Xiaodan Zhang

zhang349@illinois.edu · http://xiaodanzhang.com

EDUCATION

University of Illinois at Urbana-Champaign   **Urbana-Champaign, IL**
- Major: Computer Science   **Master of Science,** December 2015 (expected)
- GPA: Available upon request
- Concentration: Text mining, Information Retrieval

University of Illinois at Urbana-Champaign   **Urbana-Champaign, IL**
- Major: Statistics and Computer Science   **Bachelor of Science,** May 2014
- Minor: Informatics, Mathematics
- GPA: Available upon request
- **Departmental Honor**: Distinction STAT/CS

WORK EXPERIENCE

Department of Computer Science, *Teaching Assistant*, University of Illinois, Champaign, IL   **August 2014 – Present**
- Help monitor student teams and their progress in Senior Project I (CS 492) and II (CS493) class
- Meet with teams regularly to talk about their technical designs, plans and the software development process

Bazaarvoice, Inc., *Software Engineer Intern*, Austin, TX   **June 2014 – August 2014**
- Worked on Data Services team to help rebuild the Universal Brand Catalog that enables content moderators to view brand candidates and validate their legitimacy alongside validating their associated synonyms
- Worked as a data wrangler to clean, transform, group and analyze brand data using Python and text mining techniques
- Web Crawled brand name lists using Scrapy framework
- Worked on back-end to design and implemented API definitions using Java and Dropwizard framework
- Built, released and deployed the project using Jenkins

NeuStar, Inc., *Data Science Intern*, Neustar Innovation Center, Champaign, IL   **June 2013 – May 2014**
- Analyzed large scale DNS traffic data stored in the Hadoop cluster using HiveQL, Python and Shell scripts
- Wrote Hive UDFs (e.g. substring_index function in SQL, etc.) in Java to process the domain name field
- Visualized the traffic data in Tableau by creating geographical Heat maps, Bar charts, Bubble graphs, etc.
- Built a DNS Dashboard for displaying plots generated by Tableau of major characteristics and descriptive statistics like Traffic, Response time, Response rate based on different Domains and Nodes
- Utilized Machine Learning techniques to build models for detecting DDoS attacks

CS357 Numerical Analysis I, *Undergraduate Grader,* University of Illinois, Champaign, IL   **January 2012 – December 2012**
- Graded the written homework, programming assignments (in Python/Matlab), and projects of over 150 students who were enrolled in CS357 for two semesters
- Answered students' questions on the Piazza website

Irwin Academic Services Center (University of Illinois), *Tutor*, Champaign, IL   **September 2011- May 2012**
- Tutored for the Division of Intercollegiate Athletics in CHEM, CS, STAT, and PHYS courses
- Helped improve student-athlete's performance in the classroom by teaching them efficient ways of learning materials in class and preparing for exams

RESEARCH AND PROJECT EXPERIENCE

Automatic Extraction and Ranking of Menu Items, University of Illinois, Champaign, IL   **Fall 2014**
- Created a framework for extracting candidate dishes, ranking and visualizing the most popular dishes in a restaurant
- Three steps: Candidate dish extraction, Candidate dish filtering with reference models, Ranking dishes by rating and novelty
- Utilized Phrase Mining technique, NLP noun-phrase extraction algorithm and smoothed background language model

Social Visualization Project - PortraitOfStrongMan, University of Illinois, Champaign, IL   **Fall 2014**
- Data Visualization for training and competition data of Powerlifting and Strongman from 2012 to 2014
- Created quantified-selfie line charts, bubble charts and comparison line charts for comparing the Bench Press, Squat and Deadlift training data of multiple trainers using Python, Highcharts JS and D3.js
- Zoom-in effect supported with clickable icons linked to Strongman competition videos on YouTube

Social Visualization Project – AccentDiff, University of Illinois, Champaign, IL   **Fall 2014**
- Data Visualization for comparing the accent of different countries including both native and non-native English speakers
- Web crawled data using Scrapy framework, edited and transformed audio data using FFMPEG and AudioWaveform, and D3.js for visualizing the data

Social Visualization Project – TemporalPoliMap, University of Illinois, Champaign, IL   **Fall 2014**
- Data Visualization for exploring events happened between politically significant countries during specific time periods
- Used Google BigQuery for querying and extracting data from the GDELT database, and D3.js for visualizing the data

Text Information Systems Project - ReviewsHub, University of Illinois, Champaign, IL   **Spring 2014**
- A review Analyzer for Best Buy Reviews. It has a following three functions:
- Search Engine: Search in all the reviews about product properties that users queried
- Compare Visualization: Search and visualize comparison patterns in reviews
- Summarizer: Summarize long reviews to a few sentences

Programming Studio Assignments Portfolio, University of Illinois, Champaign, IL   **Fall 2013**
- SVN code viewer with Comment Filter System (Link: http://web.engr.illinois.edu/~zhang349/Portfolio)
- All Java codes for a Chess Library and Python codes for a CSAir Flight System can be shown with jQuery slide down effects

Department of Computer Science, *Research Student*, University of Illinois, Champaign, IL   **August 2012 – August 2013**
- A scientific computing project about Topology of Relations; read papers about Topological Data Analysis in sports
- Did data analysis and integration by writing scripts in Python and plotting graphs with Matplotlib and Numpy packages

Yummy! Web-app, University of Illinois, Champaign, IL   **Spring 2013**
- This is an IM-like web application based on Google App Engine. It provides a place for people to discuss where to eat with others in a specific group. (Link: http://yummy-webapp.appspot.com)

Web Programming Course Project, University of Illinois, Champaign, IL   **Spring 2013**
- A Client-side TodoList that is based on HTML5 Local Storage
- The functions include adding/removing list items, supporting multiple lists with names, marking list items as done, deleting all don items in a list, lists persist between browser sessions, animations on main actions, and sorting items by deadlines

Statistical Computing Course Project, University of Illinois, Champaign, IL   **Spring 2013**
- Analyzed a breast cancer dataset to find genetic markers by using Principal Component Analysis.

Personal Website, University of Illinois, Champaign, IL   **Winter 2012**
- Wrote in HTML5, CSS, jQuery, JavaScript and with 960 Grid System, Google Chart Tools, Leaflet library
- Contain my recent accomplished projects (Link: http://web.engr.illinois.edu/~zhang349/old)

Applied Regression and Design Course Project, University of Illinois, Champaign, IL   **Fall 2012**
- Explored the "Titanic Disaster Data" on *Kaggle.com* about the factors that affect a passenger's survival
- Fitted several Logistic Regression models, Tree models and Random Forest model, and did predictions in R

Time Series Analysis Course Project, University of Illinois, Champaign, IL   **Fall 2012**
- Fitted a forecasting SARIMA model for the monthly industry sales of printing and writing papers (1963-1972) in R

Database Systems Course Project, University of Illinois, Champaign, IL   **Spring 2012**
- Created a website that offers campus-based book exchange service with LAMP package
- Worked on MySQL database and Ranking System in the aspects of manipulating experiential matrix

LEADERSHIP AND ACTIVITIES

Illini Statistics Club, *Webmaster*, Champaign, IL   **May 2013 – Present**
- Create electronic signup sheets with HTML5, PHP, MySQL to store members' information in an efficient way
- Create and manage the club's website: get more and more undergraduate/graduate students of Statistics major, Stat&CS major involved in career events, club social events, mentor/peer advisor programs, and company tech talks
- (Link: http://publish.illinois.edu/illinistatistics)

Women in Computer Science, *Active Member*, Champaign, IL   **September 2010- Present**
- Actively take part in general meetings, tech talks and networking events

Society of Asian Scientists and Engineers, *Outreach Chair*, Champaign, IL   **September 2010 – May 2012**
- Helped plan volunteer events for SASE to promote science and engineering in the local community
- Took charge of the Engineer in the Classroom (EITC) Program for 2011

Silicon Valley Entrepreneurship Workshop, *Attendee,* Silicon Valley, CA   **January 2012**
- Selected as one of the 25 students to a one-week Silicon Valley Entrepreneurship Workshop and visit several start-ups
- Interacted with corporate leaders to study technology commercialization and new venture creation
- Understood the prevailing trends of Big Data technology

TECHNICAL SKILLS AND QUALIFICATIONS
- **Computer Languages:** Java, Python, C/C++; R, Matlab, SAS; SQL, HTML5, CSS, JavaScript, jQuery; Ocaml
- **Other:** Hadoop, HIVE, MapReduce; Tableau, Weka; Github, Agile Software Development; Maven
- **Languages Spoken:** English, Mandarin Chinese, French

HONORS AND CERTIFICATIONS
- **2014 Fall Conference Travel Grant from UIUC CS department**   **Fall 2014**
- **2014 Grace Hopper Celebration Sponsorship from Bazaarvoice Inc.**   **Summer 2014**
- **Graduated with Departmental Honor: Distinction STAT/CS**   **Spring 2014**
- **2013 Grace Hopper Celebration Scholarship from UIUC CS department**   **Spring 2013**
- **University of Illinois Dean's List**   **Fall 2010, Spring 2012**
- **SAS Certified Base/Advanced Programmer for SAS 9**   **Spring/Summer 2012**
- Certifications of Accomplishment for Exploratory Data Analysis on coursera.com   **Summer 2014**
- Certifications of Accomplishment for Reproducible Research on coursera.com   **Summer 2014**
- Certifications of Accomplishment for Statistical Inference on coursera.com   **Summer 2014**
- Certifications of Accomplishment for R Programming on coursera.com   **Spring 2014**
- Certifications of Accomplishment for Getting and Cleaning Data on coursera.com   **Spring 2014**
- Certifications of Accomplishment for Data Scientist Toolbox on coursera.com   **Spring 2014**
- Certifications of Accomplishment for Computing for Data Analysis on coursera.com   **Winter 2012**
- Certifications of Accomplishment for Machine Learning on coursera.com   **Fall 2012**