# Raghav Somani

**Second year Ph.D. student at**
**Paul G. Allen School of Computer Science & Engineering**
**University of Washington**
*Advisor: **Prof. Sewoong Oh***

October 29, 2020
raghavs@cs.washington.edu
raghavsomani1995@gmail.com
raghavsomani.github.io

## Research Interests

I seek to solve theoretical machine learning problems using ideas from mathematical optimization, probability theory and statistics.

| | |
|---|---|
| **Research Interests** | Machine Learning, Large Scale Optimization, Probability theory |
| **Other Learning Interests** | Differential Geometry |

## Education

- **University of Washington**
  *Ph.D.* in Computer Science and Engineering     *Sept '19 - Present*
  *Advisor: Prof. Sewoong Oh*
  **GPA: 3.9/4**
- **Indian Institute of Technology Guwahati**
  *Bachelor of Technology* in Mathematics and Computing     *July '13 - June '17*
  **GPA: 9.10/10** (9.30/10 in major courses)

## Work Experience

- **University of Washington - Graduate Research Assistant**     *Sept '19 - Present*
  *Advisor: Prof. Sewoong Oh*
  Machine Learning Lab
- **Microsoft Research India - Research Fellow**     *July '17 - July '19*
  *Advisors: Dr. Praneeth Netrapalli & Dr. Prateek Jain*
  Project group - Provable Non-convex Optimization for Machine Learning Problems
- **Microsoft Research India - Research Intern**     *May '16 - July '16*
  *Advisor: Dr. Sreangsu Acharyya*
  Project - Recommendation systems
- **CAFRAL, Reserve Bank of India - Summer Research Intern**     *June '15 - July '15*
  *Advisor: Prof. Nagpurnanand R. Prabhala*
  Project - Modeling 'Economic Policy Uncertainty Index' for India

## Publications

1. **Robust Meta-learning for Mixed Linear Regression with Small Batches** [arXiv]
   W. Kong, **R. Somani**, S. M. Kakade, and S. Oh.
   *Advances in Neural Information Processing Systems (NeurIPS), December 2020.*

2. **Meta-learning for mixed linear regression** [paper]
   W. Kong, **R. Somani**, Z. Song, S. M. Kakade, and S. Oh.
   *International Conference on Machine Learning (ICML), July 2020.*

3. **Soft Threshold Weight Reparameterization for Learnable Sparsity**. [paper]
   A. Kusupati, V. Ramanujan[*], **R. Somani**[*], M. Wortsman[*], P. Jain, S. M. Kakade, and A. Farhadi.
   *International Conference on Machine Learning (ICML), July 2020.*

4. **Non-Gaussianity of Stochastic Gradient Noise** [arXiv]
   A. Panigrahi, **R. Somani**, N. Goyal, and P. Netrapalli.
   *SEDL* workshop, *Neural Information Processing Systems (NeurIPS), Vancouver, Canada, December 2019.*

5. **Support Recovery for Orthogonal Matching Pursuit: Upper and Lower bounds.** [paper]
   **R. Somani**[*], C. Gupta[*], P. Jain, and P. Netrapalli.
   *Advances in Neural Information Processing Systems (NeurIPS), Montréal, Canada, December 2018.* **Spotlight**
   Presentation (168/4856 submissions).

6. **Clustered Monotone Transforms for Rating Factorization** [paper]
   **R. Somani**[*], G. Hiranandani[*], O. Koyejo, and S. Acharyya.
   *ACM International Conference on Web Search and Data Mining (WSDM), Melbourne, Australia, February 2019.*

   * - equal contribution

---

## RESEARCH PROJECTS

**Robustness and Meta Learning for Mixed Linear Regression (MLR)** *Nov '19 - Present*
*Advisors: Prof. Sewoong Oh & Dr. Weihao Kong, University of Washington*

- Analyzed conditions under which abundant tasks with small data can compensate for lack of tasks with large data in the context of MLR. [arXiv]
- Tightened the conditions of the above exchange using the Sum-of-squares method. Worked on making the subspace estimation step robust for MLR. [arXiv]
- Currently working on Robust Subspace estimation for mixture distributions under the batch setting.
- ★ Works on Meta Learning and Robust Meta Learning with small batches got accepted at **ICML '20** and **NeurIPS '20** respectively.

**Implicit bias of Gradient Descent (GD) on Deep Generative CNNs** *Aug '19 - Nov '19*
*Advisor: Prof. Sewoong Oh, University of Washington*

- Understanding the dynamics and implicit bias of GD on a deep linear convolutional network under the least squares loss.

**Optimization and Generalization in Deep Neural Networks** *July '18 - Sept '19*
*Advisors: Dr. Prateek Jain, Dr. Praneeth Netrapalli & Dr. Navin Goyal, Microsoft Research* [Page 1, Page 2]

- Understanding the dependence of batch-size (stochasticity), over-parameterization, and optimization on the generalization properties of a variety of neural networks on real world data distributions for classification tasks.
- Analyzing the dependence of support separation, number of hidden neurons, ambient dimension of data distribution, number of training points on optimization and generalization of neural networks.
- ★ A work on distributional characterization of SGD got accepted at the workshop **SEDL '19** at **NeurIPS '19**. [arXiv]

**Sparse Regression and Optimal Bounds for Orthogonal Matching Pursuit (OMP)** *Sept '17 - June '18*
*Advisors: Dr. Prateek Jain & Dr. Praneeth Netrapalli, Microsoft Research* [Page]

- Analyzed Accelerated IHT, trying to strengthen Jain et al.'s results for better support expansion and generalization.
- Analyzed OMP for the Sparse Linear Regression problem under Restricted Strong Convexity (RSC) assumptions obtaining its support recovery and generalization guarantees. Also provide tight lower bounds for OMP. Our results are the first such matching upper and lower bounds (up to log factors) for *any* Sparse Regression algorithm under RSC assumption.
- ★ Accepted for a **Spotlight** paper presentation at **NeurIPS '18**. [paper, spotlight video]

**Clustered Monotone Transforms for Rating Factorization (CMTRF)** *May '16 - Aug '18*
*Advisors: Dr. Sreangsu Acharyya (MSR India) & Prof. Oluwasanmi Koyejo (UIUC)* [Page]

- Implemented and analyzed CMTRF for recommendation systems which performs regression under shared low-rank structure up to unknown monotonic transforms for a better fit. CMTRF recovers a unique solution under mild conditions and also outperforms other state-of-the-art baselines on 7 real-world and 2 synthetic datasets.
- ★ Accepted for an oral presentation at **WSDM '19**. [arXiv]

## Relevant Courses and Electives

| | |
|---|---|
| **Machine Learning** | Machine Learning, Learning from Data (Caltech MOOC), Statistical Learning (MOOC) |
| **Optimization** | Theory of Optimization, Variational Analysis*, Convex Optimization (Stanford MOOC) |
| **Probability** | Probability & Random Processes, Advanced Probability, Monte Carlo Simulations |
| **Statistics** | Advanced Statistical Algorithms, Statistical Methods and Time Series Analysis |
| **Computer Science** | Design and Analysis of Algorithms, Discrete Mathematics |
| **Mathematics** | Linear Algebra, Real & Complex Analysis, Stochastic Calculus, Modern Algebra, Matrix Computation, Scientific Computation, Symplectic Geometry (MOOC)* |

* - ongoing

## References

1. *Prof. Sewoong Oh*, Associate Professor, University of Washington
2. *Dr. Praneeth Netrapalli*, Senior Researcher, Microsoft Research Lab - India
3. *Dr. Prateek Jain*, Senior Principal Researcher, Microsoft Research Lab - India