

(720) 761-1716
Amherst, MA
ttripathi@umass.edu

TUHINA TRIPATHI

linkedin.com/in/tuhina-tripathi
github.com/tuhina2313
github.io/tuhina2313

EDUCATION

University of Massachusetts Amherst

Doctor of Philosophy, Computer Science | GPA: 3.85/4

Amherst, MA

SEP 2023

- Coursework: Machine Learning, Human-centric machine learning, Advanced Natural Language Processing, Advanced Information Assurance
- Advised by [Prof. Scott Niekum](#)

University of Colorado Boulder

Master of Science in Computer Science | GPA: 3.98/4

Boulder, CO

AUG 2021 - MAY 2023

- Coursework: Advanced Robotics, Robotic Manipulation, Distributed & Datacenter-Scale Systems, Decision Making under Uncertainty, Theoretical Foundations of Autonomous Systems
- Advised by [Prof. Bradley Hayes](#)

Delhi Technological University

Bachelor of Technology in Information Technology | GPA: 8.0/10

Delhi, India

AUG 2015 - MAY 2019

PUBLICATIONS

- **Pairwise or Pointwise? Evaluating Feedback Protocols for Bias in LLM-Based Evaluation(2025) | COLM 2025**
Tuhina Tripathi, Manya Wadhwa, Greg Durrett, Scott Niekum
- **Breaking the Tie: Evaluating Human Preferences in Reinforcement Learning (2023) | Masters Thesis**
Tuhina Tripathi, Bradley Hayes

EXPERIENCE

Research Assistant

Emotive Computing Lab, University of Colorado

AUG 2021 - JAN 2022

Boulder, CO

- Developed scalable preprocessing pipelines for multi-modal data (speech, vision, physiological signals), enabling robust affective computing and HRI research across distributed experimental datasets.

Software Development Engineer

Citicorp Services India Pvt. Ltd.

JUL 2019 - JUN 2021

Pune, India

- Led development of a production-grade data enrichment pipeline using SpringBoot and Angular to auto-patch trade discrepancies, reducing data inconsistencies by 65%.
- Designed and deployed Elasticsearch-based search infrastructure that cut average query latency by 80%, enabling faster access to high-volume financial data.

Research Intern

Indian Institute of Technology Delhi | Advised by [Prof. PK Kalra](#)

AUG 2018 - OCT 2018

Delhi, India

- Implemented GAN-based image enhancement system for latent fingerprints, surpassing state-of-the-art baselines with 3x improved NFIQ perceptual quality scores.

Research and Development Intern

Nucleus Software Exports Ltd.

JUN 2018 - JUL 2018

Noida, India

- Created a screenshot-to-HTML code generator deployed across internal teams.
- Developed a RASA-based multilingual chatbot in Hindi and Punjabi.

PAST RESEARCH

Emergent Misalignment with Narrow RL Tuning

Spring 2025

- Applied RL-tuning to a multi-turn LLM dialogue assistant in a simulated sales domain with profit-driven objectives.
- Demonstrated that narrow task optimization produced persuasive strategies in-domain but led to emergent misalignment on broader, open-ended dialogue tasks.

Prompt Recovery in LLMs

Spring 2024

- Fine-tuned open-weight language models to recover hidden prompts from rewritten text, building a synthetic dataset with diverse rewriting strategies.
- Achieved 73% strict and 87% loose accuracy, demonstrating strong generalization to unseen prompt styles.

Bounded-Risk IRL from Suboptimal Data

Spring 2022

Research project with *Prof. Zachary Sunberg*

- Developed a maximum entropy inverse reinforcement learning algorithm that learns from imperfect trajectories while ensuring safety via dynamic risk clipping at the sub-trajectory level.
- Validated performance on autonomous driving and gridworld navigation, achieving lower error than baseline IRL approaches.

Dynamic Obstacle Avoidance in Shared Human-Robot Workspaces

Spring 2022

Research project with *Prof. Nikolaus Correll*

- Implemented real-time POMDP-based motion planning with probabilistic goal prediction to coordinate with human collaborators.
- Improved task efficiency and safety in dynamic, sensor-driven environments, demonstrating scalable decision-making under uncertainty for interactive systems.

UNDERGRADUATE PROJECTS

BIO-METRIC IDENTIFICATION AND FINGERPRINT PERCEPTIVITY ENHANCEMENT

Undergraduate Major Project

2019

- Performed fingerprint enhancement using short-term Fourier Transform and Contextual filtering, to help intensify the ridges and minutiae [Dataset: 12k images from Optical and Capacitive sensors]. Enhanced images fed into a CNN for feature extraction. Achieved an accuracy of 98.32% with a significantly less False Acceptance Rate(FAR).

AUTOMATED HATE SPEECH DETECTOR

Undergraduate Minor Project

2018

- Implemented a solution to categorize tweets into hate speech, offensive language, and normal text using a dataset of 25k labeled tweets. Applied extensive preprocessing with TF-IDF scores and POS tags, and utilized Regression with L2 Regularization achieving a precision of 0.91 and F1 score of 0.90.

TEACHING

Object Oriented Programming (CS160)

Spring'25

Teaching Assistant

Introduction to Robotics (CSCI 3302/ECEN 3303)

Fall'22

Teaching Assistant

Starting Computing (CSCI 1300)

Spring'22, Summer'22

Teaching Assistant

SKILLS

Languages

Python, C++, C, Julia, SQL, Java, Javascript

Software & Tools:

PyTorch, TensorFlow, Transformers, OpenAI Gym, ROS, Webots, SpringBoot, Elasticsearch, Linux, Angular, Git, Docker

ACTIVITIES

Organizer, Machine Learning & Friends Lunch (MLFL) talk series at UMass Amherst

2025-2026

PhD Chair of the CSWomen club at UMass Amherst

2024-2025

Undergraduate Research Volunteer (URV) mentor at UMass Amherst

2023 - 2025

Graduate Peer Mentor at CU Boulder

Fall 2022

CitiCorp **Bronze award** for Enrichment Tool deployment on Production

2021

Best Innovation Award at Nucleus Software for 'Code Generator'

2018

Among the Top 5 teams in SIH' 17 conducted by Govt. of India

2017