

## 基于用户静动态兴趣和去噪隐式负反馈的新闻推荐算法

魏金生, 周苏, 卢官明, 丁佳伟

引用本文

魏金生, 周苏, 卢官明, 丁佳伟. [基于用户静动态兴趣和去噪隐式负反馈的新闻推荐算法](#)[J]. 计算机科学, 2026, 53(2): 152-160.

WEI Jinsheng, ZHOU Su, LU Guanming, DING Jiawei. [News Recommendation Algorithm Based on User Static and Dynamic Interests and Denoised Implicit Negative Feedback](#) [J]. Computer Science, 2026, 53(2): 152-160.

---

## 相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

**Similar articles recommended (Please use Firefox or IE to view the article)**

[融合对比学习的掩码图自编码器](#)

Contrastive Learning-based Masked Graph Autoencoder

计算机科学, 2026, 53(2): 145-151. <https://doi.org/10.11896/jsjcx.250100155>

[DCL-FKT:融合双重对比学习与遗忘机制的个性化知识追踪模型](#)

DCL-FKT: Personalized Knowledge Tracing via Dual Contrastive Learning and Forgetting Mechanism

计算机科学, 2026, 53(2): 99-106. <https://doi.org/10.11896/jsjcx.250600002>

[基于方向感知孪生网络的知识概念先序关系预测方法](#)

Direction-aware Siamese Network for Knowledge Concept Prerequisite Relation Prediction

计算机科学, 2026, 53(2): 39-47. <https://doi.org/10.11896/jsjcx.250600005>

[基于双层级对比学习的健壮知识追踪模型](#)

Robust Knowledge Tracing Model Based on Two-level Contrastive Learning

计算机科学, 2026, 53(2): 31-38. <https://doi.org/10.11896/jsjcx.250700196>

[结合超图学习的多注意力机制新闻推荐方法](#)

Multiple Attention Mechanism News Recommendation Approach with Hypergraph Learning

计算机科学, 2025, 52(11A): 250200067-7. <https://doi.org/10.11896/jsjcx.250200067>

# 基于用户静动态兴趣和去噪隐式负反馈的新闻推荐算法

魏金生<sup>1,2</sup> 周 苏<sup>1</sup> 卢官明<sup>1,2</sup> 丁佳伟<sup>1</sup>

1 南京邮电大学通信与信息工程学院 南京 210003

2 智能信息处理与通信技术省高校重点实验室(南京邮电大学) 南京 210003

(weijs@njupt.edu.cn)

**摘 要** 在现有的新闻推荐系统中,通常将用户历史未点击新闻视为隐式负反馈,对隐式负反馈进行建模能够引导推荐模型过滤掉用户不感兴趣的新闻。未点击的新闻中可能会存在用户感兴趣的内容,即存在偏好噪声,导致对隐式负反馈建模的干扰;此外,由于用户兴趣的多样化和多变性,现有的新闻推荐系统往往存在“信息茧房”的问题。为了解决上述问题,提出了一种基于用户静动态兴趣和去噪隐式负反馈的新闻推荐算法。通过对用户的静态和动态兴趣进行融合建模,并对静态兴趣隐式负反馈进行去噪,来构建动态更新的用户偏好模型。首先,构建基于正交映射的静态兴趣去噪模块,对静态兴趣中的隐式负反馈进行去噪;然后,将门控循环单元(Gated Recurrent Unit,GRU)与正交映射相融合,构建基于改进 GRU 的动态兴趣去噪模块,充分建模用户的兴趣变化并实现对动态兴趣中的隐式负反馈进行去噪;最后,通过引入对比学习技术,增强模型对隐式正反馈和负反馈的区分能力,以提升个性化新闻推荐性能。在 MIND 数据集上的实验表明,与基线方法相比,该模型在 AUC, MRR, NDCG@5, NDCG@10 这 4 种评价指标上分别提升了 1.18%, 1.84%, 2.75%, 1.67%, 验证了该模型的有效性。

**关键词:** 新闻推荐;信息茧房;正交映射;门控循环单元;对比学习

**中图分类号** TP391

## News Recommendation Algorithm Based on User Static and Dynamic Interests and Denoised Implicit Negative Feedback

WEI Jinsheng<sup>1,2</sup>, ZHOU Su<sup>1</sup>, LU Guanming<sup>1,2</sup> and DING Jiawei<sup>1</sup>

1 School of Communication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China

2 Jiangsu Key Laboratory of Intelligent Information Processing and Communication Technology, Nanjing University of Posts and Telecommunications, Nanjing 210003, China

**Abstract** In the existing news recommendation system, the news that users have not clicked on in the past is usually regarded as implicit negative feedback, and modelling the implicit negative feedback can guide the recommendation model to filter out the news that users are not interested in. Because there may be content of interest to users in the news that is not clicked, that is, there is preference noise, which leads to interference with modelling implicit negative feedback. In addition, due to the diversification and variability of user interests, the existing news recommendation system often has the problem of an “information cocoon”. To solve the above problems, this paper proposes a news recommendation algorithm based on users’ static and dynamic interests and the denoised implicit negative feedback. By fusing and modeling the static and dynamic interests of users, and the denoised implicit negative feedback of static and dynamic interests, the dynamically updated user preference model is constructed. Firstly, a static interest denoising module based on orthogonal mapping is constructed to denoise the implicit negative feedback in static interest. Then, the GRU and orthogonal mapping are fused to construct a dynamic interest denoising module based on the improved GRU, which fully models the user’s interest change and realizes the denoising of the implicit negative feedback in the dynamic interest. Finally, by introducing contrastive learning technology, the model’s ability to distinguish between implicit positive and negative feedback is enhanced to improve the performance of personalized news recommendations. Experiments on the MIND dataset show that compared with the baseline method, the model improves by 1.18%, 1.84%, 2.75% and 1.67% on the four evaluation in-

到稿日期:2024-12-24 返修日期:2025-05-10

基金项目:国家自然科学基金(72074038);江苏省高等学校基础科学(自然科学)研究项目(24KJB520022);南京市留学人员科技创新项目(RK002NLX23004);南京邮电大学引进人才自然科学研究启动基金(NY223030);江苏省研究生科研与实践创新计划项目(KYCX22\_0950)

This work was supported by the National Natural Science Foundation of China(72074038), Natural Science Foundation of the Higher Education Institutions of Jiangsu Province (24KJB520022), Nanjing Science and Technology Innovation Foundation for Overseas Students (RK002NLX23004), Natural Science Research Start-up Foundation of Recruiting Talents of Nanjing University of Posts and Telecommunications (NY223030) and Postgraduate Research & Practice Innovation Program of Jiangsu Province(KYCX22\_0950).

通信作者:卢官明(lugm@njupt.edu.cn)

dexes of AUC, MRR, NDCG@5 and NDCG@10, respectively, which verifies the effectiveness of the proposed model.

**Keywords** News recommendation, Information cocoon, Orthogonal mapping, Gated recurrent unit, Contrastive learning

## 1 引言

随着互联网时代的发展,众多用户接收新闻的媒介从传统报纸转变为在线新闻平台(如今日头条等)。为了从海量新闻中找到用户可能感兴趣的新闻,个性化新闻推荐系统应运而生。此外,个性化新闻推荐系统对在线新闻平台缓解信息过载、提升用户阅读体验起着至关重要的作用。

与电影推荐或电子商品推荐不同,人们通常不会对推荐的新闻进行评分或评论。因此,用户对新闻的反馈总是隐式的(点击行为、阅读时间等)。用户在浏览新闻时,每次点击行为都会被系统记录下来,用户历史点击的新闻被视为隐式正反馈,而未点击的新闻被视为隐式负反馈。现有的一些新闻推荐系统<sup>[1-2]</sup>在隐式正反馈的基础上引入隐式负反馈,通过对隐式正反馈和负反馈的建模,分别引导模型学习用户感兴趣的新闻和过滤用户不感兴趣的新闻。然而,在隐式负反馈中可能存在用户感兴趣的内容,即隐式负反馈中存在偏好噪声,例如,用户同时被推荐多个感兴趣的新闻,却只点击了其中一个,而当时未点击的新闻,也可能在稍后会再次引起他们的兴趣。偏好噪声的存在造成了对隐式负反馈建模的干扰,进而降低了推荐模型对用户感兴趣内容的建模能力。

用户可能会习惯性地阅读某一类新闻,即用户拥有静态的长期兴趣(静态兴趣)。如果近期有全民关注的时事热点,用户则倾向于点击该话题相关的新闻。由于新闻具有时效性,用户对这类新闻的兴趣会随着热度的降低而减弱,即用户拥有动态的短期兴趣(动态兴趣)。因此,静态兴趣和动态兴趣对新闻推荐的性能至关重要。另外,从用户的静态兴趣角度来看,隐式负反馈中应该去除其稳定的长期偏好;从用户的动态兴趣角度来看,由于用户的兴趣是动态变化的,某一时刻隐式负反馈中的偏好噪声就是该用户当时感兴趣的新闻。例如,用户当时对“美国大选”话题感兴趣,那么对该用户隐式负反馈进行去噪时应该去除有关“美国大选”的新闻。然而,现有新闻推荐算法没有充分利用和挖掘动态兴趣和静态兴趣,且未有效去除动态兴趣和静态兴趣中隐式负反馈的偏好噪声,容易推荐与用户之前阅读内容相似的新闻,导致出现“信息茧房”问题。

针对上述问题,本文提出了一种基于用户静态兴趣和去噪隐式负反馈的新闻推荐算法,通过对用户的静态和动态兴趣进行融合建模,并对静态兴趣隐式负反馈进行去噪,构建动态更新的用户偏好模型。本文的主要贡献如下:

1)提出了一种基于正交映射的静态兴趣去噪方法,以隐式正反馈为正交映射的基,去除静态兴趣隐式负反馈中的偏好噪声。

2)提出了一种基于改进门控循环单元(Gated Recurrent Unit, GRU)的动态兴趣去噪方法,将 GRU 和正交映射相融合,同时捕捉用户隐式正反馈和隐式负反馈的兴趣变化,并对每一个时间步动态兴趣中的隐式负反馈进行去噪。

3)提出了一种基于对比学习的隐式正、负反馈判别方法。

计算隐式正反馈和隐式负反馈分数,并通过引入对比损失函数,最大化隐式正反馈与隐式负反馈分数的差值,增强模型对隐式正反馈和负反馈的区分能力。

4)在 MIND 数据集上的实验结果表明,本文算法在 AUC, MRR, NDCG@5, NDCG@10 这 4 种评价指标上均优于其他基线算法,验证了其有效性。

## 2 相关工作

### 2.1 新闻推荐

传统的新闻推荐系统主要采用基于内容的算法<sup>[3]</sup>、基于协同过滤<sup>[4]</sup>(Collaborative Filtering, CF)的算法以及结合不同推荐算法的混合推荐算法<sup>[5]</sup>。这些算法往往受数据稀疏性的影响,没有充分考虑新闻的时效性以及用户兴趣的多样化和多变性,具有一定的局限性。

随着深度学习的发展,研究人员将深度学习方法应用于新闻推荐任务中,取得了显著进展。基于深度学习的新闻建模旨在从原始输入中自动学习新闻嵌入表示。Wang 等<sup>[6]</sup>使用知识感知的卷积神经网络(Convolutional Neural Network, CNN)学习新闻的表示。为了更有效地捕捉新闻文本序列中的语义依赖关系,Wu 等<sup>[7]</sup>采用词级注意力机制识别新闻标题中重要的单词。Wu 等<sup>[8]</sup>使用多头自注意力机制和加性注意力机制捕捉不同语义空间中新闻文本的多种特征,学习更准确的新闻表示。上述文献主要聚焦于建模新闻的标题。为了构建更丰富的新闻表示,Qi 等<sup>[9]</sup>综合利用新闻的标题、摘要、类别等信息,将用户点击的新闻串联成一个长文档,采用 Fastformer 同时捕捉新闻内部和新闻之间的词级交互。Tran 等<sup>[10]</sup>建模了新闻的标题、正文、类别、子类别以及从 WikiData 知识图谱中提取的实体信息。

用户建模通常是从学习到的新闻表示中推断用户的兴趣和偏好。Wu 等<sup>[7]</sup>采用新闻级注意力机制识别用户点击历史中重要的新闻。由于用户历史点击序列具有顺序性,An 等<sup>[11]</sup>通过引入用户 ID 的嵌入来捕捉用户的长期兴趣,采用 GRU 从用户最近浏览的新闻中学习用户的短期表征。Qi 等<sup>[12]</sup>采用候选新闻感知的 CNN 学习用户的短期兴趣。Wu 等<sup>[2]</sup>使用用户 ID 向量作为用户的静态兴趣表征,基于候选新闻感知的注意力网络动态生成用户兴趣表征,通过在静态兴趣和动态兴趣之间加入对比损失函数,实时更新静态兴趣。Tran 等<sup>[10]</sup>通过关注用户近几天最频繁的阅读主题来推断用户的长期偏好,采用 GRU 学习用户的短期偏好。

为了缓解推荐模型建模时的数据稀疏问题,研究人员考虑引入其他类型的特征。Wu 等<sup>[13]</sup>结合了点击行为、跳过或完成阅读、快速关闭、分享、不喜欢等显式和隐式反馈,虽然在一定程度上缓解了数据稀疏性的问题,但未能充分挖掘用户行为数据中潜藏的完整信息。现有工作大多聚焦于已发生的正向行为和部分常见的反馈而忽视了用户历史未点击新闻背后蕴含的丰富信息,这部分数据同样对理解用户兴趣偏好、规避不感兴趣内容有着重要意义。Wu 等<sup>[2]</sup>提出对比注意力

机制,借助用户曝光未点击新闻融合建模用户兴趣,突出用户兴趣偏好,并构建多维度兴趣,通过对比学习实现不同维度兴趣相互促进更新。用户没有点击某个新闻并不一定是对其不感兴趣,可能是浏览过类似的新闻,所以直接建模用户历史未点击新闻并不能代表用户不感兴趣的内容。

新闻标题作为新闻内容的高度概括,能够为模型提供足够的信息来捕捉用户的兴趣和反馈特征。Hu等<sup>[1]</sup>设计了一个新的模型,将新闻标题作为输入,去除用户历史未点击新闻中包含的噪声。该研究工作将未点击的新闻作为隐式负反馈,并认为不仅正反馈中包含噪声,负反馈中也包含噪声。该模型计算新闻与正负反馈的相似性以判断是否为噪声,并进行隐式负反馈的去噪。上述方法虽然引入了隐式负反馈,但是只从用户的静态兴趣角度去除隐式负反馈中包含的偏好噪声。与这些方法不同的是,本文模型不仅去除静态兴趣隐式负反馈中的偏好噪声,而且同时捕捉用户隐式正反馈和负反馈的兴趣变化,并对每一个时间步动态兴趣中的隐式负反馈进行去噪,更加充分地为用户历史未点击新闻进行建模。

## 2.2 对比学习

对比学习自主地从大量数据中学习同类实例之间的共同特征,区分非同类实例之间的不同之处。对比学习的训练目标是从同类实例数据中学到的特征是相似的,在嵌入空间中特征向量尽量靠近,从不同类实例数据中学习到的特征向量在嵌入空间中尽量远离。

近些年,对比学习技术在计算机视觉和自然语言处理等领域取得了重要的进展。对比学习的主要内容就是确定如何划分正负样本,它可以灵活地适用于不同的任务,因此现有工作也尝试将对比学习技术应用于新闻推荐任务中,以提高推荐系统的性能。例如,Liu等<sup>[14]</sup>认为流行度较高的事件可能会被多个出版商报道,因此按照文本相似性聚合文档并通过文档日期区分正负样本,采用对比学习在嵌入空间中学习增强的文档三元组表示。Yang等<sup>[15]</sup>认为知识图结构一致性得分越高的项目包含的噪声越少,对用户真实兴趣的建模贡献更大,因此使用两个掩码向量为知识图生成两个子图,在两个子图之间引入对比学习技术。Yu等<sup>[16]</sup>考虑到直接用下游新闻推荐任务对预训练语言模型(Pre-trained Language Model, PLM)进行微调不足以填补一般语料库和新闻文本之间的领域空白,因此利用新闻文章不同部分的自然匹配关系,在新闻标题和新闻正文之间引入自监督对比匹配任务。Wang等<sup>[17]</sup>构建了多个兴趣原型,设计了一个多兴趣用户编码器以学习不同兴趣条件下的多个用户表示,同时采用图增强的用户编码器,通过聚合邻居的相关信息来丰富用户在每个兴趣背景下对应的语义表示,对用户多兴趣表示和兴趣原型向量进行对比,以优化用户多兴趣表示。Shin等<sup>[18]</sup>为多个任务构建用户表征,在不同任务间引入对比学习。Xie等<sup>[19]</sup>分别引入特征级、语义级、会话级项目对比学习。特征级项目对比学习采用特征随机扰动生成特征增强作为正样本,优化项目特征表示;语义级项目对比学习将同类别下的项目视为正样本;会话级项目对比学习将经常在同一会话中出现的项目视为正样本。本文将用户历史点击新闻标题视为正样本,将去噪后的用户历史未点击新闻标题视为负样本。通过引入对比学习

技术,使得模型在特征空间更好地区分隐式正反馈和负反馈。

## 3 本文方法

本文模型的整体结构如图1所示。模型主要由新闻编码器、用户编码器和预测层3部分构成,用户编码器中又包含内容编码器、静态兴趣去噪模块、动态兴趣去噪模块和聚合层。新闻编码器对新闻标题进行建模得到新闻表示,用户编码器用于建模用户表示,预测层根据候选新闻表示与用户表示来预测用户点击候选新闻的概率。

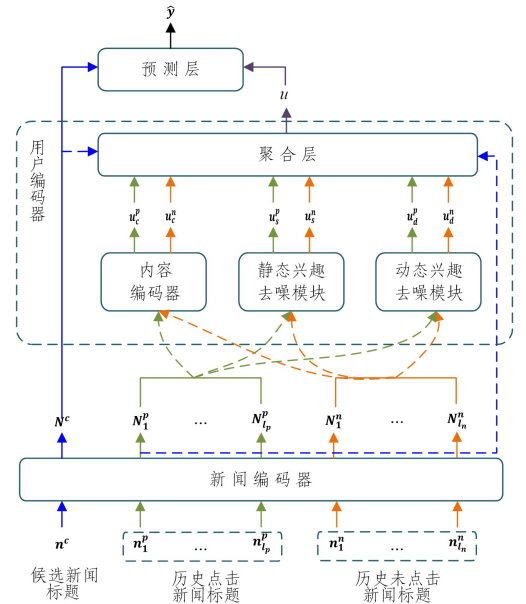


图1 本文模型结构

Fig. 1 Framework of the proposed model

### 3.1 问题定义

定义用户集合为  $U$ , 新闻集合为  $N$ , 设用户对新闻的反馈矩阵为  $Z = \{z_{u,n} | u \in U, n \in N\}$ , 其中:

$$z_{u,n} = \begin{cases} 1, & \text{如果用户 } u \text{ 点击新闻 } n \\ 0, & \text{新闻 } n \text{ 未被推荐给用户 } u \\ -1, & \text{如果用户 } u \text{ 未点击新闻 } n \end{cases} \quad (1)$$

对于每一个用户  $u$ , 可以从反馈矩阵  $Z$  中得到其历史隐式正反馈序列为  $[n_1^p, n_2^p, \dots, n_p^p]$ , 历史隐式负反馈序列为  $[n_1^n, n_2^n, \dots, n_n^n]$ , 其中  $l_p$  和  $l_n$  分别为隐式正反馈和隐式负反馈序列的长度。模型的目标是根据用户  $u$  的历史反馈计算其点击候选新闻  $n^c$  的概率。

### 3.2 新闻编码器

新闻编码器是新闻推荐的核心部件之一,目的是学习新闻的表示。本文选择新闻标题作为输入。对于每个新闻标题  $t$ , 使用嵌入矩阵  $E_w \in R^{l_w \times d}$  将标题中的每个单词转换成  $d$  维向量, 其中  $l_w$  为标题中的单词数量。新闻标题  $t$  经过嵌入层转换成矩阵  $T_{title}$ 。

本文的新闻编码器由多头自注意力机制和门控聚合层组成。将新闻标题的嵌入表示  $T_{title}$  输入多头自注意力机制中, 以捕捉新闻标题中单词的上下文表示。假设多头自注意力层的头部数量为  $l_h$ , 将输入  $T_{title}$  映射到不同的子空间, 分别从不同角度学习新闻标题中单词的上下文表示:

$$head_i = Attention(QW_i^q, KW_i^k, VW_i^v) \quad (2)$$

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d}}\right)\mathbf{V} \quad (3)$$

$$\text{MH}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = [\text{head}_1; \text{head}_2; \dots; \text{head}_{l_h}] \mathbf{W}^l \quad (4)$$

其中,  $\mathbf{W}_q^i, \mathbf{W}_k^i, \mathbf{W}_v^i \in R^{d \times d/l_h}$ ,  $\mathbf{W}^l \in R^{d \times d}$  均为可训练的参数矩阵;  $[\cdot; \cdot]$  表示矩阵按列拼接。

此外,模型使用残差连接和层归一化函数  $LN$  来融合新闻标题中单词的原始表示和上下文表示。

$$\tilde{\mathbf{T}}_{\text{title}} = LN(\mathbf{T}_{\text{title}} + \text{MH}(\mathbf{T}_{\text{title}}, \mathbf{T}_{\text{title}}, \mathbf{T}_{\text{title}})) \quad (5)$$

在得到新闻标题中单词上下文感知的表示后,将其输入门控聚合层,以选择重要的信息来生成信息丰富的新闻标题表示:

$$\hat{\mathbf{T}}_{\text{title}} = \text{Aggregate}(\tilde{\mathbf{T}}_{\text{title}}) \\ = (\text{softmax}(\tanh(\tilde{\mathbf{T}}_{\text{title}} \mathbf{W}_a + \mathbf{b}_a) \mathbf{W}_g))^T \tilde{\mathbf{T}}_{\text{title}} \quad (6)$$

其中,  $\mathbf{W}_a \in R^{d \times d}$ ,  $\mathbf{b}_a \in R^d$ ,  $\mathbf{W}_g \in R^{d \times 1}$  为可训练的参数矩阵;  $d'$  为门控聚合层的输出维度。

最后,将隐式正反馈新闻标题序列、隐式负反馈新闻标题序列和候选新闻输入到新闻编码器中得到隐式正反馈新闻标题表示  $\mathbf{N}^p = [\mathbf{N}_1^p, \mathbf{N}_2^p, \dots, \mathbf{N}_{l_p}^p]$ 、隐式负反馈新闻标题表示  $\mathbf{N}^n = [\mathbf{N}_1^n, \mathbf{N}_2^n, \dots, \mathbf{N}_{l_n}^n]$  和候选新闻表示  $\mathbf{N}^c$ 。

### 3.3 用户编码器

用户编码器通常用于从用户历史反馈中编码用户的兴趣向量。为了缓解信息茧房问题,模型同时编码用户的静态兴趣和动态兴趣,并对动静兴趣中的隐式负反馈进行去噪。本文的用户编码器分别由内容编码器、静态兴趣去噪模块和动态兴趣去噪模块组成。内容编码器用于从新闻编码器得到的新闻标题表示中编码出原始的用户表示;静态兴趣去噪模块用于编码用户的静态兴趣并从用户的静态兴趣角度对隐式负反馈进行去噪;动态兴趣去噪模块用于编码用户的动态兴趣并从用户的动态兴趣角度去除隐式负反馈中的偏好噪声。

#### 3.3.1 内容编码器

内容编码器就是从内容视图评估反馈序列中不同新闻标题的重要性,然后聚合重要新闻来表示用户的兴趣。内容编码器由多头自注意力层和门控聚合层组成,如图2所示。

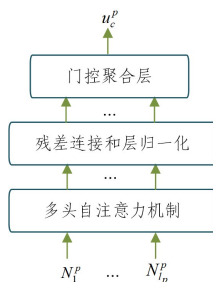


图2 内容编码器结构(以隐式正反馈为例)

Fig. 2 Framework of content encoder(taking implicit positive feedback as an example)

首先将用户隐式正反馈和隐式负反馈新闻标题表示分别输入到多头自注意力机制中捕捉新闻标题的上下文表示;然后使用残差连接和层归一化函数融合原始和上下文表示;最后使用门控聚合层选择更能反映用户兴趣的新闻来生成用户的表示。内容编码器生成的隐式正反馈用户表示和隐式负反

馈用户表示分别为  $u_c^p$  和  $u_c^n$ 。详细公式参考新闻编码器部分。

#### 3.3.2 静态兴趣去噪模块

静态兴趣去噪模块结构如图3所示。用户在浏览新闻网站的过程中被新闻标题吸引从而点击新闻,因此用户历史点击的新闻标题能够很好地反映用户兴趣,本文将用户历史点击新闻标题视为隐式正反馈。一方面,用户同时被推荐几个感兴趣的新闻,但只点击了其中一个;另一方面,一些用户此刻未点击的新闻,也可能在以后吸引他们的注意。因此历史未点击这一隐式负反馈中含有偏好噪声。为了从隐式负反馈中去掉用户可能感兴趣的部分,采用正交映射技术<sup>[20]</sup>,以隐式正反馈为基对隐式负反馈进行去噪。对于隐式正反馈序列,直接将其输入内容编码器中获得隐式正反馈用户表示,记为  $u_c^p$ 。对于隐式负反馈序列,首先将其输入正交映射模块,得到去噪后的隐式负反馈序列:

$$\mathbf{N}_p^n = \text{project}(\mathbf{N}^n, \mathbf{N}^p) \quad (7)$$

$$\text{project}(a, b) = \frac{a \cdot b}{|b|} \frac{b}{|b|} \quad (8)$$

然后将去噪后的隐式负反馈序列  $\mathbf{N}_p^n$  输入内容编码器,得到基于静态兴趣角度去噪后的隐式负反馈用户表示  $u_c^n$ 。

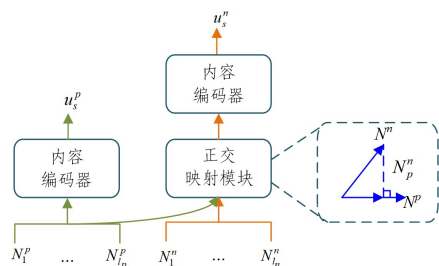


图3 静态兴趣去噪模块结构

Fig. 3 Framework of static interest denoising module

#### 3.3.3 动态兴趣去噪模块

动态兴趣去噪模块结构如图4所示。用户的兴趣是动态变化的,直接对输入新闻序列使用正交映射实现了从用户静态兴趣角度对隐式负反馈进行粗粒度去噪。但是,不同时刻的新闻对用户兴趣的影响程度是不同的,比如用户此刻对体育类新闻感兴趣,那么对此刻的隐式负反馈去噪就应该着重去除体育类新闻。如果用户对体育类新闻兴趣降低,那么对该时刻的隐式负反馈去噪就应该减少去除体育类新闻。因此在动态兴趣去噪模块采用GRU同时编码用户的隐式正反馈和隐式负反馈动态兴趣变化,对每个时刻的隐式负反馈以该时刻的隐式正反馈为基做正交映射,得到去噪后的隐式负反馈。对于隐式正反馈序列,首先将其输入到GRU中,得到每个时刻的隐藏状态。

$$\mathbf{r}_t = \sigma(\mathbf{W}_r [\mathbf{h}_{t-1}, \mathbf{x}_t]) \quad (9)$$

$$\mathbf{z}_t = \sigma(\mathbf{W}_z [\mathbf{h}_{t-1}, \mathbf{x}_t]) \quad (10)$$

$$\tilde{\mathbf{h}}_t = \tanh(\mathbf{W}_h [\mathbf{r}_t \odot \mathbf{h}_{t-1}, \mathbf{x}_t]) \quad (11)$$

$$\mathbf{h}_t = \mathbf{z}_t \odot \mathbf{h}_t + (1 - \mathbf{z}_t) \odot \tilde{\mathbf{h}}_t \quad (12)$$

其中,  $\sigma$  为 sigmoid 函数;  $\odot$  为哈达玛积;  $\mathbf{W}_r, \mathbf{W}_z, \mathbf{W}_h$  为 GRU 中的参数矩阵。此时,经过 GRU 编码的隐式正反馈序列为  $[\mathbf{h}_1^p, \mathbf{h}_2^p, \dots, \mathbf{h}_{l_p}^p]$ 。然后,将其输入门控聚合层得到动态兴趣去噪模块下的隐式正反馈用户表示,记为  $u_d^p$ 。

对于隐式负反馈序列,同样将其输入 GRU 中,对于  $t$  时刻的隐式负反馈隐藏状态,以  $t$  时刻的隐式正反馈隐藏状态为基进行正交映射,将得到的  $t$  时刻去噪后的隐式负反馈隐藏状态传递到下一个时刻。此时,经过 GRU 编码和正交映射去噪后的隐式负反馈序列为  $[h_1^-, h_2^-, \dots, h_n^-]$ 。然后,将其输入门控聚合层得到动态兴趣去噪模块下的隐式负反馈用户表示  $u_d^-$ 。

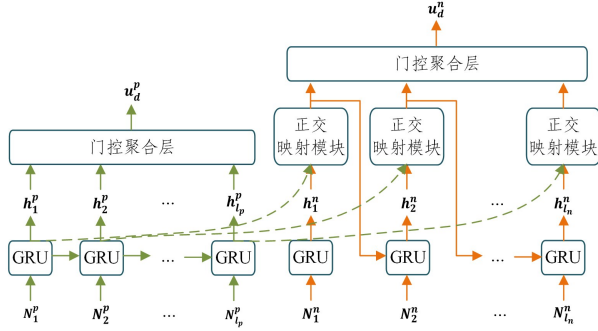


图4 动态兴趣去噪模块结构

Fig. 4 Framework of dynamic interest denoising module

### 3.3.4 聚合层

已知通过内容编码器、静态兴趣去噪模块和动态兴趣去噪模块得到 6 个用户隐式正反馈和隐式负反馈兴趣向量,分别为  $u_c^p, u_c^-, u_s^p, u_s^-, u_d^p, u_d^-$ , 聚合层旨在将这 6 个用户兴趣向量聚合成统一的用户表示,使用 6 个全连接层计算每个用户表示的权重,以  $u_c^p$  为例计算其权重  $\beta_c^p$ :

$$\beta_c^p = \tanh(f^T W_1^{p,c} + b_1^{p,c}) W_2^{p,c} + b_2^{p,c} \quad (13)$$

$$f^T = [u_c^p; N^c] \quad (14)$$

$$u_c^p = \text{Aggregate}([N^p | N^n]) \quad (15)$$

其中,  $W_1^{p,c} \in R^{2d \times d}$ ,  $W_2^{p,c} \in R^{d \times 1}$ ,  $b_1^{p,c} \in R^d$ ,  $b_2^{p,c} \in R$  均为可学习参数;  $[|]$  表示矩阵逐行拼接。同理可得  $u_c^-, u_s^p, u_s^-, u_d^p, u_d^-$  的权重分别为  $\beta_c^-, \beta_s^p, \beta_s^-, \beta_d^p, \beta_d^-$ 。因此,最终用户表示为:

$$u = \beta_c^p * u_c^p + \beta_c^- * u_c^- + \beta_s^p * u_s^p + \beta_s^- * u_s^- + \beta_d^p * u_d^p + \beta_d^- * u_d^- \quad (16)$$

### 3.4 预测层

预测层用于预测用户点击候选新闻的概率分数。受 Okura 等<sup>[21]</sup>的启发,通过将用户编码器得到的最终用户表示  $u$  和经过新闻编码器编码得到的候选新闻表示  $N^c$  做内积来计算预测分数  $\hat{y}$ :

$$\hat{y} = u^T N^c \quad (17)$$

### 3.5 对比学习模块

对比学习通过比较数据点的相似性和差异性来学习无标签数据集的特征。为了进一步在特征空间区分隐式正反馈和隐式负反馈,增强模型对是否点击候选新闻的判别能力,在隐式正反馈和隐式负反馈之间引入对比学习技术。对比学习技术的关键是构建正负样本,首先根据聚合层得到的隐式正负反馈用户表示的权重计算得到隐式正反馈用户表示  $u^p$  和隐式负反馈用户表示  $u^n$ :

$$u^p = \beta_c^p * N_c^p + \beta_s^p * N_s^p + \beta_d^p * N_d^p \quad (18)$$

$$u^n = \beta_c^- * N_c^- + \beta_s^- * N_s^- + \beta_d^- * N_d^- \quad (19)$$

然后分别将隐式正反馈用户表示  $u^p$  和隐式负反馈用户表示  $u^n$  与候选新闻表示  $N^c$  做内积,得到隐式正反馈分数  $\hat{y}^p$  和隐式负反馈分数  $\hat{y}^n$ 。在隐式正反馈分数和隐式负反馈分数之间加入对比损失函数  $\mathcal{L}^c$ ,最大化隐式正反馈分数与隐式负反馈分数之间的差值:

$$\mathcal{L}^c = -\log \sigma(\hat{y}^p - \hat{y}^n) \quad (20)$$

### 3.6 损失函数

本文的损失函数由交叉熵损失函数和对比损失函数两部分构成。对于每个正样本,随机从同一个用户中抽取  $l_k$  个负样本,构建为  $l_k + 1$  分类任务。将正样本的点击概率得分记为  $\hat{y}_i^+$ , 负样本的点击概率得分记为  $\hat{y}_i^-$ , 则交叉熵损失函数  $\mathcal{L}^e$  的计算式为:

$$\mathcal{L}^e = -\sum_{i \in P} \log \left( \frac{\exp(\hat{y}_i^+)}{\exp(\hat{y}_i^+) + \sum_{j=1}^{j=l_k} \exp(\hat{y}_{i,j}^-)} \right) \quad (21)$$

其中,  $P$  为正样本的集合。因此,本文的总损失函数为:

$$\mathcal{L} = \lambda \mathcal{L}^e + (1 - \lambda) \mathcal{L}^c \quad (22)$$

其中,  $\lambda$  为可学习的参数。

## 4 实验及其结果分析

### 4.1 实验设置

#### 4.1.1 数据集

本文采用 MIND<sup>[22]</sup>数据集,该数据集为大型新闻推荐数据集,被广泛用于新闻推荐模型的实验。MIND 数据集是从 Microsoft News 的用户行为日志中收集的,随机抽取了 100 万名用户在 2019 年 10 月 12 日至 11 月 22 日 6 周内的新闻点击记录,并将其格式化为 impression logs,记录了用户在特定时间访问新闻网站首页时显示给用户的新闻文章,以及用户对这些新闻文章的点击行为。MIND 数据集的每一篇文章都包含了新闻标题、摘要、正文、类别等文本信息。

本文选取 MIND 数据集前 5 天的用户行为日志作为用户历史行为序列,将用户点击的新闻放入隐式正反馈序列中,用户浏览但未点击的新闻放入隐式负反馈序列中,这样用户历史行为序列就由隐式正反馈序列和隐式负反馈序列两部分构成。对于数据集的划分,将原始训练集第 6 天的用户行为日志用于训练;将原始验证集前 10% 的用户行为日志作为验证集;在划分测试集时,将原始测试集的后 90% 的用户行为日志作为测试集。处理后的数据集统计数据如表 1 所列。

表 1 MIND 数据集统计信息

Table 1 Statistics of MIND dataset

统计信息	统计数量
用户数	654870
新闻数	104153
标题平均长度	12.66
隐式正反馈数量	1048414
隐式负反馈数量	25145229

#### 4.1.2 参数设置

模型采用新闻的标题作为输入,将新闻标题的最大长度设置为 15。采用 GloVe<sup>[23]</sup>词向量初始化新闻标题的嵌入,嵌入向量的维度  $d$  设置为 300。多头自注意机制的头部数量  $l_h$

为6。门控聚合层的输出维度 $d'$ 为200。训练时将负采样率 $l_n$ 设置为4,即每个正样本对应4个负样本。参考Hu等<sup>[1]</sup>的工作,将隐式正反馈序列长度 $l_p$ 设置为30。实验时,分别将隐式负反馈序列长度设置为 $\{0,30,60,90,120\}$ ,由于长度为60时模型效果最佳,因此将隐式负反馈序列长度 $l_n$ 设置为60。模型训练的 batch size 为128,学习率为0.0001,优化器为Adam。损失函数初始权重 $\lambda$ 在 $[0,1]$ 中取值,值为0.9时模型效果最佳,因此将 $\lambda$ 设置为0.9。

#### 4.1.3 评价指标

为了衡量推荐模型的性能效果,本文选择 AUC, MRR, NDCG@5, NDCG@10 这4个评价指标。AUC 是 ROC 曲线与坐标值所围成的面积,用于评价正样本排名在负样本前面的概率,AUC 的值越大,模型的分类效果越好;MRR 为平均倒数排名,用正确检索结果的排名倒数平均值来反映推荐系统的性能,MRR 越大,模型的性能越好;NDCG@ $k$  为归一化折损累积增益,用于对搜索排序的结果进行评价,NDCG@ $k$  越大,推荐效果越好,其中, $k$  为排序列表的长度。

#### 4.1.4 基线模型

本文选取了以下8个新闻推荐模型作为基线模型。

- 1) LibFM<sup>[24]</sup>:一种基于因子分解机的经典推荐方法。
- 2) DeepFM<sup>[25]</sup>:融合了因子分解机与神经网络。
- 3) LSTUR<sup>[11]</sup>:使用 GRU 建模用户的长期和短期兴趣。
- 4) NPA<sup>[7]</sup>:根据用户的偏好选择重要的词汇和新闻生成更具信息性的新闻和用户表示。
- 5) NRMS<sup>[8]</sup>:使用多头自注意力机制和加性注意力机制学习新闻和用户的表示。
- 6) FUM<sup>[9]</sup>:采用 Fastformer 建模细粒度的用户兴趣。
- 7) DRPN-G<sup>[1]</sup>:联合建模隐式正反馈和隐式负反馈。
- 8) Cupmar<sup>[10]</sup>:编码新闻的多个属性,通过用户历史点击新闻来学习长期兴趣,使用 GRU 网络捕捉近期兴趣变化。

本文模型采用基于注意力机制的新闻编码器,并且编码了用户的静态兴趣,因此选取 NPA 和 NRMS 这两个基于注意力机制的新闻推荐模型以及 LSTUR 和 Cupmar 这两个编码了用户长短期兴趣的新闻推荐模型。除此之外,为了对比基于人工构建特征的新闻推荐算法,选取 LibFM 和 DeepFM 作为基线模型。由于本文模型仅编码了新闻的标题,因此选取 FUM 和 Cupmar 这两个编码新闻多信息的新闻推荐模型。DRPN-G 和本文模型都关注隐式正负反馈建模以及反馈信息去噪问题,因此将其选作为基线模型。

## 4.2 实验结果与分析

### 4.2.1 基线模型对比

不同新闻推荐模型的实验结果如表2所列。实验结果表明,本文模型的所有指标均优于 DRPN-G,说明本文提出的改进算法是有效的。本文模型分别从静态兴趣和动态兴趣角度对隐式负反馈中包含的噪声进行去噪,进一步优化了隐式负反馈的表示,提高了推荐的准确性。此外,从表2中还可以看出,LibFM 和 DeepFM 这两个基于人工构建特征的推荐方法效果明显比其他基于深度学习的新闻推荐算法差,说明深层的神经网络可以学习到更准确的新闻表示。

表2 不同模型实验结果对比

Table 2 Experimental results comparison of different models

模型	AUC	MRR	NDCG@5	NDCG@10
LibFM	61.32	27.43	29.17	35.54
DeepFM	63.28	28.24	30.40	36.72
LSTUR	65.81	30.82	34.10	40.33
NPA	65.12	30.61	33.64	39.92
NRMS	66.00	31.01	34.28	40.71
FUM	66.93	31.58	35.01	41.41
DRPN-G	66.34	31.51	34.85	41.25
Cupmar	<b>67.83</b>	<b>33.32</b>	<b>34.35</b>	<b>41.02</b>
本文模型	67.21	32.09	<b>35.81</b>	<b>41.94</b>

### 4.2.2 参数影响分析

基于 MIND 数据集,在其他参数不变的情况下,记录了取不同损失函数权重 $\lambda$ 的初始值时,模型在 AUC 和 NDCG@10 这两个指标上的实验结果,如图5所示。当 $\lambda$ 值为1时,交叉熵损失函数的权重为0,此时 AUC 的值为46.52, NDCG@10 的值为26.41,模型效果最差。为了改善这一情况,将 $\lambda$ 取值设定为0.99进行实验。由图5可以看出,当 $\lambda$ 的初始值为0.9时,模型的效果最好,说明该值可以有效地平衡对比学习损失和交叉熵损失,使模型能够学习更具区分性的特征,并提高模型的泛化能力。

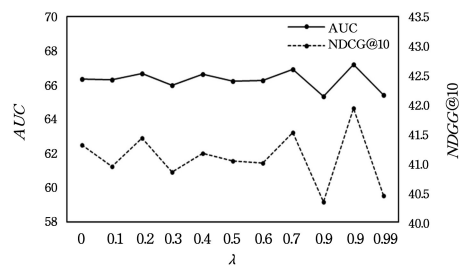


图5 不同 $\lambda$ 取值对结果的影响

Fig. 5 Effects of different  $\lambda$  values on the results

同样地,图6记录了隐式负反馈序列长度 $l_n$ 分别取0,30,60,90,120时,在 AUC 和 NDCG@10 这两个指标上的实验结果。从图中可以看出,当隐式负反馈序列长度为60时,模型的效果最佳。因为采用较少的隐式负反馈序列时,经过去噪处理后的隐式负反馈表示并不足以充分建模用户不感兴趣的部分,而采用更多的隐式负反馈序列时引入的噪声也更多。

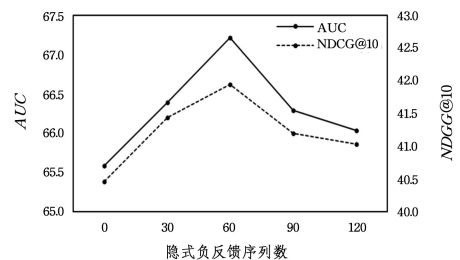


图6 不同隐式负反馈序列长度对结果的影响

Fig. 6 Effects of different sequence lengths of implicit negative feedback on the results

### 4.2.3 消融实验

为了验证本文模型各个模块的有效性,分别去掉了模型中的静态兴趣去噪模块、动态兴趣去噪模块和对比损失模块,在 MIND 数据集上对其进行消融实验,在模型参数不变的情况下,对比结果如表3所列。

表3 消融实验结果

Table 3 Results of ablation experiments

模型	AUC	MRR	NDCG@5	NDCG@10
本文模型	67.21	<b>32.09</b>	<b>35.81</b>	<b>41.94</b>
w/o 静态兴趣去噪模块	66.32	31.54	34.94	41.11
w/o 动态兴趣去噪模块	65.33	31.19	34.71	40.82
w/o 对比学习模块	<b>67.31</b>	31.89	35.51	41.74

从表3可以看出,去除静态兴趣去噪模块和去除动态兴趣去噪模块,在各个指标上模型的性能均有明显下降,说明了这两个模块的有效性,并且静态兴趣去噪模块和动态兴趣去噪模块是相辅相成的。其中,去除动态兴趣去噪模块的性能下降更为明显,说明从用户的实时兴趣角度出发,动态地对隐式负反馈进行去噪,可以捕获更加贴近用户真实兴趣的用户表示。

此外,去除对比学习模块的模型的 AUC 指标提高了

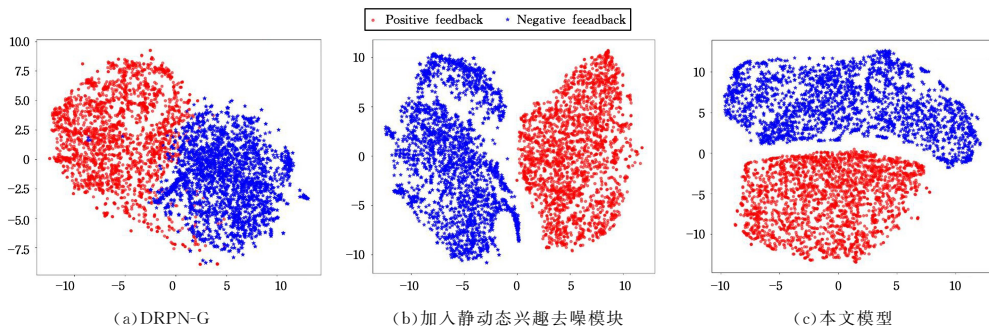


图7 不同模型隐式正、负反馈特征的可视化对比

Fig. 7 Visualization comparison of implicit positive and negative feedback features of different models

图7(a)展示了 DRPN-G 模型的隐式正反馈与隐式负反馈在特征空间的分布情况。从图中可以看出,隐式正反馈和隐式负反馈在特征空间呈现高度混杂的状态,没有明显的分离迹象,模型难以对其进行区分,无法精准捕捉用户偏好。

图7(b)是加入静态兴趣去噪模块后的特征空间分布。隐式正反馈和隐式负反馈的特征点各自变得相对集中,两者之间开始出现分离趋势,说明静态兴趣去噪模块成功减少了隐式负反馈中的噪声,增强了隐式正负反馈的特征区分度。

图7(c)是加入静态兴趣去噪模块和对比学习模块后的特征空间分布。相较于图7(b),隐式正反馈和隐式负反馈各自的簇变得更加聚集,二者之间的分离也更加明显,说明在隐式正负反馈之间引入对比损失模块,能够引导模型学习更具区分性的隐式正负反馈特征表示。

#### 4.2.5 样例分析

为了直观展示本文模型的有效性,从测试集中选择用户

0.1,但是 MRR, NDCG@5 和 NDCG@10 指标分别下降了 0.2, 0.3, 0.2, 说明在未明显影响 AUC 指标的情况下,对比学习模块在提升模型性能方面是有效的。AUC 指标衡量模型分类效果,对比损失模块虽然在模型训练过程中拉近了隐式正反馈之间的距离,扩大了隐式正反馈与隐式负反馈之间的距离,但是候选新闻中的负样本除了包含用户不感兴趣的新闻以外,还可能包含重复的用户感兴趣的新闻。经过对比损失模块,模型更趋向于推荐用户感兴趣的新闻,但是当用户感兴趣的新闻出现重复时,用户点击它的概率可能会降低, AUC 指标随之降低。

#### 4.2.4 可视化分析

为了直观地评估本文提出的创新模块对模型性能的影响,对不同模块配置下的模型隐式正反馈和隐式负反馈在特征空间的分布进行了可视化分析,如图7所示。

U541020 的新闻点击预测为例进行分析,结果如表4和表5所列。表4展示了用户历史点击新闻和未点击新闻的新闻id、类别等信息。表5以用户 U541020 候选新闻中“sports”类新闻为例,展示采用不同模型对该候选新闻的预测结果。可以看出,用户历史点击新闻中包含了类别为“sports”的新闻,说明用户对“sports”是感兴趣的。用户历史未点击新闻中同样包含了多个“sports”类新闻,可能是用户被同时推荐多个该类别的新闻,但是只点击了其中一个。由此可见,用户历史未点击新闻中存在用户可能感兴趣的偏好噪声,需要对其进行去噪。采用 DRPN-G 模型得到的点击预测分数为负数,说明用户倾向于不点击该候选新闻,与实际点击情况不符。采用本文模型后点击预测分数大幅提升,故认为用户倾向于点击该候选新闻。该样例说明本文模型能够有效去除隐式负反馈中包含的偏好噪声,优化模型对用户兴趣的捕捉能力,从而提高推荐的准确性。

表4 以用户 U541030 为例的模型有效性样例分析

Table 4 Sample analysis of model effectiveness with user U541030 as an example

新闻 id	类别	子类别	标题
N18515	sports	football_nfl	Trent Williams fails physical, experiencing discomfort with helmet
N26935	tv	tv-celebrity	Former Deadliest Catch Star Jerod Sechrist Arrested, Charged with Heroin Possession
N62019	tv	tvnews	Survivor pays tribute to 'icon' Rudy Boesch on show
N112751	foodanddrink	foodnews	Bill Murray Applied To Work At An Airport P. F. Chang's And They're Totally Down To Hire Him
N60755	sports	football_nfl	Eagles to sign wide receiver Jordan Matthews

用户 U541030  
的历史点击  
新闻

(续表)

	新闻 id	类别	子类别	标题
	N90968	sports	baseball_mlb	Mets' Alonso, Astros' Alvarez named Rookies of the Year
	N67459	news	elections-2020-us	AP sources: Deval Patrick mulling Democratic White House run
	N27097	news	newscime	Utah death-row inmate featured in best-selling book dies
	N25075	sports	baseball_mlb	Nationals GM Mike Rizzo on White House visit: 'We weren't trying to make a political statement'
用户 U541030 的历史未 点击新闻	N113741	news	newsworld	Deer-like animal thought lost to science photographed for first time in 30 years
	N93254	foodanddrink	tipsandtricks	What's the Best Oil for Frying?
	N31958	entertainment	celebrity	PEOPLE's Sexiest Men Alive; Still Sexy After All These Years
	N40664	finance	finance-savemoney	Is This Expense Busting Your Budget?
	N67199	movies	movies-celebrity	Jane Fonda Opens Up About the Possibility of Spending Her 82nd Birthday in Jail
	N425	sports	football_nfl_videos	Every touchdown from every Week 10 game

表5 “sports”类候选新闻的预测结果对比

Table 5 Comparison of the prediction results of “sports” class candidate news

模型	点击预测分数	预测是否点击	实际是否点击
DRPN-G	-0.132728844881057	否	是
本文模型	1.09048426151275	是	是

**结束语** 本文引入用户历史未点击新闻作为用户隐式负反馈,提出了一种基于用户静动态兴趣和去噪隐式负反馈的新闻推荐算法,分别从用户静态兴趣和动态兴趣角度对隐式负反馈进行去噪,并且同时编码了用户静态和动态的隐式正反馈和隐式负反馈兴趣,增强了用户的兴趣表示,缓解了“信息茧房”现象。在此基础上,本文采用对比损失函数进一步区分隐式正反馈和隐式负反馈,提取更加精确的隐式负反馈特征。实验结果表明,提出的算法能够有效去除隐式负反馈中包含的噪声,提高推荐模型的准确性,取得了具有竞争力的性能。

本文模型虽然通过建模隐式正负反馈并改进隐式负反馈去噪算法提升了模型指标,但仍存在一些不足之处。具体来说,本文模型仅利用新闻标题进行编码,未充分利用新闻的其他信息;模型结构的复杂度和表达能力相对有限;对比学习策略在样本选择和损失函数设计上较为简单。未来工作将探索融合新闻的多种信息并引入预训练语言模型提升语义表达能力,同时改进对比学习策略以更好地挖掘样本间的语义关系。

## 参考文献

- [1] HU Y F, QIU Z P, WU X. Denoising Neural Network for News Recommendation with Positive and Negative Implicit Feedback [C]// Findings of the Association for Computational Linguistics; NAACL 2022. Association for Computational Linguistics, 2022; 2320-2329.
- [2] WU J L, CUI X H. News Recommendation Algorithm Based on Users' Implicit Feedback Signals and Multi-Dimensional Interests [J]. Computer Engineering and Applications, 2024, 60(12): 101-110.
- [3] KOMPAN M, BIELIKOVÁ M. Content-based News Recommendation [C]// E-Commerce and Web Technologies; 11th International Conference. Berlin: Springer, 2010; 61-72.
- [4] DAS A S, DATAR M, GARG A, et al. Google News Personalization: Scalable Online Collaborative Filtering [C]// The 16th International Conference on World Wide Web. 2007; 271-280.
- [5] LIU J, SONG J B, LI C, et al. A Hybrid News Recommendation Algorithm based on K-means Clustering and Collaborative Filtering [J]. Journal of Physics: Conference Series. IOP Publishing, 2021, 1881: 032050.
- [6] WANG H W, ZHANG F Z, XIE X, et al. DKN: Deep Knowledge-aware Network for News Recommendation [C]// The 2018 World Wide Web Conference. 2018; 1835-1844.
- [7] WU C H, WU F Z, AN M X, et al. NPA: Neural News Recommendation with Personalized Attention [C]// The 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2019; 2576-2584.
- [8] WU C H, WU F Z, GE S Y, et al. Neural News Recommendation with Multi-Head Self-Attention [C]// The 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP). 2019; 6389-6394.
- [9] QI T, WU F Z, WU C H, et al. FUM: Fine-Grained and Fast User Modeling for News Recommendation [C]// Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2022; 1974-1978.
- [10] TRAN D H, SHENG Q Z, ZHANG W E, et al. CupMar: A Deep Learning Model for Personalized News Recommendation Based on Contextual User-Profile and Multi-Aspect Article Representation [J]. World Wide Web, 2023, 26(2): 713-732.
- [11] AN M X, WU F Z, WU C H, et al. Neural News Recommendation with Long- and Short-term User Representations [C]// The 57th Annual Meeting of the Association for Computational Linguistics. 2019; 336-345.
- [12] QI T, WU F Z, WU C H, et al. News Recommendation with Candidate-Aware User Modeling [C]// Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2022; 1917-1921.
- [13] WU C H, WU F Z, QI T, et al. Feedrec: News Feed Recommendation with Various User Feedbacks [C]// The ACM Web Conference 2022. 2022; 2088-2097.
- [14] LIU J L, LIU T Q, YU C. Newseembed: Modeling News Through Pre-trained Document Representations [C]// The 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining. 2021; 1076-1086.
- [15] YANG Y H, HUANG C, XIA L H, et al. Knowledge Graph Contrastive Learning for Recommendation [C]// The 45th Inter-

national ACM SIGIR Conference on Research and Development in Information Retrieval, 2022:1434-1443.

- [16] YU Y, WU F Z, WU C H, et al. Tiny-newsrec: Effective and Efficient PLM-based News Recommendation[C]// The 2022 Conference on Empirical Methods in Natural Language Processing, 2022:5478-5489.
- [17] WANG S C, GUO S, WANG L H, et al. Multi-Interest Extraction Joint with Contrastive Learning for News Recommendation [C] // Joint European Conference on Machine Learning and Knowledge Discovery in Databases, Cham: Springer, 2022: 606-621.
- [18] SHIN K, KWAK H, KIM S Y, et al. Scaling Law for Recommendation Models: Towards General-purpose User Representations[C] // The AAAI Conference on Artificial Intelligence, 2023:4596-4604.
- [19] XIE R B, QIU Z J, ZHANG B, et al. Multi-Granularity Item-based Contrastive Recommendation[C] // International Conference on Database Systems for Advanced Applications, Cham: Springer, 2023: 406-416.
- [20] BIAN Z, ZHOU S J, FU H, et al. Denoising User-aware Memory Network for Recommendation[C]// The 15th ACM Conference on Recommender Systems, 2021:400-410.
- [21] OKURA S, TAGAMI Y, ONO S, et al. Embedding-based News Recommendation for Millions of Users[C] // The 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2017:1933-1942.
- [22] WU F Z, QIAO Y, CHEN J H, et al. Mind: A Large-scale Dataset for News Recommendation[C]// The 58th Annual Meeting

of the Association for Computational Linguistics, 2020: 3597-3606.

- [23] PENNINGTON J, SOCHER R, MANNING C D. Glove: Global Vectors for Word Representation[C]// The 2014 Conference on Empirical Methods in Natural Language Processing(EMNLP), 2014:1532-1543.
- [24] RENDLE S. Factorization Machines with LibFM [J]. ACM Transactions on Intelligent Systems and Technology, 2012, 3(3):1-22.
- [25] GUO H F, TANG R M, YE Y M, et al. DeepFM: a Factorization-machine based Neural Network for CTR Prediction[C] // The 26th International Joint Conference on Artificial Intelligence(IJCAI'17). AAAI, 2017:1725-1731.



**WEI Jinsheng**, born in 1995, Ph.D, lecturer. His main research interests include multimedia information processing, machine learning, intelligent recommendation and affective computing.



**LU Guanming**, born in 1965, Ph.D, professor, Ph.D supervisor. His main research interests include intelligent information processing, intelligent recommendation and machine learning.

(责任编辑:何杨)