验中,在Lumosity数据集上,部分基准模型在引入PsyINN后,个别评价指标出现了轻微波动。该现象并不意味着方法本身失效,而更可能源于数据分布特征与心理先验建模假设之间的尺度不一致。为此,有必要从学习行为的时间结构角度,对PsyINN的适用条件进行进一步分析。

从符号回归初始化的理论假设来看,PsyINN所采用的经典记忆与巩固练习(如指数遗忘律、幂律练习模型)均以“交互次数”和“时间间隔”为核心自变量,其隐含前提是:学习过程发生在相对连续、可感知的认知时间尺度内。在这一前提下,时间间隔刻画的是记忆强度随时间衰减的过程,而非学习任务的中断或重启。这一假设在多数真实教育场景(如课堂学习、作业练习、间隔复习等)中是合理且普遍成立的。

然而,如图7所示,Lumosity数据呈现出一种较为特殊的结构特征:一方面,存在大量短时间内的高频交互行为;另一方面,同时伴随一定比例的极端长尾时间间隔样本。当复习间隔达到数百天量级时,其认知含义已明显偏离经典遗忘过程,更接近于“重新学习”或“任务切换”行为。在这种情况下,若仍将该类样本直接纳入基于遗忘假设构建的心理先验中,可能会引入与认知物理意义不一致的噪声,从而产生扰动。这种扰动并非来源于符号先验结构本身,而是由数据中混入的非连续学习过程所导致。

为验证上述分析,本文进一步在Lumosity数据集中对复习间隔超过200天的极端长尾样本进行剔除,

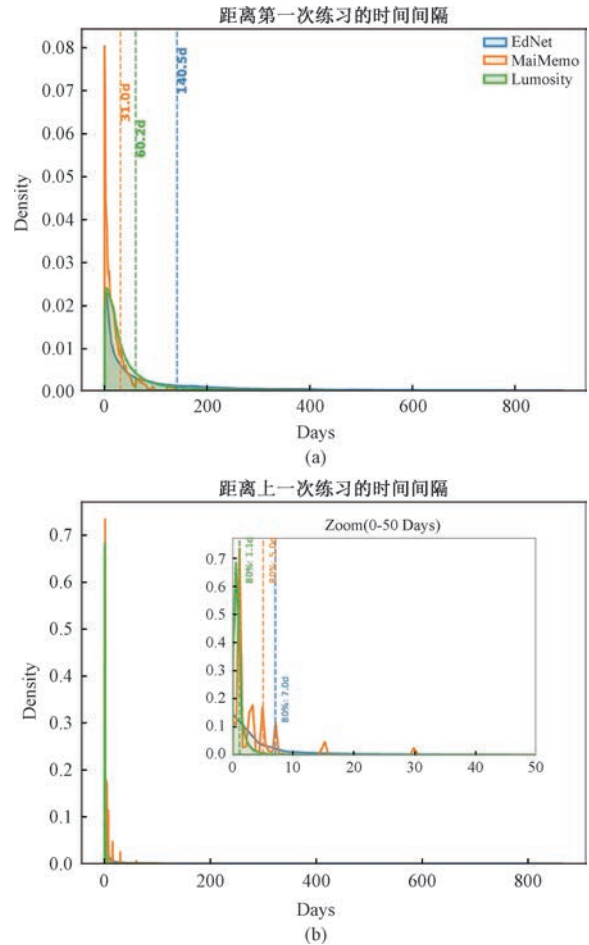


图7 不同数据集时间间隔特征画像

仅保留符合连续学习假设的数据子集,并在该设置下重新评估PsyINN及其变体在所有基准模型上的表现。

如表3所示,在剔除极端长尾时间间隔样本后,PsyINN在所有基准模型上的AUC、Precision与G-Means均取得一致的正向提升,且提升幅度在不同模型间表现出良好的稳定性。尤其是在Precision指标上,各基准模型均获得显著增益,表明在连续学习假设成立的条件下,心理先验引导能够有效提升

表3 Lumosity数据集去除掉长尾数据后基准的实验结果

| 模型 | AUC | Precision | G-Means |
|------------|---------------|---------------|---------------|
| DKT | 0.8697 | 0.7819 | 0.7830 |
| +PsyINN | 0.8723 | 0.8287 | 0.7852 |
| DKT-Forget | 0.8669 | 0.7971 | 0.7821 |
| +PsyINN | 0.8727 | 0.8311 | 0.7847 |
| FIFAKT | 0.9427 | 0.8567 | 0.8734 |
| +PsyINN | 0.9465 | 0.8968 | 0.8739 |
| SimpleKT | 0.8457 | 0.7499 | 0.7634 |
| +PsyINN | 0.8472 | 0.7952 | 0.7650 |
| QIKT | 0.8635 | 0.7887 | 0.7857 |
| +PsyINN | 0.8796 | 0.8354 | 0.7914 |
| MIKT | 0.8714 | 0.7766 | 0.7907 |
| +PsyINN | 0.8732 | 0.8295 | 0.7936 |

模型对正类样本的判别置信度,而不会引入额外的判别噪声。与此同时,G-Means 指标在所有模型上保持稳定或小幅上升,说明模型在正负样本整体平衡性方面未受到不利影响。

进一步地,从图 8 的消融实验可以观察到,在数据净化后,“Init SR”和“Dynamic”两个组件在所有基准模型上均持续带来稳定的性能增益,未再出现方向性波动。这表明此前在原始数据中观察到的个别指标异常,主要源于极端长尾时间间隔样本对心理先验假设的干扰,而非 PsyINN 框架中理论初始化或动态协同机制本身存在不稳定性。

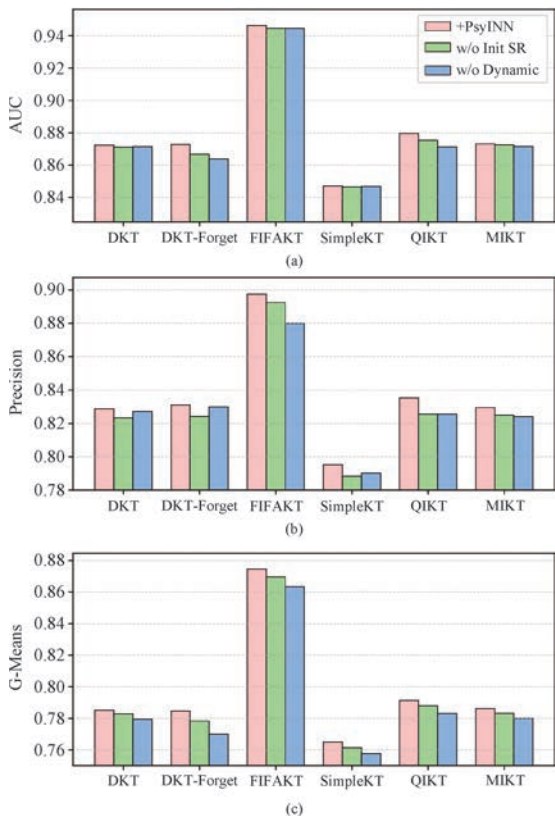


图 8 Lumosity 数据集裁剪长尾数据后的消融实验结果

综上,在符合连续学习认知时间尺度的条件下, PsyINN 能够在不同模型结构下稳定发挥作用,其经典心理理论初始化并未限制方法的适用性,反而为模型提供了可靠且一致的结构归纳偏置,并通过动态协同机制实现理论引导与数据驱动之间的自适应平衡。

4.7 学习场景感知的符号结构演化能力分析

为了验证 PsyINN 框架在不同任务数据集上的通用性与灵活性,本实验设计了对同一基础模型(QIKT)在数学练习、词汇记忆及技能习得上的方程演化过程进行系统观察和分析(如图 9 所示)。在本应用实验中,本文采用符号回归方式对不同认知

任务中的学习过程进行结构建模,并引入心理理论相关的初始表达式以增强可解释性。具体而言,本文在训练初期为符号回归器设定了 5 个结构函数作为初始化候选,分别来自三类典型心理机制:记忆理论(2 个)、巩固练习(2 个)与技能习得(1 个)。这些结构被共同置于符号回归结构池中,作为 Epoch 0 的起始模型,并在后续训练过程中不断演化。

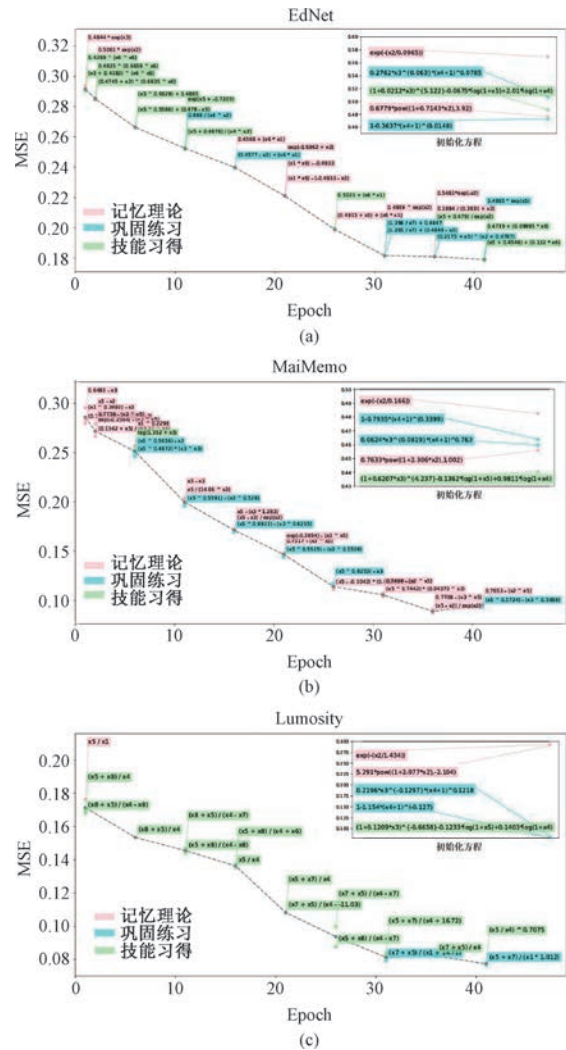


图 9 QIKT+PsyINN 在不同数据集上方程演化示意图

EdNet 数据集中学生学习行为展现出高度的复杂性多层次动态特征,这种复杂性促使符号回归模型在训练过程中快速演化出多样化的表达结构。可以观察到,在训练初期,模型更倾向于选择结构简单、拟合稳定的记忆与练习类表达式,如 $0.5061 \cdot e^{x^2}$,其体现了典型的指数型遗忘规律,并在前几个 epoch 中保持较低的损失。但随着训练的深入,模型迅速生成并保留了结构更复杂的技能习得类表达式,如 $x_1 \cdot x_6 + 0.4933 + x_3$ 与 $(x_5 + 0.4876)/(x_4/x_2)$,反映出对变量交互、非线性过程及学习状态演

变的更强建模需求。这一从简单记忆函数向高度非线性表达的快速演化趋势,揭示了知识追踪任务中多源认知机制交织的本质,也说明 PsyINN 框架需具备动态结构适应能力,才能有效表征真实学习过程中的认知机制。在 MaiMemo 数据集中,由于其原始设计就面向节奏调控与回忆机制建模,模型在结构演化过程中呈现出较为明显的记忆机制偏好。初始的记忆类函数(如 $e^{-x_2/0.166}$)在 epoch 0-1 中被长期保留,并在多个迭代轮次中演化出其非线性变体,如 $1-a \cdot (x_4+1)^b$ 或 $\log(1+x_5)$ 参与组合的记忆表达式。这些结构普遍具有较低的损失,显示出对记忆强度与时间间隔变化建模的有效性。与 EdNet 相比, MaiMemo 的演化过程更聚焦于强化单一心理机制表达的深度建模,体现了符号回归器对任务特性识别能力的适应性。

在 Lumosity 数据集中,模型演化呈现出与技能习得任务高度一致的特征。初始技能表达式虽较少,但在训练早期即迅速演化出大量幂律型技能增长模型,如 $(x_5/x_4)^{0.6017}$ 、 $(x_5/x_4)^{x_7-0.1358}$ 等。这类表达式不仅频次极高(多次在相邻迭代轮次出现相近结构),而且损失值稳定,说明模型从数据中识别

并重构了经典的技能提升过程——即随着练习次数增多,行为表现以减缓速率持续优化。与此同时,多个幂律变体(如 $(x_5/x_4)^a$ 与带指数项的 $(x_5/x_4)^{e^{-c}}$)也反复出现,进一步表明该任务中,技能习得理论与巩固练习共同构成了主导表达,而记忆机制则相对较少参与。

PsyINN 在三个任务中展现出良好的结构适应能力;在结构初始化阶段有效利用心理机制方程提供合理起点,在演化过程中可据任务特性逐步调整主导机制;尤其在 Lumosity 与 MaiMemo 数据集上,所演化出的表达式与各自任务对应的心理理论高度吻合。这验证了模型不仅具备符号演化能力,更具备从数据中自动提取与任务契合的理论结构表达的能力,为后续构建可解释性更强的认知模型提供了可行路径。

4.8 符号规律的结构构成与心理机制解析

在 PsyINN 框架中,符号回归模块在不同类型的任务上,从行为数据中自动发现了若干形式各异的学习规律方程。这些表达式在结构形态与参数特征上,体现出与经典记忆理论、巩固练习、技能习得理论的高度一致性,揭示出潜在的学习认知机制(如表 4 所示)。

表 4 PsyINN 框架挖掘的学习规律方程

| 模型 | EdNet | MaiMemo | Lumosity |
|-------------------------------|--|---|---|
| Wickelgren ^[12] | $0.678 \cdot (1+0.714 \cdot x_2)^{3.919}$ | $0.763 \cdot (1+2.306 \cdot x_2)^{1.003}$ | $5.291 \cdot (1+2.077 \cdot x_2)^{-2.103}$ |
| Wozniak ^[13] | $e^{-x_2/0.097}$ | $e^{-x_2/0.166}$ | $e^{-x_2/1.434}$ |
| Newell A ^[52] | $1-0.364 \cdot (x_4+1)^{0.015}$ | $1-0.794 \cdot (x_4+1)^{0.34}$ | $1-1.154 \cdot (x_4+1)^{-0.127}$ |
| Pavlik Jr P I ^[53] | $0.276 \cdot x_2^{0.063} \cdot (x_4+1)^{0.079}$ | $0.062 \cdot x_2^{0.082} \cdot (x_4+1)^{0.763}$ | $0.219 \cdot x_2^{-0.129} \cdot (x_4+1)^{0.122}$ |
| Pavlik P I ^[54] | $(1+0.021 \cdot x_2)^{5.121}$ $-0.066 \cdot \log(1+x_4)$ $+2.01 \cdot \log(1+x_3)$ | $(1+0.621 \cdot x_2)^{4.237}$ $-0.136 \cdot \log(1+x_4)$ $+0.981 \cdot \log(1+x_3)$ | $(1+0.121 \cdot x_2)^{-0.666}$ $-0.123 \cdot \log(1+x_4)$ $+0.14 \cdot \log(1+x_3)$ |
| FIFAKT+PsyINN | $0.557 \cdot e^{x_5}$ | $0.755 - x_2^{x_5}$ | $(x_5/x_4)^{0.4542}$ |
| SimpleKT+PsyINN | $e^{x_5} - 0.251$ | $0.720 - x_3^{x_5}$ | $(x_5/x_4)^{0.4950}$ |
| QIKT+PsyINN | $0.546 \cdot e^{-x_2}$ | $(x_5 - x_2) \cdot e^{x_2}$ | $(x_5/x_4)^{0.4625}$ |

在学习建模任务 EdNet 数据集中,PsyINN 所提取的函数结构普遍围绕 x_5 (历史正确次数)展开,展现出以 e^{x_5} 或 $0.557 \cdot e^{x_5}$ 为代表的指数增长趋势,反映了记忆强度随着正确回忆次数呈指数增强的规律,这与 Wozniak 等人^[13]提出的指数律相符,体现了记忆巩固与强化学习过程中的非线性累积效应,揭示了随着成功回忆的不断积累,个体记忆保持能力得到显著提升。同时,QIKT+PsyINN 中出现的 e^{-x_2} 项引入了时间因子的指数衰减项,反映了随

着练习间隔时间增长,记忆保持的指数衰减特性,体现了记忆随着时间推移自然遗忘的规律,这与经典的遗忘理论相吻合,强调了时间对记忆状态的关键影响。

在注重复习节奏与回顾策略的记忆建模任务 MaiMemo 数据集中,学习得到的函数形式更加突出练习间隔与回忆强度之间的交互作用。幂律结构如 $x_2^{x_5}$ 、 $x_3^{x_5}$ 表明:随着练习正确次数的增加,系统对回顾时间间隔的容忍度提升,揭示了一种“习得稳定

性”随经验积累增强的机制。这种结构契合了分布式学习理论^[50]与广义幂律^[12]的核心思想。特别值得注意的是,表达式 $(x_5 - x_2) \cdot e^{x_2}$ 展示了系统对“回忆稳定性缺口”的强敏感性:当回顾间隔偏长而回忆积累不足时,预测输出会急剧下降,体现出 DASH 模型^[16]中对复习节奏的动态调控要求,即通过反馈机制加快节奏以避免遗忘发生。

在技能习得类任务 Luminosity 数据集中,Psy-INN 所提取的规律方程更贴近经典技能增长理论。以 $(x_5/x_4)^{0.4542}$ 为代表的结构表明,学习效果受到单位练习正确率的调节,且提升速度呈边际递减。这种幂结构与 Anderson 等人^[51]在技能习得研究中提出的“幂律学习曲线”高度一致。

值得注意的是,如表 5 所示,在原始 Luminosity 数据中,SimpleKT+PsyINN 所学习到的符号方程中包含了时间间隔变量。该现象主要源于数据中存在的极端长尾时间间隔样本,符号回归在拟合过程中对异常时间尺度产生响应,从而将其引入表达式中。随着极端长尾样本被剔除,SimpleKT+Psy-INN 所演化得到的符号方程不再包含该变量,并回归到以交互次数为核心的幂律形式,表明符号结果的变化主要由数据分布差异所驱动,而非方法本身的不稳定性。

表 5 Luminosity 数据集上 PsyINN 学习规律方程对比

| 模型 | 处理后 | 处理前 |
|-----------------|----------------------|-------------------------|
| FIFAKT+PsyINN | $(x_5/x_4)^{0.4542}$ | $(x_5/x_4)^{0.667}$ |
| SimpleKT+PsyINN | $(x_5/x_4)^{0.4950}$ | $x_4^{0.214}/\log(x_1)$ |
| QIKT+PsyINN | $(x_5/x_4)^{0.4625}$ | $(x_5/x_4)^{0.601}$ |

综上所述,PsyINN 框架通过符号回归机制,能够自动发现具有清晰心理学解释的学习规律表达式,并在不同任务场景中展现出对时间因素、正确率、练习频次等变量交互作用的精准建模能力。其符号模块可适应任务目标与数据结构的差异,持续演化出结构精炼、理论合理的函数形式,所提取的表达式不仅在数学结构上具备高度简洁性,更高度贴合经典的记忆模型与技能习得规律,体现出 Psy-INN 作为可扩展学习建模工具的强泛化能力与建模灵活性。

5 总结与展望

本文提出的心理信息神经网络(PsyINN)框架,融合符号回归与神经网络建模机制,通过“双模动态

协同优化”策略,有效解决了学习认知理论难以结构化建模与神经网络缺乏可解释性的问题。在三个具有代表性的学习任务上,实验结果表明:PsyINN 在提升预测性能的同时,能够学习出具有理论一致性的表达结构。特别是在记忆与技能习得任务中,符号回归模块自动演化得到的表达式呈现出经典心理学理论特征,如指数衰减、幂律学习过程等,验证了符号建模在学习认知机制表达上的可行性与有效性。此外,动态蒸馏与伪标签调节机制基于模型性能动态转化教师—学生角色,推动了结构信息与深度表示之间的双向蒸馏,有效提升了泛化性能。框架本身还展现出良好的任务适应性与泛化能力,具有一定的跨任务建模潜力。

未来研究需进一步提升模型的结构适应性与跨任务泛化能力,以应对复杂学习场景的动态需求。同时,优化符号表达搜索效率,以支撑大规模学习行为数据的实时建模。该框架未来可进一步拓展至智能教育场景相关任务,如学习路径建模、个性化反馈生成和知识状态预测等,助力模型更深入刻画学习过程中的认知规律与行为演化机制。随着人工智能辅助理论建模在社会科学领域的快速发展,PsyINN 框架能够扩展应用至金融行为建模、社会行为建模等领域,通过融合领域理论知识与大规模数据,实现对人类行为模式的有效捕捉与内在规律揭示,提升模型的解释深度与泛化能力。

致 谢 本文得到新一代人工智能国家科技重大专项课题“多时空跨尺度认知与情感规律挖掘方法”(No. 2022ZD0117103)、国家自然科学基金重大项目课题“智能教学策略设计与调节机制”(No. 62293554)、国家自然科学基金面上项目“理论方程-神经网络协同演进的记忆演化计算与规律发现”(No. 62577028)、中国自动化学会青少年人工智能创新后备人才培养工程 2025 年度教科研重大项目“数智驱动的青少年人工智能素养测评”(No. HBRC-JKYZD-2024-310)、湖北省自然科学基金创新群体项目“智能导学机理与关键技术”(No. 2023AFA020)资助。最后,诚挚感谢审稿专家提出的宝贵意见和建议,这些意见对本文的完善与提升具有重要意义。

参 考 文 献

[1] Liu San-Nyu-Ya, Yang Zong-Kai, Li Qing. Computational ed-

- ucation; Connotations and approaches. *Educational Research*, 2020, 41(3):152-159 (in Chinese)
(刘三女牙, 杨宗凯, 李卿. 计算教育学:内涵与进路. 教育研究, 2020, 41(3):152-159)
- [2] Raissi M, Perdikaris P, Karniadakis G E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 2019, 378: 686-707
- [3] Chen Zhi-Min, Wu Li-Bo. The future is here: Blue book on the development of AI for humanities and social sciences. Shanghai: Fudan University Press, 2025 (in Chinese)
(陈志敏, 吴力波. 未来已来——人文社会科学智能发展蓝皮书. 上海:复旦大学出版社, 2025)
- [4] Li Guo-Jie. AI4R: The fifth scientific research paradigm. *Bulletin of Chinese Academy of Sciences*, 2024, 39(1):1-9 (in Chinese)
(李国杰. 智能化科研(AI4R):第五科研范式. 中国科学院院刊, 2024, 39(1):1-9)
- [5] Cheng Xue-Qi, Mei Hong, Zhao Wei, et al. Data science and computing intelligence: Concept, paradigm, and opportunities. *Bulletin of Chinese Academy of Sciences*, 2020, 35(12): 1470-1481 (in Chinese)
(程学旗, 梅宏, 赵伟等. 数据科学与计算智能:内涵、范式与机遇. 中国科学院院刊, 2020, 35(12):1470-1481)
- [6] Liu San-Nyu-Ya, Hao Xiao-Han, Li Qing. A new paradigm of educational research: A research into AI-driven educational science. *Educational Research*, 2024, 45(3): 147-159 (in Chinese)
(刘三女牙, 郝晓晗, 李卿. 教育科研新范式:人工智能驱动的教育科学研究. 教育研究, 2024, 45(3): 147-159)
- [7] Sun Jian-Wen. AI4LS: A new research paradigm of learning sciences in the age of intelligence. *Journal of Central China Normal University (Humanities and Social Sciences)*, 2025, 64(3): 154-162 (in Chinese)
(孙建文. AI4LS: 智能时代学习科学研究新范式. 华中师范大学学报(人文社会科学版), 2025, 64(3): 154-162)
- [8] Liu S, Li Q, Shen X, et al. Automated discovery of symbolic laws governing skill acquisition from naturally occurring data. *Nature Computational Science*, 2024, 4(5): 334-345
- [9] Sun J, Chen Q, Huang Z, et al. Combining denoised neural network and genetic symbolic regression for memory behavior modeling via dynamic asynchronous optimization//*Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. Toronto, Canada, 2025: 2735-2746
- [10] Cheng Xue-Qi, Liu Sheng-Hua, Zhang Ru-Qing. Thinking on new system for big data technology. *Bulletin of Chinese Academy of Sciences*, 2022, 37(1):60-67 (in Chinese)
(程学旗, 刘盛华, 张儒清. 大数据分析处理技术新体系的思考. 中国科学院院刊, 2022, 37(1):60-67)
- [11] Ebbinghaus H. Memory: A contribution to experimental psychology. *Annals of Neurosciences*, 2013, 20(4): 155-156
- [12] Wickelgren W A. Single-trace fragility theory of memory dynamics. *Memory & Cognition*, 1974, 2(4): 775-780
- [13] Woźniak P, Gorzelańczyk E, Murakowski J. Two components of long-term memory. *Acta Neurobiologiae Experimentalis*, 1995, 55(4): 301-305
- [14] Anderson J R, Bothell D, Byrne M D, et al. An integrated theory of the mind. *Psychological Review*, 2004, 111(4): 1036-1060
- [15] Pashler H, Cepeda N, Lindsey R V, et al. Predicting the optimal spacing of study: A multiscale context model of memory//*Proceedings of the 23rd International Conference on Neural Information Processing Systems*. Vancouver, Canada, 2009: 1321-1329
- [16] Lindsey R V, Shroyer J D, Pashler H, et al. Improving students' long-term knowledge retention through personalized review. *Psychological Science*, 2014, 25(3): 639-647
- [17] Settles B, Meeder B. A trainable spaced repetition model for language learning//*Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (volume 1: long papers)*. Berlin, Germany, 2016: 1848-1858
- [18] Thorndike E. *Animal intelligence: Experimental studies*. London: Routledge, 2017
- [19] TFitts P M, Posner M I. *Human performance*. California: Brooks/Cole, 1967
- [20] Singley M K, Anderson J R. *The transfer of cognitive skill*. Cambridge, USA: Harvard University Press, 1989
- [21] Bransford J D, Brown A L, Cocking R R. *How people learn*. Washington, DC: National Academy Press, 2000
- [22] Cortes R A, Peterson E G, Kraemer D J M, et al. Transfer from spatial education to verbal reasoning and prediction of transfer from learning-related neural change. *Science Advances*, 2022, 8(31): eabo3555
- [23] Schmucker R, Mitchell T M. Transferable student performance modeling for intelligent tutoring systems. *arXiv*, 2022, 2022:2202.03980
- [24] Shen S, Liu Q, Huang Z, et al. A survey of knowledge tracing; Models, variants, and applications. *IEEE Transactions on Learning Technologies*, 2024, 17: 1858-1879
- [25] Corbett A T, Anderson J R. Knowledge tracing: Modeling the acquisition of procedural knowledge. *User Modeling and User-Adapted Interaction*, 1994, 4: 253-278
- [26] Piech C, Bassen J, Huang J, et al. Deep knowledge tracing//*Proceedings of the 29th International Conference on Neural Information Processing Systems*. Montreal, Canada, 2015, 505-513
- [27] Ma B, Hettiarachchi G P, Fukui S, et al. Each encounter counts; Modeling language learning and forgetting//*Proceedings of the 13th International Learning Analytics and Knowledge Conference*. Arlington, USA, 2023: 79-88
- [28] Sun J, Yu F, Wan Q, et al. Interpretable knowledge tracing with multiscale state representation//*Proceedings of the 33rd ACM Web Conference*. Singapore, 2024: 3265-3276

- [29] Liu Z, Liu Q, Chen J, et al. SimpleKT: A simple but tough-to-beat baseline for knowledge tracing//Proceedings of the 11th International Conference on Learning Representations. Kigali, Rwanda, 2023; 20276-20288
- [30] Han K, Yi M Y, Gweon G, et al. Understanding the difficulty factors for learning materials: A qualitative study//Proceedings of the Artificial Intelligence in Education; 16th International Conference. Memphis, USA, 2013; 615-618
- [31] Shu S, Wang L, Tian J. Improving knowledge tracing via considering students' interaction patterns//Proceedings of the 28th Pacific-Asia Conference on Knowledge Discovery and Data Mining. Taipei, China, 2024; 397-408
- [32] Chen J, Liu Z, Huang S, et al. Improving interpretability of deep sequential knowledge tracing models with question-centric cognitive representations//Proceedings of the 37th AAAI Conference on Artificial Intelligence. Washington, USA, 2023; 14196-14204
- [33] Hoffer J G, Ofner A B, Rohrhofer F M, et al. Theory-inspired machine learning—towards a synergy between knowledge and data. *Welding in the World*, 2022, 66(7): 1291-1304
- [34] Wang H, Cao Y, Huang Z, et al. Recent advances on machine learning for computational fluid dynamics: A survey. *arXiv*, 2024, 2024:2408.1217
- [35] Moseley B, Markham A, Nissen-Meyer T. Finite basis physics-informed neural networks (FBPINNs): A scalable domain decomposition approach for solving differential equations. *Advances in Computational Mathematics*, 2023, 49(4): 62
- [36] Cuomo S, Di Cola V S, Giampaolo F, et al. Scientific machine learning through physics-informed neural networks: Where we are and what's next. *Journal of Scientific Computing*, 2022, 92(3): 88
- [37] Park J, Knight B G, Liao Y, et al. CFD-based design optimization of ducted hydrokinetic turbines. *Scientific Reports*, 2023, 13(1): 17968
- [38] Koza J R. *Genetic programming: On the programming of computers by means of natural selection* Cambridge. Cambridge, USA; MIT Press, 1992
- [39] Schmidt M, Lipson H. Distilling free-form natural laws from experimental data. *Science*, 2009, 324(5923): 81-85
- [40] Xu Y, Liu Y, Sun H. Reinforcement symbolic regression machine//Proceedings of the 12th International Conference on Learning Representations, Vienna, Austria, 2024; 20552-20576
- [41] Yu Z, Ding J, Li Y. Discover network dynamics with neural symbolic regression. *Nature Computational Science*, 2025; 1-13
- [42] Cranmer M. Interpretable machine learning for science with PySR and symbolic regression. *arXiv*, 2023, 2023:2305.01582
- [43] Udrescu S M, Tegmark M. AI Feynman: A physics-inspired method for symbolic regression. *Science Advances*, 2020, 6(16): eaay2631
- [44] Hinton G, Vinyals O, Dean J. Distilling the knowledge in a neural network. *arXiv*, 2015, 2015:1503.02531
- [45] Moslemi A, Briskina A, Dang Z, et al. A survey on knowledge distillation: Recent advancements. *Machine Learning with Applications*, 2024, 18: 100605
- [46] Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Computation*, 1997, 9(8): 1735-1780
- [47] Ashish V. Attention is all you need//Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach, USA, 2017; 5998-6008
- [48] Gu A, Dao T. Mamba: Linear-time sequence modeling with selective state spaces. *arXiv*, 2023, 2023: 2312.00752
- [49] Nagatani K, Zhang Q, Sato M, et al. Augmenting knowledge tracing by considering forgetting behavior//Proceedings of the 28th World Wide Web Conference. San Francisco, USA, 2019; 3101-3107
- [50] Cepeda N J, Pashler H, Vul E, et al. Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, 2006, 132(3): 354-380
- [51] Anderson J R. *How can the human mind occur in the physical universe?*. New York: Oxford University Press, 2009
- [52] Newell A, Rosenbloom P S. *Mechanisms of skill acquisition and the law of practice*. New York: Psychology Press, 2013; 1-55
- [53] Pavlik P I, Anderson J R. Practice and forgetting effects on vocabulary memory: An activation-based model of the spacing effect. *Cognitive Science*, 2005, 29(4): 559-586
- [54] Pavlik P I, Anderson J R. Using a model to compute the optimal schedule of practice. *Journal of Experimental Psychology: Applied*, 2008, 14(2): 101-117



SUN Jian-Wen, Ph. D., professor.

His research interests include educational data mining, and computational learning sciences.

CHEN Qi-Rong, M. S. candidate.

His research focuses on educational data mining.

LI Fu-Qing, M. S. candidate. His research interests include intelligent education and learning analytics.

LUO Ya-Wei, Ph. D., researcher. His research interests include computer vision and intelligent education.

LI Qing, Ph. D., researcher. Her main research areas include intelligent education and data mining.

SHEN Xiao-Xuan, Ph. D., associate professor. His research interests focus on deep neural networks and intelligent education.

LIU San-Nyu-Ya, Ph. D., professor, Yangtze River Scholar. His research focuses on intelligent education and educational big data.

Background

This study focuses on a core issue in the field of Educational Data Mining (EDM): how to effectively integrate cognitive psychological theory into EDM tasks. Although deep neural networks have demonstrated strong performance in tasks such as knowledge tracing, learning performance prediction, and learner profiling, their “black-box” nature often results in limited interpretability and weak theoretical alignment, thereby constraining their practical application in educational decision-making and personalized interventions.

To fill this gap, this paper proposes the Psychology-Informed Neural Network (PsyINN), a two-stage modeling framework that integrates interpretable symbolic expressions with data-driven neural optimization. The first stage initializes modeling structures using parameterized formulations grounded in cognitive theory (e.g., forgetting curves, practice laws). The second stage refines these structures through a neural network jointly optimized with symbolic components via dynamic knowledge alignment. Experiments across three real-world educational datasets—MaiMemo (vocabulary memory), EdNet (knowledge tracing), and Lumosity (skill ac-

quisition)—demonstrate the robustness, task generalizability, and theoretical alignment of PsyINN.

This work represents a novel paradigm in AI for education by bridging cognitive theory and neural modeling within a unified and evolvable framework. It provides both symbolic interpretability and empirical adaptability, offering methodological advances for mechanism-oriented modeling in complex learning environments. It also contributes to the long-term vision of AI for Social Sciences and Humanities, where domain knowledge and machine learning are co-optimized to support structural reasoning and scientific understanding.

This research was financially supported by the National Science and Technology Major Project of China (No. 2022ZD0117103), the National Natural Science Foundation of China (NSFC) under Major Program (No. 62293554), the National Natural Science Foundation of China (NSFC) under General Program (No. 62577028), the Youth AI Talents Fund of the Chinese Association of Automation under Major Program (No. HBRC-JKYZD-2024-310), and the Hubei Provincial Natural Science Foundation of China (No. 2023AFA020).