

Yaoyuan Hu

Phone Number: +852 92077602 | Email Address: huyaoyuanq@gmail.com | Personal Website: yaoyuan-hyy.github.io

EDUCATION

Chinese University of Hong Kong (CUHK)

Hong Kong SAR, China

BSc in Computer Science

Sep 2021 – Jun 2026

- **CGPA:** 3.52/4.0
- **Award:** ELITE Stream, Wong Lee Sau Ying Memorial Scholarship (9000 HKD)
- **Concentration:** Deep Learning, Large Language Model, AI4Math, Formal Mathematics Reasoning
- **Key Modules:** Data Structures, Artificial Intelligence, Machine Learning, Software Engineering, Database Systems, Design and Analysis of Algorithms, Formal Languages and Automata Theory, Operating Systems, Human-computer Interaction

RESEARCH EXPERIENCE

THE CHINESE UNIVERSITY OF HONG KONG

Hong Kong SAR, China

Natural Language and Formal Language Reasoning Enhancement for Lightweight Model

May 2025 – Aug 2025

- Developed FormalMath-Cross, a novel framework enabling 1.5B-parameter models to perform automated theorem proving in Lean4 via dual-modal reasoning synergy.
- Engineered a dual stage fine-tuning pipeline:
Stage 1: Adapted Qwen2.5-Math-1.5B to Lean4 syntax using DeepSeek-Prover-V1 (LoRA: 0.1% new params, r=16).
Stage 2: Aligned natural language reasoning with formal proofs via Goedel-Prover-annotated MiniF2F data.
- Achieved 21.3% formal proof accuracy on MiniF2F benchmark (vs. 0% baseline), reducing candidate proofs by 4× vs. 7B models.

PROFESSIONAL EXPERIENCE

THAI WAH PUBLIC COMPANY LIMITED

Bangkok, Thailand

Data Analytics and Strategy Internship

Jun 2024 – Dec 2024

- Applied the Long Short-Term Memory (LSTM) model to forecast commodity prices, enhancing prediction accuracy by approximately 15% to support operational and strategic decision-making.
- Developed a web application utilizing Principal Component Analysis (PCA) and Partial Least Squares Regression (PLSR) to analyze product composition and the effects of additives on product properties, enabling interactive visual analysis.
- Fine-tuned GPT-4o to create an internal chatbot deployed via a Line official account, facilitating queries on product knowledge and market data, and handling an average of 200+ user inquiries daily.
- Participated in designing and implementing an A/B testing framework to compare the effectiveness of different algorithmic strategies in real business scenarios, driving data-driven iterative optimization.

QI-AN XIN TECHNOLOGY GROUP INC.

Chengdu, China

Data Analyst Intern

Jul 2023 – Aug 2023

- Researched recent data leakage and trafficking activities by developing and deploying Python/Scrapy scripts to automate monitoring and data collection from specific sources like dark web forums and social platforms, building a foundational raw data pool.
- Performed data extraction, transformation, and loading (ETL) on collected unstructured and multi-source heterogeneous data, including deduplication, format standardization, and filtering of invalid information, to provide high-quality datasets for subsequent analysis.
- Established a data classification and labeling system based on project goals; applied basic Natural Language Processing (NLP) methods (e.g., keyword extraction, topic modeling) and rule engines for automated tagging and categorization to prioritize threat intelligence.
- Aggregated and classified data sources by relevance to the project, enhancing team prioritization and organizational efficiency.
- Contributed to developing data visualization dashboards (using Power BI) to intuitively present analytical results and project progress; automated the generation of daily summary reports, reducing manual report consolidation time by approximately 40%.

EXTRACURRICULAR ACTIVITIES

CU HUB

Hong Kong SAR, China

Project Programmer

Jan 2023 – Apr 2023

- Led the front-end development for all administrator interfaces and the user profile page using React.js within a five-person Agile team. Employed a component-based architecture to build reusable UI elements, significantly accelerating the development process and ensuring consistency across the application.
- Collaborated with back-end developers to define and consume RESTful APIs, enabling dynamic data rendering and form submissions on the user profile page. Focused on creating an intuitive and responsive user experience.

ADDITIONAL INFORMATION

- **Languages:** Mandarin (Native), English (Fluent), Cantonese (Elementary)
- **Technical Skills:** Python (PyTorch, Scikit-learn), C, C++, Java, JavaScript, React.js, R, SQL, MS Office, Power BI
- **Interests:** Music, Travel, Photography