Karthik Raman

CONTACT INFORMATION	349 Gates Hall Department of Computer Science Cornell University Ithaca, NY 14853 USA	*Mobile:* +1-607-592-1561 *E-mail:* karthik@cs.cornell.edu *Website:* www.cs.cornell.edu/~karthik

RESEARCH INTERESTS

Machine Learning, Web Search and Information Retrieval, Education at Scale, Data Mining

EDUCATION

Cornell University   August 2010 - July 2015 (*expected*)

Ph.D. in Computer Science

- Graduate Minor in Applied Math
- Advisor: Thorsten Joachims
- Area of Study: Machine Learning, Web Search
- GPA (after 9 semesters): 4.14/4.30
- Committee: Thorsten Joachims, Johannes Gehrke, Robert Kleinberg, Paul Bennett

M.S. in Computer Science

- Awarded in August 2013.

Indian Institute of Technology, Bombay   August 2006 - May 2010

B.Tech., Computer Science & Engineering, May 2010

- Thesis Topic: *Improving Pseudo-Relevance Feedback : Multilingual Feedback and Irrelevance-Based Feedback*
- Advisor: Pushpak Bhattacharyya
- Co-Advisor: Raghavendra Udupa, MSR-India
- Area of Study: Information Retrieval, Natural Language Processing
- Ranked second with CPI of 9.54/10

AWARDS & ACHIEVEMENTS

Cornell

- SIGIR 2013 Best Student Paper Award.
- Google PhD Fellowship in Search and Information Retrieval, 2013.
- Yahoo! Key Scientific Challenge Award, 2011.
- Cornell-Olin Fellowship 2010-11.
- WSDM 2015 Outstanding Reviewer Award.
- Cornell Teaching Assistant Excellence Award 2011, 2013.
- Travel Awards: *Cornell Grant* (ECML 2013, KDD 2012), *Microsoft* (ICML 2013), *SIGIR Grant* (SIGIR 2010, 2013), *ACL* (ACL 2010), *IITB Grant* (ACL 2010, SIGIR 2010).

Undergraduate and before

- All India Rank 34 (out of 300,000 students) in the IIT-JEE Examination (2006).
- All India Rank 4 (out of 500,000 students) in the AIEEE Examination (2006).
- Awarded CBSE Merit Scholarship (2006).
- Awarded National Talent Search Scholarship (2004).
- Awarded Maharashtra Talent Search Competition (2003).
- Awarded Gold Medal for Excellence in Mathematics in the University of New South Wales Examination (2001).

Actuarial Science

- Completed 2 core technical exams of the Actuarial Society of India.

JOURNAL PUBLICATIONS	Raman, K., Bennett, P.N. and Collins-Thompson, K. *Understanding Intrinsic Diversity in Web Search: Improving Whole-Session Relevance.* In ACM Transactions on Information Systems (*TOIS*) 32, 4, Article 20 (October 2014), 45 pages.
CONFERENCE PUBLICATIONS	Raman, K. and Joachims. T. *Bayesian Ordinal Peer Grading.* In *Learning@Scale 2015*, March 2015, Vancouver, BC, Canada.
	Raman, K. and Joachims. T. *Methods for Ordinal Peer Grading.* In *KDD 2014*, August 2014, New York, NY, USA.
	Raman, K. and Joachims. T. *Learning Socially Optimal Information Systems from Egoistic Users.* In *ECML 2013*, September 2013, Prague, Czech Republic.
	Raman, K., Swaminathan, A., Gehrke, J. and Joachims. T. *Beyond Myopic Inference in Big Data Pipelines.* In *KDD 2013*, August 2013, Chicago, IL, USA.
	Raman, K., Bennett, P.N. and Collins-Thompson, K. *Toward Whole-Session Relevance: Exploring Intrinsic Diversity in Web Search.* In *SIGIR 2013*, July 2013, Dublin, Ireland. (**Best Student Paper**)
	Raman, K., Joachims. T., Shivaswamy. P., and Schnabel, T. *Stable Coactive Learning via Perturbation.* In *ICML 2013*, June 2013, Atlanta, GA, USA.
	Raman, K., Svore, K., Gilad-Bachrach, R and, Burges, C. *Correctable Learning: Learning from Past Mistakes.* In *CIKM 2012*, October 2012, Maui, HI, USA.
	Raman, K., Shivaswamy, P. and Joachims, T. *Online Learning to Diversify from Implicit Feedback.* In *KDD 2012*, August 2012, Beijing, China.
	Raman, K., Joachims. T. and Shivaswamy, P. *Structured Learning of Two-Level Dynamic Rankings.* In *CIKM 2011*, October 2011, Glasgow, Scotland.
	Chinnakotla, M., Raman, K., and Bhattacharyya. P. *Multilingual PRF: English Lends a Helping Hand.* In *SIGIR 2010*, July 2010, Geneva, Switzerland.
	Chinnakotla, M., Raman, K., and Bhattacharyya. P. *Multilingual Relevance Feedback: Performance Study of Assisting Languages.* In *ACL 2010*, July 2010, Uppsala, Sweden.
	Raman, K., Udupa. R., Bhattacharyya. P, and Bhole, A. *On Improving Pseudo-Relevance Feedback Using Pseudo-Irrelevant Documents.* In *ECIR 2010*, March 2010, Edinburgh, Scotland.
WORKSHOP PUBLICATIONS	Raman, K.and Joachims, T. *Bayesian Ordinal Peer Grading.* In: NIPS-HPML 2014 - Human Propelled Machine Learning, December 2014, Montreal, QC, Canada.
	Raman, K., Shivaswamy, P. and Joachims, T. *Structured Prediction via Coactive Learning.* In: IUI-IMLW 2013 - Interactive Machine Learning Workshop, March 2013, Santa Monica, CA, USA.
	Raman, K., Shivaswamy, P. and Joachims, T. *Learning to Diversify from Implicit Feedback.* In: WSDM-DDR-2012 - 2nd Workshop on Diversity in Document Retrieval, February 2012, Seattle, WA, USA.
OTHER PUBLICATIONS	Raman, K., Joachims. T. and Shivaswamy, P. *Structured Learning of Two-Level Dynamic Rankings.*, ArXiv, August 2011.
	Raman, K.. *Improving Pseudo-Relevance Feedback.* Bachelor's thesis, Indian Institute of Technology - Bombay, May 2010.

PAPERS IN SUBMISSION	Raman, K., Dahlquist, J., Dalton, J., Gabriloivch, E., Murphy, K. and Zhang, W., *RAKE: Robust Automatic Knowledge Extraction from Semi-Structured Web Pages.* In Submission to *WWW-2015.*

PATENTS PUBLISHED

- *Prediction and Information Retrieval for Intrinsically Diverse Sessions*
  Inventors: Karthik Raman, Paul N. Bennett and Kevyn B Collins-Thompson
  Filed: February, 2013. Published: August, 2014
  Publication number: US-2014-0244610-A1.

INVITED TALKS

- *"By the User, For the User, With the Learning System": Learning From User Interactions*
  - NIPS 2014 Workshop on Personalization: Methods and Applications, Montreal, QC, Canada, December 2014
  - Los Alamos National Laboratories, Los Alamos, NM, March 2014
  - Cornell University, Ithaca, NY, March 2014

- *Ordinal Peer Grading*
  - Cornell University, Ithaca, NY, May 2014

- *Learning to Diversify From Implicit Feedback*
  - Microsoft Bing, Redmond, WA, June 2012
  - Cornell University, Ithaca, NY, September 2012

- *Structured Learning of Diverse Rankings*
  - Yahoo! Research, Santa Clara, CA, August 2011
  - Cornell University, Ithaca, NY, October 2011

TEACHING EXPERIENCE

- *Teaching Assistant* for CS 4780/5780: Machine Learning   **Fall 2014**
  - Instructor: Thorsten Joachims
  - Responsible for final projects and peer grading.

- *Teaching Assistant* for CS 4780/5780: Machine Learning   **Fall 2013**
  - Instructor: Thorsten Joachims
  - Responsible for assignments and final projects.
  - Awarded Teaching Assistant Excellence award for performance.

- Taught lecture of Machine Learning (CS 4780/5780)   **Fall 2012**
  - Lectured on overfitting, model validation and hypothesis testing.

- Head *Teaching Assistant* for CS 4780/5780: Machine Learning   **Fall 2011**
  - Instructor: Thorsten Joachims
  - Responsible for assignments and final projects.
  - Awarded Teaching Assistant Excellence award for performance.

- *Grader* for CS 6780: Advanced Machine Learning   **Fall 2010**
  - Instructor: Ashutosh Saxena

Students Advised	• *Ashudeep Singh* (IIT-Kanpur)	**Fall 2014**

  • *Ashudeep Singh* (IIT-Kanpur)   **Fall 2014**
   • Co-advised with: Thorsten Joachims, Adith Swaminathan.
   • Project: *Embedding user interactions while accounting for presentation bias.*

  • *Ziyu Fan* (Cornell)   **Spring 2014**
   • Co-advised with: Thorsten Joachims.
   • Project: *Embedding arXiv user sessions.*

  • *Akhilesh Potti* (Cornell)   **Spring 2014**
   • Co-advised with: Thorsten Joachims.
   • Project: *Embedding arXiv document sequences as playlists.*

  • *Ziyu Fan, Akhilesh Potti* (Cornell)   **Fall 2013**
   • Co-advised with: Thorsten Joachims.
   • Project: *Analyzing co-accessed documents on arXiv.*

  • *Tobias Schnabel* (Cornell)   **Spring 2012**
   • Co-advised with: Thorsten Joachims, Pannaga Shivaswamy.
   • Project: *Coactive learning for arXiv text search.*

  • *Diego Accame* (Cornell)   **Spring 2012**
   • Co-advised with: Thorsten Joachims, Pannaga Shivaswamy.
   • Project: *Diversifying arXiv text search results.*

Representative Graduate Level Course-work

Cornell:
• Advanced Topics in Machine Learning
• Optimal Learning
• Advanced Machine Learning
• Advanced Language Technologies
• NLP and Social Interaction
• Matrix Computations
• Algorithmic Game Theory
• Structure of Information Networks
• Advanced Database Systems
• Analysis of Algorithms
• Advanced Programming Languages

IIT-Bombay:
• Advanced Machine Learning
• Foundations of Machine Learning
• Information Retrieval and Mining for Hypertext and the Web
• Organization of Web Information
• Machine Learning: Theory and Methods
• Statistical Relational Learning
• Topics in AI Programming/NLP
• Program Analysis

Coursera:
• Big Data in Education

Reviewing

Program Committee
• *WWW* 2014, 2015
• *WSDM* 2015
• *ICML* 2014, 2015
• *KDD* 2015
• *SIGIR* 2014
• *CIKM* 2013
• *ECML* 2013, 2014
• *MOD* 2015
• *CaRR* 2013
• *IKDD* 2014

Reviewer
• *ICML* 2013

- *KDD* 2014
- *ICML* 2013
- *CIKM* 2012
- *SIGIR* 2012
- *AAAI* 2012
- *IJCNLP* 2011

Journal Reviewer

- JMLR (Journal of Machine Learning Research)
- MLJ (Machine Learning Journal)
- ACM TOIS (Transactions on Information Systems)

PROFESSIONAL
EXPERIENCE

Google, Mountain View USA

Intern in the Knowledge Vault Group.   **Summer 2014**

- Worked with Evgeniy Gabriloivch, Kevin Murphy, Jeff Dalton and Wei Zhang.
- Created and deployed the RAKE system for automatic extraction of factual knowledge from semi-structured webpages.

Microsoft Research, Redmond, Seattle USA

Intern in the CLUES Group.   **Summer 2012**

- Worked with Paul Bennett, Kevyn Collins-Thompson and Susan Dumais.
- Worked on intrinsic diversity in web search.

Intern in the Machine Learning and Intelligence Group.   **Summer 2011**

- Worked with Krysta Svore, Ran Gilad-Bachrach and Chris Burges.
- Worked on correctable learning for web search.

Microsoft Research-India, Bangalore, India

Intern in the Multi-Linguistic System Group.   **Summer 2010**

- Worked with Raghavendra Udupa.
- Studied applications of Non-Negative Matrix Factorization and Document-Specific Topic Models.
- Explored different models for Pseudo-Relevance Feedback.

Mulval Technologies Inc., USA

Part-Time Software Engineer   **January-May 2008**

- Involved in development of network security and vulnerability analysis tools.
- Primary developer for the UNIX platform tools.

SERVICE AND
PARTICIPATION

- Lead the Machine Learning Discussion group at Cornell (from 2010).
- Active member of the NLP Discussion group at Cornell (from 2010) and IITB (2009).
- Participant at Machine Learning Summer School (2013) at MPI, Tubingen.
- Participant of Yahoo! Machine Learning School (2010) at Bangalore.
- Served in the National Cadet Corps in India from 2006-2007.
- Volunteered for Asha (a non-profit dedicated to basic educational causes in India).

SYSTEMS AND
SOFTWARE

Public Systems

- Launched and manage the peer grading website `peergrading.org`.
- Manage the experimental text search service for arXiv at `search.arxiv.org`.
- Developed algorithms for scientific article recommendation service for arXiv at `my.arxiv.org`.

Software and Toolkits
- Peer Grading Toolkit: Toolkit for performing ordinal peer grading.
- OL-Diversity: Toolkit with different algorithms for online learning of diverse sets and rankings via implicit feedback.
- SVM-Dyn: A software used to learn dynamic rankings from training data.
- HotelReview-Scraper: Tool for scraping hotel reviews from Tridavisor/Orbitz.
- RateMyProf-Scraper: Tool for scraping reviews from RateMyProf.

PROGRAMMING SKILLS

Programming and Scripting Languages:
- C, C++, C#, Java, Python, OCaML, Scheme, Matlab
- JavaScript, PHP, AWK

Machine Learning, Information Retrieval & NLP Software:
- SVM-Light/Struct and others, LibLinear, Weka, Lucene, Lemur, NLTK

Productivity Applications:
- TeX (LaTeX, BibTeX), Vim, Microsoft Office suite

Operating Systems:
- Microsoft Windows family, Linux and other UNIX variants

REFERENCES AVAILABLE TO CONTACT

Thorsten Joachims (E-mail: tj@cs.cornell.edu)
- Professor, Department of Computer Science, Cornell University
- ⋄ 418 Gates Hall, Ithaca, NY 14853-7501
- ⋆ *Dr. Joachims is my graduate advisor.*

Johannes Gehrke (E-mail: johannes@cs.cornell.edu)
- Professor, Department of Computer Science, Cornell University
- ⋄ 437 Gates Hall, Ithaca, NY 14853-7501

Paul Bennett (E-mail: paul.n.bennett@microsoft.com)
- Senior Researcher, Microsoft Research
- ⋄ One Microsoft Way, Redmond WA 98052-6399

Evgeniy Gabrilovich (E-mail: gabr@google.com)
- Senior Staff Research Scientist, Google
- ⋄ 1600 Amphitheatre Parkway, Mountain View, CA 94043