Abhimanyu Lad

Language Technologies Institute,
School of Computer Science,
Carnegie Mellon University,
Pittsburgh, PA 15213

Phone: (412) 596-1580
Email: alad@cs.cmu.edu
Web: http://www.cs.cmu.edu/~alad

Areas of Interest

Information retrieval and statistical machine learning
Novelty and diversity-based retrieval over web documents and news streams, adaptive filtering, retrieval system evaluation, probabilistic topic modeling, active learning, multi-task learning

Education

Pursuing Ph.D. in Language and Information Technologies **2005–Present**
Carnegie Mellon University, Pittsburgh, USA (Expected graduation: Jan. 2011)
Advisor: Yiming Yang
Thesis Topic: A Framework for Evaluation and Optimization of Relevance and Novelty-based Retrieval
Current GPA: 3.99 (Yahoo! Ph.D. Fellowship recipient for 2007–2009)

B.Tech. in Information Technology **2001–2005**
Indian Institute of Information Technology, Allahabad, India
GPA: 3.91 (University Gold Medalist for 2005)

Select Publications

- A. Lad, Y. Yang. *Learning to Rank Relevant and Novel Documents through User Feedback.* In Proceedings of the 19th ACM International Conference on Information and Knowledge Management (CIKM 2010).
- A. Lad, Y. Yang. *Active Ordering of Interactive Prediction Tasks.* In Proceedings of the 10th SIAM International Conference on Data Mining (SDM 2010).
- Y. Yang, A. Lad. *Utility-Based Information Distillation.* Book Chapter in Text Mining: Classification, Clustering, and Applications (Chapman & Hall 2009).
- Y. Yang, A. Lad. *Modeling Expected Utility of Multi-session Information Distillation.* In Proceedings of the 2nd International Conference on the Theory of Information Retrieval (ICTIR 2009).
- A. Lad, Y. Yang, R. Ghani, B. Kisiel. *Toward Optimal Ordering of Prediction Tasks.* In Proceedings of the 9th SIAM International Conference on Data Mining (SDM 2009).
- K. Salomatin, Y. Yang, A. Lad. *Multi-field Correlated Topic Modeling.* In Proceedings of the 9th SIAM International Conference on Data Mining (SDM 2009).
- Y. Yang, A. Lad, H. Shu, B. Kisiel, C. Cumby, R. Ghani, K. Probst. *Graph Structure Learning for Task Ordering.* In Proceedings of the 11th International Conference on Enterprise Information Systems (ICEIS 2009).
- A. Lad, Y. Yang. *Generalizing from Relevance Feedback using Named Entity Wildcards.* In Proceedings of the 16th ACM International Conference on Information and Knowledge Management (CIKM 2007).
- Y. Yang, A. Lad, N. Lao, A. Harpale, B. Kisiel, M. Rogati. *Utility-Based Information Distillation over Temporally Sequenced Documents.* In Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2007).

(Full list and papers available at http://www.cs.cmu.edu/~alad)

Professional Experience and Volunteer Services

Teaching Assistant: Graduate course in Information Retrieval **Spring 2009**
Teaching Assistant: Undergraduate course in Information Retrieval **Fall 2008**
Designed, implemented, and graded homework assignments. Created a script for automatically and efficiently detecting (word-by-word) plagiarism in students' submissions.

Summer Intern at Yahoo! Inc. **Summer 2008**
Worked with Jan Pedersen on analysis and comparison of various machine learning approaches to ranking. Created a tool for generating artificial ranking functions with varying model complexities and noise levels, which were then used to assess the effectiveness of different machine learning algorithms for ranking.

Reviewer for Research Conferences and Journals **2005–2010**
ACM Conference on Research and Development in Information Retrieval (SIGIR 2006, 2008, 2009, 2010), International Conference on Machine Learning (ICML 2008), ACM Transactions on Information Systems (TOIS 2006), SIAM Conference on Data Mining (SDM 2009).

Webmaster of College Website **2003–2004**
Led a team of five students to develop from scratch and maintain the college website at IIIT Allahabad. Developed a content management system in PHP/MySQL, allowing different levels of access to various administrators, and enabling students to store and modify their user profiles.

Select Research Projects

Evaluation and Optimization of Novelty and Diversity-based Retrieval Systems
Developed a unified framework for evaluation and optimization of retrieval systems with respect to relevance, novelty, and diversity of information. The framework incorporates a probabilistic model of users' browsing behavior and provides a flexible approach for modeling different tolerances towards redundancy of information as well as different levels of persistence.

Optimal Ordering of Prediction Tasks in a Multi-step Interactive Process
Explored the problem of ordering multiple interrelated prediction tasks (classification and regression) in an interactive setting so that user feedback is optimally utilized for maximizing prediction performance. The proposed approach is analogous to applying multi-task active learning in an online setting.

Utility-based Information Distillation
Designed and implemented a retrieval system that effectively combines adaptive filtering, passage retrieval, and novelty detection for information distillation over news streams. Proposed a new evaluation scheme and created a reusable dataset for automatically measuring the expected utility of such a distillation system.

Computer Skills

Languages: Python, C, C++, R, Java, Javascript, PHP, SQL
Tools and Frameworks: Git, Cython, Indri/Lemur, Hadoop
Platforms: GNU/Linux, Mac OSX

Software Released

Python Library for IR and ML (`github.com/alad/mekano`)
A Python library for rapid prototyping and experimentation in information retrieval and machine learning. Supports reading common file formats, tokenizing text, representing text documents as sparse vectors, creating inverted indexes. Common operations (like dot-products) implemented in C for speed.

Expectation Maximization Demo (`www.cs.cmu.edu/~alad/em`)
A Java applet, later ported to HTML5, to demonstrate the use of EM algorithm for solving a Gaussian Mixture. Used as a demo in a course in Technion - Israel Institute of Technology.

Awards and Achievements

- Awarded the Yahoo! Ph.D. Fellowship for 2007–2009.
  (Five recipients chosen per year from US graduate schools).
- Offered the Commonwealth Scholarship for graduate studies in UK for 2005–2007.
- Awarded merit scholarship for outstanding academic performance in undergraduate program at IIIT Allahabad for 2001–2005.
- Received second prize in all-India software designing contest at IIT Chennai.
  (Designed and implemented a full-fledged image editor with pluggable filters in a team of three.)

References

Yiming Yang
Professor
LTI, Carnegie Mellon University
Email: yiming@cs.cmu.edu
Web: http://www.cs.cmu.edu/~yiming

Jaime Carbonell
Professor & Dept. Head
LTI, Carnegie Mellon University
Email: jgc@cs.cmu.edu
Web: http://www.cs.cmu.edu/~jgc

Jamie Callan
Professor
LTI, Carnegie Mellon University
Email: callan@cs.cmu.edu
Web: http://www.cs.cmu.edu/~callan

Jan Pedersen
Chief Scientist
Microsoft, Inc.
Email: jpederse@yahoo.com
Web: http://www.jopedersen.com