Journal of Combinatorial Mathematics
and Combinatorial Computing
www.combinatorialpress.com/jcmcc

Combinatorial Press

Construction of Communication Strategies Based on Natural Language Processing in the Discourse Model of Doctor-Patient Interaction

Shui Cao[1], Chunjun Cheng[2,✉], Guangyan Tang[3], Fang Ma[3], Yu Sun[4], Di Cui[4], SAGGELLA MADHUMITHA[2]

[1] *College of Medical Humanities, Jinzhou Medical University, Jinzhou, Liaoning, 121000, China*

[2] *College of International Education, Jinzhou Medical University, Jinzhou, Liaoning, 121000, China*

[3] *School of Computer Science, Jinzhou Normal College, Jinzhou, Liaoning, 121000, China*

[4] *Department of Radiology, The First Affiliated Hospital of Jinzhou Medical University, Jinzhou, Liaoning, 121000, China*

ABSTRACT

This paper defines doctor-patient interaction from the perspectives of interaction form and maintenance of patients' health respectively, and also constructs a doctor-patient interaction discourse model. Based on the data mining technology to obtain the research data, the acquired data are preprocessed and stored in the form of dataset. Bi-LSTM is used to extract topic sentence features from the dataset, and the unsupervised pattern is transformed into a self-supervised pattern through the training and learning of auxiliary tasks to complete the construction of the discourse model of doctor-patient interaction based on topic structure. Combined with the processing flow of natural language processing and semantic technology, the communication strategy generation system for doctor-patient interaction discourse is designed, and finally the communication strategy based on natural language technology is researched and analyzed. There are significant differences between the experimental group and the control group in terms of expression ability and cognitive level ($P<0.05$), which concludes that compared with the traditional discourse model, the doctor-patient interactive discourse model has a higher priority, and it can effectively improve the expression ability and cognitive level of the patients' medical terminology. On the CMedQA2.0 dataset, the average performance of this paper's model is improved by 46.34% compared with the baseline model GPT-2, indicating that this paper's model has excellent performance. Under the condition of Chinese participle and topic extraction fusion, the average accuracy of this paper's system is as high as 85.02%, which indicates that the system can provide doctors with precise communication strategies based on patients' medical-related information, thereby effectively enhancing the discourse communication skills in doctor-patient interactions.

✉ Corresponding author.
E-mail address: jzneil@163.com(C. Cheng).
Received 01 January 2024; Revised 10 March 2024; Accepted 20 December 2024; Published Online 15 April 2025.
DOI: 10.61091/jcmcc127a-175

Keywords: natural language technology, topic sentence extraction, strategy generation system, interactive discourse models

1. Introduction

Doctor-patient communication is a crucial part of modern medical service. Good doctor-patient communication can enhance the trust and understanding between doctors and patients, so as to better serve the patients. However, doctor-patient communication also faces many challenges, one of the biggest challenges is the language barrier [1-4]. Doctors and patients come from different cultural and linguistic backgrounds, which often leads to communication difficulties. With the rapid development of the Internet and artificial intelligence technology, the healthcare field also faces new challenges and opportunities [5-7]. In traditional healthcare services, patients usually need to go to the hospital in person for consultation, which not only wastes time and money, but also increases the burden of the hospital. Therefore, it has become an urgent task to construct a doctor-patient communication strategy based on natural language processing [8-10].

Natural language processing technology is a kind of artificial intelligence technology that analyzes and understands human language. It can transform human language into a form that can be understood and processed by computers, thus realizing the interaction and communication between computers and humans [11-13]. Natural language processing-based doctor-patient communication strategies can utilize advanced artificial intelligence technology to help patients communicate better with doctors and get timely and accurate medical guidance and advice [14-16]. This can not only improve the efficiency and quality of medical services, but also reduce the burden of hospitals and realize the rational allocation of medical resources. Therefore, the construction of communication strategies based on natural language processing has important practical significance and application value [17-19].

This paper designs the doctor-patient interaction discourse model to provide theoretical support for the implementation of the subsequent research work. Data mining technology is used to collect the research data of this paper, and data preprocessing is carried out to avoid the existence of interfering information in the data. Using Bi-LSTM to encode the topic structure and select the candidate abstract with the highest probability as the initial abstract, the construction of the doctor-patient interaction discourse model based on the topic structure is completed through the training and learning of auxiliary tasks, and the language modeling loss is also added to the model to make the model output discourse more fluent and easy to understand. Combining natural language processing and the overall design structure of the system to realize the doctor-patient interaction discourse communication strategy generation system. The research object, environment configuration, and experimental tools are identified to explore the communication strategies based on natural language technology.

2. Communication strategy construction based on natural language technology

2.1. Definition of doctor-patient interaction and its discourse patterns

2.1.1. Definition of doctor-patient interaction. In terms of the form of interaction, doctor-patient interaction refers to an interpersonal relationship formed between doctors and patients through verbal, behavioral, and psychological interactions based on the health and interests of the patient, which is not a unidirectional, independent process, but rather a weekly process [20-21]. The current main goals in the field of doctor-patient interaction include the following three aspects: to establish a good interpersonal relationship between the interacting subjects, to promote the adequate exchange of information between doctors and patients, and to allow patients to actively participate in relevant decisions. From the perspective of maintaining patients' health, doctor-patient interaction

significantly affects the establishment of harmonious mutual trust between doctors and patients, and it integrates the interactions of emotional, psychological, cognitive, attitudinal and behavioral interactions between doctors and patients.

2.1.2. Discourse Patterns of Doctor-Patient Interaction. Traditional doctor-patient interaction is the offline face-to-face communication between doctors and patients, although face-to-face interaction can achieve verbal and non-verbal forms of interaction, to enhance the affinity between doctors and patients, pulling the distance between doctors and patients, but there are certain drawbacks: not able to overcome the time and geographic constraints, especially in areas where medical and health care resources are relatively scarce, it takes a lot of time and financial resources to get access to medical care. Due to the specificity of the disease itself, patients have more or less sense of shame, face-to-face consultation rather make it more difficult for patients to accurately and objectively describe their own conditions, missing the best treatment period. Offline consultation through verbal language interaction, it is inevitable that there are patients' memory bias and can not get the doctor's professional answer again, there is the possibility of ambiguous medical advice, which is not conducive to the patient's treatment and recovery.

Compared with the traditional doctor-patient interaction, the doctor-patient interactive discourse mode has obvious characteristics. Interaction content is richer and more objective. In traditional doctor-patient interaction, the content of face-to-face conversation between doctors and patients is only based on the memory of each doctor and patient, and the interactive content of the medical record is richer and more objective. In traditional doctor-patient interaction, the content of face-to-face conversations between doctors and patients is only based on the memory of each doctor and patient, and the characteristics of the disease and medication of the patient are recorded in the medical record, which can be accessed by the doctor and patient at any time to make up for the problem that the content of the diagnosis and treatment cannot be clearly remembered by the patient during the offline consultation and to realize a more accurate and comprehensive information exchange and knowledge exchange between doctors and patients.

2.2. Data Acquisition and Preprocessing

2.2.1. Data acquisition. In this paper, we choose a well-known online medical community in China as the research object, and write a Python crawler program to obtain the relevant data from September 13, 2022 to October 13, 2022 based on the front-end structure design of the platform's website. Due to the nature of the website, the personal information related to doctors and patient-generated review data are presented on the doctors' personal homepage or an interface is provided on the personal homepage. Therefore, to achieve large-scale crawling of the data in this platform, firstly, we need to obtain the URL of the doctor's personal homepage, and the best way to obtain the URL of the doctor's personal homepage is to batch obtain the URL of the doctor's personal homepage from the recommended doctor page of a certain disease according to the type of the disease as the classification criterion; secondly, we enter the page through the obtained homepage URL, analyze the webpage structure of the doctor's personal homepage, and write regular expressions to obtain part of the doctor's Secondly, we analyze the structure of the doctor's homepage through the obtained homepage URL and write a regular expression to obtain some data of the doctor, including the doctor's title, education and hospital level, as well as the number of online consultations, the number of visits to the homepage, the recommendation popularity and the number of awards. All the data obtained above are divided into two datasets: doctor's personal information and patient's comments, and in order to protect doctor's privacy, avoid problems such as repetition of doctor's name affecting the results of the subsequent study, and to facilitate the merging of the relevant data of the two datasets, this paper replaces the doctor's name with the doctor's ID as a primary key that can uniquely represent the doctor's individuality.

This study used Python web crawlers to collect publicly accessible doctor-patient interaction data, strictly complying with platform service agreements and relevant laws. The collected data did not involve personal privacy or sensitive information. The crawler program was optimized with frequency limits to avoid disrupting platform operations. All data were used solely for academic research, ensuring legality and compliance.

2.2.2. Data pre-processing. After all data acquisition is completed, data cleaning and pre-processing is performed. This study utilizes MySQL database for data storage, cleaning, querying and providing the interface that Python needs to call when analyzing the data. While ensuring the accuracy of data processing, it greatly improves the efficiency of each link. On the one hand, the data in the doctor's personal information dataset with multiple empty or zero key fields are deleted, and 6232 doctor's personal information data records are finally saved. On the other hand, it deletes comments in the patient comment data set that have less than 6 Chinese characters or contain illegal characters or emoji expressions that interfere with subsequent text processing. Moreover, the same comments made by the same patient to the same doctor are deleted based on the patient ID and other fields that can uniquely represent the data, and 217,605 text comment data are ultimately obtained.

2.2.3. Textual disambiguation. In this study, we use Jieba, a lexicon for Chinese natural language processing, to perform lexical processing on review data. The principle of lexical processing is to convert text into a probabilistic language model, traverse the word map to find all word combinations, and then form a directed acyclic graph (DAG), which is used to further maximize the cut combinations based on the frequency of the words. jieba allows the developer to set up his own lexicon, add deactivation and forced participles, and expand words and phrases that are not included in the lexicon to improve the accuracy of lexical processing. And forced segmentation, expanding words not included in the lexicon and words and phrases that cannot be segmented to improve the accuracy of segmentation. Therefore, in this study, a special dictionary was constructed to add doctor's name, patient's name, polite expressions, onomatopoeic words and other high-frequency participles not related to this study to the deactivation dictionary, etc., to reduce word missegmentation and unsegmentation.

2.3. A Discourse Model of Doctor-Patient Interaction Based on Thematic Structure

The self-supervised learning model of doctor-patient interaction discourse based on topic structure will be highlighted, and the structure of the model is shown in Fig. 1, which divides the process into two modules. The topic sentence extraction module labels the topic of each sentence, and due to the characteristics of the doctor-patient scenario, there are only a limited number of topic types (personal information, symptoms, present medical history, past medical history, examination results, medication). Sentences with the same topic are put together to form a topic-structured text. The self-supervised summary generation module constructs two auxiliary tasks, diagnosis extraction generation (Task1) and diagnosis classification (Task2). The unsupervised model is transformed into a self-supervised model by training and learning from the auxiliary tasks. Among them, the initial abstracts need to be generated when the self-supervised abstract generation module is initialized. The extraction of initial summaries can be viewed as a binary classification task of sentences, which will be extracted as a replicative summary, using Bi-LSTM to encode the topic structure and select the candidate summary with the highest probability as the initial summary. After obtaining the initial summary, the optimization of the summary is guided by the learning of the two auxiliary tasks mentioned above.

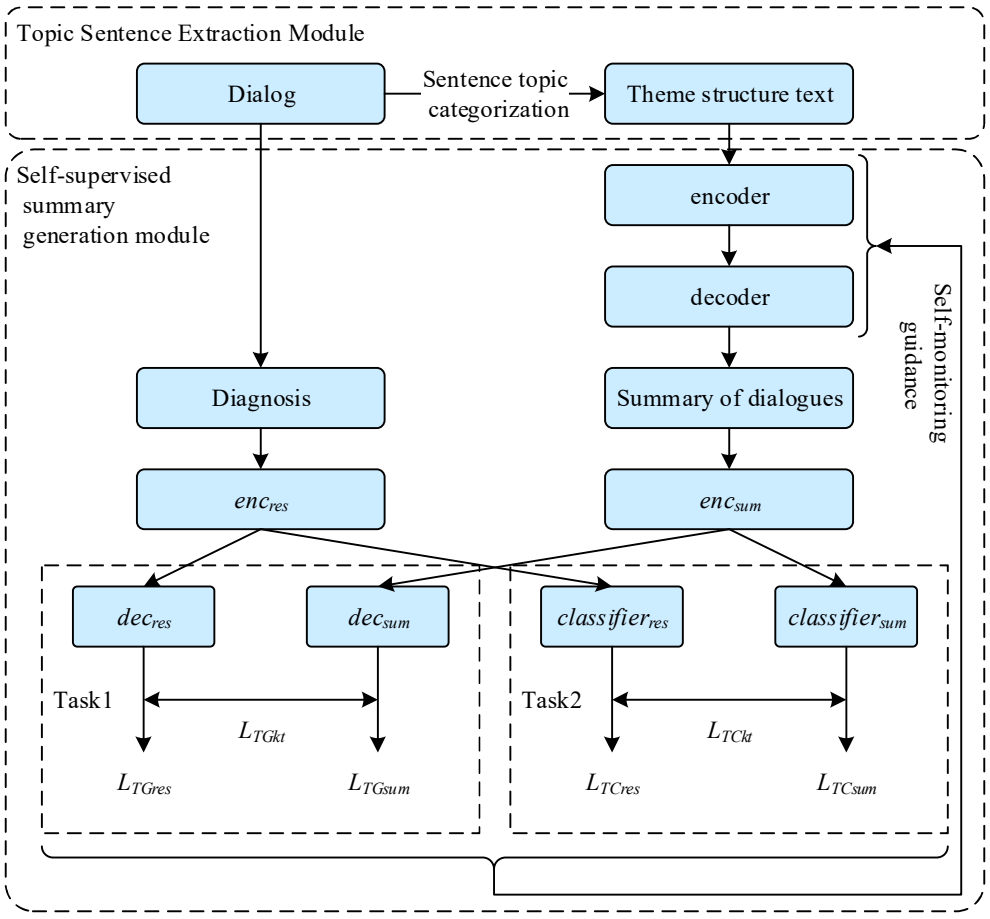


Fig. 1. Model structure

2.3.1. Topic Sentence Extraction Module. In this subsection, a multilayer LSTM is used for sentence topic classification because it can handle both long-term and short-term temporal dependencies well. The specific structure of the multilayer LSTM is shown in Fig. 2. After the input layer, there are two LSTM layers which are then forwarded to the final output Dropout layer and Dense layer. The first LSTM layer generates sequence vectors that will be used as input for the next LSTM layer. In addition, the current LSTM layer receives feedback from its previous time step, so it can capture some of the previous features and obtain features at a different scale [22]. The Dropout layer excludes 5% of the neurons to avoid overfitting, and the number of nodes in the fully-connected layer is the number of topics to be classified.

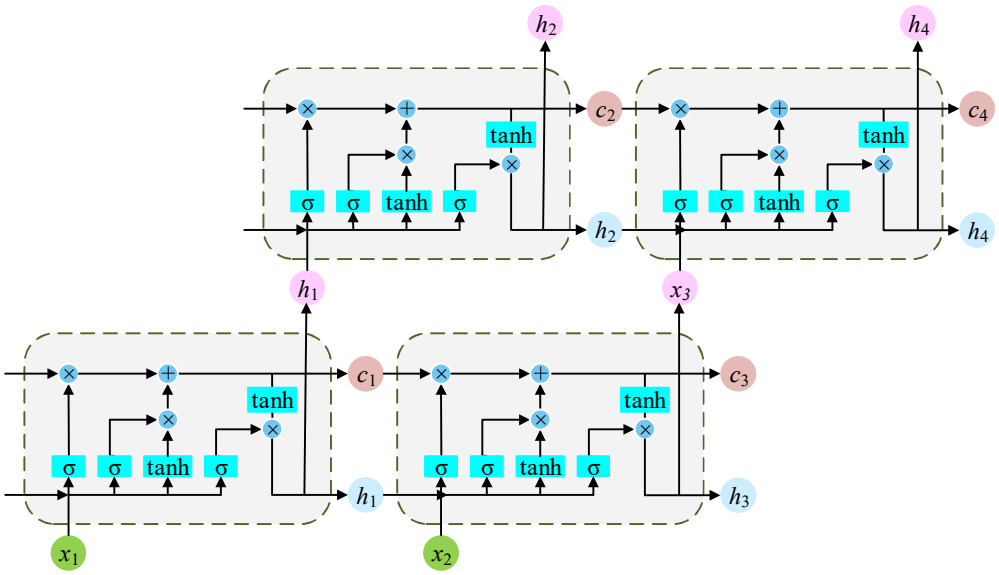


Fig. 2. Multi-layer LSTM concrete structure

2.3.2. Self-supervised summary generation module. The self-supervised summary generation module is shown in Fig. 3, and the self-supervised strategy stems from the idea that the diagnostic results corresponding to the original dialog and the diagnostic results corresponding to the summary should be similar, and that they are roughly equivalent to completing the auxiliary task. That is, the summary generation module aims at generating summaries from the original dialogs and requires an intrinsic supervision mechanism. In this generation process, two auxiliary tasks, diagnostic result extraction generation and diagnostic result categorization, are constructed to transform the unsupervised dialog summarization task into a self-supervised model through the learning of the auxiliary tasks.

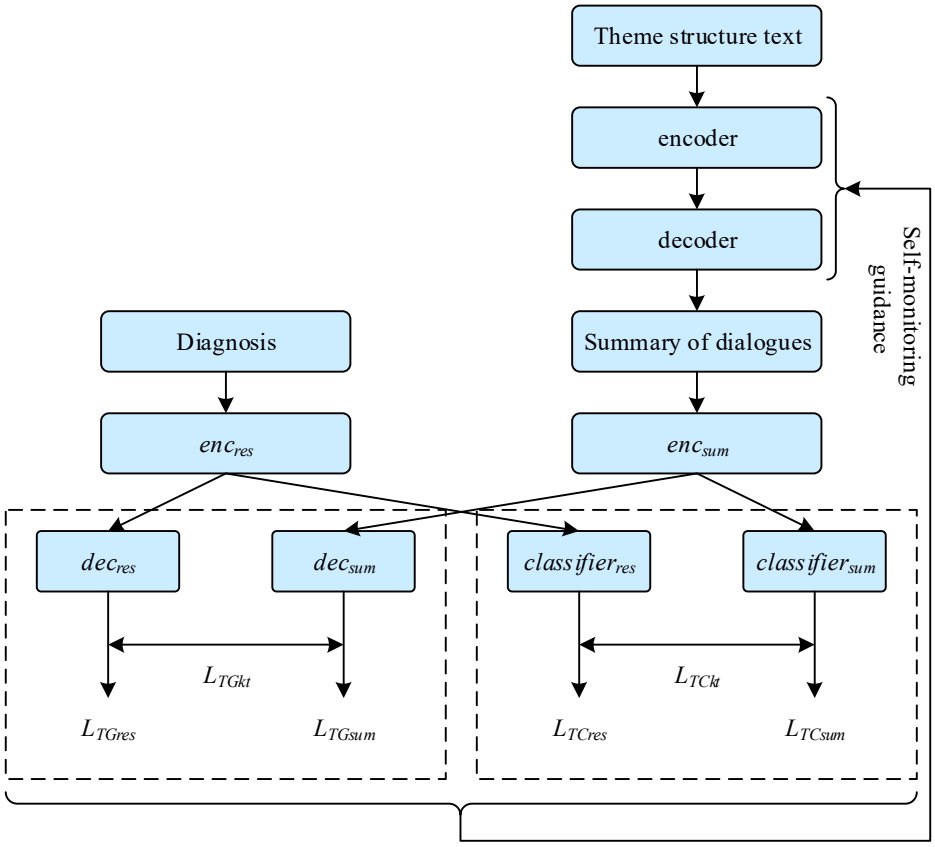


Fig. 3. Self-supervised summary generation module

1) Auxiliary Tasks. Task1: Diagnosis result extraction. Since the content of the "disease" label in the dataset is not standardized, in order to achieve better self-supervision, the content needs to be extracted. In this section, a commonly used encoder-decoder structure is adopted, and the entire dialog is connected and encoded using Bi-LSTM, where each word is represented by a concatenation of forward and backward LSTM states, $h_i = \left[h_i^{fwd}, h_i^{bwd} \right]$.

As for the decoder, a unidirectional LSTM with an attention mechanism is used in this section. the attention distributions at context vectors c_t and a^t are shown in Eqs. (1) and (2) below, respectively:

$$c_t = \sum_{i=1}^{n} a_i h_i \tag{1}$$

$$a_i^t = \sigma\left(h_i W_a s_t\right) \tag{2}$$

where W_a denotes the learnable parameters and σ denotes the softmax function.

For the context vector and the current decoder state s_t, the probability distribution used to predict the output word y_t on the vocabulary list is shown in equation (3) below:

$$p\left(y_t\right) = \sigma\left(W_p\left(\phi\left(W_k\left[y_{t-1}; s_t; c_t\right] + b_k\right)\right) + b_p\right) \tag{3}$$

where W_p, W_k, b_k, and b_p are learnable parameters. σ is the $soft\max$ function and ϕ is the $\tanh$ function.

In this section, the negative log-likelihood is used as the loss function and the loss of the task is generated by the path $enc_{res} \rightarrow dec_{res}$, which is expressed as shown in equation (4) below:

$$\mathrm{L}_{TG_{res}} = -\sum_{t=1}^{q} \log p\left(l_t \left| l_{<t}; enc_{res}\right.\right) \tag{4}$$

where $l = \{l_1, l_2, \ldots, l_q\}$ is the extraction of the generated diagnosis. Similarly, the generated dialog-based summaries $\mathrm{L}_{TG_{sum}}$ are obtained by $enc_{sum} \rightarrow dec_{sum}$ computing. In this section, the KL scatter function is added to ensure that the diagnostic results based on the original dialogs and the diagnostic results of the generated summaries have similar performance, and the effect is achieved by reducing the difference between the predicted probability distributions of each time step, which is computed as shown in Eq. (5) below:

$$\mathrm{L}_{TG_{kl}} = \sum_{t=1}^{q} KL\left(p\left(l_t \left| l_{<t}; enc_{res}\right.\right) \middle\| p\left(l_t \left| l_{<t}; enc_{sum}\right.\right)\right) \tag{5}$$

The diagnostic results are extracted to generate the task loss expressed as shown in equation (6) below:

$$\mathrm{L}_{TG} = \alpha_0 \mathrm{L}_{TG res} + \alpha_1 \mathrm{L}_{TG sum} + \alpha_2 \mathrm{L}_{TG_{kl}} \tag{6}$$

where α_0, α_1, and α_2 are each loss part weight.

Task2: Select the correct diagnostic result among the K candidate results, similar to the coding process in Task1 above, Bi-LSTM is selected as the encoder in this section. The dialog diagnostic result, denoted h_d, is the average of the hidden states of each word. In addition, each candidate is also encoded by Bi-LSTM and projected by the logic layer into a dense vector, which is then connected to h_d, denoted $\left[f\left(uc_i\right); h_d\right]$. The loss formula based on the dialog diagnostic result is shown in equation (7) below:

$$\mathrm{L}_{TC_{res}} = -\sum_{n=1}^{K} \log\left(p\left(z_n\right)\right) \tag{7}$$

where $p\left(z_n\right)$ is the output probability of the candidate outcome. $\mathrm{L}_{TC_{sum}}$ The diagnostic results are selected through $enc_{sum} \rightarrow classifier_{sum}$ paths.

This section also uses KL dispersion to measure the difference between what the diagnostic results are extracted from and the diagnostic results selected by the summary, which is calculated as shown in equation (8) below:

$$\mathrm{L}_{TC_{kl}} = KL\left(p\left(uc^{res}\right) \middle\| p\left(uc^{sum}\right)\right) \tag{8}$$

where $p\left(uc^{res}\right)$ and $p\left(uc^{sum}\right)$ are K candidate outcome probability distributions. The loss function of the final classification result is shown in equation (9) below:

$$\mathrm{L}_{TC} = \alpha_3 \mathrm{L}_{TC_{res}} + \alpha_4 \mathrm{L}_{TC_{sum}} + \alpha_5 \mathrm{L}_{TC_{kl}} \tag{9}$$

where α_3, α_4, and α_5 are each loss component weight.

2) Self-supervised summarization. For the initial summary, this subsection treats the section as a sentence binary classification task, which means that the extraction of the conforming summary results is treated as a replicative summarization. Specifically, this section encodes the topic-structured text using Bi-LSTM and represents them as hidden states $h_1, h_2, \ldots, h_{n-1}$ and then as a representation of the content by averaging the pooling of cascaded hidden states across the text, expressed as shown in equation (10) below:

$$d = \phi\left(W_d \frac{1}{n-1} \sum_{i=1}^{n-1} \left[h_i^{fwd}; h_i^{bwd}\right] + b_d\right) \tag{10}$$

where W_d and b_d are learning parameters and ϕ is a $\tanh$ function. For topic-structured text categorization, each topic-structured text is connected to a dialog representation d. And the logical layer predicts the generation probability as shown in equation (11) below:

$$p\left(u_i=1\right)=\psi\left(W_h h_i+h_i W_{hd} d+b_h\right) \tag{11}$$

Where, W_h, W_{hd}, b_h are learning parameters and ψ is sigmoid function.

In this section, the candidate summary with the highest probability is selected as the initial summary. After obtaining the initial summary, the extraction of the summary can be guided under self-supervision. Specifically, in order to obtain effective results and similar performance, the auxiliary task optimizes the first few generated summaries. Thus the training loss for the diagnostic result extraction generation task and the summary classification task is expressed as shown in equation (12) below:

$$\mathrm{L}_{ext}=\mathrm{L}_{TG}+\mathrm{L}_{TC} \tag{12}$$

The abstraction process for generating the final summary follows the traditional encoder-decoder structure, and for each time step, the word prediction probabilities are similar to Eq. (3), which are calculated as shown in Eq. (13) below:

$$p\left(y_t\right)=\sigma\left(W_p\left(\phi\left(W_k\left[y_{t-1};s_t;c_t\right]+b_k\right)\right)+b_p\right) \tag{13}$$

During forward training and testing, this section uses the reparameterization technique as an approximation to the variance of sampling from the original probabilities, specifically, converting the sampled words to obtain $\arg\max$ constraints from the new probabilities and discretizing the $\tilde{y}$ using $\arg\max$, which reduces to the following equations (14) and (15):

$$\tilde{y}_t=\arg\max\left(\log\left(p\left(y_t\right)\right)+g\right) \tag{14}$$

$$g=-\log\left(-\log\left(\xi\right)\right),\xi\sim U\left(0,1\right) \tag{15}$$

where g is the ST contribution and U is the contribution after unification. As for computing the gradient in the backpropagation, this section uses the continuous differentiable approximation of $argmax$ as shown in Eq. (16) below:

$$P\left(y_t^i\right)=\frac{\exp\left(\left(\log\left(p\left(y_t^i\right)\right)+g_i\right)/\tau\right)}{\sum_{j=1}^{|V|}\exp\left(\left(\log\left(p\left(y_t^j\right)\right)+g_j\right)/\tau\right)} \tag{16}$$

where $|V|$ is the vocabulary size and $\tau\in\left(0,\infty\right)$ is a temporary parameter.

Fluent summaries are generated by adding a language modeling loss, which predicts outputs more closely to the fluent text, as computed in Eq. (17) below:

$$\mathrm{L}_{lm}=KL\left(p\left(y_t\right)\middle\|p_{lm}\left(y_t\right)\right) \tag{17}$$

where $p_{lm}\left(y_t\right)$ denotes the distribution probability of the output word in the language model.

Therefore, the final training loss for self-supervised dialog summarization is shown in equation (18) below:

$$\mathrm{L}=\mathrm{L}_{ext}+\alpha_6\mathrm{L}_{lm} \tag{18}$$

where α_6 is the normalized weight.

2.4. Design and Implementation of Strategy Generation System

The system design of XACML (eXtensible Access Control Markup Language) policy generation system based on natural language processing needs to consider many components of policy editing, both the processing flow of natural language processing and the implementation details of semantic technology, and the use of these two technologies in combination to realize the access control policy needs to be logically technologically relevant, so that in the process of implementation it can make the close joining of functional units in the process.

2.4.1. System architecture. This subsection will focus on the design of the strategy generation system, which includes both the overall structural design of the system, as well as the detailed design of each functional module and the technical problems that need to be solved. The structural design of the system is based on the natural language processing process to achieve the Chinese word segmentation, semantic annotation, ontology storage and the final strategy file generation, the system takes the functional requirements as the design goal, combines the existing technology and tools to carry out the combination and scheduling of the tool functions, so as to make the required functional modules linkage and combination, and to realize the design goal of the strategy generation based on the natural language processing.

2.4.2. Composition Logic of Strategy Generation System. Combining XACML technology and ontology technology, which are two different technologies, needs to have similar feasibility on the logical level and composition basis, ontology technology and XACML are the same on the bearing basis, both relying on XML as the bearing basis, so in this paper, this idea of using semantic technology to solve the problem of XACML's strategy generation is feasible on the basis of theory and realization. The logical framework of the strategy generation system designed in this paper is shown in Figure 4. Logically using XACML in combination with OWL to link the strategy editing with the internal communication execution system. The strategy system combining XACML and OWL is divided into three levels:

1) Editing environment: the editing environment supports technical specifications such as access control policies, role hierarchies and constraints. This is also the design goal of the system in this paper, and the main focus here is on describing and modeling the restrictions and constraints in the policy content.

2) System Information Repository: The system information repository stores OWL ontologization relations and XACML policy files. The system information base cannot be directly accessed by the user, but the user can manage the system information base by editing the environment.

3) Execution system: policies and ontologies are utilized by communication requests at the execution level. The execution system contains the core engine of XACML, which is extended with additional modules for semantic processing duties and obligations according to the reference architecture of XACML. The execution system receives the incoming communication requests and the execution strategy evaluates the semantic functions and obligations of the ontology invoked through the query. The result is a final approval decision for the execution process.

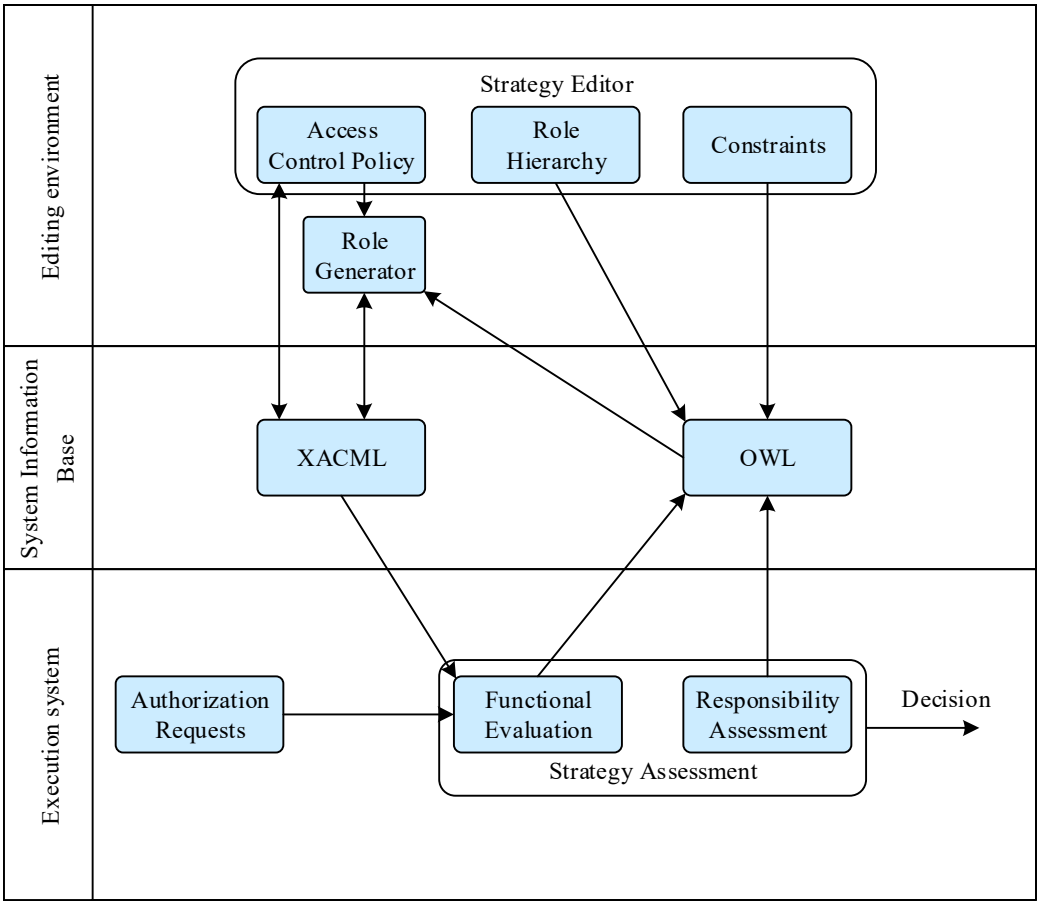


Fig. 4. The logical framework of the policy generation system

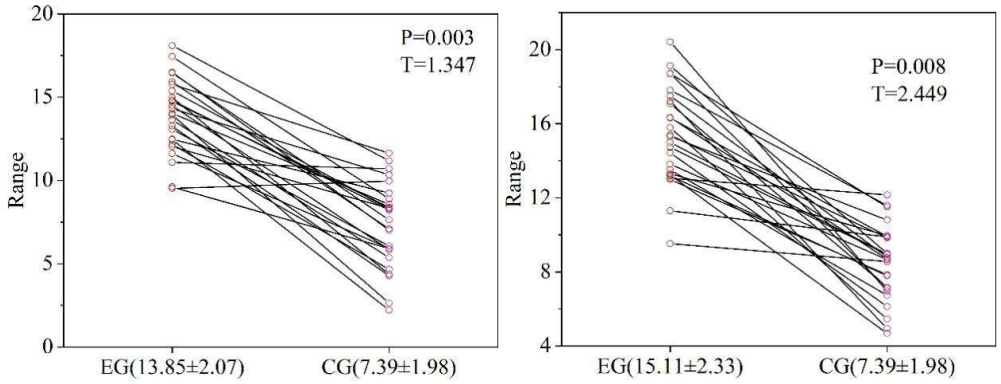
3. Example analysis of communication strategies based on natural language technologies

3.1. Analysis of the effects of the doctor-patient interactive discourse model

3.1.1. Subject of the study. In order to verify the effectiveness of the doctor-patient interactive discourse model, 50 doctors in a hospital were selected as the sample of this study, and the sample was divided into an experimental group (this paper's discourse model: 25) and a control group (traditional discourse model: 25). The experimental group and the control group were respectively subjected to experimental intervention for a period of time, and the effectiveness of the doctor-patient interactive discourse model was confirmed by analyzing the difference in effect between the experimental group and the control group after the intervention.

3.1.2. Analysis of variances. Based on the questionnaire to obtain the quantitative value of the intervention effect of the experimental group and the control group, the intervention effect is mainly reflected in the expressive ability and cognitive level, and the analysis of the difference between the intervention effect of the experimental group and the control group is shown in Fig. 5, in which (a)~(b) are the expressive ability and the cognitive level respectively, and the EG~CG sub-tables in the figure indicate the experimental group and the control group. It can be clearly seen that after a period of teaching intervention, there is a significant difference between the experimental group and the control group in terms of expression ability and cognitive level ($P<0.05$), which indicates that compared with the traditional discourse mode, the doctor-patient interactive discourse mode has a higher priority, just as described in subsection 2.1.2 (the doctor-patient interactive discourse mode compensates for

the difficulties in the communication and exchange between the doctor and the patient and realizes a more accurate communication and exchange of information and knowledge between the doctor and the patient, comprehensive information exchange and knowledge exchange between doctors and patients, and improves the doctors' ability to express medical terminology and cognitive level).



(a)Expressive ability (b)Cognitive level

Fig. 5. Difference analysis of intervention effect

3.2. Model validation analysis

3.2.1. Experimental design. For topic-structured doctor-patient interaction discourse based on topic structure, this section also chooses to conduct experiments on MedDialog, CMedQA2.0, and CovidDialog Chinese doctor-patient dialogues datasets, and employs several models for experimental comparisons, including Transformer, GPT2, and BERT-GPT2 models. The effectiveness and generalizability of the modeling approach in this paper is verified by comparing the performance advantages and disadvantages between the models in detail.

3.2.2. Experimental results. In the experiments, this section is evaluated using several automated metrics. These include Perplexity, which measures how difficult it is for the model to predict the next word in a given corpus, with lower perplexity indicating greater coherence of the generated text, BLEU, a similarity metric measuring the similarity between the generated text and the reference text (which measures the accuracy between the actual sequences on the ground and the sequences generated by the model's response, and mainly considers the results of BLEU-2 and BLEU-4.) as well as the generated text quality metrics Meteor (an exact match is the portion of the automatically generated summary that is exactly the same as the words in the reference summary), NIST (the BLEU metric differs in that it weights phrase matches with a decreasing factor, which means that longer phrase matches are more important than shorter phrases). As well as the Distinct metric (a metric for evaluating lexical diversity in the generated discourse, which essentially measures the number of unique words in the generated discourse as a proportion of the total vocabulary. When n=1, the ratio of the number of distinct individual characters in the generated text to the total vocabulary is evaluated, in which case the evaluation metric only takes into account the diversity of individual words without considering the way the words are combined with each other, and when n>1, the way n consecutive words are combined is taken into account, which better captures the contextual and linguistic information of the text, and in the experiments in this paper, the cases of n=1 and 2 are considered).

The performance results on the MedDialog dataset are shown in Fig. 6, where this paper's model significantly outperforms the other compared models in the NIST_4 and Dist_1 metrics, where it improves 73.4% and 248% compared to the Transformer comparison model, and 54.9% and 175% on average across all model comparisons. The performance in the BLEU evaluation metrics decreased compared to the BERT-GPT2 model, and it is possible that the BERT-GPT2 model adopts the BERT

coding module, which extracts more obvious semantic information. However, the average improvement over the baseline model GPT-2 is 4.73%.

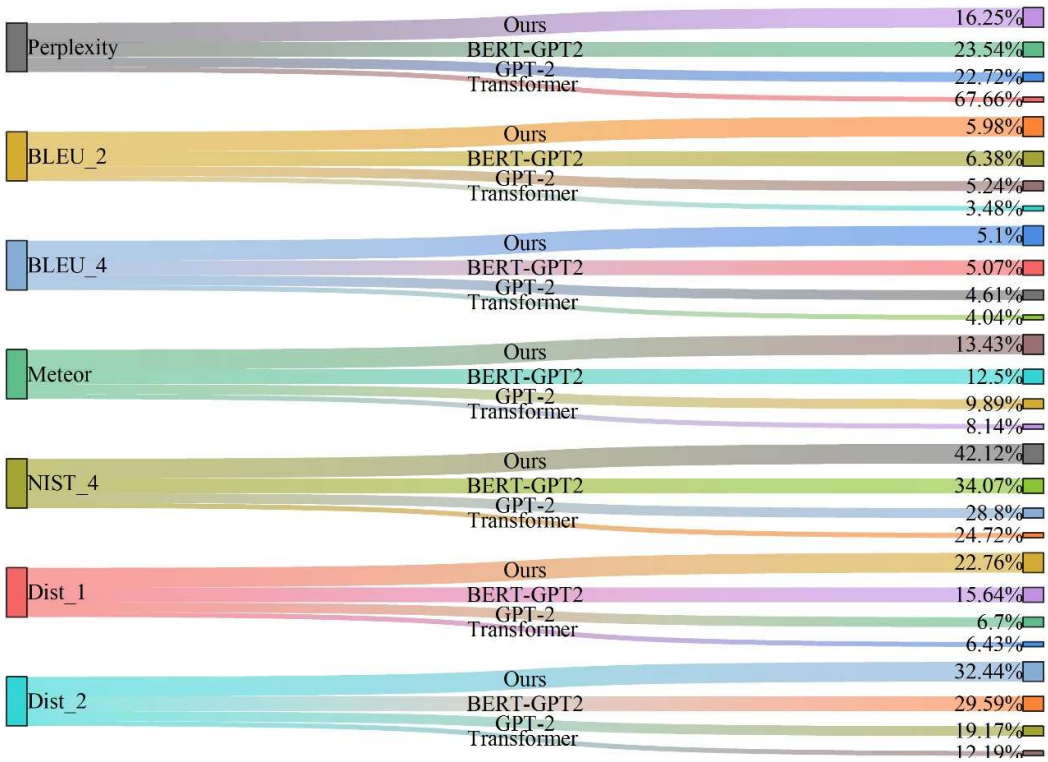


Fig. 6. The results are displayed on the MedDialog dataset

Figure 7 shows the comparison of the results of the evaluation metrics for the CMedQA2.0 dataset. The model in this paper outperforms all other comparison models in all comparison metrics. Compared with the baseline model GPT-2, the average performance is improved by 46.34%. Compared with the BERT-GPT2 model, which has close performance in metrics, the performance is improved by 11.75% on average.

Fig. 7. Comparison of evaluation index results of CMedQA2.0 dataset

Figure 8 shows the comparison of evaluation metrics results for the CovidDialog dataset. This paper's model outperforms the other compared models in the BLEU_2, BLEU_4, and Dist_1 metrics, outperforming the BERT-GPT2 model, which has close experimental results, by 1.87%, 0.17%, and 3.26%, respectively. In the comparison of NIST_4 and Dist_2 metrics, the BERT-GPT2 model produced optimal results compared to the other models, outperforming this paper's model by 2.68% and 1.6%, respectively, and it is possible that the BERT-GPT2 model is more suitable for the experimental dataset in both metrics. Comparing to the baseline model GPT-2, this paper's model has an advantage in the performance of all the compared evaluation metrics with an average improvement of 45.4%. In terms of perplexity, this paper's model outperforms the other compared models in all relevant experimental datasets, with an average of 44.9% higher compared to the baseline model GPT-2. Lower perplexity indicates that the model is better at predicting a given sequence and is able to produce a smoother as well as more coherent output when generating discourse. The model in this paper is able to more fully utilize the expertise in the medical domain, which improves its semantic comprehension and generation ability in the task of generating medical dialogues and facilitates interactive communication between doctors and patients.

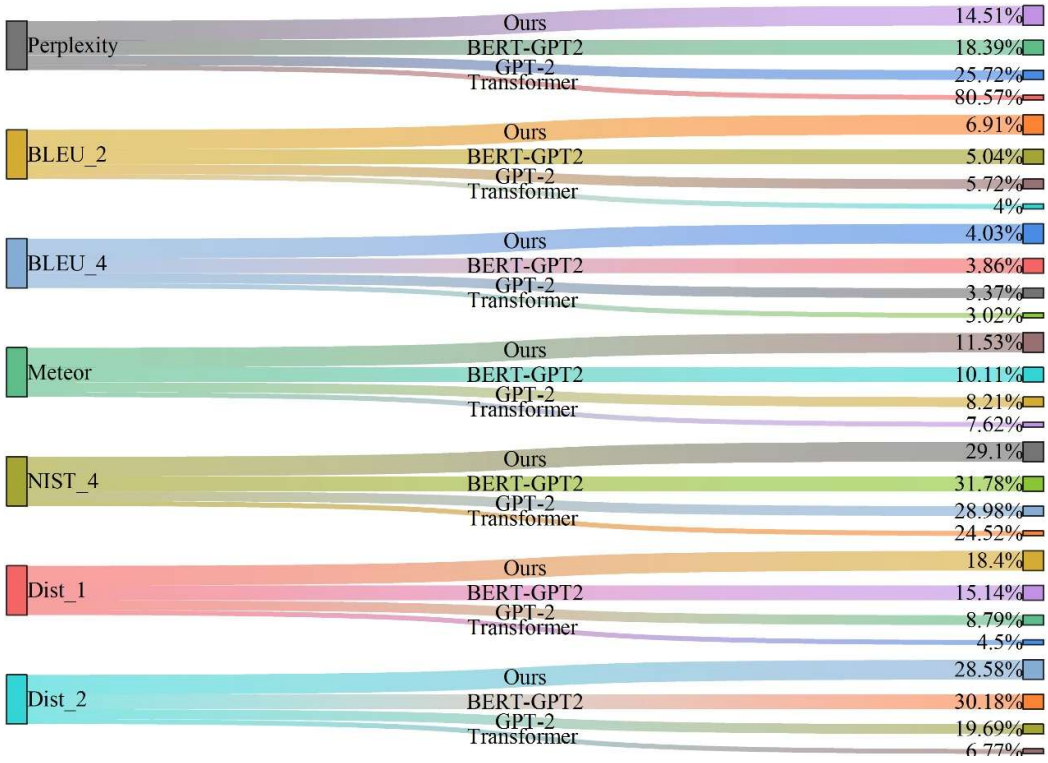
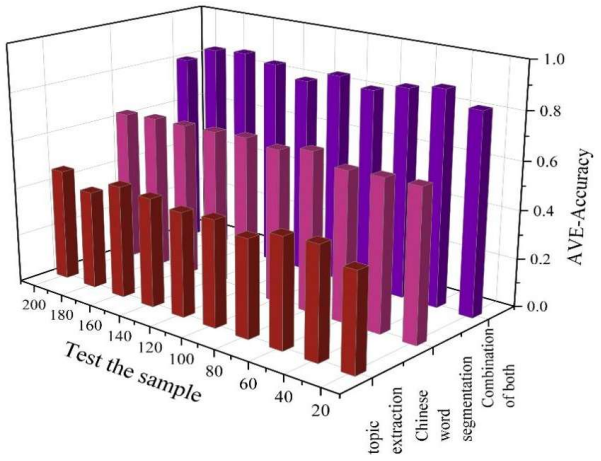


Fig. 8. Comparison of results of evaluation indicators of the CovidDialog dataset

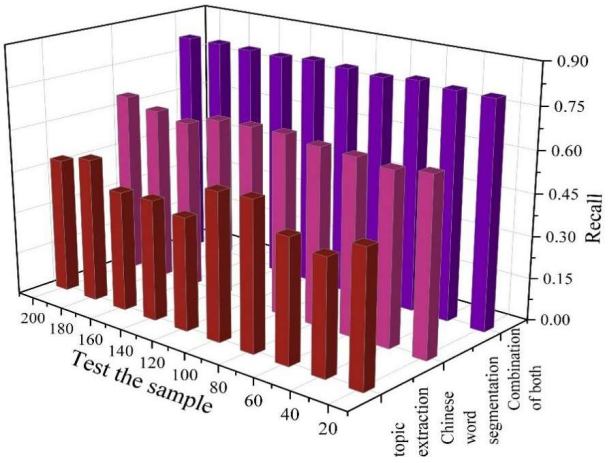
3.3. System testing

3.3.1. Environment configuration. After in-depth understanding of the realization mechanism of the dialogue system, it is first necessary to make relevant configurations of the running environment to ensure the stable operation and efficient performance of the system, in which the deep learning framework uses Pytorch, the operating system Win10, the use of the language python3.7, and the running environment Anaconda3.

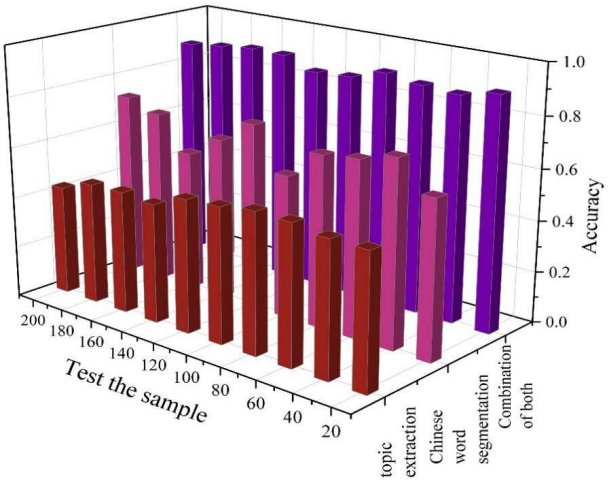
3.3.2. Test results. First, we randomly selected 200 test samples from the previous dataset, and conducted a comprehensive test of the entire strategy generation system. The test results of the strategy generation system are shown in Fig. 9, where (a) to (c) denote the average precision, recall, and accuracy, respectively. The testing process uses the three evaluation indexes of precision, recall, and average precision to test the effect of Chinese segmentation and topic extraction on strategy generation, and the average precision based on Chinese segmentation is as high as 63.57%, the average precision based on topic extraction is as high as 43.68%, and the average precision of the two together is as high as 85.02%. The doctor-patient dialog generates responses that are consistent with the context of the conversation and relevant knowledge to the real medical context, allowing users to engage in multiple rounds of dialog with the smart doctor to ensure the completeness and accuracy of the information. It solves the problems of uneven distribution of medical resources and difficult communication between doctors and patients due to limited medical resources, and realizes the interconnection of information between doctors and patients, enabling the smart doctor to assist doctors in providing patients with relevant medical information and advice, and satisfying some of the patients' inquiries.



(a)AVE-Accuracy



(b)Recall



(c)Accuracy

Fig. 9. Policies generate system test results

4. Conclusion

In this paper, we collect research data with the help of data mining technology, and perform pre-processing operations such as text segmentation, de-duplication, and de-duplication to complete the

construction of the discourse model of doctor-patient interaction based on topic structure. After that, natural language processing technology is used to design and implement the doctor-patient communication strategy generation model. Synthesizing the above theories, the communication strategy based on natural language technology is analyzed by example. On the CMedQA2.0 dataset, the model in this paper improves the average performance by 46.34% compared with the baseline model GPT-2. Compared with the BERT-GPT2 model, which is close in metric performance, the performance is improved by 11.75% on average. It shows that the model in this paper can more fully utilize the expertise in the medical field to improve the communication ability of doctor-patient interactive discourse. The average accuracy of the fusion of Chinese word splitting and topic extraction is as high as 85.02%, which realizes the interconnection of doctor-patient information, enabling the smart doctor to assist the doctor to provide relevant medical information and strategies to the patient, and to better promote the interactive communication and development of the doctor-patient interaction.

Funding

This research was supported by the Liaoning Provincial Social Science Planning Fund Project (Project Number: L24CYY013).

References

[1] Tara, & Beck. (2018). The importance of doctor-patient communication. Md Advisor A Journal for New Jersey Medical Community.

[2] Meyer, B. , & Bührig, Kristin. (2014). Interpreting risks. medical complications in interpreter-mediated doctor-patient communication. European Journal of Applied Linguistics, 2(2).

[3] Peter Lapsley. (2014). Re: improving doctor-patient communication. BMJ.

[4] Kashgary, A. , Alsolaimani, R. , Mosli, M. , & Faraj, S. . (2016). The role of mobile devices in doctor-patient communication: a systematic review and meta-analysis. Journal of Telemedicine and Telecare.

[5] Gulbrandsen, P. . (2020). Shared decision making: improving doctor-patient communication. BMJ (online), 368, m97.

[6] Ioan, B. G. , Ciuhodaru, T. , Velnic, A. A. , Crauciuc, D. , Hanganu, B. , & Manoilescu, I. S. . (2017). The role of doctor-patient communication in preventing malpractice complaints. International Journal of Communication Research, 7.

[7] Pino-Postigo, A. . (2017). Challenges in doctor-patient communication in the province of malaga: a multilingual crossroads. Procedia - Social and Behavioral Sciences.

[8] Hanadi, H. , Geoffrey A., S. , Sinyoung, P. , Jing, X. , & Zhigang, X. . (2024). Hospital patient experience: exploring hospitals as shifters and sustainers over time. Quality management in health care(3), 33.

[9] Singh, K. V. . (2021). Hospital services vs. patient care: an overview. Annals of Medical and Health Sciences Research.

[10] Farver-Vestergaard, I. . (2024). The use of podcasts as patient preparation for hospital visits—an interview study exploring patients' experiences. International Journal of Environmental Research and Public Health, 21.

[11] Donnelly, L. F. , Grzeszczuk, R. , & Guimaraes, C. V. . (2022). Use of natural language processing (nlp) in evaluation of radiology reports: an update on applications and technology advances. Seminars in Ultrasound, CT, and MRI(2), 43.

[12] Roh, T. , Jeong, Y. , Jang, H. , & Yoon, B. . (2019). Technology opportunity discovery by structuring user

needs based on natural language processing and machine learning. PLOS ONE, 14.

[13] Wang, X. , Tian, J. , & Li, F. . (2022). Text data mining of power based on natural language processing technology. Journal of Physics: Conference Series, 2221(1).

[14] Singh, G. B. , Kumar, R. , Ghosh, R. C. , Bhakhuni, P. , & Sharma, N. . (2024). Question Answering in Medical Domain Using Natural Language Processing: A Review. International Conference on Data Management, Analytics & Innovation. Springer, Singapore.

[15] Cornell, R. H. , Greve, K. A. , & Semick, M. T. . (2020). Systems and methods for processing natural language queries for healthcare data. US10649985B1.

[16] Locke, SaskiaBashall, AnthonyAl-Adely, SarahMoore, JohnWilson, AnthonyKitchen, Gareth B. (2021). Natural language processing in medicine: a review. Trends in anaesthesia and critical care, 38(1).

[17] Nair, I. , & Aswathyp, R. . (2019). Natural language processing in medicine: a review. International journal of engineering research and technology, 7.

[18] Riedl, A. B. . (2011). Applications of simple natural language processing techniques to improve data quality and semantic interoperability in medicine. Dissertations & Theses - Gradworks.

[19] Wang, S. , Ren, F. , & Lu, H. . (2018). A review of the application of natural language processing in clinical medicine. IEEE.

[20] Lauren N Hum,Chris Daniel & Gargi Mukherji. (2024). Standardized patient interactions improve dental student confidence following procedural errors. Journal of dental education

[21] Hryciw Brett N.,Fortin Zanna,Ghossein Jamie & Kyeremanteng Kwadwo. (2023). Doctor-patient interactions in the age of AI: navigating innovation and expertise. Frontiers in Medicine1241508-1241508.

[22] Jyotin Kateshia & Vikas J Lakhera. (2024). Productivity forecasting of the solar still with phase change material using LSTM and GRU neural networks. Engineering Research Express(4),045531-045531.