

基于标签卷积神经网络的文本推荐算法^①

马晓焯, 张 谧

(复旦大学 计算机科学与技术学院, 上海 201203)

通讯作者: 马晓焯, E-mail: 15210240042@fudan.edu.cn

摘 要: 在面向用户的文章收集系统中, 用户会将自己喜欢的文章收集起来构成自己的偏好文章集合, 理解用户为何喜欢特定文章、如何精确的找到用户喜欢的文章目前成为了一个重要的研究课题. 本文通过基于面向用户的文章收集系统中的一些相关信息, 如文本信息、标签等, 来辅助推荐系统更好的进行文章的推荐. 文中提出了基于标签卷积神经网络的文本推荐算法, 结合神经网络和协同过滤算法的同时, 将标签加入到神经网络的设计中. 通过在真实的 citeulike 数据集进行的实验和验证, 使用本文的模型可以有效的提高对用户偏好文章预测的准确性.

关键词: 推荐系统; 基于内容; 卷积神经网络; 深度学习; 标签

引用格式: 马晓焯, 张谧. 基于标签卷积神经网络的文本推荐算法. 计算机系统应用, 2018, 27(8): 132-137. <http://www.c-s-a.org.cn/1003-3254/6494.html>

Personalized Academic Article Recommendation with Tagged Convolutional Nets

MA Xiao-Yang, ZHANG Mi

(School of Computer Science, Fudan University, Shanghai 201203, China)

Abstract: In user-oriented article collection system, users construct their collection sets by adding articles which they are interested in, studying why users desire specific articles and find the specific articles is particularly an interesting issue in social science. In this paper, we present the prediction users' preference on articles by considering content information, i.e. semantic information and tags. In this study, we propose a new model which jointly performs deep representation learning for the content information and collaborative filtering for the collection matrix. Extensive experiments on the real-world datasets show that it can significantly advance the state-of-the-art.

Key words: recommender system; content-based; convolutional nets; deep learning; tags

当用户在查找文章或者论文的时候, 他们会更多的考虑和自己兴趣爱好、学术研究等的相关性, 比如在 citeulike 或者 Netflix 网站中, 一个正在学习深度学习的用户, 会更多的查找关于回归、神经网络等相关的文章. 通常, 用户会借助一些搜索引擎工具来帮助查找文章, 但是这仍然需要不少时间. 自动化推荐用户感兴趣的文章, 目前已经成为研究中的热点问题.

在这篇文章中, 我们研究的对象是在一个特定的系统中即面向用户的文章在线收集系统, 比如 citeulike 等. 在这样的系统中, 我们拥有用户、文章和标签等相

关实体. 每个用户都有自己已经收集的文章集合, 同时, 用户和文章共享相同标签集合, 用以表明用户或者文章的相关属性信息.

传统的推荐算法可以大致分为三类: 基于内容的方法^[1]、基于协同过滤^[2,3]的方法以及两者结合^[4,5]的方法. 然而实际情况中, 由于用户和文章之间的交互关系通常来说十分的稀疏, 传统的推荐算法比如协同过滤很难很好的应用^[6,7]. 为了解决这个问题, 通常的做法是通过挖掘更多的辅助信息, 充分利用这些信息来提高算法的效果. 在面向用户的文章收集系统中, 不但拥有

① 收稿时间: 2017-12-30; 修改时间: 2018-01-16; 采用时间: 2018-01-25; csa 在线出版时间: 2018-07-28

交互信息,如用户对于文章的收集等,还有文章中包含的文本信息以及标签信息可以充分利用。

为了做出更好的推荐,学习到一个准确的、完整的特征表示将会十分的有用。而在最近的研究中表明,不论是有监督还是无监督的情况下,深度学习模型都在特征表示学习中展示出了其强大的能力,尤其在计算机视觉^[8,9]和自然语言处理领域^[10]。然而,单纯利用深度学习虽然可以很好的学习到特征的表示,但是无法捕捉到用户和文章之间的交互关系,从而难以形成一个良好的文章推荐系统。综上所述,结合深度学习和协同过滤算法将会是最佳的选择。

尽管已经有很多关于协同过滤结合深度学习模型的工作,并且获得了效果上的提升^[11,12]。但是大部分工作在特征表示的学习过程中都忽视了存在于文章和标签中的具体语义信息。我们知道,不同领域的文章不论是在写法还是词语使用上都有很大的差别,利用同样的神经网络结构并不能很好的捕捉这其中的差别。不同于之前的工作,在这篇文章中,我们设计了一个新颖的推荐算法框架—标签卷积神经网络 (Tagged Convolutional AutoEncoder, TCAE)。在这个框架中,我们对文本信息进行编码以挖掘语义信息,同时利用标签信息对卷积神经网络的卷积核进行分组,使得具有不同标签的文章会使用不同的卷积核进行特征的学习。通过这样的设计,我们可以保证不同的卷积核可以学习到不同的语言提取模式,从而保证特征的准确性。

本文在真实的数据集 citeulike 进行试验证明 TCAE 算法对比其他算法在效果上有着显著的提升。

1 相关背景

1.1 学术文章推荐

推荐系统的本质是利用用户的历史数据,预测用户在特定商品上的偏好程度。目前研究表明,基于内容的推荐算法获得了广泛的应用和成功^[13]。越来越多基于内容的推荐算法被设计出来,核心部分的设计都比较类似,即利用一些算法表示出用户和商品的画像,然后进行比较。然而仅仅基于内容的推荐算法面临一些本质上的问题,比如无法挖掘新类型的用户喜好等。一个良好的解决办法就是结合协同过滤和基于内容的推荐算法^[12],有效的结合用户-商品矩阵和他们的基本信息可以获得更好的推荐效果。

目前推荐系统更多的应用在商品上,比如电影,音乐,生活用品等。最近,推荐系统已经扩展到学术文章的推

荐上来,大多数工作都着眼于引用的推荐^[14]。Sugiyama^[15]则利用引用关系等直接进行学术文章的推荐。

1.2 深度学习在推荐系统中的应用

最近,深度学习方法已经被广泛的应用到推荐系统中,而且很多有很多工作^[11,12,16,17]发表出来。Wang^[12]提出 CDL (Collaboration Deep Learning) 算法,把深度学习和推荐系统有效的结合起来,基于 CDL,可以很好的将辅助信息利用到推荐模型中来,但是一个严重的问题是,该算法利用词袋模型来表示文章,这样是设计忽视了文章中词语和语句的顺序,无法捕捉到词语和词语之间的关系,从而导致信息的缺失。Zhang^[11]提出了 CKE (Collaborative Knowledge based Embedding) 模型,该模型同时对结构化和非结构化的信息进行编码以用来特征提取。同样的,该模型存在语义信息缺失的问题。

同时,标签包含了丰富的信息,高耀东^[18]等则将文本标签作为普通标注,丢失掉其中的语义信息,还有一些工作则大多简单的把标签当做普通文本来处理,利用相同的结构处理具有不同标签集合的文章^[19-21],尽管不同类别的文章共享一些相同的信息或者语言模式,但是每个不同类别同样拥有自己独特的语言模式。Simon^[22]利用对于不同分组的用户分别训练了不同的回归模型,受启发于此,我们设计标签神经网络,利用标签将卷积神经网络中的卷积核进行分组,充分利用标签中包含的语义信息,从而获得更好的推荐效果。

2 问题及符号定义

在本文中,我们推荐任务是为特定的用户推荐其可能喜欢的文章。这该部分,我们首先声明一些符号的定义,然后明确的定义我们任务。

一个面向用户的文章收集系统通常具有文本信息和标签信息,该数据可以被定义为集合 $\{U, A, T\}$,其中 $U = \{u_i\}_{i=1}^{|U|}$, $A = \{a_i\}_{i=1}^{|A|}$, $T = \{t_i\}_{i=1}^{|T|}$ 分别表示用户集合,文章集合和标签集合。

如图 1 所示,用户收集文章构成了用户和文章之间的交互矩阵 $R_{u,a} \in R^{M \times N}$, M 和 N 分别表示用户和文章的总体数目, R 的取值区间 $\{0, 1\}$, 0 表示用户没有收集该文章, 1 表示用户收集了该文章。同时,文章和用户同时有用自己的标签集合,分别用 T_u 、 T_a 表示。文章/用户的内容信息我们利用一个矩阵 $X_{a/u} \in R^{d \times l}$, 该表示我们会在后面的内容详细讲解。

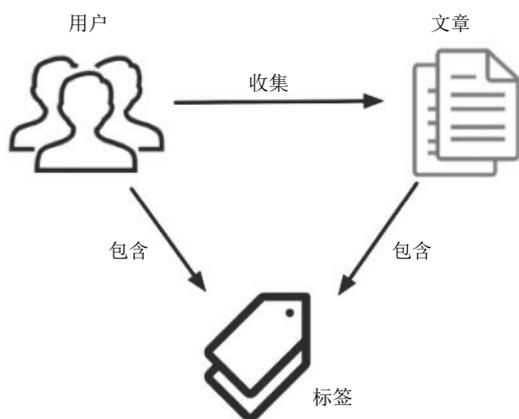


图1 面向用户的文章收集系统

然后本文中我们的主要任务就是在一个面向用户的文章收集系统中, 给定用户 u 的文章集合, 文章的文本信息、及标签信息, 去预测该用户可能喜欢的文章列表.

3 TCAE 模型设计

在该部分, 我们将详细介绍我们我们的模型设计. 首先我们介绍基于标签的卷积神经网络设计用以学习用户和文章的特征表示, 然后介绍和协同过滤结合进行最后的推荐设计.

3.1 用户和文章的特征学习

正如之前我们我们所说的, 我们拥有用户和文章的文本信息 (在这里我们把用户的文章集合作为用户的文本信息), 我们模型的核心就是去学习用户和文章的特征表示. 一个好的特征可以用于多种监督学习的任务比如用户做推荐.

对于文本信息的特征提取工作, 最常用的方法就是自编码器 (AutoEncoder), 自编码器通过重构输入获得对于输入的编码. Wang 等^[12,23,24]采用降噪自编码器用来学习文章的特征表示, 该文章采用词袋模型作为自编码器的输入, 然后利用前馈神经网络构建模型. 然而这样的设计存在一些问题, 首先, 利用词袋模型的输入会损文章中词语之间的顺序信息, 这个对于理解语义十分重要. 而且, 线性的神经网络模型不适合文本信息的特征提取. 最近一些研究表明, 卷积神经网络 (本文选择的方法) 和循环神经网络可以很好的应对文本信息的处理.

为了解决上述问题, 我们改进之前的降噪自编码器. 首先, 我们将文章表示成词语的序列 $\{w_1, w_2, \dots, w_l\}$ 每个 w_i 都是从单词表矩阵 V 中获得, 矩阵 V 我们通过

预训练的 word2vec 模型获得. 所以, 对于每一篇文章, 我们可以构建编码矩阵 $X_{a/u}$ 如下所示:

$$X_{a/u} = \begin{bmatrix} | & | & \dots & | \\ w_1 & w_2 & \dots & w_l \\ | & | & \dots & | \end{bmatrix}$$

其中, w_i 表示每个词的编码形式, 长度为 d , l 表示文本中词语的数目.

然后, 我们将简单的前馈神经网络改进为已经被证实对于处理文本信息具有良好效果的卷积神经网络. 最后, 为了充分利用标签信息, 我们将标签加入到神经网络的设计中去, 我们把其命名为标签卷积自编码 (Tagged Convolutional AutoEncoder, TCAE). 标签卷积自编码也是一个自编码器的变种, 它将标签信息引入作为卷积核设计标准, 从而可以学习到更加精准的特征表示.

如图2所示, 自编码器的本质是通过重构输入获得良好的编码表示, 所以可以分为编码和解码两个部分. 同时为了增加提取特征的鲁棒性, 在输入端加入噪声. 自编码器的目标即优化式 (1):

$$L_{TCAE} = \min_{\{W\}, \{b\}} \|X - X'\|_2^2 + \lambda \|W\|_2^2 \quad (1)$$

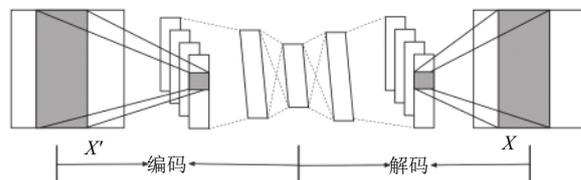


图2 基于标签的卷积自编码器

其中, X' 是加入噪声的版本, X 是真实输入的数据, W 是权重矩阵.

图3展示了我们如何将标签应用到编码器的设计细节. 在卷积层, 卷积操作可以表示为式 (2):

$$f_{m,t,j} = \begin{cases} 0, & \text{otherwise} \\ \sigma(X_{[j:j+m-1]} \times filter_t), & t \in T_u/T_a \end{cases} \quad (2)$$

其中, m 是卷积核宽度, 标签用 t 表示, T_u, T_a 是用户或者文章的标签集合. 激活函数在这里我们选择 ReLU, 定义为 $\max(0, x)$, 保证我们的特征图总是正数. Nair^[25]表明利用 ReLU 激活函数比 sigmoid 和 tanh 效果要好.

从式 (2) 我们可以看到, 我们将每一个卷积核和一个特定的标签 t 绑定, 当用户/文章拥有标签 t 的时候, 该卷积通道则正常进行, 否则该通道输出为 0. 所以我们可以保证每个卷积核可以提取该标签下特定的特征模

式. 当然, 我们仍然保留一些通用的卷积核, 保证一些泛化的特征的提取, 而且可以在标签稀疏的情况下改善实验的结果.

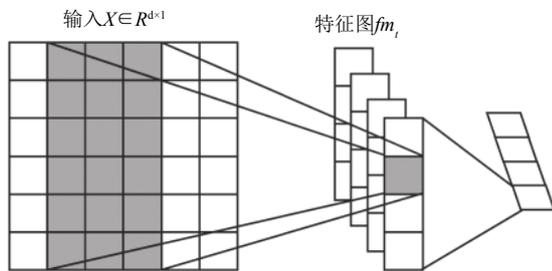


图3 编码过程细节

在卷积层之后, 得到的特征图作为池化层的输入, 实验表明, 选择最大化池化层, 在这里可以获得最好的效果. 最后, 我们加了3层全连接层去捕获各个特征图之间的关系.

在解码过程中, 我们选择反卷积操作, 并且取消上采样操作, 因为 Turchenko^[26]表明, 无论我们采用复制或变量开关的方式, 都会导致很多0变量的产生, 所以我们采用和 Masci^[27]一样的操作:

$$y = \sigma(fm_i \times \tilde{filter}_i + b^i) \quad (3)$$

其中, fm_i 是特征图, b^i 是偏置项, \tilde{filter}_i 是卷积核的逆.

最终标签卷积自编码的中间层 f_{Xu}/f_{Xa} 就是我们学习到的用户和文章的特征表示.

3.2 预测用户对于文章的偏好

通过标签卷积自编码器, 我们可以获得用户和文章的特征表示 f_{Xu}/f_{Xa} , 然后, 我们可以根据以下几步对于特定用户做出预测:

(1) 对于每个用户 u

1) 通过 TCAE, 我们得到用户的特征表示 f_{Xu}

2) 用户的隐藏偏置向量 $\epsilon_u \sim \mathcal{N}(0, \lambda_u^{-1} I_u)$, 则用户 u 的隐藏变量可以表示为:

$$f_u = f_{Xu} + \epsilon_u^T \quad (4)$$

(2) 对于每篇文章 a

1) 通过 TCAE, 我们得到用户的特征表示 f_{Xa}

2) 文章的隐藏偏置向量 $\epsilon_a \sim \mathcal{N}(0, \lambda_a^{-1} I_a)$, 然后文章 a 的隐藏变量可以表示为:

$$f_a = f_{Xa} + \epsilon_a^T \quad (5)$$

(3) 对于每个二元组 $(u, a) \in R$, 偏好预测可以表示为:

$$R_{ua} \sim \mathcal{N}(f_u^T f_a, \sigma^2) \quad (6)$$

参数学习: 我们利用最大后验概率来进行参数的更新和学习.

4 实验结果及分析

4.1 数据集

我们在两个真实的数据集 (citeulike-a, citeulike-t) 上进行了我们的实验. 这两个数据集都来自 citeulike. citeulike 是一个社交化的文章收集系统, 可以帮助用户管理感兴趣的文章. citeulike-a 数据集有约 5551 用户和 16 980 篇文章, 用户和文章的交互数为 210 537, 而 citeulike-t 分别是 7947 用户, 25 975 文章和 134 860 交互数. 可以看出, 两个数据集都非常稀疏, 这使得传统的基于协同过滤算法的推荐系统很难生效. 同时, 我们可以看出, citeulike-t 比 citeulike-a 要更加稀疏, 两个数据集的有效数据分别为 0.07% 和 0.22%, 后面的实验也验证了以上数据. 合理地利用数据集中包含的文章文本信息以及标签信息, 可以帮助我们的推荐系统设计获得更好的效果.

4.2 评价指标介绍

我们选择 80% 数据作为我们的训练集, 把剩下的部分作为测试集. 为了评价推荐算法的效果, 我们选择准确率 (Precision) 和 NDCG. 和所有的推荐算法预测一样, 我们把预测结果进行排序, 并选择最高的 M 个作为我们的推荐结果.

准确率 (Precision): 在评价推荐算法的时候, 准确率是一种十分常见的评价标准, 和大多数推荐算法一样, 我们返回偏好程度最高的 M 篇文章作为我们的推荐结果.

NDCG: NDCG (Normalized Discounted Cumulative Gain) 也是一种常见的评价指标, 不同于准确率, NDCG 更加注重排序的准确性. NDCG 的值是 DCG 和 Ideal CDG 的比值.

$$NDCG_k = \frac{DCG_k}{IDCG_k} \quad DCG_k = \sum_{i=1}^k \frac{2^{rel_i} - 1}{\log_2(i+1)} \quad (7)$$

4.3 对照算法介绍

本文采用了如下算法作为我们实验的对照算法:

UB-CF: 基于用户的协同过滤算法. 通过将于目标用户类似的近邻信息进行聚合, 然后预测出目标用户的喜好.

CMF: 集合矩阵分解是一种矩阵分解模型的扩展^[28], 通过将不同的矩阵信息进行同步分解, 然后进行最后的预测. 在本文中, 这些信息包括了用户-文章矩阵和文章之间的引用矩阵.

DeepMusic: 该模型是一个音乐推荐模型^[29], 利用音乐相关的信息进行基于内容的推荐算法设计.

CDL: 基于神经网络的协同过滤算法^[12], 基本思想是利用自编码器对文章进行编码, 然后集合协同过滤进行推荐算法.

CKE: 基于知识编码的协同过滤算法^[11], 和 CDL 大体类似, 区别在于, CKE 将包括结构化和非结构的信息统一进行编码, 然后结合协同过滤进行预测.

在实验过程中, 我们首先将所有的文章通 word2vec 进行编码, 编码的长度我们设置为 50. 然后我们将这些编码作为已知的知识, 用以后面的模型训练的输入. 对于 UB-CF 对照算法, 我们设置邻居数目为 10, 然后利用近邻信息进行预测. 对于 CMF 算法, 我们设置对于不同上下文的隐藏变量的正则化超参数为 10. 对于 DeepMusic, CDL 和 CKE, 我们采用和发表文章中相同的设置.

4.4 实验结果分析

本文在两个数据集 citeuliek-a 和 citeulike-t 上将我们的模型和对照算法进行了对比.

从图 4 和图 5 中可以看出, 本文提出的 TCAE 算法在两个数据集上均有明显提升. 例如在图 2 中, TCAE 模型在 citeulike-a 上相比于最好的对照实验 CKE 模型有着 16% 以上的提升, 在图 3 中, NDCG 则有超过 20% 的效果提升. 由于 citeulike-t 更加稀疏, 整体效果比 citeulike-a 稍差. 上述实验表面, 有效的利用文本信息可以提高推荐系统的准确性. 相对于表现稍好的对照模型 CDL 和 CKE, 两者都有文本信息结合神经网络的设计, CDL 基于词袋模型的设计不利于文本信息的挖掘, 而 CKE 对所有的文本信息等同对待, 同样不能挖掘更加良好的文本特征表示. 我们设计的 TCAE, 不但可以中分挖掘文本信息中语义知识, 同时因为标签卷积通道的设计, 使得我们可以对不同的文本信息采用不同结构的特征提取模型, 保证了我们最后提取到的特征的准确性, 下一节的内容也充分验证了我们设计具有如此效果.

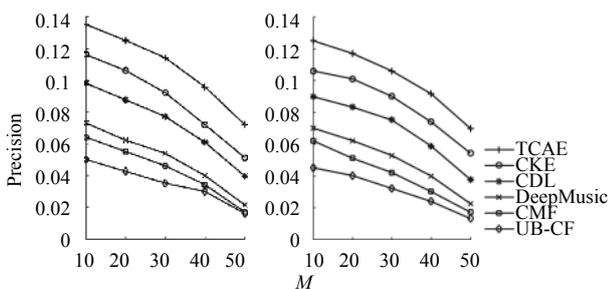


图 4 各模型准确率在 citeulike-a, citeulike-t 上结果

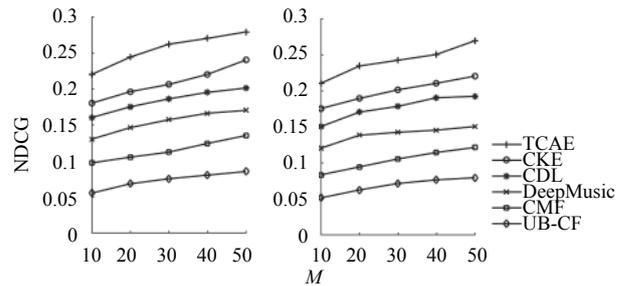


图 5 各模型 NDCG 在 citeulike-a, citeulike-t 上结果

4.5 标签使用效果分析

同时为了验证我们我们的效果提升的确来源于标签的高效使用, 我们同时在 citeulike-a 数据集上做了以下几个实验: 1) 使用普通的卷积神经网络; 2) 使用普通的前馈神经网络分别代替标签卷积神经网络.

从图 6 中可以看出, 使用标签进行卷积神经网络的设计可以有效的提升实验效果, 这也说明标签的合理使用可以帮助我们提取到更加良好的特征, 做出更好的推荐效果. 同时可以看出, 当使用普通的前馈神经网络时, 其效果和 CDL 模型效果类似.

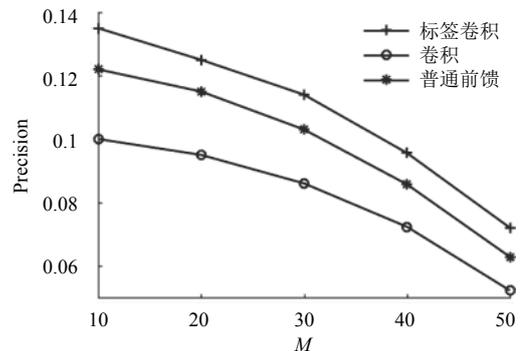


图 6 标签在实验效果中的验证

5 结论

本文主要研究了在面向用户的文章收集系统中对特定用户进行文章偏好的预测从而进行推荐. 文中提出基于标签的卷积神经网络设计, 有效的将系统中的标签信息利用起来, 而且将深度学习和协同过滤进行有效的结合. 同时在真实的数据集 citeulike 上进行了实验验证, 证明了本文提出的算法可以获得更加优秀的推荐结果. 在未来的工作中, 将会继续关注关于学术文章的推荐算法, 以及深度学习在推荐领域的应用, 尤其是结合自然语言处理相关内容, 来进行更加有效的挖掘隐藏在文章中的潜在知识, 设计效果更好的推荐系统算法.

参考文献

- Lang K. Newsweeder: Learning to filter netnews. Proceedings of the 12th International Conference on International Conference on Machine Learning. Tahoe City, CA, USA. 1995. 331–339.
- Rendle S, Freudenthaler C, Gantner Z, *et al.* BPR: Bayesian personalized ranking from implicit feedback. Proceedings of the 25th Conference on Uncertainty in Artificial Intelligence. Montreal, Canada. 2009. 452–461.
- 徐蕾, 杨成, 姜春晓, 等. 协同过滤推荐系统中的用户博弈. 计算机学报, 2016, 39(6): 1176–1189. [doi: [10.11897/SP.J.1016.2016.01176](https://doi.org/10.11897/SP.J.1016.2016.01176)]
- Hu L, Cao J, Xu GD, *et al.* Personalized recommendation via cross-domain triadic factorization. Proceedings of the 22nd International Conference on World Wide Web. Rio de Janeiro, Brazil. 2013. 595–606.
- Li WJ, Yeung DY, Zhang ZH. Generalized latent factor models for social network analysis. Proceedings of the 22nd International Joint Conference on Artificial Intelligence. Barcelona, Spain. 2011. 1705–1710.
- Beierle F, Aizawa A, Beel J. Exploring choice overload in related-article recommendations in digital libraries. arXiv: 1704.00393, 2017.
- Xia F, Liu HF, Lee I, *et al.* Scientific article recommendation: Exploiting common author relations and historical preferences. IEEE Transactions on Big Data, 2016, 2(2): 101–112. [doi: [10.1109/TBDDATA.2016.2555318](https://doi.org/10.1109/TBDDATA.2016.2555318)]
- Liu YF, Qin ZC, Luo, *et al.* Auto-painter: Cartoon image generation from sketch by using conditional generative adversarial networks. arXiv: 1705.01908, 2017.
- Zhang K, Zuo WM, Chen YJ, *et al.* Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising. IEEE Transactions on Image Processing, 2017, 26(7): 3142–3155. [doi: [10.1109/TIP.2017.2662206](https://doi.org/10.1109/TIP.2017.2662206)]
- Rath T. Word and relation embedding for sentence representation [Masters Thesis]. AZ, USA: Arizona State University, 2017.
- Zhang FZ, Yuan NJ, Lian DF, *et al.* Collaborative knowledge base embedding for recommender systems. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco, CA, USA. 2016. 353–362.
- Wang H, Wang NY, Yeung DY. Collaborative deep learning for recommender systems. Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Sydney, Australia. 2015. 1235–1244.
- Chen T, Hong LJ, Shi Y, *et al.* Joint text embedding for personalized content-based recommendation. arXiv: 1706.01084, 2017.
- He Q, Pei J, Kifer D, *et al.* Context-aware citation recommendation. Proceedings of the 19th International Conference on World Wide Web. Raleigh, NC, USA. 2010. 421–430.
- Sugiyama K, Kan MY. Scholarly paper recommendation via user's recent research interests. Proceedings of the 10th Annual Joint Conference on Digital Libraries. Gold Coast, Australia. 2010. 29–38.
- Kim D, Park C, Oh J, *et al.* Convolutional matrix factorization for document context-aware recommendation. Proceedings of the 10th ACM Conference on Recommender Systems. Boston, MA, USA. 2016. 233–240.
- 王涛, 李明. 深度学习下的推荐系统研究. 电脑知识与技术, 2017, 13(25): 180–181.
- 高耀东, 侯凌燕, 杨大利. 基于多标签学习的卷积神经网络的图像标注方法. 计算机应用, 2017, 37(1): 228–232. [doi: [10.11772/j.issn.1001-9081.2017.01.0228](https://doi.org/10.11772/j.issn.1001-9081.2017.01.0228)]
- Chen W, Li ZM. A study of tag-based recipe recommendations for users in different age groups. Proceedings of the 1st International Symposium on Emerging Technologies for Education. Rome, Italy. 2016. 315–325.
- Hu X, Wang WD, Gong XY, *et al.* Social recommendation with tag side information. IEEE International Conference on Data Science in Cyberspace. Changsha, China. 2016. 27–32.
- Rolnick D, Veit A, Belongie S, *et al.* Deep learning is robust to massive label noise. arXiv: 1705.10694, 2017.
- Simon N, Friedman J, Hastie T, *et al.* A sparse-group lasso. Journal of Computational and Graphical Statistics, 2013, 22(2): 231–245. [doi: [10.1080/10618600.2012.681250](https://doi.org/10.1080/10618600.2012.681250)]
- 张敏, 丁弼原, 马为之, 等. 基于深度学习加强的混合推荐方法. 清华大学学报(自然科学版), 2017, 57(10): 1014–1021.
- 周洋, 陈家琪. 基于栈式降噪自编码器的协同过滤算法. 计算机应用研究, 2017, 34(8): 2336–2339. [doi: [10.3969/j.issn.1001-3695.2017.08.023](https://doi.org/10.3969/j.issn.1001-3695.2017.08.023)]
- Nair V, Hinton GE. Rectified linear units improve restricted boltzmann machines. Proceedings of the 27th international Conference on Machine Learning (ICML-10). Haifa, Israel. 2010. 807–814.
- Turchenko V, Chalmers E, Luczak A. A deep convolutional auto-encoder with pooling-unpooling layers in caffè. arXiv: 1701.04949, 2017.
- Masci J, Meier U, Cireşan D, *et al.* Stacked convolutional auto-encoders for hierarchical feature extraction. Proceedings of the 21th International Conference on Artificial Neural Networks. Espoo, Finland. 2011. 52–59.
- Singh AP, Gordon GJ. Relational learning via collective matrix factorization. Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Las Vegas, NV, USA. 2008. 650–658.
- van den Oord A, Dieleman S, Schrauwen B. Deep content-based music recommendation. Proceedings of the 26th International Conference on Neural Information Processing Systems. Lake Tahoe, NV, USA. 2013. 2643–2651.