# Abhishek Panigrahi

Research Fellow, Microsoft Research India
https://abhishekpanigrahi1996.github.io ◇ @ abhishekpanigrahi034@gmail.com

## EDUCATION

- **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 – Present
  *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.

## PUBLICATIONS

- **How the choice of Activation affects training of Overparametrized Neural Nets**
  **Abhishek Panigrahi**, Abhishek Shetty and Navin Goyal *[arxiv]*
  In submission at *International Conference on Learning Representations (ICLR 2020)*.
- **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)*.
- **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)*.
- Analysis on gradient propagation in batch normalized residual networks *[arxiv]*
  **Abhishek Panigrahi**, Yueru Chen, C.-C. Jay Kuo
- 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)*.

## RESEARCH EXPERIENCE

- **Non-Gaussianity of Stochastic Gradient Noise** July'19 – Present
  *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.
  · Currently, trying to relate the difference in behavior of SGNs and generalization gap between small batch and large batch training.
  · Initial manuscript accepted in *SEDL workshop*, *NeurIPS 2019*.

- **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).
  · This work is under submission at *ICLR 2020*.

- **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.
  · The work got accepted for oral presentation at *ACL 2019*.

- **Regularization of GANs** Sept'17 – May'18
  *Bachelor Thesis, Advisor:* Dr. Pabitra Mitra*, IIT Kharagpur*
  · Worked on regularizing Wasserstein GANs and MMD GANs, to address disparity in the support of real data distribution and generated data samples in the form of heavy tails around real data clusters.
  · Approach was to boost the weights of generated points far from the real data manifold and construct a weighted MMD formulation using those weights. In low dimensional data with multiple independent clusters, MMD GANs output interconnected clusters, while weighted MMD GANs were (partially) effective in separating them.

- **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.

## RELEVANT COURSES

• Advanced Machine Learning • Advanced Linear Algebra • Foundations of cryptography • Operations Research • Probability and Statistics • Matrix Algebra • Deep Learning • Machine Learning + Practicum • Natural Language Processing + Practicum • Artificial Intelligence • Computer Graphics + Practicum • Image Processing* + Practicum • Discrete Structures • Algorithms I + Practicum • Algorithms II