Original Article

Entity relation extraction from electronic medical records based on improved annotation rules and BiLSTM-CRF

Tingyin Chen[1,2]^, Yongmei Hu[3]

[1]Department of Network and Information, Xiangya Hospital, Central South University, Changsha, China; [2]Mobile Health Ministry of Education-China Mobile Joint Laboratory, Changsha, China; [3]Department of Oncology, Xiangya Hospital, Central South University, Changsha, China

Contributions: (I) Conception and design: T Chen; (II) Administrative support: T Chen; (III) Provision of study materials or patients: Y Hu; (IV) Collection and assembly of data: Y Hu; (V) Data analysis and interpretation: T Chen; (VI) Manuscript writing: Both authors; (VII) Final approval of manuscript: Both authors.

Correspondence to: Yongmei Hu. Department of Oncology, Xiangya Hospital, Central South University, Changsha, China. Email: 366395856@qq.com.

Background: Extracting entities and their relationships from electronic medical records (EMRs) is an important research direction in the development of medical informatization. Recently, a method was proposed to transform entity relation extraction into entity recognition by using annotation rules, and then solve the problem of relation extraction by an entity recognition model. However, this method cannot deal with one-to-many entity relationship problems.

Methods: This paper combined the bidirectional long- and short-term memory-conditional random field (BiLSTM-CRF) deep learning model with an improvement of sequence annotation rules, hided relationships between entities in entity labels, then the problem of one-to-many named entity relation extraction in EMRs was transformed into entity recognition based on relation sets, and entity extraction was carried out through the entity recognition model.

Results: Entity extraction was achieved through the entity recognition model. The result of entity recognition was transformed into the corresponding entity relationship, thus completing the task of one-to-many entity relation extraction by the improved annotation rules, the accuracy rate of proposed method reaches 83.46%, the recall rate is 81.12%, and the value of comprehensive index F1 is 0.8227.

Conclusions: Through the annotation analysis of EMRs, our experimental results show that the improved annotation rules can effectively complete the task of one-to-many medical entity relation extraction from EMRs.

Keywords: Electronic medical record (EMR); relation extraction; annotation rules; entity recognition; deep learning

Submitted Jun 09, 2021. Accepted for publication Aug 26, 2021.
doi: 10.21037/atm-21-3828
View this article at: https://dx.doi.org/10.21037/atm-21-3828

Introduction

Electronic medical records (EMRs) consist of a series of digital graphic data that have been generated during patient diagnosis and treatment by hospitals and other medical institutions. They are stored in hospital databases for easy management and application (1). With the growing digitization of modern health care, the use of big data and artificial intelligence-related technologies to extract clinical information from EMRs while building a medical knowledge base has become an important method in smart medical projects. Entity relation extraction in EMRs is a major research area in information extraction and is an important

^ ORCID: 0000-0002-2365-8281.

technology for building medical knowledge bases.

As computer hardware has improved, deep learning has demonstrated amazing capabilities in various research fields. In the field of relation extraction, more and more researchers have researched deep learning technology. Zeng *et al.* (1) first used a convolutional neural network (CNN) to perform relation extraction tasks on public data sets, which obtained better results than non-neural-network methods. Nguyen and Grishman (2) designed a number of different convolution kernels for experiments and achieved better success on multiple data sets than had previously been attained. Zeng *et al.* (3) further improved the characteristics of the CNN model by expanding it through use of the segmentation method. The multi-instance learning method (3) and the multi-instance multilabel learning method were then combined by Riedel *et al.* and proved to be superior to traditional methods on data sets employed (4). Jiang *et al.* (5) combined word vectors to form sentence vectors and then used CNNs to solve the problem of multiple entity relationships for the same entity pair. Wu *et al.* (6) proposed a feature learning method based on deep learning, using deep sparse automatic coding to re-represent the vector representation of entity contexts, and achieved better entity relation extraction results. dos Santos *et al.* (7) completed a new CNN model by designing a new loss function. When using this model to extract entity relationships, the distinction between different relationship categories can be enhanced. Xu *et al.* (8) added the shortest path method based on the CNN model to improve the effect of relationship classification. Yan *et al.* (9) proposed replacing the traditional recurrent neural network with long- and short-term memory (LSTM) systems for relation extraction. Based on this, Zhang *et al.* (10) used bidirectional LSTM (BiLSTM) to obtain contextual information, thereby achieving better completion of relation extraction tasks. Miwa and Bansal (11) proposed a method of using BiLSTM systems and tree LSTM systems to construct a neural network model for entity relation extraction. Lin *et al.* (12) used the attention mechanism during entity relation extraction tasks and proposed assigning different attention mechanisms to different text contents so that relatively useful information would not be lost. Ning *et al.* (13) proposed a recurrent + transformer neural network architecture based on a multichannel self-attention mechanism to enhance the model's ability to capture sentence-level semantic features, thus improving its ability to learn the characteristics of specialized text found in EMRs. Zeng *et al.* (14) used a third entity as an intermediate entity when extracting the relationship between two entity pairs, by separately constructing the relationship between the two entities and the third entity. The relationship between the two entities was then inferred, thus establishing the reasoning path for the relationship between the two entities. Zhang *et al.* (15) combined CNNs with support vector machines and conditional random fields (CRFs) to construct a joint neural network model, which achieved extremely good results when used on a corpus of medication instructions. Ye *et al.* (16) modeled entity relationships at the semantic level. Zan *et al.* (17) proposed starting from the relevant concepts of entity relation extraction in the medical field to classify deep learning models from different perspectives and then analyzed and discussed the multi-instance learning models of supervised learning and remote supervision based on the construction method for the data sets. Huang *et al.* (18) proposed an entity recognition and entity relation extraction method based on the combination of BiLSTM networks and CRFs, which were then used in the construction and application of medical knowledge graphs. Zhang *et al.* (15) proposed a bidirectional gated recurrent unit (GRU) and dual attention mechanism to identify the medical entity relationship in Chinese EMRs, by using the bidirectional GRU to learn contextual information from words and obtain more fine-grained features.

The combined extraction method for entities and relationships proposed by Zheng *et al.* (19) based on a new annotation mode has expanded the thinking about extracting entity relationships. This method transforms entity relationships into annotation rules and completes the combined extraction of entities and relationships through entity recognition models. This paper first proposes directly modeling the relationship triplets (E_1, R, E_2) and designing a label that includes entity category and relationship category. Using this annotation mode, the relation extraction task is transformed into an annotation task. However, this method cannot solve the one-to-many entity relation extraction task. This article further improves the annotation rules, converting one-to-many relationship annotation into an entity labeling issue. The results from these experiments show that this method can be used to effectively complete the one-to-many entity relation extraction tasks on EMRs.

Methods

Concept definition

To clearly describe the annotation method for entity

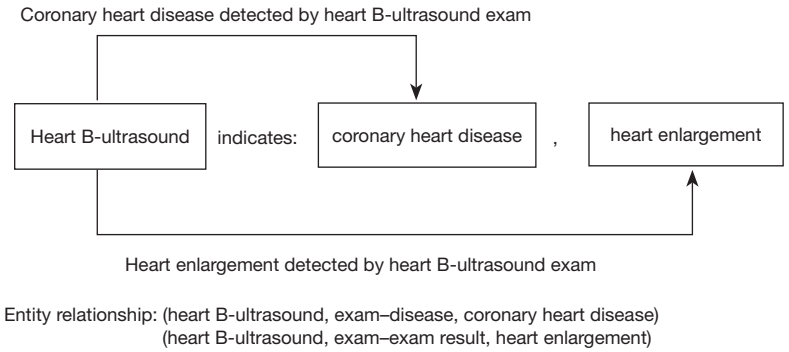

Figure 1 Entity relation extraction.

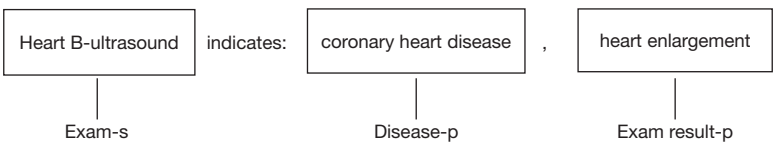

Figure 2 Annotation of relation entities.

relationships, this article first provides the following definitions:

Concept set: medical-related concept set *C* mentioned in the EMRs, such as "disease", "symptom", and "location".

Entity set: All entity sets *E* extracted from EMRs belonging to concept *C*. That is, for any entity *e* ($e \in E$), *e* is an instance of concept *c* ($c \in C$).

Relationship set: the set of relationships *R* between medical concept entities described in EMRs.

Entity relationship: a specific relationship description extracted from EMRs, usually expressed in the form of triplets (subject, predicate, object), such as (e_i, *r*, e_j), where $r \in R$, e_i, $e_j \in E$.

One-to-many relationship: entity set $E=\{e_1,e_2,\ldots,e_n\}$ is extracted from the EMR, if for one of the entities e_i there exists (e_i, r_1, e_a), (e_i, r_2, e_b), (e_i, r_3, e_c), where r_1, r_2, $r_3 \in R$, and e_a, e_b, $e_c \in E$, then entity e_i is said to have a one-to-many entity relationship.

Relation entity category: within the relationship instances (e_i, *r*, e_j), e_i is an instance of c_m, e_j is an instance of c_n ($c_n \in$C), so we have defined that, within the entity relationship (e_i, *r*, e_j), the relation entity type of e_i is c_m-*s*, and similarly, the relation entity type of e_j is c_n-*p*. That is, the agreed relation entity category is composed of the concept that the entity belongs to and *s*/*p*, where *s* (subject) represents the subject in the relationship instance and *p* (predicate) represents the predicate in the relationship instance.

Annotation analysis

Relationships are usually expressed in the form of a relationship triple as (e_1, *r*, e_2), where e_1 is the subject entity of the triplet, e_2 is the object entity of the triplet, and *r* is the relationship between the two entities. As shown in *Figure 1*, in the description from an EMR "*Heart B-ultrasound indicates*: *coronary heart disease*, *heart enlargement*", a relationship of "exam-disease" exists between the examination entity "heart B-ultrasound" and the disease entity "coronary heart disease". Thus, (heart B-ultrasound, exam-disease, coronary heart disease) is an entity relationship, and the relationship category is exam-disease. Likewise, (heart B-ultrasound, exam-exam result, heart enlargement) is also an entity relationship, and the relationship category is "exam-exam result".

By formulating specific annotation rules, the relationship between entities is hidden in entity labels, so that the relation extraction is converted into an entity recognition issue, and the one-to-many relation extraction is solved. As shown in *Figure 2*, the relation entity category label is used as the entity label.

The above figures show that, using the annotation method for the relation entity category, "heart B-ultrasound" will be annotated as "exam-s", "coronary heart disease" will be annotated as "disease-p", and "heart enlargement" will be annotated as "exam result-p". The

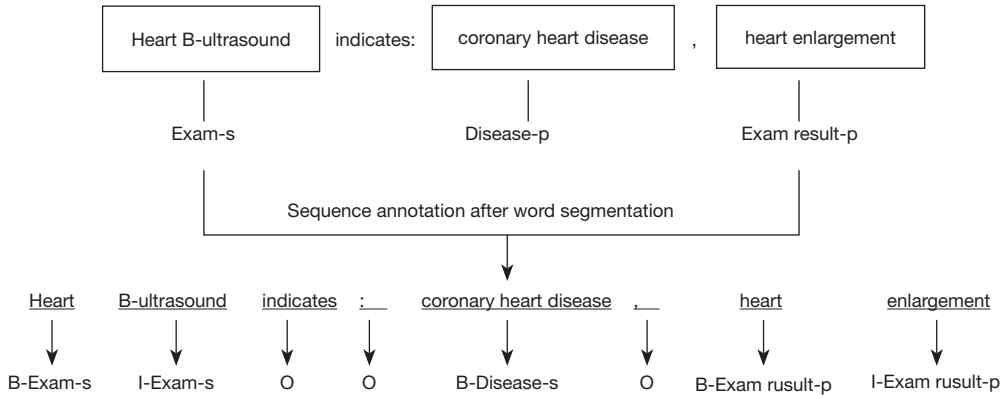

Figure 3 Sequence annotation of the entity relationship.

final entity relation recognition is formed by assembling the entities annotated c_m-s and c_n-p. According to the entity recognition results in the graph, two entity relations can be assembled to complete the one-to-many relationship recognition. At this point, this paper has converted the problem of relation extraction into an entity extraction problem based on relation entity category annotations. That is, specific annotation rules are used to achieve relation extraction, particularly for the problem of one-to-many relation extraction.

Sequence annotation based on BIO

In this article, the BIO (Begin-Intermediate-Other) sequence annotation set method was used in sequence annotation of the EMR text, where B represents the first character/word of the entity, I represents the other characters/words of the entity, and O represents any characters/words other than the entity. Unlike BIEO's (Begin-Intermediate-End-Other) sequence annotation, BIO does not require the existence of the E label (in the BIEO annotation system, E represents the end of the entity). From the perspective of multilabel prediction, the number of prediction categories is reduced, which may improve prediction accuracy. Thus, enhancing prediction accuracy is possible. Before annotation, word segmentation processing is performed on the text, and the BIO annotation set is integrated with the related entity category to complete the sequence annotation of the words in the EMR. The annotation effect is shown in *Figure 3*.

As shown in *Figure 3*, after word segmentation, "Heart B-ultrasound" is divided into two words: "Heart" and "B-ultrasound". "Heart" is annotated as "B-exam-s", and this label indicates that the currently annotated sequence belongs to the partial sequence of the "exam" entity and is the first character/word of the sequence. The entity in which the sequence is located is the subject of the relation entity; similarly, "B-ultrasound" is annotated as "I-exam-s", which indicates that the currently annotated sequence is a partial sequence of the "exam" entity and is a non-header sequence. The entity in which the sequence is located is the subject of the relation entity. After undergoing the BIO sequence annotation process, each character/word is marked with a sequence label.

Character and word vector training

Word vector technology converts words into word embedding vectors. Word vectors can be used as input into the deep network model for making calculations. The theoretical basis of Word2Vec (20) is that when two words have the same or similar meaning, the distance of the corresponding word vector in the vector space will be very close. For example, "China-Beijing" should have a similar spatial distance to the value of "England-London". Word2Vec maps words and word strings into low-dimensional vector spaces by training the word embedding matrix.

To obtain the corpus required for word vector training, we crawled the descriptions of diseases and symptoms in web pages with higher data quality based on rankings in the Baidu Medical Health Network and then used keywords such as "disease" and "symptoms" as the criteria for crawling data in the Baidu Encyclopedia about related symptoms and diseases. After the data were obtained, data cleaning along with other processing was carried out, which enabled the data to meet the criteria for use. Chinese words

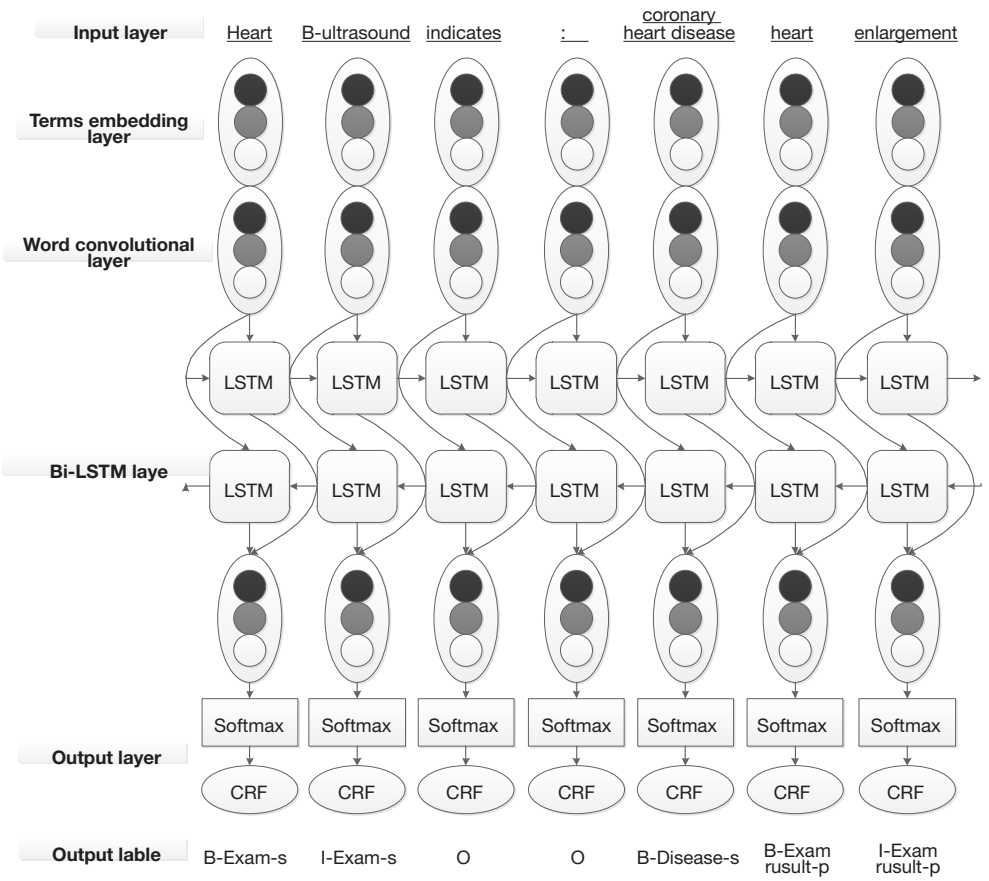

Figure 4 Model structure.

are made up of Chinese characters, and each character that forms part of a word has its own meaning. Moreover, to a certain degree, each individual character can often reflect the meaning of the word as well. Therefore, word vectors play an important role when used to represent Chinese words.

The characters and word vectors used in this article were trained based on the character-enhanced word embedding model of Chen *et al.* (21). Based on the special features involved in the relationship between Chinese words and characters, the continuous bag of words model in Word2Vec was used to train the words and characters at the same time, and the word vectors used in this article were obtained.

Entity recognition model construction

This article mainly uses the BiLSTM-CRF deep learning network model as its entity recognition algorithm. BiLSTM networks can effectively learn the characteristic information of the sequence phrase in context, and the CRF layer can improve the effective combination of the recognition entity sequence using conditional probabilities. When performing entity recognition, a CNN layer is first used for word convolution. After the word information is extracted, it is combined with the word vector as further input, and the sequence information is then extracted using BiLSTM. Finally, the sequence determination is performed through the CRF layer to obtain the prediction result. The structure of the model is shown in *Figure 4*.

The word embedding layer uses the sentences input into the model and converts them into the matrix form required in the neural network based on word vector representation. Assuming that there are n words in the input sentence, the word embedding layer combines all the pretrained word vectors into matrix $S=[w_1w_2\ldots w_n]^T$, where $w_1,w_2,\ldots,w_n$ are the vectors of each word in the sentence, the number of rows in the matrix is the number of words in the input sentence, and the columns of the matrix are the dimensions

of the word vector. Since the sentences can have different numbers of words, during processing the maximum sentence length in the training data set needs to be set to the number of matrix rows, and sentences shorter than this are supplemented with padding to ensure the same matrix dimensions.

Since Chinese words are made up of characters, the meanings of the individual characters will affect the information derived from the words to a certain extent. Therefore, character features can be obtained by performing convolution on the characters, and the features of the characters and the words can be combined to improve the information contained in the word vector, thereby improving the recognition effect of the model. During data preprocessing, as training is conducted to obtain word vectors, character vectors are also obtained. When convolution operations are used to extract the information of character vectors, the convolution kernels used are 2, 3, and 4, and the step size is 1. After performing the related operations of convolution pooling, the features extracted by convolution are obtained. After obtaining the corresponding word convolution features, according to the attention mechanism processing method (22), a dynamic weight matrix trained by a model is used to combine features obtained from the word vector and character convolution, and the newly obtained vector is used as the input for the next layer.

The BiLSTM layer obtains contextual semantic information required by the input words by modeling the contextual information. The output layer of this model is composed of a softmax layer and a CRF layer. The softmax layer normalizes the results passed in by the BiLSTM layer, and then the CRF further performs category constraints, handles some of the more obvious category errors, and finally outputs the label results predicted by the model. After the label for each word is output, it is restored to the corresponding relationship category based on the implementation-defined relationship set and the relation extraction category.

To prevent model overfitting, a dropout mechanism is added to the model. Specifically, a control mechanism is added to the hidden layer of neurons. During the training of neurons, the work of some neurons is randomly halted in order to create a network structure during training that is different. This is equivalent to training a combination model of multiple neural networks, and the parameters of each model are fewer than the total model parameters, thereby effectively preventing overfitting.

Statistical analysis

The chi-square test was used to compare the corpus entities of different symptoms, different parts, different examination results and different diseases, $P<0.05$ was statistically significant. Statistical analysis of the characteristics of electronic medical records was performed using SPSS software (version 22.0, IBM Corporation).

Results

The experimental data in this article come from 200 articles each from the fields of nephrology, cardiology, gastroenterology, respiratory medicine, and gynecology, totaling 1,000 EMRs. All the data have been desensitized. Annotated concept set *C={"symptom", "disease", "exam", "exam results", "location"}* extracts relationship set *R={"exam-disease", "exam-exam result", "location-symptoms", "location-examination results"}*, such that the annotated related entity categories are *{"exam-s", "disease-p", "examination result-p", "location-s", "symptom-p"}*. After the annotation and review of 1,000 medical records, this article divides the annotation results into two parts; 80% of the corpus results are used for training, and 20% of the corpus results are used for verification. Moreover, they are randomly divided into 10 experiments, and their average value was obtained.

Experimental parameter settings

Many parameters need to be set in the neural network model, and these parameters will affect the training results of the model. In accordance with the requirements of the control variable method, when testing a variable, other variables remain unchanged and are then compared to obtain the parameter data (*Table 1*). It is assumed that all parameters are independent of each other, and mutual influence is not taken into consideration.

Experimental evaluation

After annotating all the data sets, we gather statistics on the relevant information of the data sets (*Table 2*), where sentences refers to the number of sentences counted after the medical record text is segmented; words refer to the number of words counted after word segmentation of the medical record text; characters are the number of words in the medical records; entities refer the count of all the entries in the medical record text that are marked as

relation entities; and relationships are the number of entity relationships in the medical records.

The experiment uses general classification evaluation indicators: accuracy (P), recall (R), and F1 value:

$$P = \frac{TP}{TP + FP} \quad [1]$$

$$P = \frac{TP}{TP + FN} \quad [2]$$

$$F1 = \frac{2 \times P \times R}{P + R} \quad [3]$$

For a certain type of entity, TP is the number of entities that are correctly classified into this type of entity, FP is the number of entities that are incorrectly classified into this type of entity, and FN is the number of entities that are classified into other types of entities.

Experimental results

By serving as the recognition factor for relation entities, the entity itself contains the entity category while also concealing the relational attributes. For example, the entity recognized as "exam-s" is not only an entity with the concept of "exam" but is also an entity that is the subject of the relationship in the current sentence. Therefore, we first analyze the recognition results of entity categories (*Table 3*).

In *Table 3*, "All" means all entities identified. The accuracy P of the entities identified by the algorithm is 90.76%, the recall rate R is 91.40%, and the value of comprehensive index F1 is 0.9108. From the analysis of comprehensive indicators, the recognition effect of the entity with the concept of "location" is good, the F1 value reaches 0.9447, but the recognition effect of "exam result" is rather poor, with an F1 value of 0.8240, which is mainly because "location" in the EMRs is generally more standardized and the description range of the location is relatively narrow, while the description of the "exam result" entity is generally more complex.

After identifying these relationship entities, the entities in each sentence are then associated and integrated with subject, predicate, and object based on the definitions of the relationship categories so as to extract the relationship instances between the entities, including one-to-many entity relationships.

According to the experimental methods applied to the

Table 1 Parameter adjustment table

Parameter type	Optimal	Test range
Word embedding dimension	100	50–300
Character embedding dimension	100	50–300
CNN convolution kernel size	2, 3	2–7
CNN output size	200	100–300
LSTM output size	300	100–300
Learning rate	0.001	0.1–0.001
Minibatch size	20	10–50
Dropout value	0.5	0.5–1

CNN, convolutional neural network; LSTM, long- and short-term memory.

Table 3 Analysis of entity recognition results

Category	P (%)	R (%)	F1
All	90.76	91.40	0.9108
Disease	83.76	84.09	0.8491
Symptoms	93.36	93.33	0.9335
Location	94.26	94.68	0.9447
Exam	88.27	89.04	0.8865
Exam results	82.29	82.51	0.8240

Table 2 Statistics of the experimental data sets

Data category	200 articles	800 articles	1,000 articles	Average
Sentences	6,799	27,931	34,610	34.61
Words	39,016	154,112	192,980	192.98
Characters	191,332	761,332	950,600	950.60
Entitles	10,198	40,329	50,190	50.19
Relationships	2,566	9,688	12,130	12.13

training set and the verification set, the entity relationship in the verification set is used as the verification standard. In this article, the extracted entity relationship and the entity relationship in the verification set are matched and analyzed. The analysis and statistics results are shown in *Table 4*.

From the experimental results in *Table 4*, it can be seen that the overall effect of entity relation extraction is lower than that of entity extraction because the recognition of the relationship is based on further assembly, which in turn is based on the extraction of the relation. However, during the assembly process, certain discrepancies exist, so the recognition accuracy of entities serves as a prerequisite for relation extraction. From the overall effect of entity relation extraction, the accuracy rate reaches 83.46%, the recall rate is 81.12%, and the value of comprehensive index F1 is 0.8227. The results from this experiment show that the extraction of entity-to-entity relationships can be achieved through specific annotation rules, and the problem of one-to-many entity relation extraction can be effectively solved.

Table 4 Analysis of entity relation extraction results

Category	P (%)	R (%)	F1
All	83.46	81.12	0.8227
Exam-disease	82.11	81.21	0.8166
Exam-exam result	80.21	79.92	0.8006
Location-symptoms	85.19	86.23	0.8571
Location-exam results	81.14	80.66	0.8090

Discussion

Table 5 compares the entity recognition performance of different models under different annotation systems. From the comparison results, the effects of the models under the BIO annotation system are better than those under the BIEO annotation system. Under different deep learning systems, the effect of the iterated dilated CNNs and CRFs (IDCNN + CRF) model is slightly better than the effect of BiLSTM + CRF. As shown in *Table 6*, in order to raise the effectiveness of entity relation extraction, a very important approach involves improving the effect of entity extraction. This article has focused on improving the effect of entity extraction by enhancing annotation systems and selecting different models. From the overall comparison result, the annotation system using BIO is better than the annotation system of BIEO. Under the same annotation system, the effect of using the IDCNN + CRF model is better than that of BiLSTM + CRF.

Conclusions

This paper proposes a method of transforming entity relation extraction into entity recognition. At the same time, by improving annotation rules, the one-to-many entity relationship is transformed into an entity annotation problem based on relation sets. Annotation and recognition experiments were carried out using EMR data sets. The results of these experiments show that this method can effectively extract specific medical entity relations from EMRs. The proposed method also provides an effective solution to the many-to-many relation extraction requirements for future projects.

Table 5 Performance comparison of entity recognition results under different systems

Category	BIO						BIEO					
	BiLSTM + CRF			IDCNN + CRF			BiLSTM + CRF			IDCNN + CRF		
	P (%)	R (%)	F1	P (%)	R (%)	F1	P (%)	R (%)	F1	P (%)	R (%)	F1
All	92.03	91.94	0.9201	91.62	91.52	0.9183	90.76	91.40	0.9108	90.92	91.75	0.9178
Disease	86.52	85.87	0.8671	86.76	85.95	0.8701	83.76	84.09	0.8491	83.92	85.12	0.8508
Symptoms	94.06	93.83	0.9425	94.26	93.97	0.9375	93.36	93.33	0.9335	93.86	92.53	0.9401
Location	96.06	96.88	0.9577	96.26	96.68	0.9527	94.26	94.68	0.9447	95.16	95.08	0.9497
Exam	88.78	90.04	0.8975	89.21	89.84	0.8885	88.27	89.04	0.8865	88.36	89.46	0.8950
Exam results	85.25	84.95	0.8397	84.19	84.07	0.8382	82.29	82.51	0.8240	83.01	82.94	0.8259

BIO, Begin-Intermediate-Other sequence annotation; BIEO, Begin-Intermediate-End-Other sequence annotation; BiLSTM, bidirectional long- and short-term memory; CRF, conditional random field; IDCNN, iterated dilated convolutional neural network.

Table 6 Performance comparison of entity relation extraction results under different systems

Category	BIO						BIEO					
	BiLSTM + CRF			IDCNN + CRF			BiLSTM + CRF			IDCNN + CRF		
	P (%)	R (%)	F1	P (%)	R (%)	F1	P (%)	R (%)	F1	P (%)	R (%)	F1
All	84.76	82.62	0.8387	84.36	82.31	0.8371	83.46	81.12	0.8227	83.61	82.02	0.8269
Exam-disease	84.14	82.72	0.8286	83.91	82.61	0.8296	82.11	81.21	0.8166	82.82	81.98	0.8206
Exam-exam results	82.12	81.42	0.8260	81.81	80.92	0.8206	80.21	79.92	0.8006	80.81	79.72	0.8010
Location-symptoms	85.99	87.51	0.8691	86.79	87.93	0.8691	85.19	86.23	0.8571	84.89	86.53	0.8681
Location-exam results	83.42	82.86	0.8197	81.82	82.46	0.8195	81.14	80.66	0.8090	81.74	81.06	0.8079

BIO, Begin-Intermediate-Other sequence annotation; BIEO, Begin-Intermediate-End-Other sequence annotation; BiLSTM, bidirectional long- and short-term memory; CRF, conditional random field; IDCNN, iterated dilated convolutional neural network.

Acknowledgments

Funding: Mobile Health Ministry of Education, China Mobile Joint Laboratory Project, Research and Application of DRGs Grouping System Based on Big data (No. 2020MHL02015).

Footnote

Conflicts of Interest: Both authors have completed the ICMJE uniform disclosure form (available at https://dx.doi.org/10.21037/atm-21-3828). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. Zeng D, Liu K, Lai S, et al. editors. Relation classification via convolutional deep neural network. COLING 2014 - 25th International Conference on Computational Linguistics, Proceedings of COLING 2014; 2014.
2. Nguyen TH, Grishman R. editors. Relation extraction: Perspective from convolutional neural networks. Workshop on Vector Space Modeling for Natural Language Processing; 2015.
3. Zeng D, Liu K, Chen Y, et al. editors. Distant Supervision for Relation Extraction via Piecewise Convolutional Neural Networks. Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing; 2015 Sep; Lisbon, Portugal: Association for Computational Linguistics.
4. Riedel S, Yao L, McCallum A. editors. Modeling Relations and Their Mentions without Labeled Text. Proceedings of Joint European Conference on Machine Learning and Knowledge Discovery in Databases; 2010; Berlin, Heidelberg: Springer Berlin Heidelberg.
5. Jiang X, Wang Q, Li P, et al., editors. Relation Extraction with Multi-instance Multi-label Convolutional Neural Networks. Proceedings of COLING 2016,the 26th International Conference on Computational Linguistics; 2016 Dec; Osaka, Japan: The COLING 2016 Organizing Committee.
6. Wu J, Guan Y, Lu X. A Deep Learning Approach in Relation Extraction in EMRs. Intelligent Computer and Applications 2014;4:35-38+41.
7. dos Santos C, Bing X, Zhou B. Classifying Relations by Ranking with Convolutional Neural Networks. Computer Science 2015;86:132-7.
8. Xu K, Feng Y, Huang S, et al. Semantic Relation Classification via Convolutional Neural Networks with Simple Negative Sampling. Computer Science 2015;71:941-9.

9. Yan X, Mou L, Li G, et al. Classifying Relations via Long Short Term Memory Networks along Shortest Dependency Paths. Computer Science 2015;42:56-61.
10. Zhang S, Zheng D, Hu X, et al. editors. Bidirectional Long Short-Term Memory Networks for Relation Classification. Proceedings of the 29th Pacific Asia Conference on Language, Information and Computation; 2015 Oct; Shanghai, China.
11. Miwa M, Bansal M. editors. End-to-End Relation Extraction using LSTMs on Sequences and Tree Structures. Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics; 2016 Aug; Berlin, Germany: Association for Computational Linguistics.
12. Lin Y, Shen S, Liu Z, et al. editors. Neural Relation Extraction with Selective Attention over Instances. Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics; 2016 Aug; Berlin, Germany: Association for Computational Linguistics.
13. Ning S, Teng F, Li T. Multi-Channel Self-Attention Mechanism for Relation Extraction in Clinical Records. Chinese Journal of Computers 2020;43:916-29.
14. Zeng W, Lin Y, Liu Z, et al. editors. Incorporating Relation Paths in Neural Relation Extraction. Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing; 2017 Sep; Copenhagen, Denmark: Association for Computational Linguistics.
15. Zhang Y, Liu M, Hu H. Chinese medical entity classification and relationship extraction based on joint neural network model. Computer Engineering and Science 2019;41:1110-8.
16. Ye H, Chao W, Luo Z, et al. Jointly Extracting Relations with Class Ties via Effective Deep Ranking. ACL 2017:1810-20.
17. Zan H, Guan T, Zhang K, et al. Review of entity relation extraction for medical text. Journal of Zhengzhou University (Natural Science Edition) 2020;52:1-15.
18. Huang M, Li M, Han H. Research on entity recognition and knowledge graph construction based on electronic medical records. Application Research of Computers 2019;36:3735-9.
19. Zheng S, Wang F, Bao H, et al. editors. Joint Extraction of Entities and Relations Based on a Novel Tagging Scheme. Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics; 2017 July 30 - August 4; Vancouver, Canada.
20. Mikolov T, Sutskever I, Chen K, et al. Distributed Representations of Words and Phrases and their Compositionality. Advances in Neural Information Processing Systems 2013:3111-9.
21. Chen X, Xu L, Liu Z, et al. editors. Joint learning of character and word embeddings. IJCAI'15: Proceedings of the 24th International Conference on Artificial Intelligence; 2015. AAAI Press.
22. Rei M, Crichton G, Pyysalo S. Attending to Characters in Neural Sequence Labeling Models. Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers, 2016:309-18.

Cite this article as: Chen T, Hu Y. Entity relation extraction from electronic medical records based on improved annotation rules and BiLSTM-CRF. Ann Transl Med 2021;9(18):1415. doi: 10.21037/atm-21-3828

