

# Abhishek Panigrahi

Graduate Student, Princeton University

🔗 <https://abhishekpanigrahi1996.github.io> ◊ @ [ap34@princeton.edu](mailto:ap34@princeton.edu)

## EDUCATION

---

- **Princeton University** Jan'21 – Present  
Ph.D. student in Computer Science  
Advisor: *Prof. Sanjeev Arora*
- **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

## WORK EXPERIENCE

---

- **Google New York** Jun'23 – Feb'24  
*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.
- **Electrical and Computer Engineering, University of Southern California** May'17 – July'17  
*Research Intern*
  - Advisor: *Dr. C.-C. Jay Kuo*.
  - Project - Mathematical model for gradient back propagation in batch normalized models.

## RESEARCH INTERESTS

---

I seek to solve problems that involve (a) analysis of existing Machine Learning algorithms and their interaction with Deep Learning models, and (b) proposing new algorithms involving large scale optimization, which have either theoretical or interpretability guarantees.

## CONFERENCE PUBLICATIONS

---

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

- **Trainable Transformer in Transformer** [\[arxiv\]](#)  
Abhishek Panigrahi, Sadhika Malladi, Mengzhou Xia, Sanjeev Arora  
In submission
- **Do Transformers Parse while Predicting the Masked Word?** [\[arxiv\]](#)  
Haoyu Zhao\*, Abhishek Panigrahi\*, Rong Ge, Sanjeev Arora  
In submission
- **Task-Specific Skill Localization in Fine-tuned Language Models** [\[PMLR \(coming soon\)\]](#)  
Abhishek Panigrahi\*, Nikunj Saunshi\*, Haoyu Zhao, Sanjeev Arora  
Accepted at *International Conference on Machine Learning (ICML 2023)* [\[arxiv\]](#)
- **On the SDEs and Scaling Rules for Adaptive Gradient Algorithms** [\[OpenReview\]](#)  
Sadhika Malladi\*, Kaifeng Lyu\*, Abhishek Panigrahi, Sanjeev Arora  
Accepted at *Neural Information Processing Systems (NeurIPS 2022)* [\[arxiv\]](#)

- **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)*  
Initial manuscript presented at TOPML workshop 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  
Accepted for an **Oral** (270/3000 submissions) in *Association for Computational Linguistics (ACL 2019)*.
- **DeepTagRec: A Content-cum-User Based Tag Recommendation Framework for Stack Overflow** [\[paper\]](#)  
Suman Kalyan Maity, **Abhishek Panigrahi**, Sayan Ghosh, Arundhati Banerjee, Pawan Goyal, Animesh Mukherjee  
In *European Conference on IR Research (ECIR 2019)*.
- **Book Reading Behavior on Goodreads Can Predict the Amazon Best Sellers** [\[paper\]](#)  
Suman Kalyan Maity, **Abhishek Panigrahi** and Animesh Mukherjee  
In *ACM International Conference on Social Networks Analysis and Mining (ASONAM 2017)*.

## WORKSHOPS

---

- **Non-Gaussianity of Stochastic Gradient Noise** [\[arxiv\]](#)  
**Abhishek Panigrahi**, Raghav Somani, Navin Goyal and Praneeth Netrapalli  
In *Science meets Engineering of Deep Learning workshop, NeurIPS 2019 (SEDL 2019)*.

## PREPRINTS

---

- **Analysis on gradient propagation in batch normalized residual networks** [\[arxiv\]](#)  
**Abhishek Panigrahi**, Yueru Chen, C.-C. Jay Kuo

## TALKS

---

- 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

---

- Reviewer in JMLR, COLT'20, ICLR('21-), NeurIPS('21-), ICML('22-).
  - Top (10%) reviewer in NeurIPS '22.
- Teaching Assistant at Princeton University
  - COS 445: Economics and Computing - Spring '22
  - COS 324: Introduction to Machine Learning – Fall '22
- Organizer, Princeton Alg-ML Lunch Seminar (2022-).
- Organizer, Microsoft Research India Theory Lunch Seminar (2019-20).

## ACADEMIC ACHIEVEMENTS

---

- **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.