Arun Sharma

P: +1 612-946-5123 | arunshar@umn.edu | [Website](#) | [Google Scholar](#) | [Github](#) | [LinkedIn](#) | [ORCiD](#)

EDUCATION

UNIVERSITY OF MINNESOTA, TWIN CITIES — Minneapolis, MN

Doctor of Philosophy (Ph.D.) in Computer Science — July 2018 - August 2025
Advisor: Prof. Shashi Shekhar | Spatial Computing Research Group
Thesis Committee: Prof. Vipin Kumar, Prof. Ravi Janardan, and Prof. Ying Song

STATE UNIVERSITY OF NEW YORK AT BUFFALO — Buffalo, NY

Master of Science (M.S.) in Computer Science — August 2016 - June 2018

WORK EXPERIENCE

UNIVERSITY OF MINNESOTA, TWIN CITIES — Minneapolis, MN

Graduate Research Assistant - II — **May 2024 – Present**

- Proposed a physics-informed diffusion probabilistic model incorporating kinematic constraints to detect anomalous trajectories indicative of GPS spoofing (e.g., fake trajectories) in domain-specific areas, addressing challenges from AI-generated deep fakes, data sparsity, and lack of fine-scale spatiotemporal dependencies.
- Proposed a novel Kriging-informed Conditional Diffusion Probabilistic Model for downscaling coarser-resolution climate projections (from global models or satellites) to finer-resolution regional data, capturing spatial variability and fine-scale features while addressing heterogeneity and generalization challenges.
- Proposed Physics-guided Reasoning Models (PgRM) to incorporate broad-domain physical knowledge (e.g., conservation laws) while distinguishing broad universal principles from narrow task-specific ones.
- Proposed a novel Surrogate-based Spatial Neural Networks to capture spatial autocorrelation and long-range dependencies.

Graduate Research Assistant — **August 2018 – May 2024**

- Designed and implemented the Geo-Lucid Conditional Diffusion Model, integrating digital road map information to generate synthetic vehicle trajectories with superior geo-distribution similarity and dynamics fidelity.
- Proposed a Geometry and Drive-cycle Contrastive Learning framework outperforming SOTA for Physics-Guided Trajectory Representation Learning, explicitly modeling driving cycle dynamics (e.g., velocity profiles) in vehicle trajectories to enhance downstream tasks like travel time and energy estimation.
- Investigated and proposed a spatially lucid deep neural network for multi-category point set classification in non-Euclidean space via a spatial ensemble framework with place-calibration parameters and flexible training strategies.
- Leveraged domain-adapted AI classifier within a multi-task architecture featuring spatially-oriented self-supervised learning tasks divided into sub-modules for spatial mix-up masking and spatial contrastive predictive coding.
- Led a project analyzing data distortion using physics-aware methods to identify denial-based anomalous patterns (i.e., signal gap or clandestine rendezvous) with a 30% reduction in false positives and cutting processing time by 40%.
- Investigated and proposed a spatially lucid deep neural network for multi-category point set classification in non-Euclidean space via a spatial ensemble framework with place-calibration parameters and flexible training strategies.

ESRI - ENVIRONMENTAL SYSTEMS RESEARCH INSTITUTE — Redlands, CA

Research Intern - Ph.D. — **May 2023 – August 2023**

- Designed and implemented a scalable Graph-based Traffic Representation and Association (GTRA) framework for maritime route optimization, leveraging PySpark and GeoAnalytics APIs and reducing query latency by 40%.
- Developed an anomaly detection pipeline integrating Transformer-based models, Evidential Deep Learning (EDL), and AWS SageMaker, improving real-time anomaly classification accuracy from 54.76% to 73.02%.
- Developed an automated model training pipeline and optimized inference for GIS-based AI models on a large-scale maritime dataset by leveraging AWS services (Lambda, ECS, SageMaker Pipelines, Multi-Model Endpoints (MME), Step Functions, SQS, CloudWatch) and implementing model quantization, reducing model retraining time by 35% and API latency by 30%.

TECHNICAL SKILLS

Languages: Python, Java, R, SQL, Scala, C/C++, Julia, Rust, Go, Shell (Bash), MATLAB, JavaScript (Node.js), TypeScript.
Machine Learning: MLflow, XGBoost, MLlib, LLMs (fine-tuning & inference), RAG, LangChain, VectorDB (FAISS, Chroma, Pinecone), Hugging Face Transformers, LoRA, PEFT, RLHF, OpenAI API, Triton Inference Server, ONNX.
Big Data & Distributed Systems: PySpark, Hadoop, Flink, Presto, Trino, Hive, Kafka, Delta Lake, Docker, Kubernetes.
ML Frameworks & Deployments: PyTorch, TensorFlow, JAX, ONNX, SageMaker, Lambda, ECS, Bedrock.

HONORS AND ACHIEVEMENTS

DOCTORAL DISSERTATION FELLOWSHIP — **2022 - 2023**

University of Minnesota, Twin Cities

REVIEWER

NeurIPS, ICML, ICLR, CVPR, ECCV, AAAI, IJCAI, SIGKDD, CIKM, ICDM, SDM, MLSys, SIGSPATIAL, TKDE, JMLR