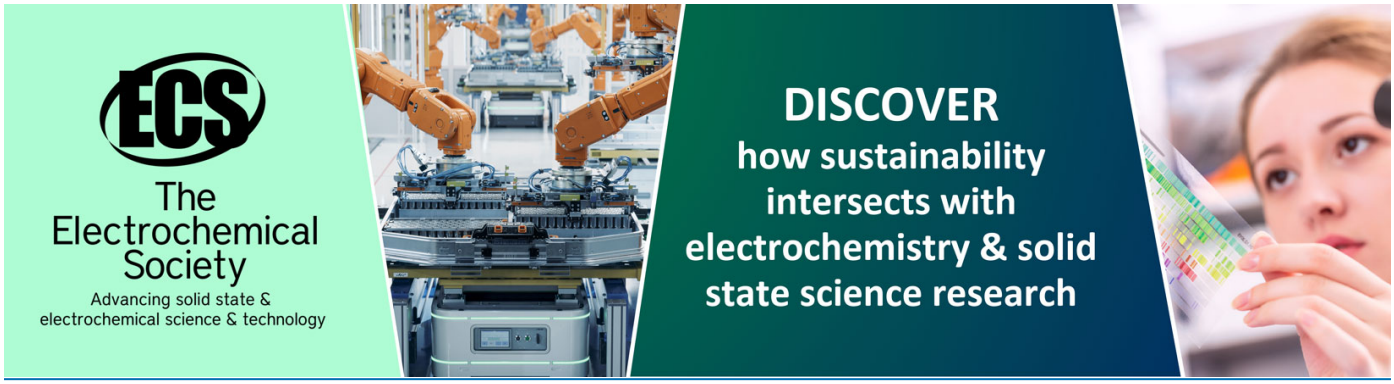
Intelligent chat robot in digital campus based on deep learning



Intelligent chat robot in digital campus based on deep learning

Yujie Liang[1], Ying Yu[2], Wenhao Ouyang[3]

[1]School of College of Software, Jiangxi Normal University, Jiangxi 330022, China
[2]School of College of Software, Jiangxi Normal University, Jiangxi 330022, China
[3]School of College of Software, Jiangxi Normal University, Jiangxi 330022, China

Corresponding author and e-mail: Yujie Liang, luck6393@163.com

Abstract. In this paper, two different technologies are adopted to realize the digital campus intelligent chatbot. The robot has three core modules: chat module, campus information module and virtual teaching assistant module. For chat module, this paper uses Encoder-Decoder framework. For the campus information module and virtual teaching assistant module, this paper pursue high precision and adopt retrieval mode. In practice, the two modules implemented by different technical means complement each other, greatly improving the user experience of students.

1. Introduction

With the advent of the era of big data, the integration of artificial intelligence and education will be promoted with the gradual advancement of digital campus and the development of online learning platforms [1]. In recent years, chatbot has become a popular product in colleges and universities. "Where is the report card opened?" "Who is my counselor?" "How many books can the library borrow?" "What time is the next class?" "How do I understand this knowledge point?" "Have my scores met the entrance criteria?" This requires a new intelligent chatbot to provide a new channel for students to solve problems.

According to the implementation technology, the types of chatbots are mainly divided into two categories: retrieval chat robot and generative chat robot. The retrieval chat robot is to search and match the corresponding answer content in the database after the user input the statement, and output it; while the generative chatbot is to generate a sentence answer through certain technical means after the user input the statement.

The main technologies to realize chat robot are: search based chat robot, artificial template based chat robot, machine translation based chat robot, deep learning based chat robot [2-3]. Chatbots based on artificial templates are different scenarios that are artificially set, and different dialog templates are written for each different scenario. For example, Siri, the intelligent voice assistant of apple, uses a lot of artificial templates. Its advantage is that compared with other technologies, this technology has higher accuracy. The chatbot based on the retrieval technology is the same as the search engine, which first stores the dialogue database and establishes the index. After the user enters the statement, search and match the corresponding answer content in the database, and then output. At present, the Encoder-

Decoder framework is used in most technologies of chatbots based on deep learning. The chatbot implemented by this framework is simple and easy to expand.

Through reading the previous articles, we can understand that the Encoder-Decoder framework is a way to deal with problems. It was first applied in the field of machine translation, inputting one sequence and outputting another sequence. We can understand it intuitively when we input one sentence to generate another. As shown in the figure 1, it is the most abstract representation of the Encoder-Decoder framework:

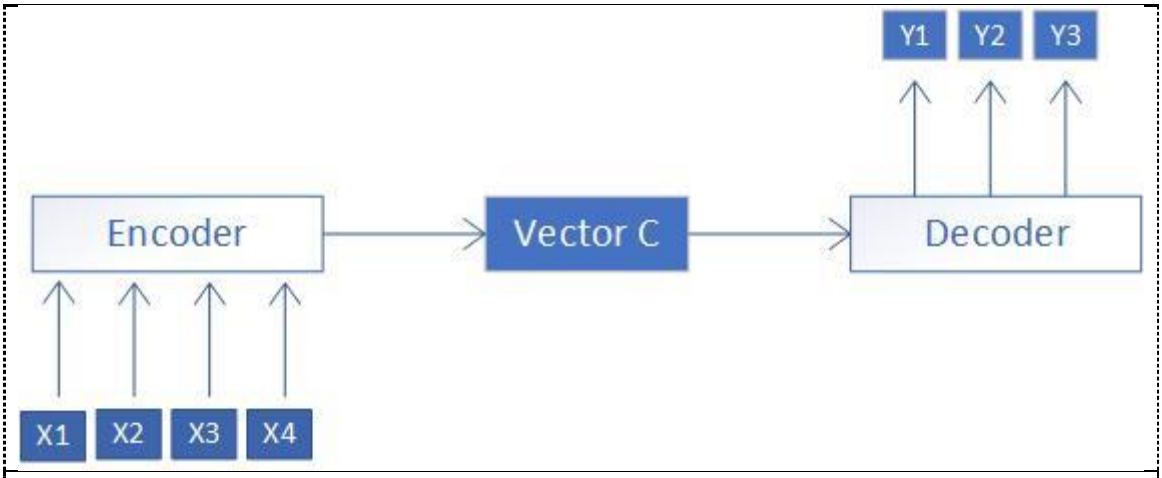


Figure 1. Encoder Decoder framework process.

2. Intelligent chat robot based on deep learning

This paper uses the Encoder-Decoder framework to realize the intelligent chat module of the digital campus chatbot. Specifically, for < X, Y >, X is the statement entered by the student. We treat it as a message. Y is the statement that the chatbot needs to answer, that is response. The basic process is as follows: After the user enters message, the message is encoded by Encoder to form the intermediate semantic vector C, and the Decoder generates the response that the chatbot needs to output according to the intermediate semantic vector C. In this way, whenever users input different messages, the robot can generate a response and form a dialogue system.

2.1. Model adopted by encoder-decoder

In the implementation of chat system, RNN model is the most commonly used deep learning model, which means Encoder and Decoder are written by RNN algorithm. RNN has an input X_t at each time, and then the output value H_t is calculated according to the current node state A_t, which is determined by the last time state A_{t-1} and the current input X_t. The improved model LSTM and GRU based on RNN are also frequently used. Especially for the long sentence situation, the actual response effect of the LSTM and GRU model is better than that of the RNN model. However, when the sentence length exceeds 30, the effect of the LSTM model will also decline. Attention model needs to be introduced at this time. The feature of Attention model is that Encoder will no longer encode the whole input sequence as a fixed length intermediate vector, but as a vector sequence.

According to the actual situation, the sentence length range of the robot designed in this paper will not exceed 30. Therefore, in the chat module, both encoder and decoder adopt LSTM algorithm. The chatbot built by this model is most commonly used to collect chat information for training. A large amount of chat information is used to train the neural network connection parameters corresponding to LSTM in Encoder-Decoder model, so as to improve the accuracy of answers.

3. Data acquisition

In terms of school common sense, we collect relevant questions through online questionnaires, interviews with freshmen and other forms. Most of these questions also contain the basic information

freshmen need to know about the school. Regarding campus information, we obtain relevant notification documents and establish answers through the school's official website, the school notification bar, and the school's official WeChat public account (notices need to be updated in real time). In terms of virtual teaching assistants, we set up answers by importing the Internet course question bank. We will classify the collected problems, a small part of which is shown in the figure below.

Table 1. Encoder decoder framework process.

Classification of problems	Examples of questions	Answer
School schedules	What is the daily schedule?	8 - 12 a.m. 2 - 5.10 p.m.
	What is the weekly break?	Every Tuesday afternoon for public holidays, there must be no curriculum, weekends generally no curriculum, specific table arrangements.
Activities at the school	What are the activities of the first semester of the school?	Basketball, football
	What are the school's special activities?	Red Culture and Arts Festival, weekly flag-raising ceremony (more grand)
The school's internal environment	What's the special view of the school?	Changshu, Water fir forest, blue lake, Mingde Garden, auditorium, red square
School rules	Does the school allow itself to leave school?	Longer time need to find a counselor leave, a short time can decide for itself
	How do I join a club?	The society will promote the new students, and in red square to perform open enrollment, you can according to their own interests to voluntarily sign up to join.

3.1. The data processing

For modules implemented by different technologies, we treat data differently. First of all, the data of the retrieval module should be processed according to the organization of the answers and questions. First, in txt file, multiple questions and answers are matched one by one, and questions with the same semantic are separated by "&&" separator, and "--" is separated with the answer, as shown in Figure 2. Questions and answers are stored as key-value pairs. The problem is the dictionary key, and the answer is the dictionary value.

```
class break time&&What's the break time&&home time
--The last class is over at 12 a.m. and 5.10 p.m.
Outage time&&Will there be a power cut in the dormitory
--Our dormitory has 24-hours electricity, but we should go to bed early
```

Figure 2. Retrieval data preprocessing.

For Encoder-Decoder framework, we need to establish a large amount of learning data and train the neural network parameters of the chatbot through a large number of inputs and outputs.

4. Model design

As can be seen from the following figure, we first accept the user's information and process it through the input module, and then intelligently classify the user's questions. According to the user's question types, we enter into different modules to intelligently answer and output.

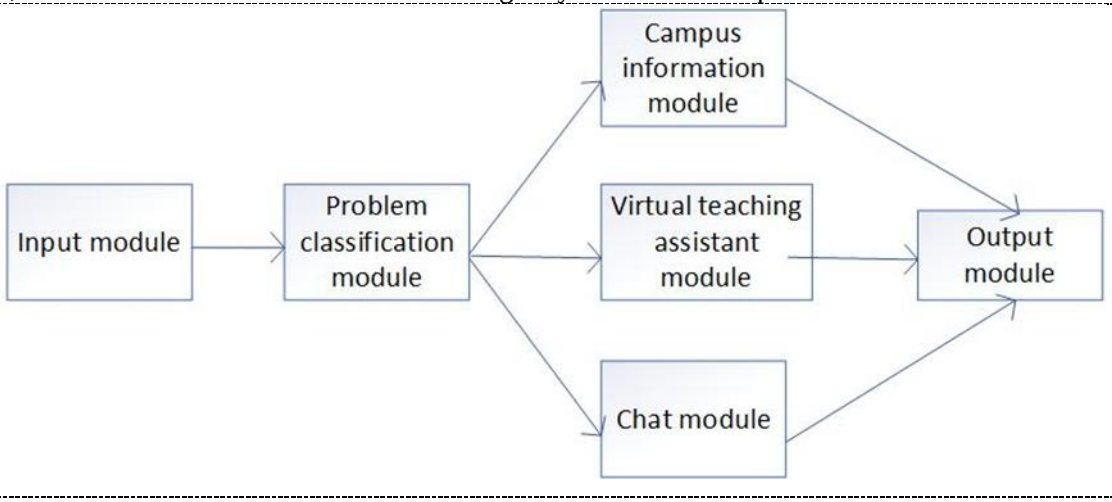


Figure 3. Overall functional structure design drawing.

4.1. Problem judgment module

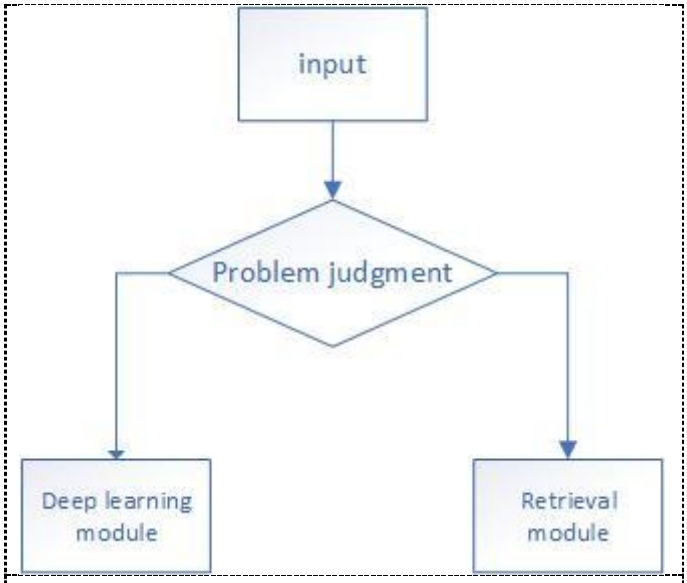


Figure 4. Process of judgment module.

As shown in Figure 4, the problem determination module is to judge and classify the user's input. In the early stage of chatbot training without extremely large data, the judgment of questions is often inaccurate, so we provide an autonomous question module selection for more accurate answers

4.2. Chat module

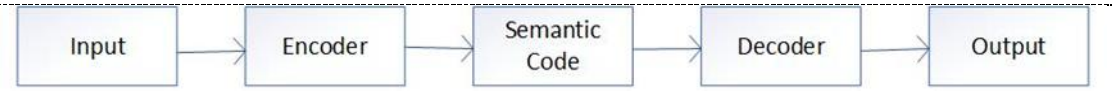


Figure 5. Flow chart of chat module.

This module is a generative answer based on deep learning. The Encoder-Decoder framework mentioned above is used. In this module, both the Encoder and Decoder adopt the LSTM algorithm. Chat module can interact with users, providing a more user-friendly design. "What's the weather going to be like tomorrow? ", "What day is it today?" for such a question we can interact through the chat module, stimulate the user's interest in using.

4.3. Campus information module and virtual teaching assistant module

The two modules are based on the retrieval technology of the chatbot. Just like a search engine, we store the dialog library and index it. These two modules are the heart modules of the model. First, the module will process the questions, match the answers in the database, and output the answers from the database. Of course, we store the database data of the two modules separately, so that the matching is more accurate after judging the problem type. In the management of database data, we need to integrate many different problems of a problem to obtain a common identity, and use this identity as the key value, so that our data can be robust.

5. Experiment and analysis

Any mass-oriented product needs repeated inspection and testing, so does the robot designed in this paper. The chatbot's conversational effect is tested by combining the two different techniques proposed in this paper. After the user enters the question, it is necessary to determine whether the answer output by the chatbot is correct. For example, input "What's the weather like today?" the robot outputs the answer related to the question "today is a cloudless day", which indicates that the robot is relatively perfect. But if you input "What's the weather like today?", the robot output "Hello!", which indicates that the robot needs to be improved. However, these are all subjective feelings of human beings. In order to objectively test the intelligence of intelligent chatbot, this part tests the chat module, campus information module and virtual teaching assistant module. According to different implementation techniques, we will divide it into two parts:

The first part is an experiment of the chat module based on deep learning. The evaluation of the framework model generally adopts BLEU value, but the basic BLEU algorithm is too simple, prone to be disturbed by common words, and does not consider the accuracy of language expression (grammar). This paper uses an improved n-gram precision. The words are divided into a tuple set, binary set, triplet set … n-tuple set. For different sets of groups, W_i weights are different, so here we take geometric weighted average. p_n is the improved multiple precision, w_n is the weight given. The specific formula is as follows:

$$Pave = \sqrt[\sum_{n=1}^{N} w_n]{\prod_{n=1}^{N} P_n^{w_n}} = \frac{1}{\sum_{n=1}^{N} w_n} \exp\left(\sum_{i=1}^{N} w_n * log^{p_n}\right) = \exp(\frac{1}{N} * \sum_{i=1}^{N} log^{p_n})$$

The second part tests the retrieval campus information module and the virtual teaching assistant module. Since retrieval robots tend to search for different answers to the same semantic questions, we need to use multiple questions to test the robustness of the answers of the modules.

5.1. The result analysis

The BLEU value is a text evaluation algorithm originally used to evaluate the accuracy of machine translation. The specific calculation formula is as follows:

$$BLEU = BP \times \exp(\sum_{n=1}^{N} W_n log P_n)$$

For the first part of the experiment, we used 1800 pieces of data for testing, and the final test result was 0.6950, with good overall function and meeting expectations.

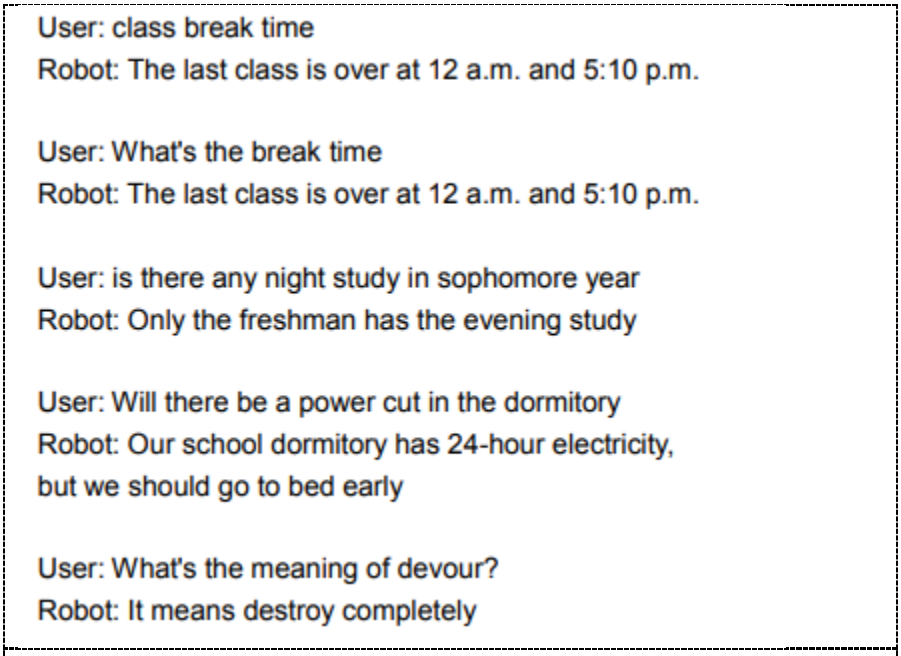


Figure 6. Module test results based on retrieval technology.

Figure 6 shows the conversation effect of the retrieval technology-based chat model. We take multiple approaches to the same question. It can be seen that the retrieval module proposed in this paper has certain robustness and can get the same answer to many different questions. The overall function is as expected.

6. Conclusions

In this paper, the chatbot designed for the digital campus can greatly reduce the human services. And in practice, the performance of the robot in this paper meets the expectation. The robot can effectively identify and answer all kinds of questions about campus. It is convenient to use and timely reply. It can help students to solve campus and course problems in a timely manner, enable freshmen to integrate into campus life as soon as possible, enable students to learn better, and provide a new way for students to solve questions.

References

[1] Zhang shuai, liu hailin, wang yuanyuan. Based on analysis of large data bot research [J]. Digital communication world, 2020, (1) : 138147. Doi: 10.3969 / J.I SSN. 1672-7274.2020.01.102.
[2] WANG Qian-ming, LI Yin Research on Personalized Chatbot Based on Deep Learning [J] Computer Technology and Development 2020,30(4):79-84.doi:10.3969/j.issn.1673-629X.2020.04.015.
[3] Guo qixin, yu weihong, li zhaohui The customer service system for bot design [J]. Journal of software, 2019, 40 (9) : 84-86. The doi: 10.3969 / j.i SSN. 1003-6970.2019.09.019.
[4] Zhao Hongyang Research and implementation of intelligent chatbot based on deep learning [J]. Intelligent computer and application,2019,9(6):308-311.
[5] Cao bin, wang feng, li shiyu. Design of intelligent chatbot system in the field of traditional Chinese medicine [J]. Computer knowledge and technology, 209,15 (12) : 174-175,185.
[6] Zhu jingjing, Han lixin Based on RNN sentence encoder bot [J]. Computer and modern, 2018, (1) : 32-35 doi: 10.3969 / j.i SSN. 1006-2475.2018.01.007.