

# Huayi Li

Email: [lhymvp@gmail.com](mailto:lhymvp@gmail.com)

Cell: (312) 927-1433

<https://www.cs.uic.edu/~hli/>

---

## Summary

Ph.D. in Computer Science with background in Machine Learning and Data Mining.  
Solid research papers in premier international conferences including AAAI, WWW, ICDM, and ACL.  
Strong expertise in Machine Learning, Spam Detection, Text Mining and Social Networks.  
Responsible team player with great troubleshooting skills and enthusiasm for innovation.  
Extensive experience of tools and programs for analysis and visualization.

## Experiences

### Google, Mountain View, CA

#### Software Engineer

06/2016 ~ present

- **Action Based Better Advertising (Value Based Bidding)**
  - Working on Google Display Ads auto-bidding algorithms & infrastructures.

### Microsoft, Bellevue, WA

#### Data Scientist Intern

06/2015 ~ 08/2015

- **Improving Ads relevance models via Active Learning**
  - Proposed and implemented Active Learning models to train robust ranking models with less data.
  - Built interactive tools to analyze and compare model performance.
  - Launched and optimized Ads campaigns in Bing Ads and Google Adwords for Advertisers.

### Microsoft, Bellevue, WA

#### R&D Intern

05/2014 ~ 08/2014

- **Interactive Big Data Analysis platform**
  - Developed a toolkit called PinQ that translates a LinQ style python script to Scope/iScope scripts.
  - Created efficient caching strategies to optimize interactive/semi-interactive big data analysis.
  - Analyzed scorecards and conducted various experiments for Bing A/B test.

### HERE Map, Chicago, IL

#### R&D Intern

05/2013 ~ 08/2013

- **Anomaly Detection of Map Data Repository**
  - Built and extracted analytics of map data with map-reduce.
  - Applied statistical machine learning techniques to evaluate the changes of map data.
  - Visualized the quarter-to-quarter changes in the map by various dimensions.
- **Hot Area of Interest (Hot-AOI)**
  - Applied spatial clustering algorithm on geo-tagged social feeds and GPS probes.
  - Analyzed the temporal and spatial pattern of clusters.
  - Generated real-time trends of people/traffic flow.

Chinese Academy of Science, Shenzhen, China

Research Intern

05/2012 ~ 08/2012

- **Sentiment Analysis System**
  - Extracted entities and features from product reviews
  - Built feature-based opinion summarization system.
  - Visualized and analyzed social happiness index of China.

## Education

University of Illinois at Chicago, IL

08/2010 ~ 05/2016

Ph.D. in Computer Science

Advisor: [Prof. Bing Liu](#)

GPA: 3.9/4.0

Nanjing Normal University, China

09/2005 ~ 06/2009

B.S. in Computer Science

GPA: 3.8/4.0 Rank 1/237

## Publications

1. **Huayi Li**, Geli Fei, Shuai Wang, Bing Liu, Weixiang Shao, Arjun Mukherjee, Jidong Shao. *Bimodal Distribution and Co-Bursting in Review Spam Detection*. **WWW 2017**.
2. Geli Fei, **Huayi Li**, Bing Liu. *Opinion Spam Detection in Social Networks*. Book chapter in Federico Pozzi, Elisabetta Fersini, Enza Messina and Bing Liu, editors, *Sentiment Analysis in Social Networks*, **Elsevier 2016**.
3. Rachel Harsley, Bhavesh Gupta, Barbara Di Eugenio, **Huayi Li**. *Hit Songs' Sentiments Harness Public Mood & Predict Stock Market*. The 7th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, **WASSA 2016**.
4. **Huayi Li**, Zhiyuan Chen, Arjun Mukherjee, Bing Liu, Jidong Shao. *Analyzing and Detecting Opinion Spam on a Large-scale Dataset via Temporal and Spatial Patterns*. **ICWSM 2015**.
5. **Huayi Li**, Arjun Mukherjee, Jianfeng Si, Bing Liu. *Extracting Verb Expressions Implying Negative Opinions*. **AAAI 2015**.
6. **Huayi Li**, Arjun Mukherjee, Bing Liu, Rachel Kornfield and Sherry Emery. *Detecting Campaign Promoters on Twitter using Markov Random Fields*. **ICDM 2014**.
7. **Huayi Li**, Zhiyuan Chen, Bing Liu, Xiaokai Wei and Jidong Shao. *Spotting Fake Reviews via Collective Positive-Unlabeled Learning*. **ICDM 2014**.
8. **Huayi Li**, Arjun Mukherjee, Bing Liu, Jidong Shao. *Spotting Fake Reviews using Positive-Unlabeled Learning*. *Journal of Computación Y Sistemas* **2014**.
9. Jianfeng Si, Arjun Mukherjee, Bing Liu, Sinno Jialin Pan, Qing Li and **Huayi Li**. *Exploiting Social Relations and Sentiment for Stock Prediction*. **EMNLP 2014**.
10. Jianfeng Si, Arjun Mukherjee, Bing Liu, Qing Li, **Huayi Li**, and Xiaotie Deng. *Exploiting Topic based Twitter Sentiment for Stock Prediction*. **ACL 2013**.
11. **Huayi Li**. *Detecting Opinion Spam in Commercial Review Websites*. Ph.D. Thesis (University of Illinois at Chicago, 2016).

## Skills

- **Languages:** C/C++, Python, Java, C#, R.
- **Databases:** MySQL, MongoDB, SQLite, and so on.
- **Big Data:** Flume, Dremel; Cosmos, Scope/iScope; Hadoop, Hive, Pig.
- **Web Developments:** HTML, CSS, JavaScript, XML, JSP, ASP.
- **Data Mining Tools:** Scikit-learn, Weka, Lingpipe, CRF++, SVMlight, Stanford CoreNLP, OpenNLP.

## **Honors and Achievements**

- Google Spot bonus from manager, March 2017
- Google Kudos from peers, March 2017
- UIC Student Presenter Award, June 2015
- AAAI International Conference on Web and Social Media (ICWSM) Student Grant Award, May 2015
- UIC Graduate Student Council Travel Award, February 2015
- IEEE International Conference on Data Mining (ICDM) Travel Award, October 2014
- 2013 Nokia HERE hack week finalist, June 2013 (20 out of 70)
- Outstanding student in Nanjing Normal University, November 2008 (top 10)
- Software Scholarship awarded by Jiangsu Province Government, September 2008 (top 1%)
- China's National Innovation Scholarship, December 2007 (top 5%)

## **Professional Services**

reviewer for

- AAAI'17-18 (PC member)
- EMNLP'16 (PC member)
- CIKM'16
- TKDE'16
- DMKD'16
- TWEB'15-17
- Computer Speech and Language'16
- The Computer Journal'15