

一种基于深度学习的混合推荐算法

曾旭禹^{1,2} 杨 燕^{1,2} 王淑莹¹ 何太军¹ 陈剑波¹

(西南交通大学信息科学与技术学院 成都 611756)¹

(四川省云计算与智能技术高校重点实验室 成都 611756)²

摘要 推荐系统在电子商务的发展中发挥着越来越重要的作用,但用户对物品评分数据的稀疏性往往是推荐精度较低的重要原因。目前通常采用推荐技术对辅助信息进行处理,以缓解用户评价的稀疏性,并提高预测评分精度。通过相关模型,可以利用文本数据来提取物品的隐藏特征。最近,深度学习算法快速发展,因此文中选用了一种具有强大特征提取能力的新型深度网络架构——变分自编码器(Variational AutoEncoder,VAE)。通过将无监督变分自编码器融合到概率矩阵分解(Probability Matrix Factorization,PMF)中,构建了一种感知上下文的新型推荐模型——变分矩阵分解(Variational AutoEncoder Matrix Factorization,VAEMF)。首先使用 TD-IDF 对物品的评价文档进行数据预处理,然后对处理后的数据使用 VAE 捕获物品的上下文信息特征,最后使用概率矩阵分解进一步提高预测评分精度。在两个真实数据集上的实验结果验证了所提方法相较于自编码算法及概率矩阵分解算法的优势。

关键词 推荐系统,深度学习,变分自编码,矩阵分析

中图分类号 TP391.3 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2019.01.019

Hybrid Recommendation Algorithm Based on Deep Learning

ZENG Xu-yu^{1,2} YANG Yan^{1,2} WANG Shu-ying¹ HE Tai-jun¹ CHEN Jian-bo¹

(School of Information Science and Technology, Southwest Jiaotong University, Chengdu 611756, China)¹

(Key Lab of Cloud Computing and Intelligent Technology, Sichuan Province, Chengdu 611756, China)²

Abstract Recommendation system is playing an increasingly indispensable role in the development of e-commerce, but the sparsity of user's rating data for the items in the recommendation system is often an important reason for the low recommendation accuracy. At present, the recommendation technology is usually used to process the auxiliary information to alleviate the sparsity of the user evaluation and improve the accuracy of the prediction score. Text data can be used to extract the hidden features of the item through related models. In recent years, the deep learning algorithm has developed rapidly. Therefore, this paper chose a variational autoencoder(VAE), which is a new type of network structure with powerful feature extraction capabilities. This paper proposed a novel context-aware recommendation model integrating the unsupervised method VAE into the variable matrix factorization (VAEMF) in the probability matrix factorization (PMF). Firstly, TD-IDF is used to preprocess the evaluation documents of the item. Then, the VAE is utilized to capture the context information features of the item. Finally, the probability matrix factorization is used to improve the accuracy of the prediction score. The experimental results on two real data sets show that this method is superior to the autoencoder and the probability matrix factorization recommendation methods.

Keywords Recommendation system, Deep learning, Variational autoencoder, Matrix factorization

1 引言

随着互联网的飞速发展,信息出现了爆发式增长,人类过去 30 年所产生的信息比之前几千年所产生的信息的总和还要多。在如此海量信息的环境下,用户获取到有效信息变得十分困难。为了解决“信息过载”的问题,推荐系统^[1]应运而生。

推荐系统可以通过分析用户的历史浏览记录、用户-物品评分信息以及用户的偏好等其他信息进行建模,挖掘出用户的一些潜在个性化需求。当前,许多领域广泛使用推荐系统,因此推荐系统具有重要的研究意义与应用价值。

推荐算法经过多年的发展,主要分为基于协同过滤的算法和基于内容的算法^[2]。基于协同过滤^[3]的算法主要依据用

收稿日期:2018-05-07 返修日期:2018-07-20 本文受国家自然科学基金(61572407),国家科技支撑计划(2015BAH19F02)资助。

曾旭禹 男,硕士,主要研究方向为推荐系统;杨 燕 女,博士,教授,CCF 会员,主要研究方向为人工智能、大数据分析 & 挖掘, E-mail:yyang@swjtu.edu.cn(通信作者);王淑莹 女,博士,研究员,主要研究方向为云服务平台构建及动态演化技术;何太军 男,硕士,讲师,主要研究方向为软件工程;陈剑波 男,硕士,讲师,主要研究方向为计算机网络、软件工程。

户过去的历史记录和评级评分为用户进行推荐,而基于内容^[4]的算法主要使用用户个人偏好和内容信息进行推荐。一般情况下,基于协同过滤的算法比基于内容的算法具有更好的效果,主要原因在于前者可以推荐不相关的两个物品,而后者只能推荐相似的物品。基于协同过滤的算法中最受关注的是矩阵分解^[5]的隐因子模型,但是在真实的用户-物品评价矩阵中其往往具有稀疏性,这是导致基于此方法的预测评分精度较差的原因。为了缓解用户-评价矩阵的稀疏性,同时提高准确性,引入辅助信息是近年来推荐系统研究的重点。辅助信息能够从用户资料和物品信息中获取,如社交网络、人的兴趣爱好及物品描述信息等^[6]。为了将这些辅助信息有效整合到矩阵分析中,寻找一种有效的特征提取算法十分必要。

由于深度学习在隐藏特征提取中具有优良效果^[7],许多学者把深度学习引入到推荐系统的隐藏特征学习中。Salakhutdinov 等^[8]直接利用两层受限玻尔兹曼机的无向图模型对用户评分进行建模,然后再进行推荐,但是并没有使用潜在信息。Chiliguano 等^[9]直接使用了卷积神经网络对物品本身进行特征提取,但这种提取物品潜在因素的方法一般只适用于音乐数据集。Kim 等^[10]和王海燕等^[11]先使用卷积神经网络提取相关辅助信息的潜在特征,再结合概率矩阵模型进行推荐,但其中的许多超参数需要人工调整。

针对上述问题,文中提出了一种变分自编码器融合矩阵分解的混合推荐模型,它使用变分自编码器(VAE)提取物品描述文档的辅助信息的特征,进一步提高了预测评分的精度。具体而言,是将 VAE 融合到 PMF 中,而 PMF 是推荐系统常用的隐因子模型,因此将辅助信息和矩阵分解结合起来进行预测评分。

本文的主要贡献如下:

- 1)为了解决推荐系统评价稀疏的问题,提出了一种基于变分自编码的混合推荐模型,该模型能从物品描述文档中提取有效的潜在特征,并获取用户和物品之间的潜在关系。
- 2)利用物品评价描述等辅助信息,将变分自编码整合进概率矩阵分解中。
- 3)使用不同的数据集进行实验,验证了所提出的混合推荐模型能有效提高预测评分的精度。

2 相关工作

2.1 矩阵分解

协同过滤技术中主要分为两类方法^[12]:基于内存的方法(如 KNN^[13]等)和基于模型的方法(如矩阵分解等)。

矩阵分解有很多模型,例如奇异值矩阵分解(SVD)^[14]、贝叶斯概率矩阵分解(BMF)^[15]和非负矩阵分解(NMF)^[16]等。

矩阵分解主要是通过分解用户对物品的评价矩阵,找到用户和物品共存的隐因子空间。假设一个用户物品评价矩阵为 $R_{n \times m}$,包括 N 个用户和 M 个物品,矩阵分解就是将初始评价矩阵 $R_{n \times m}$ 分解为由用户和物品潜在因子组成的两个 k 维低秩矩阵 $U_{n \times k}$ 和 $V_{k \times m}$,使得 $R \approx UV$ 。固定用户和物品的潜在向量 $U_{n \times k}$ 和 $V_{k \times m}$ 之后,通过向量的内积来预测用户对物品的评分。式(1)即为矩阵分解目标函数:

$$\arg \min_{U,V} L(R,UV^T) + \lambda(\|U\|_F^2 + \|V\|_F^2) \quad (1)$$

其中, L 是损失函数,用于评估两个大小矩阵之间的差距;两项正则化参数则用于防止训练过拟合。

2.2 自编码器与变分自编码器

自编码器^[17]也是一种神经网络架构,借助稀疏编码的思想,其目的是使用稀疏的高阶特征进行重新组合来重构自己,常见的形式如图 1 所示。自编码器由一个输入层 X 、输出层 X' 和一个或者多个隐藏层 z 组成,且输出层和输入层需拥有相同的节点,目的在于还原自己的输入,属于无监督学习算法。自编码器经过多年的发展,已有多种类型,如去噪自编码、稀疏自编码、收缩自编码以及变分自编码等。

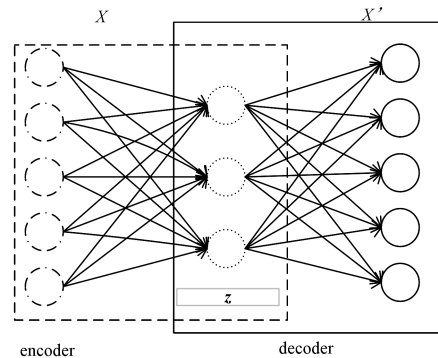


图 1 自编码器网络架构

Fig. 1 Autoencoder network structure

变分自编码作为生成模型的分支,由 Kingma 等于 2013 年提出^[18]。与传统的自编码器相比,变分自编码的核心是变分推理和参数重写。目前,变分自编码主要应用于图像生成,在文档特征提取领域的应用较少。

之前也有一些学者将自编码用于推荐系统,如 Sedhain 等^[19]直接使用普通的自编码器学习电影的评分,并没有使用电影的评价、描述文件等相关的辅助信息。Li^[20]和 Wang 等^[21]均使用了去噪自编码和堆叠去噪自编码对电影进行潜在特征的提取,由于提取的潜在特征直接用于协同过滤计算,因此效果有限。因此,本文使用变分自编码提取物品的潜在特征。

3 基于变分自编码的混合推荐方法

本节从 3 个方面分别介绍混合模型的细节:首先介绍变分矩阵分解(VAEMF)的概率模型;其次描述 PMF 模型和 VAE 模型的融合的关键思想,以便利用辅助信息;然后详细介绍 VAE 的思想,并通过分析辅助信息生成辅助描述文档的潜在特征;最后描述模型的优化。

3.1 变分矩阵分解的概率模型

图 2 给出了 VAEMF 混合模型的概述,其将 VAE 融合到了 PMF 中。首先,使用 VAE 提取文档 X_i 的特征 z ,并将其应用于物品的隐向量 V 中,再由用户的隐向量 U 和物品的隐向量 V 重建出用户-物品的评价矩阵 R 。以下给出具体说明:假定用户-物品的评价矩阵为 $R_{n \times m}$,其中包括 N 个用户和 M 个物品,本文的目标是寻找一个 k 维用户和物品隐向量矩阵,使得 $R \approx UV$,由此重建评价矩阵 R 。从概率的角度来观

察,评分的条件分布如式(2)所示:

$$P(R|U, V, \sigma^2) = \prod_{i=1}^N \prod_{j=1}^M [N(R_{ij} | U_i T_j^T, \sigma^2)]^{I_{ij}} \quad (2)$$

其中, $N(\mu, \sigma^2)$ 是高斯分布的概率密度函数, 均值为 μ , 方差为 σ^2 ; I_{ij} 为指示函数, 若用户 i 评价了物品 j , 则其函数值为 1, 否则为 0。

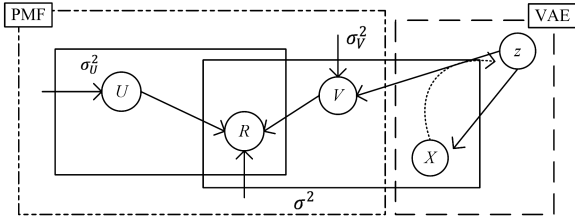


图2 变分矩阵分解架构

Fig.2 Variational matrix factorization architecture

假定用户的隐特征向量是 k 维的, 并且服从高斯分布, 其中 σ^2 是观测噪声的方差, 由人工设定, 即:

$$P(U | \sigma_U^2) = \prod_{i=1}^N N(U_i | 0, \sigma_U^2 I) \quad (3)$$

不同于 PMF 算法中物品的 k 维隐式特征向量由高斯分布生成, 本文假定物品的隐式向量源于 3 个方面: 1) 来自代表描述物品 j 的描述文档 X_j ; 2) 来自 VAE 的隐藏层 z ; 3) 由高斯噪声 ϵ_j 构成, 以此继续优化模型。物品的隐向量生成模型如式(4)所示:

$$V_j = vae(X_j, z) + \epsilon_j \quad (4)$$

其中:

$$\epsilon_j \sim N(0, \sigma_V^2 I) \quad (5)$$

由此, 最后隐向量的条件分布可以被推导为式(6):

$$P(V | z, X_j, \sigma_V^2) = \prod_{j=1}^M N(v_j | vae(X_j, z), \sigma_V^2 I) \quad (6)$$

其中, X_j 代表描述文件中第 j 个物品的信息。从 VAE 模型提取的文档的潜在向量被作为高斯分布的均值, 而该项高斯噪声则作为高斯分布的方差, 它们为实现 VAE 和 PMF 发挥了重要作用。

3.2 变分矩阵分解中的变分自编码器架构

VAE 架构的目标是从物品的辅助信息中提取出潜在向量, 这些辅助信息的潜在向量组成了物品的潜在向量模型。图 3 给出了一个变分自编码器架构。

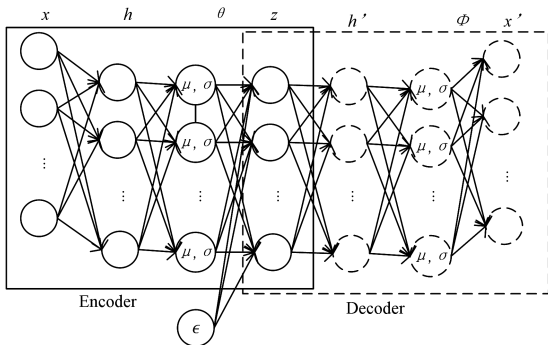


图3 变分自编码器架构

Fig.3 Variational autoencoder structure

可以看出, 变分自编码器架构与图 1 的自编码器架构类似,

均有一个编码器和解码器, 但 z 与自编码不同, 损失函数的定义使得结果趋近于概率分布。其中, x 和 x' 分别代表原始的输入和重建后的输入, h 和 h' 代表隐藏层, θ 和 ϕ 分别代表两个概率分布 $p_\theta(x^{(i)} | z)$ 和 $q_\phi(z | x^{(i)})$, ϵ 是噪声变量, 其维数与 z 一致。

在 Kingma 等^[21] 推导的 VAE 中, 为了对解码器 $p_\theta(x^{(i)} | z)$ 进行参数估计, 使用 $q_\phi(z | x^{(i)})$ 来逼近真实的后验概率 $p_\theta(x^{(i)} | z)$, 并采用 KL 散度来判断两个分布的相似程度。目标公式为:

$$\log p_\theta(x^{(i)}) = KL(q_\phi(z | x^{(i)}), p_\theta(x^{(i)} | z)) + L(\theta, \phi; x^{(i)}) \quad (7)$$

根据上面 VAE 的目标函数, 为了适应文本提取, 需要假定 ϵ 服从高斯分布, 均值为 0。 $p_\theta(x^{(i)} | z)$ 中参数 θ 的均值为 μ' , 方差是 σ'^2 。由于输入的文本向量是实值数据, 因此:

$$\begin{aligned} \mu' &= Wh' + b_1 \\ \sigma'^2 &= Wh' + b_2 \\ h' &= \tanh(Wz + b_3) \end{aligned} \quad (8)$$

其中, h' 已在上文介绍; 而 W 和 b 分别为权重矩阵和偏置量。

同理, $q_\phi(z | x^{(i)})$ 中参数 ϕ 的均值是 μ , 方差是 σ^2 , 其值为:

$$\begin{aligned} \mu &= Wh + b_4 \\ \sigma^2 &= Wh + b_5 \\ h &= \tanh(Wz + b_6) \end{aligned} \quad (9)$$

根据以上假设, 我们获取的目标公式为:

$$\begin{aligned} L(\theta, \phi; x^{(i)}) &\cong \frac{1}{2} \sum_{j=1}^{k_z} (1 + \log(\sigma_j^{(i)})^2 - (\mu_j^{(i)})^2 - (\sigma_j^{(i)})^2) + \\ &\quad \frac{1}{H} \sum_{h=1}^H \log p_\theta(x^{(i)} | z^{(i,h)}) \end{aligned} \quad (10)$$

其中, $z^{(i,h)} = \mu^{(i)} + \sigma^{(i)} + \epsilon^{(h)}$, $\epsilon^{(h)} \sim p(\epsilon)$ 。

3.3 模型优化

为了优化用户潜在特征向量的提取、偏差变量和 VAE 的隐向量, 选择最大后验估计, 如式(11)所示:

$$\max_{U, V} p(U, V | R, X, z, \sigma^2, \sigma_U^2, \sigma_V^2) = \max_{U, V} [p(R | U, V, \sigma^2) p(U | \sigma_U^2) p(V | \sigma_V^2)] \quad (11)$$

对式(11)取负对数, 重写如下:

$$\begin{aligned} L(U, V) &= \sum_i \sum_j \frac{I_{ij}}{2} (r_{ij} - u_i^T v_j)^2 + \frac{\lambda_U}{2} \sum_i \| u_i \|^2 + \\ &\quad \frac{\lambda_V}{2} \sum_j \| v_j - vae(X, z) \|^2 \end{aligned} \quad (12)$$

其中, 正则化参数 λ_U 由 σ^2 / σ_U^2 求取; 同理, λ_V 由 σ^2 / σ_V^2 求取。为了迭代地优化潜变量, 同时拟合其余变量, 我们采用了坐标下降法。具体而言, 式(12)假定 V (或 U) 为常数的同时, 成为关联 U (或 V) 的二次函数, 然后, U (或 V) 的最优解可以通过如下关于 u_i (或 v_j) 简单地优化函数 L 而以闭合形式分析计算:

$$u_j \leftarrow (VI_j V^T + \lambda_U I_k)^{-1} VR_i \quad (13)$$

$$v_j \leftarrow (UI_j U^T + \lambda_V I_k)^{-1} (UR_j + \lambda_V vae(X, z)) \quad (14)$$

其中, I_i 是 I_{ij} 的对角矩阵, $j=1, \dots, M$ 为其对角线元素, 且 R_j 是用户 i 的具有 $(r_{ij})_{j=1}^M$ 的向量。对于物品 j , I_j 和 R_j 分别被定义成 I_i 和 R_i 。式(14)主要展示了 VAE 提取文档潜在向量

在生成物品潜在模型 v_j 中的效果,其中 λ_v 作为平衡的正则化参数。

4 实验与结果分析

本文在第 3 节中提出了一个基于变分自编码的混合推荐模型,并且在本节通过个性化推荐中广泛使用的 MovieLens-1M 和 MovieLens-10M 两个数据集验证了所提模型。

4.1 度量标准

本节采用推荐系统中常用的平均绝对误差(RMSE)来评价本文所提模型在两个数据集上的效果。

$$RMSE = \sqrt{\frac{\sum_{i=1}^S (r_i - \tilde{r}_i)^2}{S}} \quad (15)$$

其中, r_i 是真实的评分, \tilde{r}_i 是预测评分, S 是测试集评分的评分数量。

4.2 数据集

本文将在 MovieLens-1M 和 MovieLens-10M 两个数据集上验证所提模型。这两个数据集的信息如表 1 所列。

表 1 MovieLens-1M 和 MovieLens-10M 数据集的信息

Table 1 Information for MovieLens-1M and MovieLens-10M dataset

Dataset	Users	Items	Ratings	Sparsity/%
MovieLens-1M	6040	3544	993482	95.349
MovieLens-10M	69878	10073	9945875	98.587

在这两个 MovieLens 数据集中用户对电影的评分均为 1~5,但是数据集中不包含关于电影的描述文档,因此我们通过 IMDB 的 api 接口获取了电影的文档。

4.3 对比算法和参数设置

本文选用了 3 种算法来与所提出的 VAEMF 模型进行比较。

1) PMF^[3]: 概率矩阵分解是将用户物品评价矩阵分解为用户因子和物品因子,其中假设用户和物品的隐向量服从高斯分布。

2) CTR^[22]: 协同主题回归模型是一种将主题建模(LDA)和概率矩阵分解(PMF)融合的新型推荐模型。

3) CDL(Collaborative Deep Learning)^[23]: 协同深度学习是另一种使用堆叠去噪自编码提取描述文档特征的模型。

对比方法中,我们使用了比例不同的训练集进行测试。其中训练集随机从数据集中抽取,剩余的数据集作为测试集,通过多次实验取 RMSE 的平均值。对于本文所提混合模型,设置参数 λ_U 和 λ_V 分别为 0.1 和 10,设置混合模型中的 VAE 的隐藏层的层数为 3,并且设置学习用户和物品的隐藏因子维度为 10,设置最大迭代次数为 200,允许误差回退次数的阈值 C 为 5。

4.4 实验和结果

表 2 列出了 PMF,CTR,CDL 和本文提出的混合模型在 MovieLens-1M 和 MovieLens-10M 数据集下多次实验的平均 RMSE。从表中可知,CTR,CDL 和本文的混合模型相比 PMF 有更好的表现,从而证明了使用辅助信息对预测评分的有效性。同时,CDL 和本文模型的效果又优于 CTR,表明神经网络具有良好的特征提取能力。另外,本文混合模型比

CDL 有更好的效果,说明 VAE 对文档信息的特征提取优于 SDAE,也说明 VAEMF 在使用了更多的辅助信息进行训练后能产生更好的预测评分效果。

表 2 模型在不同数据集上的实验结果

Table 2 Experimental results of models on different datasets

Model	MovieLens-1M	MovieLens-10M
PMF	0.8971	0.8311
CTR	0.8969	0.8275
CDL	0.8879	0.8186
VAEMF	0.8544	0.8027

从数据集 MovieLens-10M 中随机抽取 4 个不同比例的训练集和测试集,结果如表 3 所列。从中可以看出,在 4 种不同比例训练集的训练效果下,VAEMF 均优于 3 种对比算法;在训练集从 20% 增加到 80% 的情况下,VAEMF 对比 CDL 的效果一直在提升,也说明 VAEMF 在使用了更多的辅助信息进行训练后能产生更好的预测评分效果。

表 3 模型在 MovieLens-1M 中不同比例训练集下的实验结果

Table 3 Experimental results of models on MovieLens-1M with different ratio training sets

Model	20%	40%	60%	80%
PMF	1.0168	0.9497	0.9197	0.8971
CTR	1.0124	0.9481	0.9194	0.8969
CDL	1.0044	0.9377	0.9068	0.8879
VAEMF	0.9640	0.9087	0.8743	0.8544

4.5 影响推荐精度的主要参数分析

本文主要分析了 4 个参数对 VAEMF 模型性能的影响: D (物品的隐向量维数),变分自编码隐藏层的层数 h , λ_U 和 λ_V 。

图 4 给出了在数据集 MovieLens-1M 中设置不同隐向量特征 k 的维数时,VAEMF 的 RMSE 的变化。随着隐向量维数的增加, RMSE 也在增加,这说明变分自编码在提取辅助信息特征的过程中,在低维时能对特征进行良好的压缩;同时也说明变分自编码应用于稀疏数据的特征提取时,低维特征包含的信息反而可能比高维所包含的信息更有效。

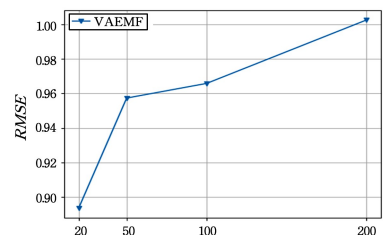


图 4 不同隐向量维数在 MovieLens-1M 数据集上的效果

Fig. 4 Effects of different dimension sizes of latent vector on MovieLens-1M

图 5 显示了参数 λ_U 和 λ_V 在数据集 MovieLens-10M 上对模型的影响。从中可以发现,当 λ_U 的值很大时,用户潜在模型的变化速度缓慢,同时也说明 λ_U 在取较高值时,物品的潜在向量更倾向于投射到用户的潜在向量中,该结论同样适用于 λ_V 。因此,当 λ_U 减小且 λ_V 增加时,其潜在空间主要由提取的辅助信息特征而不是用户评分进行构建,同时用户的潜在向量也被投射到物品的潜在向量中。

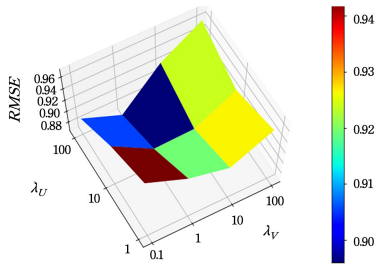


图5 参数 λ_U 和 λ_V 在MovieLens-1M数据集上的分析

Fig.5 Parameter analysis of λ_U and λ_V on MovieLens-1M

结束语 文中提出了一种基于变分自编码的混合推荐方法 VAEMF。该方法融合概率矩阵分解和变分自编码两个模型,构建了一个高效评分预测模型。为了验证 VAEMF 模型的准确性,分别将其与已有的经典推荐模型 PMF,CTR 以及 CDL 进行对比,同时还在不同稀疏性评分数据集上以及不同比例训练集的情况下进行了充分的实验,实验结果验证了本文所提出的 VAEMF 模型的评分预测精度高,推荐效果好。后续研究中我们将对变分自编码的隐藏层网络层结构进行调整,同时也将考虑其他的矩阵分解算法与变分自编码的融合。

参考文献

- [1] YANG B,ZHAO P F. Recommended algorithm review[J]. Journal of Shanxi University(Natural Science Edition),2011,34(3): 337-350. (in Chinese)
杨博,赵鹏飞.推荐算法综述[J].山西大学学报(自然科学版),2011,34(3):337-350.
- [2] SHI Y,LARSON M,HANJALIC A. Collaborative filtering beyond the user-item matrix:A survey of the state of the art and future challenges[J]. ACM Computing Surveys (CSUR),2014,47(1):1-45.
- [3] MNIH A,SALAKHUTDINOV R R. Probabilistic matrix factorization[C]// Advances in Neural Information Processing systems 20(NIPS 2007). 2008:1257-1264.
- [4] LANG K. Newsweeder: Learning to filter netnews [C]// Machine Learning Proceedings. 1995:331-339.
- [5] KOREN Y,BELL R,VOLINSKY C. Matrix factorization techniques for recommender systems[J]. Computer,2009,42(8):30-37.
- [6] SHI Y,LARSON M,HANJALIC A. Collaborative filtering beyond the user-item matrix:A survey of the state of the art and future challenges[J]. ACM Computing Surveys (CSUR),2014,47(1):3.
- [7] SUN Z J,XUE L,XU Y M, et al. Review of deep learning research[J]. Application Research of Computers, 2012, 29(8): 2806-2810. (in Chinese)
孙志军,薛磊,许阳明,等.深度学习研究综述[J].计算机应用研究,2012,29(8):2806-2810.
- [8] SALAKHUTDINOV R,MNIH A,HINTON G. Restricted Boltzmann machines for collaborative filtering[C]// Proceedings of the 24th International Conference on Machine Learning. ACM, 2007:791-798.
- [9] CHILIGUANO P,FAZEKAS G. Hybrid music recommender using content-based and social information[C]// 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE,2016:2618-2622.
- [10] KIM D,PARK C,OH J, et al. Convolutional matrix factorization for document context-aware recommendation[C]// Proceedings of the 10th ACM Conference on Recommender Systems. ACM, 2016:233-240.
- [11] WANG H Y,DONG M W. Latent group recommendation based on dynamic probabilistic matrix factorization model integrated with CNN[J]. Journal of Computer Research and Development, 2017,54(8):1853-1863. (in Chinese)
王海艳,董茂伟.基于动态卷积概率矩阵分解的潜在群组推荐[J].计算机研究与发展,2017,54(8):1853-1863.
- [12] 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 and Data Engineering,2005,17(6):734-749.
- [13] D'ADDIO R M,MANZATO M G. A sentiment-based item description approach for kNN collaborative filtering[C]// Proceedings of the 30th Annual ACM Symposium on Applied Computing. ACM,2015:1060-1065.
- [14] POLAT H,DU W. SVD-based collaborative filtering with privacy[C]// Proceedings of the 2005 ACM Symposium on Applied Computing. ACM,2005:791-795.
- [15] SALAKHUTDINOV R,MNIH A. Bayesian probabilistic matrix factorization using Markov chain Monte Carlo[C]// Proceedings of the 25th International Conference on Machine Learning. ACM,2008:880-887.
- [16] LI J,BIOUCAS-DIAS J M,PLAZA A. Collaborative nonnegative matrix factorization for remotely sensed hyperspectral unmixing[C]// 2012 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). IEEE,2012:3078-3081.
- [17] BALDI P. Autoencoders, unsupervised learning, and deep architectures[C]// Proceedings of ICML Workshop on Unsupervised and Transfer Learning. JMLR. org,2012:37-49.
- [18] KINGMA D P,WELLING M. Auto-encoding variational bayes [J]. arXiv preprint arXiv:1312. 6114,2013.
- [19] 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. ACM,2015: 111-112.
- [20] LI S,KAWALE J,FU Y. Deep collaborative filtering via marginalized denoising auto-encoder[C]// Proceedings of the 24th ACM International on Conference on Information and Knowledge Management. ACM,2015.
- [21] WANG H,SHI X,YEUNG D Y. Relational Stacked Denoising Autoencoder for Tag Recommendation [C] // Twenty-Ninth AAAI Conference on Artificial Intelligence. AAAI Press,2015: 3052-3058.
- [22] WANG C,BLEI D M. Collaborative topic modeling for recommending scientific articles[C]// Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM,2011:448-456.
- [23] WANG H,WANG N,YEUNG D Y. Collaborative deep learning for recommender systems[C]// Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM,2015:1235-1244.