# Xuanzhuo Liu

xuanzhuoliu@link.cuhk.edu.cn

## EDUCATION

**The Chinese University of Hong Kong, Shenzhen** **Sept. 2022 – May 2026**

B.S. in Data Science and Big Data Technology, Minor in Computer Science

- *GPA***:** cGPA: **3.92/4.0**, mGPA: **3.97/4.0** (rank **2/146**)
- *Scholarship***:** Full Tuition Scholarship; Class B Academic Scholarship (**top 3%**)
- *Honors***:** 2022-2023, 2023-2024 Dean's List; 2024-2025 Undergraduate Research Award

**University of Oxford** | Visiting Student (**On-Site**), Mathematics and Computer Science **Oct. 2024 – Jul. 2025**

- Modules (Master Level, all A+ grades): Theories of Deep Learning, Bayesian Statistical Probabilistic Programming, Geometric Deep Learning, Uncertainty in Deep Learning, Graph Representation Learning, Optimal Control.

## RESEARCH EXPERIENCE

**Shenzhen Research Institute of Big Data** | *Research Assistant, Advised by **Prof. Haizhou Li*** **Aug. 2023 – Jul. 2024**

- **Research Topic**: ***Human Attention Modeling for Multimodal AI Perception and Interaction***
  **Keywords**: Multimodal attention modeling, human-centered AI, Python, statistical analysis, eye-tracking, experimental design
- Conducted research on how humans integrate multimodal (visual-auditory) information to guide attention.
- Designed 5 sets of synchronized perceptual experiments using **E-Prime**, mimicking real-world sensory conditions; refined through 4 design-review cycles to minimize perceptual ambiguity and enhance temporal precision.
- Processed eye-tracking data via **Tobii Studio** and conducted statistical modeling using **Python (NumPy, SciPy, Pandas)** to identify patterns of gaze allocation and attention dynamics across conditions.
- Co-authored and presented findings at the **2024 International Conference on Social Robotics (ICSR)**; awarded **Best Student Paper**, recognizing the impact on **human-aware robotic interaction design**.

**Shenzhen Research Institute of Big Data** | *Research Assistant, Advised by **Prof. Shuang Li*** **Nov. 2024 – Feb. 2025**

- **Research Topic**: ***Neuro-symbolic Learning for Human-Aligned Medical Perception and Reasoning***
  **Keywords**: sparse attention, neuro-symbolic reasoning, reinforcement learning, decision modeling, cognitive modeling
- Built a sparse-attention Transformer that simulates doctor decision paths via rule activation sequences, improving diagnosis prediction accuracy by **+14.6%** on real-world data (MIMIC-III).
- Integrated rule-based learning with hierarchical attention to capture individual doctor preferences and cognitive variation across **100+ synthetic and real cases**.
- Designed a hybrid architecture combining **symbolic reasoning and RL**, enabling adaptive inference with **~27% fewer steps** under noisy or partial patient information.
- Prompted LLMs to generate **100+ domain rules**, constructing a rule universe and simulating decision diversity; improved generalization to unseen cases by **+11.2%** (zero-shot).

**Oxford Robotics Institute,** *Advised by **Dr. Zheng Xiong*** **Oct. 2024 – Present**

- **Research Topic:** ***Generalist Robot Policy Learning via Deep Reinforcement and Imitation Learning***
- Trained imitation learning agents in **PyTorch**, achieving **+19.3% average success rate** across 10+ MetaWorld tasks.
- Visualized **RT-1 and Octo** attention maps to analyze task-conditioned token relevance and reduce spurious activation.
- Improved few-shot task transfer by **+12.5%** via aligning sparse attention patterns with self-supervised pretraining objectives.
- Built image-conditioned deep RL agents in **JAX** using CNN encoders, maintaining **85%+ execution accuracy** under noisy visual input.
- Fine-tuned the **459M-parameter OTTER model** on **SimplerEnv** across 5 task types, enabling **1.7× better parameter reuse** in multimodal robot control.

## WORK EXPERIENCE

**Software Engineer |** Shenzhen Institute of Big Data Economics **Jun. 2024 – Oct. 2024**

- **Keywords**: Java, Cloud Platform, Docker, REST API, Git, Jenkins, CI/CD
- Contributed to the development of a secure research data transmission platform, focusing on backend module design and deployment for high-availability data sharing and access control.
- Developed the **ParIS (Participant Information Service)** component in **Java**, implementing participant state management, access policies, and RESTful APIs with **JSON Schema** validation.
- Configured **CI/CD pipelines** using **Jenkins** and **Git**, enabling automated build, test, and deployment processes.
- participated in cross-team requirement reviews and API interface alignment to support distributed system development.

**Undergraduate Teaching Assistant |** School of Data Science **Feb. 2024 – Present**

CSC3100 Data Structure, STA4001 Stochastic Process, MAT2041 Linear Algebra

- Delivered **8 tutorial sessions** for over **100 students** regarding data structures and algorithms.
- Implemented data visualization programs supplementing course materials in random walk, Brownian motion and CTMC.
- Exhibited exceptional communication and interpersonal skills in resolving students' coursework-related problems.

**Technical Volunteer | Tencent Spark Plan Challenge Week (Robotics & Embodied Intelligence Track) May. 2025 – July. 2025**

- **Keywords**: Robotics, Deep Learning, Technical Support, Embodied AI, LLM Prompting
- Selected as one of **3 technical volunteers** nationwide to support Tencent's elite robotics challenge program.
- Collaborated with researchers from **Tencent Robotics X**, assisting participants with real-world tasks involving **robot control**, **reinforcement learning**, and **multimodal LLM integration**.
- Contributed to the preparation of challenge content, simulation environments, and project walkthroughs, ensuring smooth delivery of 10+ hands-on AI tasks.
- Led Q&A sessions and offered **weekly online mentoring** on deep learning fundamentals, prompting strategies, and embodied agent design.

## SELECTED PROJECTS

**Multi-class Classification in Predicting Context Information | PyTorch, COPT** **Mar. 2024 – Apr. 2024**

*Breman Big Data Challenge 2024, top 5*

- Collaborated in a team to predict and impute **over 23,000 missing data** points based on bio-signals in a real-world setting.
- Developed CNN and decision tree models to analyze behavioral patterns and conduct multi-class classification.
- Predicted emotions and context information at specific timestamps based on large-scale movement sensor's data and biological signals, achieving a test accuracy of over 60% (**ranked top 5**).

**A Unified Framework for Interpretable Transformers Using PDEs and Information Theory Feb. 2025 - Apr. 2025**

- Developed a novel theoretical framework integrating partial differential equations, neural information flow theory, and Information bottleneck theory to analyze and interpret transformer architectures.
- Designed and implemented algorithms bridging continuous PDE formulations with discrete transformer components.
- Conducted extensive experiments across image (MNIST) and text (20 Newsgroups) domains to validate the framework's ability to simulate attention behavior, information bottleneck effects, gradient flow, and robustness.
- Demonstrated that the PDE-based model closely approximates Transformer layerwise behavior, offering new insights into interpretability, optimization, and model stability.