

DOI: 10.13718/j.cnki.xdzk.2019.11.013

基于深度神经网络的个性化推荐系统研究^①

张敏军¹, 华庆一¹, 贾伟^{1,2}, 陈锐¹, 姬翔¹

1. 西北大学 信息科学与技术学院, 西安 710127; 2. 宁夏大学 新华学院, 银川 750021

摘要: 为了有效地提升海量文献信息检索过程中的用户个性化满足程度, 该文设计了一个全新的个性化推荐系统. 在这个系统中, 核心算法是基于深度神经网络的个性化推荐方法. 此方法构建的神经网络, 包含了嵌入层、编码层、个性化特征融合层、解码层 4 个层次, 从而准确地反应用户的个性化需求并完成查询推荐. 以基于 DNN 网络的文献检索方法、基于 Segnet 网络的文献检索方法、基于 Seq2Seq 网络的文献检索方法为对比算法, 针对计算机、通信、机械、电气、建筑、历史、政治、经济、数学、英语 10 类文献数据进行查询推荐实验, 比较 4 种方法检索结果与用户需求的吻合程度. 实验结果表明: 该文提出的基于深度神经网络的检索方法, 其检索结果的用户个性化需求吻合度高于其他 3 种方法近 10 个百分点, 对于英语类文献检索结果的个性化需求吻合度, 甚至达到了 90.2%, 这充分说明了该文提出的检索方法和构建的个性化推荐系统有效.

关键词: 个性化推荐, 神经网络, 个性化融合, 推荐查询

中图分类号: TP393

文献标志码: A

文章编号: 1673-9868(2019)11-0104-06

随着互联网技术的飞速发展, 人们步入了一个知识信息爆炸式增长的年代^[1]. 每天在互联网上可以查阅到的各种信息浩如烟海, 这固然为需求者提供了丰富的信息资源, 但也让切实需要的信息检索过程变得较为困难^[2]. 以文献数据为例, 2018 年全球文献的录入量突破了 600 万篇. 按照这样的速度继续增长, 学术界可供查询的文献资料总量将达到惊人的程度^[3]. 对于一个急需在同类研究中寻找相关研究成果的研究人员而言, 如何在数量如此庞大的文献资料中找到自己的需求, 这是一个急需解决的难题^[4]. 即便是同类研究, 不同研究人员的切入点不同, 那么对于文献资料的需求也不相同. 各大搜索引擎为不同的研究者准确地提供他们所需要的信息资源, 这就涉及到一个个性化推荐问题^[5]. 所谓个性化推荐, 就是能够准确地判断需求者对于信息资源的需求特征, 并有针对性地为其选取合适的信息^[6]. 在个性化推荐的早期阶段, 一般需要借助特征标注来达到预期的目的, 即对不同的信息资源进行相关特征的标记, 以供后续查询时使用. 随着信息量爆炸式增长, 这种手工标记方法显然无法完成既定任务. 深度学习网络的出现, 使得个性化推荐寻找到一条有效的解决途径^[7]. 通过对各种个性化需求的特征表达、建模和量化, 进而将这些数据代入深度学习网络进行训练和搜索, 最终自动为客户输出个性化的查询结果, 这就是深度学习网络在个性化推荐中的应用过程^[8]. 目前, 深度学习网络在各类信息资源检索和查询领域的应用得到了比较好的验证. 在语音信息识别领域, 基于深度学习网络的识别算法将识别准确率提升了 20%; 在图像信息识别领域, 基于深度学习网络的识别算法将识别准确率提升了 10%^[9-10]. 目前, 已经出现的深度学习网络很多, 如 DNN 网络、Segnet 网络、Seq2Seq 网络等等^[11-13], 本文借助深度学习网络构建个性化推荐算法, 进而设计出一个针对文献查询的个性化推荐系统并通过实验加以验证.

① 收稿日期: 2018-09-29

基金项目: 国家自然科学基金项目(61272286); 高等学校博士学科点专项科研基金项目(20126101110006).

作者简介: 张敏军(1979-), 男, 博士研究生, 主要从事人机交互、个性化推荐及社交网络研究.

通信作者: 华庆一, 博士, 教授, 博士生导师.

1 基于深度神经网络个性化搜索算法设计

在个性化推荐系统中,个性化搜索算法的设计是核心工作.本文基于深度神经网络算法^[14-15]设计了一个4层次的搜索模型,如图1所示.

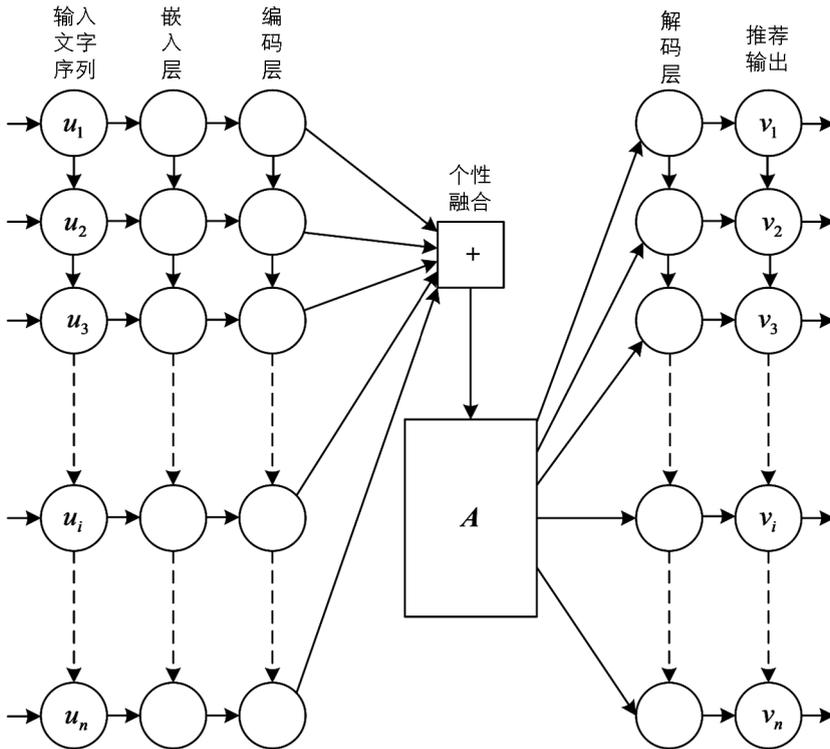


图1 基于深度神经网络的4层次搜索模型

使用的深度学习网络是长短期记忆模型(long-short term memory, 简称 LSTM 网络),这是一种特殊的 RNN(Recurrent Neural Network)模型,是为了解决 RNN 模型梯度弥散问题提出的.在传统的 RNN 中,当时间比较长时需要回传的残差会指数下降,导致网络权重更新缓慢,无法体现出 RNN 长期记忆的效果,需要一个存储单元来存储记忆,因此 LSTM 模型被提出.首先,用户输入的搜索文本信息进入深度神经网络,如 u_i 所示.然后,这些信息进入深度网络的第1个层次即嵌入层,在这里文本信息被嵌入到向量之中.最后,嵌入文本信息的向量经过 LSTM 神经网络处理,被设置成新的编码信息.在接下来的工作中,各个编码信息进入个性融合层,将被融合成个性化向量.个性融合向量再进入解码层,仍然按照 LSTM 神经网络的逆向处理完成解码.这些解码信息将在最大化任务目标的条件概率之下,作为推荐向量输出,具体处理的数学模型为

$$P(V | U; \theta) = \prod_{j=1}^N P(v_j | \{v_1, v_2, \dots, v_{j-1}\}, h; \theta) \quad (1)$$

在上述各个层次的处理过程中,个性化融合向量的生成是一个关键工作.这个融合向量对应于图1中的 A , A 的数学形式为 $A = (a_1, a_2, \dots, a_i, \dots, a_n)$.参数 n 代表了个性融合向量的维度,参数 a_i 代表了第 i 个个性化向量.假设通过之前的训练处理,用户的个性化需求特点已经被系统获悉,并且存储在兴趣向量 \bar{A} 中,实际上接下来的工作就是比较 A 和 \bar{A} 的相似性,具体处理为

$$Sim(A, \bar{A}) = \frac{\sum_{k=1}^n a_k \times \bar{a}_k}{\sqrt{\sum_{k=1}^n a_k^2 \sum_{k=1}^n \bar{a}_k^2}} \quad (2)$$

上述处理的伪代码为

```
Personality_Recommendation()
```

```
{
    Initialization(Parameters) //初始化各参数
    Input(User_Information) //输入用户信息
    LSTM_Embed(Variables) //信息嵌入变量
    LSTM_Encode(Variables) //变量编码
    LSTM_Fuse(Variables) //变量融合
    LSTM_Decompose(Variables) //变量解码
    LSTM_Recommend(User_Information) //个性化推荐
}
```

2 个性化推荐系统设计

以基于深度神经网络的推荐算法为核心, 本文设计了一个 3 层次的个性化推荐系统. 系统的框架结构如图 2 所示.

在这个系统中, 个性化推荐工作的完成需要 3 个层次之间协同配合. 最底层是数据层, 数据层链接到系统可以访问并获得文献的各类文献数据库. 中间层次是算法层, 涉及到文献数据的整理、排序、查询、推荐等一系列功能, 其核心是基于深度神经网络的个性化推荐算法. 这个算法根据本文第 1 节的神经网络处理方案, 将用户的各种个性化需求抽象成特征变量, 进而经过编码融合解码形成推荐方案. 在这一层次中, 基于深度神经网络的个性化推荐算法, 通过用户设定的检索词和文献库联系, 其中检索器和检索库成为完成检索推荐工作的媒介. 如果推荐结果不符合用户的个性化需求, 用户可以对检索关键词进行重新设定, 而算法则可以根据新的关键词或者新的关键词顺序, 重新执行个性化检索和推荐. 整个推荐系统的最上层是用户层, 系统为用户提供检索接口. 通过这个接口, 用户可以对个性化系统进行访问, 提供表征自己个性化需求的词语, 并获得来自推荐系统的个性化推荐结果.

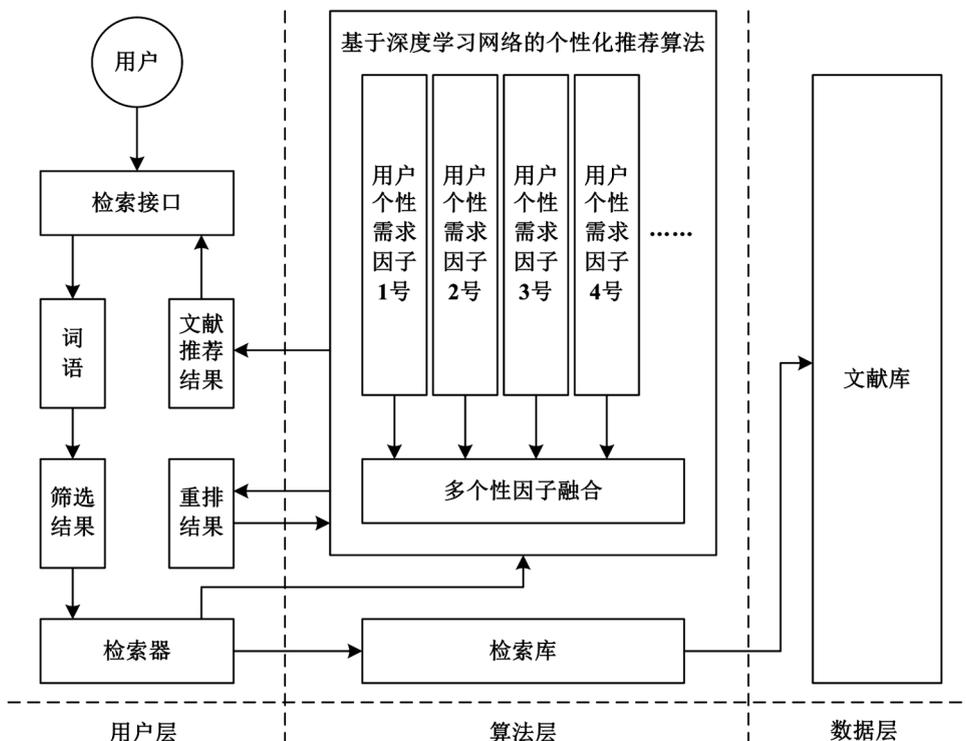


图 2 本文设计的个性化推荐系统框架结构

3 实验与实验分析

为了验证本文构建的个性化推荐系统以及基于深度神经网络个性化推荐方法的有效性,将文献查询的个性化推荐作为研究对象,展开具体的实验研究。

3.1 实验条件

在实验中选用的计算机为 HUAWEI MATEBOOK 笔记本电脑,处理器为 INTEL I7 型号,内存 8 GB,硬盘 512 GB,独立显存 2 GB。

计算机自带正版操作系统 Windows 10,算法及系统设计语言为 Java Web。

由于本文研究尚处于实验室阶段,无法与大型的文献数据库引擎进行对接,因此在各大图书资源和科技资料检索网站上下载并自行构建了文献数据库,以便于本文的个性化推荐系统访问和使用。数据库中共包含了计算机、通信、机械、电气、建筑、历史、政治、经济、数学、英语方面共 10 个类型的文献,每类文献的数量为 100 篇,总计数据库中有文献 1 000 篇。在下载文献和构建数据库的过程中,本文为了避免个性化推荐的验证效果不明显,对每类文献的 100 篇进行了差异化考量,尽可能选择那些有一定差异、能体现同类研究不同特征的文献。

为了形成与本文方法的对比,还选择了 3 类参考方法:基于 DNN 网络的文献检索方法、基于 Segnet 网络的文献检索方法、基于 Seq2Seq 网络的文献检索方法。

3.2 实验结果

为了体现 4 种方法对文献检索个性化需求的满足程度,采用了个性化需求吻合度这一指标。假设系统为用户推荐了 100 篇文献,如果其中 50 篇都符合用户的个性化需求,那么此次推荐结果的个性化需求吻合度就是 50%。

经过 10 类文献的个性化推荐测试,4 类方法获得的用户个性化需求吻合度如表 1 所示。

表 1 4 类方法的用户个性化需求吻合度

%

文献类型	个性化需求吻合度			
	DNN 方法	Segnet 方法	Seq2Seq 方法	本文方法
计算机文献	72.2	61.7	76.2	82.3
通信文献	67.9	66.5	70.4	77.5
机械文献	73.4	71.7	79.6	88.6
电气文献	52.3	55.1	68.9	81.3
建筑文献	59.2	53.4	65.1	74.9
历史文献	63.8	66.7	74.1	83.5
政治文献	47.6	44.4	50.7	65.1
经济文献	45.5	42.3	51.2	65.7
数学文献	51.8	50.2	59.8	68.4
英语文献	76.1	77.3	81.9	90.2

为了直观地显示几种文献检索方法对用户的个性化需求满足程度,将表 1 中的数据绘制成如下的柱状图,如图 3 所示。

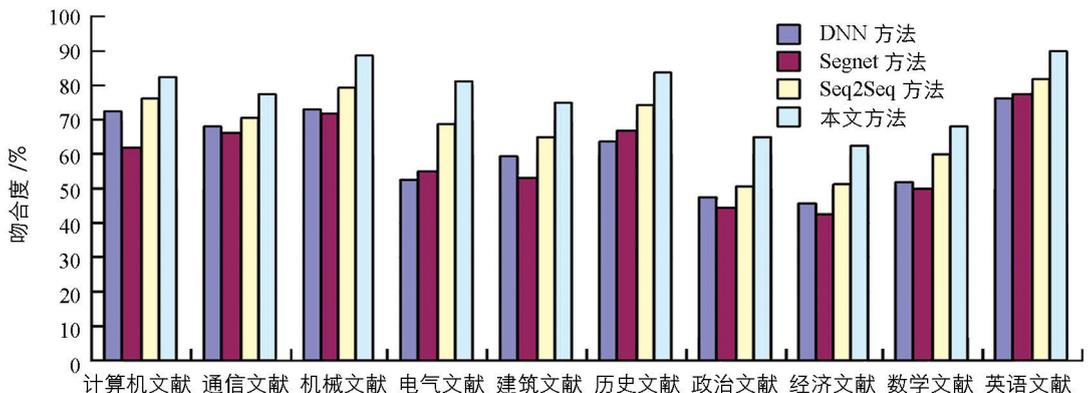


图 3 4 种方法的用户个性化需求吻合度对比

从表 1 和图 3 中的结果可以明显看出:

1) 基于 DNN 网络的文献检索方法,对于计算机、通信、机械、电气、建筑、历史、政治、经济、数学、英语 10 类文献的检索,其个性化需求吻合程度分布在 45.5%~76.1%之间,整体吻合程度不是特别理想.这种方法对于英语类文献的检索效果最好,但吻合程度也只达到了 76.1%

2) 基于 Segnet 网络的文献检索方法,对于计算机、通信、机械、电气、建筑、历史、政治、经济、数学、英语 10 类文献的检索,其个性化需求吻合程度分布在 50.2%~77.3%之间,整体吻合程度也不是特别理想.这种方法对于英语类文献的检索效果最好,但吻合程度也只达到了 77.3%.

3) 基于 Seq2Seq 网络的文献检索方法,对于计算机、通信、机械、电气、建筑、历史、政治、经济、数学、英语 10 类文献的检索,其个性化需求吻合程度分布在 50.7%~81.9%之间,整体吻合程度尚可.这种方法也是对于英语类文献的检索效果最好,吻合程度达到了 81.9%.

4) 本文提出的基于深度神经网络的文献检索方法,对于计算机、通信、机械、电气、建筑、历史、政治、经济、数学、英语 10 类文献的检索,其个性化需求吻合程度分布在 65.1%~90.2%之间,整体吻合度高,超出其他 3 类方法近 10 个百分点.尤其是对于英语类文献,吻合度更是达到了 90.2%.

5) 上述比较结果充分表明,本文提出的基于深度神经网络的文献检索方法,其检索性能明显优于其他 3 种方法,达到了更高的个性化需求吻合程度,更加准确地把握了用户的个性化需求,根据该方法设计的个性化推荐系统有效.

4 结 语

互联网技术、信息技术的飞速发展,使得用户可供选择的信息资源爆炸式增长,从而给用户准确的信息搜索带来了困难.本文基于深度学习理论,构建了一个全新的个性化推荐算法,在这个算法中共设置了 4 个层次,分别是嵌入层、编码层、个性化特征融合层、解码层,从而将用户的个性化需求准确地反映到算法中,最终为用户提供更加符合要求的推荐.

以基于深度神经网络的文献检索算法为核心,构建了一个针对文献检索的个性化推荐系统,其设计目标是充分满足用户的个性化检索需求,提供更好的用户检索服务.

为了验证本文提出的算法和个性化推荐系统设计的有效性,课题组进一步展开了实验研究.在实验中选取了基于 DNN 网络的文献检索方法、基于 Segnet 网络的文献检索方法、基于 Seq2Seq 网络的文献检索方法为对比算法,对它们检索结果的个性化吻合程度进行了比较.

实验结果表明:针对计算机、通信、机械、电气、建筑、历史、政治、经济、数学、英语等 10 类文献的检索,本文提出的基于深度神经网络的个性化推荐方法都明显优于其他 3 种方法,对于不同文献检索后的个性化吻合程度,基本超出其他 3 种方法 10 个百分点,对于英语类文献的检索,其个性化需求吻合程度甚至达到了 90.2%.这样的实验结果,充分证实了本文提出的基于深度神经网络文献的检索方法和基于此方法构建的个性化推荐系统的有效性.

参考文献:

- [1] WANG C, ZHANG Y L, GU M L. Target Tracking and Classification Algorithm for Adaptive Cruise Control System Via Internet Technology [J]. *Wireless Personal Communications*, 2018, 102(2): 1307-1326.
- [2] 王莉军. 海量数据下的文本信息检索算法仿真分析 [J]. *计算机仿真*, 2016, 33(4): 429-432.
- [3] 欧石燕, 唐振贵, 苏翡翠. 面向信息检索的术语服务构建与应用研究 [J]. *中国图书馆学报*, 2016, 42(2): 32-51.
- [4] TAJEDDINE R, GNILKE O W, EL ROUAYHEB S. Private Information Retrieval from MDS Coded Data in Distributed Storage Systems [J]. *IEEE Transactions on Information Theory*, 2018, 64(11): 7081-7093.
- [5] 叶卫根, 宋 威. 融合信任用户间接影响的个性化推荐算法 [J]. *计算机工程与科学*, 2016, 38(12): 2579-2586.
- [6] 李桃迎, 李 墨, 李鹏辉. 基于加权 Slope One 的协同过滤个性化推荐算法 [J]. *计算机应用研究*, 2017, 34(8): 2264-2268.
- [7] ZHANG L P, ZHANG L F, DU B. Deep Learning for Remote Sensing Data: A Technical Tutorial on the State of the Art [J]. *IEEE Geoscience and Remote Sensing Magazine*, 2016, 4(2): 22-40.

- [8] GREENSPAN H, VAN GINNEKEN B, SUMMERS R M. Guest Editorial Deep Learning in Medical Imaging: Overview and Future Promise of an Exciting New Technique [J]. *IEEE Transactions on Medical Imaging*, 2016, 35(5): 1153-1159.
- [9] 刘文举, 聂 帅, 梁 山, 等. 基于深度学习语音分离技术的研究现状与进展 [J]. *自动化学报*, 2016, 42(6): 819-833.
- [10] 胡二雷, 冯 瑞. 基于深度学习的图像检索系统 [J]. *计算机系统应用*, 2017, 26(3): 8-19.
- [11] HAMANAKA M, TANEISHI K, IWATA H, et al. CGBVS-DNN: Prediction of Compound-protein Interactions Based on Deep Learning [J]. *Molecular Informatics*, 2017, 36(1/2): 1600045.
- [12] O'BYRNE M, PAKRASHI V, SCHOEFS F, et al. Semantic Segmentation of Underwater Imagery Using Deep Networks Trained on Synthetic Imagery [J]. *Journal of Marine Science and Engineering*, 2018, 6(3): 93-101.
- [13] BADRINARAYANAN V, KENDALL A, CIPOLLA R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 23(12): 2481-2495.
- [14] QAWAQNEH Z, MALLOUH A A, BARKANA B D. Deep Neural Network Framework and Transformed MFCCs for Speaker's Age and Gender Classification [J]. *Knowledge-Based Systems*, 2017, 115: 5-14.
- [15] SHUAI L, SONG W, HONG Q, et al. Deep Variance Network: An Iterative, Improved CNN Framework for Unbalanced Training Datasets [J]. *Pattern Recognition*, 2018, 8(1): 242-250.

A Personalized Recommendation System Based on Deep Neural Network

ZHANG Min-jun¹, HUA Qing-yi¹, JIA Wei^{1,2},
CHEN Rui¹, JI Xiang¹

1. College of Information Science and technology, Xibei University, Xi'an 710127, China;

2. Xinhua College, Ningxia University, Yinchuan 750021, China

Abstract: In order to effectively improve the personalized satisfaction of users in the process of massive document information retrieval, this paper designs a new personalized recommendation system, whose core algorithm is the personalized recommendation method based on deep neural network. The deep neural network constructed with this method includes four layers: an embedding layer, a coding layer, a personalized feature fusion layer and a decoding layer, and can accurately reflect the user's personalized needs and complete query recommendation. In an experiment of query and recommendation with the literature data of computer, communication, machinery, electricity, architecture, history, politics, economy, mathematics and English, this method is compared with another three literature retrieval methods based on DNN, Segnet or Seq2Seq networks for the similarity between their retrieval results and users' needs. The results show that the matching degree of users' individualized needs of the proposed retrieval method is about 10% higher than that of the other three methods, its matching degree of individualized needs of English literature retrieval results being as high as 90.2%, which convincingly indicates that the retrieval method and the personalized recommendation system described in this paper is effective.

Key words: personalized recommendation; deep neural network; personalized fusion; recommendation query