

DOI:10.16356/j.1005-2615.2020.05.008

基于深度学习的欧几里得嵌入的推荐算法

余永红¹, 殷凯宇¹, 王强¹, 张文彪¹, 赵卫滨^{1,2}

(1. 南京邮电大学通达学院, 扬州, 225127; 2. 计算机软件新技术国家重点实验室(南京大学), 南京, 210023)

摘要: 推荐系统为用户推荐用户可能感兴趣的物品, 可以有效地减轻信息过载。基于欧几里得嵌入的协同过滤方法将用户和物品映射到统一的隐藏空间中, 是构建推荐系统的重要方法之一。然而, 传统的基于欧几里得嵌入的推荐方法仅考虑用户和物品隐藏特征向量之间低阶交互, 不能有效建模现实世界中用户和物品的复杂交互行为。本文提出基于深度学习的欧几里得嵌入的协同过滤算法, 利用深度学习技术学习用户和物品隐藏特征向量之间的高阶、非线性交互函数, 建模用户和物品之间复杂交互行为。在真实数据集上的实验结果表明, 基于深度学习的欧几里得嵌入的协同过滤算法性能优于传统协同过滤算法。

关键词: 推荐算法; 协同过滤; 欧几里得嵌入; 深度学习

中图分类号: TP181 **文献标志码:** A **文章编号:** 1005-2615(2020)05-0729-07

Recommendation Algorithm of Euclidean Embedding Based on Deep Learning

YU Yonghong¹, YIN Kaiyu¹, WANG Qiang¹, ZHANG Wenbiao¹, ZHAO Weibin^{1,2}

(1. Tongda College, Nanjing University of Posts and Telecommunications, Yangzhou, 225127, China;
2. State Key Laboratory for Novel Software Technology (Nanjing University), Nanjing, 210023, China)

Abstract: Recommender systems can effectively reduce the information overload by recommending items that users may be interested in. Euclidean-embedding-based collaborative filtering methods map users and items to a unified latent space, which is one of the most important methods to build a recommender system. However, traditional Euclidean-embedding-based collaborative filtering methods only consider the low-order interaction between user latent feature vectors and item latent feature vectors, and cannot efficiently model the complex interaction behavior between users and items in the real world. In this paper, we propose a deep-Euclidean-embedding-based collaborative filtering algorithm, which utilizes deep learning technology to learn the high-order and nonlinear interaction function between user latent feature vectors and item latent feature vectors. This can model the complex interaction behavior between users and items. Experimental results on real-world datasets show that our proposed algorithm outperforms traditional collaborative filtering algorithms.

Key words: recommendation algorithm; collaborative filtering; Euclidean embedding; deep learning

在信息爆炸的时代, 互联网给用户带来了丰富的信息, 满足了用户对信息的需求, 但是信息量的成倍增长使得用户很难从海量数据中找到有价值的相关信息。推荐系统^[1]可以帮助用户筛选信息, 快速地找到满足用户需求的个性化信息, 从而减轻

信息过载。

协同过滤算法^[2-5]作为目前应用最广泛的推荐技术, 根据用户的历史评分记录来预测用户的偏好。矩阵分解^[6-10]是目前最流行的协同过滤方法。矩阵分解技术是将用户-物品评分矩阵分解为两个

基金项目: 江苏省高校自然科学基金(17KJB520028)资助项目; 江苏省青蓝工程资助项目; 南京邮电大学校级科研基金(NY217114)资助项目。

收稿日期: 2019-08-10; **修订日期:** 2020-09-02

通信作者: 余永红, 男, 副教授, E-mail: yuyh@njupt.edu.cn。

引用格式: 余永红, 殷凯宇, 王强, 等. 基于深度学习的欧几里得嵌入的推荐算法[J]. 南京航空航天大学学报, 2020, 52(5): 729-735. YU Yonghong, YIN Kaiyu, WANG Qiang, et al. Recommendation algorithm of Euclidean embedding based on deep learning[J]. Journal of Nanjing University of Aeronautics & Astronautics, 2020, 52(5): 729-735.

低维隐藏特征矩阵:用户隐藏特征矩阵和物品隐藏特征矩阵,然后利用这两个低维隐藏特征矩阵的乘积来近似原始的用户-物品评分矩阵,从而预测原始的用户-物品评分矩阵中的缺失项。

但是,矩阵分解技术是将用户-物品评分矩阵分解到两个不同的隐藏特征空间,所以用户和物品之间的相关性并不能被直接观察到。为了解决此问题,基于欧几里得嵌入协同过滤算法^[11]在传统的矩阵分解技术上改善了用户和物品之间的相关性表示,它将用户和物品嵌入到统一的隐藏特征空间中,并用欧式距离来直观地表示用户和物品之间的相关性。另外,在为用户进行个性化推荐时,基于欧几里得嵌入协同过滤算法可以直接利用该用户节点在隐藏特征空间中最近范围内的物品来进行推荐,有效提高了生成推荐列表的效率。但是,传统的基于欧几里得嵌入的推荐方法仅考虑了用户和物品隐藏特征向量之间的低阶线性交互,不能有效地建模现实世界中用户和物品之间的复杂交互行为。

针对以上问题,本文提出了一种基于深度学习的欧几里得嵌入的协同过滤算法,称为DEE算法。该算法利用深度学习技术学习用户和物品隐藏特征向量之间的高阶、非线性交互函数,建模用户和物品之间复杂交互行为。非线性模型比线性模型具备更强的函数拟合能力。与低阶线性交互相比,高阶非线性交互赋予模型更高的灵活性和非线性建模的能力,使得模型更具表达力。在真实数据集上的实验结果表明,基于深度学习的欧几里得嵌入的协同过滤算法性能优于传统协同过滤算法。

1 相关研究工作

1.1 协同过滤推荐算法

协同过滤推荐算法通过分析用户的历史活动记录(一般为用户-物品评分矩阵)来进行推荐,独立于具体的应用领域,是目前应用最广泛的推荐技术。协同过滤推荐算法主要分为基于内存的协同过滤推荐算法、基于模型的协同过滤算法和混合推荐算法3大类^[1]。

基于内存的协同过滤推荐算法使用整个用户-物品评分矩阵进行推荐。它通过相似用户对目标物品的评分权重平均或者活动用户对相似物品的评分权重预测当前活动用户对目标物品的评分。典型的基于内存协同过滤算法包括基于用户的协同过滤算法^[2]和基于物品的协同过滤算法^[3-4]。

基于模型的协同过滤算法使用统计和机器学习技术从用户-物品评分矩阵中学习一个预测模

型,预测模型刻画用户的行为模式,然后利用预测模型进行推荐。典型的基于模型的协同过滤算法包括:基于贝叶斯网络推荐算法^[5]、基于聚类模型推荐算法^[12-13]以及基于潜在语义分析推荐算法^[14-15]等。

由于在处理大规模数据方面良好的可扩展性和准确的预测能力,基于矩阵分解的推荐算法^[6-7]受到学术界和工业界的广泛关注。基于矩阵分解的推荐算法将用户和物品的特征向量同时映射到两个低维的隐藏因子空间中,利用用户和物品的低维特征向量的内积来预测用户对物品的评分。典型的基于矩阵分解的推荐算法包括PMF^[7],NMF^[8],SVD⁺⁺^[9]和MMMF^[10]等。

1.2 基于欧几里得嵌入的推荐算法

基于欧几里得嵌入技术的推荐算法将用户和物品嵌入到统一的欧几里得空间中,并用欧式距离来直观地表示用户和物品之间的相关性。典型的基于欧几里得嵌入技术的推荐算法有EE^[11],TEE^[16]和SREE^[17]等。EE^[11]实现了协同过滤的隐藏因子模型,它将用户和物品嵌入到统一的欧几里得空间中。在此空间中,用户和物品之间的距离与评分成反比。与矩阵分解模型相比,基于欧几里得嵌入技术的协同过滤算法具有较强的可解释性,并且有利于快速生成推荐列表。TEE^[16]在EE模型中集成了时域影响,将时间因子向量嵌入到欧几里得空间。SREE^[17]在EE模型的基础上,融合了社交关系网络信息。它利用社会信任关系来解决传统欧几里得嵌入模型的缺陷。SREE模型使用户与其信任用户具有共同的偏好,并且可以利用信任朋友的偏好来为用户进行个性化推荐,从而有效地解决数据稀疏和冷启动问题。

2 预备知识

2.1 推荐问题的形式化描述

协同过滤算法主要利用用户-物品评分矩阵来预测用户对物品的评分。用户-物品评分矩阵 $R^{M \times N}$ 由2种实体集合组成: M 个用户集合 $U = \{u_1, u_2, \dots, u_M\}$ 和 N 个物品集合 $I = \{i_1, i_2, \dots, i_N\}$ 。 $R^{M \times N}$ 中的每项 r_{ui} 是用户 u 对物品 i 的评分。评分数据 r_{ui} 可以为任意的实数,但通常情况下,评分数据为整数,而且 $r_{ui} \in \{0, 1, 2, 3, 4, 5\}$,其中0值表示用户未对此物品进行评分。评分值越高意味着用户对当前物品越满意。由于用户通常只对少量的物品进行评分,用户-物品评分矩阵 $R^{M \times N}$ 通常非常稀疏。

推荐系统的任务主要利用用户-物品评分信息

预测当前活动用户 u 对目标物品 i 的评分值 \hat{r}_{ui} , 并根据预测评分为用户提供物品推荐列表。

2.2 基于欧几里得嵌入技术的系统协同过滤算法

基于欧几里得嵌入技术的协同过滤算法将用户和物品嵌入到统一的隐藏特征空间中,并用欧式距离来直观地表示用户和物品之间的相关性。形式化地,设用户的隐藏特征矩阵和物品隐藏特征矩阵分别为 $P \in R^{K \times N}$ 和 $Q \in R^{K \times M} (K \ll \min(N, M))$, 其中 K 为隐藏特征向量的维度。

基于欧几里得嵌入技术的协同过滤算法通过在欧几里得空间中用户和物品之间的欧式距离来预测评分,形式化为

$$\hat{r}_{ui} = \mu + b_u + b_i - (\mathbf{p}_u - \mathbf{q}_i)(\mathbf{p}_u - \mathbf{q}_i)^T \quad (1)$$

式中: μ 表示全局平均值, b_u 和 b_i 分别表示用户 u 和物品 i 的偏差, \mathbf{p}_u 和 \mathbf{q}_i 分别表示用户 u 和物品 i 的隐藏特征向量。

基于欧几里得嵌入技术的协同过滤算法通过最小化误差平方和损失函数来学习隐藏特征矩阵 P 和 Q , 即

$$L = \sum_{(u, i) \in \Omega} [(r_{ui} - \mu - b_u - b_i + (\mathbf{p}_u - \mathbf{q}_i)(\mathbf{p}_u - \mathbf{q}_i)^T)^2] +$$

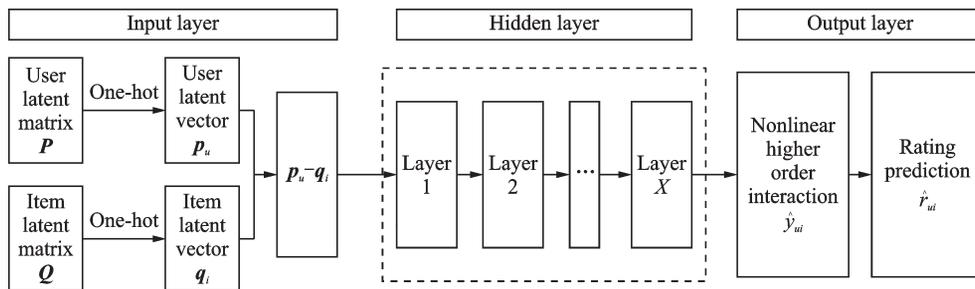


图 1 基于深度学习的欧几里得嵌入的协同过滤算法框架

Fig.1 Framework of collaborative filtering based on deep Euclidean embedding

基于深度学习的欧几里得嵌入的协同过滤算法主要由输入层、隐藏层和输出层 3 个部分组成。

(1) 输入层: 输入层包括两个低维特征矩阵, 分别表示用户特征矩阵 P 和物品特征矩阵 Q 。本文利用 one-hot 的方法可以得到用户特征向量 \mathbf{p}_u 和物品特征向量 \mathbf{q}_i 。

(2) 隐藏层: 传统的基于欧几里得嵌入的推荐算法是将用户和物品嵌入到统一的欧几里得空间中, 但是其学习的是用户和物品之间的线性低阶交互, 不能有效建模用户与物品的复杂交互行为。因此利用多层感知器来学习用户和物品之间的非线性高阶交互。

在多层感知器中, 将用户特征向量 \mathbf{p}_u 和物品

$$\lambda (\|\mathbf{p}_u - \mathbf{q}_i\|^2 + b_u^2 + b_i^2) \quad (2)$$

式中: Ω 为用户-物品评分矩阵 R 中可观测(用户, 物品)对集合, $\|\cdot\|^2$ 为 Frobenius 范式。正则化项 $(\|\mathbf{p}_u - \mathbf{q}_i\|^2 + b_u^2 + b_i^2)$ 用来避免过拟合。 λ 为正则化项参数, 用来控制正则化项对隐藏特征向量的影响。

可以采用随机梯度下降的方法求解目标函数的局部最小解, 学习用户隐藏特征向量 \mathbf{p}_u 、物品的隐藏特征向量 \mathbf{q}_i 、用户偏差 b_u 和物品偏差 b_i 。参数的更新公式为形式化为

$$b_u \leftarrow b_u + \gamma(e_{ui} - \lambda b_u) \quad (3)$$

$$b_i \leftarrow b_i + \gamma(e_{ui} - \lambda b_i) \quad (4)$$

$$\mathbf{p}_u \leftarrow \mathbf{p}_u - \gamma(\mathbf{p}_u - \mathbf{q}_i)(e_{ui} + \lambda) \quad (5)$$

$$\mathbf{q}_i \leftarrow \mathbf{q}_i + \gamma(\mathbf{p}_u - \mathbf{q}_i)(e_{ui} + \lambda) \quad (6)$$

式中: γ 为学习步长, e_{ui} 表示评分 r_{ui} 的当前误差。

3 基于深度学习的欧几里得嵌入的协同过滤算法

3.1 算法框架

基于深度学习的欧几里得嵌入的协同过滤算法框架如图 1 所示。

特征向量 \mathbf{q}_i 的差 $\mathbf{p}_u - \mathbf{q}_i$ 作为多层感知器的第 1 层输入, 经过多层感知器学习, 最终得到用户和物品之间的非线性高阶交互值 \hat{y}_{ui} 。

多层感知器隐藏层的具体过程形式化为

$$\begin{aligned} D &= [\mathbf{p}_u - \mathbf{q}_i] \\ z_1 &= \text{ReLU}(\mathbf{w}_1 D + b_1) \\ z_2 &= \text{ReLU}(\mathbf{w}_2 z_1 + b_2) \\ &\vdots \\ z_X &= \text{ReLU}(\mathbf{w}_X z_{X-1} + b_X) \\ \hat{y}_{ui} &= \text{ReLU}(\mathbf{w}_H z_X) \end{aligned} \quad (7)$$

式中: \mathbf{w}_X 表示 X 层的权重矩阵; b_X 表示 X 层的偏差; \mathbf{w}_H 为输出层的权重矩阵; z_X 表示经过激活函数处理后的第 X 层的输出向量。

(3) 输出层: 在输出层中, 基于由式(7)得到的

距离值 \hat{y}_{ui} , 预测评分值 \hat{r}_{ui} 表示为

$$\hat{r}_{ui} = \mu + b_u + b_i - \hat{y}_{ui} \quad (8)$$

式中: μ 为推荐系统所有评分的平均值; b_u 为用户 u 的偏差评分; b_i 为物品 i 的偏差评分。

最后将预测评分值 \hat{r}_{ui} 集成到欧几里得嵌入模型中进行训练。

3.2 算法模型及参数学习

与传统的基于欧几里得嵌入的推荐算法相比, 本文提出的方法不仅继承了传统的基于欧几里得嵌入的推荐算法的优点, 还考虑了用户和物品之间的非线性高阶交互, 能有效建模用户和物品之间的复杂交互行为。

为了学习欧几里得空间中的用户特征向量 p_u 和物品特征向量 q_i , 目标函数为最小化预测评分和已知评分之间的差值, 即

$$L^{\text{DEE}} = \sum_{(u,i) \in \Omega} (r_{ui} - \hat{r}_{ui})^2 \quad (9)$$

为了防止过拟合, 加上正则化项, 基于深度学习的欧几里得嵌入的协同过滤算法的目标函数形式化为

$$L^{\text{DEE}} = \sum_{(u,i) \in \Omega} [(r_{ui} - \mu - b_u - b_i + \hat{y}_{ui})^2] + \lambda (\|p_u - q_i\|^2 + b_u^2 + b_i^2) \quad (10)$$

本文采用 ADAM^[18] 优化器求解 L 的局部最小解学习模型参数。

4 实验与分析

为了验证基于深度学习的欧几里得嵌入的协同过滤算法的性能, 本文在真实的数据集上与其他流行的推荐算法进行了对比分析。

4.1 数据集

本文选择 MovieLens100K 数据集和 FilmTrust 数据集验证基于深度学习的欧几里得嵌入的协同过滤算法的性能。MovieLens100K 数据集包含 943 名用户对 1 682 部电影的 100 000 个评分记录, 用户-物品评分矩阵的稀疏度为 93.706%。FilmTrust 数据集由文献[19]提供, 包含 1 642 个用户对 2 071 部电影的 35 497 个评分记录, 用户-物品评分矩阵的稀疏度为 98.86%。

4.2 度量指标

本文采用在推荐算法中被广泛使用的评价指标: 均方根误差(Root mean squared error, RMSE)和平均绝对误差(Mean absolute error, MAE)来评价推荐算法的性能。

RMSE 的定义为

$$\text{RMSE} = \sqrt{\frac{\sum_{(i,j) \in R_{\text{test}}} |r_{ij} - \hat{r}_{ij}|^2}{|R_{\text{test}}|}} \quad (11)$$

MAE 的定义为

$$\text{MAE} = \frac{\sum_{(i,j) \in R_{\text{test}}} |r_{ij} - \hat{r}_{ij}|}{|R_{\text{test}}|} \quad (12)$$

式中: r_{ij} 和 \hat{r}_{ij} 分别表示实际的评分值和推荐系统预测的评分值; $|R_{\text{test}}|$ 表示测试数据集中的记录条数。RMSE 和 MAE 的值越小, 表明推荐算法的推荐性能越好。

4.3 实验设置

为了验证本文提出推荐算法的有效性, 选取如下推荐算法作为对比方法。

(1) UserKNN: 基于用户的协同过滤算法^[2]认为相似的用户具有相似的兴趣, 相似用户之间可以共享喜欢的物品。

(2) MF: 基于矩阵分解的推荐算法^[6]是将用户-项目矩阵分解为两个矩阵, 再通过向量内积进行预测并推荐。

(3) PMF: 概率矩阵分解^[7]由 Mnih 和 Salakhutdinov 提出。PMF 可以看作是 SVD 模型的概率扩展。

(4) U-AutoRec: U-AutoRec^[20]算法利用自编码器从用户对部分物品的评分信息中学习中间隐藏层表示, 并根据隐藏层表示重构用户评分数据。

(5) EE: 基于欧几里得嵌入技术的协同过滤算法^[11]将用户和物品嵌入到统一的欧几里得空间中, 用户和物品之间的距离与评分成反比。

本文随机将用户-物品评分矩阵分割 5 次, 其中每次抽取 80% 的数据作为训练集, 剩余 20% 的数据作为测试数据, 以推荐算法在 5 个测试集上的平均性能作为最后的运行结果。为了公平对比, 参照对比算法的相应文献或者实验结果设置不同算法的参数, 在这些参数设置下, 各对比算法取得最优性能。在 UserKNN 中, 采用余弦计算用户之间的相似度, 邻居数量为 20; 在 MF 中, $\lambda_u = \lambda_i = 0.001$; 在 PMF 中, $\lambda_u = \lambda_i = 0.001$; 在 U-AutoRec 中, $\lambda = 0.01$, 其次, 设置自编码器中间隐藏层的维度等于基于潜在因子分解模型(MF, PMF, EE)中的隐藏特征向量维度; 在 EE 中, $\lambda_u = \lambda_i = \lambda_b = 0.005$; 对于本文提出的方法 DEE, $\lambda = 0.01$ 。

实验运行环境为: 4 核 Intel(R) Core(TM) i7-4710MQ CPU, 主频 2.50 GHz, 内存 16 GB, Windows 10 操作系统和 J2SE1.8; 深度学习的框架为 Tensorflow 框架。

4.4 性能对比

在 MovieLens100K 数据集和 FilmTrust 数据集上,设置 DEE 的隐藏层层数 $X=2$ 。在隐藏特征向量维度 $K=8$ 和 $K=16$ 的情况下,所有对比算法在 2 个数据集上的实验结果如表 1 和表 2 所示。

从表 1—4 可以发现:UserKNN 算法、MF 算法和 PMF 算法的性能较差,这是由于传统协同过滤

表 1 在 MovieLens100K 数据集上的 RMSE 对比
Table 1 RMSE comparison on MovieLens100K

Algorithm	RMSE	
	$K = 8$	$K = 16$
UserKNN	0.963 767	
MF	1.149 126	1.142 656
PMF	0.966 922	1.050 630
U-AutoRec	0.972 055	0.959 055
EE	0.955 314	0.988 065
DEE	0.939 286	0.939 490

表 2 在 MovieLens100K 数据集上的 MAE 对比
Table 2 MAE comparison on MovieLens100K

Algorithm	MAE	
	$K = 8$	$K = 16$
UserKNN	0.761 761	
MF	0.947 268	0.945 449
PMF	0.742 923	0.803 363
U-AutoRec	0.764 963	0.751 951
EE	0.760 782	0.772 560
DEE	0.740 968	0.741 296

表 3 在 FilmTrust 数据集上的 RMSE 对比
Table 3 RMSE comparison on FilmTrust

Algorithm	RMSE	
	$K = 8$	$K = 16$
UserKNN	0.841 072	
MF	1.974 924	1.903 592
PMF	1.042 606	1.126 781
U-AutoRec	1.002 370	1.027 592
EE	0.813 190	0.826 568
DEE	0.806 536	0.807 158

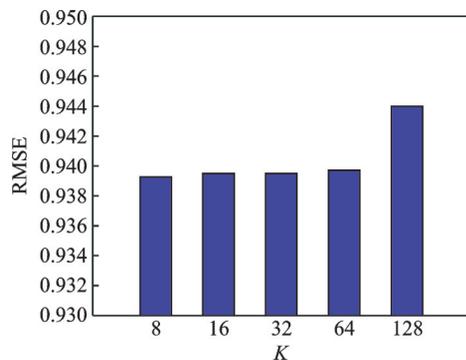
表 4 在 FilmTrust 数据集上的 MAE 对比
Table 4 MAE comparison on FilmTrust

Algorithm	MAE	
	$K = 8$	$K = 16$
UserKNN	0.641 6	
MF	1.497 560	1.475 964
PMF	0.752 228	0.825 673
U-AutoRec	0.745 328	0.762 830
EE	0.628 255	0.636 354
DEE	0.622 964	0.623 360

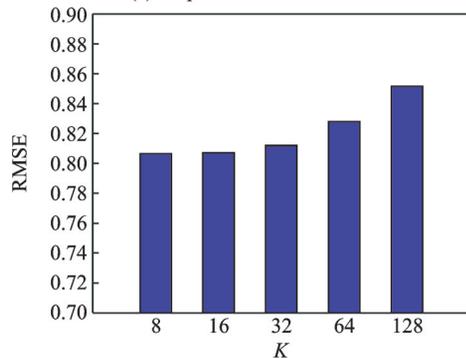
推荐算法存在的局限性,而 EE 算法和 DEE 算法的性能优于传统协同过滤推荐算法,说明将用户和物品嵌入到统一的隐藏特征空间是有效的。另外,本文提出的 DEE 算法性能优于传统的欧几里得嵌入模型,验证了本文利用深度学习技术学习用户和物品隐藏特征向量之间的高阶、非线性交互函数的有效性。在 MovieLens100K 数据集和 FilmTrust 数据集上,当 $K = 16$ 时,与对比算法中的最优 RMSE 结果对比,本文提出算法相比于传统算法,优化改进的幅度分别为 2.04% 和 2.35%。

4.5 隐藏特征向量维度 K 的影响

在本文提出的推荐算法中,隐藏特征向量维度 K 是影响推荐性能的重要参数。在本节中,从 8 到 128 变化 K 的值,观察 RMSE 在 2 个数据集上的变化情况,探索隐藏特征向量维度对 DEE 推荐性能的影响。另外,隐藏层层数 $X=2$ 。实验结果如图 2 所示。



(a) Impact of K on MovieLens100K



(b) Impact of K on FilmTrust

图 2 参数 K 对推荐性能的影响

Fig.2 Impact of parameter K on recommendation performance

从图 2 中可以看出:在 2 个数据集上,随着参数 K 的增大, RMSE 的值呈现逐渐上升的变化趋势,并且上升得十分缓慢。这个结果表明:虽然 K 值增大使得 DEE 模型可以表示较多的隐藏特征, DEE 模型的表示能力增强,但较大维度的隐藏特征向量并不能改进推荐算法的性能,这与矩阵分解模型的假设类似,即仅用少量的隐藏因子来刻画用

户的偏好和描述物品的特征。另外,在 MovieLens100K 数据集和 FilmTrust 数据集上,DEE 都是在 $K=8$ 时取得最佳性能。

4.6 隐藏层层数 X 的影响

隐藏层层数 X 是影响本文提出算法的另一个重要参数。隐藏层层数 X 控制着多层感知器中用户和物品之间的非线性高阶交互的次数。较大的 X 意味着非线性高阶交互越复杂。从 1 到 3 改变 X 值,递增步长为 1,观察 RMSE 在 2 个数据集上的变化情况。另外,隐藏特征向量的维度 $K=8$,实验结果如图 3 所示。

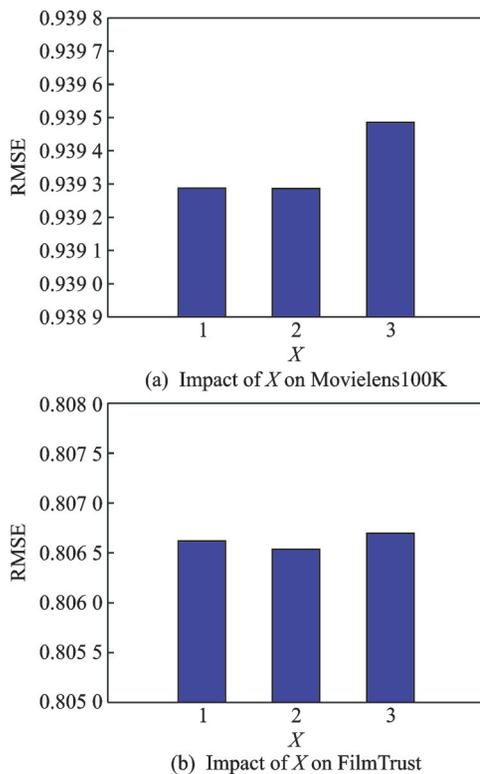


图3 参数 X 对推荐性能的影响

Fig.3 Impact of parameter X on recommendation performance

从图 3 可以看出,在 MovieLens100K 数据集和 FilmTrust 数据集上,RMSE 具有类似的变化趋势:随着 X 的增加,RMSE 逐渐降低,达到最小值,此时的推荐准确性最高,然后逐渐上升,整体的推荐性能降低。这个观测结果表明,较低的多层感知器的隐藏层层数 X 有益于 DEE 模型改进推荐性能,因为较高的 X 值意味着模型拥有更多的参数,需要大量的训练数据来学习模型参数。在 MovieLens100K 数据集和 FilmTrust 数据集上,本文提出的 DEE 模型都是在 $X=2$ 时取得最佳性能。

4.7 模型训练时间对比

在本节中,通过对比不同推荐算法的模型训练时间,以检验本文提出推荐算法的效率。在

ML100K 和 FilmTrust 数据集上,对于 DEE 模型, $\lambda=0.01, X=2$ 。其他对比推荐算法的参数设置同 4.3 节。另外,在 $K=8$ 的情况下,对比不同推荐算法在 2 个数据集上的模型训练时间,实验结果如表 5 所示。

表 5 时间性能对比

Algorithm	ML100K	FilmTrust
MF	00:00:03	00:00:04
PMF	00:00:01	00:00:01
U-AutoRec	00:11:21	00:05:10
EE	00:00:30	00:00:19
DEE	00:10:48	00:04:00

从表 5 可以看出,DEE 的模型训练时间高于其他非深度学习的方法。这是因为 DEE 需要根据训练数据额外学习隐藏层权重矩阵和偏差向量,增加了模型的训练时间。另外,DEE 模型的训练时间与 U-AutoRec 模型训练时间相当。其他非基于深度学习的协同过滤方法的模型训练时间相当。

5 结 论

欧几里得嵌入技术为解决传统协同过滤推荐算法的缺陷带来了契机,基于欧几里得嵌入的协同过滤方法将用户和物品映射到统一的隐藏空间中。然而,传统的基于欧几里得嵌入的推荐方法仅考虑用户和物品隐藏特征向量之间低阶交互,不能有效建模现实世界中用户和物品的复杂交互行为。本文提出基于深度学习的欧几里得嵌入的协同过滤算法,利用深度学习技术学习用户和物品隐藏特征向量之间的高阶、非线性交互函数,建模用户和物品之间复杂交互行为。在真实数据集上的实验结果表明,基于深度学习的欧几里得嵌入的协同过滤算法性能优于传统协同过滤算法。

参考文献:

- [1] ADOMAVICIUS G, TUZHILIN A. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions[J]. IEEE Transactions on Knowledge & Data Engineering, 2005, 17(6): 734-749.
- [2] RESNICK P, IACOVU N, SUCHAK M, et al. GroupLens: An open architecture for collaborative filtering of netnews[C]//Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work. [S.l.]: ACM, 1994: 175-186.
- [3] SARWAR B M, KARYPIS G, KONSTAN J A, et al. Item-based collaborative filtering recommendation

- algorithms[C]//Proceedings of the 10th International Conference on World Wide Web (WWW'01). New York, USA: ACM, 2001: 285-295.
- [4] LINDEN G, SMITH B, YORK J. Item-to-item collaborative filtering [J]. IEEE Internet Computing, 2003, 7(1): 76-80.
- [5] BREESE J S, HECKERMAN D, KADIE C. Empirical analysis of predictive algorithms for collaborative filtering [C]//Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence. [S.l.]: Morgan Kaufmann Publishers Inc., 1998: 43-52.
- [6] KOREN Y, BELL R, VOLINSKY C. Matrix factorization techniques for recommender systems[J]. Computer, 2009, 42(8): 30-37.
- [7] MNIH A, SALAKHUTDINOV R. Probabilistic matrix factorization[C]//Proceedings of the 21st Annual Conference on Neural Information Processing Systems. New York, USA: Curran Associates, 2007: 1257-1264.
- [8] LEE D D, SEUNG H S. Learning the parts of objects by non-negative matrix factorization[J]. Nature, 1999, 401(6755): 788-791.
- [9] KOREN Y. Factorization meets the neighborhood: A multifaceted collaborative filtering model[C]//Proceedings of the 14th ACM SIGKDD Int Conference on Knowledge Discovery and Data Mining (KDD'08). New York, USA: ACM, 2008: 426-434.
- [10] RENNIE J D M, SREBRO N. Fast maximum margin matrix factorization for collaborative prediction [C]//Proceedings of the 22nd International Conference on Machine Learning. [S.l.]: ACM, 2005: 713-719.
- [11] KHOSHNEESHIN M, STREET W N. Collaborative filtering via euclidean embedding[C]//Proceedings of the Fourth ACM Conference on Recommender Systems. [S.l.]: ACM, 2010: 87-94.
- [12] XUE G R, LIN C, YANG Q, et al. Scalable collaborative filtering using cluster-based smoothing [C]//Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. [S.l.]: ACM, 2005: 114-121.
- [13] YU Y, WANG C, GAO Y, et al. A coupled clustering approach for items recommendation[C]//Proceedings of the Pacific-Asia Conference on Knowledge Discovery and Data Mining. Berlin, Heidelberg: Springer, 2013: 365-376.
- [14] HOFMANN T. Collaborative filtering via Gaussian probabilistic latent semantic analysis[C]//Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. [S.l.]: ACM, 2003: 259-266.
- [15] HOFMANN T. Latent semantic models for collaborative filtering[J]. ACM Transactions on Information Systems (TOIS), 2004, 22(1): 89-115.
- [16] WANG Y, YU Y. Collaborative filtering via temporal Euclidean embedding[C]//Proceedings of Asia-Pacific Web Conference. Berlin, Heidelberg: Springer, 2012: 513-520.
- [17] LI W, GAO M, RONG W, et al. Social recommendation using Euclidean embedding[C]//Proceedings of 2017 International Joint Conference on Neural Networks (IJCNN). [S.l.]: IEEE, 2017: 589-595.
- [18] KINGMA D P, BA J. ADAM: A method for stochastic optimization [C]//Proceedings of the 3rd International Conference on Learning Representations. [S.l.]: [s.n.], 2015: 1-15.
- [19] GOLBECK J, HENDLER J. Filmtrust: Movie recommendations using trust in web-based social networks[C]//Proceedings of the 3rd IEEE Consumer Communications and Networking Conference. Las Vegas, USA: IEEE, 2006: 282-286.
- [20] SEDHAIN S, MENON A K, SANNER S, et al. Autorec: Autoencoders meet collaborative filtering [C]//Proceedings of the 24th International Conference on World Wide Web (WWW'15 Companion). New York, USA: ACM, 2015: 111-112.

(编辑:刘彦东)