

Zongmin Zhang

+86-17740192857 / zongzhang6-c@my.cityu.edu.hk

Education Background

➤ City University of Hong Kong	Hong Kong
MSc in Computer Science (Current GPA: 3.19/4.3)	Aug. 2022 – Oct.2023
➤ The University of New South Wales	Sydney, Australia
BSc in Computer Science – DataBaseSystems (GPA: Credit/ Second Class Honor)	Aug. 2018 - Sep. 2021

Research Experiences

➤ Hong Kong Institute of Science & Innovation, Chinese Academy of Sciences (CAIR)	Oct.2023 – Present
--	--------------------

Research Assistant

- Trying to train our LLM from scratch, crawling and cleaning medical books and content as a corpus for continued pre-training, using GPT3.5 to build the corpus for supervised fine tuning, and trained our own embedding model for RAG system to better extract medically relevant content.
- Working on how large language models can accurately invoke different tools from a large number of tools in order to satisfy the user's query without affecting the normal dialog functionality.
- Working on image segmentation of the multi-label abnormal and surgical instruments on the respiratory microscopy, using nnUNet, TransUNet, and SAM to train and predict.
- Do data washing and cleaning for more than thousands of patient cases of respiratory microscopy and manually label hundreds of patient cases of respiratory microscopy for different types of abnormal.

➤ AI in Fintech	Jan. 2023 – July 2023
------------------------	-----------------------

Supervisor: CityU - Dr.Linqi Song

- Working on a Financial Machine Translation project aimed at document-level machine translation in the financial sector.
- Try some large language pre-trained models like mBart-large, mBart-base, mT5-base, and LLaMa with different sizes of tokenizers and decoder strategies to get the SOTA performance in machine translation in the financial area.
- Responsibilities include collecting training data like crawling financial Chinese-English pairs data from Financial Times and other Financial websites, fine-tuning different models, and testing different parameters for the best performance. Try decoding strategies, like nucleus sampling, beam search with A star algorithm, and reweighting attributes for various financial and general situations to improve machine translation performance.

Publications

- Zhen Chen*, Zongmin Zhang* , WenWu GUO, Xingjian LUO, Long Bai, Jinlin Wu, Hongliang Ren, Hongbin Liu, **ASI-Seg: Audio-Driven Surgical Instrument Segmentation with Surgeon Intention Understanding**, *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. - Submitted on 2024
- Xingjian Luo*; You Pang*; Zhen Chen; JinLin Wu, Zongmin Zhang; Zhen Lei; Hongbin Liu, **SurgPLAN: Surgical Phase Localization Network**, *The IEEE International Symposium on Biomedical Imaging (ISBI)*. -Accepted on 2024

* means equal contribution

Work Experiences

- **Hong Kong Institute of Science & Innovation, Chinese Academy of Sciences (CAIR)** Jun.2023 – Aug.2023

Vision Algorithm Intern

- Supervised-finetuning a medical surgery large language model, using LLaMa-2-70b-chat as the base model to chat, and using Lora, QLora mechanisms to train the model, VLLM mechanism to accelerate the inference speed while occupying less graphics memory.
- Collect thousands of medical textbooks and top medical publications as raw data, then use the raw data with GPT-4's API to generate question-answer pairs as training data.
- Build and maintain a website, let users talk with our LLM, and develop new functions: frond-end built by React and back-end made by Flask.

- **Research Assitant in CityU** Sep.2022 – June.2023

Supervisor: CityU - Prof. Antoni B. Chan

- Working on a Plagiarism Detection System project.
- Developed basic features for the system's backend (Python Flask) and front end (React).
- Using the OCR mechanism and manually correcting the content to pre-process data for training.

- **Beijing Sankuai Online Technology Co., Ltd. (known as Meituan)** Feb. 2021 – Aug. 2022

System Development Intern

- Developed a Java SpringBoot-based financial reporting system for the company's business developer.
- Conducted essential backend development tasks, including writing scheduled tasks and new features' code.
- Participated in project iterations, functional development, and design of new features, mainly responsible for data monitoring.
- Created and maintained unit and integration tests for new and existing functions.
- Adding a dialogue function to our system allows business developers to query our system's status and information.

Selected Projects

- **BookIt ([link](#))** May 2021 - Jul. 2021

- Developed a one-stop platform for booking resources, featuring a Python Flask backend and local MySQL storage and using HTML, AJAX, and CSS as frontend.
- Implemented backend functionality and improved corresponding frontend pages.
- Realized essential features like booking, commenting, liking, favoriting, and resource management.

- **Signature Index ([link](#))** Mar. 2021 - Apr. 2021

- Implemented a signature-indexed file in C, including creation, insertion, and partial-match retrieval query functionality.
- Developed two signature algorithms and a hash function for accurate database 'Select' operation.
- Achieved 80% accuracy in data insertion and retrieval.

Additional Information

- Know Java, React, Python, TypeScript, MySQL, SpringBoot, Flask, HuggingFace, PyTorch, and ES.
- Speak Mandarin (native) and English (fluent).