

Ke Yang

PHD CANDIDATE, COMPUTER SCIENCE, NEW YORK UNIVERSITY

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

EDUCATION

Ph.D. in Computer Science, New York University, New York, U.S.A
Advisor: Julia Stoyanovich *January, 2019 - March, 2021 (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 and knowledge management, non-discriminatory machine learning, data and knowledge mining, social network analysis
Current work focus on fairness and diversity in ranking-related applications, with the goal to mitigate the effect of historical discrimination against legally-protected groups.

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	<p>Ke Yang, Vasilis Gkatzelis, and Julia Stoyanovich, “Balanced ranking with diversity constraints”, in <i>Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI 2019)</i>, Macao, China, DOI:10.24963/ijcai.2019/836</p> <p>Ke Yang, Julia Stoyanovich, Abolfazl Asudeh, Bill Howe, HV Jagadish, and Gerome Miklau, “A Nutritional Label for Rankings” (demo) to appear in <i>Proceedings of the 2018 ACM International Conference on Management of Data (SIGMOD 2018)</i>, Houston, USA</p> <p>Julia Stoyanovich, Ke Yang, HV Jagadish, “Online Set Selection with Fairness and Diversity Constraints” in <i>Proceedings of the 21th International Conference on Extending Database Technology (EDBT 2018)</i>, Vienna, Austria, DOI:10.5441/002/edbt.2018.22</p> <p>Ke Yang, Julia Stoyanovich, “Measuring Fairness in Ranked Outputs” in <i>Proceedings of the 29th International Conference on Scientific and Statistical Database Management (SSDBM 2017)</i>, Chicago, USA, DOI:10.1145/3085504.3085526</p> <p>Zhongming Han, Ke Yang, Xusheng Tan, “Analyzing Spectrum Features of Weight User Relation Graph to Identify Large Spammer Groups in Online Shopping Websites” in <i>Chinese Journal of Computers</i>, 40(4): 939-954 (2017) (in Chinese), DOI:10.11897/SP.J.1016.2017.00939</p> <p>Zhongming Han, Ke Yang, Fengmin Xu, Dagao Duan, “Probabilistic Graphical Model for Detecting Spammers in Microblog Websites” in <i>Journal of Embedded System</i>, 8(1): 12-23 (2016), DOI:10.1504/IJES.2016.073747</p> <p>Qian Mo, Ke Yang, “Overview of Web spammer detection” in <i>Journal of Software</i>, 25(7): 1505-1526 (2014) (in Chinese), DOI:10.13328/j.cnki.jos.004617</p> <p>Kai Wang, Ke Yang, “A method and application system of detecting web spammers” in <i>China Invention Patent</i>, (2015) (in Chinese), No. CN201510012860</p>
TEACHING EXPERIENCE	<p>Teaching Assistant - Drexel University</p> <ul style="list-style-type: none"> - CS500 Fundamentals of Databases: January, 2017 - March, 2017 - CS461 Database Systems: March, 2017 - May, 2017
WORK EXPERIENCE	<p>Intern at Elite & Resource (start-up company)</p> <p>Research engineer, Supervisor: Peng Sun November, 2014 - August, 2015</p> <p>Analyzed flood disaster data from China Institute of Water Resources and Hydro-power Research</p> <p>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</p> <p>Research Intern at AT&T Research Lab (New York Office)</p> <p>Research intern, Supervisor: Emily Dodwell, Ritwik Mitra, and Balachander Krishnamurthy</p> <p>June, 2019 - August, 2019</p> <p>Worked on a project called SIFT: Steps to Insure Fairness and Transparency</p> <p>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</p>

that assists internal AT&T projects through the full ML project pipeline

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