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

SELECTED PUBLICATIONS

[1] Towards Physics-informed Diffusion for Anomaly Detection in Trajectories: A Summary of Results
Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Geospatial Anomaly Detection (GeoAnomalies '25)
Arun Sharma, Mingzhou Yang, Majid Farhadloo, Subhankar Ghosh, Bharat Jayaprakash, and Shashi Shekhar

[2] Geo-lucid Conditional Diffusion Models for High Physical Fidelity Trajectory Generation
Proceedings of the 33rd ACM International Conference on Advances in Geographic Information Systems (SIGSPATIAL '25)
Mingzhou Yang, **Arun Sharma**, Majid Farhadloo, Bharat Jayaprakash, and Shashi Shekhar

[3] Towards Surrogate Models with Hybrid Spatial Neural Networks: A Summary of Results
Proceedings of the 8th ACM SIGSPATIAL International Workshop on Geospatial Simulation (GeoSIM '25), pp. 57–69, 2025
Shengya Zhang, **Arun Sharma**, Majid Farhadloo, Mingzhou Yang, Ruolei Zeng, Subhankar Ghosh, Yao Zhang, Mu Hong, Licheng Liu, David Mulla, and Shashi Shekhar

[4] Spatially-Delineated Domain-Adapted AI Classification: An Application for Oncology Data
SIAM International Conference on Data Mining, 2025
M Farhadloo, **Arun Sharma**, A. Leontovich, S N. Markovic, and Shashi Shekhar

[5] Towards Spatially-Lucid AI Classification in Non-Euclidean Space: An Application for MxIF Oncology Data
SIAM International Conference on Data Mining, 2024.
M Farhadloo, **Arun Sharma,** J. Gupta, A. Leontovich, S N. Markovic, and Shashi Shekhar

[6] Spatial Computing Opportunities in Biomedical Decision Support: The Atlas-EHR Vision
ACM Transactions on Spatial Algorithms and Systems 10, no. 3 (2024): 1-36.
M Farhadloo, **Arun Sharma**, A. Leontovich, S N. Markovic, and Shashi Shekhar

[7] Towards Kriging-informed Conditional Diffusion for Regional Sea-Level Data Downscaling: A Summary of Results
32nd International Conference on Advances in Geographic Information Systems, 2024
Authors: S Ghosh[*], **Arun Sharma**[*], J. Gupta, A. Subramanian and Shashi Shekhar
([*]Both authors contributed equally to this paper)

[8] Towards Fine-Tuning-Based Site Calibration for Knowledge-Guided Machine Learning: A Summary of Results
5th Annual AAAI Workshop on AI to Accelerate Science and Engineering (AI2ASE)
Ruolei Zeng, **Arun Sharma**, Shuai An, Mingzhou Yang, Shengya Zhang, Licheng Liu, David Mulla, and Shashi Shekhar

[9] Physics-based Abnormal Trajectory-Gap Detection
ACM Transactions in Intelligent Systems and Technology, 2024.
Arun Sharma, Subhankar Ghosh, and Shashi Shekhar

[10] Analyzing Trajectory Gaps for Possible Rendezvous Regions
ACM Transactions in Intelligent Systems and Technology, 2022
Arun Sharma and Shashi Shekhar

[11] Towards a Tighter Bound on Possible-Rendezvous Areas: Preliminary Results
30th International Conference on Advances in Geographic Information Systems, 2022
Arun Sharma, Jayant Gupta, Subhankar Ghosh, and Shashi Shekhar

[12] Mining taxonomy-aware colocations: A Summary of Results
30th International Conference on Advances in Geographic Information Systems, 2022
Jayant Gupta, **Arun Sharma**, and Shashi Shekhar

[13] Spatiotemporal Data Mining: A Survey
Handbook of Spatial Analysis for the Social Sciences, Edward Elgar, 2022
Arun Sharma, Zhe Jiang, and Shashi Shekhar

[14] Understanding COVID-19 effects on mobility: A community-engaged approach
25th AGILE Conference on Geographic Information Science, 2022
Arun Sharma, Majid Farhadloo, Yan Li, Jayant Gupta, Aditya Kulkarni, and Shashi Shekhar

[15] WebGlobe: A cloud-based framework for interacting with climate data
International Workshop on Analytics for Big Geospatial Data (SIGSPATIAL) 2018
Arun Sharma, SM Arshad Zaidi, Varun Chandola, Melissa R Dumas, and Budhendra L Bhaduri

TEACHING EXPERIENCE

SPATIAL DATA SCIENCE RESEARCH **Spring 2024**

- Guest lectures on scientific methodology, fostering research skills, and effectively communicating research ideas.
- Supervised diverse projects and mentored multiple graduate and undergraduate students from interdisciplinary fields.

SPATIAL DATA SCIENCE **Fall 2019**

- Responsible for handling class lectures, queries, homework assignments, labs, exams, lecture slides, etc.
- Guest lecture topics: Spatial Indexing, Networks, and Data Mining.

ADVANCED DATABASE SYSTEMS **Spring 2019**

- Responsible for handling class lectures, queries, homework assignments, labs, exams, lecture slides, etc.
- Guest lecture topics: Concurrency Control, Database Security, and Data Mining.

DATA STRUCTURES AND ALGORITHMS **Fall 2018**

- Instructed weekly recitation sessions with over 40+ students and graded 400+ students.

SERVICES AND LEADERSHIP

SUST 4096 SUSTAINABILITY INTERNSHIP, UNIVERSITY OF MINNESOTA, MN (Mentor) **Spring 2024**

- Supervised an undergraduate student on a course project based on KGML for Sustainable and Precision Agriculture.
- Presented a lightning talk and a poster at the Annual AI-CLIMATE Review Meeting in UMN Saint Paul Campus, MN[1].

HONORS MENTORS CONNECTION, WAYZATA HIGH SCHOOL, MN (Mentor) **Fall 2023 - Spring 2024**

- Advised a high school student to analyze a real-world case study of signal spoofing behavior in the maritime domain.
- The student helped analyze circular patterns, enabling us to create a taxonomy of spoofing behavior in open waters.

MINNESOTA DEPARTMENT OF MANAGEMENT AND BUDGET, MN (Service) **May 2020 - June 2021**

- Reported county-level mobility traffic to epidemiologists, analysts, and policymakers for informed decision-making.[2]
- Advised multiple high school and undergraduate students who were considering a research career.

INVITED PRESENTATIONS AND POSTERS

Poster: Towards Physics-guided Generative Foundation Models **2025**
The 1st ACM SIGSPATIAL International Workshop on Generative and Agentic AI for Multi-Modality Space-Time.
Authors: **Arun Sharma**, Majid Farhadloo, Mingzhou Yang, Bharat Jayaprakash, William Northrop, and Shashi Shekhar

Poster: Spatial Distribution-Shift Aware Knowledge-Guided Machine Learning **2025**
AAAI Bridge on Knowledge-Guided ML: Bridging Scientific Knowledge and AI, 2025.
Authors: **Arun Sharma**, Majid Farhadloo, Mingzhou Yang, Subhankar Ghosh, and Shashi Shekhar

Blue Sky Ideas: Towards Physics-Guided Foundation Models **2025**
AAAI Bridge on Knowledge-guided ML Bridging Scientific Knowledge and AI, 2025
Authors: Majid Farhadloo*, **Arun Sharma***, Mingzhou Yang, Bharat Jayaprakash, William Northrop, and Shashi Shekhar
(*Both authors contributed equally to this paper)

From KGML-AG to KGML Precision-Ag: A Spatial Variability Approach **2024**
AI-CLIMATE Annual Review Meeting, 2025
Presenters: Ruolei Zheng* and **Arun Sharma***
(*Both authors contributed equally)

Addressing Data Distortion: A Physics-based Approach **2023**
MIDAS Future Leaders Summit: University of Michigan, Ann Arbor
Presenter: Arun Sharma

Rendezvous Pattern Detection From AIS Ship Trajectories **2020**
Center of Geospatial Information Science: University of Maryland, College Park
Presenter: Arun Sharma

---

[1] https://cse.umn.edu/aiclimate/news/2024-ai-climate-annual-review-meeting

[2] Sharma, Arun, Majid Farhadloo, Yan Li, Jayant Gupta, Aditya Kulkarni, and Shashi Shekhar. "Understanding covid-19 effects on mobility: A community-engaged approach." *AGILE: GIScience Series* 3 (2022): 14.