Jethro Au

Hong Kong • +852-6336-5920 • jethrocsa@gmail.com • https://jethrocsau.github.io/

Education

(Incoming) Msc. in Health Data Science - Harvard University **December 2027 (Anticipated)**
Department of Biostatistics | Area of Study: Computational Biology, Bioinformatics , Reinforcement learning (MIT), Computer Vision (MIT)

Msc. in Big Data Technology - Hong Kong University of Science and Technology **In Progress - June 2025 (Anticipated)**
GPA: 4.0/4.3 | TPG Fellowship | HKUST School of Engineering (CS & Math)
Courses: Research in AlphaFold, LLM, Data Mining, Parallel Programming, Math. Foundations, Image Processing, Social Computing

B.S. Industrial Engineering - Northwestern University **2018**
McCormick School of Engineering (IE) , Certificate in Managerial Analytics - Kellogg School of Management
Courses: Stats & Machine Learning, Optimization Algorithms, Accelerated Physical Chemistry, Multivariate Calculus, Data Management

Research

Independent Research on AlphaFold for Cyclic Peptide Design **Jul 2024 – Now**
Hong Kong University of Science and Technology

- Under supervision of Prof. Nevin Zhang from the CSE department and Dr. Tao Wang (inventor of HIV antiretroviral drug Rukobia), researched use of AlphaFold for de-novo design of cyclic peptides for Hepatitis-C and HIV.
- Developed a novel technique that improved precision and control of generative drug design by backpropagating amino acid sequences from a targeting loss functions that generated binders that conformed to the CD4-binding site pocket.
- Conducted hyperparameter tuning on AlphaFold2 AfCycDesign and explored in-silico cyclic peptide designs methods using ColabDesign, comparing AlphaFold accuracy with traditional molecular simulation methods (Rosetta / Pymol / Autodock).
- Conducted literature review of generative drug design architectures , including RFDiffussion and AlphaFold (2 & 3).

Professional Work

R&D / Head of Business Development **Jun 2023 – Present**
C-POLAR Asia, Hong Kong (Health Technology)

- Digital infectious disease prevention: Researching and prototyping use of IoT sensors for continuous infectious disease monitoring with digital twins at hospitals with Dr. Anthony Law (HK PolyU) and industry partners. Researched methods for spatial and temporal analysis, such as statistical and machine learning methods to calibrate low-cost multi-array indoor air quality & occupancy sensor data for real-time aerosolized microorganism risk-scoring.
- Initiated research projects with universities & hospital laboratories on non-drug based methods to combat antibiotic-resistant organisms w/ cationic polymers that selectively disrupt protein membranes and envelopes.
- Developed antimicrobial products and coordinated comparative studies at long-term care facilities that statistically validate effectiveness of polymer-embedded healthcare applications in reducing Hospital Acquired Infection (HAI) risk.

Strategy & Operations Manager (Data Analytics & Special Projects) **Jun 2021 – May 2023**
Prenetics, Hong Kong (Biotech Unicorn - Listed on NASDAQ 2022)

- (COVID-19 2nd wave & onwards) Led RT-PCR diagnostic tests for all inbound travellers during COVID-19, a high-profile project implementing decentralized point-of-care screening as part of HK's COVID-19 health policy.
- Rapidly iterated, developed and implemented health data solutions to solve challenges during time of crisis:
  a. Digitally-augmented remote-testing practices with a telehealth mobile app and backend analytics for healthcare workers, such as automating data tasks via computer vision (OpenCV)
  b. Improved service delivery quality through a NLP-driven performance scoring algorithm, use of sentiment analysis with BERT (HuggingFace API) from personalized SMS satisfaction surveys (Twilio SMS API).
  c. Developed prediction models for 7-day caseloads to improve manpower rostering, combining data crawling, scheduling, and daily residual bias prediction with multivariate regressions, decisions trees, and statistics (Poisson/MA/EMA).
  d. Improved data quality of over +2M reported results and mitigated false-positive reporting risks due to noisy electronic health records through ETL & reporting infrastructure on dockerized Python scripts deployed on AWS server.
  e. Trained a mask-detection application (Flask) to augment PPE training by fine tuning YOLOv5 *(used for trial only)*
- Collaborated with the Department of Health to set the gold operational standard for RT-PCR testing and served as consultative expert on COVID-19 health policy changes.

Innovation Consultant (Data Analytics) **Sept 2018 – May 2021**
PwC

- (COVID-19 1st wave) Resolved bottlenecks in public health administration of the 1st tranche of a HKD$80B emergency relief scheme for all companies during COVID-19 lockdowns. Introduced a data-driven operations with data pipelines and dashboards to automate handling of voluminous applications (Python, PowerBI, UIPath)

Jethro Au

Academic Work

XAI - Detecting and Reducing Hallucinations in Factual Question Answering w/ Recursion **Feb 2025 – May 2025**
Hong Kong University of Science and Technology

- Designed, and implemented a multi-turn recursive framework for detecting and mitigating hallucinations in large language models (LLM) responses to factual questions was implemented using the Open-llama-7B model.
- Developed a sensitivity-tuned hint prefix method to calibrate LLM without requiring an external oracle. Tested and evaluated the effectiveness on TriviaQA dataset, achieving an improved reduction in hallucination rate

Graph NNs: Exploring Cross-domain task generalization of graphs with language embeddings **Mar 2025 – May 2025**
Hong Kong University of Science and Technology

- Explored scaling and generalizing cross-domain tasks amongst graph datasets using language embeddings and Mixture-of-Feature models, which are inspired from the MoE architecture.
- Evaluated model performance of word2vec, E5-Small-v2, and the GraphAlign on classification tasks from OGB graph datasets.

DNA K-mer Parallel Programming: CUDA, Pthreads, MPI **Feb 2025 – May 2025**
Hong Kong University of Science and Technology

- Implemented parallelized algorithms for k-mer mining of DNA sequences with CUDA, Pthreads, and MPI across both CPU and GPU clusters (HKUST SuperPod). Evaluated wall-clock performance optimizations across different algorithms.

Training Multi-Modal Generative Models with LLMs **Oct 2024 – Dec 2024**
Hong Kong University of Science and Technology

- Researched and employed text-to-music generation models for multimodal music generation with only an album image.
- Trained LoRa adaptors for Qwen language models to dynamically generate generation prompts for assistive music production.

Random Forest Parallelization with PySpark **Feb 2024 – May 2024**
Hong Kong University of Science and Technology

- Optimized the Random Forest algorithm using parallel computing in Spark on Microsoft Azure to measure how to improve training time of random forest through the divide and conquer approach. Analyzed computational overhead, proposing enhancements for dynamic partitioning and resource allocation.

Stroke Prediction with Machine Learning **Oct 2023 – Dec 2023**
Hong Kong University of Science and Technology

- Benchmarked performance by cross validating models on F1 score tuned w/ grid & bayesian search, comparing effect of synthetic data vs non-synthetic data on: Naive Bayes, Support Vector Machines and Deep Neural Networks (DNNs).

Publications & Achievements

- Co-Author in "Financial Considerations in Primary Healthcare after COVID-19", *Springer Nature*
- TPG Fellowship for MSc. Big Data Technology
- Co-Founder of Northwestern Blockchain Group
- Northwestern Dean's List (2017)
- Certificate in Deep Learning Specialization - DeepLearning.AI
- Certificate in Professional Scrum Master 1 (PSM1)
- Certificate in UIPath Professional Developer

Community & Philanthropy

- Charity Photobook Exhibition (Project Director): Led the product development, planning, marketing, and publication of "FORCES AT WORK", a charity photobook exhibition. Raised HKD $100K for Médecins Sans Frontières (Doctors Without Borders) in book sales.
- ACS Sprint Hackathon Pitch Winner (2018): Represented PwC in a non-profit hackathon event hosted by the Asian Charity Services (ACS). Pitched and won HKD $25,000 for non-profit organization AuDeAHK by pitching the use of haptic feedback gloves to enhance the sensory experience of visually impaired individuals.

Skills

- Programming: C++, Python | Packages: sklearn, torch/tf, transformers, huggingface, pd, np, JAX, matplotlib, sns, opencv
- Data Analytics: R, MATLAB, STATA, Alteryx, Power Query, PowerBI, Tableau, UIPath, AMPL
- Database: SQL, Spark, Hadoop | Performance Optimization: Cuda, MPI, Pthreads
- Biomedical computation: AlphaFold, ColabFold, ColabDesign, PyMol, PyRosetta
- Languages: Native in English, Mandarin, Cantonese | Beginner in German
- Hobbies: Golf, Disc Jockey, Pleasure Vessel Operator Certificate Grade 2 (PVOC2)