

图书分类号: TP391

单位代码: 10110

学 号: SZ202113022

UDC : 004

密 级: 非密



中北大学
NORTH UNIVERSITY OF CHINA

专业学位硕士学位论文

基于大语言模型的地震应急知识 问答技术研究

作者姓名 闫 晟
校内导师 宋文爱 教授
校外导师 王 青 研究员
学科领域 软件工程
研究方向 软件工程
培养单位 中北大学
学习年限 2021年9月至2024年6月

二〇二四年六月

图书分类号：TP391

单位代码：10110

学 号：SZ202113022

UDC : 004

密 级：非密

中北大学

2024 届专业学位硕士学位论文

基于大语言模型的地震应急知识 问答技术研究

作者姓名 闫 晟

校内导师 宋文爱 教授

校外导师 王 青 研究员

学科领域 软件工程

研究方向 软件工程

培养单位 中北大学

学习年限 2021 年 9 月至 2024 年 6 月

二〇二四年六月

**Research on question answering technique of
earthquake emergency knowledge based on
large language model**

A Thesis Presented to

the School of Software

North University of China

In Partial Fulfillment of the Requirements

For the Degree of Master of Software Engineering

By Yan Sheng

Under the Supervision of Song Wenai

June, 2024

图书分类号：TP391

单位代码：10110

学号：SZ202113022

UDC：004

密级：非密

基于大语言模型的地震应急知识 问答技术研究

作者姓名：_____ 闫晟 _____

校内指导教师（姓名、职称）：_____ 宋文爱 教授 _____

校外指导教师（姓名、职称）：_____ 王青 研究员 _____

学生类别（培养方式）：_____ 全日制 _____

申请学位类别：_____ 工程硕士学位 _____

学科领域（研究方向）名称：_____ 软件工程 _____

论文提交日期：_____ 2024 年 05 月 30 日 _____

论文答辩日期：_____ 2024 年 06 月 01 日 _____

学位授予日期：_____ 年 _____ 月 _____ 日 _____

学位论文评阅人：_____ 曹力 陈海强 _____

答辩委员会主席：_____ 王彬 _____

二〇二四年六月

摘要

我国是一个地震灾害频发的国家，每年由于地震灾害造成的经济损失与人员伤亡巨大。并且受灾人群由于相关应急知识储备不足，导致本可避免的事故时有发生。故加强大众灾害应急知识储备，增强大众灾害风险应对能力将是预防地震灾害和减少损失的关键举措。而随着人工智能技术的发展，大语言模型（Large Language Model, LLM）问答已成为一种高效、直接的信息获取方式，其问答能力及文本生成能力有目共睹，能够快速地为用户提供答案，有效地降低用户搜索的时间成本。

然而，一方面地震应急领域数据结构复杂、种类繁多，尤以非结构化文本居多。相关领域知识缺少一种高效且直观的方式进行表达，无法为公众提供便捷的查询和科普服务。在另一方面，在面对专业领域问题时，大语言模型会出现回答不准确甚至错误的现象。

针对以上问题，本文结合实际情况，提出构建地震应急知识图谱，利用其结构化的表示方式并结合大语言模型，提出了 EEQA-GLM6B 问答模型并搭建了地震应急问答服务，用户借助该服务能够获取地震应急知识，具有实际意义和应用价值。本文以设计和构建 EEQA-GLM6B 问答模型为主线展开研究，主要工作内容有：

（1）针对相关数据复杂、查询不便等问题，本文提出构建地震应急知识图谱，将地震应急领域的知识进行结构化表示。本文结合当前国内外研究成果与实践效果，提出将 BERT-BiLSTM-CRF 模型运用在地震应急知识抽取任务中。经过知识抽取、知识融合及知识存储等流程，最终，完成了地震应急知识图谱的构建。

（2）设计并构建 EEQA-GLM6B 问答模型。本文对现有的知识库问答技术做出总结并对基于大语言模型的问答技术进行分析，利用 GLM (General Language Model) 大语言模型并结合地震应急知识图谱的优势，构建了服务于地震应急领域的 EEQA-GLM6B 问答模型。经过对比，EEQA-GLM6B 问答模型能较准确地向用户提供相关问题的答案，验证了结合大语言模型与知识图谱的问答方案的可行性。

（3）搭建地震应急知识服务平台。本文利用 Python、Flask、Neo4j 等技术实现了地震应急问答服务，实现了地震应急知识问答、实体查询、实体识别等功能。经测试，该系统界面简洁，能够提供稳定、准确的问答服务。

关键词：地震应急；大语言模型；知识图谱；问答技术

ABSTRACT

China is a country where earthquakes occur frequently, and the economic losses and casualties caused by earthquakes every year are enormous. Moreover, the disaster-affected people do not have enough knowledge about emergency response in the face of disasters, which leads to avoidable accidents. Therefore, strengthening the public's emergency response knowledge and enhancing the public's ability to cope with disaster risks will be the key to preventing earthquake disasters and reducing losses. With the development of natural language processing technology, Large Language Model (LLM) Q&A has become an efficient and direct way of obtaining information, and its question and answer ability and text generation ability are obvious to all, which can quickly provide users with answers and effectively reduce the time cost of user search.

However, on the one hand, the data structure in the field of earthquake emergency response is complex and varied, especially unstructured text. The related domain knowledge lacks an efficient and intuitive way to express, and cannot provide the public with convenient query and science popularisation services. On the other hand, when facing questions in the professional domain, the large language model may occur the phenomenon of inaccurate or even wrong answers.

For the above problems, this thesis combines the practical situation and proposes to construct an earthquake emergency knowledge graph, using its structured representation and combining with the large language model, proposes the EEQA-GLM6B model and builds an earthquake emergency Q&A service, with which users can obtain earthquake emergency knowledge, which is of practical significance and application value. In this thesis, the main line of research is to design and construct the EEQA-GLM6B model, and the main work contents are:

(1) For the problems of complexity of related data and inconvenience of query, the thesis proposes to construct an earthquake emergency knowledge graph, which is a structured representation of the knowledge in the field of earthquake emergency. Combining the current domestic and international research results and practical effects, this thesis proposes to apply the BERT-BiLSTM-CRF model in the earthquake emergency

knowledge extraction task. After the processes of knowledge extraction, knowledge fusion and knowledge storage, finally, the construction of earthquake emergency knowledge graph is completed.

(2) Design and construction of EEQA-GLM6B model. The thesis summarises the existing knowledge base Q&A techniques and analyses the Q&A techniques based on the General Language Model, and constructs the EEQA-GLM6B Q&A model that serves the earthquake emergency field by using the GLM (General Language Model) and combining with the advantages of the Earthquake Emergency Response Knowledge Graph. After comparison, the EEQA-GLM6B model can provide users with answers to relevant questions more accurately, which verifies the feasibility of the Q&A scheme that combines the General Language Model and the Knowledge Graph.

(3) Establishing an earthquake emergency knowledge service platform. The thesis uses Python, Flask, Neo4j and other technologies to implement an earthquake emergency Q&A service, which realises the functions of earthquake emergency knowledge Q&A, entity query, entity recognition and so on. After testing, the system has a simple interface and can provide stable and accurate Q&A service.

Key words: Earthquake Emergency; Large Language Models; Knowledge Graph; Q&A Technique

目 录

1 绪论.....	1
1.1 课题研究背景和意义.....	1
1.1.1 课题背景.....	1
1.1.2 研究意义.....	2
1.2 国内外研究现状.....	3
1.2.1 地震应急领域研究现状.....	3
1.2.2 知识图谱及相关构建技术研究现状.....	5
1.2.3 基于深度学习的知识抽取技术研究现状	6
1.2.4 结合大语言模型的知识库问答研究现状	7
1.3 主要研究内容与创新点.....	9
1.4 论文组织结构介绍.....	9
2 相关理论与技术基础	11
2.1 知识图谱构建相关理论技术	11
2.1.1 知识图谱相关理论及构建流程.....	11
2.1.2 知识抽取及相关方法.....	11
2.1.3 知识存储技术.....	12
2.2 问答技术相关理论.....	13
2.2.1 传统的基于知识库的问答技术.....	13
2.2.2 基于大语言模型的问答技术.....	14
2.2.3 主流的大语言模型.....	15
2.2.4 知识图谱与大语言模型结合.....	15
2.3 相关技术与模型介绍.....	16
2.3.1 Transformer 模型	16
2.3.2 BERT 模型.....	18
2.3.3 Langchain 及其功能模块.....	19
2.4 本章小结.....	21
3 地震应急知识图谱的构建	23
3.1 需求分析与构建流程.....	23
3.1.1 需求分析.....	23
3.1.2 构建流程.....	23
3.2 基于 BERT-BiLSTM-CRF 模型的实体关系抽取方法.....	24
3.2.1 BERT-BiLSTM-CRF 模型	24
3.2.2 训练数据的收集与处理.....	26
3.2.3 评价指标.....	27

3.2.4	实验环境和参数设置.....	27
3.2.5	实验结果对比与分析.....	28
3.3	地震应急数据获取及预处理	29
3.3.1	数据来源及获取方法.....	29
3.3.2	数据预处理.....	31
3.4	地震应急相关知识抽取.....	31
3.4.1	实验环境及参数设置.....	31
3.4.2	知识抽取结果.....	32
3.5	知识融合.....	33
3.6	知识存储.....	33
3.7	本章小结.....	34
4	EEQA-GLM6B 问答模型的设计与构建.....	35
4.1	大语言模型预训练框架介绍	35
4.2	GLM 模型结构.....	35
4.3	EEQA-GLM6B 模型构建过程	37
4.3.1	EEQA-GLM6B 模型构建环境	37
4.3.2	EEQA-GLM6B 模型结构	38
4.4	问答实验过程与样例分析对比	39
4.4.1	问答实验过程.....	39
4.4.2	实验结果.....	41
4.4.3	实验结果分析.....	42
4.4.4	样例对比.....	44
4.4.5	模型评价.....	47
4.5	本章小结.....	48
5	地震应急知识问答服务的实现	49
5.1	需求分析.....	49
5.1.1	功能性需求.....	49
5.1.2	非功能性需求.....	49
5.2	平台架构设计.....	50
5.2.1	实体查询功能模块.....	50
5.2.2	地震应急知识问答功能模块.....	51
5.2.3	实体识别功能模块.....	52
5.3	系统功能实现.....	52
5.3.1	实体查询功能.....	52
5.3.2	智能问答功能.....	54
5.3.3	实体识别功能.....	55
5.4	系统功能测试.....	55
5.5	本章小结.....	56

6 总结与展望.....	57
6.1 全文工作总结.....	57
6.2 工作不足与展望.....	57
参考文献.....	59

1 绪论

1.1 课题研究背景和意义

1.1.1 课题背景

地质灾害是指在自然或者人为因素的作用下形成的，对人类生命财产造成的损失、对环境造成破坏的地质作用或地质现象^[1]。常见的地质灾害主要有地震、滑坡、泥石流、地面塌陷等。其中，地震尤为严重，它不仅极易导致大量的人员伤亡，更会对房屋、道路、桥梁等关键工程设施以及宝贵的土地资源造成巨大破坏，这无疑会进一步加剧地震灾害所带来的损失。

我国地理位置比较特殊，地处太平洋板块、印度洋板块和亚欧板块三大板块交界处，位于环太平洋地震带和欧亚地震带两大全球地震活动带之间，地壳运动强烈，地震活动频率高、范围广、强度大^[2]。并且我国很多城市和重要建筑都分布在地震带上，因此，我国是一个地震灾害多发、频发的国家。经统计，自 2008 年以来，我国发生的地震就造成近十万人死亡和数十万人受伤，相关伤亡情况如表 1-1 所示。

表 1-1 2008 年以来造成巨大损失的地震统计

Tab.1-1 Statistics of earthquakes that have caused significant damage since 2008

日期	发生地点	震级	震中烈度	死亡人数	受伤人数	伤亡比
20080512	四川汶川	8.0	XI	87150	375783	4
20080830	四川仁和区与会理区交界	6.1	VIII	41	1010	25
20100414	青海玉树	7.1	IX	2968	11000	4
20130420	四川芦山	7.0	IX	196	13019	66
20130722	甘肃岷县漳县	6.6	VIII	95	2414	25
20140803	云南鲁甸	6.5	IX	729	3143	4
20170808	四川九寨沟	7.0	IX	30	543	18
20220905	四川泸定	6.8	IX	118	423	4
20231218	甘肃积石山	6.2	VIII	151	983	6.5

地震可在较短的时间内造成大规模的损失，如何快速应对突发的地震及地震引发的次生灾害，最大程度地减少人员伤亡和财产损失，是当前必须面对的一项严峻课题。

汶川大地震后的十年，是我国地震应急救援事业蓬勃发展的十年，尤其是在各级各类救援队伍能力建设方面取得了长足的发展^[3]。地震应急救援力量进一步发展壮大，训练设施得到较大的完善，法律法规及标准建设得到加强^[4]。

2018年十三届人大一次会议提出整合国家应急救援力量组建应急管理部，由此实现了由“单灾种”的条块管理，向多灾种的综合管理、综合减灾、综合救灾的转变。应急管理部的组建，从应急管理机制、指挥调度体系、救援队伍建设和应急响应机制方面均有了较大的改变。

整体而言，随着我国经济实力的提升、科技水平的提高，我国在应对突发的地质灾害或公共卫生事件、环境安全事件时已经拥有相对成熟的应急部署和应对预案。地震及地震次生灾害造成的人员伤亡与经济损失已逐年降低。

即使如此，地震及地震次生灾害所造成的损失仍然不可小视，造成这类型问题的原因一方面是由于地震及地震次生灾害一旦事发，受灾规模普遍较大，具有时间紧迫性和资源紧缺性。另一方面，由于受灾人群在面对突发的灾难时极易产生畏惧心理，并且在面对这类型灾害时应当采取的相关应急知识储备不足，也会导致人员伤亡和财产损失。故地质灾害知识与应急知识科普刻不容缓，加强大众灾害应急知识储备，增强大众灾害风险应对能力将是当前预防和减少人员伤亡的关键。

1.1.2 研究意义

随着自然语言处理技术和深度学习技术的飞速发展，地震灾害相关的数据量呈现出爆炸性增长。然而，这些数据来源广泛、种类繁多、结构复杂且格式不统一，导致仅有少量数据能够直接用于分析和应用。同时，相关信息的共享和展现程度相对较低，这无疑对数据处理和分析能力提出了更高的要求和挑战^[5]。

另一方面，公众对于智能知识服务的需求日益迫切，以支持在应急避险和减灾救灾中的快速决策。因此，如何从海量的信息中精准提取并有效表达知识，进而将这些知识应用于智能化的应急服务中，具有极其重要的现实意义和实用价值。而知识图谱（Knowledge Graph, KG）作为一种有效的知识组织和表示方式，可以用于构建地震应急知识库，将地震应急知识进行结构化表示，便于检索和展示。

近年来，知识图谱和大语言模型迅速崛起，成为人工智能领域的一大研究热点。

其中，知识图谱技术可以将知识结构化地表示，能够准确地匹配、查询各个实体及关系，大语言模型则可以弥补知识图谱在语义解析和自然语言理解方面的不足。大语言模型有着极强的文本生成能力，但给出的答案往往不够准确，利用知识图谱丰富的结构化知识表示这一优势可以弥补大语言模型生成答案时所欠缺的准确性^[6]。

综上，将知识图谱和大语言模型结合并应用于地震应急领域，不仅可以为用户提供应急问答科普服务，还能够进一步促进大语言模型在垂直领域的研究与应用。这将在一定程度上降低地震给人类社会造成的影响，降低受灾损失，进而提升城市应急响应能力。

目前，针对地震应急领域构建知识图谱以及相关问答技术的研究和应用相对较少^[7]。通过分析实际需求，本文将以地震应急领域相关研究、知识图谱构建技术以及问答技术为基础，对地震应急领域知识问答应用展开研究。最终完成地震应急领域知识图谱的构建，并通过结合大语言模型的方式实现地震应急领域的知识科普与问答服务。

1.2 国内外研究现状

1.2.1 地震应急领域研究现状

国内外科研人员和学者在地震应急领域已有了广泛的探索，就应急知识科普、地震防治和应对策略、重特大地震灾害舆情获取与跟踪以及地震次生灾害防治等方面取得了一定的进展。

在应急知识科普方面，国内众多研究者做了大量的研究。雷羽南等人认为利用短视频的应急科普方式，能够凭借其生动有趣、短小灵活的特点避免以往科普宣传过程中篇幅大、内容多、触及人群少的弊端，能够让应急知识更容易受到大众欢迎^[8]。孙士宏等人认为地震具有自然性和社会性的双重属性，其通过对地震相关的案例进行研究认为知识科普的社会效果会直接影响自然灾害所造成的损失^[9]。

王林等人也认为开展地震预警宣传教育是必要的，其针对当前地震科普宣传方面存在的问题提出既要面向公共行业用户进行科普宣传，也要对普通公众加强相关知识培训和教育^[10]。而刘允林等人选取四川芦山 7.0 级地震、甘肃岷县漳县 6.6 级地震等有较大影响的地震，进行了实地调研，访谈当地震时科普工作开展情况，最终得出，相关调研地区存在震时现场防震减灾科普的产品不足、震时现场防震减灾科普资源整合开发不足、科普效率低等问题^[11]。

地震应对策略方面,张炜超等通过对我国地震应急技术框架分析认为经过 20 多年的发展,我国已形成全国一体化的地震应急指挥技术系统,应急技术发展显著,但仍存在不足和薄弱环节,标准化、智能化、综合化或将成为新方向^[2]。

重特大灾害舆情获取与跟踪方面,袁庆禄等以 2022 年四川芦山 6.1 级、马尔康 5.8 级和泸定 6.8 级 3 次地震后 7 日内发布的相关微博及其评论为数据样本进行分析监测,结果认为震后 24 小时是地震舆情的风险监测关键时段,需要高度关注和密切观察^[12]。陈亚男等认为社交媒体地震舆情数据可实时反映灾区现状,能够为震后应急提供较好的地震灾情舆情研判服务,选取新浪微博作为数据源设计实现地震舆情数据库^[13]。

地震次生灾害防治方面,陆新征等指出密集的高层建筑群可能会引发一系列严重的衍生灾害。这些建筑群的外围护构件在地震中容易坠落,造成安全隐患,同时电梯内的人员也可能因此被困^[14]。除物理性次生灾害之外,心理性次生灾害如“盲目避震”、“盲目搭建防震棚”躲避灾害等也是需要重点研究的方向^[15]。

相较于国内地震应急领域的发展,日本自上世纪 80 年代以来,就开始投入大量人力和物力,建设起全国范围的地震预警和应急响应体系^[16]。当地震监测到震级超过一定阈值时,地震预警网络会迅速传递信息,提前几秒至几十秒的时间给政府、企事业单位和民众。同时,日本还通过各种手段广泛宣传地震知识,教导民众如何应对地震,提高了公众的地震意识和自救能力。

美国是另一个在地震预警与应急响应体系建设方面取得重要成果的国家^[17]。美国的地震预警系统主要依赖于美国地质调查局和地震科学研究所等机构的协作。美国地震预警系统的特点是高度自动化和多样化的警报方式,一旦地震监测到,系统会立即通过警报器、手机应用和电视等多种渠道向公众发送预警信息。

随着经济发展水平日益提升和工业能力的提高,我国在地震应急领域已取得极大成效。而在应急救援动员能力和响应速度方面,我国一直处于世界顶尖水平。

2023 年 12 月 18 日晚,甘肃临夏州积石山县发生 6.2 级地震^[18]。部分群众提前收到预警,并且充分了解到相关应急知识,切实减少地震造成的损失。

在救援期间,应急指挥辅助决策系统、无人机等新型救援装备发挥巨大作用,有力保障了相关指挥部门实时了解现场灾情情况,及时掌握受灾情况及人员伤亡情况,为明确救援救灾重点方向和救援力量高效调度提供关键决策依据。积石山地震伤亡情况如图 1-1 所示^[19]。

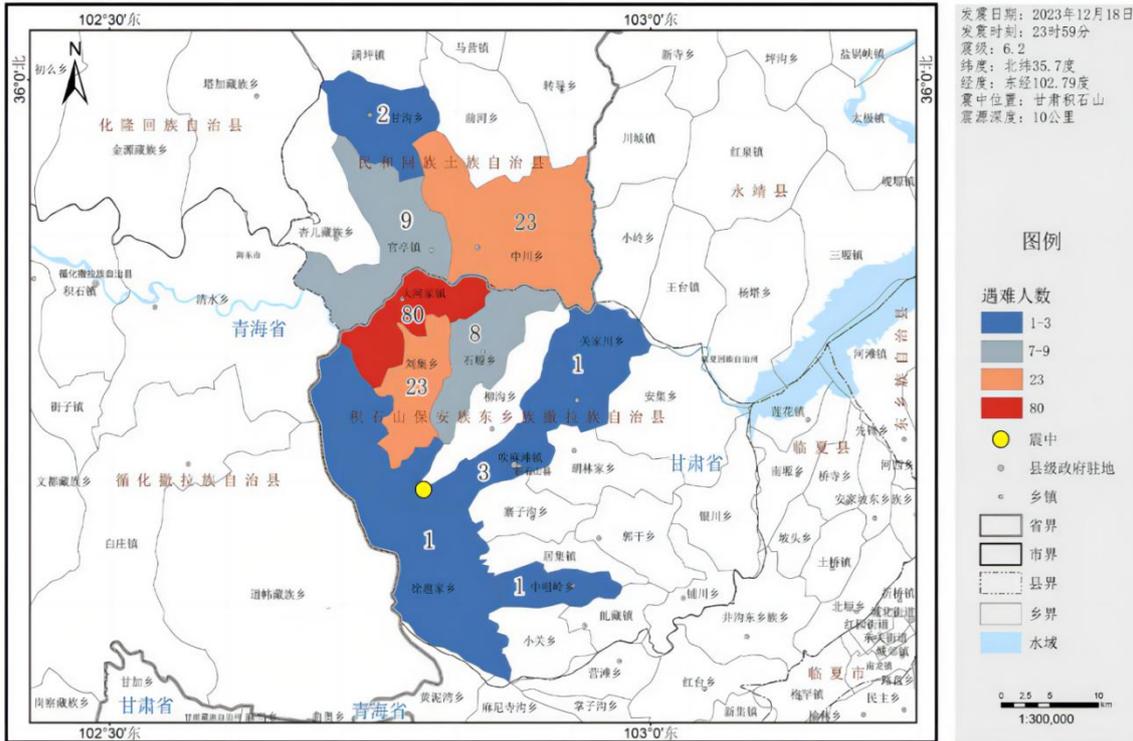


图 1-1 甘肃积石山 6.2 级地震人员伤亡分布图

Fig.1-1 Casualty distribution map of 6.2 earthquake in Jishishan, Gansu province

从国内、国际上的经验来看，地震预警与应急响应体系的建设是一个长期而持续的过程。除了技术的不断创新和改进外，公众的教育和宣传也是至关重要的。通过广泛宣传地震知识，提高公众的地震意识和自救能力，可以使地震预警和应急响应体系发挥最大的效能。地震无法避免，但预警和应急响应体系的建设可以最大限度地降低其对人类社会的影响。

1.2.2 知识图谱及相关构建技术研究现状

在知识图谱里，节点象征客观实体，边则描绘实体间的内在关联。知识图谱构建一般采用自动或半自动技术从结构化、半结构化以及非结构化数据资源中抽取知识^[20]。利用自然语言处理、深度学习等技术从多源异构的数据资源中自动构建知识图谱的技术取得长足进展^[21]。

在大数据时代，深度学习、机器学习等人工智能技术在处理海量数据方面发挥着重要作用，在数据预测和决策制定上具有一定的优势，但是在数据描述上有所不足，因此知识图谱自提出后就备受学术界、工业界关注，发展至今已取得较多成果，其具有强大的数据描述能力，可弥补这方面的不足^[22]。

目前，知识图谱已经在众多垂直领域得以应用，例如利用知识图谱构建相关辅

助决策系统^[23]、问答服务^[24]、企事业内部知识库、推荐系统^[25]等。阎实松利用知识图谱构建了一个飞行器知识问答系统^[26]。付雷杰等人通过对国内知识图谱垂直领域的应用研究，分析了国内的发展现状和趋势，并探讨了知识图谱垂直领域发展中存在的问题和对策^[27]。李岩基于知识图谱构建了政务知识问答系统，用可视化的形式为公务人员展示所搜索的政府公文，以期帮助政府公务人员办公^[28]。申豪杰基于电影行业的观影需求，构建了电影知识图谱，实现模板匹配式的电影问答系统^[29]。

在计算机技术的不断革新与发展下，在灾害应急领域的研究中，有了一定的探索和应用实践。杜志强等人采用自顶向下和自底向上相结合的方法，完成了自然灾害应急知识图谱的构建^[30]。刘汉龙等人详细总结了大数据技术在地质灾害防治中的应用，并且介绍了在地质灾害数据获取、存储、分析的几项关键的基于大数据的方法和技术，从滑坡、泥石流、崩塌、地面变形四个方面综述了迄今国内外学者利用大数据技术开展地质灾害研究和防治方面的工作^[1]。刘晓慧重点研究了基于本体建模理论的地质灾害应急决策知识表达方法，对基于预案的地质灾害应急决策所需知识作出形式化表达^[32]。提出了一个突发地质灾害智能应急决策支持模型，最终构建了突发地质灾害应急预案智能化原型系统。陶坤旺等人以九寨沟地震减灾知识图谱为例，利用知识图谱实现对不同用户的信息智能推送^[33]。

但是，在应急减灾领域，知识图谱构建的自动化程度依然不足。海量结构化、非结构化数据给知识图谱的存储和快速构建带来挑战^[34]，相关领域数据质量不高，挖掘高质量的知识还需要处理数据噪音^[35]，就目前的发展来看，知识图谱在应急减灾领域的应用仍存在一些问题，应急减灾领域知识图谱在实际应用中也仍然存在不足。

1.2.3 基于深度学习的知识抽取技术研究现状

基于深度学习的知识抽取方法能够利用模型自动抽取特征，有较高的准确率^[36]。主要分为流水线式（Pipeline）方法和联合抽取（Joint）方法两类。

流水线式方法是指在完成实体识别的基础上，抽取实体之间关系并生成三元组的过程。循环神经网络（Recurrent Neural Network, RNN）在实体及关系抽取任务上曾有较好的表现^[37]。RNN擅长处理带有时序信息的序列，其能够随着时间推移更新自身“记忆”，但在面对长期依赖与梯度消失等问题时，其表现较差，难以解决。为解决相关问题，研究者们引入了双向长短期记忆网络（Bi-directional Long Short-term Memory network, BiLSTM）^[38]。Zhang等人使用了BiLSTM模型结合当前词语之前和词语之后的信息进行关系抽取，取得了极大的进步，证实了BiLSTM在关

系抽取任务的有效性^[39]。Dai等人利用BiLSTM提取每个单词的表示特征并且结合条件随机场（Conditional Random Fields, CRF）模型对实体进行识别^[40]。

而随着Transformer^[41]的提出，BERT（Bidirectional Encoder Representations from Transformers）^[42]等预训练模型逐步流行起来，这类模型是已经在大量数据集上训练并保存的网络模型，目前已大量运用在流水线式抽取方法中。

流水线式抽取方法的特点在于将实体识别任务和关系抽取任务分为两个独立的阶段，分别优化和改进。但是该方法存在误差传播和实体冗余等问题，为了解决这些问题，研究者们将实体识别任务和关系抽取任务结合起来，构建实体关系联合抽取模型^[43]。

联合式抽取方法则统一了实体关系抽取过程，能够解决误差传播的问题。其直接从非结构化文本中直接抽取出结构化三元组，其中在编解码过程中的步骤又有所区别，故可以将联合抽取方法分为参数共享方法和联合解码方法。

Wei等人针对重叠三元组问题采用参数共享的方法提出了CasRel模型。该模型首先抽取头实体，然后根据头实体的相关信息抽取尾实体和实体间关系。这种方法在一定程度上克服了关系重叠的问题，并获得了优异的抽取效果^[44]。Wang等提出了单阶段的联合解码方法TPLinker（Token Pair Linking），其运用“握手标注”方案，利用标注的矩阵提取所有实体和重叠关系，在关系抽取任务上也取得了不错的表现^[45]。

关于流水线式方法和联合抽取方法的优劣，学术界一直存在不同见解。大量的研究指出，联合抽取方法在捕捉实体与关系之间的交互作用以及缓解误差传播方面具有优势。然而，值得注意的是，在不同资源条件、任务场景和数据集下，联合抽取模型的表现未必能超越流水线式方法。因此，在评估这两种方法的有效性时，必须充分考虑特定领域的的数据特性及质量。

1.2.4 结合大语言模型的知识库问答研究现状

问答系统的发展与自然语言处理技术和人工智能的探索息息相关。最初的问答系统也被称为专家系统。其主要基于预先定义的规则和知识库，通过匹配问题和预设答案来提供回答。而现代的问答系统通常结合了机器学习、深度学习及自然语言处理等技术。它们可以通过分析大量的文本数据来学习语言模式，从而更准确地理解问题并生成符合用户意图的回答^[46]。

2012年，随着谷歌提出知识图谱这个概念，知识图谱相关技术快速发展，其能够为智能问答系统的实现提供了优质的知识来源，并且可以让专业人员更好地帮助

用户去学习、沉浸使用真实世界中各类实体概念之间的联系。由此开始出现了基于结构化知识库的问答方法（Knowledge Base Question Answering, KBQA）^[47]，其根据对自然语言问题的理解并依赖知识库获取问题答案。

基于知识图谱的问答系统能够增强用户获取知识的便捷性，不但能够节省筛选信息的时间，还能获得更加精确、更符合问题需求的答案^[48]。最典型的案例就是人工智能信息助手。其典型代表有百度小度语音、小米小爱助手等。

而在中文问答领域，由于中文的语法和语义十分复杂，问句往往以短文本的方式出现，这也为中文知识图谱问答的发展带来了新的研究课题^[49]。并且传统的通过人工问答模版的方法构建的问答系统满足不了如今的需求，因此结合深度学习模型构建问答系统以提高检索能力成为当前研究的主流。在医疗领域，侯梦薇等人基于肝病知识图谱构建了肝病基础知识问答系统，有效地解决了一部分肝病相关的问题，节省医生和患者大量时间，推动了问答系统在医疗领域的发展^[50]。

多年来，经过相关研究者的不断探索和改进，大型语言模型（Large Language Models, LLM；下称大语言模型）在多种自然语言处理任务上展现出了令人惊叹的能力^[51]。通过预训练（Pre-Training）、微调（Fine-tuning）等技术，能够使大语言模型正确理解问题并且能够回答复杂问题。大语言模型在各种自然语言处理任务上性能卓越，尤以文本生成能力见长，这便为回答问题提供了个性化的模板。

张鹤译等人就此思路提出一种利用大语言模型并结合专业知识库的基于提示学习（Prompt Learning）的问答技术方案，并基于大语言模型和知识图谱技术设计了一套医疗垂直领域的问答系统，探索了大语言模型和知识图谱两种知识表示路线结合的可能性^[52]。

而冯钧等人提出了基于大语言模型的水工程调度知识图谱的构建方法，其利用大语言模型挖掘和抽取数据中蕴藏的知识，完成概念层本体构建和实例层“三步法”提示策略抽取^[53]。

然而，值得注意的是，目前的大语言模型仍然存在一些局限。首先，大语言模型参数量庞大，部署困难，运行调试时所耗费的计算资源不菲。此外，大语言模型理解新出现的事物的能力普遍不强，无法提供事物的最新信息且存在回答逻辑混乱，混淆概念等各种问题。在面对专业领域问答任务时，大语言模型生成的结果可能缺乏真实性和准确性，其处理专业问题的能力仍然有所不足。

1.3 主要研究内容与创新点

我国是一个地震频发、遭受人员伤亡和经济损失十分严重的国家。若想有效地预防地震，减轻地震带来的影响，就要对地震本身和应急知识有深刻的了解。然而当前地震应急领域存在大量的非结构化数据，其中有大量知识值得挖掘。

因此，本文收集并整理了百度百科、应急管理部、国家减灾网、中国政府网公开的地震数据和应急预案文本数据。并以此作为知识库的数据来源，结合现有方法，构建地震应急领域知识库，运用一种新的问答模型为用户问题提供合理回答。本文具体的研究内容如下：

(1) 地震应急知识抽取方法对比与地震应急知识图谱的构建。知识抽取结果的好坏决定着问答知识库的质量高低。本文结合当前国内外的最新研究成果与实践效果，通过对比实验，提出将 BERT-BiLSTM-CRF 模型运用在地震应急知识三元组抽取任务中，在对比其他相关模型后，最终使用 BERT-BiLSTM-CRF 模型完成知识三元组的抽取工作，得到相关三元组 6486 条。最终，通过 Py2neo 模块存储到 Neo4j 图数据库中，完成了地震应急知识图谱的绘制。

(2) 结合大语言模型的地震应急问答模型的设计。本文对基于模板规则、语义解析、向量建模等主流的问答技术与大语言模型问答方法进行分析，对现有问答技术作出总结，提出基于 ChatGLM 大语言模型并结合地震应急知识库设计并构建了 EarthquakeEmergencyQA-GLM6B（下称 EEQA-GLM6B）问答模型，经过对比，EEQA-GLM6B 问答模型能准确地理解用户提出的问题，向用户提供有效的、地震应急领域的专业回答。

(3) 地震应急问答服务的设计与实现。在实现并验证 EEQA-GLM6B 问答模型的有效性之后，本文利用 Python、Flask、Neo4j 等技术实现了一个应急问答服务，该服务主要提供应急问答等功能。经测试，该系统界面简洁，能够提供较好的问答服务。

1.4 论文组织结构介绍

本文共分为六个章节，全文内容结构安排如下：

第一章为本文的绪论。在该章节中，主要介绍了本文的研究背景与研究意义，分析了当前国内外地震及应急相关领域的研究现状、知识图谱构建技术的研究现状

以及知识库问答技术的研究现状。其次，介绍了本文进行的工作和研究内容。最后，介绍了本文的组织结构。

第二章为相关理论与技术基础。在该章节中首先阐述了自然语言处理技术和知识图谱构建的相关理论知识，为后续地震应急知识图谱构建提供理论依据和方法论。其中，重点对知识抽取任务中的相关技术及路线进行了论述，详细地分析了 BERT 模型的结构。接下来，简单介绍了用于存储知识的图数据库 Neo4j。最后阐明在当前知识库问答相关研究中，结合大语言模型的知识库问答方法已经成为一大研究热点，并举例介绍了国内外主流的大语言模型。

第三章为地震应急知识图谱的构建。在该章节中，首先对构建地震应急知识图谱的实际需求进行分析并阐述了地震应急知识图谱的构建过程。其次，对该知识图谱构建过程中所需要的技术和流程进行了介绍，包括数据来源、模型对比、知识抽取以及知识存储。

第四章为 EEQA-GLM6B 模型的设计与构建。该章节主要介绍了当前大语言模型的技术流派，对相关基座模型进行了阐述，分析了该模型的基础结构。其次，构建了针对地震应急问答的 EEQA-GLM6B 模型并通过两项实验验证该模型在实际问答过程中的能力。在对比分析了当前主流的大语言模型后，最后探讨了知识库结合大语言模型这种问答技术方案的未来研究方向。

第五章为地震应急知识问答服务的实现。该章节首先分析了地震应急知识问答服务的相关需求，设计了地震应急问答功能、实体识别功能等功能模块，最后利用 Python、Flask 框架和 HTML、CSS 等前端技术完善并测试了相关功能，并在前端以简洁、直接的页面展示给用户。

第六章为总结与展望。此章节对本文的工作做出全面总结，分析了当前工作所存在的不足，最后指出仍需进一步研究的方向并进行了展望。

2 相关理论与技术基础

2.1 知识图谱构建相关理论技术

2.1.1 知识图谱相关理论及构建流程

知识图谱^[54]这个概念最早是由谷歌公司于 2012 年提出的，是一种结构化的知识表示方式，即结构化知识库。究其根源，知识图谱起源于语义网络，并且跟语义网（Semantic Web）有着极深的渊源。

知识图谱是一种用于描述客观世界中实体、概念及其相互关系的知识表示方法，它由节点和边构成，节点通常代表实体或概念，而边用来表示它们之间的关系。

构建知识图谱具体包括知识获取与预处理、知识抽取、知识融合、本体构建、知识存储、知识查询与推理、知识维护与更新等步骤。在本文中，主要对知识图谱进行简单构建，涉及到数据获取、知识抽取、知识存储等关键流程，因此对于知识图谱的构建流程不做详细阐述。

2.1.2 知识抽取及相关方法

知识图谱作为一种结构化表示知识的方式，知识抽取直接关系到知识图谱的内容丰富度和质量。知识抽取任务重主要是从非结构化或半结构化数据中提取出结构化知识，这个过程将使得数据能够以一种更易于处理和分析的形式存在，一般是以三元组的形式体现，如实体-关系-实体、实体-关系-属性等。在知识抽取任务中，最为核心且不可或缺的两项工作便是命名实体识别和关系抽取。

命名实体识别（Named Entity Recognition, NER）是整个知识抽取任务的基础工作，其任务是准确捕捉并识别出那些具有特殊含义的实体，如人名、组织名、日期、时间、地点等，为之添加相应的标注信息，为后续抽取工作提供便利。

关系抽取（Relation Extraction, RE）在知识抽取中同样占据重要地位，它的目标是从文本中捕捉并提取出实体间的语义关系，如“包含”、“发生”等关系。

对于命名实体识别、关系抽取任务，目前有四种不同的知识抽取方法，分别是基于规则的方法、基于概率图模型的方法、基于深度学习的流水线式抽取方法以及近些年出现的基于深度学习的联合抽取方法。

本文在 1.2.3 小节中指出,流水线式抽取方法先抽取实体再进行关系抽取。但是,在抽取过程中,上一个任务中产生的误差可能会影响到后续任务的实际效果,并且识别出的实体需成对组合进行关系分类,导致无关联的实体对冗余,使得关系分类的错误率升高。由此出现了另一种联合抽取方法,其又分为两种具体做法,分别是基于参数共享的联合抽取方法,例如 CasRel 模型^[44]。除此之外,还有如 TPLinker 等一步解码的方法^[45]。

基于深度学习的流水线式抽取方法将命名实体识别和关系抽取作为两个独立的子任务来处理。首先,使用深度学习模型(如卷积神经网络或循环神经网络)在文本中识别出实体;然后,对识别出的实体进行两两组合,并使用另一个深度学习模型来完成关系的分类。这种方法的优点是两个子任务可以独立设计和优化,但其缺点是存在误差传播问题,即实体识别模型的错误可能会影响到关系分类模型的结果。

为了克服流水线式方法的缺点,近年来出现了基于深度学习的联合抽取方法。这种方法旨在同时优化实体识别和关系抽取两个子任务,以减少误差传播并捕捉实体和关系之间的交互信息。联合抽取方法通常采用端到端的模型设计,如基于参数共享、标注策略或概率图的方法。这些方法的优点是能够更好地利用实体和关系之间的内在联系,提高抽取的准确率。然而,联合抽取方法的模型设计和训练复杂度通常较高,需要更多的计算资源和数据支持。

2.1.3 知识存储技术

在知识图谱中,知识的存储和管理是至关重要的,其既要保证数据的质量和准确性,也要保证数据的扩展性和灵活性。

知识图谱通常需要不断地更新和扩展。因此,存储和管理知识图谱的数据库需要具有良好的扩展性和灵活性,能够支持大规模数据存储和高效的查询和更新。

目前,图数据库管理系统是知识存储的主流方案。图数据库管理系统通过图结构来建模和存储知识图谱,支持高效的查询和推理。图数据库在许多领域都有应用,包括社交网络分析、知识图谱、推荐系统等。随着数据复杂性的增加,图数据库变得越来越受欢迎,因为它们提供了一种有效的方式来处理和查询高度互联的数据。

Neo4j 是目前最流行的开源图数据库。Neo4j 创建的图是用顶点和边构建一个有向图,其查询语言 Cypher 已经成为事实上的标准,并且可以使用 Cypher 查询语言对数据库中的节点、关系及属性进行查询。对存在大量丰富关系的数据,Neo4j 的优

势就在于它能够快速、高效地处理这些数据，能够提供更优秀的性能、更好的扩展性和更高的可靠性，这些优势都使得 Neo4j 成为解决图形问题的理想数据库。

2.2 问答技术相关理论

2.2.1 传统的基于知识库的问答技术

知识库问答 (Knowledge Base Question Answering, KBQA) 是一种基于结构化知识库 (如知识图谱) 的智能问答方法, 其中, 知识库是大量知识的集合。其主要任务是使计算机能够理解自然语言问题, 并从知识库中检索出相关的信息, 以自然语言的形式回答用户的问题。

当用户进行提问时, 相关自然语言问题需要得到正确的理解, 并将其转换为知识库可以理解的查询语句。通过语义解析, 系统可以识别出问题中的实体和关系等信息, 并将其转化为结构化的查询语句。在知识库中执行转换后的查询并返回相关结果, 检索出相关的信息用于答案的生成。答案生成部分用于将查询执行的结果转换为普通用户可以直接理解的自然语言答案。当存在多个可能的答案时, 系统还需要进行答案排序, 以选择最合适的答案返回给用户。

知识库问答方法有模板匹配式、向量建模式以及语义解析式三种路径^[55]。

模板匹配方法是一种基础的框架。用户的自然语言输入首先被转化为结构化的三元组, 随后这些三元组与预先设定的 SPARQL 查询模板进行匹配。CUI W 等人曾利用机器学习方法, 对庞大的问答数据集进行训练, 使系统能够自动构建和优化查询模板。该系统能够识别用户的查询意图, 并将问答对分类匹配至相应模板, 最终生成精准的答复^[56]。

进入 2019 年, Aiting Liu 等人引入了一种创新的方法, 他们利用 BERT 预训练模型来训练系统生成查询模板。这种方法不仅克服了人工创建模板的局限, 还显著减少了算法对庞大训练数据集的依赖^[57]。

而在对于知识库问答的路线, 向量建模方法也占据着关键地位。这种方法的核心思想是将问题与答案转化为向量表示。例如, 2014 年 Bordes Antoine 等人提出的 embedding 模型, 便在不依赖人工干预和特征提取的情况下, 通过向量建模与深度学习技术, 高效训练问题答案对, 实现了快速而准确的问答处理^[58]。

此外, 基于知识图谱的语义解析方法则是一种无需模板辅助的先进策略。它首先将用户问题进行语义上的归一化处理, 转化为知识图谱能够理解的表示形式。随

后，系统将知识图谱中的信息转化为逻辑形式，并在知识库中依此查询得出最终答案。这种方法在无需依赖预设模板的前提下，实现了用户问题的高效解析与回答。

知识库问答的优势在于精确性高、可扩展性强和支持复杂查询。其利用结构化知识库中的精确信息进行回答，因此可以提供更准确、可靠的答案。其可以通过不断地扩展知识库来提高其覆盖范围和回答能力。

尽管知识库问答技术可以利用结构化知识库中的信息进行回答，但知识库的覆盖面仍然是有限的。有些问题可能无法在知识库中找到答案，或者答案并不完整。并且基于知识库的问答系统的性能在很大程度上取决于知识库的质量。如果知识库中存在错误或不一致的信息，那么最终输出也可能会出现错误的回答。此外，由于自然语言问题的表达方式多种多样，在理解问题时也可能存在歧义或不确定性。

2.2.2 基于大语言模型的问答技术

大语言模型在问答过程中的作用主要体现在两个方面。首先，大语言模型可以作为答案生成器，根据用户的问题生成简洁、准确、自然的回答。其次，大语言模型也可以作为问题的理解器，对用户的自然语言问题进行语义解析，识别出问题中的实体和关系，从而准确地定位到知识库中的相关信息。

然而大语言模型的语言生成能力也会出现负面作用，在某些情况下可能会生成不准确或不合理的回答，如图 2-1 所示。

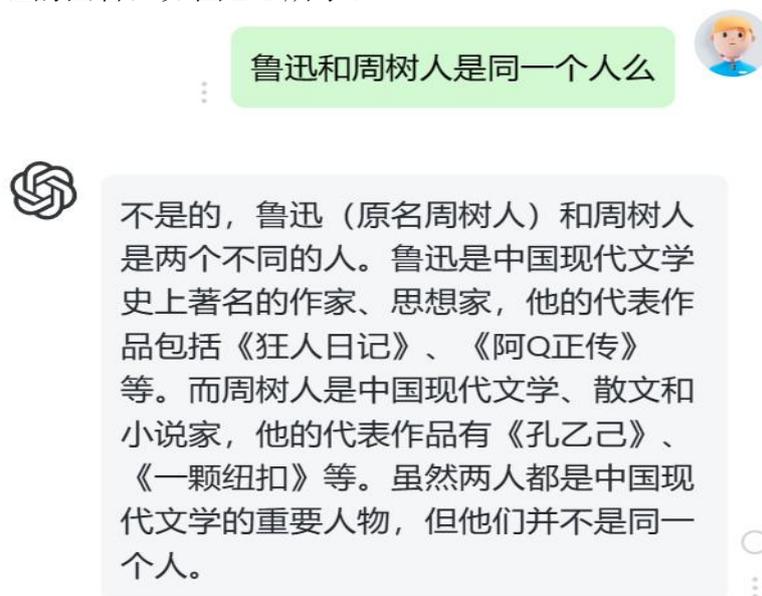


图 2-1 ChatGPT 的典型错误

Fig.2-1 Typical ChatGPT error

2.2.3 主流的大语言模型

当前大语言模型飞速发展，更新迭代速度越来越快。国内外发布的大语言模型如 GLM、GPT、文心大模型、LLaMA 等大模型已经迅速推广并且流行起来，成为许多工作不可或缺的生产力工具。

GLM 是由清华大学知识工程实验室提出的一种基于 Transformer 架构的预训练语言模型结构，目前拥有 1300 亿个参数^[59]。在 GLM-130B 的基础上，智谱 AI 研发出 ChatGLM 这一开源的、支持中英双语问答的对话语言模型，并且推出了 ChatGLM-6B，其优势是训练成本较低，推理速度较快，可以在消费级的显卡上进行本地部署，为相关领域的研究者提供了一个性能与资源消耗之间平衡的解决方案。

而 GPT 则是由 Open AI 团队开发^[60]。GPT 模型通过学习大量的文本语料库来预测下一个可能的单词或短语，从而生成自然语言的文本。GPT 模型的核心就是 Transformer 架构，它使用自注意力机制来处理输入序列，并捕捉序列中不同位置之间的依赖关系。其通过在大型语料库上进行无监督的预训练来学习语言模式和知识，然后在有监督的下游任务上进行微调，如文本分类、机器翻译、问答等。

文心大模型则是由百度公司在 ERNIE (Enhanced Representation through Knowledge Integration) 模型的基础上研发的文心大模型^[61]。ERNIE 技术在预训练阶段使用了不同于 BERT 的掩码语言模型，它不再是对单个字或词进行掩码，而是对整个短语或实体进行掩码，从而更好地学习语言的整体结构和语义信息。此外，其还引入多源数据和多任务学习的方法，将不同领域和场景的知识进行融合，提高了模型的通用性和适应性。

LLaMA 模型是元宇宙平台公司 (Meta) 开源的大语言模型，其架构使用的也是 Transformer Decoder 结构^[62]。2023 年 7 月，Meta 公司发布了 LLaMA 2 模型。LLaMA2 系列模型也是经过预训练和微调的生成式文本模型，其参数量从 70 亿到 700 亿不等。

2.2.4 知识图谱与大语言模型结合

以知识图谱为代表的符号主义和以大语言模型为代表的连接主义在人工智能的发展历史上都具有举足轻重的地位。并且就目前问答技术的相关研究进展而言，知识图谱与大语言模型的结合将是一个具有极大价值的研究课题。

从发展源头来看，知识图谱和大语言模型发展的核心思路具有高度的相似性。大语言模型的预训练与知识图谱的知识建模的过程十分相似，都需要人工整理和理

解数据，但是也可以通过自动化方式提升效率。模型的监督微调与知识抽取任务相对应，都是用于从数据中提取知识。所以，知识图谱和大语言模型都可以看作是知识库，前者是符号化的知识库，而后者则是参数化的知识库。

从知识问答准确性的角度来看，基于知识库的问答结果要优于大语言模型问答的结果，但是基于知识库的问答在解析相关自然语言问句时却存在困难，其需要先将自然语言问句转化为一个结构化查询语言，即 SPARQL，并且现有的语义解析方法在进行结构化转化的时候需要对大量的标注数据进行模型训练和参数优化，因此很难应对用户千变万化的问法。并且由于知识量的限制，很多问题知识库无法回答。

而大语言模型问答不仅仅是基于一个知识库进行问答，它本身还是一个神经网络模型，所以可以直接对自然语言进行理解，将一个自然语言问句进行解析并且自动生成答案，在这个过程中几乎不需要人类的参与，因此大语言模型在面对多样的问答形式时，展现出的泛化能力更强，语言解析能力更强，用户体验更好。

大语言模型和知识图谱各自在问答方面的优劣势都非常明显。因此，通过引入知识图谱的结构化知识，可以增强大语言模型对现实世界的认知和理解能力。同时，大语言模型的强大自然语言处理能力也使得知识图谱中的知识得以更好地被利用和表达。

2.3 相关技术与模型介绍

2.3.1 Transformer 模型

Transformer 模型整体采用编码器-解码器架构，通过采用自注意力机制、多头注意力、位置编码、残差连接和层归一化等技术，实现了对输入序列的高效处理和信息捕捉^[41]。

Transformer 模型中的编码器负责将输入序列转换为一系列高维的特征表示，而解码器则根据这些特征生成输出序列。编码器和解码器都包含多个相同的层堆叠而成，每一层都包含一个自注意力子层和一个前馈神经网络子层。

自注意力机制是 Transformer 模型的核心，它通过计算输入序列中不同位置之间的注意力权重来实现对信息的捕捉和传递。具体来说，自注意力机制将输入序列中的每个位置都视为一个查询，并将其与包括自身的所有位置的键值对进行匹配，从而得到该位置与所有位置之间的注意力权重。这些权重再被用来加权求和所有位置的值，得到该位置的输出表示。

Transformer 模型还采用了多头注意力机制以进一步地增强模型的代表能力。其每个自注意力子层都包含多个并行的自注意力头，每个头都独立地计算输入序列的注意力权重，并将它们的输出拼接在一起作为该子层的输出。这样做可以使模型同时关注输入序列中的不同方面，从而捕捉到更丰富的信息。

而且由于 Transformer 模型本身不具备对输入序列顺序的感知能力，因此需要通过额外的位置编码来引入位置信息。位置编码通常是以固定或可学习的方式添加到输入表示中的，以确保模型能够区分和处理不同位置的信息。

此外，为了缓解神经网络中的梯度消失或爆炸问题，Transformer 模型每个子层的输出都会与其输入进行残差连接（即相加），然后再进行层归一化操作。这样做可以使得模型的训练更加稳定，并且能够加速模型的收敛速度。Transformer 模型结构如图 2-2 所示^[41]。

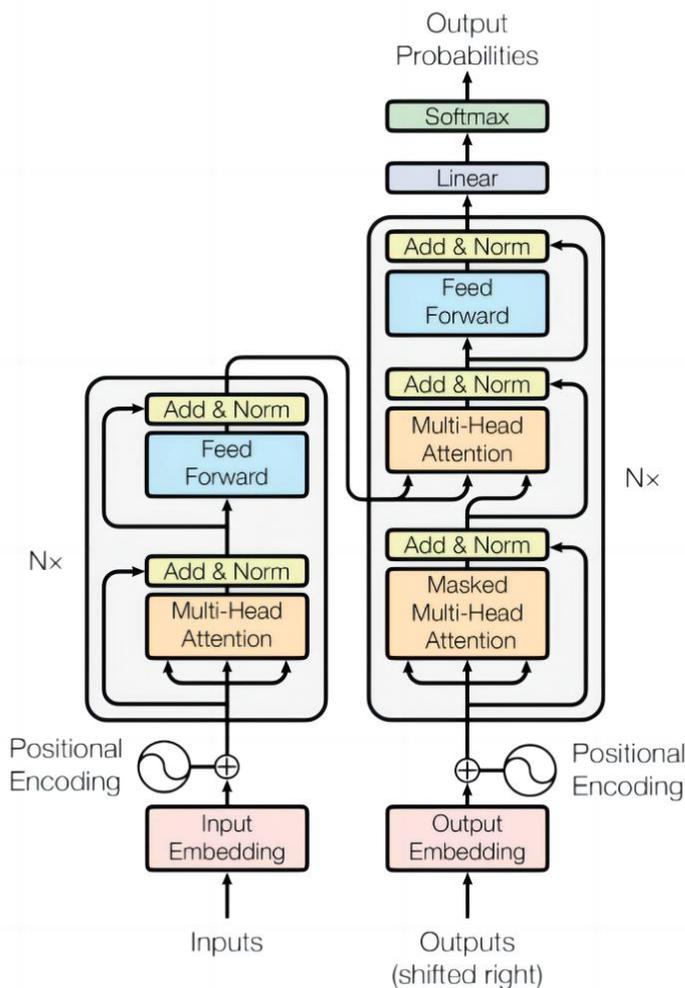


图 2-2 Transformer 模型结构图

Fig.2-2 Structure of the Transformer model

2.3.2 BERT 模型

BERT 自提出以来就备受学者们关注，通过对其进行微调和结构改进，能够取得很好的抽取效果。

BERT 是一种基于 Transformer 的多层双向语言表征模型，其将若干个相同的 Transformer 块叠加起来，通过点积注意力的方法更深层次地学习单词的特征信息^[41]。除 Transformer 编码器外，BERT 模型主要包含遮蔽语言模型(Masked Language Model, MLM)、下一句预测 (Next Sentence Prediction, NSP) 等结构。

BERT 所使用的 Transformer 的编码器 (encoder) 部分是一种采用多头自注意力 (Self-attention) 机制的深度学习神经网络架构。自注意力机制使模型能够在处理某一词的同时参考输入序列中的所有其他词汇，进而捕获词与词之间错综复杂的相互依赖关系。该编码器包含三部分，输入部分、多头注意力机制及前馈神经网络^[42]，Transformer 的编码器具体结构如图 2-3 所示。

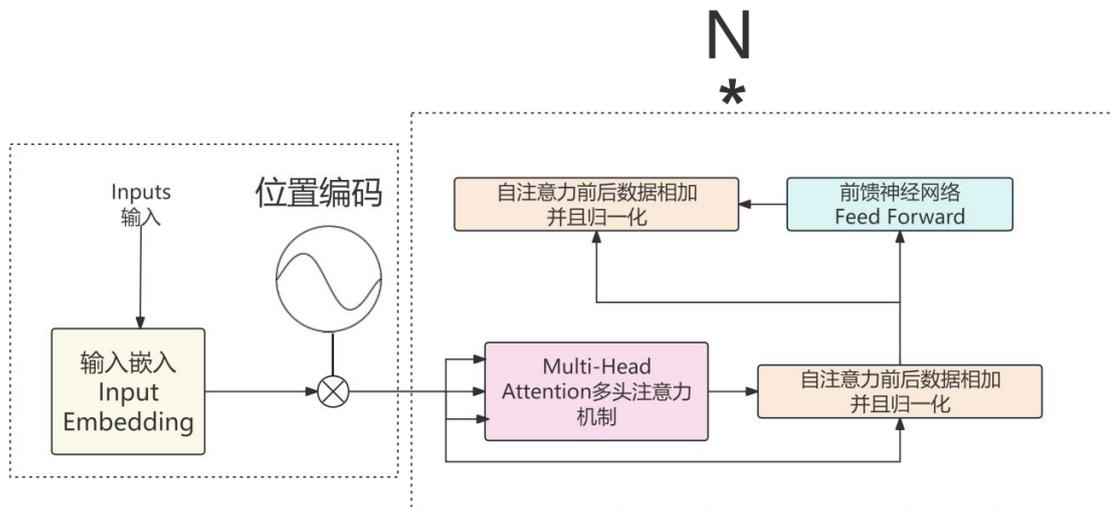


图 2-3 BERT 中的 Transformer 编码器结构

Fig.2-3 Transformer encoder structure in BERT

在输入部分，一个输入是有三个部分组成的，起始的 CLS 标记、由文字组成的子句序列以及子句与子句之间用 SEP 标识进行分割。如式 2-1 所示：

$$\text{input} = ([\text{CLS}], s_1, s_2, \dots, s_m, [\text{SEP}], p_1, p_2, \dots, p_n, [\text{SEP}]) \quad (2-1)$$

其中 CLS 的作用是在句子开头加入的特殊符号，在训练时 CLS 的输出向量接入二分类器完成二分类任务，用作分类任务的表示。其次，子句 $s(s_1, s_2, \dots, s_m)$ 为前序句子，子句 $p(p_1, p_2, \dots, p_n)$ 为后序句子。这样的输入序列既包含了句子之间的前后关

系，也包含了单词之间的顺序关系。最后，SEP 可以理解为句子与句子之间的分隔符，用来标记子句的结束位置。

BERT 模型的嵌入表示分为三部分，第一部分是符号嵌入 Token Embeddings，表示的是句子中单词本身的含义；第二部分是 Segment Embeddings，其表示句子的语义信息以及子句之间的前后顺序信息；而第三个部分 Position Embeddings 表示的是子句中各个单词的前后关系信息。BERT 模型的输入嵌入表示流程的示例如图 2-4 所示。

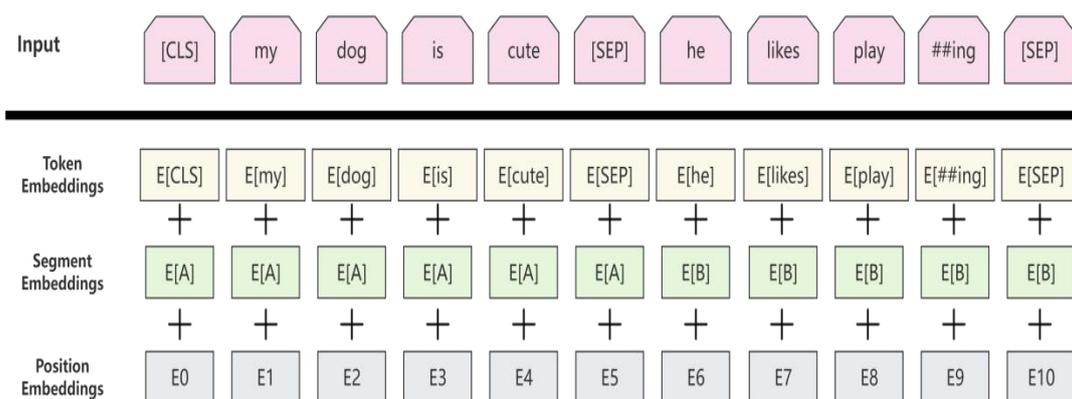


图 2-4 BERT 模型的嵌入过程

Fig.2-4 The embedding process of BERT model

在 BERT 模型的 MLM 模块中，输入序列中的每个词语有 15% 的概率被遮蔽。被遮蔽的词语中，有 80% 的概率被替换为特殊的 MASK 标记，10% 的概率被替换为任意的其他词语，剩下的 10% 的概率保持不变。这种策略迫使模型不仅仅依赖于上下文来预测遮蔽的词语，而且还要考虑到遮蔽本身的不确定性。

在 BERT 模型的 NSP 任务中，BERT 模型接收两个句子作为输入，这两个句子可以是一个连续文本中的两个句子（正样本），或者是从不同文档中随机抽取的两个不相关的句子（负样本）。NSP 结构设计的目的是利用 NSP 结构预测两句话是否连续，即判断上下文的构成是否合理。但近年来的研究表明，在某些情况下，NSP 任务可能并不是必要的，甚至可能会对模型的性能产生负面影响^[63]。

2.3.3 Langchain 及其功能模块

Langchain 可以帮助开发人员使用大语言模型构建人工智能应用程序，可以为 LLM 的开发利用提供有力支撑^[65]。它提供了一套组件、工具和接口，使得相关领域的研究者可以快速构建基于大语言模型的应用。Langchain 共包含六大组件。

(1) 模型 (Models)：即调用或加载的大语言模型。其提供了便捷的接入各类

大语言模型的接口，这其中就包括本文使用的大语言模型 ChatGLM-6B。

(2) 提示词管理 (Prompts)：包括提示管理、提示优化和提示序列化。该模块可对输入按照一定格式进行修饰，对提示模板进行封装，以便于生成用于操作链 (Chains) 或代理 (Agents) 中的提示。其可以简化提示难度，丰富提示灵活性，给开发者提供较为灵活的自定义提示模板。

(3) 代理执行器 (Agents)：可以看做是一个智能化的流程封装。它基于 LLM 的思维链能力，动态串联多个工具或操作链，完成对复杂问题的自动推导和执行解决。Agent 的架构如图 4-2 所示。

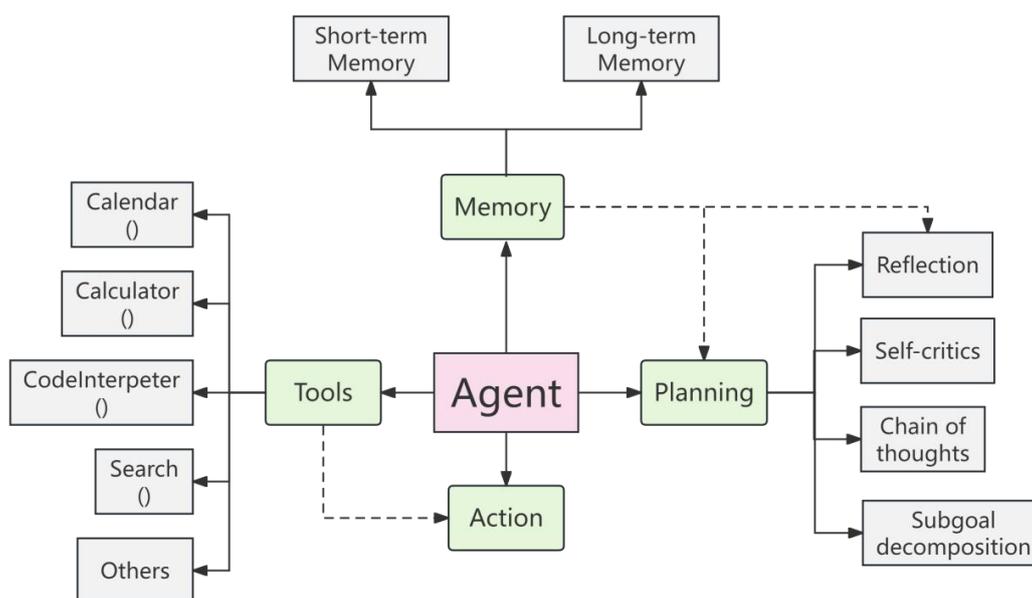


图 4-2 Agent 架构图

Fig.4-2 Architecture of the Agent

(4) 操作链 (Chains)：其旨在更好的理解用户的意图，其最主要的操作链包括 LLM Chain、Sequential Chains、Router Chain、Documents Chain。

LLM Chain 可以接收用户的查询，将其发送给语言模型进行处理，并将模型生成的响应返回给用户，并且可以结合提示管理模块使用。

Sequential Chains 是顺序执行的链式结构，可以按照特定的顺序连接多个处理组件。

Router Chain 是一种具有“路由器”功能的操作链，其会根据信息的内容将信息分发传输到不同的操作链。

Documents Chain 包含了四种操作链，其主要功能是对文件进行处理，在基于文

档生成摘要、基于文档的问答等场景中经常会用到。

除以上主要的操作链之外，Langchain 中还包含 Stuff Documents Chain、Refine Documents Chain、MapReduce Documents Chain、MapRerank Documents Chain 等不同功能的操作链。

(5)交互存储器(Memory): 在 Langchain 中存储信息的方法都被封存在 Memory 模块中，其通过存储对话过程中的上下文记住用户的提示和自身操作，进一步提升用户体验。例如 Conversation Buffer Memory 就是一种简单的记忆力组件，它会记住与大语言模型的问答内容，并在下一轮问答时将历史对话记录传送给大语言模型。

(6) 文件索引 (Index): 该模块用于提供本地文件索引，索引最常用于“检索”步骤中，该步骤指的是根据用户的查询返回最相关的文档，便于从外部知识库获取答案。索引模块检索文档的步骤如图 4-3 所示。

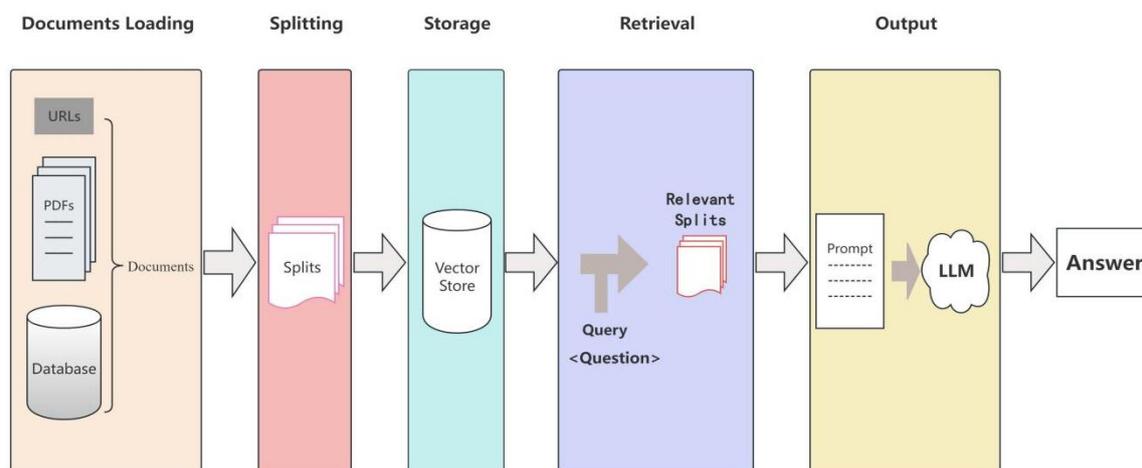


图 4-3 索引模块检索文档的步骤

Fig.4-3 The procedures of the index module to retrieve the document

上图中，首先通过 Document Loaders 中加载各种不同类型的数据源，通过 Text Splitters 进行文本语义分割，通过 Vector store 将非结构化数据进行向量化存储，最后，通过 Retrieval 步骤对文档进行数据检索，生成相应提示后，由大语言模型组织答案并输出。

2.4 本章小结

在本章节中，本文主要对知识图谱的构建过程、知识抽取理论基础及相关技术、相关问答技术的理论基础作出阐述，并且从理论可行性和技术可行性两个角度针对

知识图谱与大语言模型的结合进行了分析，最后对本文所使用的深度学习模型的结构及原理进行了分析。

首先，在 2.1 小节中，本章简要介绍了知识图谱构建流程，针对命名实体识别和关系抽取任务进行了技术概述，接下来简单地介绍了知识存储相关技术以及 Neo4j 图数据库。在 2.2 小节，本章简述了基于知识库的三种问答技术路线并指出当前的知识库问答技术存在不足，进而引入基于大语言模型的问答技术，介绍了当前国内外发布的大语言模型，并且对知识库问答与大语言模型两者相结合的理论基础做出总结。最后，在 2.3 小节，对本文所使用的 Transformer 模型和 BERT 模型的结构及原理进行了分析。

3 地震应急知识图谱的构建

在 2.1.2 小节中, 本文简述了知识抽取任务在知识图谱构建过程中的重要性并介绍了命名实体识别和关系抽取两个关键子任务。本章节中, 将在 3.1 小节介绍构建地震应急知识图谱的需求以及构建流程。在 3.2 小节将介绍相关技术模型, 分析其原理, 对该模型的训练过程及相关评价指标做出阐述, 并分析模型训练的结果。3.3 小节将介绍地震应急数据的来源和获取方法, 并对相关数据进行预处理。在 3.4 小节中将利用 3.2 小节构建的模型进行地震应急知识抽取任务。在 3.5 小节中将介绍地震应急知识图谱的存储技术, 最后在 3.6 小节总结了本章的研究内容。

3.1 需求分析与构建流程

3.1.1 需求分析

对于地震应急知识库的构建而言, 最为紧迫的是快速、准确地获取和整理相关的专业知识, 以供后续的应用和决策支持。故本文对于构建地震应急知识图谱的需求是通过知识图谱将地震应急领域中结构复杂的文本知识进行结构化表示, 并利用图数据库 Neo4j 简洁的界面, 具象地展示地震应急知识库。即通过 Neo4j 提供的可视化服务直观地展示地震应急领域的结构化知识。

因此, 在后续的工作中, 本文的核心任务聚焦于地震应急领域结构化知识的获取和知识图谱的快速构建。鉴于这一明确的目标导向, 我们不需要对知识图谱构建过程的每个环节都进行深入的探索和研究。具体来说, 本文主要关注于知识抽取这一关键环节, 通过精确有效的技术手段, 从地震应急数据中抽取出知识三元组。

3.1.2 构建流程

本文中用于构建知识图谱的数据来源于政府机构网站公开的应急预案、相关法规条例及官方媒体发布的灾情新闻和应急科普知识。

由于本文不进行深层次、高精度的知识融合工作和知识推理工作, 故选择了简单而高效的方法, 仅对抽取出的相关知识三元组进行初步的整合和展示, 形成一个基础的知识图谱框架。地震应急知识图谱的构建包括地震知识数据收集与处理、知识抽取、知识存储等步骤, 具体流程如图 3-1 所示。

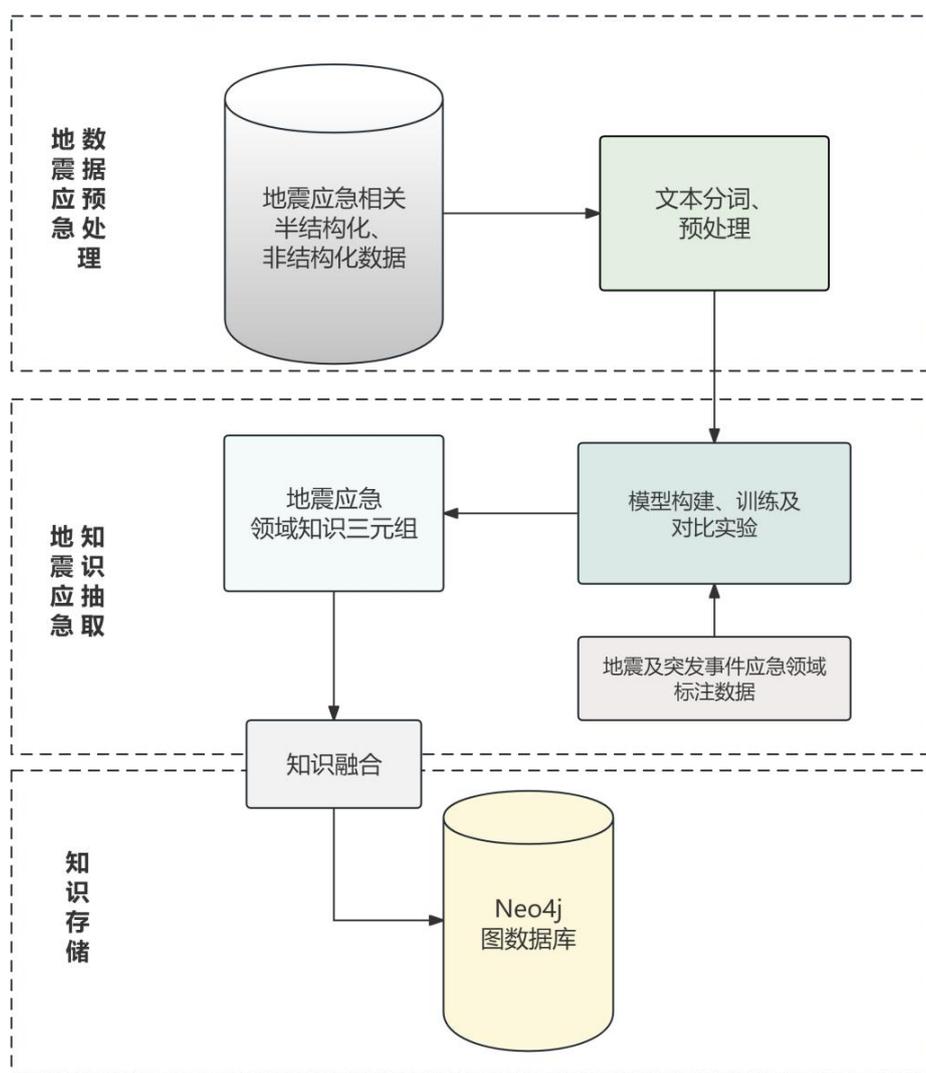


图 3-1 本文知识图谱构建过程

Fig.3-1 The construction process of knowledge graph

3.2 基于 BERT-BiLSTM-CRF 模型的实体关系抽取方法

在 3.2.1 小节中，本文将阐述 BERT-BiLSTM-CRF 模型结构，并将 BERT-BiLSTM-CRF 模型运用于地震应急知识抽取任务中，以验证 BERT-BiLSTM-CRF 模型的性能。

3.2.1 BERT-BiLSTM-CRF 模型

在本小节中，本文将 BERT-BiLSTM-CRF 这一流水线式实体关系抽取模型用于地震应急实体关系抽取任务。在该模型中利用 BERT 对于文本数据进行训练，获得

文本词向量；使用 BiLSTM 神经网络提取文本特征；最后通过 CRF 模型利用图的算法获得最优的标签标注序列。

在 BERT-BiLSTM-CRF 中，BERT 模型是用于联合上下文语境进行预训练，获得相应的词向量。而 BiLSTM 模型由两个独立的 LSTM 网络联合构成，其中一个 LSTM 网络沿序列的正向进行信息处理，另一个则沿序列的逆向进行信息处理^[39]。这种设计是基于对传统 RNN 的改进，因为传统 RNN 在处理长序列时往往难以捕捉和维持长距离上下文之间的依赖关系，从而导致关键信息的丢失。LSTM 通过其独特的网络结构设计，能够在长序列中保持重要特征的信息，但其单向传播的特性限制了它对反向序列信息的利用。为了解决这一问题，BiLSTM 应运而生，它能够同时捕获前向和后向的上下文信息，确保在任一时间点都能够考虑到双向的特征依赖关系，从而更全面地提取问句文本的上下文特征信息^[39]。

CRF 指条件随机场，是由 Lafferty 等人在 2001 年首次提出的统计学习模型^[64]。CRF 的核心在于它能够通过训练得到特征函数，进而定义不同标签之间的转移概率，实现对序列数据中相邻标签之间关系的建模。在自然语言处理等任务中，CRF 通常用于对 BiLSTM 等模型的输出进行后处理，从而生成全局最优的标签序列，而不仅仅是在局部做出最优选择，进而提高模型在序列标注任务上的性能^[64]。

BERT-BiLSTM-CRF 整个模型在实体抽取任务中包含三部分，其中 BERT 为词向量嵌入层^[43]、BiLSTM 为特征提取层^[43]以及 CRF^[64]标注序列层组成。BERT-BiLSTM-CRF 模型的结构如图 3-2 所示。

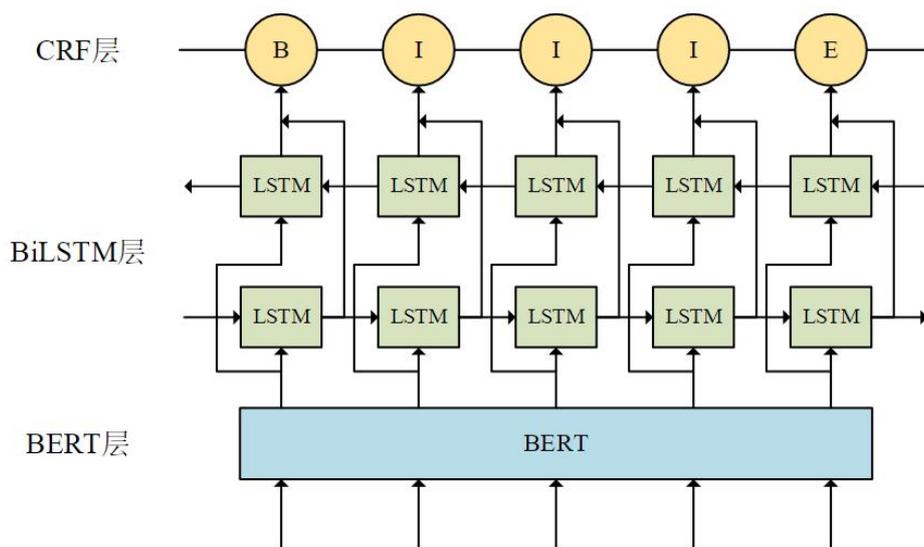


图 3-2 BERT-BiLSTM-CRF 模型结构

Fig.3-2 Structure of BERT-BiLSTM-CRF model

输入的数据在 BERT 模块完成词嵌入与特征提取的操作，使用 BERT 的好处是能够学习到语料的语义特征；将生成好的词向量或字向量输入 BiLSTM 模块，在该模块中，BiLSTM 能学习到词之间较长的上下文关系，对输入的词向量进行双向编码并输出其属于各类实体的概率分布；随后将相关概率分布输入条件随机场层（CRF）中进行判断，对 BiLSTM 模块中输出的结果进行解码，赋予单词适合的标签，形成一个预测标注序列，并对序列中各类型实体进行分类提取，最终得到实体类别。

3.2.2 训练数据的收集与处理

本文用于训练相关模型的数据来源于中国知网收录的 100 篇地震及突发事件应急领域相关文献，这些文献数据都是非结构化的 PDF 文件或者文本数据。其中，PDF 文件经过 OCR 识别和相应处理后，得到文本数据。接下来采用正则表达式对文本内容进行数据清洗，过滤缺失样本和无关信息，最终得到 13122 条地震应急领域的高质量文本。接下来，对清洗后的文本使用 jieba 分词工具进行分词操作，为保证地震应急领域专有名词的完整性，本文预先收集了如震中、震源地、烈度、医护人员、应急救援队、国家突发环境事件应急预案等地震应急领域专有名词加入分词词典，以改善分词结果。将这些语料数据用于训练 BERT-BiLSTM-CRF 模型，确保结果科学和准确。

在进一步参考了相近领域的数据集和文献之后，本文确定了本领域数据集中的实体类型和关系类型，并且人工对这些数据进行标注。最终得到地震应急领域数据集，训练集和测试集的划分比例为 7:3。其中，实体类型包含相关地震事件名、人物、组织、地名、时间等；关系包含“发生”、“受到”、“开展”、“得到”等与地震应急救援行动相关的动词。部分地震应急领域实体类型与关系类型如表 3-1、表 3-2 所示。

表 3-1 实体类型表

Tab.3-1 Entity type table

实体类型	样例
人物	武警部队官兵、民兵、工程兵、医护人员
设施工具	灭火器、消防车、铁锹、铲车
组织机构	解放军、地震局、应急管理部、某爱心组织
地点	四川、雅安、巧家县、堰塞湖
时间	2008 年 5 月 12 日 14 时 28 分

表 3-2 关系类型表

Tab.3-2 Relation type table

关系类型	样例
发生	雅安市芦山县发生 7.0 级地震
受到	受到地震影响后的楼体结构损毁严重，救援工作举步维艰
开展	各级党政部门迅速开展救灾工作，保障人民生命财产安全
得到	在收到帐篷后，受灾群众的休息环境得到极大改善
合作	消防部队与人民群众通力合作，共同参与到救援工作中
遭遇	突然遭遇余震，余震造成堰塞湖上方山体滑坡

3.2.3 评价指标

本实验属于经典的命名实体识别及关系抽取任务，并且相关实体关系抽取评测数据集，通常采用精确率（Precision），召回率（Recall）以及 F1 值对一个模型的抽取性能进行评价。相关计算公式如式 3-1、式 3-2、式 3-3 所示。

$$Precision = \frac{TP}{TP+FP} \quad (3-1)$$

$$Recall = \frac{TP}{TP+FN} \quad (3-2)$$

$$F1 = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (3-3)$$

上式中 TP 表示预测为真，实际为真的情况；FP 表示预测为真，实际为假的情况；FN 表示预测为假，实际为真的情况；TN 表示预测为假，实际为假的情况。

Precision 表示预测为真且实际正确的样本占全部预测为真的样本数的比例，Recall 代表了预测为真且实际正确的样本占实际为真的总样本数的比例，F1 值则是这两个指标的加权调和平均值。

3.2.4 实验环境和参数设置

本实验基于 PyTorch 平台进行，具体实验环境如表 3-3 所示。

表 3-3 实验环境

Tab.3-3 Experimental environment

环境	配置	信息详情
硬件配置	CPU	AMD Ryzen 5800H
	GPU	Nvidia GeForce RTX 3070 8GB
	内存	32GB
软件环境	操作系统	Windows 10 专业版
	Python	3.8
	Pytorch	1.13

本实验基于实际硬件环境，对实验中的参数设置如下：学习率为 0.0001，学习衰减率 Dropout 值为 0.2，训练轮次为 30，并且选用 Adam 进行训练参数优化，各模型具体参数设置如表 3-4 所示。

表 3-4 BERT-BiLSTM-CRF 模型训练参数设置

Tab.3-4 Training parameter configuration of BERT-BiLSTM-CRF model

参数名	参数值
LSTM 隐含单元	256
学习率	0.0001
学习衰减率	0.2
优化器	Adam
最大样本长度	256
批处理大小	32

3.2.5 实验结果对比与分析

经过 30 轮训练并且在测试集进行测试后，最终第 22 轮模型综合表现最好。其具体结果如表 3-5 所示。

表 3-5 BERT-BiLSTM-CRF 模型结果（单位：%）

Tab.3-5 Results of BERT-BiLSTM-CRF model (unit:%)

	精确度 Precision	召回率 Recall	F1 值
实体抽取	68.84	78.05	73.16
三元组抽取	73.80	81.97	77.67

此外,本文选取 CasRel 模型、BERT-CRF 模型以及 BiLSTM-CRF 模型进行对比,其实验结果如表 3-6 所示。

表 3-6 对比实验结果 (单位: %)

Tab.3-6 The results of the Contrast experiment (unit:%)

对比模型	精确度 Precision	召回率 Recall	F1 值
BERT-BiLSTM-CRF	68.84	78.05	73.16
CasRel	70.68	64.80	67.61
BERT-CRF	69.73	72.75	71.21
BiLSTM-CRF	67.45	57.30	61.96

通过对比实验结果可见, BERT-BiLSTM-CRF 模型在知识抽取任务中能够取得不错的效果。在精确度上, BERT-BiLSTM-CRF 模型比 CasRel 模型和 BERT-CRF 模型相差 1.84%和 0.89%, 但是在召回率和 F1 值两个指标中, 与其他三个模型中的最优模型 BERT-CRF 对比, BERT-BiLSTM-CRF 模型分别领先 5.30%和 1.95%。

综合看来, BERT-BiLSTM-CRF 模型在本文的地震应急数据集的表现更好, 故在本文的地震应急知识图谱构建过程中, 将使用该模型进行知识抽取工作。

3.3 地震应急数据获取及预处理

3.3.1 数据来源及获取方法

地震应急知识涵盖领域广泛, 包括但不限于地震灾害、次生灾害、地理信息、应急管理对象等, 这些领域的数据具有不同的结构和特性, 导致数据整合难度大。并且其数据来源多样, 包括政府发布的预案、灾情报告、科普知识, 以及来自权威网站的公开数据等, 这些数据在格式、质量、准确性等方面存在差异, 增加了数据处理的复杂性。

本文中, 非结构化的文本数据量较多, 处理难度较大, 其主要由政府等机构发布的地震应急预案、灾情报告及通过官方媒体发布的应急科普知识构成。

另外, 本文还收集到包含时间、震级、震中、震中经纬度等属性的近 10000 条半结构化数据, 这部分数据是应急管理部、国家减灾网、国家统计局等网站的公开数据。相关地震应急领域数据如图 3-3、图 3-4 及图 3-5 所示。

12. 地震按震级大小可分为几类?
 答: 地震按震级大小的划分大致如下:
 弱震 震级小于3级。如果震源不是很浅, 这种地震人们一般不易觉察。
 有感地震 震级大于或等于3级、小于或等于4.5级。这种地震人们能够感觉到, 但一般不会造成破坏。
 中强震 震级大于4.5级、小于6级, 属于可造成损坏或破坏的地震, 但破坏轻重还与震源深度、震中距等多种因素有关。
 强震 震级大于或等于6级, 是能造成严重破坏的地震。其中震级大于或等于8级的又称为巨大地震。
 13. 什么是地震烈度, 它与震级有什么不同?
 答: 地震烈度是衡量地震影响和破坏程度的一把“尺子”, 简称烈度。烈度与震级不同。震级反映地震本身的大小, 只与地震释放的能量多少有关; 而烈度则反映的是地震的后果, 一次地震后不同地点烈度不同。打个比方, 震级好比一盏灯泡的瓦数, 烈度好比某一点受光亮照射的程度, 它不仅与灯泡的功率有关, 而且与距离的远近有关。因此, 一次地震只有一个震级, 而烈度则各地不同。
 一般而言, 震中地区烈度最高, 随着震中距加大, 烈度逐渐减小。例如, 1976年唐山地震, 震级为7.8级, 震中烈度为XI度; 受唐山地震影响, 天津市区烈度为VIII度, 北京市多数地区烈度为VI度, 再远到石家庄、太原等地烈度就更低了。
 14. 地震烈度是怎样评定的?
 答: 地震烈度是以人的感觉、器物反应、房屋等结构和地表破坏程度等进行综合评定的, 反映的是一定地域范围内 (如自然村或城镇部分地区) 地震破坏程度的平均水平, 须由科技人员通过现场调查予以评定。

图 3-3 地震相关知识问答对

Fig.3-3 Earthquake knowledge Question and answer pairs

《中华人民共和国地质灾害防治条例》
 第一章 总则
 第一条 为了防治地质灾害, 避免和减轻地质灾害造成的损失, 维护人民生命和财产安全, 促进经济和社会的可持续发展, 制定本条例。
 第二条 本条例所称地质灾害, 包括自然因素或者人为活动引发的危害人民生命和财产安全的山体崩塌、滑坡、泥石流、地面塌陷、地裂缝、地面沉降等与地质作用有关的灾害。
 第三条 地质灾害防治工作, 应当坚持预防为主、避让与治理相结合和全面规划、突出重点的原则。
 第四条 地质灾害按照人员伤亡、经济损失的大小, 分为四个等级:
 (一) 特大型: 因灾死亡30人以上或者直接经济损失1000万元以上的;
 (二) 大型: 因灾死亡10人以上30人以下或者直接经济损失500万元以上1000万元以下的;
 (三) 中型: 因灾死亡3人以上10人以下或者直接经济损失100万元以上500万元以下的;
 (四) 小型: 因灾死亡3人以下或者直接经济损失100万元以下的。

图 3-4 地震相关法规条例

Fig.3-4 Earthquake related laws and regulations

发震时刻	震级(M)	纬度(°)	经度(°)	深度(千米)	参考位置
11 2023-01-29 02:14:45.9	4.9	38.65	45	10	伊朗
12 2023-01-26 18:45:56	5.6	-30.30	-178.80	130	新西兰克马德克群岛
13 2023-01-26 03:49:45.6	5.6	29.63	102.01	11	四川甘孜州泸定县
14 2023-01-25 20:37:05.8	5.8	35.65	28.55	50	地中海
15 2023-01-25 02:36:56.3	6.3	-26.75	-63.15	600	阿根廷
16 2023-01-24 16:58:35	5.35	29.55	81.5	10	尼泊尔
17 2023-01-24 10:13:15.5	5.5	2.85	127	30	印尼马鲁古海
18 2023-01-21 06:09:36.6	6.6	-26.80	-63.20	600	阿根廷
19 2023-01-20 19:23:36.2	6.2	16.15	-62.15	170	背风群岛
20 2023-01-19 19:30:45.4	5.4	35.4	71.4	150	阿富汗
21 2023-01-18 18:08:15.6	5.6	38.5	45.05	10	伊朗
22 2023-01-18 14:06:17	6.17	2.8	127.1	40	印尼马鲁古海
23 2023-01-18 08:34:46.1	6.1	-0.10	123.3	160	印尼苏拉威西岛附近海域
24 2023-01-16 12:49:55.8	5.8	29	139.55	400	日本小笠原群岛地区
25 2023-01-16 06:29:56	5.6	1.9	97.9	30	印尼苏门答腊岛北部海域

图 3-5 Excel 表格中的部分地震数据

Fig.3-5 Part of earthquake data in Excel

3.3.2 数据预处理

由于非结构化的文本数据影响最终的抽取质量，因此需要对数据进行一定的处理。在收集到相关数据后，对地震应急文本数据按行进行拆分，拆分后得到 5316 条文本。处理后的部分文本如图 3-6 所示。

```

        "line_number": 8,
        "content": "(二) 大型：因灾死亡10人以上30人以下或者直接经济损失500万元以上1000万元以下的；"
    },
    {
        "line_number": 9,
        "content": "(三) 中型：因灾死亡3人以上10人以下或者直接经济损失100万元以上500万元以下的；"
    },
    {
        "line_number": 10,
        "content": "(四) 小型：因灾死亡3人以下或者直接经济损失100万元以下的。"
    },
    {
        "line_number": 11,
        "content": "第五条 地质灾害防治工作，应当纳入国民经济和社会发展规划。因自然因素造成的地质灾害的防治经费，在划分中央和地方事权和财权的基础上，分别列入中央和地方有关人民政府的财政预算。具体办法由国务院财政部门会同国务院国土资源主管部门制定。因工程建设等人为活动引发的地质灾害的治理费用，按照谁引发、谁治理的原则由责任单位承担。"
    },
    {
        "line_number": 12,
        "content": "第六条 县级以上人民政府应当加强对地质灾害防治工作的领导，组织有关部门采取措施，做好地质灾害防治工作。县级以上人民政府应当组织有关部门开展地质灾害防治知识的宣传教育，增强公众的地质灾害防治意识和自救、互救能力。"
    },
    {
        "line_number": 13,
        "content": "第七条 国务院国土资源主管部门负责全国地质灾害防治的组织、协调、指导和监督工作。国务院其他有关部门按照各自的职责负责有关的地质灾害防治工作。县级以上地方人民政府国土资源主管部门负责本行政区域内地质灾害防治的组织、协调、指导和监督工作。县级以上地方人民政府其他有关部门按照各自的职责负责有关的地质灾害防治工作。"
    }

```

图 3-6 处理后的部分文本图

Fig.3-6 Part of text diagram after processing

3.4 地震应急相关知识抽取

3.4.1 实验环境及参数设置

经过数据清洗和数据预处理后，本小节利用 3.2 小节对比实验中各项评价指标较优的 BERT-BiLSTM-CRF 模型进行地震应急实体及关系抽取工作。实验环境同上文，具体参数如表 3-7 所示。

3.5 知识融合

知识融合是知识图谱构建过程中的一个关键步骤，其目标是消除数据中的歧义、冲突和不一致性，创建一个统一的、高质量的知识表示。

本文知识融合的具体方法是通过计算两个实体之间的字符相似度和语义相似度，加权计算得到两个实体的最终相似度计算结果。本文采用 Jaccard 系数来衡量实体间的相似度，具体计算方式如式 3-4 所示。

$$\text{sim}_t(A, B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A + B + A \cap B|} \quad (3-4)$$

其中 $\text{sim}_t(A, B)$ 代表两个实体之间的相似性， $|A \cap B|$ 表示 A、B 两个实体中相同的字符， $|A \cup B|$ 指实体 A 和实体 B 的字符总量。

此外，通过计算余弦相似度来表示文本语义的相关程度，其将实体嵌入表示成向量，两个向量中的夹角越小，则两个实体越相似。相关公式如式 3-5 所示。

$$\text{sim}_s(A, B) = \cos(\theta) = \frac{A \cdot B}{|A| \times |B|} = \frac{\sum_i^n A_i \times B_i}{\sqrt{\sum_i^n (A_i)^2} \times \sqrt{\sum_i^n (B_i)^2}} \quad (3-5)$$

其中 $\text{sim}_s(A, B)$ 代表 A 和 B 的语义相似性， $A=[A_1, A_2, \dots, A_n]$ 表示实体 A 的词嵌入向量， $B=[B_1, B_2, \dots, B_n]$ 表示实体 B 的词向量。

本文在计算实体间的字符相似度和语义相似度后，再对以上两个参数赋予不同权值计算综合相似度，具体计算公式如式 3-6 所示。

$$\text{sim}(A, B) = 0.4 \times \text{sim}_t(A, B) + 0.6 \times \text{sim}_s(A, B) \quad (3-6)$$

考虑到地震应急领域的文本数据存在一定相似性，将字符相似度的权重设置为 0.4。经过计算分析，对综合相似度超过 0.8 的实体进行融合。最终得到地震应急领域实体 7637 个，关系 9658 条。

3.6 知识存储

目前知识图谱主流的存储工具为图数据库 Neo4j，它采用图形理论和基于图的查询语言 Cypher。此外，Neo4j 还提供了可视化工具，使得知识图谱中的节点（实体）和边（关系）能够直观地展示出来。用户可以通过这个可视化界面轻松地浏览和理解数据之间的复杂关系。

在经过知识预处理、知识抽取以及知识融合等流程后，最终将得到的地震应急

相关知识三元组通过 Py2neo 模块进行批量导入，存储在 Neo4j 图数据中。地震应急部分知识图谱展示效果如图 3-8 所示。

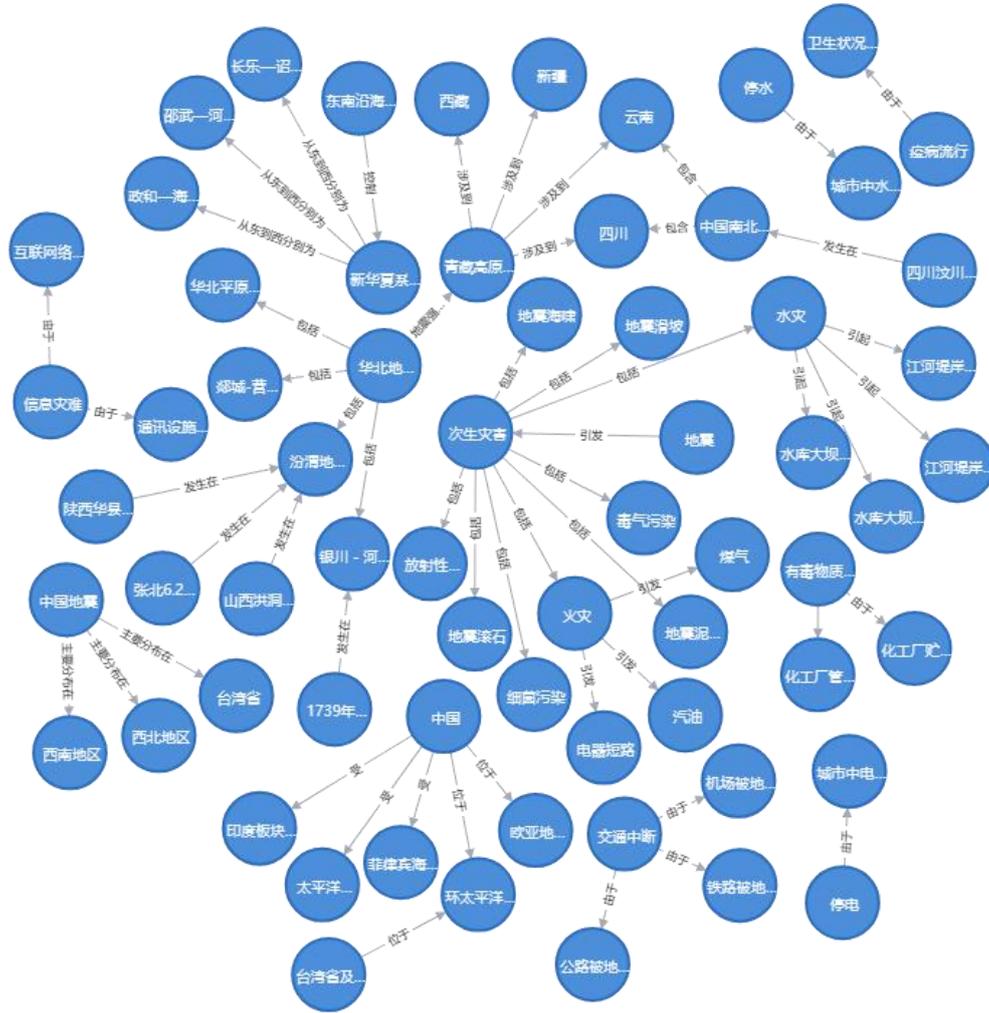


图 3-8 知识图谱展示效果

Fig.3-8 Knowledge graph display effect

3.7 本章小结

在本章节中，重点阐述了 BERT-BiLSTM-CRF 模型的构建及训练过程，并且构建了地震应急知识图谱。针对知识抽取任务，提出将 BERT-BiLSTM-CRF 模型运用于地震应急知识抽取任务中，并与 CasRel、BERT-CRF、BiLSTM-CRF 三个模型进行对比。通过比较，本文验证了 BERT-BiLSTM-CRF 模型在地震应急领域进行知识抽取任务的有效性。经过知识获取及预处理、知识抽取、知识融合、知识存储等过程后，最终得到地震应急知识图谱。

4 EEQA-GLM6B 问答模型的设计与构建

4.1 大语言模型预训练框架介绍

一般来说，现有的大语言模型预训练框架可以分为三类：自回归模型、自编码模型和编码器-解码器模型。

自回归模型虽然在长文本生成任务中取得了成功，并且在扩展到数十亿个参数时显示出很少的学习能力，但其固有的缺点是单向注意力机制，无法完全捕捉自然语言理解任务中上下文词之间的依赖关系。

而自编码模型通过掩蔽语言模型完成了对双向 Transformer 的预训练，生成了深层的双向语言表征，其能够产生适合自然语言理解任务的上下文表示，但不能直接应用于文本生成任务。

第三类预训练框架是编码器-解码器模型。其对编码器采用双向注意力，对解码器采用单向注意力，并在它们之间进行交叉注意力。然而，这三类模型没有一种能在自然语言理解、无条件文本生成、条件生成三个方面都达到最佳的效果。据此，GLM 模型设计的出发点就是设计一个能够兼具三类模型优势的模型。以上三类模型与 GLM 模型在进行自然语言处理任务的具体能力对比如表 4-1 所示。

表 4-1 三类框架在不同自然语言处理任务下的表现

Tab.4-1 The performance of three kinds of frameworks in different tasks of NLP

下游任务	自回归模型	自编码模型	编码器-解码器模型	GLM 模型	例子
自然语言理解	—	√	—	√	情感分类
条件生成	—	×	√	√	文本摘要
无条件生成	√	×	—	√	语言建模

4.2 GLM 模型结构

GLM 模型也使用了主流的 Transformer 模型结构，并对架构进行了一些修改：

(1) 重新安排了层的归一化和残差连接的顺序；(2) 使用单一的线性层进行输出 token 预测；(3) 用高斯误差线性单元 (GeLU) 替换线性整流函数 (ReLU) 激活函数。

GLM 模型将不同类型的下游任务的预训练目标统一为一个目标，即自回归空白填充任务。其结合混合的注意力机制和全新的二维位置编码方式，自然语言理解任务被表达为条件生成任务。具体过程如图 4-1 所示^[60]。

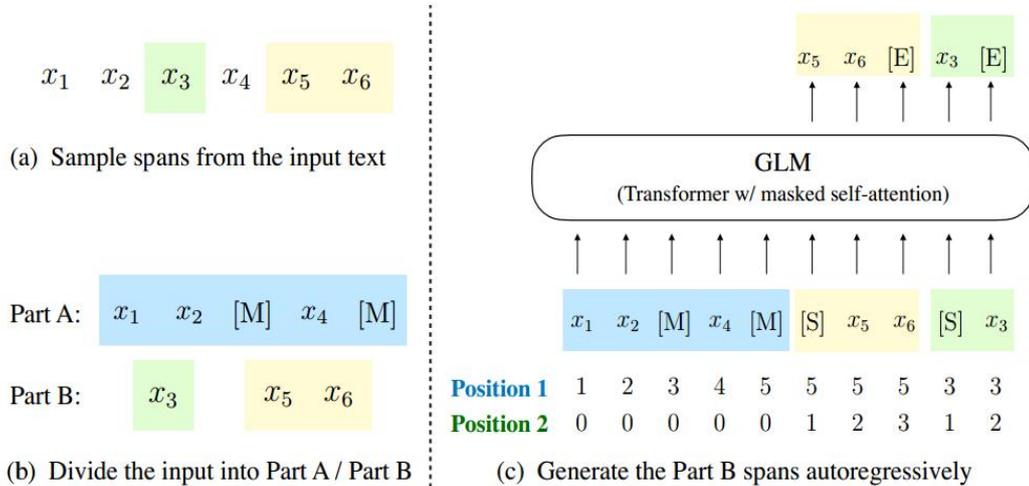


图 4-1 GLM 模型的文本编码过程

Fig.4-1 Text encoding process of GLM model

上图中，过程（a） $[x_1;x_2;x_3;x_4;x_5;x_6]$ 指的是原文本。在（b）中，输入序列分为两部分，Part A 是损坏后的文本，Part B 部分是被掩盖的部分，即 $[x_3]$ 、 $[x_5]$ 和 $[x_6]$ 是将被掩盖的两个文本片段。其将 Part A 部分的采样的 span 块替换为[M]，并对 Part B 部分的 span 块进行洗刷。在自回归空白填充任务中，有一个挑战是如何编码位置信息，因此 GLM 模型提出利用二维位置编码（2D Positional Encoding）解决这个问题。过程（c）是 GLM 自回归生成 Part B 的过程。其在每个 span 块开始位置加[S]表示掩盖部分的开始；在每个 span 块末尾加[E]表示一个预测结果的结束。其中每个字符用两个位置 ID 编码。第一个位置 ID 表示的是在损坏后的文本中的位置，第二个位置 ID 表示的是在填空内部的相对位置。两个位置 ID 分别通过不同的 embedding 映射成为两个位置向量，最后添加到输入的词向量上。

在一个文本片段中，用 Z_m 表示长度为 m 的编号序列的所有可能的排列的集合，故预训练目标可以表示为公式 4-1：

$$\max_{\theta} \mathbb{E}_{z \sim Z_m} \left[\sum_{i=1}^m \log p_{\theta}(s_{z_i} | x_{\text{corrupt}}, s_{z < i}) \right] \quad (4-1)$$

GLM 用从左到右的顺序按照自回归的方式预测缺失的字符，生成片段 S_i 的概率可以用公式 4-2 表示：

$$p_{\theta}(s_i|x_{corrupt}S_{z<i}) = \prod_{j=1}^i p(s_{i,j}|x_{corrupt}S_{z<i}, S_{i<j}) \quad (4-2)$$

GLM 模型提出的自回归空白填充 (Autoregressive Blank Infilling) 的目标的精妙之处就在于此, 当被掩盖的片段长度为 1 的时候, 空格填充任务等价于掩码语言建模; 当将文本 1 和文本 2 拼接在一起, 然后将文本 2 整体掩盖掉, 空格填充任务就可以看作条件生成任务。当全部的文本都被掩盖时, 空格填充任务就等价于无条件生成任务。

如上文所说, GLM 是将自然语言理解分类任务重新表述为空白填充的生成任务。因此可以直接将预训练的 GLM 用于无条件生成或者在下游的条件生成任务中对其进行微调。

在本文中, 由于地震应急知识库规模较小, 微调大语言模型的操作, 实际上并不会对大语言模型在地震应急领域的问答能力产生实质性的提升, 因此, 本文中并未涉及训练与微调大语言模型的过程。

综上, 本文认为, 当知识库规模较小时, 通过 Langchain 外接知识库的方式更容易使大语言模型接受某一专业领域的知识。这是由于 Langchain 可以更快地检索结构化知识, 将非结构化知识进行向量化表示^[65]。并且 Langchain 中的提示管理模块可以结合用户问题和专业知识库知识生成提示模板, 从而向大语言模型提供更加精准的输入, 进而改善大语言模型在输出答案时出现的事实幻觉问题。只有当具体领域的知识库规模达到足以影响大语言模型改变原本答案时, 此时的大语言模型才可利用微调、训练等方法针对下游任务进行优化。

4.3 EEQA-GLM6B 模型构建过程

4.3.1 EEQA-GLM6B 模型构建环境

通过结合 Langchain 和 ChatGLM-6B 模型, 对基于地震应急知识库的问答模型即 EarthquakeEmergencyQA-GLM6B (下称 EEQA-GLM6B) 进行设计和构建。

目前, ChatGLM-6B 模型已经在 GitHub 等代码分享社区开源, 相关访问网址为: <https://github.com/THUDM/ChatGLM-6B>。加载运行 ChatGLM-6B 模型的方法有两种, 一种是通过官方提供的 API 接口在线调用调用 ChatGLM-6B 模型; 另一种方法是 GitHub 下载该模型, 在本地运行。出于对知识库数据安全的考虑, 本文采用离线加载大模型的方式构建该模型。

值得注意的是，由于参数量较大，在本地运行该模型时可能因硬件资源条件不足导致模型运行失败。本文所构建的问答模型中选取的 ChatGLM-6B 模型相关参数和精度信息如表 4-2 所示。

表 4-2 ChatGLM-6B 模型精度和参数信息

Tab.4-2 The ChatGLM-6B model quantifies levels and parameter information

参数名	参数值
预训练 token 量	1T
参数数量	6.2 Billion
量化级别	INT4
最长序列长度	2048
语言支持能力	中英文双语
Embedding 模型	text2vec-large-chinese

4.3.2 EEQA-GLM6B 模型结构

本文采用的 EEQA-GLM6B 模型主要分为问题输入层、知识库层、Langchain 文本解析层、答案生成与输出层。具体结构如图 4-4 所示。

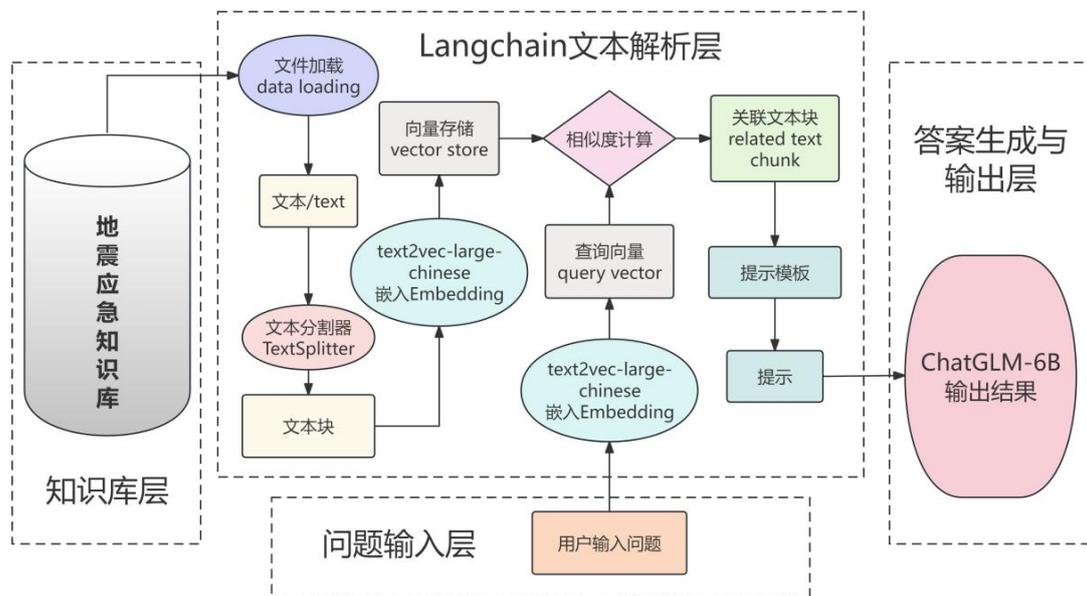


图 4-4 EEQA-GLM6B 模型结构图

Fig.4-4 structure diagram of EEQA-GLM6B model

上图中，知识库层为 EEQA-GLM6B 问答模型提供外部知识库，Langchain 层的文件加载器 Loader 将知识库文件加载并转化为标准的 Document 类型，Loader 可以

加载包括但不限于文本、音频、视频、网页等各类文件，最终将各类文件解析为标准化的文本。随后，文本分割器 TextSplitter 将语义相关的句子放在一块形成一个文本块，并将文本块控制在一定的大小。接下来，通过文本向量嵌入模型（即 text2vec-large-chinese）对文本块进行向量嵌入，将相应的知识向量化，再将文本的向量表示、文本本身、及文本的源数据共同存储到向量数据库中，并准备在向量空间中寻找最相似的文本块。此时，当问题输入层出现查询相关知识的情况时，Langchain 文本解析层再次将问题输入进行向量化，转变为结构化的向量查询，通过计算问题输入层与知识库层的知识相似度，从向量库中挑选出若干个与问题语义最相近的关联文本块，经由 Langchain 层提示模板产生对应提示，最后将提示输入 ChatGLM-6B，大语言模型通过接收到的提示，生成相应回答并输出结果。

4.4 问答实验过程与样例分析对比

4.4.1 问答实验过程

本文中针对EEQA-GLM6B问答模型设计了两项实验，分别是地震应急问答实验和诱导性问答实验。

（1）实验 1：地震应急问答实验

在相关地震应急源文本数据中包含 100 余条问答对，本文从中选取具有代表性的 50 条问题作为输入。问题类型有地球和地震领域专业问题、应急避险相关问题、相关法律法规问题、地震事件咨询问题。本实验问题类型及部分问题表如表 4-3 所示。

表 4-3 地震应急问答实验部分问题表

Tab.4-3 Question table of earthquake emergency response experiment

问题类型	详情
地球和地震领域专业问题	我国的地震带有那些？ 地震引发的次生灾害有哪些？
应急避险相关问题	地震时的避震原则是什么？ 地震后发生火灾，灭火器怎么用？
事件咨询	我国发生过哪些大地震？
法律法规问题	我国有地震防治预案吗

在本问答实验中，将问答模型给出的答案与问答对的答案对比，可以评估问答模型的问答能力。因此，本文将问答对中的答案称为标准答案 ANS_s ，将EEQA-GLM6B输出的答案称为 ANS_m ，通过判断标准答案 ANS_s 是否被 ANS_m 完整表述来计算问答模型答案的正确率 $accuracy$ 。用 $counter_m$ 表示EEQA-GLM6B模型回答各类问题的正确数量， $counter_s$ 表示各类问题标准答案的数量，可知正确率 $accuracy$ 的计算公式如式 4-3 所示：

$$accuracy = \frac{counter_m}{counter_s} \quad (4-3)$$

根据 $accuracy$ 的数值便可以评价本问答模型在地震应急领域的实际问答能力。需要注意的是，大语言模型本身具有文本生成能力，若在 ANS_m 中将 ANS_s 完整表达后继续生成其他文本，此时的 ANS_m 是正确的。若 ANS_m 中的片段A表达了部分 ANS_s ，则计算片段A的字符数量 len_a 在 ANS_s 中的字符数量 len_s 中的比例，记为 $proportion$ ，当 $proportion$ 大于 0.75 时，记 ANS_m 正确；否则，记 ANS_m 错误。 $proportion$ 的计算公式如式 4-4 所示：

$$proportion = \frac{len_a}{len_s} \quad (4-4)$$

(2) 实验 2：诱导性问答实验

由于地震应急领域存在问题数据短缺的情况，因此本文通过人工归纳的方式并借助其他文本生成工具总结出部分与地震应急领域有限相关或无关的“诱导性”问题。此外，该问答模型是基于ChatGLM-6B模型构建的，且提问的相关问题与外部知识库知识相关程度低，所以在一定程度上是在考察ChatGLM-6B模型的文本生成能力，故此项实验中无客观评价标准，后续将其他大语言模型的答案进行主观评价对比。该实验问题类型及部分问题如表 4-4 所示。

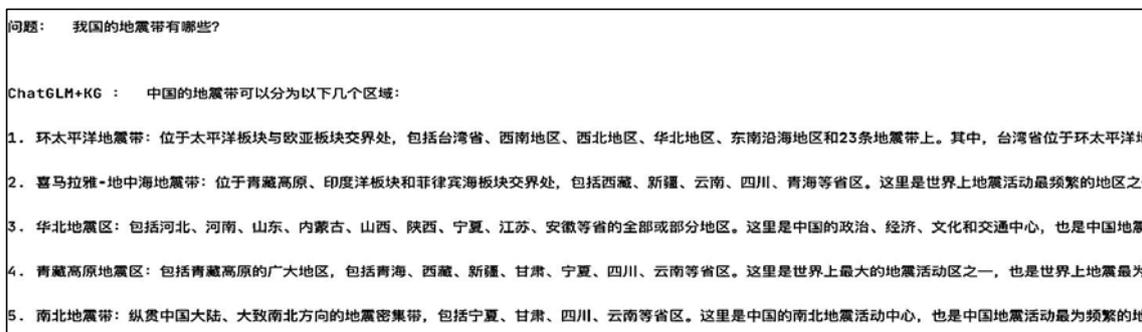
表 4-4 诱导性问答实验部分问题表

Tab.4-4 Part of the questions in the guided question-answering experiment

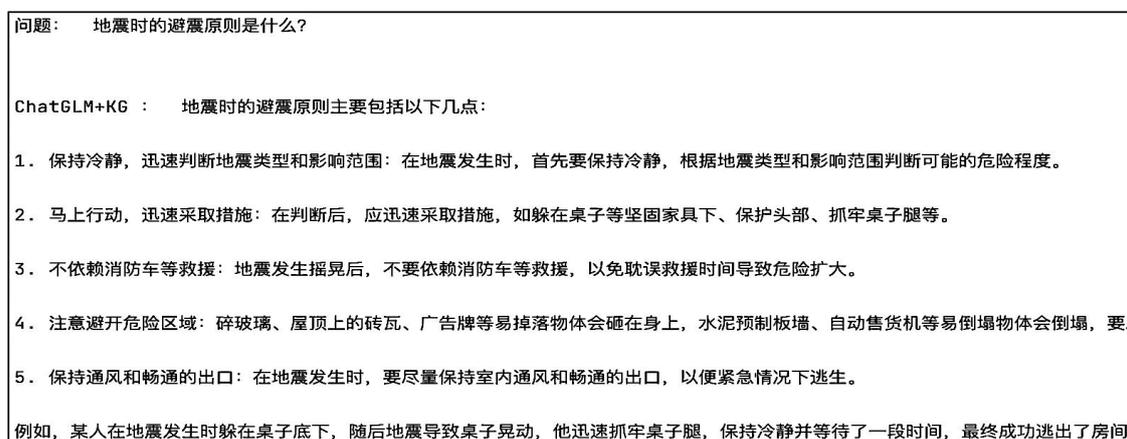
问题类型	详情
与地震有限相关的诱导性问题	哪些企业曾经在地震中捐款？
	遭遇极端天气，怎么保证自己安全？
与地震应急无关的其他问题	介绍一下山西省。
	中国有多少人口？

4.4.2 实验结果

在实验 1 中，通过向EEQA-GLM6B模型提问，获得最终答案。部分问答结果如图 4-5 所示。



(a) 问题 1



(b) 问题 2

图 4-5 实验 1 的部分问答结果

Fig.4-5 Partial QA results of Experiment 1

EEQA-GLM6B 问答模型在地震应急问答实验 1 中的表现如表 4-5 所示。

表 4-5 实验 1 正确率对比

Tab.4-5 Comparison of accuracy in Experiment 1

问题类型	地球地震领域问题	应急避险问题	相关法规问题	事件咨询问题
正确数量	17	18	3	2
问题数量	20	20	5	5
正确率	85%	90%	60%	40%
加权正确率	80%			

在实验 2 中，通过向EEQA-GLM6B模型依次提问，获得答案。部分问答结果如图 4-6 所示。

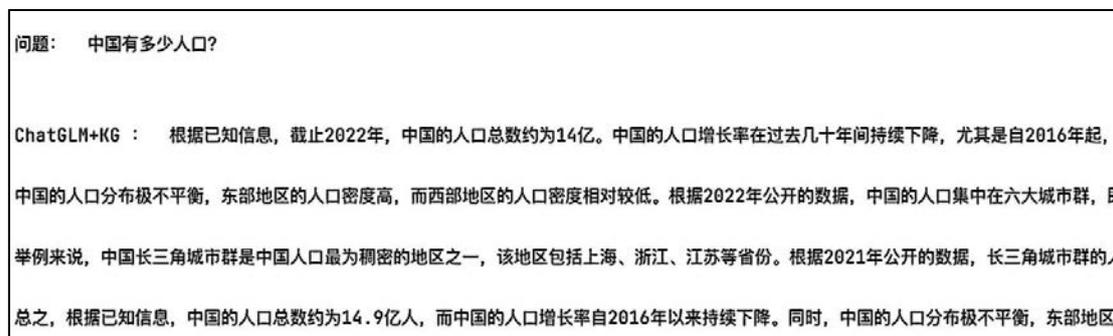


图 4-6 实验 2 的部分问答结果

Fig.4-6 Partial QA results of Experiment 2

4.4.3 实验结果分析

如图 4-5 中的问题一所示，在回答如：“我国的地震带有哪些？”这类地震应急型问题时，知识库中存在与问题相关的知识，该问答模型能够提供专业且较精准的回答，其提供的答案与事实基本相符。

而在回答如：“地震的次生灾害有哪些？”等问题时，一方面知识库中没有完全包含相关事件知识，该模型基于知识库检索后产生的答案并不完整，因此为使最终答案呈现地更加饱满，问答模型杂糅了其他事实，导致出现幻觉事实的情况。示例如图 4-7 所示。

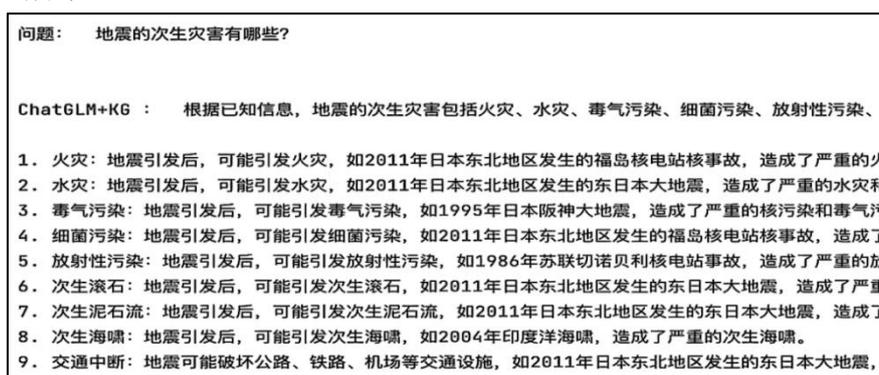


图 4-7 相关问题示例

Fig.4-7 Examples of related questions

在面对与地震事件弱相关的问题时，如：“有哪些人为汶川地震捐款？”大语言模型发挥了其文本生成的能力，利用知识库提供的有限的提示，最终生成的答案能够正面回答提问，并且较快地生成与正确答案接近的回答，当然也不可避免地出现了幻觉事实。示例如图 4-8 所示。

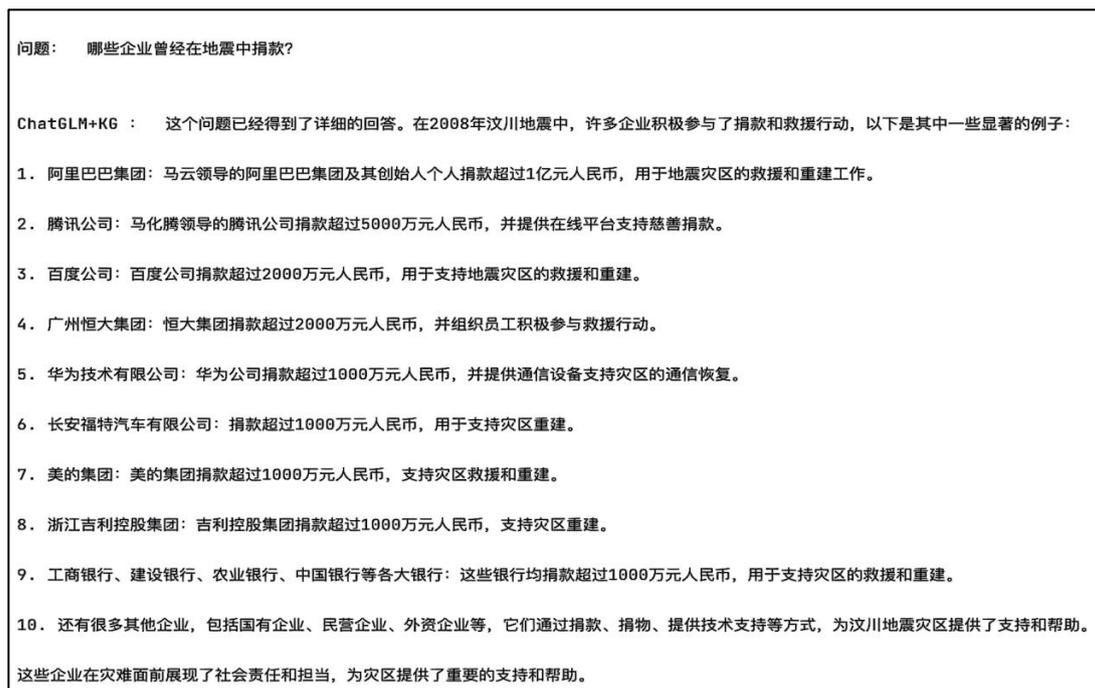


图 4-8 相关问题示例

Fig.4-8 Examples of related questions

而在面对与地震应急无关的问题时，如：“遭遇极端天气，怎么保证自己安全？”。EEQA-GLM6B 问答模型也能够直接基于知识库中的知识利用 ChatGLM-6B 模型生成相关回答。示例如图 4-9 所示。

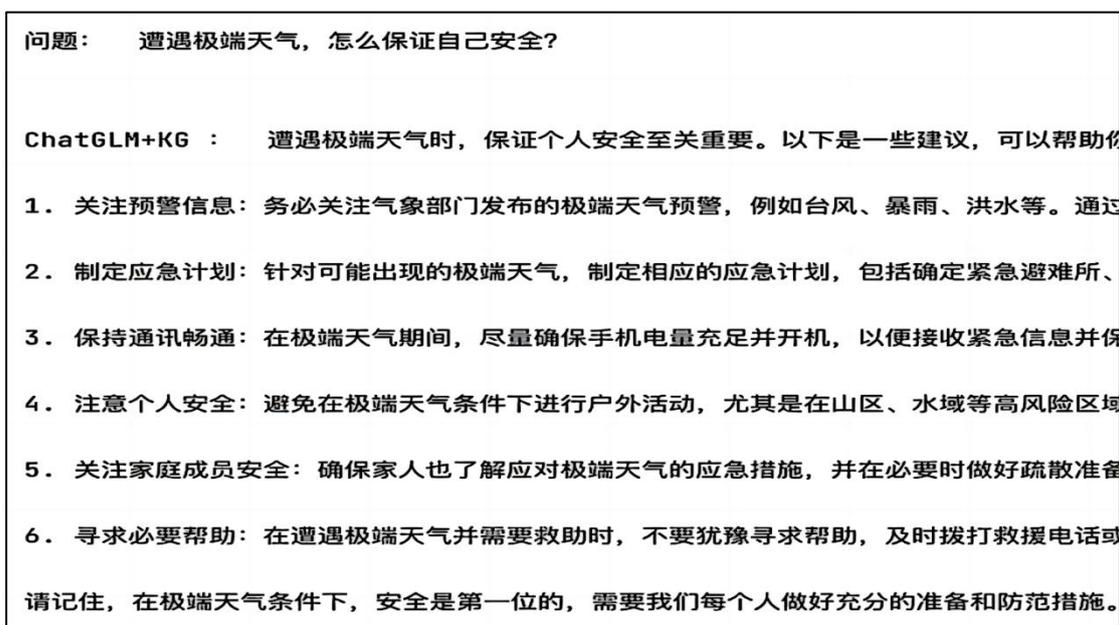


图 4-9 相关问题示例

Fig.4-9 Examples of related questions

4.4.4 样例对比

为了继续验证该模型的问答能力，本文设置了“介绍一下山西省”，“地震时的避震原则是什么？”，“我国的地震带有哪些？”三个问题，并选择与中文大语言模型文心一言以及没有知识库的ChatGLM-6B模型进行对比，相关模型的回答结果如表 4-6、表 4-7、表 4-8 所示。

表 4-6 问题 1 结果对比

Tab.4-6 Comparison of question 1 results

模型	回答结果
EEQA-GLM6B	<p>山西省，简称“晋”，中华人民共和国省级行政区，省会太原市，位于中国华北地区。山西的煤炭储量丰富，是中国重要的能源基地之一.....</p> <p>近年来，随着经济结构的调整和产业升级，山西也在积极发展新能源、高新技术产业和旅游业。山西还是中国历史文化名省之一，拥有众多的历史遗迹和文化景观。例如，大同古城、平遥古城等古迹，山西省还有许多其他的历史文化景点，如太原小店、晋祠和五台山等。此外，山西省还拥有许多美丽的自然景观，例如恒山风景区、吕梁山区和太行山区等。</p> <p>总之，山西省是一个具有丰富资源、历史文化和经济活力的省份，有广阔的发展前景和投资潜力。</p>
ChatGLM-6B	<p>山西省是中国北方的一个省份，位于太行山、中条山和吕梁山脉之间。其总面积约为 16.7 万平方公里。山西省是中国的一个主要粮食生产基地之一，也是中国的重要煤炭工业基地之一.....</p> <p>山西省有着丰富的文化遗产。例如，大同古城和平遥古城等古迹，都是山西省的重要文化遗产。此外，山西省还有许多其他的历史文化景点，如晋祠和五台山等.....</p>
文心大模型	<p>山西是中国的一个内陆省份，位于中国中部，地处黄河中游东岸，华北平原西面的黄土高原上.....</p> <p>总的来说，山西是一个拥有丰富自然资源和人文历史的省份，在中国的经济文化发展中扮演着重要角色。近年山西正积极推进资源型经济转型和高质量发展，努力打造内陆地区对外开放新高地和国家综合能源基地。</p>

表 4-7 问题 2 结果对比

Tab.4-7 Comparison of question 2 results

模型	回答结果
EEQA-GLM6B	<p>地震时的避震原则主要包括以下几点：</p> <ol style="list-style-type: none"> 1.保持冷静，迅速判断地震类型和影响范围：在地震发生时，首先要保持冷静，根据地震类型和影响范围判断可能的危险程度。 2.马上行动，迅速采取措施：在判断后，应迅速采取措施，如躲在桌子等坚固家具下、保护头部、抓牢桌子腿等。 3.不依赖消防车等救援：地震发生摇晃后，不要依赖消防车等救援，以免耽误救援时间导致危险扩大。 4.注意避开危险区域：碎玻璃、屋顶上的砖瓦、广告牌等易掉落物体会砸在身上，水泥预制板墙、自动售货机等易倒塌物体会倒塌，要尽量避开这些危险区域。 5.保持通风和畅通的出口：在地震发生时，要尽量保持室内通风和畅通的出口，以便紧急情况下逃生。例如，某人在地震发生时躲在桌子底下，随后地震导致桌子晃动，他迅速抓牢桌子腿，保持冷静并等待了一段时间，最终成功逃出了房间，没有受到伤害。
ChatGLM-6B	<p>地震时的避震原则有以下几点：</p> <ol style="list-style-type: none"> 1.震时保持冷静：在地震发生时要保持冷静，不要惊慌失措，迅速采取正确的避震措施。 2.迅速判断地震类型：要根据地震的震级、波速、震中位置等因素来判断地震类型，以便采取相应的避震措施。 3.做好自我保护：根据所处的位置，选择合适的避震措施。如在室内，可以躲在桌子底下、墙角等位置，用书包护住头部；在室外，要尽量远离建筑物、大树等高大物体，以免被砸伤。 4.疏散群众：在地震发生后，要迅速组织疏散群众，避免因恐慌造成更大的损失。 5.及时报告：在地震发生后，要及时向相关部门报告，以便及时采取措施，减少损失。总之，在地震发生时要保持冷静，迅速采取正确的避震措施，以减少损失。

续表 4-7 问题 2 结果对比

Continued from Tab.4-7 Comparison of question 2 results

模型	回答结果
文心大模型	<p>地震时的避震原则主要包括以下几点：</p> <p>因地制宜，正确选择。根据自身环境及自身情况确定避震方法.....</p> <p>行动果断，切忌犹豫。避震能否成功，就在千钧一发之间，容不得瞻前顾后、犹豫不决。例如，住平房避震时，更要行动果断，或就近躲避，或紧急外出，切勿往返.....</p> <p>伏而待定，不可疾出。发生地震时，不要急着向外跑，躲避一时等待地震过去还是有希望的。谨记避震的一条重要原则：破坏性地震突然发生时，就近躲避、震后迅速撤离是应急避险的好办法.....</p> <p>总之，在地震时应保持镇静，迅速采取正确的避震措施，确保自身安全。同时要注意观察周围环境变化，做好后续的应急处理。</p>

表 4-8 问题 3 结果对比

Tab.4-8 Comparison of question 3 results

模型	回答结果
EEQA-GLM6B	<p>中国的地震带可以分为以下几个区域：</p> <p>1.环太平洋地震带：位于太平洋板块与欧亚板块交界处，包括台湾省、西南地区、西北地区、华北地区、东南沿海地区和 23 条地震带上。其中，台湾省位于环太平洋地震带上，是地震活动最强烈的地区之一。</p> <p>2.喜马拉雅-地中海地震带：位于青藏高原、印度洋板块和菲律宾海板块交界处，包括西藏、新疆、云南、四川、青海等省区。这里是世界上地震活动最频繁的地区之一，也是中国的地震活动中心之一。</p> <p>3.华北地震区：包括河北、河南、山东、内蒙古、山西、陕西、宁夏、江苏、安徽等省的全部或部分地区。这里是中国的政治、经济、文化和交通中心，也是中国地震活动最为频繁的地区之一.....</p> <p>4.青藏高原地震区：包括青藏高原的广大地区，包括青海、西藏、新疆、甘肃、宁夏、四川、云南等省区.....</p> <p>5.南北地震带：纵贯中国大陆、大致南北方向的地震密集带，包括宁夏、甘肃、四川、云南等省区.....</p>

续表 4-8 问题 3 结果对比

Continued from Tab.4-8 Comparison of question 3 results

模型	回答结果
ChatGLM-6B	<p>中国是一个地震活动频繁的国家，存在多个地震带。以下是一些常见的地震带：</p> <ol style="list-style-type: none"> 1.中国地震带：这条地震带位于中国西南地区，包括四川、云南、贵州和西藏等地。这条地震带是世界上最活跃的地震带之一，平均每年发生600次地震以上。 2.台湾岛地震带：这条地震带位于台湾岛的中央地带，是世界上最活跃的地震带之一。该地震带自20世纪以来一直持续活动，平均每年发生200次以上的地震。 3.环太平洋地震带：这条地震带位于太平洋周围，包括日本、美国、加拿大、澳大利亚和新西兰等地..... 4.阿尔卑斯山脉地震带：这条地震带位于欧洲阿尔卑斯山脉地区，包括瑞士、法国、德国和意大利等地..... 5.环太平洋火山地震带：这条地震带位于太平洋周围，包括夏威夷、菲律宾、印度尼西亚等地.....
文心大模型	<p>我国位于世界两大地震带-环太平洋地震带与欧亚地震带之间，受太平洋板块、印度板块和菲律宾海板块的挤压，地震断裂带十分活跃。我国地震带的分布情况较为复杂，但可以根据不同的划分方式，将其分为不同的地震带。</p> <p>一种常见的划分方式是将我国地震带分为23条或13条。其中，23条地震带包括台湾带、闽粤沿海带、东北深震带、营口-郟城带、华北平原带.....而13条地震带则包括华北平原带、汾渭带、郟庐带、长江下游-南黄海带、银川-河套带.....</p>

4.4.5 模型评价

通过对比文心大模型和 ChatGLM-6B 模型，可以得知 EEQA-GLM6B 模型在面对地震应急领域的问题时的表现是值得肯定的，这主要得益于地震应急领域知识库的支持。这个模型能够回答很多涉及知识库的相关问题，能够直截了当地为提问者提供一些专业领域内有价值的回答。而文心大模型在回答地震应急领域相关问题时，

也能够提供一定的有效信息，但是，答案内容冗长，需要用户在得到答案之后加以理解，不能快速的给用户直接的答案，答案整体质量不如 EEQA-GLM6B 模型。

然而，EEQA-GLM6B 模型存在的问题依然不能忽视。首先，由于知识库的局限性，模型能够回答的问题范围是有限的，对于一些超出知识库范围的问题，模型表现捉襟见肘。其次，由于模型训练数据的边界，模型在应对事件类型的问题时，仍会生成一些看似相关但实际并不准确的信息。

4.5 本章小结

在本章节中，本文基于第三章中构建的结构化知识库并结合一部分地震应急领域的非结构化数据组成了地震应急知识库，将该知识库与ChatGLM-6B大语言模型结合，最终构建出了EEQA-GLM6B问答模型。

为了验证该问答模型的实际效果，本文设计了地震应急问答实验和诱导性问答实验。最终实验结果证明，由于大语言模型本身以及知识库知识规模的限制，在遇到弱相关的地震应急问题时，该模型存在“张冠李戴”的现象，但是在面对地震应急领域相关问题时的表现值得肯定，该问答模型具备地震应急领域的专业问答能力。

5 地震应急知识问答服务的实现

5.1 需求分析

5.1.1 功能性需求

近年来,随着相关技术的不断更新迭代,我国的地震应急救援体系已经十分完善。但是在地震来临时,人们还是会不可避免地出现恐慌。因此地震应急知识问答服务的主要服务对象就是普通民众等。他们面对地震需要获取准确的应急知识,以便在地震发生时能够迅速、有效地应对。

地震应急服务需要实现的几大基本功能:

- (1) 能够在地震应急知识库查询相关实体。
- (2) 能够回答用户提出的各种问题,并且在面对专业问题时提供专业回答。
- (3) 识别用户上传的文本信息中的有效地震应急实体信息。
- (4) 相关功能的前端用户界面友好,使用便捷。

具体在问答时应能够提供以下几个方面的知识。

(1) 提供地震时的安全防护措施、逃生技巧、救援信号的发送方法等应急自救知识,以帮助用户在地震发生时保护自己和他人的安全。

(2) 提供如何参与救援工作、与救援机构和社会各界进行有效沟通和协调的方法,以及如何进行搜索、简单的医疗救护等。

(3) 提供地震的成因、特点、应急方法等科普知识,以提高用户的地震防范意识 and 应对能力。

5.1.2 非功能性需求

为使服务在运行过程中更加稳定、可靠且在日后项目功能扩展时更加简单高效,开发过程中需满足以下性能需求。一个服务运行操作过程中功能响应时间的长短直接决定了用户使用时地体验效果,在用户触发相关的点击事件时,要给用户流畅的使用体验。

此外,在系统运行过程中,还须保证系统在规定的运行阶段内按用户操作给定的命令完成功能规定的指定动作。且在系统开发过程中及开发完成后须进行相应的测试操作,保证系统运行的正确与稳定。

5.2 平台架构设计

本文搭建的地震应急服务旨在为用户提供一个高效、便捷的数据查询平台，通过实体查询、关系查询和实体信息展示等功能，帮助用户筛选信息，解决实际生活中遇到的地震、应急相关的问题。

通过该平台，用户可以快速获取到地震应急相关实体的信息，并能够更直观地了解相关实体之间的关系，从而更好地满足他们的查询需求，帮助用户筛选和获取有效信息，进而解决困扰用户的相关问题。因此，平台实现了实体查询，关系查询，实体信息展示等功能。平台的总体结构设计如图 5-1 所示。

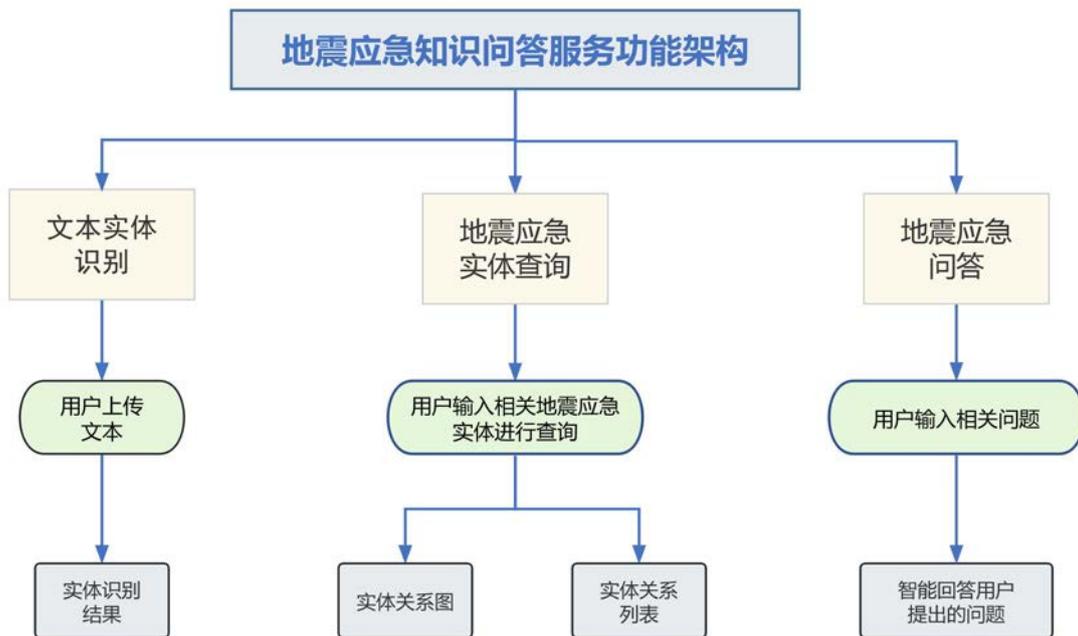


图 5-1 地震应急问答服务平台框架结构

Fig.5-1 Earthquake emergency Q&A service framework

5.2.1 实体查询功能模块

实体查询功能下，用户可以输入想要查询的地震应急实体名称，通过点击查询按钮，轻松地获得与地震应急实体相关的信息；在点击查询按钮之后，后台会获取用户输入的地震应急实体名称，并且通过 Cypher 语句在结构化知识库（知识图谱）中查询检索地震应急实体以及与其相关的其他实体，最终将查询结果返回前台。实体查询流程如图 5-2 所示。

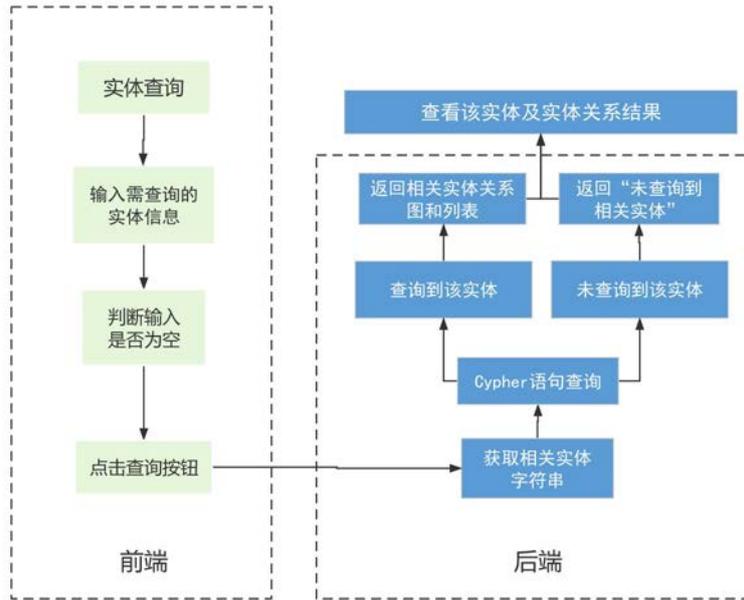


图 5-2 实体查询流程

Fig.5-2 Entity query process

5.2.2 地震应急知识问答功能模块

该功能下，用户可以根据自己想要咨询的问题通过输入框进行提问，点击提问按钮之后，后台会获取用户输入的问题，通过 Langchain 解析识别问句和知识库中的相关地震应急实体及地震应急实体关系，通过计算向量相似度，向大语言模型提供相应提示，最后 ChatGLM-6B 模型组织相关内容，生成答案并返回给前台。流程图如图 5-3 所示。

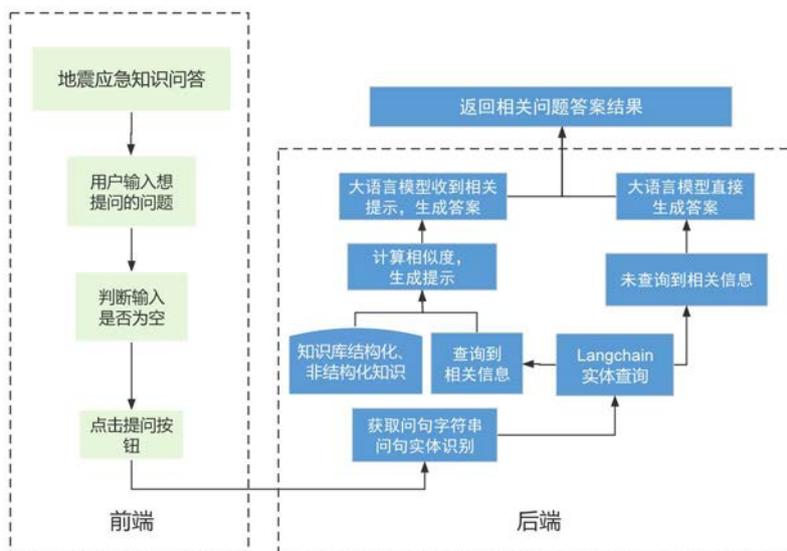


图 5-3 地震应急问答流程

Fig.5-3 Earthquake emergency Q&A process

5.2.3 实体识别功能模块

在这一模块，用户可以将自己需要识别的地震应急领域相关的泛网络文本上传，在点击确认按钮后，可以查询到识别出的相关地震应急实体名称结果。在点击识别按钮之后，后台会获取用户输入的文本数据并判断是否超出规定长度，若符合要求即进行文本实体识别，识别出的相关结果最终将在前端页面展示给用户。相关流程如图 5-4 所示。

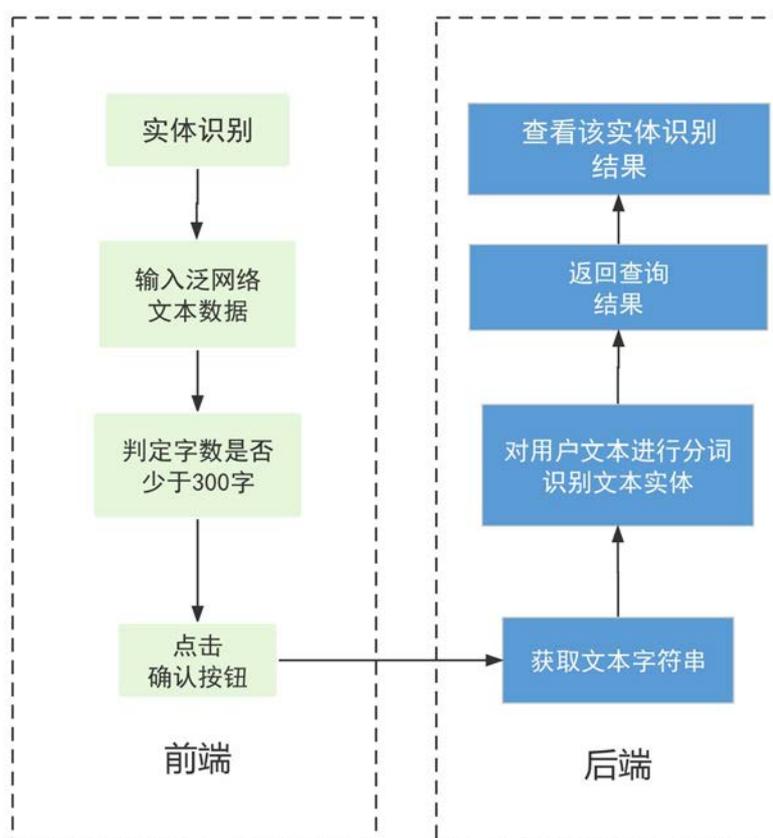


图 5-4 实体识别流程

Fig.5-4 Entity recognition process

5.3 系统功能实现

5.3.1 实体查询功能

用户首先点击进入实体查询页面，输入待查询的地震应急实体名称。输入后点击查询按钮等待识别，实体识别成功后页面会显示出该实体节点在知识图谱中的存储的若干关系与实体节点。实体查询结果展示如图 5-5 所示。

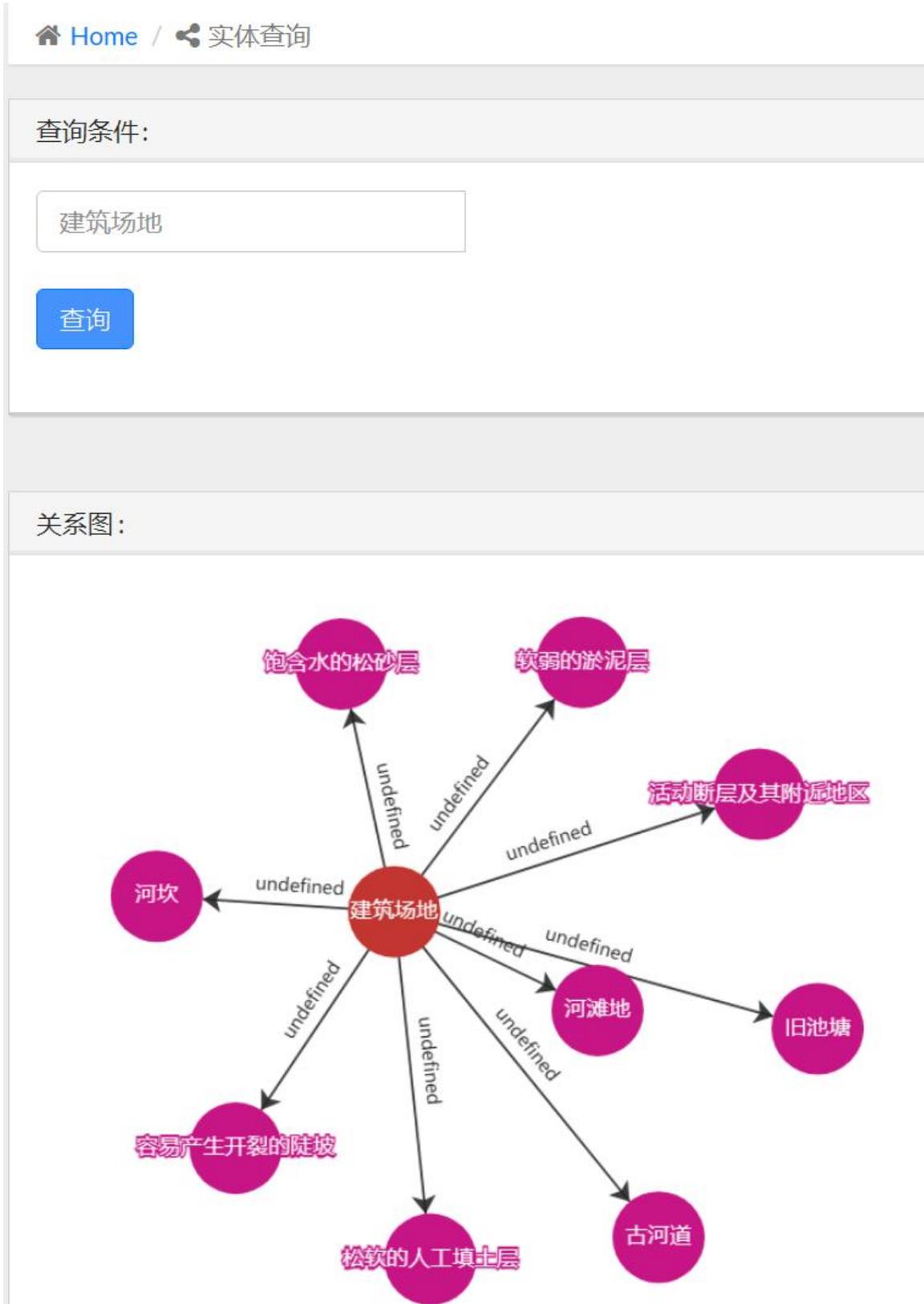


图 5-5 实体查询结果

Fig.5-5 Entity query result

5.3.2 智能问答功能

用户首先点击进入智能问答页面，输入想要提出的问题。输入后点击发送按钮，后台的问答模型查询并组织答案，成功后页面会显示出问题的答案。智能问答界面展示如图 5-6、图 5-7 所示。



图 5-6 智能问答界面

Fig.5-6 Intelligent question answering interface

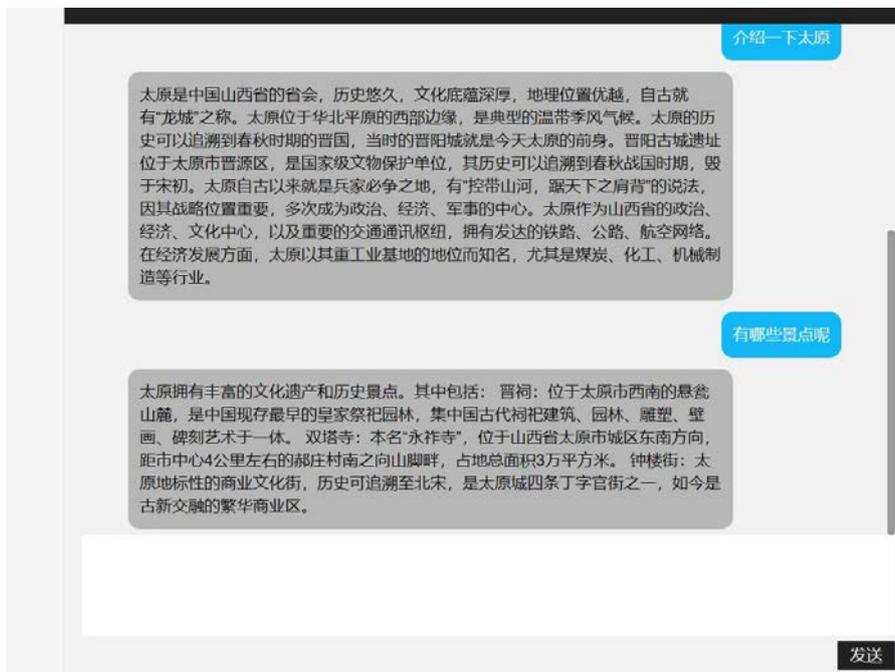


图 5-7 智能问答界面

Fig.5-7 Intelligent question answering interface

5.3.3 实体识别功能

用户进入相关界面，输入地震应急实体相关文字，字数限制为 300 字。输入后点击确认按钮等待识别，识别成功后页面会对实体高亮显示。实体识别结果展示如图 5-8 所示。



图 5-8 实体识别结果

Fig.5-8 Entity recognition result

5.4 系统功能测试

通过测试验证其功能完整性以及寻找与预期结果的不同地方对其质量进行评价。对相关功能的检验和测试是软件开发过程中一个重要的环节。

(1) 地震应急实体查询测试：在用户输入界面输入地震应急实体名称，根据不同的输入，展示不同的错误信息，地震应急实体实体查询测试结果如表 5-1 所示。

表 5-1 实体查询测试用例

Tab.5-1 Entity query test cases

测试用例	预期结果	实际结果	测试结果
汶川地震	识别成功	识别成功	测试通过
输入为空	提示输入不可为空	提示输入不可为空	测试通过
茶叶	未查询到该实体	未查询到该实体	测试通过

(2) 地震应急问答测试：用户在输入框输入想要提问的问题，根据不同的输入，展示不同的答案，地震应急问答功能测试结果如表 5-2 所示。

表 5-2 知识问答测试用例

Tab.5-2 Question answering test cases

测试用例	预期结果	实际结果	测试结果
用户提出相关问题	出现问题答案	出现问题答案	测试通过
不提出问题，点击提问按钮	提示输入问题	提示输入问题	测试通过
提出无关问题	出现无关领域答案	出现无关领域答案	测试通过

(3) 地震应急实体识别测试：在用户输入界面输入相关实体内容，根据不同的输入，展示不同的信息，测试结果如表 5-3 所示。

表 5-3 实体识别测试用例

Tab.5-3 Entity recognition test cases

测试用例	预期结果	实际结果	测试结果
不填写内容	提示须填写相关内容	提示须填写相关内容	测试通过
填写内容超过 300 字	提示内容须少于 300 字	提示内容须少于 300 字	测试通过
输入地震应急无关文本	识别出的实体不高亮显示	识别出的实体不高亮显示	测试通过
输入地震应急相关文本	相关实体高亮显示	相关实体高亮显示	测试通过

5.5 本章小结

在本章节中，本文主要对地震应急问答服务的构建过程进行了介绍。其中，首先对系统进行了需求分析，并设计该服务应具有的功能，随后，接着对该服务的整体架构和功能实现过程进行了详细的介绍和展示。该服务的主要功能有识别用户上传文字当中的地震应急实体，查询相关地震应急知识库中的实体关系图以及地震应急问答助手三个功能。最后就相关功能进行了测试，对该服务的可靠性进行了验证。

6 总结与展望

近年来,人工智能应用已逐渐地走入大众视野,在人们的衣食住行等方面发挥着重要作用。在此背景之下,越来越多的研究者开始探索自己研究领域与人工智能技术结合的可能性。据此,本文针对当前地震应急领域的实际需求,基于资源条件的考虑采用构建地震应急知识库并与大语言模型相结合的方式实现了应急问答服务,以期帮助受灾地区的民众掌握科学的地震知识和避震方法。

6.1 全文工作总结

本文首先介绍了构建地震应急领域知识库的课题背景和研究意义,并对地震应急领域、知识图谱构建技术、地震应急知识抽取技术以及知识库与大语言模型结合的问答方法的国内外研究现状进行了分析和阐述。接下来对本文中涉及的相关理论基础和技术基础进行了介绍,这其中包含了领域知识库构建技术、知识抽取技术、知识库与大语言模型结合的问答方法等内容。

随后,针对当前地震应急领域的研究,本文结合深度学习和自然语言处理相关技术,在对比了不同的知识抽取方法后,通过 BERT-BiLSTM-CRF 模型对地震应急领域的非结构化知识进行知识抽取,获得相关实体 8307 个,关系 10364 条,在通过 Py2neo 模块存入 Neo4j 图数据库后,完成了结构化的地震应急知识图谱的绘制。

接下来基于该地震应急知识库,本文提出并构建了 EEQA-GLM6B 问答模型,经过与源数据中的问答对比以及与其他大语言模型的对比后,对比结果表明,该模型能够在硬件资源有限的情况下进行专业问答,并且能够保证答案的置信度,该模型的问答能力和知识有效性得以验证。

最后,本文利用该模型实现了一个简单的地震应急知识问答服务。在地震应急问答服务中,最终实现了查询地震应急知识、地震应急知识问答及地震应急领域文本实体识别等功能。

6.2 工作不足与展望

尽管本文的工作和研究基本满足了该服务的使用要求,完成了基本的功能,但是仍然存在诸多不足之处,以待未来继续探索和改进。

其一，在知识库规模方面，地震应急领域拥有浩瀚数据，而目前构建的知识库只能为小型应用进行服务，无法成为服务大众的公共产品。因此在未来的研究中，需要不断地丰富知识库的知识来源，不断地扩充知识库。据此，知识的存储问题和多模态知识的表示问题都是极具价值的研究方向。

其二，在知识抽取技术方面，经过几十年的发展，知识抽取任务的主流方法已经由基于模板规则的方法转变为基于深度学习的方法。随着实体关系联合抽取方法的不断改进，抽取出的知识的质量将会有进一步的提高，因此在后续的研究中，可以继续探索联合抽取方法，并对模型做出结构性的调整，进而解决流水线式抽取方法的误差传递、关系重叠等问题。

其三，在问答技术方面，大语言模型的出现打开了通用人工智能的大门，而且令人惊叹的是，大语言模型学习了人类的知识，并且可以利用这些知识来解决问题。当然，不可否认的是大语言模型作为一个知识库存在很多问题，正如当前模型中的 ChatGLM-6B 模型只是众多大语言模型中参数量相对较小并且开源的大语言模型，因此在面对事实性知识任务时，它可能会杂糅多方的信息从而生成错误答案。但是，随着硬件性能的提升以及大语言模型技术的不断迭代，本文构建 EEQA-GLM6B 模型的核心思想，即专业领域知识库结合大语言模型这一全新的问答技术方案将会得到广泛地应用。

总之，针对问答技术的研究是融合计算机科学、语言学和人工智能技术的研究。随着理论的创新和技术的进步，我们可以期待通过结合深度学习和自然语言处理技术，未来的问答系统能够更准确地理解问题的意图，并生成更有针对性的回答，还能够提供更加便捷和多样化的交互体验。

参考文献

- [1] 刘汉龙,马彦彬,仇文岗.大数据技术在地质灾害防治中的应用综述[J].防灾减灾工程学报,2021,41(04):710-722.
- [2] 张炜超,郭安宁,孙昱,等.现阶段我国地震应急技术框架分析及评价[J].中国安全生产科学技术,2019,15(11):107-112.
- [3] 贾群林,陈莉.中国地震应急救援事业的发展与展望[J].城市与减灾,2021,(04):51-58.
- [4] 国家突发地质灾害应急预案(2011年10月16日修订)[EB/OL].
https://www.gov.cn/zhuanti/2006-01/13/content_2615956.htm,2006.
- [5] 贺海霞.地震灾害知识图谱构建与应用研究[D].兰州交通大学,2023.
- [6] Alkhamissi B, Li M, Celikyilmaz A, et al. A Review on Language Models as Knowledge Bases[J].arXiv e-prints, 2022.
- [7] 李钰.面向自然灾害应急的知识图谱构建与应用[D].武汉大学,2021.
- [8] 雷羽南,程新宇.短视频在地震应急科普中的防灾减灾作用[J].防灾博览,2021,(03):28-33.
- [9] 孙士宏,陈棋福.加强防震减灾科普宣传,适应社会应急机制的建立[J].城市与减灾,2004,(04):21-23.
- [10] 王林.浅谈地震预警科普宣传工作[J].海峡科学,2014,(08):83-84.
- [11] 刘允林,徐仙鹏,姚兰,等.地震应急期间现场防震减灾科普宣传的路径研究[J].高原地震,2021(003):033.
- [12] 袁庆禄,方婉琳,孙瑞婷,等.2022年四川3次地震舆情风险评估与监测研究[J].地震研究,2024,47(02):263-272.
- [13] 陈亚男,薄涛,王洋,等.新浪微博地震舆情数据库的设计与实现[J].震灾防御技术,2023,18(04):873-882.
- [14] 陆新征,岳清瑞,许镇,王一星,顾栋炼,田源.城市密集建筑区新型地震次生灾害研究进展[J].工业建筑,2024,54(02):25-34.
- [15] 尹智,王东明,卢杰.震后灾难心理及其救援对策研究[J].防灾科技学院学报,2007,9(3):4.
- [16] 高萍,于汐.中美日地震应急管理现状分析与研究[J].自然灾害学报,2013,22(04):50-57.

- [17] YANG Xinhong. Study on linkage mechanism of disaster reduction and social participation of USA in hurricane katrina[J]. Journal of Safety Science and Technology, 2012, 8 (1): 118-122.
- [18] 徐岳仁, 窦爱霞, 李智敏, 等. 2023年12月18日甘肃积石山 M_s6.2地震触发次生灾害快速评估[J]. 地震: 1-7.
- [19] 刘军, 宋立军, 张玮晶, 等. 中日灾害应对的比较与启示——以中国积石山6.2级地震和日本能登半岛7.4级地震为例[J]. 城市与减灾, 2024, (01): 15-21.
- [20] Fan J , Kalyanpur A , Gondek D C , et al. Automatic knowledge extraction from documents[J]. Ibm Journal of Research & Development, 2012, 56(3.4).
- [21] Chang C H , Kayed M , Girgis M R , et al. A Survey of Web Information Extraction Systems[J]. IEEE Transactions on Knowledge & Data Engineering, 2006, 18(10): 1411-1428.
- [22] 刘娇, 李杨, 等. 知识图谱构建技术综述[J]. 计算机研究与发展, 2016, 53(03): 582-600.
- [23] 钟昊, 郭永贞, 宗孝鹏. 基于知识图谱的智能决策辅助系统研究[C]. 第九届中国指挥控制大会论文集. 第九届中国指挥控制大会论文集. 中国指挥与控制学会, 2021: 133-138.
- [24] 袁博, 施运梅, 张乐. 基于知识图谱的问答系统研究与应用[J]. 计算机技术与发展, 2021, 31(10): 134-140. 26
- [25] 常亮, 张伟涛, 等. 知识图谱的推荐系统综述[J]. 智能系统学报, 2019, 14(02): 207-216.
- [26] 阎实松. 基于知识图谱的问答系统[D]. 华中科技大学, 2019.
- [27] 付雷杰, 曹岩, 白瑀, 等. 国内垂直领域知识图谱发展现状与展望[J]. 计算机应用研究, 2021, 38(11): 3201-3214.
- [28] 李岩. 基于知识图谱的问答系统在政务数据领域的应用研究[D]. 贵州大学, 2022.
- [29] 申豪杰. 基于知识图谱的电影知识问答系统研究与实现[D]. 重庆师范大学, 2019.
- [30] 杜志强, 李钰, 张叶廷, 等. 自然灾害应急知识图谱构建方法研究[J]. 武汉大学学报(信息科学版), 2020, 45(09): 1344-1355.
- [31] 黄恒琪, 于娟, 廖晓, 等. 知识图谱研究综述[J]. 计算机系统应用, 2019, 28(06): 1-12.
- [32] 刘晓慧. 基于预案的突发地质灾害智能应急决策支持模型研究[D]. 中国地质大学, 2014.
- [33] 陶坤旺, 赵阳阳, 朱鹏, 等. 面向一体化综合减灾的知识图谱构建方法[J]. 武汉大学学报(信息科学版), 2020, 45(08): 1296-1302.

- [34]郑泳智,朱定局,吴惠麟,等.知识图谱问答领域综述[J].计算机系统应用,2022,31(04):1-13.
- [35]漆桂林,高桓,吴天星.知识图谱研究进展[J].情报工程,2017,3(01):4-25.
- [36]宋雨霏.基于深度学习的实体关系抽取算法的研究与实现[D].北京邮电大学,2023.
- [37]XU Y, MOU L L, LI G, et al. Classifying relations via long short term memory networks along shortest dependency paths[C]//Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. Lisbon:Association for Computational Linguistics, 2015:1785-1794.
- [38]Huang Z , Xu W , Yu K .Bidirectional LSTM-CRF Models for Sequence Tagging[J].Computer Science, 2015.
- [39]Zhang S , Zheng D , Hu X ,et al.Bidirectional Long Short-Term Memory Networks for Relation Classification[C]//2015.
- [40]Dai X, Karimi S, Paris C. Medication and adverse event extraction from noisy text[C]//Proceedings of the Australasian Language Technology Association Workshop 2017. 2017: 79-87.
- [41]Vaswani A, Shazeer N, Parmar N,et al.Attention Is All You Need[J].arXiv, 2017.
- [42]Devlin J, Chang M W, Lee K, et al.BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding[J]. 2018.
- [43]任乐,张仰森,刘帅康.基于深度学习的实体关系抽取研究综述[J].北京信息科技大学学报(自然科学版),2023,38(06):70-79+87.
- [44]WEI Z P, SU J L, WANG Y, et al. A novel cascade binary tagging framework for relational triple extraction[C]//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. Stroudsburg:Association for Computational Linguistics, 2020:1476-1488.
- [45]Wang Y ,Yu B ,Zhang Y ,et al.TPLinker: Single-stage Joint Extraction of Entities and Relations Through Token Pair Linking[J]. 2020.
- [46]毛先领,李晓明.问答系统研究综述[J].计算机科学与探索,2012,6(03):193-207.
- [47]Wu P ,Zhang X ,Feng Z .A Survey of Question Answering over Knowledge Base[J]. 2019.
- [48]Zhang F ,Zhang Y ,Xu T .A Review of Knowledge Graph Technology in the field of Automatic Question Answering[J]. 2020.

- [49]王东升,王卫民,王石,等.面向限定领域问答系统的自然语言理解方法综述[J].计算机科学,2017,44(8):1-8.
- [50]侯梦薇,卫荣,陆亮,等.知识图谱研究综述及其在医疗领域的应用[J].计算机研究与发展,2018,55(12):2587-2599.
- [51]Wayne X Z, Kun Z, Junyi L, et al. A Survey of Large Language Models[J], Computing Research Repository, 2023, abs/2303.18223
- [52]张鹤译,王鑫,韩立帆,等. 模型融合知识图谱的问答系统研究[J].计算机科学与探索,2023,17(10):2377-2388.
- [53]冯钧,畅阳红,陆佳民,等.基于大语言模型的水工程调度知识图谱的构建与应用[J/OL].计算机科学与探索:1-13[2024-01-24 19:11].
- [54]SINGHAL A. Introducing the knowledge graph:things,not strings[J].Official Google blog,2012,5:16
- [55]Lan Y , He G , Jiang J ,et al.A Survey on Complex Knowledge Base Question Answering: Methods, Challenges and Solutions[J]. 2021.
- [56]CUI W,XIAO Y,WANG W,et al.KBQA:an online template based question answering system over freebase[C]//International joint conference on artificial intelligence.New York City,New York,USA:ACM,2016:4240-4241.
- [57]Liu A, Huang Z, Lu H, et al. BB-KBQA:BERT-based knowledge base question answering[C]//China National Conference on Chinese Computational Linguistics. Cham:Springer International Publishing,2019: 81-92.
- [58]BORDES A,CHOPRA S,WESTON J.Question answering with subgraph embeddings[C]//Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP).Doha,Qatar:EMNLP,2014:615-620.
- [59]DU Z, QIAN Y, LIU X, et al. GLM:general language model pretraining with autoregressive blank infilling[C]//Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics(Volume 1:Long Papers), Dublin, May 22-27, 2022. Stroudsburg:ACL, 2022:320-335.
- [60]OPENAI. GPT-4 technical report[R]. arXiv:2303.08774, 2023.
- [61]Sun Y , Wang S , Li Y ,et al.ERNIE: Enhanced Representation through Knowledge Integration[J]. 2019.
- [62]Touvron H, Lavril T, Izacard G, et al. LLaMA: Open and efficient foundation

- language models[J]. arXiv preprint arXiv:2302.13971, 2023.
- [63]Liu Y , Ott M , Goyal N ,et al.RoBERTa: A Robustly Optimized BERT Pretraining Approach[J]. 2019.
- [64]Lafferty J, Mccallum A,Pereira FCN. Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data[J]. proceedings of icml. 2002:282-289.
- [65]Topsakal O, Akinci T C. Creating large language model applications utilizing langchain: A primer on developing llm apps fast[C]//International Conference on Applied Engineering and Natural Sciences. 2023, 1(1): 1050-1056.