

同行专家业内评价意见书编号： 20250854468

附件1

浙江工程师学院（浙江大学工程师学院） 同行专家业内评价意见书

姓名： 赵佳乐

学号： 22260313

申报工程师职称专业类别（领域）： 电子信息

浙江工程师学院（浙江大学工程师学院）制

2025年05月27日

填表说明

一、本报告中相关的技术或数据如涉及知识产权保护、军工项目保密等内容，请作脱密处理。

二、请用宋体小四字号撰写本报告，可另行附页或增加页数，A4纸双面打印。

三、表中所涉及的签名都必须用蓝、黑色墨水笔，亲笔签名或签字章，不可以打印代替。

四、同行专家业内评价意见书编号由工程师学院填写，编号规则为：年份4位+申报工程师职称专业类别(领域)4位+流水号3位，共11位。

一、个人申报

(一) 基本情况【围绕《浙江工程师学院(浙江大学工程师学院)工程类专业学位研究生工程师职称评审参考指标》，结合该专业类别(领域)工程师职称评审相关标准，举例说明】

1. 对本专业基础理论知识和专业技术知识掌握情况(不少于200字)

在临床试验资格标准信息抽取项目的专业实践中，我显著提升了自身在专业基础理论知识和技术知识方面的掌握程度。该项目融合了医学、计算机科学等多个领域的知识，通过实践探索，我学会了如何有效整合这些看似独立的学科知识，为解决复杂问题提供了重要的方法论支撑。

1. 信息抽取领域专业知识：

基础理论学习：通过阅读大量信息抽取领域的综述报告，系统地了解了该领域内的主要方法、技术流派和发展趋势。针对临床试验资格标准信息抽取的具体任务，深入调研了相关的学术论文，掌握了该任务的特定要求、挑战和现有解决方案。

模型理解与应用：

通过复现经典的信息抽取任务和方法，深入理解了信息抽取模型的内部结构、工作原理和关键参数。将所学知识应用于临床试验资格标准信息抽取项目，提升了模型在特定任务上的性能。

学术表达能力提升：学习并实践了科研结果的撰写方法，包括实验设计、结果分析和论文撰写等，提高了学术表达能力。

2. 跨学科知识融合能力：

医学知识：

在项目实践中，我深入学习了与临床试验资格标准相关的医学知识，包括医学术语、临床流程等，为信息抽取任务提供了必要的领域知识支持。

计算机科学知识：

进一步巩固了计算机科学领域的基础知识，特别是自然语言处理、机器学习和深度学习等方面的知识，并将其应用于信息抽取模型的开发和应用中。

知识融合方法：

通过实践探索，掌握了将医学、计算机科学等多个领域知识进行有效融合的方法，提升了解决复杂问题的能力。

通过临床试验资格标准信息抽取项目的专业实践，我系统地掌握了信息抽取领域的基础知识和前沿动态，并提升了跨学科知识融合的能力。

2. 工程实践的经历(不少于200字)

在临床试验资格标准信息抽取项目中，我针对召回率低和分布抽取问题，设计并实现了基于多轮问答的BERT-BiLSTM-CRF模型。

该模型融合了BERT的语义理解、BiLSTM的上下文捕捉和CRF的序列标注优势，通过多轮问答机制，有效解析复杂文本中的实体及关系。我基于BERT引入BiLSTM和CRF层以增强长距离依赖建模和标签优化。为处理实体名称的精准边界，我设计了注意力机制驱动的多轮问答框架，迭代细化抽取结果。模型性能显著提升，召回率提高7%，F1达91%，已形成专利并受理中。此实践锻炼了我的工程技术能力，深化了对自然语言处理应用挑战与机遇的理解，为未来研究和工作奠定了坚实基础。

3. 在实际工作中综合运用所学知识解决复杂工程问题的案例(不少于1000字)

首先，在参与临床试验资格标准信息抽取项目的过程中，我深入分析了现有信息抽取技术在实际应用中面临的挑战，特别是召回率不足和流水线式抽取误差累积的问题。为了解决这些问题，我提出并实现了一种基于多轮问答机制的BERT-BiLSTM-CRF模型。该模型结合了预训练语言模型BERT强大的语义理解能力、双向长短期记忆网络BiLSTM的上下文捕捉能力以及条件随机场CRF的序列标注能力，通过精心设计的多轮问答流程，能够有效地解析复杂临床试验资格标准文本中的关键信息。

我引入BiLSTM层来增强模型对长距离依赖关系的建模能力。BiLSTM是一种双向循环神经网络，能够有效地捕捉文本中的上下文信息。在临床试验资格标准文本中，许多关键信息可能分布在较长的句子或段落中，BiLSTM能够通过其双向注意力机制，捕捉到这些长距离的依赖关系，从而提高信息抽取的准确性。我设计了一种基于多轮问答的联合抽取框架，在这个框架中，模型通过多轮问答的方式，逐步抽取出可能存在的实体，每一轮问答都基于前一轮的结果，并利用注意力机制聚焦于当前最相关的文本片段。通过这种方式，模型能够有效地解析复杂文本结构，并提高召回率。

实验结果表明，该模型在临床试验资格标准信息抽取任务上取得了显著的性能提升，召回率较传统方法提高了7%，F1达到了91%，大大提高了信息抽取的效率和准确性。这一结果不仅验证了模型的有效性，也为后续的产业化应用奠定了基础。

这一创新成果已经形成专利申请，并正在积极推动其产业化应用。该技术的应用将显著降低临床试验信息处理的成本，提高临床试验的入组效率，为医疗科研和药物开发提供强有力的技术支持，预期将产生显著的经济效益和社会效益。

此外，在抽取出资格标准实体后，需要对实体名称进行标准化映射，为临床试验匹配等工程奠定研究基础。具体来说，在构建双语层级化术语词表的过程中，我面临了诸多挑战。首先，肿瘤资格标准中存在大量的专业术语和复杂的层级关系，如何准确地提取和表示这些信息是一个难题。中英文术语之间存在着语义鸿沟，如何建立准确的对应关系也是一个挑战。为了解决这些问题，我采取了以下步骤：

1、英文层级化术语词表的构建：我利用开源的NCIt术语词表作为基础，结合合作医院提供的数据库，构建了针对不同字段的英文层级化术语词表。NCIt是一个权威的医学术语词表，包含了大量的肿瘤相关术语和它们的层级关系。我根据肿瘤资格标准的特点，对NCIt中的术语进行了筛选和整理，并补充了一些新的术语。同时，我还利用合作医院的数据，对术语词表进行了验证和扩展，确保其覆盖了肿瘤资格标准中的关键层级概念。

2、双语术语映射：为了实现中英文术语之间的准确映射，我采用了多种方法。首先，我利用谷歌翻译对英文术语进行翻译，谷歌翻译在专业医学术语翻译方面具有较高的精度。然后，我利用有道词典对翻译结果进行验证和补充，有道词典在通用术语名称提供方面具有广泛性。此外，我还结合了人工标注的中英文肿瘤资格标准实例，对映射结果进行人工校对和修正，确保映射的准确性。

3、术语扩展词表的构建：为了进一步提高术语词表的覆盖率和准确性，我利用大量人工标注的中英文肿瘤资格标准实例，开发了双语层级化术语扩展词表。我通过分析标注实例中的术语使用情况，提取出一些新的术语和它们的层级关系，并将其添加到术语扩展词表中。

4、评估与优化：为了评估双语层级化术语词表的性能，我进行了两项实验。首先，我在人工标注的实体集上分别对中文层级化术语词表和英文层级化术语词表的映射精度进行了测试。实验结果表明，双语层级化术语词表在映射精度方面显著优于现有的通用术语词表。其次，我进行了一项消融实验，以检验双语层级化术语扩展词表中的层级概念对试验匹配效果的

提升。实验结果一致显示，双语层级化术语词表及其术语扩展词表在试验匹配任务中显著提升了匹配的有效性。

通过这一项目的实践，我不仅提升了自己的工程技术能力，也对自然语言处理在实际应用中的挑战和机遇有了更深刻的认识。

(二) 取得的业绩(代表作) 【限填3项, 须提交证明原件(包括发表的论文、出版的著作、专利证书、获奖证书、科技项目立项文件或合同、企业证明等) 供核实, 并提供复印件一份】

1. 公开成果代表作 【论文发表、专利成果、软件著作权、标准规范与行业工法制定、著作编写、科技成果获奖、学位论文等】

成果名称	成果类别 [含论文、授权专利(含发明专利申请)、软件著作权、标准、工法、著作、获奖、学位论文等]	发表时间/ 授权或申请时间等	刊物名称/ 专利授权或申请号等	本人排名/ 总人数	备注
基于多轮问答的临床试验资格标准中实体关系自动抽取方法和系统	发明专利申请	2024年08月05日	申请号: 2024110649125	2/4	已进入实质审查

2. 其他代表作 【主持或参与的课题研究项目、科技成果应用转化推广、企业技术难题解决方案、自主研发设计的产品或样机、技术报告、设计图纸、软课题研究报告、可行性研究报告、规划设计方案、施工或调试报告、工程实验、技术培训教材、推动行业发展中发挥的作用及取得的经济社会效益等】

(三) 在校期间课程、专业实践训练及学位论文相关情况	
课程成绩情况	按课程学分核算的平均成绩: 85 分
专业实践训练时间及考核情况(具有三年及以上工作经历的不作要求)	累计时间: 1 年(要求1年及以上) 考核成绩: 83 分
本人承诺	
<p>个人声明: 本人上述所填资料均为真实有效, 如有虚假, 愿承担一切责任, 特此声明!</p> <p>申报人签名: 赵佳乐</p>	

工作单位考核评价：
合格
(公章) 刘厚松 2015 年 1 月

<p>日常表现考核评价</p>	<p>非定向生由德育导师考核评价、定向生由所在工作单位考核评价：</p> <p> <input checked="" type="checkbox"/> 优秀 <input type="checkbox"/> 良好 <input type="checkbox"/> 合格 <input type="checkbox"/> 不合格 </p> <p>德育导师/定向生所在工作单位分管领导签字（公章）  年 月 日</p>
<p>申报材料审核公示</p>	<p>根据评审条件，工程师学院已对申报人员进行材料审核（学位课程成绩、专业实践训练时间及考核、学位论文、代表作等情况），并将符合要求的申报材料在学院网站公示不少于5个工作日，具体公示结果如下：</p> <p> <input type="checkbox"/> 通过 <input type="checkbox"/> 不通过（具体原因： </p> <p>工程师学院教学管理办公室审核签字（公章）： 年 月 日</p>

浙江大学研究生院
攻读硕士学位研究生成绩表

学号：22260313		姓名：赵佳乐		性别：女		学院：工程师学院			专业：电子信息			学制：2.5年			
毕业时最低应获：24.0学分				已获得：26.0学分					入学年月：2022-09			毕业年月：			
学位证书号：					毕业证书号：					授予学位：					
学习时间		课程名称		备注	学分	成绩	课程性质	学习时间		课程名称		备注	学分	成绩	课程性质
2021-2022学年春季学期		研究生英语基础技能			1.0	免修	公共学位课	2022-2023学年冬季学期		产业技术发展前沿			1.5	88	专业学位课
2021-2022学年春季学期		研究生英语			2.0	免修	公共学位课	2022-2023学年冬季学期		生理系统仿真与建模			2.0	86	专业选修课
2022-2023学年秋季学期		现代信号处理技术与应用			1.0	88	专业学位课	2022-2023学年春季学期		智能医疗技术应用案例分析			1.0	98	专业学位课
2022-2023学年秋季学期		新时代中国特色社会主义思想理论与实践			2.0	88	公共学位课	2022-2023学年春夏学期		工程伦理			2.0	95	公共学位课
2022-2023学年秋季学期		工程技术创新前沿			1.5	84	专业学位课	2022-2023学年夏季学期		自然辩证法概论			1.0	86	公共学位课
2022-2023学年冬季学期		生物医学工程方法学			1.0	88	专业学位课	2022-2023学年夏季学期		研究生论文写作指导			1.0	92	专业学位课
2022-2023学年秋冬学期		高阶工程认知实践			3.0	78	专业学位课	2023-2024学年秋冬学期		数据分析的概率统计基础			3.0	61	专业选修课
2022-2023学年冬季学期		智能化仪器软硬件系统设计与应用			1.0	87	专业学位课			硕士生读书报告			2.0	通过	

说明：1. 研究生课程按三种方法计分：百分制，两级制（通过、不通过），五级制（优、良、中、及格、不及格）。

2. 备注中“*”表示重修课程。

学院成绩校核章：
成绩校核人：张梦依
打印日期：2025-06-03



国家知识产权局

310013

浙江省杭州市西湖区竞舟路1号筑品金座501室 杭州天勤知识产权代理有限公司
曹兆霞(0571-87755912)

发文日:

2024年08月05日



申请号: 202411064912.5

发文序号: 2024080501210130

专利申请受理通知书

根据专利法第28条及其实施细则第43条、第44条的规定,申请人提出的专利申请已由国家知识产权局受理。现将确定的申请号、申请日等信息通知如下:

申请号: 2024110649125

申请日: 2024年08月05日

申请人: 浙江大学

发明人: 吕旭东,赵佳乐,陈辉,段会龙

发明创造名称: 基于多轮问答的临床试验资格标准中实体关系自动抽取方法和系统

经核实,国家知识产权局确认收到文件如下:

权利要求书 1份 4页,权利要求项数: 10项

说明书 1份 17页

说明书附图 1份 3页

说明书摘要 1份 1页

专利代理委托书 1份 2页

发明专利请求书 1份 5页

实质审查请求书 文件份数: 1份

申请方案卷号: 24162F1480

提示:

1.申请人收到专利申请受理通知书之后,认为其记载的内容与申请人所提交的相应内容不一致时,可以向国家知识产权局请求更正。

2.申请人收到专利申请受理通知书之后,再向国家知识产权局办理各种手续时,均应当准确、清晰地写明申请号。

审查员: 自动受理

联系电话: 010-62356655

审查部门: 初审及流程管理部



200101
2023.03

纸件申请,回函请寄: 100088 北京市海淀区前门桥西土城路6号 国家知识产权局专利局受理处收
电子申请,应当通过专利业务办理系统以电子文件形式提交相关文件。除另有规定外,以纸件等其他形式提交的文件视为未提交。