PsyChatbot: A Psychological Counseling Agent Towards Depressed Chinese Population Based on Cognitive Behavioural Therapy

TIANTIAN CHEN, School of Software Engineering, Tongji University, Shanghai, China
YING SHEN*, School of Software Engineering, Tongji University, Shanghai, China
XURI CHEN, School of Humanities, Tongji University, Shanghai, China
LIN ZHANG, School of Software Engineering, Tongji University, Shanghai, China

Nowadays, depression has been widely concerned due to the growing depressed population. Depression is a global mental problem, the worst case of which can lead to suicide. However, factors such as high treatment costs and social stigma prevent people from obtaining effective treatments. Chatbot technology is one of the main attempts to solve the problem. But as far as we know, existing chatbot systems designed for depressed people are still sporadic, and most of them have some non-negligible limitations. Specifically, existing systems simply guide users to release their negative emotions or provide some general advice. They cannot offer personalized advice for users' specific problems. In addition, most of them only support English speakers, despite the fact that depressed Chinese constitute a large population. Psychological counseling systems for the depressed Chinese population with improved responsiveness are temporarily lacking. As an attempt to fill in the research gap to some extent, we design a novel Chinese psychological chatbot system, namely PsyChatbot. First, we establish a counseling dialogue framework based on Cognitive Behavioral Therapy (CBT), which guides users to reflect on themselves and helps them discover their negative perceptions. Then, we propose a retrieval-based Q&A algorithm to provide suitable suggestions for users' specific problems. Last but not least, we construct a large-scale Chinese counseling Q&A corpus, which contains nearly 89,000 psychological Q&A triples. Experimental results have demonstrated the effectiveness of PsyChatbot. The source code and data has been released at https://github.com/slptongji/PsyChatbot.

CCS Concepts: • **Human-centered computing** → **Human computer interaction (HCI)**; **Empirical studies in HCI**; *Natural language interfaces*;

Additional Key Words and Phrases: Chatbot, conversational agent, language design, mental health, psychiatry, depression, CBT

1 Introduction

Depression (major depressive disorder) is one of the most common and serious mental illnesses, the worst case of which can induce self-mutilation and suicide [2, 18]. Nowadays, the depressed population is increasing year by year. A recent study reveals that there were nearly 100 million people suffering from depression in China, and the lifetime prevalence of depression was 6.8% [24]. More than 280,000 Chinese people commit suicide every year, 40%

*Corresponding author

This work was supported by the Fundamental Research Funds for the Central Universities.
Authors' Contact Information: Tiantian Chen, School of Software Engineering, Tongji University, Shanghai, Shanghai, China; e-mail: 2111287@tongji.edu.cn; Ying Shen, School of Software Engineering, Tongji University, Shanghai, Shanghai, China; e-mail: yingshen@tongji.edu.cn; Xuri Chen, School of Humanities, Tongji University, Shanghai, Shanghai, China; e-mail: xurichen@tongji.edu.cn; Lin Zhang, School of Software Engineering, Tongji University, Shanghai, Shanghai, China; e-mail: cslinzhang@tongji.edu.cn.

ACM 2375-4702/2024/7-ART
https://doi.org/10.1145/3676962

of whom are suffering from depression [30]. Although depression can be treated by medication, psychotherapy, and physical therapy, 35.5%~50.3% of serious cases in developed countries and 76.3%~85.4% in less-developed countries received no treatment [18]. There are three main reasons accounting for the low treatment rates:

(1) Shortage of medical resources. Compared with the considerable size of the depressed population, very few social workers are engaged in psychological services [28]. In addition, patients in less-developed regions may fail to receive professional services within their reach [39].
(2) Stigma towards depression. The public has prejudice and discrimination against patients with depression, which makes them ashamed to admit their illness and refuse to look for psychological assistance [53].
(3) High cost of psychological intervention. A study in 2009 [13] shows that depressive inpatients spent an average of more than 70,000 yuan for the disease. By contrast, the Chinese residents' per capita disposable income in 2021 was only 35,128 yuan [27].

To improve the treatment rates of the patients, over recent years, researchers have continuously developed Internet-based Cognitive Behavioural Therapy (ICBT) to provide high-quality online mental health services [3, 46]. ICBT-based applications can be categorized into two types. The first type of applications only provides a chatting platform, which connects the depressed users and the psychotherapists [7, 14, 44]. However, the online treatments are still expensive for the users. In addition, the treatments should be reserved with the psychotherapists beforehand, which is inflexible considering the patients' needs. The second type of applications integrates a psychological chatbot agent in addition to the chatting platform [11, 14, 25]. The psychological chatbot agent can automatically carry out psychotherapy to the depressed users and consequently relieve their anxiety or depression. The psychotherapy services are much cheaper and can be provided to the depressed users at anytime and anywhere. However, most of these applications adopt predefined response templates to implement psychotherapy, which is far from intelligent enough to deal with various mental problems proposed by the users.

Despite the weakness of the existing ICBT-based applications, experimental results indicate that these applications could timely intervene and improve users' psychiatric states, and their intervention effects on anxiety and depression were close to traditional face-to-face psychotherapy [1, 4, 43]. Compared with traditional face-to-face psychotherapy, they have the advantages of not being restricted by time and space, breaking the shackles of patients' self-stigma, and reducing their economic burdens. Therefore, the promotion of ICBT-based applications can make it easier for individuals to access psychological assistance. However, most of the ICBT-based applications simply guide users to release their negative emotions and provide some comfort [9, 29, 36] instead of providing professional psychotherapy. In addition, there are few ICBT-based applications towards Chinese users despite that depressed Chinese constitute a large population.

In this work, we investigate the shortcomings of the existing psychological intervention systems and propose a novel psychological counseling agent based on cognitive behavioural therapy (CBT) for depressed Chinese people. Our contributions can be summarized as follows:

- We implemented a CBT-based Chinese psychological counseling agent, namely PsyChatbot. PsyChatbot provides online interventions for the depressed people based on the CBT model. It can reshape users' faulty cognitions and alleviate their depressive feelings.
- We designed a CBT-based guided dialogue framework. The framework guides users to describe their emotions and problems by raising a series of questions. At the end of the dialogue, it provides appropriate professional psychological advice to users' problems.
- To provide professional advice, we constructed a Chinese psychological Q&A dataset, namely cPsychQASet. cPsychQASet consists of 89,000 professional and individualized suggestions against the problems raised by the depressed people. It can greatly benefit the development of related applications for mental state intervention.

- We proposed a topic-aware psychological Q&A retrieval algorithm. It calculates the similarity between the user's input and the candidate questions stored in cPsychQASet, and then retrieves suitable psychological advice to the questions that are most similar to the user's problems. In this way, PsyChatbot can provide users with more targeted psychological advice.

2 Related Work

2.1 Computer-aided Psychological Therapy

In recent decades, chatbot techniques have been widely used in psychological intervention. Bickmore et al. [9] designed an empathic chatbot with a virtual avatar for depressed people. The chatbot introduced therapeutic schedules and precautions to the patients in the form of text and speech. Shim [25] was a text-based chatbot that utilized the CBT method to alleviate users' anxiety. Gardiner et al. [12] provided lifestyle-related suggestions to participants through a chatbot application. The most successful chatbot applications for depression intervention are Woebot [11] and Wysa [14]. Woebot implemented the CBT method to alleviate users' depression and anxiety through short daily conversations and emotion tracking. Wysa combined CBT, behavioral reinforcement, and other approaches to help patients with depression (as shown in Fig. 1). In addition to the effectiveness of these applications, recent research reveals that users believed it was safe to disclose sensitive information to chatbot agents [8].

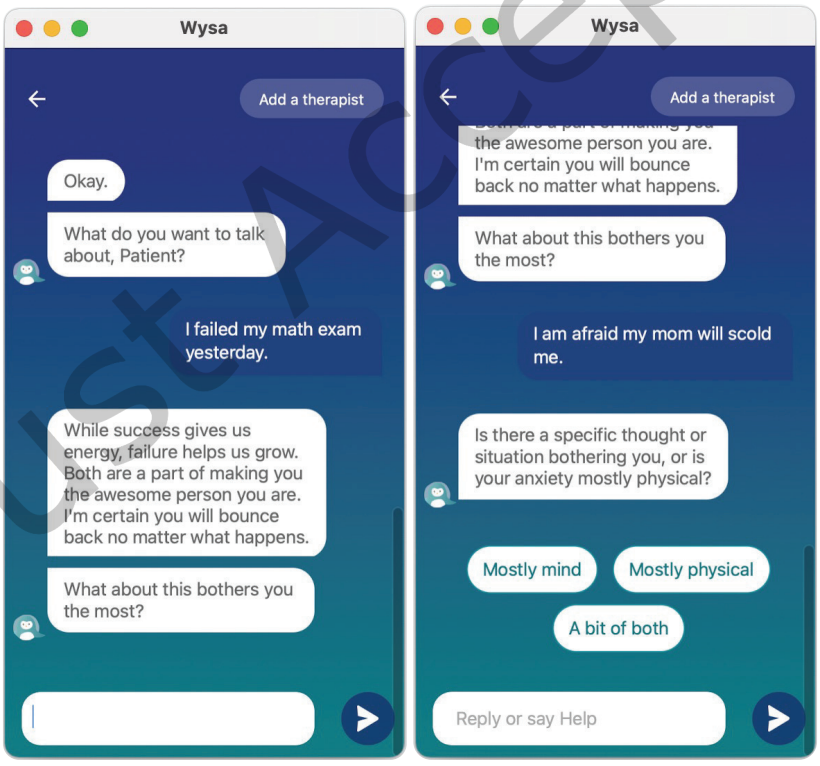

Fig. 1. Psychological intervention interface of Wysa [14].

Although these applications have achieved the purpose of psychological intervention to some extent, they still have some limitations. Firstly, the number of mental health applications is quite limited. Despite the abundance of smart agents (such as Siri, Cortana, and Xiao Ice) in our daily lives, few chatbot applications are specifically designed for psychological intervention. Furthermore, many claimed "psychological intervention" applications only offer users functionalities for daily emotion tracking and basic psychological knowledge. Woebot and Wysa are the only two publicly available mental health applications, but they exclusively cater to English-speaking users. The applications designed for Chinese users are even rarer. Nowadays, only a few studies have explored psychological chatbot applications in Chinese contexts, such as the Chinese version of MoodPYM [35] and Emohaa [37]. However, there is still a lack of publicly accessible psychological intervention applications that support Chinese. Secondly, the psychological suggestions provided by these applications tend to be overly general. These applications typically adhere to a fixed process, restricting user expression with pre-defined options. As illustrated in Fig. 1, Wysa recognized the word "*failed*" mentioned by the user and reassured him/her not to worry too much about the failure. Such advice, while reasonable, did not directly respond to the exam failure. As a result, chatbot systems like Wysa often struggle to empathize with users effectively. In addition, some applications like Emohaa attempt to utilize generative algorithms to generate flexible responses. However, they may inevitably generate incoherent and less informative responses, hindering their ability to provide informative and thought-provoking suggestions for users. Lastly, the existing psychological intervention systems are mainly proposed by researchers from psychology. These systems primarily explore various online interaction modalities from a psychological standpoint but often overlook the intelligence of the systems themselves. Consequently, the current psychological intervention applications exhibit limited semantic understanding capabilities, serving merely as "tree holes" that guide users to release their negative feelings, rather than comprehending their thoughts.

2.2 Chatbot Technology

The chatbot technology can be categorized into two main classes: the rule-based methods and the learning-based methods. The former ones generate responses based on predefined rules. The latter ones generate responses based on the learned pragmatic rules from the large-scale dialogue corpus.

The rule-based methods are simple in principle and do not require extensive data for training. For example, ELIZA [49] was a representative rule-based chatbot. It recognized the input keywords and matched templates to generate predefined responses. Although ELIZA did not really appreciate the users' inputs, users still felt that it could understand them [49]. The rule-based methods are simple in principle and do not require extensive data for training. However, rule customization is time-consuming and labor-intensive. Besides, they are inflexible when dealing with complicated problems.

The learning-based chatbots can be further divided into retrieval-based and generative-based chatbots. The retrieval-based chatbots use heuristic algorithms to select appropriate replies from the predefined corpus. Lowe et al. [22] were the first to utilize neural networks in retrieval-based dialogue systems. The method retrieved candidate replies using the TF-IDF (Term Frequency–Inverse Document Frequency) algorithm [23] and calculated the matching scores of the candidate replies based on an LSTM model. SMN [50] further fused the candidate replies with the context information by accumulating the matching information through a GRU model. The existing retrieval-based chatbot technology has the advantages of being simple and controllable, and it can generate natural and grammatical responses. However, it lacks the capability of flexible response generalization.

The generative-based chatbots learn response generation rules from the large-scale dialogue corpus and generate words one by one according to the learned rules. They do not depend on the predefined corpus when generating responses. In addition, the generated responses can be utterances that do not exist in the training corpus. Vinsal et al. [47] constructed an end-to-end dialogue model which generated responses by maximizing the probability of word generation. Serban et al. [40] added a context encoder to the basic encoder-decoder structure,

which could model the context more effectively. The generative methods can generate diverse responses. However, they need to learn the generation rules from massive amounts of data and are prone to generate grammatical errors and incoherent sentences.

The rule-based and retrieval-based methods are widely used in the real world due to their mature technologies and the natural responses generated. Besides, existing psychological intervention systems all adopt these kinds of methods. For example, the overarching methodology adopted by Woebot [11] is a decision tree which determines the routing to the subsequent conversation nodes based on the user inputs. The nodes on the decision tree store various response templates which are used to generate suitable responses. Wysa's so-called AI models were trained in-house by clinicians. Unlike its name, it did not use AI-generated responses as psychological advice. Instead, it utilized the predefined interventive conversations created by an internal team to implement psychological intervention [5]. Therefore, in this work, we integrated the rule-based methods and the retrieval-based methods to implement a psychological counseling agent under the CBT framework for Chinese depressed people, which can automatically recognize users' problems and generate specific psychological suggestions for them.

2.3 Cognitive Behavioral Therapy

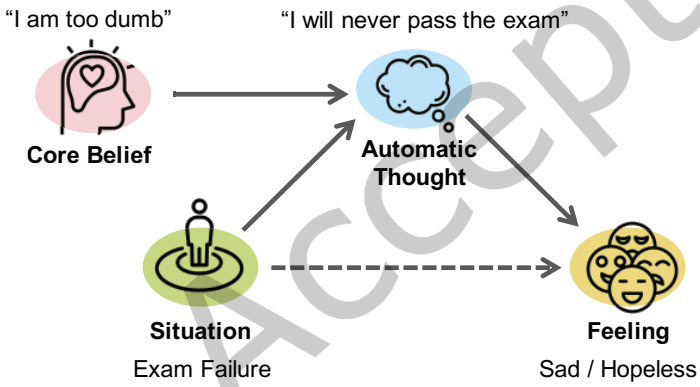

Fig. 2. An example of maladaptive thoughts on the exam failure. The student exaggerated the consequences of an exam failure and developed the false perception "*I will never pass the exam*", which led him/her to a state of depression.

Cognitive behavioral therapy [6] is a type of short-term psychotherapeutic treatment. It focuses on "*here and now*" problems and looks for ways to improve patients' mental states. The cognitive-behavioral model suggests that it is one's thoughts about his/her situation that intrinsically contribute to his/her emotions [15]. For example, as shown in Fig. 2, a student failed his exam and fell into a depressive state. According to the cognitive-behavioral model, the situation of exam failure could not arouse depressive feelings. Instead, the automatic thoughts about the exam failure ("*I will never pass the exam*") were the main reason that contributed to the depressive feelings. In addition, these automatic thoughts are often induced by inappropriate core beliefs ("*I am too dumb*").

As shown in Fig. 2, the maladaptive thoughts distort and exaggerate real-life situations, inducing individuals pessimistic about themselves, the world, and their lives. Therefore, identifying and changing maladaptive thoughts is the central focus of CBT. CBT methods utilize the *situation* → *thought* → *feeling* triangle to identify and change automatic thoughts [15]. In the face-to-face treatment, the therapist guides the patient to describe his/her emotions and situations, observes and records his/her psychiatric states, and assists his/her in identifying the automatic thoughts associated with the emotion shift. An example is shown in Table 1. The therapist first guides

the patient to describe his/her recent mood ("*sad*") and the triggered event ("*failed the math exam*"). Then the therapist assists the patient to understand that his/her hopelessness stems from his/her maladaptive thoughts ("*I will never pass the exam*") instead of the event itself. In this way, the therapist helps the patient identify their maladaptive automatic thoughts and further improve their emotional state by transforming their way of thinking.

Speaker	Content
Therapist	How have you been feeling this week?
Patient	Just really sad…as usual. It seems like I'm always feeling that way.
Therapist	Did anything in particular trigger this sad feeling this weekend?
Patient	Yes, I failed the math exam this week. I think I'm so stupid that I'll never pass the exam.
Therapist	So, let's write down this automatic thought that you are having. I will never pass the exam. Your exam failure was the situation that triggered the thought, 'I will never pass the exam'.
Patient	Yes, that's true.
Therapist	When you got your math grade and that thought came to you, how did you feel?
Patient	I felt really sad and hopeless.
Therapist	So, can you see how our thoughts can affect our mood and change the way we are feeling? Although you failed this exam, you can pass it if you prepare well. Would you still feel so bad if you thought this way?
Patient	Yeah, I guess if I hadn't had that thought, I wouldn't have felt so bad.

Table 1. An example of CBT therapeutic session. The therapist helps the patient recognize that his/her hopelessness stems from his/her maladaptive thoughts.

3 PsyChatbot: A Chinese Psychological Counseling Agent Based on CBT Framework

We have developed a new Chinese psychological intervention system, namely **PsyChatbot**, which can guide users to build a positive way of thinking under the CBT framework and alleviate their depressive feelings. In this work, we have solved three key problems, which include:

- How to implement CBT principles using the chatbot technology?
- How to accurately retrieve psychological advice from a predefined professional counseling corpus for the user's specific question?
- How to construct the predefined professional counseling corpus?

In this work, a CBT-based guided dialogue framework is proposed to implement CBT principles. Besides, a Chinese psychological counseling corpus, cPsychQASet, is constructed for providing professional psychological advice. In the end, a topic-aware psychological Q&A retrieval algorithm is proposed to accurately retrieve psychological advice from cPsychQASet. The details of our work have been illustrated in the following subsections, respectively.

3.1 CBT-based Psychological Intervention Framework

As mentioned in Sec. 2.3, the core therapeutic technique of CBT is knowledge reconstruction, i.e., to identify and reshape the maladaptive thoughts of the depressed people. In the clinical treatment, a stage of the rapid CBT method consists of the following three steps [45] (as shown in Fig. 3(a)):

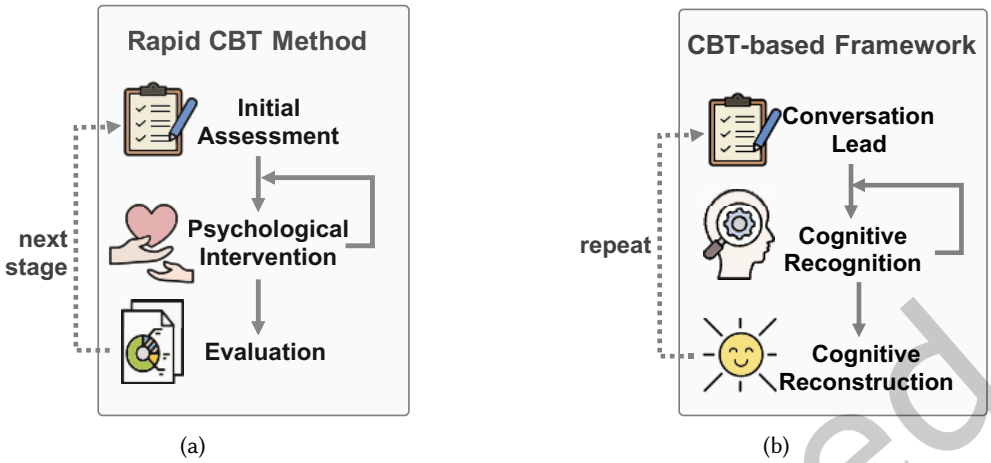

Fig. 3. (a) The rapid CBT approach used in the clinical psychotherapy; (b) The proposed CBT-based psychological intervention framework.

(1) Initial assessment. Therapists focus on patients' mental states, disease progression, appeals and so on. They assess the causes, types, severity, and physical states of patients' mental problems.
(2) Psychological intervention. During the therapy session, therapists listen to patients' problems and help them get a thorough understanding of the situations and their thoughts. At the end of the session, therapists execute individualized treatment plans to help the patients establish positive self-perceptions.
(3) Evaluation. Therapists summarize the treatment results, analyze the reasons for the effectiveness or the failure of the treatments, and record the plans for the next stage of the treatments. If necessary, the therapists schedule the next stage of treatments with the patients.

According to the rapid CBT approach mentioned above, we proposed a CBT-based intelligent psychological intervention framework. As shown in Fig. 3(b), the CBT-based psychological intervention framework consists of three steps:

(1) Conversation lead. When an intervention session starts, users are led to talk about their recent problems and feelings. These problems may relate to their health status, marital problems, academic stress, etc. After the users have fully described their problems, they are encouraged to share their thoughts on these problems.
(2) Cognitive recognition. The system automatically identifies users' emotions and maladaptive thoughts based on the users' inputs, helps them identify the ways of thinking and behaviors that may cause the problems.
(3) Cognitive reconstruction. The system attempts to rectify the users' maladaptive thoughts by providing them with professional advice and solutions, which are retrieved from a professional counseling corpus.

The CBT-based psychological intervention framework is implemented with AIML (Artificial Intelligence Markup Language) [48]. AIML is an XML dialect for creating natural language software agents. It has high grammatical accuracy and process controllability, which is capable to simulate CBT interventions. Fig. 4 shows the workflow of the CBT-based dialogue framework. The framework consists of three stages: Conversation Lead, Cognitive Recognition and Cognitive Reconstruction. Conversation Lead and Cognitive Recognition, which correspond to *Initial Assessment* in the rapid CBT approach, are implemented using AIML to guide users to describe their feelings and problems. Cognitive Reconstruction, which corresponds to *Psychological Intervention* in the rapid CBT approach, is implemented by a retrieval algorithm to retrieve specific psychological advice

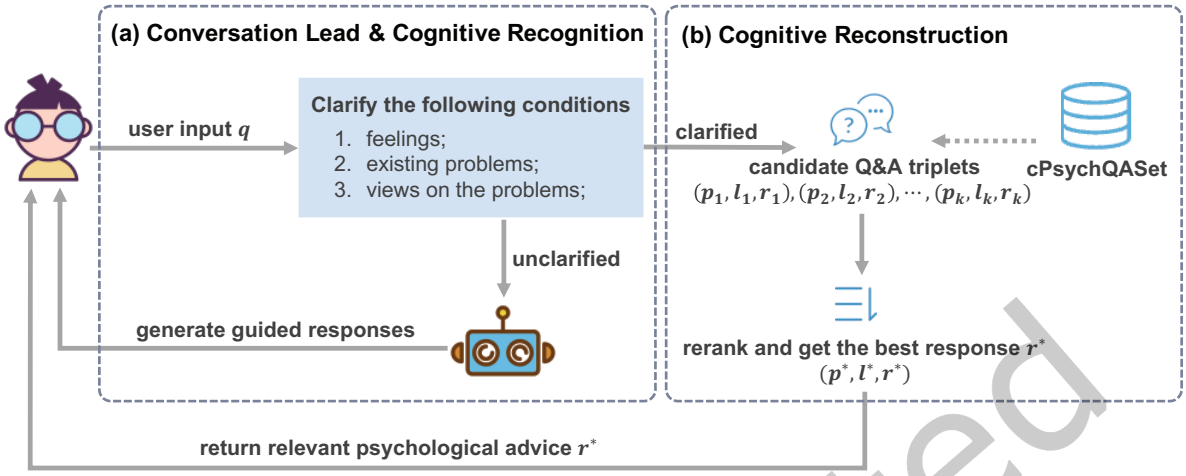

Fig. 4. The workflow of CBT-based dialogue framework. (a) When the conversation starts, PsyChatbot switches to the stage of Conversation Lead or Cognitive Recognition according to the user's input. It guides the user to describe his/her feelings, existing problems and his/her thoughts on these problems. (b) When the user clarifies his/her problems and seeks psychological assistance, PsyChatbot switches to the stage of Cognitive Reconstruction and retrieves relevant psychological advice from cPsychQASet, a professional psychological counseling corpus, to provide psychological support.

for users' problems to rectify users' cognitions. The *Evaluation* step in the rapid CBT approach is omitted in PsyChatbot because it is only used for the treatment summarization for the therapists.

It should be noted that the transitions between three stages are conducted by the AIML interpreter automatically. The interpreter constructs a decision logic tree to determine which type of responses should be generated. After the user has initiated the conversation, the interpreter searches the logic tree based on the user's input. If the pattern of the input matches one of the tree nodes, PsyChatbot generates responses using the template attached to the node. If the user's input doesn't match any pattern stored on the tree, the interpreter assumes that a specific question is raised and will retrieve suitable psychological advice from cPsychQASet. In this way, PsyChatbot completes the transitions from different stages.

3.1.1 Conversation Lead & Cognitive Recognition. In the stages of *Conversation Lead* and *Cognitive Recognition*, users are guided to share their situations, thoughts and feelings using a set of predefined response templates. Each time the user's input is received, the system normalizes the input text and matches it with the AIML rules which are designed based on CBT conceptions.

A simple example of the rules is shown in Fig. 5. In this example, *<aiml></aiml>* defines the beginning and the end of the XML file, *<category></category>* defines the CBT-based response rules, *<pattern></pattern>* defines the patterns for input matching, *<template></template>* defines the response templates under the rule, and * is a wildcard character. An AIML file usually contains several rules. An AIML rule can correspond to multiple input patterns and multiple response templates. While matching a rule, the system reads the response template, processes the special labels (e.g., * or some variables) which appear in the template and generates a response according to the rules. Finally, the system returns the response to the user and waits for the new utterances.

As shown in Fig. 6, PsyChatbot starts the conversation and guides the user to describe his/her current feelings. During the conversation, PsyChatbot identifies the user's emotions (keywords such as "*unhappy*" or "*sad*") and specific reasons (keywords such as "*because* $\cdots$") from user's utterances and then verifies the user's troubles and problems in life.

```
<aiml version = "1.0.1" encoding = "UTF-8"?>
<category>
        <pattern> HELLO </pattern>
        <template>
          哈喽，你今天感觉怎样呀？
        </template>
</category>
<category>
         <pattern>我感觉 *</pattern>
         <template>
          你经常感到 <star index="1"/> 吗？
         </template>
</category>
</aiml>
```

```
<aiml version = "1.0.1" encoding = "UTF-8"?>
<category>
        <pattern> HELLO </pattern>
        <template>
          Hello, how are you feeling today?
        </template>
</category>
<category>
         <pattern>I feel *</pattern>
         <template>
          Do you often feel <star index="1"/>?
         </template>
  </category>
</aiml>
```

(a) An AIML file containing different CBT-based AIML rules.

Speaker	Response	Translated Response
PsyChatbot	哈喽，你今天感觉怎样呀？	Hello, how are you feeling today?
User	我感觉很不开心。	I feel unhappy.
PsyChatbot	你经常感到不开心吗？	Do you often feel unhappy?
User	时不时。	Sometimes.

(b) The dialogue example that corresponds to the CBT-based AIML rule.

Fig. 5. A simple CBT-based AIML rule and its corresponding dialogue example.

If the emotions and reasons are not complete enough for the retrieval algorithm to retrieve suitable suggestions, PsyChatbot will continue to generate guided responses (such as "*Do you often feel unhappy?*" or "*What happened recently that made you think in this way?*"), guiding the user to talk more about his/her moods and the occurred events. By posing questions like "*What is the cause do you think?*", PsyChatbot encourages the user to reflect on himself/herself and find the inappropriateness in his/her thoughts. Finally, PsyChatbot retrieves suitable answers (the proposed retrieval algorithm will be illustrated in Sec. 3.3) from cPsychQASet and returns them as psychological advice for the user. In this way, the system can reshape user's maladaptive thoughts and guide him/her to think from a positive perspective.

3.1.2 Cognitive Reconstruction. Once the user's emotions and problems are obtained, the user's input $\boldsymbol{q}$ is matched with the Q&A triplet $(\boldsymbol{p}_i, \boldsymbol{l}_i, \boldsymbol{r}_i)$ retrieved from cPsychQASet, where $\boldsymbol{p}_i$ is the i-th problem stored in cPsychQASet, $\boldsymbol{r}_i$ is the corresponding response to $\boldsymbol{p}_i$, and $\boldsymbol{l}_i$ is the topic label vector of $\boldsymbol{p}_i$. The matching scores are calculated based on the topic similarity and the text similarity between $\boldsymbol{q}$ and $\boldsymbol{p}_i$. Finally, k candidate responses $\{\boldsymbol{r}_i\}_{1\leq i\leq k}$ retrieved from cPsychQASet are sorted by the matching scores and the one with the highest score is regarded as the system response. The details will be illustrated in Sec. 3.3. In this way, PsyChatbot can provide users with suitable psychological advice based on the users' inputs and consequently alleviate their depressive feelings.

Speaker	Response	State	Origin
PsyChatbot	哈喽，你今天感觉怎样呀？ Hello, how are you feeling today?	Conversation Lead	(1) CBT-based AIML rules
User	很不好。 Terrible.		
PsyChatbot	你能详细描述一下你今天的心情吗？ Can you describe your mood in detail?	Conversation Lead	(1) CBT-based AIML rules
User	沮丧，嫉妒，无能为力。 Frustrated, jealous, and powerless.		
PsyChatbot	你经常产生这种情绪吗？ Do you often feel this way?	Conversation Lead	(1) CBT-based AIML rules
User	有时候。 Sometimes.		
PsyChatbot	你觉得是什么原因引起的？ What's the cause, do you think?	Cognitive Recognition	(2) CBT-based AIML rules
User	对同学的嫉妒，觉得他很受欢迎。 Jealousy to my classmates. I think he is very popular.		
PsyChatbot	嗯嗯，我明白了。你觉得还有其他原因使你产生这种情绪吗？ Hmm, I get it. Are there other reasons that make you have such feelings?	Cognitive Recognition	(2) CBT-based AIML rules
User	我有点自卑，觉得自己一无是处。 I have low self-esteem and I think I am nothing.		
PsyChatbot	每个人身上都有太阳，关键是如何让它发光。 In every person there is a sun. Just let them shine.	Cognitive Reconstruction	(3) Retrieved from Q&A corpus

Fig. 6. An example of the CBT-based guided dialogue framework. PsyChatbot first asks about the user's recent emotions to discover the problems occurring in user's life (1). Then, by asking the trigger events of the emotions, PsyChatbot obtains the user's views on their passive emotions (2). Finally, PsyChatbot uses a topic-aware Q&A retrieval algorithm to retrieve suggestions from cPsychQASet, reshaping their negative thoughts (3).

3.2 Chinese Psychological Counseling Corpus

A professional psychological counseling corpus is essential for PsyChatbot when the system tries to provide psychological advice to the user in the step of *Cognitive Reconstruction*. However, it is difficult to obtain counseling information directly from the therapists due to privacy protection. Therefore, we turned to collecting publicly available data from three professional psychological counseling forums, i.e., *xinli001* (https://www.xinli001.com/), *yidianling* (https://www.ydl.com/), and *xinliwenda* (https://wenda.mrzcl.com/). The questions from Q&A sections of psychological counseling forums were proposed by various depressed users who sought for mental helps.

A Q&A example has been shown in in Fig. 7. In this example, a user raised an issue about how to maintain a love relationship. The therapist from the counseling forum emphasized the importance of tolerance in a relationship and advised the user to understand and be tolerant to his/her partner. Through the Q&A interaction, the user obtains the psychological support from the therapist. The advice can help him/her out and build a better relationship with his/her partner. Each question from Q&A sections of psychological counseling forums was assigned several topic labels (e.g., "*love*" shown in Fig. 7) so that experts in the related areas could answer the question timely and properly. The question $\boldsymbol{p}$ together with its topic label string $\boldsymbol{l}_s$ and the corresponding best answer $\boldsymbol{r}$ constitute a psychological Q&A triplet ($\boldsymbol{p}$,$\boldsymbol{l}_s$,$\boldsymbol{r}$). Based on these Q&A triplets, we construct a Chinese psychological Q&A dataset named **cPsychQASet**.

Title	如何经营爱情？ How to manage a love relationship?
Description	没谈恋爱的时候，以为自己什么都懂，以为自己会如鱼得水…真的碰到那个人，才发觉自己什么都不会，手足无措，像个小学生，感情的道路上磕磕碰碰…爱情，要怎么经营呢？ When I was not in a relationship, I thought I knew everything and I would be like a fish in water...But when I met that person, I realized I knew nothing, I was at a loss, like a schoolboy, stumbling on the road of relationships...Love, how to manage it?
Label	恋爱婚姻 love
Best Answer	爱情不是1+1=2，而是0.5+0.5=1。两个人各削去一半自己的个性和缺点，然后凑合在一起才完整。如果没有相互的忍让和包容，两个人很难成为一体而最终各分东西。 Love is not 1+1=2, it is 0.5+0.5=1. Two individuals cut out half of their personalities and flaws, and then come together as a whole. Without mutal forbearance and tolerance, it is hard for them to become one and the relationship will eventually break down.

Fig. 7. A Q&A example obtained from the psychological counseling forum.

3.2.1 Q&A Triplet Construction. To construct cPsychQASet, approximate 160,000 raw Q&A triplets were collected from three psychological counseling forums. The question in the raw Q&A triplets often corresponds to multiple answers which were given by different therapists. In order to construct the Q&A triplet, only the best answer was kept and associated with the corresponding question. we noticed that each answer contains an attribute named "*like*", which recorded the number of approvals to the answer and indicated its appropriateness to the raised question. Therefore, the answer with the maximum number of "*like*" was regarded as the best advice to the raised question. Those questions which were unanswered were filtered out directly. In addition, the questions whose answers received no "*like*" were also removed.

Each question from the the psychological counseling forums is associated with a "*label*" attribute, which indicates the topics that the question belongs to. Considering that the labels reveal the category information of the questions, they were kept in the Q&A triplets. By examining the topics of all the questions, we summarized seven categories of the labels, which include: *love*, *emotion*, *growth*, *family*, *relationship*, *treatment* and *career*. Some infrequent topics were manually assigned to one of the label categories mentioned above. For example, "*mother-in-law*" were merged into "*family*".

3.2.2 Data Cleaning. After obtaining a set of Q&A triplets, the contents of questions and answers in Q&A triplets were further processed. As shown in Fig. 7, the question consists of two parts: the "*title*" and the "*description*". The "*title*" contains the summary information and the "*description*" contains a detailed description of the problem. Therefore, they were concatenated to give a full description of the user's problem. Some problem descriptions are too long to be a suitable response in the dialog context. They were truncated with an upper length limit (e.g., 1,000 characters) and the exceed part was removed. In addition, some answers are too short or too long, which are inappropriate to be directly used as chatbot responses in conversational scenarios. Consequently, short answers which contain less than ten characters were also filtered out. For long answers whose lengths exceed the length limit, they were split into a set of discrete sentences by terminal punctuation. The longest sentence within the upper limit (e.g., 1,000 characters) was selected as the best answer.

Considering the large size of the dataset, it is difficult for us to verify the Q&A triplets one by one. Therefore, we performed a sampling evaluation to find potential issues in the initially constructed dataset and designed filtering rules to rectify these issues accordingly. Specifically, we randomly selected 250 Q&A examples from the initially constructed dataset (obtained from Sect. 3.2.1) to build a sampling set. Then, we verified the Q&A triplets in the sampling set one by one and recorded the appeared issues, such as garbled characters, self-introductions,

and advertisements. We manually designed filtering rules to fix these issues, such as removing garbled characters and self-introductions. These rules were applied to the entire dataset to ensure that the dataset was ethical and its contents were well-written.

In the end, nearly 89,000 valid Q&A triplets were obtained. Each Q&A triplet $(\boldsymbol{p}, \boldsymbol{l_s}, \boldsymbol{r})$ contains a question description $\boldsymbol{p}$, topic label string $\boldsymbol{l_s}$ of question $\boldsymbol{p}$, and the best answer $\boldsymbol{r}$ to question $\boldsymbol{p}$. These Q&A triplets consist of the psychological counseling corpus cPsychQASet, which will be used to provide professional advice and guidance in the Q&A retrieval module.

3.3 Topic-Aware Psychological Q&A Retrieval Algorithm

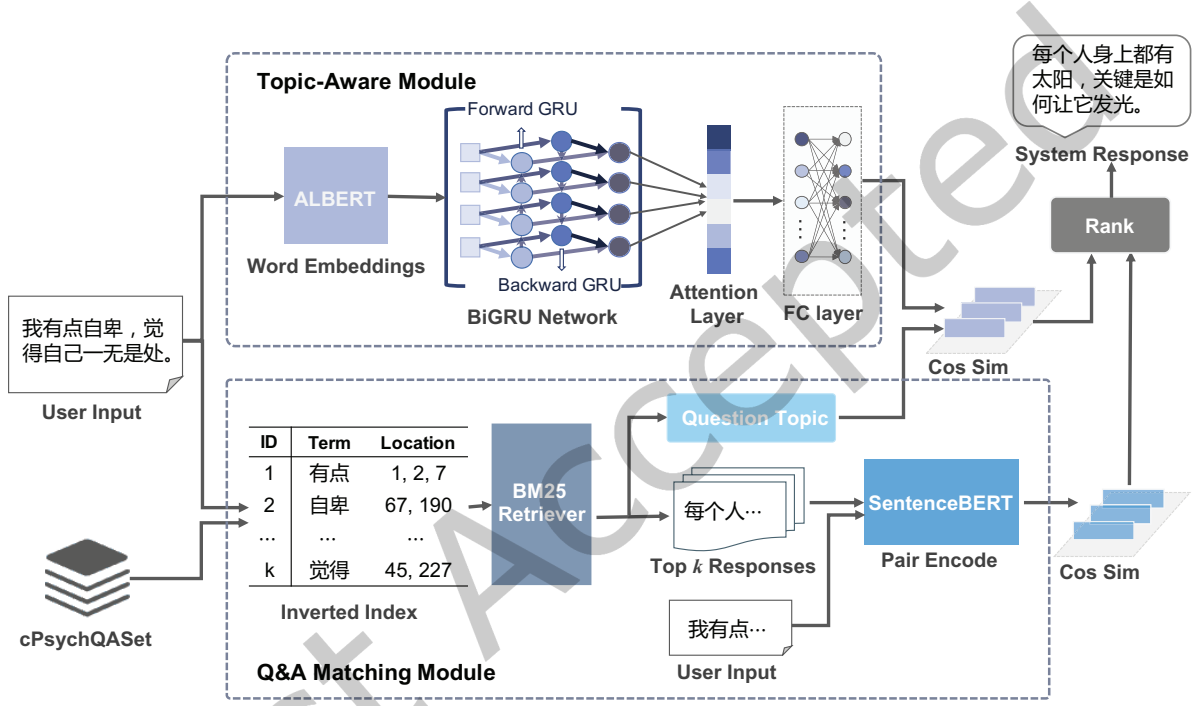

Fig. 8. The architecture of the topic-aware psychological Q&A retrieval algorithm.

The AIML-based guided dialogue framework solves the conversation guidance problem. It guides users to talk about their situations and feelings, and encourages them to reflect on their negative thoughts. When the conversation moves into the *Cognitive Reconstruction* stage, the system needs to provide psychological advice for user's specific questions. Therefore, a topic-aware psychological Q&A retrieval algorithm is proposed to provide appropriate psychological advice to the user in the *Cognitive Reconstruction* stage. When the user's input contains a specific question that cannot be solved by the predefined templates, PsyChatbot utilizes the topic-aware Q&A retrieval algorithm to retrieve a suitable answer from cPsychQASet and returns the answer to the user as the psychological advice.

As shown in Fig. 8, the algorithm consists of a *Topic-Aware Module* and a *Q&A Matching Module*. First, the user input $\boldsymbol{q}$ is compared with each question $\boldsymbol{p}'_i$ from the candidate Q&A triplet set $\{(\boldsymbol{p}'_i, \boldsymbol{l}'_i, \boldsymbol{r}'_i)\}_{1 \le i \le k}$ which are retrieved from cPsychQASet based on their content similarity and the topic similarity. Then the two similarities are weighted and added up to obtain the matching scores between $\boldsymbol{q}$ and $\boldsymbol{p}_i$. The answer $\boldsymbol{r}^*$ to the question $\boldsymbol{p}^*$

which corresponds to the highest matching score is outputted as the system response. The algorithm will be elaborated in the following subsections.

3.3.1 Topic-Aware Module. Generally speaking, PsyChatbot can retrieve suitable psychological advice from cPsychQASet merely based on the contents of the user's inputs. However, in our daily life, conversations are often made around certain topics. The topic information can help listeners better understand the speaker's thoughts and intents. Similarly, the topic information can help PsyChatbot better understand the user's situations and problems. If the topic information can be extracted from the user's input beforehand, PsyChatbot can retrieve more suitable psychological advice based on both the input contents and its potential topics.

Topic acquisition is typically treated as an unsupervised learning problem [20, 52, 55]. For example, Xing et al. [52] acquired topic word embeddings of messages from a pre-trained Twitter LDA model [54]. They estimated the parameters of Twitter LDA with the collapsed Gibbs sampling algorithm [54] and selected the top n words with the highest probabilities as the topic words. Luckily, the labels in cPsychQASet provide the known topics, such as *love*, *career* and *family*, etc. We assume that the new problems raised by the users also belong to the existing topics. Therefore, in this work, the topic extraction from the user's inputs can be regarded as a 7-topic classification task. To solve this problem, a bidirectional GRU model with an attention mechanism is proposed, the structure of which is shown in Fig. 9.

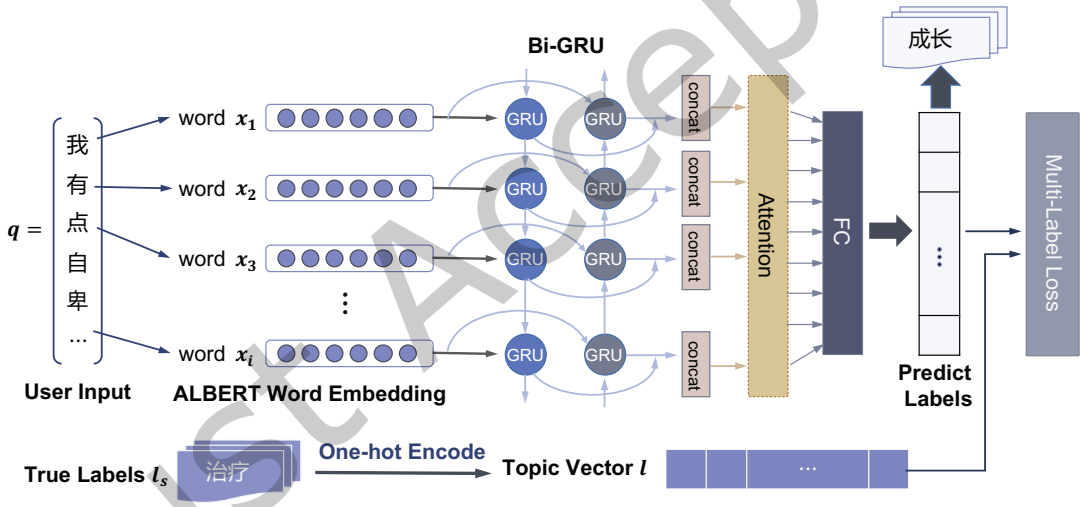

Fig. 9. Illustration of overall architecture of the Topic-aware Module.

The workflow of the topic-aware module is described as follows. First, the topic label string $\boldsymbol{l_s}$ in $(\boldsymbol{p}, \boldsymbol{l_s}, \boldsymbol{r})$ are one-hot encoded into a 7-dimensional vector $\boldsymbol{l} = (l_0, l_1, \cdots, l_6)$. Then the question description $\boldsymbol{p}$ is encoded by ALBERT [17] to extract a set of critical textual features $\boldsymbol{x}_i$ which represents the embedding of the i-th word in $\boldsymbol{p}$. ALBERT is a lightweight BERT variant with similar performance to BERT but less communication overhead. Finally, a BiGRU model with an attention layer is constructed to learn the association between $\boldsymbol{p}$ and $\boldsymbol{l}$. The configuration of the proposed BiGRU model has been summarized in Table 2.

The BiGRU network consists of a forward GRU unit and a backward GRU unit. As shown in Eqs. 1-2, $\boldsymbol{X}_t = [\boldsymbol{x}_{1t}, \boldsymbol{x}_{2t}, \cdots, \boldsymbol{x}_{nt}]$ is the input vector at time t, the hidden layer of the forward GRU unit is expressed as $\overrightarrow{\boldsymbol{h}_t}$ and

Layer Name	Parameter Settings
BiGRU	Hidden:128 Layer: 2 Dropout: 0.5
Attention	
FC	Output feature: 7 Activation: Sigmoid

Table 2. Parameter settings of the BiGRU model with an attention layer.

the hidden layer of the backward GRU unit is expressed as $\overleftarrow{\boldsymbol{h}_t}$. The hidden layer output of BiGRU at time t is concatenated by the hidden layer output of the forward and the backward GRU unit, as shown in Eq. 3.

$$\overrightarrow{\boldsymbol{h}_t} = \text{GRU}(\boldsymbol{X}_t, \overrightarrow{\boldsymbol{h}_{t-1}}) \tag{1}$$

$$\overleftarrow{\boldsymbol{h}_t} = \text{GRU}(\boldsymbol{X}_t, \overleftarrow{\boldsymbol{h}_{t-1}}) \tag{2}$$

$$\boldsymbol{h}_t = [\overrightarrow{\boldsymbol{h}_t}, \overleftarrow{\boldsymbol{h}_t}] \tag{3}$$

The weight coefficients of the attention layer are computed as

$$\begin{aligned} e_{i,t} &= \text{att}(\boldsymbol{s}_{t-1}, \boldsymbol{h}_i) \\ \alpha_{i,t} &= \text{softmax}(e_{i,t}) \\ \boldsymbol{c}_t &= \sum_{i=1}^{n} \alpha_{i,t} \boldsymbol{h}_i \end{aligned} \tag{4}$$

where $\boldsymbol{h}_i$ is the encoded hidden state at position i, $\boldsymbol{s}_{t-1}$ is the previous decoder output at time t, $\boldsymbol{c}_t$ is the final context vector, and $\text{att}(\cdot)$ is the attention weight function.

The model finally outputs a 7-dimensional vector $\boldsymbol{o} = (o_0, o_1, \cdots, o_6)$ and o_i represents the probability that question $\boldsymbol{p}$ belongs to the corresponding topic l_i, $i \in (0, 6)$. The trained BiGRU model with an attention layer will be used to predict the potential topics for the user's input problems. With the predicted topics and contents of the problem, PsyChatbot can retrieve more related psychological advice from cPsychQASet. The details will be introduced in Sec. 3.3.2.

3.3.2 Q&A Matching Module. This module aims to retrieve the most similar question $\boldsymbol{p}^*$ to the user's problem $\boldsymbol{q}$ from cPsychQASet and to take the corresponding answer $\boldsymbol{r}^*$ to $\boldsymbol{p}^*$ as the response of PsyChatbot. Following the STC framework [16] which formalized the conversation generation task as an information retrieval problem, the workflow of the Q&A matching module consists of three steps: *Candidate Retrieval*, *Candidate Matching*, and *Candidate Reranking* (as shown in Fig. 10).

In the *Candidate Retrieval* step, a quick search is adopted to recall a batch of candidate triplets $\{(\boldsymbol{p}_i, \boldsymbol{l}_i, \boldsymbol{r}_i)\}_{1 \le i \le k}$ as follows. Firstly, the user input $\boldsymbol{q}$ is split into words $(w_1, \cdots, w_n)$ by jieba[1]. Secondly, an inverted index structure [57] is constructed to retrieve all the questions $\{\boldsymbol{p}_i\}$ in cPsychQASet that have overlapping words with $\boldsymbol{q}$. The structure establishes a mapping between the words and the questions, which can reduce the time cost of retrieval. Finally, The relevance between $\boldsymbol{q}$ and each question $\boldsymbol{p}_i$ is computed using BM25 algorithm [26]. BM25 algorithm

[1]https://github.com/fxsjy/jieba

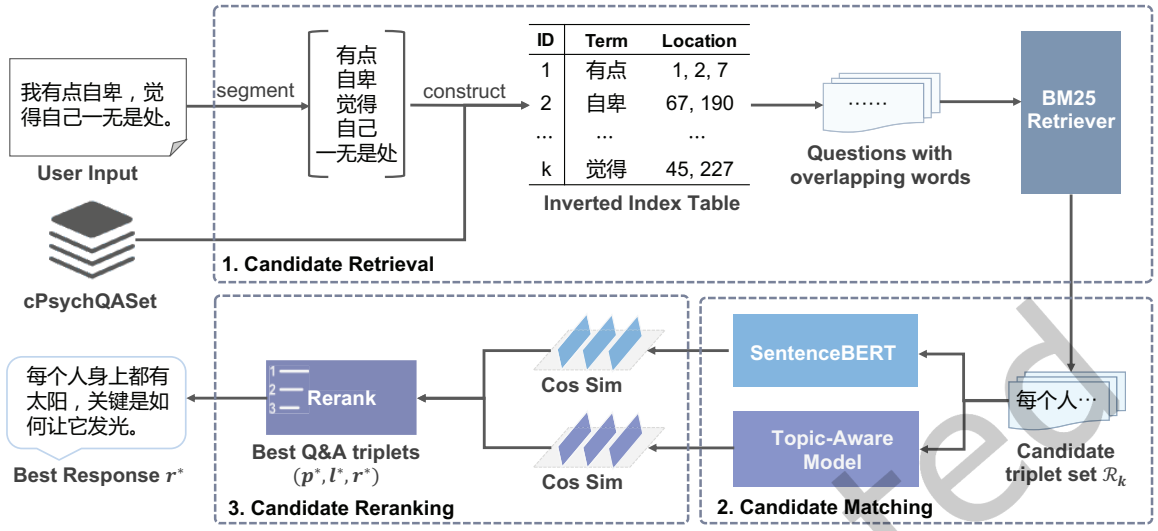

Fig. 10. Illustration of overall architecture of the Q&A Matching Module.

is an information retrieval framework, which is used to estimate the relevance of a document to a given query. It is proposed based on two assumptions:

(1) When the document is represented as a vector, only the existence of a term is considered, not the location. Specifically, when document D is represented as vector $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$, if the i-th word appears in document D, $x_i = 1$; otherwise, $x_i = 0$. In this work, document D corresponds to the question $\boldsymbol{p}_i$ in cPsychQASet and the query corresponds to $\boldsymbol{q}$.

(2) The occurrence of words in a document is independent of each other, i.e., $P(D) = \sum_{i=1}^{n} P(x_i)$. $P(D)$ and $P(x_i)$ are the occurrence probabilities of D and x_i, respectively.

Based on the above two assumptions, the relevance score between $\boldsymbol{p}_i$ and the input $\boldsymbol{q}$ can be computed as

$$\text{Score}_{\text{bm25}}(\boldsymbol{q}, \boldsymbol{p}_i) = \sum_{j=1}^{m} W_j \cdot R(w_j, \boldsymbol{p}_i) \tag{5}$$

where w_j is the j-th word in $\boldsymbol{q}$, W_j is the weight of w_j, and $R(w_j, \boldsymbol{p}_i)$ stands for the relevance of w_j to $\boldsymbol{p}_i$. In this work, W_j is calculated using IDF algorithm (Inverse Document Frequency) [38], which tends to filter out common words and retain important words. It can be computed as

$$W_j = IDF(w_j) = \log \frac{N - count(w_j) + 0.5}{count(w_j) + 0.5} \tag{6}$$

where N is the total number of questions in cPsychQASet, $count(w_j)$ is the total number of questions that contains word w_j. For a given set of questions in cPsychQASet, the more questions that contain word w_j, the lower the weight of w_j. As a result, meaningless words such as "*ah*" and "*umm*" will be filtered out.

$R(w_j, \boldsymbol{p}_i)$ can be computed as

$$R(w_j, \boldsymbol{p}_i) = \frac{f_j \cdot (k_1 + 1)}{f_j + k_1 \cdot (1 - b + \frac{b \cdot L}{L_{avg}})} \tag{7}$$

where f_j is the frequency of w_j occurring in $\boldsymbol{p}_i$, L is the length of $\boldsymbol{p}_i$, and L_{avg} is the average length of all the questions in cPsychQASet. k_1, b are hyperparameters. Based on Eqs. 5-7, we can preliminarily calculate the relevance of question $\boldsymbol{p}_i$ to $\boldsymbol{q}$. Then all the Q&A triplets $\{(\boldsymbol{p}_i, \boldsymbol{l}_i, \boldsymbol{r}_i)\}$ in cPsychQASet are sorted based on $\text{Score}_{\text{bm25}}(\boldsymbol{q}, \boldsymbol{p}_i)$. The top k ($k = 15$) Q&A triplets are selected to construct the candidate triplet set $\mathcal{R}_k = \{(\boldsymbol{p}'_i, \boldsymbol{l}'_i, \boldsymbol{r}'_i)_{1 \le i \le k}\}$.

In the *Candidate Matching* step, a fine-grained evaluation is performed on the candidate triplet set $\mathcal{R}_k$, and the response with the highest matching score will be served as the system response. First, $\boldsymbol{q}$ and the candidate questions $\boldsymbol{p}'_i$ are pair-encoded by the SentenceBERT model [34]. SentenceBERT is a modification of the pretrained BERT network proposed by Nil et al. [34]. SentenceBERT uses siamese and triplet network structures to derive semantically meaningful sentence embeddings. It can deal with the time-consuming problem of BERT model in semantic similarity retrieval scenes and is suitable for sentence similarity calculation. The sentence pairs are fed into two BERT models whose parameters are shared (siamese structure) and the word embeddings of the two sentences are derived from them. Then an average pooling is performed on all the word embeddings and the representations of the two sentences are derived. SentenceBERT uses cosine similarity to evaluate the similarity of the two sentence embeddings $(\boldsymbol{u}, \boldsymbol{v})$ as

$$\text{sim}(\boldsymbol{u}, \boldsymbol{v}) = \frac{\boldsymbol{u} \cdot \boldsymbol{v}}{||\boldsymbol{u}|| \cdot ||\boldsymbol{v}||} \tag{8}$$

PsyChatbot has predicted the topic vector $\boldsymbol{o}$ for the user's input $\boldsymbol{q}$ in the *Topic-Aware Module*. Similarly, the query topic $\boldsymbol{o}$ and the question topics $\boldsymbol{l}'$ are calculated their cosine similarities in this step.

In the *Candidate Reranking* step, we take a weighted sum of the sentence similarity and topic similarity between the user's input $\boldsymbol{q}$ and the candidate question $\boldsymbol{p}'_i$. The weighted sum is regarded as the final matching score of $(\boldsymbol{p}'_i, \boldsymbol{l}'_i, \boldsymbol{r}'_i)$ to $\boldsymbol{q}$. Finally, $\{(\boldsymbol{p}'_i, \boldsymbol{l}'_i, \boldsymbol{r}'_i)_{1 \le i \le k}\}$ are sorted by the matching scores and the response $\boldsymbol{r}^*$ with the highest matching score is chosen as the system response. Through these two modules, PsyChatbot gives comfort and advice to users within the CBT-based dialogue framework, enhancing the diversity and appropriateness of the system responses.

4 Experiments and Results

Based on the proposed CBT-based psychological intervention framework, PsyChatbot can guide users to reflect on themselves and to provide professional psychological advice. The performance of PsyChatbot mainly depends on the effectiveness of the CBT-based guided conversation framework and the appropriateness of retrieved psychological advice. Therefore, in this section, the system will be evaluated from these two aspects.

4.1 Case Study

Fig.11 exhibits two conversation examples from PsyChatbot. In the first case, PsyChatbot started the conversation and asked about user's daily life. When it received the user's inputs "*I just feel exhausted and don't want to do anything*", PsyChatbot asked "*Any more? I am listening.*" in anticipation of revealing more of the user's thoughts. When the user made a negative statement such as "*I feel life is very meaningless*", PsyChatbot called the psychological Q&A retrieval algorithm to retrieve professional suggestions. PsyChatbot reminded the user that he/she should take his/her possible mental problems seriously, adjust the mood state and make some positive changes to his/her life.

In the second case, PsyChatbot guided the user to describe more about her mental state and the problems. When the user said that her boyfriend cheated on her and she was considering whether to forgive him, PsyChatbot analysed the pros and cons of continuing the relationship or breaking up for her. It politely advised her to stop the relationship in time but left the final decision in her hands. When the user said that she couldn't help crying, PsyChatbot comforted her to move on and to wait for the next relationship.

Speaker	Response
PsyChatbot	你好，很高兴你来找我聊天。最近过得怎么样呢？ Hello, I am glad you're here. How's life these days?
User	就那样。 Not so hot.
PsyChatbot	最近发生了什么事情，使你觉得产生这种想法？ What happened recently that made you think this way ?
User	没发生什么事情，就是觉得很疲惫，什么事情都不想干。 Nothing happened. I just feel exhausted and don't want to do anything.
PsyChatbot	还有吗？我在听。 Any more? I am listening.
User	觉得人生很没有意思。 I feel life is very meaningless.
PsyChatbot	这样的状态有多久了呢？这是一种抑郁的状态，提示你生活中存在一些需要改变的地方。它也是一种实实在在的问题，就像感冒发烧一样客观，它影响你看问题的角度，看什么都悲观，所以要高度重视，及时调整。 How long has this state been going on? It is a state of depression which reminds you that some changes should be made in your life. It is also a real problem which affects the way you look at things and makes you pessimistic about everything. It is just as objective as having a cold or fever. So you should take it seriously and adjust your state in time.

Speaker	Response
PsyChatbot	别来无恙？ Long time no see. Is everything alright?
User	很不好。 Terrible.
PsyChatbot	最近有什么事情发生吗？ Did anything happen recently?
User	男朋友出轨了，我很难过。 My boyfriend treated on me. I feel depressed.
PsyChatbot	你能详细描述一下吗？我想多多了解你。 Can you describe it in detail? I want to know more about you.
User	我发现我的男朋友在和别人聊骚，我不知道该不该原谅他，我因此每天都过得很痛苦。 I found out that my boyfriend was flirting with someone else. I don' know whether I should forgive him and I suffer every day because of it.
PsyChatbot	你有两个选择，继续和他交往，然后有可能再被劈腿，意识到他根本不值得你喜欢。他再回来找你，你告诉他，人走茶凉了。第二个选择，就是直接拒绝他，因为你根本无法容忍他的劈腿。与其冒着在一次被他伤害的风险，倒不如就此打住。哪个你觉得更能接受就选择哪个。我无法帮你选择，因为是你在恋爱。 You have two choices: one is to continue to date him, risk getting cheated on again, and finally realize that he's not worthy of your affection at all. When he comes back to you, you tell him that the relationship has come to an end. The other is to reject him directly because you can't tolerate his cheating. Instead of risking being hurt by him again, stop there. Choose the one you are more comfortable with. I can't choose it for you, because you are the one in love.
User	或许你是对的。但每当我想到他的背叛和曾经的爱，我还是会忍不住流泪。 Maybe you're right. But every time I think of his betrayal and the love I once had, I can't help crying.
PsyChatbot	很多人一生都不会只爱一个人，而每一个爱人都会心中留有一个固定的位置，把这份感情放在那里就好，偶尔想起也会有甜蜜和痛苦，这样很好。但进入下一段感情时，希望已经处理好了上一段感情带给你的伤害，不要将对前任的感情移植到现任上，而是以全新的感觉认知新人，培养你和现任的感情，祝你幸福！ Many people will not love just one person in their lifetime, and every lover will have a fixed place in their heart. It's good to keep the feeling there, and occasionally recall the sweetness and pain that come with it. As you move into your next relationship, you should have dealt with the hurt that the last relationship brought you. Do not transplant your feelings for your ex onto your current partner, instead nurture the relationship with your current one by perceiving the new person with a new sense of purpose. Best wishes for your happiness!

Fig. 11. Two conversation examples of interacting with PsyChatbot.

The above two cases demonstrate that our psychological intervention framework can effectively simulate the CBT treatment process. PsyChatbot leads the conversation using predefined response templates. It first guides the users to describe their moods and situations by questioning, such as "*Can you describe your mood in detail?*" and "*How's life these days?* ". Then it guides the users to describe their frustration and thoughts (e.g., "*What's the cause, do you think?*"), leading them to summarise the items that trigger their negative emotions. Finally, PsyChatbot offers suggestions on the problems proposed by the users. These suggestions are retrieved from cPsychQASet which were provided by the professional therapists. Therefore, the suggestions are helpful to the users who are stuck in the similar situations and can help them correct their misconceptions.

4.2 Subjective Evaluation

We conducted a subjective evaluation experiment to evaluate the appropriateness of the retrieved psychological advice. Based on the same set of the input problems, the psychological advice was retrieved by five algorithms which include C-TFIDF, R-TFIDF, BM25, Dual Encoder, and the proposed topic-aware psychological Q&A retrieval algorithm, respectively. The appropriateness of the retrieved advice are evaluated manually by the volunteers.

4.2.1 Baseline Models.

- **C-TFIDF and R-TFIDF.** TF-IDF[23] is a statistics-based information retrieval algorithm which is often used as a measurement of the relevance between a document D and a user query $\boldsymbol{q}$. In this work, $\boldsymbol{q}$ is the user input and D is the question in cPsychQASet. The relevance between $\boldsymbol{q}$ and D can be computed with Eq. 9:

$$\text{tf-idf}(\boldsymbol{q}, D) = \sum_{i=1}^{n} \text{tf}(q_i, D) * \log \frac{N}{df_i + 1} \tag{9}$$

  where $\text{tf}(q_i, D)$ indicates the number of times word q_i occurring in document D, N is the number of documents in the corpus, df_i is the number of times word q_i occurring in the corpus. Similar to [21], we set up two baseline models related with TF-IDF. The first one is C-TFIDF which calculates the relevance between the user input and the question in cPsychQASet. The second one is R-TFIDF which calculates the relevance between the user input and the answer in cPsychQASet.
- **BM25.** BM25[26] is a classical information retrieval algorithm. It is also the key component of our topic-aware psychological Q&A retrieval module. Therefore, we also set it as a baseline model. The algorithm is implemented using Eqs. 5-7 in Sec. 3.3.2.
- **Dual Encoder.** Dual Encoder (DE) [23] is an RNN-based dialogue retrieval algorithm. It consists of two RNNs which compute the representation of the input problem $\boldsymbol{q}$ and the response $\boldsymbol{r}$ in cPsychQASet respectively. The model calculates the probability that the given response $\boldsymbol{r}$ is the true response of $\boldsymbol{q}$:

$$P(\boldsymbol{r} \text{ is true response} | \boldsymbol{q}, \boldsymbol{r}; \boldsymbol{M}) = \sigma(\boldsymbol{q}^T \boldsymbol{M} r + \boldsymbol{b}) \tag{10}$$

  where $\boldsymbol{M}$ is a matrix of parameters and $\boldsymbol{b}$ is a bias, both of which are learned from the training data.

4.2.2 Human Rating. First of all, 30 questions related with various mental problems were randomly selected from an independent psychological counseling website[2]. Then, the five models mentioned above retrieved their corresponding answers from cPsychQASet so that each models corresponded to 30 questions and 30 answers. Finally, 20 volunteers were invited to evaluate the responses of the five models.

The appropriateness of the retrieved psychological advice is considered from two aspects: the question relevance and the advice helpfulness. For each retrieved psychological advice, the volunteers were asked to rate its appropriateness to the corresponding question on a scale of 1 to 5 by considering the above two aspects:

- **+5**: Excellent. The response is very helpful and conforms to the question topic well.
- **+4**: Good. The response copes with the topic and is helpful to some extent.
- **+3**: Fair. It is hard to determine whether the response is useful or in line with the question content.
- **+2**: Poor. The response is a little relevant to the question but without much help.
- **+1**: Bad. The response has nothing related to the question and is unhelpful.

[2]https://www.jiandanxinli.com/questions

Models	+5	+4	+3	+2	+1	Average
Topic-Aware	**36.9%**	**28.2%**	13.1%	9.8%	12.0%	**3.68**
BM25	21.1%	17.1%	**14.0%**	13.1%	34.7%	2.77
C-TFIDF	14.2%	16.2%	9.8%	13.1%	46.7%	2.38
R-TFIDF	14.2%	12.2%	10.9%	**15.6%**	47.1%	2.31
DE	8.7%	14.9%	9.1%	14.0%	**53.3%**	2.12

Table 3. Subjective evaluation results. Grade "+1" to grade "+5" shows the score percentage of the models, and the models with the largest proportion are in bold.

4.2.3 Evaluation Results. Table 3 shows the human rating results over 150 retrieved psychological suggestions. The proposed topic-aware psychological Q&A retrieval algorithm received 600 ratings, of which **36.9%** are "+5" grade and **28.2%** are "+4" grade. It has the highest percentage of "+5", which exceeds the second place by **15.8%**. Its percentage of "+4" is also the highest and exceeds the second by **11.1%**. BM25 has the second highest performance. Its average score is 0.91 points lower than our topic-aware Q&A retrieval algorithm, which reveals that our algorithm is effective in capturing latent information of users' questions. C-TFIDF performs slightly better than R-TFIDF, which indicates that the question-based matching works better than the response-based matching in Q&A retrieval problem. DE shows the worst performance. Over half of its responses (53.3%) were rated as "+1". It may due to the fact that the performance of neural network models like DE heavily depends on the scale of data. cPsychQAset is not large enough to for it learning the intrinsic response rules. In addition, the average score of our topic-aware psychological Q&A retrieval algorithm is **3.68**, which closes to "Good" level. It means that in most cases, the proposed model can provide helpful psychological advice to users. By contrast, none of the average rating scores of other models are above 3.0, indicating that the performances of all the baseline models do not even reach the "Fair" level. Therefore, the answers selected by the topic-aware psychological Q&A retrieval algorithm are more relevant to the questions.

Question	我们真的能得到一个人无条件的爱吗？ Can we really have someone's unconditional love?
Topic-Aware	爱是自由的，只是你需要承担选择爱的代价，可能他会爱上你，也会不爱你。因为爱与被爱正是自由，所以未知。勇敢面对才有被爱的可能。 Love is free, only you need to bear the cost of choosing to love. He may or may not fall in love with you because loving and being loved is precisely freedom, it is unknown. It is the courage to face it that makes it possible to be loved.
BM25	无条件接纳是反人性的。而且应该是失控的。 Unconditional acceptance is anti-human. And it should be out of control.
C-TFIDF	"我平常也是个爱说话的人" "我觉得爱说话的人不够靠谱"难道……？真…真的吗!?你的疑虑更多指向TA还是自己呢？ "I'm usually a talker too" "I don't think talkers are reliable enough" Could it be .....? Really...really! Do your doubts point more towards TA or yourself?
R-TFIDF	无条件接纳是反人性的。而且应该是失控的。 Unconditional acceptance is anti-human. And it should be out of control.
DE	具体来说，让你受够了的内容是什么呢？ Specifically, What is the content that makes you fed up?

Fig. 12. An example of responses retrieved by five psychological advice retrieval models.

Fig. 12 shows an example of the responses retrieved by five different models. The user raised a question "*Can we really have someone's unconditional love?*". Among five models, only the topic-aware psychological Q&A algorithm answered the question properly around the topic of *love*. It explained to the user that love was free and encouraged them to be brave to pursue love. BM25 and R-TFIDF algorithms capture partial keywords (i.e. "*unconditional*"). Their responses were somewhat relevant to the question but were not related to the topic of love. The responses of C-TFIDF and DE were completely off-topic and could not be regarded as the proper answers to the question.

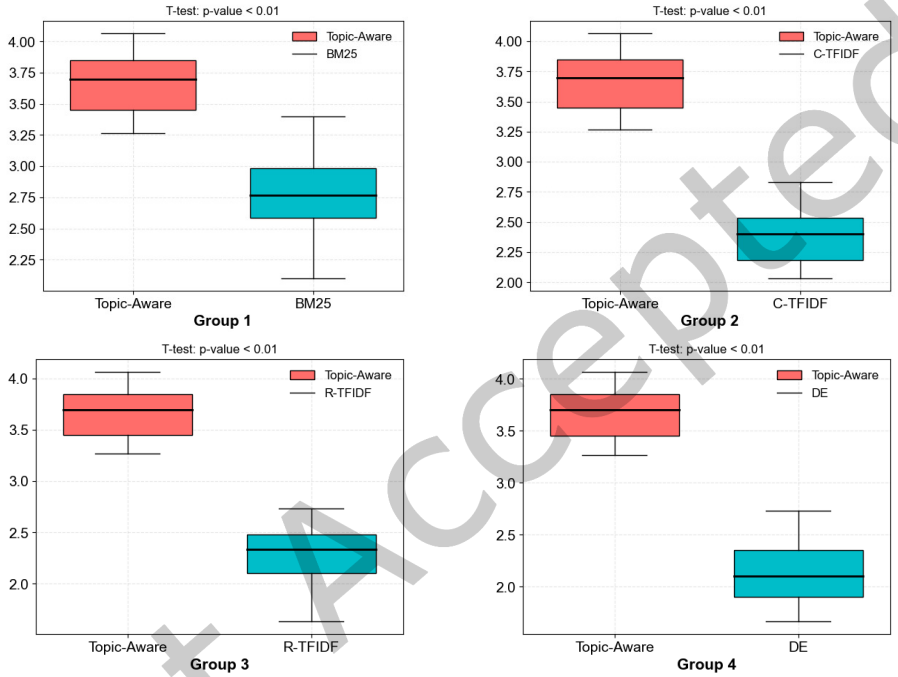

Fig. 13. Boxplots of subjective evaluation results for the Topic-Aware retrieval algorithm and baseline models. The line in the box represents the median, the interquartile range box represents the middle 50% of the data and the whiskers represent the ranges for the bottom 25% and the top 25% of the evaluation results.

To confirm the validity of statistics, we conducted one-tailed T-test between the topic-aware psychological Q&A retrieval algorithm and the baseline models. The results indicates that the improvement of PsyChatbot is statistically significant (p-value <0.01). Fig. 13 illustrates the distribution of the subjective evaluation results for four groups of compared models. It can be seen that, the median line and the box plot of the topic-aware psychological Q&A retrieval algorithm are both much higher than other baseline models, indicating that the responses of the topic-aware psychological Q&A retrieval algorithm are rated significantly higher than other models. The topic-aware psychological Q&A retrieval algorithm and C-TFIDF have a concentrated distribution of scores, indicating that their scores are consistent among volunteers. In contrast, BM25 has a scattered distribution of scores, indicating that volunteers hold different views of BM25.

Metrics	Topic-Aware	BM25	C-TFIDF	R-TFIDF	DE
R@1	0.88	0.8	0.74	0.54	0.6
R@2	0.96	0.98	0.88	0.62	0.64
R@5	0.98	1	0.9	0.64	0.9

Table 4. Results on automatic metrics.

4.3 Objective Evaluation

We use the Recall@k [21] (denoted by R@1, R@2, R@5 below) to measure the performance of the topic-aware Q&A retrieval algorithm and the baseline models. Recall@k is the percentage of the top k relevant results selected out of all the relevant results in the corpus. It is one of the critical metrics to evaluate the performance of dialogue retrieval algorithms. We constructed 50 questions with one true response and nine responses randomly selected from cPsychQASet. The models computed the matching scores of the 10 responses and ranked them according to their scores. Table 4 gives the results of automatic metrics, i.e. R@1, R@2, and R@5. The topic-aware Q&A retrieval algorithm performs best on the R@1 metric, with a recall of **88%**. It indicates that the algorithm has an 88% probability of selecting the correct answer from ten candidate responses. On the evaluation metrics of R@2 and R@5, the topic-aware Q&A retrieval algorithm and BM25 show similar performances which are significantly higher than the other three models. For example, on R@2 metric, C-TFIDF has a recall of only 0.88, which is almost 0.1 lower than the topic-aware Q&A retrieval algorithm and BM25. In practice, we pay more attention to the model performance on R@1 metric in the dialog retrieval scenarios. Therefore, the topic-aware Q&A retrieval algorithm outperforms other baseline models based on the evaluation results.

5 Limitations and Future Work

In this work, we design and implement a Chinese psychological counseling agent, namely PsyChatbot, which guides users to engage in self-reflection and offers psychological advice for them. Nevertheless, our work exhibits certain limitations. Firstly, although PsyChatbot employs a rule-based framework to facilitate users' engagement in counseling dialogues, it struggles to handle complex dialogue scenarios and sometimes mismatches templates. Secondly, PsyChatbot should improve the accuracy of its Q&A retrieval algorithm. The algorithm fails to leverage historical dialogue information, which may result in poor retrieval performance. In addition, PsyChatbot can only provide suggestions derived from existing answers in the cPsychQASet corpus, lacking the flexibility to generate proper and unrestricted suggestions based on user inputs.

Based on the aforementioned limitations, we propose several directions for further improvement of PsyChatbot in future work. Firstly, the CBT-based guided framework of PsyChatbot can be implemented by generative algorithms instead of rule/retrieval-based algorithms, aiming to produce more flexible responses closely associated with user inputs. For example, this framework can be built upon large language models (LLMs) as well as techniques like prompt engineering and in-context learning, as demonstrated in previous works [37, 51]. Secondly, the Q&A retrieval module of PsyChatbot should consider more dialogue histories to provide contextually relevant psychological advice. In addition, research in related fields, such as emotion recognition [31, 33] and cause extraction [19, 32] in conversations, can contribute to the effectiveness of PsyChatbot. The recognized emotions and causes can assist counseling systems like PsyChatbot in clarifying users' problems and addressing their maladaptive thoughts. Lastly, the integration of multimodality into dialogue systems is an emerging trend that can significantly improve the understanding of users' emotions and intentions [10, 41, 42, 56]. Multimodal dialogue systems not only analyze textual inputs but also interpret non-verbal information such as speech, facial expressions, and gestures, thereby enhancing the effectiveness and appropriateness of interactions. Therefore,

exploring the integration of multimodality into counseling dialogue systems is a promising area for further investigation.

In the future, we plan to deploy this counseling agent PsyChatbot on a self-developed mental health platform, with the goal of providing assistance to users and further evaluating the performance of PsyChatbot. Additionally, we will establish a feedback channel on the platform, allowing users to provide feedback and ratings for PsyChatbot.

6 Conclusions

Traditional face-to-face psychotherapy has proved its effectiveness in the psychological intervention. However, factors such as insufficient regional psychological resources and social stigma prevent Chinese people from obtaining effective treatments. To solve the problems mentioned above, this work develops a Chinese psychological counselling agent named PsyChatbot to provide counselling services for the depressed users. PsyChatbot leads the conversation for the psychological intervention under CBT framework and offers psychological suggestions to users' specific problems. The main contributions of this work can be summarized as:

- We build a CBT-based psychological intervention framework with AIML. The intervention framework simulates the rapid CBT method in clinical psychotherapy, guides users to reflect on themselves, discovers users' negative perceptions and reshapes them.
- We construct a Chinese psychological counselling Q&A corpus, namely cPsychQASet. To construct cPsychQASet, nearly 160,000 psychological question-answer pairs are collected and finally nearly 89,000 valid Q&A triples are obtained after data cleaning. cPsychQASet can provide professional psychological advice for users' specific mental problems.
- We implement a topic-aware Q&A retrieval algorithm. The algorithm first uses a BiGRU model with an attention layer to predict the potential topics for the user's problem. Then it retrieves the suitable response from cPsychQASet based on both the content similarity and the topic similarity between the user's problem and the problem stored in cPsychQASet.

Online psychological intervention system is of great importance for the depressed people. In this work, we design and implement a Chinese psychological counselling agent, PsyChatbot, which automatically guides users to reflect on themselves and provides psychological advice for their specific problems. The experimental results indicate the effectiveness and helpfulness of PsyChatbot in the psychological intervention. In practice, it can comfort users, relieve their pressure, and improve their life quality.

Ethical Considerations

First of all, all the raw question-answer pairs are collected from the publicly available counselling platforms. cPsychQASet only keeps the question contents, the question labels and the corresponding best responses. It does not contain personal information about the questioners and the answerers. Besides, we have cleaned the sensitive information which may appear in the questions and the answers, such as nicknames, phone numbers, links, etc. What's more, the counselling platforms themselves provide privacy protection mechanisms for the users. Users can choose to ask questions anonymously when they are reluctant to reveal their identity. Most users asked questions anonymously, with only a few of them displaying their nicknames. In the end, cPsychQASet can only be used for academic purposes, which means that data users should sign an End User License Agreement before using cPsychQASet.

References

[1] Gerhard Andersson and Pim Cuijpers. 2009. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. *Cognitive behaviour therapy* 38, 4 (2009), 196–205. https://doi.org/10.1080/16506070903318960

[2] American Psychiatric Association. 2013. *Diagnostic and statistical manual of mental disorders* (5th ed.). American Psychiatric Association, Washington, D.C.
[3] David Bakker, Nikolaos Kazantzis, Debra Rickwood, Nikki Rickard, et al. 2016. Mental health smartphone apps: review and evidence-based recommendations for future developments. *JMIR mental health* 3, 1 (2016), e7. https://doi.org/10.2196/mental.4984
[4] Azy Barak, Liat Hen, Meyran Boniel-Nissim, and Na'ama Shapira. 2008. A comprehensive review and a meta-analysis of the effectiveness of internet-based psychotherapeutic interventions. *Journal of Technology in Human services* 26, 2-4 (2008), 109–160. https://doi.org/10.1080/15228830802094429
[5] Clare C Beatty, Tanya Malik, S Meheli, and Chaitali Sinha. 2022. Evaluating the Therapeutic Alliance With a Free-Text CBT Conversational Agent (Wysa): A Mixed-Methods Study. *Frontiers in Digital Health* 4 (2022), 847991. https://doi.org/10.3389/fdgth.2022.847991
[6] Judith S Beck and JH Wright. 1997. Cognitive therapy: Basics and beyond. *J. Psychother. Pract. Res* 6 (1997), 71–80.
[7] BetterHelp. 2022. *Affordable, private therapy with BetterHelp anytime, anywhere.* Retrieved March 22, 2022 from https://www.betterhelp.com/
[8] Roy Bhakta, Maggi Savin-Baden, and Gemma Tombs. 2014. Sharing Secrets with Robots. In *Proceedings of EdMedia 2014–World Conference on Educational Media and Technology*. AACE, Tampere, Finland, 2295–2301. Retrieved March 22, 2022 from https://www.learntechlib.org/primary/p/147797/
[9] Timothy W. Bickmore, Suzanne E. Mitchell, Brian W. Jack, Michael K. Paasche-Orlow, Laura M. Pfeifer, and Julie O'Donnell. 2010. Response to a relational agent by hospital patients with depressive symptoms. *Interacting with computers* 22, 4 (2010), 289–298. https://doi.org/10.1016/j.intcom.2009.12.001
[10] Elisabetta Fersini, Francesca Gasparini, Giulia Rizzi, Aurora Saibene, Berta Chulvi, Paolo Rosso, Alyssa Lees, and Jeffrey Sorensen. 2022. SemEval-2022 Task 5: Multimedia Automatic Misogyny Identification. In *Proceedings of the 16th International Workshop on Semantic Evaluation (SemEval-2022)*. Association for Computational Linguistics, Seattle, United States, 533–549. https://doi.org/10.18653/v1/2022.semeval-1.74
[11] Kathleen Kara Fitzpatrick, Alison M Darcy, and Molly Vierhile. 2017. Delivering Cognitive Behavior Therapy to Young Adults With Symptoms of Depression and Anxiety Using a Fully Automated Conversational Agent (Woebot): A Randomized Controlled Trial. *JMIR Mental Health* 4, 2 (2017), e19. https://doi.org/10.2196/mental.7785
[12] Paula M Gardiner, Kelly D McCue, Lily M Negash, Teresa Cheng, Laura F White, Leanne Yinusa-Nyahkoon, Brian W Jack, and Timothy W Bickmore. 2017. Engaging women with an embodied conversational agent to deliver mindfulness and lifestyle recommendations: A feasibility randomized control trial. *Patient education and counseling* 100, 9 (2017), 1720–1729. https://doi.org/10.1016/j.pec.2017.04.015
[13] Haihong Huang. 2010. *Economic Burden of Depression: an empirical study based on inpatients from on tertiary mental health hospital in Shanghai.* Master's thesis. Fudan University.
[14] Becky Inkster, Shubhankar Sarda, and Vinod Subramanian. 2018. An Empathy-Driven, Conversational Artificial Intelligence Agent (Wysa) for Digital Mental Well-Being: Real-World Data Evaluation Mixed-Methods Study. *JMIR mHealth and uHealth* 6, 11 (2018), e12106. https://doi.org/10.2196/12106
[15] A. Cully Jeffrey and L. Teten Andra. 2008. *A Therapist's Guide to Brief Cognitive Behavioral Therapy.* MIRECC, ebook. https://depts.washington.edu/dbpeds/therapists_guide_to_brief_cbtmanual.pdf
[16] Zongcheng Ji, Zhengdong Lu, and Hang Li. 2014. An Information Retrieval Approach to Short Text Conversation. *ArXiv* abs/1408.6988 (2014). https://doi.org/10.48550/arXiv.1408.6988
[17] Zhenzhong Lan, Mingda Chen, Sebastian Goodman, Kevin Gimpel, Piyush Sharma, and Radu Soricut. 2020. ALBERT: A Lite BERT for Self-supervised Learning of Language Representations. In *International Conference on Learning Representations*. OpenReview.net, Addis Ababa, Ethiopia. https://doi.org/10.48550/arXiv.1909.11942
[18] Jean-Pierre Lépine and Mike Briley. 2011. The increasing burden of depression. *Neuropsychiatric disease and treatment* 7, Suppl 1 (2011), 3–7. https://doi.org/10.2147/NDT.S19617
[19] Wei Li, Yang Li, Vlad Pandelea, Mengshi Ge, Luyao Zhu, and Erik Cambria. 2023. ECPEC: Emotion-Cause Pair Extraction in Conversations. *IEEE Transactions on Affective Computing* 14, 3 (2023), 1754–1765. https://doi.org/10.1109/TAFFC.2022.3216551
[20] Yanxiang Ling, Fei Cai, Xuejun Hu, Jun Liu, Wanyu Chen, and Honghui Chen. 2021. Context-Controlled Topic-Aware Neural Response Generation for Open-Domain Dialog Systems. *Information Processing & Management* 58, 1 (2021), 102392. https://doi.org/10.1016/j.ipm.2020.102392
[21] Chia-Wei Liu, Ryan Lowe, Iulian Serban, Mike Noseworthy, Laurent Charlin, and Joelle Pineau. 2016. How NOT To Evaluate Your Dialogue System: An Empirical Study of Unsupervised Evaluation Metrics for Dialogue Response Generation. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*. ACL, Austin, Texas, 2122–2132. https://doi.org/10.18653/v1/D16-1230
[22] Ryan Lowe, Nissan Pow, Iulian Serban, and Joelle Pineau. 2015. The Ubuntu Dialogue Corpus: A Large Dataset for Research in Unstructured Multi-Turn Dialogue Systems. In *Proceedings of the 16th Annual Meeting of the Special Interest Group on Discourse and Dialogue*. ACL, Prague, Czech Republic, 285–294. https://doi.org/10.18653/v1/W15-4640
[23] Ryan Lowe, Nissan Pow, Iulian V Serban, and Joelle Pineau. 2015. The Ubuntu Dialogue Corpus: A Large Dataset for Research in Unstructured Multi-Turn Dialogue Systems. In *SIGDIAL Conference*. ACL, Prague, Czech Republic, 285–294. https://doi.org/10.18653/v1/

w15-4640
[24] Jin Lu, Xiufeng Xu, Yueqin Huang, Tao Li, Chao Ma, Guangming Xu, Huifang Yin, Xiangdong Xu, Yanjuan Ma, Limin Wang, et al. 2021. Prevalence of depressive disorders and treatment in China: a cross-sectional epidemiological study. *The Lancet Psychiatry* 8, 11 (2021), 981–990. https://doi.org/10.1016/S2215-0366(21)00251-0
[25] Kien Hoa Ly, Ann-Marie Ly, and Gerhard Andersson. 2017. A fully automated conversational agent for promoting mental well-being: A pilot RCT using mixed methods. *Internet interventions* 10, C (2017), 39–46. https://doi.org/10.1016/j.invent.2017.10.002
[26] Christopher D. Manning, Prabhakar Raghavan, and Hinrich Schütze. 2008. *Introduction to Information Retrieval.* Cambridge University Press, USA. https://doi.org/10.5555/1394399
[27] National Bureau of Statistics of China. 2022. *Households' Income and Consumption Expenditure in 2021.* National Bureau of Statistics of China. Retrieved March 22, 2022 from http://www.stats.gov.cn/english/PressRelease/202201/t20220118_1826649.html
[28] World Health Organization. 2021. *Mental health atlas 2020.* World Health Organization, Geneva.
[29] Agata Pasikowska, Abolfazl Zaraki, and Nicole Lazzeri. 2013. A Dialogue with a Virtual Imaginary Interlocutor as a Form of a Psychological Support for Well-Being. In *Proceedings of the International Conference on Multimedia, Interaction, Design and Innovation (MIDI '13, 16).* ACM, New York, USA, 1–15. https://doi.org/10.1145/2500342.2500359
[30] Michael R Phillips, Gonghuan Yang, Yanping Zhang, Lijun Wang, Huiyu Ji, and Maigeng Zhou. 2002. Risk factors for suicide in China: a national case-control psychological autopsy study. *The Lancet* 360, 9347 (2002), 1728–1736. https://doi.org/10.1016/S0140-6736(02)11681-3
[31] Soujanya Poria, Devamanyu Hazarika, Navonil Majumder, and Rada Mihalcea. 2023. Beneath the Tip of the Iceberg: Current Challenges and New Directions in Sentiment Analysis Research. *IEEE Transactions on Affective Computing* 14, 1 (2023), 108–132. https://doi.org/10.1109/TAFFC.2020.3038167
[32] Soujanya Poria, Navonil Majumder, Devamanyu Hazarika, Deepanway Ghosal, Rishabh Bhardwaj, Samson Yu, Pengfei Hong, Romila Ghosh, Niyati Chhaya, Alexander F. Gelbukh, and Rada Mihalcea. 2020. Recognizing Emotion Cause in Conversations. *Cognitive Computation* 13 (2020), 1317 – 1332. https://doi.org/10.1007/s12559-021-09925-7
[33] Soujanya Poria, Navonil Majumder, Rada Mihalcea, and Eduard Hovy. 2019. Emotion Recognition in Conversation: Research Challenges, Datasets, and Recent Advances. *IEEE Access* 7 (2019), 100943–100953. https://doi.org/10.1109/ACCESS.2019.2929050
[34] Nils Reimers and Iryna Gurevych. 2019. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing.* ACL, Hong Kong, China, 3982–3992. https://doi.org/10.18653/v1/D19-1410
[35] Zhihong Ren, Xianyun Li, Lingbo Zhao, Xianglian Yu, Zhenghan Li, Lizu Lai, Yijun Ruan, and Guangrong Jiang. 2016. Effectiveness and mechanism of internet-based self-help intervention for depression: The Chinese version of MoodGYM. *Acta Psychologica Sinica* 48, 7 (2016), 818–832. https://doi.org/10.3724/SP.J.1041.2016.00818
[36] Albert A Rizzo, Belinda Lange, John G Buckwalter, Eric Forbell, Julia Kim, Kenji Sagae, Josh Williams, Barbara O Rothbaum, JoAnn Difede, Greg Reger, et al. 2011. An intelligent virtual human system for providing healthcare information and support. *Studies in health technology and informatics* 163, 4 (2011), 503–509. https://doi.org/10.1515/IJDHD.2011.046
[37] Sahand Sabour, Wen Zhang, Xiyao Xiao, Yuwei Zhang, Yinhe Zheng, Jiaxin Wen, Jialu Zhao, and Minlie Huang. 2023. A chatbot for mental health support: exploring the impact of Emohaa on reducing mental distress in China. *Frontiers in Digital Health* 5 (2023), 1133987. https://doi.org/10.3389/fdgth.2023.1133987
[38] Gerard Salton and Christopher Buckley. 1988. Term-weighting approaches in automatic text retrieval. *Information Processing and Management* 24, 5 (1988), 513–523. https://doi.org/10.1016/0306-4573(88)90021-0
[39] Shekhar Saxena, Graham Thornicroft, Martin Knapp, and Harvey Whiteford. 2007. Resources for mental health: scarcity, inequity, and inefficiency. *The lancet* 370, 9590 (2007), 878–889. https://doi.org/10.1016/S0140-6736(07)61239-2
[40] Iulian Serban, Alessandro Sordoni, Yoshua Bengio, Aaron Courville, and Joelle Pineau. 2016. Building end-to-end dialogue systems using generative hierarchical neural network models. In *Proceedings of the AAAI Conference on Artificial Intelligence.* AAAI Press, Phoenix, Arizona, 3776–3783. https://doi.org/10.5555/3016387.3016435
[41] Chhavi Sharma, Deepesh Bhageria, William Scott, Srinivas PYKL, Amitava Das, Tanmoy Chakraborty, Viswanath Pulabaigari, and Björn Gambäck. 2020. SemEval-2020 Task 8: Memotion Analysis- the Visuo-Lingual Metaphor!. In *Proceedings of the Fourteenth Workshop on Semantic Evaluation.* International Committee for Computational Linguistics, Barcelona (online), 759–773. https://doi.org/10.18653/v1/2020.semeval-1.99
[42] Aakash Singh, Deepawali Sharma, and Vivek Kumar Singh. 2024. MIMIC: Misogyny Identification in Multimodal Internet Content in Hindi-English Code-Mixed Language. *ACM Trans. Asian Low-Resour. Lang. Inf. Process.* (2024). https://doi.org/10.1145/3656169
[43] Viola Spek, PIM Cuijpers, Ivan Nyklíček, Heleen Riper, Jules Keyzer, and Victor Pop. 2007. Internet-based cognitive behaviour therapy for symptoms of depression and anxiety: a meta-analysis. *Psychological medicine* 37, 3 (2007), 319–328. https://doi.org/10.1017/S0033291706008944
[44] Talkspace. 2022. *Feeling better starts with a single message.* Retrieved March 22, 2022 from https://www.talkspace.com/
[45] Barbra Teater. 2019. *An Introduction to Applying Social Work Theories and Methods 3e.* McGraw-Hill Education (UK), London, UK.

[46] John B Torous, Steven Richard Chan, Shih Yee-Marie Tan, Jacob Behrens, Ian Mathew, Erich J Conrad, Ladson Hinton, Peter M. Yellowlees, and Matcheri S. Keshavan. 2014. Patient Smartphone Ownership and Interest in Mobile Apps to Monitor Symptoms of Mental Health Conditions: A Survey in Four Geographically Distinct Psychiatric Clinics. *JMIR Mental Health* 1, 1 (2014), e5. https://doi.org/10.2196/mental.4004

[47] Oriol Vinyals and Quoc Le. 2015. A Neural Conversational Model. *CoRR* abs/1506.05869 (2015). https://doi.org/10.48550/arXiv.1506.05869

[48] Richard S. Wallace. 2009. *The Anatomy of A.L.I.C.E.* Springer Netherlands, Dordrecht, 181–210. https://doi.org/10.1007/978-1-4020-6710-5_1

[49] Joseph Weizenbaum. 1966. ELIZA—a Computer Program for the Study of Natural Language Communication between Man and Machine. *Commun. ACM* 9, 1 (1966), 36–45. https://doi.org/10.1145/365153.365168

[50] Yu Wu, Wei Wu, Chen Xing, Ming Zhou, and Zhoujun Li. 2017. Sequential Matching Network: A New Architecture for Multi-turn Response Selection in Retrieval-Based Chatbots. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*, Vol. 1. ACL, Vancouver, Canada, 496–505. https://doi.org/10.18653/v1/P17-1046

[51] Mengxi Xiao, Qianqian Xie, Ziyan Kuang, Zhicheng Liu, Kailai Yang, Min Peng, Weiguang Han, and Jimin Huang. 2024. HealMe: Harnessing Cognitive Reframing in Large Language Models for Psychotherapy. arXiv:2403.05574 [cs.HC]

[52] Chen Xing, Wei Wu, Yu Wu, Jie Liu, Yalou Huang, Ming Zhou, and Wei-Ying Ma. 2017. Topic Aware Neural Response Generation. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*. AAAI Press, San Francisco, USA, 3351–3357. https://doi.org/10.5555/3298023.3298055

[53] Shoji Yokoya, Takami Maeno, Naoto Sakamoto, Ryohei Goto, and Tetsuhiro Maeno. 2018. A brief survey of public knowledge and stigma towards depression. *Journal of clinical medicine research* 10, 3 (2018), 202–209. https://doi.org/10.14740/jocmr3282w

[54] Wayne Xin Zhao, Jing Jiang, Jianshu Weng, Jing He, Ee-Peng Lim, Hongfei Yan, and Xiaoming Li. 2011. Comparing Twitter and Traditional Media Using Topic Models. In *Advances in Information Retrieval*, Vol. 6611. Springer Berlin Heidelberg, Berlin, Heidelberg, 338–349. https://doi.org/10.1007/978-3-642-20161-5_34

[55] Shijie Zhou, Wenge Rong, Jianfei Zhang, Yanmeng Wang, Libin Shi, and Zhang Xiong. 2021. Topic-Aware Dialogue Generation with Two-Hop Based Graph Attention. In *2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, Toronto, Canada, 7428–7432. https://doi.org/10.1109/ICASSP39728.2021.9414472

[56] Zhouan Zhu, Chenguang Li, Jicai Pan, Xin Li, Yufei Xiao, Yanan Chang, Feiyi Zheng, and Shangfei Wang. 2023. MEDIC: A Multimodal Empathy Dataset in Counseling. In *Proceedings of the 31st ACM International Conference on Multimedia*. Association for Computing Machinery, New York, NY, USA, 6054–6062. https://doi.org/10.1145/3581783.3612346

[57] Justin Zobel, Alistair Moffat, and Kotagiri Ramamohanarao. 1998. Inverted files versus signature files for text indexing. *ACM Transactions on Database Systems* 23, 4 (1998), 453–490. https://doi.org/10.1145/296854.277632

Received 29 November 2023; revised 13 May 2024; accepted 2 July 2024