情报科学
第38卷 第3期 2020年3月

基于深度学习的信息组织与检索研究综述

熊回香,杨梦婷,李玉媛

(华中师范大学 信息管理学院,湖北 武汉 430223)

摘 要:【目的/意义】深度学习是近几年来人工智能领域的研究热点之一,了解深度学习在信息组织与检索方面的研究现状,能为信息组织与检索的深入研究提供参考和借鉴。【方法/内容】通过对国内基于深度学习的信息组织与检索方向的相关文献进行梳理,剖析深度学习相关模型、阐述深度学习在信息组织与检索中的研究热点主题,并结合深度学习技术的特点和信息组织与检索的研究内容,对深度学习在信息组织与检索方向的应用前景进行预测。【结果/结论】研究表明,当前深度学习在信息组织与检索中的研究热点主要集中在智能信息抽取、自动文本分类、情感分析和文本聚类这四个主题,预测未来深度学习在信息组织与检索方向会朝着对异构信息处理、智能信息检索、个性化信息推荐等方向发展。

关键词:深度学习;神经网络;信息组织与检索

中图分类号:G250.2 **DOI**:10.13833/j.issn.1007-7634.2020.03.001

A Survey of Information Organization and Retrieval Based on Deep Learning

XIONG Hui-xiang, YANG Meng-ting, LI Yu-yuan

(School of Information Management, Central China Normal University, Wuhan 430223, *China)*

Abstract:【Purpose/significance】Deep learning is one of the research hotspots in the field of artificial intelligence in recent years. Understanding the research status of deep learning in information organization and retrieval can provide theoretical reference and new study perspective for the in-depth study of information organization and retrieval.【Method/process】Through analyzing relevant literature on information organization and retrieval based on deep learning in China, this paper analyzes some deep learning models, and then elaborates on the hot topics of deep learning in information organization and retrieval. Combined with the characteristics of deep learning and the research content of information organization and retrieval, it forecasts the application prospects of deep learning in the future.【Result/conclusion】Research shows that the current research topics mainly focuses on intelligent information extraction, automatic text classification, sentiment analysis and text clustering. It is predicted that deep learning will study deeply towards heterogeneous information processing, intelligent information retrieval and personalized information recommendation.

Keywords: deep learning ; neural networks information; organization and retrieval

1 引 言

随着大数据、云计算和物联网等互联网信息技术的迅猛发展,网络数据规模呈爆发式增长,网络信息发布也具有很大的随意性,信息的纯净度和可信度大大降低,信息质量难以得到保障[1]。在海量繁杂的数据信息面前,用户的信息查询变得十分困难,如何快速高效地从繁杂的数据中获取所需信息成为大数据发展的一大难题。高效的信息检索依赖于有序的信息组织,只有对网络信息资源进行有效地组织和管理,才能实现网络资源利用最大化[2]。传统的信息组织与检索方法依赖于人工设计特征,无法根据数据深层语义特征对信息进行组织,也不能处理多源异构数据信息,无法满足用户高层次多元化的信息需求。2006年Hinton提出深度学习

收稿日期:2019-05-17

基金项目:国家社会科学基金年度项目"融合知识图谱和深度学习的在线学术资源挖掘与推荐研究"(19BTQ005);中央高校基本科研业务费重大培育项目"基于语义网的在线健康信息的挖掘与推荐研究"(CCNU19Z02004)

作者简介:熊回香(1966-),女,湖北武汉人,博士,教授,博士生导师,主要从事网络信息组织与检索研究.

概念[3]，通过多层神经网络将低维特征组合成更加抽象的高层特征或属性类别，来发现数据的分布式表示[4]。目前，国内外众多学者已经对深度学习做了很多相关研究，主要集中在语音识别[5-6]、图像识别[7-8]、自然语言处理[9]和计算机视觉[10]等领域并取得了不错的成果。深度学习为信息组织与检索研究提供了新的契机，一方面深度学习技术可以通过深层非线性网络结构从海量样本中自动学习数据的深层语义特征；另一方面深度学习能够将多源异构数据映射到同一隐空间来学习数据的统一特征表示。基于深度学习的信息组织与检索研究已经成为图书情报领域的研究热点之一。

本文尝试对国内近年来基于深度学习的信息组织与检索研究的研究成果进行梳理，主要从深度学习相关模型、深度学习在信息组织与检索方向的研究热点以及未来发展趋势这三个方面进行论述，在总结分析深度学习在信息组织与检索方向的研究现状的同时，为未来深度学习在信息组织与检索的深度研究提供参考和借鉴。

2 深度学习相关模型

深度学习是近几年来机器学习领域的新起之秀，已经成为学术界的一个研究热潮。深度学习能够通过多层神经网络将低维特征组合成更加抽象的高层语义特征，从而自动学习数据的分布式表示，解决了传统机器学习中浅层模型无法提取数据深层特征和需要人工设计特征等问题。目前在信息组织与检索研究中应用较多的深度学习模型主要是神经网络语言模型、深度置信网络模型、卷积网络模型和递归神经网络模型。

2.1 神经网络语言模型

语言模型就是指根据一定的训练集按照某种算法来计算任意词序在文本中出现的概率的统计模型。基于神经网络生成的语言模型就称为神经网络语言模型(Neutral Network Language Model，NNLM)。Wei 等人[11]最早提出利用人工神经网络来训练语言模型，之后 Bengio[12]对其进行深入研究，提出一个简单的三层神经网络语言模型。如图 1 所示，该模型是仅有一层隐藏网络，隐藏层使用 tanh 函数(双曲正切)进行特征学习，在输出层使用 softmax 函数计算词序的分布概率，同时使用随机梯度上升法(SGA)训练模型参数。

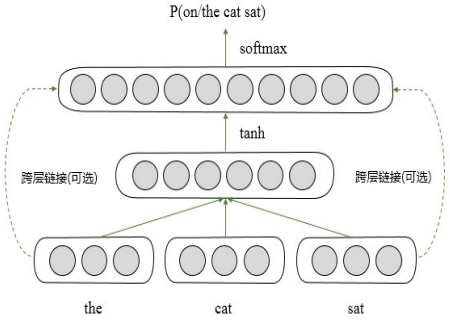

图 1 基于神经网络的语言模型

在 Bengio 的研究基础上，Mnih 等人[13]结合受限玻尔兹曼机(RBM)，提出一个分层分布式神经网络语言模型 HLBL 来自动从语料中学习出词的层次结构。从 HLBL 模型开始，研究 NNML 主要目的是为了生成神经网络模型的输入——词嵌入[14]。

词嵌入(Word embedding)，又称词向量，是通过训练神经网络语言模型生成的一种词的分布式特征表示。词的分布式表示(Distributed Representation)是基于 Harris 的分布假说[15] (Distributional Hypothesis):“上下文相似的词，其语义也相似”，将词的语义信息转化成稠密的低维向量(通常是 50-100 维)。目前已经公开且知名度较高的生成词向量的模型有六种：HLBL、SENNA、Huang’s、Turian’s、Glove 和 Word2vec[14]。其中在信息组织与检索领域基于深度学习的研究中，主要使用 Mikolov 等人[16]训练出来的 Word2vec 模型所生成的词向量作为神经网络的输入。郑文超等人[17]提出利用 Word2vec 将处理过后的中文文本转换成词向量，计算词之间的相似度，以完成中文词的聚类任务；曾颖黎[18]对网络舆情短文本使用 Word2vec 进行特征拓展，以降低短文本数据稀疏，实现网络舆情文本分类；来斯惟[19]从模型、语料和训练参数三个维度对不同的词向量训练方法进行比较，实验发现语料规模越大，词向量的语义表示能力越强；张晓娟等人[20]利用词向量获得查询词的上下文信息构建语义模型，计算词之间的语义相似度来重构查询式，生成符合用户偏好的候选查询推荐，发现基于词向量的查询式推荐要优于已有的相关方法。

2.2 深度置信网络模型

深度置信网络(Deep Belief Network，DBN)由 Hinton 等人[21]在 2006 年提出，它是由多层受限玻尔兹曼机层(Restricted Boltzmann Machine,RBM)[22]和一层反向传播神经网络(Back Propagation，BP)组合而成的一种概率生成式模型。DBN 可以看成是多层 RBM 单元的堆叠，并由低到高逐层训练。由图 2 可以看出，在 DBN 结构中，单个层内单元之间不连接，相邻层间单元全连接，前一个 RBM 的可见层作为下一个 RBM 的隐层。DBN 通过训练其神经元之间的权重来让整个神经网络模型能按照最大概率生成训练数据，其训练过程使用的是贪婪逐层算法[23]。首先以最底层的 RBM 作为可见层，通过对比散度算法[24]对原始数据进行特征学习，然后将学习到特征信息作为到下一层 RBM 的输入，通过这种方式反复训练，直到整个深度置信网络模型训练完成。不恰当的初始化权重会影响整个模型的性能，而 DBN 模型所预训练的初始化权重比随机权重更接近最优权重，这不仅提升了最终训练出来的神经网络模型的性能，还提高了调优阶段的收敛速度。Bengio[25]验证并扩展了 DBN 网络模型，提出使用自动编码器(auto-encoder, AE)来替换传统 DBN 中的每一层 RBM，通过堆叠多个 AE 来构建堆叠自编码网络模型(stacked auto-encoders,SAE)，但是实验效果并不理想。Vincent 等人[26]改进 SAE 模型，提出堆叠降噪自编码器，将噪音

加入到SAE模型的编码过程中，不仅提升了SAE模型的抗噪性，其特征学习能力还优于传统的DBN。同时，针对DBN难以图像多维结构信息的问题，Lee等人[27]改进了深度置信网络，提出卷积深度置信神经网络模型，并在视觉识别和音频识别中取得不错效果。

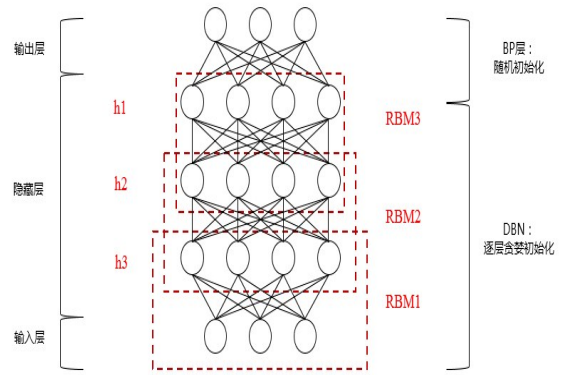

图2 深度置信网络模型

2.3 卷积神经网络模型

卷积神经网络(Convolutional Neural Network,CNN)最早由Yann Lecun等人[28]提出并成功应用于手写数字识别。CNN实际上是一种多层感知机模型，主要用于处理二维图像数据，在语音识别、人脸识别、图像识别和自然语言处理等领域取得重大突破。CNN的结构如图3所示，主要由输入层、卷积层、池化层、全连接层和输出层构成。其中卷积层是CNN的核心，在卷积层中的每个神经元只连接部分邻层神经元，通过卷积核提取图片数据特征，接着使用池化操作降低数据维度，卷积层和池化层配合使用，组成多个卷积组来逐层提取二维图像的特征信息。此外，与传统的感知机相比，CNN通过卷积核生成的权值共享减少了网络模型中各层之间的连接，降低了模型中的自由参数的数量级，极大地简化了模型的复杂度。

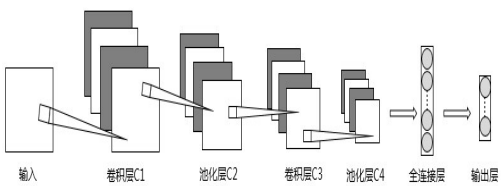

图3 卷积神经网络模型

2012年Krizhevsky等人[29]基于深度的CNN提出的Alexnet模型，并在ImageNet大规模视觉识别挑战赛(ImageNet Large Scale Visual Recognition Challenge, LSVRC)夺得冠军，这让CNN模型名声大噪，不仅奠定其在计算机视觉领域的地位，还使得越来越多的研究者重视和扩展CNN模型。例如Google公司推出的人脸识别系统FaceNet[30]就使用卷积神经网络直接优化嵌入本身；Gong等人[31]将注意力机制引入卷积神经网络(CNN)来优化微博中Hashtag的推荐。

2.4 递归神经网络模型

循环神经网络(RNN)是时间递归神经网络(Recurrent Neural Network，RNN)和结构递归神经网络(Recursive Neural Network，RNN)这两种人工神经网络的总称[32]。时间递归神经网络的各神经元之间连接构成有向图(图4)，主要用于处理序列化数据，而结构递归神经网络利用树型神经网络递归构造更加复杂的深度神经网络，主要用于处理树结构、图结构等复杂的信息结构。信息组织与检索方向最为常用的是时间递归神经网络，所以下文的RNN均指时间递归神经网络。深度置信网络或卷积神经网络模型都是由输入层、隐藏层和输出层构成，层与层之间的神经元是全连接，而隐藏层内的神经元相互独立，这种神经网络模型难以处理序列化数据。而RNN模型最大的特点在于神经网络各隐藏层内的神经元是具有连接的，它通过获取输入层和前一个神经元的输出来计算当前时刻隐层的输出，也就是说RNN能够对前面的信息状态进行记忆并用于当前的计算[33]。理论上来讲RNN能够处理任意长度的序列数据，但是在实际应用中为了降低神经网络模型的复杂度，往往假设当前的状态仅仅与前面几个时刻的状态相关。

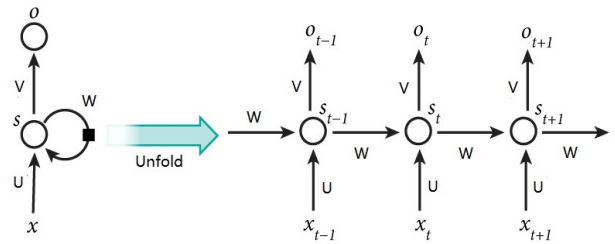

图4 时间递归神经网络模型

同时，普通的RNN结构存在着梯度消失和梯度爆炸的问题，无法解决学习序列化数据之间的长程依赖关系。针对这一难题，研究者们相继推出了一些改良版的RNN模型。如Hochreiter等人[34]提出的长短时记忆网络(Long Short-Term Memory，LSTM)和Cho等人[35]提出来的门控循环单元(Gated Recurrent Unit，GRU)，通过增加门单元来控制信息输出，能有效地学习序列数据之间的长程依赖关系。

通过上述模型的剖析可以看出，深度学习的深层非线性网络结构能够从海量繁杂的数据中自动学习数据信息的隐特征，并根据其语义特征对信息进行规范化处理。而信息组织与检索主要目的是将所搜集到的信息按其形式特征和内容特征有序化，进行重新组织和存储，以便用户查找符合其信息需求的信息资源。深度学习的特点与信息组织与检索的需求不谋而合，将深度学习模型引入信息组织过程中，可以使信息处理不再局限于简单的形式特征和内容特征，而是更高层次的语义特征，这能让后面的信息检索更加精准化，信息服务更加个性化。同时深度学习模型的自动特征学习还简化了信息组织的工作流程，提高了信息组织的工作效率。因此，深度学习模型为智能化的信息组织与检索提供了强有力的技术支持。

3 深度学习在信息组织与检索中的研究主题分析

深度学习因其多源异构数据的自动特征提取和高层次的隐藏语义特征表示等优点，成为人工智能和机器学习领域发展最为迅速的研究方向，受到了工业界和学术界的高度关注。学术界中基于深度学习的研究主要集中在计算机领域和工程控制领域，其在信息组织与检索中的应用研究还处于起步阶段，主要集中在智能信息抽取、自动文本分类、情感分析、文本聚类和智能信息检索这五个主题。

3.1 基于深度学习的智能信息抽取研究

智能信息抽取(Information Extraction, IE)是从自然语言文本中自动抽取出特定的实体、关系、事件和事实信息，并以结构化数据形式保存到数据库中以便用户查询和利用的文本处理技术[36]。智能信息抽取主要包括命名实体识别、实体消歧和实体关系抽取，深度学习模型在信息抽取方面均有较好的表现。

(1)命名实体识别。命名实体识别是指识别自然语言文本中人名、地名、机构名、时间表达式和数字表达式等具有特定意义的实体。朱丹浩等人[37]利用双向长短时记忆(Bi-LSTM)对中文机构进行命名实体研究，提出以字特征为输入的汉字级别循环神经标注模型，发现汉字级别的中文机构名标注模型在机构名识别能力上有明显提高；朱娜娜等人[38]研究深度神经网络(DNN)对微博短文本中的图书名识别的效果，将微博文本词序列化成语义词向量，再使用深度学习模型自动学习图书名特征表示以完成命名实体识别任务，研究结果发现该方法不仅能完成微博文本中的图书名识别任务，还能准确识别书名号中的图书名和非图书名。

(2)实体消歧。实体消歧是指根据上下文将有歧义的实体正确地映射到知识库中相应实体上，解决同名实体产生歧义的过程[39]。比如根据上下文语义，识别“苹果”指代的是一种水果还是某一手机品牌。Zhou等人[40]融合群体认知和深度学习技术，提出一个“众包特征”的命名实体消歧模型，通过动态卷积神经网络(DCNN)训练群体用户标签和来实现命名实体消歧；秦越[41]分别选取双向长短时记忆网络(Bi-LSTM)、堆叠降噪自动编码机(SDAE)和深度置信网络(DBN)来研究维吾尔语中的人称代词消歧，根据实体之间的深层语义关联来构建消歧框架。

(3)实体关系抽取。实体关系抽取是指从自然语言文本中抽取已标注实体之间的语义关系，一般可以形式化地描述为<实体，关系，实体>三元组，是构建语义网络和知识图谱的基础[42]。王仁武等人[43]利用基于特征的方法，采用深度置信网络(DBN)自动识别领域信息中实体之间的语义关系，并使用图数据库来存储实体关系图谱，从而构建领域知识图谱；李枫林等人[44]比较不同神经网络模型对于实体关系抽取的效果，发现卷积神经网络模型(CNN)擅长提取句子局部关键信息，时间递归神经网络更适合抽取句子多个实体之间的语义关系，结构递归神经网络在短文本中抽取实体关系表现更优，而且将神经网络模型与自然语言的先验知识相结合，实体关系抽取性能会显著提升。

信息抽取是信息组织和检索的基础性环节，信息抽取的精度和粒度直接影响信息组织和检索的质量。深度学习能够基于实体之间语义关联来建立信息抽取模型，充分考虑实体之间的语义关系，使得信息抽取的精准度和细粒度有了很大的提升。基于深度学习的智能信息抽取模型还能融合自然语言的先验知识来提升信息抽取效果，具有较高的适用性，在信息组织与检索中得到广泛应用。

3.2 基于深度学习的文本分类研究

文本分类是指在给定训练文本分类体系的情况下，对新输入的无标注文本自动确定相关类别的过程[45]。深度学习技术在图像识别、语音识别和计算机视觉等领域取得显著的成果，这些都为信息组织与检索中基于深度学习的文本分类研究打下了坚实的基础。 陈庆一[46]选取受限玻尔兹曼机(RBM)作为独立文本分类器，并改进特征选择算法，将基于类专属词的特征选择算法融入RBM模型以构建文本分类模型，研究发现基于类专属词的RBM分类器效果要优于传统基于文档频率的RBM分类器；袁彬[47]比较不同神经网络模型的文本特征表示对文本分类的影响，研究发现基于卷积神经网络(CNN)模型的文本特征表示对文本分类的效果更好；郭利敏等人[48]在自动文献分类中引入卷积神经网络模型，将文献分类问题转化为基于神经网络的自动学习和预测的问题，构建基于题名、关键词的多层次卷积神经网络模型来建立文献分类系统模型，自动学习文献的中图分类号，并对7000多条未加工文献进行预测以论证模型的可行性，其研究为文献半自动化分类的实现提供新的方法；邓三鸿等人[49]提出了一个解决中文图书多标签分类的方法，将图书的多标签分类转化成多个二元分类问题，结合字嵌入和长短时记忆(LSTM)训练多个二元分类模型，从而为书目标注多个标签；黄磊等人[50]选取双向递归神经网络模型(RNN)来提取文本特征，发现该方法虽然较传统基线分类模型有很大的提升，但存在比较严重的长尾效应；蓝雯飞等人[51]提出注意力机制结合深度学习技术来构建文本分类模型，该模型以Word2vec模型学习到的文本语义特征向量作为输入，将注意力机制引入到卷积神经网络(CNN)模型中来训练文本自动分类器，研究发现与传统的文本分类模型相比，引入注意力机制的卷积神经网络模型分类效果提升显著。

深度学习技术有效地缓解了大规模文本分类中语义特征提取困难和文本表示问题，利用神经网络语言模型获得文本的分布式表示方法，同时再利用CNN/RNN等神经网络结构自动获取文本的特征表达能力，简化了人工特征设计，实现端到端的信息处理。但是深度学习技术在文本分类中还存在文本信息过度丢失、独立性较差和语料库限制等问题，有待于进一步探索深度学习技术在文本分类中的应用

和发展。

3.3 基于深度学习的情感分析研究

情感分析，又称为倾向性分析，是指借助于额外的文本资源（如情感词表、词汇本体等）对主观性文本的情感信息进行极性识别，实现对文本信息的积极(Positive)、消极(Negative)和中立(Neutral)等情感分类，从而挖掘和分析用户的情感倾向[52]。

如何利用深度学习技术从用户的评价、日志、问答等文本中获悉用户的情感倾向引起了研究者的广泛关注。吴鹏等人[53]将认知情感评价模型——OCC模型与深度学习技术相结合来构建网络舆情情感识别模型。该模型使用OCC情感规则对网络舆情文本进行情感标注，并采用卷积神经网络(CNN)对突发事件网络舆情的微博文本进行情感分类；张连彬[54]针对汽车用户短评论和长评论文本，分别选取卷积神经网络(CNN)模型和双向长短时记忆网络模型(Bi-LSTM)进行情感识别，以构建一个完整的汽车用户评论自动情感分类系统；庄须强等人[55]将注意力机制融入循环神经网络来对视频弹幕评论文本进行情感分析，根据弹幕评论文本对片段视频标注主题词，并引入注意力机制突出情感关键词，通过LDA模型对主题片段视频进行聚类，并利用长短时记忆模型(LSTM)计算片段视频之间的情感语义相似性以完成弹幕评论文本的情感分析任务，该研究基本解决了视频情感标签标注问题；余传明等人[56]将深度学习技术与结构对等学习方法相结合，使用卷积神经网络模型进行跨领域情感分析，并在针对跨领域学习中特征漂移的问题进一步提出基于深度循环神经网络的跨领域情感分析模型(CD-DRNN)[57]，利用长短时记忆模型(stacked-LSTM)训练源领域的标注数据以构建情感分类模型，并将学习到的源领域情感分类模型迁移到目标领域以对其无标注数据进行情感分析，有效地解决了跨领域学习中特征漂移的问题。

总的来说，基于深度学习的情感分析方法有效地缓解了人工情感标注问题，循环网络模型(RNN)特别是长短时记忆(LSTM)在情感分析中的表现尤为突出。但是是单一的神经网络模型并不能很好地完成情感分析任务，将深度学习技术与传统的研究方法相结合，并引入自然语言的先验知识进行辅助，能使得情感分析模型具有更高的准确率和更强的泛化能力。

3.4 基于深度学习的文本聚类研究

文本聚类是指在划分类别未知的情况下，将文本档集合自动划分为若干个簇，使得簇之间文本相似性尽可能大，而簇内文本相似性尽量小[58]。文本聚类的基本思想就是“将相似的文档归为同一类别”，相似度计算则是判断文本之间的相似或相异的重要手段。逯万辉[59]引入深度学习模型来构建期刊选题同质化测度模型，通过计算单篇期刊论文的特征相似度和语义相似度来评估期刊研究主题相似性；曹祺等人[60]将深度学习引入专利文献相似度测量，并与传统的基于TF-IDF、LIS和LDA模型的专利相似度测量方法进行比较，发现基于深度神经网络的专利相似度检测方法在获得同样相似度的情况下，简化了相似度测量流程；文奕等人[61]以用户访问网站的序列日志信息为研究对象，借鉴深度学习的Word2Vec词嵌入模型，并融合频繁序列挖掘算法，提出了一个专利聚类的PatentFreq2Vec模型来计算专利之间的相似度；李杰等人[62]从用户评价的情感倾向出发来提出产品特征，利用卷积神经网络(CNN)对用户短文本评论进行情感分类训练，以得出未标注评论文本的情感标签，并按照情感标签进行产品特征词聚类；谢宗彦等人[63]选取深度学习模型对旅游短文本评论进行主题聚类，利用Word2Vec模型获取用户短文本评论的语义特征向量，再使用卷积神经网络模型(CNN)抽取评论文本的高阶语义特征，并将其输入到SOM模型中进行评论文本主题聚类，发现该模型在查准率、回召率和F值上均有提高。

基于深度学习的文本聚类能够深层次地评估文本之间的语义关系，尤其是利用神经网络言语模型所获得分布式文本表示，使得文本能基于语义关联进行聚类。但是对文本之间的相似度评估主要还是采用的传统相似度计算方法，其聚类方法还有待进一步探索和扩展。

4 深度学习在信息组织与检索中的研究趋势

随着大数据时代的不断发展，深度学习技术在信息组织与检索方向的研究越来越受到重视。尽管在智能信息抽取、文本分类、情感分析和文本聚类等方面取得了一定的进展，但是整体来说深度学习在信息组织与检索方向的应用研究还处于初级阶段，在未来必定会有更多的尝试和突破，根据深度学习模型的海量多源数据处理和高层抽象特征表示的特点，预测了以下四个深度学习在信息组织与检索方向未来可能的研究方向。

4.1 多源异构信息处理

在大数据时代，用户的偏好特征获取不再仅限于文本数据，图像、标签、音频、视频等多媒体数据也都蕴含着丰富的用户行为偏好和个性化需求信息。而且相较于文本信息，图像、音频和视频等多源异构信息蕴藏的数据信息更加直观和准确，更能深层次地描绘用户的认知结构。但是多源异构数据具有多模态、异构性、分散性等特点，难以对其进行系统化处理。深度学习作为一种多层次非线性的深度神经网络结构，具有强大的特征学习能力，能够多源异构数据映射到统一的隐空间中，自动学习到异构信息的语义特征。Guan等人[64]尝试利用堆叠自动编码器神经网络模型(SAE)从产品图像、描述和评论文本等多个信息源中来抽取产品信息，并取得了不错的效果。因此，利用深度学习网络模型提取多源异构数据中的辅助信息特征，为后续的信息分析、信息检索和个性化信息推荐等研究提供额外信息辅助这个研究方向值得信息组织与检索方向学者深入研究。

4.2 智能信息检索

深度学习技术因其多源异构数据的自动特征提取和高层次的抽象语义特征表示等优点,已经在自然语言处理领域取得不错的进展。自然语言处理(Natural Language Processing NLP)与信息检索(Information Retrieval IR)在很多方面是重叠的,自然语言处理(NLP)的主要目的是让计算机理解自然语言文本,并对其进行结构化的分析和处理,而IR主要目的是理解用户需求和文档内容,并根据用户查询式将最符合用户需求的候选文档集合反馈给用户。大数据时代,深度神经网络模型能够获取多源异构媒体信息的抽象特征,并映射到共享空间进行语义相似度计算,建立跨媒体之间的语义检索关联[66]。同时深度神经网络能从多角度对图像、音视频等多媒体进行语义标注,提升检索结果的准确率。同时,深度学习在NLP领域中的语音识别和智能问答等方向的突破和发展也为信息组织与检索方向的语音检索[67]提供了研究基础。目前基于深度学习的视频检索[68]、跨媒体语义检索[69]和语音检索等智能信息检索的研究相对较少,未来还需要深入地探索和研究。

4.3 个性化信息推荐

信息推荐是指通过用户的历史行为数据以及一些隐含用户偏好的辅助信息,建立用户画像来分析预测用户的兴趣倾向和潜在信息需求,然后将数据库中满足用户偏好或需求的信息内容推荐给用户。人工智能时代,挖掘信息深层次的语义关联备受关注,基于语义的个性化信息推荐也成为研究热点[70]。深度学习模型能够深层次的学习到数据的隐藏语义特征,充分挖掘和分析用户的偏好特征,及时为用户推荐其感兴趣或所需的信息。比如根据用户兴趣爱好和职业来为用户推荐个性化咨询新闻[65];根据学者的研究领域和浏览日志为其推荐相关的学术论文[71]、期刊、合作学者和研究机构等[72-73];用户生病时根据用户的症状描述和诊断情况,及时为用户推荐相关的专家、医院等健康信息。基于深度学习的个性化信息推荐能够基于语义信息更加准确地预测用户的信息需求,及时为用户推送精准信息服务。虽然目前已有相关研究出现,但这个基于深度学习的个性化信息推荐这个研究向未来仍然有待更加深入地探索和更广泛地拓展。

5 结 语

随着大数据时代的不断深入,各个领域的数据信息呈指数增长,用户信息需求也日益提高。如何对海量纷繁复杂的数据信息进行组织和检索是大数据发展的关键问题之一。基于深度学习的信息组织与检索研究能融合各种类型的多源异构数据,自动学习数据的高层次语义特征表示,实现智能化的信息组织与检索。目前,基于深度学习的信息组织与检索在智能信息抽取、自动文本分类、情感分析和文本聚类等方面取得不错的成果,但深度学习在信息组织与检索中的研究仍然处于初级阶段,其可以深入并可以取得成果的方向还有很多,如异构信息处理、智能信息检索、个性化信息推荐等。本文在介绍深度学习在信息组织与检索中常用模型的基础上,梳理和分析了基于深度学习的信息组织与检索的研究现状和进展,并预测了今后可能的研究趋势,希望能为相关领域的研究人员提供借鉴和参考。

参考文献

1 熊回香. 面向Web3.0的大众分类研究[D]. 武汉:华中师范大学, 2011.
2 张佩云,吴 江.信息资源的组织与检索模式研究[J].计算机技术与发展,2006,(1):132-134,157.
3 Hinton GE, Salakhutdinov RR. Reducing the dimensionality of data with neural networks[J].Science, 2006,313(5786):504-507.
4 Schmidhuber, Jürgen. Deep learning in neural networks: An overview[J]. Neural Networks, 2015,(61): 85-117.
5 Hinton G , Deng L , Yu D , et al. Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups[J]. IEEE Signal Processing Magazine, 2012, 29(6):82-97.
6 Amodei D , Ananthanarayanan S , Anubhai R , et al. Deep Speech 2: End-to-End Speech Recognition in English and Mandarin[C]//Curran Associates, Inc..33rd International conference on machine learning: ICML 2016, New York City, New York, USA, 19-24 June 2016.
7 Krizhevsky, Alex,Sutskever, Ilya,Hinton, Geoffrey E.. ImageNet Classification with Deep Convolutional Neural Networks[J].Communications of the ACM, 2017,60(6):84-90.
8 Szegedy C , Liu W , Jia Y , et al. Going Deeper with Convolutions[C]//Institute of Electrical and Electronics Engineers.2015 IEEE Conference on Computer Vision and Pattern Recognition: 2015 28th IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2015), 7-12 June 2015, Boston, MA, USA,2015.
9 Young T, Hazarika D, Poria S, et al. Recent Trends in Deep Learning Based Natural Language Processing Review Article[J].IEEE Computational Intelligence Magazine, 2017, 13(3):55-75.
10 Wang X , Shrivastava A , Gupta A . A-Fast-RCNN: Hard Positive Generation via Adversary for Object Detection [C]//Institute of Electrical and Electronics Engineers.2017 IEEE Conference on Computer Vision and Pattern Recognition: CVPR 2017, Honolulu, Hawaii, USA,2017.

11 Wei Xu,Alex Rudnicky.Can Artificial Neural Networks Learn Language Models?[C]//6th International Conference on Spoken Language Processing v.1,2000.
12 Bengio Y, Ducharme R, Vincent P, et al. A neural probabilistic language model[J]. Journal of machine learning research, 2003, 3(Feb): 1137-1155.
13 Mnih A , Hinton G . A Scalable Hierarchical Distributed Language Model[C]// International Conference on Neural Information Processing Systems. Curran Associates Inc. 2008.
14 林奕欧,雷 航,李晓瑜,等. 自然语言处理中的深度学习:方法及应用[J]. 电子科技大学学报, 2017,46(6):115-121.
15 Harris Z S. Distributional Structure[J]. Word,1981,10(2-3): 146-162.
16 Mikolov T, Chen K, Corrado G, et al. Efficient Estimation of Word Representations in Vector Space [EB/OL]. https://arxiv.org/pdf/1301.3781.pdf,2019-04-19.
17 郑文超,徐 鹏.利用word2vec对中文词进行聚类的研究[J].软件,2013,34(12):160-162.
18 曾颖黎.网络舆情文本分类系统研究与开发[D].成都:电子科技大学,2014.
19 来斯惟.基于神经网络的词和文档语义向量表示方法研究[D].北京:中国科学院自动化研究所,2016.
20 张晓娟.利用嵌入方法实现个性化查询重构[J].情报学报,2018,37(6):621-630.
21 Hinton GE, Osindero S, Teh, YW. A Fast Learning Algorithm for Deep Belief Nets[J].Neural Computation,2006,18 (7):1527-1554.
22 Salakhutdinov R , Mnih A , Hinton G . ACM Press the 24th international conference - Corvalis, Oregon (2007.06.20-2007.06.24) [C]//Proceedings of the 24th international conference on Machine learning - ICML \" 07 - Restricted Boltzmann machines for collaborative filtering,2007:791-798.
23 Bengio Y. Learning Deep Architectures for AI[J]. Foundations & Trends® in Machine Learning, 2009, 2(1):1-127.
24 Hinton G . A practical guide to training restricted boltzmann machines[J]. Momentum, 2010, 9(1):926-947.
25 Bengio Y, Lamblin P, Dan P, et al. Greedy layer- wise training of deep networks[J]. Advances in Neural Information Processing Systems, 2007,(19):153-160.
26 Vincent P, Larochelle H, Bengio Y, et al. Extracting and composing robust features with denoising autoencoders [C]//in:Proceedings of the Twenty- Fifth International Conference on Machine Learning,2008.
27 H Lee, R Grosse, R Ranganath et al..Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations[C]//ACM Press.Proceedings of the 26th Annual International Conference on Machine Learning,2009.
28 LeCun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
29 Krizhevsky A , Sutskever I , Hinton G . ImageNet Classification with Deep Convolutional Neural Networks[C]//in: Neural Information Processing Systems.Advances in Neural Information Processing Systems 25, vol. 2: 26th annual conference on Neural Information Processing Systems 2012, December 3-6, 2012, Lake Tahoe, Nevada, USAC ,2012.
30 Schroff F, Kalenichenko D, Philbin J. FaceNet: A unified embedding for face recognition and clustering[C]//in:Institute of Electrical and Electronics Engineers.2015 IEEE Conference on Computer Vision and Pattern Recognition: 2015 28th IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2015),Boston, MA, USA,2015.
31 Gong Y , Zhang Q . Hashtag Recommendation Using Attention- Based Convolutional Neural Network[C]//AAAI Press.Proceedings of the Twenty- Fifth International Joint Conference on Artificial Intelligence: IJCAI- 16, New York City, New York, USA,2016.
32 刘 琴.基于深度学习的短文本分类研究综述.中国计算机用户协会网络应用分会.中国计算机用户协会网络应用分会2017年第二十一届网络新技术与应用年会论文集[C].中国计算机用户协会网络应用分会:北京联合大学北京市信息服务工程重点实验室,2017:11-15.
33 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.
34 Hochreiter S, Schmidhuber J. Long Short- Term Memory [J]. Neural Computation, 1997, 9(8):1735-1780.
35 K Cho, B Merrieenboer, D Bahdanau, et al,.On the Properties of Neural Machine Translation: Encoder- Decoder Approaches[C]//in:Association for Computational Linguistics.8th Workshop on syntax, semantics and structure in statistical translation 2014: 8th Workshop on syntax, semantics and structure in statistical translation 2014 held at the conference on empirical methods in natural language processing (EMNLP 2014),Doha, Qatar,2014.
36 郭喜跃,何婷婷.信息抽取研究综述[J].计算机科学,2015, 42(2):14-17,38.
37 朱丹浩,杨 蕾,王东波.基于深度学习的中文机构名识别研究——一种汉字级别的循环神经网络方法[J].现代图书情报技术, 2016,(12):40-47.
38 朱娜娜,景 东,薛 涵.基于深度神经网络的微博图书名识别研究[J].图书情报工作,2016,60(4):102-106,141.
39 曾维新,赵 翔,冯 滔,等.面向领域的命名实体消歧方法改进研究[J].计算机工程与应用,2018,54(17):126-134.

40 Le-kui ZHOU,Si-liang TANG,Jun XIAO,Fei WU,Yue-ting ZHUANG.Disambiguating named entities with deep supervised learning via crowd labels[J].Frontiers of Information Technology & Electronic Engineering,2017,18(01):97-107.
41 秦 越. 维吾尔语人称代词指代消歧研究[D]. 乌鲁木齐: 新疆大学,2018.
42 鄂海红,张文静,等.深度学习实体关系抽取研究综述[EB/OL].https://doi.org/10.13328/j.cnki.jos.005817,2019-04-11.
43 王仁武, 袁 毅, 袁旭萍. 基于深度学习与图数据库构建中文商业知识图谱的探索研究[J]. 图书与情报, 2016,(1):110-117.
44 李枫林,柯 佳.基于深度学习框架的实体关系抽取研究进展[J].情报科学,2018,36(3):169-176.
45 奉国和.自动文本分类技术研究[J].情报杂志,2007,(12):108-111
46 陈庆一. 基于RBM的文本分类算法研究[D]. 长春:吉林大学,2015
47 袁 彬. 基于语义特征的文本分类算法研究[D]. 北京:北京邮电大学,2016.
48 郭利敏.基于卷积神经网络的文献自动分类研究[J].图书与情报,2017,(6):96-103.
49 邓三鸿,傅余洋子,王 昊.基于LSTM模型的中文图书多标签分类研究[J].数据分析与知识发现,2017,1(7):52-60.
50 黄 磊,杜昌顺.基于递归神经网络的文本分类研究[J].北京化工大学学报(自然科学版),2017,44(1):98-104.
51 蓝雯飞,徐 蔚,王 涛.基于卷积神经网络的中文新闻文本分类[J].中南民族大学学报(自然科学版),2018,37(1):138-143.
52 李光敏,许新山,熊旭辉.Web文本情感分析研究综述[J].现代情报,2014,34(5):173-176.
53 吴 鹏, 刘恒旺, 沈 思. 基于深度学习和OCC情感规则的网络舆情情感识别研究[J]. 情报学报,2017,(36):980.
54 张连彬.基于汽车品牌评论的情感分类系统研究[D].合肥:合肥工业大学,2017.
55 庄须强,刘方爱.基于AT-LSTM的弹幕评论情感分析[J].数字技术与应用,2018,36(2):210-212.
56 余传明,冯博琳,安 璐.基于深度表示学习的跨领域情感分析[J].数据分析与知识发现,2017,1(7):73-81.
57 余传明.基于深度循环神经网络的跨领域文本情感分析[J].图书情报工作,2018,62(11):23-34.
58 夏立新, 金 燕, 方 志. 信息检索原理与技术[M]. 北京: 科学出版社, 2009.
59 逯万辉. 基于深度学习的学术期刊选题同质化测度方法研究[J]. 情报资料工作, 2017,(5):107-114
60 曹 祺, 赵 伟, 张英杰,等. 基于Doc2Vec的专利文件相似度检测方法的对比研究[J]. 图书情报工作, 2018,62(13):74-81.
61 文 奕,陈文杰,张 鑫,等. 利用词嵌入模型实现基于网站访问日志的专利聚类研究[J]. 现代情报, 2018, 38(4):112-117.
62 李 杰,李 欢. 基于深度学习的短文本评论产品特征提取及情感分类研究[J]. 情报理论与实践,2018,41(2):143-148.
63 谢宗彦,黎 嶸,周纯洁.基于CNN和SOM的评论主题发现[J].情报科学,2018,36(6):30-34.
64 Yue Guan, Qiang Wei, Guo-qing Chen. Deep learning based personalized recommendation with multi-view information integration[J]. Decision Support Systems,2019,118:58-69.
65 彭 欣.基于深度学习的数字图书馆跨媒体语义检索方法研究[J].情报探索,2018,(2):16-19.
66 Hung-Yi L , Pei-Hung C , Yen-Chen W , et al. Interactive Spoken Content Retrieval by Deep Reinforcement Learning[J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2018,26(12):2447-2459.
67 李 想.基于深度学习的视频敏感信息检索的研究[J].电子设计工程,2017,25(21):137-140.
68 王兆凯,李亚星,冯旭鹏,刘利军,黄青松,刘晓梅.基于深度信念网络的个性化信息推荐[J].计算机工程,2016,42(10):201-206.
69 吴 丹,程 磊.信息组织与检索的研究热点与动向:语义、交互与社群[J].图书情报知识,2017,(4):4-12.
70 王 妍,唐 杰.基于深度学习的论文个性化推荐算法[J].中文信息学报,2018,32(4):114-119.
71 Wang X , Yu L , Ren K , et al. Dynamic Attention Deep Model for Article Recommendation by Learning Human Editors´ Demonstration[C]// Acm Sigkdd International Conference on Knowledge Discovery & Data Mining. ACM, 2017.
72 Hassan H A M . Personalized Research Paper Recommendation using Deep Learning[C]// Conference on User Modeling. ACM, 2017.
73 李志勇,李鹏伟,高小燕,等.人工智能医学技术发展的聚焦领域与趋势分析[J].中国医学装备,2018,15(7):136-145.

(责任编辑:孙晓明)