

融合门控注意力机制与双线性特征交互的推荐模型

何昌隆 文斌

(成都信息工程大学通信工程学院 四川 成都 610225)

摘要 为了在影视、书籍等单一类型推荐中准确地表达用户的真实偏好,充分地捕获到推荐数据中的有效特征,研究并提出一种融合门控注意力机制与双线性特征交互的推荐模型。使用融入门控机制的注意力单元来对用户的局部显性偏好建模,使用双线性特征交互层来对用户的长期泛性偏好进行挖掘,以提升深度推荐模型的学习能力。在 Amazon(Books)和 MovieLens-1M 两个公开数据集中进行实验,实验结果表明所提模型相比于其他推荐模型,推荐效果有一定程度的提升。

关键词 推荐系统 深度学习 注意力机制 双线性函数 多层感知机

中图分类号 TP391

文献标志码 A

DOI:10.3969/j.issn.1000-386x.2024.04.043

A RECOMMENDATION MODEL COMBINING GATING ATTENTION MECHANISM AND BILINEAR FEATURE INTERACTION

He Changlong Wen Bin

(College of Communication Engineering, Chengdu University of Information Technology, Chengdu 610225, Sichuan, China)

Abstract In order to accurately express the users' true preferences in a single type of recommendation such as movies and books, and fully capture the effective features in the recommendation data, a recommendation model that integrates the gating attention mechanism and bilinear feature interaction is proposed. This model used the attention unit integrated into the gated mechanism to model the user's local explicit preferences and used bilinear feature interaction layer to mine the long-term general preferences of users to improve the learning ability of the deep recommendation model. Experiments were conducted on two public data sets, Amazon (Books) and MovieLens-1M. The experimental results show that the proposed model has a certain degree of improvement in recommendation effect compared with other recommendation models.

Keywords Recommendation system Deep learning Attention mechanism Bilinear function Multilayer perceptron

0 引言

推荐系统能够在纷繁复杂的数据中挖掘到数据之间的联系,将用户的真正需求信息进行推送。现如今已广泛应用在电子商务和社交媒体等领域^[1-2]。传统的推荐方法主要从用户项目属性(基于内容)^[3]、用户的历史交互(协同过滤)^[4]两个方向进行建模并推荐。但在信息爆炸和信息类别繁多的今天,传统的推荐算法难以兼顾用户与项目的所有特征并进行深层次学习,推荐效果始终有限。近年来,深度学习快速发展并在计算机视觉、自然语言处理等领域得到了广泛的应用。于是将深度学习引入到推荐领域来以得到更好的

推荐效果,成为当前推荐领域研究的重点。同时得益于推荐系统能够直接提高企业的收益能力,在实际需求下,许多优秀的推荐模型得以涌现。

Guo 等^[5]将多层感知机与因子分解机结合,分别对高阶特征与低阶特征进行学习并提供推荐,让推荐模型具有泛性和记忆的特点;Qu 等^[6]认为在推荐系统中的特征之间多是一种乘积的关系,因此在模型中提出 Product Layer 来更加充分地处理特征的交叉组合;Zhou 等^[7]认为不同于以往模型对嵌入向量一视同仁的处理,人的兴趣是十分多样的,并将注意力机制引入推荐模型以此挖掘用户的显性偏好并应用于电商领域;Song 等^[8]使用 Multi-head Self-Attention 机制自动构建高阶交叉特征,来进一步提高推荐效果。尽管这

些推荐模型在推荐场景中取得了不错的成果,但往往不能同时兼顾用户历史交互信息中的局部显性偏好和长期泛性偏好的捕捉。在类如影视推荐和图书推荐等单一类型推荐中难以全面而充分地挖掘出用户真实兴趣。因此本文提出一种融合门控注意力机制与双线性特征交互的推荐模型(GABIN),对用户特征中的局部显性偏好与长期泛性偏好进行建模,来更加全面而充分地挖掘用户兴趣,以得到更好的推荐质量,并在 Amazon(Books) 和 MovieLens-1M 两个公开数据集中与其他推荐模型进行对比实验,验证其推荐效果。

1 相关工作

随着计算机硬件水平的提升以及对于深度学习研究的深入,越来越多的深度学习模型被引入到了推荐系统当中,在上文已做了许多相关介绍。其中文献[7]对于以往的深度学习在推荐算法中应用的研究总结出了一个基准的模型,名为 Embedding & MLP 模型,推荐领域中大多数热门的推荐模型是由此模型改进而来的。

如图 1 所示,Embedding & MLP 模型主要由嵌入层(Embedding Layer)和全连接层(MLP)构成。在推荐系统中可以将用户与项目的多种信息作为输入,由模型学习并产生推荐结果,这些信息既包括用户的特征信息,比如职业、年龄等;或是项目的特征信息,比如项目类别、项目的生产时间等;还包括用户和项目的历史行为信息,比如用户的评分、购买和点击等交互行为。然而推荐算法所使用的这些多源异构信息通常具有高维和稀疏的特点,直接编码输入神经网络需要大量的参数进行学习,得到的推荐效果也不理想。嵌入层(Embedding Layer)的作用就是将这些长度不一的高维稀疏信息投射成长度相同低维稠密的嵌入向量^[9]。再把通过嵌入层处理后的嵌入向量作为输入,在全连接层构建多层感知机(MLP)进行参数学习。最后将多层感知机的输出通过激活函数计算,以产生最终的推荐结果。

2 模型设计

提出的 GABIN 模型结构如图 2 所示。为了在影视书籍等单一类型推荐领域中更加全面而充分地挖掘用户的偏好,提高推荐系统的推荐效果,GABIN 模型在上文提到的 Embedding & MLP 基准模型的基础上,在嵌入层之后加入了一层特征捕获层,其将带有门控机制的注意力单元来对历史交互信息中的局部显性特征进行捕获,再选用一种双线性特征交互层来对用户的长期泛性特征进行捕获,然后将处理后的特征互相拼接输入全连接层,进而对用户个性化推荐。

©element-product

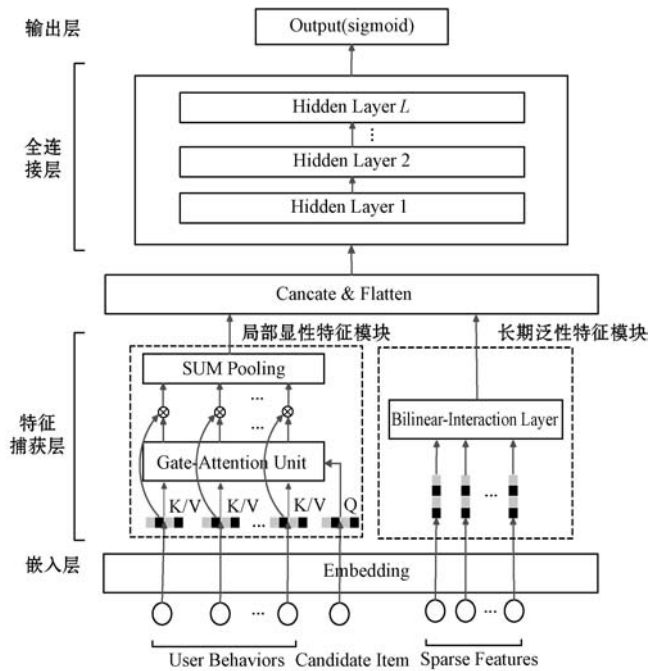


图 2 GABIN 模型结构

整个模型由嵌入层、特征捕获层、全连接层、输出层构成,其中特征捕获层对于用户交互特征的处理则分为两个部分,即图 2 中的局部短期偏好模块与一般长期偏好模块。

2.1 嵌入层

推荐模型的作用是在系统项目列表中找到用户真实需要的部分并推荐。因此需要将用户与项目的相关特征进行有效的表示。首先将所有数据进行预处理得到用户、项目的特征信息以及用户项目的历史交互信息。接着分别通过 one-hot 与 multi-hot 编码对得到的用户项目的相关信息与历史交互信息进行特征向量的表征。由于此类编码方式产生的特征空间维度过大,模型需要在嵌入层通过一个全连接网络将输入的高维稀疏特征转化为低维稠密特征,进而有效地降低特征

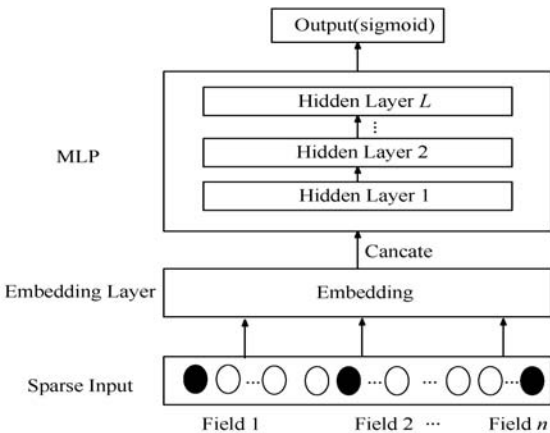


图 1 Embedding & MLP 基准模型结构

空间的维度,为后面的特征捕获层的输入做好准备。

2.2 局部显性偏好模块

模型特征捕获层的局部显性偏好模块在嵌入层后对用户的历史行为序列特征进行处理。在通常推荐算法的处理中,对于用户所有的历史行为特征都是平等对待的,使用这种方式得到的用户交互特征是用用户一种全面、泛性而长期的特征。而模型中的局部显性偏好模块希望捕获到用户历史交互特征中局部、显性而短期的特征。对此本文设计一种门控注意力单元来处理用户的历史行为特征,其使用融合门控机制的注意力单元来为每个用户行为的嵌入向量分配一个权值系数 W_j ,其表示目标推荐项目下,用户对此历史交互项目的偏好程度大小,以此挖掘出用户的局部显性偏好。其中本文设计的门控注意力单元构成如图 3 所示。

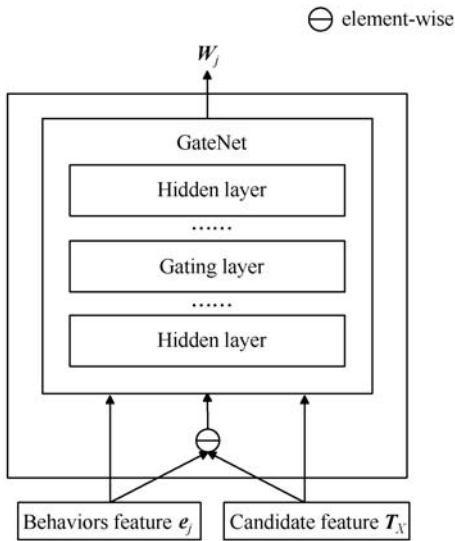


图 3 门控注意力单元结构

在门控注意力单元中,用户的历史行为编码和目标推荐项目编码过嵌入层后,模块得到了该用户总计 n 个行为的嵌入向量特征列表 $\{e_1, e_2, \dots, e_n\}$ 与目标项目特征向量 T_x 。接着见式(1),将单个行为嵌入向量与目标特征向量与两者相减的新特征向量拼接,得到 h_0 作为门控网络的输入并在门控网络中进行训练,最后经过线性层计算就得到该单个交互特征的最终的偏好权值 W_j 。

$$a_0 = \text{Concate}(e_j, T_x, (e_j - T_x)) \quad (1)$$

式中: e_j 代表单个历史交互特征向量; T_x 代表目标项目特征向量; a_0 是门控网络的输出。

单元中所使用的门控网络的整个网络结构是在多层感知机的基础上融合门控机制^[10]改进的,即在多层感知机中的多个隐藏层之间加入门控层 (Hidden Gating layer),进而在上一层隐藏层结果中挑选出更加重要的交互特征传递到下一层网络,以此更加准确地挖

掘用户的短期显性特征。门控网络中隐藏层计算式为:

$$a_{i+1} = \text{PreLU}(w_{i+1}a_i + b_{i+1}) \quad (2)$$

式中: w_{i+1} 和 b_{i+1} 是此隐藏层的参数; a_i 是上一个隐藏层的输入; a_{i+1} 是这一层的输出。

门控网络中门控层的计算式为:

$$g_i = a_i \odot \text{PreLU}(w_g a_i) \quad (3)$$

式中: w_g 是该门控层参数; a_i 是上一个隐藏层的输入; g_i 是此门控层的输出。

在经过门控注意力单元得到了单个行为嵌入向量的偏好权值 W_j 后,再将其与该行为嵌入向量相乘,就得到了所需的此历史行为嵌入向量的局部显性特征。当数据充足时,模型可以规定输入的用户交互行为个数,直接将所有经过门控注意力单元的行为嵌入向量拼接,再作为整个局部显性模块的输出。但考虑到推荐系统中用户项目的交互特征矩阵是十分稀疏的,用户可能有不同的行为数量,导致拼接后的不同用户的模块输出长度不一,而模型的全连接层只能处理固定长度输入向量。因此在最后模块使用 Sum 池化层,将所有的经过门控注意力单元加权后的单个历史交互特征进行累加操作,进而得到整个局部显性偏好模块的最终输出 q_{out} ,其计算式为:

$$q_{out} = \sum_{j=1}^n w_j e_j \quad (4)$$

2.3 长期泛性偏好模块

该模块的主要功能是为了挖掘出用户的长期泛性偏好,这种偏好不是表现在某些局部特征里,而是隐藏在用户的性别、年龄、历史交互记录等所有相关信息中。在以往的推荐算法中用户的长期一般偏好仅仅将嵌入层得到的用户相关信息的特征向量组直接拼接表示,而这种方式显然不够精确。因为直接将各个特征相拼接的操作表达的是特征之间的一种“和”的关系,而在用户兴趣的特征表达中更多的是一种“且”的关系^[6]。因此长期泛性偏好模块使用一种特征乘积交互的方式来更加准确地挖掘用户历史行为特征中的长期泛性偏好。

其中点乘 (\cdot) 广泛应用在 FM^[11]、FFM^[12] 等推荐模型,哈达玛积 (\odot) 常应用在 AFM^[13]、NFM^[14] 等推荐模型。这两种交互层建模方式不能很好地表征特征之间的组合,本文使用一种双线性特征交互层^[15] (Bilinear-Interaction Layer) 来进一步学习用户属性、交互信息等所有的嵌入特征向量 $\{e_1, e_2, \dots, e_m\}$ 中的长期泛性偏好。

如图 4 所示,整个双线性特征交互层,在处理特征之间的乘积操作时将特征 e_i 和特征 e_j 之间引入了一个 $k \times k$ 的权值参数矩阵 W 来学习特征之间的交互组合。先将特征向量 e_i 与矩阵 W 做点乘,再将得到的结果与特征向量 e_j 做哈达玛乘,最终得到交叉特征向量 p_{ij} ,其具体计算式为:

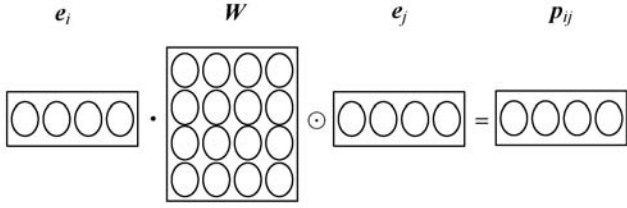


图 4 双线性乘积

$$p_{ij} = e_i W \odot e_j \quad (5)$$

式中: $W \in \mathbf{R}^{k \times k}$, e_i 和 e_j 分别是第 i 场和第 j 场的用户特征向量, $e_i, e_j \in \mathbf{R}^k$ 。设共有 n 条特征,避免重复计算这里规定 $i < j$,因此整个模块会有 $L = m(m-1)/2$ 次计算,最后得到总的用户二阶交互特征 $\{p_1, p_2, \dots, p_L\}$,在长期泛性模型的最后,将所有的经双线性特征交互层得到的用户二阶交互特征向量进行拼接操作,得到最终模块的输出 p_{out} ,其计算式为:

$$p_{out} = \text{Concat}(p_1, p_2, \dots, p_L) \quad (6)$$

2.4 全连接层与输出层

全连接层主要使用多层感知机来学习特征之间的高阶交互信息。如式(7)所示,全连接层将在特征捕获层得到的局部显性偏好模块输出 q_{out} 与长期泛性偏好模块输出 p_{out} 进行拼接操作,生成最终的模型整体特征嵌入向量 h_0 并作为多层感知机的输入,进行参数训练。

$$h_0 = \text{Concat}(q_{out}, p_{out}) \quad (7)$$

多层感知机中的隐藏层计算公式如下:

$$h_{l+1} = \text{ReLU}(w_{l+1}h_l + b_{l+1}) \quad (8)$$

式中: w_{l+1} 和 b_{l+1} 是此隐藏层的参数; h_l 是上一个隐藏层的输入; h_{l+1} 是这一层的输出。

接着整个模型在输出层得到最后的预测结果,如式(9)所示,输出层将全连接层中的多层感知机输出结果 h_L 作为输入,经过线性层计算,最后在激活函数的作用下输出最终的预测结果 \hat{y} 。

$$\hat{y} = \text{sigmoid}(wh_L + b) \quad (9)$$

式中: $\hat{y} \in (0, 1)$ 作为整个模型的预测输出; w 和 b 是线性层的参数; h_L 是全连接层的输出。

3 实验

为了验证提出的融合门控注意力机制与双线性特

征交互的推荐模型(GABIN)的推荐效果,采用 PyTorch 1.4.0 作为深度学习框架,在 Ubuntu 16.04.6、GTX 1080 Ti 显存 11 GB、Intel Xeon E5-2658 v3 @ 2.20 GHz、32 GB 内存、PyThon 3.6 的实验环境下进行对比实验。

3.1 实验数据集和评价指标

由于 GABIN 模型主要针对单一类型领域推荐,本实验使用了 MovieLens-1M 和 Amazon (Books) 两个推荐领域常使用的针对电影与图书这种单一领域的公开数据集作为验证算法效果的实验数据。

Amazon 数据集是依据亚马逊网站真实业务数据而收集的。本文使用其 Books 子集作为实验数据集,其中包括了 603 668 位用户对 367 982 个物品的 8 898 041 条评论信息。每个用户至少对数据集中的五个项目进行打分。模型的任务是根据用户物品相关特征和用户前 n 个正面评价过的物品来预测第 $n+1$ 个物品是否应该被推荐。

MovieLens1M 数据集是由明尼苏达大学的 GroupLens 研究项目收集整理。该数据被广泛应用于推荐系统的实验中,其中包含了 6 040 个用户对 3 952 部电影的 1×10^6 条评分记录,评分范围是 0~5,每个用户至少对数据集中的二十个项目进行打分。模型的任务是根据用户电影相关属性特征和用户前 n 个正面评价的电影来预测第 $n+1$ 部电影是否应该被推荐(此数据集实验中设置 $n=10$)。

实验过程中采用 F 值、AUC 和 RelImpr 三种指标作为对比不同算法推荐质量的评价依据。

F 值指标是兼顾精确率 (Precision) 和召回率 (Recall) 的分类评价指标,常用于评估推荐系统的综合推荐性能^[16],范围为 0~1,值越大越好,F 值 (F-Score) 的计算式为:

$$F_{\text{Score}} = \frac{2 \times P_{\text{recision}} \times R_{\text{ecall}}}{P_{\text{recision}} + R_{\text{ecall}}} \quad (10)$$

AUC 指标在分类问题中被广泛使用,表示正例排在负例前面的概率,指标对分类阈值不敏感,范围为 0~1,值越大越好,以往的一些工作验证了推荐领域中 AUC 指标的必要性^[17]。RelImpr 用来衡量对比模型相对于基准模型的推荐效果提升的程度,RelImpr 的计算式为:

$$R_{\text{elImpr}} = \left(\frac{A_{UC_{\text{Measured_Model}}}}{A_{UC_{\text{Embedding\&MLP}}}} - 0.5 \right) \times 100\% \quad (11)$$

3.2 对比模型

为了对本文提出的 GABIN 模型的推荐效果进行验证,实验除了上文所提到的 Embedding & MLP 基准

模型,还选取了四个以往研究中具有代表性的推荐算法模型进行对比,对比模型如下:

(1) Wide & Deep^[18]: Wide & Deep 模型对基准模型的全连接层进行了改进,其全连接层由多层感知机和基础线性模型并行组成,共同对用户特征进行学习。

(2) DeepFM^[5]: Deep & FM 模型是在 Wide & Deep 模型基础上的一种改进,DeepFM 模型将 Wide & Deep 模型中基础线性模型(Wide)换成因子分解机(FM)来对低阶的线性特征进行学习。

(3) PNN^[6]: Product Network 模型在基准模型的基础上,在 Embedding 层之后加入一层 Product Layer 来对每一个 field 特征的交互信息进行捕获。

(4) DIN^[7]: Deep Interest Network 模型在基准模型的基础上引入了注意力机制,通过 Embedding 层之后的局部激活单元来挖掘用户历史交互特征中的兴趣偏好。

3.3 实验结果与分析

首先将上述章节所介绍的 Embedding & MLP 基准模型和在此基础上改进的融合门控注意力机制与双线性特征交互的推荐模型(GABIN)以及 Wide & Deep、DeepFM、PNN 和 DIN 共六种推荐模型在 Amazon 数据集中的 Books 子集与 MovieLens1M 数据集上进行实验,实验数据集随机分为训练集和测试集用作模型的训练和评估,其分别占整个数据集的 80% 和 20%。评价指标使用的是上文提到的 F 值指标、AUC 指标和 RelImpr 指标,用来评估对比实验模型的推荐质量,其中由于推荐系统巨大的用户基数,相关指标 1% 的提升也能给企业带来巨大的收益。针对不同的公开数据集,实验在大量测试中选取了相关的最佳参数值如表 1 所示。

表 1 不同数据集选用的实验参数

参数	Amazon(Books)	MovieLens1M
嵌入层向量维度	18	16
DNN 层数	3	3
DNN 隐藏层节点数	200,80	32,16
Dropout	0.1	0.3
优化器	Adam	Adam
学习率	0.001	0.001
训练批次	10	10

实验中各模型得到的评价指标 F 值(F-Score)的对比结果如表 2 所示。

表 2 不同模型的 F-Score 指标实验结果

模型	Amazon(Books)	MovieLens1M
Embedding & MLP	0.706 6	0.689 2
Wide & Deep	0.709 8	0.692 8
DeepFM	0.718 3	0.701 6
PNN	0.715 0	0.701 0
DIN	0.731 6	0.717 3
GABIN	0.742 4	0.725 7

实验中各模型得到的评价指标 AUC 的对比结果如表 3 所示。

表 3 不同模型的 AUC 指标实验结果

模型	Amazon(Books)	MovieLens1M
Embedding & MLP	0.752 9	0.604 3
Wide & Deep	0.754 7	0.604 8
DeepFM	0.763 1	0.607 4
BHPNN	0.760 3	0.606 5
DIN	0.775 6	0.611 2
GABIN	0.785 9	0.614 7

实验中各模型相对于 Embedding&MLP 基准模型的相对提升指标 RelImpr 的对比结果如图 5 所示。

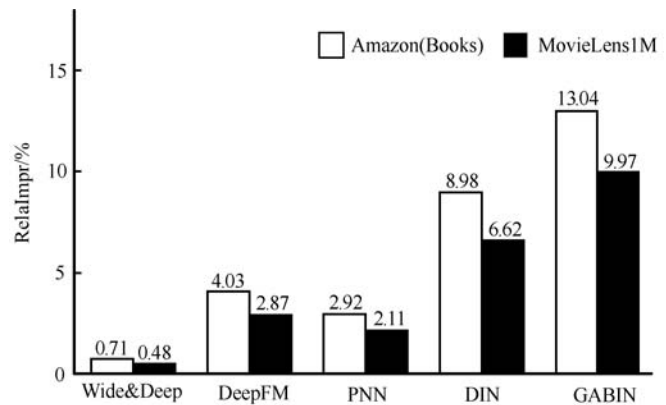


图 5 不同模型的 RelImpr 指标实验结果

表 2、表 3 和图 5 展示了对比实验中的各模型在 Amazon(Books) 和 MovieLens1M 两种单一数据集上的推荐性能,在同样参数的训练环境下,本文提出的 GABIN 模型相比于基准模型 Embedding & MLP 和针对全连接层进行改进的 Wide&Deep、DeepFM 模型,以及针对嵌入层改进的 PNN、DIN 模型,在不同数据集集中的推荐效果都有明显的提升。

在 Amazon(Books)数据集上,本文提出的 GABIN 模型对比 Embedding & MLP、Wide & Deep、DeepFM、PNN、DIN 五种经典的推荐模型,在 F 值指标上分别提高了 0.035 8、0.032 6、0.024 1、0.027 4 和 0.010 8,在 AUC 指标上分别提高了 0.033、0.031 2、0.022 8、

0.025 6 和 0.010 3,在相对提升指标 RelImp_r 上分别提高了 13.04 百分点、12.33 百分点、9.01 百分点、10.12 百分点和 4.06 百分点。在 MovieLens1M 数据集上,本文提出的 GABIN 对比 Embedding & MLP、Wide & Deep、DeepFM、PNN 和 DIN 五种经典的推荐模型,在 F 值指标上分别提高了 0.036 5、0.032 9、0.024 1、0.024 7 和 0.008 4,在 AUC 指标上分别提高了 0.010 4、0.009 9、0.007 3、0.008 2 百分点和 0.003 5,在相对提升指标 RelImp_r 上分别提高了 9.97 百分点、9.49 百分点、7.1 百分点、7.86 百分点和 3.35 百分点。这表明了本文提出的融合门控注意力机制与双线性特征交互的推荐模型(GABIN)是有效的,使用门控注意力单元与双线性交互层能够更加全面而充分地捕获用户历史交互特征中的局部显性偏好与全局泛性偏好,并在单一类型的推荐场景中相对于其他经典的推荐算法,进一步提高推荐系统的推荐效果。

4 结 语

为了提高推荐模型在影视、书籍等单一类型推荐场景中的推荐效果,提出一种融合门控注意力机制与双线性特征交互的推荐模型(GABIN)。该模型设计一种门控注意力单元来对用户行为中的局部显性特征进行捕获,同时选用双线性层来对用户行为中的全局泛性特征进行捕获,接着将所有特征进行拼接作为多层感知机的输入进行训练学习,最后激活函数计算下产生推荐结果,这在一定程度上全面而充分地挖掘出了用户的兴趣偏好,提高了推荐模型的推荐效果。通过在 Amazon(Books)和 MovieLens1M 不同公开数据集之间的实验表明,本文提出的 GABIN 推荐模型在推荐性能上相比于 Embedding & MLP、Wide & Deep、DeepFM、PNN 和 DIN 这几种经典的推荐算法有明显的提升,验证了该模型在推荐系统中实践的可行性,具有实际意义。今后的工作会将社交关系、项目信息、时间间隔等更多的系统异构信息引入到模型中,来提升推荐系统的推荐质量。

参 考 文 献

[1] Shuai Z, Yao L, Sun A, et al. Deep learning based recommender system: A survey and new perspectives[J]. ACM Computing Surveys, 2019, 52(1): 1-38.

[2] 黄立威,江碧涛,吕守业,等.基于深度学习的推荐系统研究综述[J].计算机学报,2018,41(7):1619-1647.

[3] 王光,张杰民,董帅含,等.基于内容的加权粒度序列推荐算法[J].计算机工程与科学,2018,40(3):564-570.

[4] 李昆仑,万品哲,张德智.基于改进用户相似性度量和评分预测的协同过滤推荐算法[J].小型微型计算机系统,2018,39(3):567-571.

[5] Guo H F, Tang R M, Ye Y M, et al. DeepFM: A factorization-machine based neural network for ctr prediction[C]//26th International Joint Conference on Artificial Intelligence, 2017:1725-1731.

[6] Qu Y R, Cai H, Zhang W N, et al. Product-based neural networks for user response prediction[C]//2016 IEEE 16th International Conference on Data Mining, 2016:1149-1154.

[7] Zhou G R, Zhu X Q, Song C R, et al. Deep interest network for click-through rate prediction[C]//4th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2018:1059-1068.

[8] Song W P, Shi C, Xiao Z P, et al. AutoInt: Automatic feature interaction learning via self-attentive neural networks[C]//28th ACM International Conference on Information and Knowledge Management, 2019:1161-1170.

[9] 练建勋.基于多样化内容数据的个性化推荐系统[D].合肥:中国科学技术大学,2018.

[10] Huang T W, She Q Y, Wang Z Q, et al. GateNet: Gating-enhanced deep network for click-through rate prediction[EB]. arXiv:2007.03519, 2020.

[11] 燕彩蓉,周灵杰,张青龙,等.因子分解机模型的宽度和深度扩展研究[J].软件学报,2019,30(3):822-844.

[12] Juan Y C, Zhuang Y, Chin W S, et al. Field-aware factorization machines for CTR prediction[C]//10th ACM Conference on Recommender Systems, 2016:43-50.

[13] Xiao J, Ye H, He X N, et al. Attentional factorization machines: Learning the weight of feature interactions via attention networks[C]//26th International Joint Conference on Artificial Intelligence, 2017:3119-3125.

[14] He X N, Chua T S. Neural factorization machines for sparse predictive analytics[C]//40th International ACM SIGIR Conference on Research and Development in Information Retrieval, 2017:355-364.

[15] Huang T W, Zhang Z Q, Zhang J L, et al. FiBiNET: Combining feature importance and bilinear feature interaction for click-through rate prediction[C]//13th ACM Conference on Recommender Systems, 2019:169-177.

[16] 朱郁筱,吕琳媛.推荐系统评价指标综述[J].电子科技大学学报,2012,41(2):163-175.

[17] Graepel T, Candela J Q, Borchert T, et al. Web-scale Bayesian click-through rate prediction for sponsored search advertising in Microsoft's Bing search engine[C]//27th International Conference on International Conference on Machine Learning, 2010:13-20.

[18] Cheng H Z, Koc L, Harmsen J, et al. Wide & deep learning for recommender systems[C]//1st Workshop on Deep Learning for Recommender Systems, 2016:7-10.