Arun Sharma

Minneapolis MN | P: +1 612-946-5123 | sharm485@umn.edu | Website | Google Scholar | Github | LinkedIn | ORCiD

EDUCATION

UNIVERSITY OF MINNESOTA, TWIN CITIES — Minneapolis, MN
Doctor of Philosophy (Ph.D.) in Computer Science — Expected June 2025
Advisor: Prof. Shashi Shekhar | Spatial Computing Research Group, UMN
Thesis Committee: Prof. Vipin Kumar, Prof. Ravi Janardan, and Prof. Ying Song

STATE UNIVERSITY OF NEW YORK AT BUFFALO — Buffalo, NY
Master in Science (M.S.) in Computer Science — August 2016 - June 2018
Advisor: Prof. Varun Chandola

WORK EXPERIENCE

ESRI - ENVIRONMENTAL SYSTEMS RESEARCH INSTITUTE — Redlands, CA
Research Intern - Ph.D. — June 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, 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%.

UNIVERSITY OF MINNESOTA, TWIN CITIES — Minneapolis, MN
Graduate Research Assistant — January 2020 – Present

- Led a project analyzing data distortion using physics-aware methods, improving anomaly detection accuracy by ~25% and identifying complex mobility patterns (e.g., clandestine rendezvous) with a 30% reduction in false positives.
- Designed and implemented scalable detection algorithms, cutting processing time by 40% while boosting precision by 20%.
- Proposed state-of-the-art generative models for climate downscaling, achieving a 35% reduction in error for abnormal sea-level rise predictions, and enhanced physics-informed trajectory generation by augmenting realistic simulations across various scenarios, leading to a 25% improvement in trajectory fidelity.

STATE UNIVERSITY OF NEW YORK AT BUFFALO — Buffalo, NY
Graduate Project Assistant — June 2017 – August 2017

- Developed an end-to-end framework on Apache Spark and Hadoop to analyze climate change impacts using NetCDF data, integrating a Gaussian process-based algorithm for detecting extreme temperature events (e.g., excessive heat, winter storms).
- Proposed and scaled an efficient GP regression on Apache Spark, reducing complexity from $O(N^3)$ to $O(N^2)$ via matrix factorization, with geo visualizations, intuitive UX/UI, and reproducible Jupyter Notebooks.

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

[2] Towards Physics-informed Diffusion for Anomaly Detection in Trajectories: A Summary of Results
31st SIGKDD Conference on Knowledge Discovery and Data Mining - Research Track 2025 (Under Review).
Arun Sharma, Subhankar Ghosh, Mingzhou Yang, Majid Farhadloo, Jayant Gupta, Bharat Jayaprakash, and Shashi Shekhar

[2] Geo-lucid Conditional Diffusion Models for High Physical Fidelity Trajectory Generation
31st SIGKDD Conference on Knowledge Discovery and Data Mining - Research Track 2025 (Under Review).
Mingzhou Yang, **Arun Sharma**, Majid Farhadloo, Bharat Jayaprakash, and Shashi Shekhar

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

[4] 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

[5] Spatial Computing Opportunities in Biomedical Decision Support: The Atlas-EHR Vision
ACM Transactions on Spatial Algorithms and Systems, 2024
M Farhadloo, **Arun Sharma**, A. Leontovich, S N. Markovic, and Shashi Shekhar

[6] 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)

[7] Towards Pareto-optimality with Multi-level Bi-objective Routing: A Summary of Results."
17th ACM SIGSPATIAL International Workshop on Computational Transportation Science GenAI and Smart Mobility Session
Mingzhou Yang, Ruolei Zeng, **Arun Sharma**, Shunichi Sawamura, William F. Northrop, and Shashi Shekhar

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

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

[10] 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

[11] 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

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

[13] 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

[14] 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, 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 grading of 400+ students.

SERVICES AND LEADERSHIP

MINNESOTA DEPARTMENT OF MANAGEMENT AND BUDGET, MN **2020**

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

MINNESOTA DEPARTMENT OF MANAGEMENT AND BUDGET, MN **2020**

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

MINNESOTA DEPARTMENT OF MANAGEMENT AND BUDGET, MN **2020**

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

INVITED PRESENTATIONS AND POSTERS

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, 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