Mahdi Qezlou

mahdiqezlou@gmail.com
github.com/qezlou
linkedin.com/in/qezlou/
phone : (951)565-6687

Employment

2024-present **Data Scientist**, *University of Texas, Austin*.
Leading machine learning and AI projects in astronomy for accurate causal inference and forecasting.

2018-2024 **Research fellow**, *University of California, Riverside*.
Led three major machine learning projects to infer the behavior of complex systems.

Education

2018–2024 **Ph.D. Machine Learning for Astrophyscis**, *University of California, Riverside*.
Developing Machine learning and *Statistical tools to infer properties of astrophysical systems*

2018–2019 **M.Sc. Physics**, *University of California, Riverside*.

2013-2018 : **B.Sc. in Physics** , *Sharif University of Technology*.

Skills

Programming: `Git, Bash, Python, C, SQL, TensorFlow, PyTorch, XGBoost, Cloud Architecture (AWS/Google Cloud), Apache Spark (Databricks), Tableau, Matplotlib`

ML & stat Bayesian/Frequentist statistics, stochastic processes (Markov/Gaussian), Deep Learning (CNN, NLP), Generative AI. Strong background in mathematics and computation, with active engagement in machine learning literature.

Languages: English (Fluent), Persian (Fluent), Turkish (Intermediate)

Highlighted Project leadership:

2018-present **Leading Experimental Design, Causal Inference, and Forecast Success Metrics**.
Objectives: Designing a high-dimensional (20+ D) feature space to fully characterize and forecast the observables of complex systems in astrophysics.

Challenges: Forward modeling of the observed data is computationally expensive, and the feature space is typically sparsely sampled. For causal inference, an affordable but accurate surrogate model is needed for frequent evaluations at arbitrary points of parameter space using Markov Chain Monte Carlo (MCMC).

My contribution: I identified the issue and performed experimental design on the feature space. I trained Gaussian Process (GP) and generative AI models on a small set of high-fidelity forward models. These machine learning models replicated the expensive computational models with 1% accuracy at 1/100th of the cost. I applied these fast, nonlinear surrogate models during MCMC sampling for accurate causal inference across the 20+ dimensional feature space.

Techniques: Experimental design, A/B testing, frequentist and Bayesian statistical modeling, deep learning, generative AI (Diffusion Models, Normalizing Flow)

Programming Languages and Platforms: `Python, SQL, TensorFlow, Cloud Architecture (AWS), Apache Spark (Databricks)`

Collaboration with Technical and Non-Technical Teams: Regular discussions with Software/ML engineers at the Texas Advanced Computing Center (TACC) | Presentations to non-technical stakeholders, including the Department of Energy, NASA, and the Air Force.
Github : `Lya-Emulator`

Relevant courses/Certifications

UCR Deep Learning, Probabilistic models in AI

Coursera Deep Learning specialization