Xiaowei Huang

(+86) 135-4020-5963 · 1446269148@qq.com · GitHub@kitaharatomoyo ·

Personal Summary

I'm optimistic, responsible, self-motivated, love to try new things, committed to the intensive study of the core technology of natural language processing and its application in products. I have been engaged in medical dialogue diagnosis research based on reinforcement learning in school, I am very interested in the implementation of intelligent dialogue system. **Now studying in Harbin Institute of technology (Shenzhen)**。

Educational Background

Harbin Institute of technology (Shenzhen), Computer science and technology, *Academic Postgraduates* 2019.9 - 2022.1

Ranking Top 10%, Expected to graduate in January 2022

Sichuan University, Computer science and technology,*Bachelor of Engineering* 2015.9 - 2019.6

Ranking Top 5%, National Scholarship

Professional Skills

- Familiar with Linux platform, C++, Python
- Familiar with common machine learning algorithm theory and NLP related deep learning model (word2vec, Bert, etc.)
- Familiar with common machine learning and deep learning algorithm package (scikit learn, pytorch)
- Basic English literature reading ability (CET-6)

Internship experience

Guangdong Laboratory of cyberspace science and technology, Shenzhen, PCL, Wise Information Technology of Med Department 2020.7-2020.12

- **Overall Project**: Dialogue system related to the new coronavirus epidemic prevention. The recall rate reached TOP1 84%， TOP3 98%， TOP5 100%
- Independently responsible for the cleaning and expansion of data sets and the construction of negative samples. Through short text clustering, design template problem, Label data according to template questions. Problems in the same template can be replaced with each other. Through short text clustering, the samples of different classes are defined as "pseudo" negative sample pairs.
- Independently responsible for the training of candidate recall and problem matching model. Using BM25 recall candidate question and answer pairs. Using abcnn model for training.

Beijing Megvii Co., Ltd., Logistics algorithm Department 2018.10-2018.12

- **Overall Project**: Research on path planning algorithm of robot in logistics warehouse
- Responsible for order, shelf, robot simulator design and implementation. Use C + + to design business scenario simulator to test the effect of the algorithm.
- Responsible for exploring advanced path planning algorithm. This paper mainly explores how to design a flexible path planning algorithm for dynamic orders.

Competition Awards / Project Works

- JDDC2020 Jingdong E-commerce Multi-modal Multi-round Dialogue Competition**3/53** (*http://jddc.jd.com/rank*) 2020.7-2020.10
  - Independently responsible for the design and training of retrieval based question answering module. BM25, DSSM and Polyencoder are used to retrieve candidate QA pairs.
  - Independently responsible for the design and training of generative question answering model based on pretrain model. GPT-2 is used as the generation model, and entropy gain is used to select the generation and retrieval answers.
  - Responsible for exploring the generative question answering model combined with retrieval answers. Explore the use of copy mechanism combined with the retrieval of the question and answer pairs to build the model.

- ACM-ICPC Regional competition Nanning competition area **Silver medal**, November 2017
- ACM-ICPC Regional competition Xi'an competition area **Silver medal**, October 2017
- CCF-CCPC Regional competition Changchun competition area **Silver medal**, October 2016