

基于矩阵分解的多模态军事职业教育在线学习成果预测

邵东春,李康,杜英,赵涵

(陆军装甲兵学院蚌埠校区,安徽 蚌埠 233000)

摘要:军事职业教育的在线学习成果预测因其教育意义而受到广泛关注.在线学习成果预测方法中,基于推荐的方法凭借其突出的个性化优势占有重要地位.虽然现有研究通过利用辅助信息来解决推荐系统中冷启动和合并侧面信息的问题,但基于多模态推荐在线学习成果预测方法的潜力尚未研究.首次使用多模态辅助学习者在线学习成果预测,使用大语言模型作为信息混合器,为矩阵分解产生额外的引导信号.具体来说,所提的语言引导矩阵分解模型,使用语言化的多模态信息,为教育交互产生丰富的语义嵌入,并将它们作为矩阵分解的辅助信号.提供了两种文本嵌入方法,嵌入初始化方法计算数据的特定先验以初始化分解矩阵,嵌入蒸馏方法用于对齐矩阵分解的潜在特征和嵌入特征,使语言模型更充分地指导矩阵分解.在一个军事职业教育平台收集的数据集上评估了所提出的模型,包括 2.5 万个学习者的 35 万个互动和 5 万个不同的成果表现.大量实验表明,当前模型在各种军事职业教育在线学习成果预测场景,包括冷启动中优于已有方法.

关键词:军事职业教育;在线学习成果预测;多模态;冷启动;矩阵分解

中图分类号:TP315.69

文献标志码:A

文章编号:1000-2367(2025)06-0058-08

目前在线和自主学习是军队人员参加军事职业教育学习的主要形式.在线学习成果预测对于解决军事职业教育学习中的学习者积极性不高、学习管理难度大等问题具有极大潜力.在线学习成果预测的主要任务是预测学习者成绩、活动评价或课程成效^[1-2].在军事职业教育中,学习成果预测可为学习者提供高质量个性化指导、进行成长心态干预,可根据学习者的学术成就早期识别高危学习者,因而长期受教育工作者的关注.目前,已有工作开发出基于机器学习(machine learning, ML)的可扩展自动化在线学习成果预测系统^[3-5].

学习成果预测模型的一部分工作采用传统的 ML 算法.文献[6]使用了 3 种具有代表性的 ML 算法(决策树、最近邻算法和神经网络)对学习者的学习成果进行了研究,展示了 ML 在在线学习成果预测任务中的潜力.同样, HUY-LE 等^[7]给出了产生最佳学习成果预测结果的模型及参数选择. AL-SHEHRI 等^[8]在一个大学教育数据集上使用支持向量机和最近邻算法实现了在线学习成果预测,得到较好的预测性能. LUO 等^[9]使用聚类算法实现混合课程中学生成果预测.此外,随着深度学习(deep learning, DL)的发展,基于 DL 的在线学习成果预测方法也随之展开^[10-12].在线学习成果预测模型的另一部分工作将学习成果预测问题转化为推荐系统(recommendation systems, recSys)^[13-14]问题,以追求高个性化的学习成果预测. SWEENEY 等^[15-16]探索了基于推荐的代表性算法,矩阵分解(matrix factorization, MF)和因子分解机,与传统 ML 模型相比性能更优. ZHONG 等^[17]进一步加强了基于 MF 的学习成果预测方法,通过引入课程平均得分,控制课程

收稿日期:2024-08-12; **修回日期:**2024-09-02.

基金项目:军队科研资助重点项目;2024 年度全国教育科学国防军事教育学科规划军队重点课题(JYKY-B2024005).

作者简介:邵东春(1978—),男,安徽蚌埠人,陆军装甲兵学院教授,主要研究方向为军事教育理论研究,军事职业教育与深度学习.

通信作者:赵涵(1993—),男,山东淄博人,主要研究方向为军事职业教育学习数据分析,知识追踪, E-mail: hanz@aaaf.nt-esmail.com.

引用本文:邵东春,李康,杜英,等.基于矩阵分解的多模态军事职业教育在线学习成果预测[J].河南师范大学学报(自然科学版),2025,53(6):58-65.(Shao Dongchun, Li Kang, Du Ying, et al. Multi-modal military vocational education online learning outcome prediction based on matrix factorization[J]. Journal of Henan Normal University(Natural Science Edition), 2025, 53(6): 58-65. DOI:10.16366/j.cnki.1000-2367.2024.08.12.0002.)

选择率的影响.本文关注基于 MF 的学习成果预测方法,以更适用于个性化指导.尽管目前基于 MF 的学习成果预测方法表现出显著性能^[18-19],但其在推荐系统中遭遇了众所周知的冷启动问题^[20-21]——模型由于稀疏观察采样不能充分代表新的或不常见的学习者.PAL 等^[22]通过分析学习者提交的问题获取重要信息来缓解推荐系统中的冷启动问题.亦有工作曾将辅助信息整合到学习的 MF 嵌入中来处理这一问题^[23].但这些方法只涵盖同构的辅助信息,尚未有工作探索使用异构(多模态)辅助信息的联合框架.

基于此,本文提出一个基于 MF 的多模态辅助信息在线学习成果预测系统.提出一个语言模型引导的在线学习成果预测矩阵分解方法,该方法通过语言模型(language model, LM)将多模态辅助信息作为上下文嵌入向量,并将其输入 MF 算法用作学习指导.该方法分为 3 个步骤:语言化、提示和知识迁移,如附录图 S1 所示.首先将混合的表格(学习者和成果的元数据)和文本(成果的详细描述和概念)数据转换为带有预定义提示的简单句子(语言化).接下来,这些句子被输入一个预先训练好的 LM,以产生信息丰富的句子嵌入向量(提示).随后,嵌入向量作为静态教师动态学习 MF(知识迁移).同时,可以进一步分析成果和学习者在 LM 的嵌入空间上的关系(如图 S1 右侧所示).最后,通过在一个军事职业(JZ)教育平台收集的数据集上进行实验,结果表明,本文所提出的方法取得了优于已有方法的性能,特别是在冷启动的情况下.

1 方法

1.1 基于矩阵分解的在线学习成果预测

1.1.1 问题描述

当学习者进行评估时,如果回答是有效的,则观察到结果的值为 1,否则为零.本文的目标是预测学习者对未知评估的预期结果.这一过程可以通过推荐系统类比学习者-项目交互实现.现有的一些方法采用传统的 ML 算法,将模型拟合到给定的学习者特征、评估,以及相应的结果.虽然这些方法适用于预测新同学的结果,但本文关注于对每个学习者的个性化预测,这对现实世界的教育应用更为重要.为此,本文将个性化的在线学习成果预测作为一个推荐问题,并采用最流行的协同过滤方法矩阵分解(MF).

带有偏置项的 MF 的基本过程如下:

$$L_{MF}(\mathbf{S}, \mathbf{A}, b^*) = \sum_{(i,j) \in \text{obs}} (O_{i,j} - \langle \mathbf{S}_i, \mathbf{A}_j \rangle - (b_i^S + b_j^A + b^G))^2 + \beta (\|\mathbf{S}\|_F^2 + \|\mathbf{A}\|_F^2 + \|b^S\|^2 + \|b^A\|^2), \quad (1)$$

通过最小化式(1),MF 估计了 n_s 个学习者 n_A 个成果的两个潜在嵌入矩阵 $\mathbf{S} \in \mathbf{R}^{n_s \times D_{MF}}$ 和 $\mathbf{A} \in \mathbf{R}^{n_A \times D_{MF}}$.这一结果与 $i=1, \dots, n_s$ 和 $j=1, \dots, n_A$ 的全部观察结果 $\{O_{i,j}\}$ 相匹配.其中 $b^G \in \mathbf{R}^{n_s}$ 是全局偏差, $b^S = \{b_i^S\}_{i=1}^{n_s} \in \mathbf{R}^{n_s}$ 和 $b^A = \{b_j^A\}_{j=1}^{n_A} \in \mathbf{R}^{n_A}$ 表示学习者和成果的潜在偏差, β 控制了正则化强度.推理时,可以通过学习嵌入与附加采样的点积来估计未观察到的学习者-成果的结果 (i, j) , 例如 $\hat{O}_{i,j} = \mathbf{S}_i \cdot \mathbf{A}_j + b_i^S + b_j^A + b^G$. 对在线学习成果预测使用偏差项的有效性之前已经验证过^[24], 本文额外将 LM 嵌入到 MF 框架中, 以利用多模态辅助信息.第 1.2 节详细介绍该方法.

1.1.2 数据集描述

本文从一个 JZ 在线教育平台收集了数据集,目的是开发军事职业教育领域的机器学习算法,用于诊断学习者的脆弱性,并为其提供个性化指导.附录表 S1(a)和(b)展示了每个学习者和成果的数据示例.对于学习者,提供了身份类别、每次考试的理解程度和平均分数等属性.为学习成果提供了难度、辨别力、猜测和知识结构等标签属性.表 S1(c)展示了学习者-成果交互的(正确或不正确的二进制标签)数据例子.作为附加的评估信息,为每个知识结构标签配对的主要概念和详细描述以自然语言的形式提供(见表 S1(d)).本文合并了表 S1(b)和(d)上的标签数据,并将它们作为每个评估的多模态(表格和文本)元数据.同时数据集中使用“测量时间”作为评估特征.本文 1.2 节中阐述的多模态信息融合技术利用了这个辅助特性.

1.2 语言模型引导的在线学习成果预测(LMF)

在使用 MF 建模学习者-成果交互(表 S1(c))时,首先利用多模态辅助信息与自然语言(表 S1(d))和属性形式(表 S1(a),(b)),将每个实例的元数据集特征转换成自然语言句子格式(语言化),然后只需获得 LM

中的句子嵌入(提示),最后将句子嵌入用作指导 MF(知识迁移).

1.2.1 通过提示学习进行多模态信息融合

语言化和提示描述为:(1)通过合并标签列上的成果元数据和数学知识结构数据,构建了针对个体评估的多模态元数据.(2)语言化:学习者元数据(表格)和评估元数据(表格、文本)基于预定义的模板转换为纯文本形式.由于期望 LM 比数值更好地理解纯文本,所以本文将连续属性,如平均分数和难度,离散到 5 个类别,最低、低、中、高和上.因此学习者元数据的文本模板为:“该用户属于{身份类别},她的成绩是{平均分},她把握问题的能力是{理解能力}”,评估元数据的文本模板为:“这个问题是难度{难度}、辨别力{辨别力}、评估正确率{评估正确率},内容是{主要概念},在{章}单元{节}节”.(3)提示:将语言化的元数据实例输入到一个预训练的 LM 中,产生嵌入向量(序列的最后一个标记嵌入).通过这种方式,可以将每个成果的多模态辅助信息融合到一个密集的嵌入向量中.在此过程中,本文可以采用任何可提取嵌入的开源大语言模型.

1.2.2 语言模型引导的矩阵分解

现在,将 LM 嵌入到基于 MF 的在线学习成果预测中,例如知识迁移.为此,本文考虑了两种方法:(1)嵌入初始化(IN),使用 LM 嵌入作为 MF 的潜在嵌入矩阵初始值,以给出学习者和成果的特定先验.LMF-IN 的过程与标准 MF 相同,如式(1)所示,唯一的区别是使用信息初始化,而不是随机初始化.(2)嵌入蒸馏(DI),通过一个正则化项,每个学习者和成果的 MF 潜在嵌入被迫与相应的 LM 嵌入相似.LMF-DI 的过程如下:

$$L_{\text{LM}_{\text{g}}\text{MG-DI}} = L_{\text{mf}}(\mathbf{S}, \mathbf{A}, b^*) + \gamma L_{\text{distill}}(b^S, b^A, \mathbf{u}, \mathbf{v}),$$
$$L_{\text{distill}}(b^S, b^A, \mathbf{u}, \mathbf{v}) = \sum_{i=1}^{n^S} (b_i^S - \mathbf{u} \cdot \mathbf{Z}_i^S) + \sum_{j=1}^{n^A} (b_j^A - \mathbf{v} \cdot \mathbf{Z}_j^A), \tag{2}$$

给定学习者和成果的 LM 嵌入 $\mathbf{Z}_i^S \in \mathbf{R}^D$ 和 $\mathbf{Z}_j^A \in \mathbf{R}^{D_{\text{LM}}}$,其中 D_{LM} 表示 LM 的嵌入空间维数,LMF-DI 首先使用可学习投影向量 $\mathbf{u} \in \mathbf{R}^{D_{\text{LM}}}$ 和 $\mathbf{v} \in \mathbf{R}^{D_{\text{LM}}}$ 投影 LM 嵌入到标量值.然后,鼓励每个学习者的偏差项和成果通过 γ 倍大小的 $L_{\text{distill}}(b^S, b^A, \mathbf{u}, \mathbf{v})$ 更接近预测的 LM 嵌入值.本文实验中,LMF-DI 显示出了优于 LMF-IN 的结果,所以使用 LMF-DI 作为下一节的模型.

值得注意的是,本文方法不需要真正的 LM 参数的可访问性,这可能会损害模型的所有权^[25],本文假设语言模型可以提供给定输入句子的嵌入向量.同时,本文既没有完全或部分地训练 LM,只是使用它并利用它的输出.因此,与没有 LM 引导的标准 MF 相比,所需的计算成本可以忽略不计,但带来了显著的性能改进.表 1 进一步提供了跨基线 MF 和本文方法的计算复杂度分析,表 1 中 Wall-clock 训练时间是在单个 NVIDIA A100 GPU 上使用批大小(N)=1 024 和在 100 个 epochs 上运行的平均值.对于给定的学习者与成果之间的条交互数据,MF 在学习者和评估的 $N \times D_{\text{MF}}$ 嵌入矩阵之间进行矩阵乘法,得到 $O(N^2 \cdot D_{\text{MF}})$,其中 D_{MF} 是潜在嵌入维数.LMF-IN 要求 MF 的潜在维数为 D_{LM} ,最终得到 $O(N^2 \cdot D_{\text{LM}})$ 训练时间和推理时间的复杂度,其中 D_{LM} 表示所采用的语言模型的嵌入维数.最后,本文的 LMF-DI 产生了 $N \times D_{\text{LM}}$ 的额外计算,通过 D_{LM} 维投影向量和嵌入矩阵 $N \times D_{\text{LM}}$ 之间的向量-矩阵乘法进行.就每种方法的实际训练时间而言,LMFs 仅比 MF 增加了 13%到 31%.最重要的是,推理时,LMF-DI 采用的计算量与 MF 的计算量相同,证实了本文方法的实用性.

表 1 MF 与本文模型的计算复杂度比较

Tab. 1 Comparison of the computational complexity of the MF and the model presented in this paper

方法	训练时间	推理时间	Wall-clock 训练时间/min
MF	$O(N^2 \cdot D_{\text{MF}})$	$O(N^2 \cdot D_{\text{MF}})$	93
LMF-IN	$O(N^2 \cdot D_{\text{LM}})$	$O(N^2 \cdot D_{\text{LM}})$	106
LMF-DI($\mathbf{S}$)or($\mathbf{A}$)	$O(N^2 \cdot D_{\text{MF}} + N \cdot D_{\text{LM}})$	$O(N^2 \cdot D_{\text{MF}})$	115
LMF-DI($\mathbf{S}, \mathbf{A}$)	$O(N^2 \cdot D_{\text{MF}} + N \cdot D_{\text{LM}})$	$O(N^2 \cdot D_{\text{MF}})$	122

2 实验

2.1 设置

为进行在线学习成果预测任务,实验中采用了 1.1 节讨论的用于开发军事职业教育领域机器学习算法的数据集,包括 2.5 万个学习者的 35 万个互动和 5 万个不同的成果评估.本文使用 1.5 万名学习者的数据,为稳定建模,只保留了参加 5 次以上考试的学习者,因此训练中保留了 1 万名学习者和 16 万个交互活动.与已有工作类似,本文为每个学习者保留了一个活动的评估结果作为测试集,并使用交叉验证法评估所有模型,实验中按照 7 : 3 随机划分训练集和测试集.下文表格中所有结果均取 3 次训练的平均值,* 代表使用本文数据集复现的结果.

对于指导矩阵分解的语言模型,本文采用 7 B 大小的 Linly-Chinese-LLaMA 模型,给定语言化的学习者和成果元数据实例,将最后一层的标记嵌入作为学习指导信号.构建了一个(9 454,2 048)形状的成果 LM 嵌入矩阵和一个(9 834,2 048)维数的学习者 LM 嵌入矩阵作为一个查找表.

和已有工作一样,将 4 个基于经验均值的预测器作为简单的基线,假设交互的二元结果为伯努利分布,并将其参数设置为结果的全局均值(GM)、每个学习者的平均结果(SM)和成果(AM)的平均值,以及全局、学习者、成果均值(MoM).接下来,作为实验的主要目的,比较了标准矩阵分解(MF)和 LM 引导的变体的性能.根据是否使用关于学习者(S)和成果(A)的辅助信息,在括号中将它们表示为(S)、(A)和(S,A).另外,实验与最新的工作^[7,9]做了对比.实验中使用 Adam 优化器,epochs 为 100,以 10 作为{10、20、30、100}中验证精度的潜在维度.所有方法,对每个 epoch 进行评估,并报告 100 个 epochs 内的最高性能.

2.2 在线学习成果预测的主要结果

附录表 S2 提供了 13 个模型对数据集不同身份类别的在线学习成果预测结果.IN(*)和 DI(*)表示 1.2 节中的 MF 方法.从表 S2 中可以看出,虽然简单的基线模型通常显示出优于随机模型的结果(超过 0.5 Acc),但基于 MF 的方法与简单基线之间存在较大的性能差距.相比 MF 和 LMFs,LMF-DI 在 4 个指标上均提高了性能(所有身份类别 Acc.均增加 1%~2%).验证了 LM 指导多模态辅助信息的有效性.对于基于初始化(IN)和基于蒸馏(DI)的方法,DI 明显优于 IN 对应的方法,因此在接下来的实验中采用 DI 方法作为默认方法.另外,与最新的工作^[7,9]相比,本文均达到最好结果.

2.3 流行度鲁棒性分析

本节从流行度鲁棒性的角度来研究 MF 的有效性,这在冷启动问题中至关重要.为此,首先将流行度定义为每个学习者和成果在训练数据集中的出现频率,然后将 1 万个测试样本分成 10 个流行度分位数,并测量每个分位数的预测性能.图 1 提供了 LMF(S,A)与 MF 的平均联合流行度学习曲线对超参数的敏感性分析,这也进一步验证了 LMF 的优势.表 2~4 分别展示了学习者、学习成果和联合流行度性能,结果为所有身份类别的平均值.图 2~4 提供了每个身份类别的结果.

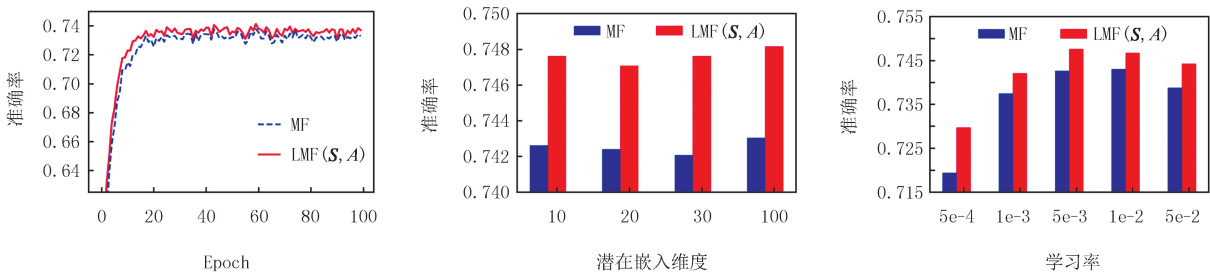


图1 训练时的学习曲线和指标分析(学习率和潜在嵌入维度)

Fig.1 Learning curve and index analysis with training(learning rate and latent embedding dimension)

总的来说,由于冷启动问题,这些方法预测正确的学习成果有些困难.例如,表 3 中,学习者在 0.1 分位数上的成果比 1.0 分位数差约 2%,同样 MF(A)的成果也下降了约 1%.此外,如果分别观察每个身份类别的结果,如图 2~4 所示,在某些情况下,性能下降更为严重(约 6%~7%).

表 2 学习者在数据集上特定流行度的准确性

Tab. 2 Accuracy of the user-specific prevalence on the dataset

方法	学习者流行度分位数										Avg.
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	
GM	0.573 5	0.563 0	0.556 7	0.558 4	0.561 9	0.562 5	0.562 7	0.563 4	0.565 7	0.565 9	0.563 4
SM	0.632 7	0.635 9	0.639 7	0.636 8	0.637 6	0.637 5	0.637 1	0.636 6	0.637 1	0.639 1	0.637 0
AM	0.598 2	0.602 1	0.604 8	0.603 3	0.607 6	0.610 1	0.611 9	0.610 4	0.612 1	0.613 7	0.607 4
MoM	0.625 5	0.617 9	0.621 1	0.621 9	0.623 5	0.624 5	0.625 8	0.627 6	0.630 2	0.632 1	0.625 0
MF	0.751 9	0.761 1	0.762 5	0.764 4	0.767 9	0.768 7	0.769 9	0.770 6	0.771 4	0.774 6	0.766 3
LMF(S)	0.766 6	0.7637	0.769 9	0.770 4	0.773 9	0.775 1	0.777 9	0.779 6	<u>0.781 9</u>	<u>0.785 1</u>	0.774 4
LMF(A)	0.776 8	0.774 5	0.776 7	0.775 4	0.778 7	0.780 2	0.780 7	0.781 4	0.783 5	0.786 5	0.779 4
LMF(S,A)	<u>0.768 6</u>	<u>0.764 7</u>	<u>0.770 2</u>	<u>0.771 2</u>	<u>0.775 6</u>	<u>0.775 7</u>	<u>0.778 1</u>	<u>0.779 8</u>	0.781 4	<u>0.785 1</u>	<u>0.775 0</u>

表 3 学习成果在数据集上特定流行度的准确性

Tab. 3 Accuracy of learning outcomes for specific prevalence on datasets

方法	学习成果流行度分位数										Avg.
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	
GM	0.587 5	0.582 8	0.563 2	0.564 4	0.563 1	0.564 9	0.565 0	0.564 7	0.564 5	0.565 9	0.568 6
SM	0.648 3	0.649 8	0.645 3	0.644 0	0.639 8	0.638 2	0.638 0	0.639 4	0.641 0	0.639 1	0.642 3
AM	0.656 2	0.645 8	0.636 8	0.625 3	0.622 1	0.622 7	0.617 6	0.615 5	0.614 0	0.613 7	0.627 0
MoM	0.637 9	0.645 9	0.632 9	0.631 4	0.633 1	0.632 6	0.630 9	0.632 7	0.631 9	0.632 1	0.634 1
MF	0.765 4	0.762 1	0.771 5	0.775 3	0.775 9	0.777 1	0.777 6	0.778 3	0.776 2	0.774 6	0.773 4
LMF(S)	0.765 8	<u>0.769 4</u>	<u>0.778 5</u>	<u>0.783 1</u>	<u>0.784 8</u>	<u>0.786 6</u>	<u>0.787 1</u>	0.787 3	0.787 7	<u>0.785 1</u>	<u>0.781 5</u>
LMF(A)	<u>0.767 0</u>	0.766 5	0.773 2	0.779 3	0.783 1	0.785 5	<u>0.787 1</u>	0.787 9	0.787 7	0.786 5	0.780 4
LMF(S,A)	0.771 8	0.773 9	0.782 1	0.785 6	0.788 0	0.790 7	0.788 4	<u>0.787 8</u>	<u>0.787 0</u>	<u>0.785 1</u>	0.784 0

表 4 混合特征在数据集上特定流行度的准确性

Tab. 4 Accuracy of mixed features for specific prevalence on the dataset

方法	联合流行度分位数									Avg.
	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	
GM	0.567 2	0.553 6	0.551 3	0.556 3	0.561 3	0.561 6	0.561 4	0.564 7	0.565 9	0.560 4
SM	0.656 3	0.645 7	0.635 6	0.631 9	0.632 5	0.634 9	0.636 6	0.638 7	0.639 1	0.639 0
AM	0.646 8	0.615 9	0.605 4	0.613 3	0.617 5	0.613 8	0.612 5	0.612 6	0.613 7	0.616 8
MoM	0.636 3	0.612 7	0.623 4	0.621 4	0.625 8	0.626 3	0.629 6	0.629 8	0.632 1	0.626 4
MF	0.756 1	<u>0.767 4</u>	0.761 4	0.766 1	0.769 9	0.772 9	0.774 4	0.773 3	0.774 6	0.768 5
LMF(S)	0.760 6	0.772 8	<u>0.762 2</u>	0.770 6	<u>0.774 9</u>	0.780 2	0.782 7	0.784 9	<u>0.785 1</u>	0.774 9
LMF(A)	<u>0.757 7</u>	0.755 8	0.754 9	<u>0.771 3</u>	0.774 2	<u>0.780 9</u>	<u>0.782 4</u>	0.784 9	0.786 5	0.772 1
LMF(S,A)	0.752 4	0.761 9	0.766 5	0.774 9	0.779 0	0.781 0	0.782 7	<u>0.783 4</u>	<u>0.785 1</u>	<u>0.774 1</u>

这些情况下, MF 成功减轻了不少流行学习者的成果下降. 特别是, 联合利用学习者和成果信息的 MF(S, A) 在学习者(表 2)、评估(表 3)和联合(表 4)中普遍稳定. 有趣的是, 当学习者的流行程度被简化为表 2 时, MF(A) 的效果优于表 4 的 MF(S).

如图 2~4 所示, 在身份类别成果方面, 从左下角到右上角呈总体上升趋势, 符合身份类别平均趋势. 然而, 一些结果显示流行度低的身份类别有更好的准确性, 如图 3 中的身份类别 3, 这是因为实例的数量太少(少于 100 个), 从而使得抽样的不确定性影响了数据集背后的广泛趋势.

一方面, 除少数情况外, LMF 的结果始终优于 MF, 证明了 LM 引导辅助信息的有效性. 另一方面, 成果

最好的 MF 在不同身份类别之间有所不同,意味着在线学习成果预测的首选信息在不同身份类别之间是不同的。由此,还可以推测出,当采样的不确定性和预测难度较高时,辅助信息的预先注入有时是有害的。

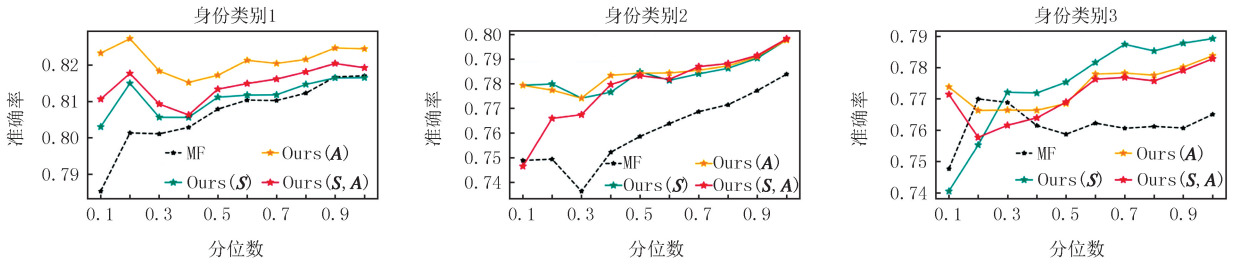


图2 数据集中学习者流行度对不同身份类别的性能

Fig.2 Performance of user prevalence for different identity categories in the dataset

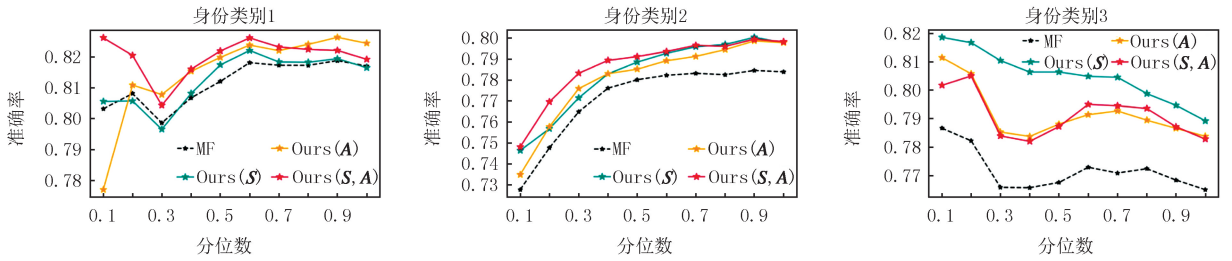


图3 数据集中学习成果流行度对不同身份类别的性能

Fig.3 Summary of learning outcomes prevalence performance for different identity categories in the dataset

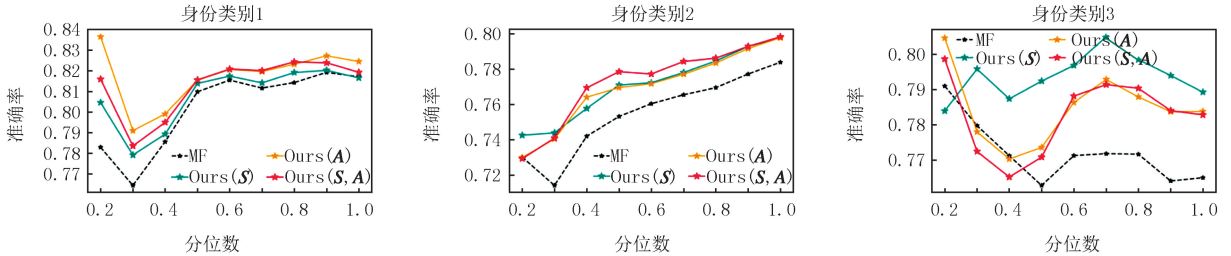


图4 数据集中联合流行度对不同身份类别的性能

Fig.4 Performance of joint prevalence for different identity categories in the dataset

2.4 提示选择和聚类分析

MF 基于预定义提示模板的语言化将多模态辅助信息转换为纯文本形式,会产生一个问题:如何确定最佳提示?为解决该问题,本文提供了一个半自动的过程来选择模板。首先,手动设计具有不同数量的学习者和成果的候选对象,其中包含属性、句子结构和表达模式。然后,将这些候选对象转换元数据,并生成相应的 LM 嵌入候选对象。最后,在这些嵌入候选算法上拟合了一个现成的聚类算法(例如 K-means),并使用轮廓系数等度量来衡量聚类结果的质量。实验中,选择轮廓系数最佳的归一化度量对应的候选对象进行嵌入。

选定模板后,对表现进行聚类分析,并作为副任务进行评估。如图 5(a)给出了一些评估集群的例子,来自数据集的实例评估和学习者嵌入被很好地分成了不同集群。说明聚类是基于身份类别和主要内容构建的,而不是在提示选择之后,也可以对学习者的表现进行聚类分析,并作为副任务进行评估,如图 5(b)中给出了一些例子,来自数据集的实例表现和学习者嵌入被很好地分成了不同的集群。说明集群是基于等级和主要概念,而不是难度、辨别能力和概率水平。例如,这里的主要概念包括了“关于物理中的质点运动学”“关于三角函数的课程”以及“关于高等数学中的极数”3 类。这意味着可以在预训练的 LM 嵌入空间上使用多模态元信息分析每个学习者和表现,这将触发新的应用,如个性化多模态教育内容推荐。

3 总结

本文率先使用多模态辅助信息来预测学习者的成果,并提出了一个新的框架 LMF。LMF 首先将多模态元数据转换为自然语言形式,提供给一个预训练的大语言模型,从而诱导语义丰富的嵌入。然后,利用矩阵分

解算法中的嵌入作为附加的学习信号,为每个军事职业教育学习者和成果注入信息丰富的先验.本文在一个包括 2.5 万个学习者的 35 万个互动和 5 万个不同的成果表现的数据集上验证了所提出的方法.在正常和冷启动环境下的广泛实验结果表明,与基线相比,MF 可以持续提高军事职业教育在线学习成果预测性能,并在学习者和/或成果不流行时普遍有效.此外,对学习者和 LM 嵌入多模态元数据的学习者表示进行聚类分析,提出了人工智能在军事职业教育领域的关键应用蓝图,如个性化多模态教育内容推荐.本文希望能阐明:(1)利用大规模的训练模型进行军事职业教育应用.(2)利用多模态辅助信息进行高度个性化的决策.虽然本文假设只使用语言模型的输出而绕过对语言模型的参数访问,但未来工作可以通过对 LM 微调,用 LM 自身来提高 LM 指导的质量.此外,本文方法在涵盖军事职业教育 JZ 在线教育平台学习者数据集上呈现结果时,也应该探索真实课堂环境数据集或非二进制评分数据集上的效果,以扩大其影响.

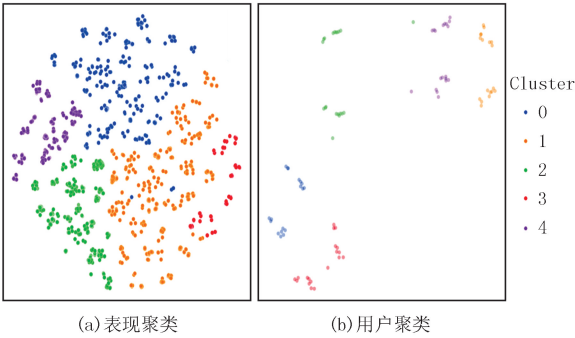


图5 成果(a)和学习者(b)的LM嵌入t-SNE图
Fig. 5 Results(a) and User(b) LM embedded t-SNE diagram

附录见电子版(DOI:10.16366/j.cnki.1000-2367.2024.08.12.0002).

参 考 文 献

[1] 彭涛,丁凌云.基于教育数据挖掘学生表现预测模型构建研究[J].黑龙江高教研究,2015,33(11):55-58.
PENG T,DING L Y.Performance prediction model based on data mining based education students[J].Heilongjiang Researches on Higher Education,2015,33(11):55-58.

[2] 江培超,王川,胡富珍,等.基于阅读认知诊断的学生表现预测[J].计算机工程与应用,2022,58(11):160-170.
JIANG P C,WANG C,HU F Z,et al.Student performance prediction based on reading cognitive diagnosis[J].Computer Engineering and Applications,2022,58(11):160-170.

[3] OSKOU EI R J,ASKARI M.Predicting academic performance with applying data mining techniques(generalizing the results of two different case studies)[J].Computer Engineering and Applications Journal,2014,3(2):79-88.

[4] NGHE N T,JANECEK P,HADDAWY P.A comparative analysis of techniques for predicting academic performance[C]//2007 37th Annual Frontiers In Education Conference-Global Engineering:Knowledge Without Borders,Opportunities Without Passports.Milwaukee: IEEE,2007:T2G-7-T2G-12.

[5] ASIF R,MERCERON A,ALI S A,et al.Analyzing undergraduate students' performance using educational data mining[J].Computers & Education,2017,113:177-194.

[6] KABAKCHIEVA D. Student performance prediction by using data mining classification algorithms[J]. International Journal of Computer Science and Management Research,2012,1(4): 686-690.

[7] HUY-LE Q,PHAM-THI T,HUYNH-TUONG N,et al.Student's academic performance prediction in the programming techniques course by using machine learning methods[C]//2024 IEEE 30th International Conference on Telecommunications(ICT).Amman:IEEE,2024: 1-7.

[8] AL-SHEHRI H,AL-QARNI A,AL-SAATI L,et al.Student performance prediction using Support Vector Machine and K-Nearest Neighbor[C]//2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering(CCECE).Windsor:IEEE,2017:1-4.

[9] LUO Y Y,HAN X B,ZHANG C Y.Prediction of learning outcomes with a machine learning algorithm based on online learning behavior data in blended courses[J].Asia Pacific Education Review,2024,25(2):267-285.

[10] HU Q,RANGWALA H. Academic performance estimation with attention-based graph convolutional networks[EB/OL]. [2024-07-12]. arXiv:2001.00632. 2019. <https://arxiv.org/abs/2001.00632>.

[11] NIU K,CAO X,YU Y. Explainable student performance prediction with personalized attention for explaining why a student fails[EB/OL].[2024-07-12]. arXiv:2110.08268. 2021. <https://arxiv.org/abs/2110.08268>.

[12] REN Z Y,NING X,LAN A S,et al.Grade prediction with neural collaborative filtering[C]//2019 IEEE International Conference on Data Science and Advanced Analytics(DSAA).Washington,DC:IEEE,2019:1-10.

[13] ISINKAYE F O,FOLAJIMI Y O,OJOKOH B A.Recommendation systems:Principles,methods and evaluation[J].Egyptian Informatics Journal,2015,16(3):261-273.

[14] RESNICK P,VARIAN H R.Recommender systems[J].Communications of the ACM,1997,40(3):56-58.

[15] SWEENEY M,LESTER J,RANGWALA H.Next-term student grade prediction[C]//2015 IEEE International Conference on Big Data (Big Data).Santa Clara:IEEE,2015:970-975.

[16] SWEENEY M,RANGWALA H,LESTER J,et al. Next-term student performance prediction: A recommender systems approach[J]. Journal of Educational Data Mining,2016,8(1): 22-50.

[17] ZHONG S T,HUANG L,WANG C D,et al.Constrained matrix factorization for course score prediction[C]//2019 IEEE International Conference on Data Mining(ICDM).Beijing:IEEE,2019:1510-1515.

[18] GUO Z X,LIU T Y,CHANG L,et al.Deep matrix factorization based on self-attention mechanism for student grade prediction[J].Journal of Physics:Conference Series,2020,1651(1):012019.

[19] DRACHSLER H,VERBERT K,SANTOS O C,et al.Panorama of recommender systems to support learning[M]//Recommender Systems Handbook.Boston:Springer,2015:421-451.

[20] 李璐,张志军,范钰敏,等.面向冷启动用户的元学习与图转移学习序列推荐[J].山东大学学报(工学版),2024,54(2):69-79.

LI L,ZHANG Z J,FAN Y M,et al.Sequential recommendation for cold-start users with meta graph transitional learning[J].Journal of Shandong University(Engineering Science),2024,54(2):69-79.

[21] 张子扬,刘小洋.基于元学习的多视图对比融合冷启动推荐算法[J].计算机应用研究,2024,41(7):2025-2032.

[22] PAL S,PRAMANIK P K D,CHOUDHURY P.Learner's intention analysis to mitigate the cold start problem in personalized learning recommendation systems[J].Multimedia Tools and Applications,2025,84(19):21273-21327.

[23] FERNÁNDEZ-TOBIÁS I,BRAUNHOFER M,ELAHI M,et al.Alleviating the new user problem in collaborative filtering by exploiting personality information[J].User Modeling and User-Adapted Interaction,2016,26(2):221-255.

[24] LIU T Q,WANG Z W,TANG J L,et al.Recommender systems with heterogeneous side information[C]//The World Wide Web Conference.San Francisco:ACM,2019:3027-3033.

[25] THAI-NGHE N,DRUMOND L,HORVÁTH T,et al. Using factorization machines for student modeling[C]//Proceedings of UMAP Workshops. Montreal:[s.n.],2012:16.07.-20.07.

Multi-modal military vocational education online learning outcome prediction based on matrix factorization

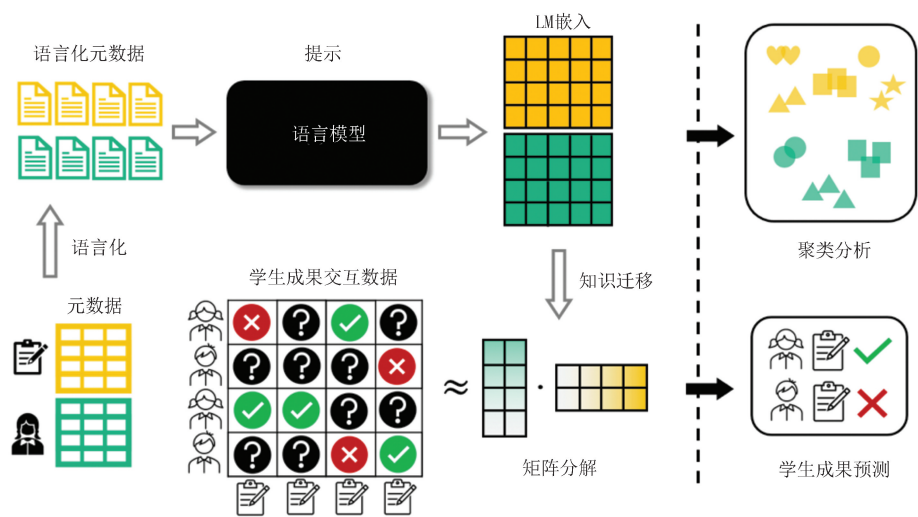
Shao Dongchun¹, Li Kang¹, Du Ying¹, Zhao Han¹

(Army Academy of Armored Forces, Bengbu 233000, China)

Abstract: Online learning outcome prediction in military vocational education has received wide attention due to its educational significance. In online learning outcome prediction methods, recommendation-based methods are more important. However, the recommendation system has problems with cold-start and incorporating side information. Although the existing work addresses the above problems by utilizing auxiliary information, the potential of multi-modal based recommended online learning outcome prediction methods has not been investigated. For the first time, this paper uses multi-modal assisted student online learning outcome prediction, using a large language model as an information mixer to generate additional guidance signals for matrix factorization. Specifically, the language-guided matrix factorization model in this paper, using linguistic multi-modal information, produces rich semantic embeddings for educational interactions and uses them as auxiliary signals for matrix factorization. This paper provides two text embedding methods, embedding initialization method calculates specific priors of the data to initialize the decomposition matrix, and embedding distillation method is used to align the underlying features and embedding features of the matrix decomposition, making the language model more fully guide the matrix decomposition. This paper evaluates the proposed model on a dataset collected on an online education platform, including 3.5×10^5 interactions for 2.5×10^4 users and 5×10^4 different evaluations of outcomes. A lot of experiments show that the current model outperforms the existing methods in various military vocational education online learning outcome prediction scenarios, including cold-start.

Keywords: military vocational education; online learning outcomes predication; multi-modal; cold-start; matrix factorization

[责任编辑 赵晓华 刘洋]



图S1 在线学习成果预测框架
Fig. S1 Online learning outcome prediction framework

表 S1 学习者数学能力测量数据集示例
Tab. S1 Example of user mathematical ability measurement dataset

(a)学习者元数据				(b)成果元数据				
学号	身份类别	理解能力	平均分	学号	评估难度	辨别能力	概率	标签
A010000727	1	0.196 8	0.967 8	A010001001	−3.719 7	0.721 1	0.010 1	5 844
A010000542	3	−0.321 5	0.939 8	A010003005	−1.717 3	1.011 8	0.590 1	6 307
A040000160	2	−0.162 9	0.789 2	A060066004	0.129 0	1.425 5	0.001 1	610
A060001349	3	−0.337 9	0.624 2	A080080003	−0.453 6	1.958 1	0.000 0	26
A060002328	2	0.910 2	0.904 6	A090020003	0.001 5	1.475 4	0.001 0	974
(c)学习者—成果交互数据				(d)数学知识结构数据				
学号	成果编号	交互编码	标签	主要概念	描述			
A010000727	A010001001	1	5844	加速度	速度变化量与所需时间的比值			
A010000727	A010003005	0	610	判断级数	将数列的项依次用加号连接起来的函数			
A010000542	A010003005	0						
A060001349	A060066004	1	26	行列式计算	任一行(列)的各元素与其对应的代数余子式的乘积之和			
A060002328	A060066004	1						

表 S2 13 个模型和 4 个指标上不同身份类别的在线学习成果预测

Tab. S2 Prediction of online learning outcomes for different identity categories on 13 models and 4 indicators

方法	身份类别 1				身份类别 2				身份类别 3			
	Acc.	Prec.	Rec.	F_1	Acc.	Prec.	Rec.	F_1	Acc.	Prec.	Rec.	F_1
GM	0.672	0.778	0.808	0.792	0.597	0.703	0.738	0.720	0.577	0.667	0.719	0.692
SM	0.694	0.801	0.805	0.802	0.655	0.754	0.755	0.754	0.632	0.718	0.729	0.723
AM	0.713	0.803	0.834	0.819	0.658	0.748	0.774	0.761	0.636	0.711.	0.757	0.733
MoM	0.718	0.810	0.832	0.821	0.656	0.754	0.757	0.755	0.638	0.714	0.751	0.732
MF	0.822	0.843	0.951	0.894	0.794	0.831	0.902	0.865	0.775	0.801	0.888	0.843
HUY-LE Q ^[7]	0.751	0.731	0.894	0.821	0.798	0.762	0.801	0.792	0.712	0.751	0.725	0.765
LUO Y Y ^[9]	0.819	0.820	0.911	0.824	0.809	0.821	0.899	0.851	0.752	0.786	0.797	0.801
IN(S)	0.806	<u>0.849</u>	0.918	0.882	0.791	0.839	0.881	0.860	0.781	0.819	0.860	0.839
IN(A)	0.820	0.856	0.929	0.891	0.804	0.846	0.895	0.870	<u>0.796</u>	<u>0.827</u>	0.885	0.855
IN(S,A)	0.804	<u>0.849</u>	0.915	0.880	0.793	<u>0.840</u>	0.881	0.860	0.786	0.831	0.857	0.844
DI(S)	0.827	0.848	<u>0.953</u>	<u>0.898</u>	<u>0.807</u>	0.834	0.918	<u>0.874</u>	0.797	0.813	<u>0.907</u>	0.857
DI(A)	0.830	0.847	0.958	0.899	0.812	0.839	0.918	0.877	0.792	0.808	0.910	<u>0.856</u>
DI(S,A)	<u>0.829</u>	0.846	0.958	0.899	<u>0.807</u>	0.835	<u>0.917</u>	<u>0.874</u>	0.789	0.808	0.905	0.854

注: Acc.代表准确率;Prec.代表精度;Rec.代表召回率.