

DOI:10.19651/j.cnki.emt.2107720

基于注意力的 R-GCN-GRU 的在线学生绩效预测*

崔立志^{1,2} 何泽彬^{1,2} 李璇^{1,2}

(1.河南理工大学 电气工程与自动化学院 焦作 454000; 2.河南理工大学 河南省智能装备直驱技术与控制国际联合实验室 焦作 454000)

摘要:针对传统的绩效预测方法没有区别性地对待各属性特征对学生成绩的重要程度、学生在线学习的低完成率的问题,提出了一种融入注意力机制的关系图卷积神经网络和门控循环单元(AR-GCN-GRU)的学生绩效预测方法。其融入的注意力机制用于捕获学生之间的关系属性特征,同时提取学生重要属性特征并进行可视化,且该方法综合了关系图卷积神经网络(R-GCN)和门控循环单元(GRU)的优点,既能捕捉节点之间的内部关联,又能很好地抽取最具代表性的学生行为属性特征信息。在公开数据集上对模型进行了对比验证和消融实验,模型 F 值和精确率分别达到了 99.00% 和 99.73%,实验结果表明所提方法较其他算法有明显提升,验证了注意力机制的有效性。

关键词: 绩效预测;注意力机制;属性特征;R-GCN;GRU

中图分类号: TP399 **文献标识码:** A **国家标准学科分类代码:** 520.99

Online student performance prediction of R-GCN-GRU based on attention

Cui Lizhi^{1,2} He Zebin^{1,2} Li Xuan^{1,2}

(1. School of Electrical Engineering and Automation, Henan Polytechnic University, Jiaozuo 454000, China; 2. Henan International Joint Laboratory of Direct Drive and Control of Intelligent Equipment, Henan Polytechnic University, Jiaozuo 454000, China)

Abstract: Aiming at the problem that the traditional performance prediction method does not treat the importance of each attribute feature to student's scores and the low completion rate of student's online learning differently, a convolutional neural network of relational graph and gated recurrent unit integrated into the attention mechanism (AR-GCN-GRU) score prediction method for students is proposed. The integrated attention mechanism is used to capture the relationship attribute characteristics between students, and at the same time, extract the important attribute characteristics of students and visualize them, and the method integrates the advantages of the convolutional neural network of relational graph (R-GCN) and the gated recurrent unit (GRU), and can not only capture the internal correlation between nodes, but also extract the representative information of students' behavioral attributes well. The model was contrasted and ablation experiment on a public data set. The F value and accuracy of the model reached 99.00% and 99.73%, the experimental results show that the method has been significantly improved than other algorithms, and the effectiveness of the attention mechanism is verified.

Keywords: performance prediction; attention mechanism; attribute characteristics; R-GCN; GRU

0 引言

近年来,教学方式呈现多样化形式,其中大规模开放在线课程的兴起引起了传统教育界的争议。Yin 等^[1]提出了与传统的教育系统相比,在线学习中学生的辍学率和挂科率更高。针对这一问题,教育者试图开发出更有效的预测模型,模型可以预测在线课程中表现欠佳的学生,并能根据系统提供的数据进行及时的干预,以达到对学生的监督作

用并改善学生通过率、修正教学策略的目的。

目前,陈曦等^[2]提出了关于现有的学生绩效预测的研究方法主要分为两类:一类是基于机器学习的成绩预测,另一类是基于推荐方法的成绩预测。不同的研究方法中研究者所构建模型的侧重点各不相同。Hu 等^[3]提出了采用由 300 名学生组成的具有 13 种在线属性的数据,使用决策树分类器来开发预警系统以识别学生是否通过或挂科。Pandey 等^[4]提出了在影响学生成绩的 18 种属性特征中

收稿日期:2021-08-29

* 基金项目:国家自然科学基金项目(61573129)、河南理工大学教改项目(2019JG033)资助

过计算各个属性特征的信息增益率挑选出 8 种重要属性, 并利用所挑选的 8 种重要属性构建决策树模型对学生成绩进行预测。张麒麟等^[5]提出了利用由 5 种属性特征组成的学生数据, 采用过采样技术配合神经网络模型以预测学生成绩是否及格。从研究方法上来看, 现有的成绩预测模型大多采用机器学习方法, 预测模型泛化能力大多不强或历史数据无法更全面地构成学生画像, 无法准确捕获学生学习的复杂互动, 例如传统算法和神经网络无法直接对图结构的数据进行预测。

受深度学习的影响, Kipf 等^[6]提出了图卷积神经网络 (graph convolutional network, GCN), 它将卷积神经网络拓展到图结构形式中, GCN 因可以很好地融合图结构数据的结构特征和属性特征并且有较好的组合泛化能力而被广泛使用。在 GCN 的基础上, Schlichtkrull 等^[7]提出了关系图卷积神经网络 (relational-graph convolutional network, R-GCN), 它通过链接预测和实体分类等任务对非对称关系进行建模, 从而显著地提高了模型的性能。Velickovic 等^[8]提出了图注意力网络 (graph attention networks, GAT), 它的出现为给定实体的邻居节点分配不同的权重 (注意力), 通过迭代的方式以层传播注意力达到目标节点, 进一步提高了模型利用率。李梦莹等^[9]提出了基于两组课程数据, 采用双路注意力机制的学生成绩预测模型预测学生成绩, 通过数据挖掘从 30 个属性特征中挑选出 15 种对学生期末成绩存在影响的属性特征并对学生进行个性化分析。Liu 等^[10]提出了通过机器学习自动捕获最具代表性的学生特征, 然后在循环神经网络结构中加入注意力机制来预测学生成绩。Wang 等^[11]提出了一种根据学生的时间序

列数据信息来由时间控制的长短期记忆^[12] (long short-term memory, LSTM) 预测模型, 其中的时间控制单元能捕获学生学习过程的长期和短期信息, 从而提高模型预测性能。尽管上述方法都取得了较好的分类效果, 但它们没有从图结构的角度考虑学生在线课程的交互, 未充分利用节点间的信息。

根据上述分析, Cho 等^[13]提出了 GRU 作为 LSTM 的一种变体, 它通过重置门和更新门来捕捉序列间长距离依赖关系, 其计算过程相比 LSTM 更高效。为此, 本文提出了一种 AR-GCN-GRU 的学生绩效预测方法, 该方法能准确识别出表现不佳的学生并及时地对其恰当干预与指导。本文贡献如下: 1) 基于在线学习环境中的特征信息构建图结构表示, 与传统的非结构化数据相比, 属性信息更加丰富; 2) 本文所提方法能充分学习学生之间特征信息的相似性, 有效标识属性特征对学生绩效的影响程度, 同时提高了模型的预测能力; 3) 该方法基于公开的数据集 (open university learning analytics dataset, OULAD) 展开实验, 实验结果验证了本文所提模型的有效性。

1 模型整体设计

整体框架如图 1 所示。所提方法首先把学生和课程信息构建成图结构表示; 其次采用加入注意力机制的 R-GCN 来提取学生重要属性特征信息; 然后利用 GRU 输出最具代表性的学生行为属性特征信息; 最后在全连接层中对两个网络输出的关键信息进行融合并进行学生绩效分类。其中图结构在章节 1.1 中给出分析。其具体内容描述如下。

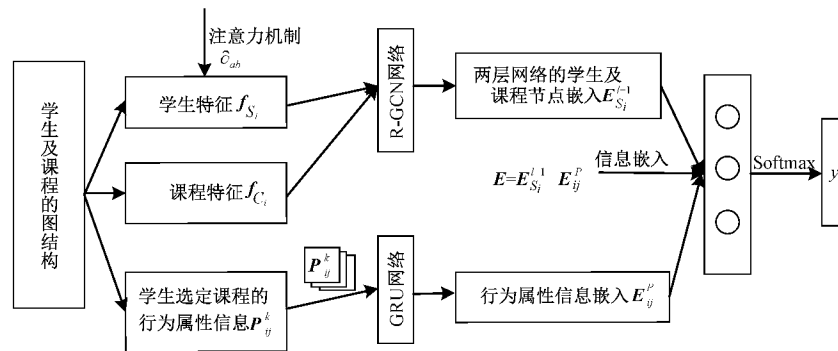


图 1 AR-GCN-GRU 模型框架

1.1 图结构分析

传统方法无法对图结构的数据信息进行预测, 因此不能基于相关数据集挖掘更深层次的信息。为了解学生之间的复杂互动, 有利于模型充分学习学生的信息, 本文将学生及课程之间的交互信息建模为图数据结构, 如图 2 所示。本文将数据集中的每个学生或课程都当成一个节点, 学生选定的课程以及有互动的学生被当作是有向图, 使用 $S_i \in V, i \in \{1, \dots, n\}$ 表示第 n 个学生节点, 使用 $C_j \in V,$

$j \in \{1, \dots, m\}$ 表示第 m 个课程节点, 图中的总结点数为 $m+n, m$ 代表课程总数, n 代表学生总数, 它们的特征分别用 f_{C_j} 和 f_{S_i} 来表示, 为便于计算, 课程特征数据与学生特征数据维度不一致的用 0 补全, $f_{S_i}, f_{C_j} \in \mathbb{R}^{n \times d}$ 中 d 代表特征数据的维度, 学生在线参与课程时产生的连续的行为数据用 $P_{ij}^k \in \mathbb{R}^q, k \in \{1, \dots, k\}$ 表示, 其中 k 代表学生行为数据的周数, q 代表每周的学生行为数据的维度。图结构用 $G(V, \epsilon)$ 来表示, V 是节点集合, ϵ 为节点-边-节点的三

元组集 (V_i, r, V_j) , 关系 r 代表学生注册指定的课程, 其中有注册关系的 $r = 1$, 反之为 $r = 0$ 。有互动关系的学生节点的邻接矩阵 $A \in \mathbb{R}^{n \times n}$, 其中学生节点有互动的设定为 1, 反之为 0。

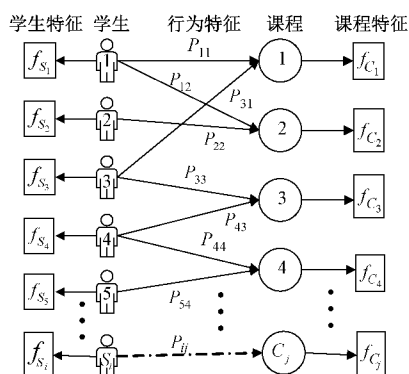


图 2 学生及课程的图结构表示

1.2 注意力机制模块

学生在线绩效受学习条件以及个人表现等多种因素的影响, 加入注意力机制的主旨是捕获不同学生之间的关系特征, 根据其重要度不同对其赋予不同的权重, 并从学生特征中捕捉出关键信息^[14-15]。本文把学生节点特征用 f_{s_a} 和 f_{s_b} 表示, 经过特征权重计算, e_{ab} 表示学生节点 b 对节点 a 的注意力权重, 即节点之间的关联程度, 此时的注意力被分配到所有的节点上却未凸显出有互动关系节点的结构信息。 θ 为注意力机制参数(权重), 维度为 $\mathbb{R}^{n \times n}$, W 为权重参数, 维度为 $\mathbb{R}^{d \times d}$ 。

$$e_{ab} = \theta(Wf_{s_a}, Wf_{s_b}) \quad (1)$$

θ_{ab} 由引入的 Softmax 函数对节点 a 的所有相邻节点进行正则化得到, 表示相邻节点 b 占节点 a 的所有相邻节点 c 的比重, 且通过邻接矩阵 A , 注意力机制关注到了有互动关系的节点, 公式如下。

$$\theta_{ab} = \text{Softmax}_b(e_{ab}) = \frac{\exp(e_{ab})}{\sum_{i=1}^{s_i} \exp(e_{ac})} \quad (2)$$

$$F_{s_i} = \theta_{ab} f_{s_i} \quad (3)$$

式(3)将注意力机制 θ_{ab} 与 f_{s_i} 进行点积, 得到的 $F_{s_i} \in \mathbb{R}^{n \times d}$ 能增强或削弱特征信息的表达, 从而自适应优化相关特征的影响力, 提高了模型提取学生重要特征因素的能力^[16]。整体过程如图 3 所示。

1.3 关系图卷积网络(R-GCN)

R-GCN 是 GCN 的一个扩展形式, 利用 R-GCN 能挖掘出课程节点与融入注意力机制的学生节点的内部关联, 从而能抽取更关键的节点信息。为了减少节点信息误差, 在对邻居节点信息融合时与学生选定课程的课程节点的平均特征信息进行拼接。考虑到训练时间问题故而采用两层隐藏层, 两层隐层神经元个数分别为 (15, 15), 即第 1 层由输入的 35 维特征降维为 15 维, 第 2 层输出为 15 维,

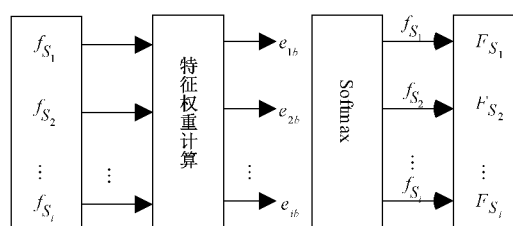


图 3 特征注意力机制

为探索节点间的关联性, 以下对其进行了重点描述。在进行第 1 层 R-GCN 网络时, 学生节点将其携带的信息传递到其相邻节点, 输出的学生节点则学习到了课程节点所附加的信息, 再经过第 2 层 R-GCN 网络, 学生节点信息及其携带的课程信息传递到下一层学生节点, 从而得到学生 S_i 与其所有同学学习的所有课程信息, 其学生节点信息聚合过程如图 4 所示。经过两层的 R-GCN 后, 强化了融入注意力的学生节点与课程节点间的相关性, 得到的输出 $E_{S_i}^{t+1}$ 自适应融合了丰富的节点信息, 突出了节点的关键影响因素, 公式如下。

$$E_{S_i}^{t+1} = \sigma(W_0^t F_{S_i}^t + \sum_{r=1} \frac{1}{|N_S^r(S_i)|} \sum_{C_j \in N_S^r(S_i)} W_r^t f_{C_j}^t) \quad (4)$$

其中, 激活函数 σ 采用 ReLU 函数进行激活, $F_{S_i}^t$ 与 $f_{C_j}^t$ 为输入的特征, W_0^t, W_r^t 分别为自身特征转换参数与关系转换参数, 第 1 隐层维度为 $\mathbb{R}^{d \times d'}$, 第 2 隐层维度为 $\mathbb{R}^{d' \times d'}$, d' 为变换后的特征维度, $|N_S^r(S_i)|$ 为代表关系为 r 的邻居节点个数的正则化常量, $E_{S_i}^{t-1}$ 为输出的具有代表性的特征, 其维度为 $\mathbb{R}^{n \times 2d'}$ 。

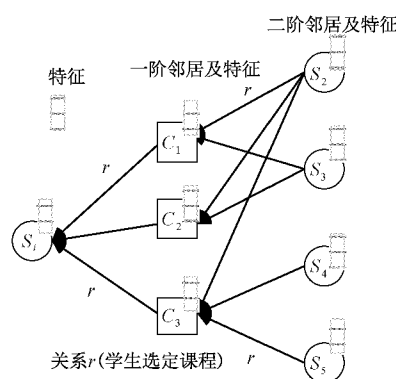


图 4 学生及课程节点信息聚合过程

1.4 GRU 网络

与 LSTM 不同, GRU 只包含更新门和重置门, 因此训练参数更少。本文 GRU 网络包括输入向量、1 层隐藏层、输出层, 隐藏层神经元个数为 10, 即由输入的 20 维特征降维为 10 维。GRU 的主要作用是对学生行为特征长期依赖关系进行识别和提取, 进一步提取学生的代表性特征。网络输入为学生 S_i 在第 k 周(共 40 周)线上学习课程 C_j 后产生的行为特征序列 $P_{ij} = [P_{ij}^1, P_{ij}^2, P_{ij}^3, \dots, P_{ij}^k]$, 输入特

征结合前一时刻 GRU 网络的隐藏状态通过赋予权重计算经激活函数后得到修正后的输入特征信息,再经过一层 GRU 网络学习特征,通过对历史时刻输出信息结合当前时刻网络输出赋予权重后得到改进后的当前时刻隐层状态输出 h_t , 以提取关键时刻的影响因素,最后输出层得到的是具有代表性的学生行为特征。更新门 Z_t 帮助 GRU 单元决定上一隐藏层的状态和当前时刻状态信息的更新程度,重置门 R_t 决定了过去信息的遗忘程度^[17]。本文设计的 GRU 网络如图 5 所示。

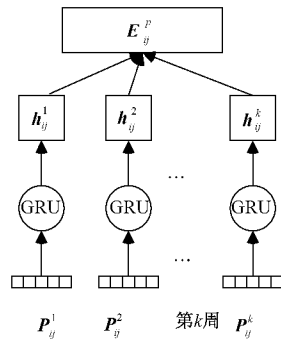


图 5 GRU 网络

在 t 时刻,GRU 的状态通过下列公式计算:

$$Z_t = \sigma_1(W_z x_t + U_z h_{t-1}) \quad (5)$$

$$R_t = \sigma_1(W_r x_t + U_r h_{t-1}) \quad (6)$$

$$\tilde{h}_t = \tanh(W_h x_t + U_h (R_t \odot h_{t-1})) \quad (7)$$

$$h_t = Z_t \odot \tilde{h}_t + (1 - Z_t) \odot h_{t-1} \quad (8)$$

其中,式(5)~(8)分别代表更新门、重置门、候选隐藏层、隐藏层, t 时刻代表第 k 周,输入的 x_t 维度为 $\mathbb{R}^{n \times 10 \times q}$, $W_z, W_r, W_h, U_z, U_r, U_h$ 表示 GRU 网络的权重参数, σ_1 为 sigmoid 函数, \odot 为 Hadamard 乘积, h_{t-1} 表示前一时刻的状态信息, h_t 为每周输出的 h_{ij}^k , E_{ij}^p 为去维后输出的具有代表性的行为特征信息,维度为 $\mathbb{R}^{n \times q'}$, q' 为变换后的特征维度。

1.5 特征信息融合

本部分将来自 R-GCN 网络的融入注意力机制的关联

节点信息以及 GRU 网络输出的代表性的特征信息采用 PyTorch 中的 cat 函数进行拼接融合,维度为 $\mathbb{R}^{n \times (2d'-q')}$ 。如图 1 所示,融合后的 E 吸收了两个网络的优点后具有丰富的关键特征信息,从而帮助模型做出更加准确的预测,公式如下。

$$E = E_{S_i}^{t-1} \parallel E_y^p \quad (9)$$

然后把得到的输出 E 经过线性变换后使用交叉熵损失函数来优化损失值,最终得到学生通过课程的预测精度。

1.6 优化目标函数

本文使用反向传播算法来训练网络模型,用交叉熵作为分类损失,通过迭代求解损失值和随机梯度下降来优化模型,使得损失函数的值收敛到最小。考虑到模型的复杂性,避免模型在训练的过程中出现过拟合,本文引用了 L_2 正则项对参数进行约束,故模型的最终损失函数为:

$$L = -\frac{1}{G} \sum_{i=1}^G y_i \log p_i + \lambda \| \theta \|^2 \quad (10)$$

其中, G 为训练集数据量; y_i 为第 i 个学生样本的标签; p_i 为第 i 个学生样本的预测概率; $\lambda \| \theta \|^2$ 为 L_2 正则项; θ 为模型的所有参数集合。

2 实验结果与分析

2.1 数据集

本文使用 Open University 学习分析数据集 (OULAD)^[18],该数据集包含 2013 年和 2014 年的 22 门课程、32 593 名学生的信息以及学生的评估结果和 VLE(日常交互信息),每年两个学期 40 周的学生交互数据分别在 B(2 月份)和 J(10 月份)被获取,每个学生的课程结果分为通过和挂科两种类别,本文选择了其中一个社会科学课程对其展开实验,训练数据采用 2 013 J,测试数据采用 2 014 J,学生总人数为 735 人。针对 35 种属性特征包含信息的差异性,本文对其进行进一步类别划分,属性类型共分为基本信息、学生行为、学习条件和家庭情况 4 类,属性类型对应的特征属性划分,具体如表 1 所示。

表 1 各属性特征分类

属性类型	特征属性
基本信息	Gender, Nationality, Semester, Health
学生行为	Dualpane, external quiz, quiz, folder, forum, glossary, homepage, html activity, collaborate, content, elluminate, page, questionnaire quiz, repeatactivity, sharedsubpage, subpage, activities
学习条件	Dataplus, wiki, resource, url, paid, Nursery, address
家庭情况	Relation, ParentAnsweringSurvey, ParentschoolSatisfaction, famsup, Medu, Fedu, Guardian

2.2 实验环境及参数设置

本文运算环境配置为:Inter Core i7-8700 CPU@3.20 GHz、16 GB 运行内存、酷睿 Inter HD Graphics 630,在实验中利用 Adam 对模型进行优化,训练迭代次数为 300 次,学习率

为 0.03,权值衰减为 0.001,损失函数使用的 L_2 正则化为 0.001。

2.3 实验评估指标

本文中采用精确率(Precision)、召回率(Recall)以及 F

值(F1-Score)进行模型分类预测性能度量。精确率表示正确分类的正例个数占预测为正例总数的比例。召回率表示正确分类的正例个数占预测为正数的比例,F值越高,分类效果越好。

$$\text{Precision} = \frac{TP}{TP + FP} \quad (11)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (12)$$

$$F = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (13)$$

其中, TP 代表了学生标签为通过被正确判定为正样本, FP 代表了学生标签为挂科被错误的判定为正样本, FN 代表学生标签为通过但被判定为负样本。

2.4 实验结果分析

1) 对比实验

为了演示所提出模型的整体性能,将本文所提出的方法同支持向量机(support vector machine, SVM)^[19]、LSTM、GRU、R-GCN 4种算法在部分数据集(OULAD)中进行对比实验。实验结果如表2所示,每列数值最佳结果用粗体表示。

表2 对比实验预测结果

分类算法	精确率	召回率	F值
SVM	0.655 8	0.656 1	0.691 5
LSTM	0.906 9	0.780 0	0.838 7
GRU	0.975 0	0.780 0	0.866 7
R-GCN	0.975 3	0.790 0	0.872 9
AR-GCN-GRU	0.997 3	0.990 0	0.990 0

从表2的实验结果可以看出,本文所提方法相比其他4种算法,在精确率(Precision)、召回率(Recall)以及F值(F1-score)这3个指标上均取得最优值。对比实验的4种算法预测准确率普遍不高的原因可能为SVM、LSTM、GRU算法不能直接对图结构的信息进行训练。SVM算法对所有的特征信息进行随机分类训练,因此在训练时没有区别性地对待各属性特征对学生绩效的重要程度。而

LSTM和GRU两种算法只捕获了学生的序列信息,后者由于参数较前者少因此训练结果更优。R-GCN算法在图结构的基础上学习学生与课程之间的属性特征,但并未考虑学生每周的行为信息。而本文引入注意力机制,注意的对象为有交互关系的学生,通过模型训练加强了对相似学生的重要属性特征的区别性对待,且模型吸收了R-GCN和GRU的优点进而大大提高了模型的预测分类效果。

2) 注意力机制的消融研究

本文进行的注意力机制的消融实验对比如表3所示。其中R-GCN-LSTM代表的是没有加注意力机制的DOPE^[20]模型,该模型在训练时对学生属性信息平等对待,而加入了注意力的模型深度关注到属性特征对学生绩效影响程度的差异性,训练复杂度增加因而时间比之较高,相较于R-GCN-LSTM模型的F值,本文算法提升了4.16%,AR-GCN-LSTM算法比之高3.64%,因GRU参数少因此本文算法较AR-GCN-LSTM算法时间更优。试验结果说明注意力机制能有效关注到不同学生的相似特征信息,并分配不同的权重信息,来强化相关特征的影响力并抑制不相关属性特征。

表3 消融实验对比

分类算法	精确率	召回率	F值	训练时间/s
R-GCN-LSTM	0.978 5	0.910 0	0.948 4	325
AR-GCN-LSTM	0.989 5	0.970 0	0.984 8	485
AR-GCN-GRU	0.997 3	0.990 0	0.990 0	457

为了确定模型的迭代次数,将本文方法与其他两种方法的F值和训练损失值作对比,如图6所示,训练开始3条损失曲线下降幅度较大,说明学习率合适,在学习到第10个epoch时,图中曲线渐趋平稳,3种方法都能较快的收敛,R-GCN-LSTM模型因未加入注意力模块,在模型收敛后F值低于其他两种模型,而本文模型较AR-GCN-LSTM模型稍好,原因为本文所采用的数据集小且GRU比LSTM模型更简单,因此能更充分的提取学生的代表性特征,从而提高了模型的准确率,经过15个epoch(300步)训练后,本文模型取得最优的收敛结果。

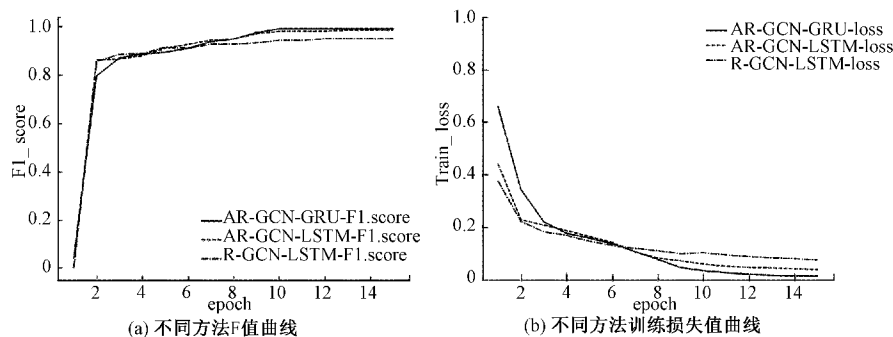


图6 不同方法的F值和训练loss曲线

3) 注意力权重分析

为了挖掘出影响不同类别学生的相对重要特征因素,本文对不同类别学生的重要特征进行了注意力结果可视化,更直观地显示出每个因素对学生分类结果的不同影

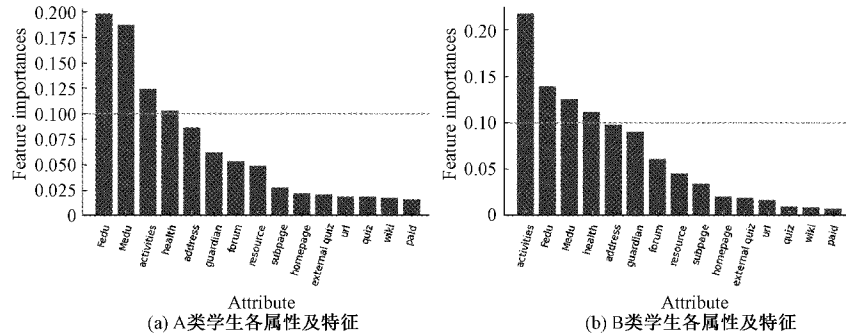


图 7 学生分类各属性及特征

如图 7 所示,课程通过的 A 类学生和挂科的 B 类学生的关键因素是父亲、母亲受教育程度(Fedu、Medu)、活动交流(activities)以及健康(Health),其比重均超过了 50%,而 B 类学生的活动交流显然更多,其次监护人(Guardian)、参与论坛讨论(forum)和学习资源(resource)也对学生有一定的影响。上述说明学生的成绩很大程度上受到来自家庭环境的影响,同时个人的努力也占据较高的比重。通过特征提取出对学生绩效影响的 15 种相对重要特征因素,其属性分类和特征属性如表 4 所示。

表 4 重要特征属性分类

属性类型	特征属性
基本信息	Health
学生行为	external quiz, forum, homepage, questionnaire quiz, subpage, activities
学习条件	wiki, resource, url, paid, address
家庭情况	Medu, Fedu, Guardian

3 结 论

本文提出了一种基于注意力的 R-GCN-GRU 在线绩效预测方法,该方法在特征提取的过程中加入注意力机制,使得模型关注到学生重要属性特征,再利用 R-GCN 网络捕获节点间的关联程度、GRU 网络对学生每学期行为信息的长期记忆能力,将 R-GCN 与 GRU 两种网络优点结合后使得模型的预测能力更强,有效提高了对学生的分类准确性。AR-GCN-GRU 在(OULAD)公开数据集上进行实验,实验结果表明,本文所提模型在二分类任务中取得了最好的预测结果,从而有效地监督在线课程中表现欠佳的学生,并给予其恰当地干预与指导。本文未来工作主要关注利用学生和课程之间的关系用细化的知识图谱来表述,更详细地知识图谱将会展示学生更全面地语义信息,

通过对不同类别的学生进行分析来进一步反映学生个体差异的情况。可视化结果如图 7 所示。横坐标代表重要属性信息,纵坐标表示各属性特征对应的注意力权重,用图中虚线 0.1 代表关键影响因素分界线。

如何在保证分类准确率的同时模型能提取多种节点关系信息是下一步工作中需要解决的问题。

参 考 文 献

- [1] YIN S J, LEI L Q, WANG H Z, et al. Power of attention in MOOC dropout prediction [J]. IEEE Access, 2020, 8: 202993-203002.
- [2] 陈曦,梅广,张金金,等. 融合知识图谱和协同过滤的学生成绩预测方法[J]. 计算机应用, 2020, 40(2): 595-601.
- [3] HU Y H, CHIALUN L, SHENGPao S. Developing early warning systems to predict student's online learning performance [J]. Computers in Human Behavior, 2014, 36: 469-478.
- [4] PANDEY M, SHARMA V K. A decision tree algorithm pertaining to the student performance analysis and prediction [J]. International Journal of Computer Applications, 2013, 61(13): 1-5.
- [5] 张麒增,戴翰波. 基于数据预处理技术的学生成绩预测模型研究[J]. 湖北大学学报(自然科学版), 2019, 41(1): 101-108.
- [6] KIPF T N, WELLING M. Semi-supervised classification with graph convolutional networks [J]. ArXiv, 2016, ArXiv: 1609.02907, 2016.
- [7] SCHLICHTKRULL M, KIPF T N, BLOEM P, et al. Modeling relational data with graph convolutional networks [C]. In European Semantic Web Conference, Heraklion, Jun 3-7, 2018, Berlin: Springer, 2018: 593-607.
- [8] VELICKOVIC P, CUCURULL G, CASANOVA A, et al. Graph attention networks [J]. ArXiv, 2017, ArXiv: 1710.10903, 2017.
- [9] 李梦莹,王晓东,阮书岚,等. 基于双路注意力机制的学生成绩预测模型[J]. 计算机研究与发展, 2020,

- 57(8): 1729-1740.
- [10] LIU D, ZHANG Y P, ZHANG J, et al. Multiple features fusion attention mechanism enhanced deep knowledge tracing for student performance prediction[J]. IEEE, 2020, 8: 194894-194903.
- [11] WANG L T, WANG H. Learning behavior analysis and dropout rate prediction based on MOOCs data[C]. 2019 10th International Conference on Information Technology in Medicine and Education (ITME), Qingdao, Aug 23-25, 2019. IEEE, 2019: 419-423.
- [12] CHUNG J, GULCEHRE C, CHO K H, et al. Empirical evaluation of gated recurrent neural networks on sequence modeling [J]. Computer Science, 2014, ArXiv:1412.3555.
- [13] CHO K, VAN MERRIENBOER B, GULCEHRE C, et al. Learning phrase representations using RNN Encoder-Decoder for statistical machine translation [J]. ArXiv, 2014, ArXiv: 1406.1078, 2014.
- [14] 朱书勤. 基于注意力融合网络的 RGB-D 目标检测算法[J]. 电子测量技术, 2021, 44(9): 110-115.
- [15] 白欢欢. 一种融合多注意力机制的服饰分割方法研究[J]. 国外电子测量技术, 2021, 40(7): 52-58.
- [16] 杨梅, 贾旭, 殷浩东, 等. 基于联合注意力孪生网络目标跟踪算法[J]. 仪器仪表学报, 2021, 42(1): 127-136.
- [17] 王震宇, 张雷. 基于深度卷积和门控循环神经网络的传感器运动识别[J]. 电子测量与仪器学报, 2020, 34(1): 1-9.
- [18] KUZILEK J, HLOSTA M, ZDRAHAL Z. Open university learning analytics dataset [J]. Scientific Data, 2017, 4:170171.
- [19] BURMAN I, SOM S. Predicting students academic performance using support vector machine[C]. Proc of the 2019 Conf on Amity Int Conf on Artificial Intelligence (AICAI), Dubai, Feb 4-6, 2019, Piscataway, NJ: IEEE, 2019: 756-759.
- [20] KARIMI H, DERR T, HUANG J, et al. Online academic course performance prediction using relational graph convolutional neural network[C]. In Proceedings of the 13th International Conference on Educational Data Mining (EDM 2020), Fully virtual conference, Jul 10-13, 2020, Ifrain: EDM, 2020: 444-450.

作者简介

崔立志, 副教授, 博士, 主要研究方向为人工智能信号处理、控制工程。

E-mail: clzh0308@126.com

何泽彬(通信作者), 硕士, 主要研究方向为信息处理与网络控制。

E-mail: shuju973815980@163.com

李璇, 硕士, 主要研究方向为信息处理与网络控制。

E-mail: lixuan592021@163.com