dasgupta.ishita@gmail.com
ishita-dg.github.io

Ishita Dasgupta

Researcher in machine learning and computational cognitive science.
25+ publications, 12 in machine learning, 14 in computational cognitive science; h-index 13.
20+ invited talks and panels; see Academic CV.

PROFESSIONAL EXPERIENCE

Senior Research Scientist — Nov 2022 – Present

Research Scientist — Dec 2020 – Oct 2022

DeepMind, New York City

- Extracting information from large language models for use in complex embodied action spaces.
  - Co-led a research team of 6 as part of a 50+ person long-term research effort.
  - Collaborating with language models for embodied reasoning.
    Dasgupta *et al.* NeurIPS LaReL 2022; **best paper award**.
  - Distilling internet-scale vision-language models into embodied agents.
    Sumers, ... Dasgupta. (senior author) ICML 2023
- Cognitive abilities (deliberation, reasoning) in large language models.
  - Developing and evaluating multimodal reasoning capabilities in mainline Gemini models.
  - Language models show human-like content effects on reasoning.
    Dasgupta*, Lampinen* et al. arXiv 2022, journal submission in prep.
  - Can language models learn from explanations in context?
    Lampinen, Dasgupta et al. EMNLP 2022.
- First Research Scientist hire at DeepMind NYC; key role in research vision, growth, and culture.
- Several other publications, details on Google Scholar.

Princeton University. Postdoctoral Fellow, Dept. of Computer Science. Jan – Dec 2020

– Analyzing and manipulating representations learned by AI systems, focus on inductive bias & abstraction.
– 8 publications (3 first author), including a NeurIPS Outstanding Main Track Paper award.

EDUCATION

Harvard University, Ph.D. in Physics. March 2020
Thesis: *Algorithms for ecological rationality in humans and machines.*

Indian Institute of Technology Bombay, B.Tech. with Honours in Engineering Physics. August 2014

RECENT AWARDS

NeurIPS Outstanding Main Track Paper. 2022
Using natural language and program abstractions to instill human inductive biases in machines.

Best Paper Award, NeurIPS Language and Reinforcement Learning Workshop. 2022
Collaborating with language models for embodied reasoning.

SKILLS

Machine learning methods and large-scale training and serving frameworks.
Experiment design, collection, and analysis for online human behavioral data.
Technical leadership (leading research teams of 5+) and mentorship (supervised 6+ graduate student projects).

SERVICE

Reviewing. Area Chair (NeurIPS 2023, NeurIPS workshops 2022, 2021), extensive peer reviewing.
Volunteering and Outreach. Harvard Women in Physics, Teach for India Mumbai, DeepMind Scholars.