

基于深度学习的协同过滤推荐模型

王莉, 徐亮

(江苏联合职业技术学院徐州医药分院, 江苏 徐州 221116)

摘要: 为了探索深度神经网络在推荐模型中的应用前景, 文中构建了一个集成了深度神经网络和矩阵分解的推荐模型, 通过与用户的反馈配合使用实现协作过滤和评分预测。用户评分被视为对商业项目的用户明确反馈, 然后将其投影到用户和项目的共享空间中, 将表示每个用户和每个项目的潜在特征向量作为嵌入形式传递到矩阵分解和多层感知器模型中, 最后将这两个结果嵌入进行组合, 并传递到另一个神经网络结构中依次为每个用户的商品项目进行评分预测。实验表明, 相对于一般的协同过滤模型, 文中提出方法的 MAE 和 RMSE 分别明显下降, 具有良好的预测性能。

关键词: 神经网络; 深度学习; 协同过滤; 矩阵分解

中图分类号: TP399 **文献标识码:** A

Collaborative filtering recommendation model based on deep learning

WANG Li, XU Liang

(Jiangsu United Vocational and Technical College Xuzhou Medical Branch, Xuzhou 221116, Jiangsu Province, China)

Abstract: In order to explore the application prospects of deep neural networks in recommendation models, a recommendation model is built by integrated deep neural networks and matrix factorization. And collaborative feedback and user feedback is used to achieve collaborative filtering and scoring prediction. User ratings are considered as explicit feedback to users of commercial projects, then projected into the shared space of users and projects. Finally, the potential feature vectors representing each user and each project are passed to the matrix decomposition and multi-layer as an embedded form. In the perceptron model, these two results are finally embedded and combined, and passed to another neural architecture to make score predictions for each user's product items in turn. Experiments show that compared with the general collaborative filtering model, the MAE and RMSE of the proposed method are significantly reduced and have good prediction performance.

Key words: neural network; deep learning; collaborative filtering; matrix decomposition

0 引言

旨在向用户提出有用信息来帮助用户进行决策的推荐模型已在电子商务网站和社交网站得到广泛应用, 对改进商业营销策略、提升消费体验起到了至关重要的作用^[1-3]。尽管深度神经网络在自然语言处理、图像识别和语音识别方面取得良好成果, 但是少有针对在推荐模型中探索深层神经网络作用的研究。为此本文提出了一种集成深度神经网络的协同

过滤推荐模型。在该模型中, 首先将矩阵分解的内点积替换为可以学习近似任意非线性神经网络的神经框架, 然后将此神经框架与矩阵分解模型组合, 以将 MF 的线性方面与深度神经模型的非线性相结合。本文还使用实验数据集进行实验以研究所提出方法的

收稿日期: 2020-02-13

基金项目: 江苏联合职业技术学院科研项目(B/2018/07/037)

作者简介: 王莉(1977-), 女, 本科, 讲师, 工程师, 研究方向为大数据技术、计算机网络技术。

性能。

1 模型设计

推荐模型通常分为内容过滤、知识过滤和协同过滤三种类型^[4-5]。内容过滤是根据用户个人资料和商业项目之间的相似性提出推荐信息。这种类型的推荐模型在新闻推荐应用中被大量采用。知识过滤是使用有关商业项目、用户偏好和过滤标准提出相关信息的推荐。作为个性化推荐模型最常用的一种信息过滤方法,协作过滤通过累积来自不同用户的信息为用户做出直观的预测^[6-7]。

协同过滤主要分为基于记忆和基于模型两种类型^[8-9]。基于记忆的协作过滤使用评分信息计算用户或信息项目的相似性,提出推荐信息。基于模型的协作过滤使用不同的机器学习和数学模型来学习和对项目进行评分^[10-11]。在基于模型的协同过滤中,最常见的算法是矩阵分解法(Matrix Factorization, MF)^[12]。MF使用内积来模拟用户与信息项目之间的交互。尽管矩阵分解方法是有效的,但由于交互功能使用用户和项目之间的向量点积来表达,而点积只是对潜在因子进行线性乘法,不足以表示相互作用函数的复杂性,因此其性能会受到阻碍^[13]。

本研究利用深度神经网络和矩阵分解算法对基于用户明确反馈的协同过滤推荐模型进行改进,以期获得能够准确预测用户偏好的内容推荐模型。下面对本文所提出的模型设计进行阐述。

1.1 数据准备

用户对信息项目的明确反馈可以表示为表格,其中每个用户对每个项目给出评分。表1为用户对信息项目的反馈评分的示例。在此示例中,用户1为项目3、5给出了3分,其他项目的反馈评分依此类推。

表1 用户对信息项目的反馈评分示例

用户	项目	评分(1-5)
1	3	3
2	5	3
4	5	1

一个用户可以对多个项目进行评分。其背后的原因是项目数量远远超过个性化推荐模型中的用户数量。由于这个原因,唯一用户和唯一项将被转换为多维空间或可以转换为向量以可视化它们之间关系的任何事物。

因此,根据潜在特征的数量分别为每个用户和每个项目制定用户嵌入和项目嵌入,以确定用户之

间关系的强度。表2为在表1中为用户进行用户嵌入的示例。表2中,用户1、2和4根据三个潜在因子转换为嵌入,因为每行具有三个值。可以根据潜在因素的数量,以类似的方式将唯一项转换为项目嵌入。

表2 在表1中为用户嵌入用户关系强度的示例

关系强度	用户1	用户2	用户4
用户1	1	0.72	0.01
用户2	0.21	1	0.03
用户4	0.60	0.03	1

1.2 预测方法

所提出协同过滤方法的评分预测流程如图1所示。首先,为系统中的所有唯一用户和项生成用户嵌入和项目嵌入,然后将这两种嵌入分别转化为矩阵分解和多层感知器,随后将合成的向量连接起来并传递到单层感知器中。最后在输出层生成对评分的预测结果。

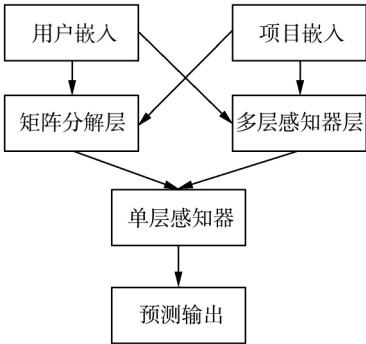


图1 推荐模型

该模型学习了将矩阵分解的线性方面与多层感知器的非线性相结合的用户与项目的交互功能,从而实现未见项目的用户评分预测。

①矩阵分解层:在矩阵分解(MF)层中,为每个用户和每个项目合并了一个潜在因子的实值向量。让用户 u 和项 i 的嵌入分别由 p_u 和 q_i 表示。然后,通过下式计算MF的 p_u 和 q_i 的点积 y_{ui} :

$$\hat{y} = f(u, i; p_i, q_i) = \sum_{k=1}^K p_{uk} q_{ik} \quad (1)$$

其中 K 表示嵌入的潜在特征数量。由于MF对用户和项目潜在向量执行线性组合,因此可以将其视为潜在向量的线性模型。

②多层感知器层:具有相应隐藏层的多层感知器(MLP)用于理解用户和项目的潜在的交互功能。隐藏层使用ReLU激活函数^[14]。这样可以了解 p_u 和 q_i 之间的非线性相互作用。因此,如果将激活函数用 h 表示输出层,则在多层感知器的输出

\hat{y} 为^[15]:

$$\hat{y} = R(h^T \odot (z_{L-1})) \quad (2)$$

其中, $\odot(z_{L-1}) = a_L(W^T z_{L-1} + b_L) \alpha$ 。

第 x 层感知器权重矩阵、偏差矢量和激活函数分别由 W_x 、 b_x 和 α_x 表示。

③合并 MF 和 MLP 层: 为了合并 MF 和 MLP 层并了解它们各自的功能, 可以通过将两个嵌入模型附加到最终的隐藏层中来组合它们, 以提供更多的可操作性。输出层用 \hat{y} 表示为:

$$\hat{y} = R\left(h^T \alpha\left(p_u \cdot q_i + W \begin{bmatrix} p_u \\ q_i \end{bmatrix} + b\right)\right) \quad (3)$$

其中 R 为输出层中的激活函数。

2 实验过程

2.1 数据集

用于实验的数据集是广泛流行的电影评分数据集。该数据集包含对 1682 部电影的 943 位用户的 100000 评分。每个电影的评级为 1 到 5。每个用户至少对 20 部电影进行评级。该数据集还包括有关用户的简单统计信息。出于实验目的, 此数据集被分为两部分。一部分用于训练, 另一部分用于测试。训练集包含 80% 的主要数据集, 这些数据是随机分配的。其余 20% 的数据为测试集。

2.2 实验设置

实验中使用的方 法, 如 MF 层、MLP 层以及这两层的融合, 在前面的部分中进行了介绍。对于 MLP, 将使用具有 200 个节点的输入层和三个隐藏层, 分别将第 100 个、第 50 个和第 20 个节点用于每个层, 而节点数较少。为了将这两个模型连接到一个单层感知器中, 首先使用式(1)训练矩阵分解, 然后使用式(2)训练多层感知器。当将它们串联到新层中时, 直接采用串联起来的最后一个隐藏层的潜在矢量, 而无需使用式(3)进行任何输出层激活。在这一层中, 使用了一个具有 200 个节点的输入层和一个具有 100 个节点的隐藏层。该层简称为神经网络分解层(NMF)。

从 NMF 到建议模型的输出层, 由于仅使用显式等级, 因此 ReLU 用作激活函数, 不会将其映射到 0 至 1 的范围内。从零开始训练 MF 和 MLP 层使用自适应估计或 Adam 对常用参数执行较小的更改, 对常用参数进行较大的更新。将预训练的参数输入 NMF 后, 随机梯度下降(SGD)用于代替 Adam 的参数优化, 因为需要存储更新动量。

2.3 评估指标

为了评估推荐模型, 使用了平均绝对误差

(MAE) 和均方根误差(RMSE)。MAE 测量了与预测评级和真实评级的平均绝对差异, 计算方法如下:

$$MAE = \frac{1}{n} \sum_{j=1}^n |y_j - \hat{y}_j| \quad (4)$$

RMSE 是 MSE(均方误差)的平方根, 测量与预测等级和真实等级的平均平方差, 其计算方法如下:

$$RMSE = \sqrt{\frac{1}{n} \sum_{j=1}^n (y_j - \hat{y}_j)^2} \quad (5)$$

其中 y_j 是实际等级; \hat{y}_j 是预测等级; n 是评分总数。为了进行性能评估, 本实验使用了精度、召回率、 F 值和准确性四个指标。

3 实验结果

为了与所提出的 NMF 模型进行比较, 这里考虑了基于记忆的协同过滤(基于用户和基于项目), 使用余弦、皮尔逊和均方差分相似度量。此外还分别与常规 MF 和基于 MLP 模型的协同过滤进行了对比实验。

本文所提出的模型基于实验数据集的实验结果如表 3 所示。

表 3 不同 CF 方法的 MAE 和 RMSE 对比

协同过滤方法	类型	MAE	RMSE
本文方法 NMF	基于模型	0.69	0.94
MLP	基于模型	0.71	0.98
MF	基于模型	0.74	0.97
基于用户的协同过滤 (均方差分相似度)	基于记忆	0.78	0.99
基于项目的协同过滤 (均方差分相似度)	基于记忆	0.77	0.98
基于用户的协同过滤 (余弦)	基于记忆	0.81	1.02
基于项目的协同过滤 (余弦)	基于记忆	0.82	1.03
基于用户的协同过滤 (皮尔逊)	基于记忆	0.80	1.01
基于项目的协同过滤 (皮尔逊)	基于记忆	0.83	1.04

表 3 表明, 将建议的 NMF 方法与其他协作过滤方法(例如具有不同相似性指标的基于记忆的 CF 方法和基于模型的方法(MF 和 MLP))进行比较时, MAE 和 RMSE 的最小值分别为 0.69 和 0.94。这两个值都是基于本文所提出的 NMF 模型获得的。原因是本文所提出的协作过滤方法通过将非线性矩阵规划与线性矩阵规划相结合来任意地学习激活函数。这种将线性和非线性结合起来学习方法可以提

供更低的 MAE 和 RMSE 分数。

为了获取所实现的基于常规模型的 CF 方法的更多信息,表 4 给出了所提出方法的混淆矩阵。

表 4 所提出方法的混淆矩阵

		实际				
预测	评分	1	2	3	4	5
	1	214	54	27	16	5
	2	279	668	333	161	40
	3	571	1245	4256	3430	1213
	4	91	256	1016	3584	2584
	5	0	0	0	1	1

精确度、召回率、 F 值和准确度以及产生结果所需的迭代次数如表 5 所示。

表 5 基于模型的 CF 的精度比较

模型	精确度	召回率	F 值	准确度	迭代次数
NMF	0.49	0.33	0.39	0.44	12
MLP	0.46	0.32	0.37	0.42	17
MF	0.44	0.28	0.34	0.31	25

一般来说,就评分预测而言,基于模型的 CF 往往比基于记忆的 CF 表现更好。表 5 表明,与其他基于模型的 CF 相比所提出的 NMF 模型在精确度、召回率、 F 值和准确度方面性能更优。应该注意的是,本文提出的模型在迭代次数上也少于另外两个模型,因此计算效率也更好。所提出的模型在速度和性能方面比其他常用的协作过滤方法更好。

4 结束语

本文探讨了基于深度神经网络和矩阵分解的反馈推荐模型。矩阵分解模型与神经网络体系相结合,明显提高了推荐模型的性能。在实验中,提议的模型为未评分项目提供了接近实际评分的预测,这可能会对电子商务、娱乐和个性化推荐产生现实影响,从而提供更好的用户体验和销售增长。该模型

仅使用了用于预测的用户、项目和评分,未使用用户统计信息(即年龄、职业)等变量,因此在未来的研究中将更改用户和项目嵌入的影响因素的数量,进一步提升模型的性能和鲁棒性,并对模型的性能进行更加深入的评估。

参考文献:

- [1] 杨凯,王利,周志平,等.基于内容和协同过滤的科技文献个性化推荐[J].信息技术,2019,43(12):11-14.
- [2] 李振波,杨晋琪,岳峻.基于协同回归模型的矩阵分解推荐[J].图学学报,2019,40(6):983-990.
- [3] 王运,倪静,马刚.基于 FunkSVD 矩阵分解和相似度矩阵的推荐算法[J].计算机应用与软件,2019,36(12):245-250.
- [4] 艾静超.基于改进协同过滤技术的个性化旅游线路推荐研究[J].现代电子技术,2019,42(23):182-186.
- [5] 张家精,夏翼鹏,陈金兰,等.基于张量分解和深度学习的混合推荐算法[J].南京大学学报:自然科学,2019,55(6):952-959.
- [6] 胡思才,孙界平,据生根,等.基于深度神经网络和概率矩阵分解的混合推荐算法[J].四川大学学报:自然科学版,2019,56(6):1033-1041.
- [7] 黎雪微,应时,周寅.基于用户兴趣迁移的网络图书推荐模型研究[J].图书馆学研究,2019(22):56-65.
- [8] 厉小军,柳虹,施寒潇,等.基于深度学习的课程推荐模型[J].浙江大学学报:工学版,2019,53(11):2139-2145,2162.
- [9] 曹银,李伟.基于信任和图聚类的协同过滤推荐算法[J].信息技术,2019,43(10):77-82.
- [10] 邱神芬,黄业文,宋杰,等.基于深度矩阵分解网络的矩阵填充方法[J].计算机科学,2019,46(10):55-62.
- [11] 邓存彬,虞慧群,范贵生.融合动态协同过滤和深度学习的推荐算法[J].计算机科学,2019,46(8):28-34.
- [12] 宋瑞雪,李国勇.基于改进的矩阵分解模型在推荐系统中的应用[J].计算机应用,2019,39(S1):93-95.
- [13] 王永贵,尚庚.融合注意力机制的深度协同过滤推荐算法[J].计算机工程与应用,2019,55(13):8-14.
- [14] 李同欢,唐雁,刘冰.基于深度学习的多交互混合推荐模型[J].计算机工程与应用,2019,55(1):135-141.
- [15] 吴彦文,刘闯.基于用户偏好和可疑度的推荐方法研究[J].计算机应用研究,2018,35(12):3632-3634.

责任编辑:杨静

(上接第 65 页)

- [12] Holt B, Ong E J, Cooper H, et al. Putting the pieces together: connected poselets for human pose estimation [C]. IEEE International Conference on Computer Vision Workshops. IEEE, 2011: 1196-1201.
- [13] Hilde E P, Sicco A B, Frans N, et al. Gait patterns in association with underlying impairments in polio survivors with calf muscle weakness [J]. Gait & Posture, 2017, 58: 146.

- [14] Andrew J J, Smith, Edward D L. Temporal-spatial gait parameter models of very slow walking [J]. Gait & Posture, 2018, 61: 125-129.
- [15] Hansen L, Siebert M, Diesel J, et al. Fusing information from multiple 2D depth cameras for 3D human pose estimation in the operating room [J]. International Journal of Computer Assisted Radiology and Surgery, 2019, 14(8): 1871-1879.

责任编辑:丁玥