Interim Report: Data: From Patient to Health Record

Thesis submitted to the University of Nottingham for the degree of
BSc Computer Science, Dec 2020.

Yiming LI

20031525

Supervised by Dr. Boon Giin LEE

School of Computer Science University of Nottingham, Ningbo, China

Abstract

Artificial intelligence has a huge potential to improve the efficiency of clinical tasks. Nowadays, Doctors are overwhelmed by typing records into computers using traditional keyboard and mouse and they usually spend more time on typing than diagnosing the patients. Voice recognition methods can help doctor reduce clerical work and improve the accuracy of Electronic Health Record (EHR). Together with language processing methods, it provides a feasible solution for doctors to fill in the EHR form and diagnose the patient simultaneously without distraction. Therefore, it can improve the efficiency of diagnose and avoid manual mistakes. However, auto-filling EHR systems based on voice recognition has no efficient voice-based EHR deployment in clinical field currently in China. This project aims to create a system capable of translating real time voice input or audio files into plain text and process and analysis the key points to fill in the EHR form. The challenge is to achieve high accuracy in voice recognition and part of speech tagging to minimize the manual inspections as many as possible.

Contents

List of Tables

List of Figures

Abbreviations

AI Artificial Intelligence.

API Application Programming Interface.

EHR Electronic Health Record.

MoH Ministry of Health.

NLP Natural Language Processing.

PoS Part of Speech.

WPM Words Per Minute.

Chapter 1

Introduction

1.1 Background

The Electronic Health Record (EHR) serves as a collection of patients' health record and health status throughout their whole life for clinical purposes. It is a computer information system which can support data collection, storage and access in both hospitals and healthcare centers to uniform the medical record format [1]. The Chinese government, as well as Ministry of Health (MoH) of China, had regarded EHR as an efficient tool to improve the safety and quality of Chinese health care service and set a goal to ensure the universal usage of EHR among the whole population in most of the hospitals and clinics by the end of 2020 [2].

Comparing to traditional paper records, EHRs offered advantages such as remote data access, unified data standard, searchable digital database and integrated patient records including medical history [3]. A research taken by Jennifer King demonstrated that over 75% of EHR adopters identified that EHR enhanced the health care service [4].

EHRs offered more efficient entry and retrieval of relevant patient information. However, a potential weakness of EHR is the discommodious input interaction using traditional keyboard and mouse. The maximum number of words per minutes (WPM) was 80, when concentrating on typing [4]. A 2016 study estimated that doctors spent between 37% and 49% of their working hours on clerical tasks [5]. Doctors were overwhelmed by this clerical work and had a great possibility to make serious mistakes by typing manually under this circumstance [6]. All that paperwork contributed to the high level of burnout and depression in the profession, according to a 2018 study [3].

To resolve the shortage of traditional computer systems such as EHR with only basic functions, Artificial Intelligence (AI) has been implemented into many aspects of human daily lives, especially in clinical workflow [7]. It shows its influence on clinicians, health systems and patients since it produce accurate image interpretation, reduce manual errors and provide high accurate health diagnosis. [8].

1.2 Motivation

To solve the inefficiency of current EHR, voice assistant can serve as clinical stenographers that transcribe doctors' observations and instructions and insert them into a patient's EHR [3]. In theory, with the assist of the speech recognition system, it will liberate doctors from tedious clerical work and improve the accuracy and quality of EHRs.

As the practice of voice recognition in the past 10 years, the accuracy of the outputs is not ideal. A research in 2010 implemented a voice recognition method and compared result with the manually translated content. The average accuracy was less than 82%, with 6.1% of incorrect recognition and 11.2% of rejected voice [9]. Speech averages about 110-150 Words Per Minute (WPM). However, due to the poor performance of voice recognition system, 70% of extra time was required to correct errors [3]. At present, with the development of the machine learning, plenty of voice recognition API was provided with a higher accuracy, such as Google voice recognition Application Programming Interface (API) [10] and Baidu voice recognition API [11]. Combined with noise reduction algorithm to obtain a clear voice input, the accuracy would achieve near 99%.

Apart from the voice recognition, an approach was necessarily required to allow machine to derive meaning from human languages, as well as decompose a sentence into independent words. Many Natural Language Processing (NLP) studies had been conducted and developed to analyze the Part of Speech (PoS) and the meaning of languages. Tested by Che et al. [12], a sufficient accuracy and speed have been attained in some of Chinese processing modules, including WordSeg (97.4% of accuracy, 185KB/s of speed), POSTag (97.80% of accuracy, 56.3KB/s of speed), NER (92.25% of accuracy, 7.2KB/s of speed) and so on. With a further training on an additional dictionary containing specific disease names and medical drug names, it can be perfectly adapted to medical segmentation analysis.

Despite of all the benefits listed above, EHR adoption in China has low prevalence as well as low quality. Nearly 30% of hospitals are still using handwriting medical records without the use of computers and another 30% only have the basic functions of EHR [1]. The application delivered in this project has high possibility to improve the efficiency and quality on processing these clerical works. However, there are numerous hospitals and health-care centers in China. High installation fee, high training cost and inefficient popularization from main hospital to tertiary hospitals are remaining obstacles. In addition, the number of research conducted on EHR in China is much lower than researches in the USA. During 2008-2017, there are 1031 publications on EHR in the USA while there are only 173 publications in China [13]. Researches on EHR is a relatively new emerging and promising field in China. This project mainly aims to explore the feasibility of implementing voice recognition methods (speech to text) and NLP methods (text analysis) in EHR system based on Chinese language to contribute to EHR in medical field.

1.3 Aims and Objectives

The main objective of this project is to create an auto-filling system using voice recognition to reduce the working load of doctors. Using this system, the speech of doctors should be able to recognized and analyzed to fill into the EHR at the same time when doctors are diagnosing the disease of patients. In this way, doctors can save those time spent on writing documents or typing in digital records and be more focused on diagnose. The system needs to be capable of translating real time voice input or audio files into plain text and process and analysis the key points to fill in the electronic health record (EHR). In other words, the system is a voice assistant can serve as clinical stenographers that transcribe doctors' observations and instructions and insert them into a patient's EHR.

The key objectives of this project are:

1. Collection of doctor diagnose prescription for testing the feasibility and accuracy of the system.
2. Voice recognition methods implementation for transcribing doctors' speech into text.
3. Chinese language processing algorithms implementation for analyzing the text and fill in the EHR. Together with objective 2, it can allow doctors use voice as input instead of keyboards and mouse.
4. Assessment of different voice recognition methods and different Chinese language processing algorithms for achieving higher accuracy.
5. Development of prototype to demonstrate the proposed work for simulating the realistic situation.

Chapter 2

Related Work

With the developemnt of AI approaches, there are several mature application or API for voice recognition and nature language processing. This chaper will introduce relative studies on voice recognition, state-of-the-art of NLP, and current status for AI-based EHR in healthcare industry.

2.1 Overview of Voice Recognition Methods

Mentioned in Chapter 1.2, Google voice recognition API [10], Iflytek voice recognition API [14] and Baidu voice recognition API [11] are the main products that provide Chinese language recognition service. All these APIs will be test in this project to figure out the accuracy of each API and find the best one.

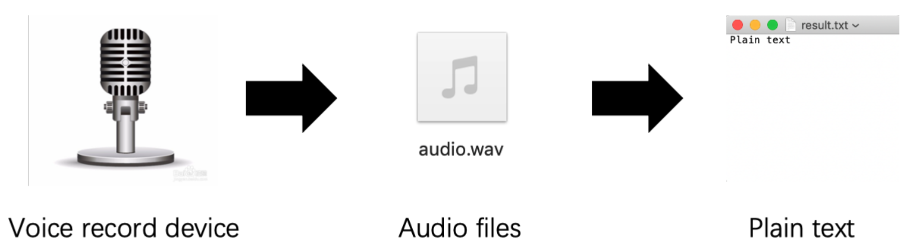

Figure 2.1: Voice Recognition Workflow

2.1.1 Baidu Voice Recognition

Baidu voice recognition API supports the self-training model on the voice self-training platform, and the training can be completed with zero code after uploading the vocabulary text in a dictionary. It can accurately improve the vocabulary recognition rate of the specific domain by 5-20%. To use Baidu voice recognition API, an access key should be applied. [11].

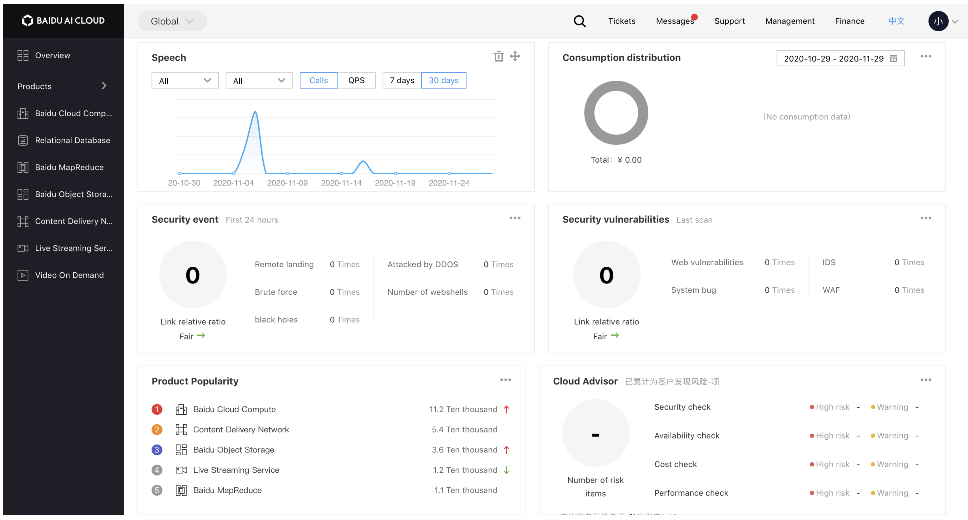

Figure 2.2: Baidu AI server Main Page

2.1.2 Google Voice Recognition

Google with voice recognition API supports voice recognition that supports more than 125 languages and variants for free and without the restriction of access key. The same as Baidu voice recognition API, it can customize speech recognition to transcribe domain-specific terms and rare words and boost your transcription accuracy of specific words or phrases. In addition, it can automatically convert spoken numbers into addresses, years, currencies, and more using classes.

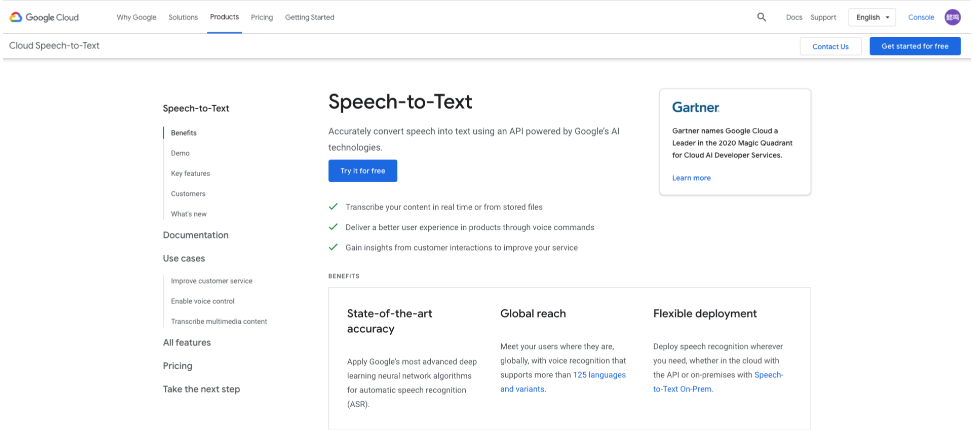

Figure 2.3: Google AI server Main Page

2.2 Overview of Nature Language Processing methods

According to a literature review [15], an NLP system includes two parts: syntactic processing modules and semantic processing modules. Several different NLP systems based on English have been utilized to extract key information from clinical contents, such as MedLEE, MetaMap, KnowledgeMap, cTAKES, HiTEX and MedTagger [15]. 65% of existing NLP systems use rule-based methodology while the rest use machine learning based methodology to do the information extraction. The table below listed five frameworks or tools on clinical information extraction, which are the most mentioned in publications from 2009 to 2016.

Table 2.1
Top 5 information extraction frameworks/tools included in publications from 2009-16.

Name	Description	No. of Papers
UIMA	software framework for analysis of unstructured contents	31
cTAKES	Open-source NLP system based on UIMA framework	26
MetaMap	National Institutes of Health developed NLP tool	12
MedLEE	NLP systems for narrative clinical notes	10
GATE	Java-based open-source software	5

There are many English language processing frameworks and tools such as NLTK, Stanfordnlp, and CoreNLP. However, similar tools based on Chinese language are relatively too little to find any research on utilizing these tools on EHR. Chinese sentence syntax and word characteristic components are quite different from English grammar, so a chinese language processing tool are required in this project.

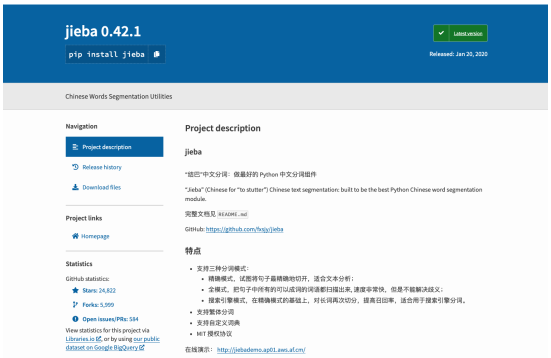

Figure 2.4: Jieba Main Page.

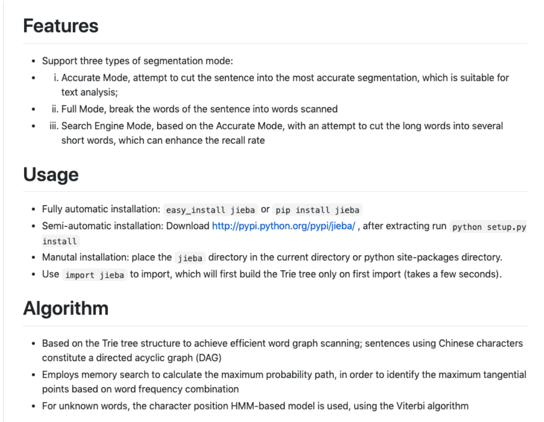

Figure 2.5: Features and Usage Documentation for Jieba.

"Jieba" (Chinese for "to stutter") [16] is one of Chinese text segmentation tools, which is built to be the best Python Chinese word segmentation module. It implemented almost all the functions similar to other English language processing frameworks and tools. Keyword Extraction, Part of Speech Tagging, Tokenize are the main functions which used in this project.

2.3 Overview of Voice Assistant for EHR

Several voice recognition and language processing systems had been developed which only supported English. Alexa voice assistant of Amazon [17], Saykara [18] and Suki [19] are famous and mature applications of AI assistant for doctors on either mobile platform or computer platform. Unfortunately, there has been no similar research or application in Chinese hospitals and clinics except IFLYTEK CO.LTD. [14] started implementing a voice EHR system in 2017.

2.3.1 Iflytek EHR

System of Iflytek [14] is designed to synchronously records the voice into patients' medical record, when doctors communicate with the patient. This project will explore feasible approach and method to implement voice recognition technology as well as NLP on automatically filling keywords into EHR to improve the efficiency of doctor's diagnose based on Chinese.

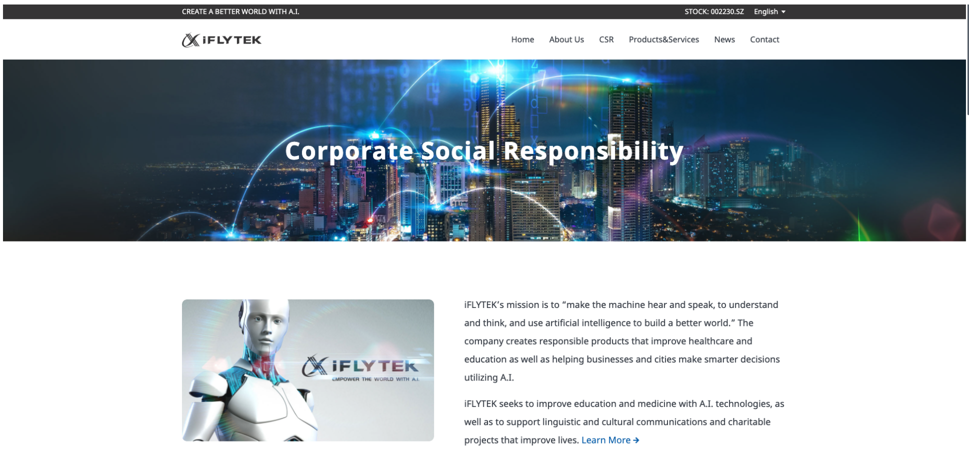

Figure 2.6: Iflytek Main Page

However, there is only description of such system on the Chinese official website of Iflytek and even not mentioned on English version of website. In this case, this project regards it as an unfinished system, which is still under development. Refer to this hypothesis, there does not exist any of developed AI-based EHR systems in China.

Chapter 3

Project Specification

This section states the main requirements for the project including functional and non-functional requirements. Functional requirements define basic system behaviour, while non-functional requirements show constraints or restrictions on the design of the system.

3.1 Functional Requirements Specification

1. The application can be started and used by doctors after logging in.
2. The application must include button to record real time speech.
3. The application must include option to load audio files from local computer.
4. The application must include voice recognition method to convert recognition into text.
5. The application must include text processing method to analyze raw text and fill the corresponding information into EHR form.
6. The application must be be available for users to manually modify the EHR form.
7. The application should include option to save the generated EHR form to local directory.
8. Quit option should be available for users in the main menu.

3.2 Non-Functional Requirements Specification

3.2.1 Usability

The user interface user-friendly, which means it should be easy to use. The buttons and text fields should be clear and readable in a neat format. Any user should understand the meaning of each button or text field that are used in the application without any training. Efficiency and satisfaction are two main categories to measure the usability to see whether the user has can achieve their goals quickly and be pleasant to the UI design. Beyond that, error rate of any user operations should be under 5%.

3.2.2 Performance

The application should load within 5 seconds of running the program. After audio recording or audio loading, the application should convert speech to text within 5 seconds, process text within 5 seconds and to predict disease within another 5 seconds. Totally maximum of 15 seconds are acceptable to give the final result in EHR form. There should be no frozen scenes or stuck moments and there should be a status bar to inform user the current status of application.

3.2.3 Operating Environment

The application should work on computer-based environment with python installed. Microphone or speaker is required if users want to use real-time speech as input.

3.2.4 Security

The application can store personal information into local directory, so that the local file should be password-protected or equipped with an encryption algorithm. In addition, The applications cannot access to personal data unless the user gives permission to the application before the start.

3.2.5 maintainability

The application should be able to recover from critical failures in a short time. Any run time error should be recorded in a log file for further maintenance. In addition, warning or error message should be popped up

3.3 Priority and Risk

Table 3.1 specifies the priority and risk of each functional requirement during the development stage. Priority is ordered from 1 to 5, with the highest priority to the least. Loss Probability is ordered from 1% to 99%. The lower the percentage is , the lower risk the system has to fail. The average loss probability of this system should be low enough to ensure the stability.

Table 3.1
Priorities and Risks of Functional Requirements.

ID	Name	Priority	Loss Probability
1	application start	1	5%
2	real time recording	2	20%
3	local audio file loading	2	10%
4	form fillment	3	20%
5	form modificatione	4	5%
6	file storage	4	10%
7	quit	5	5%

Chapter 4

Methods

4.1 Dataset Collection

There isn't any existing clinical documentation dataset in Chinese language, so the best way to prove the feasibility of this system is to collect a new dataset and use these data to test on the whole system.

All the data sources come from an online doctor diagnose website (https://www.haodf.com/). On this website, diseases are classified into different classes within different departments in the hospital. Patient can choose the specific departments to find for possible diseases in related class and ask questions in it. Patients type in the description of their symptom based on their own observation and doctors can ask further questions to diagnose the disease for the patient. The basic format of conversations between patients and doctors is text-based or speech-based.

To save the time for data collection, a crawler algorithm is implemented to grab useful information from the websites. First step is to grab the classification categories into excel files as labels. The structure is designed as demonstrated in Figure 4.1. Second step is to grab all the text in the conversation between patients and doctors. However, not every URL is useful and not every conversation is useful. In this case, useful URLs are be manually identified.

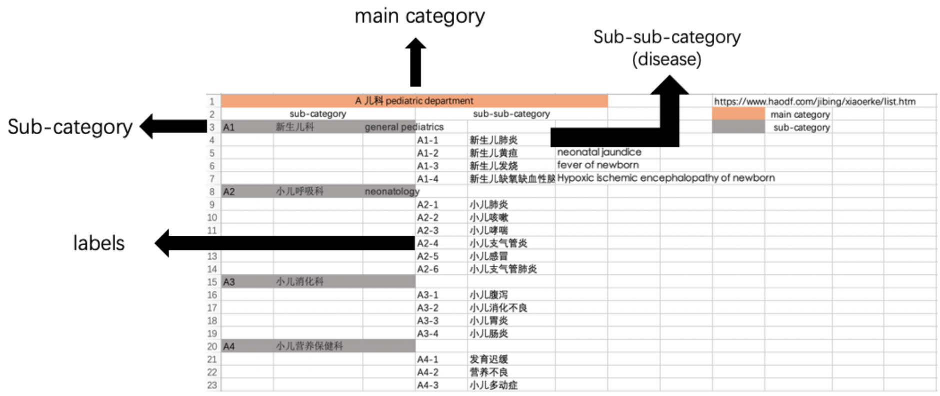

Figure 4.1: Classifications and labels in Dataset.

4.2 Data Pre-processing

After collecting the row data, data pre-processing is required to manually clean the data and label all the cases to make the data follow the classification in order to do the following training or testing. Figure 4.2 shows the samples of manually labeled cases. "Symptom description" is from the observation and description from the patients and doctors repeat these symptoms and combine with their diagnose together to be the "Doctor's diagnose". At current stage, Doctor's diagnose is read and recorded as input.

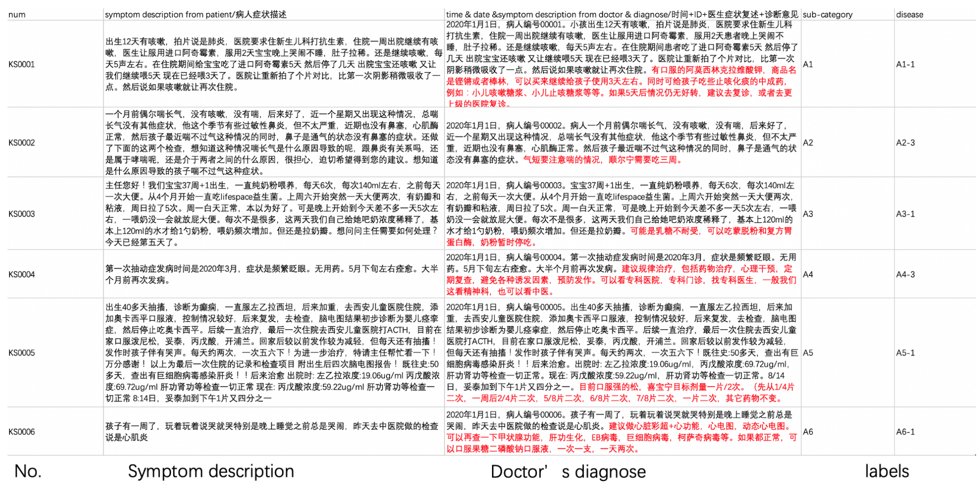

Figure 4.2: Labeled cases in Dataset.

4.3 Application

As mentioned in Chapter 1.3 - aims and objectives, main tasks in this project is to utilize all the required methods, including voice recognition methods and Chinese language processing methods. User interface is required for a better demonstration as long as it is clear for users to understand the meanings of each part.

4.3.1 Voice Recognition

The first step is to implement voice recognition methods. Two voice recognition methods are tested in this project: Baidu voice recognition method and Google voice recognition method. Shown in Figure 4.1, both of them can recognise speech in Chinese and have API for Python. Baidu voice recognition method requires an access key to pull requests from the Baidu AI server and have usage limitation on the number of times. The processing time is relatively faster than Google AI server. Besides of that, VPN is required to have access to Google server. After considering the advantage and disadvantage of both voice recognition methods, Baidu voice recognition method was selected because of the better performance.

4.3.2 NLP

Since there isn't much choices for Chinese language processing methods and Jieba library contains all the functions that are needed, Jieba library is used to do the information extraction in this project. Inside the Jieba library, there are two kinds of Part of Speech Tagging Function offered. The first one is default mode which allows developers to modify dictionary by their own and the other one is pre-trained mode called "paddle mode" with abundant lexicon. The idea is to combine these two kinds of mode to obtain a better information extraction. Incorporate the more accurate result of "paddle mode" with the definition of unfamiliar terminology with dictionary of default mode, it is believed that the result will be more accurate, which can avoid doctors to proofread contents or modify mistakes.

4. Part of Speech Tagging

- `jieba.posseg.POSTokenizer(tokenizer=None)` creates a new customized Tokenizer. `tokenizer` specifies the jieba.Tokenizer to internally use. `jieba.posseg.dt` is the default POSTokenizer.
- Tags the POS of each word after segmentation, using labels compatible with ictclas.
- Example:

```
>>> import jieba.posseg as pseg
>>> words = pseg.cut("我爱北京天安门")
>>> for w in words:
...     print('%s %s' % (w.word, w.flag))
...
我 r
爱 v
北京 ns
天安门 ns
```

Figure 4.3: Jieba Part of Speech Tagging Function.

4.3.3 GUI Design and Implementation

A user interface is designed and implemented shown as Figure 4.4. For now, it is a basic medical record with patient names, IDs, dates, symptoms, disease, treatments and also a container for the whole recognized content. The input of the application is an audio file or real-time speech recording on symptom and treatment description. Then, voice recognition method is used to recognize audio or speech into plain text. After that, use part of speech tagging method to separate the words from the sentence. Trigger words are detected to filter all the information into corresponding cell of the EHR form. Finally, users have the option to save the EHR form to local csv files.

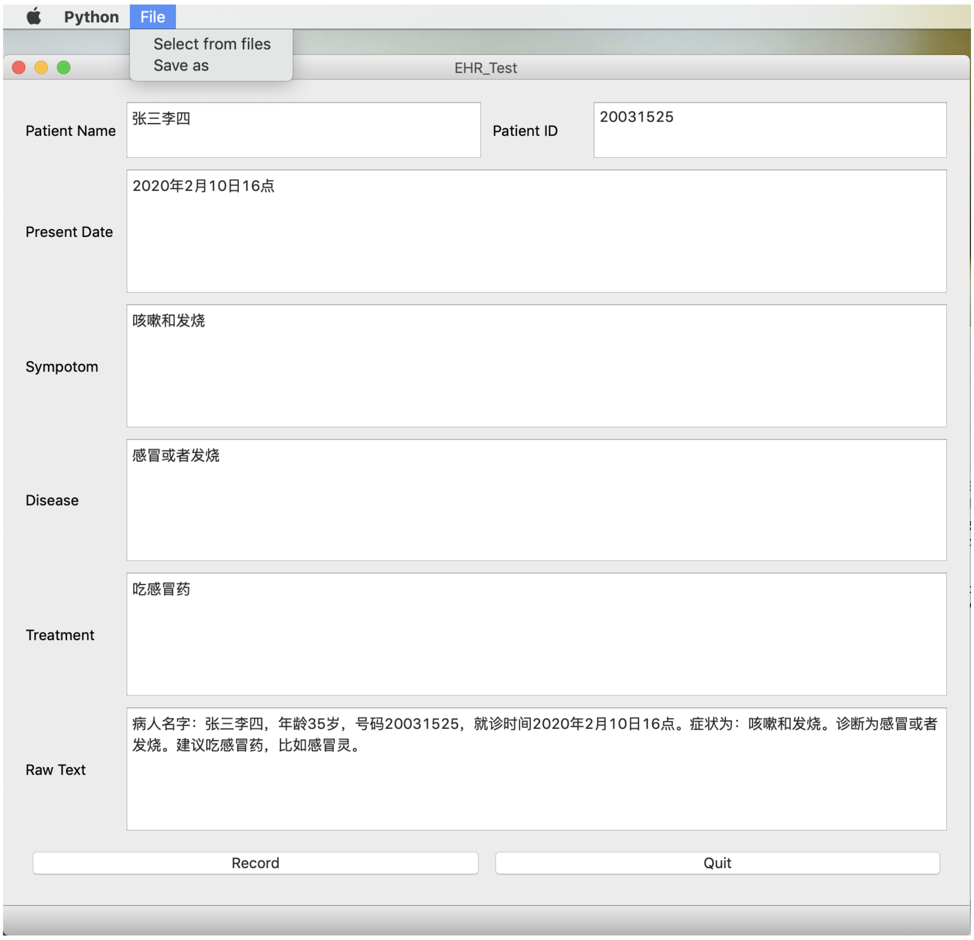

Figure 4.4: User Interface of the Application.

4.3.4 Software Specification

This chapter presents the software Specification of this project. Python is a very versatile programming language that can be used across a variety of different fields, including user interface design, number of available API as well as many mature AI methods. In this case, Python is selected to use in this project. Because the application has not been encapsulated now, PyQt5 is also required to run this application. In addition, several libraries are also required. After completing the development, it is possible to encapsulate

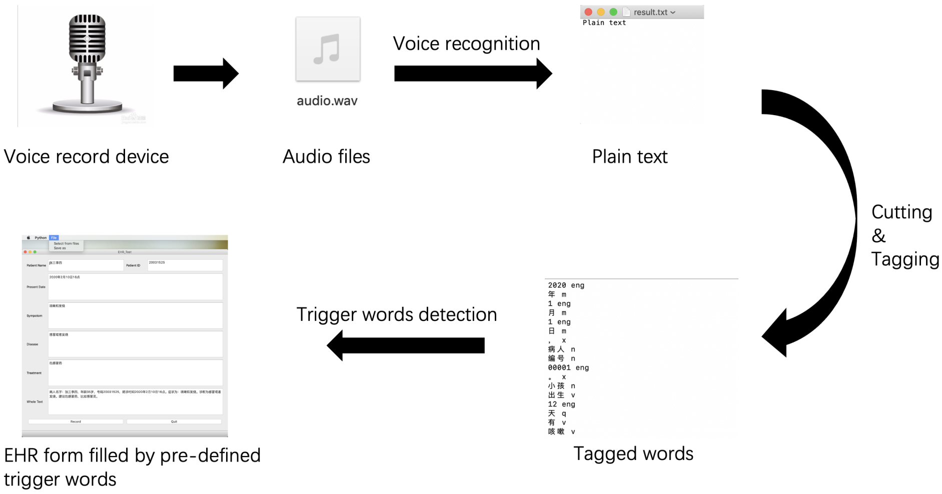

Figure 4.5: Whole Progress of the Application.

the application so that there will be no need for users to configure the environment and install those libraries.

Table 4.1
Software Specification.

Detail	**Requirement**
Platform	computer-based
Operating System	Any operating system installed Python
Configure of Environment	Python, PyQt5
Libraries	SpeechRecognition, jieba, requests, OpenCV
Language	Chinese

Chapter 5

Result and Discussion

5.1 Preliminary Result

Table 5.1 demonstrated preliminary result of Baidu voice recognition method and Google voice recognition method for different language speech. It can be clearly observed that Baidu voice recognition method shows better result on mandarin and Google voice recognition method shows better result on English. In the case of mixture of mandarin and English, Baidu voice recognition method can still recognize most of the contents in Chinese mode. The English part can be correctly recognized while Google voice recognition method has great probability to recognize it as a Chinese word. In addition, Baidu voice recognition method can correctly recognize most of the medical terminologies from the speech without adding any dictionary.

The results is calculated by average percentage of difflib, Levenshtein and fuzzywuzzy comparison methods and kept two decimal places. Baidu voice recognition method offers better performance as well as the quicker response. After considering the advantage and disadvantage of both voice recognition methods, Baidu voice recognition method was selected.

After the implementation of two NLP methods, the results will be compared to find the

Table 5.1
Comparison between Baidu and Google voice recognition methods.

Description	Methods	Mode	results
All mandarin	Baidu voice recognition	Chinese	80.04%
All mandarin	Google voice recognition	Chinese	76.31%
All English	Baidu voice recognition	English	86.53%
All English	Google voice recognition	English	89.12%
Mixture of mandarin and English	Baidu voice recognition	Chinese	81.26%
Mixture of mandarin and English	Google voice recognition	Chinese	70.75%
mandarin with terminology	Baidu voice recognition	Chinese	80.38%
mandarin with terminology	Google voice recognition	Chinese	63.23%

similar part of the same category. Since one of the methods does not separate punctuation mark from the words, it is hard to identify the boundary of the sentences. In this case, there may exists some cases no similar content is found between two methods so that the cell of EHR form will be empty. It requires doctors to fill in this cell manually. After testing, it is found that 13% for this case to happen.

5.2 Discussion

The cases in the dataset are manually identified and collected which is time-consuming, so the number of cases is not sufficient at the current stage. More data should be collected in the future work. In addition, The result given by the application cannot be quantified for now. The only accuracy we got is the result from voice recognition methods. The average accuracy is not high enough due to several reasons other than recognition error: a slip of the tongue, noises, and the speed of speech. If the speakers can speak with no oral mistakes in a proper speed, together with a noise reduction algorithm, the result is believed to be much better. Also, limited by the language requirement, insufficient terminologies are found and tested.

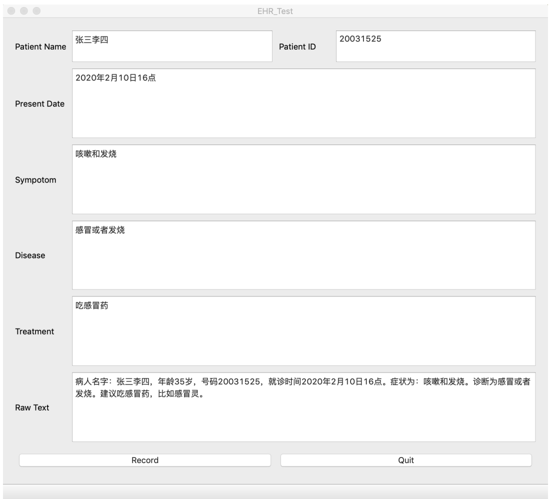

Figure 5.1: Result of Perfect Speech Input.

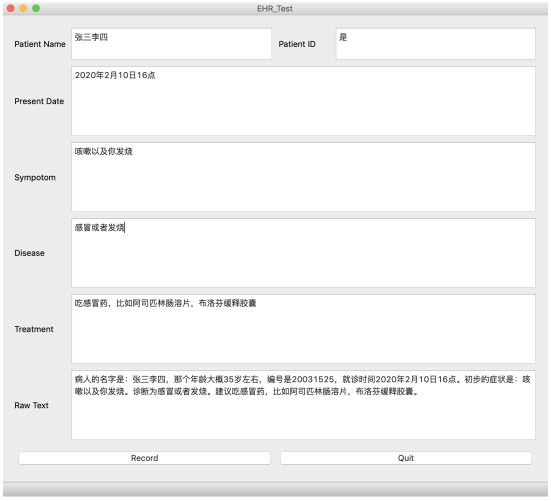

Figure 5.2: Result of Speech Input with noise.

In addition, as shown in Figure 5.1 and Figure 5.2, the result of speech input with a slip of the tongue and oral mistakes is not accurate enough and some blanks are filled in with wrong information. In this case, users should manually check and correct the form, which violates the aim and objective of this project. So, text extraction methods should be improved to achieve a better performance with higher accuracy.

Chapter 6

Conclusion

At the current stage, a basic user interface has been developed to demonstrate the result of process. Voice recognition methods are implemented to convert speech into raw text. The application allows users to use local file as well as real-time speech as input. After transforming the speech into text, text processing methods are implemented to analyze the content and filtered into corresponding blanks in EHR form. Users also have options to save filled in EHR form to local directory. All the basic functional requirements listed in Chapter 3.1 are met. However, the accuracy of voice recognition part and text extraction part can be improved to achieve a better performance. Security also needs to be considered to ensure the safety of personal information.

Chapter 7

Progress

7.1 Project Management

This project is divided into four stages: Planning and Feasibility Analysis, Project Design, Project Implementation and Integration and Test and Maintenance. Project Implementation is the most important part of the project because it requires code writing as well as interim report writing. At this stage, most of the functions should be implemented as designed and an interim report should be written to review and reflect on the process.

For the original timetable, the time period is incorrect which did not match the schedule of this module. The timing for interim report and final report is earlier than the plan. In addition, the working speed is faster than the original plan and more works have been done in a shorter period. So, the timetable is changed as shown in Figure 6.2.

Stage 1 Planning and Feasibility Analysis

1.1 Ethics form signed and approval.

1.2 Requirements identification and specification of this project.

1.3 Articles reading and accomplishment of proposal.

Stage 2 Project Design

2.1 Decision of the format and content of database.

2.2 Further articles and websites reading to find useful voice recognition and nlp libraries or API.

2.3 User interface design for the software.

Stage 3 Data Collection and Implementation

3.1 Collection of doctor diagnose prescription and/or conversation contents online for creation of audio type dataset.

3.2 Development of user interface of EHR form with all basic functions provided.

3.3 Implementation and combination of speech recording, voice recognition, and Chinese language processing.

*3.4 Interim report writing and validation.

3.5 Classification of keyword from the result of Chinese language processing.

Stage 4 Test and Maintenance

4.1 Test on speech recording, voice recognition, and Chinese language processing functions.

4.2 Test on the performance of keyword classification.

4.3 Test on the application interaction.

4.4 Maintenance of repository, implementation of advanced function and optimization of the code.

*4.4 Final report writing and validation.

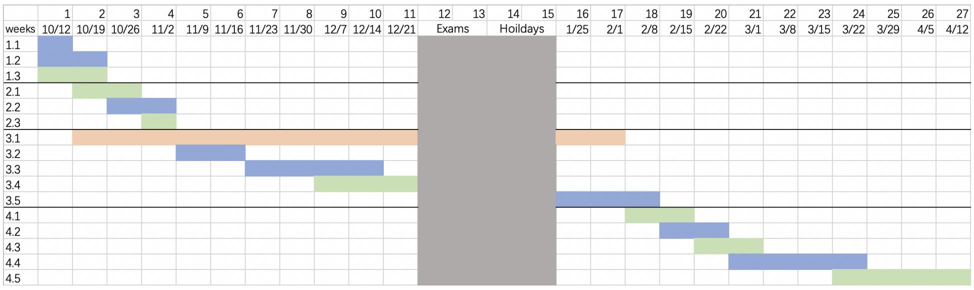

Figure 7.1: Original Timetable.

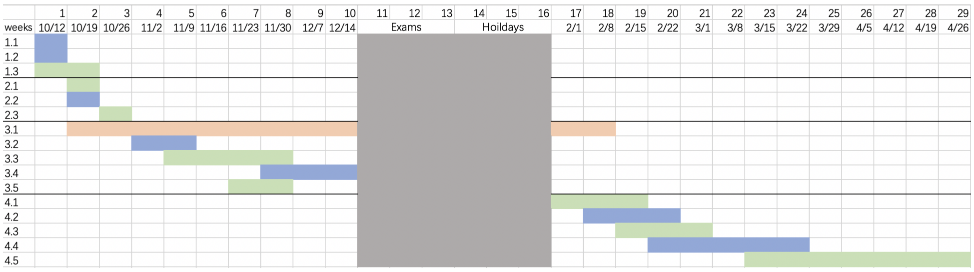

Figure 7.2: New Timetable.

7.2 Contribution

More than twenty articles have been read and database is designed and collected for more than thirty cases of doctor diagnose description for now. Speech recording, voice recognition, and Chinese language processing methods are implemented. Also, a user interface has been created to give a brief demonstration of the results form voice recognition, and Chinese language processing methods.

7.3 Future work

1. More cases of of doctor diagnose description should be collected.
2. Test on the whole application is needed.
3. Test on the UI interaction is needed.
4. Possible to implement a neural network on the prediction of disease based on the description of symptom.

7.4 Reflections

After testing, the scope of this project increased to include deep learning method for the prediction of disease. This additions was made to increase the practicability of the EHR to make it not only save the time for doctors and also save time for patients. Additionally, the level of project difficulty was increased because of those functionalities. Therefore, time management is revised to ensure more time is given to the new functionalities in the second half of the project. The speed of the project development is good and even faster than the original plan. Deadlines are strictly maintained by all deliverable submitted on time. Works that have been done during the first weeks covered most parts of the related field research including EHR usage, voice recognition researches on EHR, AI methods for medical usage and clinical information extraction methods. Later, the process followed development stages. It is now important to continue working on the project with the same speed in the next stages and try best on finish the addition.

References

[1] T. Shu, H. Liu, F. R. Goss, W. Yang, L. Zhou, D. W. Bates, and M. Liang, "Ehr adoption across china's tertiary hospitals: A cross-sectional observational study," *International Journal of Medical Informatics*, vol. 83, no. 2, pp. 113 – 121, 2014.

[2] J. Owusu-Marfo, Z. Lulin, H. A. Antwi, and M. O. Antwi, "Electronic health records adoption in china's hospitals: A narrative review," 2019.

[3] Y. A. Kumah-Crystal, C. J. Pirtle, H. M. Whyte, E. S. Goode, S. H. Anders, and C. U. Lehmann, "Electronic health record interactions through voice: a review," *Applied clinical informatics*, vol. 9, no. 3, p. 541, 2018.

[4] J. King, V. Patel, E. W. Jamoom, and M. F. Furukawa, "Clinical benefits of electronic health record use: national findings," *Health services research*, vol. 49, no. 1pt2, pp. 392–404, 2014.

[5] T. D. Shanafelt, L. N. Dyrbye, C. Sinsky, O. Hasan, D. Satele, J. Sloan, and C. P. West, "Relationship between clerical burden and characteristics of the electronic environment with physician burnout and professional satisfaction," in *Mayo Clinic Proceedings*, vol. 91, no. 7. Elsevier, 2016, pp. 836–848.

[6] G. Li, Y. Qian, Y. Huang, J. Chen, and X. Bai, "Building electronic health record using voice recognition and big data techniques," in *2nd Symposium on Health and Education 2019 (SOHE 2019)*. Atlantis Press, 2019.

[7] J. He, S. Baxter, J. Xu, J. Xu, X. Zhou, and K. Zhang, "The practical implementation of artificial intelligence technologies in medicine," *Nature Medicine*, vol. 25, 01 2019.

[8] E. Topol, "High-performance medicine: the convergence of human and artificial intelligence," *Nature Medicine*, vol. 25, pp. 44–56, 2019.

[9] S. Doyle, "Determining voice recognition accuracy in a voice recognition system," Feb. 23 2010, uS Patent 7,668,710.

[10] PyPI.org, "Google voice recognition api," https://pypi.org/project/SpeechRecognition/, accessed Dec 10, 2020.

[11] Baidu.com, "Baidu voice recognition api," https://ai.baidu.com/tech/speech, accessed Dec 10, 2020.

[12] W. Che, Z. Li, and T. Liu, "Ltp: A chinese language technology platform," in *Coling 2010: Demonstrations*, 2010, pp. 13–16.

[13] X. Chen, Z. Liu, L. Wei, J. Yan, T. Hao, and R. Ding, "A comparative quantitative study of utilizing artificial intelligence on electronic health records in the usa and china during 2008–2017," *BMC Medical Informatics and Decision Making*, vol. 18, 12 2018.

[14] Iflytek.com, "Iflytek voice ehr system," https://www.iflytek.com/health/zlwyl, accessed Dec 10, 2020.

[15] Y. Wang, L. Wang, M. Rastegar-Mojarad, S. Moon, F. Shen, N. Afzal, S. Liu, Y. Zeng, S. Mehrabi, S. Sohn, and H. Liu, "Clinical information extraction applications: A literature review," *Journal of biomedical informatics*, vol. 77, pp. 34–49, 2018.

[16] fxsjy, "Jieba python api," https://github.com/fxsjy/jieba, accessed Dec 10, 2020.

[17] Amazon.com, "Alexa voice assistant of amazon," https://www.digitaltrends.com/home/what-is-amazons-alexa-and-what-can-it-do/#:~:text=While%20Alexa%

20is%20the%20official%20name%20for%20Amazon\OT1\textquoterights,to%20be%20Alexa%20or%20something%20that%20sounds%20similar, accessed Dec 10, 2020.

[18] Saykara.com, “Saykara mobile ai assistant for physicians,” https://www.saykara.com, accessed Dec 10, 2020.

[19] Suki.ai, “Suki ai-powered, voice-enabled digital assistant for doctors,” https://www.suki.ai, accessed Dec 10, 2020.