PRIYANKA **KARGUPTA**

@ pk36@illinois.edu 🌐 pkargupta.github.io pkargupta in pkargupta

EDUCATION

University of Illinois, Urbana-Champaign
Ph.D. Student in Computer Science (2022-)
Overall GPA: 4.0/4.0
Advisor: Professor Jiawei Han

University of California, Berkeley
B.A. in Computer Science, Minor in Data Science:
Overall GPA: 3.773/4.0; Upper Division: 3.89/4.0
Advisor: Professor Ren Ng

EXPERIENCE

Data Mining Group: PhD Candidate

University of Illinois, Urbana-Champaign / 08/2022 – Current

• Funded by NSF Graduate Research Fellowship

• Designing models that exploit structured reasoning and knowledge to enhance both their own critical thinking and human users', in both the education & AI for Research domains.

Berkeley AI Research: Research Assistant

University of California, Berkeley / 08/2020 - 05/2022

• Funded by the Intel SRC Research Scholar Program

• Developed multi-object tracking and segmentation methods for self-driving applications.

• Explored dependencies in coordinate-based neural networks, applied to optimized reconstruction of neural radiance fields.

Principles and Techniques of Data Science: Student Instructor

University of California, Berkeley / 01/2021 - 05/2022

• Taught discussion and lab sections of 30 students each, hold office hours, and assist with various aspects of running the course of over 900 students.

Vyncs (under Agnik): Project Manager

05/2018 - 12/2021 (Full-time summers, part-time school year)

• Led development of sensor-based on-demand location detection and sharing mobile application w/ 70,000+ users.

• Led design & development of an Android augmented reality app which renders visual location data (e.g. nearby restaurants, traffic accidents) via user's phone camera.

Microsoft Research: Knowledge and Reasoning

Redmond, Washington / 06/2025 - 09/2025

• Training an LLM using RL to inject strategic noise into its reasoning process to spark creativity and break out of predictable generations.

• Tasks: Generating scientific research proposals & code.

Intel: Algorithms Research Intern & Product Manager

3DAT - Olympics Technology Group / 01/2022 - 07/2022

• Designed algorithms and optimize ML models for inferred 3D pose estimation for athlete tracking.

• Developed product roadmaps and manage the group's 2D skeletal product.

IBM T.J. Watson Research Center: Research Intern

Efficient & Resilient Systems @ Yorktown Heights / 05/2021 - 09/2021

• Worked on collaborative detection and tracking for autonomous vehicles

• Led demo creation & production for the 2021 DARPA Electronics Resurgence Initiative Summit

NASA Ames Research Center: Research Intern

Data Sciences Group, Intelligent Systems Division / 05/2017 - 02/2018

• Designed and implemented deep image classification algorithms for swarms of drones in loosely coupled environments.
• *Poster presentation at the Ames Center symposium.*

AWARDS

• 2024 NSF Graduate Research Fellowship
• EMNLP'2024 Outstanding Reviewer
• 2020-2021 Intel Semiconductor Research Corp. Research Scholar
• 2021 Computer Graphics (CS184) Showcase Winner
• 2020 (Team Lead) Global Top 15 Winner for Health & Wellness: The Global Hack (12,000+ participants from 100+ countries)
• 2018 National Center for Women & IT (NCWIT) Maryland Winner
• 2018 NCWIT National Honorable Mention (Top 10% out of 3,600 applications)
• 2017 National Top 15 Company & JA Maryland Company of the Year
• 2016 Rep. Elijah Cummings's Congressional App Challenge Winner
• 2016 MIT Media Lab/QuHacks Hackathon Winner: 1st Place

SELECTED PUBLICATIONS

Tree-of-Debate: Multi-Persona Debate Trees Elicit Critical Thinking for Scientific Comparative Analysis
Priyanka Kargupta, Ishika Agarwal, Tal August, Jiawei Han
ACL 2025 Main Conference (Oral)

• Determine the novelties, similarities, and equivalent ideas within papers through the reasoning induced by author personas debating. Structures the debates as a tree of rounds, augmented with fine-grained iteratively retrieved paper segments.

Synergizing Unsupervised Episode Detection with LLMs for Large-Scale News Events
Priyanka Kargupta, Yunyi Zhang, Yizhu Jiao, Siru Ouyang, Jiawei Han
ACL 2025 Main Conference (Oral)

• Proposing a new task, episode detection, which seeks to detect episodes from a news corpus containing key event articles.
• Designed unsupervised framework which automatically partitions a key event corpus into cohesive episode sequences.

Beyond True or False: Retrieval-Augmented Hierarchical Analysis of Nuanced Perspectives
Priyanka Kargupta*, Runchu Tian*, Jiawei Han
ACL 2025 Main Conference

• Scientific and political claims are often nuanced and are not strictly "true" or "false" (e.g., Vaccine A is better than B). Automatically constructs a hierarchy of subaspects and corresponding perspectives on a given claim given a corpus.

TaxoAdapt: Automatically Adapting LLM-Constructed Taxonomies to Evolving Research Corpora
Priyanka Kargupta, Nan Zhang, Yunyi Zhang, Rui Zhang, Prasenjit Mitra, Jiawei Han
ACL 2025 Main Conference

• A self-supervised framework which expands a taxonomy based on whether a conference-level corpus indicates that a node needs further *exploration*. Proposes a unified enrichment-classification-expansion method.

Instruct, Not Assist: Improving Socratic Questioning for Active Multi-Turn Code Debugging
Priyanka Kargupta, Ishika Agarwal, Dilek Hakkani-Tur, Jiawei Han
EMNLP'24 Findings; Poster @ NeurIPS'24 Workshop on Educational Assessment

• Designing a framework to re-orient a language model towards instructor-like responses and questions for helping users debug code in an educational domain. Guides conversations using a state-representation-based dynamic tree.

MEGClass: Text Classification with Extremely Weak Supervision via Mutually-Enhancing Text Granularities
Priyanka Kargupta, Tanay Komarlu, Susik Yoon, Xuan Wang, Jiawei Han
EMNLP'23 Findings

• Designed an extremely weakly-supervised text classification method that leverages multiple, mutually-enhancing text granularities. Learns a contextualized document representation that captures the most discriminative class indicators.

ONGOING PUBLICATIONS

Day & Night Entropy: Sparking Inspiration Via Adaptive Structured & Unstructured Reasoning *(Lead)*

• Using reinforcement learning to train a model to understand how and when to inject strategic high-level semantic noise (high entropy tokens) into the reasoning process for both open-ended and deterministic tasks.

Structures of Thought: Analyzing the Gap Between Human & LLM Cognitive Reasoning Behaviors *(Co-Lead)*

• Presenting a comprehensive taxonomy of cognitive reasoning behaviors and utilizing it to systematically analyze which behaviors are and are not (1) exhibited by state-of-the-art LLMs and (2) explored within the NLP reasoning community.

• Proposing a knowledge-driven mechanism to elicit reasoning behaviors better aligned to the task. This improves both quality and efficiency of generation.

SKILLS

• ***Programming:*** Python, Java, HTML5/CSS, C, SQL, JavaScript, Android Development, C++, C#, R
• ***Development Tools:*** verl, vLLM, PyTorch, Transformers, HuggingFace, scikit-learn, scikit-image, OpenCV, AWS (EC2, S3, Lambda), Alexa Skill Development, Spark, GeoSpark, Pandas, Plotly, Caffe, .NET, Linux, Firebase, Git, Docker, vLLM
• ***Graphic & Video Design:*** Blender, Adobe Photoshop, Adobe Illustrator, Adobe Premiere Pro, Camtasia