

支训涛, 刘凌杰. 融合双向 GRU 和注意力机制的混合推荐模型[J]. 智能计算机与应用, 2026, 16(4): 35-41. DOI: 10.20169/j. issn. 2095-2163. 25030902

融合双向 GRU 和注意力机制的混合推荐模型

支训涛, 刘凌杰

(安徽建筑大学 电子与信息工程学院, 合肥 230000)

摘要: 推荐系统在提供基于个人历史记录个性化信息方面至关重要。然而由于评分数据的稀疏性, 准确捕捉用户偏好仍然是一个挑战。现有的许多研究引入评论信息来缓解这一现象。评论文本作为重要的辅助信息具有多样化的特征, 包含了用户维度的属性描述, 能够有效缓解评分数据的稀疏性。本研究提出了一种融合双向门控循环单元(Gated Recurrent Unit, GRU)和注意力机制的混合深度模型(HRM-BAM), 旨在同时学习评分特征和文本特征。该模型使用两个嵌入层, 分别学习用户和物品之间交互的评分特征。此外, 两个基于注意力机制的门控循环单元(GRU)网络从用户和物品的评论中获取上下文感知表示。接着, 通过特征级联和卷积神经网络处理, 将评分特征和评论特征有机结合, 利用多层卷积和池化操作提取更深层次的特征信息。最后, 通过全连接层对用户评分进行精确预测。实验结果表明, 在三组数据集上, 该模型的均方根误差(RMSE)相较于基线方法有显著提高。

关键词: 个性化推荐; 注意力机制; 门控循环单元; 特征提取; 评分预测

中图分类号: TP391.1

文献标志码: A

文章编号: 2095-2163(2026)04-0035-07

Deep recommendation models incorporating bidirectional GRU and attention mechanisms

ZHI Xuntao, LIU Lingjie

(School of Electronic and Information Engineering, Anhui Jianzhu University, Hefei 230000, China)

Abstract: Recommender systems are crucial in providing personalized information based on an individual's history. However, accurately capturing user preferences remains a challenge due to the sparsity of rating data and the heterogeneity of review data. In this study, we propose a hybrid deep model (HRM-BAM) that incorporates bidirectional GRU (Gated Recurrent Unit) and attention mechanisms, aiming to learn both rating features and text features. The model uses two embedding layers to learn rating features for interactions between users and items, respectively. In addition, two gated recurrent unit (GRU) networks based on the attention mechanism acquire context-aware representations from user and item reviews. Next, the rating features and review features are organically combined through feature cascading and convolutional neural network processing, and deeper feature information is extracted using multilayer convolution and pooling operations. Finally, the user ratings are accurately predicted through the fully connected layer. The experimental results show that the root mean square error (RMSE) of the model is significantly improved compared to the baseline method on the three datasets.

Key words: personalized recommendation; attention mechanism; gated recurrent unit; feature extraction; score prediction

0 引言

近年来, 推荐系统在电子商务、信息检索和社交媒体等领域发挥着关键作用, 帮助用户在海量信息中高效定位所需内容^[1]。随着需求的不断扩大, 各类推荐方法快速发展, 其中协同过滤以及矩阵分

解^[2]因简单高效而备受关注。但在实际应用中, 点积运算难以拟合非线性关系且可解释性不足, 加之评分数据的高稀疏性导致相关性难以识别^[3]。研究者因此引入更多辅助信息(如标签、社交关系、评论、视觉特征等)来提升推荐效果, 其中评论不仅能反映用户偏好, 还为推荐结果提供合理解释^[4],

基金项目: 中国高校产学研创新基金——新一代信息技术创新项目(2023IT001)。

作者简介: 刘凌杰(2000—), 男, 硕士研究生, 主要研究方向: 个性化推荐, 托攻击检测。

通信作者: 支训涛(1999—), 男, 硕士, 主要研究方向: 个性化推荐, 深度学习。Email: 13155680768@163.com。

收稿日期: 2025-03-09

成为近年来改善推荐系统性能和可解释性的主要方向。

在融合评论和评分的具体方法上, HFT^[5]、CTR^[6]、CDL^[7]、RMR^[8]、RMLT^[9]等模型通过潜在话题或深度表征来补充评分信息,但多采用词袋模型,忽略了上下文关系与语序信息。随着深度学习的兴起,研究者开始利用更深层次的表征来融合评论与评分,提升推荐的准确性和可解释性。然而,现有方法往往在整合评论和评分时仍有局限:要么对评分利用不足,要么过于依赖评论而限制了其潜在价值。如 DeepCoNN^[10]主要基于辅助信息学习用户与物品表示;而 CTR、CDL 和 ConvMF^[11]有时会将评论与评分表示相互靠近,从而弱化了评论中丰富的信息。

为此,本文提出了一种深度混合模型:先利用 GRU 学习评论的高阶语义特征,并通过两个嵌入层学习评分特征,然后使用卷积神经网络融合文本与评分信息,最终将用户和物品的潜在向量连接以完成评分预测。

1 相关工作

1.1 基于深度学习的协同过滤

协同过滤领域的深度学习研究致力于通过先进的神经网络架构优化用户和物品的表示学习。Salakhutdinov^[12]等先驱性地将深度学习引入协同过滤,通过修改受限玻尔兹曼机,创新性地构建了两层无向图模型来捕捉用户偏好。然而,这种模型在处理大规模数据时计算成本高,且在面对数据稀疏时表现受限。为了应对这些挑战,Wu 等^[13]提出了协同去噪自编码器(CDAE),该方法通过重构用户-项目反馈数据,有效降低数据噪声,提升推荐性能。但 CDAE 在高维数据上可能面临过拟合问题。更为深入的研究方向是通过多层感知器提取高阶潜在特征,以对隐式反馈数据进行精细建模,这不仅扩展了传统协同过滤的表征能力,也为个性化推荐提供了更为深入的技术路径。

1.2 结合深度学习的混合推荐

在社交媒体环境下,社交关系、标签、评论和位置信息对用户偏好和行为的理解至关重要。深度学习同样被广泛应用于大规模多媒体数据的处理。Wang 等^[14]提出了一种融合贝叶斯堆叠去噪自编码器(SDAE)和概率矩阵分解的协同深度学习方法,通过 SDAE 学习项目的潜在表示,并利用概率矩阵分解分析文本与项目的信息。与此同时,Bansal 等^[15]基于门控递归单元(GRU)设计了一种混合推

荐方法,结合评级和项目描述实现多任务学习。Kim 等^[16]所提出的卷积矩阵分解模型则引入辅助文本信息,运用卷积神经网络学习项目的潜在表示,通过多层卷积操作捕捉文本中词汇之间的相关性。这些研究不仅延续了协同过滤领域的进展,也为个性化推荐提供了新的视角与方法。

1.3 基于注意力机制的推荐方法

近年来,基于注意力机制的推荐系统取得了显著进展,但可解释性仍有待提升。已有研究如 D-att 模型^[17]、基于注意力记忆网络的微博推荐^[18]、以及双重注意力机制的协作模型^[19]等,通过关注评论中的重要单词和短语,为推荐提供更直观的解释。然而,这些方法在大规模数据处理、融合非文本信息以及应对噪声方面仍存在局限。

针对上述问题,本文提出了一种融合双向 GRU 和注意力机制的混合模型(HRM-BAM)。该模型中,首先使用两个嵌入层学习用户和物品的评分特征;然后通过两个注意力门控循环单元(GRU)从用户与物品的评论中获取上下文表示;接着利用特征级联和卷积神经网络,将评分特征与评论特征融合并提取更深层次信息;最后通过全连接层进行评分预测。实验结果表明,该模型在多个数据集上显著优于现有算法,展现了其在推荐系统中的良好应用前景。

2 HRM-BAM 模型

2.1 HRM-BAM 模型框架

如图 1 所示,HRM-BAM 模型由两个并行深度神经网络构成,分别处理用户和物品评论文本,记作 N_u 和 N_i 。最终在耦合层生成预测评分。第一层为嵌入网络层,使用预训练词向量将原始文本转化为词序列矩阵;第二层为双向 GRU,用于抽取文本的语义特征;第三层引入注意力机制,关注不同词语在上下文中的重要性;第四层为池化层,对提取特征进行降维和压缩;接着通过特征映射层将特征统一到固定维度;最后在预测层输出评分。该模型针对用户和物品的网络结构对称, N_u 和 N_i 输入和输出虽不同,但整体流程相似。

2.2 基于评分的嵌入表示学习

为了从评分中学习用户和物品的潜在因素,本文使用两个嵌入层,分别为嵌入层 $\text{Embedding}^{(U)}$ 和 $\text{Embedding}^{(I)}$,将用户/物品表示为密集向量,即用户嵌入层学习用户 n 的 K 维密集向量 U_n^R ,物品嵌入层学习物品 m 的 K 维密集向量 V_m^R 。

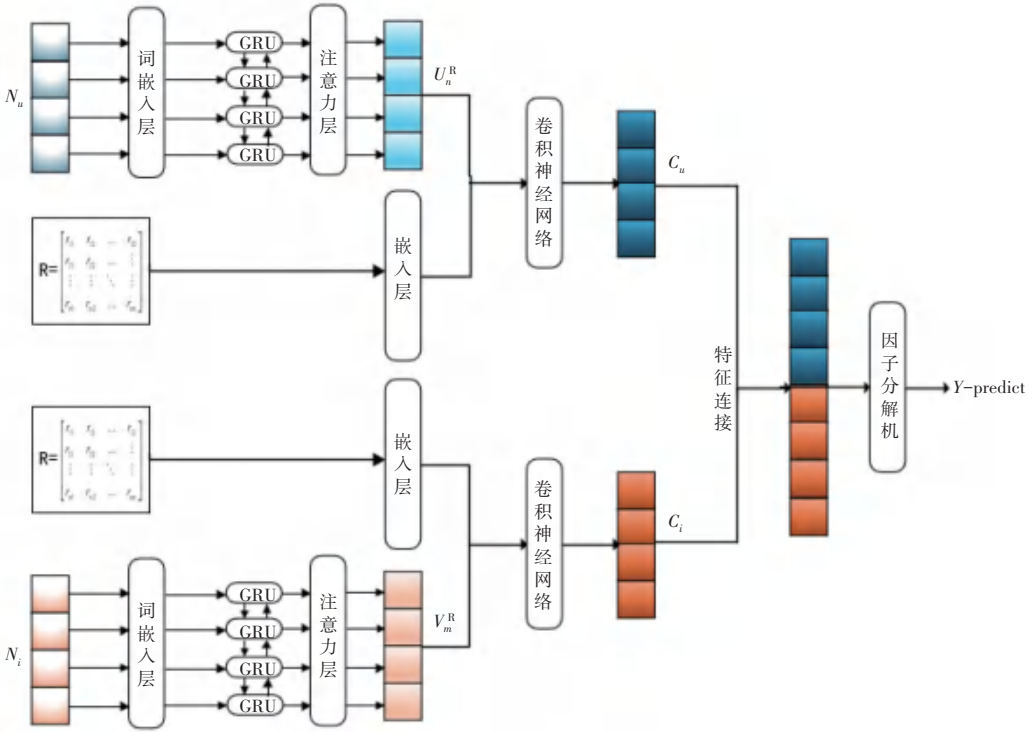


图 1 HRM-BAM 模型架构图

Fig. 1 Architecture diagram of HRM-BAM model

$$U_n^R = \text{Embedding}^{(U)}(n) \in \mathbb{R}^K \quad (1)$$

$$V_m^R = \text{Embedding}^{(I)}(m) \in \mathbb{R}^K \quad (2)$$

2.3 基于上下文的评论表示学习

本文通过结合注意力机制的双向 GRU 来提取评论中的上下文信息。相较于 LSTM, GRU 不仅参数更少、收敛更快, 而且在与注意力机制结合后能更准确地抽取潜在特征。双向 GRU 由正反两个 GRU 网络组成, 并通过拼接方式组合其结果。本节阐述的学习过程如下:

$$u_n^C = \text{GRU}^{(U)}(X_n^{(U)}) \quad (3)$$

$$v_m^C = \text{GRU}^{(I)}(X_m^{(I)}) \quad (4)$$

2.3.1 嵌入层

嵌入层将单词序列转换为稠密矩阵, 可选择随机初始化或采用预训练词嵌入 (如 Word2Vec、GloVe) 以加速收敛。由于各序列长度不同, 通常统一到最大长度并对短序列进行零填充。对于给定序列 $X = [x_1, x_2, \dots, x_M]$, 会从预训练词向量中查找每个单词的嵌入, 并将其拼接为稠密矩阵, 生成过程如下式:

$$X = \{\dots x_{m-1} x_m x_{m+1} \dots\} \quad (5)$$

2.3.2 双向 GRU 层

双向 GRU 层通过同时考虑前后文信息, 从词序列中提取上下文信息, 因此性能优于普通 GRU。其

由两个独立的 GRU 网络在相反方向运行。在实际应用中, 同一词的语义可能因上下文而异, 而标准 GRU 仅基于之前的状态计算当前隐藏状态。

步骤 t 的激活值被表示为前一激活值 h_{t-1} 的线性组合。候选激活值 \tilde{h}_t 可以表示为:

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \quad (6)$$

隐藏状态的计算如下:

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

更新门 z_t 的计算如下式:

$$z_t = \sigma(W_z x_t + U_z h_{t-1} + b_z) \quad (8)$$

重置门 r_t 的计算如下式:

$$r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r) \quad (9)$$

用户评论的最大长度为 $L(U)$, 物品评论的最大长度为 $L(V)$ 。本文讨论用户评论的处理方式。双向 GRU 包含两个模块: 前向 GRU $\overrightarrow{f}_{\text{GRU}}$ 从 $X_i^{(L(U))}$ 到 $X_{iL}^{(U)}$ 正向读取, 反向 GRU $\overleftarrow{f}_{\text{GRU}}$ 从 $X_{iL}^{(U)}$ 到 $X_i^{(L(U))}$ 正向读取数据。

$$\overrightarrow{h}_i = \text{GRU}(W_{\text{GRU}}^{\text{forward}} X_i) \quad (10)$$

$$\overleftarrow{h}_i = \text{GRU}(W_{\text{GRU}}^{\text{backward}} X_i) \quad (11)$$

其中, X_i 是评论, W_{GRU} 是 GRU 单元的参数。序列的最终注释通过连接前向隐藏状态 \overrightarrow{h}_i 和后向隐

藏状态 \overleftarrow{h}_i 得到:

$$h_i = [\overrightarrow{h}_i, \overleftarrow{h}_i] \quad (12)$$

2.3.3 分层注意力层

并非所有词在句子中的语义贡献都相同。例如,在句子“这个手机的电池真的耐用!”中,词“真的”和“耐用”更能反映用户的满意度。文档可能包含多个主题,某个词可能对特定主题更相关。为捕捉评论中关键语义信息,设计分层注意力机制:

(1)局部注意力:计算词语级注意力,衡量每个词对上下文的贡献。通过学习参数矩阵 $W_a \in \mathbb{R}^{d_a \times 2d_h}$ 和上下文向量 u_a 生成注意力得分:

$$a_t = \text{Softmax}(u_a^T \cdot W_a \cdot h_t) \quad (13)$$

其中, h_t 为第 t 个词的隐层状态。

加权后的评论表示为:

$$c_{\text{local}} = \sum_{t=1}^T a_t \cdot h_t \quad (14)$$

其中 T 为评论中的词语总数。

(2)全局注意力:引入用户-物品交互信息,通过评论评分生成全局权重,调整不同评论的重要性。设用户 u 的评分分量为 R_u , 全局注意力得分通过以下指标计算:

$$\alpha_u = \text{Softmax}(W_g \cdot c_{\text{local}}) \quad (15)$$

其中, W_g 为全局权重矩阵, α_u 为用户 u 对应评论的全局注意力权重。

最终评论特征向量由局部与全局表示的加权结合:

$$c_{\text{final}} = \sum_{i=1}^N \alpha_u^i \cdot c_{\text{local}}^i \quad (16)$$

其中, N 为评论总数, i 为第 i 评论。

2.3.4 池化与特征压缩层

为提取评论的高层次特征,针对 GRU 输出实施最大池化操作。本文对最终的评论特征 text_u 和物品评论特征 text_i , 通过连接层降维至 K 维, 与评分对齐。最终用户评论特征 c_u 与物品评论特征 c_i 通过池化操作合并为:

$$c_{\text{pooled}} = \max(\text{pool}(c_u), \text{pool}(c_i)) \quad (17)$$

2.4 基于上下文的评论表示学习

在 2.2 节和 2.3 节中,分别获得了用户与物品评分向量 U_n^R 和 V_m^R 与评论向量 c_u 和 c_i 。为融合这些特征并进行评分预测,本文先将评分与评论向量拼接:

首先,本文将用户评分特征和评论特征进行级联,形成一个统一的特征表示:

$$F_{\text{joint}} = [U_n^R \oplus C_u \oplus V_m^R \oplus C_i] \in \mathbb{R}^{4K} \quad (18)$$

然后通过卷积神经网络提取深层语义特征,输出多个特征图:

$$F_{\text{cnn}} = \text{CNN}(F_{\text{joint}}) \quad (19)$$

再利用最大池化来选择局部最优的特征:

$$F_{\text{pooled}} = \text{MaxPool}(F_{\text{cnn}}) \quad (20)$$

获取最优特征表示,最后将池化结果输入全连接层,得到预测评分

$$\hat{y} = W \cdot F_{\text{pooled}} + b \quad (21)$$

其中, W 为权重矩阵, b 是偏置项。

3 实验结果与分析

3.1 数据集

本研究采用了来自 Amazon5-core^[20] 的 3 个领域数据:Toys and Games (TG)、Digital Music (DM) 和 Kindle Store (KS),其中, TG、DM、KS 分别包含 1~5 星用户评分与真实评论。选择这些数据集是为了覆盖不同领域与规模,并利用评论文本获取更丰富的用户偏好信息,从而在评分之外捕捉更深层次的态度与需求,增强推荐的准确度与可解释性。KS 数据量最大,包含约 98 万条评论,而 DM 最小,仅有约 6 万条评论。表 1 展示了各数据集的基本统计信息。

表 1 实验数据集信息

Table 1 Experimental dataset information

数据集	用户数	项目数	样本数	稀疏度/%
DM	5 541	3 568	64 706	99.67
TG	19 412	11 924	167 597	99.92
KS	68 223	61 934	982 619	97.67

3.2 评价指标与基线方法

在评分预测场景中, RMSE^[21] 通常被视为核心衡量指标,因其直观量化了模型预测与真实评分之间的偏差,并且有成熟的优化方法可用。鉴于本研究的重点在于比较不同模型的评分预测效果, RMSE 足以满足性能评估需求。实验中对所有模型及 HRM-BAM 均采用 RMSE 作为评估指标,数值越低表示预测精度越高。RMSE 的计算公式为:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{u,i} (\hat{r}_{u,i} - r_{u,i})^2} \quad (22)$$

其中, $\hat{r}_{u,i}$ 为用户 u 对物品 i 的预测评分; $r_{u,i}$ 为真实评分; N 表示用户和物品之间的评分数。

为评估本文模型在评分预测方面的表现,本文进行了对比实验,将 HRM-BAM 与 4 个先进的模型

进行比较,即PMF、HFT、CMF、DeepCoNN。

(1)PMF^[10]:一种经典的概率矩阵分解模型,其输入仅包括评分数据。

(2)HFT^[4]:一种基于评论的推荐系统方法,其利用用户和物品评论的主题分布来进行潜在层次的评分预测。

(3)CMF^[14]:将卷积神经网络与矩阵分解相结合,通过使用积累卷积网络对项目潜在特征进行评级预测。使用词嵌入(word embedding)来学习上下文感知的项目特征信息。

(4)DeepCoNN^[10]:一种深度协同神经网络模型,首次将深度学习应用于结合用户评论和商品评论。该模型包含两个并行的卷积神经网络,分别负责处理用户评论和商品评论,以独立提取用户和商品的特征表示。此外,该模型不依赖于评分矩阵。

3.3 实验环境及参数设置

实验使用Python 3.8,在Intel Core i5-8300H CPU (2.30 GHz)、32 GB内存、NVIDIA GTX 1050Ti GPU环境下运行。数据集划分为训练集(80%)、测试集(10%)和验证集(10%)。参数采用截断高斯分布初始化,小批量大小为100,学习率为0.0001或0.001。潜在维度可选[4, 8, 16, 32, 64],双向GRU隐藏维度可选[5, 10, 20, 50, 100, 200],注意力层大小选自[1, 2, 5, 10, 20, 50],并于最终预测层前添加dropout比例[0, 0.1, 0.2, 0.5]。使用Google word2vec作为预训练词嵌入,共运行5次并报告平均结果。在HRM-BAM^[22]中,dropout比例固定为0.5,针对不同规模数据集,参数集信息见表2。

表2 参数集信息

Table 2 Parameter set information

数据集	潜在维度数	GRU隐藏层数	Dropout比率
DM	8	10	0.5
TG	16	20	0.5
KS	32	50	0.5

对于DeepCoNN,滤波器数量设置为50(用于DM和TG),其余数据集设置为100。潜在因子维度在DM数据集上为8,其余数据集为16。对于CMF,正则化参数 λ_u 设置为0.1, λ_v 设置为100。对于PMF,正则化参数 λ_u 和 λ_v 设置为0.1。对于HFT模型,少量数据集DM中隐藏层维度设置为5,而TG数据集设置为10,KS数据集的隐藏层维度设置为20。

3.4 实验结果

3.4.1 整体性能分析与比较

通过表3可知,所提出的方法在3个真实数据集上表现均优于所有基线模型,且在对比PMF(仅用评分)与HFT、DeepCoNN(仅用评论)时可以发现,辅助信息对推荐性能贡献很大;当在DeepCoNN基础上融入评分信息后,性能进一步提升,说明评分同样至关重要;相比同样结合评分与评论的CMF,DeepCoNN表现反而更好,而本方法通过在评分与评论信息间取得更佳平衡,实现了最优的推荐结果。

表3 基线模型性能对比

Table 3 Baseline model performance comparison

方法/数据集	DM	TG	KS
PMF	1.026	1.023	1.160
HFT	1.007	0.986	1.130
CMF	0.972	1.004	1.200
DeepCoNN	0.921	0.930	1.061
HRM-BAM	0.907	0.923	1.031
降低	3.4%	2.1%	0.6%

3.4.2 消融实验

为进一步验证HRM-BAM算法模型各个组成部分在提升推荐性能上的作用,本文在DM和TG数据集上展开消融实验,实验结果如图2所示。

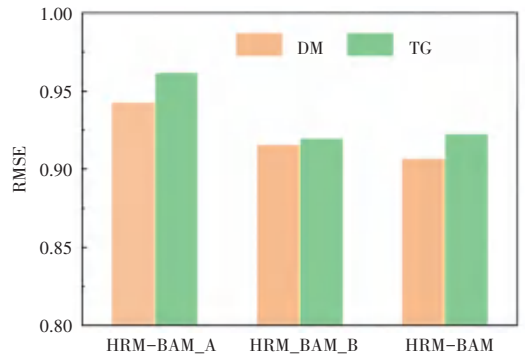


图2 消融分析

Fig. 2 Ablation analyses

HRM-BAM_A模型:简单地使用两个并行的双向GRU网络从用户和物品中学习特征,并通过将相应的特征馈送到分解机中进行评级预测。

HRM-BAM_B模型:该模型建立在前一模型的基础上,但用户和物品评级特征由用户和物品嵌入层学习。将用户和物品的上下文感知特征和评级特征连接起来,馈送到分解机进行评级预测。

HRM-BAM模型:该模型是本文提出的方法,在HRM-BAM_B模型的基础上加入了卷积神经网络

对评论特征和评分特征进行融合。

从图 2 可以看出,针对 DM 和 TG 这两个数据集,HRM-BAM_A 模型的 RMSE 明显要比 HRM-BAM_B 的值要高,这说明评分和评论之间的融合是非常必要的,在利用评论作为重要辅助信息进行预测评分的同时,也不能忽略评分对提高推荐性能的重要性。HRM-BAM_B 模型的 RMSE 值要比 HRM-BAM 的 RMSE 值略高,这表明加入卷积神经网络对评论特征和评分特征进行融合时,可以自适应地利用评分和评论之间的贡献。总的来说,消融实验的结果进一步证明了所提出方法的有效性。

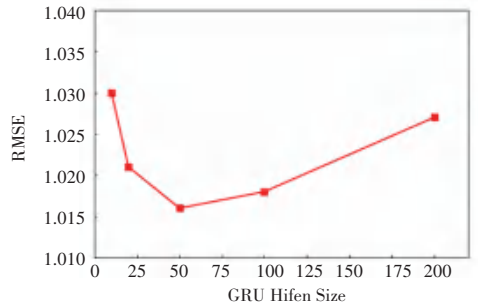
3.4.3 参数分析

在所提出的方法中,有两个重要参数:GRU 隐藏层大小、潜在维数的大小。本文将在本节中展示这些参数的影响。

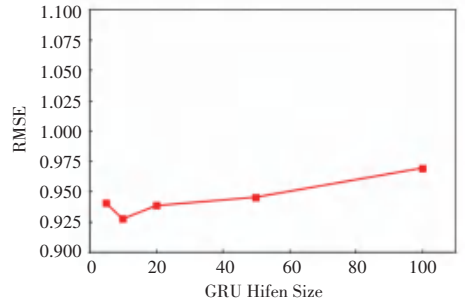
图 3 展示了 GRU 隐藏层大小在 DM 及 KS 数据集上对模型性能的影响。本文可以观察到,所提出的方法对 GRU 隐藏单元的大小非常敏感。在 DM 数据集中,当大小为 10 时获得最佳性能,而在 KS 数据集中,当大小为 50 时获得最佳性能。由于 GRU 隐藏单元的大小决定了评论中信息的发现程度,因此较大的数据集可能包含更多的信息,并且可能需要更大的 GRU 隐藏单元。这样,小的数据集可以将 GRU 隐藏单元的大小设置为较小的值。当数据相对较小时,本文可以将 GRU 隐藏单元大小设置为 10,而当数据相对较大时,可以设置为 50。这表明 GRU 隐藏单元的大小应该与数据大小相匹配,以实现最佳性能。

图 3 展示了不同潜在维数在 DM 和 KS 数据集上对模型性能影响。本文从候选集 [4, 8, 16, 32, 64] 中选择维数,之所以从 [4, 8, 16, 32, 64] 这 5 个维度中选取,一方面是因为其覆盖了从较低阶到相对较高阶的典型取值;另一方面,这样的候选范围在实际应用和先前研究中都较为常用,能够有效比较不同维度对模型复杂度、训练效率和预测精度的影响。尽管 64 维在某些深度模型中显得并不高,但对于大多数协同过滤或基于潜在因子的应用来说,已经能够提供足够的表示能力,可以更好地捕捉用户与项目之间的复杂关系,同时又不会给训练和存储带来过大的负担。

根据图 4 可得所提出的方法对潜在维数很敏感,在两个数据集上,潜在维数小的结果比潜在维数大的结果表现得更好。在实践中,可以采用小于 32 的小尺寸。



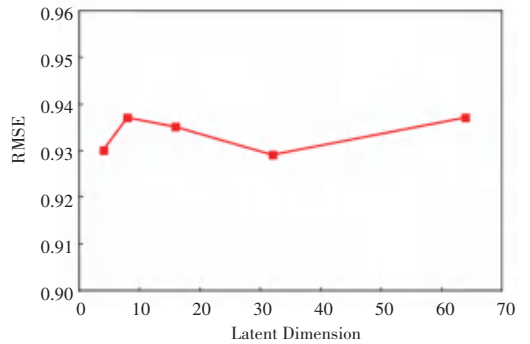
(a) Kindle Store



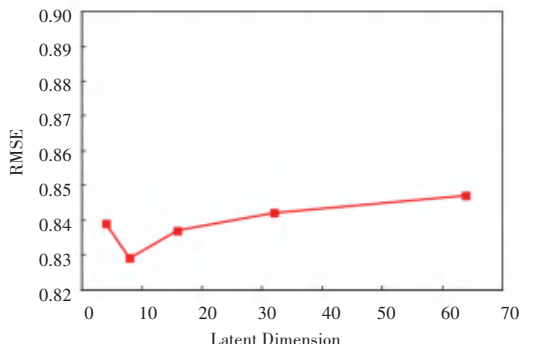
(b) Digital Music

图 3 GRU 隐藏层大小对模型性能影响

Fig. 3 Impact of GRU hidden layer size on model performance



(a) Kindle Store



(b) Digital Music

图 4 不同潜在维数对模型性能影响

Fig. 4 Impact of different potential dimensions on model performance

4 结束语

本章提出了一种融合双向 GRU 和注意力机制

的混合推荐模型(HRM-BAM),通过深度学习技术有效整合评论文本和评分数据,提升推荐系统的性能和可解释性。首先,模型在文本处理模块中利用双向GRU提取上下文语义,并在特征提取层引入分层注意力机制(词级与全局),既关注关键词重要性,也结合用户-物品交互信息。随后,模型采用特征级联与卷积神经网络,将评分和评论特征深度融合,从而缓解数据稀疏、捕捉复杂特征交互,最终通过全连接层实现评分预测。实验结果表明,HRM-BAM在多个领域的数据集上均优于传统算法和现有深度学习模型,平均RMSE较基准降低2%以上,为结合文本语义信息的个性化推荐提供了新的技术思路。

参考文献

- [1] 杨中金. 基于用户行为及知识传播的图神经网络推荐系统[J]. 智能计算机与应用, 2024, 14(2): 35-40.
- [2] LI Y, LIU K, SATAPATHY R. Recent developments in recommender systems: A survey [J]. IEEE Computational Intelligence Magazine. Piscataway, NJ: IEEE, 2024, 19(2): 78-95.
- [3] DELDJOO Y, HE Z, MCAULEY J. A review of modern recommender systems using generative models (gen-recsys) [C]// Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. New York :ACM, 2024: 6448-6458.
- [4] UTA M, FELFERNIG A, LE V M. Knowledge - based recommender systems: Overview and research directions [J]. Frontiers in Big Data, 2024, 7: 1304439.
- [5] MCAULEY J, LESKOVEC J. Hidden factors and hidden topics: Understanding rating dimensions with review text [C]// Proceedings of the 7th ACM Conference on Recommender Systems. New York: ACM, 2013: 165-172.
- [6] WANG C, BLEI D M. Collaborative topic modeling for recommending scientific articles [C]// Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York :ACM, 2011: 448-456.
- [7] WANG H, WANG N, YEUNG D Y. Collaborative deep learning for recommender systems [C]// Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York :ACM, 2015: 1235-1244.
- [8] LING G, LYU M R, KING I. Ratings meet reviews, a combined approach to recommend [C]// Proceedings of the 8th ACM Conference on Recommender Systems. New York: ACM, 2014: 105-112.
- [9] CHENG Z, DING Y, ZHU L. Aspect-aware latent factor model: Rating prediction with ratings and reviews [C]// Proceedings of the 2018 World Wide Web Conference. New York : ACM, 2018: 639-648.
- [10] ZHENG L, NOROOZI V, YU P S. Joint deep modeling of users and items using reviews for recommendation [C]// Proceedings of the 10th ACM International Conference on Web Search and Data Mining. New York :ACM, 2017: 425-434.
- [11] JAIN A, JAIN G, NAGAR S. Ratingdistribution - aware deep cognitive convolution matrix factorization for recommendation systems [J]. Arabian Journal for Science and Engineering, 2024, 53(10): 1-22.
- [12] SALAKHUTDINOV R R, MNH A. Probabilistic matrix factorization [C]// Proceedings of Advances in Neural Information Processing Systems. NeurIPS, 2007: 1257-1264.
- [13] WU Y, DUBOIS C, ZHENG A X. Collaborative denoising auto-encoders for top-n recommender systems [C]// Proceedings of the 9th ACM International Conference on Web Search and Data Mining. New York :ACM, 2016: 153-162.
- [14] WANG H, SHI X, YEUNG D Y. Relational stacked denoising autoencoder for tag recommendation [C]// Proceedings of AAAI Conference on Artificial Intelligence. AAAI, 2015: 3052-3058.
- [15] BANSAL T, BELANGER D, MCCALLUM A. Ask thegru: Multi - task learning for deep text recommendations [C]// Proceedings of the 10th ACM Conference on Recommender Systems. New York :ACM, 2016: 107-114.
- [16] KIM D, PARK C, OH J. Convolutional matrix factorization for document context-aware recommendation [C]// Proceedings of the 10th ACM Conference on Recommender Systems. New York : ACM, 2016: 233-240.
- [17] SEO S, HUANG J, YANG H. Interpretable convolutional neural networks with dual local and global attention for review rating prediction [C]// Proceedings of the 11th ACM Conference on Recommender Systems. New York: ACM, 2017: 297-305.
- [18] HUANG H, ZHANG Q, HUANG X. Mention recommendation for Twitter with end-to-end memory network [C]// Proceedings of the International Joint Conference on Artificial Intelligence. IJCAI, 2017: 1872-1878.
- [19] CHEN J, ZHANG H, HE X. Attentive collaborative filtering: Multimedia recommendation with item - and component - level attention [C]// Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval. New York :ACM, 2017: 335-344.
- [20] FANG X, WANG J, SENG D. Recommendation algorithm combining ratings and comments [J]. Alexandria Engineering Journal, 2021, 60(6): 5009-5018.
- [21] HODSON T O. Root Mean Square Error (RMSE) or Mean Absolute Error (MAE): When to use them or not [J]. Geoscientific Model Development Discussions, 2022, 2022: 1-10.
- [22] YANG Z, LIU J, LI T. A two-tiershared embedding method for review-based recommender systems [C]// Proceedings of the 32nd ACM International Conference on Information and Knowledge Management. New York :ACM, 2023: 2928-2938.