dasgupta.ishita@gmail.com
Academic CV

Ishita Dasgupta

Researcher in machine learning and cognitive science.

PROFESSIONAL EXPERIENCE

DeepMind, New York City.
Senior Research Scientist — Nov 2022 – Present
Research Scientist — Dec 2020 – Oct 2022

- First Research Scientist hire at DeepMind NYC; played a key role in the team's research vision, growth, and culture.
- Cognitive science approaches for understanding and improving large-scale language models. Selected works:
  - Language models show human-like content effects on reasoning [arXiv 2022]
  - Transformers generalize differently from information stored in-context vs in-weights [NeurIPS MemARI 2022]
  - Can language models learn from explanations in context? [EMNLP 2022]
- Extracting information from language models for use in complex embodied action spaces. Selected works:
  - Collaborating with language models for embodied reasoning [NeurIPS LaReL 2022; best paper award]
  - Distilling Internet-Scale Vision-Language Models into Embodied Agents [ICML 2023]
- 20+ publications, 20+ invited talks; details in Google Scholar profile.
- Co-led a research team of 6, represented this team's work as part of a 50+ person long-term research effort.
- Supervised 2 interns, mentored 2 DeepMind junior engineers and 2 Google student researchers.

Princeton University. Postdoctoral Fellow, Dept. of Computer Science. Advisor: Prof. Tom Griffiths. Jan – Dec 2020

- Studied learning and representation in artificial systems, emphasis on inductive biases and complexity.
- 8 journal / conference publications, including a NeurIPS Outstanding Main Track Paper award.
- Supervised 3 graduate and undergraduate researchers.

Harvard University. Graduate Researcher, Dept. of Physics. Advisor: Prof. Sam Gershman. 2015 – 2019

- Built new models of human learning and decision-making that incorporate methods from modern machine learning (e.g. amortized Monte Carlo sampling, variational auto-encoders, memory-based meta-learning, and data augmentation).
- Ran large-scale experiments on online crowdsourcing platforms like Amazon MTurk and Prolific.
- Explained a large number (15) of previously unexplained empirical findings from economics, neuroscience & psychology.
- 9 articles in peer-reviewed journals and conferences, 4 academic awards (see Selected Awards).

Stanford University. Visiting Researcher, Dept. of Psychology. Advisor: Prof. Noah Goodman. May – Aug 2017

- Investigated compositionality in vector embeddings of natural language, collaboration led to 3 peer-reviewed articles.

EDUCATION

Harvard University, Ph.D. in Physics. March 2020
Indian Institute of Technology Bombay, B.Tech. (+ Honors) in Engineering Physics. August 2014

SCIENTIFIC PRESENTATIONS

Colloquium Speaker: Broad-interest audiences of 100+ at NYU, UC Irvine, American Psychological Society.
Conference presentations: 10+ presentations at machine learning and cognitive science conferences, see publications.
Other invited talks: 15+ seminars to audiences of 20-70 at e.g. Brown, Berkeley, MIT, McGill, MPI, UCL.
Teaching: Teaching fellow for 3 courses in physics and mathematics. Student feedback here.

SELECTED 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.
Mind Brain Behavior Graduate Student Award. Fully funded research visit to Stanford University. 2017
Student Award for Outstanding Scientific Contribution. International Conference on Thinking. 2016
Amartya Sen Fellowship. Graduate School of Arts and Sciences, Harvard University. 2014

SKILLS

Python (8 years), R (3 years), C++, JavaScript (2 years).
Machine learning: probabilistic inference, natural language, meta-learning, large models (JAX, Tensorflow, PyTorch).
Experiment design, collection, and analysis for human behavioral data online (MTurk, Prolific).
Experienced in project leadership, technical communication, and mentorship.

SERVICE

Reviewing. Area Chair (NeurIPS 2023, NeurIPS workshop 2022, 2021), reviewer for 5+ other journals / conferences.
Volunteering and Outreach. Harvard Women in Physics, Teach for India Mumbai, DeepMind Scholars.