Foad Namjoo

GitHub | Hugging Face | LinkedIn | Google Scholar | Website | foad.namjoo@utah.edu | (765) 767 3525 | Salt Lake City

Summary

Ph.D. researcher in LLM interpretability with a foundation in high-dimensional geometry and statistical algorithms. I build datasets and activation steering pipelines to control LLM behaviors such as truthfulness, tone, and conciseness.

Research Experience & Publications

Graduate Research Assistant | University of Utah

- **LLM Activation Steering and Robustness under Dataset Corruption | 2025–Present**
  Controlled LLM verbosity and tone via residual-stream activation steering in Llama 3.1 and Qwen-2.5, with layer and steering-strength sweeps for mechanistic interpretability. Built a guardrail evaluation pipeline using LLM judges and JSON schema validation, reducing hallucinations. Studied robustness to noisy and adversarial corruption in steering datasets, identifying failure modes and showing that robust high-dimensional mean estimation can mitigate them.
  ***Paper:** Understanding and Mitigating Dataset Corruption in LLM Steering — arXiv preprint, 2026* [link].

- **Efficient Multidimensional Two-Sample Testing (dKS) | 2025**
  Extended the Kolmogorov–Smirnov distance to multidimensional probability distributions and robust two-sample testing. Implemented ε-approximate algorithm with a near-linear run time $O(n \log n)$ in 2D; unit-invariant integral probability metric (IPM) with ε-accurate grid method; stronger tests than unstable mdKS.
  ***Paper:** Efficient and Stable Multidimensional Kolmogorov–Smirnov Distance — submitted to SIMODS* [link].

- **MotionPI — Privacy-First Wearable Sensing for Behavior Analytics (Full-Stack) | 2024–Present**
  Built an on-device, wearable analytics stack (MotionSense HRV wristbands to app to API to database); streamed PPG/ENMO/Survey with offline-first, schema-validated sync; sustained $\sim$7.7M records/day with zero malformed writes. Designed an interactive visualization tool for analyzing and adjusting activity-trigger thresholds. Reported and verified a Flutter SDK bug (Issue #166937), later confirmed and triaged by the Flutter team.
  ***Paper:** Designing a Secure Distributed Participant Data Collection System — EAI SmartSP 2025* [link].

- **Anomaly Detection — Region-Aggregated Spatial Scan Statistics | 2023**
  Replaced centroid scans with multi-point region sampling to increase detection power under axis-aligned rectangles; reproducible experiments with a C++ backend.
  ***Paper:** Sampling for Region-Aggregated Spatial Scan Statistics — under review (double-blind).*

Graduate Research Assistant | University of Tehran

- **Computer Vision — High-Dimensional Spectral–Spatial Change Detection | 2022**
  Created a dual-stream 3D/2D CNN with SE attention; *accuracy* >96%, $\kappa > 0.9$, and lower false positive vs. baselines.
  ***Paper:** A Hyperspectral Change Detection Framework Based on Double Stream CNNs Attention Module* [link].

Education

Ph.D., Computer Science | University of Utah | Jan 2023 – Expected Jan 2027

Advisor: Prof. Jeff Phillips | GPA: 3.9/4 | Focus: LLM interpretability; high-dimensional geometric data analysis
Selected Coursework: Machine Learning; Probabilistic ML; Deep Learning; Data Mining; Visualization.

M.Sc., Algorithms and Computation (Computer Science) | University of Tehran | 2019 – 2022

Thesis: Graph-theoretic modeling of bushfire propagation | GPA: 17.8/20
Selected Coursework: Advanced Algorithms; Approximation Algorithms; Randomized Algorithms; Quantum Algorithms & Computation; Graph Algorithms; Network Science; Internet Algorithms; Distributed Systems.

Skills

Languages: Python, C++, JavaScript, Dart
ML & LLMs: PyTorch (CUDA), Hugging Face (Transformers, Datasets, Hub), TensorFlow, scikit-learn, Activation Steering, TransformerLens, Data Curation & Evaluation
Systems & Tools: Docker, Podman, Node.js, MongoDB, REST APIs, CI/CD, Git, Cloud GPUs, Google Earth Engine

Honors & Awards

Iran M.Sc. University Entrance Exam — Rank 29/20,000 (top 0.15%, 2019) | Top-Talent Scholarship — University of Tehran (2020–2021) | Graduate Fellowship — University of Utah (2023)