

Progress Report

1. Task Introduction

My role in this project is to build a conversation generation system in collaboration with other group members. Ideally, the conversation generation system should feature: a. human-like conversation; b. diagnosing function; c. be able to connected with the speech-to-text system and the text-to-speech system.

2. Activities and Outputs

In these weeks, I mainly did three things:

- Read 2 survey papers for my literature review;
- Took a corsera course to study Deep Learning(DL) systematically ;
- Attended meetings with my supervisor and other students.

Through this activities, I start comprehending what happened in the areas of natural language processing(NLP) and DL in Healthcare. Also, taking online courses is a training for my furture work in implementing NLP task in the project.

Activity 1: Literature Review

The first paper is called "A guide to deep learning in healthcare". In the paper, the authors present 3 deep-learning techniques applied a lot for healthcare in recent years. For our project, the discussion on NLP in healthcare is helpful. Detailed information can be seen in [Table 1](#). The other paper, "Pre-trained Models for Natural Language Processing: A Survey", mainly focus on pre-trained models(PTMs) paper collection and summarization. From this paper, several PTMs applications and relevant techniques can be explored for the NLP in our project. A discuss of the output of this paper reveals in [Table 2](#).

DL Techniques	Models	Tasks	Limitations
CV	CNNs	<i>Medical images classification</i> object detection * Image-level diagnostics	<i>lack of clinial context</i> Small and poor-labeled datasets
NLP	- RNNs - CNN, RNN - information extraction models, - RNN-based language translation	- EHR building - incidents prediction - clinical voice assistants	- unstructured data - how well techniques derived from ICU patinets data generalize to broader populations - classify attributes from conversation
RL	Deep RL	robotic-assisted surgery	<i>correctly localizing position</i> data collection

Table 1, A summerization for techniques discussed in the paper "A guide to deep learning in healthcare", including the referred DL techniques, most applied models in the area, possible tasks in healthcare, as well as limitations in implementing the techniques to solve healthcare problems.

NLP tasks	Applications
General Evaluation Benchmark	QA system improvement
Question Answering	Helthcare conversation
Sentiment Analysis	Not very related
Name Entity Recognition	Diagnosing information extraction
Machine Translation	Unclear

Table 2, Several NLP tasks referenced in the survey "Pre-trained Models for Natural Language Processing: A Survey" and possible applications in our project.

Activity 2 -- Training

Apart from literature review, I also attend an online [course](#) -- Neural Networks and Deep Learning. For now, I have understood basic concepts about Neural Networks(NN), Logistic Regression.

Activity 3 -- Meetings

Minutes of meeting with Professor Chang, March 12 2021

- Topic Discussed:
 - * Brief introduction **for** the project **and** Xiaoxiao's role;
 - * Possible solution for the healthcare chatbot;
 - * Dataset for the project;
- Todo List:
 - * Exploring existing projects similar like ours;
 - * Be familiar with existing researches and works;
 - * Preparing for Initial Review;

Minutes of meeting with Ruibin, Jiajun and Kavi, March 9 2021

- Topic Discussed:
 - * Ruibin introduced his work for the project;
 - * The group discussed the challenging for the project (such as algorithms, dataset collections)
- Todo List:
 - * Everyone keep doing literature review and exploration;
 - * Kavi Initiate a project on Github;
 - * Everyone upload codes to Github;

[1] Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., ... & Dean, J. (2019). A guide to deep learning in healthcare. Nature medicine, 25(1), 24-29.

[2] Qiu, X., Sun, T., Xu, Y., Shao, Y., Dai, N., & Huang, X. (2020). Pre-trained models for natural language processing: A survey. Science China Technological Sciences, 1-26.