

MENG, Rui

Tel: +86 15066282180 | Email: ruimeng@link.cuhk.edu.cn

EDUCATION

Chinese University of Hong Kong, Shenzhen (CUHK-SZ) **09/2024-06/2028 (Expected)**
 Bachelor of Computer Science and Engineering *Shenzhen, China*

- Core Courses: Data Structures, Algorithms, Database System, Machine Learning, Artificial Intelligence, Speech and Language Processing, Image Processing and Computer Vision, Bioinformatics.
- Honors: Bowen Scholarship (30% Tuition Fee Waiver) in AY2022-2026.

SKILLS

SOFTWARE: Python, C++, Java, R, MATLAB, SQL, Linux, Git, PyTorch, scikit-learn, SAS, Microsoft Office.**TECHNICAL:** Machine learning experimentation, LLM evaluation, data analysis, technical documentation, troubleshooting.**LANGUAGE:** English (Fluent), Chinese (Native)

RESEARCH EXPERIENCE

Research Assistant | Patch-Paradigm Router for Automated Program Repair **03/2026-Present**
 CUHK-SZ, School of Artificial Intelligence | Mentor: Prof. Pinjia He *Shenzhen, China*

- Developed a routing framework to classify code-change patches and cluster bug fixes into different repair paradigms for learning-based APR.
- Represented buggy code, fixed code, test information, and search/replace patches as structured features; applied code anonymization to reduce project-specific naming bias.
- Built a clustering and diversity-sampling pipeline to identify representative patch groups and train a router for selecting suitable modification paradigms.
- Evaluated routing signals with LLM likelihood, avg NLL, and perplexity, comparing REW, EDIT, and mixed strategies under Defects4J and Agentless-style settings.

Research Assistant | Triple-Stage Drug-Drug Interaction (DDI) Predictor **01/2025-05/2025**
 CUHK-SZ, Warshel Institute for Computational Biology | Mentor: Prof. Hsien-Da Huang *Shenzhen, China*

- Designed a triple-stage DDI prediction pipeline for occurrence prediction, interaction-type classification, and severity prediction.
- Integrated Mole-BERT features, graph representations, and cross-attention mechanisms to improve drug-pair feature fusion.
- Applied cross-attention mechanisms to refine feature extraction and optimized model parameters through backpropagation.
- Analyzed large-scale pharmaceutical datasets to build robust prediction models for drug interactions and computational drug repurposing.

Research Assistant | High-Performance Preoperative mRS Scoring Agent **09/2025-02/2026**
 CUHK-SZ, Shenzhen Research Institute of Big Data | Mentor: Prof. Renzhi Wang *Shenzhen, China*

- Preprocessed cranial CT data with noise reduction and threshold-based screening to improve downstream analysis quality.
- Built a multi-task model integrating CT scans and clinical data for intracranial hemorrhage diagnosis and mRS scoring.
- Investigated foundational Mamba models for CT image classification and implemented architectural improvements for medical image analysis.
- Combined visual encoders with LLMs to support medical-image interpretation and diagnostic decision support.

INTERNSHIP EXPERIENCE

Data Analyst Intern | Alzheimer's Disease Classification Using Deep Learning **06/2025-08/2025**
 Pfizer (China) Co., Ltd. | Mentor: Xiaomeng Cui *Shanghai, China*

- Developed SAS and R programs to analyze clinical-trial data and check statistical outputs against study objectives.
- Contributed to clinical-trial design and statistical analysis planning, including sample size estimation and statistical modeling.
- Performed quality-control checks on statistical outputs and collaborated with cross-functional teams to generate actionable insights.
- Applied machine-learning methods to clinical datasets and contributed to biomarker identification for Alzheimer's disease.

COURSE PROJECTS

Poetry Generation Model Design **03/2025-05/2025**
 DDA4210 Advanced Machine Learning Course Project *Shenzhen, China*

- Built a SeqGAN-based Chinese poetry generation model in PyTorch with bidirectional LSTM and attention mechanisms.
- Designed a hybrid adversarial and reinforcement-learning training framework using Monte Carlo reward estimation.
- Achieved 89% tonal compliance by incorporating trainable metrical templates and dual-attention mechanisms.