

Design of Question Answering System in Substation Fault Field Based on Knowledge Graph

He Jun, Wu Sheng Ke, Rao Fang Xi, Zhong Ke Jia, Liu Peng Zheng, Sun Jian Ming

College of Information and Engineering, Nanchang University, Nanchang, China

Email address:

577803848@qq.com (Wu Sheng Ke)

To cite this article:

He Jun, Wu Sheng Ke, Rao Fang Xi, Zhong Ke Jia, Liu Peng Zheng, Sun Jian Ming. Design of Question Answering System in Substation Fault Field Based on Knowledge Graph. *Science Discovery*. Vol. 10, No. 5, 2022, pp. 366-373. doi: 10.11648/j.sd.20221005.23

Received: September 26, 2022; Accepted: October 27, 2022; Published: October 29, 2022

Abstract: The operation and maintenance of substation equipment is of great significance to ensure the safety and stability of the power grid. In order to summarize the historical operation and maintenance experience of substation equipment and facilitate the management personnel to uniformly manage and accurately query the historical data of substation faults, this paper constructs an intelligent question and answer system for substation faults based on the knowledge map. At present, the research on knowledge atlas in power field mainly focuses on the construction and visualization of knowledge atlas. This paper combines natural language processing, deep learning, knowledge atlas, graph database and other knowledge with power. Firstly, the substation fault data is cleaned and extracted, and the substation fault knowledge map is constructed; Secondly, based on Bert+TextCNN algorithm, the intention recognition function of questions is realized; Thirdly, based on BILSTM-CRF algorithm, the named entity recognition function of questions is realized. Finally, combined with Flask deployment model, the API interface is released, and the substation fault question answering system is built through the establishment of WeChat automatic question answering robot. The experimental results show that the accuracy of the system reaches 96%. Finally, a complete question answering system for substation fault is formed, and accurate, fast and intelligent knowledge question answering service is realized.

Keywords: Knowledge Graph, Substation Faults, Named Entity Recognition, Question Answering System

基于知识图谱的变电故障领域问答系统设计

何俊, 吴慎珂, 饶方喜, 钟可佳, 刘鹏政, 孙建民

南昌大学信息工程学院, 南昌, 中国

邮箱

577803848@qq.com (吴慎珂)

摘要: 变电设备的运行维护工作, 对保障电网的安全稳定具有十分重要的意义。为总结变电设备的历史运维经验, 方便管理人员对变电故障历史数据统一管理 and 准确查询, 本文基于知识图谱构建了一个变电故障智能问答系统。目前, 电力领域知识图谱研究主要集中在知识图谱的构建及可视化。本文将自然语言处理、深度学习、知识图谱、图数据库等知识与电力进行结合: 首先, 对变电故障数据进行清洗、提取, 构建变电故障知识图谱; 其次, 基于Bert+TextCNN算法实现问句的意图识别功能; 再次, 基于BILSTM-CRF算法实现问句的命名实体识别功能; 最后, 结合Flask部署模型发布API接口, 通过建立微信自动问答机器人搭建变电故障问答系统。实验结果表明该系统准确度达到96%, 最终形成完整的变电故障问答系统, 实现准确、快速、智能的知识问答服务。

关键词：知识图谱，变电故障，命名实体识别，问答系统

1. 引言

随着电网智能化与信息化的建设，电网企业积累了庞大的电力大数据，对这些数据的存储、查询以及挖掘一直是研究的热点。其中包括的输变电故障以及故障解决记录对输变电运维和管理技术人员应对突发事件具有重要借鉴意义，如何对变电故障领域的数据信息统一管理维护并能快速准确查询相关知识是目前需要解决的问题。知识图谱的出现，为数据存储提供了一种新的方式。知识图谱是一种用图模型的形式表示事物实体、属性及不同事物之间关系的大型语义网络。随着知识图谱技术的不断发展和完善，已经在许多领域有了广泛的应用。比如刘欢等[1]从数字人文视角下构建了古汉语典籍领域的知识图谱，以《左传》为研究对象构建古汉语领域知识图谱，并结合深度学习算法设计智能问答系统，实现古文领域知识的智能检索。郑泳智等[2]以病虫害数据为基础构建农业领域荔枝和龙眼病虫害知识图谱，利用实体识别和意图识别模型识别用户问句，实现农业领域病虫害智能问答系统。陈一龙等[3]构建罕见病及专家融合型医疗领域知识图谱，并利用智能问答系统、结合大数据和人工智能方法，设计出了一款针对患者的罕见病就医决策引擎。鉴于知识图谱对领域知识管理的良好表现，目前电力领域知识图谱也取得一些成果。杨迪等[4]构建了一种基于知识图谱的电力设备缺陷检索方案，通过对现有设备缺陷记录语料的充分运用实现相应缺陷知识图谱的构建，在此基础上通过使用图搜索方法实现对设备缺陷的检索过程。蒲天骄等[5]针对我国电力物联网的数据特征与发展趋势提出了一种基于NoDKG (not only domain-specific knowledge graph)思想的电力知识图谱应用框架设计。李刚等[6]构建了电力设备健康管理知识

图谱，针对电力设备的特点，描述了电力设备健康管理知识图谱的构建方法及构建过程。

目前来看，电力领域知识图谱研究主要集中在知识图谱的构建及可视化方面。本文将自然语言处理、深度学习、知识图谱、图数据库等知识与电力相结合，构建基于知识图谱的变电故障智能问答系统，总结了变电设备的历史运维经验，方便管理人员对变电故障历史数据统一管理和准确查询，进一步提高工作人员分析、处理设备故障及缺陷的能力。通过对变电故障数据进行分析，经过概念模型设计、知识抽取和知识存储等方式构建变电故障知识图谱，将变电故障领域内结构化、半结构化和非结构化的数据信息整合，将散乱冗余的信息处理后形成数据完整、结构严谨的知识图谱。构建基于知识图谱的变电故障智能问答系统，用户可以查询变电故障相关知识，为用户提供了准确、快速、智能的知识问答服务，从而实现对变电故障知识的管理、维护和查询。

2. 变电故障知识图谱设计

2.1. 数据收集

目前在电力领域尚无公开的结构化变电故障数据，但是电网企业积累了大量的变电故障以及故障解决记录，这些数据大都存在与零散的故障检修分析报告中。因此本次任务中的变电故障数据是结合江西历年变电故障检修报告收集、整理的，以后可以根据实际需求利用知识抽取技术再次扩充知识图谱的内容，进一步丰富和完善电力故障知识图谱。经过对数据的获取、清洗、提取后数据内容主要以变电设备为核心，包括设备名称、变电站、运行维护单位、实物ID、设备类型、设备电压等级、设备厂家、设备型号、设备属性、出厂日期、故障部位、故障日期、事件原因、事件概况、整改措施等内容。部分原始数据如表1所示。

表1 部分变电故障数据示意图。

序号	运行维护单位	变电站	设备名称	设备类型	电压等级	故障日期	设备型号	故障部位
1	赣州供电公司	梅江220kV变电站	220kV莲梅线2162隔离开关	GIS	220kV	2020.12.12	IFT-252	绝缘拉杆
2	九江供电公司	庐峰220kV变电站	#2主变有载分接开关	主变	220kV	2019.12.26	SSZ-150000/220	有载开关
3	吉安供电公司	众村220kV变电站	220kV樟众线217开关电流互感器	电流互感器	220kV	2020.3.9	LB10-220W2	瓷套
4	上饶供电公司	罗桥110kV变电站	110kV上罗线114开关线路避雷器	避雷器	110kV	2018.6.6	YH10W1-108/281	阀片
5	宜春供电公司	盐化220kV变电站	110kV盐蓝线1122刀闸	隔离开关	110kV	2018.2.13	GW22B-126	传动机构

2.2. 知识建模

知识建模是基于行业的应用属性、知识特点、实际需求，依据知识图谱的模式进行业务抽象和业务建模。将知识进行形式化和结构化的抽象。

本次任务采用自顶向下的方法构建垂直领域电力知识图谱，即首先为知识定义数据模型，主要是实体定义、关系定义、属性定义。根据上述已收集的变电故障数据以及知识图谱的用途，共定义了五类实体、两类关系、余类

属性。其中将设备名称定义为实体，将设备电压等级、实物ID、设备类型、设备型号、出厂日期等定义为设备名称实体的属性；将故障事件定义为实体，将故障时间、故障部位、故障原因、故障简介、整改措施等定义为故障事件实体的属性；将设备生产厂家、变电站、运行维护单位定义为实体，将电压等级定义为变电站实体的属性；同时将故障事件与设备、设备与变电站、变电站与运行维护单位的关系定义为属于；将设备与生产厂家的关系定义为出场于。经过上述知识建模就可以将原始变电故障数据集的内

容构造为（实体，关系，实体）、（实体，属性，属性值）这两种模式的三元组进行抽取存储，同时也将原始非结构化的数据处理成结构化的数据。处理后的部分结构化的数据如表2所示。

表2 部分结构化数据。

	类别	实例
设备	实体	220kV莲梅线2162隔离开关
变电站	实体	梅江220kV变电站
运行维护单位	实体	赣州供电公司
故障日期	属性	2020.12.12
设备型号	属性	IFT-252
设备与变电站	关系	（220kV莲梅线2162隔离开关）属于（梅江220kV变电站）

2.3. 知识存储

存储知识的方法没有统一标准，主流的知识存储方式主要有以下三种：第一种是三元组形式的RDF存储；第二

种是传统关系型数据库存储；第三种就是图数据库存储[7]。而目前最受人们欢迎的就是图数据存储或者关系型数据库+图数据库存储的方式。图数据库的基本概念就是以“图”这种数据结构存储和查询数据。

本文为了方便后续设计与实现智能问答系统，选取Neo4j图数据库储存变电故障知识。Neo4j以图论为基础，通过节点和关系构建属性图为真实世界直观建模，可以快速管理、存储和遍历节点和关系，将数据结构的节点和关系形象化展示，拥有独特的声明式查询语言，具有界面友好、可靠性高、容量庞大等特点，由此可以提高知识管理维护效率。

利用pycharm编译器建立变电故障实体、属性、关系模型并通过Neo4j数据库的python驱动py2neo连接图数据库，实现变电故障数据三元组的快速导入，将变电故障知识进行存储，从而成功构建变电故障知识图谱。当完成任务后，其可视化界面如图1所示。

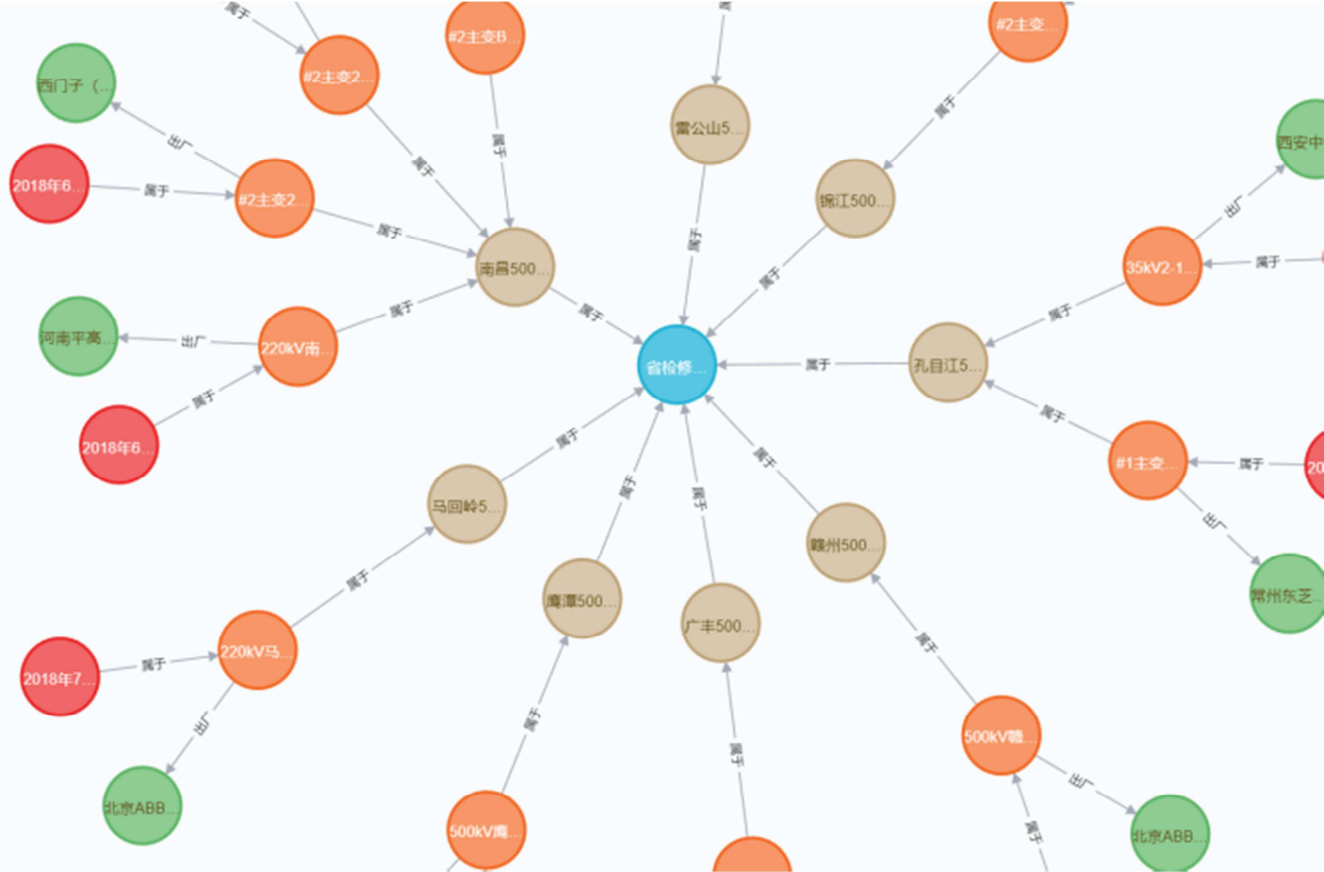


图1 变电故障知识图谱示意图。

3. 变电故障问答系统设计

3.1. 系统框架

结合实际需求，基于知识图谱的变电故障智能问答系统实现过程可分为知识图谱构建模块、自然语言处理模块、

对话管理模块以及交互模块这四部分。系统框架流程示意图如图2所示。

当系统启动后，首先进入交互界面，本文问答系统的交互界面是借助微信实现的，以微信问答机器人的形式展现。当用户在微信窗口提出问题后，系统后端自然语言处理模块运行，首先经过分词、向量化处理，再进入意图识别模型进行意图识别，识别具体意图后进入命名实体识别（NER）模型进行实体识别与提取。得到用户问题的实体

和意图后进入对话管理模块，通过遍历槽位列表和实体识别的结果填充故障意图下的槽位，然后根据意图置信度确定回复策略。如果置信度较高，则生成数据库 Cypher 查询语句，在图数据库（Neo4j）中查询检索答案，最后输

出返回给交互界面微信机器人。如果置信度没有达到要求则返回相应提示。系统实现的最大的难点在于如何快速、准确的理解用户问句的含义，本文主要采用意图识别和命名实体识别方法。

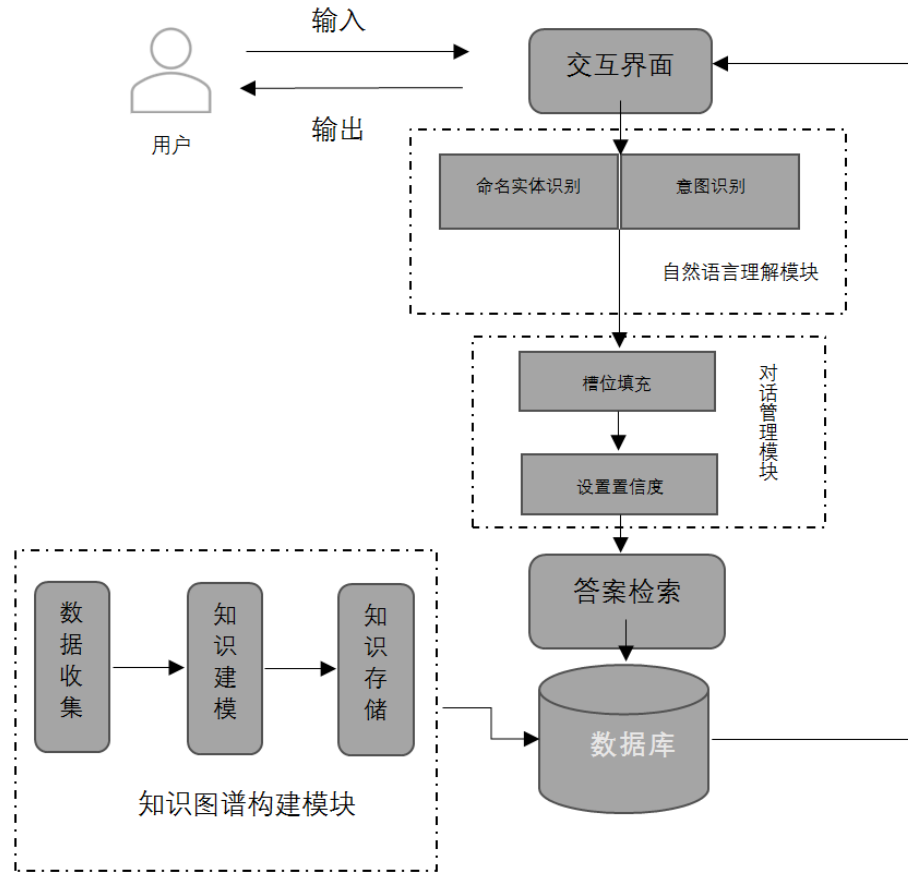


图2 变电故障问答系统框架流程图。

3.2. 意图识别

识别问句的意图本质上是对用户意图进行归类，可以将其看作一个文本分类问题。常用的意图识别方式有基于词典和模版规则的方法、基于分类模型进行意图识别的方法。这两种分类方式是当前比较主流的方法。完成意图识别任务的难点主要有两点，一是数据的匮乏，电力领域内缺乏变电故障意图分类的数据集，需要耗费大量时间进行数据标注。二是由于变电故障意图种类较多，且要分类明确。本文所研究的变电故障问答系统采用基于分类模型的方法进行意图识别，使用深度学习的方法训练意图分类模型。本文提出使用Bert+TextCNN模型进行变电故障多意图分类，针对变电故障数据和变电故障知识图谱的特点定义不同的变电意图类别。

BERT[8]是谷歌2018年提出的自然语言处理模型，采用Transformer编码器的双向编码结构作为特征抽取器，可以看作多个Transformer的encoder的堆叠。该模型基于Masked Language Model (MLM)和Next Sentence Prediction (NSP)，能够精确而全面将变电故障问句向量化表示。

TextCNN[9]是应用卷积神经网络的文本分类模型，具有网络结构简单、训练速度快、效果优良以及能够捕捉局部特征等特点。虽然BERT模型也可以用于文本分类，但是该模型缺乏提取故障问句局部特征信息的能力。双向编码结构具有的强大特征提取，会导致过拟合问题。而TextCNN稀疏连接和权值共享可以很好地清除故障问句文本噪音、提高局部特征提取能力，与BERT模型形成良好的互补，提高模型的泛化表现。所以本文提出基于Bert-TextCNN模型对故障问句进行分类训练。BERT-TextCNN模型[10]由词嵌入层、卷积层、池化层和全连接层构成，该模型如图3所示。输入层中选取经过数据清洗和预处理后的故障问句，通过BERT模型进行训练得到向量化表示，同时提取Bert输出的cls特征向量。将预训练的词向量传入TextCNN，通过设计不同大小的卷积核、经多重卷积池化操作后获得关于故障问句中的不同层次的特征信息的特征向量。然后将上述两种特征向量进行拼接后传入全连接层进行降维，经过降维处理的特征向量维度大小就是变电意图分类的标签数量。最终输入到Softmax中得到预测的意图类别标签。

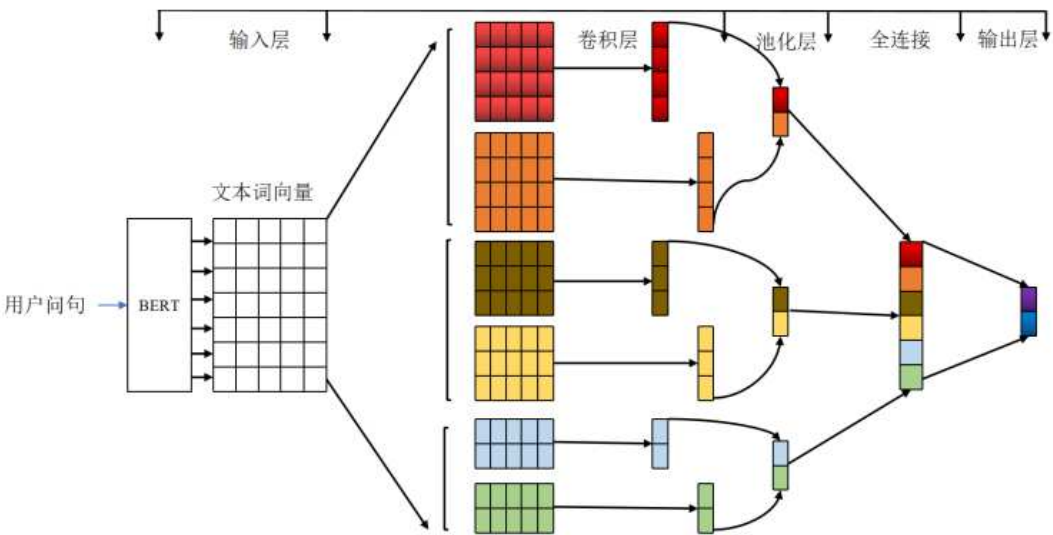


图3 基于BERT-TextCNN模型结构图。

3.3. 命名实体识别

描述实体的词汇被称为命名实体，命名实体识别就是识别句子中所存在的命名实体及其类别，本质上可以看作序列标注任务。目前于此任务表现较好的是基于深度学习的或者统计学习的方法，其难度同样在于缺少大量标注的数据作为支撑。在基于深度学习的命名实体识别模型中表现较好且最常用模型为BiLSTM-CRF[11]算法。为了识别效果更好，本研究采用的是 BiLSTM-CRF算法和AC（Aho-Corasick）自动机算法相结合的方法进行变电故障问句中实体的识别。在获取问句中的实体时，为保证用户高效问答体验，用AC算法匹配变电故障设备实体，用于快速匹配实体词。但是这种方法难以处理复杂的问答，当

AC自动机匹配失败后，此时BiLSTM-CRF算法试图在问句中提取实体词。

LSTM是RNN模型的一种优化，是具有长短期记忆功能的神经网络。LSTM在RNN结构的基础之上引入了门限机制和记忆单元，门限机制的引入不仅可以对已有信息进行遗忘还可以对进入记忆单元中的记忆信息有效的筛选和学习，记忆单元的存在使得LSTM保存信息时长大于RNN中的短时记忆，同时还解决了梯度消失和梯度爆炸等问题。BiLSTM采用双向LSTM机制，前向LSTM层可以提取输入序列中t时刻以及之前时刻的信息，而后向LSTM层可以提取输入序列中t时刻以及之后时刻的信息。两个LSTM层输出的特征信息可以使用相加、平均值或连接等方式进行融合。BiLSTM具备LSTM优点的同时又解决了其无法兼顾上下文的问题，可以有效捕捉到序列前后的特征信息。Bi-LSTM模型结构如下图4所示。

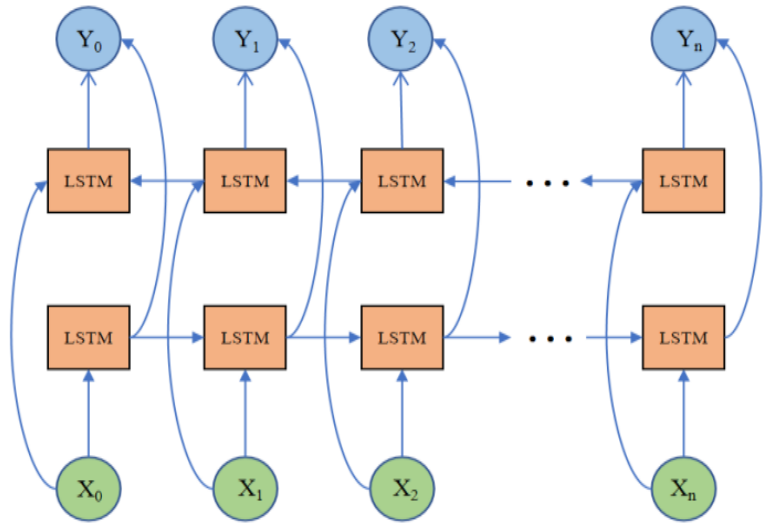


图4 Bi-LSTM模型结构图。

条件随机场（CRF）是基于概率图的无向图模型，给定输入随机变量X的值，求解概率 $P(Y|X)$ 的概率。CRF在解码时可以考虑当前字符属于的上下文关系，依据概率

转移矩阵判断当前字符所属的标签，避免了不符合序列排序要求的实体片段生成，在序列标注和实体识别等方面应用广泛。

BiLSTM-CRF模型首先通过embedding将序列文本转成词向量，然后通过BiLSTM提取每个字在其上下文的语义表达向量，输出每个字所对应的标签的概率，将得到的

概率作为CRF模型的输入，通过学习标签之间的依赖关系及考虑标签之间的限制关系，获得全局最优的标签序列。Bi-LSTM-CRF模型结构如下图5所示。

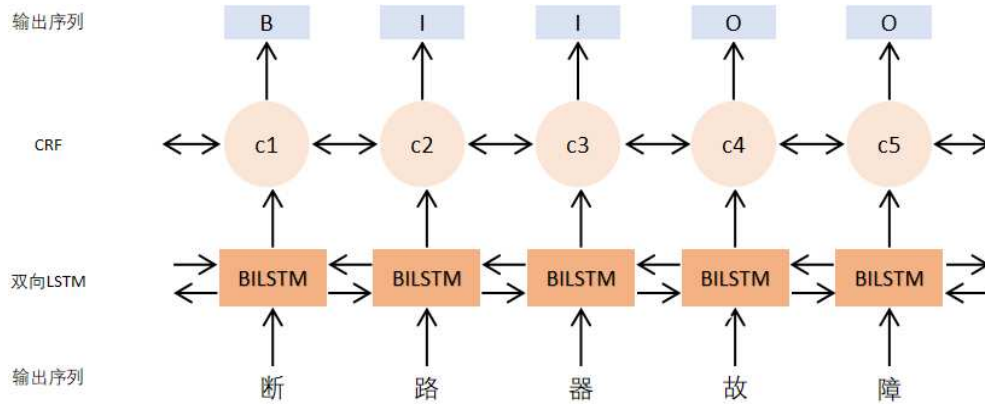


图5 Bi-LSTM-CRF模型结构图。

3.4. 对话管理模块实现

对话管理（DM）模块主要由两部分内容，一个是对话状态追踪（DST），另一个是设置置信度。对话状态追踪（DST）就是根据所有对话历史信息推断当前对话状态和用户目标，基于对话历史，为槽位列表中的每一个槽位找到一个槽位值。概率机器人一个主要的概念是置信度（belief），置信度反映了机器人有关环境状态的内部信息。本文中问答机器人的置信度是通过意图识别模型的准确度判定的，然后根据意图识别置信度确定回复策略，这里分了三种情况：

- （1）置信度 ≥ 0.8 ，构造Neo4j查询语句，并从数据库中查询答案
- （2）置信度 $0.4 \sim 0.8$ ，向用户进行问题澄清
- （3）置信度 < 0.4 ，识别失败，回复兜底语句

4. 实验

4.1. 环境配置

表3 实验环境配置。

名称	配置
硬盘大小	512 GB
内存	16 GB
CPU	AMD Ryzen 7 5800H with Radeon Graphics
GPU	RTX3060
数据库	Neo4j
开发环境	Python
编译器	PyCharm
前端	微信机器人

变电故障问答系统主要模块功能是通过Python语言编写实现的，前端交互界面是以微信机器人的方式和用户进行交互。具体开发环境见表3。

运行设备为本地主机，整个系统所使用的编译器为PyCharm。其中问答系统所需环境Python版本是3.6.2；Java开发工具包是JDK1.8.0；图数据库Neo4j版本是3.5.7；深度学习框架选用的是tensorflow1.14.0和keras2.3.1；交互功能用的是Python中的itchat模块，版本是1.3.10。

4.2. 实验结果

4.2.1. 意图识别任务

在意图识别模型[12]训练前需要设计问句种类，并且获取训练语料作为数据集。首先设计变电故障意图类型，本次任务共设计了13类问题意图，包括：电压等级、设备类型、设备型号、出厂日期、设备厂家、故障时间、故障部位、故障原因、故障简介、整改措施、故障地点、其他。本次任务的变电故障语料是人工收集、编写、标注的，共有23130条语料，其中包括训练集19412条和测试集3718条。样本示例如表4所示。

本文BERT模型使用的是谷歌官方发布的中文预训练模型，采用Bert+TextCNN模型对上述语料进行训练后，通过召回率（recall）、精确率（precision）、F1值对模型进行评测，每个变电故障意图的评测实验结果如表4所示。

实验结果由下表5所示，使用Bert+TextCNN模型训练变电故障语料数据集，13种变电故障意图的精确度、召回率、F1值均为0.9以上，其中最高可达0.99。同时实验结果表明，变电故障意图识别准确率为0.96，精确度为0.96，召回率为0.95，F1值为0.96。Bert+TextCNN模型性能良好，可以满足变电故障问答系统意图识别任务的实现。

表4 变电故障意图语料。

序号	标签	短文本
1	电压等级	能够告诉我110kV II 段母线的电压多少伏吗？
2	设备类型	500kV罗抚 II 回线50412刀闸设备的所属类型？
3	设备型号	请问#2主变220kV侧避雷器设备型号是什么？
4	出厂日期	#2主变202断路器是几月几号生产的呢？

序号	标签	短文本
5	设备厂家	220kV雷燕线222断路器生产于什么厂家？
6	故障时间	#1主变35KV侧301开关柜故障发生在哪一天？
7	故障部位	220kV石蓝Ⅱ线2142刀闸的设备故障位置在什么地方？
8	故障原因	10kV分段931开关为什么会有故障？
9	故障简介	你知道220kV新白线215开关故障吗，发生了什么？
10	整改措施	#2主变10kV侧管型母线故障后实施了哪些补救方案？
11	故障地点	罗珠Ⅱ线2222刀闸设备是在什么地方发生的故障？
12	运维单位	帮我查询珠珊220kV变电站的运行维护单位？
13	其他	请问购买5013断路器设备花费了多少钱呢？

表5 标签评测结果。

标签	Precision	Recall	F1-score
电压等级	0.96	0.99	0.97
设备类型	0.93	0.94	0.94
设备型号	0.99	0.98	0.99
出厂日期	0.91	0.92	0.91
设备厂家	0.97	0.98	0.98
故障时间	0.91	0.94	0.93
故障部位	0.99	0.99	0.99
故障原因	0.96	0.94	0.95
故障简介	0.98	0.95	0.97
整改措施	0.97	0.96	0.96
故障地点	0.99	0.99	0.99
运维单位	0.98	0.93	0.95
其他	0.90	0.91	0.90

4.2.2. 命名实体识别任务

在本系统中构建命名实体模型的目的是从变电故障问句中识别出故障设备名称（Equipments）、变电站（Transformer）、运行维护公司（Company）三大类实体，并将识别的实体提取并填充到意图下的槽位中，用于构建知识图谱查询语句。数据集序列标注任务采取的是BIO的方式进行标注的，再由人工检查核对后完成，共标注文本十万字左右。取其中80%为训练集、10%为验证集、10%为测试集。

表6 三类实体识别实验结果。

实体类别	Precision	Recall	F1-score
设备名称	0.9686	0.9686	0.9686
变电站	0.9868	0.9740	0.9804
运行维护公司	0.9615	0.9694	0.9711

同样采用召回率（recall）、精确率（precision）、F1作为模型评价指标。本次实验采用Python第三方库Kerasi作为深度学习框架搭建模型进行训练，BiLSTM隐藏层设置为128层，模型训练80个epoch。BiLSTM-CRF模型训练结果测试集Precision为0.97，Recall为0.96，F1值为0.97。其中每类实体识别实验结果如上表6所示。

4.2.3. 界面展示

为方便用户使用，本文基于python的第三方库itchat搭建交互界面，itchat是一个开源的微信号个人接口，使用python调用微信，通过二维码扫描登录，可以获取用户列表，用户详情信息，可以接收回复消息，构建微信自动聊天机器人，实现智能问答。扫描登录后，用户即可通过微信界面提出问题，文本传入后端通过本文设计的问答系统即可查询答案并返回用户。用户交互界面如图6所示。



图6 微信界面示意图。

5. 结语

本文收集整理变电故障数据,构建了变电故障知识图谱,使得变电故障领域的信息方便统一管理维护。并在此基础上通过命名实体识别、意图识别等自然语言处理知识构建了基于知识图谱的变电故障问答系统,探索了人工智能在电力领域的应用,为用户提供了准确、快速、智能的知识问答服务。同时标注了大量的变电故障领域的文本数据,构建了变电故障意图数据集,为深度学习在电力领域的应用提高了方便。本文目前所实现的基于知识图谱的变电故障问答系统仍存在一些不足,比如构建电力知识图谱所使用的故障数据是由人工收集、清洗得到的,存在花费时间长,专业要求高等问题。本文的研究对象仅仅是面向于变电故障领域,存在获取知识少,数据量有限等问题。针对上述问题,后期将考虑扩充领域范围,同时使用实体、关系抽取模型获取数据,提高构建知识图谱效率。

致谢

本文为国家自然科学基金资助项目《基于深度学习与多模态融合方法的情感识别技术研究》(62066025)的阶段性成果之一。

参考文献

- [1] 刘欢, 刘浏, 王东波. 数字人文视角下的领域知识图谱自动问答研究 [J]. 科技情报研究, 2022, 4 (01): 46-59.
- [2] 郑泳智, 吴惠彝, 朱定局, 陈炳旭, 李文景. 基于荔枝和龙眼病虫害知识图谱的问答系统 [J]. 计算机与数字工程, 2021, 49 (12): 2618-2622.
- [3] 陈一龙, 卜嘉彬, 李景宇, 李东泽, 龚力, 万智. 基于知识图谱的罕见病就医决策引擎设计研究 [J]. 华西医学, 2021, 36 (12): 1730-1733.
- [4] 杨迪, 梁懿, 王秋琳, 陈新梅, 陈恩光, 王燕蓉. 一种知识图谱的电力设备缺陷检索方案设计与实现 [J]. 微型电脑应用, 2022, 38 (01): 191-194.
- [5] 蒲天骄, 谈元鹏, 彭国政, 徐会芳, 张中浩. 电力领域知识图谱的构建与应用 [J]. 电网技术, 2021, 45 (06).
- [6] 李刚, 李银强, 王洪涛, 谢庆, 黄文琦, 侯佳萱. 电力设备健康管理知识图谱: 基本概念、关键技术及研究进展 [J]. 电力系统自动化, 2022, 46 (03): 1-13.
- [7] 郑泳智, 吴惠彝, 朱定局, 等. 基于荔枝和龙眼病虫害知识图谱的问答系统 [J]. 计算机与数字工程, 2021 (049-012).
- [8] 陈旻杰, 魏轩, 郑莹, 周世元. 基于BERT的智能问答系统 [J]. 数字技术与应用, 2022, 40 (01): 161-163.
- [9] 邵辉. 基于BERT-TextCNN的中文短文本情感分析 [J]. 信息与电脑 (理论版), 2022, 34 (01): 77-80.
- [10] 杨飞洪, 王序文, 李姣. 基于BERT-TextCNN模型的临床试验筛选短文本分类方法 [J]. 中华医学图书情报杂志, 2021, 30 (01): 54-59.
- [11] 曾勇. 基于BiLSTM-CRF模型的中文命名实体识别研究与实现 [D]. 江西财经大学, 2020.
- [12] 张耀允. 基于用户意图分析的交互式问答技术研究 [D]. 哈尔滨工业大学, 2012.