

Received 16 February 2023, accepted 4 March 2023, date of publication 8 March 2023, date of current version 16 March 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3254303

RESEARCH ARTICLE

Intelligent Recommendation for Departments Based on Medical Knowledge Graph

ZHAOJIAN CUI¹, ZHENMING YUAN^{1,2}, YINGFEI WU^{1,2}, XIAOYAN SUN¹, AND KAI YU^{1,2}

¹Department of Information Science and Technology, Hangzhou Normal University, Hangzhou 311121, China

²Hangzhou HeleTech Co. LTD., Hangzhou 311121, China

Corresponding author: Kai Yu (yk@hznu.edu.cn)

This work was supported in part by the Xinjiang Production and Construction Corps Key Scientific and Technology Project under Grant 2021ABO34.

ABSTRACT The clinical sub-speciality departments are increasing. It is usually difficult for patients without medical education to choose a suitable department when they need to make an online registration with doctor. A novel approach is presented for departments recommendation in this paper. The ICD codes and symptoms of medical health data are used to describe the patients' characteristics, diseases, and departments in the quantitative relation. The knowledge graph is built with the relations to automatically recommend departments. The MacBERT-BiLSTM-CRF medical entity recognition model (MacNER) is proposed to identify the patient's symptoms, parts, treatment methods and drug entities to structure a knowledge graph. The knowledge graph is used to provide knowledge for the intelligent department recommendation. The identified entities are mapped to nodes in the graph with the entity mapping method. Finally, an algorithm named Feature Rate Multiply Converse Disease Rate (FRMCDR) is proposed. The most appropriate department can be recommended by fusing the patient's chief complaint and past medical history. The experiment shows that our method obtained an 88.77% precision rate in the recommendation.

INDEX TERMS Intelligent recommendation, departments recommendation, knowledge graph, medical entity recognition.

I. INTRODUCTION

Various forms of medical services such as online popularization of medical, remote diagnosis, and online question-and-answer between doctors and patients are increasing with the continuous development of the Internet and medical information technology. Artificial intelligence has also entered the medical field early, such as imaging diagnosis [1], as well as disease screening and prediction using various statistical models. The online registration service has been popularized in major hospitals today in China. Many general hospitals, especially the provincial and municipal tertiary hospitals, often have dozens of departments. These departments may even be divided into hundreds of subspecialty sections according to different diseases. In large special hospitals, a department in a general hospital can be further divided into various subspecialty sections or sub-departments. For example, you can find the menopause clinic and infertility clinic in an obstetrics and gynaecology hospital, which are both

parts of gynaecology clinic. Patients can ask the nurses in the hospital which department they should go to when registering on the spot. However, because of the online reserved registration service, a large number of registration sources have been reserved in advance. Therefore, an automatic department guide for patients is an urgent need nowadays.

There are two categories of intelligent department recommendation approaches at the moment: template-based [2], [3], [4] and model-based [5], [6], [33], [34]. With the template-based method, a series of rules are established to step patients in describing their symptoms, and the system tells which department the patients should go to. For example, Cui and Liu [2] developed an offline intelligent guidance robot, which designed a graphical interface that guided patients to choose uncomfortable positions. Then the manual question-and-answer rules were introduced to have a conversation with the patient to collect information from patients. Appropriate departments were recommended by calculating the weights that match the existing medical knowledge bases of the corresponding departments. An auxiliary diagnosis and treatment system was developed by Quan et al. [3], which

The associate editor coordinating the review of this manuscript and approving it for publication was Wenbing Zhao¹.

interacted with patients in the form of an interactive interface as well. The recommendations of the department were made by analyzing the symptoms entered by the patient and the established disease database. In addition, the guidance robots developed by Chen et al. [4] and Cui and Liu [2] added sensor modules.

The model-based method does not require a set of pre-defined rules and conditions. Only the text input by the patient is required, and the traditional machine learning models or deep learning models could be used for classification. Ya [5] built an online consultation platform and used the linear support vector machine(SVM) to recommend departments for patients by vectorizing the consultation text. Similarly, Wang et al. [33] proposed a method of classifying chief complaints based on a cloud platform to achieve the prediction. The convolutional neural network (CNN) and the recurrent neural network (RNN) were used by Zhou et al. [6] to construct a model called CRNN. The recommended department was obtained by taking the output of the multi-layer perceptron (MLP). Wang et al. [34] proposed a self-learning framework to provide a treatment recommendation for the doctors. The above methods, however, did not take into account the requirement of medical interpretability and the relevance of knowledge and lacked specific descriptions between the patient's condition and the department.

An intelligent guidance method is presented to try to solve the problem. It not only relieves the strain on medical resources but also protects the privacy of patients as much as possible. At the same time, compared to the traditional template method, there is no need to manually define a series of rules in advance. Natural language processing (NLP) techniques are used to automatically extract information from text. Furthermore, in terms of interpretability, a knowledge graph is constructed to illustrate the relationship between the patient's condition, disease, and department. By using the data that are processed by the model to build a knowledge graph, a system of intelligent recommendations for departments is designed. The contributions of this paper can be summarized as follows:

- We not only consider the patient's main complaint text but also the patient's past medical history information, to describe the patient's attributes and characterize the patient's characteristics from multiple aspects.
- We propose the MacNER model and FRMCDR algorithm to improve the effectiveness of the whole system. Meanwhile, the different entity similarity methods are applied to achieve the goal of entity normalization.
- Experimental results verify the effectiveness of the proposed method. Compared to the traditional method, the precision, recall and F1-score are improved by 7.71%, 7.31% and 7.5%, respectively.

II. RELATED WORK

In this section, the development of pre-trained models and their application in medical named entity recognition tasks are introduced. Then, the construction situation of the

medical knowledge graph using the named entity recognition method and Electronic Medical Records (EMRs) gained from the hospital is given.

A. PRE-TRAINED MODEL

Since Transformer [8] is proposed, various pre-trained language models have emerged one after another. Bidirectional Encoder Representations from Transformers (BERT) pre-trained model [9] is the most classic pre-trained language model and frequently appears in many works. It had achieved significant results on multiple natural language processing tasks at that time. Then the Robustly Optimized BERT Pretraining Approach (RoBERTa) model [10] is proposed. It adjusts and improves the native BERT in pre-training processes, such as the use of a dynamic Masking mechanism, the removal of Next Sentence Prediction (NSP) tasks, etc. It exceeded the performance of BEET on some tasks.

To apply language models in the Chinese field, the joint laboratory of Harbin Institute of Technology and iFLYTEK released the pre-trained language models based on Chinese [7]. These models are released to address all kinds of Chinese tasks. Experimentally, they expand the data in the pre-training process and use the whole-word masking mechanism. At the same time, the MacBERT their named pre-trained model was released and achieved the best results on most Chinese tasks.

Related research [11], [12] using various BERT-based pre-trained models in the named entity recognition task was published. The BERT-base model added the Bi-directional Long Short-Term Memory (BiLSTM) and Conditional Random Field (CRF) layer was used to extract entities from Chinese clinical texts. For example, Zhang et al. [13] solved the breast cancer Named Entity Recognition (NER) task.

B. MEDICAL NAMED ENTITY RECOGNITION

The task of medical named entity recognition (MNER) refers to identifying various entities, such as symptoms, diseases, drugs, etc, from medical clinical texts. Currently, three main approaches have been studied to solve this problem, named dictionary-based and rule-based methods, statistics and machine learning methods, and deep learning methods.

The dictionary and rule-based approaches focus on processing the text to be processed in a matching way. There were MedKAT [14], cTAKES [15], and other systems built with the dictionary and rule-based methods. Intuitively, it can achieve very good results with small datasets. But it generally consumes a lot of manpower and is difficult to cover all medical entities due to the complexity and diversity of medical entities.

Supervised learning flourishes as the quality improvement and type enrichment of data. Various machine learning methods have been applied in entity recognition tasks, such as Hidden Markov Models (HMM) [16], Maximum Entropy Markov Model (MEMM) [17], CRF [18], and so on. Among them, CRF makes full use of contextual information and achieves a good result.

The deep learning method has been evolving with the development of computer hardware. Therefore, methods integrated deep neural networks with machine learning models are emerging to improve learning efficiency and result accuracy. The model structure of combining the CRF layer and BiLSTM layer together commonly is used. And the performance of State-Of-The-Art (SOTA) in many NER tasks is obtained [19], [20], [21], [22]. For example, Dai et al. [22] constructed the BERT-BiLSTM-CRF model that could achieve a 74.53% F1-score, which was superior to the other model without a BiLSTM layer or CRF layer.

C. MEDICAL KNOWLEDGE GRAPH

Medical knowledge graphs have attracted attention from academia and the healthcare industry in light of their powerful functions in intelligent medical applications. The development of knowledge graphs is prompted by the establishment of various public datasets and knowledge libraries. Xiong et al. [23] extracted various information from the domain-specific corpus to construct a prescription drug monitoring program for Chinese patent medicines based on knowledge graphs. This method ensures the safety of the patient's medication. In addition, there are free medical knowledge graphs such as the Chinese Medical Knowledge Graph (CmeKG¹) opened for medical applications [24], and the OpenKG² designed by the University of East China University of Science and Technology.

Electronic Medical Records (EMRs), including patients' disease course records, examination results, doctor's orders, surgical records, nursing records, and other relational information, promote the construction and development of medical knowledge graphs. Li et al. [25] used EMRs and natural language processing techniques to construct a knowledge graph containing nine entities. Rotmensch et al. [26] explored a method that could learn knowledge bases linking diseases and symptoms to automatically construct knowledge graphs from EMRs. As for the application of question and answer, some researchers constructed medical knowledge graphs by extracting entities and their relations from documents [27], [28].

III. METHOD

The overall architecture of the entire intelligent guidance system is depicted in FIGURE 1 and FIGURE 2. The named entity recognition model is proposed for building the knowledge graph. The entity normalization problem and its solution are discussed. To recommend suitable departments, an algorithm for calculating relation weights called FRMCDR (Feature Rate Multiply Converse Disease Rate) is considered in the process.

¹<http://cmekg.pcl.ac.cn/>

²<http://www.openkg.cn/>

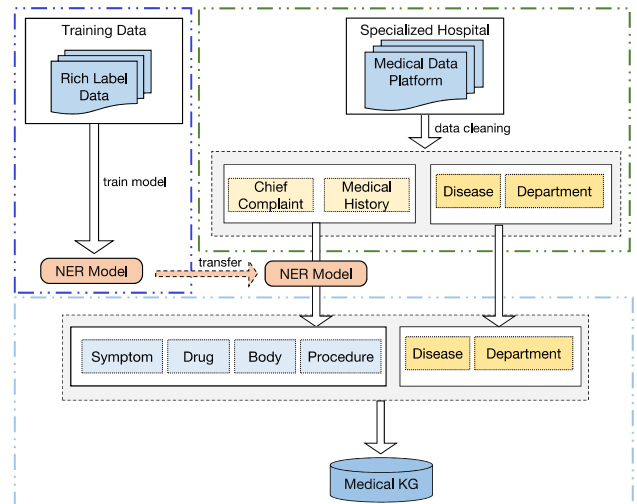


FIGURE 1. The building of knowledge graph of intelligent guidance system. The upper left part is model training, the upper right is data cleaning, and the bottom is KG building with the triplet.

A. GENERAL FRAMEWORK

The overall framework of the intelligent guidance system consists of two parts, which are illustrated in FIGURE 1 and FIGURE 2.

The first stage, as shown in FIGURE 1, is mainly composed of three parts: model training, data cleaning and graph building. Firstly, the public entity recognition dataset was used to train the model proposed in this paper. And then the trained model was transferred to predict entities in the collected dataset from the hospital. Secondly, the data obtained from a regional specialized hospital are cleaned. The detailed processing is shown in Section DATA PROCESSING. Finally, the identified entities were stored in the graph according to the defined triplet.

As shown in FIGURE 2, the reasoning process on the knowledge graph built in the first stage mainly consists of three parts. The first part, entity recognition, identifies four types of patient characteristic entities (symptom, body, procedure and drug). MacNER model and Aho-Corasick (AC) automata algorithm [29] are used to extract entities from the patient's chief complaint text and past medical history. The AC automata algorithm is used to complement symptom entities. Specifically, a post-processing way of constructing a symptom corpus from some medical websites is introduced. It could solve the problem of insufficient entities predicted after model transferring. AC automata algorithm was used to obtain symptom entities from the text. It should be noted that the symptom entity set was obtained by removing duplicating the symptom entities obtained by MacNER and AC automata algorithm. It avoided repetitive operations when retrieving. The second part is entity mapping. The problem should be considered that the entities cannot be normalized to the nodes in the graph, which is explained in detail in Section ENTITY MAPPING. The last part is the FRMCDR algorithm. It has better interpretability and accuracy than methods that only

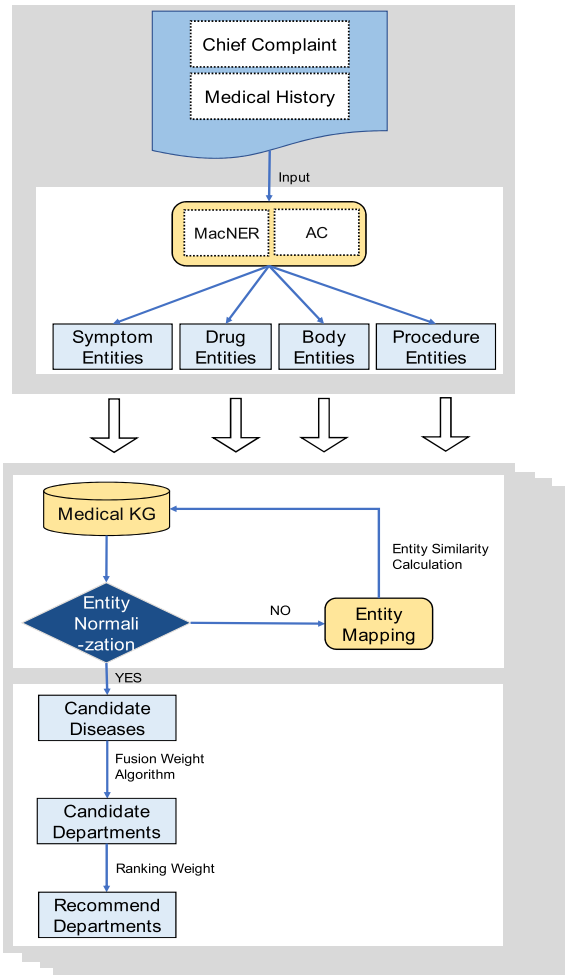


FIGURE 2. Flowchart of recommendation of intelligent guidance system, which is composed of upper, middle and lower white parts.

consider the frequency of entities as relation weights. The design of this algorithm is described in Section FRMCDR ALGORITHM. By using the three components above, the three departments with the highest scores are ultimately recommended to a patient.

B. MLM AS A CORRECTION STRATEGY

Since the BERT [9] is based on English, there are some problems with the applicability of Chinese. Moreover, experiments in the previous work [10] prove that removing the NSP pre-training task of BERT is more beneficial for downstream tasks. In addition, the Masked Language Model (MLM) pre-training task of native BERT does not consider the relation between subwords after word segmentation by using WordPiece embeddings [30]. Therefore, Google proposed the Whole Word Masking (WWM) strategy to further improve the reasoning ability of the language model in the pre-training stage.

To better exploit the knowledge from the existing pre-trained language model, a Chinese-based pre-trained model called MacBERT(MLM as correction BERT) [7] is released. Referring to the WWM strategy, the pre-training task of

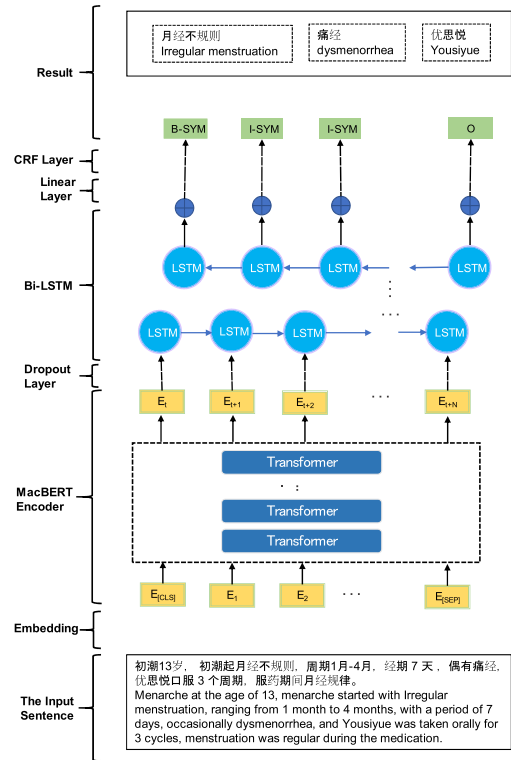


FIGURE 3. The structure of MacNER model and illustration.

TABLE 1. The different masking strategies.

Masking Strategy Name	Chinese Result
None	一般疗法为卧床休息，给予高热量，富含蛋白质的饮食，纠正电解质紊乱。
MLM*	一般疗法为[MASK]床休息，给予高[MASK]量，富含[MASK]白质的饮食，纠正电解质紊乱。
WWM	一般疗法为[MASK] [MASK] 休息，给予 [MASK] [MASK] [MASK]，富含[MASK] [MASK] [MASK] 的饮食，纠正电解质紊乱。
MLMAC	一般治疗为卧床睡觉，给予高能量，包含蛋白质的饮食，纠正电解质混乱。

The mask symbol is shown in italics [MASK]. The none represents the original sentence and masking strategy without use. MLM* represents the masking method of native BERT, and WWM is the whole word masking strategy. MLMAC represents MLM as correction method that MacBERT applies.

BERT is improved according to the Chinese characteristics. Formally, it adopts the MLM as a correction (MLMAC) task instead of the MLM task. This aims is to convert the task of predicting the original words represented by [MASK] tokens into the task of text correction. Specifically, the task of text correction is to correct words replaced by similar words without using [MASK]. The reasons are, on the one hand, that compensates for the impact of [MASK] on downstream tasks due to the [MASK] does not appear except during pre-training, and on the other hand, it is more natural than the MLM task. The different masking methods described above are shown in TABLE 1.

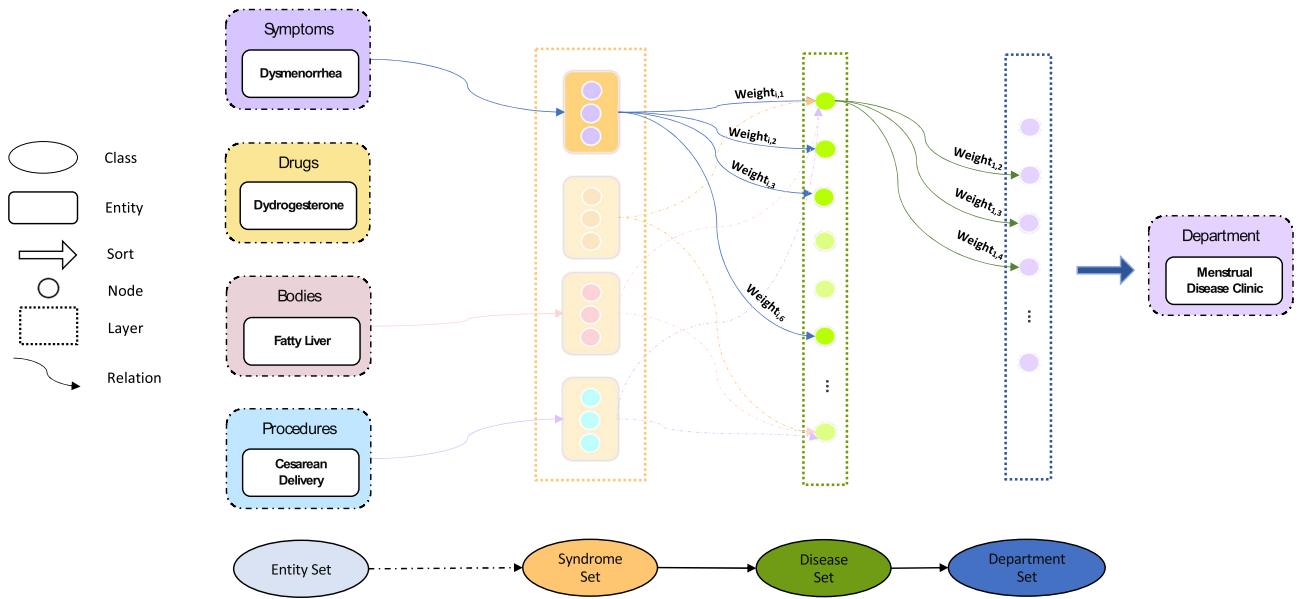


FIGURE 4. The upper part of the figure is the instance layer of the medical knowledge graph, the lower is the concept layer of the medical knowledge graph. The leftmost part of the figure is the legend. The dotted line indicates that there is a connection between the start node layer and the target node layer.

At the same time, the NSP task of BERT is changed to the Sentence Order Prediction (SOP) task. The SOP task allows the pre-training model to judge the sequence of input sentence pairs. The two improved pre-training tasks have been shown to achieve significant results in some Chinese tasks, such as machine reading comprehension, text classification, etc. So MacBERT is improved according to the Chinese characteristics and explores more knowledge than BERT. In addition, it collects more abundant corpora than BERT for pre-training.

C. MACBART FOR NER

To better explore the knowledge of the pre-trained language model, MacBERT is applied as the backbone for the NER task on the collected dataset. The structure of this paper proposed MacBERT-based entity recognition model is shown in FIGURE 3. This model consists of three parts: the pre-trained language model MacBERT, the Bi-LSTM layer, and the CRF layer. Since MacBERT is used as the backbone for the NER task, this model is named MacNER.

D. MEDICAL KNOWLEDGE GRAPH DEFINITION

The concept layer and instance layer of the medical knowledge graph are shown in FIGURE 4. The schema of the graph is represented by the triple $G = \langle S_k, W_k, O_k \rangle$, where G represents the graph, S_k represents the subject, O_k represents the object, and W_k represents the relation weight between S_k and O_k .

For the subject S_k and the object O_k , there are three types from the concept layer of FIGURE 4, called syndrome, disease and department. The syndrome set consists of four types of entities, which are the patient's symptoms, the drugs used,

the uncomfortable positions of the body, and the treatment methods. They are briefly defined as symptoms, drugs, bodies and procedures.

If there is no correlation between the patients' symptoms and diseases, it is difficult to infer which department the patients should go to. Therefore, the relationship between the syndromes and diseases must be considered. W_k describes five kinds of relations, which are the relation of the patient's symptoms with the diseases, the sick position of the body with the diseases, the drug used with the diseases, treatment methods with the diseases, and the diseases with the departments. Specifically, W_k represents the correlation between the former and the latter.

And the larger W_k is, the stronger correlation is. For example, dysmenorrhea (set its id as i) in symptom is considered. According to the relation in the knowledge graph, it can be known that dysmenorrhea is closely related to disease₁, disease₂, disease₃ and disease₆. In other words, $Weight_{i,1}$, $Weight_{i,2}$, $Weight_{i,3}$ and $Weight_{i,6}$ are larger than other disease relation (solid lines are used in FIGURE 4). The process of relation inference between entities in the remaining syndrome sets and diseases is the same. The probability relationship between diseases and departments is the same. Namely, there is a fixed relationship between diseases and departments. It is concluded that the patients should go to the menstrual disease clinic according to the FRMCDR algorithm.

E. ENTITY MAPPING

Since we transfer the model to predict the entities, entities extracted from the patient's complaint or medical

history may be slightly different. It results in empty query results when retrieving. For example, the symptom entities “跑步时漏尿 (leaking urine while running)” and “蹦跳时漏尿 (leaking urine while jumping)” are similar symptom entities. But they are classified as completely different entities when retrieved. So it is necessary to consider a solution that entities can be normalized to graph nodes.

We consider both the lexeme level and semantics level. The three methods, Jaccard-similarity, Hamming-similarity and Edit-similarity, solve the above problem by comparing the entity similarity from the lexeme aspect. To be specific, they are used to contrast the similarity between the entities in the graph and the entities in the query. In addition, the similarity of entities at the semantics aspect also needs to be considered. So the Jieba³ tool of word segmentation performs Chinese word segmentation on the patient's complaint and medical history text. And language model of word2vec [31] is trained to obtain the word representation in our dataset. The corpora obtained by entity recognition are added to the dictionary set of the Jieba tool for word segmentation. Compared to word-by-word segmentation, entity phrases can be segmented. In this way, on the one hand, the size of the generated corpora can be reduced, and on the other hand, some semantic information can be preserved. Since there are long symptom entities, they are split into several short entities by using Jieba word segmentation. The different word representations are obtained by adding the segmented short entity representations. So far, we can obtain word representations of the entities at the semantic level. Then the Cos-similarity similarity algorithm is used to compare the entity similarity. In particular, word representation is treated as an all-zero vector when the long entity after segmentation has no corresponding word representation.

F. FRMCDR ALGORITHM

We define the following five types of relationships among the entities and the disease: 1). $S = \{s_1, s_2, s_3, \dots, s_i, \dots\}$ is represented as the set of symptom entities; 2). $B = \{b_1, b_2, b_3, \dots, b_i, \dots\}$ is represented as the set of body entities; 3). $DR = \{dr_1, dr_2, dr_3, \dots, dr_i, \dots\}$ is represented as the set of drug entities; 4). $P = \{p_1, p_2, p_3, \dots, p_i, \dots\}$ is represented as the set of procedure entities; 5). $D = \{d_1, d_2, d_3, \dots, d_i, \dots\}$ is represented as the set of disease entities. We define $\text{Rel}(\cdot)$ as the relation pair, and $\text{Sum}(\cdot)$ as the number statistics. Then the relationship between symptom s_i and disease d_j can be defined as $\text{Rel}(s_i, d_j)$, and the total number of occurrences of relation in the dataset is defined as $\text{Sum}(\text{Rel}(s_i, d_j))$.

Firstly, the most intuitive method is to select the most frequent one. According to the weight calculation method of probability statistics, its definition, between symptom s_i and

disease d_j , can be expressed as follows:

$$\text{Weight}_{(s_i, d_j)} = \frac{\text{Sum}(\text{Rel}(s_i, d_j))}{\sum_{s_k \in S} \text{Sum}(\text{Rel}(s_k, d_j))} \quad (1)$$

Similarly, the relat of residual entities and the disease entities can be expressed according to equation (1).

Supposing the set of recommendation departments that should be deduced is expressed as $\widehat{\text{DEP}} = \{\text{dep}_1, \text{dep}_2, \text{dep}_3, \dots, \text{dep}_i, \dots\}$, we define $\widehat{S} = \{s_1, s_2, s_3, \dots, s_i, \dots\}$ is represented as the symptom entity set of the current patient, $\widehat{B} = \{b_1, b_2, b_3, \dots, b_i, \dots\}$ is represented as the body entity set, $\widehat{DR} = \{dr_1, dr_2, dr_3, \dots, dr_i, \dots\}$ is represented as the drug entity set and $\widehat{P} = \{p_1, p_2, p_3, \dots, p_i, \dots\}$ is represented as the procedure entity set. And $\widehat{D} = \{d_1, d_2, d_3, \dots, d_i, \dots\}$ is represented as the set of all diseases associated with the patient's syndrome. Then the probability of going to department dep_i can be expressed as follows:

$$\text{Possibility}_{(\text{dep}_i)} = \sum_{d_j \in \widehat{D}} \sum_{f_k \in \widehat{F}} (\text{Weight}_{(f_k, d_j)} \times \text{Weight}_{(d_j, \text{dep}_i)}) \quad (2)$$

In the equation (2), $\widehat{F} = \{\widehat{S}, \widehat{B}, \widehat{DR}, \widehat{P}\}$ represents the patient's syndrome set and $\text{Weight}_{(d_j, \text{dep}_i)} \in [0, 1]$. When there is a relation between disease d_i and department dep_j , $\text{Weight}_{(d_j, \text{dep}_i)} = 1$, otherwise $\text{Weight}_{(d_j, \text{dep}_i)} = 0$.

Nevertheless, there is a problem with the equation (2) resulting in unsatisfactory results during the experiment. For instance, when the patient's chief complaint text is “初潮 12 岁, 周期 30 天, 经期 5-6 天, 脸及背部痤疮, 不经常痛经 (menarche 12 years old, cycle 30 days, menstrual period 5-6 days, face and back acne, infrequent dysmenorrhea)”, there are two symptom entities “脸及背部痤疮 (face and back acne)” and “痛经 (dysmenorrhea)” in the chief complaint text. It is found that there are two diseases related to “face and back acne” and forty-eight diseases related to “dysmenorrhea” after searching in the knowledge graph. The weight of the relationship between each symptom entity and the disease entity can be calculated by the equation (1). We found the disease most related to “face and back acne” is “青春期的疾病 (pubertal disease)” with a weight of 1/9; the disease most related to “dysmenorrhea” is “宫颈上皮瘤 II 级 (cervical epithelioma grade II)” with a weight of 1/8. If the final department is selected according to the equation (2), the recommended departments may be the cervix-vaginal outpatient department, menstrual disease outpatient department, and gynaecological endocrinology outpatient department. However, the result of queried data shows that the patient is treated in the adolescent outpatient department.

The reason that affects the accuracy of the guidance effect is that the symptom entity “dysmenorrhea” often appears in the patient's chief complaint text. So it cannot well reflect which symptoms of the patient are the most characteristic, or the most distinguishing feature of the disease. Therefore, we improve the calculation method of relationship weight,

³<https://github.com/fxsjy/jieba>

reduce the influence of high-frequency feature entities, and enhance the inference effect of low-frequency feature entities. A new calculation method of relation weight named FRMCDR is proposed. As the name implies, the new relation weight method is composed of two parts. The first part is the feature rate, which specifically represents the frequency of occurrence of four types of entities: symptom, body, procedure and drug. The other part is the inverse disease frequency, which is expressed as the importance of the feature in the diseases.

In the following, we describe the FRMCDR algorithm in detail. For example, $\hat{S} = \{s_1, s_2, s_3, \dots, s_i, \dots\}$ is the symptom entity set of the current patient, and $\hat{D} = \{d_1, d_2, d_3, \dots, d_i, \dots\}$ is the set of all diseases related to the patient's syndrome. The $\|\hat{D}\|$ is represented as the number of disease entity contained in the disease set $\hat{D}$, and $\|\hat{D}_{s_i}\|$ is represented as the number of disease entity associated with the patient's symptom s_i . For the symptom s_i , its Feature Rate can be expressed as follows:

$$FR_{(s_i, d_j)} = \exp\left[\frac{\text{Sum}(\text{Rel}(s_i, d_j))}{\sum_{s_k \in \hat{S}} \text{Sum}(\text{Rel}(s_k, d_j))}\right] \quad (3)$$

Then the Converse Disease Rate of symptom s_i is expressed as follows:

$$CDR_{(s_i)} = \log_2 \left(\frac{\|\hat{D}\| + 1}{\|\hat{D}_{s_i}\|} \right) \quad (4)$$

The weight of the relationship between the patient's symptom s_i and the disease d_j is determined as follows:

$$\text{Weight}'_{(s_i, d_j)} = FR_{(s_i, d_j)} \times CDR_{(s_i)} \quad (5)$$

Finally, we reconstruct the equation (2) to the following equation and get the final recommended department score:

$$\text{Possibility}_{(dep_i)} = \sum_{d_j \in \hat{D}} \sum_{f_k \in \hat{F}} \left(\text{Weight}'_{(f_k, d_j)} \times \text{Weight}_{(d_j, dep_i)} \right) \quad (6)$$

We describe the FRMCDR algorithm in the form of pseudocode in Algorithm 1 in order to describe the above algorithm process more clearly.

IV. EXPERIMENT

In this section, we will describe the data situation and the parameter settings of the model in the experiments.

A. DATA COLLECTION

For the training dataset, we use the dataset published by Alibaba Tianchi [32], which contains 8 Chinese tasks, among which the Chinese Medical Named Entity Recognition (CMEE) task dataset was used for model training. It contains nine types of entities, Diseases (Dis), Clinical Manifestations (Sym), Drugs (Dru), Medical Equipment

⁴The CMEE dataset is a publicly accessed dataset, but test set are stored online for testing, so we repartition the dataset

Algorithm 1 FRMCDR Algorithm

Input: Symptom entity set $\hat{S}$; Body entity set $\hat{B}$; Drug entity set $\hat{D}$; Procedure entity set $\hat{P}$; Knowledge Graph G ; Department Score set DSS.

// Calculation

for $\hat{F} = \{\hat{S}, \hat{B}, \hat{D}, \hat{P}\}$ **do**

for all $f_i \in \hat{F}$ **do**

 search from G where node type is $\hat{F}$ and node name is f_i ;

 get disease result set $\hat{D}_{f_i}$;

 compute FR weight score through equation (3);

 compute CDR weight score through equation (4);

 compute Weight' overall weight score through equation (5);

 put Weight' in DSS;

end for

end for

// Sort

for dss $\in$ DSS **do**

 merge and compute possibility of department that patient should go thorough equation (6);

end for

return sort score of departments

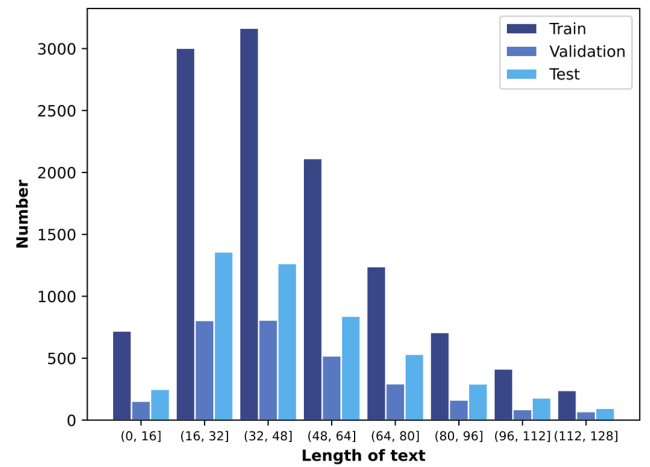


FIGURE 5. Text length statistics in CMEE. The horizontal coordinate is the interval of the text length and ordinate is the corresponding quantity.

(Equ), Medical Procedures (Pro), Body (Bod), Medical Test Items (Ite), Microorganisms (Mic) and Department (Dep). The statistics of the number of entities and the text length on the training, validation and test sets for the entire CMEE dataset are shown in TABLE 2 and FIGURE 5, respectively.

In addition, we collect the consultation data of a regional specialized hospital from January to December 2021. The data include age, gender, weight, patient complaints, brief medical history, physical examination, clinical diagnosis,

TABLE 2. The entity number in the CMeEE dataset.

Entity Class	Train set	Validation set	Test set
Dis(Disease)	12613	3230	4935
Sym(Symptom)	9703	2566	4130
Dru(Drug)	3194	736	1440
Equ(Equipment)	684	204	283
Pro(Program)	5152	1180	2057
Bod(Body)	14099	3589	5883
Ite(Item)	2050	531	923
Mic(Microorganism)	1503	405	584
Dep(Department)	274	74	110

The first column is the abbreviation and name of the entity, and the second, third and fourth columns represent the number of different entities in the training set, validation set and test set.

TABLE 3. A record of a patient in collected data.

	Chinese	English
Chief Complaint text	月经不规则，优思明口服 2 个月	Irregular menstruation, Yousiming oral for 2 months
Medical History text	2 个月前因月经不规则就诊，给予优思明口服 2 个月，复查时，药物治疗期间月经规则。	The patient saw a doctor 2 months ago because of irregular menstruation, and was given oral Yousiming for 2 months. At the time of review, menstruation was regular during drug treatment.
Clinically diagnosed diseases	D-二聚体升高	D-Dimer rise
Department	月经病门诊	Menstrual Outpatient Clinic

department and other information. The original dataset contains a total of 531,374 records in 164 departments. The goal of our experiment is that provide recommendations for the subspecialties of gynaecology. A total of 109,533 records are screened for subspecialties of gynaecology. TABLE 3 shows a sample record of patient, including the chief complaint text, the medical history text, the clinically diagnosed diseases, and the visiting department.

We obtain 239 and 258 symptom phrases of gynaecological diseases from some medical websites, such as xunyiwenyao⁵ and 39health.⁶ It served as an expanded data set to enrich symptom entities and improve the effectiveness of the intelligent department recommendation.

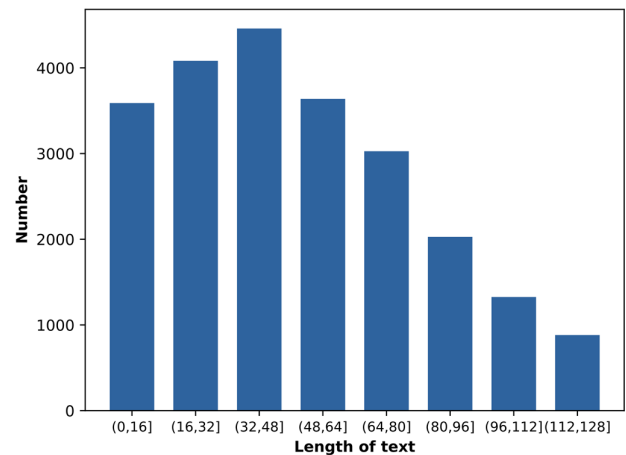
B. DATA PROCESSING

1) DATASET CLEANING

Since some medical records not contain chief complaints or the chief complaint only contains one or two characters, the collected data need to be cleaned. We use the template to filter some invalid text and delete it.

⁵<http://zzk.xywy.com/>

⁶<http://www.39.net/>

**FIGURE 6.** The length statistics of the text we collected.

2) DATASET ANNOTATION AND DIVISION

We used “BIO” (B-begin, I- inside, O-outside) labels to tag different types of entities. In addition, we separated each text with a full stop “。” in Chinese. The training and validation sets of the CMeEE dataset were used for training and validation models. The best performance was selected when validating and making predictions on our dataset.

We divided our collected data into two parts. About 80% of the data is identified by our model and then used to build a knowledge map. The rest of the data is used as a test set to evaluate our approach. In the prediction process, we obtained the predictions for four entities, Sym, Dru, Pro and Bod, and formed a triplet of them with the patient’s corresponding medical department and the ICD-10 of clinically diagnosed diseases. These triplets were used to construct the medical knowledge graph according to FIGURE 4.

3) DEPARTMENT SELECTION

The sub-department under the gynaecological outpatient service is the research goal. We further analyze the data and found that there are some similar departments in function. For example, “Gynecological Endocrinology Outpatient Clinic” and “Famous Gynecological Endocrinology Outpatient Clinic” are the same departments, but one is a non-famous doctor and the other is a famous doctor. In addition, the name of the department contains “Special”, such as “Cervical Vaginosis Special Outpatient Clinic”, which is the same as the name of the corresponding department without the word “Special”. Therefore, we consider merging similar departments whose name is slightly different after consulting experts. For example, “Uterine Fibroids Outpatient Clinic”, “Uterine Fibroids Special Outpatient Clinic” and “Adenomyosis Specialist Outpatient Clinic” were merged into “Uterine Fibroids Outpatient Clinic”. The selection and merging of all departments are shown in TABLE 4. Finally, we obtain 11 departments corresponding to 25,106 records, which detailed length statistics of the text in FIGURE 6.

TABLE 4. Selection and merging of departments.

Original departments	Merge departments
妇科内分泌门诊 Gynecological Endocrinology Outpatient Clinic	妇科内分泌门诊 Gynecological Endocrinology Outpatient Clinic
妇科内分泌名医门诊 Famous Gynecological Endocrinology Outpatient Clinic	
更年期门诊 Menopause Outpatient Clinic	
更年期联合门诊 Menopause Joint Outpatient Clinic	
宫颈阴道病门诊 Cervical Vaginosis Outpatient Clinic	宫颈阴道病门诊 Cervical Vaginosis Outpatient Clinic
宫颈阴道病特需门诊 Cervical Vaginosis Special Outpatient Clinic	
两癌筛查宫颈初诊门诊 Cancer Screening Cervical Initial Outpatient Clinic	两癌筛查宫颈门诊 Cancer Screening Cervical Outpatient Clinic
两癌筛查宫颈复诊门诊 Cancer Screening Cervical follow- up Outpatient Clinic	
子宫肌瘤门诊 Uterine Fibroids Outpatient Clinic	子宫肌瘤门诊 Uterine Fibroids Outpatient Clinic
子宫肌瘤特需门诊 Uterine Fibroids Special Outpatient Clinic	
子宫腺肌病专病门诊 Adenomyosis Specialist Outpatient Clinic	
复发性流产门诊 Recurrent Miscarriage Outpatient Clinic	复发性流产门诊 Recurrent Miscarriage Outpatient Clinic
复发性流产名医门诊 Famous Recurrent Miscarriage Outpatient Clinic	
月经病门诊 Menstrual Outpatient Clinic	
盆底整复门诊 Pelvic Recovery Outpatient Clinic	
妇科肿瘤门诊 Gynecological Tumour Outpatient Clinic	
妇科泌尿门诊 Gynecological Urology Outpatient Clinic	
青春期门诊 Puberty Outpatient Clinic	
多囊卵巢综合征联合门诊 Polycystic Ovary Syndrome Joint Outpatient Clinic	

The first column represents the original departments and the second column is merged departments consolidated. The Chinese department name follows by the corresponding English. For some departments that do not need to be merged, we only show them in the first column of the table, such as the Menstrual Outpatient Clinic.

C. EXPERIMENTAL SETTINGS

We implement our method based on HuggingFace⁷ and Pytorch.⁸ The pre-training models that we use are the base version ($L = 12$, $H = 768$, $A = 12$) as the backbone. And we use a two-layer bidirectional LSTM layer with a hidden layer of 256 dimensions and a CRF layer as a classifier. For all experiments, all models are trained on an NVIDIA TITAN V GPU with a memory size of 12G for 30 epochs and 7k steps with a batch size of 64. We use the AdamW optimizer and linear learning rate warm-up for more than 10% of the training steps. The learning rate is set to $1e-4$ and the dropout rate is 0.5. For the sequence input to the model, we set the longest sequence length of the input text to 128. The vector dimension used to calculate entity similarity is 300. For other hyper-parameter, we follow the default setting. We sample different datasets and repeated experiments on each dataset 3 times with different random seeds. Precision(P), Recall (R) and F₁-score (F₁) are taken as evaluation criteria.

V. RESULTS AND ANALYSES

We use the CMEE dataset as the rich-resource dataset and medical domain corpus to train our model and selected the best checkpoint on the test set to predict the entities on our collected data, which are used to build the knowledge graph. We verify the effectiveness of our proposed FRMCDR algorithm based on the knowledge graph. And then the results of the combination of different mapping methods are obtained based on the FRMCDR algorithm. For all ablation experiments, unless otherwise specified, we compare our model with several models for recommendation by departmental with the base settings.

A. RESULTS

In this section, we compared the effectiveness of our proposed FRMCDR algorithm with the base settings according to the equation (2), and the results were shown in TABLE 5. We also conducted relevant experiments on entity mapping methods. From FIGURE 7, we can see that considering entity mapping can improve the results of guidance.

1) RECOMMENDATION ALGORITHM

The result of different methods on the test set we divided is shown in TABLE 5. It can be seen that the FRMCDR algorithm is better than the original calculation method. Without considering the AC automata, our method can improve precision, recall, and F1-score by 4.27%, 3.63%, and 3.95%, respectively. It shows that the influence of feature frequency needs to be considered in the process of department recommendation. This shows that high-frequency entities can not accurately and directly reflect the characteristics of patients,

⁷<https://github.com/huggingface/transformers>

⁸<https://github.com/pytorch/pytorch>

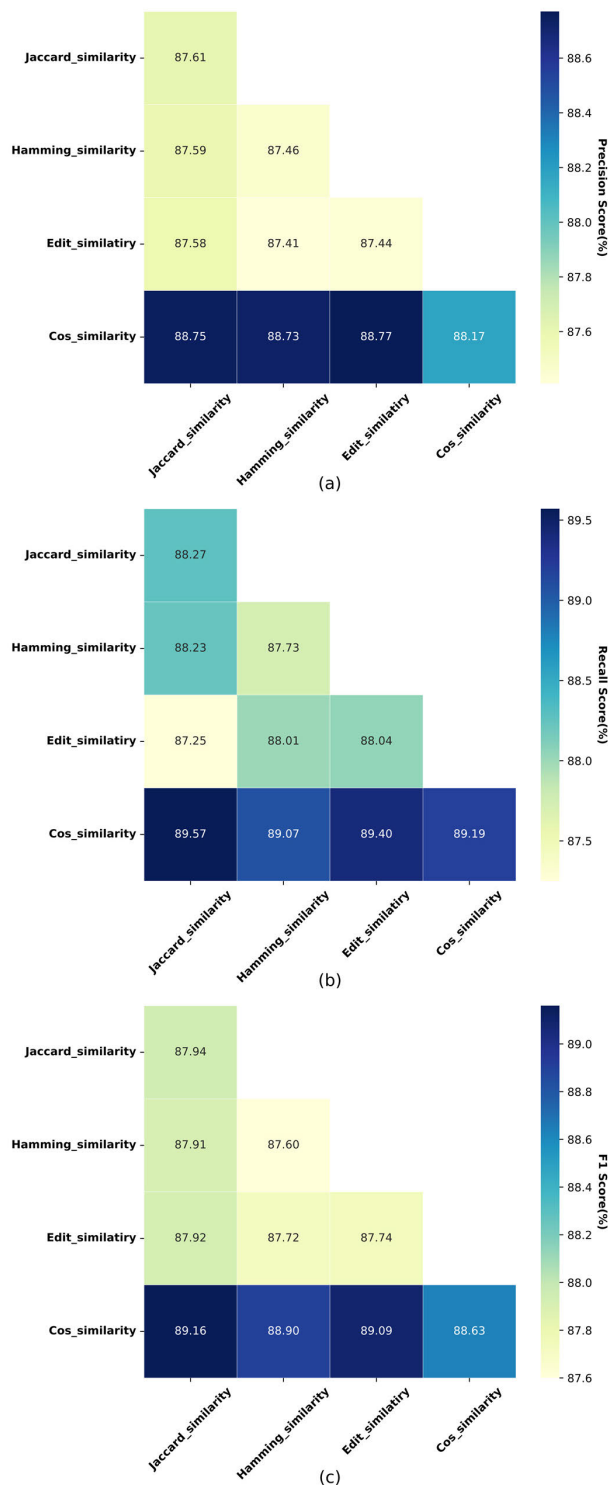


FIGURE 7. Precision-score(a), Recall-score(b), and F1-score(c) of rate for different entity mapping methods. And only one method is used on the diagonal.

but some low-frequency entities can dominate the characteristics of patients, which can be considered distinctive features of patients. After considering the distinctive characteristics of patients, the overall department recommendation effect has been improved.

TABLE 5. The performance of FRMCDR and AC automata algorithm.

AC	WS			FRMCDR		
	P	R	F ₁	P	R	F ₁
w/o	81.06%	82.26%	81.66%	85.33%	85.89%	85.61%
w	82.91%	84.29%	83.59%	87.79%	88.20%	87.99%

The result of Weight Sum (WS) is derived by using the equation (2.) to calculate the weight; the result of FRMCDR is derived by using the equation to calculate the weight; AC represents AC automata algorithm. The P, R, and F1 represent Precision, Recall, and F1-score, respectively⁹.

In addition, we use AC automata to increase symptom entities to solve the model's inability to identify any entity. After adding some entities, the overall effect can be partly improved. It can improve the precision by 2.46%, the recall by 2.31%, and the F1-score by 2.38%.

2) MAPPING METHODS

In the previous section, we discuss results without considering the failure of entity normalization. Now, we will discuss the entity mapping methods from the lexeme level and semantics level to map the entities to the graph.

We compare Jaccard-similarity, Hamming-similarity, and Edit-similarity with the Cos-similarity algorithm in combination. It can be seen from FIGURE 7 that the method that uses Edit-similarity and Cos-similarity algorithms can obtain 88.77% in precision, which is the highest score in precision.

The recall rate can achieve 89.57% by combining the Edit-similarity and Cos-similarity algorithms. When the Cos-similarity method and the Jaccard-similarity method are adopted, the F1-score can reach 89.16%, which is 1.17% higher than the F1-score without using the entity mapping method.

Compared with the methods considering semantic similarity, the three methods of Jaccard-similarity, Hamming-similarity, and Edit-similarity are not ideal, and the best precision rate is 87.61%. Although only using the Cos-similarity method precision rate can reach 88.17%, there are still more than 100 records that cannot be normalized to the entity nodes of the graph. Therefore, it is necessary to consider the similarity of lexeme-level, so that all entities can be normalized to the graph.

From the overall results, the combination of the similarity method at the lexeme level and semantics level, compared with only considering the similarity from one aspect, can further improve the effectiveness of the recommendation of the department.

B. ANALYSES

In the following section, we conduct some ablation experiments to further analyze our approach from effectiveness, adaptability, and influence of entity and feature type.

⁹We have done some post-processing for entities, such as removing "[UNK]" tokens from entities, expanding entities to add semantics, and removing entities with length less than 3 to reduce data noise

TABLE 6. Result for different models.

Model	P	R	F1
BERT-BiLSTM-CRF	78.79%	78.21%	78.50%
RoBERTa-BiLSTM- CRF	78.07%	77.79%	77.93%
BiLSTM-CRF	75.33%	74.45%	74.89%
BERT-CRF	74.97%	72.48%	73.70%
BERT	68.73%	64.57%	65.58%
Ours	79.96%	81.16%	80.56%

The results in the table are calculated based on the equation (2.) and do not consider the failure of entity normalization.

1) EFFECTIVENESS OF MODEL

In order to verify the advantages of our model, we compared the effects of our model on department recommendation with those of several basic models on the test set. We compare BERT-BiLSTM-CRF, RoBERTa-BiLSTM-CRF, BERT-CRF, BERT [9], BiLSTM-CRF and MacNER [7]. BERT, released by Google, is the backbone of BERT-BiLSTM-CRF and BERT-CRF models. And RoBERTa and MacBERT were the Chinese-based full-mask language models released by the joint laboratory of Harbin Institute of Technology and iFLYTEK. The recommended results of the different models mentioned above are shown in TABLE 6. Our model achieves the best results with a 79.96% precision rate, 81.16% recall rate and 80.56% in F1-score.

It can be seen that the precision rate of BERT, RoBERTa and MacBERT with BiLSTM and CRF modules can achieve 78.79%, 78.07% and 79.96%, respectively, which are higher than the BiLSTM-CRF without the pre-trained language model. In addition, the CRF layer is still very helpful for the final result. For BERT, the precision rate of BERT-CRF achieves 74.97%, which is 6.24% higher than only the BERT model. It can prove that BiLSTM and CRF can perform better on downstream tasks than single-language models [20], [21], [22].

2) ADAPTABILITY OF MODEL

We present the performance of different models in TABLE 7 on our test dataset that is described in Section DATASET ANNOTATION AND DIVISION to measure the adaptability and transferability of our model.

There are three metrics defined for convenience of expression. The number of Entity Mapping Failures (NEMF) represents the number of patients' records mapping failures. Similarly, the Number of Vacant Entities (NVE) represents the number of patient records without any entities identified, and the Number of Test Records (NTR) represents the number of patient records that are used to retrieve. For example, for the BERT-BiLSTM-CRF, among the 5,021 patients' records, 423 records do not contain any entity, and among the remaining 4,598 patients' records, 374 records cannot be

TABLE 7. The comparison of different models from the entity level.

Model Name	NEMF	NVE	NTR
BERT-BiLSTM-CRF	374	423	4598
RoBERTa- BiLSTM- CRF	417	491	4530
BERT-CRF	664	1504	3517
BiLSTM-CRF	502	711	4310
BERT	1225	398	4623
Ours	251	397	4624

NEMF, NVE and NTR represent the number of entity mapping failures, the number of vacant entities and the number of test dataset, respectively.

directly mapped to the four types of entity nodes in the graph. As shown in TABLE 7, in our model, only 251 patients' records cannot be directly mapped to the graph and the value of NVE is 397. Compared with the second-ranked model, BERT-BiLSTM-CRF, more than 100 records can be queried for a department recommendation. It not only indirectly explains why our model is better than BERT-BiLSTM-CRF, but also reflects adaptability and transferability. Although the BERT model and our model are similar in NVE, the value of NEMF is more than 1,200, which means that about 30% of the data in the whole test dataset cannot be queried. So the result is dissatisfactory.

3) INFLUENCE OF ENTITY AND FEATURE TYPE

FIGURE 8 shows the number of entity recognitions of different models. It can be observed that the BERT-BiLSTM-CRF model and the RoBERTa-BiLSTM-CRF model have a similar number of entity recognition. The former is only better in symptom recognition than the latter, which indirectly reflects that the effect of the BERT-BiLSTM-CRF model in TABLE 6 is better than the RoBERTa-BiLSTM-CRF model. And the BERT model alone recognizes the fewest entities, so it is worse than the other models.

The BiLSTM-CRF model could recognize 1,400 body entities, which is more than four times the number of other models. We analyze these entities and find that they contained a lot of punctuation marks, and the identification results are not accurate body parts, such as “。 宫”。 Therefore, we believe that the BiLSTM-CRF was not ideally transferable in recognizing body entities, and does not learn more representations and knowledge on another rich-resource dataset. And we also observe the same situation in other types.

We find that the guidance effect would be better if the model identified more symptom entities by comparing the number of different entities. This can also indirectly reflect the rationality of using the AC automaton to increase the symptom entity. Moreover, the effect of the CRF layer on the quantity of entity recognition is obvious. Since the BERT model has no CRF layer, the number of entities recognized is the smallest, and the guidance effect is the worst. In terms

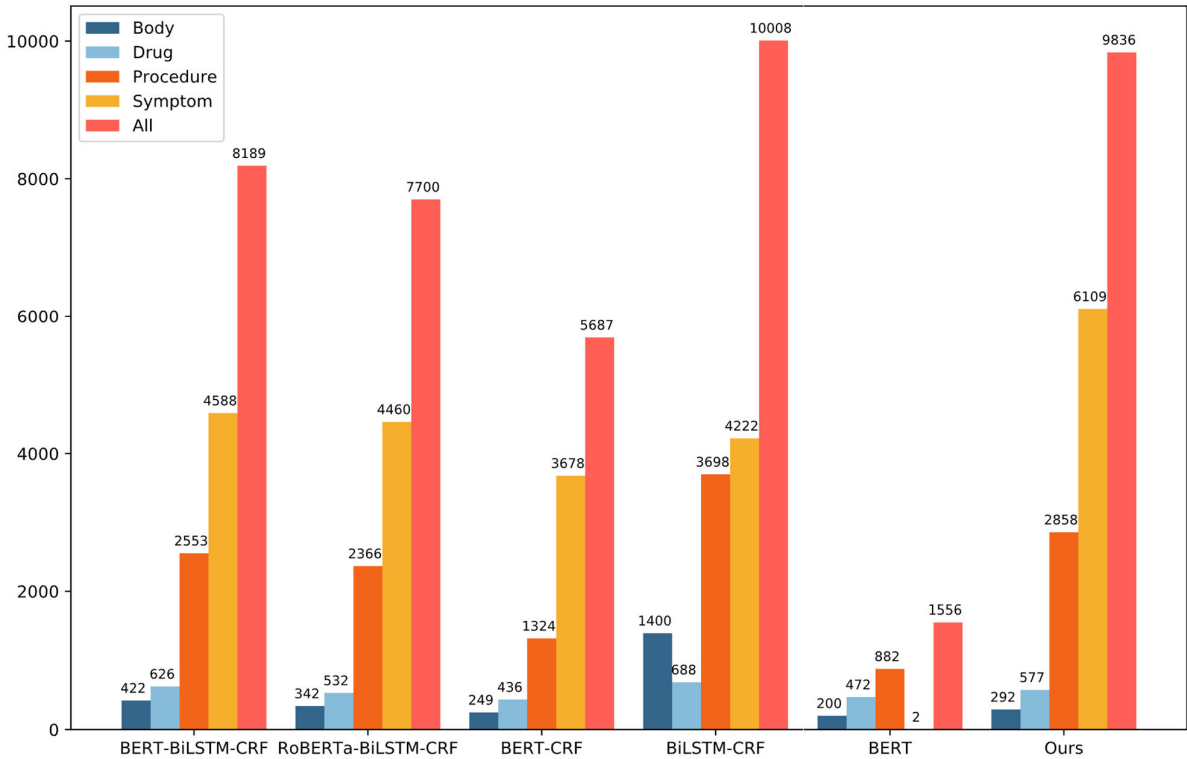


FIGURE 8. Comparison of the number of different types of entities identified by different where All is the sum of the number of all entities.

TABLE 8. Evaluation on our model using different entity types.

Entity type	P	R	F1	NVE
Symptom	86.42%	87.01%	86.72%	1410
Body	87.06%	84.46%	85.74%	1836
Procedure	86.94%	86.81%	86.88%	2345
Drug	74.51%	71.79%	73.13%	4052
All	87.79%	88.20%	87.99%	182

Every row of results represents the use of the corresponding entity type base on the FRMCDR and AC automata algorithm, and the last row uses four types.

of both the overall number and the number of single-category entities, the BERT-CRF model is better than the BERT model.

In addition, we also verified the influence of feature types on the recommendation effect. There are four types of entities extracted from the chief complaint text and medical history, which respective contributions are displayed in TABLE 8. The NVE of combining four types is better than only one type being used. The result proves combination of the chief complaint text and medical history is meaningful for the recommendation. Although the value of NVE is 4,052 when only considering the drug entity, it can obtain 74.51% in precision, reflecting the FRMCDR algorithm’s effectiveness.

VI. CONCLUSION

This paper proposes a department recommendation method based on the medical knowledge graph. The model is trained on a public entity recognition dataset, which adds BiLSTM

and CRF layers to MacBERT, and then it is transferred to predict the several types of entities on patients’ chief complaints and past medical history texts collected from a regional specialized hospital. The FRMCDR relationship weight calculation algorithm is proposed to select the appropriate departments for patients. For the cases where entities cannot be directly mapped to the nodes in the graph, different entity mapping methods are compared. The proposed method obtains a precision rate of 88.77% in the recommendation. It can recommend patients to the corresponding departments, which solves the problem that patients do not know which department to register. The findings from a series of experiments are summarised as follows: the pre-training method of MacBERT does bring certain benefits to downstream tasks, especially in the scenario of model migration; the FRMCDR algorithm captures the weight of features in the recommendation process; different mapping methods can further improve the recommendation effect, of which the Cos-similarity method is crucial; a maximum difference of approximately 13 percent in the effectiveness of recommendations when considering patient characteristics together compared to considering a single characteristic.

In future work, we will continue to consider the recommended methods of other departments because we only focus on gynaecological clinics, but there are other specialized departments. In addition, we will utilize additional medical data and mine information for department recommendations and further improve the performance and availability.

REFERENCES

- [1] D. Wang, Y. Zhang, K. Zhang, and L. Wang, "FocalMix: Semi-supervised learning for 3D medical image detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 3950–3959, doi: [10.1109/CVPR42600.2020.00401](https://doi.org/10.1109/CVPR42600.2020.00401).
- [2] H. Cui and F. Liu, "Design and implementation of intelligent guidance service robot," *Comput. Appl. Softw.*, vol. 37, no. 7, pp. 329–333, Jul. 2020, doi: [10.3969/j.issn.1000-386x.2020.07.055](https://doi.org/10.3969/j.issn.1000-386x.2020.07.055).
- [3] W. Quan, Z. Ma, L. Jie, D. Ye, and H. Zhong, "Interactive intelligent diagnosis guidance system based on medical knowledge graph," *Comput. Syst. Appl.*, vol. 30, no. 12, pp. 55–62, Dec. 2021, doi: [10.15888/j.cnki.csa.008229](https://doi.org/10.15888/j.cnki.csa.008229).
- [4] X. Chen, Y. Yang, J. Zhao, and Y. Wu, "Design and implementation of intelligent medical service robot," *Internet Things Technol.*, vol. 7, no. 11, pp. 70–73, Nov. 2017, doi: [10.16667/j.issn.2095-1302.2017.11.021](https://doi.org/10.16667/j.issn.2095-1302.2017.11.021).
- [5] S. Ya, "Research on construction of precise medical guidance model for online consultation platform," M.S. thesis, Dept. IM., Nanjing Univ., Nanjing, China, 2020.
- [6] X. Zhou, Y. Li, and W. Liang, "CNN-RNN based intelligent recommendation for online medical pre-diagnosis support," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 18, no. 3, pp. 912–921, May 2021, doi: [10.1109/TCBB.2020.2994780](https://doi.org/10.1109/TCBB.2020.2994780).
- [7] Y. Cui, W. Che, T. Liu, B. Qin, and Z. Yang, "Pre-training with whole word masking for Chinese BERT," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 29, pp. 3504–3514, 2021, doi: [10.1109/TASLP.2021.3124365](https://doi.org/10.1109/TASLP.2021.3124365).
- [8] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 1–11. Accessed: Jan. 15, 2022. [Online]. Available: <https://papers.nips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html>
- [9] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," 2018, *arXiv:1810.04805*.
- [10] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, and V. Stoyanov, "RoBERTa: A robustly optimized BERT pretraining approach," 2019, *arXiv:1907.11692*.
- [11] X. Li, H. Zhang, and X.-H. Zhou, "Chinese clinical named entity recognition with variant neural structures based on BERT methods," *J. Biomed. Informat.*, vol. 107, Jul. 2020, Art. no. 103422, doi: [10.1016/j.jbi.2020.103422](https://doi.org/10.1016/j.jbi.2020.103422).
- [12] W. Gao, X. Zheng, and S. Zhao, "Named entity recognition method of Chinese EMR based on BERT-BiLSTM-CRF," *J. Phys., Conf.*, vol. 1848, no. 1, Apr. 2021, Art. no. 012083, doi: [10.1088/1742-6596/1848/1/012083](https://doi.org/10.1088/1742-6596/1848/1/012083).
- [13] X. Zhang, Y. Zhang, Q. Zhang, Y. Ren, T. Qiu, J. Ma, and Q. Sun, "Extracting comprehensive clinical information for breast cancer using deep learning methods," *Int. J. Med. Informat.*, vol. 132, Dec. 2019, Art. no. 103985, doi: [10.1016/j.ijmedinf.2019.103985](https://doi.org/10.1016/j.ijmedinf.2019.103985).
- [14] A. Coden, G. Savova, I. Sominsky, M. Tanenblatt, J. Masanz, K. Schuler, J. Cooper, W. Guan, and P. C. de Groen, "Automatically extracting cancer disease characteristics from pathology reports into a disease knowledge representation model," *J. Biomed. Informat.*, vol. 42, no. 5, pp. 937–949, Oct. 2009, doi: [10.1016/j.jbi.2008.12.005](https://doi.org/10.1016/j.jbi.2008.12.005).
- [15] G. K. Savova, J. J. Masanz, P. V. Ogren, J. Zheng, S. Sohn, K. C. Kipper-Schuler, and C. G. Chute, "Mayo clinical text analysis and knowledge extraction system (cTAKES): Architecture, component evaluation and applications," *J. Amer. Med. Informat. Assoc. JAMIA*, vol. 17, no. 5, pp. 507–513, Sep. 2010, doi: [10.1136/jamia.2009.001560](https://doi.org/10.1136/jamia.2009.001560).
- [16] S. R. Eddy, "Hidden Markov models," *Current Opinion Struct. Biol.*, vol. 6, no. 6, pp. 361–365, 1996, doi: [10.1016/S0959-440X\(96\)80056-X](https://doi.org/10.1016/S0959-440X(96)80056-X).
- [17] A. McCallum, D. Freitag, and F. Pereira, "Maximum entropy Markov models for information extraction and segmentation," in *Proc. 17th Int. Conf. Mach. Learn.*, 2000, pp. 591–598.
- [18] J. D. Lafferty, A. McCallum, and F. Pereira, "Conditional random fields: Probabilistic models for segmenting and labeling sequence data," in *Proc. 18th Int. Conf. Mach.*, San Francisco, CA, USA, 2001, pp. 282–289.
- [19] Z. Liu, M. Yang, X. Wang, Q. Chen, B. Tang, Z. Wang, and H. Xu, "Entity recognition from clinical texts via recurrent neural network," *BMC Med. Inform. Decis. Making*, vol. 17, no. 2, p. 67, 2017, doi: [10.1186/s12911-017-0468-7](https://doi.org/10.1186/s12911-017-0468-7).
- [20] I. J. Unanue, E. Z. Borzeshi, and M. Piccardi, "Recurrent neural networks with specialized word embeddings for health-domain named-entity recognition," *J. Biomed. Inform.*, vol. 76, pp. 102–109, Dec. 2017, doi: [10.1016/j.jbi.2017.11.007](https://doi.org/10.1016/j.jbi.2017.11.007).
- [21] X. Yin, S. Zheng, and Q. Wang, "Fine-grained Chinese named entity recognition based on RoBERTa-WWM-BiLSTM-CRF model," in *Proc. 6th Int. Conf. Image, Vis. Comput. (ICIVC)*, Jul. 2021, pp. 408–413, doi: [10.1109/ICIVC52351.2021.9526957](https://doi.org/10.1109/ICIVC52351.2021.9526957).
- [22] Z. Dai, X. Wang, P. Ni, Y. Li, G. Li, and X. Bai, "Named entity recognition using BERT BiLSTM CRF for Chinese electronic health records," in *Proc. 12th Int. Congr. Image Signal Process., Biomed. Eng. Inform. (CISP-BMEI)*, Oct. 2019, pp. 1–5, doi: [10.1109/CISP-BMEI48845.2019.8965823](https://doi.org/10.1109/CISP-BMEI48845.2019.8965823).
- [23] W. Xiong, J. Cao, X. Zhou, J. Du, B. Nie, Z. Zeng, and T. Li, "Design and evaluation of a prescription drug monitoring program for Chinese patent medicine based on knowledge graph," *Evidence-Based Complementary Alternative Med.*, vol. 2021, pp. 1–8, Jul. 2021, doi: [10.1155/2021/9970063](https://doi.org/10.1155/2021/9970063).
- [24] O. Byambasuren et al., "Preliminary study on the construction of Chinese medical knowledge graph," *J. Chin. Inf. Process.*, vol. 33, no. 10, pp. 1–9, Oct. 2019, doi: [10.3969/j.issn.1003-0077.2019.10.001](https://doi.org/10.3969/j.issn.1003-0077.2019.10.001).
- [25] L. Li, P. Wang, J. Yan, Y. Wang, S. Li, J. Jiang, Z. Sun, B. Tang, T.-H. Chang, S. Wang, and Y. Liu, "Real-world data medical knowledge graph: Construction and applications," *Artif. Intell. Med.*, vol. 103, Mar. 2020, Art. no. 101817, doi: [10.1016/j.artmed.2020.101817](https://doi.org/10.1016/j.artmed.2020.101817).
- [26] M. Rotmensch, Y. Halpern, A. Tlimat, S. Horng, and D. Sontag, "Learning a health knowledge graph from electronic medical records," *Sci. Rep.*, vol. 7, no. 1, Jul. 2017, doi: [10.1038/s41598-017-05778-z](https://doi.org/10.1038/s41598-017-05778-z).
- [27] X. Huang, J. Zhang, Z. Xu, L. Ou, and J. Tong, "A knowledge graph based question answering method for medical domain," *PeerJ Comput. Sci.*, vol. 7, p. e667, Sep. 2021, doi: [10.7717/peerj-cs.667](https://doi.org/10.7717/peerj-cs.667).
- [28] L. Zhang, X. Yang, S. Li, T. Liao, and G. Pan, "Answering medical questions in Chinese using automatically mined knowledge and deep neural networks: An end-to-end solution," *BMC Bioinf.*, vol. 23, no. 1, p. 136, Apr. 2022, doi: [10.1186/s12859-022-04658-2](https://doi.org/10.1186/s12859-022-04658-2).
- [29] A. V. Aho and M. J. Corasick, "Efficient string matching: An aid to bibliographic search," *Commun. ACM*, vol. 18, no. 6, pp. 333–340, Jun. 1975, doi: [10.1145/360825.360855](https://doi.org/10.1145/360825.360855).
- [30] Y. Wu, M. Schuster, Z. Chen, Q. V. Le, M. Norouzi, W. Macherey, M. Krikun, Y. Cao, Q. Gao, K. Macherey, and J. Klingner, "Google's neural machine translation system: Bridging the gap between human and machine translation," 2016, *arXiv:1609.08144*.
- [31] D. Shen, M. Zheng, Y. Shen, Y. Qu, and W. Chen, "A simple but tough-to-beat data augmentation approach for natural language understanding and generation," 2020, *arXiv:2009.13818*.
- [32] N. Zhang, M. Chen, Z. Bi, X. Liang, L. Li, X. Shang, K. Yin, C. Tan, J. Xu, F. Huang, and L. Si, "CBLUE: A Chinese biomedical language understanding evaluation benchmark," 2021, *arXiv:2106.08087*.
- [33] J. Wang, G. Zhang, W. Wang, K. Zhang, and Y. Sheng, "Cloud-based intelligent self-diagnosis and department recommendation service using Chinese medical BERT," *J. Cloud Comput.*, vol. 10, no. 1, p. 4, Jan. 2021, doi: [10.1186/s13677-020-00218-2](https://doi.org/10.1186/s13677-020-00218-2).
- [34] Y. Wang, Y. Tian, L.-L. Tian, Y.-M. Qian, and J.-S. Li, "An electronic medical record system with treatment recommendations based on patient similarity," *J. Med. Syst.*, vol. 39, no. 5, p. 55, May 2015, doi: [10.1007/s10916-015-0237-z](https://doi.org/10.1007/s10916-015-0237-z).



ZHAOJIAN CUI was born in Shandong, China, in January 1999. He received the B.S. degree in computer science and technology from Yantai University, China, in June 2021. He is currently pursuing the M.S. degree in computer application technology with the School of Information Science and Engineering, Hangzhou Normal University, China, in 2021. His research interests include artificial intelligence, machine learning, knowledge graph, and natural language processing, which are applied to the processing and analysis of medical information.



ZHENMING YUAN received the M.S. and Ph.D. degrees in computer application technology from the College of Computer Science and Technology, Zhejiang University, China, in 2001 and 2005, respectively. He is currently a Professor with the School of Information Science and Engineering, Hangzhou Normal University, and the Vice Dean of the Engineering Research Center of Intelligent Healthcare, Ministry of Education, China. His main research interests include artificial intelligence in medical, machine learning, and big data mining. He has published more than 50 papers in major journals and conferences at home and abroad, more than 20 of which are included in related areas.



XIAOYAN SUN was born in Hangzhou, China, in 1980. She received the B.S. degree in biomedical engineering from Zhejiang University, in 2004, the M.S. degree in biomedical engineering from Peking University, in 2007, and the Ph.D. degree in electrical and computer engineering from Old Dominion University, Norfolk, VA, USA, in 2012. From 2004 to 2012, she was a Research Assistant with the MIDA Laboratory, Old Dominion University. In 2013, she joined Hangzhou Normal University, Hangzhou, where she is currently a Lecturer with the School of Information Science and Engineering. She has published more than 20 papers in peer-reviewed journals and conferences. Her research interests include computer-aided surgery and unobtrusive health monitoring.



YINGFEI WU received the B.S. degree in industrial design, the M.S. degree in computer application, and the Ph.D. degree from Zhejiang University, Hangzhou, China, in 1994, 1997, and 2009, respectively. He is currently a Lecturer with the School of Information Science and Technology, Hangzhou Normal University, China. In recent years, he has presided and attended a number of National Natural Science Foundation of China, the Ministry of Education funding in key projects, the Natural Science Foundation of Zhejiang Province, the Key Project of Zhejiang Province, and the Soft Science Project of Hangzhou.

In recent years, he has published many academic research papers in many international journals and conferences. His research interests include medical big data analysis, cross media, computer image analysis, and scientific visualization.



KAI YU received the Ph.D. degree from the College of Computer Science and Technology, Zhejiang University, China, in 2010. Since 2010, he has been with the School of Information Science and Technology, Hangzhou Normal University. His general research interests include machine learning, medical information, and digital library. In particular, his current research interest includes machine learning in medical areas.

...