Existing Coding

https://github.com/itachi9604/healthcare-chatbot (https://github.com/itachi9604/healthcare-chatbot)

The project of this healthcare-chatbob is a very simple rule-based diagnosing system. The wokd is similar with the one Ruibin has implemented in his diagnosing system. But this one focus on chat and diagnosis. Also, the project contains a small dataset of disease and relevent information.

The shortage of this one is that user can only input single word or number during the conversation. So I add a basic tokenization and stop words detection function so that the user can input sentences and the program can detect an obvious disease from the sentence.

Following is a demo conversation:

```
Your Name
                                                ->Xiaoxiao
hello  Xiaoxiao
Enter the symptom you are experiencing
                                                ->Hi I think I have so
me itching
searches related to input:
0 ) itching
1 ) internal_itching
Select the one you meant (0 - 1):  0
Okay. From how many days ? : 2
Are you experiencing any
itching ? : yes
skin_rash ? : yes
stomach_pain ? : no
burning_micturition ? : no
spotting_ urination ? : no
It might not be that bad but you should take precautions.
You may have  Drug Reaction or  Chronic cholestasis
An adverse drug reaction (ADR) is an injury caused by taking medicatio
n. ADRs may occur following a single dose or prolonged administration
of a drug or result from the combination of two or more drugs.
Chronic cholestatic diseases, whether occurring in infancy, childhood
or adulthood, are characterized by defective bile acid transport from
the liver to the intestine, which is caused by primary damage to the b
iliary epithelium in most cases
Take following measures :
1 ) stop irritation
2 ) consult nearest hospital
3 ) stop taking drug
4 ) follow up
```

Literature Review

Chatbot for healthcare system using Artificial Intelligence[1]

In this paper, keyword extraction, ranking and sentence similarity calculation are done using n-gram, TF-IDF and cosine similarity. The query posed to the bot which isn't comprehended or not present in the database isfurther processed by the third party, expert system.

The methods mentioned in this paper is well-known for NLP. These methods, such as n-gram and Tf-Idf can be also adopted by us. Using cosine similarity for NLG part is also feasible if the dataset is large enough.

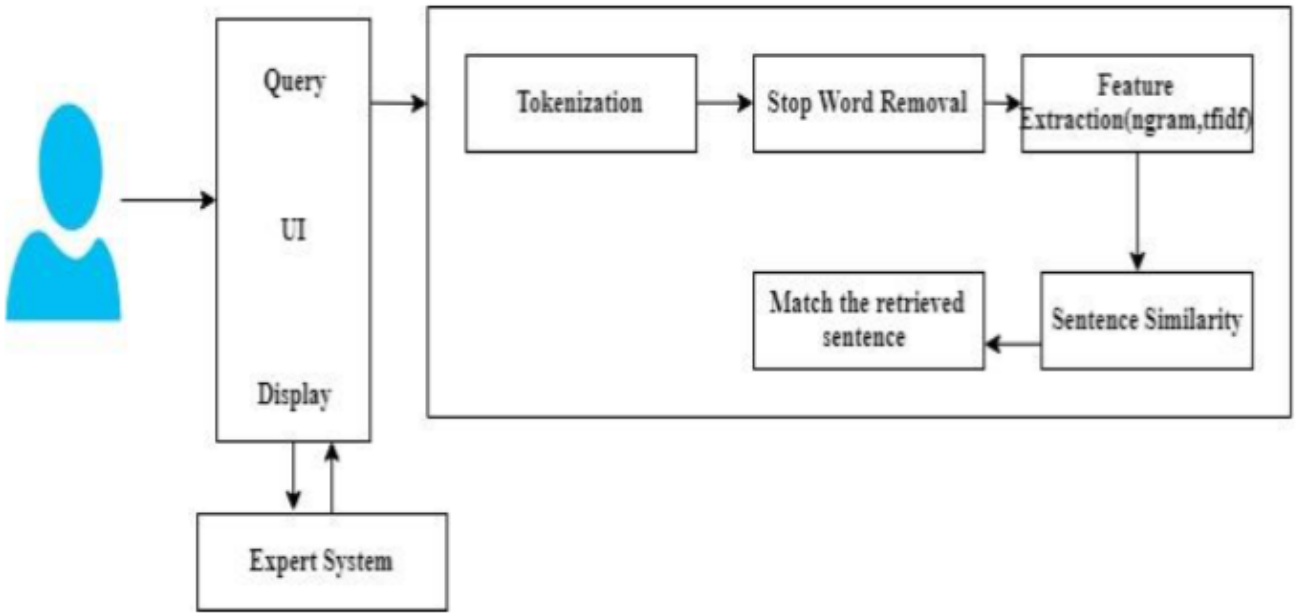

Fig. 1: System architecture

An overview of the features of chatbots in mental health: A scoping review[2]

Some highlights conclustion from this paper:

- Most chatbots were implemented in the United States.
- Chatbots were used for several purposes, namely: therapy, training, and screening.
- Majority of chatbots were rule-based and implemented in stand-alone software.
- Chatbots focused mainly on depression and autism.

Type *Markdown* and LaTeX: α^2

Conclusion

For NLU part: basic NLP methods such as tokenization, POS tagging, Key words extraction using regular expression and Tf-Idf can be done for disease words detection. After these work, Named Entity Recognition approaches can be tried.

For NLG part: apporches such as cosine similarity and Tf-Idf is feasible for a rule-based conversation generation system, along with some conversation templates. In the future work, we need to try and explore NLG models for a more human-like conversation.

Problems for this stage: We still has no dataset big enough for NLP work.

Next week, apart from Literature Review, I will keep searching datasets contain healthcare dialogue.

Reference

[1]Kavitha, B. R., and Chethana R. Murthy. "Chatbot for healthcare system using Artificial Intelligence." International Journal of Advance Research, Ideas and Innovations in Technology ISSN 5.3 (2019): 1304-1307.

[2]Abd-Alrazaq, Alaa A., et al. "An overview of the features of chatbots in mental health: A scoping review." International Journal of Medical Informatics 132 (2019): 103978.

[3]Bhirud, Nivedita, et al. "A Literature Review On Chatbots In Healthcare Domain." International Journal of Scientific & Technology Research 8.7 (2019): 225-231.