

# 创邻科技 NLP (知识图谱)

## CCF 竞赛方案

# 1. 赛题介绍

## 1.1. 赛题概述

本活动对参赛选手基于自然语言理解的对话能力进行评估，重点评估范围是知识抽取、自动构建知识图谱以及基于知识图谱的问答能力。

本规范规定了各选手需要进行的评估内容及其方法。

## 1.2. 赛题内容

本次赛题内容包括两部分。

### 1.2.1. 电力规章制度知识图谱构建能力

电力规章制度知识图谱的构建能力主要包括电力规章制度知识抽取、自动构建知识图谱。

### 1.2.2. 电力规章制度智能问答能力

电力规则制度智能问答能力包括标准问答、多轮对话及根据不同的问题进行中控路由（即：精确识别问题类型是标准问答，还是多轮对话；并且依据识别结果进行不同方法的回答）等功能。

# 2. 数据集介绍

## 2.1. 数据集说明

本次比赛使用创邻科技开放的“电力行业规章制度知识非结构化数据集”，来自我国某市供电局的真实规章制度知识数据。

数据集主要包括国家电网有限责任公司电力事故事件调查规程（2020 版）、某市电网电力调度管理规程（2020 版）。

参赛选手需要构建模型，构建电力规章制度知识图谱，并且依据知识图谱完成电力规则制度的智能问答。

## 2.2. 数据集获取方式

数据集下载链接：

[https://galaxybase.com/public/activity/task1\\_data.zip](https://galaxybase.com/public/activity/task1_data.zip)

## 2.3. 数据集文件介绍

文件一：国家电网有限责任公司电力事故事件调查规程（2020 版）.doc

该调查规程适用于国家电网有限责任公司系统电力生产安全事故事件调查管理和信息报送，为了达到规范电力生产安全事故事件调查和管理的目的。

文件二：某市电网电力调度管理规程（2020 版）.doc

该规程适用于某市中调直调系统调度管理工作，是某市电网电力调度运行工作的有关单位及人员均严格遵守的基本规程。

# 3. 答题流程

## 3.1. 平台搭建

选手需要自行搭建 Galaxybase 图平台，具体平台搭建流程可参见《附件 3-GalaxyBase 安装部署配置指南》。

具体平台搭建流程详情也可参见《Galaxybase 开发者版本文档》中“安装部署配置”全部章节。

相关内容文档链接：

<https://www.galaxybase.com/document?file=dev&docid=27>

## 3.2. 知识建模

选手需要依据抽取后的数据结果进行知识建模，具体建模流程可参见《附件

4-GalaxyBase 图构建数据导入文档》。

具体建模流程详情也可参见《Galaxybase 开发者版本文档》中“图构建数据导入文档”章节的“`schema.json` 样例与讲解”内容。

相关内容文档链接：

<https://www.galaxybase.com/document?file=dev&docid=24>

### 3.3. 知识建图

选手需要将抽取后的数据结果按本文 3.3 节“知识建模”所得的模型结果进行数据映射，从而完成数据导入流程，具体数据建图流程可参见《附件 4-GalaxyBase 图构建数据导入文档》

具体数据建图流程详情也可参见《Galaxybase 开发者版本文档》中“图构建数据导入文档”章节的“`mapping.json` 样例与讲解”内容与“Galaxybase-console-buildgraph.jar 图构建样例与讲解”内容。

相关内容文档链接：

<https://www.galaxybase.com/document?file=dev&docid=24>

### 3.4. 知识抽取

选手需要对《某市电网电力调度管理规程（2020 版）》、《国家电网有限责任公司电力事故事件调查规程（2020 版）》两篇文档进行相关知识抽取，抽取的结果中最好将实体与关系分开，从而给自动化建图提供方便。

### 3.5. 业务验证

依据所给数据，整理若干条标准问（标准问，即知识问答的问题。如：电力调度的定义是什么）及标准答（标准答，即知识问答中问题的答案。如：电力调度机构依法行使调度权，对所辖电网运行进行的组织、指挥、指导和协调）。用来验证提交的知识图谱与业务逻辑功能。

## 4. 提交项说明

选手提供的提交物的文档结构示意图如下：

example_submit > XXX_任务1基于自然语言理解的知识图谱构建及智能问答				
名称	修改日期	类型	大小	
CodeData	2021/6/15 22:23	文件夹		
FileResource	2021/6/15 22:22	文件夹		
Input	2021/6/15 22:21	文件夹		
Output	2021/6/15 22:21	文件夹		
TestData	2021/6/15 22:22	文件夹		
buildGraph.sh	2021/6/16 11:51	SH 文件	1 KB	
exeTopic1.sh	2021/6/16 11:28	SH 文件	1 KB	
kgExtraction.sh	2021/6/16 11:51	SH 文件	1 KB	
testKGAnswer.sh	2021/6/16 11:52	SH 文件	1 KB	

文件名命名规则：【选手姓名】+ “\_” + 【任务全名】

具体文件介绍：

Input：数据集存放位置

Output：执行结果（包含：知识抽取结果、关系抽取结果、schema 结果、mapping 结果、电力规章制度智能问答模型自测结果）存放位置

TestData：选手自拟测试集（电力规章制度智能问答的标准问与标准答）存放位置

FileResource：相关要求输出文档存放位置

CodeData：相关要求输出程序源代码存放位置

exeTopic1.sh：整体测试执行脚本，选手无需更改。操作实例如下：

```
bash exeTopic1.sh testQuestions.txt > 选手报名手机号.log
```

kgExtraction.sh：知识抽取模块执行脚本，选手需要替换“echo ---知识抽取源代码执行---”处。

buildGraph.sh：图构建模块执行脚本，选手需要替换“echo ---schema.json 生成源代码执行---”、“echo ---mapping.json 生成源代码执行---”、“echo ---图构建源代码执行---”三处。

testKGAnswer.sh：智能问答验证模块执行脚本，选手需要替换“echo ---智能问答模型源代码执行---”、“echo 开始验证自己的测试问题集 \$1”、“echo 开

始验证主办方的打分问题 \$2” 三处。

#### 4.1. 电力规章制度知识图谱构建能力

模块一	提交项	提交项要求
电力规章制度知识图谱	电力规章制度知识抽取说明文档	应具有的内容： 1、抽取思路及实现方法 2、抽取的合理性分析 3、抽取的测试报告 4、抽取方法的不足及改进策略
	知识抽取源代码	1、编码语言为 Java/Python 2、代码应遵循低内聚、高耦合 3、代码应具备优良的注释习惯
	知识抽取结果	结果为具有表头的 csv 文件，编码格式为 UTF-8
	关系抽取源代码	1、编码语言为 Java/Python 2、代码应遵循低内聚、高耦合 3、代码应具备优良的注释习惯
	关系抽取结果	结果为具有表头的 csv 文件，编码格式为 UTF-8
	schema.json 生成源代码	1、编码语言为 Java/Python 2、代码应遵循低内聚、高耦合 3、代码应具备优良的注释习惯
	schema.json 结果	满足相关文档要求即可
	mapping.json 生成源代码	1、编码语言为 Java/Python 2、代码应遵循低内聚、高耦合 3、代码应具备优良的注释习惯
	mapping.json 结果	满足相关文档要求即可
	电力规章制度知识图谱自动化生成程序源代码	
	电力规章制度知识图谱自动化生成程序执行说明文档	应具有的内容： 1、执行命令说明 2、执行参数说明 3、执行示例演示说明

#### 4.2. 电力规章制度智能问答能力

模块二	提交项	提交项要求
电力规章制度智能问答模型	电力规章制度智能问答模型代码说明文档	应具有的内容： 1、设计思路及实现方法 2、模型的合理性分析 3、模型的测试报告

		4、模型的不足及改进策略
电力规章制度智能问答模型代码		1、编码语言为 Java/Python 2、代码应遵循低内聚、高耦合 3、代码应具备优良的注释习惯
电力规章制度智能问答模型自拟测试集（含标准问与标准答）		1、结果为具有表头的 csv 文件，编码格式为 UTF-8 2、至少 50 条
电力规章制度智能问答验证程序源代码		要求： 1、“读取测试集”模块需要写成可配置模式，并暴露配置参数 2、每段测试需添加打印响应时长函数，并在最后打印响应总时长 3、最后需要计算本次测试的正确率，计算测试正确率耗时不算入响应总时长内。
电力规章制度智能问答验证程序执行说明文档		应具有的内容： 1、执行命令说明 2、执行参数说明 3、执行示例演示说明

## 5. 评估内容

测试步骤中的问题只做参考，实际评估时会有不同。

### 5.1. 电力规章制度知识图谱问答

#### 5.1.1. 推理

测试项	推理问答	用例编号	UC001
业务场景	没有配置相关知识的情况下，能够模糊推荐最相关的知识		
测试目的	测试系统是否支持推理，推荐相关问题		
前置条件	测试环境搭建完毕，构建电力规章制度知识图谱。		
测试步骤	例如： 输入“电力调度的定义是什么”；		
预期结果	电力调度机构依法行使调度权，对所辖电网运行进行的组织、指挥、指导和协调		
测试得分			

#### 5.1.2. 复杂条件句

测试项	复杂条件句	用例编号	UC002
测试场景	针对主配网调管范围的比较。		
测试目的	测试系统是否具备知识图谱的存储及查询能力		
前置条件	测试环境搭建完毕，构建电力规章制度知识图谱。		
测试步骤	例如： 输入“涉及 110kV 主变，主配网调管范围的不同之处？”		
预期结果	主网调管范围：110kV 主变中性点接地刀闸及主变分接头开关 配网调管范围：110kV 主变及高压侧开关及其附属的一次设备。		
测试得分			

#### 5.1.3. 多轮对话（指代消解和省略恢复）

测试项	多轮对话（指代消解和省略恢复）	用例编号	UC004
-----	-----------------	------	-------



测试场景	针对两个多轮对话场景，演示上下文识别能力。
测试目的	支持对话过程中的上下文关联，能够根据上下文识别问题。
前置条件	测试环境搭建完毕，构建电力规章制度知识图谱。
测试步骤	<p>现场演示，自主设计演示上下文识别功能。</p> <p>举例：</p> <p>Q: 电力人身事故事件有哪些分类？</p> <p>A: 人身死亡、人身重伤、人身轻伤、人身轻微伤</p> <p>Q: 人身重伤的界定内容是什么？</p> <p>A: 指使人肢体残废、毁人容貌、丧失听觉、丧失视觉、丧失其他器官功能或者其他对于人身健康有重大伤害的损伤。</p> <p>Q: 人身重伤的界定依据是什么？</p> <p>A: 事故发生后人员受伤且损失工作日</p> <p>Q: 人身重伤的界定依据阈值是多少？</p> <p>A: 等于和超过 105 日的失能伤害</p> <p>Q: 人身重伤的界定依据需要哪些机构给出？</p> <p>A: 依照国家法规及县级以上医疗机构诊断确定</p> <p>Q: 人身重伤的工作日计算标准执行文件是？</p> <p>A: 《事故伤害损失工作日标准》（GB/T 15499-1995）</p> <p>考核：</p> <p>选手：电力人身事故事件有哪些分类？</p> <p>机器：人身死亡、人身重伤、人身轻伤、人身轻微伤</p> <p>选手：第二个的界定内容是什么？</p> <p>机器：指使人肢体残废、毁人容貌、丧失听觉、丧失视觉、丧失其他器官功能或者其他对于人身健康有重大伤害的损伤。</p> <p>选手：如何界定？</p> <p>机器：事故发生后人员受伤且损失工作日</p> <p>选手：界定依据阈值是多少？</p> <p>机器：等于和超过 105 日的失能伤害</p> <p>选手：需要哪些机构给出？</p> <p>机器：依照国家法规及县级以上医疗机构诊断确定</p> <p>选手：那工作日计算标准执行文件是哪个？</p> <p>机器：《事故伤害损失工作日标准》（GB/T 15499-1995）</p>
预期结果	智能问答系统能够准确理解选手问题，并基于上下文提供答案。
测试得分	

#### 5.1.4. 实体继承

测试项	实体继承	用例编号	UC005
业务场景	多轮对话中，支持实体继承		
测试目的	测试系统多轮对话是否支持实体继承		
前置条件	测试环境搭建完毕，构建电力规章制度知识图谱。		

测试步骤	例如： 1、输入“涉及 110kV 主变，主网调管范围的不同之处”； 2、输入“那配网调管的范围呢？”
预期结果	1、110kV 主变中性点接地刀闸及主变分接头开关 2、110kV 主变及高压侧开关及其附属的一次设备
测试得分	

## 5.2. 中控路由

测试项	多轮对话和 FAQ 互转	用例编号	UC006
测试场景	测试中控路由，多轮对话和 FAQ 互转		
测试目的	验证智能问答系统多轮任务的综合能力，能够按照设计逐级引导选手去解决问题，允许选手跨节点咨询问题，能够实现上下文理解及记忆功能。		
前置条件	测试环境搭建完毕，构建电力规章制度知识图谱。		
测试步骤	<p>现场演示，自主设计演示上下文识别功能。</p> <p>举例：</p> <p>Q: 电力人身事故事件有哪些分类？</p> <p>A: 人身死亡、人身重伤、人身轻伤、人身轻微伤</p> <p>Q: 人身重伤的界定内容是什么？</p> <p>A: 指使人肢体残废、毁人容貌、丧失听觉、丧失视觉、丧失其他器官功能或者其他对于人身健康有重大伤害的损伤。</p> <p>Q: 人身重伤的界定依据是什么？</p> <p>A: 事故发生后人员受伤且损失工作日</p> <p>Q: 人身轻微伤的界定内容是什么？</p> <p>A: 指各种致伤因素所致的人身原发性损伤，造成人体组织器官结构轻微损害或者轻微功能障碍。</p> <p>Q: 人身轻微伤与人身重伤的界定依据阈值有什么不同？</p> <p>A: 人身轻微伤：受伤人员需要暂停工作进行如消毒、包扎、敷药、服药等简单医疗救治后即可恢复，或者需要至少休息 1 日；人身重伤：等于和超过 105 日的失能伤害</p> <p>考核：</p> <p>选手：电力人身事故事件有哪些分类？</p> <p>机器：人身死亡、人身重伤、人身轻伤、人身轻微伤</p> <p>选手：第二个的界定内容是什么？</p> <p>机器：指使人肢体残废、毁人容貌、丧失听觉、丧失视觉、丧失其他器官功能或者其他对于人身健康有重大伤害的损伤。</p> <p>选手：如何界定？</p> <p>机器：事故发生后人员受伤且损失工作日</p> <p>选手：第四个的界定内容又是什么？</p> <p>机器：指各种致伤因素所致的人身原发性损伤，造成人体组织器官结构轻微损害或者轻微功能障碍</p> <p>选手：那与第二种事件在界定依据阈值上有什么不同？</p>		

	机器：人身轻微伤：受伤人员需要暂停工作进行如消毒、包扎、敷药、服药等简单医疗救治后即可恢复，或者需要至少休息 1 日；人身重伤：等于和超过 105 日的失能伤害
预期结果	智能问答系统能够精确识别问题类型是标准问答，还是多轮对话；并且依据识别结果进行不同方法的回答
测试得分	

## 6. 评分标准

序号	评测指标	分值	评测指标定义及说明	测试项考察点	评分标准
1	提交项要求	10	定义：提交项能否满足要求	提交项目完整度	本项分值为 10 分。按规定要求提交且结果正确满分；未按规定要求提交酌情扣分。
2	知识图谱构建能力	40	定义：知识图谱构建能力是否满足要求	1、知识抽取说明文档 2、知识抽取结果 3、图谱构建结果	本项分值为 40 分。文档内容占 10 分，内容不完备者酌情扣分；知识抽取结果满分 20 分；图谱自动化构建项满分 10 分。
3	问答性能	10	定义：监测选手说话结束以后到机器人响应的时间间隔。	现场测试问题发出到机器人响应的的时间间隔。（以现场时间来判断）	响应时间小于 1.5s 得 10 分，每增加 0.5 秒，扣 5 分，不足 0.5 秒按线性计算，扣完为止
4	基础问答准确率	30	定义：监测在交互过程中意图识别和问答精准程度。	1、电力规章制度智能问答模型说明文档 2、基于图谱能力进行问答，范围限测试范围中能力，根据准确率进行打分	本项分值为 30 分。文档内容占 10 分，内容不完备者酌情扣分；准确率测试共 10 次，全部准确得 20 分，未定位问题或回答错误 1 次扣 2 分，扣完为止。
5	中控路由能力	10	定义：测试中控路由，多轮对话和 FAQ 互转	验证智能问答系统多轮任务的综合能力，能够按照设计逐级引导选手去解决问题，允许选手跨节点咨询问题，能够实现上下文理解及记忆功能。	本项分值为 10 分。功能具备得 10 分。
	合计	100			