

基于深度学习的正则化矩阵分解推荐系统^①

刘荣辉, 张敬普

河南城建学院 计算机与数据科学学院, 河南 顶山 467000

摘要: 针对当前推荐算法面临的冷启动、数据稀疏以及推荐准确度低等问题, 本文提出一种基于深度学习的正则化矩阵分解推荐系统, 该系统利用深度自动编码器对基于矩阵分解的用户和项目潜在特征进行初始化, 然后使用 Node2vec 网络嵌入技术在用户信任网络中捕获用户潜在特征, 用于计算用户信任度和预测用户对项目的评分. 为了使用户的兴趣与可信用户和社区中最具影响力的人兴趣相似, 本文算法采用 Louvain 和超链接诱导主题搜索 (HITS) 方法寻找社交网络中最具影响力的用户节点, 以正则化的方式将约束信息添加到矩阵分解的目标函数中. 实验结果表明: 本文算法明显优于其他对比推荐算法, 不仅可以缓解用户的冷启动问题, 还提高了推荐质量.

关键词: 推荐系统; 深度学习; 矩阵分解; 正则化; 社交信任网络

中图分类号: TP393

文献标志码: A

文章编号: 1000-5471(2021)07-0103-06

随着在线信息数量、复杂性和动态性的增长, 互联网信息的爆炸性增长常常使用户感到困惑. 推荐系统是解决此类问题的重要手段, 该系统通过有效的信息过滤工具分析用户喜好, 以个性化的方式推荐用户可能感兴趣的产品或服务^[1-2]. 在推荐系统领域, 常用的技术可分为基于协同过滤(collaborative filtering, CF)、基于内容(content-based, CB)、基于关联规则(association rule-based, ARB)和混合推荐系统 4 类. 但是, 这些模型或在处理冷启动和数据稀疏性问题时存在局限性, 或在推荐精度和效率方面有所欠缺^[3].

最近几年, 深度学习在自然语言处理、计算机视觉以及人工智能等领域取得了迅猛进展, 基于深度学习的推荐系统也引起了研究人员的广泛关注^[4]. 深度学习具有多层次非线性网络结构, 其强大的学习能力可以获得用户与项目之间更深层次的特征表示, 而且深度学习能够将不同数据映射到同一个隐空间内, 实现数据的统一表征. 因此, 将深度学习和传统推荐方法进行融合, 能够有效解决推荐系统中冷启动和数据稀疏性问题^[5]. Zhang 等^[6]提出一种基于散列的深度学习框架, 该方法利用离散深度学习将用户项目交互信息和项目内容信息统一起来, 克服了推荐系统中冷启动和效率问题. You 等^[7]提出了层次时间卷积网络, 通过用户与项目的顺序多会话交互来构造深度学习网络, 解决推荐方法的精度和效率问题. 但是, 由于深度神经网络模型的训练大多基于数据的最大似然概率, 很难从不相关的信息中筛选出用户感兴趣的信息, 从而导致推荐的多样性不足. 因此, 研究人员通过将深度学习与传统推荐算法相结合, 来解决两种方法存在的局限性. Fessahaye 等^[8]使用基于内容和协同过滤的混合模型作为深度网络的输入, 利用深度学习分类模型生成具有实时预测的精确推荐系统. 邓存彬等^[9]利用深度学习模型来学习用户和项目之间的特征信息, 将用户特征和项目特征的隐向量用于动态协同过滤算法的输入.

尽管一些研究人员以深度学习和传统推荐算法相结合的方式取得了一定的成果, 但在某一具体问题中并不是每种方法的混合都有效. 随着用户和项目数据量增大, 用户与项目之间的评分矩阵也会越来越稀疏, 推荐偏差和冷启动问题也日益突出. 为了更好地解决这些问题, 本文考虑利用辅助信息来提取用户和

① 收稿日期: 2020-06-11

基金项目: 河南省科技攻关基金项目(1521022100093).

作者简介: 刘荣辉, 博士, 副教授, 主要从事数据挖掘、进化计算及深度学习研究.

物品的共同特征表示,提出了一种基于深度学习的正则化矩阵分解推荐系统.该系统首先利用深度自动编码器初始化用户和项目的潜在特征;然后使用 Node2vec 网络嵌入技术在用户信任网络中捕获用户表示,这些提取的用户表示用于计算用户的信任度和预测用户对项目的评分.同时,本文所提模型还考虑社交信任网络中用户间的信任信息和社区内最具地位用户的影响力作用,在矩阵分解的目标函数中加入相应正则化项,使得用户兴趣与可信用户和社区中最具影响力的人兴趣相似.

1 传统推荐系统模型

1.1 基于矩阵分解的协同过滤

在经典的 CF 问题中,假设用户和项目的数量分别为 m 和 n , $m \times n$ 用户项目矩阵 $\mathbf{R}$ 中的元素 r_{ui} 描述了用户 u 给项目 i 的评分. r_{ui} 如果是空值,则表示 u 没有对 i 评分.通过对矩阵 $\mathbf{R}$ 使用均方根误差来测量性能,推断出缺失的元素.在这种情况下,推荐系统是用户兴趣和项目特征之间的双向交互.在协同过滤方法中,基于矩阵分解(matrix factorization, MF) 技术是常用的算法之一,它将评分矩阵 $\mathbf{R}$ 分解为两个低维度的矩阵.

$$\hat{\mathbf{R}}_{u,i} = \mathbf{P}_u^T \mathbf{Q}_i \quad (1)$$

式(1)中: $\mathbf{P}$ 为用户特征矩阵, $\mathbf{Q}$ 为项目特征矩阵.基于矩阵分解的协同过滤目标函数可以表示为

$$L(\mathbf{R}, \mathbf{P}, \mathbf{Q}) = \min_{\mathbf{P}, \mathbf{Q}} \frac{1}{2} \sum_{u=1}^m \sum_{i=1}^n I_{ui} (\mathbf{R}_{u,i} - \hat{\mathbf{R}}_{u,i})^2 + \frac{\lambda_P}{2} \|\mathbf{P}\|^2 + \frac{\lambda_Q}{2} \|\mathbf{Q}\|^2 \quad (2)$$

式(2)中: λ_P 和 λ_Q 是为了防止模型过拟合的正则化参数. $\mathbf{R}_{u,i}$ 表示真实评分, $\hat{\mathbf{R}}_{u,i}$ 表示预测评分.

1.2 基于社交信任网络的推荐算法

假定将社交信任网络视为有向图 $G(v, \epsilon)$, 其中 v 表示有 m 个用户的集合, ϵ 表示用户之间的信任关系.随着用户之间距离增加,用户之间信任度逐渐降低.因此,利用用户的社会偏好相似度来计算信任度, T 的值定义为

$$T_{u,v} = \frac{\text{sim}(X_u, X_v) \times \text{trust}(u, v)}{\sum_{s \in S_u} \text{sim}(X_u, X_s) \times \text{trust}(u, s)} \quad (3)$$

式(3)中: sim 函数计算两个向量之间的余弦相似度, $X_u \in \mathcal{R}^k$ 和 $X_v \in \mathcal{R}^k$ 分别表示用户 u 和 v 的特征, $\text{trust}(u, v)$ 表示用户 u 对用户 v 的信任值, S_u 表示用户 u 直接链接到的用户集合.如果存在多个信任传播路径,则选择最短路径.

2 基于深度学习的推荐模型

为了更好地解决推荐系统中存在的冷启动、数据稀疏和推荐准确度低等问题,本文所提出的推荐模型结合了 Louvain 社区检测方法、HITS 算法、深度自动编码器和 Node2vec 深度嵌入技术.该模型分为 5 个步骤:①深度自动编码器用于初始化 MF 中用户和项目的潜在特征向量;②使用 Node2vec 深度嵌入方法,从用户信任网络中提取用户潜在特征,用于计算信任值并预测用户对项目的评价;③使用 Louvain 方法识别用户信任网络的社区,并通过 HITS 算法找到每个社区中最重要的节点.④将基于每个社区中最重要的节点的正则化项添加到 MF 目标函数中,以便每个活动用户的特征与公共社区中最重要的节点的特征更加相似,同时添加到 MF 目标函数的另一个正则化项是基于用户之间的信任值,使用户的偏好更接近他们所信任的人;⑤使用梯度下降法来最小化修正后的目标函数.

2.1 问题描述

假设一个推荐系统包含 m 个用户和 n 个项目,用户项目评分矩阵表示为 $\mathbf{R} = \{r_{ij}\}_{m \times n}$, 同时 m 个用户可以用一个集合 v 表示,其社交网络视为有向图 $G(v, \epsilon)$, 用户社交矩阵表示为 $\mathbf{T} = \{t_{ik}\}_{m \times m}$.基于社交信任网络和协同过滤混合的方法通过矩阵分解技术,利用评分矩阵和用户信任网络来学习用户和项目的潜在特征.具体而言,在式(1)的基础上,将社会信任集成到矩阵分解协同过滤模型中,不仅利用社会信任信息来调整用户的潜在因子,同时还添加信任正则化和重要用户影响力正则化两项,表达式为

$$\hat{\mathbf{R}}_{u,i} = \mathbf{P}_u^T \mathbf{Q}_i + \mathbf{W}_u^T \mathbf{X}_u + avg + bias_u + bias_i \quad (4)$$

式中: $\mathbf{W}_u$ 作为用户协同过滤器的权值向量, 由下文中神经网络训练的社会网络表示转化得到. avg 表示所有项目的平均评分, $bias_u$ 表示用户 u 的评分平均值与 avg 的偏差, $bias_i$ 表示项目 i 的评分平均值与 avg 的偏差. 本文优化的目标函数为

$$L(\mathbf{R}, \mathbf{P}, \mathbf{Q}, \mathbf{W}) = \min_{\mathbf{P}, \mathbf{Q}, \mathbf{W}} \frac{1}{2} \sum_{u=1}^m \sum_{i=1}^n I_{ui} (\mathbf{R}_{u,i} - \hat{\mathbf{R}}_{u,i})^2 + \frac{\lambda_C}{2} \sum_{u=1}^m \|\mathbf{P}_u - \mathbf{P}_{s_u}\|^2 + \frac{\lambda_T}{2} \sum_{u=1}^m \sum_{k=1}^m T_{u,k} \|\mathbf{P}_u - \mathbf{P}_k\|^2 + \frac{\lambda_P}{2} \|\mathbf{P}\|^2 + \frac{\lambda_Q}{2} \|\mathbf{Q}\|^2 + \frac{\lambda_W}{2} \|\mathbf{W}\|^2 \quad (5)$$

式(5)中: $\mathbf{P}_{s_u}$ 是用户 s 在用户 u 所属社区中具有最高评分的因子向量, λ_T 和 λ_C 是正则化项参数, 分别用于控制社区中最重要用户的信任和影响力对应的项. 本文仍然使用随机梯度下降算法求解潜在特征向量 $\mathbf{P}, \mathbf{Q}$ 和 $\mathbf{W}$.

2.2 本文提出的推荐模型

本文提出的基于深度学习的混合推荐模型主要分为 MF 初始化、用户信任特征嵌入、社交网络最重要节点发现、目标函数修正和求解 5 个阶段.

2.2.1 深度自动编码器

由于矩阵分解可以随着 $\mathbf{P}$ 和 $\mathbf{Q}$ 的初始值变化而收敛到不同的局部最小值, 因此正确设置初始值对给定的 MF 非凸目标函数获得最优解至关重要. 为了构建合适的 $\mathbf{P}$ 和 $\mathbf{Q}$ 矩阵初始值, 一些文献通过引入自动编码器来优化初值. 自动编码器是实现编码器 $\text{encode}(\mathbf{x}): \mathbf{R}^n \rightarrow \mathbf{R}^k$ 和解码器 $\text{decode}(\mathbf{x}): \mathbf{R}^k \rightarrow \mathbf{R}^n$ 映射的神经网络. 自动编码器的目标是获得数据 x 的 k 维表示, 从而使 x 和 $\text{decode}(\text{encode}(\mathbf{x}))$ 之间的误差度量最小化.

本文使用基于可缩放指数型线性单元的堆栈式深度自动编码器模型来学习用户与项目的隐层特征表示, 模型中的解码器架构与编码器对称, 因此降低了参数数量. 深度自动编码器(deep auto encoder, DAE)使用 UDAE 和 IDAE 两种编码器来获得用户特征矩阵 $\mathbf{P}$ 和项目特征矩阵 $\mathbf{Q}$. 其中 U 表示用户, I 表示项目. 自动编码器的深层神经网络体系结构, 用于获取用户和项目的潜在因子, 其中 x 表示输入量, z 表示获得的潜在因子, y 则是输入 x 的重构. 在 DAE 中, 编码器和解码器部分均包含前馈神经网络, 该网络具有 n 个完全连接层以计算 $f(\mathbf{W} \times x + b)$, 其中 f 是非线性激活函数. $\mathbf{W}$ 是权重, b 是偏置. 权重 $\mathbf{W}_e^l$ 和 $\mathbf{W}_d^l$ 解分别是 l 层编码器和解码器的对应权重. e 代表编码, d 代表解码. 对于每一个输入量 x , 深度自动编码器模型中的 UDAE 和 IDAE 部分分别获得的潜在因子分别为 $\mathbf{P}_u$ 和 $\mathbf{Q}_i$. 本文采用掩模均方误差优化作为 DAE 的损失函数.

$$MMSE = \frac{m_i \times (r_i - y_i)^2}{\sum_{i=0}^n m_i} \quad (6)$$

式(6)中: r_i 和 y_i 分别是实际评分和预测评分, m_i 是掩模函数, 当 $r_i \neq 0$ 时, $m_i = 1$, 否则 $m_i = 0$.

2.2.2 基于 node2vec 嵌入的用户网络

在本文提出的方法中, 用户特征通过用户项目矩阵分解的潜在因子和预先训练用户社交网络的嵌入因子(信任特征)得到. 受神经网络嵌入模型在链路预测和节点分类中的成功启发, 该方法在半监督任务中使用预训练网络嵌入模型 node2vec^[10], 该模型可以学习任意给定网络中节点的高层稳定特征表示, 并通过在网络中多次随机游走来获得连接模式的多样性.

将给定网络视为 $G = (A, E)$, 其中 A 代表节点集合, E 代表边集合. 对于节点 $m \in A$, 令 $N_s(m) \subset A$ 为节点 m 的网络邻域, 该邻域通过 S 策略进行邻域采样生成. S 策略是一种随机游走方法, 通过广度优先或深度优先采样来检测邻居. 为了学习每个节点的高级表示, node2vec 尝试根据其特征表示来最大化节点邻域 $N_s(m)$ 的对数概率.

$$\max_f \sum_{m \in A} \log \Pr(N_s(m) \mid f(m)) \quad (7)$$

式(7)中: f 是从节点到低维特征表示的映射函数, $\Pr$ 表示概率.

由于 node2vec 输出可以解释为网络节点的高级表示, 因此通过训练 node2vec 模型可以提取用户信任网络的深层社会结构, 本文将 $X_u \in \mathcal{R}^k$ 作为从用户信任网络中提取到的用户 u 特征. 这些提取的特征反映了用户深层次的社会兴趣, 它们的线性组合能够表示用户与他人建立社会联系的程度. 此程度信息可用于

预测评分,特别是当用户对很少的项目进行评分时.社交网络和评分偏好可能会对不同类型的信息进行编码,因此将两者有机结合能够获得更佳的预测效果.在本文所提方法中,利用简单线性模型将协同过滤方法的潜在因素与提取的社会因素结合起来,因此,用户 u 对项目 i 的预测评级可以由式(1)更改为式(4).

2.2.3 基于 Louvain 方法的社区发现

社交网络中用户倾向于形成具有较高联系的群体,这些群体称为社区,社区内部成员相对于外部成员是紧密联系的,可以相互信任,共享偏好.社区结构发现技术作为社交网络分析的一项基本技术,对理解社交网络结构特性和信息传播具有重要作用.当前,基于 Louvain 的模块度优化方法是使用较为广泛的一种社区发现方法,该方法具有计算速度快,易于理解和无需监督的优点^[11].

Louvain 算法采用模块度来衡量社区网络划分程度,模块度定义为

$$Q = \frac{1}{2m} \sum_{i,j} \left(A_{ij} - \frac{k_i k_j}{2m} \right) \delta(c_i, c_j) \quad (8)$$

式(8)中: A_{ij} 表示节点 i 和节点 j 之间边的权重, k_i 和 k_j 分别表示所有与节点 i 和节点 j 相连的边的权重之和, m 表示所有边的权重之和, c_i 和 c_j 分别表示节点 i 和节点 j 所属社区, δ 为克罗内克函数. Q 值越大,说明社区内部连接越紧密.

Louvain 算法主要分为 3 个步骤:① 对网络节点进行初始化;② 将节点 i 从原来社区中取出后加入其他社区,并计算相应的模块度增益值;③ 将该点放入增益值最大的社区中.对所有节点重复上述过程,直到所有节点、社区不再发生改变.本文使用 Louvain 方法来识别用户信任网络的社区,该社区将用户 i 的社区索引返回为 c_i .

2.2.4 基于 HITS 查找社区中最重要的节点

在信任社交网络中,不同的人在一个社区中的重要性不一样,可能有一些节点更受人们信任,更容易接受他们的经验和意见.因此,在社会网络分析中,需要使用中心性分析方法来识别网络中有影响的节点.

本文所提方法使用 HITS 算法^[12]来寻找用户信任网络中每个社区最有效的节点. HITS 算法是一种基于链接分析的网页排名算法,该算法通过权威值和枢纽值两个属性对扩展集合中所有页面的重要性进行刻画.为了有效地利用 HITS 算法对社交网络节点的影响力进行评估,首先将每个用户视为一个网页,用户之间跟随或者被跟随视为网页之间的链接关系;然后基于该方法的假设条件,建立整个社交网络的有向图,利用 HITS 算法对各用户的权威值和枢纽值进行计算;最后将算法收敛时用户的权威值作为社交网络中用户所具有影响力的量化值.在信任网络的社区中通过衡量权威值找到最具影响力的用户,而在每个社区中拥有最高权威值的节点被用来影响该社区中其他用户的评分.

2.2.5 添加正则化条件

在实际应用中,用户倾向于在共享社区中与他们信任的朋友共享关于项目的类似偏好.为了获得更加精确的推荐列表,在式(2)的基础上添加信任和重要用户影响力两项正则化项.用户之间的相似度取决于用户之间的信任值,即用户之间的信任值越大,用户之间的相似度就越大.

3 实验与结果分析

3.1 数据集和实验环境

为了评估本文所提推荐模型的性能,使用 Ciao 和 Epinions 这 2 个标准数据集^[12]对算法进行验证. Ciao 数据集来自一个数百万用户的在线评论社区,该社区用户可以给产品打分,分数从 1 到 5. Epinions 数据集作为一个物品评论数据集,具有 40 163 个用户在 139 738 个物品上给出的 664 824 个评分.两个数据集都包括用户对项目的评分、时间以及用户之间的社会信任网络.在这些实验中,数据集被分为 80%和 20%两个部分,分别用于训练和测试.由于评分矩阵非常稀疏,因此这两个数据集中的推荐者问题颇具挑战性.

文章中所有的实验均在 Spyder 环境下,使用 Python 3.7 在 Intel(R) Xeon(R)CPU E5 26200@2GHz, RAM64GB 的机器上进行,Python 中的 Keras 库用于实现深层的自动编码神经网络.

3.2 评估指标

本文使用均方根误差(root mean squared error, RMSE)和平均绝对误差(mean absolute error, MAE)

作为实验结果的评价指标, RMSE 和 MAE 两种评价指标的计算方式如式(9)和式(10)所示.

$$MAE = \frac{1}{|R^{te}|} \sum_{(u,i) \in R^{te}} |R_{u,i} - \hat{R}_{u,i}| \quad (9)$$

$$RMSE = \sqrt{\frac{1}{|R^{te}|} \sum_{(u,i) \in R^{te}} (R_{u,i} - \hat{R}_{u,i})^2} \quad (10)$$

式(9)、式(10)中: R^{te} 表示测试集, $|R^{te}|$ 表示测试集中评分的数量. RMSE 和 MAE 通过计算预测评分和真实评分之间的差异来衡量结果的准确性, 其值越小, 说明推荐准确性越高.

3.3 实验结果分析

为了评估本文所提方法的性能, 将其在 Ciao 和 Epinions 这 2 个数据集上的测试结果与 MIRROR^[12], SBRNE^[13], InSRMF^[14] 和 GraphRec^[15] 等推荐算法进行比较. 本实验比较了各种推荐算法在全体用户集和冷启动用户集上的测试结果, 其中冷启动用户集是指训练集中评分数量少于 5 个的用户所构成的集合.

在两个数据集上, 本文所提出的方法都比其他方法获得了更低的 MAE 和 RMSE, 说明性能优于其他对比方法, 并具有显著的性能提升. 由于 MIRROR 算法只使用了基于矩阵分解的推荐系统, 因此在这两个数据集中它的性能都不如其他方法; InSRMF 算法将社交网络信息和矩阵分解协同过滤技术有效地结合起来, 取得了比 MIRROR 更好的效果; SBRNE 和 GraphRec 利用了用户的社交网络信息和深度嵌入技术, 也取得了较好的效果; 本文所提算法在基于矩阵分解的推荐系统中结合了 Louvain 社区检测方法、HITS 算法、深度自动编码器和 Node2vec 深度嵌入技术, 使结果优于所有其他最先进的方法.

随着冷启动用户的参与, 每种方法的性能都会下降, 在这两个数据集上, 本文所提算法的降低速度始终比其他方法慢, 这意味着本文提出的方法对冷启动用户问题更健壮.

4 结 语

本文提出一种基于深度学习的正则化矩阵分解推荐系统, 用于解决当前推荐算法面临的冷启动、数据稀疏以及推荐准确度低等问题. ①该系统通过深度自动编码器初始化用户和项目的潜在特征; ②使用 Node2vec 深度嵌入技术在用户信任网络中捕获用户特征表示, 进而计算用户的信任度并预测用户对项目的评级. 为了使用户兴趣与可信任用户和社区中最具影响力的人兴趣相似, 本文所提算法采用 Louvain 和 HITS 方法寻找社交网络中最具影响力的用户节点, 然后将其和用户间的信任信息以正则化的方式添加到 MF 目标函数中; ③使用梯度下降算法来最小化修正后的 MF 目标函数. 实验结果表明: 相对于其他推荐算法, 本文算法能够有效缓解用户的冷启动问题, 其推荐质量也得到明显提升.

参考文献:

- [1] GUO Lei, WEN Yu-fei, WANG Xin-hua. Exploiting Pre-Trained Network Embeddings for Recommendations in Social Networks [J]. Journal of Computer Science & Technology, 2018, 33(4): 682-696.
- [2] YI B L, SHEN X H, LIU H, et al. Deep Matrix Factorization With Implicit Feedback Embedding for Recommendation System [J]. IEEE Transactions on Industrial Informatics, 2019, 15(8): 4591-4601.
- [3] 曹玉琳, 李文立. 基于状态空间模型和概率矩阵分解的推荐算法 [J]. 计算机应用研究, 2020, 37(11): 3263-3266.
- [4] HUANG Z H, SHAN G X, CHENG J J, et al. TRec: an Efficient Recommendation System for Hunting Passengers with Deep Neural Networks [J]. Neural Computing and Applications, 2019, 31(1): 209-222.
- [5] 黄立威, 江碧涛, 吕守业, 等. 基于深度学习的推荐系统研究综述 [J]. 计算机学报, 2018, 41(7): 1619-1647.
- [6] ZHANG Y, YIN H Z, HUANG Z, et al. Discrete Deep Learning for Fast Content-Aware Recommendation [C]// The 11th ACM International Conference on Web Search and Data Mining (WSDM). Los Angeles: ACM, 2018.
- [7] YOU J X, WANG Y C, PAL A, et al. Hierarchical Temporal Convolutional Networks for Dynamic Recommender Systems [C]//The World Wide Web Conference. San Francisco: ACM, 2019.
- [8] FESSAHAYE F, PEREZ L, ZHAN T, et al. T-RECSYS: A Novel Music Recommendation System Using Deep Learning [C]// 2019 IEEE International Conference On Consumer Electronics(ICCE). Las Vegas: IEEE, 2019.
- [9] 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.

- [10] PALUMBO E, RIZZO G, TRONCY R, et al. Knowledge Graph Embeddings with Node2vec for Item Recommendation [C]// 2018 European Semantic Web Conference(ESWC). Anissaras: Springer, 2018.
- [11] GHOSH S, HALAPPANAVAR M, TUMEO A, et al. Distributed Louvain Algorithm for Graph Community Detection [C]//2018 IEEE International Parallel and Distributed Processing Symposium(IPDPSW). Vancouver: IEEE, 2018.
- [12] MENG X Y, WANG S H, LIU H, et al. Exploiting Emotion on Reviews for Recommender Systems [C]//The Thirty-Second AAAI Conference on Artificial Intelligence(AAAI-18). New Orleans: AAAI, 2018.
- [13] ZHAO W Z, MA H F, LI Z X, et al. SBRNE: An Improved Unified Framework for Social and Behavior Recommendations with Network Embedding [C]// The 24th International Conference on Database Systems for Advanced Applications (DASFAA). Chiang Mai: Springer, 2019.
- [14] LIU H F, JING L P, YU J, et al. Social Recommendation with Learning Personal and Social Latent Factors [J]. IEEE Transactions on Knowledge and Data Engineering, 2019, 31(9): 1-14.
- [15] FAN W Q, MA Y, LI Q, et al. Graph Neural Networks for Social Recommendation [C]// 2019 The Web Conference (WWW2019). San Francisco: ACM, 2019.

Recommender System of Regularization Matrix Decomposition Based on Deep Learning

LIU Rong-hui, ZHANG Jing-pu

School of Computer and Data Science, Henan University of Urban Construction, 467000, China

Abstract: Aiming at the problems of cold start, data sparseness and low recommendation accuracy faced by current recommendation algorithms, a recommender system of regularization matrix decomposition based on deep learning has been proposed. The depth automatic encoder is used in the system to initialize the potential features of users and projects based on the matrix decomposition. And then, node2vec network embedding technology is used to capture the potential characteristics of users in the user trust network, which is used to calculate the user trust and predict the user's score on the project. In order to make the user's interest similar to that of trusted users and the most influential people in the community, the proposed algorithm uses Louvain and HITS methods to find the most influential user nodes in the social network, and restricts information in a regularized manner add to objective function of matrix decomposition. Experimental results show that the proposed algorithm is significantly better than other comparative recommendation algorithms, which can not only ease the user's cold start problem, but also improve the quality of the recommendation.

Key words: recommender system; deep learning; matrix decomposition; regularization; social trust network

责任编辑 夏 娟