

Sonal Gupta

PhD candidate, Department of Computer Science, Stanford University

sonal@cs.stanford.edu

<http://cs.stanford.edu/people/sonal>, <https://github.com/sonalgupta>

Education

Stanford University

PhD Candidate in Computer Science

Stanford, CA
2009 – August 2015 (expected)

- Thesis: Distantly Supervised Information Extraction using Bootstrapped Patterns
- Advisor: Christopher D. Manning GPA 4.0/4.0

University of Texas at Austin

Masters in Computer Science (with thesis)

Austin, Texas
2007 – 2009

- Thesis: Activity Retrieval in Closed Captioned Videos
- Advisor: Raymond J. Mooney GPA 4.0/4.0

Indian Institute of Technology (IIT), Roorkee

Bachelor of Technology in Computer Science

Roorkee, India
2003 – 2007

- GPA: 9.0 / 10.0, Dept. Rank: 2

Research Interests

Natural language processing, information extraction, unsupervised and bootstrapped learning, information retrieval, machine learning, and multi-modal learning.

Relevant Experience

Data Science Intern, PayPal Inc.

Modeled dominant complaints and compliments in PayPal's customer feedback data by learning topic models.

Summer 2012

Research Intern, Microsoft Research

Proposed and implemented an online learning algorithm, Amnesiac Averaged Perceptron, to learn arbitrary similarity functions as features using clickthrough logs. Predicted new keywords similar to keywords provided by advertisers.

Summer 2008

Software Intern, Google R&D Center, Bangalore

Implemented boosting machine learning technique for feature selection in the MapReduce framework, using sequential and parallel update, for classifying shopping web pages.

Summer 2006

Technical Coordinator, Information Management Group, IIT Roorkee

Actively involved in developing and maintaining web applications for the university. Led teams to create a social network for the alumni, and a platform for students to create their research profiles.

2004-2007

Journal and Conference Publications

S. Gupta and C. D. Manning. *Distributed Representations of Words to Guide Bootstrapped Entity Classifiers*. In Conference of the North American Chapter of the Association for Computational Linguistics (NAACL) 2015.

- D. MacLean, S. Gupta, Anna Lembke, C. D. Manning, and J. Heer. *Forum77: An Analysis of an Online Health Forum Dedicated to Addiction Recovery*. Computer-Supported Cooperative Work and Social Computing (CSCW) 2015. (Best Paper Honorable Mention)
- S. Gupta and C. D. Manning. *Improved Pattern Learning for Bootstrapped Entity Extraction*. In Conference on Computational Natural Language Learning. (CoNLL) 2014.
- S. Gupta, D. MacLean, J. Heer, C. D. Manning, *Induced Lexico-Syntactic Patterns Improve Information Extraction from Online Medical Forums*. Journal of the American Medical Informatics Association (JAMIA) 2014.
- J. Chuang, S. Gupta, C. D. Manning, J. Heer, *Topic Model Diagnostics: Assessing Domain Relevance via Topical Alignment*. In International Conference on Machine Learning (ICML) 2013
- M. Surdeanu, S. Gupta, J. Bauer, D. McClosky, A. Chang, V. Spitzkovsky, and C. D. Manning, *Stanford's Distantly-Supervised Slot-Filling System*. In Text Analysis Conference (TAC) 2011
- S. Gupta, C. D. Manning, *Analyzing the Dynamics of Research by Extracting Key Aspects of Scientific Papers*, In International Joint Conference on Natural Language Processing (IJCNLP), 2011
- S. Gupta and R. Mooney, *Using Closed Captions as Supervision for Video Activity Recognition*, In the AAAI Conference on Artificial Intelligence (AAAI), 2010
- S. Gupta, M. Bilenko and M. Richardson, *Catching the Drift: Learning Broad Matches from Clickthrough Data*, In Knowledge Discovery and Data Mining (KDD), 2009
- S. Gupta, J. Kim, K. Grauman and R. Mooney, *Watch, Listen & Learn: Co-training on Captioned Images and Videos*, In the European Conference on Machine Learning (ECML PKDD), 2008

Awards & Honors

School of Engineering fellowship at Stanford University	2009 – 2010
Finalist for Google Anita Borg Scholarship	2009
Institute Silver Medal for the Best B. Tech. project (thesis) in CS, IIT Roorkee	2007
Next Best Girl Award, IIT Roorkee	2007
Certificate of Merit in Summer Undergraduate Research Award, IIT Roorkee	2005

Teaching Experience

Head Course Assistant, Information Retrieval and Web Search	Spring 2013
Prof. C. D. Manning, P. Raghavan, and P. Nayak at Stanford University	
Course Assistant, Machine Learning	Fall 2012
Prof. Andrew Ng at Stanford University	

Relevant Graduate Level Courses

Social and Information Network Analysis, Probabilistic Graphical Models, Convex Optimization, Machine Learning, Computer Vision, Natural Language Processing, Numerical Linear Algebra, Modern Statistical Methods

Activities

Asha for Education, Stanford	2010 – current
Financial officer (2013-current) of the student group that raises funds by organizing an yearly Holi event (around 10k attendance), and support around 15 projects to educate underprivileged children in India.	
Member, Dramatics Section, IIT Roorkee	2003 – 2004