

Ramnath Balasubramanyan

1620 Beechwood Blvd #8,
Pittsburgh PA 15217
Phone: (650) 796-4962
<http://www.cs.cmu.edu/~rbalasub>
rbalasub@cs.cmu.edu

RESEARCH INTERESTS: statistical machine learning, natural language processing, text mining.

EDUCATION

- **PhD candidate at the Language Technologies Institute**, Aug '09 – present.
School of Computer Science, Carnegie Mellon University, Pittsburgh, PA.
Adviser: Prof. William Cohen.

- **M.S. in Natural Language Processing (Computer Science)**, Aug '07 – Aug '09
GPA: 3.83
Relevant Coursework: Machine Learning, Numerical Optimization, Active Learning seminar, Language & Statistics – I, Language & Statistics – II, Algorithms for NLP, Lab in NLP, Grammar Formalisms

- **Bachelor of Engineering, Computer Science and Engineering**, Jun '99 – Jun '03.
Vishveswaraiah Technological University, PESIT, Bangalore, India.

- **Other Coursework**
 - **Text Retrieval and Web Search**, Aug '06 – Dec '06
Non Degree Option program, Stanford University, Stanford, CA
 - **Bayesian Statistics**
Conducted by Prof. David Draper (Univ. California at Santa Cruz) at Yahoo! Inc.

PUBLICATIONS

- From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series - Brendan O'Connor, Ramnath Balasubramanyan, Bryan R. Routledge, and Noah A. Smith accepted to ICWSM 2010: International AAAI Conference on Weblogs and Social Media, Washington DC, USA.
- From Episodes to Sagas: Understanding the News by Identifying Temporally Related Story Sequences - Ramnath Balasubramanyan, Frank Lin, William W. Cohen, Matthew Hurst, Noah A. Smith accepted to ICWSM-2009, International AAAI Conference on Weblogs and Social Media, San Jose, USA.
- CutOnce - Recipient Recommendation and Leak Detection in Action – Ramnath Balasubramanyan, Vitor Carvalho and William W. Cohen in AAAI-2008, Workshop on Enhanced Messaging, Chicago, USA.
- Activity-centered Search in Email - Einat Minkov, Ramnath Balasubramanyan William W. Cohen (2008) in CEAS 2008, Conference on Email and Anti-Spam, Seattle, USA
- Document Preprocessing for Naïve Bayes Classification and Clustering with Mixture of Multinomials - Dmitry Pavlov, Ramnath Balasubramanyan, Byron Dom, Shyam Kapur, Jignashu Parikh in KDD 2004, Seattle, USA.

PATENTS

- Attribute extraction using limited training data – United States Patent 7689527
- Assigning into one set of categories information that has been assigned to other sets of categories, B. Dom, H. Han, R. Balasubramanyan, D. Pavlov – Pending Docket No. 50269-0793

- Automatic product categorization, B. Dom, A. Goyal, R. Balasubramanian, D. Pavlov, B. Suresh - Pending Docket No. 50269-0579

RESEARCH PROJECTS

Querendipity project at CMU under Prof. Cohen, Sep '09 - now

- Developing stochastic block models and topic models to text from the biological domain and investigating the advantages of joint modeling.

News Sagas project at CMU under Prof. Cohen (Microsoft Live Labs), Sep '08 to May '09

- Developed topic models to cluster atomic news stories into “sagas” using parametric clustering models that model time sensitive text and nonparametric Bayesian models based on hierarchical Dirichlet processes.

RADAR project at CMU under Prof. Cohen(funded by DARPA), Aug '07 to Aug '08

- Message-Task Linking using lazy graph walks with random restarts on a specialized graph constructed for the task.
- Developed a Mozilla Thunderbird extension for email recipient recommendation and leak detection.

COURSE PROJECTS

- Using Latent Dirichlet Allocation and Probabilistic Latent Semantic Indexing to improve recipient recommendation in email systems – as a part of Machine Learning, Fall '08
- Comparing Exponentiated Gradient and Sequential Minimal Optimization to train SVMs – as a part of Optimization, Spring '08

WORK EXPERIENCE

Software Engineering Intern, Google Inc., Pittsburgh, PA, USA, Jun '09 – Aug '09

Product Search team

- Worked on matching product offers to product catalog entries.

Senior Research Engineer, Yahoo! Inc, Santa Clara, CA, USA, Jun '05 – Jun '07

Content Analysis Group headed by Byron Dom and

Natural Language Processing group headed by Hadar Shemtov

- Built a large scale text categorization application to categorize product offers for Yahoo! Shopping. Product categorization enabled key functionalities such as guided navigation and merchant cost per click pricing.
- Developed an information extraction system to extract attributes from merchant product offers using Hidden Markov Models and Conditional Random Fields for Yahoo! Shopping. Attribute extraction allows users to narrow searches for products based on attribute values such as brand, size etc.
- Extended the automatic question categorization system for Yahoo! Answers.

Software Engineer, Yahoo! Inc, Bangalore, India, Jun '03 – May '05

- Core member of a team that developed a document categorization platform.
- Won Yahoo! Ratna (Best Employee) award for the year 2004

OTHER

- Webmaster for ICWSM 2009, 2010.

References available upon request.