Raoyi (Cathy) Huang

(+1) 412-932-7290 | huangraoyi66s@gmail.com | linkedin.com/in/cathy-huang-5b7719189/ | github.com/Huang-Raoyi

EDUCATION

Carnegie Mellon University (CMU)   **Pittsburgh, PA, USA**
Master of Information Systems Management: GPA 3.96/4.0   Sep. 2022 – Dec. 2023 (Expected)

The Hong Kong Polytechnic University (PolyU)   **Hong Kong**
*Bachelor of Science in Computer Science **First Class Honours**: GPA 4.01/4.3, ranking 1/143*   Sep. 2019 – Jun. 2022
- **Honors**: Dean's Honors Award, Graduating Class Ambassador 2022.
- **Exchange Programs**: School of Computing at National University of Singapore (4.83/5), Pembroke College and King's College at University of Cambridge (A/A)

WORK EXPERIENCES

Bayer   Hanover, NJ, USA
Mobile Application Developer   Aug. 2023 – Present
- **In progress:** developing a cross-platform mobile application myRadiology360 by **Flutter** with **Firebase** backend**.**

The Hong Kong University of Science and Technology (HKUST)   Hong Kong
Full-stack Developer   June 2023 – Present
- Developed a full-stack interactive chatGPT-empowered web application to co-create digital blogs with the elderly; Built the frontend by **React**, integrated with a **FastAPI** backend and **MangoDB** database.
- Deployed the web application on **AWS** for scalability; Implemented **ELK stack**; Improved issue diagnosis speed by 40%.

Department of Computer Science and Engineering, HKUST   Hong Kong
Research assistant | Advisor: Prof. Huamin Qu   May 2023 – Present
- Built a ChatGPT-empowered automatic pipeline to extract visual and data encodings from **SVGs**, covering SVG preprocessing and compression, chart type classification, element role classification, and **Vega-lite** encoding generation.

School of Computer Science, CMU   Pittsburgh, PA, USA
Research assistant | Advisor: Prof. Tongshuang Wu Sherry   Nov. 2022 – Present
- Researched, summarized, and implemented NLP dataset characterization metrics in difficulty, disagreement, and diversity.
- Profiled and filtered NLP evaluation datasets on different metrics by **clustering**, intending to better characterize adversarial datasets.

USMLESarthi   Pittsburgh, PA, USA
Technology Developer Intern   May 2023 – Aug. 2023
- Transferred company website from **Angular** frontend to **React** frontend; Developed a web application prototype with a GLM data model by **Jakarta EE** and **TomEE** to crowdsource and predict candidate program matching probabilites based on their profile data; Improved data collection and refinement efficiency by 15% and prediction accuracy by 12%.

Institute of Software Application Technology, Guangzhou & Chinese Academy of Science   Guangzhou, China
Machine Learning Engineer Intern   Jun. 2021 – Jul. 2021
- Detected streetlight anomalies regularly and automatically via **LSTM** trained on streetlight data in **Mongo** and **MySQL** databases by **PyTorch**; Visualized real-time predictions via **Tableau** and **Matplotlib**; *Achieved an F1 score of 0.82*.
- Built structured unit tests, **Tencent ASR**, **HMM-based** part-of-speech recognition, and **DNN-based** intent recognition for a light cloud chatbot system for a Smart City platform; Integrated **Baidu hotword detection package snowboy** with **Tencent STT** and **TTS**; Developed algorithms to improved data retrieval efficiency by 19% and acoustic input range by 17%.

DataStory   Guangzhou, China
Data Engineer Intern   Jun. 2020 – Jul. 2020
- Developed **web-based dynamic sales panels** for daily necessities of multiple home-delivery platforms by **React, Node.js and D3** and an AI-enabled algorithm for automatical resource distribution to help medium to small companies get through COVID-19; Won *3rd Prize in 10th National University Student E-commerce Competition for Innovation, Creativity, and Entrepreneurship*.
- Crawled social media data via **BeautifulSoup**, designed keywords, extracted key information, analyzed the data multi-dimensionally, and optimized operation strategies for clients (iQOO, Yili); The outcome was appreciated and praised by the clients.

Department of Computing, PolyU   Hong Kong
Student Assistant | Advisor Prof. Maggie Wenjie Li   Jul. 2021 – Aug. 2021
Vision, Language, and Graphics Group: An AI-Driven Platform for Smart Living
- Curated a multi-modal conversational question-answering dataset; decomposed the complex questions into simple ones.

Reinforced Reading, Reasoning, and Writing for Question-focused Summarization   Feb. 2021 – Jun. 2021
- Summarized the evolutionary history of Neural Machine Reading Comprehension.

PROJECT EXPERIENCES

ProChat | An Elderly-companion Humanoid Chatbot on Misty II Robot, PolyU   Jun. 2021 – May 2022
Advisor: Prof. Jiannong Cao, Prof. Yuanqing Zheng, Prof. Maggie Wenjie Li
- Enabled proactive, emotional, and multi-modal communication; Won ***Champion in the Inter-departmental Capstone Project Competition*** and showcased at the ***PolyU Outstanding Work by Students Exhibition***.
- Developed a chatbot framework with a **Flask backend server** for easy integration and separation of multi-modal modules.
- Built a non-task-oriented module based on **BlenderBot1.0** and a task-oriented chatbot to send emergent messages by **Twilio** and **Imgur**, integrated a remote video streaming App in Python to the robot, and connected to **WeatherStack** for weather reporting; .
- Designed rule-based responses based on psychology consulting theory for the task-oriented chatbot module; divided interaction process into exploration, comfort, and suggestion stages, where each stage involves a different combination of strategies such as questioning, self-exposure, emotion mapping, and confirmation.

LegiFind | A Search Engine for Legal Case Retrieval, NUS   Jan. 2021 – May 2021
- Built **phrasal inverted indexes** of positional lists and zone information, compressed indexes by **gap encoding**, merged indexes by **two-way SPIMI**; *Achieved 54% compression ratio*.
- Implemented **boolean query, phrasal query, proximity query, and free text query** with **BM25 and Indri's query likelihood model** with two-level smoothing, integrated with **pseudo relevance feedback** by Indri, and **query expansion with WordNet**; *Achieved 23% higher average MAP*.

ToxicGuard | A Toxic Comment Detection System, NUS   Jan. 2021 – May 2021
- Engineered textual features via **pretrained Spacy, self-trained word2vec, BERT [CLS] token and max pooling embedding** as well as **RandomForest feature importance**.
- Classified multiple labels by **Binary Relevance, Adaptive Algorithm, OneVsRest, biLSTM,** and **stacking**, tuned hyperparameters by **RandomizedSearch and Bayesian Optimization**; *Outperformed base dummy classifier by 0.43 F1 score*.

Simulation for Asymptomatic Covid-19 Cases in Hong Kong, PolyU   Sep. 2020 – Jan. 2021
- Utilized correlation matrix, MLP, and SVM to analyze the influential factors of asymptomatic Covid-19 cases; Simulated those cases by SEIR model with Gillespie algorithm and HK-specific parameters.

HONORS AND AWARDS

- Champion in Best Project Award Competition for Inter-departmental Final Year Project competition, PolyU
- The Professional Validation Council of Hong Kong Industries Scholarship, Hong Kong
- Silver Price in the 7th China International College Students' 'Internet+' Innovation and Entrepreneurship Competition, Ministry of Education in China
- HKSAR Government Scholarship -- Reaching Out Award, Hong Kong
- Wong Tit-shing Exchange Scholarship, PolyU
- 3rd Prize in the 10th National University Student E-commerce Competition for Innovation, Creativity, and Entrepreneurship, Ministry of Education in China
- IAO Summer Abroad Subsidy for Summer@Oxbridge, PolyU

SKILLS

Programming: Python (Proficient), Java (Proficient), SQL (Advanced), C/C++, Shell Script, R, Scheme, Haskhell, Prolog, JavaScript.
Tools & Libraries: Git, Docker, Hadoop, Spark, Flask, MS Azure, MongoDB, MySQL, Navicat, React, RESTful API, Android Studio, Pytorch, TensorFlow, Scikit-Learn, Numpy, SciPy, Pandas, Seaborn, HuggingFace, Tableau, STATA
Languages: English (IELTS: 7.5, GRE: 156 + 170); Mandarin (Native); Cantonese (Basic).

HONORS AND AWARDS

Last Guard, **China Mainland Student Association Women Basketball Team**, PolyU   Sep. 2018 – Jun. 2022
- 1st runner-up in Hong Kong Mainland Undergraduate Basketball League, Hong Kong

Publication Secretary, **Hall Community Service Group**, PolyU   Sep. 2019 – May 2020
- Organized Beach Rubbish Collecting Activity in Stanley, Hong Kong.
- Organized Re-energizing Workshop to relax students before exams at PolyU.
- Organized Voluntary Teaching Activity in Guizhou, Mainland China.

Vice project manager, **Enactus Entrepreneurial Club**, PolyU   Sep. 2018 – Dec. 2019
- Participated in project *InterVictor* to build a platform to offer university candidates intensive and inclusive overseas interview suggestions.