

协同过滤算法前沿论文最新进展

2018.11.02 方建勇

1. [arXiv: 1810.08765](#)[pdf,其他] Cs. 红外

属性感知协作筛选: 调查和分类

作者:文浩陈,志驰,彝-荔,文森特,美日元叶,寿德林

摘要: 属性感知 CF 模型不仅针对用户对项目的历史评级, 还针对与用户相关的信息 (如年龄)、项目 (例如价格), 甚至评级 (例如评级时间), 以评级预测为目标。本文的研究工作在过去十年开发的属性感知 CF 系统, 并发现数学上, 他们可以分为四个不同的类别。我们不仅为读者提供了对这一领域现有工程的高层次数学解释, 而且为每类模型的数学见解。最后, 对各类主要工程的有效性进行了深入的实验研究。更少

提交 20 2018 年 10 月;最初宣布 2018 年 10 月。

2. [arXiv: 1810.08189](#)[pdf,其他] Cs. 简历

基于视频推荐系统的卷积协同滤波器网络

作者:郑康谢,米格尔,卢森堡 Taliyan,马特 Nickens, Mitkumar 潘迪亚, JJ 埃斯皮诺萨

摘要: 本分析探讨了电影预告片中物体的时序排序问题。电影预告片中对象的时序排序 (例如, 一个物体的长镜头与间歇性的短镜头) 可以传达有关电影类型、电影情节、主角角色以及电影制作者电影选择的信息。当与历史客户数据结合使用时, 排序分析可用于改进对客户行为的预测。例如, 客户购买了一部新电影的门票, 也许客户以前看过包含类似序列的电影。为了探索电影预告片中的对象排序, 我们提出了一个视频卷积网络, 用于捕获预测客户偏好的动作和场景。该模型了解不同类型对象 (例如, 汽车 vs 面) 的序列的特定性质, 以及序列在预测客户未来行为中的作用。我们展示了这样一个时间感知模型是如何优于我们以前的作品中提出的简单要素池方法, 重要的是展示了这种模型所允许的附加模型解释能力。更少

提交 22 2018 年 10 月;v1 提交 18 2018 年 10 月;最初宣布 2018 年 10 月。

评论:8 页, 3 数字, 1 表包括消融研究. 参数/结果不变

3. [arXiv: 1810.05376](#)[pdf,其他] Cs. 红外

神经变分混合协同滤波

作者:腾晓,上宋,洪沈, Zaiqiao 孟

摘要: 协同滤波 (CF) 是推荐系统最常用的方法之一。由于贝叶斯的性质和非线性, 深生成模型, 如变分自动编码器 (VAE), 已被应用到 CF 任务, 并取得了很大的性能。然而, 大多数基于 VAE 的方法都受到矩阵稀疏性的影响, 考虑到用户的潜在因素是相同的, 这会导致用户和项目的潜在表示不佳。此外, 大多数现有方法仅对用户的潜在因素进行建模, 而不是项目, 这使得它们无法向新用户推荐项目。为了解决这些问题, 我们提出了一种神经变分混合协同滤波 VDMF {}。具体来说, 我们考虑了用户和项目的生成过程, 以及用户和项目的潜在因素是 \ 公众 {侧 ~ 信息特定}, 这使我们的模型能够缓解矩阵稀疏性, 并学习更好的潜在表示用户和项目。为了推理的目的, 我们推导了随机梯度变分贝叶斯 (SGVB) 算法, 对用户和项目潜在因素的棘手分布进行了近似

分析。对两个大型数据集进行的实验表明, 我们的方法显著优于先进的 CF 方法, 包括基于 VAE 的方法。更少

提交 26 2018 年 10 月;v1 提交 12 2018 年 10 月;最初宣布 2018 年 10 月。

评论:8 页, 5 数字

4. [arXiv: 1809.06120](#)[pdf,其他] Cs. 红外

cf2vec: 基于图分布表示的协同滤波算法选择

作者:蒂亚戈,卡洛斯苏亚雷斯,安德烈 c.p.l. 卡瓦略

摘要: 使用元学习的算法选择旨在查找问题特征 (即 metafeatures) 与相对算法性能之间的映射, 以预测新数据集的最佳算法。因此, 最重要的是 metafeatures 所使用的信息。在协作筛选中, 最近的研究已经创建了大量的这种 metafeatures 的集合。然而, 由于这些是基于从业者对问题的理解而创建的, 因此他们可能无法捕捉到正确描述问题所必需的最相关方面。我们建议通过利用制图表达学习来克服这一问题, 通过让数据指导制图表达的设计而不是实践者的意见, 能够创造出一种另类的问题表征。我们的假说指出, 这种替代性表述可以用来取代标准 metafeatures, 从而导致对元学习采取更加稳健的做法。针对协同滤波算法的选择, 提出了一种新的设计方法。该过程将协作筛选作为图形进行建模, 并使用 graph2vec 提取分布式表示。实验结果表明, 所提出的程序创建了与最先进的 metafeatures 具有竞争力的制图表达, 同时需要的数据少得多, 而且几乎没有任何人工输入。更少

提交 17 2018 年 9 月;最初宣布 2018 年 9 月。

5. [arXiv: 1809.05662](#)[pdf,其他] Cs. 红外

用于协作筛选的沃瑟斯坦自动编码器

作者:郭景彬中,张凤

摘要: 本文对推荐系统进行了长期的研究。最近, 用户的隐含反馈, 如 "点击" 或 "浏览" 被认为能够提高推荐性能。因此, 已进行了一些尝试来解决此问题。其中, 变分自动编码器 (VAE) 方法已经达到了优异的性能。然而, 被编码的潜在变量的分布重叠了很多, 可能会限制其推荐能力。为了应对这一挑战, 本文试图扩展沃瑟斯坦自动编码器 (挫伤) 进行协同过滤。特别是, 通过引入两个额外的损失条件, 重新设计了适应挫伤的损失函数: (1) 潜在变量分布与假定地真值分布之间的相互信息丢失, 以及 (2) L1 正则化损失引入来限制编码的潜在变量是稀疏的。两种不同的成本函数用于测量隐式反馈数据与其重新生成的数据版本之间的距离。实验计价了三广泛采用的数据集, 即 ML-20M、Netflix 和 LASTFM。比较基准和最先进的方法是选择的性能比较, 多 DAE, 多 VAE, CDAE 和苗条。所提方法的性能优于 Recall@1、Recall@5 和 NDCG@10 评价标准的比较方法, 这表明了所提出方法的有效性。更少

提交 19 2018 年 9 月;v1 提交 15 2018 年 9 月;最初宣布 2018 年 9 月。

6. [arXiv: 1809.00999](#)[pdf, ps,其他] Cs. 红外

面向协同过滤的自动编码器大规模训练

作者:阿卜杜拉穆萨维

摘要: 本文应用一种基于小批量的负抽样方法, 有效地培养了一种用于隐式反馈协同滤波的大规模稀疏数据的潜因子自动编码器模型。我们将我们的工作与不同实验数据集上最先进的基准模型进行比较, 并表明此方法可以使基线模型性能的良好和快速逼近。源代码在 <https://github.com/amoussawi/recoder> 中可用。更少

提交 23 2018 年 10 月;v1 提交 30 2018 年 8 月;最初宣布 2018 年 9 月。

评论:2 页, ACM RecSys 2018 晚期结果跟踪 (海报)

7. [arXiv: 1808.10523](#)[pdf, ps,其他] Cs. 红外

[doi:10.1145/3240323.3240343](#)

频谱协同过滤

作者:雷锋,春塔路,飞江, 加味张,菲利普 s. Yu

摘要: 尽管协作过滤 (cf) 的普及, 基于 CF 的方法被 textit {冷启动} 问题困扰, 这对用户使用推荐系统 (RS) 的体验产生了显著的负面影响。为了克服上述缺点, 首先提出了用户与项目之间的关系。然后, 我们提出了一种新的谱卷积运算, 直接在 textit {频谱域} 中执行, 其中不仅显示了图的接近信息, 而且还揭示了图中隐藏的连通性信息。利用所提出的频谱卷积运算, 建立了一种称为频谱协同滤波 (SpectralCF) 的深度推荐模型。得益于 textit {频谱域} 中存在的丰富的连接信息, SpectralCF 能够发现用户和项目之间的深层连接, 因此缓解了 CF 的 textit {冷启动} 问题。我们最好的知识, SpectralCF 是第一个基于 CF 的方法, 直接从用户项二部图的 textit {频谱域} 中学习。我们在几个标准数据集上应用我们的方法。结果表明, SpectralCF 显著优于最先进的模型。代码和数据可在 \url {https://github.com/lzheng21/SpectralCF}。更少
提交 30 2018 年 8 月;最初宣布 2018 年 8 月。

评论:RecSys2018

8. [arXiv: 1808.09785](#)[pdf,其他] Cs. 红外

使用味觉组进行协作筛选

作者:Farhan 哈瓦尔,内文. 张

摘要: 隐式反馈是用户反馈的最简单形式, 可用于项目建议。它很容易收集和域独立。然而, 缺乏消极的例子。现有的工程通过对未使用项进行各种假设来规避这个问题, 当用户没有使用某个项目, 因为她不知道它时, 就无法容纳这些项目。本文提出了一种新的方法来解决隐性反馈中的否定例子。其动机是, 如果有一大群用户共享相同的味道, 他们都没有消耗一个项目, 那么很可能该项目是不相关的这种味道。我们使用分层潜伏树分析 (HLTA) 来识别基于味觉的用户组, 并根据她在组中的成员身份为用户提出建议。更少

提交 28 2018 年 8 月;最初宣布 2018 年 8 月。

评论:RecSys 2018 LBRS. arXiv 管理说明: 大量文本与 arXiv 重叠: 1704.01889

9. [arXiv: 1808.04288](#)[pdf,其他] Cs. 红外

通过协作过滤器组成的自动播放列表继续

作者:艾琳 Teinemaa,尼克税,卡洛斯本特斯

摘要: RecSys 挑战 2018 专注于自动播放列表, 即, 任务是根据播放列表的标题和/或它已经包含的曲目的子集, 为播放列表推荐额外的音乐曲目。挑战基于 Spotify 百万播放列表数据集 (MPD), 其中包含 100 万真实播放列表中的曲目和元数据。本文介绍了拿铁的自动播放列表继续解决方案, 这是基于协作筛选器的组合, 每个捕获播放列表的不同方面, 这些协作筛选器的最佳组合是使用树结构 Parzen 估计器 (TPE) 确定。该解决方案在最终排行榜中获得了 112 个参与团队的第十二名。拿铁车队参加了 RecSys 挑战 2018 的主要轨道。更少

提交 13 2018 年 8 月;最初宣布 2018 年 8 月。

10. [arXiv: 1808.03912](#)[pdf,其他] Cs. 红外

基于外部产品的神经协同滤波

作者:[湘南](#)、[小雨杜](#)、[祥王](#)、[丰田](#)、[金辉塘](#)、[锦盛蔡](#)

摘要: 在这项工作中, 我们贡献了一个新的多层神经网络架构, 名为 ONCF 来执行协作筛选。其目的是使用外部产品显式地模拟嵌入空间维度之间的成对相关性。与现有的神经推荐器模型相比较, 通过简单的串联或元素的产品结合用户嵌入和项目嵌入, 我们建议在嵌入层上方使用外部产品会导致二维交互映射, 即更具表现力和语义上的合理性。在外部产品所获得的交互映射之上, 我们建议采用卷积神经网络来学习嵌入维度之间的高阶相关性。对两个公共隐式反馈数据的广泛实验表明了我们建议的 ONCF 框架的有效性, 特别是使用外部产品来模拟低级别嵌入维度之间相关性的积极效果。多层神经推荐模型。实验代码可在: <https://github.com/duxy-me/ConvNCF> 少

提交 12 2018 年 8 月;最初宣布 2018 年 8 月。

评论:IJCAI 2018

11. [arXiv: 1808.03298](#)[pdf,其他] Cs. 红外

协同滤波器的概率集合

作者:[蛭英](#), [大华林](#)

摘要: 协同过滤是推荐的重要技术。虽然它在以前的工作中一再证明是有效的, 但在许多实际应用中, 其性能仍然不理想, 尤其是那些项目或用户高度多样化的应用程序。在本文中, 我们将探讨一个基于集合的框架, 以提高推荐处理不同数据的能力。具体来说, 我们制定了一个概率模型, 将项目、用户以及它们之间的关联整合成一个生成过程。在此公式的顶部, 我们进一步推导出一个渐进的算法来构建协同滤波器的集合。在每次迭代中, 一个新的筛选器派生自重加权条目并合并到集成中。值得注意的是, 虽然我们的算法的算法程序是明显类似于提高, 它是从本质上不同的公式派生, 因此不同于几个关键技术方面。我们在三个大型数据集上测试了所提出的方法, 并观察到了大量的改进, 包括 L2Boost, 这是一种基于提升的有效方法。更少

提交 14 2018 年 8 月;v1 提交 26 2018 年 6 月;最初宣布 2018 年 8 月。

评论:8 页。在 AAAI-2018 的诉讼中

12. [arXiv: 1808.03233](#)[pdf,其他] Cs. Lg

双簧管: 用于 AutoML 初始化的协作过滤

作者:[Chengrun](#), [本健](#), [大获金](#), [马德琳尤戴尔](#)

摘要: 算法选择和超参数优化仍然是机器学习中最具挑战性的任务之一。机器学习应用程序的数量比机器学习专家的数量快得多, 因此我们看到对学习过程的高效自动化的需求越来越大。在这里, 我们介绍了双簧管, 一种时间约束模型选择和超参数调谐算法。利用数据集之间的相似性, 双簧管通过协同过滤找到了有前途的算法和超参数配置。我们的系统在时间约束下对这些模型进行了探索, 以便快速初始化可以提供热启动的更细粒度优化方法。本文研究的一个新方面是基于优化实验设计的时间约束矩阵完成中主动学习的一种启发式方法。我们的实验表明, 双簧管提供的最先进的性能比竞争的方法在测试床上的监督学习问题更快。更少

提交 9 2018 年 8 月;最初宣布 2018 年 8 月。

13. [arXiv: 1808.01006](#)[pdf,其他] Cs. 红外

协同滤波的混合变分自动编码器

作者: Kiloi 古普塔, 阿联酋 Yelahanka Raghuprasad, Pankhuri

摘要: 在当今的时代, 几乎每个行业都有在线市场的用户, 个性化的建议变得非常重要。传统上, 利用线性化的矩阵分解方法解决了协同滤波问题。通过提出一种混合的多模态方法, 我们将 [11] 的工作扩展到使用变分自动编码器 (放瓦伊斯·多斯拉克) 与隐式反馈进行协同滤波。我们的方法将电影嵌入 (从同级 VAE 网络学习) 与 Movielens 20M 数据集的用户评分相结合, 并将其应用于电影推荐任务。我们通过整合电影嵌入的经验证明了 VAE 网络是如何获得授权的。我们还通过对从 VAE 获得的潜在表示进行聚类来可视化电影和用户嵌入。更少

提交 23 2018 年 9 月; v1 提交 14 2018 年 7 月; 最初宣布 2018 年 8 月。

14. [arXiv: 1807.09097](#) [pdf, 其他] Cs. 红外

协同滤波算法的选择: 图 metafeatures 和多准则 metatargets 的影响

作者: 蒂亚戈, 卡洛斯苏亚雷斯, 安德烈 c.p.l. 卡瓦略

摘要: 为 为 新问题选择最佳算法是一项昂贵而艰巨的任务。然而, 有解决这个问题的自动方案: 使用元学习, 利用问题特征 (即 metafeatures), 一个能够预测算法的相对性能。在协作过滤范围内, 最近的作品提出了不同的 metafeatures 描述这个问题的几个方面。尽管有有趣和有效的发现, 但这是否是最有效的 metafeatures 尚不得而知。因此, 本文提出了一套新的图 metafeatures, 从图论的角度探讨了协同滤波问题。此外, 为了了解多维 metafeatures 是否更适合, 我们研究了综合 metafeatures 的影响。这些 metafeatures 是从所有现有协作筛选 metafeatures 中选择的最佳 metafeatures。在受控实验装置中, 研究了最具代表性的 metafeatures 的影响。我们目前的另一项贡献是使用帕累托高效的排名程序来创建多准则 metatargets。这些新的算法排名, 考虑到多种评估措施, 允许以更公平和更详细的方式探索算法选择问题。根据实验结果, 图 metafeatures 是 metafeatures 相关工作的较好替代方法。但是, 结果表明, 用于创建综合 metafeatures 的特征选择过程是不有效的, 因为预测性能没有增加。最后, 进行了广泛的元知识分析, 以确定最具影响力的 metafeatures。更少

提交 23 2018 年 7 月; 最初宣布 2018 年 7 月。

15. [arXiv: 1807.06839](#) [pdf, 其他] cs SI

基于信任的协同过滤: 用正则等价处理冷启动问题

作者: 托米斯拉夫 Duricic, 伊曼纽尔 Ladic, 多米尼克 Kowald, 伊丽莎白 Lex

摘要: 基于用户的协作过滤 (CF) 是创建推荐系统最常用的方法之一。此方法基于查找最相关的 k 用户从其评级历史我们可以提取项目推荐。但是, CF 由于数据稀疏性和冷启动问题而受到影响, 因为用户通常只对可用项目的一小部分进行评价。一种解决方案是将附加信息纳入推荐过程, 例如用户分配给他人的显式信任分数或用户之间的社交连接导致的隐式信任关系。这种关系通常形成一个非常稀疏的信任网络, 可用于根据用户信任的人生成建议。在我们的工作中, 我们探索使用网络科学 (即常规等价) 的度量值, 应用于信任网络来生成相似性矩阵, 用于选择推荐项的 k 最近邻域。我们对 Epinions 的方法进行评估, 我们发现在推荐准确性方面, 我们可以优于处理冷启动用户的相关方法。更少

提交 18 2018 年 7 月; 最初宣布 2018 年 7 月。

16. [arXiv: 1807.05853](#) [pdf, ps, 其他] Cs. 红外

一种基于多数据源的分布式协同滤波算法

作者:穆罕默德商会 Bouadjene,埃丝特帕奇蒂,马克西米连·罗伯 Servajean,佛罗伦 Massegia, Amr El Abbadi

摘要: 协作过滤 (CF) 是最常用的推荐方法之一。CF 包括通过利用用户首选项以及其他用户的知识来预测用户是否喜欢 (或不喜欢) 某个项目。在实践中, 用户仅在一小部分项目上进行交互和表达意见, 这使得相应的用户项评级矩阵非常稀疏。这种数据稀疏性会给推荐系统带来两个主要问题: (1) 缺乏有效地模拟用户偏好的数据, 以及 (2) 缺乏有效地模拟项目特征的数据。但是, 通常有许多其他数据源可供推荐程序系统提供商使用, 可以描述用户兴趣和项目特征 (例如用户的社交网络、与项目相关的标签等)。这些有价值的信息可能提供有用的信息, 以加强推荐系统, 以更准确地建模用户的喜好和项目特征, 从而希望使推荐人更精确。由于各种原因, 这些数据源可能由不同数据中心的集群管理, 因此需要开发分布式解决方案。本文提出了一种新的分布式协同滤波算法, 它利用并结合多种不同的数据源来提高推荐质量。我们使用真实数据集进行的实验评估显示了算法与最先进的推荐算法相比的有效性。更少

提交 16 2018 年 7 月;最初宣布 2018 年 7 月。

评论:第十数据库、知识和数据应用进展国际会议, DBKDA 2018 2018 年 5 月 20 日至 2018 年 5 月 24 日-尼斯, 法国

17. [arXiv: 1806.06875](#)[pdf,其他] Cs. Lg

[doi:10.1145/2792838.2800192](#)

通过协作筛选的审查来学习分布式制图表达

作者:阿姆贾德 Almahairi,凯尔卡斯特纳, Kyunghyun 町,亚伦库维尔

摘要: 最近的工作表明, 基于协作过滤器的推荐系统可以通过将诸如自然语言评论等侧信息作为规范化派生产品表示形式的方式来改进。通过这种方法的成功, 我们引入了两种不同的评价模型, 并研究了它们对协同过滤性能的影响。虽然以前的最先进的方法是基于一个潜在的不分配 (LDA) 模型的审查, 我们探索的模型是基于神经网络: 一个包词的专家模型和一个递归神经网络。我们证明, 专家产品模型提供的灵活性增强, 使其能够在 Amazon 审阅数据集上实现最先进的性能, 优于基于 LDA 的方法。然而, 有趣的是, 周期性神经网络提供的更大的建模能力似乎破坏了模型作为产品表示的拉普拉斯的能力。更少

提交 18 2018 年 6 月;最初宣布 2018 年 6 月。

评论:在 RecSys 2015 会议上发表

18. [arXiv: 1806.06446](#) Cs. 红外

自注意神经协同滤波

作者:一泰,帅张, Luu,兆祥惠

摘要: 本文已被撤回, 因为我们在我们的 tensorflow 实现中发现了一个错误, 涉及跨批次的矢量混合。这导致不同的推断结果给出不同的批次大小, 这是完全陌生的。业绩得分仍然保持不变, 但我们得出结论认为, 这不是对业绩作出贡献的自我关注。我们之所以撤回这份文件, 是因为这份文件的主要说法是错误的。感谢新加坡国立大学的关鑫宇在我们以前的开源代码中发现这个问题。更少

提交 19 2018 年 7 月,v1 提交 17 2018 年 6 月;最初宣布 2018 年 6 月。

评论:我们在我们的 tensorflow 实现中发现了一个 bug, 其中涉及跨批次的矢量混合, 从而导致纸张的主要声明不正确。我们正在撤回这份文件, 直到找出原因

19. [arXiv: 1806.06192](#)[pdf,其他] Cs。红外

利用强化学习处理冷启动协同过滤

作者:[喜玛朝龙 Dureddy](#), [Kaden](#)

抽象: 推荐系统的一个主要挑战是处理新用户, 他们也被称为冷启动用户。本文针对电影推荐系统, 提出了一种新的学习方法, 用于访谈冷启动用户的一系列问题。我们建议使用深度 Q 网络学习面试问题, 创建用户配置文件, 以便对冷启动用户提出更好的建议。虽然我们的建议系统是使用电影推荐系统进行培训的, 但我们的深度 Q 网络模型应该在各种类型的推荐器系统中进行推广。更少
提交 16 2018 年 6 月;最初宣布 2018 年 6 月。

20. [arXiv: 1805.06563](#)[pdf,其他] Cs。红外

用于协作过滤的神经个性化嵌入

作者:[ThaiBinh 阮](#), [曝露 Takasu](#)

摘要: 矩阵分解是推荐系统中最有效的方法之一。但是, 这种算法依赖于用户和项目之间的交互, 对 "冷用户" (具有此类交互的历史记录很少的用户) 和捕获紧密相关项目之间的关系表现不佳。针对这些问题, 提出了一种神经个性化嵌入 (NPE) 模型, 提高了冷用户的推荐性能, 并可以学习项目的有效表示。它以两个术语对用户的单击进行建模: 用户对项目的个人喜好, 以及此项目与用户单击的其他项目之间的关系。我们表明, NPE 优于竞争方法的顶级建议, 特别为冷用户建议。我们还进行了定性分析, 显示模型所学到的制图表达的有效性。更少

提交 16 2018 年 5 月;最初宣布 2018 年 5 月。

评论:IJCAI-ECAI 2018 (显示)

21. [arXiv: 1805.00472](#)[pdf,其他] Cs。简历

基于协同双域补丁滤波的图像去噪

作者:[Muzammil 贝赫扎德](#)

摘要: 本文提出一种利用空间和变换域特征的图像去噪算法。针对协同支持无关稀疏重建, 我们实现了基于强度不变性的改进分组。首先, 我们通过强度不变相关度量来叠加相似结构的补丁。分组的修补程序协作为噪声过滤产生理想的稀疏估计值。这是因为类似的修补程序在已转换的域中共享相同的支持, 因此类似的支持可以用作主动点击的概率来细化稀疏估计。这最终会产生一个非常有用的补丁估计, 从而通过丢弃噪声导致的组件来提高恢复图像的质量。然后应用基于区域生长的后处理器, 通过提取空间域特征进一步增强平滑区域。我们还扩展了我们提出的彩色图像去噪方法。比较结果与最先进的算法的峰值信噪比 (毛里塔尼亚) 和结构相似性 (独运) 指数从广泛的实验, 通过广泛的情况下, 证明了我们提出的算法的优越性。更少

提交 1 2018 年 5 月;最初宣布 2018 年 5 月。

评论:14 页, 14 数字, 4 表格, 文章待定

22. [arXiv: 1804.08891](#)[pdf] Cs。红外

[doi](#)10.1016/j. eswa. 2015.11.023

改进建议的多级协作筛选方法

作者:[欧斯尼古拉斯 Polatidis](#), [克里斯托斯欧亚 k. 乔治亚迪斯](#)

摘要: 协作筛选是在各种在线环境中提供建议的最常用方法之一。尽管协作推荐方法因其简单易用而得到广泛应用, 但准确性仍然是一个问题。本文提出了一种多层次的推荐方法, 其主要目的是通过提供更好的质量建议来帮助用户做出决策。该方法可应用于使用协作推荐系统的不同在线域, 从而提高了整体用户体验。通过使用五个实际数据集进行广泛的实验评估, 并与备选方案进行比较, 可以看出该方法的效率。更少

提交 24 2018 年 4 月;最初宣布 2018 年 4 月。

期刊 ref:Polatidis、欧斯尼古拉斯和克里斯托斯欧亚 k. 乔治亚迪斯。改进建议的多级协作筛选方法。应用专家系统 48 (2016): 100-110

23. [arXiv: 1804.06201](#) [pdf,其他] Cs. 红外

LCMR: 用于与非结构化文本进行协作筛选的本地和集中内存

作者:光能,余张,羌阳

摘要: 协同滤波 (CF) 是推荐系统的关键技术。纯 CF 方法仅利用用户项交互数据 (例如单击、赞和视图), 并遭受稀疏问题的影响。项目通常与内容信息 (如非结构化文本) 相关联 (例如, 文章摘要和产品评论)。CF 可以扩展以利用文本。本文开发了一个统一的神经网络框架来无缝地利用交互数据和内容信息。建议的框架, 称为 LCMR, 是基于内存网络, 并由本地和集中的记忆, 以开发内容信息和交互数据, 分别。通过将内容信息建模为本地记忆, LCMR 用心学习如何利用用户项目交互的指导。在实际数据集上, LCMR 在命中率和 NDCG 指标方面与各种基线进行比较, 显示出更好的性能。我们进一步进行分析, 了解本地和集中的记忆如何为拟议的框架工作。更少

提交 20 2018 年 4 月,v1 提交 17 2018 年 4 月;最初宣布 2018 年 4 月。

24. [arXiv: 1803.09551](#) [pdf,其他] Cs. 红外

[doi](#)10.1145/3127873

协作式包含隐含反馈的主题过滤与社会隐性因素

作者:广力湖、鑫宇傣族、凤峪秋、瑞夏、陶里、蜀建黄、嘉君陈

摘要: 推荐系统 (RSs) 为不同用户选择个性化项目提供了缓解信息超载问题的有效途径。基于潜在因素的协同滤波 (CF) 由于其准确性和可扩展性, 已成为 RSs 的热门方法。最近, 在线社交网络 and 用户生成的内容为超出评级的推荐提供了不同的来源。尽管 {em 社会矩阵分解} (社会性 MF) 和 {em 主题矩阵分解} (主题 MF) 成功地利用了社会关系和项目审查, 他们都忽略了一些有用的信息。本文结合上述方法对有效数据融合进行了研究。首先, 我们提出了一种新的模型 {mbox{MR3}}, 通过对潜在因素和隐藏主题的调整, 对三种信息源 (即评分、项目审查和社会关系) 进行有效的联合建模。其次, 我们将来自评级的隐含反馈纳入建议模型, 以增强其能力并证明其灵活性。我们通过各种最先进的方法实现对真实数据集的更准确的评级预测。此外, 我们还测量了三个数据源中每一个的贡献, 以及从评级中隐含反馈的影响, 接着是超参数的灵敏度分析。实证研究表明了该模型的有效性和效果, 并提出了推广建议。更少

提交 26 2018 年 3 月;最初宣布 2018 年 3 月。

评论:27 页, 11 数字, 6 表, ACM TKDD 2018

25. [arXiv: 1803.03502](#) [pdf,其他] Cs. 红外

协作式基于图的隐式反馈滤波

作者:领牛,渭南张,田静区,血脂曹,瑞明塘,王秀强他,永宇

摘要: 在矩阵分解 (MF) 方法中引入消费项作为用户的隐式反馈, SVD 是个性化推荐系统最有效的协同过滤方法之一。虽然功能强大, SVD++ 有两个限制: (i). 仅使用用户端隐式反馈, 而项目端隐式反馈 (也可以丰富项目表示) 不被利用; (二). 在 SVD++ 中, 当结合隐式反馈时, 相互作用的项目是同等加权的, 不能准确反映用户的真实喜好。为解决上述局限性, 本文提出了基于图的协同滤波 (GCF) 模型、基于加权图的协同滤波 (W gcf) 模型和基于图的协同过滤模型 (i)。基于用户项目二部图对项目侧隐式反馈进行概化; (ii). 灵活地了解个体在隐含反馈中的权重, 从而提高模型的容量。综合实验表明, 我们提出的模型优于最先进的模型。对于稀疏隐式反馈方案, 通过利用步骤两个隐式反馈信息进一步改进。更少

提交 9 2018 年 3 月;最初宣布 2018 年 3 月。

评论:8 页, 7 数字

26. [arXiv: 1803.02250](#)[pdf,其他] Cs。红外

[doi](#)10.1145/3240323.3240378

CF4CF: 使用协作筛选建议协作筛选算法

作者:蒂亚戈,卡洛斯苏亚雷斯,安德烈 c.p.l. 卡瓦略

摘要: 在 可为新问题选择最佳算法的自动解决方案通常在文献中找到。最近得到大量努力的一个研究领域是协作过滤。现有的工作包括使用元学习的几种方法, 它将数据集的特性与算法的性能相关联。这项工作探讨了解决这一问题的另一种方法。从本质上讲, 两者都是建议问题, 因此该工作使用协作过滤算法选择协同滤波算法。我们的方法集成了次像素采样界标, 它是元学习中常用的一种数据表征方法, 采用标准协同滤波方法。实验结果表明, CF4CF 在协同滤波算法选择问题上与标准元学习策略竞争。更少

提交 6 2018 年 3 月;最初宣布 2018 年 3 月。

27. [arXiv: 1801.06605](#)[pdf,其他] cs.SE

[doi](#)10.1145/3167132.3167299

Web 应用程序中测试用例优先级的协同过滤推荐系统

作者:马鹿阿齐兹, Hyunsook 做

摘要: 软件系统的相关指标的使用可以改进各种软件工程任务, 但确定指标之间的关系并不简单, 而且非常耗时。推荐系统可以帮助这个决策过程, 许多应用程序利用这些系统来提高其应用程序的性能。为了研究推荐系统在回归测试中的潜在优势, 我们实现了一个基于项目的协同过滤推荐系统, 它使用用户交互数据和应用程序更改历史信息来开发测试用例优先级技术。为了评估我们的方法, 我们使用三个具有多个版本的 web 应用程序进行了实证研究, 并比较了四种控制技术。我们的结果表明, 我们的推荐系统可以帮助提高测试优先级的有效性。更少

提交 19 2018 年 1 月;最初宣布 2018 年 1 月。

评论:在 SAC 2018 的议事录: 应用计算研讨会

28. [arXiv: 1711.11458](#)[pdf,其他] Cs。红外

协作式社会暴露的过滤: 社会推荐的模块化方法

作者:赵梦晗,杨杨, 坤张

摘要: 本文关注如何有效利用社会信息来改进建议。大多数现有的社交推荐系统都假定人们与他们的社交朋友分享相似的喜好。然而, 这可能不真实, 由于各种动机, 使在线

朋友和动态的在线社交网络。受最近基于因果流程的建议的启发，第一个模型用户对项目的暴露，然后使用这些曝光来指导评级预测，我们利用社交信息来捕获用户曝光，而不是用户偏好。我们假设人们从他们的在线朋友那里获得产品信息，他们不需要分享类似的喜好，这是限制性较低，似乎更接近现实。在这一新的假设下，本文提出了一种新的推荐方法（名为 SERec），将社会接触集成到协同过滤中。我们提出了实施 SERec 的两种方法，即社会正规化和社会促进，各有不同的方式来构建社会暴露。对四个真实世界数据集的实验表明，我们的方法优于最先进的方法。进一步的研究比较了两种方法的鲁棒性和可扩展性。更少

提交 30 2017 年 11 月;最初宣布 2017 年 11 月。

评论:在第三十二人工智能会议 (AAAI 2018), 路易斯安那州新奥尔良接受出版

29. [arXiv: 1711.10816](#)[pdf,其他] Cs. 红外

协同过滤的潜在因子解释

作者:[匿名达拓](#),[索菲亚娜塔莉娅](#),[彼得亚雷](#) [Mardziel](#), [Shayak](#) 森

摘要: 许多机器学习系统利用潜在因素作为预测的内部表述。由于这些潜在因素主要未解释，但是，使用它们进行的预测是不透明的。通过矩阵分解进行协作过滤是使用未解释潜伏特征的这种算法的一个主要例子，但在许多推荐任务中却被广泛采用。我们提出了潜在因素解释 (LFI)，一种解释模型的方法，利用潜在因素对人理解特征的解释。因此，潜在因素的解释可以取代未解释的潜在因素，从而产生一个新的模型来表达对解释特征的预测。然后，可以使用最近开发的模型解释技术来解释这个新模型。本文开发了基于协同滤波的推荐系统 LFI。我们说明了 LFI 解释在 MovieLens 数据集上的使用，结合了比喻和 DB 的辅助特征，并表明潜在的因素可以有足够的准确性来预测真实模型的预测。更少

提交 9 2018 年 4 月;v1 提交 29 2017 年 11 月;最初宣布 2017 年 11 月。

30. [arXiv: 1708.05031](#)[pdf,其他] Cs. 红外

神经协同滤波

作者:[湘南](#),[栗子辽](#),[汉王张](#),[李强](#),[下胡](#), [Tat](#)

摘要: 近年来，深度神经网络在语音识别、计算机视觉和自然语言处理等方面取得了巨大的成功。然而，对推荐系统的深度神经网络的探索却受到了相对较少的审查。在这项工作中，我们努力开发基于神经网络的技术来解决推荐中的关键问题--协同过滤--在隐含反馈的基础上。虽然最近的一些工作已采用深度学习的建议，他们主要使用它来模拟辅助信息，例如项目的文本描述和音乐的声学特征。在对协同过滤的关键因素进行建模时--用户与项目特征之间的相互作用，它们仍然采用矩阵分解法，并将内部产品应用于用户和项目的潜在特征。通过将内部产品替换为可从数据中学习任意函数的神经结构，我们提出了一个名为 NCF 的通用框架，简称为基于神经网络的协作筛选。NCF 是通用的，可以在其框架下表达和推广矩阵分解。为了用非线性来增压 NCF 建模，我们建议利用多层感知器来学习用户-项目交互功能。在两个实际数据集上进行的大量实验显示了我们建议的 NCF 框架在最先进的方法上的显著改进。经验证据表明，使用更深的神经网络层可以提供更好的推荐性能。更少

提交 25 2017 年 8 月;v1 提交 16 2017 年 8 月;最初宣布 2017 年 8 月。

评论:10 页, 7 数字

31. [arXiv: 1708.03658](#)[pdf, ps,其他] Cs. 红外

iTrace 信任协作过滤的隐式信任推理方法

作者:徐、刘、柯家陈

摘要: 互联网商业的增长刺激了使用协作过滤 (CF) 算法作为推荐系统。协作筛选 (CF) 算法通过利用其他类似用户提供的选票, 建议目标用户感兴趣的项目。在标准 CF 框架中, 假定每个投票用户的可信度与目标用户完全相同。这种假设不满足, 因此可能会导致许多实际应用中的误导性建议。一种自然的对策是设计一种信任感知 cf (终端) 算法, 它可以考虑到在执行 CF 时投票用户 credibilities 的差异。为此, 本文提出了一种信任推理方法, 它可以预测隐式来自稀疏显式信任矩阵的每个投票用户的目标用户信任。然后提出了一种改进的 cf 算法 iTrace, 利用显式和预测隐式信任来为 cf 框架提供建议。对公共数据集的实证评估表明, 所提出的算法在平均绝对误差 (MAE) 方面对推荐质量有显著提高。更少

提交 15 2017 年 8 月;v1 提交 11 2017 年 8 月;最初宣布 2017 年 8 月。

评论:6 页, 4 数字, 1 表

32. [arXiv: 1708.03058](#)[pdf,其他] Cs。红外

使用从属武器的多臂强盗在线交互式协作过滤

作者:庆王,春秋,伍佰,陶李,拉里萨 Shwartz, Genady 雅。Grabarnik

摘要: 在线互动推荐系统力求根据当前环境 (包括消费者和项目内容信息) 及时向消费者推荐合适的项目 (如电影、新闻文章)。但是, 此类上下文信息在实践中通常不可用, 因为该建议仅可利用用户对项目的交互数据。此外, 缺乏互动记录, 特别是对于新用户和项目, 进一步恶化了建议的表现。为了解决这些问题, 协同过滤 (CF) 是仅依赖于交互数据的推荐技术之一, 以及能够实现开发与探索之间平衡的在线多武装强盗机制, 是通过在在线互动推荐设置中, 假设独立项目 (即武器)。然而, 这种假设很少在现实中存在, 因为现实世界中的项目往往相互关联 (例如, 两篇具有类似主题的文章)。本文通过考虑项目间的依赖性, 研究了在线交互式协同过滤问题。我们明确地将项目依赖项作为武器上的集群, 其中单个集群内的武器共享类似的潜在主题。在主题建模技术的基础上, 我们想出了一个生成模型, 从它们的底层主题产生项目。此外, 还开发了一种基于粒子学习的高效在线算法, 用于推断模型的潜在参数和状态。此外, 我们的推断模型可以在在线交互式协作设置中与现有的多武器选择策略自然集成。对两个实际应用的实证研究, 电影和新闻的在线建议, 证明了所提出方法的有效性和效率。更少

提交 11 2017 年 8 月;v1 提交 9 2017 年 8 月;最初宣布 2017 年 8 月。

评论:推荐系统;交互式协作过滤;主题建模;冷启动问题;粒子学习;10 页

ACM 类:H. 3.3;i. 2。6

33. [arXiv: 1708.02867](#)[pdf,其他] Cs。Lg

基于快速矩阵分解的协同滤波模拟退火与征税分布

作者:Mostafa a. 谢哈塔,穆罕默德 Nassef, Amr a. 巴德尔

摘要: 矩阵分解是协同滤波的最佳方法之一, 它在呈现用户和项目潜在因素方面具有较高的准确性。矩阵分解的主要缺点是它的复杂性, 并且很难并行化, 特别是非常大的矩阵。本文介绍了一种基于矩阵分解的协同滤波新方法, 将模拟退火与征税分布相结合。采用此方法, 与随机梯度下降、交替最小二乘法和加权非负矩阵分解等其他方法相比, 在可接受的时间内实现了良好的求解。更少

提交 9 2017 年 8 月;最初宣布 2017 年 8 月。

34. [arXiv: 1708.00417](#)[pdf] Cs。红外

社会网络推荐系统与协同过滤技术研究

作者:玛丽亚姆 [Nayebzadeh](#), 巴克穆阿扎姆,埃米尔穆罕默德. 萨巴,哈迪 [Abdolrahimpour](#),伊勒姆"流星-

摘要: 如今, 随着互联网信息的显著扩展, 用户更喜欢通过朋友或个人资料中的一些建议来获得所需的确切信息, 以节省时间和金钱。建议基于不同算法的系统是通过互联网实现这一目标的基本途径之一, 但它们各自都有各自的优缺点。在本研究中, 我们选择并实施了两种协同过滤 (CF) 和社会网络推荐系统 (SNRS) 的方法。基于一些限制来查找包含友情、分级和项目类别的数据集, 我们为 10 个类别、10 个项目和 100 个用户生成了它, 并比较了两种方法。我们用平均绝对误差 (MAE) 和精度来比较两个提到的方法的结果, 发现 SNRS 方法, 因为它被声称是改进的 CF 的版本工作更有效。更少提交 1 2017 年 8 月;最初宣布 2017 年 8 月。

35. [arXiv: 1706.07881](#)[pdf,其他] Cs。Lg

基于神经网络协同滤波的抽样策略研究

作者:亮杰, 岳孙,悦仕,洪

摘要: 神经网络的最新进展启发人们设计混合推荐算法, 它们可以同时包含 (1) 用户项交互信息和 (2) 内容信息, 包括图像、音频和文本。基于神经网络的推荐算法尽管具有很大的应用前景, 但其计算成本也很高, 因此在规模和改进上具有挑战性。本文提出了一种基于神经网络的通用推荐框架, 包容了几种现有的最先进的推荐算法, 并通过对随机梯度抽样策略的研究来解决效率问题。框架的下降训练。我们首先在损失函数和用户项交互二部图之间建立一个连接, 在链路上定义损失函数术语, 同时主要计算负担位于节点上, 从而解决这个问题。我们将这种类型的损耗函数称为 "基于图的" 损耗函数, 不同的小批量采样策略可以有不同的计算成本。在深入研究的基础上, 提出了三种新的抽样策略, 可以显著提高建议框架的训练效率 (最多 $\times 30$ 在我们的实验中加速了速度), 并提高了推荐性能。同时还为计算成本和收敛性提供了理论分析。我们认为, 抽样策略的研究对一般基于图的损耗函数有进一步的影响, 并能在基于神经网络的推荐框架下进行更多的研究。更少

提交 23 2017 年 6 月;最初宣布 2017 年 6 月。

评论:这是 KDD'17 纸的较长版本 (附附加)

36. [arXiv: 1706.00061](#)[pdf,其他] Cs。Lg

在线一级协同过滤的示例复杂性

作者:莱因哈德海格尔,坎南 [Ramchandran](#)

摘要: 我们考虑在线的一级协作过滤 (CF) 问题, 包括在一段时间内以在线方式向用户推荐项目, 仅基于正面评级。当用户偶尔响应具有正数的建议, 而不使用负数时, 就会出现此问题。我们研究了用户对建议的响应概率 (p_f) 对示例复杂性的影响, 即: 做出 "好" 建议所需的评分数量, 并询问是否接收正和负评级, 而不是正额定值只会提高示例复杂性。在推荐系统的设计中出现了两个问题。介绍了一种简单的概率用户模型, 分析了一种基于用户的在线 CF 算法的性能。我们证明, 在最初的冷启动阶段后, 建议被投资于探索用户的喜好, 此算法使---到更新用户首选项所需的建议的一小部分---完美的建议。冷启动阶段所需的分级数量几乎与 $1/p_f$ 成正比, 对于更新用户首选项, 本质上与 p_f 无关。因此, 我们发现, 接受阳性和阴性评级而不是只有阳性值, 可以提高初始勘探所需的评分数量 ($1/p_f$), 这可能很重要。更少

提交 31 2017 年 5 月;最初宣布 2017 年 6 月。

评论:ICML 2017

37. [arXiv: 1705.08804](#)[pdf, ps,其他] Cs. 红外

超越奇偶:协同过滤的公平目标

作者:斯瑞姚,伯特黄

摘要: 我们在协作过滤推荐系统中学习公平性, 这对历史数据中存在的歧视非常敏感。有偏见的数据可以引导协作过滤方法, 对少数群体的用户做出不公平的预测。我们确定现有公平指标的不足, 并提出四个新的指标来解决不同形式的不公平。通过向学习目标添加公平性条件, 可以优化这些公平性指标。关于合成和真实数据的实验表明, 我们的新指标比基线更能衡量公平性, 公平目标有效地帮助减少不公平。更少

提交 30 2017 年 11 月;v1 提交 24 2017 年 5 月;最初宣布 2017 年 5 月。

38. [arXiv: 1705.07051](#)[pdf,其他] Cs. 红外

加速基于内存的协同过滤与地标

作者:利马,卡洛斯 e. 梅洛,杰拉尔多 Zimbrão

摘要: 推荐系统在许多情况下都扮演着重要的角色, 在这些场景中, 用户无法做出太多选择。在这种情况下, 协作过滤 (CF) 是通过为个性化推荐提供一种简单而广泛使用的方法而产生的。基于内存的 CF 算法主要依赖于用户或项目对的相似性, 这些向在 k 最近邻 (kNN) 等分类器中使用, 以便对未知的评级进行一般化。这个方法的一个主要问题是建立相似矩阵。根据评级矩阵的维数, 相似性计算可能会变得难以计算。为了克服这个问题, 我们建议用户与预选用户的距离, 即地标。此过程允许大幅降低与相似性矩阵相关的计算成本。我们对两个截然不同的数据库的建议进行了评估, 结果表明, 我们的方法在计算性能方面一贯且显著优于八 CF 算法 (包括基于内存和基于模型的)。更少

提交 19 2017 年 5 月;最初宣布 2017 年 5 月。

ACM 类:H. 2.8;H. 3.3;高 4. 0

39. [arXiv: 1704.05735](#)[pdf,其他] cs SI

协作式使用社交本地模型进行筛选

作者:刘全明, 羚羊, 詹姆斯.郭,李先生

摘要: 矩阵分解 (MF) 是推荐系统的一种非常流行的方法。它假定下面的评级矩阵是低级的。但是, 这种假设可能过于严格, 无法捕获用户和项目之间的复杂关系和交互。最近, 局部低秩矩阵逼近 (LLORMA) 已被证明是非常成功的解决这个问题。它只是假设评级矩阵是由一些低级子矩阵由类似的用户和项目的子集构造的。虽然 LLORMA 优于 MF, 但如何构造这种子矩阵仍然是一个大问题。在今天的推荐系统中, 我们提出了一种新的框架, 即社会局部低级矩阵逼近 (SLOMA) 来解决这个问题。根据我们的知识, SLOMA 是第一个将社会联系融入当地低级框架的工作。此外, 我们通过将社会正规化应用于子矩阵分解 (表示为 SLOMA++) 来增强 SLOMA。因此, 提出的模型可以从社会建议和地方低级假设中获益。从两个真实世界的数据集、Yelp 和豆瓣的实验结果, 证明了所提出的模型优于 LLORMA 和 MF 的优越性。更少

提交 14 2017 年 9 月;v1 提交 19 2017 年 4 月;最初宣布 2017 年 4 月。

评论:10 页

40. [arXiv: 1704.02552](#)[pdf,其他] Cs. 红外

"冷启动" 预测的嵌入式协同过滤

作者: [优博, 阿里非洲](#)

摘要: 由于仅使用隐式数据, 许多推荐器系统一般无法向具有有限交互历史记录的用户提供一组精确的建议。此问题被视为 "冷启动" 问题, 通常通过切换到需要额外昂贵信息的基于内容的方法来解决。本文采用一种维数约简算法, Word2Vec (W2V), 最初应用于协同滤波 (CF) 框架下的自然语言处理问题, 仅使用隐式数据处理 "冷启动" 问题。这种组合方法被命名为嵌入式协同过滤 (ECF)。对两种不同的隐式数据集进行了实验, 确定了 ECF 的性能。我们表明, ECF 方法在 "冷启动" 方案中优于其他流行和最先进的方法。更少

提交 8 2017 年 4 月;最初宣布 2017 年 4 月。

41. [arXiv: 1704.00551](#) [pdf, 其他] Cs. 红外

[doi](#)10.1145/3077136.3080689

AutoSVD: 一种基于压缩自编码器的高效混合协同滤波模型

作者: [帅张, 莉娜瑶, 西威许](#)

摘要: 协作过滤 (CF) 已成功地用于为用户提供个性化的产品和服务。然而, 处理用户项矩阵的增加稀疏仍然是一个挑战。为了解决这一问题, 我们广泛研究了混合 CF, 如结合基于内容的过滤和利用用户和项目的侧信息, 以提高性能。然而, 这些方法大多依赖于手工制作的特征工程, 它们通常容易产生噪声, 并有不同的特征提取和选择方案的偏差。本文提出了一种新的混合模型, 将压缩自编码器范式推广到矩阵分解框架中, 具有良好的可扩展性和计算效率, 将内容信息作为效率和紧凑性, 并利用隐性用户反馈来做出准确的建议。三多个大型实际数据集进行的大量实验表明, 建议的方法优于项目建议的比较方法。更少

提交 12 2017 年 6 月;v1 提交 3 2017 年 4 月;最初宣布 2017 年 4 月。

评论:4 页, 3 数字

42. [arXiv: 1703.01093](#) [pdf] Cs. 红外

采用频谱域特征实现高效协作过滤

作者: [Doaa m Shawky](#)

摘要: 协作过滤 (CF) 是一款功能强大的推荐系统, 可根据类似用户的评级为活动用户生成推荐项目列表。本文首先通过自组织映射 (SOM), 其次是 k-均值聚类, 提出了一种新的 CF 方法, 即先找到与主动用户相似的用户组。然后, 最接近活动用户的群集中的每个项目的分级将使用离散傅立叶变换 (DFT) 映射到频域。给出了映射额定值的功率谱, 并计算了基于这些功率谱相干性的新相似测度。建议的相似性度量比当前最先进的措施更有效。此外, 它还可以捕获用户配置文件之间的全局相似性。实验结果表明, 该方法克服了现有 CF 算法中存在的主要问题: 首先, 通过创建相似用户群并应用时间效率相似性测度, 减轻了可伸缩性问题。其次, 其基于频率的相似性度量对稀疏性问题不太敏感, 因为 DFT 即使使用稀疏数据也能有效地执行。第三, 它在精度方面优于标准相似性度量。更少

提交 3 2017 年 3 月;最初宣布 2017 年 3 月。

43. [arXiv: 1703.00397](#) [pdf, ps, 其他] Cs. 红外

基于模型协同过滤的冷启动用户问题的研究

作者: [Sampoorna 比斯瓦斯, 利器 v.s. 拉克什曼南, Senjuti 巴苏射线](#)

摘要: 为了解决基于模型的推荐系统中众所周知的冷启动用户问题, 一种方法是向冷启动用户推荐一些项目, 并使用反馈来学习配置文件。然后, 可以使用学习的配置文件向冷用户提出好的建议。在没有良好的初始配置文件的情况下, 建议类似于随机探针, 但如果不明智地选择, 建议和太多建议都可能会关闭用户。我们通过问什么是冷启动用户问题形式化 B 最好的项目, 我们应该推荐给冷启动用户, 为了了解她的个人资料最准确, 其中 B, 一个给定的预算, 通常是一个小数目。我们将这个问题形式化为一个优化问题, 并提出了多个非平凡的结果, 包括 NP 硬度以及近似的硬度。此外, 我们还表明, 目标函数, 即已学习配置文件的最小平方误差 w.r.t. 真实用户配置文件, 既不是脱机也不是模, 这表明不太可能存在有效逼近。最后, 我们讨论了几种可扩展的启发式方法来识别 B 推荐给用户的最佳项目, 并在 4 实际数据集上对其性能进行了实验性评估。我们的实验表明, 我们所提出的加速算法在奔跑时间内显著优于现有技术, 同时在学习的用户配置文件和评级预测中实现类似的错误。更少
提交 17 2017 年 2 月;最初宣布 2017 年 3 月。

44. [arXiv: 1702.01713\[pdf\]](#) Cs. 红外

[doi:10.1016/j.csi.2016.10.014](#)

一种改进建议的动态多级协同过滤方法

作者: 欧斯尼古拉斯 Polatidis, 克里斯托斯欧亚 k. 乔治亚迪斯

摘要: 在各种在线环境 (如电子商务) 中提供建议的最常用方法之一是协作筛选。虽然, 这是一个简单的方法来推荐项目或服务, 准确性和质量问题仍然存在。因此, 我们提出了一种动态的多层次协同过滤方法, 提高了建议的质量。所提出的方法是基于正和负的调整, 可用于不同的领域, 利用协同过滤来提高用户体验的质量。此外, 该方法的有效性通过提供基于三个实际数据集的广泛的实验评估以及与替代方法的比较来显示。更少

提交 6 2017 年 2 月;最初宣布 2017 年 2 月。

期刊 ref: 计算机标准 & 接口, 51, 14-21 (2017)

45. [arXiv: 1612.07146\[pdf,其他\]](#) Cs. Lg

协作式使用用户项共自回归模型进行筛选

作者: 潮杜, 常崇煊, 殷郑, 君朱, 博张

摘要: 深度神经网络在协同滤波 (CF) 中表现出了一定的前景。但是, 现有的神经方法要么是基于一用户的, 要么基于一项目的, 不能显式利用所有基础信息。本文提出了一种用于 cf 任务的神经共自回归模型, 它利用了用户和项目域中的结构相关乌伊卡。自回归允许为不同的任务合并额外的所需属性。此外, 我们还开发了一种有效的随机学习算法来处理大规模数据集。我们评估乌伊卡在两个流行的基准: MovieLens 1M 和 Netflix, 并获得最先进的性能在评级预测和前 N 推荐任务, 这表明了 cf 乌伊卡的有效性。更少
提交 5 2018 年 7 月;v1 提交 21 2016 年 12 月;最初宣布 2016 年 12 月。

评论: 发布于 AAAI 2018

46. [arXiv: 1612.07025\[pdf,其他\]](#) Cs. 红外

在前 N 项建议中进行协作筛选的布尔内核

作者: 米尔科 Polato, 法比奥赫泥

摘要: 在许多个性化建议问题中, 可用数据仅包含用户和项目之间的正交互 (隐式反馈)。此问题也称为一级协作筛选 (OC)。线性模型通常在 OC 问题上达到最先进的性

能,许多努力都致力于建立更具表现力和更复杂的表述,从而改进建议。最近的分析表明,协同滤波(CF)数据集具有高稀疏性和长尾分布的特征。在本文中,我们提出了一个布尔核,称为分离核,它的表达性比线性的少,但它能够缓解CF上下文中的稀疏性问题。此内核的嵌入由输入变量的某个arity d 的所有组合组成,这些组合要素在语义上解释为输入变量的析取。对几个CF数据集的实验表明了所提出的内核的有效性和效率。更少

提交 18 2017 年 7 月;v1 提交 21 2016 年 12 月;最初宣布 2016 年 12 月。

评论:24 页, 28 数字, 2 表格

47. [arXiv: 1612.05729](#)[pdf,其他] Cs。红外

利用稀疏性构建高效的基于内核的协同过滤算法

作者:米尔科 Polato,法比奥赫泥

摘要: 隐式反馈数据集的越来越多的可用性提高了开发有效的协作过滤技术的兴趣,能够以非对称方式处理明确的正反馈和不明确的负反馈。本文提出了一种基于原则内核的协同滤波方法,用于顶 N 项推荐和隐式反馈。我们提出了一个使用线性内核的高效实现,并展示了如何将其推广到点产品系列的内核中,以保持效率。我们还研究了影响标准余弦核稀疏性的因素。此分析表明,内核的稀疏性强烈依赖于数据集的属性,尤其是长尾分布。我们将我们的方法与最先进的算法相比较,在效率和有效性方面取得了良好的效果。更少

提交 17 2016 年 12 月;最初宣布 2016 年 12 月。

评论:神经计算(爱思唯尔日志)下的修订版

48. [arXiv: 1611.00384](#)[pdf] Cs。红外

从内容到协作过滤的深度旅程

作者:奥伦巴尔肯,诺姆·乔姆斯基克尼格斯泰因,奥地利 Yogev

摘要: 在推荐系统研究中,算法通常被描述为协作过滤(CF)或基于内容的(CB)。CF算法使用用户显式或隐式首选项的数据集进行训练,而CB算法通常基于物料配置文件。这些方法利用非常不同的数据源,因此所得到的推荐项通常也非常不同。本文提出了一种新的模型,作为从项目内容到其CF表示的桥梁。引入了一种基于文本描述和元数据的多输入深度回归模型来预测项目的CF潜伏嵌入向量。通过对基于文本描述的电影和应用的CF向量进行预测,我们展示了所提出模型的有效性。最后,我们通过合并元数据(如电影发布年份和标记)来进一步改进模型,从而提高准确性。更少

提交 1 2016 年 11 月;最初宣布 2016 年 11 月。

49. [arXiv: 1610.06633](#)[pdf,其他] Cs。Hc

通过协作近距离过滤进行新奇学习

作者: Schrater, 保罗。

摘要: 绝大多数推荐系统模型首选项为静态或缓慢变化由于可观察的用户体验。但是,用户首选项中的自发更改在许多域(如媒体消耗)中无处不在,而在首选项中驱动更改的关键因素是不可直接观察到的。这些潜在的偏好变化来源带来了新的挑战。当系统不跟踪和适应用户的喜好时,用户会失去信心和信任,从而增加用户流失的风险。我们通过开发一种新颖的偏好模型来学习和跟踪潜在的用户品味,从而应对这些挑战。结合三项创新:基于消费共产模式的商品相似度新测度;在首选项中的{em 自发}更改

的模型;和一个学习代理,跟踪每个用户的动态偏好并学习多样化的个性化策略。由此产生的框架自适应地为用户提供了根据自己的喜好更改本身的新颖性.较少

提交 20 2016 年 10 月;最初宣布 2016 年 10 月。

50. [arXiv: 1610.04850](#)[pdf,其他] Cs。红外

协同滤波中有效的矩形最大体积算法在分级启发式中的实现

作者:[亚历山大 Fonarev](#),[亚历山大 Mikhalev](#),[帕维尔·彼得罗维奇谢尔久科夫](#),[格莱布 Gusev](#),[伊万 Oseledets](#)

摘要: 协同过滤中的冷启动问题可以通过要求新用户对代表性项目的小种子集进行评价,或者要求代表用户对新项目进行评价来解决。问题是如何建立一个种子集,可以给出足够的偏好信息来提出好的建议。最成功的方法之一,称为基于矩阵分解,是基于 Maxvol 算法。不幸的是,这种方法有一个重要的限制---一个特定大小的种子集需要与该大小一致的固定秩的分级矩阵分解。这在一般情况下不一定是最佳的。在本文中,我们引入了一种快速算法来分析推广这种方法,我们称之为矩形 Maxvol。它允许分解的秩低于所需的种子集的大小。此外,本文还包括对方法误差的理论分析、现有方法的复杂性分析以及先进方法的比较。更少

提交 16 2016 年 10 月;最初宣布 2016 年 10 月。

评论:[IEEE 国际数据挖掘会议 \(ICDM\) 2016](#)

51. [arXiv: 1608.08646](#)[pdf, ps,其他] Cs。红外

里拉: 一种新的基于可能性的协同过滤相似性评分

作者:[纯美 Stmadova-尼利](#),[Aydin Buluc](#),[约翰 r. 吉尔伯特](#),[狮子座 Olier](#),[欧阳](#)

摘要: 推荐系统数据为数据挖掘、机器学习和算法社区带来了独特的挑战。高缺失数据速率与典型的推荐系统数据的大规模和高维度相结合,需要新的工具和方法来进行高效的数据分析。在这里,我们解决了评估推荐系统中两个用户之间的相似性的挑战,其中每个用户只有一小部分评级可用。我们提出了一个新的相似性分数,我们调用里拉,基于用户相似性的统计模型,对于具有许多缺失值的大规模离散值数据。我们表明,基于可能性比率的这一评分比传统的基于用户的协作过滤(如皮尔逊相关系数)的相似程度得分更有效。我们认为,我们的方法具有显著的潜力,可以提高协作筛选的准确性和可扩展性。更少

提交 20 2017 年 3 月,v1 提交 30 2016 年 8 月;最初宣布 2016 年 8 月。

评论:-添加确认-修正的打字错误(结果不变) 8 页

52. [arXiv: 1608.08182](#)[pdf,其他] 手语 Lg

基于因子分解的协同过滤数据中毒攻击

作者:[博里](#),[一宁王](#),[阿尔蒂辛格](#),[制片人 Vorobeychik](#)

摘要: 推荐和协作过滤系统在现代信息和电子商务应用中非常重要。由于这些系统在行业中越来越受欢迎,其产出可能会影响商业决策,为敌对方提供激励措施,以损害此类系统的可用性或完整性。介绍了协同过滤系统的数据中毒攻击。我们演示了一个强大的攻击者如何充分了解学员,可以生成恶意数据,从而最大限度地提高其恶意目标,同时模仿正常用户行为以避免被检测到。尽管完整的知识假设似乎很极端,但它能够对协作筛选方案的脆弱性进行可靠的评估,以实现高度主动的攻击。我们提出了两种流行的基于分解的协作过滤算法的有效解决方案: 公众 {替代最小化} 公式和 \ln 公众 {核

范数最小化} 方法。最后, 我们测试了我们所提出的实际数据算法的有效性, 并讨论了潜在的防御策略。更少

提交 5 2016 年 10 月,v1 提交 29 2016 年 8 月;最初宣布 2016 年 8 月。

53. [arXiv: 1608.07400](#)[pdf, ps,其他] Cs. 红外

协作式用递归神经网络进行滤波

作者:罗宾 Devooght,雨果 Bersini

摘要: 我们表明, 协同滤波可以看作是一个序列预测问题, 在这种解释下, 递归神经网络提供了非常有竞争力的方法。我们特别研究了长短期记忆 (LSTM) 如何应用于协同滤波, 以及它如何与标准近邻和矩阵分解方法在电影推荐中的比较。我们表明, LSTM 在各个方面都具有竞争力, 在项目覆盖范围和短期预测方面, 在很大程度上优于其他方法。更少

提交 3 2017 年 1 月,v1 提交 26 2016 年 8 月;最初宣布 2016 年 8 月。

54. [arXiv: 1608.04839](#)[pdf,其他] Cs. Lg

复合泊松分解的动态协同滤波

作者:Ghassen Jerfel,穆罕默德 e. 巴什布,芭芭拉 e. 恩格尔哈特

摘要: 基于模型的协作筛选分析了用户项交互, 以推断出表示用户首选项和项特性的潜在因素, 以便预测未来的交互。大多数协作筛选算法都假定这些潜在因素是静态的, 尽管已显示用户首选项和项目感知随时间推移而漂移。本文提出了一种基于复合泊松矩阵分解的共轭和数值稳定的动态矩阵分解 (DCPF) 模型, 利用伽马马尔可夫链对平滑漂移的潜在因子进行建模。提出了一种数值稳定的伽玛链结构, 然后提出了一种随机变分推理方法来估计模型的参数。我们将我们的模型应用于时间戳评级数据集: Netflix、Yelp 和最后.fm, 其中 DCPF 实现了比最先进的静态和动态分解模型更高的预测精度。更少

提交 1 2016 年 11 月,v1 提交 16 2016 年 8 月;最初宣布 2016 年 8 月。

55. [arXiv: 1607.00223](#)[pdf, ps,其他] 手语红外

基于内存的协同滤波与 Lucene

作者:真纳罗

摘要: 基于内存的协作筛选是一种广泛使用的方法来提供建议。它通过形成加权投票来预测未观测到的评级, 从而利用用户群体之间的相似性。通常采用定制解决方案来处理大型数据集的高质量建议问题。然而, 这种方法的缺点是一般协作过滤系统的通用性和灵活性丧失。在本文中, 我们开发了一种方法, 允许在传统的全文搜索引擎 (如 Apache Lucene) 之上构建可扩展且有效的协作过滤系统。更少

提交 1 2016 年 7 月;最初宣布 2016 年 7 月。

ACM 类:H. 3.3;高 3. 1

56. [arXiv: 1606.07674](#)[pdf,其他] 手语红外

[doi:10.1145/2988450.2988453](#)

隐式反馈的神经自回归协同滤波

作者:王才亮, 刘,杨邦胜唐,汉宁

摘要: 本文提出了一种利用隐式反馈 (如单击、监视、浏览行为) 进行协同过滤任务的神经自回归模型的隐式 CF 名手。我们首先将用户隐式反馈转换为类似向量和置信向

量, 然后对像向量的概率进行建模, 并由置信向量加权。隐式 CF 名手的训练目标是最大化加权负对数似然。我们在从流行的数字电视流服务收集的数据集上测试隐式 CF 名手的性能。更具体地说, 在实验中, 我们描述了如何将监视计数转换为隐式相对等级, 并将其送入隐式 CF 名手。然后将隐式 CF 名手模型的性能与常用的隐式矩阵分解方法进行了比较。实验结果表明, 隐式 CF 名手显著优于基线。更少

提交 12 2016 年 9 月;v1 提交 24 2016 年 6 月;最初宣布 2016 年 6 月。

评论:5 页, 2 数字, DLRS2016 <http://dlrs-workshop.org/>接受

57. [arXiv: 1605.09477](#)[pdf,其他] Cs。红外

协同滤波的神经自回归方法

作者:杨邦胜,文魁进出口丁,汉宁周

摘要: 本文提出了一种基于有限玻尔兹曼机 (RBM) 的 cf 模型和神经自回归分布估计 (名手) 的名手协同滤波 (cf) 任务的神经自回归体系结构。我们首先介绍了 cf 任务的基本 cf 名手模型。然后, 我们建议通过在不同等级之间共享参数来改进模型。此外, 还提出了名手的分解版本, 以实现更好的可扩展性。此外, 我们考虑了偏好的序号性质, 并提出了优化 CF 名手的序号成本, 从而显示出卓越的性能。最后, 名手可以扩展到一个深度模型, 只有适度增加的计算复杂性。实验结果表明, 带有单个隐藏层的 CF 名手在 MovieLens 1M、MovieLens 10M 和 Netflix 数据集上都优于以前的所有先进方法, 添加更多隐藏图层可以进一步提高性能。更少

提交 30 2016 年 5 月;最初宣布 2016 年 5 月。

评论:1CML2016 接受

58. [arXiv: 1604.05813](#)[pdf,其他] Cs。红外

福尔摩斯: 用于视觉感知的一级协作筛选的稀疏层次嵌入

作者:毁了他, Chunbin 林,建国王,朱利安麦克奥利

摘要: 构建成功的推荐系统需要揭示描述项目属性的基础维度以及用户对它们的偏好。在 "服装推荐" 等领域, 解释用户偏好需要对所讨论项目的视觉外观进行建模。由于人们的 "视觉偏好" 的复杂性和微妙性, 以及所涉及的数据和要素的规模和维度, 这使得推荐特别具有挑战性。最终, 成功的模型应该能够在不同的类别和样式中捕获相当大的差异, 同时仍然建模 "全球" 结构解释的共性, 以对抗稀疏性 (例如冷启动)、变异性和真实世界数据集的规模。在这里, 我们通过构建这样的结构来对不同产品类别的视觉尺寸建模来解决这些挑战。采用一种新颖的分层嵌入体系结构, 我们的方法同时兼顾了高层次 (丰富多彩、黑暗等) 和微妙 (例如随意性) 视觉特征。更少

提交 20 2016 年 4 月;最初宣布 2016 年 4 月。

评论:7 页, 3 数字

59. [arXiv: 1603.04259](#)[pdf] Cs。Lg

Item2Vec: 用于协作过滤的神经项目嵌入

作者:奥伦巴尔肯,诺姆·乔姆斯基克尼格斯泰因

摘要: 许多协作筛选 (CF) 算法是基于项目的, 因为它们分析项目项关系以产生项目相似性。最近, 自然语言处理 (NLP) 领域的一些作品建议使用神经嵌入算法来学习词汇的潜在表示。其中, 带负抽样 (SGNS) (也称为 word2vec) 的跳克被证明可以为各种语言学任务提供最先进的结果。本文介绍了基于项目的 CF 可以在相同的神经字嵌入框架内进行转换。在 SGNS 的启发下, 我们描述了一种我们为基于项目的 CF 命名

item2vec 的方法, 为潜在空间中的项目生成嵌入。即使用户信息不可用, 该方法也能够推断项目项关系。本文给出了实验结果, 证明了 item2vec 方法的有效性, 表明它具有与 SVD 的竞争优势。更少

提交 20 2017 年 2 月;v1 提交 14 2016 年 3 月;最初宣布 2016 年 3 月。

60. [arXiv: 1603.00806](#)[pdf,其他] 手语红外

自动编码器混合协同过滤

作者:Strub,阿利亚玛丽,岁 Gaudel

摘要: 协作筛选旨在利用用户的反馈来提供个性化建议。此类算法在一个大型稀疏矩阵中查找潜在变量。通过添加侧面信息来解决已知的冷启动问题, 可以增强它们的功能。虽然在图像和语音识别方面, 网络具有巨大的成功, 但在协作过滤方面却受到较少的关注。更令人惊讶的是, 神经网络能够发现大型和异构数据集中的潜在变量。本文介绍了一种基于稀疏额定值输入和侧信息的协同滤波神经网络结构, 即 CFN 计算非线性矩阵分解。我们在实验上展示了 CFN outper 的 MovieLens 和豆瓣数据集, 它从侧面信息中形成了艺术的状态和好处。我们提供的算法的实现作为一个可重用的插件火炬, 一个流行的神经网络框架。更少

提交 19 2016 年 7 月;v1 提交 2 2016 年 3 月;最初宣布 2016 年 3 月。

61. [arXiv: 1602.01585](#)[pdf,其他] 手语艾

[doi:10.1145/2872427.2883037](#)

起伏: 通过一级协作过滤来模拟时尚趋势的视觉演进

作者:毁了他,朱利安麦克奥利

摘要: 建立一个成功的推荐系统取决于理解人们的喜好的维度以及他们的动态。在某些领域, 如时尚, 建模这样的喜好是非常困难的, 因为需要同时模拟产品的视觉外观以及随着时间的推移而演变。时尚进化的微妙语义和非线性动态引发了独特的挑战, 尤其是考虑到底层数据集的稀疏性和大规模程度。在本文中, 我们为一级协作过滤设置建立了新的模型, 我们的目标是根据他们过去的反馈来估计用户的时尚感知个性化排名功能。为了揭示人们在评估产品时考虑到的复杂和不断变化的视觉因素, 我们的方法结合了从深度卷积神经网络、用户过去反馈以及不断变化的趋势中提取的高层次视觉特征。社区。实验中, 我们对来自 Amazon.com 的两个大型真实世界数据集进行了评估, 在这里我们展示了它优于最先进的个性化排名措施, 并使用它来可视化 11 年跨度内的高级时尚趋势。数据。更少

提交 4 2016 年 2 月;最初宣布 2016 年 2 月。

评论:11 页, 5 数字

62. [arXiv: 1601.04745](#)[pdf,其他] Cs。红外

冷启动协同滤波两阶段推荐的理论分析

作者:小雪赵,君王

摘要: 本文提出了一种解决冷启动协同过滤问题的理论框架, 其中未知目标 (项目或用户) 不断进入系统, 并且有数量有限的资源 (用户或项目), 可以分配和相关的。解决方案需要在开发和勘探之间权衡, 因为建议机会有限, 一方面, 我们需要立即分配最相关的资源, 但另一方面, 也必须分配资源这对于学习目标的属性非常有用, 以便在将来推荐更多相关的特性。在本文中, 我们研究了一个简单的两阶段建议, 将顺序和批次解决方案结合在一起。我们首先用部分可观测马尔可夫决策过程 (POMDP) 对问题进行

行建模, 并提供一个精确的解决方案。然后, 通过对 POMDP 值迭代解决方案的深入分析, 我们确定一个精确的解决方案可以被抽象为选择资源, 它不仅根据初始阶段信息与目标高度相关, 而且还高度与下一阶段的其他潜在资源有正面或负面的关联。通过这个发现, 我们提出了一个近似的解决方案来缓解难解性的精确解。我们对合成数据和电影镜头 100K 数据集的初步结果证实了我们理论发展和分析的性能提升。更少

提交 18 2016 年 1 月;最初宣布 2016 年 1 月。

63. [arXiv: 1601.02377](#)[pdf, ps,其他] 手语 Lg

展示广告中的隐式相似建模: 将协作过滤转换为点击率估计

作者:渭南张,灵溪,君王

摘要: 在网络广告中, 用户行为定位至关重要。与赞助商搜索关键字定位和上下文广告页面内容定位相比, 用户行为定位通过跟踪其在线行为来构建用户的兴趣配置文件, 然后根据每个用户提供相关广告兴趣, 这将导致更高的定位精度, 从而提高广告性能。当前用户分析方法包括生成关键字和主题标记或将用户映射到分层分类。然而, 根据我们的知识, 没有以前的工作, 明确调查用户在线访问的相似性, 并将这种相似性纳入他们的广告响应预测。在这项工作中, 我们提出了一个通用框架, 它根据他们的在线浏览行为学习用户配置文件, 并将所学知识转移到他们的广告响应预测中。从技术上讲, 我们提出基于概率潜伏因子图形模型的转移学习模型, 用户的广告响应配置文件是从他们的在线浏览配置文件生成的。基于真实世界数据的大规模实验表明, 我们的解决方案在一些强大的基线上有了显著的改进。更少

提交 11 2016 年 1 月;最初宣布 2016 年 1 月。

64. [arXiv: 1511.00792](#)[pdf,其他] Cs. Lg

从隐式反馈到可证明保证的快速协作过滤

作者:Sayantan 达斯古普塔

摘要: 构建推荐算法是机器学习中最具挑战性的任务之一。尽管大多数推荐系统都是基于用户在评级或文本方面提供的明确反馈构建的, 但大多数应用程序没有收到此类反馈。在这里, 我们考虑的建议任务, 唯一可用的数据是用户项目交互的记录, 随着时间的推移, 在订阅或购买项目;这称为隐式反馈建议。对于任何 web 应用程序, 通常都有大量这样的用户项交互。像 PLSI 或矩阵分解这样的算法通过数据集运行多个迭代, 对于大型数据集来说可能非常昂贵。本文提出了一种基于矩量法的推荐算法, 它涉及数据集的二阶和三阶矩的分解。利用 PAC 学习理论可以证明我们的算法是全局收敛的。此外, 我们还演示了如何仅使用三次传递整个数据集来提取参数。这就产生了一种高度可扩展的算法, 即使是在具有单核处理器和 8 GB RAM 的机器上, 也可以扩展到百万用户, 并与现有算法相比产生竞争性能。更少

提交 21 2016 年 8 月,v1 提交 3 2015 年 11 月;最初宣布 2015 年 11 月。

65. [arXiv: 1507.05371](#)[pdf,其他] 手语 Lg

项目项协作筛选的遗憾保证

作者:家伙布雷斯勒, Devavrat,路易斯 f. Voloch

摘要: 有很多经验证据表明, 项目项目协同过滤在实践中效果良好。为了理解这一点, 我们提供了一个框架来设计和分析各种推荐算法。设置相当于在线二进制矩阵完成, 在每次随机用户请求推荐时, 该算法选择要在用户行中显示的条目。目标是尽量减少遗憾, 或等同于在任何时候显示的 +1 个条目的最大数量。与用户-用户协作过滤相比, 我们

分析了一个项目项目协同过滤算法, 可以实现更佳的性能。该算法实现了良好的 "冷启动" 性能 (适当定义), 快速向新用户提出了一些信息很少的好建议。更少
提交 8 2016 年 1 月;v1 提交 19 2015 年 7 月;最初宣布 2015 年 7 月。

66. [arXiv: 1502.03473](#)[pdf,其他] Cs. Lg

协作式过滤土匪

作者:帅里, Karatzoglou, 外邦教徒

摘要: 经典协作筛选和基于内容的筛选方法尝试学习给定训练数据的静态推荐模型。这些方法远不是理想的高动态推荐域, 如新闻推荐和计算广告, 其中一组项目和用户是非常流畅的。在这项工作中, 我们研究了基于上下文多臂强盗设置中的探索开发策略的内容推荐的自适应聚类技术。我们的算法考虑到由于用户与项目之间的交互而产生的协作效果, 通过根据正在考虑的项目动态分组用户, 同时根据 clusterings 诱导了用户。由此产生的算法以类似于协作过滤方法的方式利用数据中的偏好模式。我们提供了对中型实际数据集的实证分析, 显示了可扩展性和更高的预测性能 (通过点击率来衡量), 对集群土匪的最先进方法进行了研究。我们还在标准线性随机噪声设置中提供了一个遗憾分析。更少

提交 31 2016 年 5 月;v1 提交 11 2015 年 2 月;最初宣布 2015 年 2 月。

评论:第三十九 SIGIR (SIGIR 2016)

67. [arXiv: 1406.2431](#)[pdf,其他] 手语红外

基于优化设计的协同过滤推荐人中的预算约束项目冷启动处理

作者:奥伦 Anava,沙哈尔戈兰,土耳其 Golbandi, Zohar Karnin,罗尼伦佩尔, 罗赫连科, 奥伦哈密

摘要: 众所周知, 基于协作过滤 (CF) 的推荐系统提供了更好的用户和项目与相当的评级历史相关的建模。缺乏历史评级导致用户和项目冷启动问题。后者是这项工作的主要重点。当前的大多数文献通过集成基于内容的推荐技术来为新项目建模来解决这个问题。然而, 在许多情况下, 这种内容是不可用的, 问题出现的是, 是否只能使用 CF 技术来缓解此问题。我们将此问题形式化为优化问题: 给定新项、可用用户池和预算约束, 选择要分配给新项目评级的任务的用户, 以最小化模型的预测误差。我们证明了目标函数是单调模的, 并提出了一种基于最优优化设计的有效算法。我们的研究结果通过使用 Netflix 数据集的实证研究进行验证, 在这里, 建议的算法优于手头问题的几个基线。更少

提交 20 2016 年 9 月;v1 提交 10 2014 年 6 月;最初宣布 2014 年 6 月。

评论:11 页, 2 数字

MSC 课程: 62K05