LUZHE HUANG

PHONE: +1-424-402-2604 **EMAIL:** LZHUANG0324@UCLA.EDU
LINKEDIN: /LUZHE-HUANG **GITHUB:** /PORPHURA **PAGE:** /PORPHURA.GITHUB.IO

SUMMARY

- Senior Machine Learning Scientist with a Ph.D. from UCLA and over **70** research articles (1,100+ citations) in top-tier journals like *Nature Machine Intelligence*, *Nature Methods*, and *Science Advances*.
- Expert in architecting **Multi-modal LLMs (MLLM)** and **Generative Recommendation Systems** at industry-scale, specializing in self-supervised alignment and long-sequence modeling.
- Proven track record of bridging research and production, from pioneering physics-informed neural networks to deploying high-throughput generative recall frameworks.

CAREER EXPERIENCE

TikTok — San Jose, USA
Machine Learning Scientist — Oct 2024 – Present

- **MLLM for Large-Scale Recommendation**
  - **Led development** of a joint **Next-Token Prediction (NTP) & Noise Contrastive Estimation (NCE)** training paradigm, producing unified embedding models with superior understanding and prompt-following abilities.
  - Architected a specialized fine-tuning process using **rs-LoRA**, bridging the performance gap between PEFT and full-parameter training for region-specific deployments.
- **Generative Recommendation Engine**
  - Engineered a SOTA recall framework using **multi-query multi-token prediction** to model complex user interest profiles, resulting in a **0.14%** online user stay duration improvement and **0.16%** online interest diversity gain.
  - Innovated on position embeddings with **Timedelta RoPE**, capturing periodic user behaviors and boosting long-history recommendation performance.
  - Resolved transformer scaling bottlenecks by integrating **sequence compression** and **gated attention**, increasing model throughput by **45%** and model recall by **2%**.

Autowise.ai — Shanghai, China
Software R&D Engineer Intern — Feb 2019 – Jun 2019

- Developed 3D point cloud-based object detection and LiDAR calibration models for autonomous driving platforms.

SELECTED RESEARCH PROJECTS

Physics-Informed Self-Supervised Learning (GedankenNet) — Feb 2022 – Aug 2023

- Pioneered a self-supervised neural network using **physics-informed learning** and synthetic data, achieving superior out-of-distribution (OOD) generalization compared to supervised models.
- Synergistically designed a generalizable optical system and backend image processing pipeline.

Cycle-Consistency Uncertainty Quantification (UQ) — Dec 2022 – Oct 2023

- Built a novel UQ framework for inverse problems using forward-backward cycles and developed an **autonomous hallucination monitor** to detect OOD data.

EDUCATION

University of California, Los Angeles | Ph.D., Electrical and Computer Engineering — 2019 – 2024

- GPA: 3.96/4.0; UCLA **Dissertation Year Fellowship** and **Amazon Fellowship** recipient.

Zhejiang University | BEng, Optical Science and Engineering, Statistics — 2015 – 2019

- GPA: 3.96/4.0; Received **National Scholarship** (2x) and **Special Scholarship**.

TECHNICAL SKILLS

Programming: Python (PyTorch, JAX, TensorFlow), C++, SQL, R, MATLAB.
ML Infrastructure: FlashAttention, DeepSpeed, vLLM, Triton, Docker, Spark.

PROFESSIONAL SERVICE

- **Reviewer:** Reviewed 20+ articles for *Nature*, *Nature Machine Intelligence*, and *Science Advances*.
- **Mentorship:** Supervised 10+ undergraduate and master students in research projects at UCLA.