Zihan Qi

- **Mobile: (852) 61948792**
- **Email: normalman743@gmail.com**
- **GitHub: github.com/normalman743**
- **LinkedIn: zihan-qi-115b15301**

Education

The Chinese University of Hong Kong

BSc in Statistics, Minor in Computer Science | ***2022 - Present (Expected: July 2026)***

Academic Performance:

- Cumulative GPA: 3.616 / 4.0
- Major GPA (Statistics): 3.787 / 4.0

Honors:

- Dean's List, Faculty of Science, 2023–2024 & 2024–2025
- Head's List, Chung Chi College, 2024–2025

Awards:

- Alex Chung Shun Chan Science Faculty Scholarship (Chung Chi College Departmental Prize, HKD 10,000), 2025/26 – competitive merit-based scholarship.
- Dept. of Statistics Scholarship： 2023/24 & 2024/25
- Dept. of Statistics & Data Science Scholarship： 2025/26

Relevant Coursework:

Computer Science & Machine Learning:

- Data Structures (CSCI 2100, A), Computer Organization (CSCI 2510, A), Intro to Operating Systems (CSCI 3150, A), Building Web Applications (CSCI 2720, A).
- Data Mining & Statistical Learning (STAT 4001, A-), Introduction to Data Science (SEEM 2460, A).
- Current Enrollment: Bayesian Learning (STAT 4010).

Statistical & Mathematical Foundation:

- Statistical Computing (STAT 3006, A-), Intro to Stochastic Processes (STAT 3007, A-), Programming Lang for Stat (STAT 2005, A), Linear Algebra I (MATH 1030, A-).

Research Interests:

- Artificial Intelligence, Data Science, Multimodal / Vision–Language Models, Large Language Models

Technical Skills

Programming: Python, R, C/C++, SQL, JavaScript, HTML/CSS
ML & DL Frameworks: TensorFlow, scikit-learn, PyTorch
Data Analysis & Visualization: pandas, NumPy, statsmodels, ggplot2, seaborn, Excel, Tableau
Databases: MySQL, MongoDB
Tools & Platforms: Git, Linux, Jupyter Notebook, VS Code, Terminal
LLM Tools: OpenAI (GPT), Anthropic Claude, Google Gemini; FAISS, Chroma, basic RAG pipelines

Research & Project Experience

Research Preparation | Multimodal Alignment & Foundation Models **CUHK · Dec 2023 – Present**

- Started systematic preparation in deep learning in late 2023, including the NVIDIA "Fundamentals of Deep Learning" workshop at CUHK and course-based/self-directed study of CNNs and basic deep learning concepts.
- Since Feb 2025, have been studying multimodal and transformer-based foundation models under the academic guidance of Prof. Xiangyu Yue (CUHK, MMLab), with focused reading on Transformers, BERT/GPT, ViT, CLIP, BLIP and related vision–language frameworks.
- In the current semester, began implementing preliminary experiments on representation alignment in CLIP and OneLLM (layer-wise cosine similarity analyses of matched vs. mismatched image–text embeddings), and received informal feedback on research direction from senior researchers, including the first author of OneLLM.

Backend Developer | ICU (Intelligent CUHK) Campus Q&A Assistant **CUHK · Jun 2025 – Dec 2025**

- Designed and implemented a modular FastAPI backend for a CUHK-specific study assistant and Q&A platform, following a layered architecture (API routers, services, models, schemas).
- Built a course–semester–folder–file hierarchy with local file storage and background processing (Celery) to extract, chunk and embed PDFs/HTML into per-course and global Chroma vector stores.
- Developed RAG services and chat/message handling to support course-specific and global question answering over uploaded lecture materials, with basic authentication, invite codes and audit logging.

Student Assistant | Auto Essay Grading Project **CUHK · Jul 2023 – Jan 2025**

- Built preprocessing pipelines (PDF → structured text) using AI-assisted extraction and manual checks to improve LLM input quality.
- Ran a user study comparing several LLM feedback variants, collected student ratings and backend logs, and analyzed clarity/accuracy/helpfulness scores.
- Iterated on prompt and workflow design, evolving from an early RAG prototype to a more robust prompt-based essay evaluation system.

RAG Backend Developer (Intern) | OPTISM ASD Support Platform **Remote · Apr 2025 – Sep 2025**

- Helped upgrade an autism-support Q&A platform with a new Flask-based RAG backend using OpenAI embeddings, FAISS and a MySQL knowledge base.
- Implemented streaming chat API (SSE), conversation logging and API-key based access control to improve reliability and monitoring.
- Designed multi-threaded FAISS index initialization and scheduled hot updates to support a growing bilingual question bank.

Coursework Projects | Statistics, Machine Learning & Software Engineering **CUHK · 2023–2024**

- Mainland China housing price forecasting and multi-disease prediction: led modelling using ARMA and smoothing methods on multi-year housing data, and deep models on a 4GB clinical dataset, including cleaning, feature engineering and evaluation.
- Full-stack cultural event management platform： built backend APIs (Node.js/Express, MongoDB) and worked with a React/TypeScript frontend for authentication, event management and map-based browsing.