

居家护理场景下用户护理需求命名实体识别研究*

张卓越¹ 杨天赋² 左美云¹

(¹ 中国人民大学信息学院 北京 100872 ² 华南师范大学阿伯丁人工智能与数据科学学院 广州 510631)

[摘要] **目的/意义** 探讨应用深度学习模型在居家护理场景识别用户需求文本中的护理需求实体,以期通过自动化手段精准识别用户需求,为提升居家护理服务的效率和质量提供技术支持。**方法/过程** 选取 560 条用户护理需求文本客观数据,基于《国际功能、残疾和健康分类》对文本中的护理需求实体进行分类标注,采用 BERT-BiLSTM-CRF 模型进行实体识别,通过消融实验验证模型效果,分析实验结果,评估模型性能。**结果/结论** BERT-BiLSTM-CRF 模型实体级别微平均准确率、召回率、F1 值分别为 0.752 9、0.775 8、0.764 2,表明该模型可以为居家护理场景下自动化挖掘用户需求、优化护理服务流程和提高护理质量提供有力支持。

[关键词] 居家护理;命名实体识别;自然语言处理;护理需求

[中图分类号] R-058 **[文献标识码]** A **[DOI]** 10.3969/j.issn.1673-6036.2024.12.012

Study on Named Entity Recognition of User Care Needs in Home Healthcare Scenarios

ZHANG Zhuoyue¹, YANG Tianfu², ZUO Meiyun¹

¹ School of Information, Renmin University of China, Beijing 100872, China; ² Aberdeen School of Artificial Intelligence and Data Science, South China Normal University, Guangzhou 510631, China

[Abstract] **Purpose/Significance** To explore deep learning models can identify care needs entities in user needs texts in home healthcare scenarios, so as to accurately identify user needs through automated means to provide technical support for improving the efficiency and quality of home care services. **Method/Process** 560 pieces of objective data of user care needs are selected to classify and label the care needs entities in the texts based on the International Classification of Functioning, Disability and Health. The BERT-BiLSTM-CRF model is used to perform the entity recognition task. The effect of the model is verified by the ablation experiment, the experimental results are analyzed, and the performance of the model is evaluated. **Result/Conclusion** The model's precision, recall, and F1-score at the entity-level micro-average are 0.752 9, 0.775 8, and 0.764 2, respectively, indicating that the model can provide strong support for automatic mining of user needs, optimization of care service process and improvement of care quality in home healthcare scenarios.

[Keywords] home healthcare; named entity recognition (NER); natural language processing (NLP); care needs

[修回日期] 2024-06-08

[作者简介] 张卓越,博士研究生,发表论文 2 篇;通信作者:左美云,教授,博士生导师。

[基金项目] 国家社会科学基金重大项目(项目编号:22&ZD328)。

1 引言

老年居家护理,旨在为具有养老护理服务需求的用户(通常为失能、慢性病、高龄、残疾、出院后仍需持续医疗照护的老年人)提供专业化的上门护理服务^[1]。随着我国人口老龄化趋势不断加剧,国内对于居家护理服务的需求大幅提升^[2]。挖掘居家护理用户未被满足的护理需求,围绕用户需求设计解决方案,正成为紧迫的公共卫生优先事项^[3]。挖掘居家护理用户的护理需求,必须从用户的角度确定和了解护理与需求偏好。国内已有多项居家护理用户的护理需求研究^[4-7],其通常采用访谈或者问卷调查的方式挖掘不同地区不同疾病的用户护理需求,往往需要大量人力资源和时间成本收集并处理数据,且研究人员设计的问卷和访谈也可能受到主观偏见的影响^[8]。命名实体识别(named entity reorganization, NER)任务能够自动化处理并提取文本中的实体信息,不仅避免依赖研究人员的主观判断,减少主观偏见的影响,而且能够近乎实时地赋能居家护理机构,优化平台护理流程。

医学领域命名实体识别任务的目标是从给定文本(通常为临床电子病历、医疗文本)中识别与临床相关的症状名称、药物名称等实体^[9],该任务在临床实践中具有应用价值^[10]。在医学信息抽取任务中,BERT-BiLSTM-CRF模型^[11]是实体识别有效性最优的模型之一。其中,双向编码器表征(bidirectional encoder representations from transformers, BERT)^[12]对上下文信息全局建模的能力能够帮助模型更好地理解实体在文本中的语义。双向长短期记忆网络(bidirectional long short-term memory, BiLSTM)^[13]能够捕捉序列数据长距离依赖关系,增强特征表示。条件随机场(conditional random fields, CRF)能够建模序列标注任务中不同标签间的依赖关系,提高识别任务的准确性和一致性。因此,采用BERT-BiLSTM-CRF模型进行居家场景下的用户护理需求识别。

2 《国际功能、残疾和健康分类》评估框架

《国际功能、残疾和健康分类》^[14](International Classification of Functioning, Disability and Health, ICF)是由世界卫生组织(World Health Organization, WHO)发布的健康和功能评估框架。其目的是识别、描述和评估健康和与健康相关的个体状态,被WHO推荐用于居家护理场景下的用户护理需求识别和评定工作^[15]。其基于生物-心理-社会模型定义了身体功能、身体结构、活动和参与、环境因素4部分组件,适用于不同人群患者康复全周期^[16]。

ICF框架是康复医学^[17-18]、老年病学^[3,16]研究中识别各类患者医疗保健需求的分析工具,故选取其作为居家护理场景中识别患者护理需求的分析框架。区别于WHO发布的《国际疾病分类》,ICF不包含疾病、症状、病因等代表患者健康状况的概念术语。然而,患者的健康状况(如患病名称、治疗方案、症状)影响其护理需求。因此,添加健康状况维度,作为除ICF 4个分类维度之外的第5个护理需求维度。

综上,本研究护理需求的5个维度分别为健康状况、身体功能、身体结构、活动和参与、环境因素。各维度定义为:健康状况指个体的患病名称、症状以及治疗方案;身体功能指人体系统的生理功能(包括心理功能),如听力、肢体活动能力等;身体结构指个体的身体解剖部位,如器官、四肢及其他组成部分;活动和参与指个体执行的行动任务以及对社会生活的参与,如吃饭、行走、聊天等;环境因素指个体所处环境的外部因素对用户造成的影响,如患者使用的设备、所接受的服务等。

3 研究方法

3.1 数据与数据标注

数据来自于互联网居家护理机构“一号护工”。用户可利用智能手机等访问该护理机构的线上平台,通过输入文本信息向平台发送自身护理需求,随后,该机构会根据需求文本中的关键词建立用户

画像，并向用户推荐合适的护理人员，使用户获得及时便捷的上门护理服务。

排除重复数据和质量较差的数据，共得到 560 条代表老年人护理需求的非结构化、脱敏文本客观数据。数据实体类别按照 5 个维度进行标注。标注团队由 1 名信息学（计算机应用技术专业）研究者与 2 名具有临床护理实践经验的护士共同完成。其中，2 名护士主要负责需求实体的分类问题，信息学研究者主要负责需求实体的粒度与边界问题。3 名标注人员使用 label - studio^[19] 标注平台进行标注，遵循 BIO 标注法和平台标注设置、数据导出规范，

确保标注的准确性和一致性。BIO 标注使用“B -”表示实体的起始位置字符，“I -”表示实体的中间位置字符，“O”表示非实体位置字符。身体功能、身体结构、活动和参与、环境因素和健康状况分别以“F”“S”“A”“E”“H”表示。在用户护理需求文本“倒地腿摔了，暂时轮椅、卧床需要女护理人员喂饭喂药”中，“腿”在 ICF 中属于“下肢结构”，单字被标注为“B - S”；“轮椅”在 ICF 中属于“用于人员室内外移动和交通的产品和技术”，起始字符和中间字符分别被标注为“B - A”和“I - A”，见图 1。

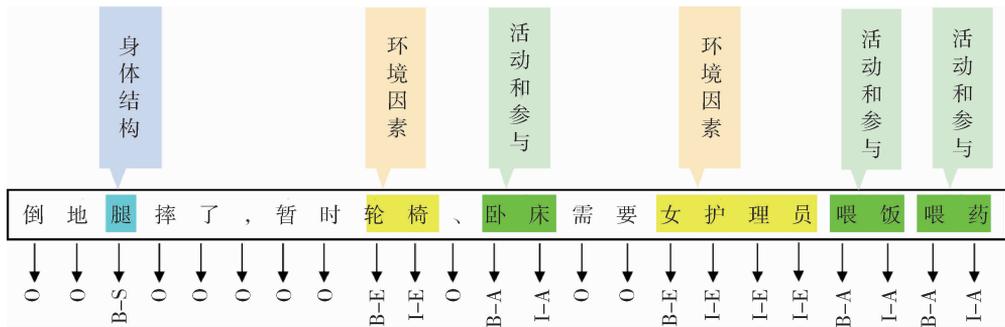


图 1 居家护理用户护理需求命名实体识别示例

3.2 模型介绍

采用 BERT - BiLSTM - CRF 模型，整体架构，见图 2。一是 BERT 预训练层。BERT 是基于 Transformer 架构的预训练语言表示模型，通过自注意力机制，同时考虑词语前后两侧的上下文信息。向 BERT 层输入序列中的每个字符 (token) 均会被转换为维度 768 的向量。二是 BiLSTM 层。BiLSTM 能够有效捕捉输入序列中上下文信息并处理长距离依赖关系。BiLSTM 接受从 BERT 层传递来的向量，在正向和逆向运行 LSTM 单元，根据当前字符的向量及前一个字符的隐藏状态，更新当前字符的隐藏状态。随后，BiLSTM 将正向和逆向的隐藏状态拼接。BiLSTM 层的维度设为 256。三是 CRF 层。CRF 作为一种概率图模型，能够对 BiLSTM 层输出的特征序列联合建模，在考虑标签之间的依赖关系后进行标签预测。CRF 层的维度为 512。CRF 层最终输出预测序列中每个字符的预测标签。

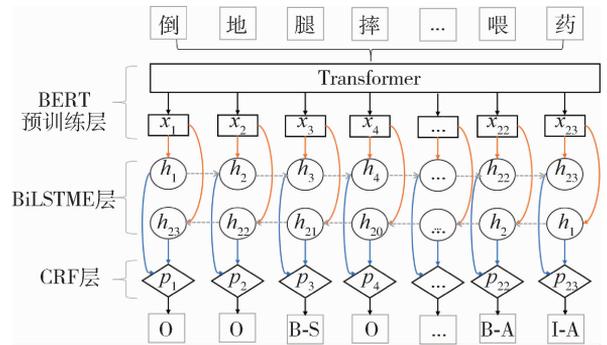


图 2 BERT - BiLSTM - CRF 模型

3.3 模型训练

本研究训练集、验证集和测试集的比例为 1:1:1。将输入文本与 BIO 标签送入模型训练。在每个训练批次结束后调用验证集对模型性能进行初步评估，如果当前训练批次在验证集上的损失值小于历史最佳损失值且当前训练批次在验证集中的准确率高于历史最佳准确率，则更新最佳模型为当前模型。完成训练阶段后，使用更新后的最佳模型对测试集的

数据进行实验，得到测试集预测序列中每个字符的预测 BIO 标签。为验证 BERT - BiLSTM - CRF 模型效果，分别将 BERT 预训练层与 BiLSTM 层去掉，进行消融实验。

PyTorch 可提供简洁的调用接口、直观的语法和庞大而活跃的社区支持，选用 PyTorch 深度学习框架构建实验模型^[20]。模型超参数包括：单训练批次样本数量 (batch_size)、学习率 (learning_rate)、训练批次 (epochs) 和最大文本长度 (max_sequence_length)。这些参数分别被设置为：64、0.001、200、256。选取 HuggingFace 社区中的 bert-base-Chinese 模型作为 BERT 预训练层的预训练语言模型。

3.4 评价指标

采用准确率 (precision)、召回率 (recall) 以及 $F1$ 值 ($F1$ -score) 的宏平均 (macro) 指标与微平均 (micro) 指标作为评价指标以全面评估命名实体识别模型的性能^[21]。在宏平均指标中，每个实体类别对评估指标的贡献相等；在微平均指标中，每个实体对评估指标的贡献相等。此外，采用基于字符 (token-level) 的度量和基于实体 (entity-level) 的度量两种性能指标方法。基于字符的度量方法将每个标记 (B-D, I-D 等) 与真实标签逐一比较，而基于实体的度量方法要求整个实体的起始到结束标记均被正确预测。

4 实验结果分析

本研究需求文本的文本长度最大为 73 字符，最小为 2 字符。在单个需求文本中被标记的实体数量最大为 8 个，最小为 1 个，平均出现的实体数量为 5.12 个。实验中各类别被标记的实体数量为：环境因素 1 159 个、活动和参与 856 个、健康状况 484 个、身体结构 193 个、身体功能 179 个。使用 BERT - BiLSTM - CRF 模型在型号为 RTX 3090Ti 的单个 GPU 上运行，平均单条需求文本实体抽取预测时间为 0.3 秒，远远短于人工查询 ICF、确认类型

并标注实体所需的时间 (根据 label-studio 工具统计，专家标注单条需求文本的平均时间为 228 秒)。这极大缩短了对用户进行健康管理构建用户画像或知识图谱的时间。

实体级别和字符级别的评价指标结果，见表 1。由于数据集存在实体类别分布不均衡问题，以环境因素和身体功能两类实体为例，实体数量相差 6.4 倍，环境因素的准确率较身体功能的准确率高近 20 个百分点。一般而言，训练的实体数量越多，模型预测效果越好。然而，尽管环境因素被标记的实体数量最多，但该类别未能表现出最优的预测效果。这可能是由于环境因素各实体的语义关系较为松散，包括为残疾用户提供的设备和技术 (e110—e199)、他人对用户的态度 (e410—e499) 以及外界对用户提供的服务 (e510—e599) 等，模型不能很好地学习环境因素类别中实体的特征。根据 $F1$ 值的预测结果，各实体类别预测效果从高到低分别为：活动和参与、环境因素、健康状况、身体功能和身体结构。

表 1 BERT - BiLSTM - CRF 模型在不同实体类别中的实验结果

分类	准确率	召回率	$F1$
健康状况	0.719 5	0.766 2	0.742 1
身体功能	0.571 4	0.545 5	0.558 1
身体结构	0.518 5	0.538 5	0.528 3
活动和参与	0.820 4	0.835 4	0.827 8
环境因素	0.760 6	0.786 4	0.773 3
实体级别微平均	0.752 9	0.775 8	0.764 2
实体级别宏平均	0.678 1	0.694 4	0.685 9
字符级别微平均	0.824 8	0.834 8	0.829 4
字符级别宏平均	0.764 3	0.773 4	0.767 6

为进一步了解模型性能，分析模型在护理需求识别任务中存在的不足。预测错误的抽样案例，见表 2，原因主要集中于实体边界预测问题。这与用户口语化表述的特征相关，如文本片段 1 中模型预测“足”为身体结构，而“足部”才是被标注的正确实体。另外，模型的查全率仍有提升空间，以文本片段 3 为例，模型未能识别“开颅手术”这类健康状况的实体。

表 2 错误示例分析

文本片段	标注说明	预测说明	结果分析
1. 足部溃烂	标注“足部”为“身体结构”	预测“足”为“身体结构”	模型在实体边界的预测中存在模糊性
2. 看护输液	标注“输液”为“环境因素”	预测“液”为“环境因素”	模型在实体边界的预测中存在模糊性
3. 开颅手术	标注“开颅手术”为健康状况	未能预测出实体	模型未能很好地学习到类似实体类型的特征

为评估每个组件的贡献，分别移除 BERT 层和 BiLSTM 层进行消融实验，使用相同数据集和训练参数以保证实验结果的可比性，见表 3。移除 BERT 层时，基于实体的微平均 $F1$ 值降低至 0.720 9，宏平均 $F1$ 值降至 0.606 5。BERT - CRF 模型的 $F1$ 值与召回率有所降低，但准确率与 BERT - BiLSTM -

CRF 基本持平。这突显了 BERT 在生成高质量嵌入方面的核心作用。当 BERT 层与 BiLSTM 层、CRF 层联合使用时，通过三者的协同作用能有效提升模型识别所有相关实体的能力。实验证明，BERT - BiLSTM - CRF 达到最佳的护理需求实体识别性能。

表 3 消融实验结果

度量方式	模型	微平均			宏平均		
		准确率	召回率	$F1$	准确率	召回率	$F1$
基于字符	BERT - BiLSTM - CRF	0.824 8	0.834 8	0.829 4	0.764 3	0.773 4	0.767 6
	BERT - CRF	0.814 9	0.803 5	0.808 7	0.754 9	0.740 7	0.745 9
	BiLSTM - CRF	0.800 1	0.772 3	0.784 6	0.727 8	0.655 4	0.685 0
基于实体	BERT - BiLSTM - CRF	0.752 9	0.775 8	0.764 2	0.678 1	0.694 4	0.685 9
	BERT - CRF	0.757 6	0.751 5	0.754 6	0.669 7	0.663 4	0.665 1
	BiLSTM - CRF	0.735 3	0.707 1	0.720 9	0.635 9	0.583 0	0.606 5

最后，对本研究中标注的用户需求进行排序。在共计 560 条数据中，最常被用户提及的 10 个护理需求如下。环境因素：女护理员（238 次）；活动和参与：做饭（145 次）、大小便（99 次）、自理（89 次）；环境因素：有护理经验（86 次）；活动和参与：卧床（76 次）；健康状况：脑梗死（66 次）；环境因素：男护理员（58 次）；活动和参与：半自理（55 次）；健康状况：手术（54 次）。在居家护理服务中，用户最常提及的护理需求主要集中在 ICF 的环境因素、活动和参与以及本研究增设的健康状况 3 个维度。用户对护理员的性别和经验有较高的关注度，特别是对女护理员和有护理经验的护理员的需求较为突出。上述信息为护工的培养提供有针对性的证据，对于提升护工的专业技能和改善机构服务质量具有重要意义。自动化护理需求抽取任务作为智能化健康管理的前置工作，对优化护理服务流程具有关键作用。

5 结语

本研究以真实居家护理平台的客观数据作为实验数据，采用 BERT - BiLSTM - CRF 模型深入探索居家护理场景下的护理需求命名实体识别。通过自动化的方式抽取居家护理用户的护理需求，不仅有效减少主观偏见对结果的影响，还实现了近乎实时的数据支持，为居家护理机构或平台构建企业级用户画像、知识图谱以及优化“护工 - 患者”匹配等下游任务提供强有力的决策依据。BERT - BiLSTM - CRF 模型借助 BERT 模块的语义建模能力、BiLSTM 模块的序列建模能力以及 CRF 模块的概率分布建模能力，实现了对用户护理需求的有效识别。

然而，本研究也存在一定的局限性。由于居家护理需求文本数据通常属于护理机构的内部订单数据，具有非公开属性，本研究仅针对单一护理预约

平台的数据进行需求挖掘, 数据量相对有限, 研究结果可能存在片面性。为克服这一局限, 未来研究可进一步获取多个平台甚至多个国家的数据, 以更准确地反映不同地区和平台用户的多样化需求, 从而提高研究的应用价值和实用性。此外, 针对居家护理场景数据资源稀缺的问题, 设计基于小样本学习的命名实体识别解决方案也是值得研究的方向, 可推动居家护理服务的智能化和个性化发展, 为老年人群提供更加精准、高效的护理服务。

作者贡献: 张卓越负责研究设计、实验实施、结果分析、论文撰写与修订; 杨天赋负责数据预处理、实验实施、论文修订; 左美云负责研究设计、研究管理、论文修订。

利益声明: 所有作者均声明不存在利益冲突。

参考文献

- 1 居家和社区医养结合服务指南(试行) [EB/OL]. [2023-11-01]. https://www.gov.cn/zhengce/zhengceku/202311/content_6914596.htm.
- 2 湛扬, 赵锐祗, 兰美娟, 等. “互联网+”居家护理质量评价指标体系的构建 [J]. 中国护理管理, 2022, 22 (11): 1722-1727.
- 3 ABDI S, SPANN A, BORILOVIC J, et al. Understanding the care and support needs of older people: a scoping review and categorisation using the WHO international classification of functioning, disability and health framework (ICF) [J]. BMC geriatrics, 2019 (19): 1-15.
- 4 ATAMAN G, ERBAYDAR T. Unmet home healthcare needs and quality of life in cancer patients: a hospital-based Turkish sample [J]. Health & social care in the community, 2017, 25 (4): 1347-1354.
- 5 BANG K S, TAK S H, OH J, et al. Health status and the demand for healthcare among the elderly in the rural Quoc-Oai district of Hanoi in Vietnam [J]. Biomed research international, 2017 (9): e4830968.
- 6 袁艳, 陈洪波, 杨华, 等. 医养结合背景下居家养老护理服务需求调查 [J]. 上海护理, 2020, 20 (1): 23-26.
- 7 GONG Y, ZHOU J, DING F. Investigating the demands for mobile internet-based home nursing services for the elderly [J]. Journal of investigative medicine, 2022, 70 (3): 844-852.
- 8 GAO H, LU S, KOU X. Research on the identification of medical service quality factors: based on a data-driven method [J]. Internet research, 2022, 32 (5): 1617-1645.
- 9 KUNDETI S R, VIJAYANANDA J, MUJJIGA S, et al. Clinical named entity recognition: challenges and opportunities [C]. Washington D. C.: 2016 IEEE International Conference on Big Data (Big Data), 2016.
- 10 汤昊成, 苏万春, 冀秀元, 等. 面向淋巴水肿疾病的电子病历命名实体识别应用研究 [J]. 医学信息学杂志, 2024, 45 (2): 52-58.
- 11 谢腾, 杨俊安, 刘辉. 基于 BERT-BiLSTM-CRF 模型的中文实体识别 [J]. 计算机系统应用, 2020, 29 (7): 48-55.
- 12 EVLIN J, CHANG M W, LEE K, et al. Bert: pre-training of deep bidirectional transformers for language understanding [J]. Computation and language, 2018 (10): 4171-4186.
- 13 LAMPLE G, BALLESTEROS M, SUBRAMANIAN S, et al. Neural architectures for named entity recognition [EB/OL]. [2023-03-04]. <https://arxiv.org/abs/1603.01360>.
- 14 World Health Organization. International classification of functioning, disability and health (ICF) [EB/OL]. [2024-06-02]. <https://www.who.int/standards/classifications/international-classification-of-functioning-disability-and-health>.
- 15 FLORIN J, STRANDBERG E, JANSSON I, et al. A comparison between the ICNP and the ICF for expressing nursing content in the electronic health record [J]. International journal of medical informatics, 2021, 154 (10): 4544.
- 16 燕铁斌, 章马兰, 于佳妮, 等. 国家标准《功能、残疾、健康分类的康复组合评定》践行指南 [J]. 中国康复医学杂志, 2023, 38 (6): 724-729.
- 17 O'HALLORAN R, WORRALL L, HICKSON L. Stroke patients communicating their healthcare needs in hospital: a study within the ICF framework [J]. International journal of language & communication disorders, 2012, 47 (2): 130-143.
- 18 PERIN C, BOLIS M, LIMONTA M, et al. Differences in rehabilitation needs after stroke: a similarity analysis on the ICF core set for stroke [J]. International journal of environmental research and public health, 2020, 17 (12): 4291.
- 19 TKACHENKO M, MALYUK M, HOLMANYUK A, et al. Label studio: data labeling software [EB/OL]. [2024-06-02]. <https://github.com/HumanSignal/label-studio>.
- 20 PASZKE A, GROSS S, MASSA F, et al. PyTorch: an imperative style, high-performance deep learning library [J]. Advances in neural information processing systems, 2019, 32 (12): 8026.
- 21 SANG E F, DE MEULDER F. Introduction to the CoNLL-2003 shared task: language-independent named entity recognition [EB/OL]. [2023-06-12]. <https://arxiv.org/abs/cs/0306050>.