Abhishek Panigrahi

Graduate Student, Princeton University
https://abhishekpanigrahi1996.github.io ◇ @ abhishekpanigrahi034@gmail.com

EDUCATION

- **Princeton University** Jan'21 – Present
  Ph.D. student in Computer Science
  Current Advisor: *Prof. Sanjeev Arora*
  Cum. GPA: 4.0/4.0
- **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

- **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.

CONFERENCE PUBLICATIONS

- **Learning and Generalization in RNNs** *[arxiv]*
  Abhishek Panigrahi, and Navin Goyal
  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** *[paper]*
  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

RESEARCH EXPERIENCE

- **Learning and Generalization in RNNs** Feb'20 – Dec'20
  Advisors: Dr. Navin Goyal, *Microsoft Research India*
  · Theoretically analyzed the learning capacity of overparametrized RNNs with SGD algorithm and improved upon the previous generalization results of RNNs.
  · Showed that overparametrized RNNs can learn concept classes consisting of one-hidden-layer neural networks that take the entire sequence of tokens as input, using SGD. By the universality theorem for one-hidden-layer networks, such RNNs can compute all continuous functions of the input sequence.
  · Illustrated the strength of our concept class on some regular languages like PARITY.
- **Non-Gaussianity of Stochastic Gradient Noise** July'19 – Dec'19
  Advisors: Dr. Navin Goyal *and* Dr. Praneeth Netrapalli, *Microsoft Research India*
  · Conducted an experimental study for analyzing difference in behavior of Stochastic Gradient Noise (SGN) vectors, at different training batch sizes in Stochastic Gradient Descent.
- **Effect of Activation Functions on the Training of Overparametrized Neural Nets** July'18 – May'19
  Advisor: Dr. Navin Goyal, *Microsoft Research India*
  · Analyzed activation functions' impact on training convergence rate of 2-layer overparameterized neural networks.
  · Provided worst case polynomial rate, in number of samples, for functions with jump discontinuity e.g. ReLU.
  · For smooth activations e.g. polynomials, swish and tanh, we showed that convergence rates can be exponentially small if the data dimension is small. Also showed that both larger data dimension and sufficient neural network depth lead to polynomial convergence rate (in number of sample).
- **Word2Sense: Sparse Interpretable Word Embeddings** July'18 - March'19
  Advisors: Dr. Harsha Vardhan Simhadri *and* Dr. Chiranjib Bhattacharyya, *Microsoft Research India*
  · Proposed an unsupervised method to generate sparse and interpretable single prototype word embeddings, that contains explicit information about all the senses of a word.
  · Proposed an optimization method, that uses sparsity of Word2Sense, to explicitly specify the meaning of a word in a context. Performs the best in word sense disambiguation task.
- **Mathematical model for gradient back propagation in batch normalized models** May'17 – Oct'17
  Advisor: Dr. C.-C. Jay Kuo , *University of Southern California*
  · Developed a complete mathematical model for modeling the gradient back propagation in batch normalized models and residual networks.
  · Showed that both residual branches and batch normalization in each residual block work together to maintain variance of gradients across a batch, stabilizing gradients across the network atleast in initial training phase.

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.

PROFESSIONAL RESPONSIBILITIES

- Reviewer in JMLR, COLT'20, ICLR'21, Neurips'21.
- Organizer, Microsoft Research India Theory Lunch Seminar (2019-20).