Computer Science and Application 计算机科学与应用, 2025, 15(12), 100-111
Published Online December 2025 in Hans. https://www.hanspub.org/journal/csa
https://doi.org/10.12677/csa.2025.1512325

Hans汉斯

基于Lora微调的阿尔兹海默症知识图谱构建

金　瀚，李　莉

天津职业技术师范大学电子工程学院，天津

收稿日期：2025年10月29日；录用日期：2025年11月27日；发布日期：2025年12月4日

摘　要

阿尔茨海默症是一种不可逆的神经退行性疾病，利用知识图谱对阿尔茨海默症和轻度认知障碍患者进行准确的辅助诊断具有重要意义。然而传统知识图谱构建方法通常依赖大量人工标注，成本高且领域适应性有限。近年来，人工智能技术特别是大语言模型的快速发展，为此提供了新的技术支撑。本文提出一种基于大语言模型与Lora微调的阿尔茨海默症知识图谱构建方法，旨在为低资源、低成本场景下高效构建知识图谱提供参考。该方法通过设计合理的信息抽取提示模板并构建指令数据集，分别采用中英文语料的5个大语言模型进行少样本的Lora微调，对比分析实体关系联合抽取的不同表现。实验结果表明，Llama-3.1-Tulu-3-8B在实体关系联合抽取方面表现最优，在60个训练轮次下精确率达到82.5%，并进一步实现了从相关文献中自动抽取阿尔茨海默症知识，并完成知识图谱的构建与可视化分析。

关键词

大语言模型，LoRA微调，阿尔兹海默症，Neo4j

Construction of Alzheimer's Disease Knowledge Graph Based on LoRa Fine-Tuning

Han Jin, Li Li

School of Electronic Engineering, Tianjin University of Technology and Education, Tianjin

Received: October 29, 2025; accepted: November 27, 2025; published: December 4, 2025

Abstract

Alzheimer's disease is an irreversible neurodegenerative disorder. Utilizing knowledge graphs for accurate auxiliary diagnosis of patients with Alzheimer's disease and mild cognitive impairment holds significant importance. However, traditional knowledge graph construction methods often rely on

文章引用: 金瀚, 李莉. 基于 Lora 微调的阿尔兹海默症知识图谱构建[J]. 计算机科学与应用, 2025, 15(12): 100-111.
DOI: 10.12677/csa.2025.1512325

extensive manual annotation, which is costly and has limited domain adaptability. In recent years, the rapid development of artificial intelligence technology, especially large language models, has provided new technical support for this purpose. This paper proposes a method for constructing an Alzheimer's disease knowledge graph based on large language models and Lora fine-tuning, aiming to provide a reference for efficiently constructing knowledge graphs in low-resource and low-cost scenarios. This method involves designing reasonable information extraction prompt templates and constructing an instruction dataset. Five large language models in both Chinese and English corpora are employed for few-shot Lora fine-tuning, and the different performances of joint entity relation extraction are comparatively analyzed. Experimental results show that Llama-3.1-Tulu-3-8B performs optimally in joint entity relation extraction, achieving an accuracy rate of 82.5% after 60 training epochs. Furthermore, it automatically extracts Alzheimer's disease knowledge from relevant literature, completes the construction of the knowledge graph, and performs visual analysis.

Keywords

Large Language Model, LoRA Fine-Tuning, Alzheimer's Disease, Neo4j

1. 引言

阿尔茨海默症(Alzheimer's Disease, AD)是一种由德国神经病理学家 Alois Alzheimer 于 1906 年首次发现并命名的神经退行性疾病，其主要发病群体为 65 岁及以上的老年人[1]。随着病情进展，患者可能出现语言能力下降、空间定向障碍、肌肉萎缩甚至全身功能衰竭，最终导致死亡[2]。研究预测，到 2050 年，全球阿尔兹海默症患者人数将再增两倍以上，其导致的死亡可能占美国老年人口死亡总数的 43% [3]。阿尔兹海默症的诊断过程通常分为多个阶段，包括早期轻度认知障碍 MCI、晚期 MCI 以及完全发展的阿尔兹海默症。其中，MCI 阶段被视为正常衰老与病理性认知衰退的关键过渡期，也是临床干预的黄金窗口[4]。当前，阿尔兹海默症的研究多集中于医学实验领域，但随着信息技术的快速发展，尤其是互联网所带来的数据流通性增强，大量与阿尔兹海默症相关的临床、遗传与病理数据得以积累[5]。在此背景下，知识图谱(Knowledge Graph, KG)作为一种结构化语义知识表示方法，逐渐展现出其在整合领域知识、挖掘潜在关联方面的独特价值。知识图谱的概念最早由 Google 于 2012 年提出，其核心理念是将世界理解为“事物”而非“字符串”，旨在揭示知识之间深层次的语义关联[6]。通用知识图谱如 Google Knowledge Graph、YAGO [7]、DBpedia [8]等。

然而，传统知识图谱构建方法高度依赖人工标注语料，成本高昂且难以适应低资源条件下的领域扩展需求[9]。近年来，生成式大语言模型(Large Language Models, LLMs)的快速发展为知识图谱的自动化构建提供了新的技术路径[10]。此类模型通过参数高效微调技术(Parameter-Efficient Fine-Tuning, PEFT)，如低秩自适应(Lora)，能够在仅更新少量参数的情况下，使通用大模型快速适应特定领域任务，有效克服了全参数微调对数据量与算力的高要求。

本文聚焦于探索在低资源条件下，如何结合大语言模型与 Lora 微调方法，实现阿尔兹海默症领域的高质量知识图谱构建。具体研究路径包括：设计适用于阿尔兹海默症实体关系抽取的提示模板与指令数据集，对包括 Qwen、DeepSeek、GLM、Llama、Gemma 在内的多款中英文大模型进行少样本 Lora 微调，系统比较各模型在实体关系联合抽取任务上的性能差异。

2. 研究方法原理

2.1. 知识图谱及其构建方法

知识图谱(Knowledge Graph, KG)作为一种结构化的语义知识库，其概念由 Google 公司在 2012 年正式提出，并宣布以此为基础构建下一代智能化搜索引擎，标志着信息检索从“字符串匹配”向“事物理解”的根本性转变。

知识图谱本质是一个语义网络，其核心是以“实体–关系–实体”的三元组形式对知识进行符号化表示。一个知识图谱 G 可以形式化地定义为 G = (E, R, S)，其中 E 是实体集合，R 是关系集合，S 是由三元组构成的事实集合[11]。

早期方法主要依赖于基于规则、字典与统计机器学习的技术，在构建流程上通常遵循“数据获取→知识抽取→知识融合→知识加工”[12]。然而传统方法，特别是其中的命名实体识别模型，面临诸多固有挑战：如只能识别预定义的有限实体类型、严重依赖专家知识进行手动特征工程、对文本上下文语义的理解能力薄弱等。而后以 BERT 为代表的预训练语言模型的兴起，为知识图谱构建带来了革命性变化，通过在大规模语料上进行自监督预训练，这些模型掌握了深层的语言表征能力，研究者通过简单的微调即可使模型适应特定的知识抽取任务，显著降低了特征工程的复杂度，提升了自动化水平[13]，典型的应用模式是将 BERT 与 BiLSTM、CRF 等模块结合，形成如 BERT-BiLSTM-CRF 等混合架构，在科技报告、生物医学、智慧图书馆等多个领域的知识抽取中取得了优异效果[14]。

近年来，生成式大语言模型的迅猛发展标志着知识图谱构建范式的新一轮转变。LLMs 凭借其强大的通用知识、上下文理解能力及指令遵循能力，为低资源条件下的知识图谱构建开辟了新路径，经过多年发展，知识图谱领域已形成了丰富的资源生态。全球范围内存在诸多大型开放知识库，如 Freebase [15]、Wikidata [16]等，它们构成了通用知识图谱的核心。在国内，OpenKG 等项目致力于推动以中文为核心的知识图谱数据的开放与互联[17]，此外为应对多模态信息处理的需求，也涌现出如 MarKG [18]、MARs [19]等多模态知识图谱基准数据集。

知识图谱技术从概念提出到生态繁荣，其构建方法经历了从依赖人工规则到利用预训练模型，再到当前由生成式大语言模型与参数高效微调技术驱动的演进历程。当前，生成式大模型与 Lora 等高效微调技术的结合，为解决专业领域知识图谱构建的“低资源”困境提供了更加优秀的方案。它使得研究人员能够在不依赖海量标注数据的前提下，快速、灵活地构建出高质量的领域知识图谱，这对于加速如阿尔茨海默症等复杂疾病的科研发现与临床辅助诊断具有重大意义。

2.2. 大语言模型及其在知识图谱中的应用

大语言模型(Large Models)，特指基于 Transformer 架构、拥有超大规模参数量的神经网络模型，通过在海量文本或多模态数据上进行预训练，展现出强大的语言理解与内容生成能力[20]。

在知识图谱构建领域，大模型的应用正呈现出多元化与专业化的趋势。在应用层面，大模型凭借其强大的自然语言处理能力，已在中医、生物医学、电力工程等垂直领域的知识抽取、表示与挖掘中展现出巨大潜力，有效提升了从非结构化文本中提取实体与关系的效率，并改善了对于长尾知识的覆盖能力。在技术路径上，研究探索呈现多样化，包括将大模型与零样本学习、迁移学习等范式结合，以增强其在低资源场景下的适应性。其中，参数高效微调技术(Parameter-Efficient Fine-Tuning, PEFT)，特别是低秩自适应(Lora)，因其能通过在预训练模型旁注入少量可训练参数，实现在低资源条件下高效适配特定任务，而成为研究热点。然而，该领域研究总体上仍处于探索阶段，存在明显不足。现有工作相对零散，尚未形成系统化的技术框架；尤其在医学这类高专业度领域，仍缺乏针对非结构化文献进行精准、可靠信息抽取的有效方案。

3. 研究设计构建

本研究旨在探索一种基于大模型与参数高效微调技术(Lora)的领域知识图谱构建方法，以期在低资源条件下，实现从阿尔兹海默症文献中高效、准确地抽取结构化知识。本研究提出一套基于大模型与 Lora 微调的阿尔兹海默症知识图谱构建框架，包含四个核心模块：数据收集与处理、提示模板设计、模型训练、知识图谱构建与可视化。具体的构建框架如图 1 所示：

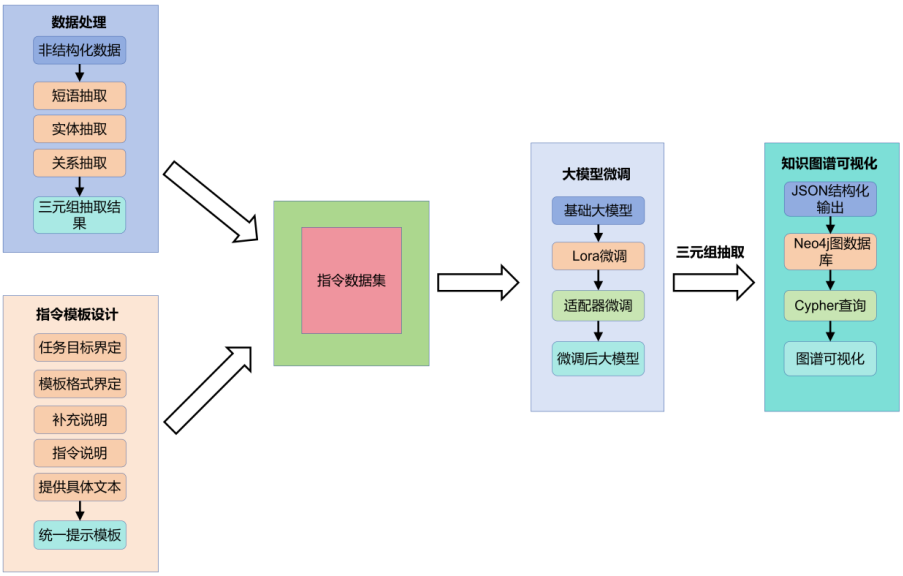

Figure 1. Framework structure diagram
图 1. 框架结构图

3.1. 提示模板设计

为克服大模型的“幻觉”问题并将其泛化能力引导至阿尔兹海默症领域，针对阿尔兹海默症领域文本跨度大、实体关系复杂、因果关系隐含等特点，本文设计了一套综合性的指令模板，旨在约束模型输出、提升事实准确性并实现结构化表达。具体的提示模板设计如图 2 所示：

```
"system"系统提示词：
"任务目标界定": "你现在是一个用于抽取三元组的模型。你的任务是根据原文中的句子和上下文关系抽取实体和关系，并将实体和关系组合成三元组"
"模板格式界定": "三元组具体格式如下：<id(label:name), relation, id(label:name)>。其中"relation"是实体之间的关系。每个三元组应包括两个实体和一个关系。实体格式如下：<label，id，
name>的形式进行实体的定义，其中"label"为实体的类别，在本图谱中总共有 16 类实体，"id"是一个自定义的实体 id，用来唯一标志实体，形式为大写字母+数字，不同实体类型的大写字母不同，
"name"为实体的名字每个实体由实体类型和实体名称组成，实体的名称为需提取段落中的单词或短语，关系由关系类型和关系值组成，关系类型要根据语句内关系和上下文关系进行抽取"
"补充说明":"16类实体类别如下：A:Analytical，Diagnostic and Therapeutic Techniques，and Equipment;B:Biomolecules;C:Chemicals and Drugs;D:Diseases;E:Disciplines and
Occupations;F:Publication Characteristics;G:Geographicals;H:Health Care;I:Information Science;N:Named Groups;O:Organisms;P:Phenomena and Processes;Q:Psychiatry and
Psychology;R:Humanities;S:Anthropology，Education,Sociology，and Social Phenomena;T:Technology，Industry，andAgriculture。17类关系类别如下：treat（药物治疗疾病）;cause（原
因引发疾病，疾病引发症状，实验分析效果）;associate（并发疾病,有联系症状）;diagnosis（诊断工具诊断分析疾病）;research_location（疾病研究的区域）;research_field（研究领
域）;psychological_behaviour（疾病对应的心理及行为）;effect（影响疾病）;increase（加剧，恶化）reduce（减少）;include（包括）;require（需要）;attribute（属性）;function（作
用）;differ（不同于）;better（更有效）;worse（更有害）。如果关系中不存在关系值，则无需抽取该关系值,除了diagnosis，实体关系顺序皆是从前往后的因果关系"
"指令说明":"以下是要抽取的文本"

"user"抽取的文本内容
"待抽取的阿尔兹海默症文献": "In this Seminar, we highlight the main developments in the field of Alzheimer's disease. The most recent data indicate that, by 2050, the
prevalence of dementia will double in Europe and triple worldwide, and that estimate is 3 times higher when based on a biological (rather than clinical) definition of
Alzheimer's disease. The earliest phase of Alzheimer's disease (cellular phase) happens in parallel with accumulating amyloid β, inducing the spread of tau pathology.
The risk of Alzheimer's disease is 60-80% dependent on heritable factors, with more than 40 Alzheimer's disease-associated genetic risk loci already identified, of which
the APOE alleles have the strongest association with the disease. Novel biomarkers include PET scans and plasma assays for amyloid β and phosphorylated tau, which show
great promise for clinical and research use. Multidomain lifestyle-based prevention trials suggest cognitive benefits in participants with increased risk of dementia.
Lifestyle factors do not directly affect Alzheimer's disease pathology, but can still contribute to a positive outcome in individuals with Alzheimer's disease. Promising
pharmacological treatments are poised at advanced stages of clinical trials and include anti-amyloid β, anti-tau, and anti-inflammatory strategies."

"assistant"三元组抽取结果
"实体关系三元组抽取结果": "<D001(Diseases:Alzheimer's disease), cause,B001 (Biomolecules:amyloid β)>,<D001(Diseases:Alzheimer's disease), cause,P001 (Phenomena and
Processes:the spread of tau pathology)>,<B002 (Biomolecules:heritable factors), cause(60-80%),D001(Diseases:Alzheimer's disease)>,<B002 (Biomolecules:heritable factors),
include,B003 (Biomolecules:APOE alleles)>,<A001 (Analytical, Diagnostic and Therapeutic Techniques, and Equipment:PET scans),diagnosis,B001 (Biomolecules:amyloid β)>,
<A001 (Analytical, Diagnostic and Therapeutic Techniques, and Equipment:PET scans),diagnosis,B004 (Biomolecules:phosphorylated tau)>,<A002 (Analytical, Diagnostic and
Therapeutic Techniques, and Equipment:plasma assays),diagnosis,B001 (Biomolecules:amyloid β)>,<A002 (Analytical, Diagnostic and Therapeutic Techniques, and
Equipment:plasma assays),diagnosis,B004 (Biomolecules:phosphorylated tau)>,<S001 (Anthropology, Education,Sociology, and Social Phenomena:Lifestyle
factors),effect,D001(Diseases:Alzheimer's disease)>,<C001 (Chemicals and Drugs:anti-amyloid β),treat,D001(Diseases:Alzheimer's disease)>,<C002 (Chemicals and Drugs:anti-
tau),treat,D001(Diseases:Alzheimer's disease)>,<C003 (Chemicals and Drugs:anti-inflammatory strategies),treat,D001(Diseases:Alzheimer's disease)>"
```

Figure 2. Template design diagram
图 2. 模板设计图

该设计重点涉及了以下四个方面：

精确的实体识别与分类：模板明确定义了阿尔兹海默症领域的关键实体类型，模板设计强调以疾病成因、病理机制和临床表现为核心，要求模型准确识别并归类关键实体，引导模型对专业术语进行精准识别与归类。

复杂关系网络的构建：阿尔兹海默症的病理过程涉及多因素的复杂相互作用。为全面捕捉实体间复杂的语义关联，指令模板需能够引导模型不仅识别实体，更能深入理解并推理实体间错综复杂的语义关系。通过预设多样化的关系类型，指导模型挖掘跨领域的深层关系，构建出能够反映疾病全貌的深层关系网络。

因果顺序的强调与提取：在阿尔兹海默症的疾病进展中，许多关系具有明确的因果性与时序性。指令模板设计特别注重对这类逻辑顺序的提取，要求模型能够精确判断实体间相互作用的先后顺序与方向性这种对因果关系的显式约束，是生成具备准确逻辑的三元组、避免事实性错误的关键。

信息的结构化与标准化输出：为实现与下游知识图谱构建流程的无缝衔接，指令模板强制要求模型将所有抽取出的信息以标准化的结构化格式输出。本设计采用 JSON 作为统一输出格式，明确规定了实体和关系的类型体系，并要求以(头实体，关系，尾实体)的三元组形式呈现结果。这种规范化的输出模式，极大地保障了信息的一致性、可处理性与可集成性，直接服务于后续的图数据库存储与应用。

3.2. 大模型 Lora 微调

为实现通用大模型在阿尔茨海默症领域知识抽取任务上的精准适配，本研究采用以低秩自适应 Lora 为核心的参数高效微调技术[21]。该方法的实施关键在于，利用前期构建的结构化指令数据集，对选定的基座大模型进行定向优化，从而显著提升其在专业文本中的实体、关系及事件抽取能力。具体的 Lora 框架如图 3 所示：

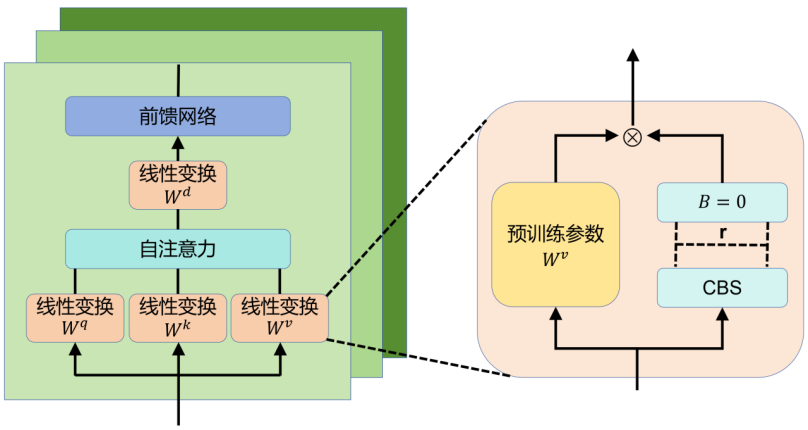

Figure 3. Lora structure diagram
图 3. Lora 结构图

与传统提示学习(Prompt Learning)仅通过设计外部模板来引导模型输出不同，Lora 是一种参数高效的微调方法。其核心思想是冻结预训练模型原有的权重参数 W，不进行直接更新，转而通过在模型的线性变换层旁路引入一组可训练的低秩矩阵 A (CBS)和 B 来模拟参数增量，即 W′ = W + BA。在初始化时，矩阵 A 采用随机高斯分布，矩阵 B 初始化为零，从而确保训练起始时旁路对模型输出的贡献为零，维持其原有状态。

此架构拥有两大优势：首先，该架构将需要优化的参数量降低了数个量级，利用有限的、经过专家

标注的阿尔兹海默症领域数据，快速捕捉其中的专业词汇、表达习惯与知识关联模式，从而大幅减少了训练过程中的内存占用与计算消耗，实现模型在低资源条件下的高效领域自适应；其次，由于主体参数被冻结，模型在预训练阶段获得的通用语言理解与推理能力得以完整保留，在此基础上通过微调少量新增参数，模型被专门优化以处理阿尔兹海默症领域中的复杂语境，并挖掘生物因素之间潜在的隐含关系，基于 Lora 微调后的模型在多项下游任务中能够达到与全量微调相媲美的性能，同时因其仅对极小部分参数进行更新，有效缓解了过拟合风险，增强了模型的泛化能力。这些优势使其在数据稀缺的垂直领域中展现出巨大应用价值，能够高效捕捉领域特有的语言风格、知识结构及隐含的语义关系。

4. 实验结果与分析

4.1. 数据集预处理

本研究选取国际权威生物医学文献数据库 PubMed 作为核心数据来源。PubMed 收录了全球范围内超过 5300 种核心生物医学期刊，是获取高质量、前沿阿尔兹海默症研究成果的理想平台[22]。本实验以“阿尔茨海默症”为核心主题，分别将检索条件限定在标题/摘要字段与 MeSH 主题词字段，系统性地采集了 2020 年至 2025 年间公开发表的文献。经过初步检索与人工筛选，最终确定并下载了 100 篇相关研究的摘要文本，构成后续知识抽取与图谱构建的原始语料库。

4.2. 指令数据集构建

为实现大模型在阿尔茨海默症领域的精准知识抽取，本研究构建了一个高质量的少样本指令微调数据集。标注体系的建立以 PubMed 的医学主题词表 MeSH 为基础，并结合阿尔兹海默症领域文献的语言特点，最终定义了 16 种实体类型(如基因、药物、病理过程等)与 17 种语义关系类型(如抑制、表达、关联等)，形成了结构化的分类体系，具体类别如表 1 和表 2 所示。

Table 1. Table of 16 entity types
表 1. 16 种实体类型表

类别标号	类别名称	类别标号	类别名称
A	Analytical, Diagnostic and Therapeutic Techniques, and Equipment	I	Information Science
B	Biomolecules	N	Named Groups
C	Chemicals and Drugs	O	Organisms
D	Diseases	P	Phenomena and Processes
E	Disciplines and Occupations	Q	Psychiatry and Psychology
F	Publication Characteristics	R	Humanities
G	Geographicals	S	Anthropology, Education, Sociology, and Social Phenomena
H	Health Care	T	Technology, Industry, and Agriculture

Table 2. Table of 17 relationship types
表 2. 17 种关系类型表

类别标号	类别名称	类别标号	类别名称
1	treat (药物治疗疾病)	10	reduce (减少)
2	cause (原因引发疾病，疾病引发症状，实验分析效果)	11	include (包括)

续表

3	associate (并发疾病，有联系症状)	12	require (需要)
4	diagnosis (诊断工具诊断分析疾病)	13	attribute (属性)
5	research_location (疾病研究的区域)	15	function (作用)
6	research_field (研究领域)	15	differ (不同于)
7	psychological_behaviour (疾病对应的心理及行为)	16	better (更有效)
8	effect (影响疾病)	17	worse (更有害)
9	increase (加剧，恶化)		

在标注质量保障方面，研究采用双重预标注生成数据集样本后，随机选取 40 篇超过 300 字的阿尔兹海默症论文摘要进行人工标注，提取结构化三元组信息，并以 JSON 格式存储。最后，邀请领域专家对全部 50 篇标注结果进行人工复审与校正，确保知识的准确性与一致性，最终形成 50 个高质量的标注样本作为训练语料，构建出了大模型指令微调数据集。指令微调数据集如图 4 所示：

Figure 4. Instruction fine-tuning dataset
图 4. 指令微调数据集

4.3. 实验环境与训练参数

本实验基于远程服务器上进行，操作系统为 Linux，所使用的 GPU 为 NVIDIA GeForce RTX 4090，24G 显存。实验环境配置的 Python 版本为 3.9，基于 pytorch 框架开，cuda 版本为 12.1，Pytorch 版本为 2.1.0。

本实验在训练前需要进行超参数配置，训练轮数 epoch 设为根据实验要求分别设置 30，60，100 三组对照组，初始学习率 Learning rate 设置为 5×10^{-5}，截断最大长度 cutoff_len 设置为 8192，保证内容充分训练，由于训练的模型较大，内存占用高，设定训练批次 Batch_size 设置为 1，Lora 秩 Lora_rank 设定为 8，Lora 缩放系数 Lora_alpha 设定 32。而在推理阶段则同时使用温度采样策略和 Top-P 采样策略，温度设定为 0.95，Top-p 设定为 0.7。

4.4. 评价指标

为系统验证本文所提出的方法在阿尔兹海默症知识三元组抽取任务中的准确性与有效性，本研究设计了严谨的对比实验和消融实验，选取了当前具有代表性的国内外开源大语言模型作为基线，其中国内模型

包括 GLM4-9B [23]、Qwen3-8B [24]及 DeepSeek-Rl-Distil1-Qwen-7B [25]，国外模型则涵盖 Llama-3.1-Tulu-3-8B [26]与 Gemma2-9b-it-SimPO [27]，采用 Lora 方法进行少样本微调，重点对比分析各模型在实体关系联合抽取任务上的性能差异，旨在全面评估不同模型架构与预训练策略在本领域任务上的适应性。

模型性能评估采用信息抽取领域的标准指标：精确率(Precision, P)、召回率(Recall, R)与 $F1$ 值($F1$-Score)。其计算公式如下：

$$P = \frac{T_p}{T_p + F_p}$$

$$R = \frac{T_p}{T_p + F_p}$$

$$F_1 = \frac{2P \cdot R}{P + R}$$

其中，T_p (True Positive)指被模型正确预测的正例三元组数量，F_p (False Positive)为被模型错误预测为正例的三元组数量，F_n (False Negative)则表示实际存在但未被模型识别出的正例三元组数量。精确率 P 衡量了模型预测结果的可靠性，即所有被预测为正例的样本中真正为正例的比例；召回率 R 则反映了模型对数据中全部正例样本的覆盖能力；F_1 值是精确率与召回率的调和平均数，用以综合评价模型在阿尔兹海默症知识抽取任务中的整体性能。

4.5. 对比实验

为评估不同大语言模型在低资源条件下对阿尔兹海默症领域知识抽取的性能表现，本研究在训练样本量为 50 例、训练轮次为 60 轮的统一设置下，对四个开源大模型进行了系统性对比实验。具体的实验结果下表 3 所示：

Table 3. Summary table of comparative experiments for different models
表 3. 不同模型对比实验情况一览表

model	抽取类别	精确度(P)	召回率(R)	$F1$ 值
Llama-3.1-Tulu-3-8B	三元组	0.766	0.894	0.825
	实体	0.779	0.916	0.842
	关系	0.766	0.894	0.825
Qwen3-8B	三元组	0.741	0.828	0.782
	实体	0.774	0.822	0.797
	关系	0.741	0.828	0.782
Gemma2-9b-it-SimPO	三元组	0.817	0.741	0.777
	实体	0.864	0.844	0.854
	关系	0.817	0.741	0.777
GLM4-9b-chat-hf	三元组	0.786	0.677	0.727
	实体	0.791	0.697	0.741
	关系	0.786	0.677	0.727
DeepSeek-R1-Distill-Qwen-7B	三元组	0.463	0.626	0.532
	实体	0.463	0.626	0.532
	关系	0.463	0.626	0.532

通过对表 3 分析可以得出，在相同的训练条件下，各模型在阿尔兹海默症三元组抽取任务上的性能存在显著差异。从综合评价指标 *F*1 值来看，Llama3.1 表现最佳，其 *F*1 值达到 0.825，展现出最优的综合性能。Qwen3 与 Gemma2 表现相当，*F*1 值分别为 0.782 和 0.777，位列第二梯队。GLM4 以 0.727 的 *F*1 值紧随其后，而 DeepSeek-Rl 的性能则远逊于其他模型，*F*1 值仅为 0.532。

经过全面综合分析可以看出，Llama3.1 在精确率与召回率上取得了最佳平衡，其三元组抽取的精确率为 0.766，召回率高达 0.894。这一结果表明，该模型不仅能够准确地识别正例，还具备强大的上下文信息覆盖能力，能最大限度地减少信息遗漏。其优异表现可能归因于其在高质量英文语料上的充分预训练，使其对生物医学文献的语境具有出色的综合理解能力。而同样以英文语料训练的 Gemma2 也具有语料的优势，同时对特定领域复杂语境的理解能力较强，使其拥有全模型第一的精确度 0.817，然而其召回率相对较低，导致其综合 *F*1 值略低于 Llama 模型。值得关注的是 Qwen3 虽然作为中文模型，却可以通过实验看出其在英文文献的抽取能力也十分优秀，其 *F*1 值和 Gemma2 旗鼓相当甚至略胜一筹，精确度达到了 0.741，召回率达到了 0.828，在英文文献的抽取任务中展现出了卓越的跨语言迁移能力，这使得其在中英文结合的综合文献分析抽取中应能取得比两个外语模型更好的精确度和结果，具有非常高的潜在开发价值。另外，GLM4 在三元组抽取任务上的精确度名列前茅，达到了 0.786，仅次于 Gemma2，这应当归功与二者的模型大小，使其拥有了更强的局部语义解析能力，但其召回率仅有 0.677 明显偏低，表明模型对英文上下文的整体理解存在局限，可能导致大量有效三元组被遗漏。最后 DeepSeek-Rl 作为依据 Qwen 架构构建的蒸馏模型，在性能和抽取效果上均不佳，可以看出侧重于逻辑推理的 DeepSeek-Rl 即使经过了基于 Qwen 框架的模型蒸馏，对于非结构化英语文本的抽取任务能力依旧不足，以及其对于提示词相对敏感，需要更多的样本训练，而本文所研究的少样本提示训练会降低其性能；同时 DeepSeek-Rl 在信息抽取任务中依赖复杂逻辑推理的机制导致抽取效率显著低于其他模型，在处理复杂语义关系时性能受限，表明其在面向非结构化文本的快速知识抽取应用场景中存在明显局限。

本次对比实验有效评估了各模型在低资源场景下的少样本学习能力，这对于阿尔兹海默症这类涉及多学科交叉、标注数据往往稀缺的研究领域具有重要实践意义，验证了大模型在医学专业领域少样本学习中的可行性，为在标注数据稀缺的特殊医学子领域，如罕见病或新兴研究方向中应用大模型技术提供了实践参考。不同模型在精确率与召回率上的差异化表现，也为针对不同应用场景的模型选型提供了灵活选择空间，验证了选择 Llama-3.1-Tulu-3-8B 作为后续研究核心模型的合理性，同时 Qwen3-8B 所展现的出色跨语言能力，也为未来面向多语言医学文本的智能处理研究提供了有价值的参考。

4.6. 知识图谱可视化分析

基于前期实验结果，本研究采用性能最优的 LoRA 微调后 Llama-3.1-Tulu-3-8B 模型，对从 PubMed 数据库获取的 150 篇阿尔兹海默症相关文献摘要进行了系统性的信息抽取与知识图谱构建，最终形成的知识网络包含约 2200 个实体和 3000 对三元组，构成了一个多维度的阿尔兹海默症知识体系。

在知识存储与可视化方面，本研究采用 Neo4j 图形数据库作为知识存储基础架构，通过 Python 接口将抽取的三元组数据导入数据库系统。借助 Neo4j 提供的原生可视化工具与 Cypher 查询语言，实现了知识图谱的多维度展示。鉴于阿尔兹海默症知识图谱的庞大规模，本文仅展示了通过 Cypher 语句查询得出的部分阿尔兹海默症的知识图谱结构，查询响应时间在 0.5 毫秒，拥有良好的查询效率。具体知识图谱如图 5 所示。

从可视化结果分析，该知识图谱呈现出清晰的层级化网络结构，以“阿尔兹海默症”为核心节点，向外辐射出多个关系网络：病因网络(如“淀粉样 β 蛋白沉积”)、治疗网络(如“抗炎症策略”)以及症

状网络(如“tau 蛋白病理传播”)等不同领域的网络关系。这种可视化呈现不仅通过该图谱可清晰得到阿尔兹海默症的与多个领域之间的复杂关联，如阿尔兹海默症与各生物大分子之间的关系，何种因素会加剧或者削弱阿尔兹海默症，以及阿尔兹海默症会带来何种症状，会并发哪些疾病等重要节点，还突破了传统文献阅读的线性思维局限，使得隐藏在大量文献中的深层次病理机制得以立体化、网络化呈现。

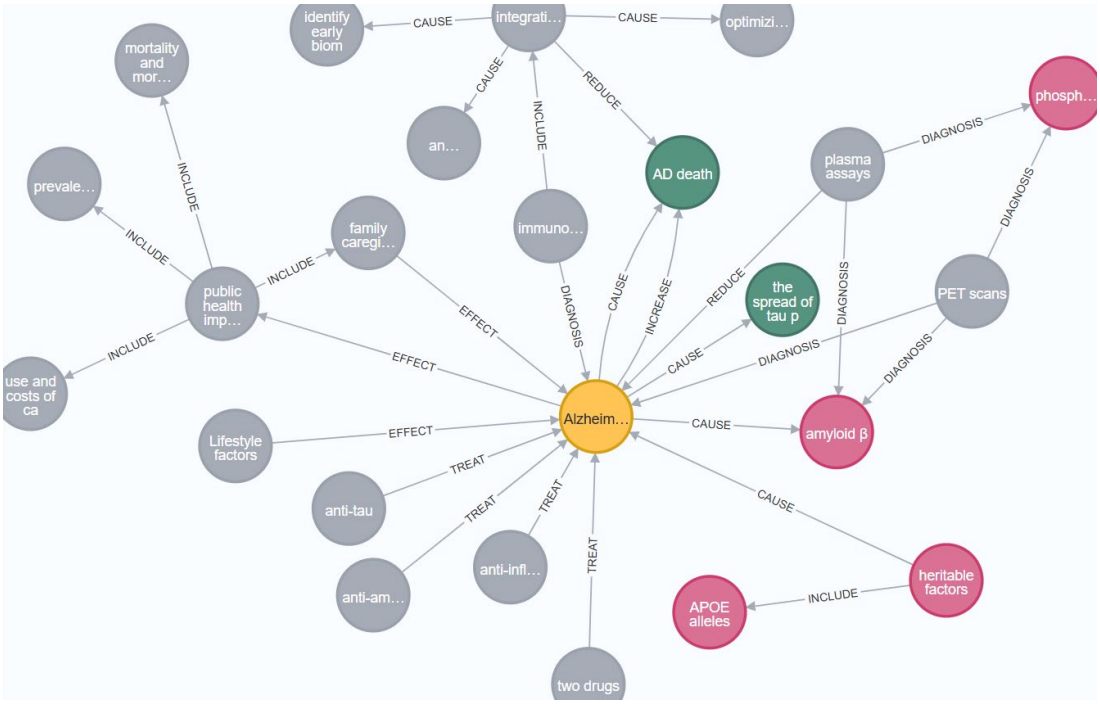

Figure 5. Partial knowledge graph of Alzheimer’s disease in Neo4j
图 5. Neo4j 阿尔兹海默症部分知识图谱

5. 结语

本文针对阿尔茨海默症领域知识结构化与可视化的需求，提出了一种基于大语言模型与提示学习的知识图谱构建方法。该方法突破了传统依赖人工标注与规则模板的知识抽取范式，通过大模型对非结构化医学文献进行端到端的信息抽取，实现了高效、低成本的知识图谱构建。在模型选型与优化方面，本研究系统对比了多个主流开源大模型，实验表明 Llama-3.1-Tulu-3-8B 在阿尔兹海默症实体关系联合抽取任务中表现最优。进一步采用参数高效的 LoRA 微调技术，在仅使用 50 个样本、训练 60 轮的情况下，模型准确率达到 0.817，性能显著优于未微调的基础模型，验证了少样本条件下领域自适应的有效性。通过实验，本文提出了融合大模型与 LoRA 微调的阿尔兹海默症知识图谱构建框架，为低资源场景下的领域知识结构化提供了可行路径；并通过实证分析确定了阿尔兹海默症信息抽取任务中的最优模型与微调参数，为医学自然语言处理任务提供了重要的模型选型依据；最后完成了阿尔兹海默症知识图谱的构建、可视化与初步应用分析，为疾病机制研究与临床知识服务提供了数据基础与工具支持。尽管本研究取得了预期成果，仍存在一定局限。例如，模型对复杂语义关系与隐含知识的推理能力尚且不足，尤其是当此语义关系之间有大量复杂逻辑关系且中间间隔大量文本的时候尤为明显，在此方面有明显提升空间；同时在大规模应用中的系统效率与扩展性也有待进一步验证。后续研究将致力于引入更复杂的推理机制，如探索将知识图谱与外部逻辑规则库相结合，实现多步推理，以及扩展多模态知识来源，结合如 YOLO

等快速视觉模型达到视觉语言共同应用，并探索在更广泛医学领域的迁移应用，以持续提升知识图谱的深度与实用性。

参考文献

[1] 张雷, 范占芳, 张作鹏, 程卯生, 刘洋. 阿尔兹海默症发病机制及相关治疗药物的研究进展[J]. 中国药物化学杂志, 2021, 31(6): 438-446+469.

[2] 邓青芳, 马风伟. 阿尔兹海默病的发病机制及药物治疗研究进展[J]. 贵州师范大学学报(自然科学版), 2020, 38(1): 104-111.

[3] 王威丽, 宋沧桑. 阿尔兹海默病发病机制的研究进展及临床用药[J]. 中国药物评价, 2019, 36(3): 204-209.

[4] 曾安, 贾龙飞, 潘丹, 等. 基于卷积神经网络和集成学习的阿尔茨海默症早期诊断[J]. 生物医学工程学杂志, 2019, 36(5): 711-719.

[5] 楚阳, 徐文龙. 基于计算机辅助诊断技术的阿尔兹海默症早期分类研究综述[J]. 计算机工程与科学, 2022, 44(5): 879-893.

[6] Hogan, A., Blomqvist, E., Cochez, M., D'amato, C., Melo, G.D., Gutierrez, C., *et al.* (2020) Knowledge Graphs. *ACM Computing Surveys*, **54**, 1-37. https://doi.org/10.1145/3447772

[7] Rebele, T., Suchanek, F., Hoffart, J., Biega, J., Kuzey, E. and Weikum, G. (2016) YAGO: A Multilingual Knowledge Base from Wikipedia, Wordnet, and Geonames. In: Groth, P., *et al.*, Eds., *Lecture Notes in Computer Science*, Springer International Publishing, 177-185. https://doi.org/10.1007/978-3-319-46547-0_19

[8] Auer, S., Bizer, C., Kobilarov, G., Lehmann, J., Cyganiak, R. and Ives, Z. (2007) DBpedia: A Nucleus for a Web of Open Data. In: Aberer, K., *et al.*, Eds., *Lecture Notes in Computer Science*, Springer, 722-735. https://doi.org/10.1007/978-3-540-76298-0_52

[9] 陈涛, 刘炜, 单蓉蓉, 等. 知识图谱在数字人文中的应用研究[J]. 中国图书馆学报, 2019, 45(6): 34-49.

[10] Minaee, S., Mikolov, T., Nikzad, N., Chenaghlu, M., Socher, R., Amatriain, X. and Gao, J. (2024) Large Language Models: A Survey. ArXiv, abs/2402.06196.

[11] Zhang, Z., Cao, L., Chen, X., Tang, W., Xu, Z. and Meng, Y. (2020) Representation Learning of Knowledge Graphs with Entity Attributes. *IEEE Access*, **8**, 7435-7441. https://doi.org/10.1109/access.2020.2963990

[12] Nickel, M., Murphy, K., Tresp, V. and Gabrilovich, E. (2015) A Review of Relational Machine Learning for Knowledge Graphs. *Proceedings of the IEEE*, **104**, 11-33. https://doi.org/10.1109/jproc.2015.2483592

[13] Kim, B., Hong, T., Ko, Y. and Seo, J. (2020) Multi-Task Learning for Knowledge Graph Completion with Pre-Trained Language Models. *Proceedings of the* 28*th International Conference on Computational Linguistics*, Barcelona, 8-13 December 2020, 1737-1743. https://doi.org/10.18653/v1/2020.coling-main.153

[14] Gao, X. and Li, Q. (2021) Named Entity Recognition in Material Field Based on Bert-Bilstm-Attention-CRF. 2021 *IEEE Conference on Telecommunications*, *Optics and Computer Science* (*TOCS*), Shenyang, 10-11 December 2021, 955-958. https://doi.org/10.1109/tocs53301.2021.9688665

[15] Pellissier Tanon, T., Vrandečić, D., Schaffert, S., Steiner, T. and Pintscher, L. (2016) From Freebase to Wikidata: The Great Migration. *Proceedings of the* 25*th International Conference on World Wide Web*, Montréal, 11-15 April 2016, 1419-1428. https://doi.org/10.1145/2872427.2874809

[16] Vrandečić, D. and Krötzsch, M. (2014) Wikidata. *Communications of the ACM*, **57**, 78-85. https://doi.org/10.1145/2629489

[17] Chen, H., Hu, N., Qi, G., Wang, H., Bi, Z., Li, J., *et al.* (2021) OpenKG Chain: A Blockchain Infrastructure for Open Knowledge Graphs. *Data Intelligence*, **3**, 205-227. https://doi.org/10.1162/dint_a_00095

[18] Venugopal, V. and Olivetti, E. (2024) Matkg: An Autonomously Generated Knowledge Graph in Material Science. *Scientific Data*, **11**, Article No. 217. https://doi.org/10.1038/s41597-024-03039-z

[19] Sun, K., Yu, S., Peng, C., Wang, Y., Alfarraj, O., Tolba, A., *et al.* (2022) Relational Structure-Aware Knowledge Graph Representation in Complex Space. *Mathematics*, **10**, Article 1930. https://doi.org/10.3390/math10111930

[20] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A., Kaiser, L. and Polosukhin, I. (2017) Attention Is All You Need. 31*st Conference on Neural Information Processing Systems* (*NIPS* 2017), Long Beach, 5998-6008.

[21] Hu, J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S. and Chen, W. (2021) LoRA: Low-Rank Adaptation of Large Language Models. ArXiv, abs/2106.09685.

[22] Wen, J., Thibeau-Sutre, E., Diaz-Melo, M., Samper-González, J., Routier, A., Bottani, S., *et al.* (2019) Convolutional Neural Networks for Classification of Alzheimer's Disease: Overview and Reproducible Evaluation. *Medical Image*

Analysis, **63**, Article 101694. https://doi.org/10.1016/j.media.2020.101694

[23] Zeng, A., Xu, B., Wang, B., *et al.* (2024) ChatGLM: A Family of Large Language Modelsfrom GLM-130B to GLM-4 All Tools. arXiv:2406.12793.

[24] Yang, A., *et al.* (2025) Qwen3 Technical Report. ArXiv, abs/2505.09388.

[25] Guo, D., Yang, D., Zhang, H., *et al.* (2025) Deepseek-R1: Incentivizing Reasoningcapability in Llms via Rein-Forcement Learning. arXiv:2501.12948.

[26] Meta (2024) Introducing Llama 3.1: Our Most Capable Models to Date. https://ai.meta.com/blog/meta-llama-3-1/

[27] Gemma Team (2024) Gemma 2: Improving Open Language Models at a Practical Size. ArXiv, abs/2408.00118.