Zhiting Xu

Room 5390
Department of Computer Science
University of Wisconsin-Madison

Cell:(608)320-6110
Email:zhiting@cs.wisc.edu
http://www.cs.wisc.edu/~zhiting

Objective 2009 Summer Internship

Interests Machine Learning, Natural Language Processing, Information Retrieval

Education

- Graduate Student — Sept. 2008 -
  Computer Science, University of Wisconsin-Madison
  GPA: 4.0
- B.S. in Computer Science, Fudan University — Sep. 2005-July 2008
  Top 3%
  -Finished the four-year program in three years
- Preparatory Student — Sep, 2004 till Sep, 2005
  Computer Science, Fudan University
  -Won an information contest, and was recommended to the college, so that took courses at Fudan during the third year of high school

Skills

Languages: C/C++, Java, Python, Pascal, Matlab
Operating System: Linux(Arch), Windows
Machine Learning and NLP packages: CRF++,Mallet,SVMLight,SGTLight,Weka openNLP,The Stanford Parser,WordNet,HowNet
Algorithms: Experienced programming of Conditional Random Fields with exact inference, Maximum Entropy, miscellaneous feature engineering of natural language processing. Familiar with graphical model algorithms, spectral graph cut algorithms and other popular machine learning models and algorithms.

Work Experience

Visiting Student — 07/07-08/07
Microsoft Research Aisa (MSRA)
Finished the semi-supervised learning project

Publications

- Andrew Goldberg, Nathanael Fillmore, David Andrzejewski, **Zhiting Xu**, Bryan Gibson, and Xiaojin Zhu. May all your wishes come true: A study of wishes and how to recognize them. In *North American Chapter of the Association for Computational Linguistics-Human Language Technologies(NAACL HLT)*, 2009.
- Andrew Goldberg, Xiaojin Zhu, Aarti Singh, **Zhiting Xu** and Robert Nowak. Multi-manifold semi-supervised learning. In *Twelfth International Conference on Artificial Intelligence and Statistics(AISTATS)*, 2009.
- Yuejie Zhang, **Zhiting Xu**, Tao Zhang. Fusion of Multiple Features for Chinese Named Entity Recognition based on CRF Model. *Asia Information Retrieval Symposium(AIRS 2008), 2008*
- **Zhiting Xu**, Xian Qian, Yujie Zhang, Yaqian Zhou. Sighan Bakeoff 2008: CRF-based Models for Word Segmentation, Named Entity Recognition and Part-Of-Speech Tagging. *SIGHAN Workshop on Chinese Language Processing 2008,* 2008.

- Yaofeng Wang, Yuejie Zhang, **Zhiting Xu**, Tao Zhang. Research on Dual Pattern Of Unsupervised And Supervised Word Sense Disambiguation. *Proceedings of 2006 International Conference on Machine Learning and Cybernetics(ICMLC 2006),* 2006

Projects

- **A study of wishes**
  With Professor Xiaojin(Jerry) Zhu el al 09/08-

  People from around the world offered up their wishes to be printed on confetti and dropped from the sky during the famous New Year's Eve "ball drop" in New York City's Times Square. We present an in-depth analysis of this collection of wishes. We then leverage this unique resource to conduct the first study on building general "wish detectors" for natural language text.

- **Multi-Manifold Separation for Semi-Supervised Learning**
  With Professor Xiaojin(Jerry) Zhu et al 09/08-

  Given n labeled points, $\{\mathbf{x}_i, y_i\}_{i=1}^{n} \in \mathbb{R}^D \times \mathbb{R}$ and m unlabeled points $\{\mathbf{x}_i\}_{i=n+1}^{n+m} \in \mathbb{R}^D$ sampled iid from a mixture of manifolds, aims to separate the manifolds.

- **Research On the Hierarchical Condition Random Fields**
  National Undergraduate Innovational Experience Program 05/07-05/08

  This research focuses on developing Hierarchical Conditional Random Fields to address Chinese Part-Of-Speech Tagging on coarse documents (without pre-segmentation). Proposed a method to compute inference of HCRF efficiently under some assumptions. To compare the results with other approaches on Chinese POS Tagging.

- **Semi-Supervised Learning for Relation Extraction**
  UCLA IPAM RIPS Program 2007, MSRA 07/07-08/07

  Found a new approach based on pairwise distance to do semi-supervised.Proposed a new method and compared it with Trasductive SVM provided by SVMLight and SDP.

- **Research On Chinese Segmentation and Named Entity Recognition**
  Wangdao Research Program for Undergraduate 11/06-10/07

  Compare Maximum Entropy model with Conditional Random Fields combining several local features and global features along with human knowledge to classify Chinese OOV words. Showed that human knowledge not only can improve the performance of the classifier, but also can smooth the model, and thus the training data size might be reduced. Also, demonstrated that global feature can overcome the problem of limited available local features during the second or latter appearance of a Named Entity.
  Participated in SIGHAN 2008 bakeoff on Segmentation, Named Entity Recognition and Part-Of-Speech.

- **Research On Word Sense Disambiguation**
  Media Computing and Web Intelligence Lab 12/05-11/06

  Used Weka to compare several popular classifiers and found that SVM performed best among all classifiers using our feature set. In addition, investigated unsupervised for this problem, and showed that the use of WordNet to do an extended lesk could improve the performance.

Teaching

- Teaching Assistant for CS635: Tools and Environments for Optimization with Prof Michael Ferris. Spring, 2009
- Teaching Assistant for CS540: Introduction to Artificial Intelligence with Prof Xiaojin(Jerry) Zhu. Fall, 2008

Honors

- Prize for Innovations in Research Autumn, 2007
  -1 all over the School of Information Science and Engineering
- Wangdao Scholar, Autumn, 2006
  -Around 50 students in the university each year are awarded this title after they finished a research project
- 2005 ACM International Collegiate Programming Contest Regional Oct. 2005
  Chengdu site 15^{th} Place
- 2005 ACM International Collegiate PRogramming Contest Regional Nov. 2005
  Beijing site 11^{th} Place
- National Olympic Information Contest (NOI 2004), Bronze Medal, Aug. 2004
  -Won the 67th Place in China

Reference Available Upon Request