Yi Yang

yiyang@gatech.edu ♦ (404) 263-6322 ♦ http://www.cc.gatech.edu/~yyang319/

EDUCATION

Aug. 2011 – present **Georgia Institute Technology, Atlanta, GA, USA**
PhD student in School of Computer Science, College of Computing

Sep. 2007 – Jun. 2011 **Beihang University (formerly Beijing University of Aeronautics and Astronautics), China**
B.S. Major in Software Engineering, GPA: 3.70/4.0, Major GPA: 3.88/4.0, Rank: 2/144

RESEARCH EXPERIENCE

Areas of Research Experience: Natural Language Processing, Data Mining, Information Retrieval, Information Security

School of Computer Science, College of Computing, Georgia Institute Technology

Aug. 2011 – present *Research Assistant, Data Intensive Systems Lab, CERCS*

- Detecting Fake Follower in Twitter
  - Developed techniques to detect fake followers on Twitter.
  - Implemented profile-based features, graph-based features, behavior features and content features.
  - Utilized directed and undirected graph-based semi-supervised learning.
- Learning to Rank Bad Reviewers
  - Built a bad word collection.
  - Tried to find bad reviewers according to how often they use such bad words.
  - Employed the machine learning method to rank reviewers.

National Lab of Intelligent Technology & Systems, Tsinghua University

Dec. 2009 – Jul. 2011 *Research Assistant, Intelligent Information Processing and Bioinformatics Group*

- Identifying Review Spam
  - Built a review spam collection by crawling over 30,000 reviews.
  - Determined the influence of different features in the supervised learning framework.
  - Proposed a co-training strategy based on review features and reviewer features.
- Ranking Short Texts using Quality Measures
  - Proposed and implemented a regression model to recommend high-quality tweets.
  - Implemented various features including query relevance, content quality, user authority (based on page-rank algorithm) and profile data to determine the quality measures.
  - Proposed several regularization factors to produce prediction more consistently and reliably.
- Patent Similarity Measurement and Patent Classification
  - Determined similar patents to a given patent, based on a term weighted model (ELT).
  - Hierarchical Multi-label Classification of patents using Two-way Poisson mixture models.

Department of Computer Science and Technology, Tsinghua University

Jul. 2010 – Nov. 2010 *Research Assistant, Knowledge Engineering Lab (Group)*

- Query Understanding for Web Search over Heterogeneous Networks
  - Proposed an iterative reinforcement approach which combined query segmentation (using a bigram model) and entity classification (using a unigram model) to mutually benefit each other.
  - Applied the iterative reinforcement approach to a real academic search system, Arnetminer, which improves the utility of the system.

PUBLICATIONS

- Fangtao Li, Minlie Huang, **Yi Yang**, Xiaoyan Zhu. Learning to Identify Review Spam. In Proceedings of the Twenty-Second International Joint Conference On Artificial Intelligence (IJCAI 2011). (Regular Paper) (17% acceptance rate)
- Minlie Huang, **Yi Yang**, and Xiaoyan Zhu. Quality-biased Ranking of Short Texts in Microblogging Services. In Proceedings of the 5th International Joint Conference on Natural Language Processing (IJCNLP 2011). (Regular Paper)

- Rezarta Islamaj Doğan, **Yi Yang**, Aurélie Névéol, Minlie Huang, Zhiyong Lu. Identifying protein-protein interactions in biomedical text articles. In Proceedings of the Third BioCreative Challenge Evaluation Workshop

SELECTED SOFTWARE PROJECTS

Beijing Key Laboratory of Network Technology

Mar. 2010 *Intern*

- Network Node Mobility Model Simulator based on MANET
  - Used a cloud model to simulate node movement for optimization of mobile network nodes.
  - Simulate scenarios under this model, such as a fire and the division and convergence of the emergency response.

School of Software of Beihang University (BUAA)

Sep. 2008 – Dec. 2009 *Intern*

- Undergraduate Students BBS System
  - Provided as a platform for sharing resources modeled on the functions of basic BBS system.
- Graduated Students Thesis Answer System
  - Improve working efficiency of answer flow system.
- Programming Online Judge System
  - Provided a platform of online program development which has the functions of editing, saving, compiling and also support multiplayer develop and share the same program at the same time.

WORK EXPERIENCE

ChinaMobile Company

Jun. 2008 – Aug. 2008 *Intern*

- Helped communicate with clients, handled IT affairs and maintained the company inner system.

COMPUTER SKILLS

- Programming Language: C, C#, C++, Java, R
- Web Technologies: HTML, CSS, ASP, JavaScript, JSP
- Databases: SQL Server, Oracle, MySql
- Tools: Weka, Eclipse, Indri, Lucene

AWARDS AND HONORS

Sep. 2010 Excellent Student of Beihang University (Top 1%)

Sep. 2010 First Prize for Academic Excellent Student (Top 2%), Beihang University

Sep. 2009 Third Prize for ACM Programming Contest (Top 7 out of 160), Beihang University

Sep. 2009 First Prize for Innovative Student (Top 2%), Beihang University

Sep. 2009 Second Prize for Academic Excellent Student (Top 5%), Beihang University

Sep. 2009 National Scholarship (Top 2%), P.R.China

Sep. 2009 Merit Student of Beihang University (Top 10%)

Sep. 2008 Third Prize for Academic Excellent Student (Top 10%), Beihang University

Sep. 2008 Second Prize for Innovative Student (Top 10%), Beihang University