



计算机科学

COMPUTER SCIENCE

融合IRT的图注意力深度知识追踪模型

董永峰, 黄港, 薛婉若, 李林昊

引用本文

董永峰, 黄港, 薛婉若, 李林昊. [融合IRT的图注意力深度知识追踪模型](#)[J]. 计算机科学, 2023, 50(3): 173-180.

DONG Yongfeng, HUANG Gang, XUE Wanruo, LI Linhao. [Graph Attention Deep Knowledge Tracing Model Integrated with IRT](#) [J]. Computer Science, 2023, 50(3): 173-180.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[深度学习模型的后门攻击研究综述](#)

Backdoor Attack on Deep Learning Models:A Survey

计算机科学, 2023, 50(3): 333-350. <https://doi.org/10.11896/jsjcx.220600031>

[特征增强损失与前景注意力人群计数网络](#)

Crowd Counting Network Based on Feature Enhancement Loss and Foreground Attention

计算机科学, 2023, 50(3): 246-253. <https://doi.org/10.11896/jsjcx.220100219>

[基于深度学习的可视化仪表盘生成技术研究](#)

Study on Visual Dashboard Generation Technology Based on Deep Learning

计算机科学, 2023, 50(3): 238-245. <https://doi.org/10.11896/jsjcx.230100064>

[极化自注意力约束颜色溢出的图像自动上色](#)

Polarized Self-attention Constrains Color Overflow in Automatic Coloring of Image

计算机科学, 2023, 50(3): 208-215. <https://doi.org/10.11896/jsjcx.220100149>

[基于特征融合的边缘引导乳腺超声图像分割方法](#)

Segmentation Method of Edge-guided Breast Ultrasound Images Based on Feature Fusion

计算机科学, 2023, 50(3): 199-207. <https://doi.org/10.11896/jsjcx.211200294>

融合 IRT 的图注意力深度知识追踪模型

董永峰^{1,2,3} 黄港^{1,2,3} 薛婉若¹ 李林昊^{1,2,3}

1 河北工业大学人工智能与数据科学学院 天津 300401

2 河北省大数据计算重点实验室 天津 300401

3 河北省数据驱动工业智能工程研究中心 天津 300401

(dongyf@hebut.edu.cn)

摘要 知识追踪,旨在根据学生的历史答题表现实时追踪学生的知识状态(知识的掌握程度)并且预测学生未来的答题表现。目前的研究仅仅探索了问题或概念本身对学生答题表现的直接影响,而往往忽略了问题及包含的概念中存在的深层次信息对学生答题表现的间接影响。为了更好地利用这些深层次信息,一种融合项目反应理论的图注意力深度知识追踪模型 GAKT-IRT 被提出。模型将图注意力网络应用于知识追踪领域,取得了显著的提升效果,并使用 IRT 增加了模型的可解释性。首先,通过图注意力网络层获得问题的深层次特征表示;接着,根据结合了深层次信息的学生历史答题序列对学生的知识状态进行建模;然后,使用 IRT 对学生未来的答题表现进行预测。在 6 个公开真实在线教育数据集上的对比实验结果证明了,GAKT-IRT 模型可以更好地完成知识追踪任务,在预测学生未来答题表现上具有明显的优势。

关键词: 知识追踪;图注意力网络;项目反应理论;深度学习;可解释性

中图法分类号 TP391.6

Graph Attention Deep Knowledge Tracing Model Integrated with IRT

DONG Yongfeng^{1,2,3}, HUANG Gang^{1,2,3}, XUE Wanruo¹ and LI Linhao^{1,2,3}

1 School of Artificial Intelligence, Hebei University of Technology, Tianjin 300401, China

2 Hebei Key Laboratory of Big Data Computing, Tianjin 300401, China

3 Hebei Engineering Research Center of Data-Driven Industrial Intelligent, Tianjin 300401, China

Abstract Knowledge tracing aims to trace students' knowledge state(the degree of knowledge) based on their historical answer performance in real time and predict their future answer performance. The current research only explores the direct influence of the question or concept itself on the performance of students' answering questions, while often ignores the indirect influence of the deep-level information in the questions and the concepts contained on the performance of students' answering questions. In order to make better use of these deep-level information, a graph attention deep knowledge tracing model integrated with IRT (GAKT-IRT) is proposed, which integrates item response theory(IRT). The graph attention network is applied to the field of knowledge tracing and uses IRT to increase the interpretability of the model. First, obtain the deep-level feature representation of the problem through the graph attention network layer. Next, model students' knowledge state based on their historical answer sequence that combines the in-depth information. Then, use IRT to predict students' future answer performance. Results of comparative experiments on 6 open real online education datasets prove that the GAKT-IRT model can better complete the knowledge tracing task and has obvious advantages in predicting the future performance of students in answering questions.

Keywords Knowledge tracing, Graph attention network, Item response theory, Deep learning, Interpretability

1 引言

在 MOOC, Udemy 和 Lynda 等智能教学系统和大规模在线开放课程平台中,知识追踪(Knowledge Tracing, KT)是

一项不可或缺的任务^[1]。其旨在捕获学生在答题过程中知识状态的变化并做出预测。近年来,国内外学者展开了对知识追踪模型的研究,并提出了各种方法以完成知识追踪任务^[2-3]。

到稿日期:2021-12-13 返修日期:2022-06-01

基金项目:国家自然科学基金(61902106, 61806072);河北省高等教育教学改革研究与实践项目(2020GJJG027);河北省自然科学基金(F2020202028)

This work was supported by the National Natural Science Foundation of China(61902106, 61806072), Hebei Province Higher Education Teaching Reform Research and Practice Project(2020GJJG027) and Natural Science Foundation of Hebei Province(F2020202028).

通信作者:李林昊(lilinhao@hebut.edu.cn)

深度知识追踪模型(Deep Knowledge Tracing,DKT)^[4]是首次将神经网络应用于知识追踪的方法,在性能上相较于传统方法有较大的提升。后续的研究同样基于深度学习的方法提出了多种相应的改进框架。DKVMN^[5]使用一个静态矩阵 *key* 存储知识概念的嵌入,使用一个动态的矩阵 *value* 存储并更新学习者对知识概念的掌握程度。卷积知识跟踪模型(Convolutional Knowledge Tracing,CKT)^[6]考虑到学生的先验知识与学习率的不同,引入卷积神经网络对学生的个性化进行建模。Deep-IRT^[7]将 DKVMN 模型与项目反应理论(Item Response Theory,IRT)进行结合,进一步提升了模型的性能。其他的改进模型,如 EERNN^[8],EKT^[9]和 SKVMN^[10]等,性能又得到了进一步提升。

真实的在线教育场景中通常包含多种问题与多种概念。现有 KT 模型通常将问题或者概念的 one-hot 编码格式作为输入来构建预测模型,但往往忽略了问题及包括的概念之间的深层次信息;同时,由于深度模型的黑盒性质,人们难以理解学生的知识状态表示。对于真实的在线教育系统,学生只有在看到一个令人信服的认知诊断后才会接受系统所推荐的项目,因此人们更倾向于具有可解释性的 KT 算法。

鉴于此,本文提出了一种新的知识追踪模型——融合 IRT 的图注意力深度知识追踪模型(Graph Attention Deep Knowledge Tracing Model Integrated with IRT,GAKT-IRT)。该模型探索了图注意力网络聚合深层次信息的能力和 IRT 建模学生未来答题表现的可解释性能力,使用图注意力层得到新的问题表示,并结合学生的答题情况,通过知识状态追踪层对学生的知识状态进行建模,然后通过难度分析层得到问题的难度表示,最后经过 IRT 预测层对学生未来的答题表现进行预测。本文的主要创新与贡献如下:

(1)将图注意力网络模型(Graph Attention Network,GAT)引入知识追踪领域,通过 GAT 学习到问题之间、概念之间以及两者之间的依赖关系,有助于精确捕获学生的知识状态,从而提升了模型的性能;

(2)结合认知心理学理论 IRT,从学生的能力、问题的难度和项目区分度 3 个方面对学生的答题表现进行预测,提升了模型的性能并增强了模型的可解释性;

(3)在 6 个真实公开的在线教育数据集上评估 GAKT-IRT 模型,结果表明 GAKT-IRT 模型在预测学生未来答题表现方面明显优于现有的流行模型。

2 相关工作

2.1 知识追踪

知识追踪任务的研究中,模型主要可以分为两大类。一类是基于传统方法的模型,另一类则是基于深度学习方法的模型。

传统方法注重使用认知与心理学理论完成知识追踪任务,主要分为贝叶斯模型与因子分析模型。贝叶斯知识追踪(Bayesian Knowledge Tracing,BKT)^[11]将每个技能视为一个二元变量,并使用贝叶斯规则更新状态。后来的研究拓展了普通的 BKT 模型,将更多的信息纳入其中,例如猜测概率、问题难度和学生的个性等。因子分析模型侧重从历史数据中

学习特定参数以进行预测。性能因子分析(Performance Factors Analysis,PFA)^[12]考虑了学生的答题情况对问题表示的影响,知识追踪机器(Knowledge Tracing Machines,KTm)^[13]利用分解机理论^[14]将问题与学生的交互特征加入到参数模型中进行讨论。

在基于深度学习的模型中,深度知识追踪模型(Deep Knowledge Tracing,DKT)是首次将深度学习引入知识追踪领域的模型,使用循环神经网络(Recurrent Neural Network,RNN)追踪学生的知识状态。后来的研究对该模型进行了改进和完善。Chen 等^[15]考虑到输入数据的稀疏性,提出了将知识结构的信息纳入模型。Minn 等^[16]根据学生能力的不同,对学生进行动态分组。Ghosh 等^[17]针对当前深度模型解释性不足的问题,提出了注意力知识追踪模型(Attentive Knowledge Tracing,AKT)。Li 等^[18]考虑了 4 个知识遗忘因素,建模学生的学习与遗忘行为。Wang 等^[19]引入时间卷积神经网络,减少了模型计算时间。Liu 等^[20]使用二部图表示问题与概念之间的复杂关系。GKT^[21]模型建立了技能关系图,并使用图神经网络(Graph Neural Network,GNN)探索技能之间的复杂关系。Wang 等^[22]针对现有工作未充分利用数据集中其他数据的问题,提出了一种融合多特征的知识追踪模型 DKTwMF。Huang 等^[23]针对 BKT 容易忽视记忆遗忘现象以及未考虑学习行为对表现结果产生影响的问题,提出一种融合学习者的行为和遗忘因素的贝叶斯知识追踪模型 BF-BKT。

2.2 项目反应理论

项目反应理论^[24],也被称为潜在特质理论,是一系列心理统计学模型的总称,广泛应用于各种领域。IRT 通过分析学生与问题之间的互动关系,建模学生的能力,进而预测学生正确回答问题的概率。学生正确回答问题的概率 p 可以由标准的 2 参数项目反应函数^[25]表示:

$$p(\theta) = \frac{1}{1 + e^{-\alpha(\theta - \beta)}} \quad (1)$$

其中, α 为项目区分度系数, θ 表示学生的潜在能力, β 表示问题的难度。一般来说,问题的难度越高,学生就需要具备更高的能力才能正确回答该问题。项目区分度系数是针对不同的测试项目而设置的影响系数。

2.3 图注意力神经网络

近年来,卷积神经网络(Convolutional Neural Networks,CNN)^[26]在图结构领域的应用越来越广泛,但其只能处理规则图结构的数据。为了处理不规则图结构数据,研究者提出了图卷积神经网络(Graph Convolutional Network,GCN)^[27]。GCN 使用拉普拉斯矩阵变换将空间域信息映射到谱域上进行卷积运算。具体来说,卷积运算将每个图节点的所有相邻特征向量聚合起来后放入全连接层,并使用非线性激活函数进行变换,得到更新之后的图节点表示。但是,GCN 聚合邻居节点信息时对每个邻居节点按照相同权重进行计算,不能关注到邻居节点对中心节点的影响程度。为解决这一问题,Veličković 等结合注意力机制提出了 GAT^[28]。GAT 更加关注影响力较大的节点而忽略影响力较小的噪声节点,并且为了提高注意力机制的泛化能力,使用了多头注意力机制。

GAT 已在多个领域中取得了良好的效果^[29-31]。类似地,根据问题与概念之间的关系对学生未来答题表现的影响,GAT 也适用于知识追踪领域。

3 GIKT-IRT 算法

3.1 问题定义

定义 1(知识追踪任务定义) 在知识追踪任务中,假设 $S = \{s_i\}_{i=1}^{|S|}$ 是长度为 $|S|$ 的学生集合, $Q = \{q_i\}_{i=1}^{|Q|}$ 是长度为 $|Q|$ 的问题集合, $C = \{c_i\}_{i=1}^{|C|}$ 是长度为 $|C|$ 的概念集合,且每个学生独立地完成 Q 中的问题。给定一个学生的历史互动序列 $X = \{(q_1, r_1), (q_2, r_2), (q_3, r_3), \dots, (q_t, r_t)\}$,其中 r_t 表示学生在 t 时刻是否能正确回答 q_t 。知识追踪任务中的输入数据是学生历史答题序列的 one-hot 编码格式。知识追踪任务基于给定的互动序列 X 追踪学生的知识状态,进而预测学生在 $t+1$ 时刻的表现,即当学生面对新的问题 q_{t+1} 时,预测其是否能正确回答。该过程可表示为 $p_{t+1} = \{r_{t+1} = 1 | q_{t+1}, X\}$ 。

定义 2(问题概念关系图定义) 为了探索问题之间、概念之间以及两者之间存在的关系,本文建立了问题概念关系图来进行探索。给定问题与概念的合并集合 $T = \{t_i\}_{i=1}^{|T|}$, $t_i \in \{q_j, c_m\}$, $j \in |Q|$, $m \in |C|$ 。定义问题概念关系图 $G = \{t_i, t_j, b_{ij}\}$,其中 $b_{ij} \in \{0, 1\}$ 。 $b_{ij} = 1$ 表示节点 t_i 与节点 t_j 之间存在边,反之不存在。

3.2 GAKT-IRT 模型

GAKT-IRT 模型框架如图 1 所示,主要包括嵌入层、图注意力网络层、LSTM 训练层、难度分析层以及 IRT 预测层。首先,嵌入层将问题映射到高维空间获得多维特征表示;然后,图注意力网络层结合邻居信息获得新的问题表示,其中包含问题之间、概念之间以及问题与概念之间的关系;得到问题的新表示之后,结合当前学生的答题情况,知识状态追踪层追踪到学生随时间变化的知识状态,同时难度分析层分析当前问题表示得出难度的特征信息表示;最后,IRT 预测层预测学生未来的答题表现。

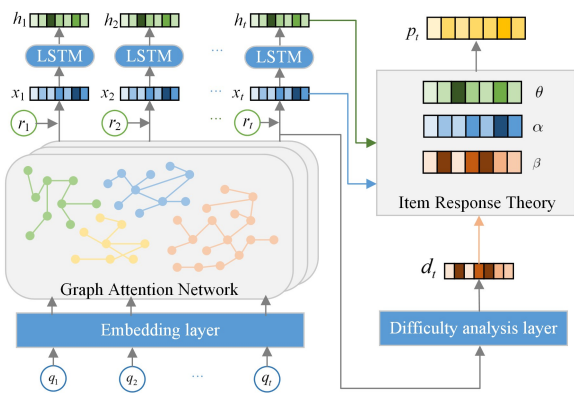


图 1 GAKT-IRT 模型框架

Fig. 1 GAKT-IRT model framework

3.2.1 嵌入层

由于知识追踪任务的输入数据是一个较为单一、信息量较少且稀疏的数据,因此可供模型学习的信息就较少。根据以往研究的经验,通常采用嵌入的方式将数据映射到高维空间以获得多维特征信息。为了获得数据的更多信息,本文也

采用嵌入的方式。本层将学生历史序列中的问题部分提取出来,对每个问题进行嵌入操作。具体过程如下:

$$\bar{q}_t = W_q \cdot q_t + b_q \quad (2)$$

其中, $W_q \in \mathbb{R}^{N \times N}$, $b_q \in \mathbb{R}^N$, N 表示嵌入的维度, \bar{q}_t 表示经过嵌入之后问题的新表示。

3.2.2 图注意力网络层

GAT 是基于 GCN 并融入注意力机制的改进模型,能够注意到邻居节点对中心节点的影响程度并为邻居节点分配不同的权重。当前的知识追踪研究并未将 GAT 引用到知识追踪领域,但 GAT 模型的特征较为契合于知识追踪领域。一般地,为了精确地追踪到学生知识状态的变化,全面的问题信息是不可或缺的。在真实的教育环境中,一个问题通常包含多种概念,且各种问题之间和概念之间存在联系。除此之外,每种联系对问题本身的影响程度又是不尽相同的,而 GAT 能聚合邻居节点的信息并根据不同的影响程度分配权重,因此本文认为 GAT 能够捕获到问题、概念以及两者之间的关系。

根据第 3.1 节中图的定义,给定一个问题概念关系图 $G = \{t_i, t_j, b_{ij}\}$ 。首先,为了计算节点之间的注意力系数,本层使用自注意力机制作用于所有图节点,使用共享的权重矩阵作用于邻居节点,然后使用非线性激活函数 LeakyReLU 获得节点之间的关系系数。具体过程如下:

$$v_{ij} = \text{LeakyReLU}(W \cdot [W_t \cdot t_i \parallel W \cdot t_j]) \quad (3)$$

其中, \parallel 表示连接操作, \cdot 表示内积操作。对 v_{ij} 进行归一化操作,得到节点的注意力权重值,具体过程如下:

$$a_{ij} = \text{softmax}(v_{ij}) = \frac{\exp(v_{ij})}{\sum_{k \in N_i} \exp(v_k)} \quad (4)$$

其中, softmax 表示激活函数, N_i 表示节点 t_i 的邻居节点的数量, a_{ij} 表示节点 t_i 与其邻居节点 t_j 的注意力权重。为了稳定地学习到注意力权重,本层也使用多头注意力层^[32] 计算每个中心节点的权重,通过设置 K 个独立的注意力机制得到 K 个不同的注意层面,然后取 K 个注意层面的平均值,最后使用非线性激活函数得到最终的节点表示。具体过程如下:

$$t_i^{(h)} = \sigma\left(\frac{1}{K} \sum_{k=1}^K \sum_{j \in N_i} a_{ij}^k \cdot W_t^k \cdot t_j\right) \quad (5)$$

其中, K 表示设置的多头注意力个数, $t_i^{(h)}$ 表示第 h 层的节点表示。为了获得更加全面的节点表示,通常都会设置多个图注意力层来进行训练,经过 h 层的图注意力层训练之后,每个中心节点可以获得 h 阶邻居节点的信息。

3.2.3 知识状态追踪层

由 3.2.2 节图注意力网络层可以得到包含更多信息的问题表示 \hat{q}_t ,而在获取学生知识状态的过程中应考虑学生的答题情况 r_t 对知识状态变化的影响。因此,本层将 \hat{q}_t 与 r_t 进行拼接操作,得到本层输入数据。本层输入数据定义为 $\bar{X} = \{\bar{x}_1, \bar{x}_2, \bar{x}_3, \dots, \bar{x}_t\}$,其中 $\bar{x}_t = \hat{q}_t \parallel r_t$, \parallel 表示拼接操作。

本层采用 LSTM^[33] 来捕获学生的知识状态随时间的变化情况,它在深度知识追踪任务中展现了出色的性能。LSTM 通过 3 个门来获取学生的知识状态 h_t ,分别是遗忘门 f_t ,输入门 i_t ,输出门 o_t ;除此之外,使用细胞状态 C_t 将

历史信息传递到每个单元,用来解决 RNN 不能捕获长期依赖的问题。

遗忘门:在真实场景中,随着时间的推移,学生会逐渐遗忘以前学习过的某些知识。遗忘门使用一个 0 到 1 的数作为学习过程中的遗忘率来模拟遗忘过程,在 t 时刻其公式为:

$$f_t = \sigma(W_f \cdot [h_{t-1}, \bar{x}_t] + b_f) \quad (6)$$

其中, σ 是 sigmoid 激活函数, $W_f \in \mathbb{R}^{N \times 3N}$, $b_f \in \mathbb{R}^N$ 是可训练的参数; h_{t-1} 表示学生在 $t-1$ 时刻的知识状态, \bar{x}_t 表示在 t 时刻的输入。

输入门:输入门模拟学生在面对一个问题时更新自身知识状态的过程,其中包含新知识的学习以及旧知识的复习。

i_t 表示模拟新知识的学习过程, \bar{C} 表示模拟旧知识掌握程度的变化过程, C_t 表示将两者相结合形成新的细胞状态。这个过程接收新的输入并更新当前的细胞状态。在 t 时刻,其公式为:

$$i_t = \sigma(W_i \cdot [h_{t-1}, \bar{x}_t] + b_i) \quad (7)$$

$$\bar{C}_t = \tanh(W_C \cdot [h_{t-1}, \bar{x}_t] + b_C) \quad (8)$$

$$C_t = f_t \otimes C_{t-1} + i_t \otimes \bar{C}_t \quad (9)$$

其中, $W_i \in \mathbb{R}^{N \times 3N}$, $b_i \in \mathbb{R}^N$, $W_C \in \mathbb{R}^{N \times 3N}$, $b_C \in \mathbb{R}^N$ 是可训练的参数, \tanh 表示双曲正切激活函数, C_{t-1} 表示 $t-1$ 时刻的细胞状态, \otimes 代表向量对应元素相乘。

输出门:输出门模拟学生的知识状态,根据当前学习的知识与历史知识遗忘的变化过程,输出当前学生的知识状态 h_t 。在 t 时刻,其公式为:

$$o_t = \sigma(W_o \cdot [h_{t-1}, \bar{x}_t] + b_o) \quad (10)$$

$$h_t = o_t \otimes \tanh(C_t) \quad (11)$$

其中, $W_o \in \mathbb{R}^{N \times 3N}$, $b_o \in \mathbb{R}^N$ 是可训练的参数, h_{t-1} 表示在 $t-1$ 时刻学生的知识状态。

3.2.4 难度分析层

在真实的教育环境中,每个问题的难度是不同的。以数学为例,简单的四则运算与微积分运算的难度是不一样的。因此,在预测学生未来的答题表现的过程中,不仅需要考虑到学生的知识状态即学生的知识掌握程度,而且需要考虑问题本身的难度。除此之外,还有一个情况也需要考虑,如学生掌握了四则运算、角度运算等基本技能,但是将多种技能融合到一个问题中,该学生很可能会回答错误。因此,考虑问题本身的难度是知识追踪中必不可少的一环。本层使用一个难度矩阵来表示问题本身的难度,具体过程如下所示:

$$d_t = W_d \cdot \hat{q}_t + b_d \quad (12)$$

其中, $W_d \in \mathbb{R}^{N \times N}$, $b_d \in \mathbb{R}^N$ 。由于问题的难度仅与问题本身的难度相关,因此本文仅针对问题本身进行矩阵操作,而没有加入学生的答题情况。除此之外,本文认为问题的难度可以从相似问题的难度或者从相关的概念进行评估,因此我们将经过图注意力层得到的新的问题表示 \hat{q}_t 进行矩阵变换来表示问题的难度。

3.2.5 IRT 预测层

IRT 作为统计心理学中经典的理论,在知识追踪中表现出强大的可解释性。本层采用的 2 参数 IRT 模型主要包含

3 个参数,分别是项目区分度系数 α 、学生的能力 θ 以及问题的难度 β 。在知识追踪任务中,学生的能力是基于学生的知识状态的,学生的知识状态越好,其能力就越强;问题的难度只与问题本身相关^[34];项目区分度系数不仅在于问题本身,也与学生回答的情况紧密相关。在知识追踪任务中,学生通常需具备多种能力才能正确回答问题,因此,我们将项目区分度系数 α 表示为学生的不同能力对正确回答问题的指标的权重系数。我们将 2 参数 IRT 模型的参数定义如下:

$$\alpha = \sigma(W_\alpha \cdot x_t + b_\alpha) \quad (13)$$

$$\theta = \tanh(W_\theta \cdot h_t + b_\theta) \quad (14)$$

$$\beta = \tanh(W_\beta \cdot d_t + b_\beta) \quad (15)$$

其中, $W_\alpha \in \mathbb{R}^{N \times 2N}$, $b_\alpha \in \mathbb{R}^N$, $W_\theta \in \mathbb{R}^{N \times N}$, $b_\theta \in \mathbb{R}^N$, $W_\beta \in \mathbb{R}^{N \times N}$, $b_\beta \in \mathbb{R}^N$, h_t 表示 t 时刻学生的知识状态。然后根据 IRT 理论对学生未来的答题表现进行预测。在 t 时刻,其公式为:

$$p_t = \sigma(W_p \cdot [\alpha(\theta - \beta)] + b_p) \quad (16)$$

其中, p_t 是一个在 $[0, 1]$ 范围内的标量,代表学生正确回答 t 时刻问题的概率。当 $p_t \in [0, 0.5]$ 时,本文判定学生回答错误;反之,判定回答正确。

3.2.6 模型优化目标

为了更好地训练模型,本文建立了一个联合训练框架来使得模型的参数达到最优。一方面,在问题概念图中,为了提取问题之间或者概念之间以及两者之间的关系,我们使用内积的方式来评估图节点之间的关系。节点间的关系表示如下:

$$\hat{b}_{ij} = \sigma(t_i^T \cdot t_j) \quad (17)$$

其中, $i, j \in [1, \dots, |T|]$; σ 表示 sigmoid 激活函数,将节点间的关系转化为 $[0, 1]$ 的值。为了使提取的节点之间的关系更加贴近真实的结果,我们构建了图的交叉熵损失函数:

$$L_1 = \sum_{i=1}^{|T|} \sum_{j=1}^{|T|} -(b_{ij} \log \hat{b}_{ij} + (1 - b_{ij}) \log(1 - \hat{b}_{ij})) \quad (18)$$

另一方面,为了使得模型能够更加精准地预测学生未来的答题表现,将预测学生答题表现的交叉熵损失函数定义如下:

$$L_2 = \sum_t -(r_t \log p_t + (1 - r_t) \log(1 - p_t)) \quad (19)$$

针对以上两个方面的训练目标,构建一个联合训练目标,具体过程如下:

$$L = \lambda_1 L_1 + \lambda_2 L_2 \quad (20)$$

其中, λ_1, λ_2 表示控制图损失与预测损失的权衡系数。

3.3 算法伪代码

GAKT-IRT 算法的伪代码如算法 1 所示。

算法 1 GAKT-IRT 算法训练过程

输入:学生的历史互动序列 $X = \{x_1, x_2, x_3, \dots, x_t\}$ 和问题概念关系图 G 及其邻接矩阵

输出:学生未来的答题表现

1. 对输入数据中的问题部分进行嵌入操作,得到问题的嵌入表示 \bar{q}_t

2. 通过图注意力网络得到包含深层次信息的表示 $t_i^{(h)}$

$$t_i^{(h)} = \text{GAT}(\bar{q}_t, t_j)$$

3. For each epoch:

4. 将问题的新表示 \hat{q}_t 与学生的答题情况 r_t 进行拼接,得到 \bar{x}_t

5. 根据 LSTM 网络捕获学生变化的知识状态 h_t

$$h_t = \text{LSTM}(\bar{x}_t, h_{t-1})$$

6. 根据问题的新表示 \hat{q}_t , 获取问题的难度表示 d_t

7. 根据 IRT 理论预测学生的未来答题表现 p_t

$$p_t = \text{IRT}(x_t, h_t, d_t)$$

8. 计算损失函数, 更新模型参数

9. End for

本算法首先使用 GAT 捕获问题及其包含的概念中的深层次信息; 然后结合学生的答题情况, 使用 LSTM 捕获学生的知识状态; 最后基于学生的知识状态, 使用 IRT 理论预测学生未来的答题表现。

4 实验

本文在 6 个公开数据集上对 GAKT-IRT 模型以及现有模型进行实验, 以验证本文模型所提框架在捕获问题及包含的概念之间的联系与知识追踪任务中的有效性。

4.1 数据集

本文在 6 个公开数据集上测试模型的性能: ASS09-up, ASSIST2012, statics2011, synthetic, AICFE-math, AICFE-phy。具体的统计信息如表 1 所列, 其中“Stu”“Qes”“Con”“Res”分别代表学生数量、问题数量、概念数量和 student 答题数据总量。

表 1 数据集的统计信息
Table 1 Statistics of datasets

Dataset	Stu	Qes	Con	Res
ASS09-up	4151	16891	110	325637
ASSIST2012	27405	47104	265	1867167
statics2011	333	1223	98	189297
synthetic	2000	50	5	200000
AICFE-math	2445	617	44	616462
AICFE-phy	964	483	54	501031

ASS09-up 数据集是 ASSISTments 在线辅导平台在 2009—2010 年所收集的。其数据中存在大量重复的记录, 知识追踪学者们重新整理之后发布了当前版本^[35]。ASSIST2012 数据集是 ASSISTments 在线辅导平台在 2012 年收集完成。statics2011 数据集来自 2011 年大学 Oli 工程静力学课程。synthetic 数据集模拟了 2000 名虚拟学生做习题的过程, 每个问题均来自 5 个隐藏的概念之一, 难度各不相同。AICFE-math 数据集是从在线教育平台——智能学习者平台 (SLP) 收集的数据, math 代表的是数学学科的数据。AICFE-phy 数据集同样来自 SLP, 是物理学科的数据。

4.2 对比方法

为了说明所提模型的有效性并显示模型相比现有 KT 模型的优势, 本文将其与当前流行的 KT 模型进行了对比实验。所有模型的参数按照研究者提供的最佳参数进行设置。对比模型的详细信息如下。

DKT^[4]: 第一个将知识追踪任务与循环神经网络结合的深度学习模型。相比传统方法, DKT 模型性能更出色。

KTM^[13]: 使用因式分解机表征学生的多种特征来进行预测, 是最新的因式分解模型。虽然 KTM 可以使用多种类型的特征, 但为了公平起见, 我们仅使用问题、概念和答案

作为其辅助信息进行比较。

TCN^[18]: 为解决 RNN 梯度消失或者梯度爆炸等问题, Li 等引入了时间卷积网络并融合了学习者个人先验基础和遗忘因素。为了方便展示, 本文简称其模型为 TCN。

LFKT^[19]: 为了准确建模知识追踪中的学习与遗忘行为, Wang 等考虑了 4 个影响知识遗忘因素, 建模了融合学习与遗忘行为的知识追踪模型 LFKT。

DKVMN^[5]: 将改进的记忆增强网络与知识追踪任务相结合。该模型设置了一个静态矩阵用于存储知识概念, 以及一个动态矩阵用于存储和更新知识的掌握水平。

4.3 评价指标与基本参数设置

本实验采用 AUC 评价指标, 它是应用广泛的评价指标之一。AUC 被定义为 ROC 曲线下与坐标轴围成的面积, 其取值范围在 0.5 和 1 之间。AUC 值越接近 1.0, 表示使用方法的真实性越高; 等于 0.5 时, 真实性最低, 无应用价值。

在本文中, 所有模型都在两块显存为 12GB 的 RTX 2080 Ti 显卡上进行实验。我们将每个数据集中的 80% 划分为训练集, 20% 划分为测试集, 其中测试集用来评价模型的性能与提前停止模型训练。在所有数据集上都使用 5-折交叉验证的方法, 并且所有模型对每个数据集都采用 5 次实验取平均值的方式进行性能评估。本文模型使用 Tensorflow 框架、Adam 优化器来训练模型, 设置最大训练次数为 500, 裁剪最大梯度范数设置为 5.0, 学习率设置范围为 [0.001, 0.01], 网络中权重矩阵与偏置矩阵的初始化使用的是平均值为 0、标准差为 0.01 的正态分布。训练的批次一般设置为 64, 但是具体会按照数据集的大小进行设置, 例如 a2012 数据集规模较大, 训练的批次设置为 32。为防止过拟合情况的发生, 使用了 Dropout, 在训练的时设置参数为 0.5。

4.4 实验结果与分析

本节对比了 5 个当前流行的模型在 6 个公开数据集上的 AUC 结果, 如图 2 所示。从结果可以看出 GAKT-IRT 模型的性能明显优于其他模型, 证明了本文模型在预测学生未来答题表现中的有效性。具体来说, GAKT-IRT 模型在所有数据集上的性能至少优于其他模型 1.1%, 其中在 ASS09-up, ASSIST2012, AICFE-math 以及 AICFE-phy 数据集上性能提升尤为明显。在图 2 中, DKVMN 模型与 LFKT 模型的性能仅次于 GAKT-IRT 模型, 而且基于深度学习的模型在所有数据集上的性能都较为优秀, 这验证了深度学习在知识追踪领域的有效性。值得注意的是, KTM 模型在某些数据集上仍具有不错的表现, 特别是在 synthetic 和 AICFE-phy 数据集上。

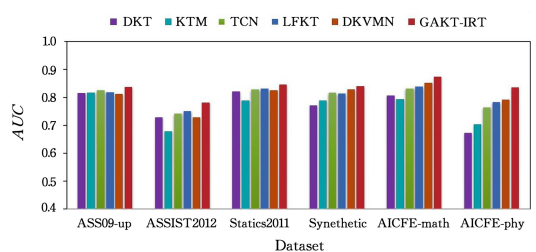


图 2 模型在所有数据集上的 AUC 结果

Fig. 2 AUC results of models on all datasets

为了进一步证明本文模型的有效性,本节展示了各模型在所有数据集上的平均性能,如表 2 所列。从表 2 可知,本文模型的平均性能要高出次优模型大概 3%,KTM 模型的平均性能与 DKT 模型的平均性能仅存在微小的差距。此外,TCN,LFKT 以及 DKVMN 模型的平均性能也相差无几。

表 2 模型在所有数据集上的平均 AUC 结果

Table 2 Average AUC results of models on all datasets

	DKT	KTM	TCN	LFKT	DKVMN	GAKT-IRT
AUC	0.7696	0.7618	0.8017	0.8065	0.8064	0.8360

此外,本文对比了不同模型的训练时间效率,如表 3 所列。具体来说,本文比较了不同模型在 ASS09-up 数据集上训练单个批次的平均运行时间。由表 3 可知,KTM 模型是最快的,LFKT 模型是最慢的,本文模型与 DKT 模型、TCN 模型的训练时间具有相似的复杂性。

表 3 各模型训练时间效率对比

Table 3 Comparison of training time efficiency of each model

(单位:s)

	DKT	KTM	TCN	LFKT	DKVMN	GAKT-IRT
Time	29.65	6.83	28.87	48.32	14.57	32.22

4.5 参数对比实验

为了进一步观察某些重要参数对模型的影响,本文制作了详细的参数对比实验,并在此过程中确定了模型的最佳参数配置。

4.5.1 不同图结构的对比

为了探索问题之间、概念之间以及两者之间的复杂关系,本文构建问题与概念图 $G = \{t_i, t_j, b_{ij}\}$,并且使用图注意力网络来探索其关系。为了验证本文所建图的有效性,本节分别构建概念图 $G_c = \{c_i, c_j, b_{ij}^c\}$ 与问题图 $G_q = \{q_i, q_j, b_{ij}^q\}$ 与本模型的图结构进行对比。由于基于概念图的知识追踪模型只讨论概念之间的联系,因此不再使用问题表示作为模型的输入,而是替换为相应的概念表示。本文在所有数据集上对最优图

结构进行了探索,如表 4 所列,其中 GAKT-IRT-S 表示基于概念图 G_c 的模型,GAKT-IRT-Q 表示基于问题图 G_q 的模型。

表 4 不同图结构的 AUC 结果

Table 4 AUC results of different graph structures

Dataset	GAKT-IRT-S	GAKT-IRT-Q	GAKT-IRT
ASS09-up	0.8268	0.8316	0.8375
ASSIST2012	0.7582	0.7734	0.7821
statics2011	0.8354	0.8315	0.8461
synthetic	0.8148	0.8296	0.8398
AICFE-math	0.8594	0.8627	0.8749
AICFE-phy	0.8187	0.8239	0.8356

从表 4 中可以看出基于图 G 的 AUC 结果明显优于其他图结构,说明本文提出的图结构不仅能够捕获问题之间和概念之间的关系,也能捕获到问题与概念之间的关系,能够提供更多的深层次特征信息并提升模型的性能。而基于问题图 G_q 的 GAKT-IRT-Q 模型的性能在大多数数据集上要优于基于概念图 G_c 的 GAKT-IRT-S 模型。经过反复实验得出的结论是概念图 G_c 中的节点数量要明显少于问题图 G_q ,因此模型能够从问题图中获得更多的信息。

4.5.2 图注意力网络层数的对比

本文创造性地引入图注意力网络来捕获问题的更多信息并得到结合深层次信息的表示,其中图注意力网络层数的设计是至关重要的一环,因此本文针对图注意力层数做了详细的实验。具体实验设计如下:采用单一变量对比实验方法,在保持模型其他参数不变的情况下,对比了不同图注意力网络层数对实验结果的影响。

表 5 列出了本文模型设置不同图注意力网络层数的 AUC 结果。从中可以看出,在所有数据集上,层数为 3 时模型的性能最优;层数为 2 的模型的性能要优于层数为 1 的模型的性能。这也说明了图注意力网络层能够聚合深层次特征信息,进而模型能够利用这些信息更好地训练并提升整体的性能。

表 5 不同图注意力网络层数的 AUC 结果

Table 5 AUC results of different graph attention network layers

Layers	ASS09-up	ASSIST2012	statics2011	synthetic	AICFE-math	AICFE-phy
1	0.8315	0.7761	0.8385	0.8247	0.8613	0.8316
2	0.8345	0.7784	0.8420	0.8362	0.8653	0.8348
3	0.8375	0.7821	0.8461	0.8398	0.8749	0.8356

4.6 图注意力网络的有效性

为了进一步验证图注意力网络在知识追踪任务中的有效性,本文将 ASS09-up 数据集中所有的概念进行聚类可视化,如图 3 所示,其中数字表示不同概念对应的标号。由于问题序列数量过于庞大,为了方便并提升可视化效果,我们随机选取了部分问题进行聚类可视化操作,其结果如图 4 所示。本文将所有概念与随机抽取的部分问题进行聚类,并且不同类别使用不同的颜色进行区分。由图 3 可知,经过本模型训练之后的概念表示能够明显地被区分为 8 个类别。本文选取了其中一个类别进行分析,如图 3 左下角橙色标签的类别。该类别包含的 6 个概念描述分别为 Median, Mean, Range, Proportion, Percent, Finding Percents。通过描述可以看出该类别是对数值的相关操作,由此可以证明

GAT 可以很好地对概念进行分类。

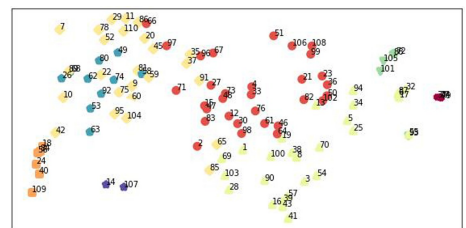


图 3 ASS09-up 数据集上所有概念的聚类结果(电子版为彩图)

Fig. 3 Results of clustering all concepts on ASS09-up dataset

除此之外,由图 4 可知,本模型能够较为明显地将不同的问题进行分类。由此可见,图注意力网络能够较好地捕获问题之间、概念之间以及两者之间的复杂关系,得到了包含更深

层次信息的节点表示,同时证明了图注意力网络在知识追踪任务中的有效性。

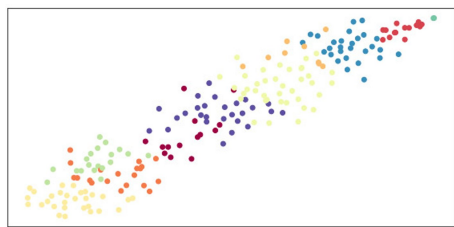


图4 ASS09-up数据集上部分问题的聚类结果(电子版为彩图)

Fig.4 Results of clustering some problems on ASS09-up dataset

4.7 IRT 模块的有效性

学生的能力诊断对于学生的个性化导学至关重要,也是评估知识追踪方法的标准之一。本文模型使用 IRT 函数中的参数 α , θ 和 β 对学生的能力进行量化描述。为了说明本 IRT 模块能够对学生的能力进行诊断,本节从 ASS09-up 数据集中随机抽取一个问题(本文模型预测学生正确回答该问题的概率为 0.37)来分析 IRT 模块中 3 个参数怎么进行学生能力的诊断。学生面对一个新问题时,3 个参数的可视化结果如图 5 所示。横坐标表示每个参数的多维特征序列,纵坐标表示每个特征对应的参数的值。

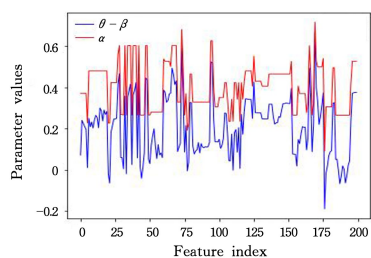


图5 ASS09-up数据集上某个问题的 IRT 模型参数的可视化结果

Fig.5 Visualization results of IRT model parameters of a problem on ASS09-up dataset

由图 5 可知,项目区分度参数 α 在 $[0.1, 0.7]$ 区间波动,显示出不同特征对学生能力的不同贡献度。 $\theta - \beta$ 表示学生的能力与问题难度的匹配程度,大多数特征维度的学生能力相对较低,甚至部分特征维度中学生的能力并不能达到问题难度的估值,说明该学生不能正确回答该问题。这也与真实的情况符合。因此,通过以上分析可以得出,IRT 模块能够较为准确地从不同维度特征去表示学生的能力,体现了其有效性。

结束语 本文创造性地将图注意力网络应用到知识追踪领域,提出了融合 IRT 的图注意力深度知识追踪模型 GAKT-IRT。在 GAKT-IRT 中,通过图注意力层聚合邻居信息得到融合新信息的问题表示,为知识状态追踪层提供更多有用的信息;通过 IRT 预测层量化学生的能力,增加了模型的可解释性。通过实验可知,GAKT-IRT 模型在 6 个真实公开的数据集上的表现优于现有流行的 KT 模型,证明了本文模型的有效性。但是知识追踪任务的输入数据较为稀疏,未来将考虑如何处理知识追踪任务中图结构稀疏的问题。

参考文献

[1] ANDERSON A, HUTTENLOCHER D, KLEINBERG J, et al.

Engaging with massive online courses[C]// Proceedings of the 23rd International Conference on World Wide Web. 2014:687-698.

- [2] VIE J J. Knowledge Tracing Machines: Towards an Unification of DKT, IRT & PFA[C]// ITS Workshops. 2018:149.
- [3] INJADAT M N, MOUBAYED A, NASSIF A B, et al. Systematic Ensemble Model Selection Approach for Educational Data Mining[J]. Knowledge-Based Systems, 2020, 200: 105992.
- [4] PIECH C, SPENCER J, HUANG J, et al. Deep knowledge tracing[J]. arXiv:1506.05908, 2015.
- [5] ZHANG J, SHI X, KING I, et al. Dynamic key-value memory networks for knowledge tracing[C]// Proceedings of the 26th International Conference on World Wide Web. 2017:765-774.
- [6] SHEN S, LIU Q, CHEN E, et al. Convolutional knowledge tracing: Modeling individualization in student learning process [C]// Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval. 2020:1857-1860.
- [7] YEUNG C K. Deep-IRT: Make deep learning based knowledge tracing explainable using item response theory[J]. arXiv:1904.11738, 2019.
- [8] SU Y, LIU Q, LIU Q, et al. Exercise-enhanced sequential modeling for student performance prediction[C]// Proceedings of the AAAI Conference on Artificial Intelligence. 2018.
- [9] LIU Q, HUANG Z, YIN Y, et al. Ekt: Exercise-aware knowledge tracing for student performance prediction [J]. IEEE Transactions on Knowledge and Data Engineering, 2019, 33(1): 100-115.
- [10] ABDELRAHMAN G, WANG Q. Knowledge tracing with sequential key-value memory networks[C]// Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval. 2019:175-184.
- [11] CORBETT A T, ANDERSON J R. Knowledge tracing: Modeling the acquisition of procedural knowledge[J]. User Modeling and User-adapted Interaction, 1994, 4(4): 253-278.
- [12] PAVLIK JR P I, CEN H, KOEDINGER K R. Performance Factors Analysis—A New Alternative to Knowledge Tracing[J]. Online Submission, 2009, 15(3): 513-521.
- [13] VIE J J, KASHIMA H. Knowledge tracing machines: Factorization machines for knowledge tracing[C]// Proceedings of the AAAI Conference on Artificial Intelligence. 2019:750-757.
- [14] RENDLE S. Factorization machines [C]// 2010 IEEE International Conference on Data Mining. IEEE, 2010:995-1000.
- [15] CHEN P, LU Y, ZHENG V W, et al. Prerequisite-driven deep knowledge tracing[C]// 2018 IEEE International Conference on Data Mining (ICDM). IEEE, 2018:39-48.
- [16] MINN S, YU Y, DESMARAIS M C, et al. Deep knowledge tracing and dynamic student classification for knowledge tracing [C]// 2018 IEEE International Conference on Data Mining (ICDM). IEEE, 2018:1182-1187.
- [17] GHOSH A, HEFFERNAN N, LAN A S. Context-aware attentive knowledge tracing [C]// Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2020:2330-2339.
- [18] LI X G, WEI S Q, ZHANG X, et al. LFKT: Deep knowledge

- tracing model with learning and forgetting behavior merging[J]. *Journal of Software*, 2021, 32(3): 818-830.
- [19] WANG C, LIU Z H, WANG B, et al. TCN-KT: temporal convolutional knowledge tracking model based on fusion of personal basis and forgetting[J]. *Application Research of Computers*, 2022, 23(2): 1-6.
- [20] LIU Y, YANG Y, CHEN X, et al. Improving knowledge tracing via pre-training question embeddings[J]. *arXiv*: 2012. 05031, 2020.
- [21] NAKAGAWA H, IWASAWA Y, MATSUO Y. Graph-based knowledge tracing: modeling student proficiency using graph neural network[C]//2019 IEEE/WIC/ACM International Conference on Web Intelligence(WI). IEEE, 2019: 156-163.
- [22] WANG B, SHENG Y X, JI X Y. DKTwMF: A deep knowledge tracing with multiple features[J]. *Computer Technology and Development*, 2021, 31(7): 35-41.
- [23] HUANG S W, LIU Z H, LUO L Y, et al. behavior-forgetting Bayesian knowledge tracing[J]. *Application Research of Computers*, 2021, 38(7): 1993-1997.
- [24] KEAN J, REILLY J. Item response theory[C]// *Handbook for Clinical Research: Design, Statistics and Implementation*. 2014: 195-198.
- [25] BIRNBAUM A L. Some latent trait models and their use in inferring an examinee's ability[M]. New York: *Statistical Theories of Mental Test Scores*, 1968.
- [26] KETKAR N. Convolutional neural networks [M]// *Deep Learning with Python*. Apress, Berkeley, CA, 2017: 63-78.
- [27] KIPF T N, WELLING M. Semi-supervised classification with graph convolutional networks[J]. *arXiv*: 1609. 02907, 2016.
- [28] VELIČKOVIĆ P, CUCURULL G, CASANOVA A, et al. Graph attention networks[J]. *arXiv*: 1710. 10903, 2017.
- [29] QIN C, ZHANG Y, LIU Y, et al. A visual place recognition approach using learnable feature map filtering and graph attention networks[J]. *Neurocomputing*, 2021, 457: 277-292.
- [30] XIAO Y, PEI Q, XIAO T, et al. MutualRec: Joint friend and item recommendations with mutualistic attentional graph neural networks[J]. *Journal of Network and Computer Applications*, 2021, 177: 102954.
- [31] LI Q, LIN W, LIU Z, et al. Message-aware graph attention networks for large-scale multi-robot path planning[J]. *IEEE Robotics and Automation Letters*, 2021, 6(3): 5533-5540.
- [32] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]// *Advances in Neural Information Processing Systems*. 2017: 5998-6008.
- [33] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. *Neural Computation*, 1997, 9(8): 1735-1780.
- [34] SU Y, CHENG Z, LUO P, et al. Time-and-Concept Enhanced Deep Multidimensional Item Response Theory for interpretable Knowledge Tracing[J]. *Knowledge-Based Systems*, 2021, 218: 106819.
- [35] XIONG X, ZHAO S, VAN INWEGEN E G, et al. Going deeper with deep knowledge tracing[C]// *Proceedings of the 9th International Conference on Educational Data Mining*. 2016: 545-550.



DONG Yongfeng, born in 1977, Ph.D, professor, is a member of China Computer Federation. His main research interests include intelligent information processing, big data technology, robotics and intelligent control, and software engineering.



LI Linhao, born in 1989, Ph.D, lecturer, is a member of China Computer Federation. His main research interests include quantization and hash learning, sparse signal recovery, background modeling and foreground detection.

(责任编辑:柯颖)