



计算机科学

Computer Science

ISSN 1002-137X, CN 50-1075/TP

《计算机科学》网络首发论文

题目：融合词间句间多关系建模的评论推荐算法
作者：邓策渝，李段腾川，胡奕仁，王晓光，李志飞
网络首发日期：2024-10-30
引用格式：邓策渝，李段腾川，胡奕仁，王晓光，李志飞. 融合词间句间多关系建模的评论推荐算法[J/OL]. 计算机科学.
<https://link.cnki.net/urlid/50.1075.tp.20241030.0915.004>



网络首发：在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

出版确认：纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

融合词间句间多关系建模的评论推荐算法

邓策渝¹ 李段腾川^{2,*} 胡奕仁² 王晓光¹ 李志飞³

¹ 武汉大学信息管理学院 武汉 430072

² 武汉大学计算机学院 武汉 430072

³ 湖北大学计算机与信息工程学院 武汉 430062

(dtelee1222@whu.edu.cn)

摘要 评论作为最常见的辅助信息，能够直观地表达用户的喜好和物品的属性，被研究者们广泛运用以不断优化推荐算法的预测精度。然而，目前评论推荐算法仍存在不足，主要体现在现有模型忽略了对评论文本多粒度特征的建模和对用户偏好、物品属性这一对异质特征的关联性交互学习，导致模型无法充分提取评论信息，影响了模型精度。因此，本文提出了融合词间句间多关系建模的评论推荐算法（MR4R），通过分析评论文本的词间重要性关系和句间时序动态关系，抽取不同层次的特征信息；并设计了融合预测层，对用户偏好与物品属性特征间的关联性挖掘过程进行优化并通过高阶非线性计算进行评分预测。研究挑选了四个不同场景的数据集，并将本文提出的模型与目前主流的7个推荐算法进行比较。实验结果表明，融合词间句间多关系建模的推荐算法能够充分提取评论中蕴含的信息，显著提升了平均推荐精度，具有更强的推荐性能。

关键词： 评论推荐算法；深度学习；多关系建模；时序特征建模；异质特征融合与交互

中图法分类号 TP391.3

Joint Inter-word and Inter-sentence Multi-relationship Modeling for Review-based Recommendation Algorithm

DENG Ceyu¹, LI Duantengchuan^{2,*}, HU Yiren², WANG Xiaoguang¹ and LI Zhifei³

¹ School of Information Management, Wuhan University, Wuhan 430072, China

² School of Computer Science, Wuhan University, Wuhan 430072, China

³ School of Computer and Information Engineering, Hubei University, Wuhan 430062, China

Abstract Reviews, a prevalent form of auxiliary information, directly reflect user preferences and item characteristics, extensively utilized by researchers to enhance the predictive accuracy of recommendation algorithms. However, the current review recommendation algorithm still has shortcomings, which are mainly reflected in the fact that the existing model ignores the modeling of multi-granularity feature extraction of review text and the relational interactive learning of user preferences and item attributes, a pair of heterogeneous features. This oversight leads to insufficient extraction of review information, compromising model accuracy. Thus, Joint Inter-word and Inter-sentence Multi-relationship Modeling for Review-based Recommendation Algorithm (MR4R) is introduced in the study. Firstly, multi-relational modeling strategy is adopted to analyze inter-word and inter-sentence relationships in review texts to extract layered feature information, thereby enriching the model's grasp of user preferences and refining item attribute representations. The fusion and prediction layer is designed to optimize the correlation mining process between user preferences and item attributes, and the score prediction is carried out by high order nonlinear calculation. The proposed model is compared with seven current mainstream recommendation algorithms on four distinct datasets. The results demonstrate that the recommendation algorithm, which incorporates

基金项目：国家社会科学基金重大项目(21&ZD334)；国家自然科学基金(62207011)

This work is supported by Major Project of the National Social Science Fund (21&ZD334) and National Natural Science Foundation of China (62207011).

通信作者：李段腾川 (dtelee1222@whu.edu.cn)

multi-relational modeling between words and sentences, effectively extracts information embedded in reviews, significantly enhancing average recommendation accuracy and exhibiting superior performance.

Keywords Review-based recommendation system; Deep learning; Multi-relation modeling model; Temporal feature modeling; Heterogeneous feature fusion and interaction

1 引言

随着时代的发展和信息技术的普及, 互联网蕴含的信息及数据呈爆炸式增长, 为社会和企业带来巨大效益的同时, 也带来了“信息过载”、难以精准搜索等挑战。推荐系统通过构建用户画像和分析物品属性寻找用户和物品的关联特征, 可以有效对信息进行过滤和筛选, 在电子商务^[1]、社交媒体^[2]、新闻推荐^[3]、休闲娱乐^[4]、工业应用^[5]等领域均取得了广泛的应用。

传统的推荐算法主要包括协同过滤 (Collaborative Filtering, CF)^[6]和矩阵分解 (Matrix Factorization, MF)^[7]两种方法。这两种方法通常使用单一交互信息, 即用户对物品的评分信息, 实现基于用户历史行为记录或用户-物品相似关系的推荐算法。近些年来, 许多学者在用户-物品交互矩阵的基础上, 引入额外信息来增强模型的特征表达能力, 例如: 物品标签、物品描述文字、用户评论等内容, 这一类推荐模型被定义为内容增强推荐系统^[8]。其中, 用户评论不仅能直观地诠释用户评分的原因, 表达用户对物品最直接的看法, 而且可以反映用户的喜好信息和物品的特征信息。在大量工作中已验证其能有效缓解数据稀疏、长尾效应和冷启动问题^{[9][10]}。

近十年来, 随着深度学习的发展, 涌现了许多用于文本上下文感知的神经网络^{[11][12]}。这些深度学习网络具备强大的特征学习和拟合能力, 能够从文本信息中建模复杂的用户偏好和物品属性特征, 进一步推进了评论推荐系统的发展。然而, 目前基于评论的推荐算法仍存在对用户和物品特征建模的能力不足以及用户-物品特征交互不充分等缺陷, 具体问题如下: (1) 我们发现

现有大部分模型仅对词间或句间单方面特征进行建模, 造成了部分信息的损失。并且对于句间特征建模来说, 现有模型虽然能区分不同评论对于用户特征构建的重要性, 但是忽略了用户的喜好和物品所表现出的属性会随着时间的变化而呈现动态性^[13], 即评论之间的时序动态关系; (2) 现有模型将提取出的用户偏好和物品特征进行线性拼接, 然而这一对特征为异质特征^[14], 现有方法忽略了对两者之间关联性的建模, 可解释性较差。

针对以上问题我们提出了融合词间句间多关系建模的评论推荐算法 (Joint Inter-word and Inter-sentence Multi-Relation for Review-based Recommender, MR4R)。该模型通过嵌入层、多关系建模层和融合预测层建模评论特征并预测用户喜好。具体来说, 在嵌入层中 MR4R 将预处理后的评论文本转化为低维稠密向量的形式, 输入进多关系建模层; 随后, 通过词间和句间多角度地对评论文本进行建模, 从中抽取用户倾向和物品属性特征; 最后, MR4R 使用基于多头注意力的融合层对异构特征进行交互并使用基于注意力的因式分解机 (Attentional Factorization Machines, AFM)^[15]对融合特征进行多阶计算并进行评分预测。本文的贡献可总结为以下几点:

(1) MR4R 采用多关系建模方法, 同时考虑评论文本中不同粒度的特征, 即通过融合词间特征和句间多关系特征, 丰富用户喜好和项目属性的特征表示。

(2) MR4R 引入评论的时序动态关系的概念作为建模评论句子之间关系的依据, 通过双向长短期网络 (Bi-directional Long Short-Term

Memory, Bi-LSTM)^[16]和局部稀疏注意力机制^{[17][18]}, 以及评论与评论之间在时间序列上的相对距离, 构建基于时序动态关系的评论句间关系特征。

(3) MR4R 对用户偏好特征和物品属性特征这一对异质特征的关联性挖掘进行优化, 即通过对两者特征进行多头注意力融合以及高阶非线性交互, 进一步增强了模型的预测能力。

2 相关工作

2.1 基于深度学习的推荐系统

近年来, 深度学习已经应用于各个领域并取得了显著成效。学者在传统推荐算法的基础上引入神经网络, 实现对大量用户-物品复杂关系的特征表示, 提升了模型的预测效果, 例如 Wide & Deep^[19]和 Deep & Cross^[20]通过对历史交互信息进行建模, 增强模型对历史数据的记忆能力和对不活跃用户或稀有物品的泛化能力。DeepFM^[21]与 FNN^[22]利用因式分解机 (Factorization Machines, FM)^[23]和深度神经网络分别建模低维和高维特征, 提升了 FM 的学习速度和表达能力。然而, 这些方法并没有从根本上解决推荐系统冷启动、可解释性差等问题。

近几年的研究中, 学者们通过深度学习进一步探究新形态的推荐系统, 有效缓解了上述问题。GraphMLP-Mixer^[24]是一种序列推荐网络, 通过构造全局物品图来增强模型对序列间协同信号的建模, 并利用图-感知机混合器模型对用户兴趣进行充分挖掘。对话推荐系统通过多轮的交互实现对用户偏好的捕捉, UCCR^[25]以用户为中心, 强调历史对话和当前对话的相似程度, 通过历史和当前信息建模用户偏好; UniMIND^[26]设计了多目标对话系统, 通过用户的多个偏好及依赖关系对用户特征进行建模。在社交推荐系统中, SMIN^[27]提炼了来自用户和项目两方面的关系结构, 以聚合来自不同类型的用户和物品的依赖关系, 构建了自监督的社交推荐系统。

以上基于深度学习的模型均取得了较好的推荐效果, 证明了将推荐系统与深度学习相结合的

可行性。

2.2 基于评论的推荐系统

随着深度学习在自然语言处理任务中取得巨大成功, 许多学者将评论引入推荐系统, 有效缓解了推荐系统的冷启动, 可解释性差等问题^[8], 通过深度学习模型提取评论文本中反映出的用户行为和物品属性信息。一方面利用主题模型从评论文本中抽取主题特征, 再使用 CF 或 MF 进行预测, 例如 HFT^[28]、RMR^[29]、TopicMF^[30]等方法使用该思路提升了推荐系统的准确度。另一方面, 学者们将提取出的评论上下文特征作为用户和物品特征进行预测, 如 DeepCoNN^[31]、ConvMF^[32]利用卷积神经网络 (Convolutional Neural Network, CNN) 从评论以及评分中中提取用户倾向和物品属性特征, 有效解决了数据稀疏的问题。

注意力机制^[33]的提出进一步增加了模型对文本的分析处理能力, 于是推荐领域的学者开始将注意力机制应用于基于评论的推荐系统。NARRE^[34]、TARMF^[35]在提取评论特征后, 进一步使用注意力机制学习评论对用户-物品的重要性; MPCN^[36]仅仅使用相似度矩阵计算注意力权值, 进而挑选出对用户-物品最重要的文本特征信息。CARCA^[37]通过多头自我关注块来捕捉用户文档在上下文特征和物品属性方面的动态特性, 进而通过特征和目标项目之间的交叉注意力来预测物品的分数。

部分学者在利用用户评论进行预测的同时, 通过额外工作进一步提升了推荐系统的可解释性。Sun 等^[38]将评论推荐任务和评论生成任务相结合进行推荐, 通过生成可解释性评论进一步获取用户信任, 提高系统准确度。RPRM^[39]通过分析评论属性和评论有用性之间的关联, 提升推荐系统捕获用户偏好和物品特征的能力。RGCL^[40]构建了一个以评论为基础的评论感知用户-项目图模型, 以辅助用户和项目的表示学习。

可见, 目前基于评论的推荐算法一定程度上缓解了由于数据稀疏性大、可解释性差、冷启动

等问题对推荐精度带来的影响。基于此, MR4R 通过多关系建模框架捕获评论中反映的用户偏好和项目属性, 该方法旨在对评论词语之间的重要性和评论之间的时序动态关系进行建模以缓解用

户评论中存在的冗余或噪声因素对模型造成的潜在负面影响。实验结果证明, MR4R 比 SOTA 取得了更高的推荐精度。

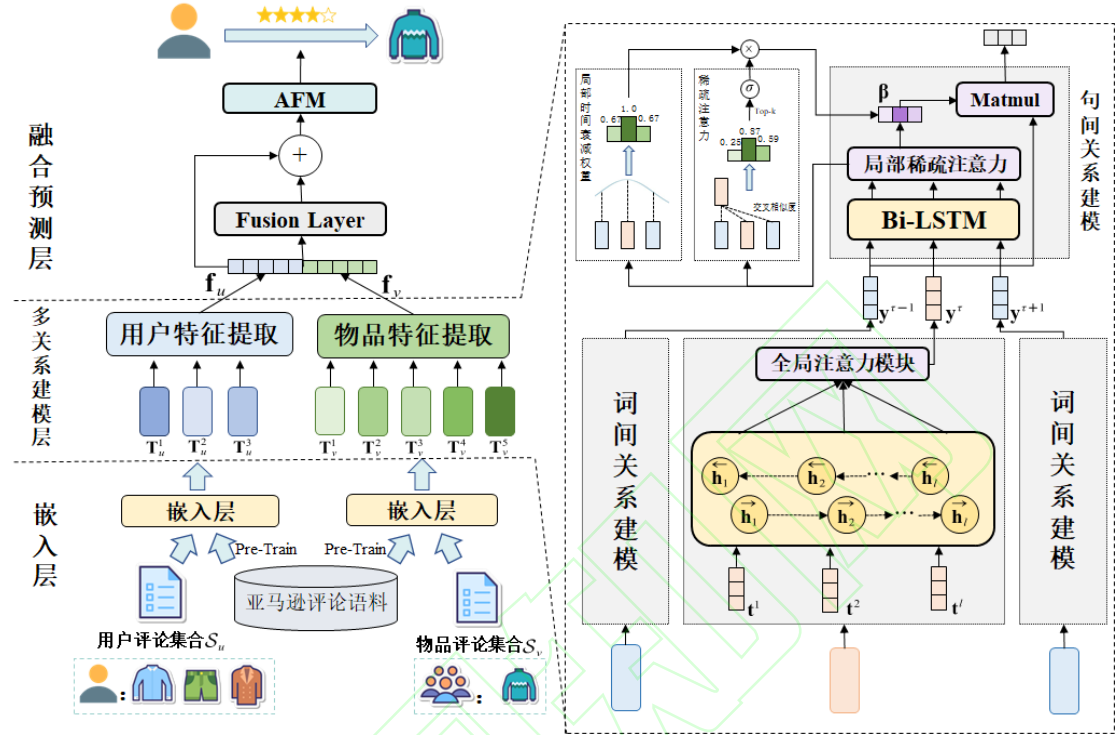


图 1 MR4R 框架流程图

Fig. 1 MR4R framework flowchart

3 融合词间句间多关系建模的评论推荐算法

本文提出了 MR4R 模型, 其整体结构如图 1 所示, 具体包含评论数据预处理模块、词间-句间多关系建模模块以及融合预测模块。首先, 本文将评论长度标准化后按照其发布的时间顺序进行排序, 得到每个用户和每个物品的评论集合。随后, 为了解决现有模型对用户和物品特征表达能力不足的问题, 本文设计了词间-句间多关系建模模块。具体来说, 对于词间关系建模, 该方法采用 Bi-LSTM 来提取评论上下文特征, 并通过全局注意力机制学习不同词语对评论文本的重要性, 并将其融合建模为每条评论的句子向量; 在句间关系建模方面, 本文使用另一个 Bi-LSTM 模型来捕获评论之间的时序关系, 并结合局部稀疏注意力机制对评论语义和相对位置关系

进行建模, 得到用户和物品的特征向量。最后, 为弥补现有模型中对用户特征和物品特征交互不足的问题, 本文设计了基于多头注意力的融合层实现用户和物品特征的融合, 并使用 AFM 预测模块对用户-物品特征向量进行高阶非线性交互计算预测评分, 以提高模型的性能。

3.1 评论数据预处理

在本文中, 我们选取了来自 Amazon 数据集的评论数据作为提取用户和物品特征的数据源。具体而言, 用户特征 f_u 用于刻画用户 u 在其所有评论中所反映的兴趣偏好和情感极性, 而物品特征 f_v 用于描述物品 v 在其对应的评论集合中所呈现的属性。为了对用户倾向和物品属性进行有效建模, 本文对评论文本进行适当的预处理, 将其转化为低维度且密集的词向量矩阵。

首先, 本文使用累计分布函数 (Cumulative Distribution Function, CDF) 缓解评论文本长度

极端值对平均评论长度的影响, 将统一长度记为 l 。具体地, 将评论文本长度为 l_i 对应评论数量记为 ζ_i , 则 l_i 的 CDF 值计算为:

$$\text{CDF}(l_i) = \frac{\sum_{j=0}^i \zeta_j}{N}, \quad (1)$$

其中, N 为所有评论的数目, 本文取 $\text{CDF}(l_i) = 0.6$ 时的评论文本长度 l_i 为统一长度 l , 并进行标准化处理 (对文本长度大于 l 的评论进行截断操作, 小于 l 的文本进行零值填充)。随后, 分别用 n 和 m 代表用户 u 和物品 v 在一段时间内发表的评论总量, 并按照发布时间进行排序, 其时序评论集合分别表示为 $S_u = \{s_u^1, s_u^2, \dots, s_u^r, \dots, s_u^n\}$ 和 $S_v = \{s_v^1, s_v^2, \dots, s_v^r, \dots, s_v^m\}$ 。其中 $s_u^r = s_v^r$, 表示用户 u 对物品 v 的评论, 该评论既被包含在用户的评论集合中也被包含在物品的评论集合中, 在 3.2.3 节作为局部注意力模块的注意力中心。

最后, 使用在 Amazon 数据集上预训练后的 Bert^[41] 语料库和分词器将评论文本进行分词处理, 转换为词语级别编码序列 S 。 S 经过嵌入层被映射为 d 维稠密矩阵 $\mathbf{T} \in \mathbb{R}^{l \times d}$ 。词向量矩阵初始化后模型不会对其进行训练, 而是为其赋予权重 $\omega \in \mathbb{R}^{l \times d}$, 表示每一个词向量的重要程度, 并在训练时更新。词向量矩阵的完整形式如下:

$$\mathbf{T} = \{\omega_0 \mathbf{t}_0, \omega_1 \mathbf{t}_1, \dots, \omega_l \mathbf{t}_l\}, \quad (2)$$

其中, \mathbf{t}_i 和 ω_i 分别代表第 i 个词的词向量和词向量权重。该模块将目标用户的所有评论和其他用户针对目标物品的评论进行处理, 分别转化为两个低维稠密向量表示, 作为用户和物品词间-句间多关系建模模块的输入。

3.2 词间句间多关系建模模块

词间-句间多关系建模模块的目的是从词间和句间两个不同的角度构建用户和物品评论信息的特征表示。首先通过对词语之间关系进行分析, 得到每一个句子内部词语级别重要性特征; 再引入局部稀疏注意力机制分析不同句子之间的时序和语义关联性, 最终得到关于用户行为和物品属性的隐向量表达。

3.2.1 词间关系建模

本文采用了 Bi-LSTM 模型和注意力机制来建模词语之间的双向依赖关系以及词语中的关键语义信息。以用户特征 \mathbf{r}_u 的构建步骤为例, 我们将经过嵌入层的第 i 条评论文本 s_u^i 的嵌入矩阵表示为 $\mathbf{T}_u^i \in \mathbb{R}^{l \times d}$, 则该评论文本中第 j 个词语对应的词向量可表示为 $\mathbf{t}_u^j \in \mathbb{R}^{1 \times d}$ 。将单词所处的不同位置视作不同的输入时间步, 即将 \mathbf{t}_u^j 作为第 j 个 Bi-LSTM 模块的输入, 得到的输出 $\mathbf{z}_u^j \in \mathbb{R}^{1 \times d}$ 用公式表示为:

$$\mathbf{z}_u^j = [\overrightarrow{\phi}_{LSTM}(\mathbf{t}_u^j, \overrightarrow{\mathbf{h}}_{t-1}), \overleftarrow{\phi}_{LSTM}(\mathbf{t}_u^j, \overleftarrow{\mathbf{h}}_{t-1})] \mathbf{W}_{xz} + \mathbf{b}_z, \quad (3)$$

其中, $\overrightarrow{\phi}_{LSTM}$ 和 $\overleftarrow{\phi}_{LSTM}$ 表示两个独立相反的 LSTM 网络, $\overrightarrow{\mathbf{h}}_t, \overleftarrow{\mathbf{h}}_t \in \mathbb{R}^{1 \times d}$ 分别代表正向和逆向隐藏向量, $\mathbf{W}_{xz} \in \mathbb{R}^{2d \times d}$ 为权重参数, $\mathbf{b}_z \in \mathbb{R}^{1 \times d}$ 为偏置参数。并将所有时间步的输出拼接起来作为第 i 条评论文本经过 Bi-LSTM 输出的句矩阵:

$$\mathbf{Z}_u^i = [\dots, \mathbf{z}_u^{i-1}, \mathbf{z}_u^i, \mathbf{z}_u^{i+1}, \dots], \quad j = (1, 2, \dots, l), \quad (4)$$

随后, 我们引入词汇级别的注意力机制, 目的是动态地为每个词汇分配权重, 以突出在给定评论上下文中更具信息量的词汇, 第 j 个词向量对应的注意力权重 $\alpha_u^{i,j}$ 的计算公式如下:

$$\alpha_u^{i,j} = \sigma(\omega_{az} [\mathbf{Z}_u^i]^\top + b_{az}), \quad (5)$$

其中, $\omega_{az} \in \mathbb{R}^{1 \times d}$, b_{az} 为线性层参数, $[\mathbf{Z}_u^i]^\top \in \mathbb{R}^{d \times l}$ 为 \mathbf{Z}_u^i 的转置。将 l 个权重的线性拼接后, 与句矩阵做内积得到最终词语级别的向量表达, 如公式所示:

$$\alpha_u^i = [\dots, \alpha_u^{i,j-1}, \alpha_u^{i,j}, \alpha_u^{i,j+1}, \dots], \quad (6)$$

$$\mathbf{g}_u^i = \alpha_u^i \mathbf{Z}_u^i, \quad (7)$$

其中, $\mathbf{g}_u^i \in \mathbb{R}^{1 \times d}$ 为评论文本 s_u^i 的词语级别向量表达, 作为其句向量。同理, 将物品 v 评论集合中包含的评论经过词间关系模型建模, 得到其不同的评论句向量表示 $\mathbf{g}_v^i \in \mathbb{R}^{1 \times d}$ 。

3.2.2 句间关系建模

一个用户或者一个物品拥有多个评论, 包含了不同的用户倾向和物品属性的描述信息, 对于预测结果有着不同的重要性, 并且这些评论随时间的变化展现出明显的时序动态特性。对于用户来说, 其偏好会随时间的推移而变化; 对于物品来说, 不同时间段的评论可能针对其不同特性

进行着重描述。因此，本文通过 Bi-LSTM 网络捕获评论内容的时间序列变化，并使用局部稀疏注意力机制结合评论之间的相对位置对评论句间关系进行建模。

以用户 u 为例，其时序评论集合 $S_u = \{s_u^1, s_u^2, \dots, s_u^r, \dots, s_u^n\}$ 通过词间关系建模模块后得到了 n 个句向量 \mathbf{g}_u^i 。将其按照时间顺序输入不同的 Bi-LSTM 模块中，即句向量 \mathbf{g}_u^i 为第 i 个 Bi-LSTM 模块的输入，利用公式(3)得到目标时间步的输出 $\mathbf{y}_u^i \in \mathbb{R}^{1 \times d}$ 。将所有时间步的输出以公式(4)的形式拼接起来作为句子级别的特征矩阵 $\mathbf{Y}_u \in \mathbb{R}^{n \times d}$ ，输入进局部稀疏注意力层。

首先，计算带有高斯分布乘积项的局部权重。具体而言，本文假设评论之间存在时序动态关系，即与目标评论 s_u^i 在时间序列上距离越近的评论，与其关联性越强，其相对距离可表示为：

$$\text{dis}(i) = \frac{|\tau - i|}{n}, \quad (8)$$

特别地， $\text{dis}(\tau) = 0$ 代表目标评论的自身距离并将其作为高斯分布的均值，即 $\mu = 0$ ，将标准差设置为 $\sigma = 0.5$ 。因此， s_u^i 的局部权重 lo_i 可表示为：

$$lo_i = e^{-\frac{(\text{dis}(i) - \mu)^2}{2\sigma^2}}. \quad (9)$$

随后，计算语义相关的稀疏注意力机制。对 \mathbf{Y}_u 与 \mathbf{y}_u^i 进行相似度计算：

$$\mathbf{o}_u^i = \mathbf{y}_u^i [\mathbf{Y}_u]^T, \quad (10)$$

得到其他评论与目标评论的语义相似特征向量 $\mathbf{o}_u^i \in \mathbb{R}^{1 \times n}$ 。在 \mathbf{o}_u^i 中，按照特征值大小选取前 k 个评论，表示与 s_u^i 在语义上最相关的评论，并赋予 0-1 稀疏权重，如公式所示：

$$sp_i = \begin{cases} 1, & s_u^i \text{ in } \mathcal{K}. \\ 0, & \text{else.} \end{cases}, \quad (11)$$

其中， \mathcal{K} 表示前 k 个与 s_u^i 在语义上最相关评论的集合。

综上，局部注意力机制通过赋予不同时间片段不同的权重，帮助模型识别出哪些时间片段对于评论的整体情感或评价有更大的贡献；而稀疏注意力允许模型在处理大量评论时，聚焦于与当前评论相似性高或具有相关性的其他评论，从而

减少不必要的信息干扰。局部稀疏注意力权重用公式表示为：

$$\beta_u^i = \sigma \left(sp_i \frac{\mathbf{o}_u^i}{\sqrt{d}} \right) lo_i, \quad (12)$$

将拼接后的局部稀疏注意力权重与 \mathbf{Y}_u 做内积得到最终用户特征 $\mathbf{f}_u \in \mathbb{R}^{1 \times d}$ ，公式表示为：

$$\beta_u = [\dots, \beta_u^{i-1}, \beta_u^i, \beta_u^{i+1}, \dots], \quad (13)$$

$$\mathbf{f}_u = \beta_u \mathbf{Y}_u. \quad (14)$$

同理，将多个物品属性特征向量 \mathbf{g}_v^i 经过句间关系建模，得到特征最终的物品特征的向量表达 $\mathbf{f}_v \in \mathbb{R}^{1 \times d}$ 。

3.3 融合预测模块

在 3.2 节中得到的用户特征 \mathbf{f}_u 和物品特征 \mathbf{f}_v 来自于一对并行模型且数据源存在差异，因此， \mathbf{f}_u 和 \mathbf{f}_v 应当被视作一对异质特征。于是，本文设计了特征融合和预测两大模块。前者旨在使用多头注意力机制对用户特征 \mathbf{f}_u 和物品特征 \mathbf{f}_v 这一对异质特征进行融合，后者在融合结果的基础上进行高阶非线性计算，得到预测评分 \hat{r} 。

首先，融合模块旨在捕捉用户和物品之间的关联信息，即对来自不同源的用户特征和物品特征进行融合，生成带有两者关联信息的融合特征。具体来说，将 \mathbf{f}_u 和 \mathbf{f}_v 线性拼接成为 $\mathbf{f}_c \in \mathbb{R}^{1 \times 2d}$ ，输入进基于多头注意力的特征融合模块，进行多头注意力计算并进行残差和归一化处理，最后经过全连接层得到融合特征 \mathbf{f} ：

$$\mathbf{f}_M = \text{MultiHead}(\mathbf{f}_c, \mathbf{f}_c, \mathbf{f}_c) + \mathbf{f}_c, \quad (15)$$

$$\mathbf{f} = \left(\frac{\mathbf{f}_M - \mu_{f_M}}{\sigma_{f_M}^2} + \mathbf{f}_M \right) \mathbf{W}_f + \mathbf{b}_f, \quad (16)$$

其中， $\mathbf{W}_f \in \mathbb{R}^{2d \times 2d}$, $\mathbf{b}_f \in \mathbb{R}^{2d}$ 为全连接层参数， $\mu_{f_M}, \sigma_{f_M}^2$ 为 \mathbf{f}_M 的均值和方差。 $\mathbf{f}_M \in \mathbb{R}^{1 \times 2d}$ 为多头注意力机制的计算结果，计算过程如下：

$$\text{MultiHead}(\mathbf{q}, \mathbf{k}, \mathbf{v}) = [\text{head}_1, \text{head}_2, \dots, \text{head}_k], \quad (17)$$

$$\text{head}_i = \text{Softmax} \left(\frac{\mathbf{q}_i \mathbf{k}_i^T}{\sqrt{d_k}} \right) \mathbf{v}_i, \quad (18)$$

其中， $\mathbf{q}, \mathbf{k}, \mathbf{v}$ 为多头注意力机制的输入， $\mathbf{q}_i, \mathbf{k}_i, \mathbf{v}_i \in \mathbb{R}^{1 \times d_k}$ 为 $\mathbf{q}, \mathbf{k}, \mathbf{v}$ 被线性层映射到第 i 个低维空间的子向量。

MR4R 在预测模块中使用了基于注意力机制的因式分解机 (Attentional Factorization Machine, AFM)，是在传统因式分解机

(Factorization Machine, FM) 的基础上引入了注意力机制, 使其能够对不同特征交互的重要性进行加权, 从而增强模型对异质特征之间复杂关系的学习能力。具体来说, 首先将 \mathbf{f} 输入进全连接层, 得到 AFM 的一阶线性特征; 引入辅助向量 $\mathbf{V} \in \mathbb{R}^{2d \times p}$, 为关联特征中的每一个元素设置维度为 p 的可学习参数, 以解决由于数据稀疏导致模型学习不充分的问题, 将辅助向量 \mathbf{V} 与融合特征 \mathbf{f} 同时输入 Pari-wise 交互层得到 AFM 的二阶交互特征, 再利用注意力机制对二阶特征进行优化, 学习每一部分特征对预测结果的重要性。在评分预测阶段, 该模块对高阶非线性深度交互的计算结果, 即对一阶线性特征和基于注意力机制的二阶交互特征分别求和, 再将求和结果线性相加得到预测评分值 \hat{r} 。公式表示为:

$$\hat{r} = \omega_0 + \sum_{i=1}^{2d} \omega_i \mathbf{f}^i + \varpi^T \sum_{j=1}^{2d} \sum_{k=j+1}^{2d} a_{jk} (\mathbf{V}^j \odot \mathbf{V}^k) \mathbf{f}^j \mathbf{f}^k, (19)$$

其中, $\omega, \varpi \in \mathbb{R}^{1 \times 2d}$ 分别为线性层 (一阶交互层) 和二阶交互层的权重, $\mathbf{V}^j \in \mathbb{R}^{1 \times p}$ 表示第 j

个特征值的辅助向量, a_{jk} 表示融合特征 \mathbf{f} 第 j 个和第 k 个特征值的注意力权重。

4 实验

4.1 数据集

本文使用来自 Amazon 四个公开数据集对 MR4R 进行实验, 分别是汽车用品 (Automobile)、乐器 (Music Instrument)、即时视频 (Instant Video) 和玩具 (Toys and Games)。其中, 乐器数据集评论数量少, 评论长度短, 被许多研究应用于评估模型推荐准确性^{[31][36][42]}。汽车用品评论数量多且稀疏性高, 而即时视频和玩具数据集每个用户的平均评论数多且平均评论长度长, 可以用于评估模型在复杂情况下的性能^[36]。四个数据集评论语言均为英文, 评分为[1,5]范围内的整数, 其具体信息如表 1 所示。其中, 评论统一长度取 CDF 值为 0.6 时的评论长度, 四个数据集的 CDF 图像如图 2 所示。

表 1 亚马逊数据集数据统计
Table 1 Data statistics of Amazon dataset

数据集	#用户	#物品	#评论	#用户平均评论	评论统一长度	稀疏度
Automobile	2928	1835	20473	7.0	349	99.62%
Music Instrument	1429	900	10261	7.2	362	99.20%
Instant Video	690	903	8685	12.6	479	98.61%
Toys and Games	921	871	18637	20.2	917	97.68%

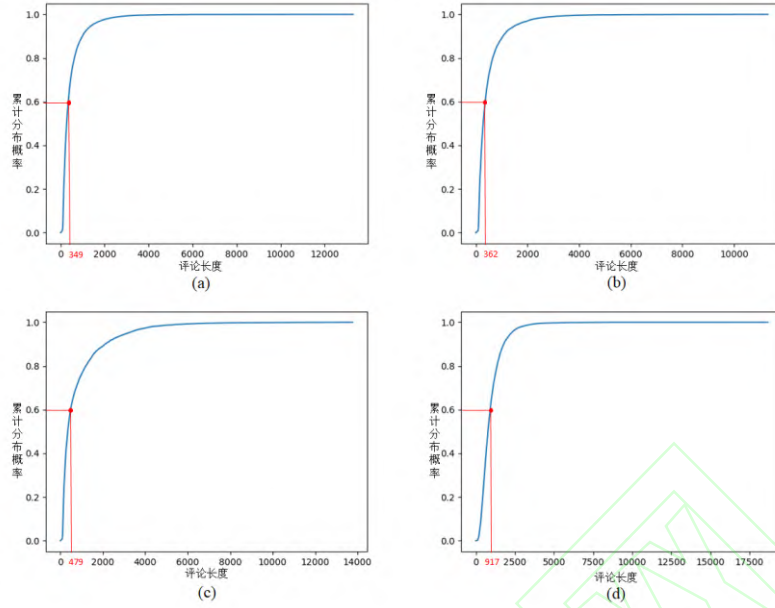


图2 汽车 (a)、乐器 (b)、即时视频 (c)、玩具 (d) 数据集的累计分布函数图像

Fig. 2 Cumulative distribution function images of car (a), musical instrument (b), instant video (c), and toy (d) datasets

4.2 基线方法

为了验证所提出模型的有效性, 将 MR4R 与以下基线方法进行对比:

PMF^[43]: 是最为经典的基于评级进行预测的矩阵分解模型。

ConvMF^[32]: 该方法结合了卷积神经网络 (CNN) 和概率矩阵分解模型 (PMF), 同时使用文本深度信息和评分概率信息预测用户喜好。

DeepCoNN^[31]: 基于两个并行 CNN 分别从用户评论和物品评论中学习用户和物品的潜在特征向量, 并使用 FM 进行评分预测。

NARRE^[34]: 利用两个并行神经网络和注意机制来学习评论的潜在特征, 然后将评论和项目结合起来预测评分。

MPCN^[36]: 通过指针机制实现评论之间的直接匹配, 并利用单词级注意力交互来学习固定的维度表示。并引入一个多指针学习方案, 提取用户和项目评论之间的多个层次交互, 完成预测。

CARL^[42]: 利用 CNN 来学习用户-物品对的相关特征, 利用动态线性融合机制来获得最终评分。

DAML^[44]: 由两个并行的 CNN 组成, 从评论中学习用户行为和项目属性, 并利用神经因子分解机网络融合用户和项目特征并预测最终评分。

MR4R: 本文提出的模型。使用多关系建模模块提取用户和物品特征, 随后利用融合预测层完成对两个异质特征的深层交互, 并进行非线性高阶评分预测计算。

4.3 评估指标

为了评估模型的有效性, 本文利用 RMSE (Root Mean Square Error) 和 MAE (Mean Absolute Error) 作为评估指标, 将预测与真实值进行根均平方差运算比较 MR4R 与现有模型的预测效果, 相关公式如下:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{r}_i - r_i)^2}, \quad (20)$$

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |\hat{r}_i - r_i|, \quad (21)$$

其中, \hat{r}_i 为第 i 个数据对应的预测评分, r_i 为真实评分, N 为数据集的大小。

4.4 实验设置

文本预处理。(1) 将累计分布值为 0.6 时对应的评论长度作为所有数据的统一长度 l , 对

于文本长度小于 l 的评论, 用[PAD]填充至统一长度; 对于文本长度大于 l 的评论, 进行截断处理, 将该评论长度缩短至 l 。(2) 使用 Bert-Amazon 词汇表作为参考, 将评论转化为单词级别的数值序列。(3) 将数据进行随机排序后, 按照 8:2 的比例划分训练集和测试集。

训练超参数设置。首先, 本文选用 Adam 算法作为优化器。为了进一步优化训练过程, 我们使用余弦退火策略以动态调整学习率, 并将初始学习率设置为 1×10^{-4} , 使其随着训练周期的进展而进行细微调整。公式如下:

$$\eta_t = \eta_{\min} + \frac{1}{2}(\eta_{\max} - \eta_{\min}) \left(1 + \cos\left(\frac{T_{\text{cur}}}{T_{\max}} \pi\right) \right), \quad (22)$$

其中, η 代表学习率, T_{cur} 为当前训练的轮次, T_{\max} 为迭代的半周期。为了使学习率周期在每一个轮次都周期变化两次, T_{\max} 需要满足:

$$T_{\max} = \left\lceil \frac{N}{4b} \right\rceil, \quad (23)$$

其中, N 为数据集长度, b 为单次训练的样本个数, 本文设置为 $b = 16$ 。此外, 本文将权重衰减系数设定为 0.001, Dropout 概率设置为 0.5。在训练策略上, 本文采用基于性能的动态停止策略, 即本文设定的停止阈值为 1×10^{-4} , 当连续两个训练周期的损失函数差值小于停止阈值时, 就可认为模型已训练完成。最后, 本文对每个模型均执行了三次以上的训练, 并将最小的

均方根误差 (RMSE) 分数和平均绝对误差 (MAE) 分数上作为模型的最优分数。

模型特征参数设置。对于基线模型, 本文选取其中关键超参数进行实验, 其余特征参数与原文中的最优参数保持一致。参数细节如表 2 所示。

表 2 模型的参数细节

Table 2 Parameter details of models		
基线模型	参数名称	参数值
PMF	特征值维度	20,30,40
ConvMF	共享窗口大小	50,100,200
DeepCoNN	特征值维度	20,50,80
NARRE	特征值维度	16,32,64
MPCN	指针数目	1,2,5
CARL	特征值维度	15,25,50
DAML	特征值维度	8,16,64
MR4R	特征值维度	64,128,256
MR4R	稀疏注意力 k 值	2,3,5

对于 MR4R, 词间关系模块采用的局部注意力中高斯分布乘积项的均值 $\mu = 0$, 标准差 $\sigma = 0.5$; 在基于多头注意力的融合模块中, 多头注意力头数设置为 4; 基于 AFM 的预测模块中的隐含因子矩阵维度设置为 20。Bi-LSTM 的特征参数在 {64,128,256} 中进行优化, 稀疏注意力 k 值在 {2,3,5} 中进行优化。此外, 我们在后文针对 MR4R 在不同数据集上进行了参数实验以选取最适合的参数。

表 3 不同模型在四个数据集上的对比结果

Table 3 Comparison results of different models on four datasets

	Automobile		Music Instr.		Instant Video		Toys & Games	
	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
PMF	3.1213	2.4787	3.6544	2.8871	3.7286	2.7313	3.7880	3.0181
ConvMF	<u>0.9010</u>	<u>0.5336</u>	0.8709	<u>0.5145</u>	1.0039	0.7135	0.8081	0.6145
DeepCoNN	0.9036	0.5587	0.8650	0.5521	<u>0.9884</u>	0.7098	0.8046	0.6214
NARRE	0.9264	0.5559	0.8855	0.5584	1.0007	<u>0.7044</u>	<u>0.8027</u>	0.6116
MPCN	0.9204	0.5511	<u>0.8647</u>	0.5393	1.0005	0.7123	0.8179	0.6330
CARL	0.9423	0.5386	0.9292	0.5786	1.0154	0.7275	0.8204	0.6248
DAML	0.9702	0.5788	0.9098	0.5303	1.0082	0.7373	0.8123	<u>0.6109</u>
MR4R	0.8665	0.5063	0.8465	0.4913	0.9370	0.6421	0.7783	0.5987
$\Delta_{\text{vs. SOTA}} \%$	3.83%	5.12%	2.10%	4.51%	5.20%	8.84%	3.04%	2.00%

4.5 对比实验

为了验证 MR4R 的有效性, 本节将其与基

线模型在四个亚马逊公开数据集上进行对比实验, 如表 3 所示。下划线部分表示最佳基线方法所取得的评估分数, 加粗部分表示 MR4R 取得改进后的 RMSE 和 MAE 分数, $\Delta_{vs. SOTA} \%$ 表示 MR4R 相较于最优基线的提升率。

根据实验结果可知 MR4R 的预测精度明显高于各个基线模型。其中 PMF 仅使用评分信息进行预测, 难以从高稀疏性的数据集中评估用户偏好。ConvMF 将 CNN 与 PMF 相结合, 为评论数据提供评分统计信息, 在汽车用品数据集上取得了较好的效果。DeepCoNN 是经典的双塔模型, 使用并行卷积神经网络同时从用户评论和物品评论中学习特征信息, 并通过因式分解机进行预测, 取得了基线模型中最优的平均 RMSE 分数。NARRE 能够学习不同评论的重要性, 进而筛选出代表性评论对用户偏好进行建模。虽然其在十万和百万的数据集上取得了优异的推荐效果, 但是在本文中四个小数量级数据集上的实验

结果较差, 可能的原因是 NARRE 难以从有限的数据中提取有用信息。MPCN 利用相似度矩阵对用户评论和物品评论进行匹配, 忽略了其他评论中的潜在信息导致预测效果不理想。CARL 与 DAML 注重用户评论和物品评论的交互学习, 两者结构相似, 然而 DAML 使用神经因式分解网络进行预测, 取得了更好的平均评估分数。

整体来说, 对比实验验证了 MR4R 在提升推荐精度方面的有效性, 原因可总结如下:

- (1) 本文对评论条目而非评论文档进行融合词间句句多关系建模, 允许 MR4R 在多粒度层面上从评论文本中建模用户和物品的特征信息;
- (2) 本文引入了评论间的时序动态关系表征, 并采用局部稀疏注意力机制整合句间特征, 从而更有效地捕捉用户偏好和物品属性的动态变化;
- (3) 所提出的融合预测模块能够深度学习用户-物品特征的关联并执行高阶交互计算, 使得模型能够准确地预测用户评分。

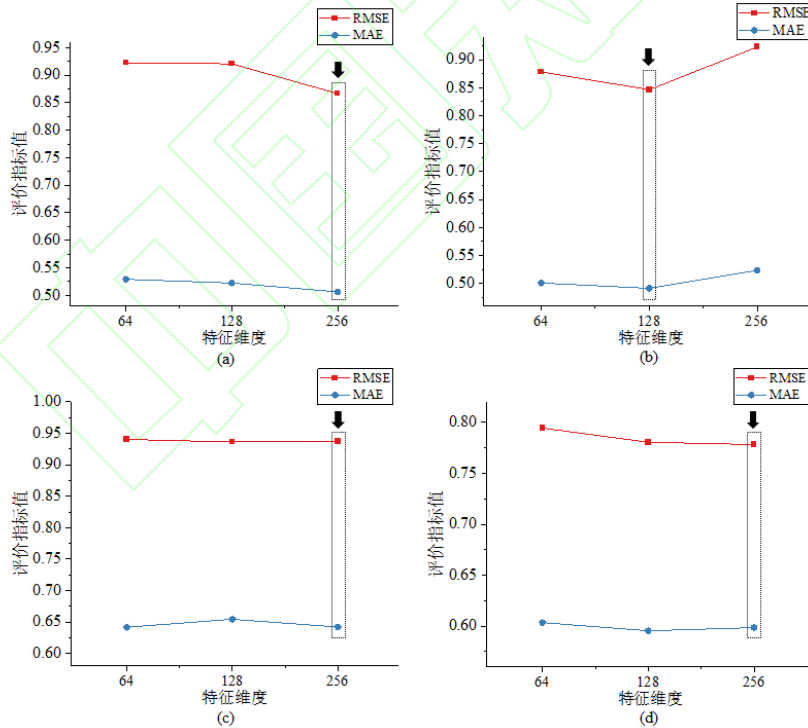


图 3 MR4R 精度在汽车用品 (a)、乐器 (b)、即时视频 (c)、玩具 (d) 数据集上随特征维度变化趋势图

Fig. 3: The trend of MR4R accuracy with respect to feature dimensions on automotive supplies (a), musical instruments (b), real-time videos (c), and toys (d) datasets

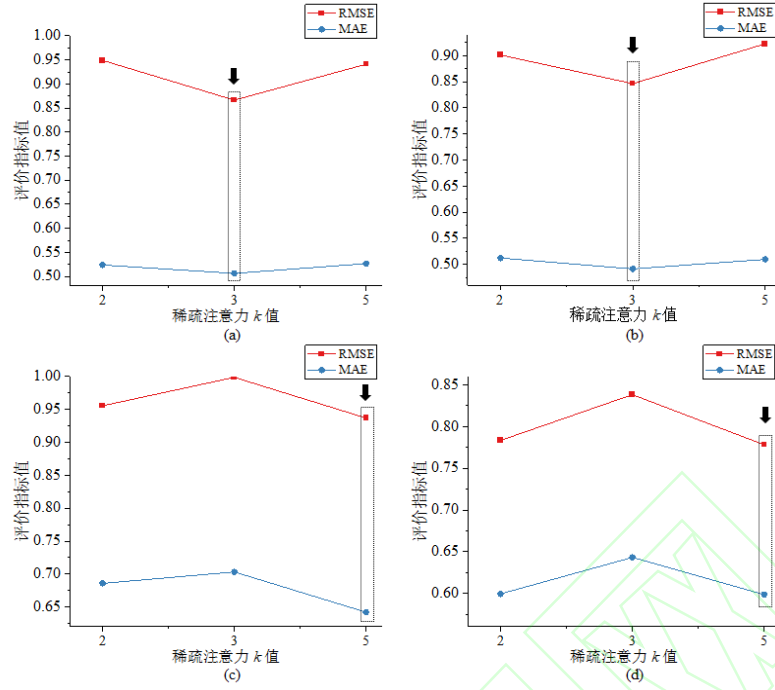


图 4 MR4R 精度在汽车用品 (a)、乐器 (b)、即时视频 (c)、玩具 (d) 数据集上随稀疏注意力 k 值变化趋势图

Fig. 4: Trends of MR4R accuracy with sparse attention k -values on automotive supplies (a), musical instruments (b), real-time videos (c), and toys (d) datasets

4.6 超参数分析

4.6.1 MR4R 特征维度的影响

在许多研究中, 特征维度被视为影响模型性能关键参数。本文通过调整 MR4R 的特征维度在四个数据集上进行超参数实验, 并对 RMSE 和 MAE 分别绘制了不同的曲线变化图, 其变化趋势如图 3 所示。

实验结果表明, 不同数据集上的最优参数不同。具体来说, 当特征维度为 128 时, MR4R 在乐器数据集上即可获得较好的预测效果, 然而当取值提升至 256 时, 模型的预测性能大幅下降, 可能是出现了过拟合情况导致对用户偏好的预测出现了较大误差。除此之外, MR4R 在其他三个数据集上的最优特征维度参数为 256, 可能的原因是三者均为复杂数据集: 汽车用品数据集稀疏度高、即时视频和玩具数据集评论数量多、长度长, 模型需要更多的特征元素来表征用户喜好和物品属性。

4.6.2 MR4R 稀疏注意力 k 值的影响

公式(11)中, 我们根据与目标评论的语义相关性挑选前 k 个最相关评论, 以保留能够体现用户目前喜好和物品属性的评论信息。图 4 展示了不同的稀疏注意力 k 值对于模型性能的影响。

结合本小节实验结果和 4.1 节中表 1 数据集信息可以看出, 稀疏注意力 k 值的最优取值很可能与数据集中每个用户的平均评论数有关。汽车用品和乐器数据集中每个用户平均拥有 7 个评论, 当 $k=3$ 时, 即选取三个最相关评论即可有针对性地建模该用户的喜好和目标物品的属性特征; 而当 $k=5$ 时会不可避免的引入冗余信息, 对 MR4R 的评分预测产生负面影响。对于即时视频和玩具数据集来说, k 值为 5 是最优稀疏注意力超参数取值, 可能的原因是两个数据集中每个用户都拥有上 10 个评论, 当保留的相关评论数为 5 时才能够较为全面地建模用户和物品特征。

4.7 消融分析

本文的主要创新点是词间-句间多关系建模模块和融合预测模块, 为了验证本文所提出的各

个模块对于整体推荐精度提升的有效性，本小节在 4.6 节所讨论的最优参数的基础上，对模型分别进行以下更改，构成 4 个变体：去除句间时序关系建模模块（without TimeRela，简称 w/o TimeRela），去除局部稀疏注意力机制并使用全局注意力机制代替（w/o LoSp），去除基于多头注意力机制的特征融合模块并使用线性连接代替（w/o MHA），去除基于 AFM 的预测层并使用哈达玛积（Hadamard）代替（w/o AFM）。

实验结果如表 4 所示。从实验结果可以看出，去除句间时序动态关系模块后（w/o TimeRela），模型性能受到了负面影响，证明了

通过评论之间时序关系建模用户倾向和物品属性的可行性，但下降幅度较小，在后续工作中可以针对此进行改进。去除局部稀疏注意力后（w/o LoSp），模型精度下降幅度达到了 5%，表明局部稀疏注意力相较全局注意力能够更有效地保留相关信息而过滤无关评论。分别去除融合模块（w/o MHA）和预测模块（w/o AFM）后的两个变体，与原模型 MR4R 相比，性能出现了显著下降，表明学习异质特征之间的关联信息的重要性以及对用户和物品特征进行高阶交互计算的必要性。

表 4 MR4R 及各个变体的对比结果

Table 4 Comparison results of MR4R and its variants

变体	Automobile		Music Instruments		Instant Video		Toys & Games		$\Delta_{avg} \%$	
	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
w/o TimeRela	0.9340	<u>0.5269</u>	<u>0.8846</u>	0.5174	0.9341	<u>0.6699</u>	0.7930	<u>0.6042</u>	3.43%	3.45%
w/o LoSp	0.9588	0.5401	0.8847	0.5167	0.9863	0.6915	0.7942	0.6077	5.40%	4.99%
w/o MHA	0.9453	0.5376	0.9198	0.5387	0.9824	0.6838	0.7933	0.6084	5.84%	5.49%
w/o AFM	<u>0.8897</u>	0.5682	0.903	0.5794	0.9421	0.6738	<u>0.7910</u>	0.6098	2.77%	7.93%
MR4R	0.8665	0.5063	0.8465	0.4913	<u>0.9370</u>	0.6432	0.7783	0.5987	\	\

4.8 实例分析

本节以即时视频数据集中某一用户的评论为例，分析局部稀疏注意力机制的权重参数，如表 5 所示。首先，我们观察到，局部注意力会为不同距离的评论赋予不同的局部权重 lo ，且距离目标评论（在该例子中为评论 2）越近的评论， lo 取值越大。其次， o_u^i 和 $sp@3$ 为稀疏注意力机制的参数，前者为不同评论与目标评论 s_u^i 的交互特征向量，表示语义相似程度；后者根据语义相似度大小选取了前 3 个语义最相关的评论。结合用户的具体评论来看，**加粗**的三条评论均体现了

用户“兴奋”的情绪。相较其他评论而言，评论 1 和评论 5 与目标评论的交互特征值更高，表明其语义更为相似，因此其稀疏值（ sp ）为 1。随后，根据公式(12)计算得到每条评论的局部稀疏注意力权重 β ，其中 β 值为非 0 对应的句子级别特征参与后续用户偏好表征的计算。因此，MR4R 使用局部稀疏注意力机制不仅能够挑选出时间关系上最为重要的评论，还能有效突显出用户关于特定方面的显著倾向，有助于构建包含丰富语义信息的特征向量。

表 5 局部稀疏注意力机制的权重参数实例

Table 5 Examples of weight parameters of local sparse attention mechanism

序号	用户评论	lo	o_u^i	$sp@3$	β
1	...I watch it whenever I want a great laugh!	0.6703	0.6784	1	0.3332
2	... it's not that hard to follow and it's exciting ...	1.0000	0.7284	1	0.3722
3	Great movie if you are into the Roman rule of Time...	0.6703	0.5296	0	0
4	...it would be a good series to watch...	0.4493	0.6432	0	0
5	...I am so excited that I can't wait to watch it more!	0.3011	0.6912	1	0.2947

5 结束语

为解决现有模型对用户和物品表征能力不足和交互能力不足的问题,本文提出了融合词间句间多关系建模的评论推荐算法。我们发现了评论文本之间的时序性,并将其用于建模用户和物品一段时间内喜好和属性特征。本文基于此提出词间-句间多关系建模方法,从不同的层面学习评论上下文特征作为用户和物品的特征表示。除此之外,我们验证了用户特征和物品特征为一对异质特征,对其进行融合交互能够增强模型对两者特征的关联性学习,并提升模型的预测能力。

我们认识到 Bi-LSTM 对于评论文本特征提取存在一定局限性,在后续研究中,我们将采用更有效的方法提取评论上下文特征,例如 Transformer、Bert 等模型。除此之外,我们还将进一步探索通过关键情感词或功能词建模用户和物品之间关联的方法,以进一步捕获用户细粒度喜好和物品细粒度属性。

参考文献

- [1] ALMAHMOOD R, TEKEREK A. Issues and solutions in deep learning-enabled recommendation systems within the e-commerce field[J]. *Applied Sciences*, 2022, 12(21): 11256.
- [2] ZHANG Y, YAMASAKI T. Style-aware image recommendation for social media marketing[C]// *Proceedings of the 29th ACM International Conference on Multimedia*. Virtual Event China: ACM, 2021: 3106–3114.
- [3] BAO J, ZHANG K, WU L, et al. Conformity-Aware Debaised Neural News Recommendation with Causal Reasoning [J/OL]. *Journal of Computers*, 1-20[2024-09-04].
鲍纪敏, 张琨, 吴乐, 等. 从众性感知的因果去偏新闻推荐方法[J]. *计算机学报*, 1-20[2024-09-04].
- [4] HUANG F. Personalized marketing recommendation system of new media short video based on deep neural network data fusion[J]. *Journal of Sensors*, 2021: 1–10.
- [5] HU H, TANG B, ZHANG Y, et al. Vehicular ad hoc network representation learning for recommendations in internet of things[J]. *IEEE Transactions on Industrial Informatics*, 2020, 16(4): 2583–2591.
- [6] MARLIN B. Modeling user rating profiles for collaborative filtering[C]// *In Proceedings of the 16th International Conference on Neural Information Processing Systems (NIPS'03)*. MIT Press, 2003: 627–634.
- [7] DUECK D, FREY B. Probabilistic sparse matrix factorization[J]. *Technical Report PSI TR*, 2004: 023.
- [8] WU L, HE X, WANG X, et al. A Survey on Accuracy-Oriented Neural Recommendation: From Collaborative Filtering to Information-Rich Recommendation[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2023, 35(3): 4425–4445.
- [9] PENG Y, MENG X, DU Z. Survey on Diversified Recommendation [J]. *Journal of Computer Research and Development*, 1-27[2024-09-04].
彭迎涛, 孟小峰, 杜治娟. 多样化推荐综述[J/OL]. *计算机研究与发展*, 1-27[2024-09-04].
- [10] ZHOU Y, CHU H, ZHU F, et al. A Survey on Deep Learning-based Personalized Learning Resource Recommendation[J/OL]. *Computer science*. 1-26[2024-09-04].
周洋涛, 褚华, 朱非非, 等. 基于深度学习的个性化学习资源推荐综述[J/OL]. *计算机科学*, 1-26[2024-09-04].
- [11] KIM Y. Convolutional Neural Networks for Sentence Classification[C]// *Empirical Methods in Natural Language Processing (EMNLP)*. Doha, Qatar: Association for Computational Linguistics, 2014: 1746–1751.
- [12] ZAREMBA W, SUTSKEVER I, VINYALS O. Recurrent Neural Network Regularization[J/OL]. *ArXiv*, abs/1409.2329, 2014.

- [13] WANG Z, SHAO Y, HE J, et al. Sequential Recommendation Based on Multi-space Attribute Information Fusion[J]. Computer Science, 2019, 51(03): 102-108.
- 王子泓, 邵莹侠, 何吉元, 等. 基于多空间属性信息融合的序列推荐[J]. 计算机科学, 2024, 51(03): 102-108.
- [14] ZHOU L, WANG J, WANG L, et al. Heterogeneous Information Network Representation Learning: A Survey. Journal of Computers, 2022, 45(01): 160-189.
- 周丽华, 王家龙, 王丽珍, 等. 异质信息网络表征学习综述[J]. 计算机学报, 2022, 45(01): 160-189.
- [15] XIAO J, YE H, HE X, et al. Attentional factorization machines: learning the weight of feature interactions via attention networks[C]//International Joint Conference on Artificial Intelligence (IJCAI'17). AAAI Press, 2017: 3119-3125.
- [16] HUANG Z, XU W, YU K. Bidirectional LSTM-CRF Models for Sequence Tagging[J/OL]. ArXiv, abs/1508.01991. 2015
- [17] LUONG T, PHAM H, MANNING C. D. Effective Approaches to Attention-based Neural Machine Translation[C]//In Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. 2015: 1412-1421.
- [18] CHEN W, XING X, XU X, et al. Key-Sparse Transformer for Multimodal Speech Emotion Recognition[C]//IEEE International Conference on Acoustics, Speech and Signal Processing. 2022: 6897-6901.
- [19] CHENG H, KOC L, HARMSSEN J, et al. Wide & deep learning for recommender systems[C]//Proceedings of the 1st Workshop on Deep Learning for Recommender Systems. Association for Computing Machinery, 2016: 7-10.
- [20] WANG R, FU B, FU G, et al. Deep & cross network for ad click predictions[C]//Proceedings of the ADKDD'17. Association for Computing Machinery, 2017.
- [21] GUO H, TANG R, YE Y, et al. DeepFM: a factorization-machine based neural network for ctr prediction[C]//Proceedings of the 26th International Joint Conference on Artificial Intelligence. AAAI Press, 2017: 1725-1731.
- [22] ZHANG W, DU T, WANG J. Deep learning over multi-field categorical data[M]. Ferro N, Crestani F, Moens M-F, et al., eds.//Advances in Information Retrieval. Cham: Springer International Publishing, 2016: 45-57.
- [23] RENDLE S. Factorization Machines[C]//2010 IEEE International Conference on Data Mining. IEEE Press, 2010: 995-1000.
- [24] LU X, FENG J, HAN Y, et al. GraphMLP-Mixer: A Graph-MLP Architecture for Efficient Multi-Behavior Sequential Recommendation Method[J]. Journal of Computer Research and Development. 2024, 61(08): 1917-1929.
- 卢晓凯, 封军, 韩永强, 等. GraphMLP-Mixer: 基于图-多层感知机架构的高效多行为序列推荐方法[J]. 计算机研究与发展, 2024, 61(08): 1917-1929.
- [25] LI S, XIE R, ZHU Y, et al. User-centric conversational recommendation with multi-aspect user modeling[C]//Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval. ACM, 2022: 223-233.
- [26] DENG Y, ZHANG W, XU W, et al. A unified multi-task learning framework for multi-goal conversational recommender systems[J]. ACM Trans. Inf. Syst.: 2023, 41(3): 25.
- [27] LONG X, HUANG C, XU Y, et al. Social recommendation with self-supervised metagraph informax network[C]//Proceedings of the 30th ACM International Conference on Information & Knowledge Management. Association for

- Computing Machinery,2021: 1160-1169.
- [28] 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.Hong Kong China: ACM,2013:165-172.
- [29] LING G,LYU M,KING I.Ratings meet reviews, a combined approach to recommend[C]//Proceedings of the 8th ACM Conference on Recommender Systems.ACM,2014:105-112.
- [30] BAO Y,FANG H,ZHANG J.TopicMF: simultaneously exploiting ratings and reviews for recommendation[J].Proceedings of the AAAI Conference on Artificial Intelligence,2014,28(1).
- [31] ZHENG L,NOROOZI V,YU P.Joint deep modeling of users and items using reviews for recommendation[C]//Proceedings of the Tenth ACM International Conference on Web Search and Data Mining. Cambridge United Kingdom: ACM,2017:425-434.
- [32] KIM D,PARK C,OH J,et al. Convolutional Matrix Factorization for Document Context-Aware Recommendation[C]// Proceedings of the 10th ACM Conference on Recommender Systems,2016:233-240.
- [33] VASWANI A,SHAZEER N,PARMAR N,et al. Attention is all you need[C]//International Conference on Neural Information Processing Systems (NIPS'17).2017:6000-6010.
- [34] CHEN C,ZHANG M,LIU Y,et al.Neural attentional rating regression with review-level explanations[C]//Proceedings of the 2018 World Wide Web Conference on World Wide Web - WWW '18. ACM Press,2018:1583-1592.
- [35] LU Y,DONG R,SMYTH B.Coevolutionary recommendation model: mutual learning between ratings and reviews[C]//Proceedings of the 2018 World Wide Web Conference on World Wide Web - WWW '18. ACM Press,2018:773-782.
- [36] TAY Y,TUAN L,HUI S.Multi-pointer co-attention networks for recommendation[C]// Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining.2018:2309 - 2318.
- [37] RASHED A,ELSAYED S,SCHMIDT-THIEME L.CARCA: context and attribute-aware next-item recommendation via cross-attention[C]//Sixteenth ACM Conference on Recommender Systems.2022.
- [38] SUN P, WU L, ZHANG K, et al.Dual learning for explainable recommendation: towards unifying user preference prediction and review generation[C]//Proceedings of The Web Conference 2020. Taipei Taiwan: ACM,2020:837-847.
- [39] WANG X, OUNIS I, MACDONALD C.Leveraging review properties for effective recommendation[C]//Proceedings of the Web Conference 2021. Association for Computing Machinery.2021:2209-2219.
- [40] SHUAI J, ZHANG K, WU L, et al.A review-aware graph contrastive learning framework for recommendation[C]//Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval. Association for Computing Machinery,2022:1283-1293.
- [41] DEVLIN J,CHANG M-W,LEE K, et al.BERT: pre-training of deep bidirectional transformers for language understanding[C]//North American Chapter of the Association for Computational Linguistics. Minneapolis, Minnesota: Association for Computational Linguistics,2019:4171-4186.
- [42] WU L,QUAN C,LI C,et al.A Context-Aware User-Item Representation Learning for Item Recommendation[J]. ACM Trans. Inf.

Syst,2019,37(2):29.

[43] SALAKHUTDINOV R,MNIH A.Probabilistic Matrix Factorization[C]//Neural Information Processing Systems (NIPS).Curran Associates Inc, 2007:1257–1264.

[44] LIU D,LI J,DU B,et al.DAML: Dual Attention Mutual Learning between Ratings and Reviews for Item Recommendation[C]//Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining.2019:344-352.

全部作者中文简介:

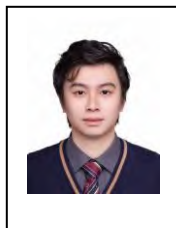
邓策渝, 2002 年生, 硕士研究生, CCF 学生会员(编号:P0293G)。主要研究方向为推荐系统和自然语言处理。

李段腾川, 1994 年生, 博士, CCF 会员。主要研究方向为推荐系统、知识图谱、计算机视觉。

胡奕仁, 1995 年生, 现任职于湖北省烟草公司武汉市公司, 湖北武汉 430000。他与武汉大学计算机学院合作聚焦推荐算法的研究。他的主要研究兴趣包括推荐系统和大数据分析。

王晓光, 1978 年生, 武汉大学信息管理学院特聘教授、副院长。研究方向为知识发现、知识组织。

李志飞, 1993 年生, 博士, 副教授, CCF 会员。主要研究方向为知识图谱和图神经网络。

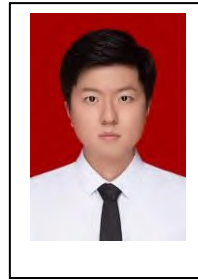


DENG Ceyu, born in 2002, postgraduate, is a student member of CCF (No.P0293G). His main research interests include recommender system and natural language processing.



LI Duantengchuan, born in 1994, Ph.D, is a member of CCF. His main research interests include recommender system,

knowledge graph and computer vision.



HU Yiren, born in 1995, He is currently working in Wuhan City Corporation of Hubei Province Tobacco Company, Wuhan 430000, China. And he has

collaborated with the School of Computer Science, Wuhan University on research related to recommendation algorithms. His main research interests include recommendation systems and big data analytics.



WANG Xiaoguang, born in 1978, is a distinguished professor and vice President of the School of Information Management, Wuhan University. His research interests include knowledge discovery and knowledge organization.



LI Zhifei, born in 1993, Ph.D, associate professor, is a member of China Computer Federation. His main research interests include knowledge graphs and graph neural networks.