自然语言处理技术在临床检查合理性质控中的应用研究

尹丽丽 范正磊 徐 畅 张 群 王 浩 韩 影
徐州仁慈医院 江苏 徐州 221000

摘 要 目的：针对临床检查中多开、漏开、错开及禁忌证检查等不合理现象，构建基于自然语言处理的临床检查合理性校验系统，实现检查单开具的实时智能审核。方法：以徐州仁慈医院2023年185 137人次检查数据为研究对象，联合多部门制订临床检查合理性规则库。采用临床BERT模型，即临床双向编码器表示模型（clinical bidirectional encoder representations from transformers，Clinical BERT），从病历文本中自动提取患者关键信息，基于Profile架构开发校验平台并嵌入电子病历系统，实现患者信息与规则库的实时匹配校验。结果：Clinical BERT模型实体识别F1值达94.7%，系统响应时间1.8 s，规则匹配精准率92.3%。系统应用后，试点科室检查不合理率降幅达63.2%。其中多开检查减少41.3%，漏开检查减少35.8%，禁忌证检查实现100%拦截。结论：基于自然语言处理（natural language processing，NLP）的临床检查合理性校验系统可有效识别病历关键信息，实现检查单实时智能校验，显著降低临床检查不合理率，具有良好推广应用价值。

关键词 自然语言处理；临床检查；合理性质控；Clinical BERT；实时校验

Research on the Application of Natural Language Processing Technology in the Rationality Quality Control of Clinical Examinations

Yin Lili, Fan Zhenglei, Xu Chang, Zhang Qun, Wang Hao, Han Ying
Xuzhou Renci Hospital, Xuzhou 221000, Jiangsu Province, China

Abstract Objective: To address unreasonable clinical examinations such as over-ordering, under-ordering, wrong-ordering, and examinations with contraindications, a clinical examination appropriateness verification system based on natural language processing was constructed to achieve real-time intelligent review of examination orders. Methods: A total of 185 137 examination records from Xuzhou Renci Hospital in 2023 were selected as the research subjects, and a clinical examination appropriateness rule base was developed jointly by multiple departments. The clinical bidirectional encoder representations from transformers (Clinical BERT) model was used to automatically extract key patient information from medical record texts. A verification platform was developed based on the Profile architecture and embedded into the electronic medical record system to achieve real-time matching and verification between patient information and the rule base. Results: The entity recognition F1 score of the Clinical BERT model reached 94.7%, the system response time was 1.8 s, and the rule matching precision was 92.3%. After system application, the unreasonable examination rate in pilot departments decreased by 63.2%. Among them, over-ordering of examinations decreased by 41.3%, under-ordering decreased by 35.8%, and examinations with contraindications achieved 100% interception. Conclusion: The clinical examination appropriateness verification system based on natural language processing (NLP) can effectively identify key information in medical records, achieve real-time intelligent verification of examination orders, and significantly reduce the rate of unreasonable clinical examinations, demonstrating good value for promotion and application.

Key words Natural language processing; Clinical examination; Appropriateness quality control; Clinical BERT; Real-time verification

临床检查是医学诊断的重要环节，涵盖一般类、检验类和医技类。当前检查合理性问题突出：20%～50%的高级影像学检查可能对患者无价值，不仅浪费医疗资源，还可能因辐射和造影剂带来风险[1]；过度、重复及禁忌证检查还加重患者经济负担并威胁安全。以徐州仁慈医院为例，四肢相关检查项目有220项，2023年相关检查达185 137人次。部分医生因系统操作失误、病情把握疏忽、忽视患者禁忌证等原因，出现多开、漏开、错开检查的情况。临床决策支持系统（clinical decision

support system，CDSS）结合NLP已被证实可提升决策准确性、优化诊疗流程[2]，但德国三所大学医院的多中心整群随机试验显示，ESR iGuide系统并未显著减少不适当影像学检查请求，说明单纯引入CDSS不足以保证有效性，系统设计、集成方式及与临床工作流程的融合程度至关重要[3]。Jiang-Kells等[4]开发的MiADE系统在临床实时应用中实现83.2%的精确率和85.2%的召回率，并将医生录入结构化问题列表的时间减少89%。综上，基于NLP的临床检查合理性校验系统具有重要研究价值，但其有效性高度依赖与临床工作流程的深度融合、用户体验优化及规则库科学构建。

1 临床检查合理性校验系统设计

1.1 系统总体架构

本研究设计的系统旨在实现实时性、准确性、全面性、可扩展性和易集成性五大目标，采用数据层、处理层和应用层三层架构。数据层负责存储电子病历数据库、医院信息系统（hospital information system，HIS）数据库和规则库；处理层包含NLP预处理、实体识别、规则匹配引擎及Profile并行架构；应用层包括医生工作站、检查单开具模块和提示反馈模块。数据流向为：医生开具检查单时自动触发校验流程，系统抓取病历文本，经预处理后使用Clinical BERT识别关键医学实体，将患者信息与规则库匹配，生成校验结论并实时反馈给医生工作站，同时记录校验日志。

1.2 临床检查合理性规则库构建

规则库由多部门联合制订：临床医技科室提供专业意见，医务处把关医疗规范，质控办提供质控标准，信息处负责技术支持。规则来源包括国内外权威指南、院内专家共识、历史不合理检查案例及药品说明书禁忌证信息。国内研究指出，CDSS开发需临床专家与技术人员密切协作，通过多次迭代完善规则库[5]。规则分为禁忌证、多开、漏开和重复检查四类：禁忌证规则识别检查禁忌（如碘过敏患者不宜开增强CT）；多开规则判断检查是否与病情无关；漏开规则提示补充必要检查；重复检查规则避免短期内重复相同检查。研究显示，基于规则引擎的CDSS能有效减少不合理检查[6]。规则采用JSON格式存储，包含规则唯一标识、类型、关联检查项目、触发条件、处理动作、提示信息、优先级和有效期等字段，便于解析和动态更新。

1.3 技术指标

系统需满足：响应时效不超过2 s；实体识别准确率不低于95%；规则匹配精准率不低于90%；规则覆盖率达到100%；系统应用后临床检查不合理率降幅不低于50%，禁忌证检查实现100%拦截。

2 基于NLP的病情结构化提取

2.1 医学文本预处理

病历文本包含大量医学术语和特定表达，需预处理以提高实体识别准确性。预处理流程包括文本清洗、分词、词性标注、去除停用词和词形还原。

文本清洗去除无关字符和格式信息，如标点规范化、数字格式统一、全半角转换。分词采用医学领域优化词典，包含5万余条医学术语，确保“股骨颈骨折”“冠状动脉粥样硬化”等专业词汇正确切分。词性标注为每个词语标注词性，名词通常对应实体名称，动词描述症状或检查过程。去除停用词剔除“的”“了”等无意义词语，词形还原将不同形态转换为基本形式，如“疼痛感”“疼痛”统一为“疼痛”。

2.2 Clinical BERT实体识别模型

1.模型选择与原理。医学实体识别是从病历文本中识别症状、诊断、检查项目、禁忌证等关键信息，是系统核心环节。本研究选择Clinical BERT作为核心模型，该模型在BERT基础上使用MIMIC-Ⅲ临床记录继续预训练，专门适用于临床文本处理。其优势在于：双向上下文理解、领域适配性强、迁移学习能力好，只需少量标注数据即可微调获得良好效果。

2.实体类型定义。本研究识别五类医学实体：症状实体如“发热”“左下肢疼痛”；诊断实体如“肺炎”“股骨颈骨折”；检查项目如“血常规”“胸部CT”；禁忌证如“碘过敏”“妊娠”；检查结果如“白细胞升高”“骨折线清晰”。

3.模型微调。微调采用交叉熵损失函数，优化器使用AdamW，学习率2e-5，批次大小16，训练轮数5轮。采用早停机制防止过拟合，Dropout率设置为0.1增强泛化能力。

2.3 实验结果与分析

1.数据集构建。数据集来自徐州仁慈医院2023年电子病历系统，共收集病历文本10 000份，按7∶2∶1划分为训练集、验证集和测试集。标注实体共计85 600个，其中症状实体25 600个，诊断实体21 400个，检查项目18 200个，禁忌证9 800个，检查结果10 600个。

2.实验结果。采用精确率、召回率和F1值作为评价指标。将Clinical BERT与通用BERT、BiLSTM-CRF和BioBERT进行对比实验，Clinical BERT的F1值达到94.7%，显著优于通用BERT的82.5%和BiLSTM-CRF的86.3%，也高于BioBERT的88.7%。不同实体类型识别效果：检查项目最佳F1值96.3%，症状95.2%，诊断94.8%，禁忌证93.5%，检查结果92.8%。检查项目识别效果最

好因其名称相对固定，检查结果识别难度较大因其描述多样。

3 实时校验引擎与系统实现

3.1 规则引擎设计

规则引擎是系统决策核心，负责将患者结构化信息与规则库匹配并生成校验结果。本研究采用自定义轻量规则引擎，规则匹配基于Rete算法优化实现，流程分三步：按诊断筛选规则、按症状和检查史精确匹配、按禁忌证安全校验。匹配采用优先级排序，禁忌证规则最高，拦截类次之，提示类最低。规则以JSON格式存储，支持动态加载和热更新，定时任务每分钟检查变更，过期规则自动失效。

3.2 Profile并行架构实现

系统采用Profile业务架构实现并行处理，核心是数据平台与业务逻辑分离。医生开具检查单时，系统从数据平台获取患者信息，同时启动实体识别、规则匹配、提示生成、日志记录等多线程并行处理，整体响应时间不随业务增加而线性增长。经测试，并行架构下系统平均响应时间为1.8 s，低于2 s目标阈值。

3.3 系统集成与界面实现

系统采用嵌入式软件开发工具包（software development kit，SDK）方式与HIS和电子病历系统集成，无须大规模改造。数据层通过视图访问HIS和电子病历（electronic medical record，EMR）数据库；服务层部署独立校验服务；表现层在医生工作站嵌入校验结果展示区。校验结果以非侵入方式呈现：符合规则的开单不作干预提示；提示类显示黄色警告图标，允许医生忽略并继续；拦截类弹出提示框，须医生确认并自动留痕。提示信息简洁明了，如“患者对碘过敏，不宜开具增强CT”等。

3.4 系统性能测试

性能测试在医院测试环境进行，模拟100个医生工作站并发访问，使用5 000份脱敏真实病历。结果显示，系统平均响应时间1.8 s，95%请求在2.1 s内完成，99%在2.5 s内完成，满足设计目标。回溯1 000份历史检查单，系统触发提示326次，正确301次，误报25次，准确率92.3%，高于90%目标；误报主要集中在多开规则。模拟100个医生同时在线，吞吐量达每秒35个请求，CPU使用率约65%，内存约4.5 G，未出现超时或崩溃；并发增至200个时响应时间升至3.2 s，建议支持并发数在150以内。

4 结束语

本研究选取徐州仁慈医院骨科、急诊科和内科进行为期6个月的试点。应用前3个科室平均检查不合理率为8.7%，应用后降至3.2%，降幅达63.2%；禁忌证检查实现100%拦截，未发生一例禁忌证检查开具[7]。86名医生问卷调查显示，92%认为系统对临床工作有帮助，96%认为不影响正常工作节奏。系统提示总采纳率为76.5%，其中拦截类提示采纳率达100%。系统优势体现在实时性强、准确性高、覆盖面广，将Clinical BERT与规则引擎有机结合，兼顾深度学习文本理解能力与规则系统的可解释性和可控性。与德国MIDAS试验相比，本研究干预效果更显著，原因在于更深度集成临床工作流程、多部门联合制订规则库、注重用户体验和非侵入式设计。尽管成效良好，系统仍存在规则覆盖不完整、模型泛化能力有限、提示信息层次不足、少数医生接受度不高等问题，后续将通过持续积累数据、多中心训练、优化提示设计和加强培训逐步改进。总体而言，系统具有明确临床价值：对患者减少不必要检查费用和辐射暴露，对医生提供实时决策支持，对医院规范医疗行为、降低资源浪费、提升整体医疗质量和管理水平[8]。

参考文献

[1]van Schorjenstein Lantman M. Clinical decision-making suffers from inequivalent measurement results and inadequate reference intervals[J]. Clinical Chemistry, 2024, 70(11): 1383-1392.

[2]Healthcare Informatics Research Group. Named entity recognition in electronic health records: a methodological review[J].Healthcare Informatics Research, 2023, 29(4): 286-300.

[3]Stroeder R, Bunk S, Gietzen C, et al. Challenges and insights from implementing clinical decision support systems for radiologic imaging: experience from the MIDAS trial[J].Insights into Imaging, 2025, 16: 152.

[4]Jiang-Kells J, Brandreth J, Zhu L, et al. Design and implementation of a natural language processing system at the point of care: MiADE (medical information AI data extractor)[J].BMC Medical Informatics and Decision Making,2025,25(1):365.

[5]王秀丽,李静,张伟.基于临床决策支持系统的检查合理性干预效果评价[J].中国数字医学,2022,17(5):32-36.

[6]张芳,王艳萍,刘敏.临床检查合理性评价指标体系构建研究[J].中国医院管理,2021,41(3):45-48.

[7]Domingues N S. A hybrid decision support system using rule-based and AI methods: the OnCATs knowledge-based framework[J]. International Journal of Medical Informatics, 2025, 186: 105678.

[8]Luxenburg O. Evaluating the accuracy and impact of the ESR-iGuide decision support tool in optimizing CT imaging referral appropriateness[J]. Journal of Imaging Informatics in Medicine, 2024, 37(4): 512-520.

（本文编辑 王 冬）