

基于用户信任和评分偏置的正则化推荐模型

朱爱云,任晓军

(潍坊科技学院计算机软件学院,寿光 262700)

摘要:

在传统推荐系统中,特别是协同过滤推荐系统已得到广泛的研究,但是它们忽视了用户之间的社会关系,事实上,这些社会关系能够提高推荐系统的准确度。近年来,基于社会关系的推荐系统已经成为一个热门的研究方向。为了进一步提高系统推荐的准确性,提出一个社会正则化方法,即整合用户的直接朋友关系和评分偏置信息到推荐系统中构建正则化模型,然后采用一种随机梯度下降法对该矩阵进行分解,得到用户潜在特征和物品潜在特征。最后在真实数据集上的实验结果表明所提出的算法与现有相关算法比较具有较好的性能。

关键词:

协同过滤;矩阵分解;推荐系统;梯度下降

基金项目:

博士基金项目(No.2017BS19)

0 引言

随着网络和电子商务的快速增长,推荐系统根据用户购买或历史评分信息,能够快速自动地为用户提供有用的信息,传统上推荐系统已经解决了系统对物品评分预测问题。但是,传统的推荐技术仅仅使用用户-项目评分信息,为了实现用户更个性化的推荐结果,开发更智能化的推荐系统,随着社交网络的增长,整合社会网络信息到推荐系统中已引起许多学者的广泛关注。

传统的推荐系统^[1,2]总是忽视用户间的社会关系,在现实世界中,人们的社会关系在很大程度上决定了他们的偏好,事实上,当人们面临多种选择时,他们可能会通过最好的朋友提供建议,因此为了提供更准确、更个性化的推荐结果,整合用户间的社会关系到推荐系统中是合理的,基于以上观点,多种基于信任的推荐系统开始提出,并由此提出了多种信任推荐方法^[3-6],这些方法都是利用单边信任信息来进一步提高推荐系统的性能,这些方法中显而易见的弱点是单方面的“信任关系”问题。它不同于那种用户之间相互合作的“社会关系”。另外,其他弱点也有不可行的假设和弱泛化能

力,显然,基于信任的推荐系统也不太合适。此外,社会网络的整合从理论上是可以提高传统推荐系统的性能。因为就预测的准确性而言,朋友关系能够提高对用户评分的理解,并且朋友关系也表明在某些方面有共同之处,因此,能够缓解冷启动问题^[3]。

为了解决以上问题,本文主要关注基于朋友的社会推荐,类似于文献[8]提出的方法,在矩阵分解中增加用户、物品偏置信息以及用户的社会关系,构成正则化推荐模型。通过实验验证了增加用户、项目的评分偏置和用户的社会关系能提高推荐的准确度,并且也能应用于现实生活中的大规模数据集中。因为在现实生活中,人们经常在购买一种产品或消费一种服务前,借助于他们在社交网络中的朋友的建议。从社会学和心理学研究发现也表明,人们倾向于结交跟自己的兴趣相似的人,由于稳定和持久的社会绑定,人们更愿意与他们的朋友分享他们的个人观点,因此通常在陌生人、供应商和朋友中首先会选择他们朋友的推荐。

1 相关工作

传统推荐方法主要有三类:协同过滤^[2,9,10]、内容过

滤以及混合过滤^[1]。其中协同过滤是最普遍和最成功的方法。

通常,协同过滤推荐方法分为基于模型^[9]方法和基于内存^[2]方法,基于内存的方法从评分矩阵中查找与当前用户偏好最相似的用户,是一种启发式的评分预测。基于模型的方法使用评分集合来学习,然后用来进行评级预测。

传统的推荐方法已发展的很成熟,但是它们都是假设用户是独立的,没有考虑用户间的朋友关系,基于以上考虑,许多研究者提出了许多信任推荐方法^[4-7,11,14]并广泛应用于学术和工业领域。但是推荐进程仍与真实世界的推荐过程不一致,因此,他们提出了另一种集成方法,在同一时间,通过考虑用户的喜好和受信任的用户的喜好来计算用户的评分。实验表明,该方法是可行的,能够开发更好的推荐模型。基于信任的推荐系统已被证明是有效的,并取得了巨大的进步。然而,通过分析,它们也有几个固有的局限性和弱点需要解决。

近年来,在工业界和学术界如何利用社会网络信息已成为一个研究热点。通过结合社会网络信息能够影响个人在网上的行为,如用户间的互动、标签信息等能提高推荐系统的性能。文献[12]提出了一个融合社会朋友信息的个性化推荐概率模型,并在真实数据集上验证了用户与他的朋友在许多方面有相似的偏好。文献^[8]提出了一种基于概率矩阵分解的社会正则化和因子分解法,该方法通过利用网络信息提高了数据稀疏性问题和冷启动问题。文献[8,13]整合朋友关系来提高推荐系统的性能。并且很多社会推荐技术都是基于矩阵分解来解决评分预测问题^[8,13,16],矩阵分解方法是是目前比较成功的推荐方法,它将评分/购买矩阵分解成低维的用户矩阵和项目矩阵,用户和项目特征向量的点积说明了给定用户对项目的偏好程度。假定用户 u 对物品 i 的评分用 $r_{u,i}$ 表示,用户 u 对物品 i 的预测评分用 $\hat{r}_{u,i}$ 表示,其中 $\hat{r}_{u,i}$ 是由用户特征向量 p_u 和项目特征向量 q_i 的内积得到。

即:

$$\hat{r}_{u,i} = p_u q_i^T \quad (1)$$

但是,在实际的推荐系统中,有的用户往往热衷于给用户打高分,有的项目也给予了很高的评分,因此预测评分^[15]为:

$$\hat{r}_{u,i} = b_u + b_i + e + p_u q_i^T \quad (2)$$

其中, b_u 为用户 u 的偏置评分, b_i 为项目 i 的偏置评分, e 为数据集中所有评分的平均评分。

因此,目标函数为:

$$L(p, q) = \frac{1}{2} \min \sum (r_{u,i} - \hat{r}_{u,i})^2 + \frac{\lambda_1}{2} \|p_u\|^2 + \frac{\lambda_2}{2} \|q_i\|^2 + \frac{\lambda_2}{2} (b_u^2 + b_i^2) \quad (3)$$

2 融合社会关系和评分偏置的正则化推荐模型

随着社交媒体的日益普及,使得越来越多的在线用户参与在线活动,从而产生了更为丰富的社会关系。在社会推荐系统中,除了用户的评分信息外,还有用户之间的社会关系,社会关系的有效性为推荐系统提供了一个独立的资源,也为独特的社会推荐的属性带来了新的机遇。本文将结合用户的社会网络关系来提高推荐系统性能,假设用户有不同类型的社会关系(家人、朋友、同事、同学等),如果两个人建立一种社会关系,那么就说他们存在一种社会关系。一种社会关系可能对称也可能不对称。需要从以下方面定义这种关系。因此,在这一部分描述了本文提出的方法。

定义 1 假定 $U = \{u_1, u_2, u_3, \dots, u_n\}$ 是一个用户集, $I = \{i_1, i_2, i_3, \dots, i_m\}$ 是一个项目集,则社会评分网络 $SRN = \langle U, I, \phi, \varphi \rangle$ 是一个四元组,其中 $\varphi: U \times I \rightarrow R^+ \cup \{*\}$ 是一个评分函数,用一个真实的值关联着一个用户 $u_x \in U$ 和一个项目 $i_n \in I$,即用户 u 对项目 i 的评分为 $r_{u,i}$,否则用“*”表示。

$\varphi: U \times U \rightarrow \{0, 1\}$ 是一个社会函数,即一对用户 $u_x, u_y \in U$ 存在一种社会关系,则函数值为 1,否则为 0。这种社会关系对一些用户对 (u_x, u_y) 通常是不对称的,也就是 $\varphi(u_x, u_y) \neq \varphi(u_y, u_x)$ 。

定义 2 假定 $SRN = \langle U, I, \phi, \varphi \rangle$ 是一个社会评分网络,且用户 $u_x \in U$,用户 u_x 与他的邻居 N_x 存在一种社会关系 $N_x = \{u_y \in U : \varphi(u_x, u_y) = 1\}$ 。

在现实生活中,我们通常咨询我们熟悉的朋友,因为他们熟悉我们的品味,所以来自社会信息的熟悉度和相似度证据表明包含评分信息的社会信息能够潜在的提高推荐性能,因此,本文中在文献[9,14]方法基础上增加了用户的直接朋友关系作为正则化条件构建模型。猜想如果两个用户 u, w 是直接朋友关系,那么他们应该在特征空间中会映射成一个非常接近的点,换

言之假设有三个用户 u, w, x , 分别映射到特征空间中的点为 p_u, p_w, p_x , 但如果只有 u, w 是朋友, 那么 p_u, p_w 之间的距离可能要小于点 p_w 到点 u 的距离, 事实上, 如果一个用户 p_u 在潜在特征空间中很接近于他的直接朋友 p_w , 那么用户 p_u 的观点与他的朋友 p_w 的观点将会是相似的, 为此从数学的观点考虑在潜在的特征空间中用户 u, w 之间的距离为 $\|p_u - p_w\|$, 其中 $\|\cdot\|$ 是欧几里德距离的范式, 用 $N(u)$ 表示用户 u 的直接朋友的最近邻居, 我们的目的是使 $\sum_{w \in N(u)} \text{sim}(u, w) \|p_u - p_w\|^2$ 尽可能的最小, 因此公式(6)增加惩罚因子改为如下公式:

$$L(p, q) = \frac{1}{2} \min \sum_{u=1}^m \sum_{i=1}^n (r_{u,i} - \hat{r}_{u,i})^2 + \frac{\lambda_1}{2} \|p_u\|^2 + \frac{\lambda_2}{2} \|q_i\|^2 + \frac{\lambda_3}{2} (b_u^2 + b_i^2) + \frac{\beta}{2} \sum_{w \in N(u)} \text{sim}(u, w) \|p_u - p_w\|^2 \quad (4)$$

本文称之为融合社会关系和评分偏置信息的正则化方法(简称 Social B-SVD)。其中, $(\beta, \lambda_1, \lambda_2, \lambda_3)$ 均为 >0 的常数)用于调整过拟合, $\text{sim}(u, w)$ 表示用户 u 与他的直接朋友 w 的相似度, 我们用皮尔逊相关系数(PCC)即可求出相似度:

$$\text{sim}(u, w) = \frac{\sum_{i=1}^n (r_{ui} - \bar{r}_u)(r_{wi} - \bar{r}_w)}{\sqrt{\sum_{i \in I(u) \cap I(w)} (r_{ui} - \bar{r}_u)^2} \cdot \sqrt{\sum_{i \in I(u) \cap I(w)} (r_{wi} - \bar{r}_w)^2}} \quad (5)$$

其中, 相似度 $\text{sim}(u, w)$ 值越大, 表明特征向量 p_u, p_w 之间的距离越小, 也就表明他们之间有更加相似的偏好, 反之, 相似度越小, 表明特征向量之间的距离越大。其中 \bar{r}_u 是用户 u 的平均评分, 从这个相似度公式中得到 $\text{sim}(u, w) \in [-1, 1]$, 为了约束它的范围 $[0, 1]$, 采用一个映射函数 $f(x) = (x+1)/2$ 。

为了解决这个最优化问题, 首先需要对目标函数中的参数 p_{uk}, q_{ik}, b_u, b_i 分别求偏导。

$$\frac{\partial L}{\partial p_{uk}} = -e_{ui} q_{ik} + \lambda_1 p_{uk} + \beta \sum_{w \in N(u)} \text{sim}(u, w) (p_u - p_w) \quad (6)$$

$$\frac{\partial L}{\partial q_{ik}} = -e_{ui} p_{uk} + \lambda_2 q_{ik} \quad (7)$$

$$\frac{\partial L}{\partial b_u} = -e_{ui} + \lambda_3 b_u \quad (8)$$

$$\frac{\partial L}{\partial b_i} = -e_{ui} + \lambda_3 b_i \quad (9)$$

然后利用随机梯度下降法, 沿最速下降方向递推得到如下公式:

$$p_{uk} = p_{uk} + \alpha(e_{ui} q_{ik} - \lambda_1 p_{uk} - \beta \sum_{w \in N(u)} \text{sim}(u, w) (p_{uk} - p_{wk})) \quad (10)$$

$$q_{ik} = q_{ik} + \alpha(e_{ui} p_{uk} - \lambda_2 q_{ik}) \quad (11)$$

$$b_u = b_u + \alpha(e_{ui} - \lambda_3 b_u) \quad (12)$$

$$b_i = b_i + \alpha(e_{ui} - \lambda_3 b_i) \quad (13)$$

其中, α 为学习速率, 并按每次迭代缩减为 0.9 倍的速度递减。

下面是具体的算法:

算法: 融合社会关系的矩阵分解(SocialB-SVD), 用户对项目的分 r_{ui} , 需要分解的特征维数 k

输入: 用户数 m , 项目数 n , 迭代次数为 T 。

输出: 潜在用户特征矩阵 U 和潜在项目特征矩阵 V 。

开始

随机初始化: 用户特征矩阵 U , 项目特征矩阵 V , 每个用户 u 的朋友特征矩阵 p_w , 每个用户的偏置向量 b_u 和每个项目 i 的偏置向量 b_i 。

for iter=1,2,...,T do //迭代次数从 1 到 T

for $u=1, \dots, m$ do //用户数从 1 到 m

for $i=1, \dots, n$ do //项目数从 1 到 n

$$e_{ui} = r_{ui} - \hat{r}_{ui} = r_{ui} - b_u - b_i - p_u q_i^T$$

//计算实际评分与预测评分之间的误差

$$b_u = b_u + \alpha(e_{ui} - \lambda_3 b_u) // \text{梯度更新用户 } u \text{ 的}$$

偏置评分 b_u

$$b_i = b_i + \alpha(e_{ui} - \lambda_3 b_i) // \text{梯度更新项目 } i \text{ 的}$$

偏置评分 b_i

for $k=1$ to K do //需要分解的特征维数

从 1 到 K

根据 PCC 得到用户 u 的最相似朋

友 w

$$p_{uk} = p_{uk} + \alpha(e_{ui} q_{ik} - \lambda_1 p_{uk} - \beta$$

$$\sum_{w \in N(u)} \text{sim}(u, w) (p_{uk} - p_{wk}))$$

//梯度更新用户 u 的特征向量 p_{uk}

$$q_{ik} = q_{ik} + \alpha(e_{ui} p_{uk} - \lambda_2 q_{ik})$$

//梯度更新项目 i 的特征向量 q_{ik}

End for

End for

End for

End for

End

3 实验结果

3.1 数据集和结果分析

在本文中,我们采用 Flixster 数据集作为此实验数据集,此数据集包含了用户间的社会网络和用户评分,用户间的社会网络是无向的,评分值介于[0.5,5.0]之间,且步长为 0.5。

在实验中,我们从数据集中随机抽取了 80%的评分数据作为训练集,剩余 20%作为测试集,并且从社交网络关系数据中抽取了 20000 个朋友作为训练集,学习速率参数 α 与正则化参数 $\lambda_1, \lambda_2, \lambda_3, \beta$ 通过交叉验证决定。本文采用 $\alpha=0.0002$, $\lambda_1=0.003$, $\lambda_2=0.002$, $\lambda_3=0.004$, $\beta=0.2$ 进行实验。并且采用均方根误差(RMSE)和平均绝对误差(MAE)来评价预测准确度。

取特征维数为 $k=10$,和 $k=30$,迭代次数为 30 时对比了以下这 5 种算法的预测准确度。

(1)ItemMean:此方法使用每个项目的平均值来预测缺失值。

(2)SVD:是最传统的矩阵分解推荐算法,已广泛应用于推荐系统中,但它忽视了用户间的社会关系。

(3)Bias_SVD:是 Koren^[16]提出的一种推荐方法,该方法使用户的偏置信息、项目的偏置信息整合到推荐系统中,提高了推荐性能。

(4)Social_SVD:是 Ma^[8]提出的一种信任感知推荐方法,利用矩阵分解融合用户和她朋友的品味构建正则化模型。

(5)SVD++:这种方法是 Koren^[15]提出的,他考虑了用户和项目的偏见值对评分的影响,也融合了用户评分的显式和隐式影响。

本实验中,特征维数设置为 $k=10$ 和 30,针对所有用户做了实验对比,实验结果如表 1 所示,不论 $k=10$, 30,与表中其他的方法相比较,我们所提出的 Social B-SVD 方法的性能是最好的(MAE 或 RMSE 值最小),虽然相对提高的比例很小,但是也表明融合社会网络信息的 Social B-SVD 方法会大大提高推荐系统的性能。

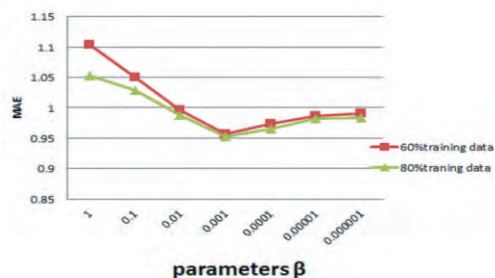


图1 参数 β 的影响($k=10$)

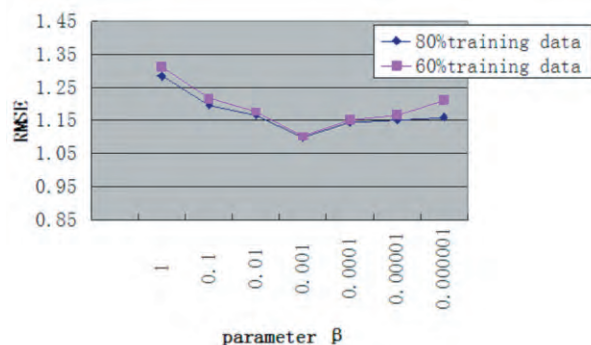


图2 参数 β 的影响($k=10$)

3.2 参数 β 的影响

在本文中参数 β 对预测准确度起着重要的作用,它表示到底应该结合多少社交网络信息才能达到最佳状态。在极端的情况下,如果我们用一个很小的 β 值,它表明仅仅使用用户自己的品味来做出推荐,相反,我们如果用一个很大的 β 值,那么社会网络信息在学习过程中就占支配地位。但在许多情况下,我们都不想设置这些极端值,从图 1,图 2 中我们看到不管用哪种数据集,随着 β 值的增加 MAE/RMSE 值开始在减少,但当达到某一个阈值(0.001 在 Flixster 数据集上)时,随着 β 值的继续增加 MAE/RMSE 又开始增加,这个拐点的存在说明融合社交网络信息到矩阵分解中进行推荐远远好于单纯使用用户项目评分或单纯使用社交网络信息。

表 1 Flixster 数据集中预测准确度的比较

k	度量	ItemMean	SVD	Bias-SVD	Social-SVD	SVD++	SocialB-SVD
10	RMSE	1.6564	1.48831	1.2332	1.1192	1.1128	1.1099
10	MAE	1.4232	1.2791	1.2791	0.9679	0.9567	0.9510
30	RMSE	1.5117	1.4146	1.4146	1.1056	1.0911	1.0360
30	MAE	1.3689	1.2120	1.2120	0.9308	0.9247	0.9179

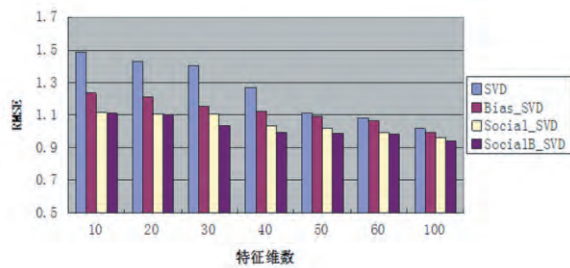


图3 特征维数 k 对 RMSE 的影响

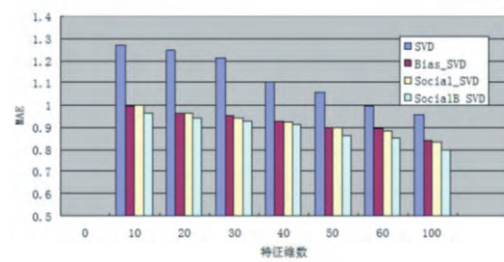


图4 特征维数 k 对 MAE 的影响

3.3 特征维数 k 的影响

在本文中通过实验验证特征维数 k 对预测准确度影响较大,我们选择 $k=10,20,30,40,50,60,100$ 进行模型训练。从图3、图4可以看到,在我们提出的方法中,随着特征维数 k 值的增加,MAE/RMSE 值开始下降很快,但是随着 k 值的增加 MAE/RMSE 下降速度变慢(由图3、图4得到 k 值从60到100,MAE 值从0.8451下降到0.7906,RMSE 值从0.9842下降到0.9439),也从侧面验证了算法每次计算出的都是最显著的特征向量。

4 结语

本文在传统的奇异值(SVD)矩阵分解模型中融合了社交网络中的直接朋友关系和用户评分偏置信息作为辅助信息,设计了基于社交网络信息的矩阵分解算法。实验表明,本文提出的算法具有较好的预测效果,其性能明显优于相比较的相关算法。

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. IEEE Trans.Knowl. Data Eng.,2005,17(6):734-749
- [2]Bellogin, A., Castells, P., Cantador, I.. Improving Memory-Based Collaborative Filtering by Neighbour Selection Based on User Preference Overlap. In:Proceedings of the 10th Conference on Open Research Areas in Information Retrieval, Lisbon, Portugal, 2013, May 15-17:145-148.
- [3]Jamali, M., Ester, M.. A Matrix Factorization Technique with Trust Propagation for Recommendation in Social Networks. In: Proceedings of the 4th ACM Conference on Recommender Systems, Barcelona, Spain, 2010, September 26-30:135-142.
- [4]Ozsoy, M.G., Polat, F.. Trust Based Recommendation Systems. In: Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM'13), Niagara, ON, Canada, 2013, August 25-29:1267-1274.
- [5]Massa, P., Avesani, P.. Trust metrics in recommender systems. Comput. Soc.Trust, 2009:259-285.
- [6]Nazemian, A., Gholami, H., Taghiyareh, F.. An Improved Model of Trustaware Recommender Systems Using Distrust Metric. In: Proceedings of the IEEE/ACM International Conference on Advances in Social,2012.
- [7]Ma, H., Yang, H., Lyu, M.R., King, I.. SoRec: Social Recommendation Using Probabilistic Matrix Factorization. In: Proceedings of the 17th ACM Conference on Information and Knowledge Management, Napa Valley, California, USA, 2008, October 26-30:931-940.
- [8]Ma H,Zhou D Y,Liu C. Recommender Systems Withsocial Regularization[C]. In Proceedings of the 4th ACM International Conference on Web Search and Data Mining,2011,287-296.
- [9]Bergner, Y., Droschler, S., Kortemeyer, G., et al.. Model-based Collaborative Filtering Analysis of Student Response Data: Machine-Learning Item Response In: Proceedings of the 5th International Conference on Educational Data Mining, Chania, Greece,, 2012, June 19-21:95-102.

- [10]Gunes, I., Bilge, A., Polat, H.. Shilling Attacks Against Memory-based Privacy-Preserving Recommendation Algorithms. Internet Inf. Syst, 2013,7(5):1272-1290.
- [11]Nazemian, A., Gholami, H., Taghiyareh, F.. An Improved Model of Trustaware Recommender Systems Using Distrust Metric. In: Proceedings of the 2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, Istanbul, Turkey, 2012, August 26-29:1079-1084.
- [12]He, J., Chu, W., PhD Dissertation. A Social Network-Based Recommender System (SNRS). University of California at Los Angeles, CA, USA,2010.
- [13]Wang, X., Huang, W.. Research on Social Regularization-Based Recommender Algorithm. Comput. Mod.,2014,1:77-80.
- [14]王瑞琴,蒋云良,李一啸. 一种基于多元社交信任的协同过滤推荐算法[J].计算机研究与发展,2016,53(6):1389-1399
- [15]Koren Y,R BeII, C Volinsky. Matrix Factorization Techniques for Recommender Systems[J]. Compute Society,2009,42(8):30-37.
- [16]Koren Y. Factor in the neighbors Scalable and Accurate Collaborative Filtering[J]. ACM Transactions on Knowledge Discovery from Data(TKDD),2010,1(4):1-11.

作者简介:

朱爱云(1977-),女,硕士,讲师,研究方向为数据挖掘、推荐系统

任晓军(1984-),男,博士,讲师,研究方向为无线传感网络、大数据分析、数据语义识别

收稿日期:2018-01-30

修稿日期:2018-04-20

Recommendation Model Regularized with User Trust and Ratings Bias

ZHU Ai-yun, REN Xiao-jun

(School of Computer and Software Engineering, Weifang University of Science and Technology, Shouguang 262700)

Abstract:

The traditional recommender systems, especially the collaborative filtering recommender systems, have been studied by many researchers in the past decade. However, they ignore the social relationships among users. In fact, these relationships can improve the accuracy of recommendation. In recent years, the study of social-based recommender systems has become an active research topic. Proposes the social regularization approach that incorporates social network information to benefit recommender, the proposed method merges social friendships, the users rating bias and items rating bias into recommender systems, designs a Stochastic Gradient Descent (SGD) algorithm to obtain potential user feature matrix and item feature matrix. Experimental results show that the proposed algorithm has good prediction accuracy, which is better than the existing algorithms.

Keywords:

Matrix Factorization; Recommender System; Gradient Descent; Social Relationship