ITM Web of Conferences **88**, 01016 (2026) https://doi.org/10.1051/itmconf/20268801016
AIBDCS 2026

Medical Text Knowledge Discovery and Clinical Decision Support Based on Natural Language Processing

Liusiyuan He[1*]

[1]Stony Brook, Anhui university, Hefei, 230000, China

Abstract. Natural language processing (NLP) technology is the key driving force to unlock the value of massive, unstructured medical text data and promote the development of smart medicine. This paper aims to systematically review the application of NLP in the field of health care, focusing on how "medical text mining" drives "knowledge discovery" and ultimately serves "clinical decision support". Firstly, this paper reviews the evolution of technology from the early rule method to the current pre training language model. Then, the core technologies such as medical information extraction, knowledge map construction, text classification and generation and their application in typical scenarios such as electronic medical record analysis, auxiliary diagnosis, prognosis prediction, and patient management are reviewed. Through the induction and comparison of existing studies, this paper summarizes the main challenges currently facing, including medical data privacy and labeling problems, interpretability and clinical credibility barriers of the model, and systemic barriers to multimodal fusion. Finally, this paper looks forward to the future research directions, such as the development of interpretable AI and the construction of NLP system for real-world evidence, in order to promote the transformation of this technology from research to safe, reliable and efficient clinical landing.

1 Introduction

With the development of artificial intelligence, NLP has become the core driving force of medical digitalization. The medical field is one of the earliest and most important professional fields of NLP application. Since the first medical expert system aaphelp for auxiliary diagnosis was released in 1972, intelligent medical research has been at the forefront of NLP development [1]. Extracting value information is the key challenge of medical informatization.

The development of NLP in the medical field is divided into three stages. The rule expert system from 1980 to 2000 can deal with specific problems, but it is difficult to acquire knowledge and has poor expansibility; From 2000 to 2015, statistical machine learning (conditional random fields, support vector machines) promoted the initial development of medical information extraction; Since 2015, the in-depth learning and pre training model (BioBERT 、 ClinicalRadioBERT) has significantly improved semantic understanding through massive medical text pre training, and has penetrated into clinical diagnosis and treatment, medical research and other scenes [2].

The current research focuses on a single technical link, and lacks a systematic analysis of the whole link of "text mining knowledge discovery decision support". The field is still facing challenges, and the general model is difficult to adapt to the professionalism of medical text; High quality labeling data is scarce (the cost of labeling a single case is hundreds of yuan); Deep learning "black box" reduces clinical trust; Data privacy and multimodal fusion also restrict applications.

This paper aims to sort out the development context of NLP medical field, analyze the core technologies, summarize the application scenarios, analyze the challenges and explore the future direction, so as to provide reference for research.

2 Natural language processing (NLP) technology: the key engine to release the value of medical text

2.1 Technological evolution and development context

The development of NLP in the medical field has experienced three stages. In the stage of rule expert system from 1980 to 2000, although it can deal with specific medical problems, knowledge acquisition depends on manual coding and has poor scalability; In the statistical machine learning stage from 2000 to 2015, conditional random fields, support vector machines and other models were applied to medical named entity recognition and relationship extraction to promote the preliminary development of medical information extraction; Since 2015, in the stage of in-depth learning and pre training model, the emergence of transformer architecture has completely changed the research paradigm - pre training models such as BioBERT and

* Corresponding author: liusiyuan.he@outlook.com

ClinicalRadioBERT have significantly improved the understanding ability of medical text semantics through pre training on a large number of medical texts. After standardization, the kappa value of the structured system increased from 0.3 to 0.51, and the F1 score increased significantly, which verified its effectiveness in reducing the subjective differences between doctors and improving the quality of reports. In addition, this study also combines the structured results with the medical knowledge map, which provides a data basis for subsequent image quality assessment and prognosis analysis. [3].

2.2 Core values and application framework

The value of NLP in the medical field is reflected in three aspects. In terms of efficiency improvement, automated text processing can liberate doctors from tedious text work [2]; In terms of quality improvement, the accuracy and consistency of medical records were improved through standardized terminology and structured management [2]; In terms of knowledge discovery, the association between disease rules and treatment in massive data is mined to provide data support for precision medicine [2].

The current application framework of medical NLP is built around four core directions, including information extraction and structure of extracting clinical entities and relationships from unstructured text, support system of assisting clinical decision-making based on extracted information, patient service of realizing personalized health management through text analysis, and knowledge discovery of mining new medical knowledge from the literature, which form a closed loop of "data information knowledge decision".

3 Core technology exploration: the technical implementation path from text mining to decision support

3.1 Medical text mining and knowledge discovery

3.1.1 Information extraction of medical text

Named entity recognition is the basic task of medical NLP, which aims to identify clinical entities such as diseases, symptoms, drugs, examinations, and operations in the text. The current mainstream methods are based on deep learning sequential annotation model (such as bilstm-crf) or pre training language model to fine tune. For example, Chen et al. evaluated the diagnostic ability of two online general NLP models, Kimi and chatgpt-4o, to achieve automatic FIGO staging of cervical cancer by interpreting pathological reports, and explored the feasibility of applying them to automatic tumor staging and assisting clinical decision-making[4].

Based on entity recognition, relationship extraction can further judge the semantic relationship between entities (such as "drug treatment disease"). The technical method has developed from early pattern matching to deep learning method based on graph neural network, which can capture complex syntactic structure and long-distance dependence. Rahman · F et al. used pre-set inclusion and exclusion criteria for screening, while relevant studies evaluated their quality using PROBAST. Extracted data on research features, datasets, annotation processes, natural language processing methods, performance indicators, population demographics, and clinical applications. Use PRISMA flowchart to illustrate the research screening process providing a text analysis basis for the comprehensive assessment of geriatric syndrome [5].

Event extraction is a higher-order extraction task, which needs to deal with more complex clinical events (such as the time, method, result and other attributes of surgical events), and needs to jointly model multiple entities and trigger words.

3.1.2 Mapping and representation of medical knowledge

The construction of knowledge map systematically integrates the triples (entity relationship entity) obtained from information extraction to construct a structured medical knowledge network. Rahman F et al.'s study aims to explore the role of NLP in identifying and analyzing GS in the field of geriatric syndrome, examining its application, methodology, and efficacy [5].

Knowledge representation learning maps the entities and relationships in the knowledge map into low dimensional vectors through embedded models such as TransE and RotatE, so that the machine can simulate logical reasoning through vector operations. For example, by calculating the vector similarity of "chest pain+dyspnea", People can find potential diseases (such as acute coronary syndrome).

3.2 Technical implementation of clinical decision support system

3.2.1 Diagnostic aids and triage recommendations

The reasoning based on knowledge map maps the patient information into the knowledge map, and generates diagnostic hypotheses through graph search, which has good interpretability. Zhang et al.'s scheme has realized the quantitative calculation of diabetes condition, symptom information matching, automatic symptom summary, disease type discrimination and intelligent recommendation of traditional Chinese medicine, and conducted empirical research. After testing and evaluation, it shows that the scheme can effectively help doctors to solve the uncertainty of diagnosis in the actual use process, and provide more diversified and intelligent selection basis [6].

The classification model based on machine learning converts the patient text information into feature vectors, and the input classifier directly outputs the disease probability prediction. These methods rely on a large number of annotation data, but can capture complex nonlinear patterns. Among the 7 deep NLP models established by Xun et al., ELECTRA model performed the best in the classification task of longitudinal imaging reports of brain tumors [7].

Annie e Bowles et al. developed and validated a natural language processing (NLP) system based on echocardiography report data from the US Department of Veterans Affairs to automatically extract heart valve morphology from echocardiography reports, with a focus on BAV detection [8].

3.2.2 Treatment scheme recommendation and medication safety

Guide driven recommendation by structuring clinical guidelines (such as diabetes and cancer guidelines) and automatically matching patient characteristics with standard schemes, Zhang et al.'s TCM system can provide more accurate treatment plans [6]. Similar case retrieval is based on text similarity to provide personalized treatment reference for cancer patients (such as matching chemotherapy schemes with the same pathological type). In addition, NLP can analyze the drug information in the electronic medical record, combined with the drug knowledge base to warn the risk of allergy or interaction (for example, amoxicillin is forbidden for penicillin allergy), so as to avoid medication errors.

3.2.3 Prognosis assessment and risk prediction

Such applications integrate text features (such as "infiltration depth" in pathological reports) and laboratory indicators to build models.Xun et al. trained a deep natural language processing model to analyze MRI imaging reports of brain tumors [7]; Wang et al. found that advanced natural language processing models can accurately extract chronic kidney disease related features from ultrasound reports, which is expected to provide scalable tools for early detection and risk stratification. Future research should focus on validating the application of these models in different healthcare systems [9]. In addition, NLP can also analyze the "operation duration" and "complication history" in the operation records, predict the risk of postoperative infection or anastomotic leakage, and provide the basis for clinical intervention.

3.2.4 Evidence based medicine support

NLP quickly extracts the core conclusions of the literature through the text summary (such as summarizing the key evidence from 10000 literatures for the "diabetes diet intervention"), so as to improve the efficiency of scientific research. Semantic retrieval breaks through the restriction of keywords, identifies related concepts such as "elderly diabetes", "hypoglycemia risk" [2]. At the same time, the system can automatically analyze the "research design" and "sample size" of the literature, and help mark the level of evidence (such as the high/medium/low quality of grade system) to help doctors quickly judge the reliability of evidence.

4 Empirical analysis and technical performance evaluation

4.1 Performance analysis of medical text processing

In terms of medical text classification and quality assessment, the deep learning model shows significant advantages. Li et al.'s research shows that the quality assessment model for nasopharyngeal carcinoma MRI reports constructed based on the Bert model significantly improves the consistency and accuracy of the reports. After system optimization, the Kappa value increased from 0.3 to 0.51, and the F1 value significantly improved [3]. The multimodal analysis conducted by Liu et al. shows that knowledge graphs are often used as a learning process for external expert knowledge injection models, assisting machine learning models in achieving accurate understanding and effective inference of data. Given the high dependence of medical imaging research on professional knowledge and clinical experience, integrating expert knowledge, especially in the form of knowledge graphs, into model architecture is a research direction with significant practical value [2].

In the task of information extraction, the pre training model has obvious advantages. Chen et al.'s research shows that NLP models have the same potential as clinical physicians in assisting clinical staging and have certain clinical application value [4].

Jonathan C Stanley et al. developed a natural language processing (NLP) method for extracting key concepts such as tumor thickness and ulcer status from melanoma pathology reports. The system achieved high accuracy and recall, addressing the complex staging criteria for melanoma and meeting the operational needs of downstream staging and queue recruitment [10].

4.2 Evaluation of clinical application effect

In terms of clinical decision support, the research of Zhang et al. can better assist doctors in diagnosis and treatment, and has a high reference value for diabetes diagnosis and medication recommendations. However, due to data size constraints and other uncontrollable factors, there are also many cases of miscarriage, omission and miscarriage of justice [6]. Cui et al. used NLP technology to analyze speech and text data, which may provide more personalized diagnostic information; At the same time, specific exercises targeting language weakness are used to improve or compensate for patients' language abilities through human-machine collaboration to adapt to and cope with the environment [11].

In Wang et al.'s study, NLP analysis in renal ultrasound reports showed high consistency with the diagnosis of chronic kidney disease (CKD), providing an effective tool for early screening of kidney disease [9].

5 Challenges, limitations and future prospects

5.1 Main challenges

The challenge at the data level is the most prominent. Medical data involves sensitive privacy and is restricted by HIPAA, gdpr and other laws and regulations, and sharing and utilization are limited; The cost of high-quality labeling is high. It takes several hours for professional doctors to label a case. Multimodal data fusion technology is not mature, and there are still obstacles in the collaborative analysis of text, image, gene and other data.

At the technical level, the lack of interpretability of the model limits the clinical acceptance. The "black box" feature of the deep learning model makes it difficult for doctors to trust their decision-making. The consistency between the results of Xun et al. RNN model and human annotation is general (k=0.46). On the one hand, the reason may be that the task of tumor state classification itself is complex, and the deep learning model has high requirements for language recognition and correlation between languages; On the other hand, the RNN model itself has the disadvantage of gradient disappearance and is not good at dealing with long-term dependence [7]

At the landing level, the system integration complexity is high, and it needs to be seamlessly connected with the existing hospital information system (his, EMR); The clinical validation cycle is long and the cost is high, which requires multi center trials to verify the safety; There is a lack of uniform ethical norms and evaluation standards, and the compliance of data use has not yet formed a consensus.

5.2 Future development direction

Multimodal learning is an important direction, combining text, image, gene and other multi-source data to build a comprehensive patient portrait. The multimodal analysis of Liu et al. Shows that the fusion of text and image data can improve the diagnostic accuracy [2].

The development of interpretable AI will make the decision-making process of the model transparent, display the text fragments concerned by the model through attention mechanism and visualization technology (such as grad CAM), and improve the clinical trust.

Real world evidence generation uses real world data (such as electronic medical records and medical insurance data) to verify the effect of the model and promote NLP from laboratory to clinic.

Federated learning can realize multi agency data sharing under the premise of privacy protection, and improve the model generalization ability through distributed training.

Technical standardization and the establishment of ethical framework will also be the focus of future development, which needs the collaborative promotion of industry, University, research and medicine, and the formulation of medical NLP terminology specifications, data use consent process and other standards.

6 Conclusion

This paper systematically reviews the technological evolution of natural language processing in the field of health care, and deeply analyzes the core technology path from medical text mining, knowledge mapping to clinical decision support. Combined with typical application scenarios and empirical data, this paper summarizes the actual results of current research in terms of efficiency improvement, quality improvement and knowledge discovery.

NLP technology shows irreplaceable core value in the field of medical and health care. NLP is profoundly changing the medical service mode through the complete technology chain of medical text mining driving knowledge discovery and serving for clinical decision support. From information extraction to knowledge mapping and decision support system, NLP not only improves medical efficiency and quality, but also provides new possibilities for precision medicine and personalized treatment.

Although it still faces challenges such as data privacy, model interpretability and system integration, with the progress of multimodal learning, interpretable AI, federated learning and other technologies, NLP will play an increasingly important role in the construction of smart medicine. In the future, collaborative breakthroughs in technology, data and specifications are needed to promote the transformation of NLP from an auxiliary tool to an intelligent partner of doctors and contribute to improving the health level of the whole people.

References

1. J. Y. Hu., et al. Natural Language Processing and Its Application in the Medical Field. Chinese Journal of Evidence-Based Medicine **24**.10(2024):1205-1211.
2. Z. L. Liu., et al. Application of Natural Language Processing in Medical Image Analysis." Journal of Central South University (Health Sciences) **47**.08(2022):981-993.
3. T. Li. Quality Evaluation Method of MRI Report for Nasopharyngeal Cancer Based on Natural Language Processing. Guilin University of Electronic Science and Technology, MA thesis. (2025) doi:10.27049/d.cnki.ggldc.2025.001053.
4. X. Chen., et al. Application of natural language processing model in cervical cancer staging and risk factor extraction." Oncology **45**.03(2025):287-296.
5. F. Rahman, I. Gailer, A. Hasan, et al. Natural language processing for geriatric syndromes: a systematic review of methods, applications, and challenges. BMC Medical Informatics and Decision Making. (2026) doi:10.1186/S12911-026-03417-0.

6. Y. J. Zhang., et al. Research on Traditional Chinese Medicine Auxiliary Diagnosis and Treatment Scheme Integrating Multi-natural Language Processing Tasks: A Case Study of Diabetes. Data Analysis and Knowledge Discovery **6**.01(2022):122-133.
7. X. P. Xun. Research on the correlation between brain tumor reporting and classification results and prognosis using deep NLP. Central South University, MA thesis. (2022) doi:10.27661/d.cnki.gzhnu.2022.003959.
8. A. E. Bowles, J. A. Lynch, F. Bermudez, et al. Detecting Bicuspid Aortic Valve From Echocardiographic Reports Using Natural Language Processing: A Veterans Affairs Study. JACC. Advances. **5**(1):102463. (2025) DOI:10.1016/J.JACADV.2025.102463.
9. C. L. Wang., et al. Natural language processing for kidney ultrasound analysis: correlating imaging reports with chronic kidney disease diagnosis. Renal failure. **47**(1): 2539938. (2025) DOI:10.1080/0886022X.2025.2539938.
10. J. C. Stanley., et al. Development and Evaluation of Natural Language Processing Methods for Extracting Key Melanoma Pathology Concepts. Studies in health technology and informatics. **329**:774-778. (2025) DOI:10.3233/SHTI250945.
11. Y. M. Cui., et al. Research progress on the application of natural language processing in aphasia rehabilitation." Medical Research and War Trauma Therapy **38**.03:327-331. (2025) doi:10.16571/j.cnki.2097-2768.2025.03.019.