

面向移动应用评分推荐的多任务图嵌入深度预测模型

李海明, 朱智衡, 刘磊, 过辰楷

引用本文

李海明, 朱智衡, 刘磊, 过辰楷. 面向移动应用评分推荐的多任务图嵌入深度预测模型[J]. 计算机科学, 2023, 50(9): 160-167.

LI Haiming, ZHU Zhiheng, LIU Lei, GUO Chenkai. [Multi-task Graph-embedding Deep Prediction Model for Mobile App Rating Recommendation](#) [J]. Computer Science, 2023, 50(9): 160-167.

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)

[融合语义和句法图神经网络的实体关系联合抽取](#)

Fusion of Semantic and Syntactic Graph Convolutional Networks for Joint Entity and Relation Extraction

计算机科学, 2023, 50(9): 295-302. <https://doi.org/10.11896/jsjcx.220700041>

[基于并行卷积网络信息融合的层级多标签文本分类算法](#)

Hierarchical Multi-label Text Classification Algorithm Based on Parallel Convolutional Network Information Fusion

计算机科学, 2023, 50(9): 278-286. <https://doi.org/10.11896/jsjcx.221200133>

[基于深度学习的红外视频显著性目标检测](#)

Deep Learning Based Salient Object Detection in Infrared Video

计算机科学, 2023, 50(9): 227-234. <https://doi.org/10.11896/jsjcx.220700204>

[基于LpTransformer网络的手语动画拼接模型](#)

Sign Language Animation Splicing Model Based on LpTransformer Network

计算机科学, 2023, 50(9): 184-191. <https://doi.org/10.11896/jsjcx.221100043>

[基于双视角纠偏的推荐模型](#)

Rectifying Dual Bias for Recommendation

计算机科学, 2023, 50(9): 152-159. <https://doi.org/10.11896/jsjcx.220900035>

面向移动应用评分推荐的多任务图嵌入深度预测模型

李海明¹ 朱智衡¹ 刘磊² 过辰楷³

1 上海电力大学计算机科学与技术学院 上海 201306

2 南开大学人工智能学院 天津 300350

3 南开大学网络空间安全学院 天津 300350

(shdianli2022@163.com)

摘要 随着智能终端设备以及移动应用程序的普及,用户对应用质量的要求和用户体验需要愈发凸显。移动应用的评分推荐作为一项有效的事前评估手段,逐渐得到市场关注。传统的应用评分推荐工作主要围绕解决数据稀疏和模型深度问题,未能对应用推荐本身的图结构、多任务形态进行准确表征。针对该问题,提出了一种面向移动应用评分推荐的图嵌入多任务模型 AppGRec,利用归纳型二部图的嵌入结构对特征中的用户交互关系进行挖掘,并使用 shared-bottom 多任务模型捕获应用评分中的多任务特点,同时兼顾了数据稀疏和模型深度的影响。在 Google Play 上收集了 16031 个有效的移动应用及其特征数据作为验证数据集,实验结果表明,AppGRec 在 MAE 和 RMSE 上相比 state-of-the-art 模型分别提升了 10.4% 和 10.9%。此外,对 AppGRec 超参和核心模块的影响做了具体分析,多角度验证其有效性。

关键词: 移动应用;推荐系统;图嵌入;深度学习;神经网络;评分预测

中图法分类号 TP391

Multi-task Graph-embedding Deep Prediction Model for Mobile App Rating Recommendation

LI Haiming¹, ZHU Zhiheng¹, LIU Lei² and GUO Chenkai³

1 College of Computer Science and Technology, Shanghai University of Electric Power, Shanghai 201306, China

2 College of Artificial Intelligence, Nankai University, Tianjin 300350, China

3 College of Cyber Science, Nankai University, Tianjin 300350, China

Abstract With the prevalence of smart terminal devices and mobile application (app for short), the requirements for application quality and user experience gradually increase. As an effective pre-assessment method, mobile app rating recommendation has gained increasing attention from app markets. The traditional app rating and recommendation works mainly focus on challenges such as data sparsity and model depth. Nevertheless, they fail to accurately capture the graph relationship within the apps and users. Furthermore, the multi-task characteristic of app recommendation is neglected. Aiming at these shortcomings, this paper proposes a graph embedding multi-task model AppGRec for mobile app rating and recommendation. AppGRec uses the embedding structure of inductive bipartite graph to mine the user interaction features. It uses the shared-bottom based model to capture the multi-task feature in app rating, while considering the effects of data sparsity and model depth. 16031 valid mobile apps and their feature data on Google Play are collected as dataset for method evaluation. Experimental results show that AppGRec achieves 10.4% and 10.9% improvement in terms of MAE and RMSE respectively comparing with the state-of-the-art models. In addition, this paper also makes quantitative analysis of the impact of hyperparameters and some core modules in AppGRec, and verifies the effectiveness from multiple perspectives.

Keywords Mobile app, Recommendation system, Graph embedding, Deep learning, Neural network, Rating prediction

到稿日期:2022-07-04 返修日期:2022-10-12

基金项目:国家自然科学基金(62002177);天津市科技计划项目(20YDTPJC01810);南开大学 2022 年实验课程改革项目(22NKSYSX05);教育部—华为产学研合作协同育人项目(202002142035)

This work was supported by the National Natural Science Foundation of China(62002177), Science and Technology Planning Project of Tianjin City(20YDTPJC01810), 2022 Experimental Course Reform Project of Nankai University(22NKSYSX05) and Ministry of Education Industry University Cooperation Collaborative Education Project(202002142035).

通信作者:过辰楷(guochenkai@nankai.edu.cn)

1 引言

随着智能终端设备以及移动应用软件的普及,移动应用市场的软件下载量与日俱增。以 Android 官方应用市场 Google Play 为例,其每周新增的应用数量超过 1 万款,总应用数量达到 260 万款¹⁾。据统计,2021 年国内应用市场的应用规模也达到惊人的 252 万款²⁾。如此庞大的数量规模下,所包含的应用软件质量参差不齐。虽然,诸如 Google Play、华为应用市场这样的权威分发市场对其上传的移动应用执行了基本的质量安全性检测,但更多的用户体验、功能易用性等方面的问题却难以排查。更不用说大量的非官方应用分发渠道掺杂其中,这些渠道并不会对上传的应用进行有效检验,使得移动应用质量问题层出不穷,各种误导用户下载、触犯个人隐私、非法获取权限的行为时有发生,侵害用户利益。

在此背景下,用户在寻找、下载和使用应用软件之前,对该应用的功能、质量安全和用户体验均会有所期待。为了避免使用不满足需求的应用,用户往往会关注应用市场中的应用评分和下载数量等特征,来对应用质量进行预先评判,从而决定是否下载和使用该应用。因此,对移动应用的评分预测,作为一项有效的事前评估手段,逐渐得到市场关注。该任务不仅能帮助应用市场初步过滤质量不合格的应用,也提示应用开发者如何有效提升应用质量和用户认可度。

移动应用推荐预测任务的实用价值使得关于该任务的研究成为一项热点。在此方向的研究工作主要包括基于传统的推荐系统改进模型^[1-3]以及基于深度学习的特征提取模型^[4-6]。然而这些工作未能深入提取移动应用和用户之间的交互关系,也无法对其多任务形态进行准确表征,在预测效果上仍存在较大提升空间。近年来,图结构模型在社交网络模拟和推荐系统任务上的发展^[7-9],证明了其具备强大的关系提取能力。一些研究者将图结构模型引入移动应用推荐预测上来,成功推理了应用和用户的隐式交互关系,取得了性能结果的提升^[10-11]。然而,这些工作在面对冷启动和数据稀疏问题时,仍会遭遇预测性能的瓶颈。

针对上述问题,本文提出了一种多任务图嵌入深度预测模型 AppGRec。AppGRec 利用二部图嵌入模型深入挖掘用户对移动应用的交互关系。该图嵌入模型采用归纳型设计,从而可以解决新数据的冷启动问题。同时,AppGRec 保留了因子分解机(Factorization Machine, FM)和深度神经网络(Deep Neural Network, DNN)的模块构成,以此解决数据稀疏和特征提取问题。再者,移动应用推荐预测场景中存在显著的多任务特点,应用的推荐倾向除了总体评分之外,还包含下载量、评语、点赞等多种交互特征,这些特征之间既有区别又相互关联,适用于多任务体系。AppGRec 利用这一特点整合多任务模型,进一步提升预测准确率。为了验证 AppGRec 的有效性,本文收集了 Google Play 市场的 16031 个有效的移动应用形成实验数据集。实验结果表明 AppGRec 在 MAE

和 RMSE 上相比 state-of-the-art 模型分别提升了 10.4% 和 10.9%。此外,本文还对 AppGRec 中的超参和核心模块的影响做了具体分析,进一步验证图嵌入和多任务模块的有效性。

总体而言,本文工作的主要贡献体现在以下 3 个方面:

1) 针对移动应用评分预测中的冷启动和数据稀疏问题,提出了一种基于归纳型图嵌入的深度学习模型,深入挖掘用户和应用的交互关系。

2) 利用移动应用评分预测中的多任务特点,构建和融合多任务学习模型,进一步提升模型预测的准确性。

3) 在收集的真实数据集上验证了本文模型的有效性。通过对比分析表明本文模型的典型模块对准确性提升具有正向作用。

2 相关工作

由于移动应用功能体验等用户需求的日益增长,近年来关于移动应用的评分推荐工作渐渐兴起。其中一些工作旨在改进传统推荐模型或文本分析方法来提升推荐的性能。Yu 等^[12]将协同过滤和主题建模技术 LDA 相结合形成应用预测模型 AppLibRec,可兼容基于文本描述和基于代码库的两种不同预测分析。类似的 LDA 改进模型还包括 Lin 等^[13]基于评论主题的移动应用推荐。在解决数据稀疏和冷启动方面,Yin 等^[14]提出了一种面向移动应用的稀疏加法生成模型(Mobi-SAGE)。该模型通过对用户兴趣和类别感知的用户隐私偏好进行综合评估来推荐移动应用。Cao 等^[15]提出了一种新颖的版本敏感的移动应用推荐框架。该框架通过发现用户与应用不同版本之间的交互来挖掘其潜在特征,利用共同正则化技术提高其推荐精度,从而缓解由版本划分引起的数据稀疏问题,很好地解决了应用新版本的冷启动问题。类似的工作还包括 Yao 等^[16]提出的一种细粒度模型版本感知矩阵分解模型(VAMF)。该模型一方面利用与评级相关的评论文本来丰富传统的矩阵分解模型;另一方面,捕获了同一应用多个版本之间的时间相关性和相似应用之间的聚合相关性。这两方面的改进被合并到矩阵分解框架中以提升推荐精度。其他的传统模型改进工作包括 Liu 等^[17]提出的层次结构化用户选择模型 SUCM 和 Tu 等^[18]提出的兴趣感知矩阵协同分解(IMCF)模型。总体而言,这些传统模型的改进工作虽然考虑到了用户偏好关系、用户评论等要素,但由于缺少深入特征挖掘方法作为支撑,推荐效果的提升程度受限。

近年来,一些研究者从模型深化层面对移动应用评分推荐和数据稀疏问题进行了有益的尝试。Bae 等^[10]的 AppTrends 是较早的基于图模型的移动应用推荐工作。AppTrends 利用构建用户历史使用图来提取移动应用与用户使用之间的关系,以此缓解应用推荐冷启动和数据稀疏问题。Xu 等^[19]提出了一种用户和移动应用嵌入表征的上下文联合学习方式,并利用神经网络进行用户偏好预测。此外,他们还引入了多维度采样方法来解决数据不平衡问题。Ouyang

¹⁾ Number of available applications in the Google Play Store: <https://www.statista.com/statistics/266210/number-of-available-applications-in-the-google-play-store/>

²⁾ 2021 年国内 App 总量: <https://finance.sina.com.cn/tech/2022-02-02/doc-ikyarmz8610748.shtml>

等^[11]提出了深度多图嵌入模型 DMGE,在跨域图的基础上构建一种多图神经网络,将图神经网络扩展到跨域应用的嵌入表征,并设计了域共享嵌入层来生成应用的共享嵌入。相关的基于图神经网络移动应用推荐工作还包括 Liang 等^[20]提出的混合图神经网络模型、Xie 等^[21]提出的多视图图结构模型、Zhang 等^[22]提出的知识图模型等。此外,我们前期的工作 KDFM^[23]提出了基于知识图的注意力深度因子分解机,一定程度提升了对应用评分的预测性能。然而,现有的应用推荐深度模型较侧重于图模型的构建和多视图的考量,尚缺少有效的设计以充分发挥图关系抽象、多维度特征提取和多任务学习预测这三者的增益作用。本文提出的 AppGRec

利用合理的模型构建巧妙地解决了这个问题,有效提升了应用评分预测的准确性。

3 AppGRec

结合移动应用预测推荐中的多任务、强交互等特征,本文提出了一种基于多任务深度图嵌入的评分预测模型 AppGRec,该模型主要分为归纳型图嵌入、深度因子分解机和多任务学习 3 个模块,分别如图 1 中橙色框、蓝色框和绿色框所示。图嵌入模块包含了图构建、采样和卷积聚合计算 3 个步骤;深度因子分解机包含了深度模块和因子分解两部分;多任务模块则是由两个子任务的独立 MLP 塔组成。

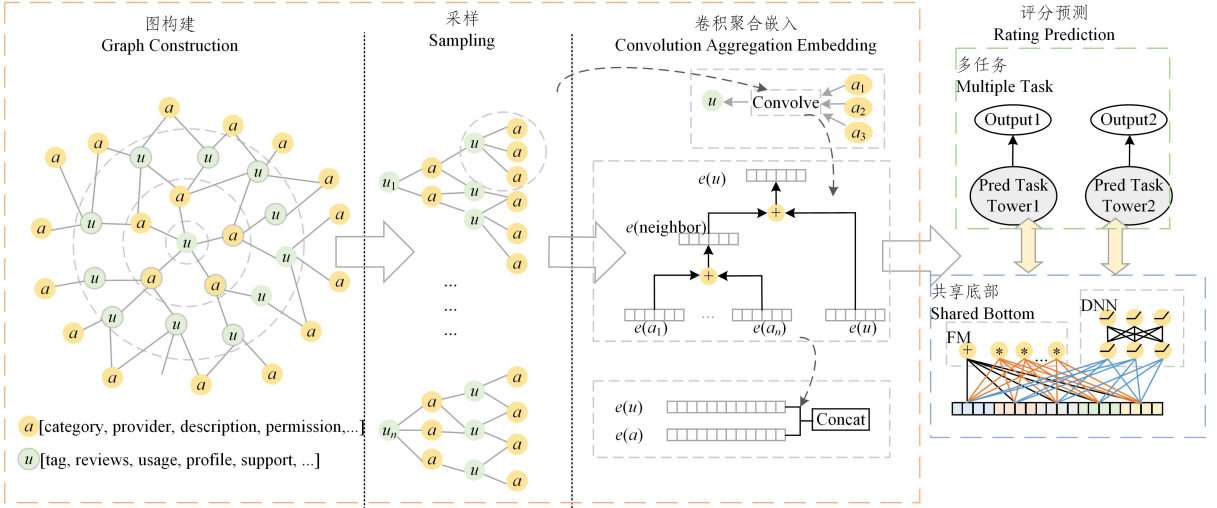


图 1 AppGRec 总体流程(电子版为彩图)

Fig. 1 Workflow of AppGRec

3.1 归纳型图嵌入

3.1.1 图构建

本文将用户集合 U 和移动应用集合 A 之间的使用关系抽象为二部图模型,称为使用图 G 。

$$G=(U, A, Q) \quad (1)$$

其中, Q 表示用户和移动应用之间使用关系集合,该关系反映了评分或下载量等评价特征。 $q_{i,j} \in Q$ 是连接用户 u_i 和应用 a_j 的一个连续值。 U 和 A 统称为使用图 G 的实体集合。图中的每个实体在训练开始时将会被初始化为一个多维向量,而各实体的向量维度保持相同,便于后续计算。该多维向量可以由每个实体的多个属性经过初始词嵌入(如 Word2Vec^[24])得到。

3.1.2 基于邻居交互的图采样

图节点的邻接节点数量没有限制,导致训练效率和内存空间难以控制,且过多的邻接节点的推理路径会引发过拟合,因此需要在图嵌入训练之前进行邻接节点的采样。AppGRec 采用一种基于重要性的随机游走方式进行节点采样。对于目标节点 t ,首先获取其所有邻居节点集合,而后从目标节点 t 开始进行固定步数的随机游走,对于节点 t 而言,其游走的概率 $p_{t,t'}$ 由节点之间的交互性 $q_{t,t'}$ 值决定,即:

$$p_{t,t'} = \frac{q_{t,t'}}{\sum_{t' \in N(t)} q_{t,t'}} \quad (2)$$

在移动应用评分推荐场景下, $q_{t,t'}$ 为综合评价长度,表明

用户和应用之间的交互热度,其计算如式(3)所示:

$$q_{t,t'} = \sum_{r \in R_{t,t'}} |r| \quad (3)$$

其中, $|r|$ 表示评论 r 的长度, $R_{t,t'}$ 为用户 t 对应用 t' 的评论集合。注意评论的交互是双向的,即 $q_{t,t'}$ 既支持用户 t 的采样又支持应用 t' 的采样。游走结束之后统计各邻居节点被访问的次数,取次数最高的 n 个邻居节点。而当满足的邻居节点未达到 n 个时, AppGRec 取重复的访问次数最高节点进行填补,便于下述卷积聚合计算。总体采样过程如算法 1 所示。

算法 1 图采样算法

输入: 采样邻居数量 n , 目标节点 t , t 的邻居集合 $N(t)$, 游走长度 k , 游走概率集合 P

输出: 采样的邻居集合 $\hat{N}(t)$

1. 初始化: $l \leftarrow 0, s \leftarrow t$ /* 从目标 t 出发 */;
2. while $l < k$ do
3. $s' \leftarrow s$. nextWalk($P, N(t)$)
4. if $t! = s'$ then
5. $C(s') \leftarrow C(s') + 1$
6. end if
7. $s \leftarrow s'$
8. $l \leftarrow l + 1$
9. end while
10. $\hat{N}(t) \leftarrow \text{top}(n, C)$

其中, nextWalk 是游走函数,可以通过概率参数 P 和

范围参数 $N(t)$ 获取到下一个游走的节点; C 为一个计数函数,用于获取参数节点的访问次数; top 是一个排序函数,可根据计数数组得到 $top\ n$ 的结果。

3.1.3 卷积聚合计算

基于使用图构建和图上的关系集合定义,本文将图卷积计算^[25]引入到应用评分推荐中来,其整体的聚合计算过程可以抽象为式(4)。假设用户集合为 $U^{l \times |U|}$,应用集合为 $A^{l \times |A|}$ 以及使用关系集合为 $Q^{U \times |A|}$,图嵌入模块的目标是构建一个模型 F ,将图 G 转换为用户和应用的嵌入向量组合。

$$[u_1, u_2, \dots, u_{|U|}], [a_1, a_2, \dots, a_{|A|}] \stackrel{F}{\leftarrow} (U, A, Q) \quad (4)$$

其中, $u_i^{1 \times l}$ 和 $a_j^{1 \times l}$ 分别表示第 i 个用户和第 j 个应用的 l 维嵌入向量。

AppGRec通过归纳式的图节点信息传播来训练图嵌入向量。在此过程中,本文使用典型的卷积聚合计算进行图信息传播,即经过初始的Word2Vec向量初始化后,信息从底层的各实体节点逐步通过卷积计算传播至其邻居节点,最终得到整个图结构的嵌入结果。本文方法与传统的直推式卷积图神经网络(如GCNs^[26-27])传播的关键不同在于,信息传播并不直接训练各邻居节点的向量,而是训练连接邻居节点的卷积函数 $conv$ 。传统GCN在训练时需要对所有图节点做拉普拉斯变换,当节点很多时,计算性能会显著降低,且每当遇到新节点时就需要重新计算,造成冷启动问题。而本文采用GraphSage^[28]的聚合函数计算思想,训练节点之间的聚合函数,以此得到嵌入表征的计算方法。当遇到新节点时,只需要按照训练好的聚合函数进行多层的嵌入计算即可得到其嵌入向量,并不需改变图中其他节点的嵌入向量值。本文的卷积计算主要分为3个步骤。首先对于目标实体 e ,得出其邻居向量的总体卷积表征 z_n 。

$$z_n = \frac{1}{|\ddot{N}(e)|} \sum_{v \in \ddot{N}(e)} (\text{ReLU}(W_1 z_v + b_1)) \quad (5)$$

其中, $\ddot{N}(e)$ 是通过随机游走采样而得到的 e 的固定数量的邻居节点, z_v 是这些邻居向量的初始嵌入向量。而后,目标实体的嵌入向量更新为与邻居节点的融合向量 z_e' 。

$$z_e' = \text{ReLU}(W_2 (z_e \oplus z_n) + b_2) \quad (6)$$

其中, \oplus 指向量求和运算。为了使训练更加稳定且效率更高,在融合计算之后再进一步标准化。

$$z_e'' = z_e' / \|z_e'\|_2 \quad (7)$$

至此,我们得到了一个局部的卷积聚合计算,即通过邻居节点更新目标节点的一步计算。而全局的卷积聚合,即目标节点嵌入向量的最终生成,则是通过多个局部卷积迭代完成,一直迭代到相邻节点出现重复运算为止。在此,每步的更新过程共用同一个参数集合 $\{W_1, b_1, W_2, b_2\}$,以达到聚合函数共享的效果。这不仅有助于效率的提升,而且可直接用于新加入实体的嵌入表征。如算法2所示。

算法2 全局卷积聚合算法

输入:实体节点集合 E

输出:嵌入向量集合 V

1. 初始化: $Ls \leftarrow \text{Null}$, $Qe \leftarrow \text{Null}$ /* 初始化访问列表和功能队列 */;
2. for each e in E do

3. $Qe.\text{Insert}(e)$
4. while(! $Qe.\text{IsEmpty}()$)
5. $Qe.\text{Delete}(e)$;
6. $Ls.\text{add}(e)$;
7. for each nb in $\ddot{N}(e)$ do
8. if($nb \notin Ls$) then
9. $nb.\text{parent} \leftarrow e$
10. $Qe.\text{Insert}(nb)$
11. else $Leaf \leftarrow nb$ /* 标记叶子节点便于寻找聚合起始点 */
12. end if
13. end for
14. end while
15. $Qe.\text{Insert}(Leaf)$
16. while(! $Qe.\text{IsEmpty}()$) do
17. if(! $isAggregated(leaf.\text{parent})$)then
18. $leaf.\text{parent} \leftarrow \text{localAggregate}(leaf.\text{parent})$ /* 局部聚合计算 $leaf.\text{parent}$, 得到其嵌入向量 */
19. $Qe.\text{Insert}(leaf.\text{parent})$
20. end if
21. end while
22. $V.\text{add}(e)$ /* 所有局部聚合结束, 得到 e 最终的嵌入向量 */
23. end for
24. return V

3.2 深度因子分解预测

由于移动应用的用户使用数据极为稀疏,即一个用户只会使用整个移动应用集合中的极少部分;相反,除极个别常见应用(如微信)之外,一个移动应用也只会所有用户中的极少部分所使用。在此数据特点下,运用传统矩阵分解推荐算法^[3,29]将面临严重的数据稀疏性问题。因此,本文选用因子分解机(Factorization Machine, FM)^[28]处理稀疏数据。然而FM对高阶特征的支持较弱,需要使用深度神经网络进行高阶交互特征捕获,形成了深度因子分解机^[30]。虽然在图嵌入阶段节点向量经历了多步的卷积训练,但这并不能代替在预测阶段的高阶特征提取。FM的预测值由一阶线性运算和二阶交互运算组合得到^[28]。

$$y_{fm} = W_0^f + \sum_{i=1}^n W_i^f x_i + \sum_{i=1}^{n-1} \sum_{j=i+1}^n \langle v_i, v_j \rangle x_i x_j \quad (8)$$

其中, n 为特征数量, x_i 为第 i 个特征, w_0, w_i 和 $\langle v_i, v_j \rangle$ 是要学习的模型参数, $\langle v_i, v_j \rangle$ 表示低阶交互特征参数。FM在实际运算时通常将二阶项转换为二次差值形式便于求梯度,在此不做赘述。需要注意的是,这里的输入特征不再是人工提取的独立特征,而是图嵌入的向量集合。因此需要在计算之前将图嵌入向量进行维度的均等划分后再输入计算。即:

$$z_e'' = x_1 \square x_2 \square \dots \square x_l \quad (9)$$

其中, \square 表示连接运算。该模块的深度部分采用多层DNN实现,假设共有 l 层,则深度部分的预测值可通过式(10)得到。

$$y_{dnn} = \text{ReLU}(W_{(l)}^d s_{(l-1)} + b_{(l)}^d) \quad (10)$$

其中,输入 $s_{(l-1)}$ 的初始值 $s_{(0)}$ 是由图嵌入模块的嵌入向量 z_e'' 给定。 $s_{(l-1)}$ 即由 $s_{(0)}$ 根据式(10)迭代而来。注意,式(10)和式(6)作用在不同的区域范围以实现不同的功能,并不是简单的重复。式(10)是对整个图嵌入进行深度特征提取,而式(6)则是对局部邻居信息进行深度传递。总体的

预测向量为两部分之和。

$$y_{bt} = \text{sigmoid}(\omega_d y_{dnn} + \omega_f y_{fm}) \quad (11)$$

3.3 多任务性能提升

移动应用的用户评价不仅通过总体评分来体现,其他交互特征如下载量也可以侧面反映用户对目标应用的欢迎程度,这些特征之间是相互关联的。基于该观察,AppGRec引入了多任务学习策略来挖掘预测潜力,进一步提升预测准确性。

多任务模块对于每个子任务使用一个独立的 MLP 多层感知机作为独立塔来进行训练。将深度因子分解机作为共享底部,在反向传播的过程中进行参数共享,利用相关任务特征提高模型整体的泛化能力,降低过拟合,提升预测准确性。假设 T 为任务集合, $\tau \in T$ 表示某个子任务,则该子任务的输出可以表示为:

$$y^{\tau} = M^{\tau}(y_{bt}) \quad (12)$$

其中, M^{τ} 是该子任务的 MLP 任务塔, y_{bt} 是深度分解机的输出向量。而各子任务的损失函数是一个带有 L2 正则项的均方差(MSE)计算式。

$$\mathcal{L}^{\tau} = \frac{1}{|Y^{\tau}|} \sum_{y \in Y^{\tau}} ((\hat{y}^{\tau} - y^{\tau}) / \|\hat{y}^{\tau} - y^{\tau}\|_2)^2 + \lambda \|\Theta\|_2^2 \quad (13)$$

其中, Θ 是子任务参数集合,而 λ 系数用于控制正则强度从而控制训练过程中的过拟合程度。由于各子任务的量纲差异,因此计算 MSE 时采用了标准化形式。模型最终的损失函数是各个子损失的加权和,如式(14)所示:

$$\mathcal{L} = \sum_{\tau \in T} \omega^{\tau} \mathcal{L}^{\tau} \quad (14)$$

其中, ω^{τ} 是各子任务权重,在模型训练时作为参数动态训练得来。整个多任务模型的共享底部(shared-bottom)参数训练通过每个子任务的反向传播得到。相较于单个任务的模型训练更具泛化能力和鲁棒性。注意,这里我们没有使用基于专家融合(Mixture-of-Experts)的多任务学习方法^[31-32],因为AppGRec的深度因子分解机足够庞大,假设MMOE模型^[32]连接若干个这样的experts并且加上相应门控逻辑,会使模型变得臃肿,严重影响运行效率。

4 实验评估

4.1 实验设置

4.1.1 数据集

实验数据集来自于 Google Play 所公开的移动应用。虽然 Google Play 包含超百万的应用,为了消除低数据量造成的分布随机性,仅收集具有丰富特征数据的应用,排除了缺少描述、评论数少于 1000、下载量低于 10000 的应用数据,最终数据集包含 16031 个 Android 市场移动应用,涉及 65483 个评分用户和 189937 个评分数据,如表 1 所列。

表 1 数据集统计

Table 1 Statistics of datasets

# App	# User	# Ratings	# AP	# AR	Sparsity
16031	65483	189937	2.9	11.85	99.982%

其中 # AP 表示平均用户的评分数, # AR 表示平均 App 的评分数。评分数侧面反应了应用使用图中的节点间交互数量及节点的邻居数量。Sparsity 列显示的是使用图的稀疏

度,显然由实验数据集所构造的使用图是极度稀疏的。此外,在统计中可以发现单连数据(即用户仅对一个 App 进行评分和评论)的占比约为 7.4%,本文在实验时将该类数据单独统计以验证 AppGRec 对冷启动的鲁棒性(见 4.2 节)。

4.1.2 实验环境与参数

对比实验在一台高性能计算机上运行,该计算机具有 Intel Xeon CPU、32G DDR4 内存和 2 个 Tesla V100 GPU。AppGRec 在深度学习框架 Tensorflow 1.14.01 上实现,其中 dropout rate 设置为 $\{0.1, 0.3, 0.5, 0.7, 0.9\}$, epoch 设置为 3000, L2 正则化系数设置为 $\{0.001, 0.01, 0.1\}$, 批量大小设置为 $\{16, 32, 64, 128, 256\}$, 学习率设置为 0.01。实验的多任务由移动应用的下载量预测和评分预测两个子任务构成。模型的超参是通过细粒度微调确定的,实验整体上采用 10 折交叉实验进行训练和测试,其中 70% 作为训练集, 20% 作为验证集, 10% 作为测试集,以平均预测表现为最终结果。

4.1.3 评价指标

模型的评价指标采用两类常用的评分推荐指标,即平均绝对误差(Mean Absolute Error, MAE)和均方根偏差(Root Mean Squared Error, RMSE),如式(15)和式(16)所示:

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (15)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (16)$$

其中, n 为测试样本规模。

4.1.4 基准方法

为了验证 AppGRec 在移动应用评分预测任务上的有效性,在实验数据集上将本文方法与当前先进的评分预测算法进行对比。本文所选取的基准对比方法涵盖了深度推荐模型、移动应用推荐模型和图嵌入模型的多个对比角度。其中 DeepFM^[23] 和 NARRE^[33] 是两类最具代表性的深度推荐模型; KDFM^[23] 是面向移动应用评分推荐任务的典型深度模型;而 NGCF^[34] 和 Pinsage^[25] 是两类先进的图嵌入评分预测模型。

DeepFM 是深度推荐任务的代表性模型,其在特征提取方面结合了深度特征提取和交互特征提取两部分的优点,将 DNN 网络和 FM 因子分解机进行广度融合形成最终的综合评分。

NARRE 在卷积神经网络(CNN)的基础上,通过用户评论集合并行学习用户行为和物品属性特征。此外,其利用注意力机制综合考虑评论的贡献程度,进一步提高模型准确度和可解释性。

KDFM 是面向移动应用推荐的典型深度模型,其首先利用主题注意力机制进行异构数据的嵌入表征,而后利用 DeepFM 的特征提取优势进行最终预测。在进行基于异构数据的移动应用评分预测时, KDFM 相较于传统 DeepFM 更具优势。

NGCF 是基于图神经网络的典型推荐算法。它通过建立用户-物品二部图,以 GCN 的消息传播思想让推荐特征从邻居节点中聚合,再通过高阶连通关系让特征在图中各层传播。

PinSage 是一种典型的归纳型图嵌入推荐模型,它利用

随机游走采样确定目标节点的邻居节点,并通过训练聚合函数得到模型预测参数,相较于传统的直推式(Transductive)图嵌入表征方法更为高效、准确。

本文在实现基准方法时,均使用 MSE 损失函数进行预测。

4.2 结果分析

对比实验的 MAE 和 RMSE 结果如表 2 所列,其中 Overall 部分为全体样本的预测结果,Single-connected 部分为单连数据预测结果,单连数据对冷启动鲁棒性进行了更为直接的验证。从全体样本结果上(Overall 部分)看,与现有的先进方法相比,AppGRec 在 MAE 和 RMSE 上均获得了最优结果,分别比表现最优的基准方法提升了 10.4%(Pinsage)和 10.9%(KDFM)。实验采用了 10 折交叉验证,表明 AppGRec 的性能提升并非偶然所得,证明其在真实数据集上的预测有效性。而从 4.1.1 节的数据集分析中可知,本文的实验数据集从图构建角度是极度稀疏的,侧面验证了 AppGRec 在稀疏数据预测上的优良性能。这主要归因于其使用了归纳型图嵌入和因子分解,这些模块对稀疏数据的适应性在现有工作中都得到了证明。在对比方法中,KDFM 和 Pinsage 的表现相对优于其他方法。这两种方法采用了不同的预测策略,其中 KDFM 侧重于文本嵌入和深度因子分解,而 Pinsage 侧重于归纳型图嵌入表征。而这两类提升模块也集成于 AppGRec 中,关于 AppGRec 中各模块的细节作用,将会在 4.4 节通过消融实验进行讨论。

表 2 实验结果对比

Model	Overall		Single-connected	
	MAE	RMSE	MAE	RMSE
DeepFM	1.164	1.491	1.581	1.994
NARRE	1.235	1.647	1.55	1.923
KDFM	1.083	1.295	1.462	2.09
NGCF	1.139	1.529	1.527	1.905
Pinsage	1.046	1.384	1.348	1.867
AppGRec	0.937	1.154	1.294	1.622

从单连数据集结果(Single-connected 部分)中可以发现,AppGRec 的预测表现依然优于对比方法,其中 MAE 和 RMSE 分别比最优方法提升了 4%和 13.1%,说明 AppGRec 对冷启动具有较强的鲁棒性,且预测的波动不大。

4.3 参数影响

模型验证时的超参取值会影响性能结果,虽然本文在训练时使用了 AdaGrad 优化器进行自适应参数优化,但参数取值分析仍有助于模型有效性的评估。为此,本文对比了全体样本在不同超参取值下的模型性能,如图 2—图 4 所示。

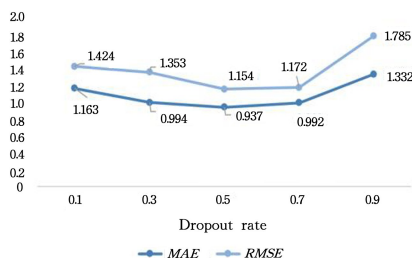


图 2 Dropout rate 对结果的影响

Fig. 2 Impact of Dropout rate on results

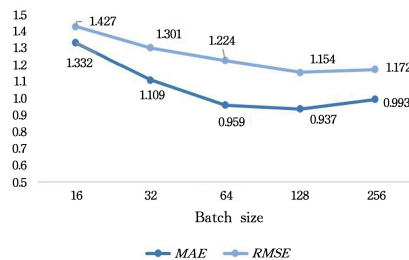


图 3 批量大小对结果的影响

Fig. 3 Impact of batch size on results

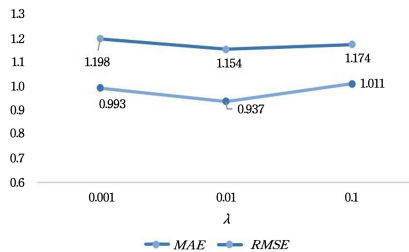


图 4 正则化系数 lambda 对结果的影响

Fig. 4 Impact of regularization coefficient lambda on results

从对比结果中可以看出,当 dropout rate 取 0.9 时对结果影响较大,此时较大的 dropout rate 可能无法缓解模型过拟合的问题。而从图 3 中可知,批量大小取值较小时,模型性能会显著降低,过小的批处理数量不仅会影响模型训练的效率,而且可能导致欠拟合的问题,使得 AppGRec 无法深入学习关键特征。图 4 表明,L2 正则化系数 lambda 的不同取值对整体结果未能造成实质影响。当 lambda 取值较大或较小时,MAE 的性能均略有下降,但并不明显,这可能是由于模型中的深度因子分解机和多任务模块的使用使得正则化的调整作用有限。

4.4 消融实验

为了验证 AppGRec 中关键模块的有效性,本文在全体样本上进行了 3 方面的消融实验,分别针对 AppGRec 中的图嵌入模块、多任务模块和深度因子分解模块,消融对比结果如图 5 所示。其中,Without GE 指去除图嵌入模块的结果,在实现中将图嵌入模块替换成传统 skip-gram^[24]的文本 Word2Vec 嵌入方式;Without MT 指去除多任务模块的结果,它使得深度因子分解机进行直接预测;而 Without DF 指去除深度因子分解模块的结果,在实现中利用多层 DNN 进行替换。

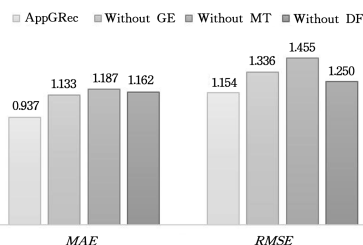


图 5 消融实验对比结果

Fig. 5 Results comparison of ablation experiments

从结果中可知,对 3 个关键模块进行消融之后的预测表现均有显著下降,验证了 AppGRec 模块设计的有效性。其中多任务模块的消融对结果影响最为显著,说明下载量预测和

评分预测之间的参数训练影响是积极的。该结果和不同任务之间相似的数据分布相关,一方面任务数据的相似性强化了模型 shared-bottom 参数的确定;另一方面相似却不相同的数据分布可以很好地规避模型训练的过拟合,提升模型的泛化能力。此外,多任务模块作用于 AppGRec 的上层,相较于图嵌入和深度因子分解仅在本模块层面实施影响,多任务模块对模型的影响是全面而显著的。本文暂时没有尝试具有其他预测子任务的多任务设置,这是因为受到移动应用评分场景的限制,缺少其他合适的可预测属性。

结束语 本文针对移动应用评分预测问题,提出了一种基于归纳型图嵌入的多任务深度预测模型 AppGRec。该模型利用图嵌入捕获稀疏矩阵中的用户使用关系特征,通过深度因子分解机提取特征间交互信息,并加入多任务模块提升模型的预测能力。通过在实际数据集上与 state-of-the-art 方法进行对比实验,验证了 AppGRec 在应用评分预测上的有效性。此外,本文利用消融实验分析了各子模块在性能提升方面的作用。AppGRec 目前存在两方面的问题,一是图嵌入采样游走虽然能够帮助选择合适的信息传播对象,但也降低了模型的训练效率,将来考虑与随机采样相结合来提升效率;二是多任务模块中缺少其他合适子任务进行对比,这是受移动应用预测场景的限制,未来可以尝试将合适的文本属性数值化以进行预测,从而进一步提升多任务的有效性。

参 考 文 献

- [1] KOREN Y. Collaborative filtering with temporal dynamics [C]//Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and data Mining. 2009: 447-456.
- [2] GUO H, TANG R, YE Y, et al. DeepFM: a factorization-machine based neural network for CTR prediction[J]. arXiv:1703.04247, 2017.
- [3] KOREN Y, BELL R, VOLINSKY C. Matrix factorization techniques for recommender systems[J]. Computer, 2009, 42(8): 30-37.
- [4] ELKAHKY A M, SONG Y, HE X. A multi-view deep learning approach for cross domain user modeling in recommendation systems[C]//Proceedings of the 24th International Conference on World Wide Web. 2015: 278-288.
- [5] WANG H, ZHANG F, XIE X, et al. DKN: Deep knowledge-aware network for news recommendation [C] // Proceedings of the 2018 World Wide Web Conference. 2018: 1835-1844.
- [6] HE X, LIAO L, ZHANG H, et al. Neural collaborative filtering [C]//Proceedings of the 26th International conference on World Wide Web. 2017: 173-182.
- [7] WANG X, HE X, CAO Y, et al. Kgat: Knowledge graph attention network for recommendation[C]//Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2019: 950-958.
- [8] FOUSS F, PIROTTE A, RENDERS J M, et al. Random-walk-computation of similarities between nodes of a graph with application to collaborative recommendation[J]. IEEE Transactions on Knowledge and Data Engineering, 2007, 19(3): 355-369.
- [9] FAN W, MA Y, LI Q, et al. Graph neural networks for social recommendation[C]//The World Wide Web Conference. 2019: 417-426.
- [10] BAE D, HAN K, PARK J, et al. AppTrends: A graph-based mobile app recommendation system using usage history[C]//2015 International Conference on Big Data and Smart Computing (BIGCOMP). IEEE, 2015: 210-216.
- [11] OUYANG Y, GUO B, TANG X, et al. Mobile App Cross-Domain Recommendation with Multi-Graph Neural Network[J]. ACM Transactions on Knowledge Discovery from Data (TKDD), 2021, 15(4): 1-21.
- [12] YU H, XIA X, ZHAO X, et al. Combining collaborative filtering and topic modeling for more accurate android mobile app library recommendation[C]//Proceedings of the 9th Asia-Pacific Symposium on Internetware. 2017: 1-6.
- [13] LIN K P, CHANG Y W, SHEN C Y, et al. Leveraging online word of mouth for personalized app recommendation[J]. IEEE Transactions on Computational Social Systems, 2018, 5(4): 1061-1070.
- [14] YIN H, CHEN L, WANG W, et al. Mobi-sage: A sparse additive generative model for mobile app recommendation [C] // 2017 IEEE 33rd International Conference on Data Engineering (ICDE). IEEE, 2017: 75-78.
- [15] CAO D, NIE L, HE X, et al. Version-sensitive mobile app recommendation[J]. Information Sciences, 2017, 381: 161-175.
- [16] YAO Y, ZHAO W X, WANG Y, et al. Version-aware rating prediction for mobile app recommendation[J]. ACM Transactions on Information Systems (TOIS), 2017, 35(4): 1-33.
- [17] LIU B, WU Y, GONG N Z, et al. Structural analysis of user choices for mobile app recommendation[J]. ACM Transactions on Knowledge Discovery from Data (TKDD), 2016, 11(2): 1-23.
- [18] TU Z, LI Y, HUI P, et al. Personalized Mobile App Recommendation by Learning User's Interest from Social Media[J]. IEEE Transactions on Mobile Computing, 2019, 19(11): 2670-2683.
- [19] XU Y, ZHU Y, SHEN Y, et al. Leveraging app usage contexts for app recommendation: a neural approach [J]. World Wide Web, 2019, 22(6): 2721-2745.
- [20] LIANG T, SHENG X, ZHOU L, et al. Mobile app recommendation via heterogeneous graph neural network in edge computing [J]. Applied Soft Computing, 2021, 103: 107162.
- [21] XIE F, CAO Z, XU Y, et al. Graph neural network and multi-view learning based mobile application recommendation in heterogeneous graphs[C]//2020 IEEE International Conference on Services Computing (SCC). IEEE, 2020: 100-107.
- [22] ZHANG M, ZHAO J, DONG H, et al. A knowledge graph based approach for mobile application recommendation[C]//International Conference on Service-Oriented Computing. Cham: Springer, 2020: 355-369.
- [23] GUO C, XU Y, HOU X, et al. Deep attentive factorization machine for app recommendation service[C]//2019 IEEE Interna-

- tional Conference on Web Services (ICWS). IEEE, 2019: 134-138.
- [24] MIKOLOV T, SUTSKEVER I, CHEN K, et al. Distributed representations of words and phrases and their compositionality [J]. arXiv:1310.4546, 2013.
- [25] YING R, HE R, CHEN K, et al. Graph convolutional neural networks for web-scale recommender systems[C]// Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2018:974-983.
- [26] KIPF T N, WELING M. Semi-supervised classification with graph convolutional networks[J]. arXiv:1609.02907, 2016.
- [27] VELÍČKOVIČ P, CUCURULL G, CASANOVA A, et al. Graph attention networks[J]. arXiv:1710.10903, 2017.
- [28] HAMILTON W, YING Z T, LESKOVEC J. Inductive representation learning on large graphs[C]// Proceedings of Advances in Neural Information Processing Systems. 2017:1024-1034.
- [29] SALAKHUTDINOV R, MNIH A. Probabilistic matrix factorization[C]// Proceedings of Advances in Neural Information Processing Systems. 2007:1257-1264.
- [30] RENDLE S. Factorization machines[C]// 2010 IEEE International Conference on Data Mining. IEEE, 2010:995-1000.
- [31] SHAZEER N, MIRHOSEINI A, MAZIARZ K, et al. Outrageously large neural networks: The sparsely-gated mixture-of-experts layer[J]. arXiv:1701.06538, 2017.
- [32] MA J, ZHAO Z, YI X, et al. Modeling task relationships in multi-task learning with multi-gate mixture-of-experts [C] // Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2018:1930-1939.
- [33] 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. 2018:1583-1592.
- [34] WANG X, HE X, WANG M, et al. Neural graph collaborative filtering[C]// Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval. 2019:165-174.



LI Haiming, born in 1964, Ph.D, professor, master supervisor. His main research interests include intelligent information processing and power informatization.



GUO Chenkai, born in 1988, Ph.D, associate professor, master supervisor, is a member of China Computer Federation. His main research interests include intelligent software engineering and code analysis of mobile app.

(责任编辑:何杨)