Visualization and Data Analytics Research Center (VIDA)
370 Jay Street
Brooklyn, NY, 11201
ky630@nyu.edu | keyang0923@gmail.com
Webpage : keyang0923.github.io
+1-215-900-5854

Ke Yang

PhD Candidate, Computer Science, New York University

Education

Ph.D. in Computer Science, New York University, New York, U.S.A
Advisor: Julia Stoyanovich *January, 2019 - December, 2020 (Expected)*
GPA: 4.0/4.0

Ph.D. in Computer Science, Drexel University, Philadelphia, U.S.A
Advisor: Julia Stoyanovich *September, 2015 - December, 2018*
GPA: 3.98/4.0

M.S.E in Computer Science, Beijing Technology and Business University, Beijing, China
Advisors: Qian Mo and Zhongming Han *September, 2012 - June, 2015*
GPA: 3.51/4.0

B.Eng. in Software Engineering, Beijing Technology and Business University, Beijing, China
Advisor: Qian Mo *September, 2008 - June, 2012*
GPA: 3.45/4.0

Research Interests

Responsible data management, non-discriminatory machine learning
Specifically my research focuses on the multiple quality aspects of algorithmic outcomes: fairness, diversity, stability, etc, with the goal to develop methods to efficiently evaluate and improve it.
Current work focus on quantifying the stability of ranked outcomes.

Research Projects

Data, Responsibly
Supervisor: Prof. Julia Stoyanovich *December, 2015 - Present*

- *Encoding fair schemes in ranking-related applications and designing models to mitigate effect of bias towards sub-groups such as females in job hunting or African-American people in the prediction of criminal activities. More: DataResponsibly*
- *Designing diverse schemes for online set selection*
- *Making ranking-related decision procedures more transparent for users. Application: RankingFacts*

Analyzing and Preventing Spam Information in Online Shopping Websites
Supervisors: Prof. Zhongming Han and Qian Mo *May, 2013 - May, 2015*

- *Analyzed the profiles and behaviors of the large spamming groups in online shopping websites, such as Taobao.com and JD.com*
- *Designed a model employing graphical features extracted from users relationships to efficiently detect the large spamming groups in online shopping websites*
- A survey paper accepted in Journal of Software, one of the Chinese premier peer-reviewed journal in Computer Science
- A paper accepted in Chinese Journal of Computers, one of the Chinese premier peer-reviewed journal in Computer Science
- Applied a China invention patent for the above detection model

Detecting and Predicting Hot Topics in Social Network based on Analyzing Human Behavior Pattern
Supervisor: Prof. Zhongming Han *December, 2012 - June, 2013*

- Designed a new probabilistic graphical model to efficiently identify Sina Weibo spamming users by combining the profiles and behaviors features from users in Sina Weibo (Chinese version of Twitter)

- Compared the performance with the baseline SVM model

- Paper accepted in Journal of Embedded System

Publications

***Ke Yang**, Vasilis Gkatzelis, and Julia Stoyanovich, "Balanced ranking with diversity constraints", in Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI 2019), Macao, China, DOI:10.24963/ijcai.2019/836*

***Ke Yang**, Julia Stoyanovich, Abolfazl Asudeh, Bill Howe, HV Jagadish, and Gerome Miklau, "A Nutritional Label for Rankings" (demo) to appear in Proceedings of the 2018 ACM International Conference on Management of Data (SIGMOD 2018), Houston, USA*

*Julia Stoyanovich, **Ke Yang**, HV Jagadish, "Online Set Selection with Fairness and Diversity Constraints" in Proceedings of the 21th International Conference on Extending Database Technology (EDBT 2018), Vienna, Austria, DOI:10.5441/002/edbt.2018.22*

***Ke Yang**, Julia Stoyanovich, "Measuring Fairness in Ranked Outputs" in Proceedings of the 29th International Conference on Scientific and Statistical Database Management (SSDBM 2017), Chicago, USA, DOI:10.1145/3085504.3085526*

*Zhongming Han, **Ke Yang**, Xusheng Tan, "Analyzing Spectrum Features of Weight User Relation Graph to Identify Large Spammer Groups in Online Shopping Websites" in Chinese Journal of Computers, 40(4): 939-954 (2017) (in Chinese), DOI:10.11897/SP.J.1016.2017.00939*

*Zhongming Han, **Ke Yang**, Fengmin Xu, Dagao Duan, "Probabilistic Graphical Model for Detecting Spammers in Microblog Websites" in Journal of Embedded System, 8(1): 12-23 (2016), DOI:10.1504/IJES.2016.073747*

*Qian Mo, **Ke Yang**, "Overview of Web spammer detection" in Journal of Software, 25(7): 1505-1526 (2014) (in Chinese), DOI:10.13328/j.cnki.jos.004617*

*Kai Wang, **Ke Yang**, "A method and application system of detecting web spammers" in China Invention Patent, (2015) (in Chinese), No. CN201510012860*

Teaching Experience

***Teaching Assistant** - Drexel University*
- CS500 Fundamentals of Databases: January, 2017 - March, 2017
- CS461 Database Systems: March, 2017 - May, 2017

Work Experience

***Research intern** at AT&T Lab Research (New York Office)* *June, 2019 - August, 2019*
Mentors: Emily Dodwell, Ritwik Mitra, and Balachander Krishnamurthy
Worked on a project named SIFT: Steps to Insure Fairness and Transparency. In this project, we are interested in ensuring that ML projects meet reasonable guidelines for fairness and avoid (even inadvertent) bias and discrimination. Towards this end we built a tool that assists internal AT&T projects through the full ML project pipeline

***Research engineer** at Elite & Resource (start-up company)* *November, 2014 - August, 2015*
Mentor: Peng Sun
Analyzed flood disaster data from China Institute of Water Resources and Hydro-power Research
Designed a model to predict the water flow rate of small watershed torrents and encoded the model as a component of national watershed data management system

Software Projects & Skills

Ranking Facts *full-stack development: Python, Django, JavaScript, JQuery*
Data Synthesizer *front-end development: JavaScript, JQuery*
Researcher Questionnaire *database design and front-end development: PHP, Drupal, MySql*
DataResponsibly *website development: Jekyll*
Other Used Languages (project-level)*: Java, C, C++, MATLAB and R*