

CMLNER: Chinese medical literature named entity recognition based on multi-feature fusion and attention mechanism

Wenqi Chen¹, Pu Han¹, Yule Zhong¹, Liang Gu¹, Jianxiang Wei¹, and Xiaoyan Li²

¹Nanjing University of Posts and Telecommunications

²Nanjing Medical University School of Basic Medical Sciences

March 7, 2023

Abstract

Background: With the rapid development of big data and artificial intelligence, the medical knowledge in medical literature has attracted more and more attention from the academic community. As the foundation for medical knowledge extraction and decision support system construction, Chinese medical literature named entity recognition (CMLNER) is the process of automatically recognizing entities in medical literature. Due to the diverse types, unclear boundary, complex composition and lack of explicit separators like space in Chinese medical entity, the task of CMLNER is more complicated compared with English medical literature named entity recognition. **Objective:** In this study, we aim to investigate novel methods to model CMLNER and analyze the results. **Methods:** This study proposes a novel neural network architectural model MFA-BERT-BiLSTM-CRF based on external medical knowledge and self-attention mechanism. **Results:** Compared with traditional NER methods, the proposed model could more effectively capture rich medical semantic features, global context information and further improve the performance of CMLNER. In addition, the key factors affecting CMLNER is also analyzed using the correlation coefficient method and the result indicates that the number and composition rules of entity are main factors. Finally, the recognition performance and error results are also analyzed in this paper. **Conclusions:** Our research makes up for the deficiency of previous frameworks and will further promote the development of medical entity recognition.

Hosted file

CMLNER Chinese medical literature named entity recognition.docx available at <https://authorea.com/users/593257/articles/628362-cmlner-chinese-medical-literature-named-entity-recognition-based-on-multi-feature-fusion-and-attention-mechanism>