Expert Systems With Applications 227 (2023) 120211

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Novel medical question and answer system: Graph convolutional neural network based with knowledge graph optimization

Xu Wang [a,c,*], Zijin Luo [b,c], Rui He [a], Yixin Shao [a]

[a] College of Communication Engineering, Jilin University, Changchun, Jilin 130000, China
[b] College of Mathematics, Jilin University, Changchun, Jilin 130000, China
[c] Chengdu Institute of Computer Applications, Chinese Academy of Sciences, Chengdu 610000, China

ARTICLE INFO

Keywords:
Medical diagnosis
Knowledge graph
Graph neural network
Speech recognition
Disease classification

ABSTRACT

In order to effectively integrate medical data and alleviate the problem of uneven distribution of medical resources. In this paper, we combine the techniques of expert systems, graph neural networks, and knowledge graphs to propose a disease guidance model combining semi-supervised graph neural networks and knowledge graphs. We use the MASR speech recognition module combined with gated convolutional units for effective text processing of different types of speech; then we use the LTP module in natural language processing for semantic analysis and segmentation matching of interrogative sentences; we combine keywords with the number of diseases and divide and construct the set of nodes with knowledge graphs. And we use semi-supervised graph neural network type analysis to give treatment results and rehabilitation suggestions effectively. We optimize the Chinese and English corpora respectively, adding consideration for local dialect audiences. We performed a comprehensive comparison of the accuracy and training time of several mainstream GCN algorithms and our GCN semi-supervised (SGS) under various graphical text datasets to validate the efficiency and accuracy of our own algorithm choices. We preprocess the number of different symptoms for classification and simplify the redundant nodes to optimize the running time while taking into account the overall convergence. The operational mechanism of the model as well as the convergence and hits under different symptom parameters are explained through hit rate and convergence rate metrics to demonstrate the effectiveness and stability of the model under proprietary medical conditions.

1. Introduction

Intelligent Healthcare aims to build a regional medical interaction platform for sharing medical information through the integrated use of technologies such as the Internet of Things, cloud computing, and big data to achieve. It aims to enable linking and data sharing between patients and medical staff, healthcare providers, and medical equipment (Ali et al., 2020). Intelligent healthcare breaks through the spatial and temporal limitations of traditional healthcare at both the technological and institutional levels, emphasizing the integration and development of network information technology with all aspects of traditional healthcare, fully alleviating the current global problem of uneven spatial and temporal span in the distribution of healthcare resources (Demirkan, 2013).

There are six main aspects of intelligent healthcare: Reception, Examination, Diagnosis, Treatment, Prescription, and Healthcare (Liu et al., 2019). The most central part of intelligent healthcare is diagnosis. The prerequisite for effective treatment in the medical process is also based on an accurate and rational diagnosis. However, traditional medical diagnosis has many limitations in time and space: traditional medical models are disconnected from each other, making it difficult to effectively integrate the exponential growth of medical data; traditional medical resources are unevenly distributed geographically, and it is also difficult to meet the needs of patients around the clock; in the rapid iteration of medical data, it is difficult for specialized doctors to accurately diagnose complex symptoms that are interdisciplinary and unpracticed (Vayena et al., 2018; Goldstein et al., 2007; Allen et al., 2003).

Faced with the dilemma of optimizing traditional diagnostic methods in intelligent healthcare, the rapid development of artificial

* Corresponding author at: College of Communication Engineering, Jilin University, Changchun, Jilin 130000, China.
E-mail addresses: wangxu2020@mails.jlu.edu.cn (X. Wang), luozj1020@mails.jlu.edu.cn (Z. Luo), herui2020@mails.jlu.edu.cn (R. He), shaoyx2020@mails.jlu.edu.cn (Y. Shao).

https://doi.org/10.1016/j.eswa.2023.120211
Received 13 July 2022; Received in revised form 16 April 2023; Accepted 16 April 2023
Available online 26 April 2023

intelligence in multiple fields has brought new solutions. This approach can be temporally divided into two main phases, knowledge-driven and data-driven (Zheng and Rodríguez-Monroy, 2015; Abidi, 2001;Cai et al.,2019; Hu et al., 2016). The two main representations are Expert-based treatment systems based on computer logic and causal inference in the second half of the last century (Calegari et al., 2020; Hasan, 2012; Shameer et al.,2018); and the data-driven machine-learning approaches that have been flourishing since the beginning of this century.

This century has seen a further boom in machine learning methods in intelligent healthcare systems, and one of the relatively mainstream branches is the derivation of Neural Networks (Liu, Faes, Kale, Wagner, Fu, 2019). It is a computational model that simulates the human nervous system and consists of a large number of artificial neurons. It is mostly used for learning and training to achieve classification, recognition, and prediction of input data.

In recent years, with the widespread use of large-scale parallel computing and Graphic Processing Unit (GPU) devices, coupled with a significant increase in data volume, the computing power of computers has been greatly improved. Traditional or variant artificial intelligence models have been widely applied in various fields such as finance and industry (Li et al., 2021; Huang et al, 2020; Zhang et al., 2021;Diniz et al.,2022).For example, Ana P. M. Diniz et al. effectively identified Submerged Entry Nozzle (SEN) blockage problems in a dataset with high-noise signals using the LSTM algorithm in the steel industry (Diniz et al.,2022). Indranil Ghosh et al. developed a model to predict stock price trends under class-imbalanced conditions (Ghosh and Chaudhuri, 2021).

Neural networks have also flourished in the application of knowledge graphs. In 2011, Bordes et al. constructed a knowledge graph by learning entity relationships based on traditional neural networks, such as MLP and RNN (Bordes et al. 2011). (Socher et al., 2013) completed the completion and inference of knowledge graphs using neural tensor networks (NTN). (Kipf and Welling, 2017) proposed a semi-supervised classification method based on Graph Convolutional Networks (GCN) for building knowledge graphs. With the rise of Convolutional Neural Networks (CNN), (Nguyen et al., 2018) used CNN to learn global relationships and transitional features between entities and relationships in knowledge graphs. (Xu et al., 2018) proposed a learning method called JK-Net based on GCN in 2018 to establish knowledge graphs, which uses multiple GCNs to learn different levels. (Wu et al., 2019) proposed a simplified graph convolutional network method called SGCN to build knowledge graphs, which simplified the weight-sharing mechanism of the original GCN and reduced the number of model parameters and computational resources (Wu et al., 2019).

In this paper, we use advanced technologies such as language recognition and natural language processing (NLP) and graphical neural networks (GNN), and knowledge graphs (KG) to focus on doctor-patient questions and assisted diagnosis. We use knowledge graphs and graph neural networks as the underlying core, by placing disease descriptions in the knowledge graph and using the connectivity and open organization capabilities of graph neural networks to achieve the extraction, fusion, inference, and quality assessment of disease and diagnostic ideas to achieve reasoning and assist in diagnosis.

2. Related works

2.1. Expert system

An expert system is an intelligent computer program system that uses the knowledge and experience of a level expert in a domain to reason about problems. The earliest expert system, DENDRAL, was born at Stanford University in 1968 (Lindsay et al., 1993; Buchanan and Feigenbaum, 1978; Feigenbaum et al., 1970) and the medical domain system MYCIN was developed in 1976 (Shortliffe, 2012; Daniel et al.,1997;Shortliffe et al, 1975). The following decades have also seen the extensive development of expert systems in healthcare, with interactive systems such as Icons, Diagnosis Pro, and the Help System, a logic-based health assessment system (Pryor et al., 1982).

Expert systems have two main components: the knowledge base and the inference engine. The development of fuzzy logic in recent years has accelerated the development of expert systems, incorporating probabilistic and uncertainty-based knowledge into decision-making applications, helping people to deal with complex problems with a high degree of uncertainty, such as the detection of chronic diseases and the prediction of cholera outbreaks (Fleming et al., 2007). In today's fast-growing data-driven approach, it is difficult to constantly build and update a knowledge base. At the same time, it can become more complex in resource-poor environments. The original expert systems based on logical judgment and knowledge bases appear somewhat underpowered. Many expert systems also lack accurate tracking mechanisms, which can undermine the trust of clinicians and patients.

2.2. Medical guidance system

The Guided Care System is functional computer software with doctor-patient questions and autonomous diagnosis. It has the function of alleviating labor costs, avoiding human errors, and effectively assisting doctors in guiding their judgments and early warning of risks. In principle, it can be broadly divided into two categories: rule-based templates and data-based models. The former is achieved by manually establishing rules for the correspondence between symptoms, diseases, and departments. The latter treats consultation as a departmental classification problem. Patient descriptions and department data are extracted and used as a guide model using traditional machine learning methods or deep neural network classification models. The principle of the two types is shown in the Fig. 1.

The process of implementing applications began as early as the conceptual and technical background of machine diagnosis proposed by Ledley in 1966 (Ledley, 1966), the judgment approach based on Bayesian theory to identify bacterial diseases, etc. proposed by Willcom in 1972 (Willcox et al., 1973; Curiac et al., 2009), the integrated feature selection approach applied by Abeel in disease diagnosis using support vector machine classification algorithms in 2010 (Abeel et al., 2009; Wosiak and Zakrzewska, 2018), and the combination of artificial bee colonies and improved Bayesian to achieve nearly 100% prediction accuracy in the field of heart disease and neurological disease diagnosis by Ahmed in 2014 (Babaoğlu, Kıran, & Ülker, 2013; Li, Cao, Wang, He, & Jin, 2020; Wang, Zhang, Liu, He, & Wang, 2021).

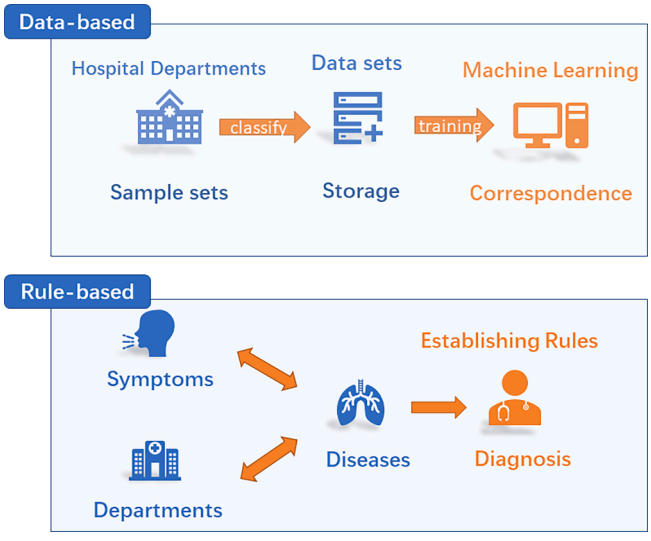

Fig. 1. Principle of Medical Guidance System.

2.3. Medical knowledge graph

Medical data have seen an exponential surge with the rapid development of the Internet, but data on the Internet is mostly characterized by heterogeneous diversity and a loose organizational framework. To better access information challenges, Google proposed the knowledge graph model in 2012, which gradually became a new structured carrier due to its powerful open organizational capability (Dong et al.,2014). In medical applications, knowledge graph construction can be divided into processes such as representation, extraction, fusion, inference, and quality assessment. The schematic diagram of the principle of graphical neural network is shown in Fig. 2

In the field of disease diagnosis application, in 2017 Rotmensch M et al. explored a method to construct medical knowledge graphs directly from electronic medical records by using Bayesian networks, but there is also the problem of not considering noise and not considering edge cases between symptom nodes (Rotmensch et al., 2017). In the same year, Weng H et al. proposed an automatic medical knowledge graph construction framework based on semantic analysis, which effectively improved the quality of knowledge graph construction (Weng et al., 2017); C Zhao et al. constructed an EMR-based medical knowledge network by extracting medical entities and proposed a clinical decision support diagnosis model that could provide the corresponding clinical decision support by using symptoms as input only (Zhao et al., 2017; Zhao et al.,2018); (Li et al., 2018) constructed a medical knowledge graph from 3,767,198 patients' electronic medical records to extract medical knowledge and innovatively proposed a quadratic structure to replace the classical triad in knowledge graphs to represent medical knowledge (Li et al., 2020).

2.4. Graph neural network

Graph neural networks are based on deep learning to process data with topological structures. It widely uses in various fields due to its good performance and traceability. It can deeply perceive the relationship between entities in the modeling process by graph structure data. The graph structure is formed using vertices and edges connecting the vertices. Specifically, in medical graphs, we consider the corresponding diseases and corresponding symptoms can be regarded as entity points, and the relative diagnostic relationships between them can be regarded as edges, which together form a kind of graph network(Scarselli et al., 2008).

By introducing a computational neural architecture, we can construct GNN models for time series prediction. In this field, there are the earliest ideas borrowed from CNN by Li zhang et al. in 2005 to design graph convolutional network (GCN) architecture to handle complex graph data (Zhang and Bai, 2005; Li et al., 2019) proposed a graph convolutional network-based disease gene prioritization (PGCN) approach that outperformed all models at that time for discovering associations of diseases; and (Ahmedt-Aristizabal et al., 2021) proposed a multi-graph neural network (GNN) model for medical diagnosis and analysis using a comprehensive methodological review.

3. Theory and methods

3.1. Overall process framework

In the process of implementing the model for the implementation of the doctor-patient question and answer session and the guided consultation, we have made the overall model patient-oriented. Concerning the process experienced by the patient in practice, is divided into three main steps as follows (Fig. 3):

1. Patient questions and describes information such as basic medical conditions. Using MASR's speech recognition module combined with gated convolution units to convert the input from speech description to text information.
2. Natural language processing is applied to the content of the text itself, using the LTP module to delineate the question types and semantic analysis. The question types can be classified to better match the question and answer results according to the question types.
3. Combining the number of keywords and diseases in the question and answer, the set of nodes and knowledge graphs are constructed according to different situations. We use graph neural networks for matching and disease-type analysis to give effective medical advice and advice on medication and diet, etc.

3.2. Patient speech recognition

3.2.1. Fundamentals

For the recognition and improvement of patient speech, we selected the MASR (Mandarin Automatic Speech Recognition) module as the Mandarin recognition module for Chinese. MASR is an automatic speech recognition framework implemented on the Pytorch platform, using an end-to-end deep neural network-based model for training, with the advantages of good model generalization and broad platform compatibility. (Ni et al., 2010). We used Gated Convolutional Neural Networks to optimize the original Wav2letter (Facebook 2016) framework by converting the activation functions from ReLU and HardTanh to GLU gated linear units and achieved better convergence in real-world tests

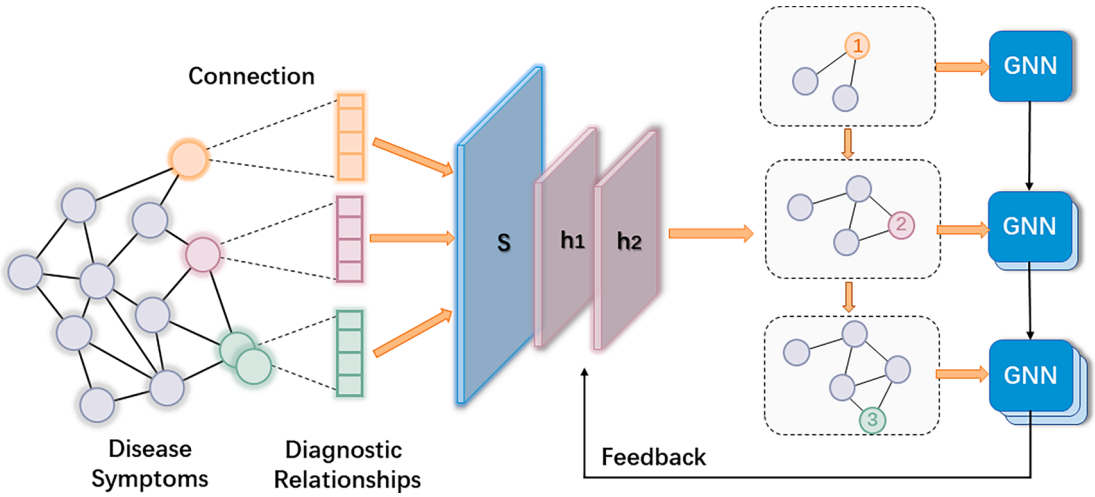

Fig. 2. Principle of Graph Neural Network Constructions.

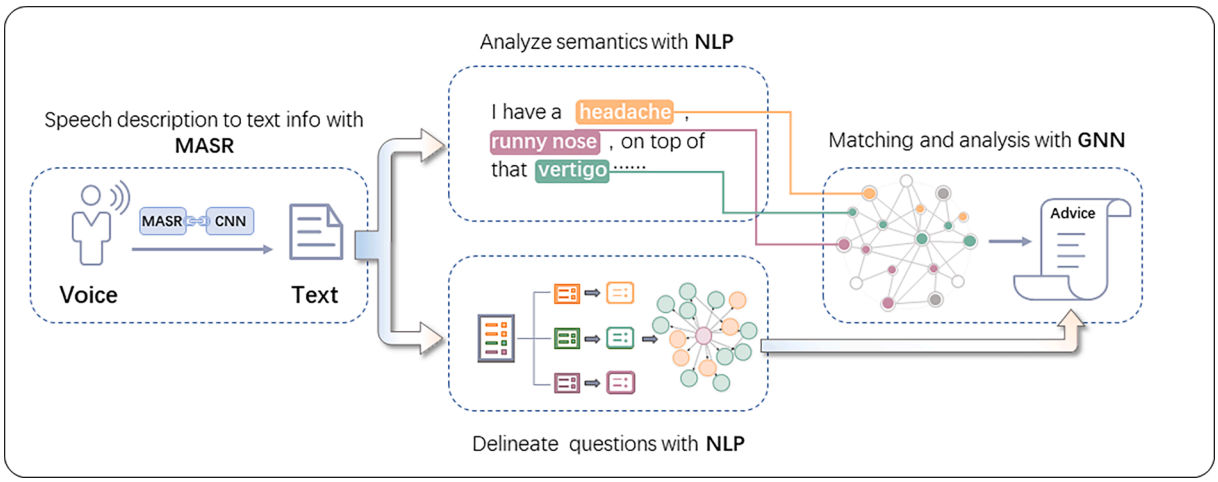

Fig. 3. Basic Schematic.

(Pratap et al., 2019; Collobert et al., 2016; Dauphin et al., 2017). The basic Construction of Wav2letter is shown in Fig. 4.

The main implementation processes are as follows (Fig. 5):

1. Recording of basic audio on the hardware side, conversion to WAV files, and normalization according to the time domain information of the amplitude.
2. The time–frequency transformation process is achieved by the Fast Fourier Transform (FFT); next we transform the signal into the recognition model (Nussbaumer, 1981; Brigham and Morrow, 1967).
3. The main body of the model consists of CNN convolutional neural networks for feature extraction, RNN recurrent neural networks for the prior iteration, and linear regression layers.
4. Finally, the processed results are decoded using the **Ctc_Beam_Search** and **Ctc_Greedy** decoders (Hori et al., 2017; Lee and Watanabe, 2021; Franke et al., 2007).

where the **Ctc_Greedy** process takes the index of the maximum value in the second dimension, takes the indexed, and removes the space. Also, adjacent values with the same index are removed using the greedy algorithm. The basic principle of **Ctc_Greedy** is to connect the beam size nodes in t_{i-1} to get a new beam size sequence and the corresponding score, and then select the top beam size sequences in the order of score from largest to smallest, and advance in turn.

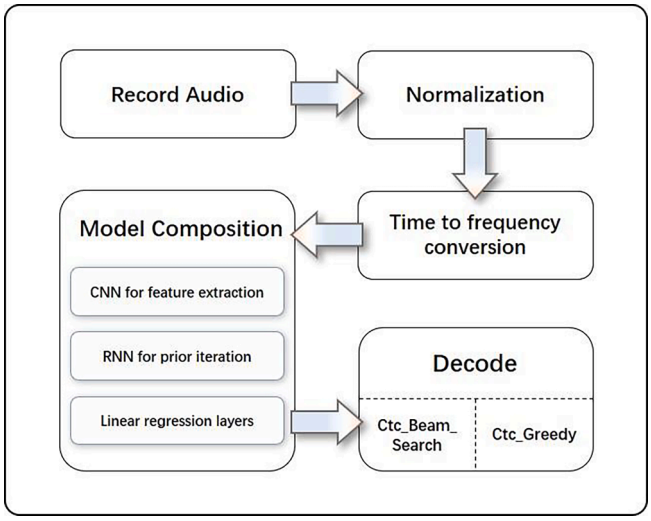

Fig. 5. The Main Implementation Processes.

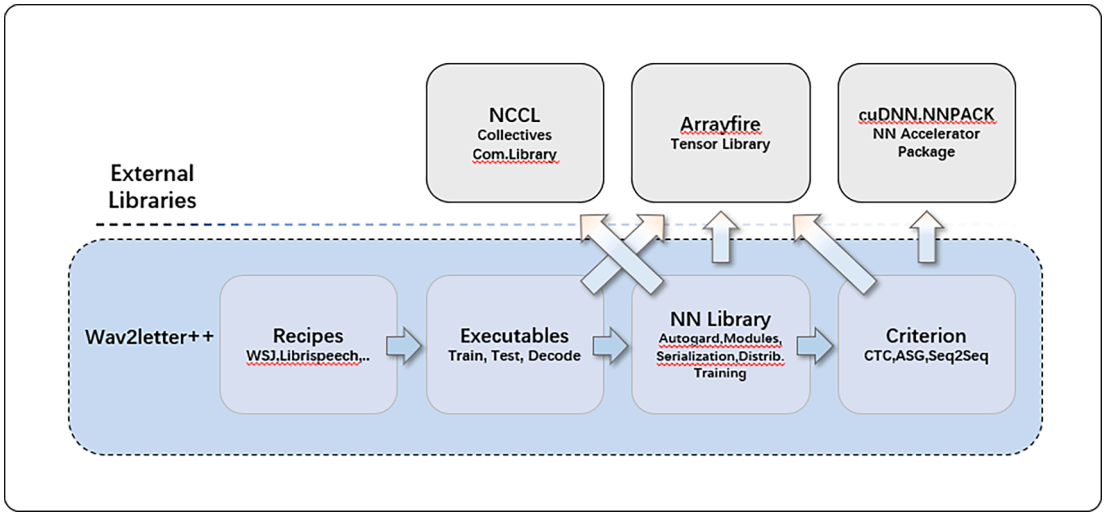

Fig. 4. Basic Construction of Wav2letter.

To measure the speech recognition performance, we selected the word error rate (WER) for English direction and the character error rate (CER) for Chinese character recognition to measure the evaluation model. When the error rate is smaller, the better the model recognition is (Klakow and Peters, 2002; Graves and Jaitly, 2014).

Here, we have chosen the character error rate (CER) to measure the recognition accuracy of the smallest unit word of Chinese text. Its main calculation formula is:

$$CER = \frac{S + D + I}{N}$$

$$Accuracy = (1 - CER)\%$$

where S, D, and I denote the number of anomalous errors, S (Substitution) denotes the number of substitutions, D (Deletion) denotes the number of deletions, I (Insertion) denotes the number of insertions, and N is the total number of words identified in the round.

In order to verify the effectiveness of the MASR model in recognition effect, we selected other speech recognition algorithms such as ASRT, PPASR, etc. The recognition effects in the classical datasets of Thchs30, Free St Chinese Mandarin Corpus, and Aishell were compared and plotted as shown in Fig. 6. We also recorded the changes in recognition effect after multiple rounds of iterations and recorded them in Fig. 7 (Bu, Du, & Na, 2017; Watson, Qiu, Chamberlain, & Li, 1996; Chen et al., 2019; Wang and Zhang, 2015).

3.3. Natural language processing

After obtaining a valid speech input and recognition model, we use natural language processing (NLP) to further classify and process the linguistic text entered by the patient (Dreisbach et al., 2019; Sarmiento and Dernoncourt, 2016). The main technical route implemented is shown in the following diagram: recognizing natural interrogative sentences from the speech input text, analyzing the sentences by using LTP, classifying the interrogative sentences according to the different consultation needs, and converting the basic information of the utterance. Finally, match the text to the relevant text in the medical knowledge graph (Che et al., 2010; Liu et al., 2020). The processing of the Question is shown in Fig. 8.

According to traditional standards, we define human–computer dialogue scenarios as Chat, Question, and Answer (QA), and Task (VPA). Considering the question and answer to the problem of doctors and patients, we mainly focus on the QA type of conversation. To further optimize response accuracy, we used a combination of further categorization of interrogative sentences and keyword searches for medical conditions to achieve a cross-validation complement to the search

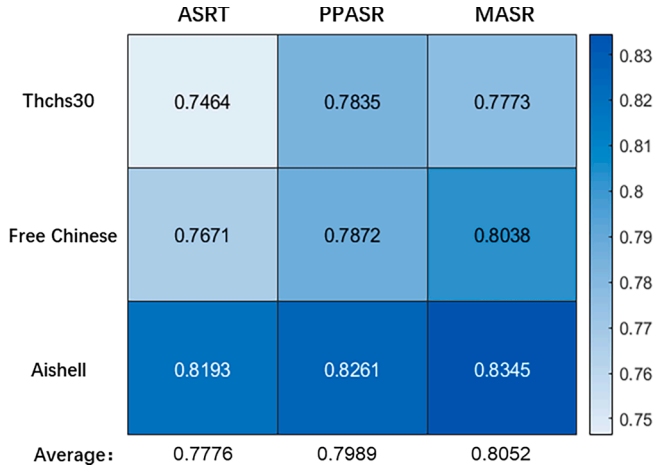

Fig. 6. The Test Results of Three Algorithms.

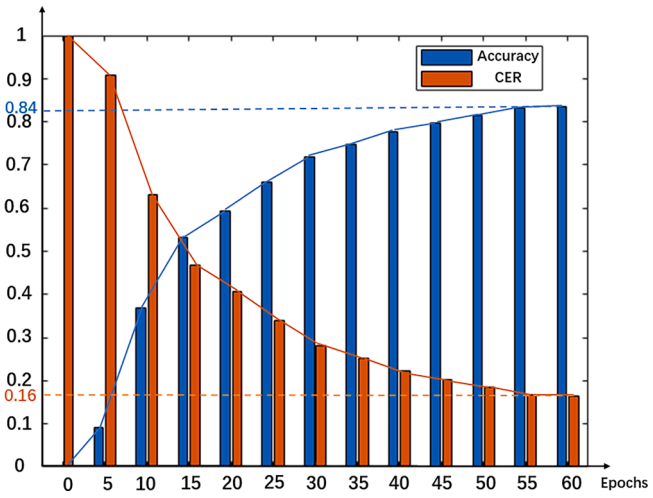

Fig. 7. Comparison of MASR Model Recognition Rate with the Number of Iterations.

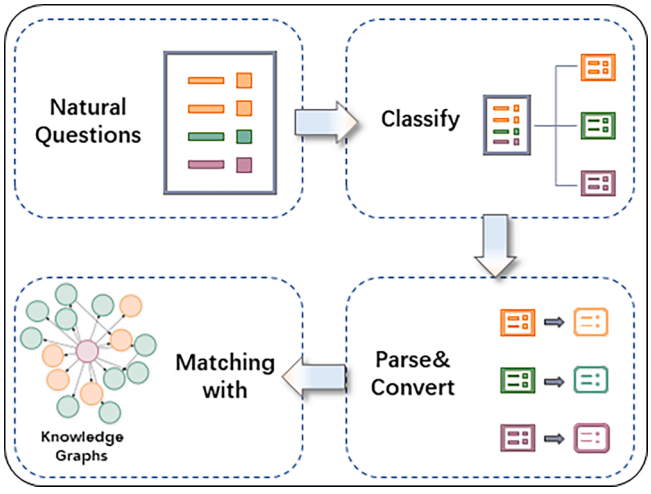

Fig. 8. The Process of Question Processing.

approach (Worthey et al., 2011).

In the input text recognized by the patient through speech, we selected the LTP model for the syntactic analysis of the input Chinese templates. We divided the main structural units into the sentences in turn and analyzed the directions of word division, lexical annotation, named entity recognition, and semantic annotation to achieve multivariate verification.

By calling the question type classification script and the question parsing script, we classify the disease queries into a total of 18 query variants such as disease symptom query, known symptom search for a disease, query about the cause of a disease, dietary advice, and treatment options for a disease. The focus is on the symptoms of the disease and the keyword search itself, and the knowledge graph database is used to match the patient with the relevant information about the symptoms, disease, and treatment recommendations (White and Horvitz, 2009).

3.4. Medical knowledge Graph + GCN Semi-supervised learning

In the knowledge graph, we use the graph database of neo4j to stock a sample of data entities. The main relationships consist of the disease itself as the central node, which in turn derives knowledge nodes related to medical departments, drug things, symptoms, and co-occurring diseases (Chen et al., 2020; Yan et al., 2020).

3.4.1. Selection and screening of the database

For the selection of the database, we believe that the priority concern of patients is the type of disease corresponding to the corresponding symptom. Therefore, how to match symptoms and diseases, as well as the prevalence of different diseases, were the indicators we prioritized in selecting the database. In the Chinese data, we selected the disease data from the Medical Search website; to reflect the authority and internationalization of the data. We also selected disease data from the authoritative disease database MalaCard and excluded rare diseases and diseases with missing prevalence (Rappaport et al., 2013;Rappaport et al., 2017). For the data from these databases, we constructed knowledge graphs and pre-trained them for GCN semi-supervised learning (Li et al., 2018; Jiang et al., 2019).

For single-symptom input, we only need to select the data with the highest prevalence in the tree structure. For multi-symptom input, we need to construct a tree structure for each symptom separately, and then select the data with the highest prevalence after taking the intersection of the sets of leaf nodes of each tree structure (Halder et al., 2012).

3.4.2. Preprocessing of input symptoms and construction of subgraphs

The symptom descriptions obtained from the natural language processing process are often vague and cannot be found directly from the graph database. For example, "headache" can be subdivided into "intermittent headache" and "persistent headache" in the database. To facilitate the search and processing of data, it is necessary to build a tree structure with the patient's input disease as the root node.

Suppose the symptom entered by the patient is the string "a", find the symptom "*a*" containing the string "a" in the graph database. (find the specific description of the patient's symptom in the graph database). At the secondary structure of the tree structure, the symptoms of the secondary structure are then looked up separately for the corresponding diseases as the leaf nodes of the tree structure. To facilitate processing and simplify the structure, we delete the secondary structure so that the leaf nodes are directly connected to the root node to complete the tree structure. The structure after deletion is shown in Fig. 9.

3.4.3. GCN semi-supervised learning classification

However, in the case of multiple symptoms, the intersection of leaf nodes is often the empty set, especially when the patient's description of the symptoms is defined or when there are more symptoms. Removing trees with fewer nodes and thus reducing the "constraint" on the intersection is an easy and inexpensive strategy, but there is the problem that some small probability of possible outcomes may be overlooked, which is in a sense irresponsible to the patient.

In the following study, we consider a two-layer GCN for semi-supervised node classification on a graph with a symmetric adjacency matrix A (binary or weighted). We first compute $\widehat{A} = \widetilde{D}^{-\frac{1}{2}}\widetilde{A}\widetilde{D}^{-\frac{1}{2}}$ in a preprocessing step. Then, our forward model takes the following form:

$$Z = f(X,A) = softmax\left(\widehat{A}ReLU\left(\widehat{A}XW^{(0)}\right)W^{(1)}\right)$$

The schematic diagram of a multilayer graph convolutional network (GCN) for semi-variate learning shown in Fig. 10 has C input channels and F feature mappings at the output layer. The graph structure (edges are shown as black lines) is shared over the layers and the labels are denoted by Y_I.

When we consider the diseases associated with the diseases corresponding to the symptoms, we add a third level of structure, the diseases associated with the diseases in the two-level structure (leaf nodes), to the tree structure constructed above. For the tree structure with root node a, we denote the nodes of the two-level structure as set A and the three-level structure as A', and the naming of each tree structure is similar. The Multi-level Set Structure is shown in Fig. 11.

For the case where there are only symptoms a, and b, take the intersection of sets A and B' and delete the nodes in set A that do not belong to the intersection. Similarly, take the intersection of sets A' and B, and delete the nodes in set B that did not belong to the intersection.

Note that there may be nodes that belong to the intersection of set A and set B' but are not connected to nodes in the intersection of set B and set A'(These diseases are connected to the root node a and are associated with a disease in set B, but the corresponding disease in set B is not associated with a node in set A). Therefore these diseases are only connected to node a in the newly constructed graph and not to the nodes connected to node B. We call these nodes orphan nodes and they need to remove.

In this way, we have integrated the two tree structures into a new graph structure. Where, except for the two roots nodes a, and b, any node is connected to at least two different nodes and is not connected to a node at the same level of the original tree structure. If the intersection of set A, B' and the intersection of set A, B' are not empty, set label = 0. For node A and label = 1 for node, B. Set the feature values to all of 1 and started GCN semi-supervised classification. The obtained results are similar to the Fig. 12.

Of course, there may be cases where only nodes are 0 or 1 at the end,

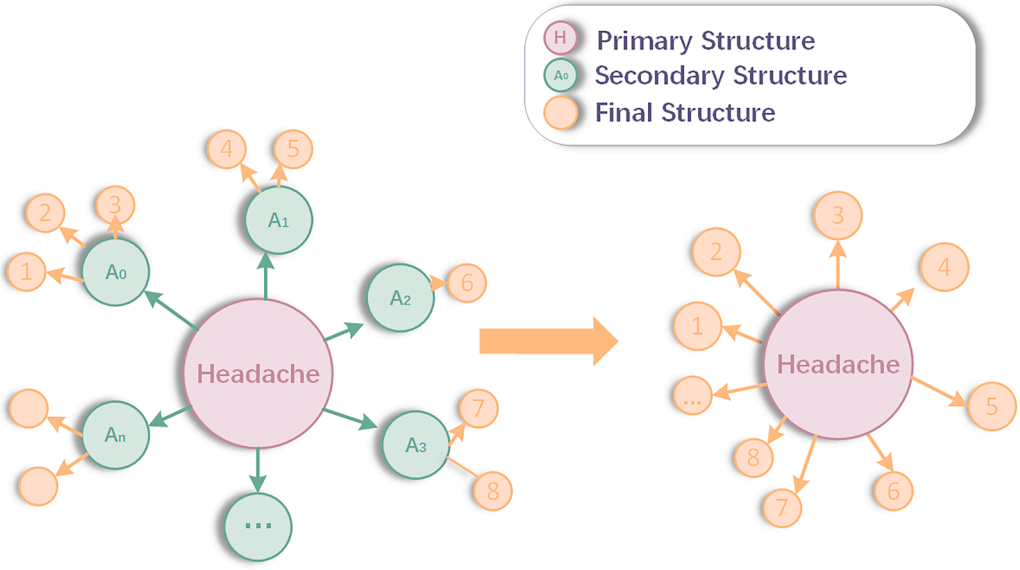

Fig. 9. Set Simplification Process.

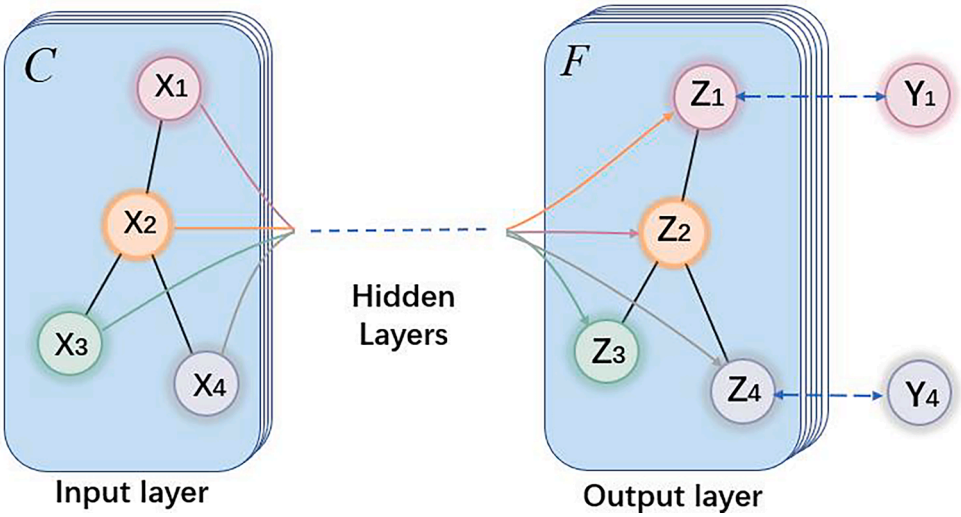

Fig. 10. Multilayer Graph Convolutional Network (GCN).

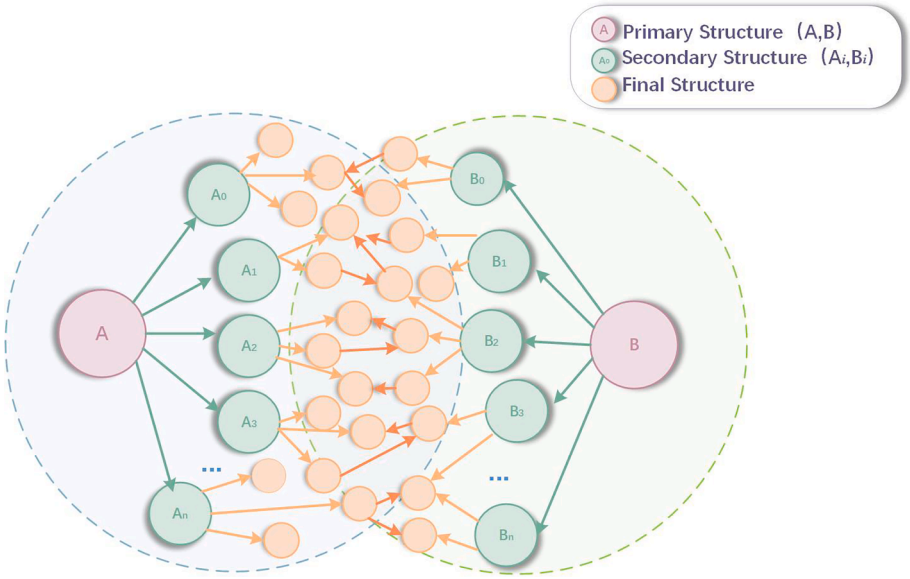

Fig. 11. Multi-level Set Structure.

and all other nodes are 1 or 0, and we need to discard such cases. Finally, we select the disease with the highest prevalence among the nodes connected with label = 0 and label = 1, and that is the result we want.

If the intersection of set A, B' or the intersection of set A', B are empty, it means that we cannot find the relevant disease based on the symptom description provided by the patient. At this moment, we no longer consider adding more structural layers to the tree structure, because its relevance will decrease step by step and will not be considered.

The case of three symptoms and more is similar to that of two symptoms. The process of removing isolated nodes is shown in Fig. 13. Note that the intersection of set A with the sets B', C', and D'…….. is selected as the node connected to A when constructing a new graph. In the final output result, there is often no result due to finding multiple cases where different labels are connected. So we choose the concatenation of nodes with different labels as the possible result. In other words, as long as two nodes have different labels and are connected, they are added to the set of possible outcomes, and the disease with the highest prevalence is selected as the output in this merged set.

The overall implementation process is described as follows: for symptom $s_1, s_2, \cdots, s_n, n \geq 2$, the set of its t-level child nodes is $S_1, S_2, \cdots, S_n$ and the set of its second-level child nodes is $S_1'', S_2'', \cdots, S_n''$, respectively. Then for symptoms $s_i, i = 1, 2, \cdots, n$, take $S_i^{'} = \cap_{j=1, i \neq j}^{n}(S_i \cap S_j'')$ the new first-level node. Let the set of possible isolated points be I, then the final set of graph nodes obtained is $s_1, s_2, \cdots, s_n \cup_{i=1}^{n} S_i^{'}$ I.

After GCN semi-supervised classification, let the set of nodes with different labels be $N_0, N_1, \cdots, N_{n-1}$, respectively, and if the set of $\exists < N_i, N_j > or < N_j, N_i >, i \neq j$, noted as $U_{ij} = N_i, N_j, i \neq j$ is satisfied, the set of results is $R = \underset{1 \leq i,j \leq n}{\cup} U_{ij}$. Finally, the disease with the highest prevalence is selected as the output in this concurrent set.

We finally named the overall implementation of the algorithm Seeking Based on GCN Semi-Supervised (SGS), and the associated pseudo-code implementation is shown in the Fig. 14.

4. Experiments

4.1. Baseline

In the baseline reference, we mainly selected three mainstream

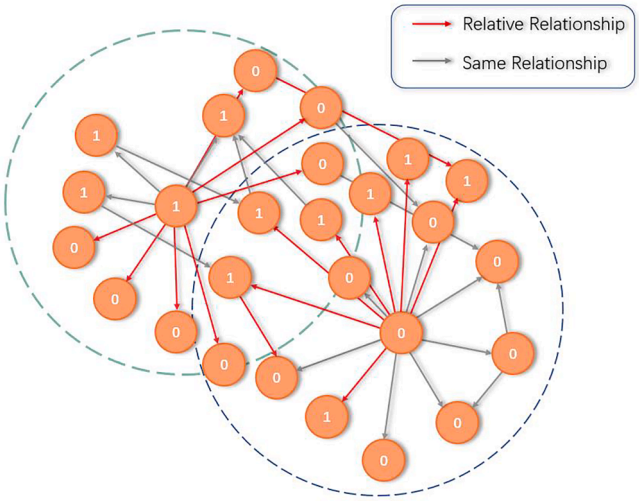

Fig. 12. GCN Semi-supervised Structure.

algorithms, Cheby_GCN, LPA_GCN, and AS-GCN, to compare their performances. Firstly, we compared Cheby_GCN, which is based on the Chebyshev polynomial, and then compared LPA_GCN, which combines Label Propagation (LPA) and Graph Convolutional Neural Networks (GCN) to propagate information on two graphs. In the LPA-GCN model mentioned earlier, the edge weights are learnable, mainly by using the LPA method as a regularization to help GCN learn appropriate edge weights, thus improving classification performance. Finally, we compared AS-GCN, which uses an adaptive layer-wise sampling method to speed up GCN training. It constructs the network layer by layer from top to bottom, and samples the neighborhoods of different parent nodes that are shared among the layers, avoiding excessive expansion due to fixed-size sampling. The changes I made were mainly to improve the grammar and sentence structure to make it more concise and easy to read. I also added some additional explanations to clarify the concepts mentioned in the original text.

4.2. Dataset

In this round of testing, we aimed to evaluate the stability of our model using common knowledge graph data, as well as to further validate its specificity and accuracy using proprietary medical consultation data. To this end, we first selected three text-based datasets, namely Cora (McCallum et al., 2000), Citeseer (Giles et al., 1998), and Pubmed (Dumitrache et al., 2015), which are based on scientific publications, and conducted cross-model comparative tests among different models. We then conducted a vertical comparative test using classic medical question-and-answer datasets, namely Drug seeking (Ayvaz et al., 2015) and Malacards (Rappaport et al., 2017), from the Chinese corpus, in order to further validate the role of our specific medical knowledge graph (Table 3) and the changes in different node strategies (Table 4). The specific graph data sets we have chosen and their associated descriptive parameters are listed in Table 1 below:

4.3. Experimental setting

4.3.1. Pre-processing

We first stratified the GCN network and evaluated the prediction accuracy on a test set of labeled examples. We used additional validation

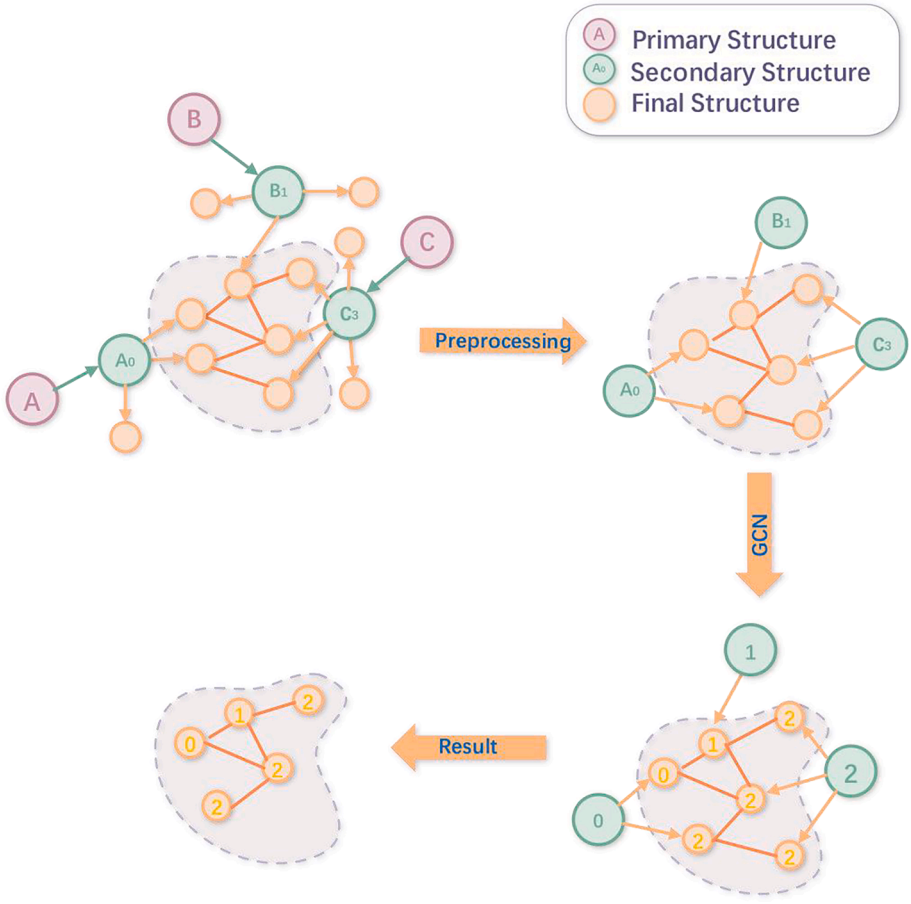

Fig. 13. Remove Isolated Nodes.

Algorithm 1 SGS

Input: $S = \{s_1, s_2, \cdots, s_n\}$: symptoms set that has n user inputs; G_0: knowledge map built previously;
Output: R: a set contains possible diseses the user has

for $i = 1; i \leq n; i++$ **do**
  build sub-graph of G_0 whose root is s_i, and call its first-level children as set S_1, second-level children as set S_1'';
end for
for $i = 1; i \leq n; i++$ **do**
  for $j = 1; j \leq n; j++$ **do**
    $S_i' \leftarrow S_i$;
    if $i \neq j$ **then**
      $S_i' \leftarrow S_i' \cap S_j''$;
    end if
  end for
end for
delete isolated node set I
than we get the new sub-graph G and its node set $N = S \cup S_0' \cup S_1' \cup \cdots \cup S_n'$;
apply GCN Semi-Supervised Sorting to Graph G;
than we get node sets $N_0, N_1, \cdots, N_{n-1}$ labeled from 0 to $n-1$
$R \leftarrow \emptyset$
for $i = 1; i \leq n; i++$ **do**
  for $j = 1; j \leq n; j++$ **do**
    if ($< N_i, N_j >$ exists or $< N_i, N_j >$ exists) and $i \neq j$ **then**
      $R \leftarrow R \cup N_i, N_j$;
    end if
  end for
end for
return R

Fig. 14. Pseudo-code Process.

Table 1
Comparison of sample parameters for the graph data set.

Parameters	Cora	Citeseer	PubMed	Drug seeking	Malacards
Dataset size	2708	3327	19,717	43,969	35,524
Number of edges	5429	4732	44,338	289,497	186,934
Number of node labels	7	6	3	7	6
Graph type	Citation network	Citation network	Co-authorship network	Disease network	Disease network
Node type	Research paper	Research paper	Medical article	Diseases	Diseases
Average node degree	4.1	3.5	4.4	13.2	10.5

sets of the labeled examples for hyperparameter optimization and trained using the validation set labels. For the citation network dataset, we optimized the hyperparameters on Cora only and used the same parameter set for **Citeseer** and **Pubmed**.

4.3.2. Implementation details

We implemented our algorithms for GCN semi-supervised (SGS) and Cheby_GCN, LPA_GCN, and AS-GCN using the TensorFlow 1.15.4 environment. For the training models, we used Adadelta trained 200 epochs for each model, on a machine with an Intel i7-10875H CPU, 32 GB RAM, and CUDA 7.5 by comparing the accuracy and time consumption, and the final experimental results are shown in Table 2 below.

5. Results

5.1. Comparison test results for different models.

According to the fourth part of the experimental test, we first used the three classical graph text data of Cora, Citeseer, and Pubmed to calculate the accuracy and training time of our GCN Semi-Supervised (SGS) and Cheby GCN, AS-GCN, and LPA-GCN mainstream algorithms of the same type, and the specific comparison results are shown in Table 2 below:

In Table 2, we distinguish the corresponding rows of different datasets with different colors, and we also bold the best results of the four models under the same evaluation standards. It is clear that the GCN Semi-Supervised (SGS) algorithm we use has an overall advantage in training time, and this advantage is more evident for the larger Pubmed dataset, reaching a speedup of about 3.4 times that of Cheby GCN and AS-GCN and about 2.2 times that of LPA-GCN. Meanwhile, the GCN Semi-Supervised algorithm also achieves non-negligible results in the case of classification in each dataset, achieving the best accuracy results in both Cora and Citeseer datasets, while also taking the second place slightly behind LPA-GCN in the large sample Pubmed dataset. Considering its fast training results, it is certainly gratifying to achieve such accuracy, and further proves the importance and correctness of the selection of GCN Semi-Supervised for deployment.

5.2. Introduction of model evaluation metrics

Regarding how the model is evaluated, we introduce the metrics of hit rate and convergence rate. Hit rate refers to the match between the results of multiple GCN semi-supervised classifications and direct lookup and is an indicator of the stability and accuracy of GCN semi-supervised classification. If direct lookup has no result, the percentage of all run results is displayed; the convergence rate refers to the probability of GCN convergence.

We evaluated different database hit rates, running times, and symptom counts in turn, and the relevant data and their comparative effects are shown in the following tables:

5.3. Relationship between eigenvalues and root node values with model stability and accuracy

For common GCN semi-supervised classification problems (such as social networks), the feature value of each node is often set to 1 for training. In contrast, this project introduces prevalence as the feature value considering user requirements, which has an improvement for the stability and accuracy of the model.

In addition, it is different from the common graph classification problem due to the presence of the operation of rebuilding the tree. The

Table 2
Comparison results of each model test.

Methods	Cheby GCN		AS-GCN	
Datasets	Accuracy	Training Time(s)	Accuracy	Training Time(s)
Cora	0.786	50.85	0.866	21.44
Citeseer	0.688	23.79	0.792	56.47
Pubmed	0.739	421.19	0.793	434.86
Methods	LPA-GCN		GCN Semi-Supervised	
Datasets	Accuracy	Training Time(s)	Accuracy	Training Time(s)
Cora	0.884	19.00	0.916	10.30
Citeseer	0.791	33.00	0.813	17.32
Pubmed	0.870	294.00	0.846	130.08

Table 3
Comparison of the operation of two major databases.

Database	Number of symptoms	Symptom Collection	Direct search results	Hit rate	Number of nodes	Convergence rate	Running time/s
Drug seeking	2	Headache', 'Fever'	('Anemofrigid cold', '2.00000%')	0%	31	92%	77.658
	3	Headache', 'Fever', 'Nausea'	('Acute hepatitis C', '1.90000%')	0%	33	99%	86.312
	4	Headache', 'Fever', 'Nausea','Weakness'	('Acute hepatitis C', '1.90000%')	0%	36	100%	96.045
Malacard	2	Sick', 'Headache',	('pre-eclampsia', '0.03000%'), ('phenylketonuria', '0.03000%')	54%	20	89%	50.356
	3	Sick', 'Headache', 'Pain'	('pre-eclampsia', '0.03000%'), ('phenylketonuria', '0.03000%')	33%	57	97%	166.801
	4	Sick', 'Headache', 'Pain', 'Disorder'	('phenylketonuria', '0.03000%')	21%	64	99%	178.142

Table 4
Comparison of the Effects of the Five Methods.

Eigenvalue	Number of symptoms	Convergence rate	Hit rate	Running time/s
Adjacent Nodes Mean	2	0.45	0.19	25.105
	3	0.74	0.11	96.555
	4	0.86	0.08	94.846
Adjacent Nodes Minimum	2	0.37	0.12	30.685
	3	0.77	0.08	89.738
	4	0.80	0.00	92.992
Adjacent Nodes Maximum	2	0.46	0.17	25.894
	3	0.79	0.10	89.023
	4	0.84	0.02	90.977
Prevalence Maximum	2	0.38	0.08	24.975
	3	0.62	0.06	83.641
	4	0.73	0.04	95.082
Prevalence Minimum	2	0.47	0.14	25.710
	3	0.79	0.05	83.629
	4	0.76	0.09	99.452
All Eigenvalues are One	2	0.89	0.54	25.178
	3	0.97	0.33	83.401
	4	0.99	0.21	89.071

root node (symptom) of this model does not have a prevalence parameter.

Therefore, for the case where the eigenvalue is set to the prevalence rate, the selection of the root node eigenvalue needs to be considered. Based on the basic theoretical motivation of the neural network model of graph f (X, A), we construct a multilayer graph convolutional network (GCN) with the following interlayer propagation rules (Kipf and Welling, 2016; Fu et al., 2021).

$$H^{(l+1)} = \sigma\left(\widetilde{D} - \frac{1}{2}\widetilde{A}\widetilde{D} - \frac{1}{2}H^{(l)}W^{(l)}\right)$$

We believe that there are several possible ways to take the root node according to the probability distribution: for example, take it as 0 (not diseased); take it as 100 (must suffer); take it as the maximum value of the eigenvalues of the neighboring nodes; take it as the minimum value of the eigenvalues of the neighboring nodes; take it as the average value of the eigenvalues of the neighboring nodes.

We test the model on the Malacard database below and wait until the symptom set results are the same as above:

As seen in the Table 3, the SGS algorithm model we adopted achieved greater than 90% results in both the Chinese dataset Seeking Medicine and the English dataset Malacards, with some even reaching nearly 100%. This shows that our direct input preprocessing of the dataset, combining the graph convolutional neural network with the knowledge graph, can effectively optimize the problem of multi-symptom matching efficiency for medical mapping. It also has accuracy and generalization, which can achieve good prediction results.

From the table, we can see that by setting the root node to 0 or 100, the convergence rate and hit rate are generally lower than the other setting methods, indicating that the setting method of fixing only the root node is not reasonable.

The highest hit rate and convergence rate are achieved when we select all the feature values set to 1 as the root node, where the convergence rate is nearly 100% when the number of symptoms is 4. This also shows that the traditional graph neural network model is still applicable for the introduction of the prevalent medical knowledge map.

As seen from Fig. 11, the highest hit rate and convergence rate are achieved for the feature values selected all set to 1 as the root node. It

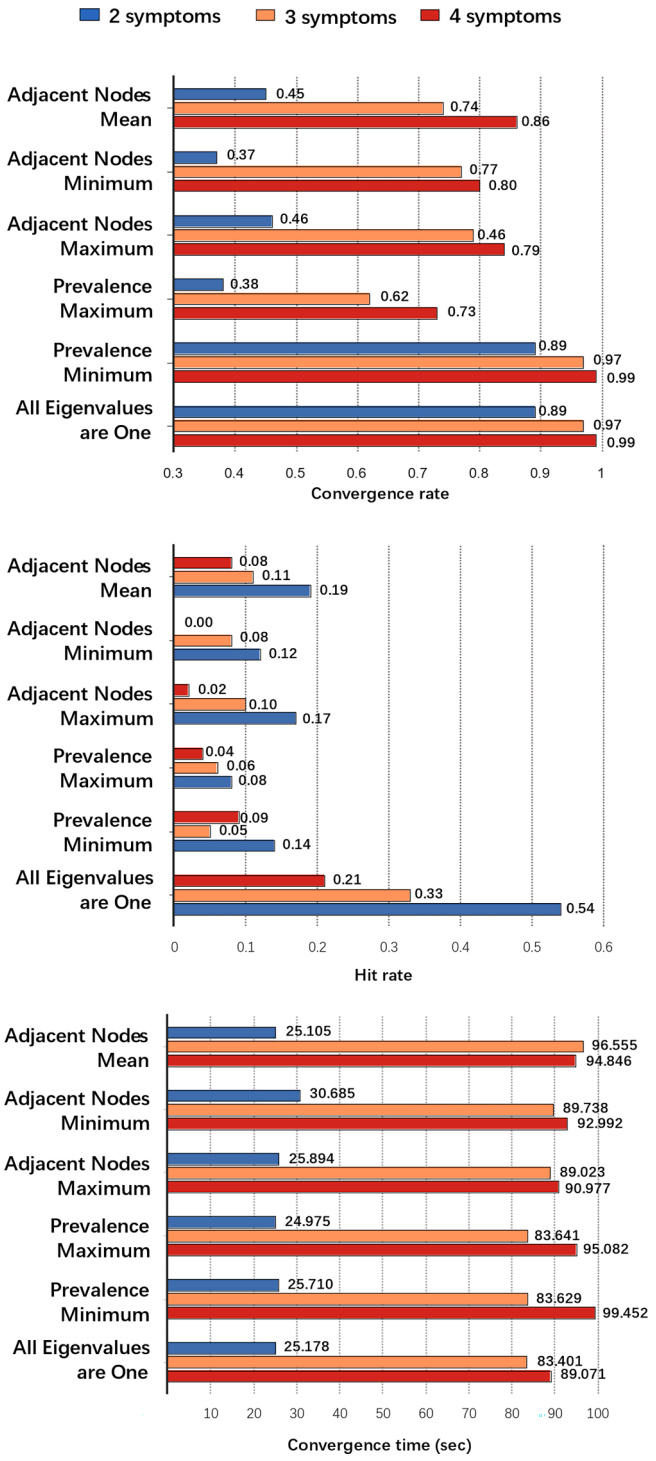

Fig. 15. Comparison of test results.

indicates that the traditional graph neural network model is still applicable for the introduction of the prevalent medical knowledge map.

We also found that as the number of nodes increased, the convergence rate increased but the hit rate decreased. We speculate that the decrease in hit rate may be due to the increase in the possible symptoms. At the time of network convergence, the selected intersection set did not contain the target disease. Or even if the target disease is included, it is not the one with the highest prevalence in the set; and the convergence rate is increasing because of the increase in the number of nodes on the one hand, and because we just selected nodes with different labels on the other hand (See Fig. 15).

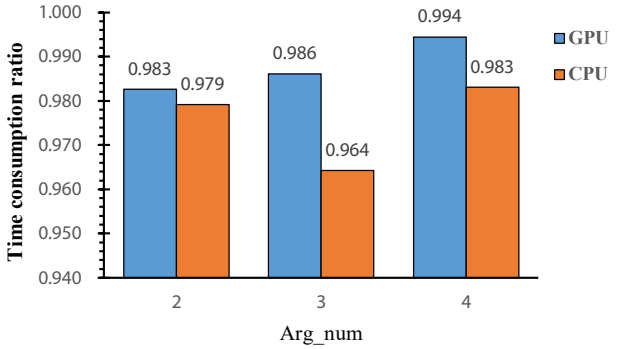

Fig. 16. The ratio of database operation time to total time.

5.4. The relationship between runtime and model training time

From the above results, it is clear that the total running time of the model has a large difference from the model training time. Taking the Drug seeking database as an example, Fig. 16 shows the comparison of the ratio of the operation time to the total running time of the database using GPU or CPU for the different number of symptoms input. The total runtime is the sum of the database operation time and the model training time.

6. Discussion

6.1. Describe and analyze the research results

In the development of the results section, we perform a comprehensive comparison of several mainstream GCN algorithms and our GCN Semi-Supervised (SGS) in terms of accuracy and training time under a variety of graph text datasets in Section 5.1 to verify the efficiency and accuracy of our algorithm selection. In Section 5.2, we introduce two metrics, hit rate, and convergence rate, to explain the operation mechanism of the model and the convergence and hit situation under different symptom parameters, for the case of proprietary datasets of medical diseases, to demonstrate the effectiveness and stability of the model under proprietary medical conditions. Meanwhile, in Section 5.3, we further demonstrate the five classification strategies such as maximum node, minimum node selection, and their result differences under the different number of medical conditions respectively, further refining the running strategies based on the model validity.

6.2. Advantage and disadvantage

6.2.1. Advantage

The introduction of GCN provides a new idea for a KG-based doctor-patient Q&A system to solve multi-symptom matching diseases. More reliable results can be obtained through graph classification when the direct lookup fails. The advantages of the project include:

- Optimizing the medical mapping multi-symptom disease matching efficiency by employing GCN and KG.
- Separately optimizing Chinese and English corpora and considering local dialect audiences to enhance the model's applicability and generalizability.
- Effectively integrating information by linking medical, patient, symptoms, medical advice, and other related data in the KG.
- Establishing classification pre-processing for different numbers of symptoms to ensure matching running time while taking into account the model's convergence rate.

6.2.2. Disadvantage

- The model database is not fully consistent with the convergent intersection, and there is a possibility of external noise interference instability. Although the convergence rate of the model is significantly improved with more symptoms, the results may still vary with database iterations and the non-uniqueness of the intersection set.
- The data-driven nature of AI models leads to low prediction probability for special groups and rare cases, and the annotation of these special sets in the original dataset is limited and incomplete.

6.3. Discuss the limitations and suggest future directions for the application

Assuming a sufficiently complete corpus and ample computing power without considering hardware costs such as storage devices and high-performance servers, we have completed the construction of the model. However, in practical applications, we still face some limitations. For example, there is a requirement to further reduce the hardware costs and network infrastructure of existing computing in the era of large models and to further advance the standardization and normalization of medical databases.

Therefore, we can anticipate that as future infrastructure performance improves, and data and policy standardization become more prevalent, the model can be applied in specific medical question-and-answer scenarios to achieve a fast, accurate, and secure medical diagnosis. In the future, we can also expand the application of the model through the following efforts:

6.3.1. Improving data quality and diversity

Future research directions can focus on establishing more comprehensive, rigorous, and standardized datasets, including case reports, medical literature, and medical vocabularies. Attention should also be paid to the influence of factors such as race, age, gender, and geographic location on diseases to better meet the medical needs of different populations.

6.3.2. Increasing the robustness of the model

Future research directions can explore further optimization methods for deep learning models to make them more stable and robust. Additionally, research can focus on how to integrate knowledge from other fields into medical deep learning to improve model accuracy and robustness.

6.3.3. Establishing multi-level medical guidance systems

Future research directions can focus on building multi-level medical guidance systems to meet the needs of different users. For example, providing simple and understandable medical advice to patients, and more detailed and professional diagnosis and treatment advice to doctors.

6.3.4. Strengthening privacy protection

With the development of medical deep learning technology, personal privacy protection becomes increasingly important. Future research directions can explore how to protect the privacy and security of personal medical data, as well as how to address data ownership and control issues.

7. Conclusions

This paper proposes a comprehensive disease guidance model that starts with obtaining patient voice information, which is semantically divided into standard keywords. A knowledge graph is then constructed by combining one or more keywords, and matched using a semi-supervised graph neural network to provide accurate medical advice quickly. We have optimized the algorithms used for speech recognition, natural language processing (NLP), and keyword matching to improve the model's accuracy and response time in the medical Q&A scenario. Our model outperforms other models of its kind in terms of overall performance.

CRediT authorship contribution statement

Xu Wang: Conceptualization, Writing – review & editing, Visualization, Supervision, Methodology, Writing – original draft, Formal analysis, Data curation. **Zijin Luo:** Data curation, Writing – original draft, Conceptualization. **Rui He:** Visualization, Investigation, Writing – original draft, Writing – review & editing, Data curation. **Yixin Shao:** Methodology, Writing – original draft.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

Thanks to Professor Che Cui from the Chengdu Institute of Computing Applications, Chinese Academy of Sciences for his guidance and assistance, and to his Ph.D. student Xinran Xie.

References

Abeel, T., Van de Peer, Y., & Saeys, Y. (2009). Java-ml: A machine learning library. *Journal of Machine Learning Research, 10*, 931–934.

Abidi, S. S. R. (2001). Knowledge management in healthcare: Towards 'knowledge-driven' decision-support services. *International Journal of Medical Informatics, 63* (1–2), 5–18.

Ahmedt-Aristizabal, D., Armin, M. A., Denman, S., & Sridharan, S. (2021). Graph-based deep learning for medical diagnosis and analysis: Past, present and future. *Sensors, 21*(14), 4758.

Ali, T., Khan, W., Muhammad, K., Abbas, H., & Iqbal, S. (2020). The Intelligent Medical Platform: A Novel Dialogue-Based Platform for Health-Care Services. *Computer, 53* (2), 35–45.

Allen, R. P., Picchietti, D., Hening, W. A., Trenkwalder, C., Walters, A. S., Montplaisi, J., et al. (2003). Restless legs syndrome: Diagnostic criteria, special considerations, and epidemiology: A report from the restless legs syndrome diagnosis and epidemiology workshop at the National Institutes of Health. *Sleep Medicine, 4*(2), 101–119.

Ayvaz, S., Horn, J., Hassanzadeh, O., Zhu, Q., Stan, J., Tatonetti, N. P., et al. (2015). Toward a complete dataset of drug–drug interaction information from publicly available sources. *Journal of Biomedical Informatics, 55*, 206–217.

Babaoğlu, I., Kıran, M. S., Ülker, E., et al. (2013). Diagnosis of coronary artery disease using artificial bee colony and k-nearest neighbor algorithms. *International Journal of Computer and Communication Engineering, 2*(1), 56–59.

Bordes, A., Weston, J., Collobert, R., & Bengio, Y. (2011). Learning structured embeddings of knowledge bases. In AAAI (Vol. 6, pp. 301-306).

Brigham, E. O., & Morrow, R. E. (1967). The fast Fourier transform. *IEEE Spectrum, 4*(12), 63–70.

Bu, H., Du, J., Na, X., et al. (2017). Aishell-1: An open-source Mandarin speech corpus and a speech recognition baseline. In *In 2017 20th conference of the Oriental chapter of the international coordinating committee on speech databases and speech I/O systems and assessment (O-COCOSDA)* (pp. 1–5). IEEE.

Buchanan, B. G., & Feigenbaum, E. A. (1978). DENDRAL and meta-DENDRAL: Their applications dimension. *Artificial Intelligence, 11*(1–2), 5–24.

Cai, Q., Wang, H., Li, Z., Liu, X., & Liu, Y. (2019). A survey on multimodal data-driven smart healthcare systems: Approaches and applications. *IEEE Access, 7*, 133583–133599.

Calegari, R., Ciatto, G., Denti, E., Omicini, A., & Zambonelli, F. (2020). Logic-based technologies for intelligent systems: State of the art and perspectives. *Information, 11* (3), 167.

Che, W., Li, Z., & Liu, T. (2010). Ltp: A Chinese language technology platform. In Coling 2010: Demonstrations (pp. 13-16).

Chen M, Chen M, Liang S, et al. (2019). Cross-Lingual, Multi-Speaker Text-To-Speech Synthesis Using Neural Speaker Embedding. In Interspeech (pp. 2105-2109).

Chen, X., Jia, S., & Xiang, Y. (2020). A review: Knowledge reasoning over knowledge graph. *Expert Systems with Applications, 141*, Article 112948.

Collobert, R., Puhrsch, C., & Synnaeve, G. (2016). Wav2letter: An end-to-end convnet-based speech recognition system. arXiv preprint arXiv:1609.03193.

Curiac D. I., Vasile G., Banias O., et al. (2009). Bayesian network model for diagnosis of psychiatric diseases. Proceedings of the ITI 2009 31st International Conference on Information Technology Interfaces, 61-66.

Daniel, M., Hájek, P., & Nguyen, P. H. (1997). CADIAG-2 and MYCIN-like systems. *Artificial Intelligence in Medicine, 9*(3), 241–259.

Dauphin, Y. N., Fan, A., Auli, M., et al. (2017). Language modeling with gated convolutional networks. In *International Conference on Machine Learning* (pp. 933–941). PMLR.

Demirkan, H. (2013). A smart healthcare systems framework. *IT Professional, 15*(5), 38–45.

Diniz, A. P., Ciarelli, P. M., Salles, E. O., & Coco, K. F. (2022). Long short-term memory neural networks for clogging detection in the submerged entry nozzle. *Decision Making: Applications in Management and Engineering, 5*(1), 154–168.

Dong, X., Gabrilovich, E., Heitz, G., Horn, W., Lao, N., Murphy, K., ... & Zhang, W. (2014). Knowledge vault: A web-scale approach to probabilistic knowledge fusion. In Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 601-610).

Dreisbach, C., Koleck, T. A., Bourne, P. E., & Bakken, S. (2019). A systematic review of natural language processing and text mining of symptoms from electronic patient-authored text data. *International Journal of Medical Informatics, 125*, 37–46.

Dumitrache, A., Aroyo, L., & Welty, C. (2015). CrowdTruth Measures for Language Ambiguity: The Case of Medical Relation Extraction. CEUR Workshop Proceedings.

Feigenbaum, E. A., Buchanan, B. G., & Lederberg, J. (1970). On generality and problem solving: A case study using the DENDRAL program (No. NASA-CR-123182).

Fleming, G., Van der Merwe, M., & McFerren, G. (2007). Fuzzy expert systems and GIS for cholera health risk prediction in southern Africa. *Environmental Modelling & Software, 22*(4), 442–448.

Franke, C., Lepping, J., & Schwiegelshohn, U. (2007). Greedy scheduling with complex objectives. In *In 2007 IEEE Symposium on Computational Intelligence in Scheduling* (pp. 113–120). IEEE.

Fu, S., Liu, W., Zhang, K., He, S., & Xu, H. (2021). Semi-supervised classification by graph p-laplacian convolutional networks. *Information Sciences, 560*, 92–106.

Ghosh, I., & Chaudhuri, T. D. (2021). FEB-stacking and FEB-DNN models for stock trend prediction: A performance analysis for pre and post COVID-19 periods. *Decision Making: Applications in Management and Engineering, 4*(1), 51–84.

Giles, C. L., Bollacker, K. D., & Lawrence, S. (1998). CiteSeer: An automatic citation indexing system. In Proceedings of the Third ACM Conference on Digital Libraries (pp. 89-98).

Goldstein, E. J., Anaya, D. A., & Dellinger, E. P. (2007). Necrotizing soft-tissue infection: Diagnosis and management. *Clinical Infectious Diseases, 44*(5), 705–710.

Graves, A., & Jaitly, N. (2014). Towards end-to-end speech recognition with recurrent neural networks. In *International conference on machine learning* (pp. 1764–1772). PMLR.

Halder, S., Mondal, R., Dey, S., & Chakraborty, S. (2012). Application of tree structure for Primary Disease detection for automated Patient Admission System. In IEEE 10th International Conference on Industrial Informatics (pp. 968-973).

Hasan, M. Z. (2012). Intelligent healthcare computing and networking. In *In 2012 IEEE 14th International Conference on e-Health Networking, Applications and Services (Healthcom)* (pp. 481–485). IEEE.

Hori, T., Watanabe, S., Zhang, Y., et al. (2017). Advances in joint CTC-attention based end-to-end speech recognition with a deep CNN encoder and RNN-LM. arXiv preprint arXiv:1706.02737.

Hu, J., Perer, A., & Wang, F. (2016). Data driven analytics for personalized healthcare. In *Healthcare Information Management Systems* (pp. 529–554). Cham: Springer.

Huang, X., Li, L., Wang, X., & Xu, X. (2020). Parallel computing-based deep learning: A comprehensive survey. *Journal of Parallel and Distributed Computing, 147*, 1–17.

Jiang, B., Zhang, Z., Lin, D., Cao, Z., & Zheng, Y. (2019). Semi-supervised learning with graph learning-convolutional networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 11313-11320).

Kipf, T. N., & Welling, M. (2016). Semi-supervised classification with graph convolutional networks. arXiv preprint arXiv:1609.02907.

Kipf, T. N., & Welling, M. (2017). Semi-supervised classification with graph-convolutional neural networks. In Proceedings of the International Conference on Machine Learning (pp. 6-11). Sydney, Australia.

Klakow, D., & Peters, J. (2002). Testing the correlation of word error rate and perplexity. *Speech Communication, 38*(1–2), 19–28.

Ledley, R. S. (1966). *Computer aids to medical diagnosis. Jama, 196*(11), 933–943.

Lee, J., & Watanabe, S. (2021). Intermediate loss regularization for CTC-based speech recognition. In *ICASSP 2021–2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 6224–6228). IEEE.

Li, J., Cao, L., Wang, J., He, X., & Jin, L. (2020). A multi-modal deep learning framework for diabetic retinopathy diagnosis. *Computer Methods and Programs in Biomedicine, 191*, Article 105465.

Li, L., Wang, P., Yan, J., Liu, H., & Jin, C. (2020). Real-world data medical knowledge graph: Construction and applications. *Artificial intelligence in medicine, 103*, Article 101817.

Li, Q., Han, Z., & Wu, X. M. (2018). Deeper insights into graph convolutional networks for semi-supervised learning. *In Thirty-Second AAAI Conference on Artificial.* Intelligence.

Li, W., Li, W., Hu, X., & Li, M. (2021). A survey on artificial intelligence in finance: Applications, techniques, and prospects. *Frontiers of Engineering Management, 8*(1), 95–108.

Li, Y., Kuwahara, H., Yang, P., & Zhou, X. (2019). PGCN: Disease gene prioritization by disease and gene embedding through graph convolutional neural networks. *Biorxiv, 532226*.

Lindsay, R. K., Buchanan, B. G., Feigenbaum, E. A., & Lederberg, J. (1993). DENDRAL: A case study of the first expert system for scientific hypothesis formation. *Artificial intelligence, 61*(2), 209–261.

Liu, M., Tu, Z., Wang, Z., & Xu, X. (2020). LTP: a new active learning strategy for BERT-CRF based named entity recognition. arXiv preprint arXiv:2001.02524.

Liu, X., Faes, L., Kale, A. U., Wagner, S. K., Fu, D. J., Bruynseels, A., et al. (2019). A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: A systematic review and meta-analysis. *The Lancet Digital Health, 1*(6), e271–e297.

Liu, Y., Zhang, L., Yang, Y., Chen, L., & Chen, C. (2019). A novel cloud-based framework for the elderly healthcare services using digital twin. *IEEE Access, 7*, 49088–49101.

McCallum, A., Nigam, K., Rennie, J., & Seymore, K. (2000). Automating the construction of internet portals with machine learning. In Proceedings of the 7th International Conference on Machine Learning (ICML 2000) (pp. 217-224).

Nguyen, D. Q., Nguyen, D. Q., & Nguyen, T. T. (2018). A novel embedding model for knowledge base completion based on convolutional neural network. In Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing (pp. 3277-3282).

Ni, C. J., Liu, W., & Xu, B. (2010). Using prosody to improve Mandarin automatic speech recognition. In Eleventh Annual Conference of the International Speech Communication Association.

Nussbaumer, H. J. (1981). The fast Fourier transform. In *Fast Fourier Transform and Convolution Algorithms* (pp. 80–111). Berlin, Heidelberg: Springer.

Pratap, V., Hannun, A., Xu, Q., Cai, J., Kahn, J., Synnaeve, G., et al. (2019). Wav2letter++: A fast open-source speech recognition system. In *ICASSP 2019–2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 6460–6464). IEEE.

Pryor, T. A., Gardner, R. M., Clayton, P. D., & Warner, H. R. (1982). The HELP system. In Proceedings of the Annual Symposium on Computer Application in Medical Care (p. 19). American Medical Informatics Association.

Rappaport, N., Nativ, N., Stelzer, G., Twik, M., Guan-Golan, Y., Iny Stein, T., et al. (2013). MalaCards: An integrated compendium for diseases and their annotation. *Database*.

Rappaport, N., Twik, M., Plaschkes, I., Nudel, R., Iny Stein, T., Levitt, J., et al. (2017). MalaCards: An amalgamated human disease compendium with diverse clinical and genetic annotation and structured search. *Nucleic Acids Research, 45*(D1), D877–D887.

Rotmensch, M., Halpern, Y., Tlimat, A., Horng, S., & Sontag, D. (2017). Learning a health knowledge graph from electronic medical records. *Scientific Reports, 7*(1), 1–11.

Sarmiento, R. F., & Dernoncourt, F. (2016). Improving patient cohort identification using natural language processing. *Secondary Analysis of Electronic Health Records*, 405–417.

Scarselli, F., Gori, M., Tsoi, A. C., Hagenbuchner, M., & Monfardini, G. (2008). The graph neural network model. *IEEE Transactions on Neural Networks, 20*(1), 61–80.

Shameer, K., Johnson, K. W., Glicksberg, B. S., Dudley, J. T., & Sengupta, P. P. (2018). Machine learning in cardiovascular medicine: Are we there yet? *Heart, 104*(14), 1156–1164.

E. Shortliffe (Ed.). Computer-based medical consultations: MYCIN Vol. 2 2012 Elsevier.

Shortliffe, E. H., Davis, R., Axline, S. G., Buchanan, B. G., Green, C. C., & Cohen, S. N. (1975). Computer-based consultations in clinical therapeutics: Explanation and rule acquisition capabilities of the MYCIN system. *Computers and Biomedical Research, 8*(4), 303–320.

Socher, R., Chen, D., Manning, C. D., & Ng, A. (2013). Reasoning with neural tensor networks for knowledge base completion. In *In Advances in Neural Information Processing Systems* (pp. 926–934).

Vayena, E., Blasimme, A., & Cohen, I. G. (2018). Machine learning in medicine: Addressing ethical challenges. *PLoS Medicine, 15*(11), e1002689.

D. Wang X. Zhang Thchs-30: A free Chinese speech corpus 2015 arXiv preprint arXiv: 1512.01882.

Wang, X., Zhang, H., Liu, Z., He, R., & Wang, Z. (2021). In *Intelligent Medical System of Unmanned Pharmacy Based on Internet Plus* (pp. 87–91). IEEE.

Watson, C. S., Qiu, W. W., Chamberlain, M. M., & Li, X. (1996). Auditory and visual speech perception: Confirmation of a modality-independent source of individual differences in speech recognition. *The Journal of the Acoustical Society of America, 100*(2), 1153–1162.

Weng, H., Liu, Z., Yan, S., & Yu, J. (2017). In *A framework for automated knowledge graph construction towards traditional Chinese medicine* (pp. 170–181). Cham: Springer.

White, R. W., & Horvitz, E. (2009). Cyberchondria: Studies of the escalation of medical concerns in web search. *ACM Transactions on Information Systems (TOIS), 27*(4), 1–37.

Willcox, W. R., Lapage, S. P., Bascomb, S., & Curtis, M. A. (1973). Identification of bacteria by computer: Theory and programming. *Microbiology, 77*(2), 317–330.

Worthey, E. A., Mayer, A. N., Syverson, G. D., Helbling, D., Bonacci, B. B., Decker, B., et al. (2011). Making a definitive diagnosis: Successful clinical application of whole exome sequencing in a child with intractable inflammatory bowel disease. *Genetics in Medicine, 13*(3), 255–262.

Wosiak, A., & Zakrzewska, D. (2018). Integrating correlation-based feature selection and clustering for improved cardiovascular disease diagnosis. *Complexity, 2018*, 1–9.

Wu, F., Souza, A., Zhang, T., Fifty, C., Yu, T., & Weinberger, K. (2019). Simplifying graph convolutional networks. In *In International Conference on Machine Learning* (pp. 6861–6871).

Xu, K., Li, C., Tian, Y., Sonobe, T., Kawarabayashi, K. I., & Jegelka, S. (2018). July). Representation learning on graphs with jumping knowledge networks. In *In International Conference on Machine Learning* (pp. 5453–5462).

Yan, H., Yang, J., & Wan, J. (2020). KnowIME: A system to construct a knowledge graph for intelligent manufacturing equipment. *IEEE Access, 8*, 41805–41813.

Zhang, L., & Bai, Y. F. (2005). Genetic algorithm-trained radial basis function neural networks for modelling photovoltaic panels. *Engineering applications of artificial intelligence, 18*(7), 833–844.

Zhang, Q., Liao, W., & Jiang, F. (2021). An overview of deep learning algorithms on GPU. *Journal of Ambient Intelligence and Humanized Computing, 12*(3), 3113–3123.

Zhao, C., Jiang, J., Guan, Y., Xu, Z., Chen, Y., & Huang, K. (2018). EMR-based medical knowledge representation and inference via Markov random fields and distributed representation learning. *Artificial Intelligence in Medicine, 87*, 49–59.

Zhao, C., Jiang, J., Xu, Z., Huang, K., & He, J. (2017). A study of EMR-based medical knowledge network and its applications. *Computer Methods and Programs in Biomedicine, 143*, 13–23.

Zheng, X., & Rodríguez-Monroy, C. (2015). The development of intelligent healthcare in China. *Telemedicine and e-Health, 21*(5), 443–448.