Mahdi (Sum) Qezlou, Ph.D.

Email: sumqezlou@gmail.com | (951) 565-6687 | linkedin.com/in/qezlou | qezlou.github.io

Innovative **AI Scientist** with a **Ph.D. in Physics** focusing on modeling complex systems with a strong foundation in **statistics**, **math**, and **programming**. Over **7 years of experience** leading end-to-end **multi-modal** (time-series, images and 3D assets) machine learning projects from **pre-training**, **fine tuning** and **Reinforcement Learning (RL).** Proficient in Python-based ML libraries and **scalable training** and **inference** using 100+ TB of real-world data on cloud (Kubernetes) and HPC (Slurm).

Highlighted projects:

More: https://qezlou.github.io/ml.html

Foundation Model Training: Signal Detection in Trillion-Element, Noisy Astrophysical Datasets

Repos: https://github.com/qezlou/BlindSight

- **Objective**: Accurately identify a **subtle signa**l from **noise and systematics** in trillion-resolution-element **image** and **time series data**, when traditional clustering methods proved ineffective.
- **Contribution:** Developed a domain-specific **foundation model** (**Variational AutoEncoder (VAE)-Transformer-**based architecture) for observed noisy **time-series** and **images** to model the underlying distribution of real signal as opposed to false positives. Leveraged advanced **self-supervised learning** techniques—including CLIP, DINOv2, MAE, BEiT, iBOT, and MoCo v3—employing contrastive learning, masked modeling, and self-distillation for robust vision and **multimodal representation learning.** This **probabilistic model** for false positive removal enabled **percent-level accuracy** in downstream analysis integral to the company's mission.

GenAI for inference: Fast surrogate models with Gen AI

Repos: https://github.com/qezlou/Goku_ELG and https://github.com/qezlou/lila

- **Objective**: Enabling **accurate inference** through **fast surrogate models** replacing computationally prohibitive Astrophysical simulations
- **Contribution:** Led the development of probabilistic **machine learning surrogate models,** including **stochastic models** (e.g., Gaussian Processes) and Deep Learning **generative models** (Normalizing Flow), to enable robust **causal inference** in **high-dimensional** parameter space and improved **forecasting accuracy by 50%.**

Scalable GPU coding: GPU accelerated cosmological simulations

Repos: https://github.com/MP-Gadget/shengi

- **Objective**: Adapt a massively parallel MPI–OpenMP CPU-based gravity and fluid solver for **NVIDIA's Grace–Hopper** heterogeneous CPU–GPU architecture deployed on the VISTA supercomputing platform. tacc.utexas.edu/systems/vista/
- **Contribution:** Led key aspects of low-level **code migration** and **performance optimization**, for multi-node GPU acceleration through novel **CUDA**-based neighbor-tree construction and FFT techniques to achieve **5X computational speedups**.

Technical Skills

- **Theoretical Knowledge:** Solid understanding of Linear Algebra, probabilistic modeling and physics simulations.
- **ML Tools:** Experienced in large-scale training (100+ TB) using **PyTorch** and **TensorFlow**, with expertise in **domain adaptation**, e.g. NVIDIA's **NeMo** framework, including **LSTM** and **Transformer encoder–decoder** architectures.
- **Programming skills:** Distributed CPU-GPU computing on Cloud (Kubernetes) and HPC (Slurm, CUDA-aware MPI), NVIDIA's SDK (NeMo, cuda, etc.), python, C/++, Git, Linux.
- **Soft Skills**: Extensive **interdisciplinary applied machine learning** research experience, with a proven ability to collaborate **cross-functionally** with **management** and business **stakeholders** to align technical solutions with strategic objectives.

Employment

AI Research Scientist | University of Texas, Austin | 2024 – Present
ML Research Scientist | Carnegie Institute for Science | 2020 – 2024
ML Research Scientist | University of California, Riverside | 2018 – 2024

Education:

Ph.D. in Physics and Machine Learning | University of California, Riverside | 2018 - 2024