

Prateek Chanda

Computer Science PhD Student, IIT Bombay · Prime Minister Research Fellow

prateekkol21@gmail.com | prateekch@cse.iitb.ac.in | prateekchanda21.github.io | Google Scholar | GitHub | LinkedIn | +91-8337055526

EXPERIENCE

Google DeepMind India

Student Researcher (two terms)

Bangalore, India
Jul–Nov 2023, Dec 2023–May 2024

Host: Pradeep Shenoy | Mentors: Dheeraj Nagaraj, Arun Suggala, Prateek Jain

- Efficient pretraining and knowledge distillation of large language models for proprietary **Gemini** and **LaMDA** models.
- Developed model-dependent data distillation to improve LLM pretraining efficiency (LaMDA).
- Implemented **MatFormer** on LaMDA with distillation; improved downstream evals (+0.40% GPT-Gen, +0.57% GPT-Rank) over baseline.
- Explored curriculum learning, meta-learned instance reweighting and skill-ordered training (**Skill-It**).

Microsoft Research India

Research Intern

Bangalore, India
Nov 2025 – Mar 2026

Mentors: Nagarajan Natarajan, Amit Sharma, Subbarao Kambhampati

- Co-developed **Interwhen** (NeurIPS 2026) [paper] [code] [MSR talk]: a single-trajectory test-time verification framework that avoids tree search and decouples verification from generation.
- Implemented asynchronous monitoring (periodic forked extraction and verifier calls) with negligible overhead on correct runs; near-perfect constraint satisfaction at a fraction of the token compute of verification baselines.
- Auto-synthesized code verifiers from natural-language policies (incl. Lean/z3); raised τ^2 -bench telecom completion from 32% to 87% (Qwen3-30B).

Microsoft Research India

SCAI Center Fellow (Research Software Engineer)

Bangalore, India
Nov 2019 – Feb 2022

Supervisors: Amit Sharma, Vageesh Chandramouli

- Built user embeddings from telemetry and response logs; shipped collaborative, content-based and causal recommenders for mental-health micro-interventions (CSCW Demo 2021).

SunPy (NASA Open Source)

Remote

Google Summer of Code

Dec 2016 – Apr 2018

- Solar data retrieval, processing and SQL-backed storage for multi-year heliophysics analyses; contributions in multiple NASA GSFC software releases.

PUBLICATIONS

- [P1] **Interwhen: A Generalizable Framework for Verifiable Reasoning with Test-time Monitors** NeurIPS 2026
Vishak K Bhat*, Prateek Chanda*, Vijal Ekbote*, Ashmit Khandelwal, Maitreyi Swaroop, Vineeth N. Balasubramanian, Subbarao Kambhampati, Nagarajan Natarajan, Amit Sharma [arXiv] [code]
- [P2] **From Task Graphs to Training Mixtures: Graph-Structured Data Allocation for Language Models** LoG 2026
Prateek Chanda, Saral Sureka, Ganesh Ramakrishnan [OpenReview]
- [P3] **Online Distributional Prediction via Latent Cluster Geometry Under Drift and Corruption** TMLR 2026
Navyansh Mahla, Prateek Chanda, Ganesh Ramakrishnan [arXiv]
- [P4] **Minibatch Selection for Language Models via Partition Matroid Constrained Gradient Matching** ICML 2026
Prayas Agrawal*, Prateek Chanda*, Ishita Khatri, Ganesh Ramakrishnan, Bamdev Mishra, Pratik Jawanpuria [arXiv] [code]
- [P5] **UniPROT: Uniform Prototype Selection via Partial Optimal Transport with Submodular Guarantees** AISTATS 2026
Prateek Chanda*, Prayas Agrawal*, Karthik S. Gurumoorthy, Ganesh Ramakrishnan, Bamdev Mishra, Pratik Jawanpuria [arXiv] [code]
- [P6] **Sculpting Subspaces: Constrained Full Fine-Tuning in LLMs for Continual Learning** ICLR 2026
Nikhil Shivakumar Nayak, Krishnateja Killamsetty, Ligong Han, Abhishek Bhandwal, Prateek Chanda, Kai Xu, Hao Wang, Aldo Pareja, Oleg Silkin, Mustafa Eyceoz, Akash Srivastava [arXiv] [code]
- [P7] **Bandit Guided Submodular Curriculum for Adaptive Subset Selection** NeurIPS 2025
Prateek Chanda*, Prayas Agrawal*, Lokesh Reddy Polu, Saral Sureka, Atharv Kshirsagar, Ganesh Ramakrishnan [arXiv] [code]
- [P8] **FairPO: Fair Preference Optimization for Multi-Label Learning** NeurIPS OPT 2025
Soumen Kumar Mondal, Prateek Chanda, Akshit Varmora, Ganesh Ramakrishnan [pdf]
- [P9] **Few Shot Chain-of-Thought Driven Reasoning to Prompt LLMs for Open-Ended Medical Question Answering** Findings EMNLP 2024
Saeel Sandeep Nachane*, Ojas Gramopadhye*, Prateek Chanda*, Ganesh Ramakrishnan, Kshitij Sharad Jadhav, Yatin Nandwani, Dinesh Raghu, Sachindra Joshi [paper]
- [P10] **Bayesian Coreset Optimization for Personalized Federated Learning** ICLR 2024
Prateek Chanda, Shrey Modi, Ganesh Ramakrishnan [arXiv]
- [P11] **Would I Regret Later Joining This Community? Temporal Neighborhood Information for Community Retention in a Game Theoretic Community Detection Framework** AAAI-22 ML4OR
Prateek Chanda, Susanta Chakraborty
- [P12] **Robust Deep Reinforcement Learning Control Framework** AAAI-22 RAISA
Prateek Chanda

- [P13] **MindNotes: A Mobile Platform to Enable Users to Break Stigma around Mental Health and Connect with Therapists** CSCW Demo 2021
Prateek Chanda, Amogh Wagh, Jemimah A. Johnson, Swaraj Renghe, Vageesh Chandramouli, George Mathews, Sapna Behar, Poornima Bhola, Girish Rao, Paulomi Sudhir, T.K. Srikanth, Amit Sharma, Seema Mehrotra
- [P14] **Distributed Anomaly Detection in Edge Streams using Frequency Based Sketch Data Structures** arXiv
Prateek Chanda, Malay Bhattacharyya
- [P15] **A Sketch Based Game Theoretic Approach to Detect Anomalous Dense Sub-Communities in Large Data Streams** arXiv
Prateek Chanda, Aadirupa Saha
- [P16] **A Novel Graph Based Clustering Approach to Document Topic Modeling** ICCCNT 2018
Prateek Chanda, Asit Kr. Das
- [P17] **The SunPy Project: Open Source Development and Status of the Version 1.0 Core Package** ApJ 2020
Stuart Mumford, Prateek Chanda, The SunPy Community. Also: SunPy: A Python Package for Solar Physics (JOSS 2020); SunPy v1.0 (NASA ADS 2019)

* denotes equal contribution.

ACHIEVEMENTS

- **Project Tapestry (2026):** Proposal on *Efficient Federated Finetuning for Language Models* accepted to [Project Tapestry](#) (AI Alliance; Chief Scientific Advisor: Yann LeCun)
- **India's Top 100 AI/ML Researchers:** Recognized by Amazon for research excellence in AI/ML
- **Fellowships:** Microsoft PhD Award (2025–2026); Prime Minister Research Fellowship (2022–2027); Qualcomm Innovation Fellowship Finalist (2025, 2026)
- **Travel and schools:** MLSS 2026, Max Planck Institute Tübingen (invited, fully funded); AISTATS Financial Award (2026); ICLR Travel Grant (2024); invited to Google PhD Networking Reception at NeurIPS (2025)
- **NASA Group Achievement Award (2024):** For contributions to the SunPy project
- **Mentorship programs:** AISTATS 2022 Mentorship Program with [Aadirupa Saha](#) (Microsoft Research NYC); AI for Science (NeurIPS) Mentorship Program
- **Earlier:** Microsoft Garage Hackathon 2020 NGO Award (AI for Social Good); MSR India internship sponsorship; GAABESU Research Award, IEST (2018); JBNSTS Scholar

INVITED TALKS

- *Introducing Interwhen: Steering Reasoning Agents with Real-Time Verification*, Microsoft Research (2026) [\[video\]](#)
- *Bandit Guided Adaptive Subset Selection*, Microsoft Research India Academic Summit (Jun 2025)
- *Adaptive Dimension towards Group DRO*, ML Collective Research Jam #25 (Nov 2024) [\[slides\]](#)

SKILLS

- **Languages:** Python, C++, SQL, Java, C#, TypeScript **Technologies:** Azure, Azure ML Studio, GitHub, GitLab, Jekyll, GCP
- **Libraries:** PyTorch, TensorFlow, scikit-learn, Pandas, Jupyter, Microsoft Graph SDK
- **Open source:** [Interwhen](#) (Microsoft Research), [Efficiency Learning](#), [DECILE](#), [SunPy](#)

EDUCATION

Indian Institute of Technology Bombay

Mumbai, India
2022 – Present

Ph.D. in Computer Science & Engineering (PMRF; Microsoft PhD Award)

Advisors: [Prof. Ganesh Ramakrishnan](#) and [Prof. Pratik Jawanpuria](#)

Coursework: Optimization in ML (AA); Advanced ML (AB); Foundations of ML (AB); Deep Learning Theory (AB); Information Retrieval & Mining (BB); Seminar (AA)

Teaching: PMRF Instructor: Subset Selection for LLMs (Aut 2025), Convex Optimization & Submodular Methods (Aut 2024), Advanced Topics in ML (Aut 2023). TA: CS 218 (Spr 2026), CS 419 (Spr 2023, Aut 2025), CS 769 (Spr 2024, Spr 2025), CS 602 (Aut 2024), CS 105 (Aut 2023), CS 335/337 (Aut 2022)

Indian Institute of Engineering Science and Technology, Shibpur

Howrah, India
2015 – 2019

B.Tech. in Computer Science & Engineering (First Class Honours)

Thesis: *Avoiding Past Choice Regrets: A Game Theoretic Community Detection using Temporal Information* [\[report\]](#)

POSITIONS OF RESPONSIBILITY

- **Reviewer:** NeurIPS (2024, 2025, 2026), ICML 2026 (**Gold Reviewer Award**), ICLR (2025, 2026; **Top 200 Reviewer**), AAAI (2025, 2027), AISTATS (2022, 2026), LoG (2024, 2026), GCLR @ AAAI 2024; TMLR (2025, 2026), IEEE Transactions on Mobile Computing, COMSNETS, JOSS
- **Program Committee:** [ODD Workshop @ SIGKDD](#) (2021)
- **Mentor:** Mentored 80+ students in Google Code-In (2018) and Hacktoberfest (2017–2018); GSoC mentor for SunPy