



OPEN ACCESS

EDITED BY

Jose Ramon Saura,
Rey Juan Carlos University, Spain

REVIEWED BY

Francisco Javier S. Lacárcel,
University of Alicante, Spain
Felipe Debasa,
Rey Juan Carlos University, Spain

*CORRESPONDENCE

Ning Li
lining@lnu.edu.cn

SPECIALTY SECTION

This article was submitted to
Educational Psychology,
a section of the journal
Frontiers in Psychology

RECEIVED 16 June 2022

ACCEPTED 20 September 2022

PUBLISHED 12 October 2022

CITATION

Yao W and Li N (2022) Construction of
artificial intelligence-assisted English
learning resource query system.
Front. Psychol. 13:970497.
doi: 10.3389/fpsyg.2022.970497

COPYRIGHT

© 2022 Yao and Li. This is an
open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which
does not comply with these terms.

Construction of artificial intelligence-assisted English learning resource query system

Wenjing Yao¹ and Ning Li^{2*}

¹Department of Foreign Languages, Jinzhong University of Shanxi, Jinzhong, China, ²College of Foreign Studies, Liaoning University, Liaoyang, China

English has become an important tool for China's opening to the outside world and exchanges with other countries. More and more people have the motivation and requirements to learn English, but under the traditional English learning mode and traditional teaching mode, the cultivation of learners' autonomous learning habits is ignored. This article aims to study the construction of artificial intelligence-assisted English learning resource query system and establish the relevant feedback mechanism of retrieval. This article applies this mechanism to the retrieval of learning resources, so as to provide learners with the learning resources they really need and improve learners' learning efficiency. This article proposes to find the relevant knowledge points by extracting the knowledge points of the retrieval content. It realizes the query expansion based on knowledge and then realizes the expansion of retrieval results. It realizes the mapping of knowledge points on the retrieval content, the query and expansion of knowledge points, and the presentation of learning resources of the knowledge point index. It also uses the relevant feedback mechanism to adjust the retrieval results to meet the retrieval needs of learners. The experimental results show that the number of knowledge points can be increased to 2–4 times by query expansion based on English resources. Thus, the number of learning resources of search results can be increased to 3–10 times, the expansion of search results can be realized, and the overall recall will be greatly improved. In this article, the related methods of artificial intelligence are applied to the construction experiment of the English learning resource query system, which has a certain promotion effect on the construction of the system.

KEYWORDS

artificial intelligence, English learning resources, query system, construction research, feedback mechanism

Introduction

At present, English learning websites are full of money, and many websites have repeated contents. It is not easy to locate English learning websites with their own characteristics that meet the special needs of learners. Then it should realize that the key to attracting learners to stay for a long time and return again is to provide truly valuable

information content with high accuracy, high credibility, and timely update. It focuses the main content of the web page on one aspect. It provides users with information references in this specific aspect. It makes it wonderful, rather than mixing all kinds of information together.

The network expands the function of computers in language teaching. As an open classroom, it enables foreign language learners to transcend the limitations of region, time, and space. It communicates with people from Britain and the United States in English. Learners can learn independently in this special classroom. Second, the network itself provides a treasure house of pure English resources. Its knowledge-based query expansion can explain and supplement the retrieved content through synonym expansion. By presenting the retrieval resources through the indexing of knowledge structure, it can mine the learning resources of the implicit index. It can reflect that there is an internal knowledge structure index between learning resources, and help learners better master the query content.

The innovations of this article are as follows: (1) this article uses the two parts of synonym expansion and knowledge index expansion of knowledge base to realize the relevant expansion of query, and integrates and confirms the two parts of expansion. (2) This article proposes a retrieval mechanism of relevant feedback on the retrieval results, which makes the retrieval results closer to the query requirements through relevant feedback.

Related work

In addition to English, a large number of multilingual content on the Internet has given birth to the impulse to develop information retrieval systems that can cross-language boundaries. Bajpai et al., discussed the development process of a complete English to Hindi cross-language information retrieval system and the contribution of each component to the system. The experimental results obtained by him confirm that adding the component “analyzer” to the CLIR architecture improves the efficiency of disambiguation algorithm (Bajpai et al., 2018). The electronic dictionary system has become very important in today's society. Yong et al., conducted research on the intelligent English electronic dictionary system. He combined the advantages of the Internet of things to design and implement an electronic dictionary system (Yong, 2021). Although the software architecture in the process of dictionary application development is the research content, the development process of the electronic dictionary is not fully discussed. The key data of the training era involve the research and improvement of English learning ability. Kaleem et al., evaluated the importance and complexity of statements used in queries, which saves the work of the automated questionnaire system, including better skill testing and formal (Kaleem et al., 2017). Although

the data direction is evaluated, there is no comprehensive test of candidates' skills. cross-language information retrieval (CLIR) system is convenient for users to query information in one language and retrieve relevant documents in another language. Thenmozhi et al., solved these problems by translating the Tamil language into English and retrieving English pages. He used the word sense disambiguation module to solve the ambiguity in Tamil queries (Thenmozhi and Chandrabose, 2018). However, the ambiguity between source query and translation query will reduce the performance of the system. Gao et al., extended and introduced KALM-QA (Gao et al., 2019) that answers more complex English problems. With the development of technology, it is easy to find information in news text. The purpose of fatmawati research is to analyze the implementation of common phrase indexing methods in information retrieval (Fatmawati et al., 2017). Although the study will be carried out in English news texts and implemented on prototypes, the degree of relevant of the generated documents is not determined. The growth of Internet-based and local applications does require that almost all Internet-based applications support global languages. Aadil and Asger has thoroughly investigated the machine transliteration models and machine learning methods (Aadil and Asger, 2017). Although machine transliteration and machine translation are used, support for local languages is not provided in all Internet-based applications.

From the relevant research on English learning resource retrieval, the construction of a knowledge ontology database, the establishment of a user interest model, situational awareness, and other means can meet the retrieval needs of specific situations to a certain extent. Building an ontology knowledge base is an effective means to realize semantic retrieval, which can realize the query expansion of retrieval content. This article is based on the construction of the knowledge base, and uses the knowledge structure in the knowledge base to realize the expansion based on knowledge, and establishes the index from learning resources to knowledge points. However, the methods related to learning resource retrieval basically analyze the needs of learners through modeling. It provides relevant learning resources. It lacks learners' judgment of retrieval results. The relevant feedback mechanism, which can be used as a means to optimize the retrieval results, has been well applied in many information retrieval systems, but it has not been well applied to learning resource retrieval.

This article uses the related methods of artificial intelligence, which is of great significance to the construction of the English learning resource query system. The innovation of this article is: using the synonym expansion of the knowledge base and the knowledge index expansion to realize the related expansion of the query, and to integrate and confirm the two parts of the expansion. And a retrieval mechanism is proposed to provide

relevant feedback to the retrieval results, and the retrieval results can be closer to the query requirements through relevant feedback.

Customer requirements and methods of English learning resource query system

User on English learning resources website

The user base has been fully popularized, and the user's mentality has also changed. Due to the wide range of activities and many opportunities, people are unprecedentedly busy, time is precious, and there are more and more things to pay attention to. Too much information input is too late to digest, and the noisy life makes people more and more impatient (Berg and Aronoff, 2021). People are eager to organize complex information orderly and turn complex problems into simple ones. The idea of information construction can just guide how to display the content and characteristics of information in a clear structure to meet the needs of users. People exchange information and engage in various activities in virtual cyberspace, and the biggest consumption is time (Ramchandani et al., 2021).

Combined with the characteristics of English learning, this article divides users into non-professional and professional English learners. Most of the existing English Learning Websites target non-professional users. Because of its rich content and various forms, it can better meet the needs of users. At the basic level, as a supplement to literature-based learning materials, it provides users with richer resources to improve their basic English skills. But at the same time, literature resources, classrooms, and training courses also have great advantages, such as being more systematic, with special teachers to provide on-site guidance and supervision.

For deep users such as professional English learners, network resources have a more irreplaceable position. Their needs are often more specialized. In the learning process, they need to grasp the knowledge as a whole, from a single knowledge point to an understanding of relevant knowledge points (Islamova, 2019). For example, when learning an English article, students should not only master the vocabulary and grammar of the text itself, but also understand the relevant background knowledge, comments, and brief introduction of the author, and even extend to the study of other articles with the same subject and author. At this level, the network has incomparable advantages. Users can retrieve the required information with the click of the mouse. However, the huge information resources also make users face difficulties in choosing and obtaining. Therefore, it is necessary to build a good English learning website based on information.

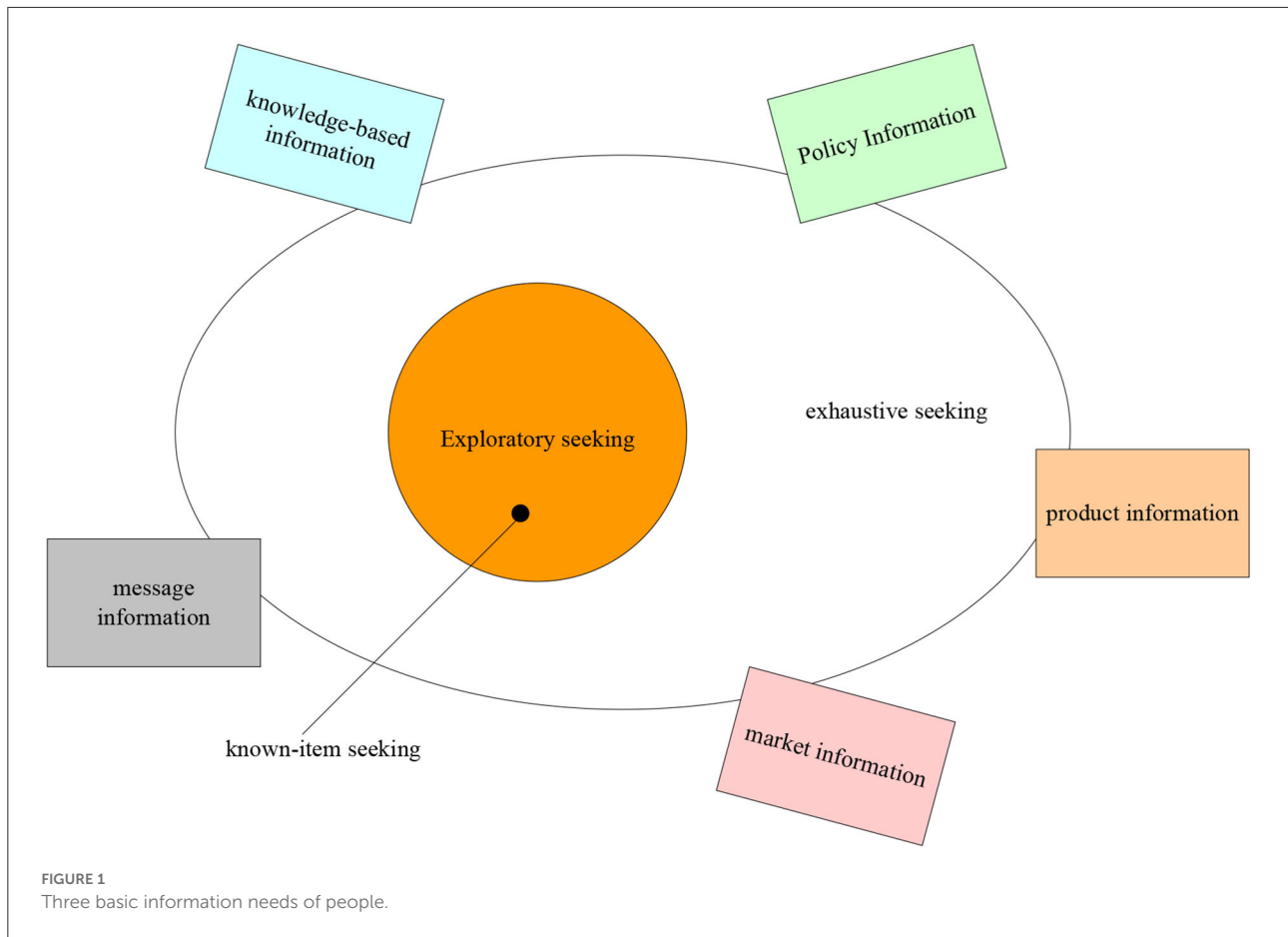
User information demand and information behavior

Although English learning website users have different ages, occupations, interests, levels, and knowledge backgrounds, their information needs for learning materials are also different. Because there are great differences in the age, interest, knowledge structure, and professional background knowledge of English learning website users, they also show some differences in the content of information needs from the level and personality. It can be divided into two levels: high level and general level (Chen and Huang, 2021; Hong and Nam, 2021). In terms of content, high-level information needs are mostly innovative, verifiable, pioneering, and practical information. It can not only reflect the professional development characteristics of disciplines but also reflect the cross penetration, comprehensive and integrated development trend of disciplines. This kind of user requires strong academic and professional information. At the same time, it is also more standardized, such as users engaged in linguistics, translation, and literary research. They need more professional information and higher academic authority. People engaged in it need computer English, those engaged in law need legal English, and those engaged in trade need business English. Users with general level information needs are those who want to practice the basic skills of English listening, speaking, reading, writing, and translation. What they need is some practical, novel, and effective learning materials (Yong, 2021). The three basic information needs of people are shown in Figure 1.

As shown in Figure 1, corresponding to the above three information needs, users' information query behavior also has three basic forms: searching, such as inputting retrieval formula in the search engine; browsing, such as viewing information from one link to another by clicking the mouse; asking is asking others for help on the Internet, such as asking others questions through email, chat system, etc. Users' information needs are closely related to information behavior. Different information needs partially determine different information query behaviors.

Neural network algorithm

Due to the uneven amount of questions students do, some students do more and some students do less. There are also differences in the focus of teaching resources and teaching requirements between colleges and universities. Moreover, when English teaching resources diagnose students, they often think intuitively. Under the characteristics of English resources, in order to make the system diagnosis more accurate and personalized, this module adopts the combination of the neural network and expert system. It makes intuitive thinking and logical thinking complementary. It transforms the experience



and knowledge of English experts into a nonlinear problem. It enables the diagnosis process to realize the nonlinear mapping of input / output. The final diagnosis result is obtained by the BP neural network.

Activation function

The judgment of students' mastery of knowledge points is actually the fault diagnosis of students' knowledge points. In the field of fault diagnosis, the forward multilayer network is the most used and effective. In neural networks, even if a linear function can be selected theoretically, such as the identity function $f(x) = x$, the nonlinear sigmoid function is usually selected (Dombrovan et al., 2021).

$$f(x) = \text{simg}(x) = \frac{1}{1 + e^{(-x)}} \tag{1}$$

The nonlinear sigmoid function and function gradient are shown in Figure 2.

As shown in Figure 2, the function value of sigmoid is between 0 and 1. The closer the function value of the independent variable is to 0, the faster the function value

changes. The larger the absolute value of the function changes, the slower the absolute value of the function changes (Nie, 2021). In addition to the nonlinear sigmoid function, the second commonly used is the hyperbolic tangent function tahn:

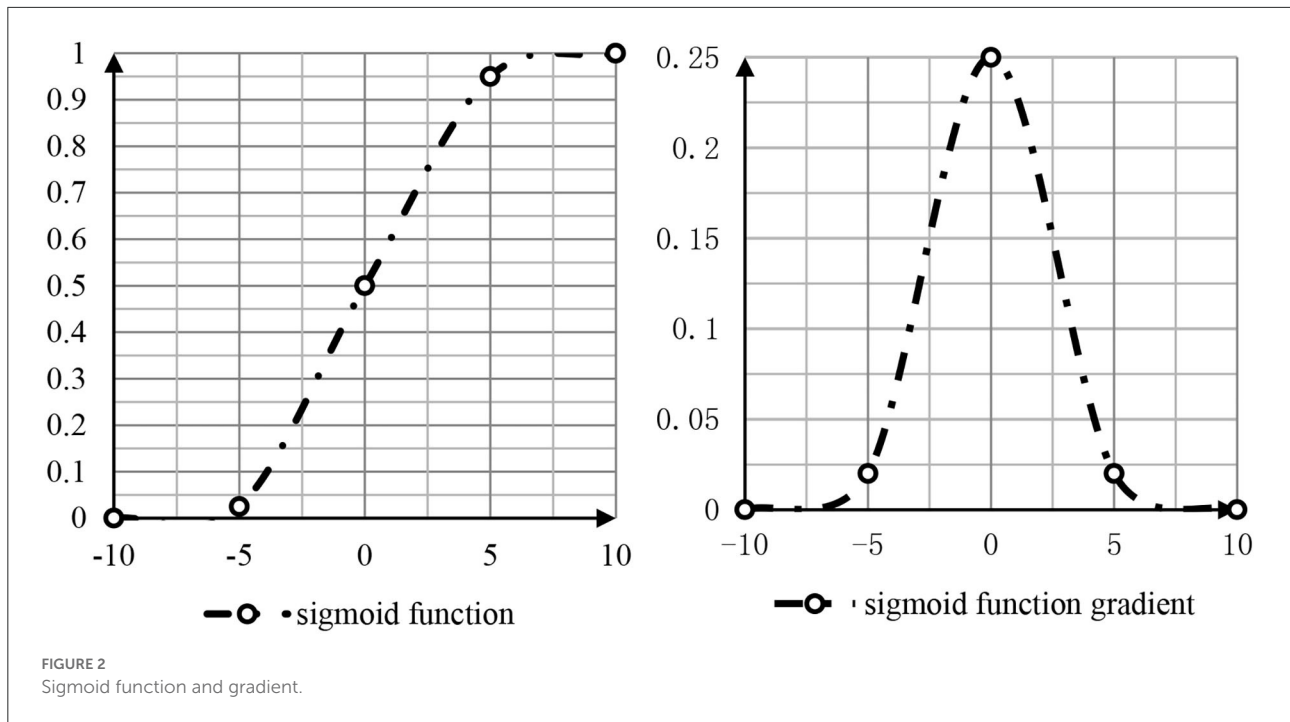
$$\text{tahn}(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{2}$$

In addition, the hard limiting function plus the skewness function can also be used to activate the function:

$$\text{hardlimit}(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x \leq 0 \end{cases} \tag{3}$$

$$\text{ramp}(x) = \begin{cases} 1, & x \geq 1 \\ x, & -1 \leq x \leq 1 \\ -1, & x \leq -1 \end{cases} \tag{4}$$

At the same time, there are two good choices, namely, the positive linear element ReLu and the leakage correction linear element.



$$\text{ReLU}(x) = \max(0, x) \tag{5}$$

$$\text{LReLU}(x) = \begin{cases} 1, & x \geq 0 \\ ax, & x \leq 0 \end{cases} \tag{6}$$

The function and gradient of ReLU are shown in Figure 3.

As shown in Figure 3, in the classical structure, the activation function using the neural network is called rectified function. In order to alleviate the “dead zone” phenomenon, the researchers adjusted the part of $x < 0$ in the ReLU function to $\psi(x) = \delta * x$, where a is a smaller positive number of the order of 0.01 or 0.001 (Narayanan and Mathew, 2021). This new activation function is called leaky ReLU:

$$\text{Leaky ReLU}(x) = \begin{cases} x(x \geq 0) \\ \delta \cdot x(x \leq 0) \end{cases} \tag{7}$$

For the setting of δ in randomized ReLU, its value is uniformly distributed in the training phase, and it is designated as the distribution expectation $(1+u) / 2$ of the uniform distribution in the testing phase.

$$\text{Randomized ReLU}(x) = \begin{cases} x(x \geq 0) \\ \delta \cdot x(x \leq 0) \end{cases} \tag{8}$$

Among them

$$\delta' \approx U(l, u), l \leq u, \text{ and } l, u \in [0, 1] \tag{9}$$

However, the exponential operation slightly increases the amount of calculation:

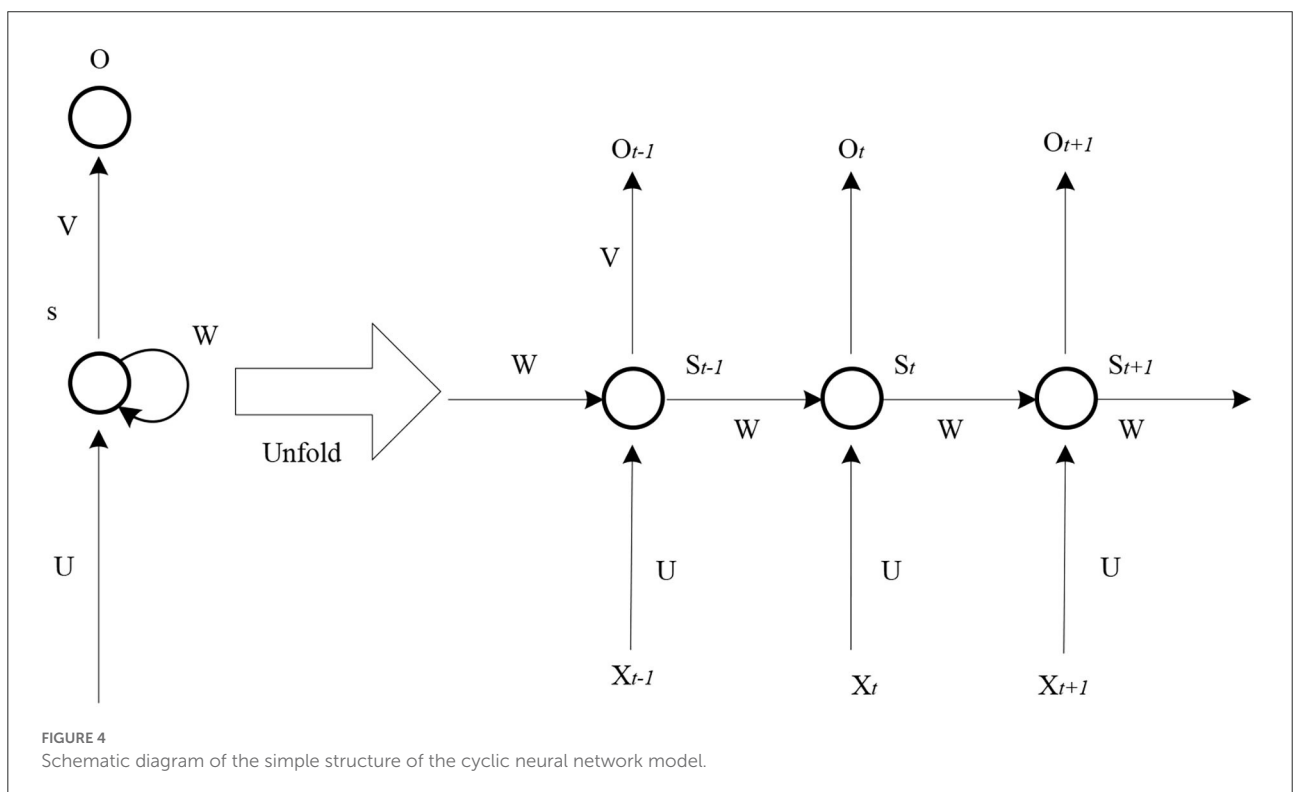
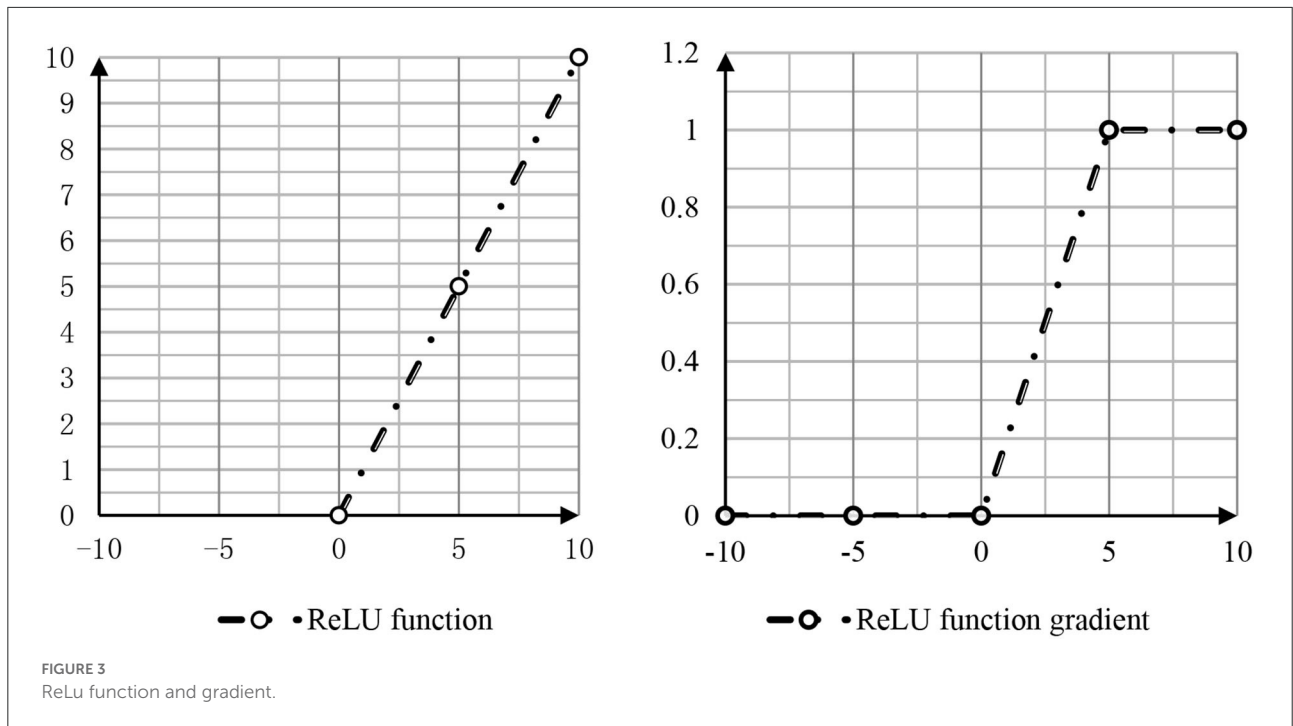
$$\text{ELU}(x) = \begin{cases} x & x \geq 0 \\ \beta \cdot (\exp(x) - 1) & x < 0 \end{cases} \tag{10}$$

Cyclic neural network

When the research field is switched to natural language processing, it is necessary to use the cyclic neural network model for modeling (Bolu et al., 2020). In the application of language, there is a great connection between the preceding and the following. When a word in a sentence is removed, people usually choose the appropriate word to fill in the air in combination with the experience of the previous and subsequent text (Fitria, 2020). The structure of the cyclic neural network model is simple, and the schematic diagram is shown in Figure 4.

As shown in Figure 4, it is difficult to solve such problems in the traditional linguistic model. How to make the computer understand “language” is a difficult thing. Therefore, it proposes a recurrent neural network (RNN). Its main feature is that the network model has memory. In theory, RNN can trace any number of words forward.

The information of RNN at a certain time can contain the information of the previous time. It can find the sequence



information of the previous and subsequent text and the semantic information of the text. Ordinary neural networks can only deal with short sequences of fixed length, and there is no relationship between the information of the previous window

and the information of the next window in the process of moving the window.

At time t , the network has an input x , and the neuron state of the network S_t at this time can be expressed by formula (11):

$$S_t = f(ux_t + ws_{t-1}) \tag{11}$$

That is, softmax processing is conducive to error backpropagation. The output O_t expression of the network at time t is (12):

$$O_t = \text{soft max}(vs_t) \tag{12}$$

RNN model is mainly used to solve time series problems, so the loss of all times must be included in the calculation of model loss (Yuan et al., 2020). The loss function of the model at time t is shown in formula (13):

$$\text{loss}_t = -[y_t \ln(o_t) + (y_t - 1) \ln(1 - o_t)] \tag{13}$$

y_t is the real label value input at time t , and o_t is the predicted value of the model. RNN is prone to gradient disappearance and explosion. In the process of model training, when the gradient is large, the function will directly skip the optimal solution, making the training unable to converge (Meng, 2020).

Based on the attention pooling BiGRU model

Attention pooling is introduced. The main feature of attention pooling is to construct an attention pooling matrix M for text matrix A and text matrix B processed by the neural network (Do et al., 2021; Flint et al., 2022). Matrix M generates the maximum pool vector of a row and the maximum pool vector of a column, respectively. Maximum pooling preserves the most significant features of the vector and reduces the computational complexity of the model (Liang, 2022; Pantanowitz, 2022). The model structure is shown in Figure 5.

As shown in Figure 5, the two vectors are normalized and multiplied by A and B to obtain the features finally integrated into the attention pool (Saura et al., 2022a). The matrix M can be obtained by formula (14).

$$M = \tanh(A^T UB) \tag{14}$$

u is the parameter matrix obtained by neural network learning, which can be adjusted and optimized in the process of model training. M actually multiplies the text pair matrix indirectly through the parameter matrix. The converted attention vector is obtained (Liu, 2021). Taking the attention vector δ^a as an example, the i th element of δ^a is expressed as formula (15):

$$[\delta^a]_i = \frac{e^{[m^a]_i}}{\sum_{1 < l < t} e^{[m^a]_l}} \tag{15}$$

t is the text length. Multiply the original matrix by the normalized attention weight vector to obtain the final text pair representation vector:

$$r^a = A \delta^a \tag{16}$$

$$r^b = B \delta^b \tag{17}$$

Finally, the similarity between text pairs is calculated by cosine, which can be expressed as:

$$\cos(\theta) = \frac{\langle r^a, r^b \rangle}{\|r^a\| \|r^b\|} \tag{18}$$

In the model based on attention pooling, its input is the word vector matrix obtained by two text pairs (Saura et al., 2022b). Because the context of the text is related, we can't ignore the backward semantics and only focus on the meaning of a single word.

Distance measurement method

When it is necessary to measure the similarity between samples, the usual method is to calculate the distance between text vectors (Abdi and Valentin, 2007; Saura et al., 2021a). There are many methods to calculate the distance. Choosing different methods will affect the accuracy of the judgment of the result.

Euclidean distance

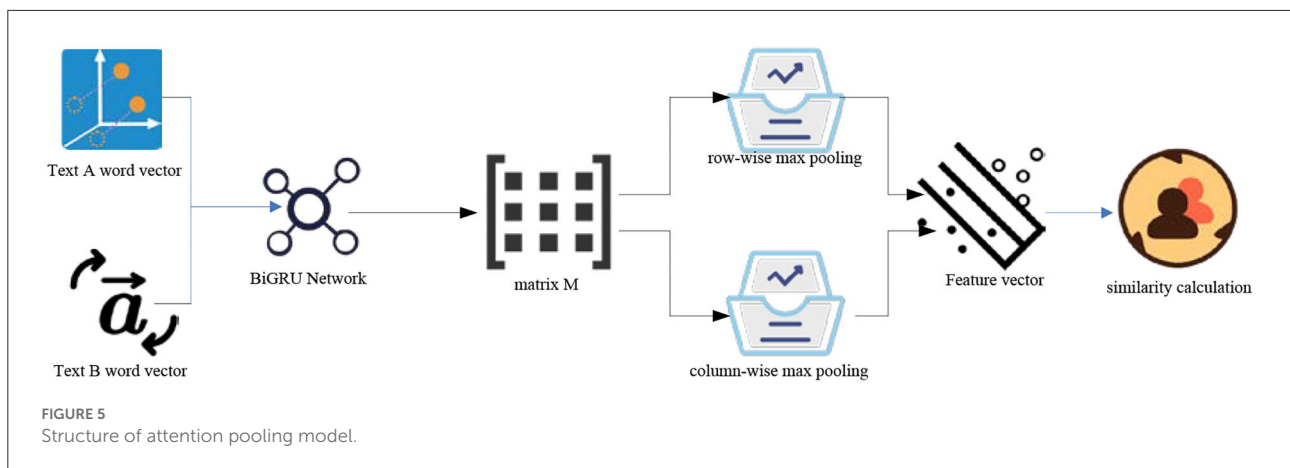
Euclidean distance uses the most intuitive method to calculate the distance, that is, to calculate the real distance between two points in space, as shown in formula (19):

$$D = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \tag{19}$$

Manhattan distance

It represents the sum of the distances between two points in the direction of the coordinate axis (Saura et al., 2021b). The calculation is as follows (20):

$$D = \sum_{i=1}^n |x_i - y_i| \tag{20}$$



Hamming distance

It represents the number of different data at the same position of the sequence. The calculation formula is as follows (21):

$$D = \sum_{i=1}^n x_i \oplus y_i \tag{21}$$

Construction of artificial intelligence-assisted English learning resource query system

Construction principles of English query system

User-oriented principle

To build an efficient English learning website, we must first consider the problems from the perspective of users. Website designers must always realize that although the construction is aimed at the construction of information content such as English learning resources, it is carried out for users. Therefore, we should organize, manage and provide information according to users' views and behavior habits. The behavior habits of English learning resource users in the network environment are different from those in the traditional environment. For example, there are obvious differences between reading learning materials on the website and reading paper documents. When reading paper materials, users' reading habits are linear reading word by word and sentence by sentence, reading one text at a time without interruption, paying equal attention to each word, and inputting other information such as music, radio, television, and conversation at the same time. In the network environment, different users have many different reading methods for the same material text. Reading is not necessarily continuous reading from left to right or from front to back. Users usually skip reading. As long as people click the mouse, different readers

will skim this part and not that part. And reading is accompanied by sound, image, and other information. In a relaxed reading environment, readers unknowingly enter new websites or texts from different parts. In addition, due to the different expression forms of information on the website, unlike the traditional black and white text, the expression forms of each word from beginning to end are not very different. So they get different attention from users.

Content first principle

After the Internet entered the field of English education, great changes have taken place in the way English teachers and learners teach and learn. Once users encounter any questions and need any information, they will first think of an online query. Many times, users visit the site with a certain purpose and specific content. The information needed by English learning resources website users is usually comprehensive, centralized, stable, fast and accurate, knowledge unit, hierarchical and personalized. With a large number of English resources and the need for resource content, there is a need for information construction. The object of information construction of English learning resources is English learning resources. Therefore, resource content is the leading and core of information construction, and everything serves the content. In addition, many English learning resources, learners' valuable time, and reading on the computer screen is easy to cause visual fatigue and other reasons, so users usually browse the content on the web instead of reading carefully. Therefore, this article proposes a knowledge-based query expansion and relevant feedback retrieval method. The overall research route of knowledge-based query expansion and relevant feedback retrieval is shown in Figure 6.

As shown in Figure 6, in the field of education, the index between knowledge points cannot be reflected at the simple text level. It makes it difficult for the retrieval system to expand the retrieval of knowledge points and keywords from the semantic

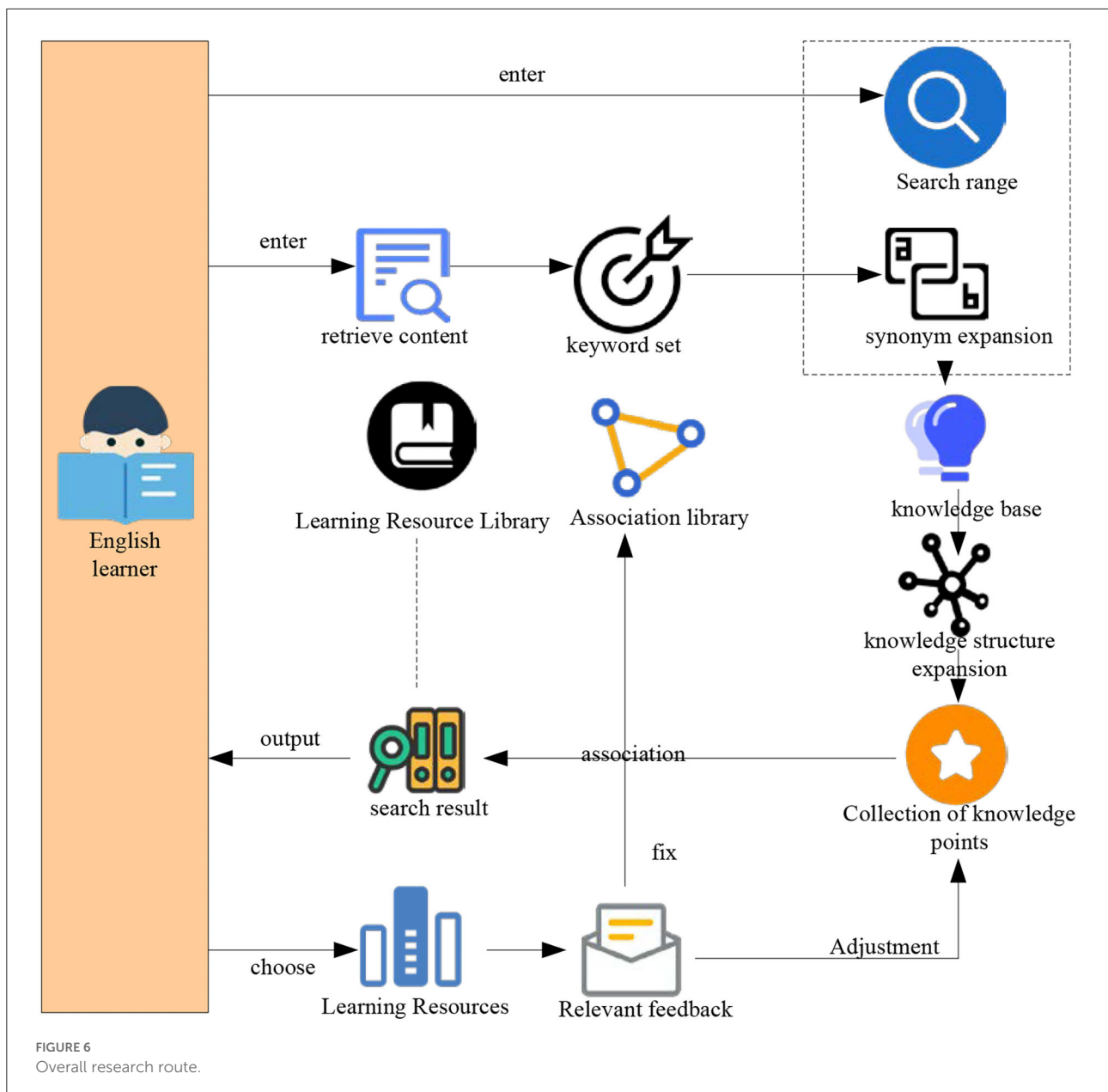


FIGURE 6 Overall research route.

level. Second, at present, it uses learners' relevant feedback information to adjust the query, which makes the query closer to learners' query intention. This is also one of the problems that the retrieval system is difficult to solve.

Implementation of English learning resource retrieval

In the retrieval of learning resources, the traditional keyword-based retrieval method has two defects: First, the retrieval results often only meet the requirements of learners

literally. Actual needs often deviate from learners. The second is to take the learning resource database as the retrieval object, which needs to match each learning resource, resulting in low retrieval efficiency. The disadvantages of limited depth and low efficiency of learning resource information retrieval systems hinder the utilization and sharing of learning resources.

To realize knowledge-based learning resource retrieval, we first need to prepare a corpus. It uses natural language processing technology and related algorithms to build the index of learning resources to knowledge points. It indexes knowledge points and learning resources. Then it determines the knowledge system through the query input, calculates the similarity between the

retrieval content and the knowledge points, and extracts the knowledge points. It takes the extracted knowledge points as query-related knowledge points. It further presents the learning resources indexed by knowledge points to learners in the reverse order.

Index construction based on LSA

The purpose of constructing a learning resource index is to extract useful information from unstructured learning resources. According to the concept and semantics of knowledge points in the knowledge base, the mapping relationship between learning resources and knowledge points is established. Thus, it can use the structured knowledge in the knowledge base to expand the semantics of the retrieval keywords and carry out the subsequent retrieval work based on this.

The mapping relationship between learning resources and knowledge points needs to use natural language processing-related technologies. It mainly includes word segmentation, removing stop words, and so on. Word segmentation refers to dividing a paragraph or sentence into several words or words according to syntactic logic and common word combination habits: removing stop words refers to removing commonly used cohesive words, meaningless words, or specific words that need to be filtered according to the stop word list after word segmentation. In the selection of natural language processing technology, this section uses Jieba word segmentation to extract keywords. Jieba word segmentation provides a variety of methods to segment text to varying degrees, and it can also obtain the weight of keywords to construct the vector model. Some fields describing information in the knowledge base and learning retrieval base need to be filtered, so it is necessary to establish a stop vocabulary based on the knowledge base and learning resource base. This can effectively improve retrieval efficiency.

The vector space model is used in the model. The vector space model considers the weight and can describe the correlation between knowledge points and learning resources. The construction process of the knowledge-based learning resource index is shown in [Figure 7](#).

As shown in [Figure 7](#), first, it takes the description information of all knowledge points in the knowledge base as the training corpus. The TF-IDF method is used to calculate the keyword weight of each knowledge point description information to form a knowledge point vector. It also uses LSA to decompose the knowledge base matrix by SVD. For each resource in the learning resource database, word segmentation and remove the stop words in the learning resource description information, extract keywords and transform them into learning resource vectors. It uses the cosine similarity calculation method to calculate the similarity between each resource and all knowledge point vectors. It selects the three knowledge points

with the greatest similarity as the knowledge to be indexed. In an implementation, this section uses the Gensim Library in Python language to realize the indexing of learning resources to knowledge points. The good scientific computing power of Python language has been widely used in artificial intelligence and other related fields. Gensim is a natural language processing library that can be called directly by python. It supports a variety of model algorithms including TF-IDF and LSA.

Query knowledge point extraction

Generally, the retrieval system directly matches the retrieval content with the resource database, mainly because the resources in the resource database are not systematic. It establishes a knowledge base to express the same learning resources and establishes an index from learning resources to knowledge points. Then we can find the learning resources that learners want according to the knowledge points in the knowledge base. Therefore, in the retrieval process, only the knowledge points of the retrieval content need to be extracted. After obtaining the relevant knowledge points, we can obtain the learning resources related to the retrieval content.

In order to realize the knowledge extraction of retrieval content, this article makes the knowledge point extraction results of retrieval content meet the needs of learners through two parts. First, this article determines the knowledge system that learners need to retrieve. All knowledge points in the knowledge system are the set of knowledge points to be selected for knowledge extraction. Then, through the method of natural language processing, it uses the vector space model to calculate the similarity between the retrieval content and each knowledge point in the set of knowledge points to be selected. It arranges the similarity between the knowledge points and the retrieval content in the descending order and takes the knowledge points of topn as the knowledge extraction result.

The accuracy of the retrieved relevant knowledge points depends entirely on the query. The actual situation is that different learners describe the same word differently. Often, documents related to query requests cannot be retrieved due to different words. In addition to retrieving relevant knowledge points, learners are often interested in the relevant knowledge points in the knowledge structure. It needs to take other knowledge points related to this knowledge point in the knowledge structure as relevant knowledge points. In order to solve these problems, this section proposes a knowledge-based query expansion method. The knowledge-based query expansion is divided into two parts: one is to expand the query by using synonyms to supplement and explain the query. One part extends the query based on the knowledge structure and provides knowledge-related extension based on the knowledge structure. After implementing the two-part extension, it needs to fuse the two-part extension

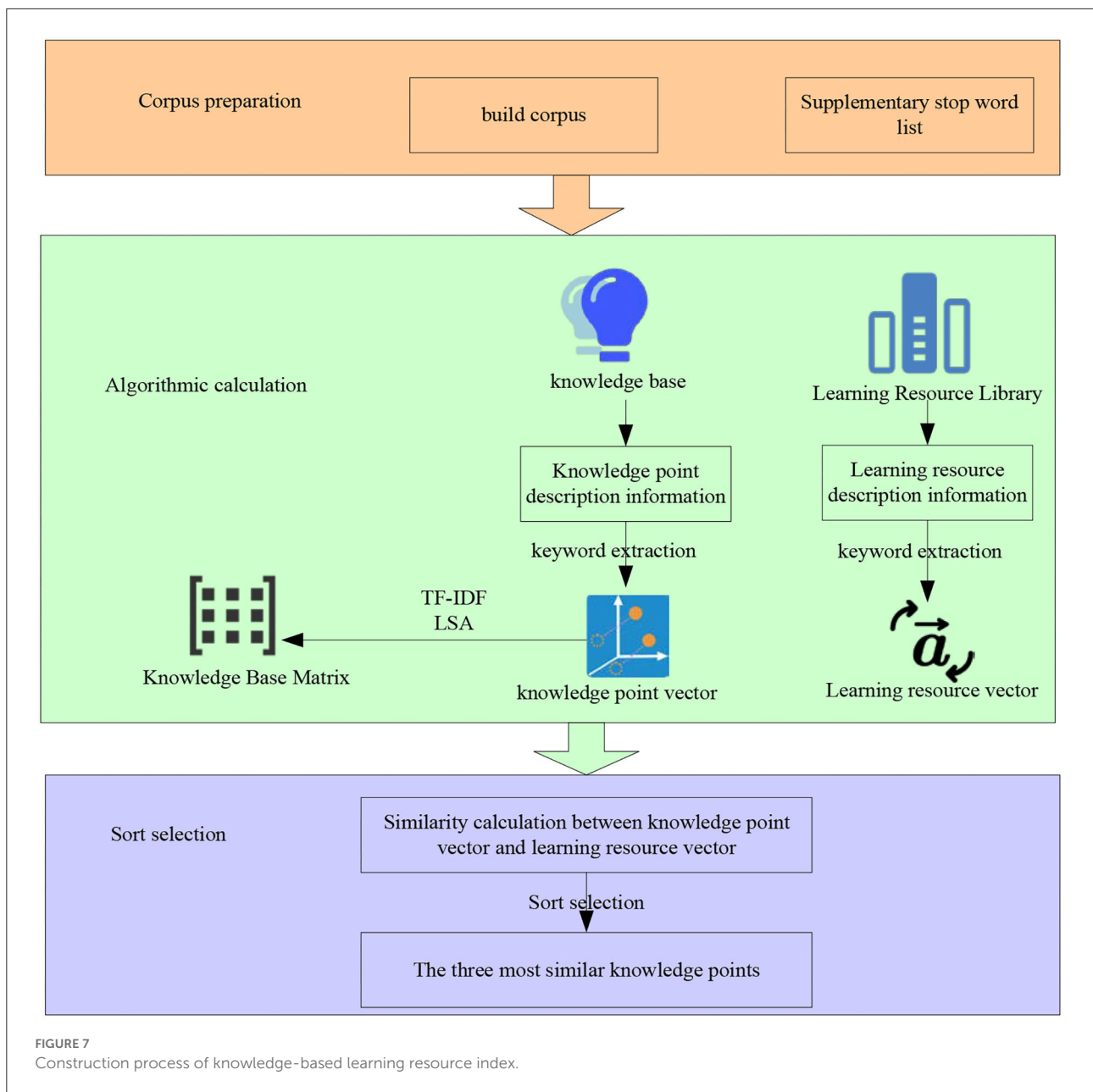


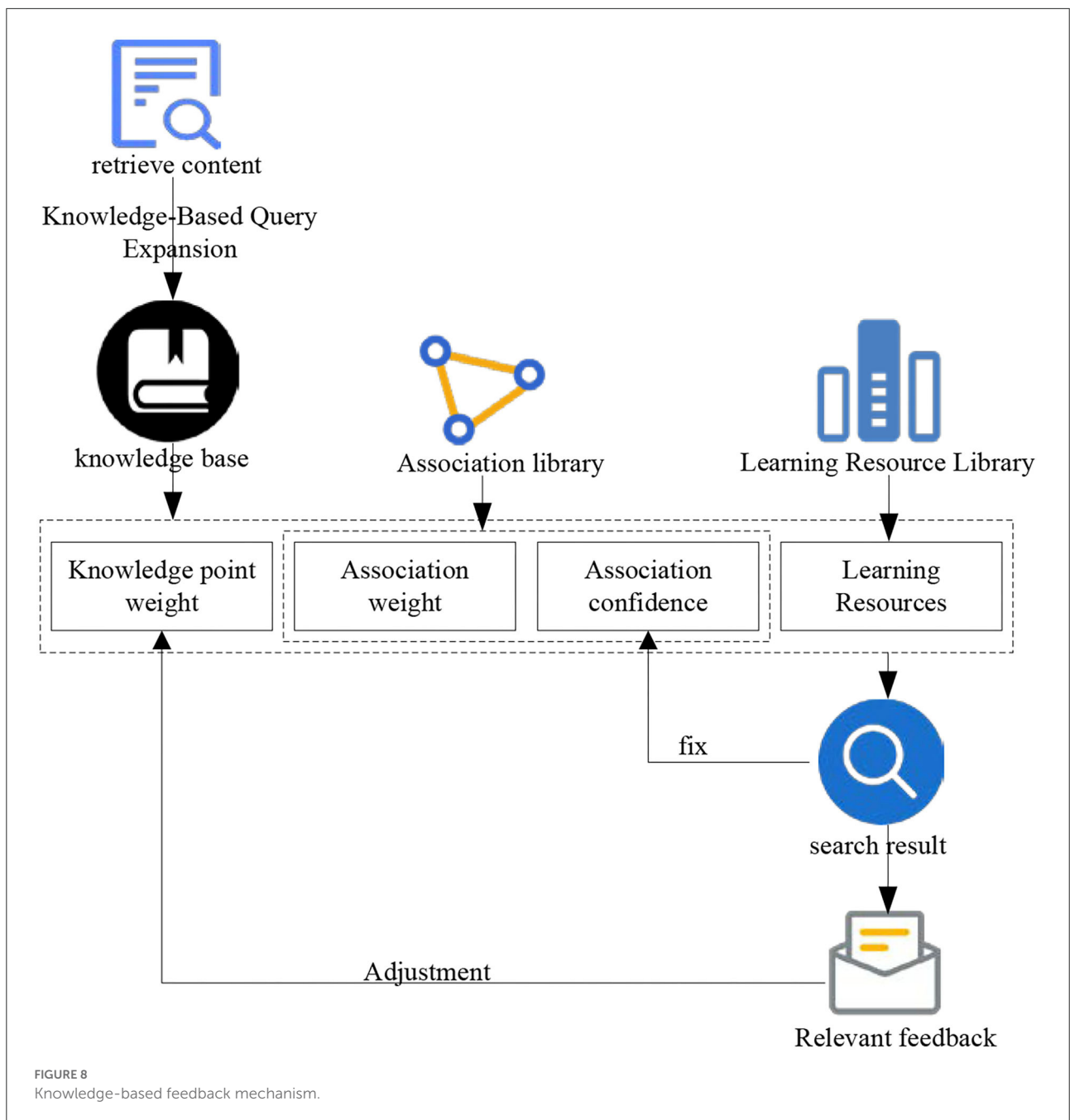
FIGURE 7 Construction process of knowledge-based learning resource index.

with different weights, and finally provide the query extension results to learners. Learners confirm the knowledge-based query expansion and present relevant learning resources according to the learners' confirmed relevant knowledge points.

Relevance feedback mechanism based on the knowledge

The essence of the knowledge-based relevant feedback mechanism is to use the user's relevant feedback to optimize the retrieval. The difference is that knowledge-based relevant

feedback is not directly used to modify the query, but to adjust the query-related knowledge points. Its working principle is that learners submit queries representing their query needs to the retrieval system. The retrieval system makes an initial query and returns the query-related knowledge points. It returns the collection of learning resources through the index of knowledge points and learning resources. It is usually arranged in the descending order according to the correlation between learning resources and queries, and then learners judge the correlation between learning resources and queries and point out which are relevant or irrelevant. The process of relevant feedback can continue until learners get more satisfactory learning resources.



The relevant feedback mechanism based on knowledge is shown in [Figure 8](#).

As shown in [Figure 8](#), the learner judges whether each learning resource retrieved is relevant or irrelevant. Each relevant judgment of learners will generate a new record in the retrieval feedback database. The record is composed of the unique identification of knowledge points, the unique identification of learning resources, the unique identification of users and whether it is relevant or not. Statistics from different dimensions can produce different feedback effects. The

use of these statistical information can help better retrieval of knowledge-based learning resources. In this section, knowledge-based relevant feedback retrieval is mainly reflected in two parts: one part is to count the relevant feedback information from different learners to the learning resource index of the same knowledge point. It proposes the concept of index confidence, which modifies the index based on natural language processing technology. The other part counts the relevant and irrelevant learning resources that the learner thinks through the feedback of a single learner. It adjusts the weight of knowledge

TABLE 1 Learner information.

Serial number	Name	Describe	Type
1	user_id	learner id	int[11]
2	username	learner name	varchar[10]
3	password	current password	varchar[20]
4	field	field of study	enum
5	grade	grade of study	enum
6	subject	Subject	enum

TABLE 2 Knowledge points.

Serial number	Name	Describe	Type
1	knowid	knowledge point id	int[11]
2	knowledge_title	Knowledge point name	varchar[50]
3	pre_knowid	parent node id	int[11]
4	description	Description	varchar[200]
5	keywords	keywords	varchar[50]
6	field	field of study	enum
7	grade	grade of study	enum
8	subject	Subject	enum

points related to queries under relevant learning resources and irrelevant learning resources, so as to adjust the retrieval results.

Database design

The MySQL database is used in database selection. It designs a learner information table for storing learner-related information, a knowledge point table for storing knowledge point information, a learning resource table for storing learning resource information, an index table for storing index information of knowledge points and learning resources, and a retrieval feedback table for storing retrieval feedback information.

In the design of table, it strictly follows the database design specification to ensure the consistency and integrity of data. The specific data sheet is designed as follows:

(1) The learner information table is used to store the basic information of learners, mainly including: learner name, level, and login password. The specific structure is shown in Table 1.

(2) The knowledge point table is used to store the basic information and knowledge structure of the knowledge point, mainly including: the name of the knowledge point, the description information of the knowledge point, the grade discipline of the knowledge point, and the parent knowledge

TABLE 3 Learning resources.

Serial number	Name	Describe	Type
1	r_id	learning resource id	int[11]
2	URL	Url address	varchar[20]
3	res_type	Resource Type	enum
4	Resource Name	Resource Name	varchar[50]
5	describe	Description	varchar[200]
6	res_difficulty	degree of difficulty	enum
7	site	field of study	enum
8	grade	grade of study	enum
9	Subject	Subject	enum

TABLE 4 Search feedback.

Serial number	Name	Describe	Type
1	id	record id	int[11]
2	knowid	knowledge point id	int[11]
3	r_id	resource id	int [11]
4	user_id	learner id	decimal
5	content	retrieve content	varchar[50]
6	relation	relevant or irrelevant	enum
7	update_time	update time	datat ime

node of the knowledge point. The specific structure is shown in Table 2.

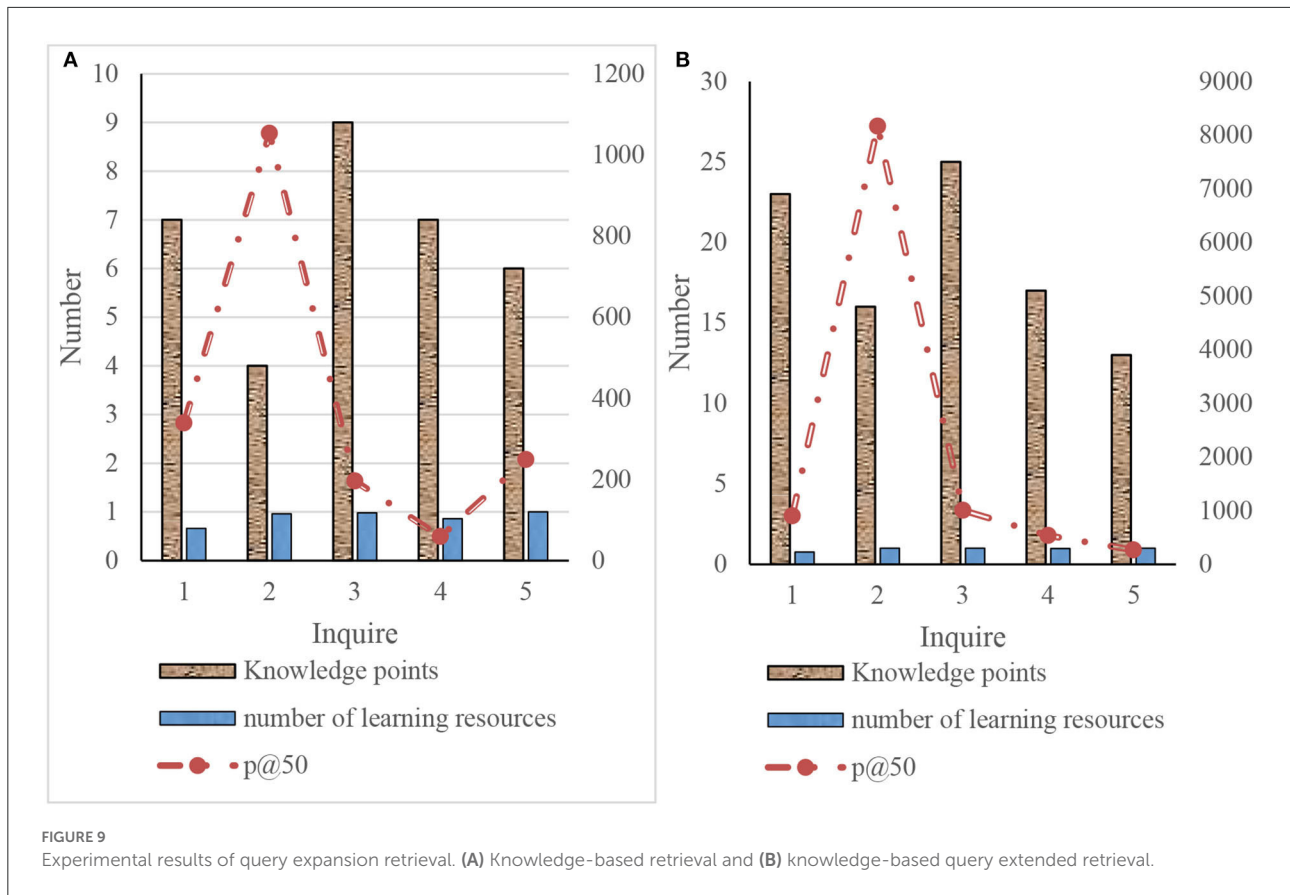
(3) Learning resource table is used to store basic information such as learning resource name, learning resource type, description information, grade, and discipline. The specific structure is shown in Table 3.

(4) The retrieval feedback table is used to store the feedback information of learners on the retrieval results. The specific structure is shown in Table 4.

Experimental on the construction of English learning resource query system

Knowledge-based query expansion experiment

Because the retrieval implementation of this article is based on a large-scale corpus, the recall is difficult to measure. The recall can only be expressed indirectly by the number of returned knowledge points and learning resources. In the measurement of precision, it is used in large-scale corpus P@50. It calculates the accuracy of the first 50 learning resources returned. This article uses the five queries established to count the number of returned knowledge points and learning resources and calculate



the accuracy of the first 50 learning resources. The specific results are shown in Figure 9.

As shown in Figure 9, through the comparison of the retrieval results of knowledge-based retrieval and knowledge-based query expansion, it can be concluded that:

(1) From the perspective of a full search, through knowledge-based query expansion, it can increase the number of knowledge points to 2–4 times the original. Thus, the number of learning resources of search results can be increased to 3–10 times, the expansion of search results can be realized, and the overall recall will be greatly improved. In the fifth query, when the number of knowledge points is doubled, the number of returned resources does not change greatly. The possible reason is that the number of resources of the extended knowledge point index is very small. The knowledge point to the learning resource index constructed by semantic analysis has a small number of learning resources.

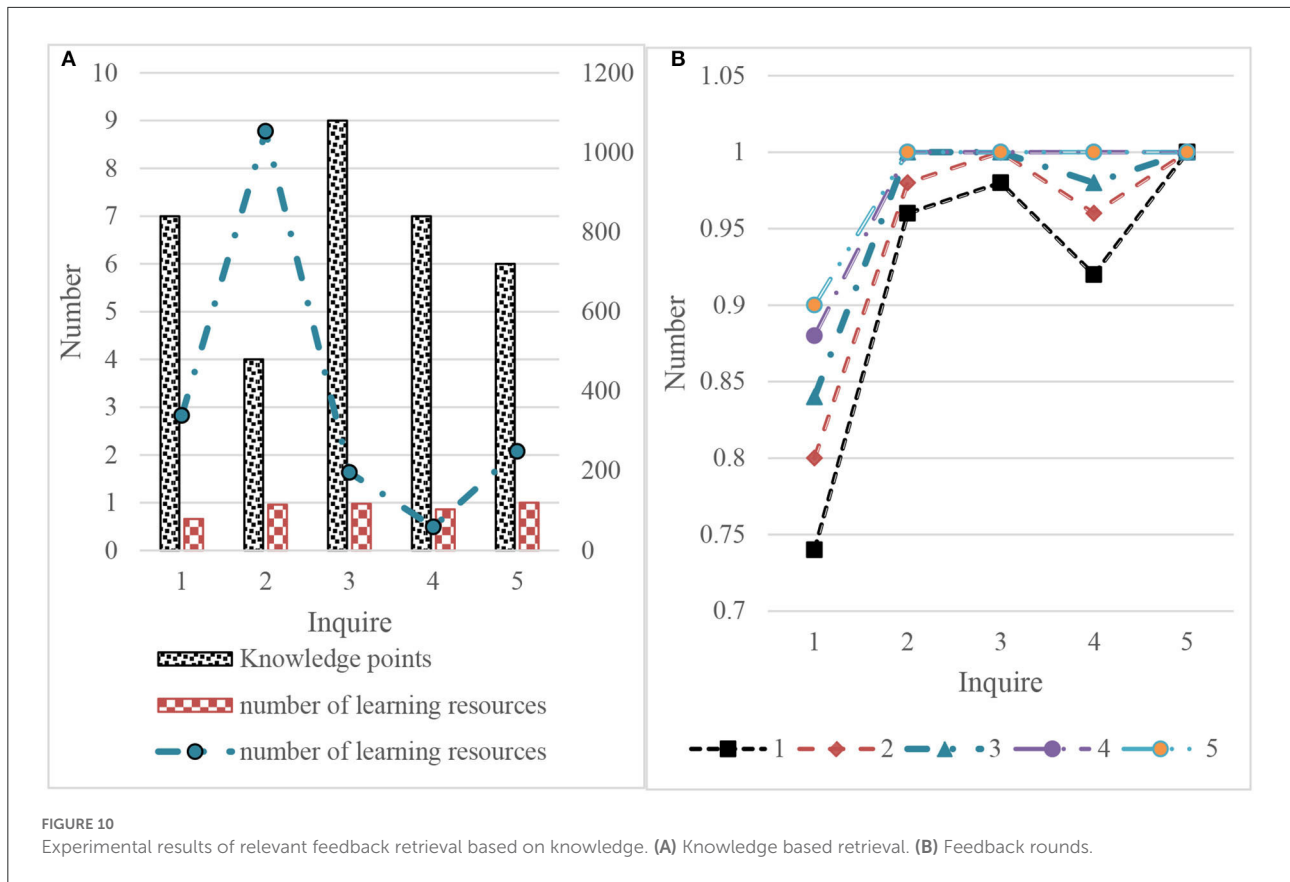
(2) From the perspective of accuracy, P@50 has been improved through knowledge-based expansion. The lower the P@50 of knowledge-based learning resources is, the higher the improvement of the P@50 of the retrieved learning resources after knowledge expansion. However, the final P@50 still largely

depends on knowledge-based learning resource retrieval. For the fifth query, the P@50 of both methods is 1. It shows that when constructing the index of learning resources to knowledge points through semantic analysis, the accuracy is high, but the number is not enough. This also confirms why the number of learning resources has not changed greatly after the expansion of knowledge points.

Relevance feedback experiment based on English resources

Experimental purpose

To verify the feasibility of the knowledge-based relevant feedback method, it should also be divided into two parts: using the relevant feedback mechanism to adjust the query and maintain the index. Index maintenance requires a large number of learners to participate in feedback in order to get more accurate results. Due to the limitations of conditions, the experiment cannot be carried out. Therefore, when the default index confidence is 0.5, the experiment using the relevant feedback mechanism to adjust the query is carried out.



Experimental data

In the experiment of knowledge-based query expansion, the relevant feedback is given to the retrieval results. It also judges the correlation between the first 50 learning resources in the search results and the query and inputs the same query again on this basis. It carries out multiple rounds of feedback and compares the changes in P@50 under different rounds of feedback. In order to avoid the mutual influence of retrieval feedback from two different methods, experiments need to be carried out with different users and independent environments.

Experimental results

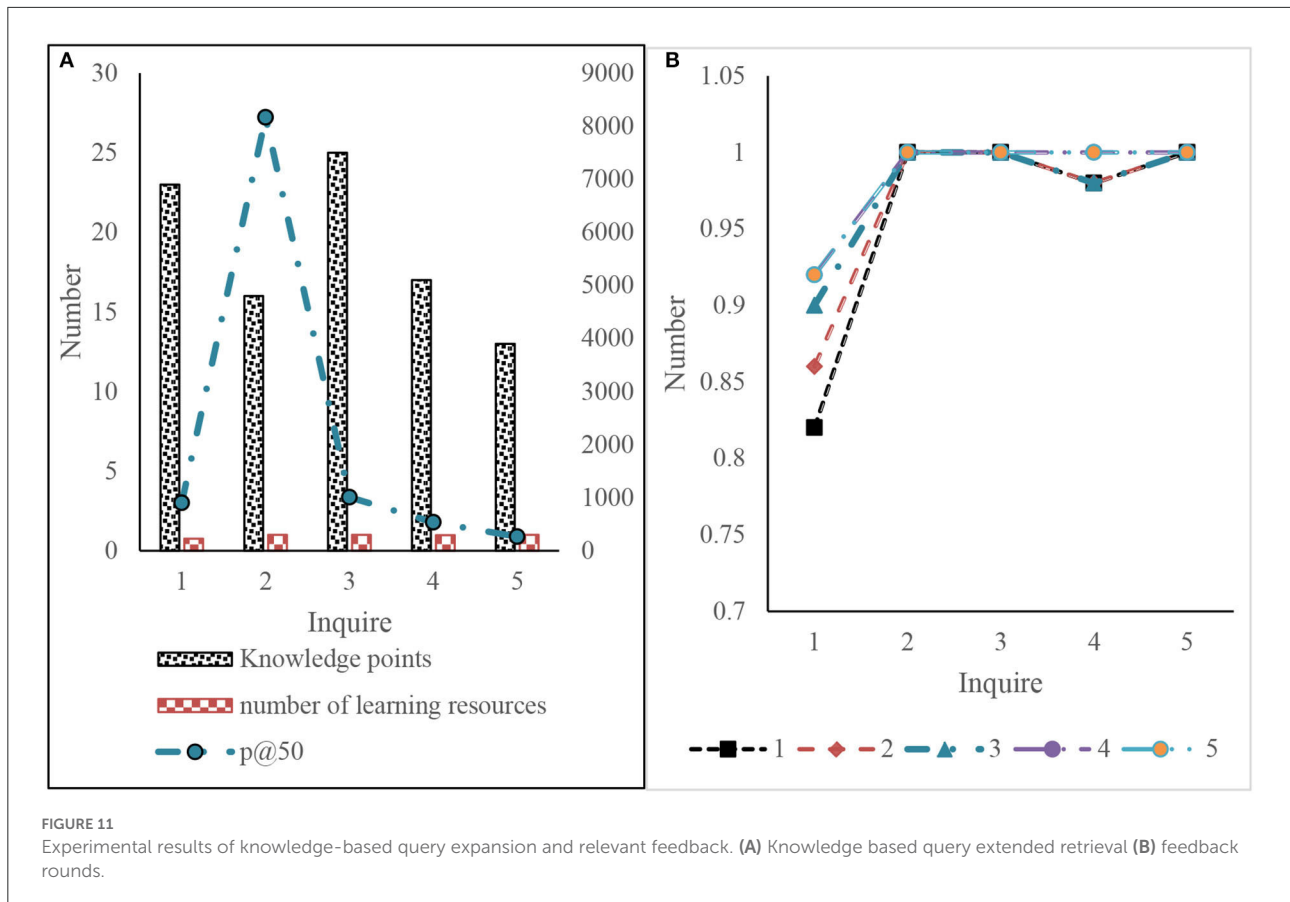
In the experiment of knowledge-based query expansion, this article adopts the same query. It carries out the relevant feedback on the retrieval results of knowledge-based retrieval and knowledge-based query expansion retrieval, respectively. In order to measure the effectiveness of relevant feedback, it carries out five rounds of feedback to see the changes in P@50 in each round. The experimental results are shown in Figure 10.

The following results can be obtained according to the initial feedback in Figure 11, which will not affect P@50 as shown in Figure 10. Therefore, the relevant feedback

can still be used when the initial accuracy is very high. For the case where the initial P@50 is high but not 1, the use of the relevant feedback can improve P@50, but multiple rounds of feedback may be required. For the case of low initial P@50, the first few rounds of relevant feedback can be greatly improved, and the later promotion becomes slow or even impossible. This may be related to the number of relevant learning resources. If the number of relevant learning resources is insufficient, there will be learning resources related to insufficient queries in the first 50 query results.

Discussion

This article mainly uses the related methods of artificial intelligence to conduct experimental analysis on the construction of the auxiliary English learning resource query system. This article first analyzes and introduces the related algorithms of artificial intelligence. Then, take the construction of the query system as the research object, and find out that the query accuracy of the query system is very high, and the relevant feedback can still be used. The full text is based on the artificial intelligence algorithm to study



the construction of the auxiliary English learning resource query system. This is not only a further expansion of the research field of research methods, but also an in-depth discussion of the research on the learning resource query system.

Conclusion

The lack of information construction of English learning resources website will lead to the reduction of website effectiveness and ease of use, and affect users' satisfaction with the website. There are not only quantitative factors such as the number of clicks, the number of views, and the number of search results but also qualitative factors such as the simplicity of information organization, the clarity of navigation, the ease of learning of the search system, and so on. From the depth of retrieval, the method of establishing a knowledge base to express learning resources is conducive to the sharing and reuse of learning resources. It can understand the knowledge semantics of the retrieval content input by learners through the retrieval system, and the retrieval of learning resources can be transformed into the retrieval of knowledge points. According to the knowledge points, it can get the learning resources related to

the retrieval needs. From the perspective of retrieval efficiency, on the premise of the huge learning resource base, online retrieval system needs to give retrieval results in real-time. If all resource databases are used as retrieval objects, the efficiency will be very low, which cannot meet the needs of learners, resulting in the loss of interest in learning. Knowledge points can be used to describe learning resources, and the number of knowledge points is far lower than that of learning resources. Knowledge points are used to represent learning resources and the resource index of knowledge points can be established offline. In this way, the efficiency of retrieval is avoided because of a large number of calculations. Because the experimental data of this experiment come from actual analysis, the objectivity is strong, the data source is more realistic, and the reliability of the experimental results is high. However, due to the existence of some defects in the experiment, the experiment needs to be further improved.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

WY put forward research ideas and editing. NL design research scheme, collecting, and analyzing data. Both authors contributed to the article and approved the submitted version.

Funding

This work was supported by the Teaching Reform and Innovation Project of colleges and universities in the Shanxi Province Research on the Construction and Application of Dynamic Discourse Corpus from the perspective of Multimodality (MDA) (No. J2021635) from the Shanxi Provincial Education Department. It was supported by the Teaching Reform and Innovation Project of Jinzhong University Research on the teaching resource database construction and application of project-based foreign language flipped classroom based on VR (No. Jg201922) from the Jinzhong University. It was supported by the program of High-quality Sharing course Cross-cultural communication (No. KC2020003) at

References

- Aadil, M., and Asger, M. (2017). English to Kashmiri transliteration system - a hybrid approach. *Int. J. Comput. Appl.* 162, 5–8. doi: 10.5120/ijca2017913418
- Abdi, H., and Valentim, D. (2007). Multiple correspondence analysis. *Encycl. Meas. Stat.* 2, 651–657.
- Bajpai, P., Verma, P., and Abbas, S. Q. (2018). English-Hindi cross language information retrieval system: query perspective. *J. Comput. Sci.* 14, 705–713. doi: 10.3844/jcsp.2018.705.713
- Berg, K., and Aronoff, M. (2021). Is the English writing system phonographic or lexical/morphological? A new look at the spelling of stems. *Morphology* 31, 315–328. doi: 10.1007/s11525-021-09379-5
- Bolu, F. O., Adefowoke, O. B., and Omomule, T. G. (2020). Development of an intelligent English auction system. *i-manager s J. Comput. Sci.* 7, 1–13. doi: 10.26634/jcom.7.3.16530
- Chen, H., and Huang, J. (2021). Research and application of the interactive English online teaching system based on the internet of things. *Sci. Program.* 2021, 1–10. doi: 10.1155/2021/5089236
- Do, P., Phan, T., and Gupta, B. B. (2021). Developing a Vietnamese tourism question answering system using knowledge graph and deep learning. *ACM Trans. Asian Low Resour. Lang. Inf. Proc.* 20, 1–18. doi: 10.1145/3453651
- Dombrovan, T. I., Mitina, O. M., Rostomova, L. M., Slobodzowa, I. V., and Pavliuk, K. T. (2021). On the historical development of the phonetic system of English: a linguosynergetic approach. *Nexo Rev. Cient.* 34, 330–345. doi: 10.5377/nexo.v34i01.11310
- Fatmawati, T., Zaman, B., and Werdiningsih, I. (2017). Implementation of the common phrase index method on the phrase query for information retrieval. *AIP Conf. Proc.* 1867, 1–9. doi: 10.1063/1.4994430
- Fitria, T. N. (2020). Teaching English through online learning system during Covid-19. *Pedagogy J. Eng. Lang. Teach.* 8, 138–148. doi: 10.32332/pedagogy.v8i2.2266
- Flint, S. W., Piotrkowicz, A., and Watts, K. (2022). Use of artificial intelligence to understand adults' thoughts and behaviours relating to COVID-19. *Perspect. Public Health* 142, 167–174. doi: 10.1177/1757913920979332
- Gao, T., Fodor, P., and Kifer, M. (2019). Querying knowledge via multi-hop English questions. *Theory Pract. Logic Program.* 19, 636–653. doi: 10.1017/S1471068419000103
- Hong, Y., and Nam, H. (2021). Evaluating score reliability of automatic English pronunciation assessment system for education. *Stud. Foreign Lang. Educ.* 35, 91–104. doi: 10.16933/sfle.2021.35.1.91
- Islamova, T. (2019). Some problems of the training english language of higher education system. *Sci. Bull. Namangan State Univ.* 1, 45–45.
- Kaleem, R., Mir, R. R., Farhan, M., Rafiq, T., and Aslam, M. (2017). Student query trend assessment with semantical annotation and artificial intelligent multi-agents. *Eurasia J. Math. Sci. Technol. Educ.* 13, 3893–917. doi: 10.12973/eurasia.2017.00763a
- Liang, S. (2022). Research on innovative thinking of ceramic art design based on artificial intelligence. *Arts Stud. Crit.* 3, 203–206. doi: 10.32629/asc.v3i2.917
- Liu, L. (2021). Construction of a 5G wireless semantic web-assisted English digital learning resource query system. *Discrete Dyn. Nat. Soc.* 1–10. doi: 10.1155/2021/8514944
- Meng, Y. (2020). Research and analysis of intelligent English learning system based on improved neural network. *J. Intell. Fuzzy Syst.* 39, 1–11. doi: 10.3233/JIFS-179946
- Narayanan, R., and Mathew, P. (2021). Teaching International English Language Testing System (IELTS) academic writing and exam strategies online to develop omani students' writing proficiency. *Arab World English J.* 2, 49–63. doi: 10.24093/awej/MEC2.4
- Nie, A. (2021). Design of english interactive teaching system based on association rules algorithm. *Secur. Commun. Networks* 2021, 1–10. doi: 10.1155/2021/9275363
- Pantanowitz, L. (2022). Improving the Pap test with artificial intelligence. *Cancer Cytopathol.* 130, 402–404. doi: 10.1002/cncy.22561
- Ramchandani, G., Millar, R., and Wilson, D. (2021). The relationship between team ability and home advantage in the English football league system. *German J. Exerc. Sport Res.* 51, 354–361. doi: 10.1007/s12662-021-00721-x
- Saura, J. R., Palacios-Marqués, D., and Ribeiro-Soriano, D. (2022b). *Exploring the boundaries of open innovation: evidence from social media mining. Technovation*, in press. doi: 10.1016/j.technovation.2021.102447
- Saura, J. R., Ribeiro-Soriano, D., and Palacios-Marqués, D. (2021a). Setting B2B digital marketing in artificial intelligence-based CRMs: a review and directions for future research. *Ind. Mark. Manag.* 98, 161–178. doi: 10.1016/j.indmarman.2021.08.006

Jinzhong University in 2020. This work was supported by the University-Industry Collaborative Education Program Construction of Practice Base for Innovative and Compound Talents of Business English under the Background of New Liberal Arts (Project Number: 202102107018).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Saura, J. R., Ribeiro-Soriano, D., and Palacios-Marqués, D. (2021b). Setting privacy “by default” in social IoT: Theorizing the challenges and directions in Big Data Research. *Big Data Res.* 25, 100245. doi: 10.1016/j.bdr.2021.100245

Saura, J. R., Ribeiro-Soriano, D., and Palacios-Marqués, D. (2022a). Assessing behavioral data science privacy issues in government artificial intelligence deployment. *Gov. Inf. Q.* 39, 101679. doi: 10.1016/j.giq.2022.101679

Thenmozhi, D., and Chandrabose, A. (2018). Ontology-based Tamil–English cross-lingual information retrieval system.

Sādhanā 43, 157–157. doi: 10.1007/s12046-018-0942-7

Yong, W. (2021). Design and implementation of intelligent english electronic dictionary system based on internet of things. *Wirel. Commun. Mob. Comput.* 2021, 1–11. doi: 10.1155/2021/5586662

Yuan, Z., Jin, C., and Chen, Z. (2020). Research on language analysis of English translation system based on fuzzy algorithm. *J. Intell. Fuzzy Syst.* 40, 1–9. doi: 10.3233/JIFS-189443