

基于对话文本智能挖掘的电力设备供应商评价

王海瑶¹,王慧芳¹,胡俊华²,徐积全²,李建红³,何奔腾¹

(1. 浙江大学 电气工程学院,浙江 杭州 310027;2. 国网浙江省电力有限公司,浙江 杭州 310007;
3. 浙江华云信息科技有限公司,浙江 杭州 310012)

摘要:目前电网企业的电力设备供应商满意度评价主要依赖于人工统计和指标计算,其准确性受评价人员和评价内容的影响较大。以电力业务平台的对话文本为研究对象,在扩充已有电力本体词典的词条和属性的基础上,建立了基于文本挖掘技术的电力设备供应商评价模型。首先提出了基于Transformer的双向编码器下句预测与余弦相似度加权的单轮对话文本下句预测分析方法,建立了对话中断交叉处理流程和供应商识别规则,实现了电力对话文本主题归纳;然后考虑对话文本语义情感的复杂性,提出了对话情感分析规则,建立了供应商评价模型。最后通过算例验证了所提方法的准确性,结果表明基于对话文本智能挖掘的电力设备供应商评价,具有可行性和有效性,可以作为目前评价方法的有益补充。

关键词:电力设备;文本智能挖掘;对话文本;下句预测;情感分析;供应商评价

中图分类号:F 274

文献标志码:A

DOI:10.16081/j.epae.202104009

0 引言

电力设备质量至关重要。我国电网企业在电力设备供应商评价过程中,通常以履约情况、售后服务和质量监督作为评价内容。国家电网公司发布《2020年第一次供应商评价评分细则(电能表、集中器、专变采集终端)》中新增了满意度评价^[1],要求项目单位通过人工统计、直接打分的模式,根据供应商质量、服务、供货情况进行评价。然而,这种评价方式不仅效率低,准确性受评价人员经验、评价材料范围等影响,还可能出现供应商仅重视评价内容而轻视其他质量所造成的评价结果失真问题。随着腾讯通RTX(Real Time eXpert)、微信、钉钉等通讯软件的广泛应用,电网企业逐渐在各类业务平台建立起了工作群,管理和技术人员通过文字(语音可转文字)在工作群中进行交流时,会提及更为细粒度的供应商的设备质量、服务等信息。这类信息来自一线人员,具有基础性、广泛性和高可信度等特点,可作为评价设备供应商的支撑数据之一。因此,对电力对话文本进行智能挖掘,对提高设备供应商评价的准确性具有重要意义,能够从日常工作实践中获得供应商的实际设备运行质量和服务水平。此外,领域对话文本中还蕴藏着事件处理应对、专家经验评估等众多信息,因此对领域对话文本挖掘技术进行探索,具有更为广泛和深厚的应用意义。

目前,文本挖掘为文本内容潜在价值的发掘与分析提供了技术手段。在电力领域,已有应用于设备缺陷、运维和告警信号文本的挖掘研究^[2-11],在一定程度上减少了人工分析缺陷的工作量,提高了分析的准确性,为电力领域文本挖掘的应用开拓了思

路。与客观描述电力设备缺陷、运维和告警信号的文本相比,对话文本具有主观性强、句法多变、内容松散的特点^[12],对话中会出现情感倾向的词语,因此对电力对话文本的挖掘需要结合情感分析。情感分析又称为评论挖掘或意见挖掘^[13],主要可以分为基于情感词典的方法和基于机器学习的方法^[14]。目前中文文本情感分析多应用于微博^[15-16]、商品评论^[17]等领域,研究多建立在以句或段为分析单位,尚未有考虑对话文本上下文特性及对话交叉中断情况的中文情感分析研究。

针对电力对话文本特点,本文提出一种考虑电力对话文本特性的电力设备供应商评价方法:在电力对话文本特性特征研究的基础上,提出基于双向编码器下句预测BERT-NSP(Bidirectional Encoder Representations from Transformers-Next Sentence Prediction)与余弦相似度加权的方法,对电力对话文本进行上、下句连贯性判断,并通过单轮对话下句匹配和对话中断交叉判断的研究,实现对话主题归纳;研究对话文本的情感分析,并设计情感判断与设备供应商评价对应规则,建立电力设备供应商评价模型。

1 电力对话文本特点及供应商评价过程

1.1 电力对话文本特点

(1)含有无关冗余内容。由于文本是以聊天的形式产生的,因此对话内容既包含电力领域工作中遇到的问题和解决办法,以及对产品质量的倾向性评价,也会掺杂无关内容,如讨论社会新闻、行业前景等。

(2)句法形式多样,口语化突出。对话文本的语法往往比较个性化、形式自由、没有固定格式,且以短句居多,口语化突出,包括有较多网络用语以及常

用简称。

(3)存在隐式评价对象。在上下句语义背景下会出现评价对象省略的情况,如连续对话文本“A厂家表计时钟异常有一串英文怎么办?”“他们的表计经常出现时钟异常情况啊”,第二句中的“他们”指代“A厂家”,因此存在隐式评价对象。

(4)对话存在交叉中断现象。定义多句对话文本为多轮对话,有且仅有上、下两句的对话文本为单轮对话。多轮对话中语义相关的内容为同一主题,其中讨论同一供应商的对话内容为同一供应商主题。多轮对话的特点是长度较长、跨度较广,由于聊天对话的自由度高、主题灵活,往往会出现对话交叉或中断现象。如在几个工作人员讨论表计故障问题时,会交叉其他人员的闲聊内容,或相隔较长对话通过链接@用户ID进行互动,此时多轮对话文本的同主题划分较为困难。

(5)语义情感丰富且存在情感转变。对话文本中的情感词丰富多样,情感表达更为灵活,且存在情感转变情况。如“轮换表计之前还挺好的,怎么换完之后就不行了”,一句中表达了2种相反的情感,对该句进行准确的情感挖掘具有一定难度。

1.2 基于对话文本的供应商评价过程

针对上述电力对话文本特点,提出如附录中图A1所示的供应商评价过程,具体说明如下。

(1)首先进行电力本体词典扩充研究。由于电力对话文本中存在专业性强的特点,与常用词库存在一定差别,为提升文本理解的准确率,需要建立对话业务领域的本体词典,并对词典新增属性条目,包括电力专有词汇、供应商名称词汇和事件关键词汇,为后续的主题归纳和情感分析评价奠定基础。本文采用半监督的分词方法,在课题组初步建立的电力设备缺陷领域本体词典^[18]和通用词典基础上,以对话的电力业务领域文本为语料库基于隐马尔科夫模型HMM(Hidden Markov Model)进行分词,按照词频排序筛选新的领域本体词汇,由人工核查确定是否成为本体词以及近义词,补充本体词典。随着语料库的不断增加,重复上述步骤进行多次校正,实现了电力领域本体词典的补充与丰富。本文研究新增的本体词典主要包括供应商名称、业务领域事件等词汇,扩充供应商名称及其常用简称以应对对话文本口语化突出的特点。

(2)然后进行主题归纳研究。由于多轮对话主题的灵活性,难以得出对供应商准确的情感倾向,需要将多轮对话拆解为多个独立的对话主题,然后通过上、下句的连贯性判断是否语义相关,故将多轮对话拆解为多个单轮对话进行下句预测判断,作为对话文本同主题归纳的重要依据。然而,单轮对话下句预测判断结果为不相关,并不一定标志着该主题

的结束。有可能是由于存在对话中断交叉情况,故应在满足对话中断交叉处理规则的基础上进行对话主题归纳。对同一主题出现多供应商或无供应商的情况,需要进行针对供应商的主题划分,识别对话集合内的隐式评价对象,同时去除对话中无关供应商的冗余内容。

(3)最后进行基于情感分析的评价建模研究。情感分析是在供应商主题归纳完成后,识别与情感相关的关键词,通过分析情感评价基础单元和辅助单元的语义情感,克服句法多样化困难,识别对话文本语义情感转变。最后基于情感分析结果,建立基于对话文本的供应商评价模型。

2 电力对话文本主题归纳技术

2.1 基于BERT-NSP与余弦相似度加权的单轮对话下句预测分析

为了便于对多轮对话文本进行同主题归类分析,先将多轮对话文本拆解为仅有上、下2句对话的单轮对话文本,以单轮对话为分析单位进行下句预测,并以此为基础进行对话中断交叉处理及供应商识别。

BERT-NSP是通过维基百科语料训练得到的一种用于判断上、下句连贯性的模型,在多项自然语言处理任务中取得了较好的结果^[19-22]。文献[19]测试了BERT在通用语言理解评估GLUE(General Language Understanding Evaluation)基准数据集上的任务,相对双向长短记忆网络BiLSTM(Bi-directional Long Short-Term Memory)和OpenAI生成式预训练GPT(Generative Pre-Training)等算法具有更好的表现,同时比较了BERT-NSP和去除NSP情况下利用左到右LTR(Left-to-Right)模型、LTR+BiLSTM等算法进行训练的表现,得到了BERT-NSP最优的训练结果。本文采用基于BERT-NSP的上、下句连贯性预测微调模型,该模型是在BERT预训练模型的基础上,在有监督的训练过程中对训练参数进行微调的模型,结构如图1所示。

模型以2条对话文本为输入,添加第1个标记[CLS],并在2条对话文本中间插入标记[SEP]作为特殊分隔符;输入层各项 T_{tok_i} ($i=1,2,\dots,N$)、 T'_{tok_j} ($j=1,2,\dots,F$)为输入语句中每个字的向量化, N 和 M 分别为2句文本的字数,由标记词嵌入、分段词嵌入、位置词嵌入3种词嵌入方式相加得到一一对应的 $E_{[\text{CLS}]}$ 、 E_1-E_N 、 $E_{[\text{SEP}]}$ 、 $E'_1-E'_F$,传入双向Transformer层中,对其进行变换输出每个字对应的隐藏向量 C 、 T_1-T_N 、 $T_{[\text{SEP}]}$ 、 $T'_1-T'_F$ 。模型输出具体如式(1)所示^[19]。

$$p = \text{softmax}(CW^T) \quad (1)$$

其中, p 为下句预测匹配概率矩阵; C 为BERT模型第1个标志[CLS]的最终隐藏状态; W 为全连接层权重

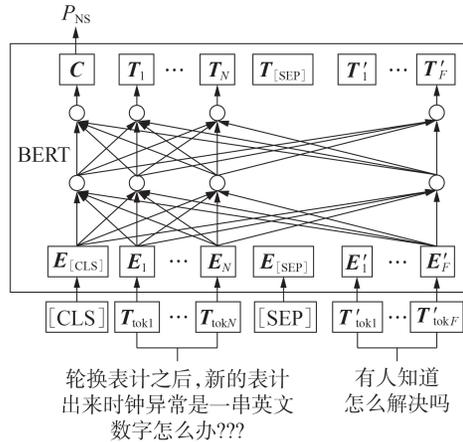


图1 基于BERT-NSP的上、下句连贯性预测微调模型

Fig.1 Fine-tuning model for prediction of coherence of upper and lower sentences based on BERT-NSP

矩阵。本模型实际是一个二分类问题,因此 p 为一个2维向量,其元素分别表示下句预测为0和1的概率值,即2句不相关和相关的概率,下句预测概率 P_{NS} 取其中表示2句相关的概率值。

由图1可知,BERT-NSP模型是以字为单位进行的深度特征预测,无形中忽略了词语级特征,而文本分析中较为常用的余弦相似度则具备从词语层面探索2句文本相关性的因素,因此可以构建基于BERT-NSP与余弦相似度加权的下句预测算法,目的是融合BERT-NSP深度分布分析的优点和余弦相似度词语级分析的优点,从而提升对话文本下句预测的准确性。

余弦相似度方法在判断对话文本上、下连贯性时,文本中重复出现的内容往往会作为连贯评判标准,因此对单轮对话文本进行分词,采用式(2)可计算得到相邻对话的余弦相似度^[23]。

$$S = \cos(A, B) = \frac{\sum_{i=1}^n a_i b_i}{\sqrt{\sum_{i=1}^n a_i^2} \sqrt{\sum_{i=1}^n b_i^2}} \quad (2)$$

其中, S 为相邻对话的余弦相似度; $A=(a_1, a_2, \dots, a_n)$ 和 $B=(b_1, b_2, \dots, b_n)$ 分别为2句文本词频向量化表示后获得的 n 维词频特征向量。在得到深度特征和语言相似度特征的基础上,本文定义单轮对话语义相关性匹配度,由式(3)对这2类特征做一融合:

$$M = (1 - \alpha)P_{NS} + \alpha S \quad (3)$$

其中, M 为单轮对话语义相关性匹配度, $M \geq 0$; α 为余弦相似度权重系数。 M 越大表示2句匹配相关性越大,参考 P_{NS} 的二分类取值标准,若 $M \geq 0.5$ 则认为上、下句相关,将上、下2句划归为同一对话主题;若 $M < 0.5$ 则认为上、下句不相关,其意义在于能够融合深度特征和相似度特征,综合考虑上、下句的语言联系,提高匹配判断的准确性。 α 的作用是平衡深度

特征和相似度特征的权重比例,对该系数进行寻优可获得单轮对话文本判断最优模型。

2.2 对话中断交叉处理

在电力对话文本中,往往会出现多轮对话文本的主题中断交叉情况,为了在单轮对话下句匹配基础上实现对话主题归纳,设计了如附录中图A2所示的对话中断交叉处理流程。分别设从对话文本集合 D 中顺序取出的2条文本 d_i 和 d_j ,判断2条文本的间隔轮次。根据对话文本研究经验,认为间隔轮次在3轮之外的对话一般无直接相关关系。如果2条文本在间隔允许范围内,则对2条文本进行单轮对话相关性匹配度 M 值判断;如果2条文本不在间隔允许范围内,则对 d_i 进行链接@用户ID信息判断。如果含有链接信息,就将含链接语句 d_i 依次和被链接用户间隔轮次最近的上、下2条对话文本分别进行相关性匹配度 M 值判断,根据匹配度判断结果进行同主题对话集合归纳;如果不含链接信息,则认为该条文本所属主题对话集合已归纳完毕。继续按照上述步骤重复递归调用,直至对话文本集合 D 为空,获得全部同主题对话集合。

2.3 供应商识别规则

获得相同主题对话集合后,依据电力业务本体词典中的供应商信息类别,提取每个主题对话集合中的供应商信息,采用向上就近原则识别其中的隐式评价对象,然后去除无关的冗余主题内容。具体采用以下3种规则。

(1)未识别到供应商信息,则认为该对话集合谈论对象为无关冗余内容,对设备供应商评价无价值,可剔除。

(2)识别到1个或多个相同供应商信息,则认为该对话集合评价对象为被识别供应商。

(3)出现2个及以上不同供应商信息,如按出现顺序分别为厂家A、B、...、X。对集合内的每条文本采用向上就近原则确定对应供应商,定义该对话集合自第一句至厂家B出现句之前(不包括厂家B出现句)的评价对象为厂家A,厂家B出现句至厂家C出现句之前(不包括厂家C出现句)的评价对象为厂家B,依此类推。若某厂家重复出现,则对该厂家的对话集进行合并。

至此,实现各供应商的主题文本归纳。

3 基于电力对话文本情感分析的供应商评价模型

3.1 对话情感分析规则建立

电力对话文本情感分析最重要的步骤是找到句子中与情感相关的关键词。通过对大量电力对话文本的观察和分析,发现影响文本中情感评价的主要词汇有事件词 I_{nc} 、情感词 S_{en} 、连词 C_{onj} 、程度副词 A_d

和否定词 N_{eg} ,上述词汇对于评价对象的作用有一定规律可循。本文在电力领域本体词典扩充时,已对事件词 I_{nc} 、情感词 S_{en} 进行了标注,并设置为情感评价基础单元;同时选择连词 C_{onj} 、程度副词 A_d 和否定词 N_{eg} 作为辅助单元用于修正文本情感分析。针对具体供应商的主题文本,对每一句进行分词、词性检索和标注,并按如下规则进行冗余句子删除和情感计分。

(1)在单句对话文本中,基础单元事件词 I_{nc} 和情感词 S_{en} 中出现至少一个时,认为该句有情感倾向和评价意见,否则认为该句为中性,对设备供应商评价无价值,可删除。

(2)当句中仅出现 I_{nc} 时,计入事件分;当句中仅出现 S_{en} 时,计入情感分;当句中 I_{nc} 和 S_{en} 并存时,各自计入事件分和情感分,同时需要确定基础单元评价优先系数 η 和辅助单元与基础单元的依附关系。

(3)事件词 I_{nc} 作为专业性且多为描述故障发生时的词汇, I_{nci} 为第 i 个事件词的情感评分,根据事件严重程度设定评分,取为-1或-2。

(4)情感词 S_{en} 参考BosonNLP情感词典,根据情感正向和负向设定评分, S_{enj} 为第 j 个情感词的情感评分,取值范围为(-2,2)。

(5)程度副词 A_d 根据极性强度设定极性系数取为0.5、1或2, A_{dk} 为第 k 个程度副词的情感极性系数,若无程度副词则取 $A_{d0}=1$ 。

(6)否定词 N_{eg} 取否定系数为-1, N_{egl} 为第 l 个否定词的否定系数,若无否定词则取 $N_{eg0}=1$ 。

(7)当句中存在多个事件词 I_{nc} 或情感词 S_{en} ,需要对句子情感语义和转变进行判断。当存在连词 C_{onj} 时,根据连词和基础单元的位置关系及连词转折关系判断情感语义评价优先级,确定连词关系语义评价优先系数 η_c 。其中,以连词为界限,连词前的基础单元优先系数为 η_{cb} ,连词后的基础单元优先系数为 η_{ca} 。对于多个基础单元间的连词,两侧的优先级设置如下:对于表示转折和因果关系的连词,如“但是”、“所以”,连词两侧基础单元的评价优先系数为 $\eta_{cb}=0.8$ 、 $\eta_{ca}=1.2$;对于表示并列和承接关系的连词,如“而且”、“然后”,连词两侧基础单元的评价优先系数 $\eta_{cb}=\eta_{ca}=1$;对于表示让步关系的连词,如“尽管”、“虽然”,连词两侧基础单元的评价优先系数为 $\eta_{cb}=1.2$ 、 $\eta_{ca}=0.8$;若无连词,则对于所有基础单元 $\eta_c=1$ 。

(8)当句中同时存在事件词 I_{nc} 和情感词 S_{en} 时,事件词作为客观陈述事实类词语,其评价信服力高于主观情感词,事件词 I_{nc} 和情感词 S_{en} 基础单元关系评价优先系数 η_v 分别为 $\eta_{v1}=1.2$ 和 $\eta_{v2}=0.8$ 。

(9)对于句中每个基础单元,综合规则(7)、(8),

其评价优先系数 η 为:

$$\eta = \eta_c \eta_v \quad (4)$$

(10)对于基础单元事件词 I_{nc} 和情感词 S_{en} 与程度副词 A_d 和否定词 N_{eg} 的依附关系,若有连词 C_{onj} 则以连词为划分界限,若无连词则以每个基础单元为划分界限。

3.2 基于情感分析的供应商评价模型

针对同主题对话集合,需结合情感评价基础单元和辅助单元的关系特征给出供应商评分。每个基础单元受该单元优先系数、程度副词和否定词系数直接影响,因此通过系数串联乘积对基础单元评分直接修正;基础单元之间互相独立,通过并联加法联系。对于同主题下同评价对象供应商,计算其以句为单位的感情评分值,该值的正、负可直观反映对话者的正面、负面评价,该值的大小则反映了情感的强度。由此,建立了如式(5)所示的单句供应商评分。

$$S_{MANu} = \sum_{i=1}^I \eta_{li} I_{nci} \prod_{k=0}^K A_{dk} \prod_{l=0}^L N_{egl} + \sum_{j=1}^J \eta_{sj} S_{enj} \prod_{p=0}^P A_{dp} \prod_{q=0}^Q N_{egq} \quad (5)$$

其中, S_{MANu} 为单句供应商评分值; I 为句中事件词的数量; η_{li} 为第 i 个事件词的优先系数,由式(4)计算得到; K 为每个事件词附属程度副词的数量; L 为每个事件词附属否定词的数量; J 为句中情感词的数量; η_{sj} 为第 j 个情感词的优先系数,由式(4)计算得到; P 为每个情感词附属程度副词的数量; Q 为每个情感词附属否定词的数量; k 、 l 、 p 、 q 取值从0开始,表示不存在对应辅助单元时该系数取其初始值1。供应商的总评价分为所有单句评分之和:

$$S_{MAN} = \sum_{u=1}^U S_{MANu} \quad (6)$$

其中, S_{MAN} 为供应商总评价分; U 为供应商评价句的数量。

4 算例分析

4.1 算例语料库及本体词典扩充情况

以电力采集运维领域对话文本为例进行验证研究。首先建立语料库,搜集了包括电力采集运维RTX工作群聊对话、《电网企业一线员工作业采集异常一本通》导则、采集异常文本在内的共23.8 MByte文本数据。然后基于HMM对话料库进行分词、整理并调整。该过程是在通用词典和已有的电力设备缺陷本体词典基础上,识别新的本体词,由人工核查后补充形成了新的领域本体词典。电力采集运维领域本体词典示例如附录中表A1所示。表中,对本体词标注了词条属性和同义词、近义词,属性包括专有领域名词、供应商名称和事件关键词汇。新增采集运

维领域本体词典共计752条,包括在采集运维语料库中出现的专业词汇但原有本体词典未包含的专有领域名词词条539条,供应商名称类词汇106条,事件关键词汇107条。

4.2 主题归纳方法验证

数据集选取电力采集运维RTX工作群聊对话文本中的347对单轮对话文本,附录中表A2列出了其中13条对话内容。BERT-NSP模型参数使用12层Transformer单元 1.1×10^8 个参数的BERT-Chinese预训练模型结构,自我注意力机制为12头,隐含层维度为768维,最大序列长度为128,学习率为 3×10^{-5} ,批量大小为32。基于BERT-NSP与余弦相似度加权的单轮对话下句预测分析对347对电力供应商的主题对话内容进行实验,对单轮对话语义相关性匹配度 M 中的余弦相似度权重系数 α 进行寻优,准确率如图2所示。图中, $\alpha=0$ 和 $\alpha=1.0$ 分别表示BERT-NSP模型和余弦相似度模型。由图可见,单轮对话判断准确率在 $\alpha=0.04$ 时最大,为80.69%, α 超过0.04后准确率单调递减。同时,选取BiLSTM-Softmax分类模型进行上、下句同主题分类效果对比实验,模型准确率指标如表1所示。由表可见,采用权重系数 $\alpha=0.04$ 的BERT-NSP与余弦相似度加权模型准确率最高。

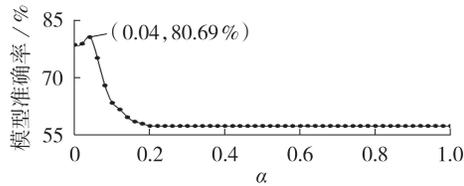


图2 α 影响下的BERT-NSP与余弦相似度加权准确率

Fig.2 Weighted accuracy of BERT-NSP and cosine similarity under influence of α

表1 单轮对话文本判断准确率

Table 1 Accuracy of single-round dialogue text judgment

模型	准确率 / %
余弦相似度($\alpha=1$)	57.35
BiLSTM-Softmax	74.64
BERT-NSP($\alpha=0$)	78.67
BERT-NSP与余弦相似度加权($\alpha=0.04$)	80.69

由图2和表1可知,本文的模型在BERT-NSP判断2句深度特征相关性的基础上与语言特征相似度加权,能够提高单轮对话判断的准确性。该加权模型的确立同时也具有一定可解释性,对话过程中出现相同文本内容更倾向于为相同主题的讨论。

在单轮对话文本下句预测分析的基础上,处理对话交叉中断情况。以附录中表A2所示的电力对话文本为例,该段对话包含2个对话主题,分别为表

计异常供应商讨论主题和陶瓷杯性价比讨论主题,比较对话交叉中断处理通过无处理、余弦相似度、BiLSTM-Softmax、BERT-NSP、BERT-NSP与余弦相似度加权模型后得到的多轮对话主题划分情况如表2所示,表中同供应商识别情况下的不同对话主题以“/”划分。从表2可以看出,在多轮对话主题划分方面,仅使用余弦相似度计算的模型主题划分准确率最低,加权模型准确率最高。相比较而言,余弦相似度模型主要看重2句对话间的文本重复率,仅依靠2句中的重复内容来判断连贯性,忽视内在联系;BiLSTM-Softmax模型由于将关键评价句的第1、2句划分为单独主题,导致其中未包含供应商信息而被划归为冗余内容;BERT-NSP模型虽然能够达到较高的准确率,但仍然出现划分不完整的情况,针对划分错误的第4、7、13句分析可知,通过相同关键词“陶瓷杯”可以判断这3句为同一主题;因此加权模型综合以上两者的优点,克服BERT模型忽略词语级联系的不足,更准确地划分各主题的范围。从加权模型的对话集合匹配结果可见,采用附录中图A2所示的对话交叉中断处理流程可以准确划分同主题内容,不仅可以正确划分对话间隔轮次在3句以内的同主题,对于类似第4、13句通过链接@用户ID进行的远距离对话也能准确划分主题。此外,供应商识别规则能够正确识别同主题下对应厂家信息,并将冗余对话从对话主题中删除。

表2 不同模型多轮对话主题划分情况

Table 2 Topic division of different models in multiple rounds of dialogue

模型	多轮对话主题划分情况		多轮对话主题划分准确率 / %
	对话编号	供应商识别	
无处理	8	A厂家	—
	11	B厂家	
	12	A厂家	
余弦相似度	1/2/3/4/5/6/ 7/9/10/13	冗余	53.85
	8—13	A厂家、B厂家	
BiLSTM-Softmax	3、5、6、8—12	A厂家、B厂家	76.92
	(1,2)/(4,7)/13	冗余	
BERT-NSP	1—3、5、6、8—12	A厂家、B厂家	84.61
	4/7/13	冗余	
BERT-NSP与余弦相似度加权	1—3、5、6、8—12	A厂家、B厂家	100
	4、7、13	冗余	

4.3 情感分析与供应商评价验证

为了展示对话文本情感分析流程,以“A厂家原来挺好的,但是现在咋经常出来时钟异常?”一句为例进行关键词提取,在电力采集运维领域本体词典和通用词典的基础上对文本进行预处理,分词并进行关键词提取标注,如表3所示。

表3 对话文本情感分析关键词
Table 3 Keywords of sentiment analysis
of dialogue text

基础单元	供应商	I_{nc}	S_{en}	C_{onj}	A_d	N_{eg}
1	A厂家		好	但是	原来、挺	
2	A厂家	时钟异常			经常	

该句共有2个基础单元,情感词“好”和事件词“时钟异常”,通过连词“但是”连接,同时辅助单元程度副词“原来”、“挺”和“经常”分别对其修正。根据式(5),对于情感词“好”和事件词“时钟异常”而言,它们的优先系数是由事件词优先系数 $\eta_{vs} < \eta_{vl}$ 和连词“但是”优先系数 $\eta_{cb} < \eta_{ca}$ 决定的,在优先级判定条件下二者为并联加法关系;情感词“好”受程度副词“原来”、“挺”影响;“时钟异常”同时受程度副词“经常”直接影响,各自具有串联乘积关系。

在此基础上进行横向对比,以语义转变、语义顺延和普通单句为例,测试本文方法对于语义转折的判断准确性,得到如表4所示的分析评分结果。语义转变句中存在着有正面评价“好”和事件词“时钟异常”,通过情感分析得到负向评分;正向语义顺延句则在正向单句评分基础上有更高的评分值。因此本文方法对于语义丰富的对话文本能够做到较为准确的评分,对于含有语义转变的内容也能在评分上体现出明显的区别。

表4 对话文本情感分析语义对比评分
Table 4 Sentiment analysis semantic comparison
score of dialogue text

分析语句	评分结果
原来挺好的,但是现在咋经常出来时钟异常?	-2.26
原来挺好的,现在也不错	1.28
原来挺好的	0.77

利用4.2节主题划分结果,对附录中表A2进行情感分析验证供应商评价效果,如表5所示。可以看出,BERT-NSP和BERT-NSP与余弦相似度加权得到的评分结果由于主题划分一致而结果相同,与无处理和余弦相似度计算模型下的评分相比差别较大。对于A厂家评分,BERT-NSP和加权模型的评分与文本表达负面情感程度更为相近,无处理模型仅选取句中含有A厂家信息的句子作为评分对象,其他相关句缺失,而余弦相似度模型同样缺失部分信息,导致评分结果不佳,BiLSTM-Softmax模型则由于缺失评价句第1、2句导致评分受到影响,负向评价明显减轻;对于B厂家评分,由于仅涉及一句对话,因此评分未受影响。由此可见,基于情感分析的供应商评价结果受主题划分正确性影响较大,加权模型在主题划分上的准确性可以有效提高供应商评价的真实性和可靠性。

表5 对话文本情感分析供应商评分
Table 5 Supplier rating based on sentiment
analysis of dialogue text

模型	评分结果	
	A厂家	B厂家
无处理	0.86	-0.80
余弦相似度	0.05	-0.80
BiLSTM-Softmax	-2.88	-0.80
BERT-NSP	-5.61	-0.80
BERT-NSP与余弦相似度加权	-5.61	-0.80

5 结论

针对电力对话文本的特点,本文提出了基于BERT-NSP和余弦相似度加权模型进行单轮对话下句预测分析,构建了基于多轮对话交叉中断的多轮对话主题划分,并识别供应商主题,在此基础上建立基于情感分析的供应商评价模型。大量算例表明,本文方法在多轮对话主题划分上具有优越性,供应商评价模型具有可靠性,为供应商评价的常态化提供了方法。

此外,基于文本挖掘技术的电力设备供应商评价模型具有很好的移植能力。本文以电力采集运维业务对话文本为例进行本体词典的补充及评价模型的建立,可结合电力其他业务文本语料库进行相应移植研究,不仅为业务对话文本的深度利用提供了新思路,而且为电网企业设备供应商满意度评价提供了新的支撑依据,进而帮助电网企业更客观地了解供应商的设备质量及服务能力,做出合理决策,同时有益于设备供应商有针对性地改善不足,提升竞争力。

然而,对话是一种与时俱进的交流系统,在时代更替的进程中,存在语言的新用法出现、旧用法淘汰的情况,因而基于电力对话文本情感分析的供应商评价模型受规则建立的动态性、完整性影响较大,如何进一步实现智能分析是后续的研究方向。

附录见本刊网络版(<http://www.epae.cn>)。

参考文献:

- [1] 国家电网有限公司营销部. 2020年第一次供应商评价评分细则(电能表、集中器、专变采集终端)[EB/OL]. (2020-04-29) [2020-12-03]. <http://shupeidian.bjx.com.cn/html/20200429/1068179.shtml>.
- [2] 邵冠宇,王慧芳,吴向宏,等. 基于依存句法分析的电力设备缺陷文本信息精确辨识方法[J]. 电力系统自动化, 2020, 44(12): 178-189.
SHAO Guanyu, WANG Huifang, WU Xianghong, et al. Precise information identification method of power equipment defect text based on dependency parsing[J]. Automation of Electric Power Systems, 2020, 44(12): 178-189.
- [3] 郑翔,王慧芳,严娴峥,等. 用于APP的缺陷文本自动分类与质量保证方法[J]. 电力系统及其自动化学报, 2020, 32(7): 131-136.
ZHENG Xiang, WANG Huifang, YAN Xianzheng, et al. Defect

- text automatic classification and quality assurance method for APP[J]. Proceedings of the CSU-EPSA, 2020, 32(7): 131-136.
- [4] 邵冠宇, 王慧芳, 何奔腾. 电网设备缺陷文本的质量评价与提升方法[J]. 电网技术, 2019, 43(4): 1472-1479.
SHAO Guanyu, WANG Huifang, HE Benteng. Quality assessment and improvement method for power grid equipment defect text[J]. Power System Technology, 2019, 43(4): 1472-1479.
- [5] 刘梓权, 王慧芳. 基于知识图谱技术的电力设备缺陷记录检索方法[J]. 电力系统自动化, 2018, 42(14): 158-164.
LIU Ziquan, WANG Huifang. Retrieval method for defect records of power equipment based on knowledge graph technology[J]. Automation of Electric Power Systems, 2018, 42(14): 158-164.
- [6] 刘梓权, 王慧芳, 曹靖, 等. 基于卷积神经网络的电力设备缺陷文本分类模型研究[J]. 电网技术, 2018, 42(2): 644-650.
LIU Ziquan, WANG Huifang, CAO Jing, et al. Power equipment defect text classification model based on convolutional neural network[J]. Power System Technology, 2018, 42(2): 644-650.
- [7] 曹靖, 陈陆桑, 邱剑, 等. 基于语义框架的电网缺陷文本挖掘技术及其应用[J]. 电网技术, 2017, 41(2): 637-643.
CAO Jing, CHEN Lushen, QIU Jian, et al. Semantic framework-based defect text mining technique and application in power grid[J]. Power System Technology, 2017, 41(2): 637-643.
- [8] 邱剑, 王慧芳, 应高亮, 等. 文本信息挖掘技术及其在断路器全寿命状态评价中的应用[J]. 电力系统自动化, 2016, 40(6): 107-112, 118.
QIU Jian, WANG Huifang, YING Gaoliang, et al. Text mining technique and application of lifecycle condition assessment for circuit breaker[J]. Automation of Electric Power Systems, 2016, 40(6): 107-112, 118.
- [9] 蒋逸雯, 李黎, 李智威, 等. 基于深度语义学习的电力变压器运维文本信息挖掘方法[J]. 中国电机工程学报, 2019, 39(14): 4162-4171.
JIANG Yiwen, LI Li, LI Zhiwei, et al. An information mining method of power transformer operation and maintenance texts based on deep semantic learning[J]. Proceedings of the CSEE, 2019, 39(14): 4162-4171.
- [10] 孙国强, 沈培锋, 赵扬, 等. 融合知识库和深度学习的电网监控告警事件智能识别[J]. 电力自动化设备, 2020, 40(4): 40-47.
SUN Guoqiang, SHEN Peifeng, ZHAO Yang, et al. Intelligent recognition of power grid monitoring alarm event combining knowledge base and deep learning[J]. Electric Power Automation Equipment, 2020, 40(4): 40-47.
- [11] 汪崔洋, 江全元, 唐雅洁, 等. 基于告警信号文本挖掘的电力调度故障诊断[J]. 电力自动化设备, 2019, 39(4): 126-132.
WANG Cuiyang, JIANG Quanyuan, TANG Yajie, et al. Fault diagnosis of power dispatching based on alarm signal text mining[J]. Electric Power Automation Equipment, 2019, 39(4): 126-132.
- [12] 周云晓. 多轮对话文本连贯性研究[D]. 上海: 华东师范大学, 2019.
ZHOU Yunxiao. Research of multi-turn dialogic discourse coherence[D]. Shanghai: East China Normal University, 2019.
- [13] 魏慧玲. 文本情感分析在产品评论中的应用研究[D]. 北京: 北京交通大学, 2014.
WEI Huiling. Application research of text sentiment analysis in product reviews [D]. Beijing: Beijing Jiaotong University, 2014.
- [14] 王科, 夏睿. 情感词典自动构建方法综述[J]. 自动化学报, 2016, 42(4): 495-511.
WANG Ke, XIA Rui. A survey on automatic construction methods of sentiment lexicons [J]. Acta Automatica Sinica, 2016, 42(4): 495-511.
- [15] 李勇敢, 周学广, 孙艳, 等. 中文微博情感分析研究与实现[J]. 软件学报, 2017, 28(12): 3183-3205.
LI Yonggan, ZHOU Xueguang, SUN Yan, et al. Research and implementation of Chinese microblog sentiment classification [J]. Journal of Software, 2017, 28(12): 3183-3205.
- [16] 黄发良, 冯时, 王大玲, 等. 基于多特征融合的微博主题情感挖掘[J]. 计算机学报, 2017, 40(4): 872-888.
HUANG Faliang, FENG Shi, WANG Daling, et al. Mining topic sentiment in microblogging based on multi-feature fusion[J]. Chinese Journal of Computers, 2017, 40(4): 872-888.
- [17] 刘玉林, 营利荣. 基于文本情感分析的电商在线评论数据挖掘[J]. 统计与信息论坛, 2018, 33(12): 119-124.
LIU Yulin, JIAN Lirong. Data mining of e-commerce online reviews based on sentiment analysis[J]. Statistics & Information Forum, 2018, 33(12): 119-124.
- [18] 邱剑. 电力中文文本数据挖掘技术及其在可靠性中的应用研究[D]. 杭州: 浙江大学, 2016.
QIU Jian. Research on chinese textual data mining techniques and reliability applications in power systems[D]. Hangzhou: Zhejiang University, 2016.
- [19] DEVLIN J, CHANG M W, LEE K, et al. BERT: pre-training of deep bidirectional transformers for language understanding [C]//North American Chapter of the Association for Computational Linguistics. Minneapolis, Minnesota, USA: ACL, 2019: 4171-4186.
- [20] SUN C, QIU X P, XU Y G, et al. How to fine-tune BERT for text classification? [EB/OL]. (2020-02-05) [2020-12-03]. <https://arxiv.org/abs/1905.05583>.
- [21] JAWAHAR G, SAGOT B, SEDDAH D. What does BERT learn about the structure of language? [C]//Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Florence, Italy: ACL, 2019: 3651-3657.
- [22] GOLDBERG Y. Assessing BERT's syntactic abilities [EB/OL]. (2019-01-16) [2020-12-03]. <https://arxiv.org/abs/1901.05287>.
- [23] 张金鹏. 基于语义的文本相似度算法研究及应用[D]. 重庆: 重庆理工大学, 2014.
ZHANG Jinpeng. Research and application on text similarity algorithm based on semantics [D]. Chongqing: Chongqing University of Technology, 2014.

作者简介:



王海瑶

王海瑶(1996—),女,陕西西安人,硕士研究生,主要研究方向为电力文本挖掘、自然语言处理(**E-mail**: zjuwanghy@zju.edu.cn);

王慧芳(1974—),女,浙江宁波人,副教授,博士,通信作者,主要研究方向为电力多模态数据挖掘、电力设备状态检修、继电保护与控制(**E-mail**: huifangwang@zju.edu.cn);

胡俊华(1981—),男,浙江丽水人,高级工程师,硕士,研究方向为电力智能化(**E-mail**: 70828809@qq.com)。

(编辑 任思思)

Evaluation of power equipment suppliers based on intelligent mining of electric power dialogue text

WANG Haiyao¹, WANG Huifang¹, HU Junhua², XU Jiquan², LI Jianhong³, HE Benteng¹

(1. College of Electrical Engineering, Zhejiang University, Hangzhou 310027, China;

2. State Grid Zhejiang Electric Power Company, Hangzhou 310007, China;

3. Zhejiang Huayun Information Technology Co., Ltd., Hangzhou 310012, China)

Abstract: Currently the satisfaction evaluation of power equipment suppliers mainly relies on manual statistics and index calculation, and its accuracy is greatly affected by evaluators and evaluation contents. The dialogue text of the electric power service platform is used as the research object. Taking the dialogue text of power service platform as the research object, the evaluation model of power equipment supplier based on text mining technology is established on the basis of expanding the entries and attributes of the existing electric power ontology dictionary. Firstly, the next sentence predictive analysis method of single-round dialogue text based on BERT (Bidirectional Encoder Representations from Transformers) NSP (Next Sentence Prediction) and cosine similarity weighted is proposed. The dialogue interruption cross-handling process and supplier identification rules are established to realize theme induction of electric power dialogue text. Then, considering the complexity of the semantic sentiment of the dialogue text, the dialogue sentiment analysis rules are proposed for establishing the supplier evaluation model. Finally, the accuracy of the proposed method is verified by an example. Results indicate that the evaluation of power equipment suppliers based on the intelligent mining of dialogue text is feasible and effective, and can be used as a useful supplement to current evaluation methods.

Key words: electric power equipment; intelligent text mining; dialogue text; next sentence prediction; sentiment analysis; supplier evaluation

(上接第 209 页 continued from page 209)

Online detection method of transformer winding deformation based on combined analysis of short circuit impedance and $\Delta U-I_1$ locus characteristics

LI Zhenhua^{1,2}, JIANG Weihui¹, YU Caiyun¹, CHEN Xingxin¹, LI Zhenxing¹, XU Yanchun¹

(1. College of Electrical Engineering & New Energy, China Three Gorges University, Yichang 443002, China;

2. Hubei Provincial Key Laboratory for Operation and Control of Cascaded Hydropower Station,

China Three Gorges University, Yichang 443002, China)

Abstract: In order to solve the problem that the short circuit impedance method cannot identify fault types and that both the short circuit impedance method and the $\Delta U-I_1$ locus method are susceptible to the interference of equipment measurement errors, an online winding deformation detection method based on the combined analysis of short circuit impedance and $\Delta U-I_1$ locus characteristics is proposed. The principle of online short circuit impedance method is introduced, and a calculation method of short circuit impedance based on the short time invariance of measurement error is put forward to reduce the measurement error. The principle of $\Delta U-I_1$ locus method is introduced, then the online detection steps and criteria of transformer winding deformation based on combined analysis of short circuit impedance and $\Delta U-I_1$ locus characteristics are given. The effectiveness of the proposed method and its accuracy considering measuring errors are verified by establishing a transformer simulation model. The results show that the proposed method can accurately identify transformer winding deformation faults when considering measurement error, and has the advantages of live detection and fault type identification, which improves the identification accuracy of winding deformation fault.

Key words: power transformers; winding deformation; short circuit impedance; $\Delta U-I_1$ locus; online detection

附录

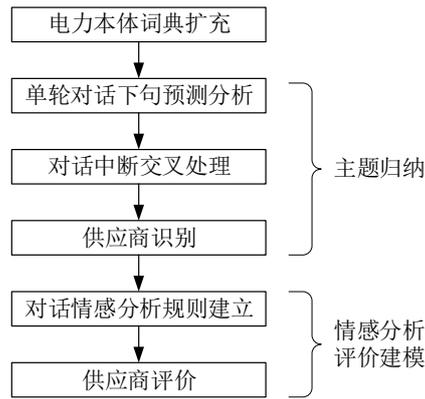


图 A1 基于电力对话文本的供应商评价过程
Fig.A1 Supplier evaluation process based on power dialogue text

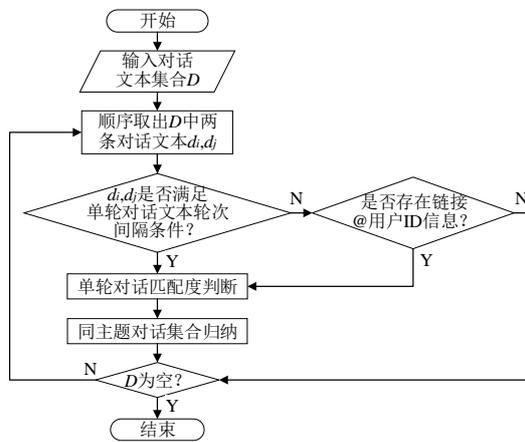


图 A2 对话中断交叉处理流程
Fig.A2 Flowchart of dialogue interrupt cross processing

表 A1 电力采集运维领域本体词典示例
TableA1 Example of ontology dictionary in power domain

本体词	类别	实例 1	实例 2
南京南瑞集团公司	供应商名称	南京南瑞	南瑞
时钟紊乱	事件关键词汇	时钟异常	紊乱
...

表 A2 电力对话文本示例
TableA2 Examples of power dialogue text

编号	用户 ID	对话文本内容	主题划分
1	68505019	轮换表计之后,新的表计出来时钟异常 是一串英文数字怎么办? ??	1
2	68505019	有人知道怎么解决吗	1
3	39021210	你什么单位啊,表多的还能轮换	1
4	68701239	这种陶瓷杯性价比最高	2
5	68505019	不是重点,重点是这表计换了之后时钟 不正常了	1
6	68505019	再换就是频繁换表了	1
7	32815174	陶瓷杯链接发我一个?	2
8	68505019	表计是 A 厂家的,原来还挺好的	1
9	68505019	换了之后表计上来的时钟一串代码,对 时都对不了	1
10	39021210	这真的没办法了,联系他们厂家吧	1
11	39021210	我们 B 厂家的相对多些,小毛病有的, 但没有大毛病	1
12	39021210	A 厂家的也有,下回招标不知道会用哪 一家	1
13	47855210	@68701239 陶瓷杯链接也发我一个	2