Trilok Padhi

Contact Information

2521 Piedmont Road NE, 2321
Atlanta, GA 30324
(470) 847-1588
tpadhi1@student.gsu.edu

GitHub
LinkedIn
trilokpadhi_
Google Scholar

Research Interests

Multimodal Agents, Large Vision Language Models (LVLMs) and Large Language Models (LLMs), Knowledge-Enhanced Grounded Models to enhance Factuality & Truthfulness of LLMs & LVLMs, AI Ethics & Safety, Reinforcement Learning to enhance Reasoning & Planning, Web AI Agents.

Education

Georgia State University, Atlanta, GA
PhD in Computer Science — Aug 2022 – *Present*
Advisors & Mentors: Prof. Ugur Kursuncu, Prof. Dr. Yi Ding, Prof. Valerie Shalin, Dr. Yaman Kumar Singla

- Research Area: Design of Multimodal Knowledge-Enhanced Neuro-Symbolic Models and Agents
- Relevant Coursework: Large Language Models, Advanced Deep Learning, Advanced Machine Learning, Advanced Human-Computer Interaction

Research Experience

Siemens GenAI R&D, Seattle, WA

Applied Scientist Intern — May 23 - August 23
Mentors: Dr. Robert Zhu, Dr. Ivo Santos
Working on pretraining and post-training of industrial foundation models as part of the Neuro-Symbolic Computing and Intelligence (NuSCI) Research Group.

- Working on development and evaluation of large-scale mutlimodal industrial foundation models, with a focus on robust pretraining strategies and post-training techniques. Project details aligned with Siemens' initiative on Industrial Foundation Models: LinkedIn post.

SRI International (formerly Stanford Research Institute), Menlo Park, CA

Research Intern — May 23 - April 24
Mentors: Dr. Sushmit Jha, Dr. Anirban Roy, Dr. Ramneet Kaur, Dr. Manoj Acharya
Working on multimodal uncertainty quantification in the Neuro-Symbolic Computing and Intelligence (NuSCI) Research Group

- Designed and proposed a new SoTA methodology for Uncertainty Quantification for LVLMs using grounding which resulted in a paper titled *Seeing is Believing: Visual Grounding for Calibrating Uncertainty Quantification of Multi-Modal LLMs. . (Arxiv preprint)*

Bosch Research, Bengaluru, India *Research Intern* — Jan 20 - Aug 20
Mentors: Dr. Sahana Prabhu

Nvidia Research, Bengaluru, India *Research Intern* — Aug 18 – Aug 19
Mentors: Dr. Christoph Angerer, Mr. Jakob Progsch, Mr. Bharatkumar Sharma, Mr. Jigar Halani

Research Publications

***Seeing is Believing*: Visual Grounding for Calibrating Uncertainty Quantification of Multi-Modal LLMs** – Authors: **Trilok Padhi**, Ramneet Kaur, Adam D. Cobb, Manoj Acharya, Anirban Roy, Colin Samplawski, Brian Matejek, Alexander M. Berenbeim, Nathaniel D. Bastian, Susmit Jha *(arXiv preprint, Under Review)* (Research Area: Uncertainty Quantification, Calibration, Multimodal LLMs)

Echoes of Human Malice in Agents: Benchmarking LLMs for Multi-Turn Online Harassment Attacks – Authors: **Trilok Padhi***, Pinxian Lu*, Abdulkadir Erol, Tanmay Sutar, Gauri Sharma, Mina Sonmez, Munmun De Choudhury, Ugur Kursuncu
(arXiv preprint, Under Review) (Research Area: LLM Agents Safety, Social Computing, Mental Health)

Co-Evolving Agents: Learning from Failures as Hard Negatives
Authors: Yeonsung Jung, **Trilok Padhi**, Sina Shaham, Dipika Khullar, Joonhyun Jeong, Ninareh

Mehrabi†, Eunho Yang† *(arXiv preprint, Under Review)* (Research Area: Self-Learning Agents)

From Actions to Understanding: Step-Wise Conformal Interpretability of Temporal Concepts in LLM Agents – Authors: **Trilok Padhi**, Ramneet Kaur, Adam D. Cobb, Manoj Acharya, Anirban Roy, Colin Samplawski, Brian Matejek, Alexander M. Berenbeim, Nathaniel D. Bastian, Susmit Jha (arXiv preprint, Under Review) (Research Area: Interpretability, Reasoning & Planning LLM Agents)

Just KIDDIN': Knowledge Infusion and Distillation for Detection of INdecent Memes
Authors: Rahul Garg*, **Trilok Padhi***, Hemang Jain, Ugur Kursuncu, Ponnurangam Kumaraguru *(arXiv preprint, Accepted at ACL Findings 2025, Acceptance Rate = 19.1%)* (Research Area: Multimodal Reasoning, Knowledge Distillation)

Improving Contextual Congruence Across Modalities for Effective Multimodal Marketing using Knowledge Graphs – Authors: **Trilok Padhi**, Ugur Kursuncu, Yaman Kumar, Valerie Shalin, Lane Peterson *(arXiv preprint, Accepted at IEEE BigData 2024, Acceptance Rate = 18.7%)* (Research Area: Multimodal Learning, Knowledge Graphs)

CoKgCoOp: Enhancing Multimodal Learning through Commonsense Knowledge-Guided Context Optimization – Authors: **Trilok Padhi**, Ugur Kursuncu, Yaman Kumar, Valerie Shalin, Lane Peterson *(arXiv preprint coming soon, Under Review)* (Research Area: Commonsense Reasoning, Multimodal Learning)

Playing Devil's Advocate: Unmasking Toxicity and Vulnerabilities in Large Vision-Language Models – Authors: Abdulkadir Erol, **Trilok Padhi**, Agnik Saha, Ugur Kursuncu, Mehmet E. Aktas *(arXiv preprint, Under Review)* (Research Area: LLM Safety, Adversarial Robustness)

From Reddit to Generative AI: Evaluating Large Language Models for Anxiety Support Fine-tuned on Social Media Data – Authors: Ugur Kursuncu, **Trilok Padhi***, Jaya Krishna Mandivarapu*, Gaurav Kumar*, Christopher R. Larrison *(arXiv preprint, Under Review)* (Research Area: Mental Health AI, LLM Evaluation)

Moth flame optimized FLANN for prediction of earthquake magnitude
Authors: Santosh Kumar Majhi, Sk. Sajid Hossain, **Trilok Padhi**
(Accepted at Evolving Systems – An Interdisciplinary Journal for Advanced Science and Technology) (Research Area: Optimization Algorithms, Predictive Modeling)
* *Equal Contribution*

Working Experience

Rakuten Inc, Bangalore, India *Data Scientist* June 2021 - July 2022
Collaborated with a diverse team of 24 members from seven countries and four continents, under the supervision of Mr. Anirban Nandi, Vice President, Mr. Gaurav Ray, Associate Director

- Led projects in the Data Science Consulting Department on *impact estimation*, *funnel analysis*, and *customer acquisition*, developing an LSTM-Attn + DistilBERT–based *churn prediction model* with interpretable attention insights. Improved retention strategy, driving an estimated revenue impact of ¥10M annually.

Wowexp Technologies, Bangalore, India, *Deep Learning Researcher* Aug 2020 - May 2021

Technical Skills

Programming & Deep Learning: Python (PyTorch, NumPy, Pandas, scikit-learn)
Multimodal & Vision-Language Models: Development with Hugging Face Transformers, OpenCV, and pre-trained models (e.g., CLIP, LLaVA, LLaMA v1, v2, v3, Mistral, GPT-OSS) with expertise in LLM finetuning, pretraining, and posttraining (RLHF, PEFT, DPO, and alignment techniques).
Knowledge Graphs & Graph Neural Networks: PytorchGeometric, NetworkX, DGL
High-Performance Computing: MPI, CUDA, and Slurm for parallel processing and resource management.

Honors and Awards

ACL Student Travel Award 2025
IEEE BigData Student Travel Award 2024
National Champion (1st/200,000), Smart India Hackathon, Jaipur, India 2018
Academic Scholarship, Veer Surendra Sai University of Technology, Burla, India 2016 – 2020