



基于生成对抗网络的评分可信推荐模型

王 永, 王淞立, 邓江洲*

(重庆邮电大学 电子商务与现代物流重点实验室, 重庆 400065)

摘要 现有的基于深度学习的推荐模型主要致力于提升推荐系统的准确性。然而, 除了推荐准确性外, 模型的推荐可靠性也备受关注。该文提出一种基于生成对抗网络的评分可信推荐模型来评估预测结果的有效性, 以实现推荐准确性与可靠性间的权衡。该模型仅利用用户显式评分信息获取预测评分的可信度, 并根据设定的可靠性阈值筛选出具有高可信度的预测评分, 以保证推荐项目的可靠性。此外, 为了提高模型的预测效果并确保训练的公平性, 设计了正样本填充策略来缓解评分可靠性矩阵中的数据不均衡问题。在 3 个真实数据集上的实验结果表明, 该模型在 Recall 和 NDCG 指标上均优于所选的对比方法, 有效提高了推荐系统的性能。

关键词 生成对抗网络; 填充策略; 可靠性; 推荐系统

中图分类号 TP391

文献标志码 A

DOI 10.12178/1001-0548.2023116

Rating-Trustworthy Recommendation Model Based on Generative Adversarial Networks

WANG Yong, WANG Songli, and Deng Jiangzhou*

(Key Laboratory of E-Commerce and Modern Logistics, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

Abstract Existing deep learning-based recommendation models have mainly focused on improving the accuracy of recommendation systems. However, beyond recommendation accuracy, the reliability of the model's recommendations is also of great concern. Therefore, a rating-trustworthy recommendation model based on generative adversarial networks (GANs) is proposed to evaluate the effectiveness of prediction results and achieve a balance between recommendation accuracy and reliability. This model solely employs explicit user rating information to gauge the credibility of predicted ratings and screens out highly credible predicted ratings based on a predefined reliability threshold, thus ensuring the trustworthiness of recommended items. Furthermore, to enhance the prediction performance of the model and ensure fairness in training, a positive sample padding strategy is designed to mitigate the data imbalance problem in the rating reliability matrix. Experimental results on three real datasets show that the proposed model outperforms selected comparison methods in both Recall and NDCG metrics, effectively improving the performance of recommendation systems.

Key words generative adversarial networks; population; reliability; recommender systems

大数据时代下, 信息资源呈爆发式增长, 同质化信息充斥网络, 导致用户难以发现准确且可信的内容, 信息过载问题日益严重。推荐系统^[1] (Recommender System, RS) 作为一种高效的信息过滤技术, 通过深度分析用户历史偏好行为信息, 为用户提供个性化推荐服务, 提升用户体验。其中基于深度学习的推荐方法^[2] 展现出良好表现。生成对抗网络^[3] (Generative Adversarial Networks, GAN) 是新兴的深度学习技术, 其被证明在解决数据噪声

和数据稀疏性方面有不错的效果^[1], 为优化推荐性能提供新的解决方案。

大多数推荐模型仅使用原始评分信息进行预测, 但由于用户行为的主观性和复杂性, 其评分信息不完全可靠, 导致预测准确性降低^[4]。文献 [5] 通过噪声管理方案进行评分修正, 以提供相对准确的信息输入。但这类方法忽略了信息输出的可靠性对推荐质量的重要性。文献 [6-8] 也证实了可信推荐对提升用户满意度有积极影响, 因此除了准确性

收稿日期: 2023-04-19; 修回日期: 2023-08-24

基金项目: 国家自然科学基金 (62272077, 72301050); 重庆市自然科学基金面上基金 (cstc2021jcyj-msxmX0557); 重庆市教育委员会科学技术研究项目 (KJQN2023000605)

作者简介: 王永, 博士, 教授, 主要从事推荐系统、人工智能安全等方面的研究。

*通信作者 E-mail: dengjz@cqupt.edu.cn

外, 推荐可靠性也应被考虑。如在酒店推荐中, 综合评分高但评分人数较少的酒店可能不如评分略低但评分人数更多的酒店更受用户信赖。因此, 推荐系统需要考虑评分可信度来提高推荐质量和用户体验。

基于此, 本文提出了一种基于生成对抗网络的评分可信推荐模型来提供可靠性预测。首先, 考虑到隐式反馈无法准确描述用户偏好, 利用条件生成对抗网络模型对用户显式评分信息进行训练, 以捕捉不同用户的偏好程度。其次, 借助基于偏好一致性的自然噪声检测来识别原始矩阵中的噪声评分, 并通过设计的生成机制产生评分可靠性矩阵; 再利用生成对抗网络框架预测出评分的可靠性概率。最后, 根据设定的可靠性阈值, 过滤掉不可信的预测评分, 以保证推送给目标用户的项目具有高可信度, 从而实现可信推荐。

1 相关工作

1.1 生成对抗网络

GAN 能够不依赖于任何先验假设, 学习到高维复杂的数据分布。GAN 最开始在图像生成领域成功应用, 后续在推荐系统领域也获得显著效果。文献 [9] 将 GAN 的对抗训练方式成功应用到 CF 中去提升推荐准确性。但由于生成器 G 会生成离散项目索引, 使得项目标签冲突。为此, 文献 [10] 利用 vector-wise (矢量对抗) 训练生成对抗网络 (GAN-Based Collaborative Filtering Framework, CFGAN) 去解决传统基于 GAN 模型中所存在的一系列问题。文献 [11] 通过增加记忆模块训练特定用户条件向量, 从而生成出更加符合用户特征的数据。文献 [12] 提出了一个融合神经协同过滤 (Neural Network-Based Collaborative Filtering, NCF) 和 GAN 的隐式推荐算法来解决用户和项目间交互信息不充足的问题。然而, 以上方法都未考虑到原始评分信息中所存在的自然噪声的问题, 这可能会对模型训练结果的准确性和可靠性带来一定影响, 使得预测效果不佳。此外, 现有基于 GAN 的方法还未充分考虑使用显式评分信息作为模型输入的问题。

1.2 可信推荐系统

为缓解噪声问题导致的推荐偏差, 研究者们提出了各种解决方案来处理自然噪声问题, 以保证模型输入的准确性。文献 [13] 介绍了一种基于评分偏好的自然噪声管理方法来识别和修正有误的用户评分。文献 [14] 借助模糊理论方法, 设计了一个模糊噪声识别框架。文献 [15] 通过将用户和物品

分类到更加细粒度的类别, 以对自然噪声进行识别。文献 [16] 针对不同推荐场景设计了相应模型来预处理自然噪声。文献 [17] 提出了一种基于概率加权的阈值方法来修正自然噪声。以上方法都是从信息源的角度去预处理用户评分信息。除了对输入信息进行噪声识别外, 对模型的输出结果进行有效性评估也相当关键, 这将直接影响到模型推荐的准确性和可靠性。

推荐模型输出结果的可信程度会影响用户的购买意愿和决策选择。因此, 引入可靠性信息来优化预测结果的可信度, 不仅可以辅助用户建模, 还能提高预测评分的准确性和可靠性。文献 [7] 引入可靠性度量方法和文本信息, 构建了一个动态信任网络来提高预测评分的准确性和可靠性。文献 [18] 利用去噪自编码器从多源数据中获取多方位潜在特征, 并使用强化学习整合特征来度量评分可靠性。文献 [8] 提出了一种基于用户项目双视角的可信推荐方法来评估预测结果的可靠性, 并修正不可信结果。文献 [19] 介绍了一种基于伯努利分布的矩阵分解推荐模型来获取预测结果的可靠性概率, 并筛选高可靠性结果输出。随后, 为了更好地捕捉和融合用户和项目的高低阶特征, 文献 [20] 利用神经协同过滤方法对用户评分矩阵进行分类处理, 以获取具有可靠性概率的离散预测评分。然而, 以上这些方法都是利用原始评分信息或其他辅助信息进行可信推荐, 未充分考虑到用户评分信息中所普遍存在的噪声问题, 这会降低模型训练的准确性以及预测结果的可信度。因此, 将信息源的可靠性考虑到推荐模型之中, 对进一步提高模型的可信推荐有积极作用。

2 本文模型

本文所提模型框架如图 1 所示, 主要由 Rating-GAN 预测评分模块、可靠性矩阵生成模块和 Reliability-GAN 评分可靠性模块构成。其中, Rating-GAN 模块使用用户显式评分信息进行偏好预测; 可靠性矩阵生成模块用于通过基于偏好一致性的自然噪声识别策略对用户-项目交互矩阵中的用户评分进行噪声检测, 并将识别出的噪声评分区域设置为 0, 否则为 1, 由此可得只含 0-1 二值的可靠性矩阵 Q ; Reliability-GAN 模块用于训练可靠性矩阵 Q , 以得到评分预测区域的可靠性概率, 即预测评分的可信度。针对两 GAN 模块的联合输出结果 $\langle \text{Rating}, \text{Reliability} \rangle$ 对, 在推荐时先按预测评

分值降序处理, 再根据设定的可靠性阈值来过滤掉一定比例的不可信推荐结果, 即对高分但其可靠性低于阈值的待推荐项目进行剔除。最后, 为了确保剔除前后推荐列表长度不变, 本文选择具有高预测评分与可靠性的未交互项目进行对应填充, 以确保高质量可信推荐。

2.1 基于用户显式评分的 Rating-GAN

考虑到隐式反馈信息无法准确描述用户的评分偏好, 因而本文的 Rating-GAN 将利用条件生成对抗网络模型对用户显式评分信息进行训练, 以此来捕捉用户对项目的偏好程度, 在一定程度上能确保推荐的项目为目标用户所喜爱。Rating-GAN 采用 vector-wise 的训练方式, 其生成器 G 和判别器 D 均由多层感知机 MLP 构成, 具体实现过程如下。

对于特定用户 u , 将随机噪声 z_u 和条件向量 c_u

进行拼接, 并输入到生成器 G 中; 再通过 GAN 拟合, 以得到虚假评分向量。条件向量有助于 GAN 去拟合数据真实分布并加速收敛。这里, 随机抽取了部分用户的真实购买向量作为用户条件向量 c_u , 其公式如下:

$$\begin{aligned} \mathbf{x}^1 &= \text{ReLU}(\mathbf{W}_1^T \{z_u, c_u\} + b_1) \\ \mathbf{x}^2 &= \text{ReLU}(\mathbf{W}_2^T \mathbf{x}^1 + b_2) \\ &\vdots \\ \mathbf{x}^n &= \text{ReLU}(\mathbf{W}_{n-1}^T \mathbf{x}^{n-1} + b_n), n = 2, 3, \dots, N-1 \\ \hat{r}_u &= \text{Sigmoid}(\mathbf{W}_{\text{out}}^T \mathbf{x}^{N-1} + b_N) \end{aligned} \quad (1)$$

式中, $\{\cdot\}$ 表示内部两向量的串联; \mathbf{W}_n 和 b_n 分别表示第 n 个隐藏层的权重矩阵和偏置矢量; \mathbf{x}^n 和 \hat{r}_u 分别表示第 n 个隐藏层的输出和输出层的输出。

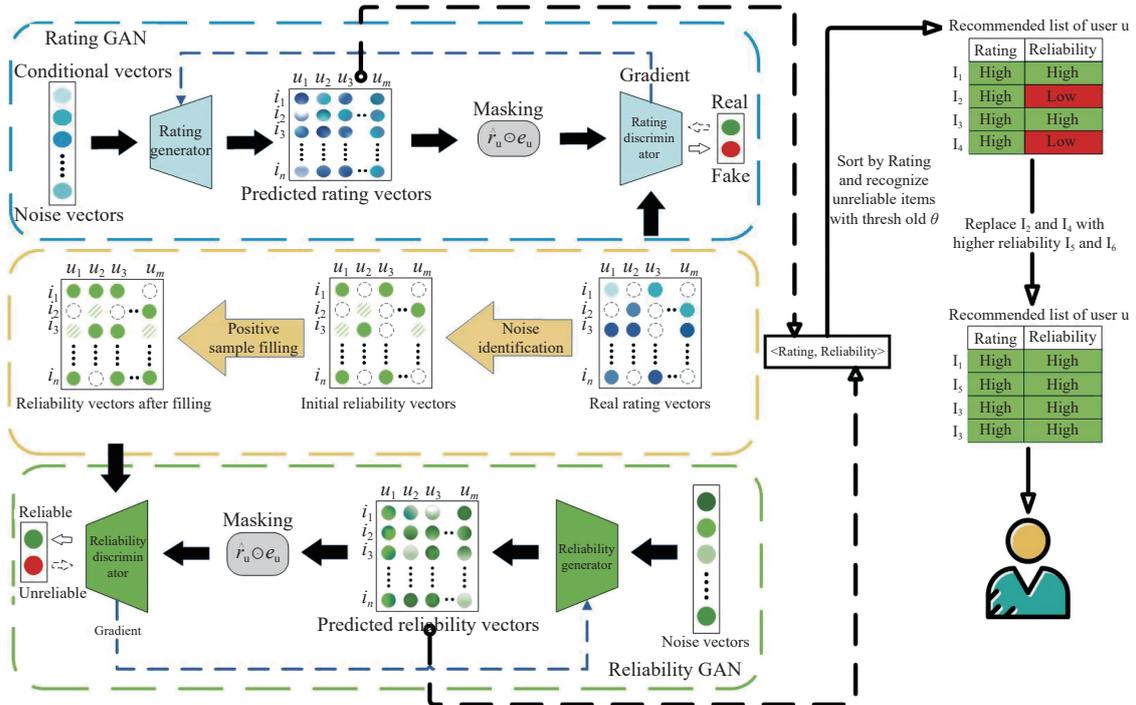


图1 DGAN+R 框架图

此外, 为了让生成器 G 能更有效地学习真实的评分分布, 将利用 Masking 层来模拟真实的数据稀疏度, 并使 G 的输出包含更多的用户偏好信息。Masking 层通过 $\hat{r}_u \odot e_u$ 实现, 其中 e_u 与 \hat{r}_u 的维度相同, 用于表示一系列项目是否已被用户 u 评过, 若其中某项目 i 被用户 u 评分过, 则 e_{ui} 为具体评分值 r_{ui} , 否则为 0; \odot 为元素积。因此, 在将 \hat{r}_u 输入到判别器 D 前, 可对未评分向量进行过滤处理并补充用户已有评分的偏好信息; 同时生成器 G 也只接受已评分项目的损失梯度, 从而能保证

G 和 D 更有效地学习。

D 的损失函数如下:

$$J^D = - \sum_{u \in U} \left(\log D(\mathbf{r}_u | c_u) + \log(1 - D(\hat{r}_u \odot e_u | c_u)) \right) \quad (2)$$

相似地, G 的损失函数如下:

$$J^G = \sum_{u \in U} \log(1 - D((\hat{r}_u \odot e_u) | c_u)) \quad (3)$$

然而, 由于评分向量通常是高维且稀疏, 使得只依赖于 D 传回损失梯度的基于 vector-wise 的对

抗训练方式无法充分指导 G，这容易导致 G 生成一个看似合理但未考虑用户相对偏好的无效解。为此，本文通过借助 CFGAN 中的负采样方法来重构正则化和部分掩蔽，以生成可靠解。

对于正则化重构，通过选取 N^t 个负样本来重构 G 的损失函数，让 G 接收到用户在未评分项目上的损失梯度，从而直接指导 G 学习。由此，将重构损失函数作为正则化项引入到 G 的损失函数中，可得：

$$J^G = \sum_{u \in U} \left(\log(1 - D(\hat{r}_u \odot e_u) | c_u) + \alpha \sum_{j \in N^t} (r_{uj} - \hat{r}_{uj})^2 \right) \quad (4)$$

式中， α 是正则项的超参数。通过最小化 J^G ，可使 G 既考虑用户的相对偏好，又防止平凡解的生成。

对于部分掩蔽，通过在 Masking 层中不剔除所选取 N^m 个负样本的输出，即 D 中的输入既有用户已评分的项目，又有未交互的项目。因此，在重构正则化基础上，G 又能接受到 D 中的负样本梯度，则 D 和 G 的损失函数更新如下：

$$J^D = - \sum_{u \in U} \left(\log D(r_u | c_u) + \log(1 - D(\hat{r}_u \odot (e_u + k_u) | c_u)) \right) \quad (5)$$

$$J^G = \sum_{u \in U} \left(\log(1 - D(\hat{r}_u \odot (e_u + k_u) | c_u)) + \alpha \sum_{j \in N^t} (r_{uj} - \hat{r}_{uj})^2 \right) \quad (6)$$

式中， k_u 为负采样矩阵。本文将采用小批量随机梯度下降和反向传播来不断更新迭代模型参数，以实现 G 和 D 的损失函数最小化。训练完成后，将原始评分输入到 G 中，可预测出用户对未评分项目的评分偏好。

2.2 评分可靠性矩阵生成机制

2.2.1 考虑偏好一致性的评分可靠性矩阵

本文生成的评分可靠性矩阵无需借助其他辅助信息，只依靠用户评分信息即可。首先，利用文献 [13] 提出的基于偏好一致性的自然噪声识别方法来找出原始评分矩阵中的噪声数据。该方法通过统计评分偏好规则，将用户和项目划分到 4 种偏好类别中。包括弱偏好类、无偏好类、强偏好类以及不确定类。以用户分类为例，根据文献 [13] 设置的两阈值 L 和 H 来区分用户评分偏好程度，其中 L 和 H 分别表示弱偏好类与无偏好类间的阈值边界和无偏好类与强偏好类间的阈值边界。如在评分区间为 1~5 时， L 和 H 分别被设置为 3 和 4。其判别标准如下。

1) 若用户 u 在项目 i 上的评分 r_{ui} 小于阈值 L ，则

该评分属于弱偏评分集 W_u ；

2) 若用户 u 在项目 i 上的评分 r_{ui} 大于等于阈值 L 且小于阈值 H ，则该评分属于无偏评分集 N_u ；

3) 若用户 u 在项目 i 上的评分 r_{ui} 大于等于阈值 H ，则该评分属于强偏评分集 S_u 。

由此，可通过以下分类标准将用户划分到 4 种不同偏好类别中。

$$\begin{cases} u \in \text{弱偏好类}, |W_u| \geq |N_u| + |S_u| \\ u \in \text{无偏好类}, |N_u| \geq |W_u| + |S_u| \\ u \in \text{强偏好类}, |S_u| \geq |W_u| + |N_u| \\ u \in \text{不确定类}, \text{其他} \end{cases} \quad (7)$$

式中， $|W_u|$ 、 $|N_u|$ 和 $|S_u|$ 分别表示用户 u 的弱偏评分集数量，无偏评分集数量和强偏评分集数量。相似地，可对项目进行偏好程度归类。

然后，使用以下识别条件来判断用户评分 r_{ui} 是否为自然噪声。若该评分是自然噪声（不可靠评分），则将其可靠性 R_{ij} 设置为 0；反之，该评分可靠性 R_{ij} 为 1。由此，可将评分矩阵转化为只含 0-1 二值的评分可靠性矩阵 Q 。

$$\begin{cases} r_{ui} \text{ 是自然噪声}, u, i \in \text{弱偏好类且 } r_{ui} > L \\ r_{ui} \text{ 是自然噪声}, u, i \in \text{无偏好类且 } (r_{ui} < L \text{ 或 } r_{ui} \geq H) \\ r_{ui} \text{ 是自然噪声}, u, i \in \text{强偏好类且 } r_{ui} < H \\ r_{ui} \text{ 是正常评分}, \text{其他} \end{cases} \quad (8)$$

2.2.2 考虑公平性的评分可靠性矩阵填充策略

推荐系统中，由于用户已评分数量通常远少于其未评分数量，使得评分矩阵极度稀疏，而且还存在一定自然噪声，这会进一步加大所生成的评分可靠性矩阵 Q 中“0”和“1”的比例失衡，其中“0”的数量远超过于“1”。为此，在利用生成对抗网络框架训练矩阵 Q 时，若不考虑模型公平性原则，会导致判别器 D 无法很好地学习到真实的评分可靠性数据分布，进而影响其损失梯度的传回，使得 D 无法有效指导生成器 G 预测出更合理的评分可信度。

为了解决上述训练公平性问题，提出一种正样本填充方法来尽可能均衡矩阵 Q 中的数据。以用户为例，假设用户 u 的已评项目集合为 I_u ，其中可靠性为 1 的项目集合为 R_u ，则该用户的可信度 C_u 为：

$$C_u = \frac{|R_u|}{|I_u|} \quad (9)$$

若用户 u 的可信度 C_u 大于所有用户可信度的均值 \bar{c} ，则认为该用户是可信用户；否则，该用户是不可信用户。相似地，可判别出项目是否为可信项目。

基于此, 设计正样本填充规则: 对于矩阵 Q 中非交互区域(该区域的评分可靠性默认是0), 若用户 u 和项目 i 都是可信的, 则认为它们间的评分 r_{ui} 也是可信的, 并将该区域的可靠性值修改为1; 否则, 可靠性值不变, 仍为0。这在一定程度上可增加矩阵 Q 中可靠性为1的比例, 模型训练公平性得以提升。

2.3 基于评分可靠性的 Reliability-GAN

Reliability-GAN 是基于 2.1 节的 Rating-GAN 框架, 但不同之处在于: 首先, 该模型将不使用用户条件向量 c_u 。这是因为在生成对抗训练中, 若将含有 0-1 二值的条件向量作为输入, 会使生成器 G 的输出分布很不均, 同时也容易造成模型快速过拟合, 从而导致训练结果不佳; 其次, 该模型 Masking 层中的 e_u 为只含 0-1 二值的多维向量。若值为 1, 则表示有交互行为; 而值为 0, 则表示无交互行为。

由此, 在 Reliability-GAN 中, 生成器 G 的输入将为随机噪声 z_u , 而所使用的负采样方法与 2.1 节相同, 其生成器 G 和判别器 D 的损失函数如下:

$$J^D = - \sum_{u \in U} \left(\log D(q_u) + \log(1 - D(\hat{q}_u \odot (e_u + k_u))) \right) \quad (10)$$

$$J^G = \sum_{u \in U} \left(\alpha \sum_{j \in N^t} (q_{uj} - \hat{q}_{uj})^2 \right) \quad (11)$$

式中, q_{uj} 为用户 u 在项目 j 上的真实评分可靠性值。通过不断更新迭代, 以实现 G 和 D 的损失函数最小化。训练完成后, 可得到预测评分的可信度。

3 实验

3.1 实验数据集

本文采用 3 个公开数据集 MovieLens 100k (ML-100K)、MovieLens 1M (ML-1M) 和 Yahoo Music (YM) 来进行实验设计与分析。数据集的详细描述信息见表 1 所示。同时, 为了全面覆盖数据集中的所有用户, 从每个用户中随机选择了 80% 的评分记录作为训练集, 剩余 20% 数据作为测试集, 用于模型有效性检验。

表 1 各数据集的详细描述信息

数据集	用户数/个	项目数/个	评分数量/个	稀疏度/%
ML-100K	943	1 682	100 000	6.3
ML-1M	6 040	3 952	1 000 209	4.2
YM	15 400	1 000	365 704	2.3

3.2 评价指标

本文使用 3 个常用的评价指标来衡量模型预测准确性以及 Top- N 推荐的性能效果, 分别是均方根误差 (RMSE)、召回率 (Recall) 和归一化折扣累计收益 (NDCG)。

RMSE: 用于衡量预测评分与真实评分差异程度, 即表示模型的预测准确性:

$$RMSE = \sqrt{\frac{1}{n} \frac{1}{m} \sum_{u=1}^n \sum_{i=1}^m (\hat{r}_{ui} - r_{ui})^2} \quad (12)$$

式中, n 和 m 分别表示用户数和项目数。

Recall: 用于衡量预测推荐结果与真实推荐结果间的项目匹配程度, 即表示模型的推荐准确性:

$$Recall@N = \frac{|I_p \cap I_a|}{|I_a|} \quad (13)$$

式中, N 为推荐列表长度; I_p 和 I_a 分别为预测推荐列表和实际推荐列表。

NDCG: 用于衡量预测推荐列表中推荐项目在真实推荐列表中排序的相关程度, 即为模型推荐质量:

$$NDCG@N = \frac{DCG@N}{IDCG@N} \quad (14)$$

式中, $DCG@N = \sum_{p=1}^N \frac{2^{rel_p} - 1}{\log_2^{(p+1)}}$ 为向目标用户推荐前 N 个项目的折损累计增益; rel_p 为已排序的推荐项目所处位置 p 的相关性, 若该项目在真实推荐列表中, 则 rel_p 为 1, 反之为 0; IDCG 是理想状态下的 DCG。考虑到数据集的划分方式以及稀疏程度等, 经过统计分析, 将固定每个用户的推荐列表长度 N 为 5。

3.3 对比方法

为验证所提模型的有效性和对比实验的可靠性, 本文从不同角度挑选了几个与所提模型高度关联的推荐模型来进行性能比较。各对比方法描述如下。

1) NCF^[2]: 是经典的神经协同过滤模型, 用于验证所提模型捕捉复杂信息的能力。

2) BeMF^[19]: 是基于伯努利的矩阵分解模型, 用于验证线性方式获取评分可靠性对推荐性能的影响。

3) CFGAN^[10]: 是基于 GAN 的推荐模型, 用于验证引入评分可信辅助信息的性能表现。

4) PRGAN^[11]: 是基于改进 GAN 的推荐模型, 用于验证条件向量对模型推荐效果的影响。

5) I_CDAE^[21]: 是基于去噪自编码器的推荐模型, 用于验证模型的推荐性能。

6) C_NCF^[20]: 是基于 NCF 的分类推荐模型, 用于验证非线性方式获取评分可靠性对推荐性能的影响。

3.4 参数设置

本文模型的参数设置如下: 对于 Rating-GAN, 生成器和判别器迭代次数设置为 4 和 2, 正则化系数设置为 0.1, 负样本总数设置为 140, 初始学习率设置为 0.000 1; 对于 Reliability-GAN, 生成器和判别器迭代次数设置为 3 和 1, 正则化系数设置为 0.05, 负样本总数设置为 80, 初始学习率设置为 0.000 15。其他对比方法的参数设置满足网格搜索的最优原则。

3.5 结果分析与讨论

3.5.1 正样本填充策略效果检验

首先, 利用 2.2.1 节提及的自然噪声检测方法去识别 ML-100K、ML-1M 和 YM 数据集的自然噪声, 检测出的噪声评分数量分别为 8 464、88 342 和 7 680, 且在对应数据集上的噪声比为 8.46%、8.83% 和 2.10%。其次, 根据 2.2.1 节, 将评分矩阵转化为评分可靠性矩阵。图 2a 显示了在不同数据集上当前评分可靠性矩阵中所含 0 和 1 数值的占比情况。由图 2a 可知, 评分可靠性矩阵中 0 和 1 的占比极度失衡, 这不利于模型训练的公平性, 会导致模型出现严重的训练偏差。

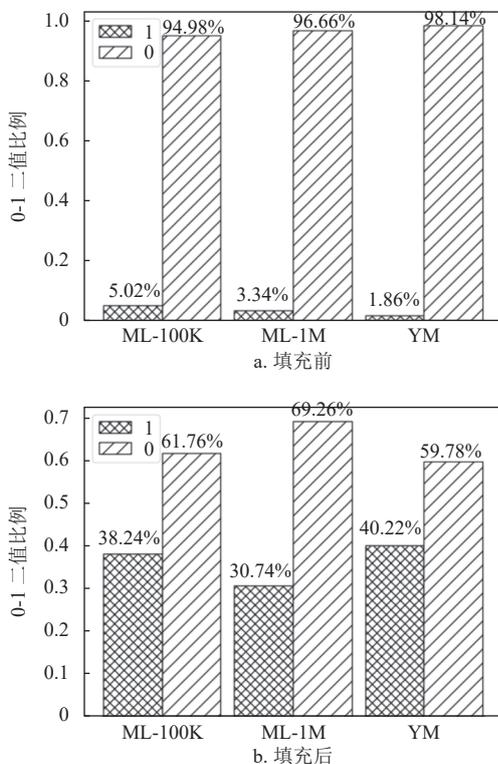


图 2 填充前后各评分可靠性矩阵中 0 和 1 占比情况

为此, 利用 2.2.2 节提出的正样本填充策略来

对当前评分可靠性矩阵进行填充。有效填充后, 0 和 1 的占比情况如图 2b 所示。从图 2b 可看出, 经过填充后可靠性矩阵中 0 和 1 的占比变得相对均匀, 但不同数据集的填充效果存在不一现象, 且在较为稀疏的数据集上表现更好。这是因为所提填充策略会受到数据稀疏度、用户偏好程度以及评分分布等因素影响。如相对于 ML, YM 的稀疏程度更高, 用户与项目交互的信息会更少, 使得所提策略在信息量较少情况下, 可根据模糊规则, 为可靠性矩阵填充更多的 1。

以 ML-100K 为例, 图 3 显示了填充前后本文模型推荐性能变化情况。由图 3 可知, 相较于填充前, 填充后模型在 Recall 和 NDCG 指标上性能分别提升了 4.2% 和 3.1%。其提升效果在其余数据集上也相似。这说明本文填充策略能有效提高模型的推荐效果。

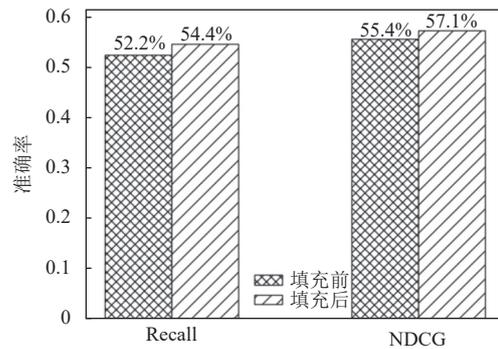


图 3 填充前后 DGAN+R 推荐性能对比

3.5.2 可靠性阈值选择

可靠性阈值 θ 的选择将直接关系到模型的推荐效果, 其目的是用于筛选出评分可信度大于可靠性阈值的预测评分, 以保证系统最终推送给目标用户的项目具有高可靠性, 从而实现可信推荐。在本实验中, 由于所提模型会过滤掉部分不可信预测评分, 可能会导致部分用户的推荐列表的长度变短, 因而设计了一种填充方法来实现剔除前后推荐长度的一致性。其方法如下: 对于目标用户 u , 从评分可靠性矩阵中找到该用户所有未交互项目中预测可信度大于系统设定阈值的项目集合 P_u 并进行内部降序, 再根据推荐列表中不可信项目被剔除的位置 p 和数量 k , 将 P_u 中前 k 个项目进行对应填充。若所填充的项目 i 的预测评分大于用户 u 的所有预测评分的均值, 则认为该项目被推荐命中, 反之亦然。

考虑到不同数据集的预测评分可靠性数值分布存在一定差异, 因此可靠性阈值设定也不同。这里, 假定使用拒绝率 q 来表示系统需过滤掉的不可

信预测项目的比例, 通过观察所提模型在不同拒绝率下的性能变化, 如图4所示, 以确定最佳的可靠性阈值。

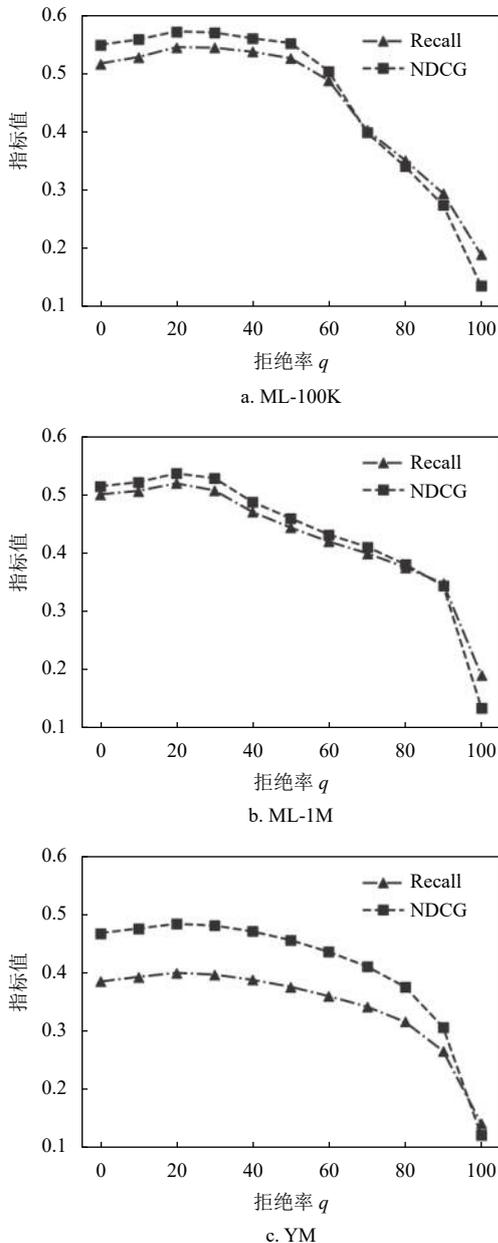


图4 不同拒绝率 q 下本文模型的性能变化情况

从图4可以看出, 当 $q=0$ 时, 所提模型将退化为 Rating-GAN; 随着 q 值的逐渐增加, 模型在3个数据集上的指标结果都有所提升, 且在 $q=0.2$ 时, 实现最佳性能, 其中 Recall 和 NDCG 指标的提升比例最高可达到 5.4% 和 4.3%; 而当 $q > 0.3$ 时, 模型指标结果开始逐渐变差, 在 $q=1$ 时表现最差。其原因在于: 首先, 评分可靠性作为辅助信息可在一定程度上消除原始评分矩阵中噪声评分对推荐结果的影响, 同时所设计的填充方法也能在

合理填充比例范围之内发挥较好的积极作用; 其次, 当拒绝率超出合理范畴, 会使 Reliability-GAN 模块的功能逐步失效, 从而退化为 Rating-GAN 的性能结果, 甚至因剔除比例过大, 导致填充方法无法找到更多满足要求的可信项目, 削弱其推荐效果。由此可知, 在推荐过程中, 合理的剔除与填充一定比例的项目, 能有效地提高模型的推荐质量。本实验中, 以剔除 20% 的预测评分来确定可靠性阈值, 能保证推荐系统中可信预测评分比例在合理范围之内实现本文模型的指标最佳。各数据集上的可靠性阈值选择如表2所示。

表2 不同数据集上的可靠性阈值选择

数据集	ML-100K	ML-1M	YM
可靠性阈值	0.435	0.475	0.395

3.5.3 对比结果分析

表3显示了在不同数据集上各对比方法的评估指标结果。从表3可知, 本文模型 DGAN+R 在3个数据集上的综合性能均优于其他对比方法, 其退化后的 Rating-GAN 也能有不错的表现。由于本文模型考虑了评分可信度, 因而相比于 PRGAN, 本文在 RMSE、Recall@5 和 NDCG@5 指标上的性能至少提升了 18%、3% 和 6%。对于提供评分可靠性的 BeMF 和 C_NCF, BeMF 的推荐性能在较为稀疏的数据集上远不如 C_NCF, 又因其预测离散评分导致预测准确性也较差; 而 C_NCF 在所有对比方法中表现较为优秀, 特别是在 ML-100K 上, 推荐准确性仅次于所提模型。I_CDAE 在预测准确性上表现突出, 尤其在 ML-1M 数据集上优于所有方法。综合3个数据集的结果可知, 相较于次优模型 I_CDAE, 本文方法在 RMSE 指标上预测准确性平均提升 11.5%, 在 Recall 和 NDCG 指标上推荐准确性平均提升 3.8% 和 5.6%。

表3 对比方法的评估指标结果

对比方法	ML-100K	ML-1M	YM
NCF	(1.419,0.513,0.532)	(1.535,0.472,0.496)	(1.823,0.375,0.446)
C_NCF	(1.172,0.526,0.541)	(1.048,0.492,0.500)	(1.898,0.369,0.426)
BeMF	(1.811,0.503,0.544)	(1.837,0.489,0.525)	(1.973,0.304,0.362)
I_CDAE	(0.984,0.523,0.541)	(0.987,0.496,0.521)	(1.686,0.389,0.446)
CFGAN	(1.382,0.494,0.527)	(1.348,0.475,0.501)	(1.837,0.361,0.442)
PRGAN	(1.331,0.527,0.537)	(1.724,0.455,0.469)	(1.783,0.380,0.446)
Rating-GAN	(1.161,0.516,0.548)	(1.197,0.500,0.514)	(1.722,0.385,0.468)
DGAN+R	(0.967,0.544,0.571)	(1.019,0.519,0.536)	(1.458,0.400,0.484)

因此, 本文在 GAN 框架下, 通过识别评分矩

阵中的自然噪声, 构建评分可靠性矩阵来预测评分的可信度, 可过滤掉不可靠预测, 有效提升推荐质量和用户满意度。

4 结束语

本文提出了一种基于生成对抗网络的评分可信推荐模型, 旨在为预测评分提供相应评分可靠性概率, 并根据可靠性阈值过滤掉具有低可靠性概率的预测评分, 以保证最终推荐给目标用户的项目既具有高预测评分值, 又具有高可信度。与基于 CF 的模型相似, 该模型无需利用除用户评分以外的其他辅助信息就能得到预测评分的可靠性值, 具有良好的通用性和可扩展性。同时, 考虑到模型训练的公平性, 本文设计了正样本填充策略来缓解评分可靠性矩阵中的数据不均衡问题, 从而提高模型输出结果的准确性。实验表明: 本文模型在 3 个公开数据集上的性能表现明显优于其他对比方法, 有效提高了推荐系统的质量。

参考文献

- [1] GAO M, ZHANG J, YU J, et al. Recommender systems based on generative adversarial networks: A problem-driven perspective[J]. *Information Sciences*, 2021, 546: 1166-1185.
- [2] HE X, LIAO L, ZHANG H, et al. Neural collaborative filtering[C]//Proceedings of the 26th International Conference on World Wide Web. Perth: ACM, 2017: 173-182.
- [3] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial networks[J]. *Communications of the ACM*, 2020, 63(11): 139-144.
- [4] DENG J, RAN X, WANG Y, et al. Probabilistic matrix factorization recommendation approach for integrating multiple information sources[J]. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2023, 53(10): 6220-6231.
- [5] JOORABLOO N, JALILI M, REN Y. Improved recommender systems by denoising ratings in highly sparse datasets through individual rating confidence[J]. *Information Sciences*, 2022, 601: 242-254.
- [6] BOBADILLA J, GUTIÉRREZ A, ORTEGA F, et al. Reliability quality measures for recommender systems[J]. *Information Sciences*, 2018, 442: 145-157.
- [7] MORADI P, AHMADIAN S. A reliability-based recommendation method to improve trust-aware recommender systems[J]. *Expert Systems with Applications*, 2015, 42(21): 7386-7398.
- [8] AHMADIAN S, AFSHARCHI M, MEGHDADI M. A novel approach based on multi-view reliability measures to alleviate data sparsity in recommender systems[J]. *Multimedia Tools and Applications*, 2019, 78: 17763-17798.
- [9] WANG J, YU L, ZHANG W, et al. IRGAN: A minimax game for unifying generative and discriminative information retrieval models[C]//Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval. Tokyo: ACM, 2017: 515-524.
- [10] CHAE D K, KANG J S, KIM S W, et al. Cfgan: A generic collaborative filtering framework based on generative adversarial networks[C]//Proceedings of the 27th ACM International Conference on Information and Knowledge Management. New York: ACM, 2018: 137-146.
- [11] WEN J, ZHU X R, WANG C D, et al. A framework for personalized recommendation with conditional generative adversarial networks[J]. *Knowledge and Information Systems*, 2022, 64(10): 2637-2660.
- [12] SUN J, LIU B, REN H, et al. NCGAN: A neural adversarial collaborative filtering for recommender system[J]. *Journal of Intelligent and Fuzzy Systems*, 2022, 42(4): 2915-2923.
- [13] BAG S, KUMAR S, AWASTHI A, et al. A noise correction-based approach to support a recommender system in a highly sparse rating environment[J]. *Decision Support Systems*, 2019, 118: 46-57.
- [14] WANG P, WANG Y, ZHANG L Y, et al. An effective and efficient fuzzy approach for managing natural noise in recommender systems[J]. *Information Sciences*, 2021, 570: 623-637.
- [15] LI W, LI X, DENG J, et al. Sentiment based multi-index integrated scoring method to improve the accuracy of recommender system[J]. *Expert Systems with Applications*, 2021, 179: 115105.
- [16] YERA R, CASTRO J, MARTINEZ L. Natural noise management in recommender systems using fuzzy tools[J]. *Computational Intelligence for Semantic Knowledge Management: New Perspectives for Designing and Organizing Information Systems*, 2020: 1-24.
- [17] LUO C, WANG Y, LI B, et al. An efficient approach to manage natural noises in recommender systems[J]. *Algorithms*, 2023, 16(5): 228.
- [18] AHMADIAN M, AHMADIAN S, AHMADI M. RDERL: Reliable deep ensemble reinforcement learning-based recommender system[J]. *Knowledge-Based Systems*, 2023, 263: 110289.
- [19] ORTEGA F, LARA-CABRERA R, GONZÁLEZ-PRIETO Á, et al. Providing reliability in recommender systems through Bernoulli matrix factorization[J]. *Information Sciences*, 2021, 553: 110-128.
- [20] BOBADILLA J, GUTIÉRREZ A, ALONSO S, et al. Neural collaborative filtering classification model to obtain prediction reliabilities[J]. *International Journal of Interactive Multimedia and Artificial Intelligence*, 2022, 7(4): 18-26.
- [21] LIU D, WANG Y, LUO C, et al. An improved autoencoder for recommendation to alleviate the vanishing gradient problem[J]. *Knowledge-Based Systems*, 2023: 110254.