

Rachit Bansal

Doctorate Student, Harvard University

📧 rachitbansal.github.io @ rachitbansal@g.harvard.edu 🎓 [Google Scholar](#)

Education

Harvard University, Cambridge

08/2024 – Present

Ph.D. in Computer Science (ongoing)

> Advised by [Prof. Sham Kakade](#) and [Prof. David Alvarez-Melis](#).

> Part of the [Harvard ML Foundations](#) research group.

> [Kempner Institute](#) Graduate Fellow.

Delhi Technological University, India

08/2018 – 07/2022

B.Tech. in Electrical Engineering

> Bachelor's thesis at the Technion, Israel (02/2022 – 07/2022).

Experience

Meta, California

05/2025 – 10/2025

Research Scientist Intern at Meta Super-intelligence Labs (MSL) with [Aston Zhang](#)

Google DeepMind, India

07/2022 – 07/2024

Pre-doctoral Researcher with [Partha Talukdar](#) and [Prateek Jain](#)

Technion, Israel

09/2021 – 07/2022

Research Intern (Bachelor's Thesis) with [Yonatan Belinkov](#)

Adobe Research, India

01/2021 – 09/2021

Research Intern with [Balaji Krishnamurthy](#)

Google Summer of Code, Remote | University of Oxford

05/2020 – 01/2021

Contributor at CDLI with [Jacob Dahl](#)

Publications

- [1] **RL Excursions during Pre-Training: Re-examining Policy Optimization for LLM training** 📄 🔗
[Rachit Bansal](#)*, Clara Mohri*, Tian Qin*, David Alvarez-Melis, Sham Kakade
Conference on Neural Information Processing Systems (under review) [NeurIPS 2026]
- [2] **Why Larger Models Learn More: Effects of Capacity, Interference, and Rare-Task Retention** 📄
Jing Huang, Daniel Wurgaft, [Rachit Bansal](#), Laura Ruis, Naomi Saphra, David Alvarez-Melis, Andrew Kyle Lampinen, Christopher Potts, Ekdeep Singh Lubana
Conference on Neural Information Processing Systems (under review) [NeurIPS 2026]
- [3] **Interleaved Head Attention** 📄
Sai Surya Duvvuri*, Chanakya Ekbote*, [Rachit Bansal](#), Rishabh Tiwari, Devvrit Khatri, David Brandfonbrener, Paul Liang, Inderjit Dhillon, Manzil Zaheer
Conference on Neural Information Processing Systems (under review) [NeurIPS 2026]
- [4] **Let's (not) just put things in Context: Test-Time Training for Long-Context LLMs** 📄
[Rachit Bansal](#), Aston Zhang, Rishabh Tiwari, Lovish Madaan, Sai Surya Duvvuri, Devvrit Khatri, David Brandfonbrener, David Alvarez-Melis, Prajjwal Bhargava, Mihir Sanjay Kale, Samy Jelassi
International Conference on Learning Representations [ICLR 2026]
- [5] **The Art of Scaling Reinforcement Learning Compute for LLMs** 📄
Devvrit Khatri, Lovish Madaan, Rishabh Tiwari, [Rachit Bansal](#), Sai Surya Duvvuri, Manzil Zaheer, Inderjit S. Dhillon, David Brandfonbrener, Rishabh Agarwal
International Conference on Learning Representations (Oral Presentation) [ICLR 2026]
- [6] **Adam or Gauss-Newton? A Comparative Study In Terms of Basis Alignment and SGD Noise** 📄
Bingbin Liu*, [Rachit Bansal](#)*, Depen Morwani*, Nikhil Vyas, David Alvarez-Melis, Sham M. Kakade
Conference on Neural Information Processing Systems (under review) [NeurIPS 2026]

- [7] **LLM Augmented LLMs: Expanding Capabilities through Composition**  
Rachit Bansal, Bidisha Samanta, Siddharth Dalmia, Nitish Gupta, Shikhar Vashishth, Sriram Ganapathy, Abhishek Bapna, Prateek Jain, Partha Talukdar
International Conference on Learning Representations [ICLR 2024]
- [8] **Linear Connectivity Reveals Generalization Strategies**  
Jeevesh Juneja, Rachit Bansal, Kyunghyun Cho, João Sedoc, Naomi Saphra
International Conference on Learning Representations [ICLR 2023]
- [9] **Measures of Information Reflect Memorization Patterns**   
Rachit Bansal, Danish Pruthi, Yonatan Belinkov
Conference on Neural Information Processing Systems [NeurIPS 2022]
- [10] **Evaluating Explanations: How much do explanations from the teacher aid students?**  
Danish Pruthi, Rachit Bansal, Bhuvan Dhingra, Livio Baldini Soares, Michael Collins, Zachary C. Lipton, Graham Neubig, William W. Cohen
Transactions of the Association for Computational Linguistics
Presented at the Annual Conference for the Association of Computational Linguistics [TACL 2022]
- [11] **CoSe-Co: Text Conditioned Generative Commonsense Contextualizer**  
Rachit Bansal, Milan Aggarwal, Sumit Bhatia, Jivat Kaur, Balaji Krishnamurthy
North American Chapter of the Association for Computational Linguistics [NAACL 2022]
- [12] **LM-CORE: Language Models with Contextually Relevant External Knowledge**  
Jivat Kaur, Sumit Bhatia, Milan Aggarwal, Rachit Bansal, Balaji Krishnamurthy
North American Chapter of the Association for Computational Linguistics (Findings) [NAACL 2022]
- [13] **How Low is Too Low? A Computational Perspective on Extremely Low-Resource Languages**   
Rachit Bansal, Himanshu Choudhary, Ravneet Punia, Niko Schenk, Jacob L Dahl, Émilie Pagé-Perron
Student Research Workshop (SRW) at ACL [ACL SRW 2021]

Awards and Honors

Kempner Institute Graduate Fellowship	Harvard University	2024 – Present
Best Research Thesis Award	Delhi Technological University	2022
Google Summer of Code	Contributor (2020) and Mentor (2022) with the Cuneiform Digital Library Initiative	

Teaching

AC297R , Harvard University.	
Reinforcement Learning , Coding Blocks.	2020
Foundations of Machine Learning & Deep Learning , Coding Blocks.	2019