

Abhishek Panigrahi

Graduate Student, Princeton University

🔗 <https://abhishekpanigrahi1996.github.io> ◊ @ ap34@princeton.edu

EDUCATION

- **Princeton University** Jan'21 – Present
Ph.D. student in Computer Science
Advisor: *Prof. Sanjeev Arora*
Apple AI/ML Ph.D. Scholar and Siebel Scholar (2025-26)
- **Indian Institute of Technology, Kharagpur** July'14 – May'18
B.Tech in [Computer Science and Engineering](#)
Cum. GPA: 9.90/10, Major GPA: 10/10, Institute Rank : **1 (Out of 1400 students)**
President of India Gold Medal and Institute Silver Medal 2018 for academic performance

INDUSTRY EXPERIENCE

- **Google New York** Jun'23 – Dec'23
Student Researcher
 - Advisors: *Dr. Sashank Reddi*, *Dr. Satyen Kale*, and *Dr. Nikunj Saunshi*.
 - Projects - Efficient and Asynchronous Pre-training of Large Language models.
- **Microsoft Research India** July'18 – Dec'20
Research Fellow
 - Advisors: *Dr. Harsha Vardhan Simhadri* and *Dr. Navin Goyal*.
 - Projects - Unsupervised Embeddings and Analysis of Deep Learning algorithms.

RESEARCH INTERESTS

I study efficient and robust training of language models using mathematical models. My research spans curriculum learning, adaptive architectures, and theoretical analysis of optimization dynamics.

SELECTED CONFERENCE PUBLICATIONS

($\alpha-\beta$) denotes author order being alphabetical, * denotes equal contribution

- **AdaptMI: Adaptive Skill-based In-context Math Instruction for Small Language Models** [Openreview]
[arxiv]
Yinghui He, Abhishek Panigrahi, Yong Lin, Sanjeev Arora
Accepted at *Conference on Language Modeling (COLM 2025)*
- **On the power of Context-Enhanced Learning in LLMs** [Openreview]
[arxiv]
Xingyu Zhu*, Abhishek Panigrahi*, Sanjeev Arora
Spotlight presentation (Top 2.6%) at *International Conference on Machine Learning (ICML 2025)*
- **Generalizing from SIMPLE to HARD Visual Reasoning: Can We Mitigate Modality Imbalance in VLMs?** [Openreview]
[arxiv]
Simon Park*, Abhishek Panigrahi*, Yun Cheng*, Dingli Yu, Anirudh Goyal, Sanjeev Arora
Accepted at *International Conference on Machine Learning (ICML 2025)*
- **Progressive distillation induces an implicit curriculum** [Openreview]
[arxiv]
Abhishek Panigrahi*, Bingbin Liu*, Sadhika Malladi, Andrej Risteski, Surbhi Goel
Oral presentation (Top 1.8%) at *International Conference on Learning Representations (ICLR 2025)*
- **Representing Rule-based Chatbots with Transformers** [ACL]
[arxiv]
Dan Friedman, Abhishek Panigrahi, Danqi Chen
Accepted at *Nations of the Americas Chapter of the Association for Computational Linguistics (NAACL 2025)*

- **Efficient Stagewise Pretraining via Progressive Subnetworks** [\[Openreview\]](#)
Abhishek Panigrahi*, Nikunj Saunshi*, Kaifeng Lyu, Sobhan Miryoosefi, Sashank Reddi, Satyen Kale, Sanjiv Kumar [\[arxiv\]](#)
Accepted at *International Conference on Learning Representations (ICLR 2025)*
- **Trainable Transformer in Transformer** [\[PMLR\]](#)
Abhishek Panigrahi*, Sathika Malladi*, Mengzhou Xia, Sanjeev Arora [\[arxiv\]](#)
Accepted at *International Conference on Machine Learning (ICML 2024)*
- **Do Transformers Parse while Predicting the Masked Word?** [\[ACL\]](#)
Haoyu Zhao*, Abhishek Panigrahi*, Rong Ge, Sanjeev Arora [\[arxiv\]](#)
Accepted at *Empirical Methods in Natural Language Processing (EMNLP 2023)*
- **Task-Specific Skill Localization in Fine-tuned Language Models** [\[PMLR\]](#)
Abhishek Panigrahi*, Nikunj Saunshi*, Haoyu Zhao, Sanjeev Arora [\[arxiv\]](#)
Accepted at *International Conference on Machine Learning (ICML 2023)*
- **On the SDEs and Scaling Rules for Adaptive Gradient Algorithms** [\[OpenReview\]](#)
Sadhika Malladi*, Kaifeng Lyu*, Abhishek Panigrahi, Sanjeev Arora [\[arxiv\]](#)
Accepted at *Neural Information Processing Systems (NeurIPS 2022)*
- **Understanding Gradient Descent on Edge of Stability in Deep Learning** [\[PMLR\]](#)
Sanjeev Arora, Zhiyuan Li, Abhishek Panigrahi ^($\alpha-\beta$) [\[arxiv\]](#)
Accepted at *International Conference on Machine Learning (ICML 2022)*
- **Learning and Generalization in RNNs** [\[OpenReview\]](#)
Abhishek Panigrahi, and Navin Goyal [\[arxiv\]](#)
Accepted at *Neural Information Processing Systems (NeurIPS 2021)*
- **Effect of Activation Functions on the Training of Overparametrized Neural Nets** [\[OpenReview\]](#)
Abhishek Panigrahi, Abhishek Shetty and Navin Goyal [\[arxiv\]](#)
Accepted at *International Conference on Learning Representations (ICLR 2020)*.
- **Word2Sense: Sparse Interpretable Word Embeddings** [\[ACL\]](#)
Abhishek Panigrahi, Harsha Vardhan Simhadri and Chiranjib Bhattacharyya
Oral presentation (Top 1%) at *Association for Computational Linguistics (ACL 2019)*.

PRE-PRINTS AND UPCOMING WORKS

- **In Good GRACES: Principled Teacher Selection for Knowledge Distillation**
Abhishek Panigrahi, Bingbin Liu, Sathika Malladi, Sham Kakade, Surbhi Goel
Initial version accepted at Foundations of Reasoning in Language Models, NeurIPS'25
- **STAT: Skill-Targeted Adaptive Training**
Yinghui He, Abhishek Panigrahi, Yong Lin, Sanjeev Arora
Initial version accepted at Foundations of Reasoning in Language Models, NeurIPS'25
- **Non-Gaussianity of Stochastic Gradient Noise** [\[arxiv\]](#)
Abhishek Panigrahi, Raghav Somani, Navin Goyal and Praneeth Netrapalli
Presented at *Science meets Engineering of Deep Learning workshop, NeurIPS 2019 (SEDL 2019)*.
- **Analysis on gradient propagation in batch normalized residual networks** [\[arxiv\]](#)
Abhishek Panigrahi, Yueru Chen, C.-C. Jay Kuo

TALKS

- (Upcoming) What Makes a Good Teacher? Principles and Metrics for Effective Knowledge Distillation
 - University of Wisconsin-Madison (12/3/25)
- (Upcoming) Teaching Language Models like Students: Skill Ontologies for Targeted Learning
 - Cornell Tech (11/18/25)

- On the power of Context-Enhanced Learning in LLMs
 - Apple ML Seminar (7/10/25); hosted by [David Grangier](#)
 - Apple Foundation Model Workshop (9/17/25); hosted by [Prof. Guillermo Sapiro](#)
- Efficient “curriculum-based” training: Theoretical modeling through synthetic testbeds
 - Young Researcher Seminar Series (5/21/25), hosted by [Toyota Technological Institute, Chicago](#)
 - Carnegie Mellon University (4/18/25); group meeting of [Prof. Andrej Risteski](#)
 - Foundations of Machine Learning seminar (4/16/25), hosted by [Kempner Institute, Harvard](#)
 - Northeastern University (4/15/25), group meeting of [Prof. Hongyang Zhang](#)
- Efficient Stagewise Pretraining via Progressive Subnetworks
 - Spotlight presentation at 15th Annual Machine Learning Symposium, [New York Academy of Sciences](#) (10/18/24)
- Trainable Transformer in Transformer
 - CSMA, Harvard (11/15/2023); hosted by [Prof. Michael Douglas](#)
 - Carnegie Mellon University (10/27/23); group meeting of [Prof. Andrej Risteski](#)
- Task-specific skill localization in Fine-tuned Language Models
 - DeepMind, London (4/12/23); hosted by [Marc’Aurelio Ranzato](#)
 - Department of Statistics, Oxford University (4/21/23); hosted by [Prof. Yee Whye Teh](#)
- Understanding Gradient Descent on Edge of Stability in Deep Learning
 - Google Research, Bangalore (8/28/22); hosted by [Praneeth Netrapalli](#)

PROFESSIONAL ROLES AND RESPONSIBILITIES

- **Workshop Organizer** of Methods and Opportunities at Small Scale ([MOSS](#)), at [ICML 2025](#)
- **Workshop Organizer** of Getting Started with Large Language Models with Princeton Language and Intelligence, at [Princeton Wintersession 2025](#)
- Reviewer in JMLR, COLT’20, ICLR(’21-), NeurIPS(’21-), ICML(’22-).
 - **Top 10% reviewer** in NeurIPS ’22, ICLR’24.
- Teaching Assistant at Princeton University
 - COS 445: Economics and Computing - Spring ’22
 - COS 324: Introduction to Machine Learning – Fall ’22
- **Seminar Organizer**, Princeton Alg-ML Lunch Seminar (2022-23).
- **Seminar Organizer**, Microsoft Research India Theory Lunch Seminar (2019-20).

ACADEMIC ACHIEVEMENTS

- **Apple AI/ML Ph.D. Scholar 2025-26**; Ph.D. funding scholarship awarded by Apple to support 20 students around the world.
- **Viterbi India program 2017**: Awarded to 20 students from India for funding their summer internship at University of Southern California, Los Angeles
- **G. Singhal Scholarship 2016-2017**, **J.C. Ghosh Memorial Endowment prize 2017**, **John Von Neuman Award 2017**, **R.M. Lalwani Award 2017** and **C. Devi Memorial prize 2017** for academic excellence
- **IIT-JEE Advanced 2014 : AIR 277** Among 1,50,000 students from across the country.
- **KVPY Fellow 2012** by the Department of Science and Technology, Government of India.