

中文在线评论文本分析对生鲜农产品电商影响研究综述

冯建英 吴丹丹 王 博 王 智 穆维松
(中国农业大学信息与电气工程学院, 北京 100083)

摘要: 深入挖掘和分析生鲜农产品的电商评论数据对于降低消费者的感知风险、辅助消费者决策具有重要意义。首先阐述了中文在线评论文本数据的获取方式、预处理方法、文本表示方法、基于不同模型和技术的文本情感分析等研究进展,然后重点分析了评论文本对生鲜农产品销量的影响机制、评论的信息属性和情感属性以及评论矛盾性对生鲜农产品网络销量影响的最新研究成果,并提出未来研究将进一步注重提升评论数据的质量、融合评论中的多模态数据和研究表情符号在情感表达中的作用。

关键词: 在线评论; 评论文本; 生鲜农产品; 产品销量; 情感分析

中图分类号: S126; TP399 **文献标识码:** A **文章编号:** 1000-1298(2021)S0-0504-09

Online Comments Analysis and Its Application Research Progress in E-commerce of Fresh Agricultural Products

FENG Jianying WU Dandan WANG Bo WANG Zhi MU Weisong
(College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China)

Abstract: Deeply mining and analysis of E-commerce online comments of fresh agricultural products is of great significance for reducing consumers' perceived risks and assisting consumers in making decisions. The research progress of online comments acquisition methods, text preprocessing methods, text representation methods, and text sentiment analysis was firstly reviewed based on different models and technologies. Then the latest research results such as the influence mechanism of the comments on the sales of fresh agricultural products, the information and emotional attributes of the comments, and the influence of the contradictory comments on online sales of agricultural products were analyzed. Finally, it was proposed that future research should further focus on improving the quality of data, fusing multi-modal data in comments, and studying the role of emoji in emotional expression.

Key words: online comments; comments text; fresh agricultural products; product sales; sentiment analysis

0 引言

《乡村振兴战略规划(2018—2022)》提出加强农村电子商务基础设施建设,截至2019年底我国农村互联网普及率已经达到了38.4%,电子商务帮助农村连接大市场,成为乡村振兴战略的强力支撑点,在解决“三农问题”中发挥着重要作用^[1]。新冠肺炎疫情影响下,线上销售生鲜农产品成为解决生鲜农产品滞销突出问题的有效手段^[2]。生鲜农产品是人们生活的必需品,由于其具有保质期短、易腐、

易损、不易保鲜保质等特征,消费者对生鲜农产品的品质和安全有很高的要求^[3]。很多消费者对于线上购买该类商品具有较高的感知风险,因而我国生鲜农产品电商销量滞后于电子产品、服装等产品。

在线评论里已购消费者发布的评价体验信息,能够帮助消费者了解生鲜农产品的内在品质和外在品质信息,有效降低消费者的感知风险,辅助消费者作出购买决策,因而生鲜农产品的在线评论信息对生鲜农产品电商发展具有重要意义。

文本评论在生鲜农产品在线评论数据中占比

收稿日期: 2021-07-13 修回日期: 2021-09-08
基金项目: 新疆杏产业技术体系专项资金项目(XJCYTX-03)
作者简介: 冯建英(1982—),女,副教授,博士,主要从事农业大数据分析与智能决策研究,E-mail: fjying@cau.edu.cn
通信作者: 穆维松(1967—),女,教授,博士,主要从事农业大数据分析与智能决策研究,E-mail: wsmu@cau.edu.cn

最大,并且生鲜农产品在线评论文本的分析方法和方向与其他产品评论信息的挖掘既有共性,又有其特殊性。因此,本文首先分析国内外学者关于在线评论文本的获取、预处理、表示、情感分析的研究进展,然后围绕评论文本对生鲜农产品销量的影响展开综述,包括影响机制、评论数据的信息属性与情感属性以及评论矛盾性对农产品销量的影响,最后讨论未来的研究趋势,以期从理论和应用两方面为今后生鲜农产品电商评论和网络销量的研究提供参考。

1 在线评论文本的获取、预处理和表示

1.1 在线评论数据获取

充足数量的生鲜农产品在线评论是进行评论情感分析和影响研究的基础。考虑到消费者网络购买生鲜农产品以后的售后服务,发表评论及追加评论的保障,为获取充足的评论数据,大多数研究者选择京东或者淘宝等综合网络销售平台。

在线评论通常使用自行编写的爬虫^[4]程序和公开发布的爬虫^[4]软件两种方式获取。两种方式各有优点,自行编写的爬虫程序可以根据研究调整获取数据的内容与存储格式等,在数据采集初期可以除去大多数的数据噪声。爬虫软件获取数据,不需要自己编写,减少了部分工作量,采集过程简单、能够节省时间,但所得数据样式固定,数据清洗时较

为麻烦。

伴随电子商务的迅速发展,生鲜农产品的线上销售渠道越来越多,研究者可以从不同的购物平台上抓取更多种类的生鲜农产品评论数据,为评论数据挖掘提供更充分的数据支撑。

1.2 在线评论预处理

获取到在线评论数据后,先人工去除数据采集过程出现的乱码以及重复采集的评论,然后对其进行分词、去停用词和词性标注预处理,为下一步文本表示作准备。

1.2.1 文本分词

相比较于英文等西方语言,中文以字为基本单位,词之间界限模糊,缺少空格的分隔,在文本分析过程中,中文分词至关重要却又存在多种困难。汉语中新词汇、地名、人名、机构名、事件名和各种专业术语逐年增加,同时,一个汉语句子常对应多个合法次序列,因而,解决未登录词语问题和消除分词歧义^[5]一直是中文分词研究致力解决的难题。

词是汉语信息处理使用的具有确定的语义和语法功能的基本单位,是最小的能够独立活动的语言成分。从20世纪80年代发展至今主要有3大类中文分词方法,分别是基于规则的分词方法^[6-7]、基于统计的分词方法^[8-9]和基于语义的分词方法^[10-12]。表1从分词方法、分词思想、分词算法、优点与缺点5方面比较汇总了3种方法。

表 1 分词方法对比				
Tab.1 Comparison of word segmentation methods				
分词方法	分词思想描述	主要算法	优点	缺点
基于规则	按照一定策略将待分析的汉字串与词典中的词条进行匹配,若在词典中找到某个字符串,则匹配成功	最大匹配算法 ^[13]	速度很快;对于歧义部分,能有效提高准确率	少数未能找出的歧义,在后面处理方法中无法再对其处理
		最小匹配算法 ^[14]		
		最短路径算法 ^[15]		
基于统计	对训练文本中相邻字的组合进行统计,计算它们之间的互现频度,频度高的形成词	n-gram ^[16] HMM ^[17]	适应性强,能够实现对日文韩文等亚洲文字和英文短语的处理	索引冗余大
基于语义	利用计算机模拟人对文本的理解,结合语义、句法等因素处理文本	CNN ^[18]	不需要人工标注的词库,当引入的规则得当时,能达到较好的分词效果	需要大量的语言知识,该方法目前还难以实施
		RNN ^[19]		

1.2.2 词性标注

词性标注的过程是赋予每个词语一个正确候选词性的过程^[20],词性有名词、动词、形容词、数词、量词等^[21]。中文词性标注目前使用较多的方式有基于统计^[22]的词性标注和基于深度学习^[23]的词性标注两种。基于统计模型将词性标注看成一个序列标注问题,先给定带有标注的词序列,再确定词最有可能的词性。该方法更加客观,适应性更强,但当训练语料达到一定规模后,不能再通过扩大语料规模来

提高性能。深度学习模型具有很强的非线性拟合能力,不依赖大规模的词典,泛化性更强,词性标注性能较高。若要训练出效果较好的深度学习模型,需要大量的标注语料数据,但目前许多研究的标注语料规模不够大^[24]。

1.2.3 停用词去除

停用词指的是在处理数据(文本等)的过程中为了提高效率而自动过滤掉一些出现频率过高且没有实际意思的词^[25],停用词还包括数字、切分标记、

数词、量词、代词、方位词、拟声词、叹词、没有实际意义的动词以及不太常用的词。停用词的数量非常少,往往就几十个。停用词表有通用停用词表与专用停用词表之分,又可分为绝对停用词和相对停用词,前者包含虚词、助动词等功能词,后者一般指在文本集中出现频率较高,但对文本信息检索没有积极作用的非功能词。目前在线评论的研究中使用较多的通用停用词表有 4 个,分别是:中文停用词表、哈工大停用词表、百度停用词表和四川大学机器智能实验室停用词库^[26],研究者通常将这 4 个停用词表整合去重复。

1.2.4 文本表示

对在线评论数据分词、词性标注和去停用词等处理后,形成一堆零散的词语,仍未变成传统的结构化数值型,不能被计算机直接识别处理,需利用数学模型将文本数据结构化表示。常用的文本表示方法有 3 种:布尔模型^[27]、向量空间模型^[28-29]和概率模型^[27,30]。布尔模型使用二进制 0 和 1 表示文本,简单易于实现,但该方法会使得文本中的信息缺失过多。向量空间模型把文本信息表示成空间中的点,通过计算向量点之间的几何距离判别需要的搜索项,返回最可能需要的结果,再对其排序。基于语义和理解的分词利用计算机模拟人对文本的理解,结合语义、句法等因素处理文本。

2 评论文本情感分析

文本情感分析,又称意见挖掘 (Opinion mining),是指对带有情感色彩的主观性文本进行采集、处理、分析、归纳和推理的过程,涉及人工智能、机器学习、数据挖掘、自然语言处理等多个研究领域^[30]。生鲜农产品在线评论的情感倾向是消费者购物体验是否满意的直观展示,是影响其他消费者判断商品是否值得购买的重要参考。生鲜农产品在线评论情感分析通常从语句粒度考虑整句评论的情感倾向,研究者往往采用基于词典的情感分析、基于 LDA 主题的情感分析和基于深度学习的情感分析。

2.1 基于词典的情感分析

基于词典的情感分析方法,通过构建情感词典与文本中的词语进行匹配,计算对比词情感倾向,通过预先设定的规则求出整个文本的情感类。分析流程如图 1 所示。

该方法判定准确的前提是确保有高质量的情感词典^[31]。建立情感词典有手动构建和自动构建两种方法^[32]。基于人工构建情感词典是对词汇进行领域区分和向量化表示。无监督词典的构建方法用于检测特定领域句子的情感极性。情感词典

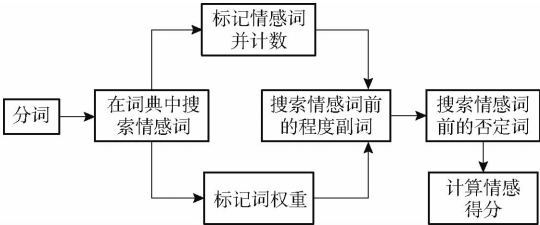


图 1 基于词典的文本情感分析流程图
Fig.1 Dictionary-based text sentiment analysis process

种类主要有 3 种,分别是:基础情感词典、扩展情感词典^[32]和领域情感词典^[33-34]。基础情感词典包含一系列被大众广泛认同的情感词,如好、差、喜欢和讨厌等,现主要有 Hownet 情感词、NTUSD 简体中文词典和中文褒贬词典 3 个。扩展情感词典由基础情感词典通过同义词词典寻找到的同义词和近义词组成,如哈尔滨工业大学信息检索实验室利用众多词语相关资源整合的《同义词词林》。领域情感词典由不属于基础情感词的在特定领域内带有情感倾向的词汇组成,需要根据领域的语料库,通过算法获取。

2.2 基于 LDA 主题模型的情感分析

LDA (Latent Dirichlet analysis) 模型在 2003 年被首次提出,是一种无监督的机器学习技术。主题模型能够深入挖掘文本词汇间的语义关系,在主题发现方面有着巨大的优势^[35]。众多学者引入监督机制对 LDA 模型进行改进,或把 LDA 主题挖掘结果作为中间层与其他模型相结合对生鲜农产品在线评论进行情感分析。

国显达等^[36]提出了一种 Gaussian LDA 在线评论主题挖掘方法,该方法克服了现有方法生成的主题稀疏性和语义导致的适应性差的局限性,应用 AP 聚类算法先将在线评论进行聚类再实现主题发现。崔雪莲等^[37]在 LDA 模型的基础上提出了一种基于主题相似性的无监督在线评论情感分类模型。该模型通过计算评论和理想评论集的主题相似度,获得在线评论情感倾向值,最终实现情感分类,该方法优于其他主题情感混合模型,并且具有领域可移植性。王珠美等^[38]结合 LDA 主题模型和直觉模糊 TOPSIS 理论,提出了一种适用于农产品在线评论的情感分析方法,对天猫商城售卖的西湖龙井茶叶特点在线评论进行分析,验证了该方法具有合理性和实际应用价值。李慧宗等^[39]基于 LDA 模型对冷链农产品在线评论的好评和差评聚类分析,结果表明从商品的分拣加工、冷链设施的改善、运输配送路程等进行优化可以吸引更多的消费者。

2.3 基于深度学习模型的情感分析

深度学习是在模拟人的神经元之间信息传递的

过程^[40],根据神经网络里计算节点(神经元)的连接方式的不同,形成不同的表达模型,基础神经网络模型有卷积神经网络(CNN)^[41]、循环神经网络(RNN)^[18]、深度神经网络(DNN)^[42]和自编码器^[43]4种。4种模型的演化算法、特点和应用领域汇总对比如表2所示。

表 2 深度学习模型对比
Tab.2 Comparison of deep learning models

模型	演化算法	特点	应用领域
RNN	LSTM ^[44]	采用网状/树状的计算图结构,考虑时间的因素,允许信息的持久化,递归神经网络还可以引入门控机制学习长距离依赖	语音分析
	Bi-directional ^[45]		文字分析
	RNN ^[45]		时间序列分析
CNN	AlexNet ^[46]	神经元是非全连接的,同一层中神经元之间的连接权重是共享的	图像
	VGG ^[47]		格状结构化数据(Grid-like data)
	GoogleNet ^[48]		
DNN	Adagrad ^[49]	逐层做无监督学习,每次学习一层网络结构,再逐步加深网络	
	ADAM ^[50]		
	FTRL ^[51]		
自编码器 (Autoencoder)	收缩自编码器 ^[52]	将输入信息作为学习目标,对输入信息进行表征学习	降维,去噪编码器
	正则自编码器 ^[53]		
	变分自编码器 ^[54]		

基于深度学习的在线评论情感分析在获取的语料中提取特征,用提取的特征训练模型,最终实现文本情感分类,流程如图2所示。

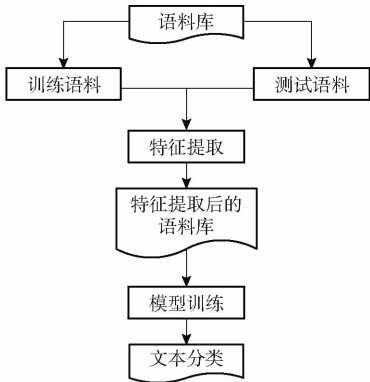


图 2 基于深度学习的文本情感分析流程图
Fig.2 Text sentiment analysis process based on deep learning

许多学者使用一种或者多种神经网络模型及其变体模型分析生鲜农产品在线评论情感倾向。许犇等^[55]提出了记忆图卷积神经网络(MGCNN)引入注意力机制的商品评论情感分析方法。叶小峰^[56]提出了基于领域情感词典与NCRF++模型的电商评论情感分类研究方法,有效改善了农产品电商评论的领域特性和普通情感词典分类不够精确的问题。李勇等^[57]从传统BSTM模型的基础上,引入位置注意力机制同时融合CNN,突出精选肉类、新鲜水果和休闲食品在线评论的情感语义极性来实现情感语义特征分类,提高了情感分类精度。姚瑶^[58]从文本语句的角度提出了基于Bi-LSTM对冷链农产品电商平台消费者评论情感分类方法,分析得出消费者对淮南“菜篮子工程”电商平台的整体评价,从而反

映消费者对整个冷链农产品电商平台的满意度。

3 评论文本对生鲜农产品销量的影响

学者们认为评论数量、评论效价、评论及时性、评论时效性、评论者专业程度、品牌效应、评论矛盾性、评论文本内容等评论特征会对生鲜农产品的网络销量产生影响^[59]。

3.1 评论文本对农产品销量的影响机制

在实施购买行为之前,消费者会考虑购买决策的不确定性及其引起的不利后果,从而影响其购买决策的制定,这就是消费者的感知风险^[60]。在线评论文本正是通过影响消费者的感知风险进而影响其购买决策。为了降低购买风险,消费者会尽量通过多种途径更多地获取商品信息,包括咨询商家、参考商品页的图文信息、在不同商家和产品之间进行比较以及参考已购者的评论信息等。由于评论者和消费者同属于买方,具有共同的买方利益,因而相较于商家提供的产品信息,消费者更倾向于相信评论者的信息,在线评论反映出的已购者的情感倾向、满意度、商品评价等信息可以帮助消费者消除部分疑虑,降低其风险感知,增强购买意愿^[61]。平均每条评论的字数越多,所携带的产品或服务信息越多,即评论深度越高,会降低消费者对生鲜农产品的感知风险,李春月^[62]研究发现初次评论的评论深度对生鲜水果销量有正向影响,而追加评论深度影响不大。涂洪波等^[63]研究发现感知风险对消费者网上购买地理标志农产品意愿有明显负向影响,商家要注重展示正面评论的内容和数量来降低感知风险,吸引消费者购买。

3.2 评论信息的可获得性及诊断性对生鲜农产品销量的影响

信息可获得性及诊断性理论认为信息线索对消费者决策影响主要取决于信息可获得性及诊断性的共同作用,其中信息可获得性是指消费者能够获取有助于评估产品质量与服务的信息,信息诊断性是指信息能够帮助消费者将产品或服务分类为独特的类别。包含产品信息详尽、态度明确的在线评论,其可获得性和诊断性相对较高,对消费者购物决策的影响也更大^[64]。信息可获得性是信息诊断性的基础,即使某条评论具有较高诊断性,但其自身信息可获得性较低,消费者也不会依据该条评论作出决策。陈在飞^[65]指出有周期性的生鲜农产品,发布时间久的在线评论对当前潜在消费者的诊断性很低,不能反映现阶段生鲜产品的质量与服务状况。陈登全^[66]认为评论深度越高,消费者阅读评论的积极性越高,从评论中获取的高诊断性信息也越多,越有助于消费者做出购买决策,研究证明滞后 1 d 的在线评论深度对生鲜农产品销量有显著的正向影响。

3.3 评论文本情感对生鲜农产品销量的影响

王娟^[67]通过对高原特色农产品特征评论进行情感分析预测了农产品销量,首先通过消费者对商品属性的评论分析出高原特色农产品 10 个指标特征与销量之间的关联关系,并用 BP 神经网络模型实现对农产品销量的预测。郭敏等^[68]通过分析京东稻米的在线评论发现:评论文本极性会对销量产生影响,对品牌声誉高的产品销量影响更为显著,而对于品牌声誉低的产品,评论高频词对销量的影响更显著。章蓬伟等^[69]通过分析新疆巴州库尔勒香梨在线评论数据,从用户情感倾向出发挖掘出香梨销售存在的优点和不足,并给出保持口感、加强售后服务以提高销量的建议。马凤才等^[70]通过基于京东生鲜中苹果、叶菜类、鱼类和猪肉在线评论的分析得出物流、包装、服务、品质及价格为影响消费者生鲜产品满意度的关键要素,认为消费者对电子商务平台销售生鲜产品的满意度影响其重购意愿。朱静^[71]提出文本正向情感倾向通过消费者感知有用性间接影响生鲜农产品网络销量,文本正向情感越强,消费者评论感知有用性程度越高,生鲜农产品网络销量越高。

3.4 评论的矛盾性对生鲜农产品销量的影响

在线评论矛盾性是指积极和消极评论同时存在并达到一定比例,使消费者无法确定所评论商品的

优劣而难以决策^[72]。不同评论者在不同时间对生鲜农产品不同属性的评价存在差异是产生评论矛盾性的原因,评论矛盾性从评论属性和评论者来源不同可分为:同一评论者对不同属性的矛盾评论,不同评论者对同一属性的矛盾评论,同一评论者在初次评论和追加评论中对同一属性的矛盾评论,不同评论者之间在初次评论和追加评论的矛盾评论。李倩等^[73]把初次评论和追加评论的矛盾分为:正-负型同一属性,正-负型不同属性,负-正型同一属性,负-正型不同属性 4 种类型,将评论文本分析和问卷调查相结合进行分析,得出负-正同一属性型评论对消费者购买意愿的影响更大的结论。王英等^[74]发现不同评论者的矛盾性评论正向影响稻米农产品销量。李林红等^[75]研究苹果的在线评论发现文本初次评价整体不一致性对销量具有显著的负向影响,而追加评价口感不一致性的离散度与销量呈现正相关关系;初次评价整体不一致性的离散度与销量呈“U”形相关,而追加评价整体不一致性的离散度与销量呈倒“U”形相关。

4 未来的研究趋势

(1)提升在线评论数据的质量。生鲜农产品网络市场逐渐扩大,市场竞争日益激烈,部分商家通过刷单等方式发布与实际购买体验不符的虚假好评,树立服务优、质量好的店铺形象来吸引消费者,以期提高销量,更有甚者去竞争者店铺故意编造恶意诋毁的差评。这些虚假评论会严重干扰消费者购买决策,因此,在文本分析之前根据生鲜农产品自身特征,如何检测删除虚假评论,提升在线评论数据的质量已成为当前需要解决的问题。

(2)注重生鲜农产品在线评论中多模态数据的融合。一条在线评论往往会包括文本、图像和视频形式的信息,下一步的研究应专注于生鲜农产品在线评论文本、图像和视频的多模态数据融合,建立多模态数据挖掘模型,发挥不同形式数据之间的互补作用,从更广泛的角度挖掘融合数据中的信息。

(3)重视表情符号在评论信息情感表达中的作用。表情符号直观易懂,很多网民在发表观点、意见和态度时,都会在文字描述中添加相应的表情符号来强调当前的情感状态,生鲜农产品评论中也会涉及大量的表情符号。但目前研究在线评论时多是将表情符号作为噪声数据去除,造成了情感分析结果的偏差,未来研究应重视表情符号在评论中所发挥的作用。

参 考 文 献

- [1] 陈德宝. 电子商务对乡村振兴的促进作用研究[J]. 江苏商论, 2021(2):29-32.
CHEN Debao. Research on the promotion effect of electronic commerce on rural revitalization[J]. Jiangsu Commercial Forum, 2021(2):29-32. (in Chinese)
- [2] 李曦方, 闫宁. 农产品新媒体营销分析——直播电商新势力[J]. 办公自动化, 2021,26(7):33-34.
LI Xifang, YAN Ning. Analysis of the new media marketing of agricultural products—new power of direct broadcasting business [J]. Office Informatization, 2021,26(7):33-34. (in Chinese)
- [3] 马若男. 生鲜农产品电子商务消费者购买意愿研究[J]. 现代营销(下旬刊), 2021(2):66-67.
- [4] 李彦. 基于 Python 的网络爬虫技术的研究[J]. 电子世界, 2021(3):39-40.
- [5] 索浩东. 中文分词[J]. 中国文艺家, 2019(2):199-200.
- [6] 马海, 马力. 基于词典的中文微博情感细粒度分析研究[J]. 计算机与数字工程, 2020,48(10):2415-2419.
MA Hai, MA Li. Research on fine-grained analysis of Chinese microblog emotion based on dictionary[J]. Computer & Digital Engineering,2020,48(10):2415-2419. (in Chinese)
- [7] 杨贵军, 徐雪, 凤丽洲, 等. 基于最大匹配算法的似然导向中文分词方法[J]. 统计与信息论坛, 2019,34(3):18-23.
YANG Guijun, XU Xue, FENG Lizhou, et al. Likelihood oriented method for Chinese word segmentation based on maximum match algorithm[J]. Statistics & Information Forum, 2019,34(3):18-23. (in Chinese)
- [8] 赵月斋. 一种基于词频统计的中文分词方法[J]. 科技展望, 2016,26(10):283.
- [9] 邹佳伦, 文汉云, 王同喜. 基于统计的中文分词算法研究[J]. 电脑知识与技术, 2019,15(4):149-150.
- [10] 王星, 李超, 陈吉. 基于膨胀卷积神经网络模型的中文分词方法[J]. 中文信息学报, 2019,33(9):24-30.
WANG Xing, LI Chao, CHEN Ji. Dilated convolution neural networks for Chinese word segmentation[J]. Journal of Chinese Information Processing, 2019,33(9):24-30. (in Chinese)
- [11] 成于思, 施云涛. 基于深度学习和迁移学习的领域自适应中文分词[J]. 中文信息学报, 2019,33(9):9-16.
CHENG Yusi, SHI Yuntao. Domain adaption of Chinese word segmentation based on deep learning and transfer learning[J]. Journal of Chinese Information Processing, 2019,33(9):9-16. (in Chinese)
- [12] 马学海. 基于神经网络的中文分词技术研究[J]. 科学技术创新, 2019(32):82-83.
MA Xuehai. Dilated convolution neural networks for Chinese word segmentation [J]. Journal of Chinese Information Processing, 2019(32):82-83. (in Chinese)
- [13] DUCOFFE G, POPA A. The use of a pruned modular decomposition for maximum matching algorithms on some graph classes [J]. Discrete Applied Mathematics, 2021,291:201-222.
- [14] 韩冰. 基于最小连通支配子图的图匹配算法[D]. 大连:大连海事大学, 2020.
- [15] WEI Liao, WEI Xiaohui, LAI Jizhou, et al. Numerical method with high real-time property based on shortest path algorithm for optimal control[J]. International Journal of Control, Automation and Systems, 2021,19(6):2038-2046.
- [16] 杨新生, 胡立生. 基于隐马尔科夫模型的古汉语词性标注[J]. 微型电脑应用, 2020,36(5):130-133.
YANG Xinsheng, HU Lisheng. Part-of-speech tagging of classical Chinese based on hidden Markovian model microcomputer application[J]. Microcomputer Applications,2020,36(5):130-133. (in Chinese)
- [17] 周潭, 莫礼平, 胡美琪, 等. 基于 MEM 和 HMM 的中文词性标注方法[J]. 吉首大学学报(自然科学版), 2020,41(2):15-18.
ZHOU Tan, MO Liping, HU Meiqi, et al. Chinese part-of-speech tagging method based on maximum entropy model and hidden Markov model[J]. Journal of Jishou University, 2020,41(2):15-18. (in Chinese)
- [18] ABID S S, MUNAF R, SAMREEN H, et al. Comparative analysis of CNN and RNN for voice pathology detection[J]. Biomed Research International, 2021(Special Issue):6635964.
- [19] KUMAR D, MATHUR H D, BHANOT S, et al. Forecasting of solar and wind power using LSTM RNN for load frequency control in isolated microgrid[J]. International Journal of Modelling and Simulation,2021,41(4):311-323.
- [20] 金国哲, 崔荣一. 一种新的朝鲜语词性标注方法[J]. 中文信息学报, 2018,32(10):53-58.
JIN Guozhe, CUI Rongyi. A novel Korean pos tagging method[J]. Journal of Chinese Information Processing, 2018,32(10):53-58. (in Chinese)
- [21] 帕丽旦·木合塔尔, 吾守尔·斯拉木, 买买提阿依甫. 基于混合模型的维吾尔文词性标注方法[J]. 计算机仿真, 2019,36(1):268-273.
MUHETAER Palidan, SILAMU Wushouer, MAIMAITAYIFU. Uyghur part of speech tagging method based on hybrid model [J]. Computer Simulation,2019,36(1):268-273. (in Chinese)
- [22] 徐艳华, 朱诗瑶. 基于统计的《HSK 词汇等级大纲》动名兼类词标注研究[J]. 云南师范大学学报(对外汉语学与研究版), 2020,18(4):51-57.
XU Yanhua, ZHU Shiyao. Study of the part-of-speech tagging of verb-noun words in outline of graded vocabulary of HSK in a statistical erspective[J]. Journal of Yunnan Normal University(Teaching & Studying Chinese as a Foreign Language Edition),

- 2020,18(4):51–57. (in Chinese)
- [23] XIE G, SHANGGUAN A, FEI R, et al. Motion trajectory prediction based on a CNN–LSTM sequential model[J]. Science China(Information Sciences), 2020,63(11):248–268.
- [24] 王连喜,钟准,丁曾强,等.融合深度神经网络与统计学习的印地语词性标注方法研究[J].湖南工业大学报,2020,34(3):17–22.
WANG Lianxi, ZHONG Zhun, DING Zengqiang. Research on Hindi part-of-speech tagging based on deep neural network and statistical learning[J]. Journal of Hunan University of Technology, 2020,34(3):17–22. (in Chinese)
- [25] 吕结红.基于文本挖掘的酒店在线评论研究[D].武汉:华中师范大学,2020.
LÜ Jiehong. Research on online hotel reviews based on text mining[D]. Wuhan: Central China Normal University, 2020. (in Chinese)
- [26] 官琴,邓三鸿,王昊.中文文本聚类常用停用词表对比研究[J].数据分析与知识发现,2017,1(3):72–80.
GUAN Qin, DENG Sanhong, WANG Hao. Chinese stopwords for text clustering:a comparative study[J]. Data Analysis and Knowledge Discovery, 2017,1(3):72–80. (in Chinese)
- [27] 骆梅柳.文本表示模型在文本挖掘中的应用[J].现代信息科技,2019,3(7):24–25.
LUO Meiliu. Application of text representation model in text mining[J]. Modern Information Technology, 2019,3(7):24–25. (in Chinese)
- [28] 杨春霞,吴佳君,李欣栩.融合实体信息的循环神经网络文本分类模型[J].小型微型计算机系统,2020,41(12):2516–2521.
YANG Chunxia, WU Jiajun, LI Xinxu. Text classification model based on recurrent neural network with entity information[J]. Mini-micro Systems, 2020,41(12):2516–2521. (in Chinese)
- [29] 高云龙,吴川,朱明.基于改进卷积神经网络的短文本分类模型[J].吉林大学学报(理学版),2020,58(4):923–930.
GAO Yunlong, WU Chuan, ZHU Ming. Short text classification model based on improved convolutional neural network[J]. Journal of Jilin University(Science Edition), 2020,58(4):923–930. (in Chinese)
- [30] 吕妹园,张永健,张永强,等.融入情感信息词向量的评论文本情感分析方法[J].河北科技大学学报,2021,42(4):380–388.
LÜ Meiyuan, ZHANG Yongjian, ZHANG Yongqiang, et al. Sentiment analysis method of comment text based on word vector with sentiment information[J]. Journal of Hebei University of Science and Technology, 2021,42(4):380–388. (in Chinese)
- [31] 高华玲,张晶.基于情感词典的酒店评论情感分析与可视化[J].软件,2021,42(1):45–47.
GAO Hualing, ZHANG Jing. Sentiment analysis and visualization of hotel reviews based on sentiment dictionary[J]. Computer Engineering & Software 2021,42(1):45–47. (in Chinese)
- [32] 景丽,李曼曼,何婷婷.结合扩充词典与自监督学习的网络评论情感分类[J].计算机科学,2020,47(增刊2):78–82.
JING Li, LI Manman, HE Tingting. Sentiment classification of network reviews combining extended dictionary and self-supervised learning[J]. Computer Science, 2020,47(Supp.2):78–82. (in Chinese)
- [33] 黄磊.基于电商网站商品评论的商品属性提取及其情感的可视化表示[D].北京:北京邮电大学,2017.
HUANG Lei. Product attributes extraction and visual representation of its emotion based on product reviews[D]. Beijing: Beijing University of Posts and Telecommunications, 2017. (in Chinese)
- [34] 胡潇涛,吴浩,杨亮,等.基于伪标注样本融合的领域分词方法[J].四川轻化工大学学报(自然科学版),2021,34(1):48–55.
HU Xiaotao, WU Hao, YANG Liang, et al. Domain word segmentation method based on pseudo-labeled sample fusion[J]. Journal of Sichuan University of Science & Engineering(Natural Science Edition), 2021,34(1):48–55. (in Chinese)
- [35] 王德川,赵瑞雪,寇远涛,等.基于LDA模型的欧盟框架计划近10年农业科研项目研究主题发现[J].农业展望,2017,13(4):69–75.
WANG Dechuan, ZHAO Ruixue, KOU Yuantao, et al. Topic discovery of agricultural scientific research projects of EU's research framework program based on LDA model in recent decade[J]. Agricultural Outlook, 2017,13(4):69–75. (in Chinese)
- [36] 国显达,那日萨,高欢,等.基于Gaussian LDA的在线评论主题挖掘研究[J].情报学报,2020,39(6):630–639.
GUO Xianda, NARISA, GAO Huan, et al. Topic mining of online reviews based on Gaussian latent Dirichlet allocation[J]. Journal of Information Science, 2020,39(6):630–639. (in Chinese)
- [37] 崔雪莲,那日萨,刘晓君.基于主题相似性的在线评论情感分析[J].系统管理学报,2018,27(5):821–827.
CUI Xuelian, NARISA, LIU Xiaojun. Sentiment analysis of online reviews based on topic similarity[J]. Journal of System Management, 2018,27(5):821–827. (in Chinese)
- [38] 王珠美,胡彦蓉,刘洪久.基于LDA主题模型和直觉模糊TOPSIS的农产品在线评论情感分析[J].数据采集与处理,2020,35(5):965–977.
WANG Zhumei, HU Yanrong, LIU Hongjiu. Emotional analysis of agricultural product online reviews based on LDA thematic mode and Intuitionistic fuzzy TOPSIS[J]. Data Acquisition and Processing, 2020,35(5):965–977. (in Chinese)

- [39] 李慧宗,姚瑶,王向前,等. 基于 LDA 的冷链农产品电商在线评论的情感分析[J]. 南阳理工学院学报, 2019,11(2): 25–30.
LI Huizong, YAO Yao, WANG Xiangqian, et al. LDA-based sentiment analysis of online reviews of cold-chain farm products e-commerce companies[J]. Journal of Nanyang Institute of Technology, 2019, 11(2): 25–30. (in Chinese)
- [40] 焦李成,杨淑媛,刘芳,等. 神经网络七十年:回顾与展望[J]. 计算机学报, 2016,39(8):1697–1716.
JIAO Licheng, YANG Shuyuan, LIU Fang, et al. Seventy years beyond neural networks:retrospect and prospect[J]. Chinese Journal of Computers, 2016, 39(8): 1697–1716. (in Chinese)
- [41] SIKAKOLLU P, DASH R. Ensemble of multiple CNN classifiers for HSI classification with superpixel smoothing[J]. Computers & Geosciences, 2021, 154:104806.
- [42] 亓慧,史颖,李灯熬,等. 基于连续型深度置信神经网络的软件可靠性预测[J]. 计算机科学, 2021,48(5):86–90.
QI Hui, SHI Ying, LI Dengao, et al. Software reliability prediction based on continuous deep confidence neural network[J]. Journal of Computer Science, 2021, 48(5): 86–90. (in Chinese)
- [43] RAO J, ZHOU X, LU Y, et al. Imputing single-cell RNA-seq data by combining graph convolution and autoencoder neural networks[J]. iScience, 2021, 24(5):102393.
- [44] LI Lifen, ZHANG Tianyu. Research on text generation based on LSTM[J]. International Core Journal of Engineering, 2021, 7(5):525–535.
- [45] 易炜,何嘉,邹茂扬. 基于循环神经网络的对话系统记忆机制[J]. 计算机工程与设计, 2019,40(11):3259–3264.
YI Wei, HE Jia, ZOU Maoyang. Memory mechanism of dialogue system based on recurrent neural network[J]. Computer Engineering and Design, 2019, 40(11): 3259–3264. (in Chinese)
- [46] 方双,赵凤霞,楚松峰,等. 基于多尺度卷积神经网络的缺陷红枣检测方法[J]. 食品与机械, 2021,37(2):158–163.
FANG Shuang, ZHAO Fengxia, CHU Songfeng, et al. Defective jujube detection technology based on multi-scale convolutional neural network[J]. Food & Machinery, 2021, 37(2): 158–163. (in Chinese)
- [47] 蒋珍存,温晓静,董正心,等. 基于深度学习的 VGG16 图像型火灾探测方法研究[J]. 消防科学与技术, 2021,40(3): 375–377.
JIANG Zhencun, WEN Xiaojing, DONG Zhengxin, et al. Research on fire detection of improved VGG16 image recognition based on deep learning[J]. Fire Science and Technology, 2021, 40(3): 375–377. (in Chinese)
- [48] HUANG Xuehua, CHEN Weihong, YANG Wangdong. Improved algorithm based on the deep integration of Googlenet and residual neural network[J]. Journal of Physics: Conference Series, 2021, 1757(1):1–8.
- [49] 盛志超,周勃,杜璞玉,等. 基于 AdaGrad 的 GNSS-R 镜面反射点自适应预测算法[J]. 无线电工程, 2019,49(11): 956–960.
SHENG Zhichao, ZHOU Bo, DU Puyu, et al. Specular point self-adaptive prediction algorithm for GNSS-R based on AdaGrad[J]. Radio Engineering, 2019, 49(11): 956–960. (in Chinese)
- [50] 雷剑. 基于 Adam 优化神经网络的电梯群控算法[J]. 智能计算机与应用, 2020,10(11):101–105.
LEI Jian. Elevator group control algorithm based on Adam optimized neural network[J]. Intelligent Computer and Applications, 2020, 10(11): 101–105. (in Chinese)
- [51] 钟小勇. 基于 FTRL 和 XGBoost 组合算法的电商销量预测系统[J]. 信息记录材料, 2020,21(1):1–3.
ZHONG Xiaoyong. E-commerce sales forecasting system based on FTRL and XGBoost algorithm[J]. Information Recording Materials, 2020, 21(1): 1–3. (in Chinese)
- [52] 邓钦文. 收缩卷积自编码器网络及其在图像去噪中的应用研究[D]. 武汉:华中师范大学, 2020.
DENG Qinwen. Contractive convolutional autoencoder net and its research on the application of image denoising[D]. Wuhan: Central China Normal University, 2020. (in Chinese)
- [53] 张重鹏. 基于正则化约束自编码器的图像聚类算法研究[D]. 阜新: 辽宁工程技术大学, 2020.
ZHANG Zhongpeng. Image clustering algorithm based on regularization constrained autoencoder[D]. Fuxin: Liaoning Technical University, 2020. (in Chinese)
- [54] 张弛,吴东,王伟,等. 不平衡样本下基于变分自编码器预处理深度学习和 DGA 的变压器故障诊断方法[J]. 南方电网技术, 2021,15(3):68–74.
ZHANG Chi, WU Dong, WANG Wei, et al. Transformer fault diagnosis method based on variational auto-encoders preprocessing deep learning and DGA for unbalanced samples[J]. Southern Power System Technology, 2021, 15(3): 68–74. (in Chinese)
- [55] 许彝,徐国庆,程志宇,等. 基于 MGCNN 的商品评论情感分析[J]. 武汉工程大学学报, 2020,42(5):585–590.
XU Ben, XU Guoqing, CHENG Zhiyu, et al. Sentiment analysis of product reviews based on memory graph convolutional neural network[J]. Journal of Wuhan Institute of Technology, 2020, 42(5): 585–590. (in Chinese)
- [56] 叶小峰. 基于 NCRF++ 模型的农产品电商推荐系统研究[D]. 合肥:安徽农业大学, 2020.
YE Xiaofeng. Research and development of agricultural product e-commerce recommendation system based on NCRF++ model[D]. Hefei: Anhui Agricultural University, 2020. (in Chinese)
- [57] 李勇,金庆雨,张青川. 融合位置注意力机制和改进 BLSTM 的食品评论情感分析[J]. 郑州大学学报(工学版), 2020,41(1):58–62.
LI Yong, JIN Qingyu, ZHANG Qingchuan. Improved BLSTM food review sentiment analysis with positional attention

- mechanisms[J]. Journal of Zhengzhou University(Engineering Science), 2020, 41(1): 58–62. (in Chinese)
- [58] 姚瑶. 基于文本挖掘的冷链农产品电商平台在线评论的情感分析研究[D]. 淮南:安徽理工大学, 2019.
YAO Yao. Online review of cold chain agricultural products e-commerce platform based on text mining emotional analysis[D]. Huainan: Anhui University of Science and Technology, 2019. (in Chinese)
- [59] 何永杰. 在线评论对汽车销量的影响研究[D]. 长春:吉林大学, 2020.
HE Yongjie. A study on the impact of online reviews on car sales[D]. Changchun: Jilin University, 2020. (in Chinese)
- [60] 杜一菲, 张全红. 感知风险对生鲜食用菌网购意愿的影响[J]. 中国食用菌, 2020, 39(11): 245–248.
DU Yifei, ZHANG Quanhong. The Influence of perceived risk on online shopping intention of fresh edible fungi[J]. Edible Fungi of China, 2020, 39(11): 245–248. (in Chinese)
- [61] 李孟涵. 图文呈现与在线评论不一致性对生鲜农产品销量的影响研究[D]. 昆明:昆明理工大学, 2020.
- [62] 李春月. 初次评论及追加评论信息对生鲜水果销量影响研究[D]. 大庆:黑龙江八一农垦大学, 2020.
LI Chunyue. Initial comment and additional comment information on fresh fruit sales impact study[D]. Daqing: Heilongjiang Bayi Agricultural University, 2020. (in Chinese)
- [63] 涂洪波, 朱清剑. 外部线索对地理标志农产品网购意愿的影响及作用机理[J]. 中国流通经济, 2020, 34(8): 37–47.
TU Hongbo, ZHU Qingjian. Research on influence and mechanism of external clues on customers' online purchase intention of agricultural products with geographical indications[J]. China Business and Market, 2020, 34(8): 37–47. (in Chinese)
- [64] 王慧. B2C 环境下在线评论对农产品销量的影响研究[D]. 深圳:南方科技大学, 2015.
WANG Hui. Study about how online reviews influence the sales volume of agricultural products in B2C environment[D]. Shenzhen: South University of Science and Technology, 2015. (in Chinese)
- [65] 陈在飞. 不同来源在线评论对网络零售商产品销量的影响研究[D]. 南京:南京大学, 2014.
CHEN Zaifei. Study on the influence of online reviews from different sources on product sales of internet retailers[D]. Nanjing: Nanjing University, 2014. (in Chinese)
- [66] 陈登全. 在线评论对生鲜农产品销量的影响研究——以柠檬为例[D]. 重庆:西南大学, 2020.
CHEN Dengquan. Research on the impact of online reviews on the sales of fresh agricultural products—take lemon as an example[D]. Chongqing: Southwest University, 2020. (in Chinese)
- [67] 王娟. 基于 BP 神经网络的高原农产品销量预测——以云南省特色水果为例[D]. 昆明:昆明理工大学, 2020.
WANG Juan. Sales forecast of plateau agricultural products based on BP neural network—taking characteristic fruits in Yunnan Province as an example [D]. Kunming: Kunming University of Technology, 2020. (in Chinese)
- [68] 郭敏, 周雅欣. 品牌声誉对农产品网络销售的调节影响——基于在线评论的内容分析[J]. 湖北农业科学, 2020, 59(5): 152–156.
GUO Min, ZHOU Yaxin. Moderating effect of brand reputation on network sales of agricultural products—based on analysis of online comments[J]. Hubei Agricultural Sciences, 2020, 59(5): 152–156. (in Chinese)
- [69] 章蓬伟, 贾钰峰, 邵小青, 等. 基于主题模型的农产品在线评论数据研究[J]. 电子技术与软件工程, 2021(5): 150–151.
- [70] 马凤才, 李春月. 消费者对电子商务平台销售生鲜产品满意度测算研究——基于京东生鲜在线评论的分析[J]. 价格理论与实践, 2020(5): 117–120.
MA Fengcai, LI Chunyue. Research on e-commerce consumer satisfaction measurement of fresh products—analysis based on online reviews of JD Fresh[J]. Price: Theory & Practice, 2020(5): 117–120. (in Chinese)
- [71] 朱静. 在线评论对生鲜农产品网络销量的影响研究[D]. 重庆:重庆工商大学, 2020.
ZHU Jing. Research on the impact of online comments on the online sales of fresh agricultural products[D]. Chongqing: Chongqing University of Technology and Industry, 2020. (in Chinese)
- [72] 曹敏, 李资博, 熊剑芳. 矛盾型在线评论情境下消费者有用性感知和购买意向评价研究——基于模糊综合评价法[J]. 绿色科技, 2021, 23(6): 261–264.
- [73] 李倩, 谢向英, 管曦. 矛盾性追加评论对青年茶叶消费者购买意愿的影响研究[J]. 茶叶通讯, 2020, 47(4): 705–711.
LI Qian, XIE Xiangying, GUAN Xi. Study on the impact of contradictory additional comments on young tea consumers' purchase intention[J]. Journal of Tea Communication, 2020, 47(4): 705–711. (in Chinese)
- [74] 王英, 杨丹, 王乐, 等. 基于品牌效应的在线评论对稻米销量影响的实证研究[J]. 中国农业大学学报, 2018, 23(1): 178–187.
WANG Ying, YANG Dan, WANG Le, et al. Impact of online reviews on the sales of rice under the influence of brand effect: an empirical study[J]. Journal of China Agricultural University, 2018, 23(1): 178–187. (in Chinese)
- [75] 李林红, 李孟涵. 农产品图文呈现与在线评论不一致性对销量的影响——以苹果为例[J]. 江苏农业科学, 2020, 48(9): 304–309.
LI Linhong, LI Menghan. Influence of graphic presentation and online reviews inconsistency of agricultural products on sales—taking apple as an example[J]. Jiangsu Agricultural Sciences, 2020, 48(9): 304–309. (in Chinese)