

基于双流融合与动态网格交互的中文医学命名实体识别

李冰¹, 龚展明¹, 张志强¹, 宋海裕¹, 孙元康²

¹浙江财经大学信息技术与人工智能学院, 中国杭州市, 310018

²东南大学计算机科学与工程学院, 中国南京市, 211189

摘要: 中文医学命名实体识别 (CMNER) 是医学信息提取中的一项基础任务, 对于构建临床知识图谱等下游应用以及实现智能临床决策至关重要。然而, 现有的主流方法, 包括词典增强、基于跨度和基于网格的标注方法, 仍难以应对临床文本中固有的自然边界缺失、复杂的嵌套实体结构以及长距离上下文依赖。为解决这些挑战, 提出一种新颖的双流融合与动态网格交互 (D2GI) 模型, 该模型通过整合互补特征流和自适应网格交互进行深度语义挖掘, 从而准确捕捉复杂实体边界和字符间关系。具体而言, 该双流融合架构利用 RoFormer 提取长距离依赖关系, 并结合 Word2Vec 提供稳定的先验语义, 从而增强模型对罕见医学术语的表示能力。此外, 为克服静态特征细化的局限性, 动态网格交互模块采用门控注意力机制自适应地融合局部和全局上下文, 促进模型对嵌套实体的准确识别。在 3 个公开数据集上的多项实验表明, D2GI 显著优于最先进的基线模型, 在 CMeEE-V2、DiaKG 和 CCKS2020 上的 F1 值分别提高 1.48、1.61 和 3.08 个百分点。

关键词: 双流融合; 动态网格交互; 医学语义挖掘; 医学信息提取

<https://doi.org/10.1631/ENG.ITEE.2025.0187>