Anshumali Shrivastava

350 Gates Hall, Cornell Univesity, Ithaca, NY 14853
anshu@cs.cornell.edu, www.cs.cornell.edu/~anshu, +1-607-793-6565

Interests

Large Scale Machine Learning, Randomized Algorithms for Big Data, Graph Mining, Information Retrieval, Hashing and Sketching.

Education

Cornell University, Ithaca, New York, USA (2010-Present)

Ph.D. student, Computer Science (Grad Minor: Applied Math)

- Advisor: Prof. Ping Li
- Thesis: Probabilistic Hashing Algorithms for Large Scale Search and Learning
- GPA: 4.0 (A+ = 4.3)
- Best Paper Award NIPS 2014
- Best Paper Award ASONAM 2014

Indian Institute of Technology (IIT), Kharagpur, West Bengal, India (2003-2008)

Integrated B.Sc.(H) and M.Sc., Mathematics and Computing

- Advisor: Prof. Somesh Kumar
- Thesis: Analysis of Email Traffic and Content Independent Spam Identification
- Institute Silver Medal for being Rank 1 in the program.

Professional Experience

Microsoft Research (MSR) Redmond, WA, USA (May 2014- August 2014)

Research Intern
Mentors: Dr. Misha Bilenko, Dr. Christian Konig, Vishal Joshi

- Learning with counts in presence of temporal drift.

FICO (Fair Issac) Core Research Team, Bangalore, India (2008-2010)

Analytic Software Scientist
Team: Dr. Stuart Crawford, Navin Doshi, Prasun Kumar.

- Segmentation tree induction and pruning (STIP) algorithm for fast automatic discovery of heterogeneous user segments

Teaching Experience

Cornell University, Ithaca, NY, USA (2010-present)

Teaching Assistant, Computer Science Department

- CS 4740: Introduction to Natural Language Processing, Spring 2011
- CS 4700: Foundations of Artificial Intelligence, Fall 2010
- CS 4701: Practicum in Artificial Intelligence, Fall 2010 (Head Teaching Assistant)

Awards and Honors

- Best Paper Award in Neural Information Processing System (NIPS) 2014 Conference.
- Best Paper Award in IEEE/ACM International Conference on Advances in Social Network Analysis and Mining (ASONAM 2014)
- FICO Spot Award, 2009 and 2010.
- IIT Kharagpur Silver Medal, 2008.
- J.C. Gosh Memorial Endowment Prize from IIT Kharagpur, 2007.
- DMMS Scholarship from IIT Kharagpur, 2003.
- Placed 5^{th} in National Science Olympiad (NSO), 2002
- Regional Mathematics Olympiad (RMO) (Maharashtra Region), 2000 and 2001.
- 1^{st} position in Maharashtra Talent Search Examination (MTSE), 1999.

Published Papers

Shrivastava, A. and Li, P. *AAsymmetric Minwise Hashing for Indexing Binary Inner Products and Set Containment*
WWW 2015

Shrivastava, A. and Li, P. *Asymmetric LSH (ALSH) for Sublinear Time Maximum Inner Product Search (MIPS).*
NIPS 2014 (**Best Paper Award**).

Shrivastava, A. and Li, P. *A New Space for Comparing Graphs.*
IEEE/ACM ASONAM 2014 (**Best Paper Award**).

Shrivastava, A. and Li, P. *Improved Densification of One Permutation Hashing.*
UAI 2014.

Shrivastava, A. and Li, P. *Densifying One Permutation Hashing via Rotation for Fast Near Neighbor Search.*
ICML 2014.

Li, P., Mitzenmacher, M., and Shrivastava, A. *Codings for Random Projections.*
ICML 2014.

Shrivastava, A. and Li, P. *In Defense of Minhash over Simhash.*
AISTATS 2014.

Shrivastava, A. and Li, P. *Beyond Pairwise: Provably Fast Algorithms for Approximate k-Way Similarity Search.*
NIPS 2013.

Shrivastava, A. and Li, P. *Fast Near Neighbor Search in High-Dimensional Binary Data.*
ECML/PKDD 2012. (**Among top papers invited for journal submission**)

Sun, X., Shrivastava, A. and Li, P. *Fast multi-task learning for query spelling correction.*
CIKM 2012.

Li, P., Shrivastava, A. and Konig, A. C. *GPU-based minwise hashing.*
WWW (Companion Volume) 2012.

Sun, X., Shrivastava, A. and Li, P. *Query spelling correction using multi-task learning.*
WWW (Companion Volume) 2012.

Li, P., Shrivastava, A., Moore, J. L. and Konig, A. C. *Hashing Algorithms for Large-Scale Learning.*
NIPS 2011.

Work Under Submission

Shrivastava, A. and Li, P. *Improved Asymmetric Locality Sensitive Hashing (ALSH) for Maximum Inner Product Search (MIPS)*

In Preparation

Fast Hashing for Conflict Casualties in Syria

Adaptive Sketches for Summarizing Temporal Data Streams

Additional Technical Reports

Shrivastava, A. and Li, P. *Graph Kernels via Functional Embedding.* 2014

Li, P., Mitzenmacher, M. and Shrivastava, A. *Coding for Random Projections and Approximate Near Neighbor Search.* 2014

Li, P., Shrivastava, A. and Konig, A. C. *b-Bit Minwise Hashing in Practice: Large-Scale Batch and Online Learning and Using GPUs for Fast Preprocessing with Simple Hash Functions.* 2012

Li, P., Shrivastava, A. and Konig, A. C. *Training Logistic Regression and SVM on 200GB Data Using b-Bit Minwise Hashing and Comparisons with Vowpal Wabbit (VW).* 2011

References

1. **Prof. Ping Li**
   Email: pingli@stat.rutgers.edu
   Affiliation: Rutgers University

2. **Dr. Mikhail Bilenko**
   Email: mbilenko@microsoft.com
   Affiliation: Microsoft Research

3. **Dr. Christian Konig**
   Email: chrisko@microsoft.com
   Affiliation: Microsoft Research

4. **Prof. Thorsten Joachims**
   Email: tj@cs.cornell.edu
   Affiliation: Cornell University

Other Activities

Admin for Machine learning discussion group, Cornell, 2012-present
Mentor for Fresher's Training at FICO Research Bangalore, 2009
President of Mathematics Colloquium, IIT Kharagpur, 2007-08
Captain of R. P. Hall of Residence Chess Team, IIT Kharagpur, 2005-08
Award of Excellence and Hall Color in Chess, IIT Kharapur, 2006