

基于物品描述和评论的多粒度注意力机制的推荐

刘世新

上海理工大学光电信息与计算机工程学院, 上海

收稿日期: 2024年4月20日; 录用日期: 2024年5月15日; 发布日期: 2024年5月22日

摘要

本研究针对传统推荐系统依赖评分信息、面临数据稀疏性的局限, 以及深度学习模型的可解释性不足, 提出了一种融合评分数据和文本信息(评论、物品描述)的多粒度注意力机制的推荐模型(Multi-Grained Attention Recommendation, MGAR)。本文采用两个并行的卷积神经网络(CNN)分别处理用户和物品的评论与描述, 并通过词级、短语级和句子级的注意力机制来提取不同层次的语义信息, 实现评论文本的深度融合, 从而更有效地捕捉用户偏好和物品特性。在Amazon上的4个子数据集的实验结果表明, 本文提出的模型在预测准确度上均优于传统的基于评分的推荐模型和近年来流行的DeepCoNN模型。

关键词

推荐算法, 多粒度注意力机制, 卷积神经网络, 评论文本, 数据稀疏性

Multi-Granularity Attention Mechanism for Recommendation Systems Based on Item Descriptions and Reviews

Shixin Liu

School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai

Received: Apr. 20th, 2024; accepted: May. 15th, 2024; published: May. 22nd, 2024

Abstract

This study addresses the challenges faced by traditional recommendation systems that rely on

rating information and contend with data sparsity, as well as the lack of interpretability in deep learning models. It proposes a Multi-Grained Attention Recommendation (MGAR) model that integrates rating data with textual information (reviews, product descriptions). Utilizing two parallel Convolutional Neural Networks (CNNs) to process user and item reviews and descriptions, the model employs word-level, phrase-level, and sentence-level attention mechanisms to extract semantic information across various granularities. This facilitates a deep integration of review text, enabling a more effective capture of user preferences and product characteristics. Experimental results on four Amazon datasets indicate that the proposed model outperforms traditional rating-based recommendation systems and the recently popular DeepCoNN model in terms of predictive accuracy.

Keywords

Recommendation Algorithms, Multi-Granularity Attention Mechanism, Convolutional Neural Networks, Review Text, Data Sparsity

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着互联网和大数据技术的迅猛发展,信息过载成为一个日益突出的问题[1]。推荐系统作为解决这一问题的关键技术,通过分析用户历史行为来提供个性化推荐,以辅助用户做出决策,已经在电子商务[2][3]、智慧城市[4]、医疗[5]等多个领域取得成功应用。尤其在电子商务领域,推荐系统通过分析用户的偏好,实现精准信息过滤,为用户呈现个性化推荐内容。然而,为了提升用户对推荐系统的满意度,如何准确预测用户对物品评分成为亟需解决的问题。

传统的推荐系统主要利用用户或物品的评分信息来进行推荐[6][7],但实际场景中这些评分数据通常较为稀疏。为了解决数据稀疏性带来的性能问题,研究者们开始关注利用评论文本信息。电子商务的快速发展不仅带来了大量物品信息,还产生了丰富的评论文本信息。这些评论文本所包含的详实信息能够揭示物品特性以及用户个人偏好,在强化推荐系统的个性化方面发挥着关键作用。因此,许多学者开始探索如何借助用户评论这一辅助信息来学习用户和物品的表示,从而弥补数据稀疏性的问题。

在早期研究中,主题建模方法得到了广泛应用[8][9]。相较于传统利用评分信息的模型,主题建模方法表现出更优的性能。然而,这些方法仅关注了评论的主题线索,使用词袋表示方式,忽略了评论的上下文信息,从而存在信息损失的问题,制约了预测准确度的提升。随着深度学习在各领域的崛起,最近的研究更倾向于采用神经网络来从评论中提取特征以进行推荐。一些模型如 DeepCoNN [10]、D-Attn [11]、和 CARL [12]等将用户和物品的多条评论连接为一个文档,利用卷积神经网络对文本进行特征提取,得到用户或物品的整体兴趣或特性。然而,这些方法在聚合所有评论时未区分不同评论的重要程度,可能会影响有效特征的提取。另一些模型如 NARRE [13]、MPCN [14]、DARMH [15]、ARPCNN [16]则专注于基于评论的建模,单独对每条评论建模,并利用注意力机制区分不同评论的重要性。随后,将评论特征汇总为用户和物品的综合特征,能够捕获用户对某一特定物品的偏好信息,从而实现了比文档级建模更高的预测准确度。

基于评论的建模关键在于对评论文本中不同粒度信息的捕获,不同用户所撰写的评论侧重点不同,

其中,评论中不同的粒度文本包含了重要的语义特征。如图 1 所展示的不同粒度的评论文本(单词、短语、句子)可以揭示用户对不同物品的偏好,例如单词: outstanding, superb 等, 短语: newfound love, great writer, star of this show, amazing actor 等, 句子: Just an outstanding job with the writing and the acting is superb all around 等。



Figure 1. Example comments

图 1. 示例评论

具体来说,我们采用两个并行的 CNN (表示为 User-Net 和 Item-Net)分别处理用户评论(用户撰写的评论)和物品评论及描述(为物品撰写的评论和描述),以提取语义信息。然后,我们在上述两个网络中提出了个性化单词级注意机制、个性化短语级注意级机制和个性化句子级注意机制,从多粒度的角度突出评论中的相关信息,提取用户和物品的特征,并采用因子分解机模型融合提取到的用户特征与物品特征,完成评分预测。

本文的贡献如下:

(1) 本文采用更稀疏的 Amazon 数据集,以及将物品描述文档与评论一起引入模型中,从而更贴切现实场景,并且在新物品缺失评论时,物品描述文档可以很好的缓解物品冷启动问题,帮助我们更好地提取物品特征,提高预测准确度。

(2) 我们提出了个性化多粒度(单词、短语、句子)注意力机制,可以通过分配高注意力权重来强化关键词、关键短语和信息性句子。并对本文提出的个性化注意力机制进行了可视化分析,以提高该模型的可解释性。

(3) 本文在四个 Amazon 子数据集上进行实验,结果显示,本文推荐算法的均方误差均优于同类算法。

2. 相关工作

作为深度学习的代表算法, CNN 一直是计算机视觉和自然语言处理领域的研究热点[17]。特别是在

推荐系统中, CNN 与各种机制的结合使用, 已经显著提高了模型的性能。ConvMF [18]模型通过融合 CNN 和概率矩阵分解(PMF)来有效提取项目描述文档中的上下文信息。进一步, 郑[10]等人通过并行结构的双 CNN 模型, 分别从用户和项目评论中提取潜在特征, 构建了 DeepCoNN 模型, 该模型能够联合学习用户和项目的特征表示, 从而提高评分预测的准确性。因此, 本文提出的模型使用两个 CNN 作为基本结构, 共同学习从用户评论和项目评论中提取的潜在特征。

近年来, 注意力机制在推荐系统中的应用已经成为一个重要的发展趋势[19] [20]。这种机制能够快速识别输入中最具信息量的内容, 从而提高推荐性能。Seo [11]等人首次在基于评论文本的推荐系统中引入注意力机制。这种机制通过计算不同单词在句子中的权重, 有助于突出重要单词, 忽视不重要的单词, 从而更精确地从评论文本中提取用户和物品的特征。接着, NARRE 模型[13]通过在 CNN 文本处理器中加入基于评论的注意力层, 使得对关键评论信息的加权更加突出, 增强了推荐的可解释性。NRPA 模型[21]使用个性化的单词级注意力为每个用户和物品选择评论中更重要的单词, 在用户和物品编码器中, 应用个性化的评论级别注意力, 通过根据其权重聚合所有评论表示来学习用户和物品表示。ARPCNN [16]提出了个性化的单词级注意机制和个性化的评论级注意机制, 通过分配高的注意力权重来强化关键性单词和信息性评论。在本文中, 我们使用多粒度注意力机制来更精细地处理单词、短语和评论文档的权重分配。旨在更全面地提取重要的语义信息, 从而进一步优化推荐系统的性能和解释能力。

3. 模型设计

3.1. 模型框架

我们的模型由两个并行的 CNN 组成, 分别称为 User-Net 和 Item-Net。User-Net 负责从用户评论中提取用户偏好, 通过卷积神经网络(CNN)对用户评论进行特征提取, 使用多粒度注意力机制来区分不同评论的重要性, 接着将提取的特征和注意力权重聚合成用户表示。Item-Net 与 User-Net 类似, 区别在于在数据输入时增加物品描述, 与物品评论一起输入模型的嵌入层, 来提取物品特征, 同样使用多粒度注意力机制, 并将提取的特征和注意力权重聚合成物品表示。最后, 将用户表示和物品表示进行融合以及交互, 以完成评分预测。模型的总体结构如图 2 所示。

3.2. 用户网络模块

- 词嵌入: 用户网络将用户撰写的评论作为输入。将评论文本中的每个单词通过 word2vec [22]映射成一个维度为 d 的词向量, 给定评论文本 $D_u = \{D_{u1}, \dots, D_{ui}, \dots, D_{um}\}$ 表示用户 u 的所写评论的集合, 其中 $D_{ui} = \{d_1, \dots, d_q, \dots, d_T\}$ 表示用户 u 的第 i 个评论, d_q 表示在此评论中第 q 个单词, T 表示一条评论的固定长度, 如果评论长度小于 T , 将用零填充处理, 若大于 T , 则截取处理。用户 u 对物品的评论的单词矩阵可以表示为:

$$X_{ui} = [x_1; \dots; x_q; \dots; x_T]$$

其中 $x_q \in R^{1 \times d}$, $X_{ui} \in R^{T \times d}$ 。

- 个性化词级注意力层: 在本文中, 单词的语义表示被定义为单词本身及其上下文。我们使用 Q 表示中心词的文本范围, 其窗口宽度可以计算为 $L = 1 + 2Q$ 。零填充用于处理评论中缺少上下文的第一个和最后一个单词。中心词 x_q 的语义表示可以表示为:

$$X_q^{att} = [x_q - Q; \dots; x_q; \dots; x_q + Q]$$

x_q 的特征 ϕ_q 可以计算为:

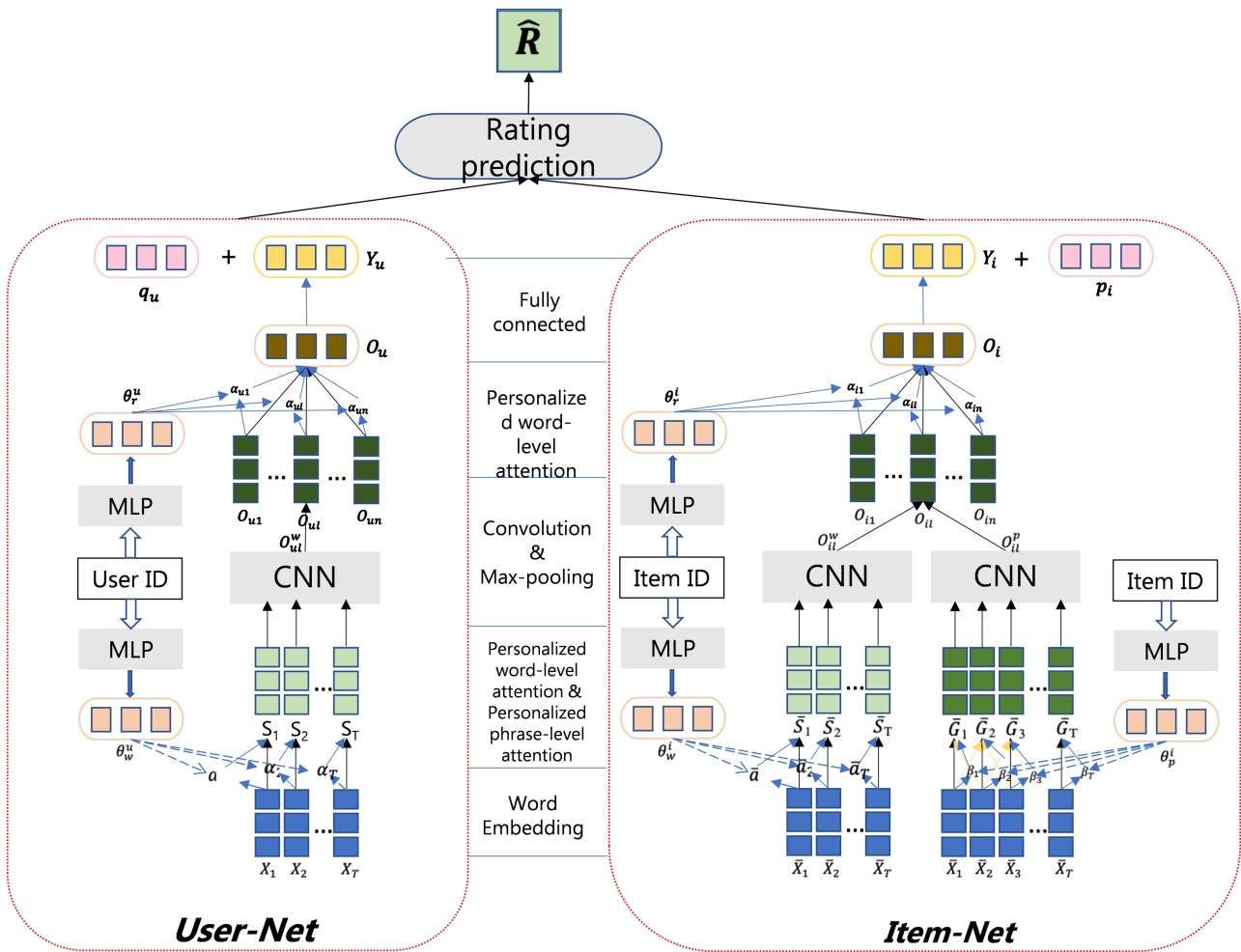


Figure 2. The neural network architecture of MGAR
 图 2. MGAR 的神经网络框架

$$\phi_q = ReLU(X_q^{att} \otimes W_u^{att} + b_u^{att})$$

其中 ϕ_q 表示评论文本中第 q 个单词的语义特征， W_u^{att} 和 b_u^{att} 分别表示加权矩阵和滑动窗口的偏置项，操作 \otimes 表示两个矩阵相乘后对应元素的和。

由于 ID 信息是每个用户和物品的唯一标识，我们将 ID 的低维特征向量视为个性化因素，并将其体现在个性化注意力机制中，如图 2 所示。给定用户 u 的嵌入 ID，多层感知器(MLP)可用于生成用户 u 的个性化单词级注意力向量。

$$\theta_w^u = ReLU(W_w^u u_{id} + b_w^u)$$

其中 W_w^u 是 MLP 的权重矩阵， b_w^u 是偏置向量， u_{id} 表示用户 u 的嵌入 ID。

为了强调关键词，每个词对应的个性化注意力值可以计算如下：

$$\alpha_q^* = \theta_w^u A_w^u \phi_q$$

$$\alpha_q = \frac{\exp(\alpha_q^*)}{\sum_{q=1}^T \exp(\alpha_q^*)}$$

其中 θ_w^u 是用户 u 个性化词级注意力向量, A_w^u 是词级注意力和谐矩阵, ϕ_q 和 α_q 分别表示评论文本中单词的语义特征和注意力值。

在将每个单词的注意力值 α_q 乘以其嵌入向量 x_q 后, 获得新的加权嵌入矩阵 S_{ui} , 如下所示:

$$s_p = \alpha_q x_q$$

$$S_{ui} = [s_1; \cdots; s_q; \cdots; s_T]$$

其中 $s_q \in R^{1 \times d}$, $S_{ui} \in R^{T \times d}$ 。

- 个性化短语级注意力层: 在本文中, 短语的长度我们定义为 3 个单词, 短语的语义表示被定义为三个单词本身及其上下文。我们使用 P 表示中心短语的文本范围, 其窗口宽度可以计算为 $L = 3 + 2P$ 。零填充用于处理评论中缺少上下文的第一个和最后一个单词。中心短语 x_p 的语义表示可以表示为:

$$X_p^{att} = [x_{q-1-P}; \cdots; x_{q-1}; x_q; x_{q+1}; \cdots; x_{q+1+P}]$$

$$X_p^{p-att} = [X_{p-1}^{att}; X_p^{att}; X_{p+1}^{att}]$$

x_p 的特征 ϕ_p 可以计算为:

$$\phi_p = \text{ReLU}(X_p^{p-att} \otimes W_u^{p-att} + b_u^{p-att})$$

其中 ϕ_p 表示评论文本中第 p 个短语的语义特征, W_u^{p-att} 和 b_u^{p-att} 分别表示加权矩阵和滑动窗口的偏置项, 操作 \otimes 表示两个矩阵相乘后对应元素的和。类似于个性化词级注意力层, 本文使用 MLP 以产生个性化短语级别注意力向量。

$$\theta_p^u = \text{ReLU}(W_p^u u_{id} + b_p^u)$$

其中 W_p^u 是 MLP 的权重矩阵, b_p^u 是偏置向量, u_{id} 表示用户 u 的嵌入 ID。

为了强调关键短语, 每个短语对应的个性化注意力值可以计算如下:

$$\beta_p^* = \theta_p^u A_p^u \phi_p$$

$$\beta_p = \frac{\exp(\alpha_p^*)}{\sum_{p=1}^T \exp(\alpha_p^*)}$$

其中 θ_p^u 是用户 u 个性化词级注意力向量, A_p^u 是词级注意力和谐矩阵, ϕ_p 和 α_p 分别表示评论文本中短语的语义特征和注意力值。

在将每个短语的注意力值 α_p 乘以其嵌入向量 x_p 后, 获得新的加权嵌入矩阵 G_{ui} , 如下所示:

$$g_p = \alpha_p x_p$$

$$G_{ui} = [g_1; \cdots; g_p; \cdots; g_T]$$

其中 $g_p \in R^{3 \times d}$, $G_{ui} \in R^{3T \times d}$ 。

- CNM 层: 词级注意力和短语级注意力的输出 S_{ui} 、 G_{ui} 被传递至卷积层, 用于提取评论中每个单词和短语的上下文特征。本文在网络中使用了 k 个滤波器且长度为 t 的滑动窗口滤波器 $K_j \in R^{t \times d}$ 应用于矩阵 S_{ui} 和 G_{ui} 上, 通过第 j 个滤波器提取 S_{ui} 和 G_{ui} 的特征可以表示为:

$$z_j = \text{ReLU}(S_{ui} * K_j^w + b_j^w)$$

$$h_j = \text{ReLU}(G_{ui} * K_j^p + b_j^p)$$

其中 $*$ 表示卷积操作, b_j^w 和 b_j^p 表示偏差项。

与普通的卷积神经网络结构一致, 在滤波器 K_j 提取特征 z_j 、 h_j 之后, 本模型使用最大池化操作来保留每个滤波器所产生的最有价值的上下文语义特征。

$$z_{uj} = \max \{z_j^1, z_j^2, \dots, z_j^{T-t+1}\}$$

$$h_{uj} = \max \{h_j^1, h_j^2, \dots, h_j^{3T-t+1}\}$$

最大池化操作后, 所有保留下来的语义特征被串联成最大池化层的输出:

$$O_{ul}^w = [z_{u1}, z_{u2}, \dots, z_{uk}]$$

$$O_{ul}^p = [h_{u1}, h_{u2}, \dots, h_{uk}]$$

$$O_{ul} = W_l^u [O_{ul}^w, O_{ul}^p] + b_l^u$$

- 个性化评论级注意力层: 类似于个性化词级注意力层和个性化短语级注意力层, 本文使用 MLP 以产生个性化评论级注意力向量。

$$\theta_r^u = \text{ReLU}(W_r^u u_{id} + b_r^u)$$

其中 W_r^u 是 MLP 的权重矩阵, b_r^u 是偏置向量。

用户 u 的第 l 条评论的个性化评论级注意力权重 a_{ul} 可以表示为:

$$a_{ul}^* = \theta_r^u A_r^u O_{ul}$$

$$\alpha_{ul} = \frac{\exp(\alpha_{ul}^*)}{\sum_{l=1}^n \exp(\alpha_{ul}^*)}$$

其中 θ_r^u 是个性化评论级注意力向量, A_r^u 是评论级注意力和谐矩阵。

通过将个性化注意力权重乘以其评论特征 O_{ul} , 整合该用户的所有评论特征, 可获得用户的总体评论特征如下:

$$O_u = \sum_{l=1,2,\dots,n} a_{ul} O_{ul}$$

- 全连阶层: 最后, 最大池化层的输出 O_u 被传递至全连接层, 以获取用户 u 的潜在表示:

$$Y_u = W_u O_u + b_u$$

式中, Y_u 可以被认为是从用户评论中提取的关于用户 u 的潜在特征表示, O_u 和 b_u 分别代表全连阶层的权重矩阵和偏置向量。至此, 已经获得用户表示 Y_u 和物品表示 Y_i 。

- 评分预测与模型学习:

$$\hat{R}_{u,i} = W^T H_{ui} + b_u + b_i + \mu$$

其中 W 表示权重矩阵, H_{ui} 表示用户和项目之间的特征交互, μ 表示全局平均评分, b_u 和 b_i 分别表示用户偏差项和项目偏差项。

考虑到从评论中获得的特征表示和从评分数据中提取的潜在特征等级, 可以对用户和项目之间的特征交互进行评估:

$$H_{ui} = (Y_u + q_u) \odot (Y_i + q_i)$$

本模型的损失函数定义为:

$$\mathcal{L} = \sum_{u,i \in \tau} (\hat{R}_{u,i} - R_{u,i})^2$$

4. 实验结果分析

4.1. 数据集和评价指标

Amazon 数据集是电子商务领域最常用的推荐系统数据集之一，包含了超过 6 亿个用户、超过 2.5 亿个物品、超过 10 亿条评论。大多数论文采用 Amazon 亚马逊的 5-core 数据集，该数据集包含了所有至少有 5 条评论的用户和项目。然而，5-core 数据集无法反映现实场景中更为稀疏的数据分布。因此，本文采用了 Amazon 的 raw review data 数据集，并通过数据库技术将其与元数据(包含 item ID 和物品描述等信息)进行合并，形成了新数据集。

本文从中选取了 Musical Instruments、Digital Music、Video Games 和 All Beauty 作为模型的实验数据。表 1 总结了这四个基本数据集的详细统计数据。为了评估，随机选取每个数据集的 80% 用于训练模型，10% 作为测试集，剩余的 10% 用于验证集以调整超参数。

Table 1. Statistical details of the datasets

表 1. 数据集的详细统计

数据集	用户数	物品数	交互数	稀疏度%
Musical Instruments	903,058	112,132	17,31,700	99.983%
DigitalMusic	127,174	66,009	194,569	99.977%
VideoGames	1,539,732	71,909	1,967,985	99.982%
All Beauty	319,335	32,486	382,532	99.963%

均方误差(Mean Squared Error, MSE)是回归模型评价指标之一，用于衡量模型预测值与真实值之间的差异程度。MSE 的计算公式如下：

$$MSE = \frac{1}{N} \sum_{u,i} (\hat{R}_{u,i} - R_{u,i})^2$$

其中， $R_{u,i}$ 表示真实值， $\hat{R}_{u,i}$ 表示预测值， N 表示样本数量。

4.2. 对比模型和参数设置

为了验证模型的有效性，本文选择了一个早期的仅利用评分矩阵的模型(PMF)、和五个利用深度学习方法来处理评论文本的先进模型(ConvMF、DeepCoNN、D_ATTN、MPCN 和 NARRE)作为对比模型。

PMF (Probabilistic Matrix Factorization)是推荐系统中常用的一种矩阵分解方法。PMF 假设用户和物品之间的评分可以由用户和物品的潜在特征表示，并通过矩阵分解的方法学习用户和物品的潜在特征。

ConvMF (Convolutional Matrix Factorization)是 PMF 的一种扩展，它采用卷积神经网络来学习用户和物品的潜在特征。ConvMF 能够有效捕捉用户和物品之间的局部关系，从而提高模型的性能。

DeepCoNN (Deep Collaborative Filtering)采用多层卷积神经网络来学习用户和物品的潜在特征。DeepCoNN 能够有效捕捉用户和物品之间的复杂关系，从而提高模型的性能。

D_ATTN (Deep Attention-Based Recommendation)采用注意力机制来学习用户和物品之间的相关性。D_ATTN 能够有效捕捉用户和物品之间的长距离关系，从而提高模型的性能。

MPCN (Multi-Perspective Collaborative Filtering)采用多视角来学习用户和物品的潜在特征。MPCN 能

够有效捕捉用户和物品之间的多种关系，从而提高模型的性能。

NARRE (Neural Attention-Based Recommendation Evolution)是一种基于注意力机制的推荐模型，它能够捕捉用户和物品之间的动态关系。NARRE 能够有效提高推荐模型的长期性能。

除了上述描述之外，表 2 展示了各模型和 MGAR 的特征。

Table 2. Comparison of features of different models

表 2. 不同模型的特征对比

	PMF	ConvMF	DeepCoNN	D_ATT	MPCN	NARRE	MGAR
评分数据	✓	✓	✓	✓	✓	✓	✓
评论文本		✓	✓	✓	✓	✓	✓
深度学习		✓	✓	✓	✓	✓	✓
双塔结构			✓	✓	✓	✓	✓
句子级注意力机制						✓	✓
短语级注意力机制							✓
词级注意力机制							✓
物品描述							✓

本文采用网格搜索方法来确定每个模型参数的最优值。对于 PMF，我们在{4, 8, 16, 32, 64, 128}中优化潜在因子的数量，并从{0.001, 0.01, 0.1, 1.0}中选择最优的正则化参数值。对于 ConvMF、DeepCoNN、D-ATTN、MPCN、NARRE、和 MGAR，分别在{0.005, 0.01, 0.02, 0.05}和{4, 8, 16, 32, 64, 128, 256}中搜索合适的学习率和潜在因子数量。我们在{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0}中微调 dropout 的值，并在{25, 50, 100, 150}中选择批处理大小。本文中使用的词嵌入模型是一个预训练模型，已经使用来自谷歌新闻语料库中超过 1000 亿词进行了训练。对于应用了注意力机制的 D_ATT、MPCN、NARRE 和 MGAR，我们从{4, 8, 16, 32, 64, 128, 256}中调整注意力大小。在 MGAR 中，我们从{4, 8, 16, 32, 64, 128, 256}中搜索用于个性化注意力机制中的 ID 嵌入的适当维度。

4.3. 实验结果分析

MGAR 与六个比对模型在四个数据集上的性能比较在表 3 中列出。如表 III 所示，PMF 有最大的均方误差，特别是当数据集非常稀疏时。相比之下，其他六个模型在四个数据集上的表现都优于 PMF，这

Table 3. Experimental result

表 3. 实验结果

	Musical Instruments	Digital Music	Video Games	All Beauty
PMF	1.401	1.211	1.629	1.091
ConvMF	1.191	1.084	1.419	0.960
DeepCoNN	0.924	1.046	1.217	0.868
D_ATT	0.901	1.027	1.192	0.846
MPCN	0.923	0.914	1.122	0.823
NARRE	0.848	0.868	1.048	0.766
MGAR	0.824	0.759	0.925	0.725

验证了具有强大处理能力的深度学习可以从文本评论中提取更多有用的语义信息。与 ConvMF 相比，DeepCoNN 有显著的性能提升，有效验证了并行结构的联合建模能力。通过对文本评论的潜在特征进行建模，NARRE、D_ATTN 和 MPCN 在四个数据集上都取得了良好的表现。这证明了这三种模型的注意力机制对于用户偏好和项目属性建模的贡献。然而，这三种方法在高稀疏性数据集上的性能提升低于其他数据集，这表明尽管现有的基于神经网络的模型在许多领域表现良好，但当评论信息不完整或对评论的有效处理不足时，它们的预测性能仍然相对较差。

根据表 3 中的实验结果，MGAR 在四个数据集上均达到了最佳性能，这验证了个性化注意力机制的应用和物品描述特征提取的有效性。与 NARRE 相比，MGAR 在词级、短语级和句子级注意力机制中都加入了个性化特征，并充分利用了物品描述中包含的丰富语义信息。

4.4. 消融实验

为了说明模型各部分对性能提升的不同影响，我们提出了 MGAR 的四种变体。它们的简要介绍如下。请注意，“非个性化”意味着我们修改了注意力权重的计算，以去除与个性化相关的信息。

MGAR1 结构为两个并行的 CNN 和扩展的潜在因子模型，既不采用个性化的多粒度(单词级、短语级和句子级)注意力机制，也不添加物品描述文档。

MGAR2 在 MGAR1 的基础上增加非个性化多粒度注意力机制。

MGAR3 在 MGAR2 的基础上增加物品描述文档。

MGAR4 在 MGAR2 的基础上将非个性化的多粒度注意力机制改进为个性化多粒度注意力机制。

MGAR 及其三种变体在四个数据集上的性能比较显示在表 4 中。如表 4 所示，每个变体在四个数据集上的均方误差都显示出明显的下降趋势，这意味着评分预测性能逐渐变好。具体来说，与 MGAR1 相比，MGAR2 的性能有了显著提升，表明仅依靠评论中包含的语义信息和 CNN 的数据处理能力是不够的，而应用注意力机制可以强化信息丰富的评论，深度挖掘评论各粒度的信息，从而提高性能。与 MGAR2 相比，MGAR3 增加了物品描述文档，改善物品冷启动问题，评分预测性能得到了进一步提升。比较 MGAR2 和 MGAR4 可以发现，MGAR4 不仅可以强化信息丰富的评论，还可以突出关键词，同时也表明个性化的 ID 特征在分配注意力权重的过程中确实发挥了有效作用。通过性能比较，我们进一步验证了并行 CNN、个性化多粒度注意力机制和添加物品描述信息协同作用，对提升推荐性能的有效性。

Table 4. Ablation experiment

表 4. 消融实验

	Musical Instruments	Digital Music	Video Games	All Beauty
MGAR1	1.041	0.974	1.251	1.007
MGAR2	1.013	0.935	1.232	0.971
MGAR3	0.995	0.894	1.212	0.938
MGAR4	0.964	0.865	1.124	0.856
MGAR	0.824	0.759	0.925	0.725

4.5. 可解释性分析

接下来，我们对本文提出的个性化多粒度注意力机制进行可视化分析，以提高模型的可解释性。表 5 展示了 MGAR 在 Digital Music 数据集上两位用户评论的注意力权重。个性化单词级和短语级注意力机制根据每个词或短语在评论中对物品属性建模的重要性，给关键词和短语分配高注意力权重。我们通过

设置不同的颜色来区分不同的注意力权重。颜色越深，权重越高。注意，为了显示方便，我们没有标记注意力权重低的词和短语。如表 5 所示，与物品相关的关键词和短语在应用个性化单词级和短语级注意力机制的用户评论中被不同程度地突出显示。

个性化句子级注意力机制根据评论对物品属性建模的有用性，给信息丰富的评论分配高注意力权重。如表 5 所示，第一位用户对物品评分为五分，详细信息在评论中得到了充分解释。因此，这篇评论获得了更高的注意力权重。与第一位用户相反，由于缺乏对物品的客观描述，第二位用户的评论获得了较低的注意力权重。

Table 5. Interpretability analysis
表 5. 可解释性分析

用户 ID	评分	用户评论
A25ZT87OMIPLNX	5	Clever, inspired and moving. This is a great album from one of the finest Christian artists ever. God took him home early, but not before leaving us with a great message and music. Keith Green makes Bible lessons come to life with his music. (注意力权重: 0.191)
A1XA8ZX068G0PJ	4	I've enjoyed this collection since I was a kid. Michael Card has a very pleasant voice & terrific songwriting talent. Listening to this brought back a lot of memories ... enough so that I'm now on the lookout for the songbook. (注意力权重: 0.056)

5. 总结

针对评分数据和评论文本的稀疏性以及物品冷启动问题，本文提出了 MGAR 模型，该模型采用并行 CNN 作为框架，并利用 User-Net 和 Item-Net 提取用户评论、物品评论和物品描述的潜在特征。应用嵌入 ID 特征的个性化多粒度注意力机制，以增强每条评论中的关键词、关键短语和信息丰富的评论。然后，将用户特征和物品特征进行融合以及交互，以完成评分预测。在四个真实世界数据集上的广泛实验结果表明，MGAR 的评分预测性能优于六个基线模型。此外，MGAR 与四个变体的性能比较以及可解释性分析更直观地展示了模型各部分在性能提升中的重要作用。在未来，我们将在努力提高推荐结果准确性的基础上，探索多个评价指标(如多样性、新颖性)，寻找新颖多样的物品列表，来提高用户的满意度。

参考文献

- [1] Anyosa, S.C., Vinagre, J. and Jorge, A.M. (2018) Incremental Matrix Co-Factorization for Recommender Systems with Implicit Feedback. *Companion Proceedings of the Web Conference 2018*, Lyon, 23-27 April 2018, 1413-1418. <https://doi.org/10.1145/3184558.3191585>
- [2] Messaoudi, F. and Loukili, M. (2024) E-Commerce Personalized Recommendations: A Deep Neural Collaborative Filtering Approach. *Operations Research Forum*, **5**, Article No. 5. <https://doi.org/10.1007/s43069-023-00286-5>
- [3] Grbovic, M., Radosavljevic, V., Djuric, N., et al. (2015) E-Commerce in Your Inbox: Product Recommendations at Scale. *Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Sydney, 10-13 August 2015, 1809-1818. <https://doi.org/10.1145/2783258.2788627>
- [4] Chen, H., Li, Z., Wang, Z., et al. (2021) Edge Data Based Trailer Inception Probabilistic Matrix Factorization for Contextaware Movie Recommendation. *World Wide Web*, **25**, 1863-1882. <https://doi.org/10.1007/s11280-021-00974-4>
- [5] Li, Z., Chen, H., Lin, K., Shakhov, V., Shi, L. and Yu, J. (2021) From Edge Data to Recommendation: A Double Attention-Based Deformable Convolutional Network. *Peer-to-Peer Networking and Applications*, **14**, 3984-3997. <https://doi.org/10.1007/s12083-020-01037-7>
- [6] Bosri, R., Rahman, M.S., Bhuiyan, M.Z.A. and Omar, A.A. (2021) Integrating Blockchain with Artificial Intelligence for Privacy-Preserving Recommender Systems. *IEEE Transactions on Network Science and Engineering*, **8**, 1009-1018. <https://doi.org/10.1109/TNSE.2020.3031179>
- [7] Wang, X., Wang, R., Shi, C., Song, G. and Li, Q. (2020) Multi-Component Graph Convolutional Collaborative Filtering. *Proceedings of the AAAI Conference on Artificial Intelligence*, **34**, 6267-6274.

- <https://doi.org/10.1609/aaai.v34i04.6094>
- [8] Wang, C. and Blei, D.M. (2011) Collaborative Topic Modeling for Recommending Scientific Articles. *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Diego, 21-24 August 2011, 448-456. <https://doi.org/10.1145/2020408.2020480>
- [9] Ling, G.L., Yu, M.R. and King, I. (2014) Ratings Meet Reviews, a Combined Approach to Recommend. *Proceedings of the 8th ACM Conference on Recommender systems*, Foster City, 6-10 October 2014, 105-112. <https://doi.org/10.1145/2645710.2645728>
- [10] Zheng, L., Noroozi, V. and Yu, P.S. (2017) Joint Deep Modeling of Users and Items Using Reviews for Recommendation. *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*, Cambridge, 6-10 February 2017, 425-434. <https://doi.org/10.1145/3018661.3018665>
- [11] Seo, S., Huang, J., Yang, H. and Liu, Y. (2017) Interpretable Convolutional Neural Networks with Dual Local and Global Attention for Review Rating Prediction. *Proceedings of the Eleventh ACM Conference on Recommender Systems*, Como, 27-31 August 2017, 297-305. <https://doi.org/10.1145/3109859.3109890>
- [12] Wu, L., Quan, C., Li, C., et al. (2019) A Context-Aware User-Item Representation Learning for Item Recommendation. *ACM Transactions on Information Systems*, **37**, Article No. 22. <https://doi.org/10.1145/3298988>
- [13] Chen, C., Zhang, M., Liu, Y., et al. (2018) Neural Attentional Rating Regression with Review-Level Explanations. *Proceedings of the 2018 World Wide Web Conference*, Lyon, 23-27 April 2018, 1583-1592. <https://doi.org/10.1145/3178876.3186070>
- [14] Tay, Y., Luu, A.T. and Hui, S.C. (2018) Multi-Pointer Co-Attention Networks for Recommendation. *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, London, 19-23 August 2018, 2309-2318. <https://doi.org/10.1145/3219819.3220086>
- [15] Zhen, L., Yuan, B. and Ma, Y. (2022) A Multi-Task Dual Attention Deep Recommendation Model Using Ratings and Review Helpfulness. *Applied Intelligence*, **52**, 5595-5607. <https://doi.org/10.1007/s10489-021-02666-y>
- [16] Li, Z., Chen, H., Ni, Z., et al. (2022) ARPCNN: Auxiliary Review-Based Personalized Attentional CNN for Trustworthy Recommendation. *IEEE Transactions on Industrial Informatics*, **19**, 1018-1029. <https://doi.org/10.1109/TII.2022.3169552>
- [17] Dezfouli, P.A.B., Momtazi, S. and Dehghan, M. (2021) Deep Neural Review Text Interaction for Recommendation Systems. *Applied Soft Computing*, **100**, Article ID: 106985. <https://doi.org/10.1016/j.asoc.2020.106985>
- [18] Donghyun, K., Park, C., Oh, J., et al. (2016) Convolutional Matrix Factorization for Document Context-Aware Recommendation. *Proceedings of the 10th ACM Conference on Recommender Systems*, Boston, 15-19 September 2016, 233-240. <https://doi.org/10.1145/2959100.2959165>
- [19] Chen, J., Zhang, H., He, X., et al. (2017) Attentive Collaborative Filtering: Multimedia Recommendation with Item-and Component-Level Attention. *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, Shinjuku, 7-11 August 2017, 335-344. <https://doi.org/10.1145/3077136.3080797>
- [20] Xiao, J., Ye, H., He, X., et al. (2017) Attentional Factorization Machines: Learning the Weight of Feature Interactions via Attention Networks. *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, Melbourne, 19-25 August 2017, 3119-3125. <https://doi.org/10.24963/ijcai.2017/435>
- [21] Liu, H., Wu, F., Wang, W., et al. (2019) NRPA: Neural Recommendation with Personalized Attention. *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, Paris, 21-25 July 2019, 1233-1236. <https://doi.org/10.1145/3331184.3331371>
- [22] Mikolov, T., Sutskever, I., Chen, K., Corrado, G. and Dean, J. (2013) Distributed Representations of Words and Phrases and Their Compositionality. *Proceedings of the 26th International Conference on Neural Information Processing Systems*, Lake Tahoe, 5-10 December 2013, 3111-3119.