

一种基于成词率和谱聚类的电力文本领域词发现方法*

杨政¹, 尹春林¹, 蔡迪², 李慧斌²

(1. 云南电网有限责任公司电力科学研究院, 云南 昆明 650217; 2. 西安交通大学 数学与统计学院, 陕西 西安 710049)

摘要: 考虑到当前电力行业仍缺少有效的领域词发现方法, 以电力行业科技项目文本为原始语料库, 将基于互信息与左右熵的统计特征与传统语言构词规则特征相融合, 提出了电力文本成词率的概念。所提方法首先利用成词率对电力文本进行无监督筛选得到初始候选词集, 然后对候选词集进行文本切片算法和常用词过滤操作, 最后进行词嵌入和谱聚类得到最终所需的电力文本领域词。实验结果表明, 所提出的方法准确有效, 为电力文本的领域词发现提供了一种新方法。

关键词: 成词率; 谱聚类; 领域词发现; 电力文本

中图分类号: TP311

文献标识码: A

DOI: 10.16157/j.issn.0258-7998.211437

中文引用格式: 杨政, 尹春林, 蔡迪, 等. 一种基于成词率和谱聚类的电力文本领域词发现方法[J]. 电子技术应用, 2021, 47(10): 29-32, 37.

英文引用格式: Yang Zheng, Yin Chunlin, Cai Di, et al. A power text domain word discovery method based on word formation rate and spectral clustering[J]. Application of Electronic Technique, 2021, 47(10): 29-32, 37.

A power text domain word discovery method based on word formation rate and spectral clustering

Yang Zheng¹, Yin Chunlin¹, Cai Di², Li Huibin²

(1. Electric Power Research Institute of Yunnan Power Grid Co., Ltd., Kunming 650217, China;

2. School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: Considering that the current power industry still lacks effective domain word discovery methods, this paper takes the power industry science and technology project text as the original corpus, combines the statistical features based on the mutual information, left entropy as well as right entropy with the features of traditional language word-formation rules, and proposes the new concept of power text word formation rate. The proposed method firstly uses the word formation rate to get the initial candidate word set by unsupervised filtering, and then performs the text slicing algorithm and common word filtering operation on the candidate word set, and finally performs the word embedding and spectral clustering algorithms to get the final power text-domain words. Experimental results show that the method proposed in this paper is accurate and effective, and provides a new method for power text domain word discovery.

Key words: word formation rate; spectral clustering; domain word discovery; power text

0 引言

针对特定领域的文本数据, 领域词的词库构建是最为关键的任务之一。传统领域词发现方法依赖互信息或邻接熵得到候选词集, 进而利用 word2vec 进行词向量转化、K-means 进行聚类^[1], 最终得到行业领域词。传统方法对词语组合规律运用得不够全面, 因此这类方法筛选的候选词集存在诸多不合理的词语。领域词发现分为候选词集筛选与字符串过滤两个步骤。

在候选词集筛选方面, 领域词发现算法主要是基于

词语统计特性的无监督方法或序列模式机器学习的有监督算法。基于无监督的方法中, 互信息和凝固度是最常见的用来筛选词语的度量, 刘伟童等^[2]提出使用互信息初步筛选词集, 随后用邻接熵对词集进行再过滤的方法。刘昱彤等^[3]使用改进的类 Apriori 算法, 通过组合、统计频率、过滤 3 个步骤来筛选候选词集。杜丽萍等^[4]提出利用改进的互信息, 同时结合一定的构词规则筛选候选词集。无监督算法泛化性优良, 但缺少规则, 会遗留有较多垃圾串与非领域词。基于监督的机器学习词集筛选方法有马建红等^[5]提出的基于 CNN 和 LSTM 抽取词特征, 随后使用半马尔科夫条件随机场 (SCRf) 来识别词语边界。Fu Guohong 等^[6]在隐马尔可夫模型 (HMM) 的框架

* 基金项目: 国家自然科学基金面上项目 (61976173); 教育部-中国移动人工智能建设项目 (MCM20190701)

下运用命名实体识别(NER)的思路,同时结合上下文筛选出候选词集。陈飞等^[7]提出运用条件随机场来判断分词的词汇边界是否为候选词边界的方法。监督方法通常需要大量标注数据进行训练,耗费高额的人工成本。此外,部分方法选择基于纯规则的构词法^[8-9](即汉语成词规则)与一些领域先验知识结合,进行候选词集的筛选。这种方式虽然准确性相对较高,但是规则维护复杂,基本无跨域能力。

在得到候选词集之后还需对垃圾字符串与非领域词进行过滤。赵志滨等^[10]定义了词语相似度计算的Words-Avg方法,通过与阈值对比来判断候选词集里的词是否为领域词。王鑫等^[11]在使用word2vec训练候选词向量时加入了词类信息,进而得到化学领域词。吉久明等^[12]运用改进的GloVe词向量模型提取词语向量,随后进行K-means聚类,最终得到领域词。

1 本文方法

图1展示了本文提出的方法,共分为5个步骤:文本预处理、候选词筛选、文本切片算法、常用词过滤以及谱聚类^[13]。

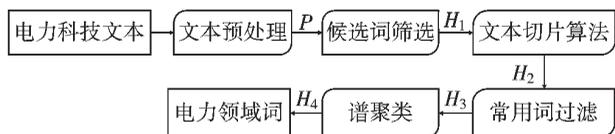


图1 本文所提方法流程图

1.1 文本预处理

在对文本进行候选词集筛选前,需对电力科技文档进行文本预处理,处理后得到文本记为P。具体处理步骤如下:

(1)筛选出电力科技文档中目的和意义、项目研究内容与实施方案三部分,避免其他电力领域词很少出现的一部分干扰最终结果,同时提升算法运行效率;

(2)去除掉停用词,即语气助词、副词、介词、连接词(如“的”)以及标点符号。

1.2 候选词筛选

在得到预处理后的文本P后,本文提出了一种新的候选词集筛选度量“成词率(Suc)”,用以对文本P中的词语进行筛选,筛选后得到初步的词集H₁。其中设定最长词语长度n为4,候选词筛选度量Suc的阈值设定为0.62,成词率由互信息(Mut)、左右熵(Adj)与构词规律(Reg)构成。

(1)互信息(Mut):指一个随机变量中包含的关于另一个随机变量的信息量,即随机变量间的联系程度,包含的信息量越高,则变量间联系程度越高,词内凝固度越高,该字符串则越容易成词。互信息Mut的计算公式如式(1)~式(3)所示,分别表示2、3、4字词的互信息,其中,x、y、z代表单个字符。

$$Mut = \log_2 \frac{P(x, y)}{P(x)P(y)} \quad (1)$$

$$Mut = \min_{x, y, z} \left\{ \log_2 \frac{P(x, y, z)}{P(x)P(y, z)}, \log_2 \frac{P(x, y, z)}{P(x, y)P(z)} \right\} \quad (2)$$

$$Mut = \min_{w, x, y, z} \left\{ \log_2 \frac{P(w, x, y, z)}{P(w)P(x, y, z)}, \log_2 \frac{P(w, x, y, z)}{P(w, x)P(y, z)}, \log_2 \frac{P(w, x, y, z)}{P(w, x, y)P(z)} \right\} \quad (3)$$

(2)左右熵(Adj):指候选词的自由程度,左右熵表示词语搭配的不确定性。左右熵越大,词语搭配的不确定性越大,即候选的词左右搭配越丰富,成词概率越高,计算公式如式(4)、式(5)所示,Adj_l、Adj_r分别为左熵和右熵,w为字符串,w_r、w_l为单前缀与单后缀。

$$Adj_l = - \sum \frac{P(w, w_r)}{P(w)} \cdot \log_2 \frac{P(w, w_r)}{P(w)} \quad (4)$$

$$Adj_r = - \sum \frac{P(w, w_l)}{P(w)} \cdot \log_2 \frac{P(w, w_l)}{P(w)} \quad (5)$$

(3)构词规律(Reg):通过对电力行业领域词的观察发现,绝大多数领域词依然遵循构词学基本原理^[14]:名词与名词、动词、形容词结合非常频繁,其他结合方式构成电力领域词的比例约四分之一,本文定义了字符串结合成词的构词规律Reg,其中a、b为搭配成词的两个字符串。

$$Reg(a, b) = \begin{cases} 0.1 & a \text{ 或 } b \text{ 为名词, 对应搭配为名、动、形容词} \\ 0.2 & \text{其他} \end{cases} \quad (6)$$

(4)成词率(Suc):用来筛选候选词集的度量,结合了上述3个变量的优点。

$$Suc = 1/m(Mut + k_1 \cdot Adj_l + k_2 \cdot Adj_r) - Reg(a, b) \quad (7)$$

其中,k₁和k₂将互信息与左右熵的值映射到相近的量子级,k₁取值为Mut均值除以Adj_l的均值,k₂取值为Mut均值除以Adj_r的均值;m为归一化常数。

1.3 文本切片算法

初步的词集H₁中仍有诸多未被筛选的电力领域词,本文提出使用文本切片算法,算法步骤分为分词和回溯,具体如下:

(1)分词:在得到初步的词集H₁以后,本文选择用H₁对预处理后的文本P进行切分,旨在得到更为合理的候选词集。在对预处理后的文本P进行分词时,遵循的原则为“最准确切分原则”,例如:“各项目经理”这个字符串,很有可能“各项”、“项目”、“经理”、“项目经理”均在H₁中,那么就选择不对此字符串进行分词,得到分词后的候选词集定义为H₂。

(2)回溯:检查候选词集H₂中的词语,如果词语长度小于n,则判断该词语是否在H₁中,若不在H₁中则从H₂中删除;如果词语长度大于n,则判断该词语内部字符串是否有一半以上位于H₁中,若没有则从H₂中删除。例如:H₂中有词语“单相接地故障”,该词语长度大于n,则判断“单相”、“单项接”、“单项接地”……是否在

H_1 中,若不在 H_1 中的词语个数小于其内部字符串组合情况总数的一半以上,则对该词语予以删除。

分词回溯举措中,分词意在最大化候选集中词语的准确性;回溯举措目的是进一步确保候选词集中的词语是内部足够“凝固”的,提高了候选词语的准确性。

1.4 常用词过滤

在得到的候选词集 H_2 中,需要进一步对非领域词进行过滤。本文从中国知网中选择经济、哲学、化工 3 个领域分别爬取 50、50、40 篇非电力领域科技文本,对这 140 篇文档重复第 1.1~1.3 小节步骤,得到词集 H_{21} ,同时将 H_2 过滤掉 H_2 与 H_{21} 的交集,进而得到词集 H_3 ,如式(8)所示。

$$H_3 = H_2 - (H_2 \cap H_{21}) \quad (8)$$

1.5 谱聚类

此时得到的候选词集 H_3 除了电力行业领域词外,还有其他领域词以及不常用的非电力行业词,诸如:“麦克斯韦方程”、“对抗样本”、“池化层”等。因此,需要对 H_3 中的词语进行聚类处理,此处选择 BERT^[15]算法对词集中的词语进行向量化处理,维度为 512,最后运用谱聚类进行降维处理并聚类,最终得到电力领域词集 H_4 。

谱聚类:把所有数据看作空间中的点,这些点可以用边连接起来,距离较远的两个点之间的边权重值较低,距离较近则较高。通过对所有数据点组成的无向图 $G=(V, E)$ 进行切图,让切图后不同的子图间边权重和尽可能低,子图内的边权重和尽可能高,从而达到聚类的目的。其中,图顶点间权重集合为 $E=\{A_{i,j}\}$,邻接矩阵为 $W=\{A_{i,j}\}$,定义度 D 为与当前图相连的所有图的权重 $A_{i,j}$ 之和。优化函数如式(9)所示:

$$\text{cut}(A_1, A_2, \dots, A_k) = \frac{1}{2} \sum_{i=1}^k W(A_i, A_i^1) \quad (9)$$

其中, k 为切图子图个数, A_i^1 代表 A_i 外其他子图, A_i^1 与 A_i 的切图权重 W 定义为式(10):

$$W(A_i^1, A_i) = \sum_{i \in A_i^1, j \in A_i} W_{ij} \quad (10)$$

其中, $W_{ij} = e^{-\frac{\|x_i - x_j\|_2^2}{2\sigma^2}}$ 。

2 实验

2.1 数据集及预处理

本文从中国南方电网有限公司云南电力科学研究院科技项目申报文档数据库中选取了 140 篇项目可研申请书作为领域词发现的语料库,研究主题主要包括:高电压与绝缘技术、电机与电气以及电力系统及自动化,表 1 给出了数据集的组成情况。其中语料中的电力领域词总数为 619 个,例如:高压、绝缘、电路、相位、电阻、继电器等。

实验环境为:8 核 16 线程 CPU, Ubuntu20.04, Python3.6, TensorFlow-GPU 2.1.0, Nvidia 2070s 显卡。

表 1 语料库中各类文本的数量

| 类别 | 高电压与绝缘技术 | 电机与电气 | 电力系统及自动化 |
|----|----------|-------|----------|
| 数量 | 50 | 40 | 50 |

在得到语料库后对每篇文档做如下处理,处理前后文档大小及本文所提算法运行时间如表 2 所示:

(1) 抽取出文档中目的和意义、项目研究内容与实施方案两部分。

(2) 对抽取出的部分做正则化处理,包括去掉通用格式、数字等对中文领域词发现无意义的字符。

(3) 将每篇文档被抽取出的部分拼接起来并存入 JSON 数据库,形成需要的数据。

表 2 语料处理前后对比

| | 空间大小/MB | 运行时间 |
|------|---------|------|
| 预处理前 | 53.8 | 约两小时 |
| 预处理后 | 1.17 | 50 s |

2.2 评价指标及参数设置

评价指标:本文选取的模型评价指标为准确率 P' 、召回率 R 以及 F 值,这些度量指标用来评判电力领域词发现模型效果,计算公式如下:

$$P' = \frac{\text{识别正确的电力领域词}}{\text{识别到的词语总数}} \quad (11)$$

$$R = \frac{\text{识别正确的电力领域词}}{\text{语料中的电力领域词总数}} \quad (12)$$

$$F = \frac{2P'R}{P'+R} \times 100\% \quad (13)$$

其中,识别到的词语总数为识别出的电力领域词集;语料中的电力领域词总数为人工筛选的电力领域词集;识别正确的电力领域词为识别出的电力领域词集与人工筛选的电力领域词集重合的部分。

参数设置: k_1, k_2 由互信息(Mut)的均值与左右熵(Adj)权重来确定。对于成词率阈值,本文选取贪心算法进行阈值的取值计算,设定成词率 Suc 阈值取值介于 0~1 之间,步长为 0.1,设定 k_1 和 k_2 取值介于 100~500 之间,步长为 1, m 为归一化常数, $n_clusters$ 与 γ 为谱聚类参数,具体设置如表 3 所示。

表 3 参数设置

| 参数 | k_1 | k_2 | m | Suc | $n_clusters$ | γ |
|----|-------|-------|--------|------|---------------|----------|
| 数值 | 300 | 100 | 15 000 | 0.62 | 2 | 8 |

2.3 实验结果

首先,使用本文方法对 140 篇电力科技文档进行领域词发现,同时将聚类效果进行可视化。如图 2 所示,圆点表示电力领域词,五角星为非电力领域词,可以发现,与预期的预测结果基本一致,谱聚类输出结果中,电力领域词基本被聚集在了一起。

为了更加直观展示本文提出的基于成词率和谱聚

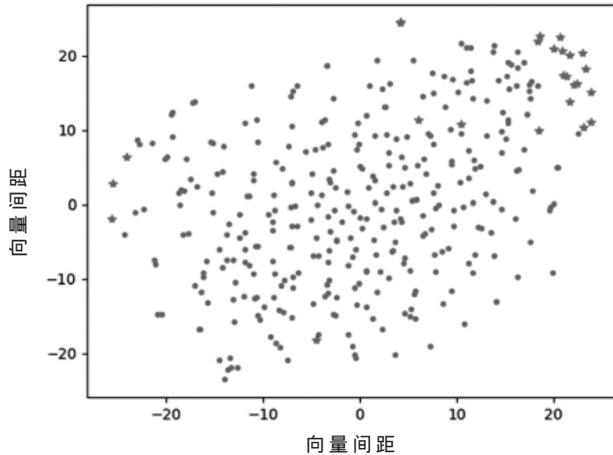


图2 谱聚类可视化效果

类的领域词发现方法的效果,进一步对聚类后的词语做部分展示,如图3所示。

电力领域词

[“变压器”, “低电压”, “配电网”, “无功补偿”, “电容器”, “分接开关”, “变电站”, “**自动化**”, “变压器绕组”, “过电压”, “电抗器”, “电信网”, “励磁器”, “避雷器”, “绝缘子”, “油纸绝缘”, “空间电荷”, “风电场”]

非电力领域词

[“控制策略”, “对抗样本”, “池化层”, “深度学习”, “智能化”, “**相位**”, “麦克斯韦方程”, “最优化”, “小波变换”, “微分”, “**母线**”, “计量”]

图3 本文方法用于电力文本领域词发现的效果展示(部分)

从图3中可以发现,筛选出的电力词汇是具有一定可信度的,但领域词中也有像“自动化”这种经常出现在电力科技文档里的词语,也有“电信网”这种带“电”字样却不是电力行业领域词的词语;非领域词中也有“相位”、“母线”这种个别被分错的电力领域词。

为了进一步验证提出方法的有效性,本文做了4个消融实验,实验结果如表4与图4所示。可以发现,本文提出的成词率相比传统运用互信息和左右熵的方法有明显优势;使用文本切分算法后,各项指标也有明显提升。谱聚类方法相比传统K-means聚类方法也有明显优势。值得注意的是,聚类算法会使召回率与聚类前相比有所降低,主要原因在于聚类操作会使部分电力领域词被错分为非领域词。

表4 实验结果展示

| 实验方法 | P'/% | R/% | F/% |
|-------------------|-------|-------|-------|
| 互信息+左右熵 | 50.15 | 78.56 | 61.22 |
| 成词率(互信息+左右熵+构词规则) | 61.23 | 81.05 | 69.54 |
| 成词率+文本切分 | 63.19 | 83.14 | 71.8 |
| 成词率+文本切分+K-means | 82.33 | 74.75 | 78.36 |
| 成词率+文本切分+谱聚类 | 84.51 | 76.64 | 80.38 |

3 结论

本文通过结合统计特征与语言规则,提出了成词率

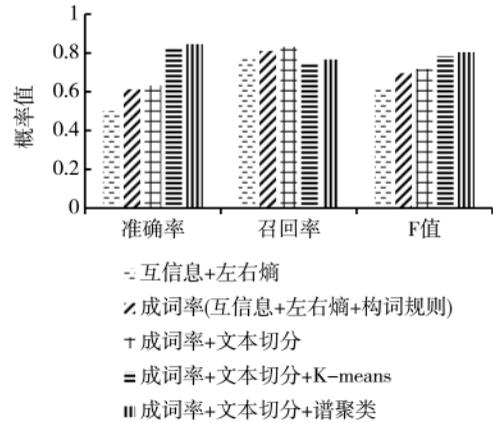


图4 实验结果展示

这一新的过滤指标,并结合谱聚类实现无监督电力科技文本领域词发现新方法。实验结果表明,所提方法与现有传统方法相比有效提升了候选词集筛选的准确性,同时在词语聚类时有效地解决了词向量维数过高的问题。

后续研究将着眼于对成词率指标的改进以及将词嵌入过程与聚类过程更加有效的融合,使其能够更准确地过滤非电力领域词,进一步提升方法效率和精度。

参考文献

- [1] 唐琳,郭崇慧,陈静锋.中文分词技术研究综述[J].数据分析与知识发现,2020,4(Z1):1-17.
- [2] 刘伟童,刘培玉,刘文锋,等.基于互信息和邻接熵的新词发现算法[J].计算机应用研究,2019,36(5):1293-1296.
- [3] 刘昱彤,吴斌,谢韬,等.基于古汉语语料的新词发现方法[J].中文信息学报,2019,33(1):46-55.
- [4] 杜丽萍,李晓戈,于根,等.基于互信息改进算法的新词发现对中文分词系统改进[J].北京大学学报(自然科学版),2016,52(1):35-40.
- [5] 马建红,张炳斐,张少光,等.基于主动MCNN-SCRF的新能源汽车命名实体识别[J].计算机工程与应用,2019,55(7):23-29.
- [6] Fu Guohong, LUKE K K.Chinese named entity recognition using lexicalized HMMs[J].ACM SIGKDD Explorations Newsletter,2005,7(1):225-226.
- [7] 陈飞,刘奕群,魏超,等.基于条件随机场方法的开放领域新词发现[J].软件学报,2013,24(5):1051-1060.
- [8] 郑家恒,李文花.基于构词法的网络新词自动识别初探[J].山西大学学报(自然科学版),2002(2):115-119.
- [9] 李明.针对特定领域的中文新词发现技术研究[D].南京:南京航空航天大学,2012.
- [10] 赵志滨,石玉鑫,李斌阳.基于句法分析与词向量的领域新词发现方法[J].计算机科学,2019,46(6):29-34.
- [11] 王鑫,吉久明,李楠,等.基于词向量的化学领域知识发

(下转第37页)

- 91-99.
- [9] LIN T, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017.
- [10] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018.
- [11] TAN M, PANG R, LE Q. EfficientDet: scalable and efficient object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2020.
- [12] QIAO S, CHEN L, YUILLE A. DetectoRS: detecting objects with recursive feature pyramid and switchable atrous convolution[J]. arXiv preprint arXiv: 2006.02334, 2020.
- [13] PENG H, DU H, YU H, et al. Cream of the crop: distilling prioritized paths for one-shot neural architecture search[J]. arXiv preprint arXiv: 2010.15821, 2010.
- [14] HU J, SHEN L, SUN G, et al. Squeeze-and-excitation networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 42(8): 2011-2023.
- [15] SANDLER M, HOWARD A, ZHU M, et al. Mobilenetv2: inverted residuals and linear bottlenecks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 4510-4520.
- [16] LIN T, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]//Proceedings of the IEEE International Conference on Computer Vision, 2017: 2980-2988.
- [17] WONG A, FAMUORI M, SHAFIEE M, et al. YOLO nano: a highly compact you only look once convolutional neural network for object detection[J]. arXiv preprint arXiv: 1910.01271, 2019.

(收稿日期: 2021-01-24)

作者简介:

赵义飞(1995-), 通信作者, 男, 硕士研究生, 主要研究方向: 深度学习、目标检测, E-mail: 723935706@qq.com。

王勇(1974-), 男, 博士, 副教授, 主要研究方向: 并行与分布式计算。



扫码下载电子文档

(上接第 28 页)

- 与实现[D]. 北京: 北京工业大学, 2018.
- [7] AI and compute[R]. OpenAI, 2018.05.16.
- [8] 盎司财经. 区块链搭建“生产关系”新模式, 较人工智能更具“革命性”[EB/OL]. (2018-03-13)[2020-12-15]. http://www.sohu.com/a/225432740_99987131.
- [9] 潘吉飞, 黄德才. 区块链技术对人工智能的影响[J]. 计算机科学, 2018, 45(11A): 53-37.
- [10] Bytom WhitePaper V1.0[R]. BYTOM, 2017, 06.
- [11] 赵泓维. 对于医疗人工智能企业算力问题, 英伟达打出关键一招[EB/OL]. (2019-01-29)[2021-01-19]. <http://www.chidaolian.com/article-23081-1>.
- [12] 中国区块链技术和应用发展白皮书(2016)[R]. 中国区块链技术和产业发展论坛, 2016.
- [13] 链门户. 什么是智能合约? 智能合约真的智能吗?[EB/OL]. (2018-08-14)[2021-01-20]. <http://www.lianmenhu.com/>

(上接第 32 页)

- 现初步研究[J]. 图书馆杂志, 2018, 37(11): 90-98.
- [12] 吉久明, 施陈炜, 李楠, 等. 基于 GloVe 词向量的“技术—应用”发现研究[J]. 现代情报, 2019, 39(4): 13-22.
- [13] FARMER W J, RIX A J. Evaluating power system network inertia using spectral clustering to define local area stability[J]. International Journal of Electrical Power and Energy Systems, 2022, 134(3): 107404.
- [14] 耿丽君. 韵律形态学研究综述[J]. 成都理工大学学报(社会科学版), 2020, 28(1): 98-104.
- [15] 杨飘, 董文永. 基于 BERT 嵌入的中文命名实体识别方

- blockchain-5574-6.
- [14] 李庆华. 智能合约——智能合约安全问题的 AI 解决方案[EB/OL]. (2018-05-04)[2021-01-14]. <https://cloud.tencent.com/developer/news/202428>.
- [15] 区块链头条. JarvisPlus 创始人兼 CEO 吴骞: 让每个人可以用自然语言来使用区块链和智能合约[EB/OL]. (2018-07-30)[2021-01-16]. https://www.sohu.com/a/244247557_100112552.

(收稿日期: 2021-02-22)

作者简介:

张伟娜(1985-), 通信作者, 女, 硕士, 工程师, 主要研究方向: 人工智能技术应用、产业发展应用, E-mail: zhwn0704@163.com。

黄蕾(1975-), 女, 博士, 高级工程师, 主要研究方向: 人工智能产业发展应用。

张箴(1985-), 男, 本科, 初级工程师, 主要研究方向: 人工智能产业发展应用。



扫码下载电子文档

法[J]. 计算机工程, 2020, 46(4): 40-45.

(收稿日期: 2021-03-01)

作者简介:

杨政(1987-), 男, 硕士研究生, 高级工程师, 主要研究方向: 电力文本分析与应用、电网数字化以及网络安全。

尹春林(1991-), 男, 助理工程师, 主要研究方向: 自然语言处理、迁移学习。

李慧斌(1984-), 男, 博士, 副教授, 主要研究方向: 计算机视觉、图像处理与模式识别、深度学习、自然语言处理。



扫码下载电子文档

版权声明

经作者授权，本论文版权和信息网络传播权归属于《电子技术应用》杂志，凡未经本刊书面同意任何机构、组织和个人不得擅自复印、汇编、翻译和进行信息网络传播。未经本刊书面同意，禁止一切互联网论文资源平台非法上传、收录本论文。

截至目前，本论文已经授权被中国期刊全文数据库（CNKI）、万方数据知识服务平台、中文科技期刊数据库（维普网）、DOAJ、美国《乌利希期刊指南》、JST 日本科技技术振兴机构数据库等数据库全文收录。

对于违反上述禁止行为并违法使用本论文的机构、组织和个人，本刊将采取一切必要法律行动来维护正当权益。

特此声明！

《电子技术应用》编辑部

中国电子信息产业集团有限公司第六研究所