Suin Kim

http://suin.kim, suin.kim@kaist.ac.kr

RESEARCH INTERESTS

My research interests lie at the intersection of machine learning and computational social science. I am particularly interested in topic modeling, a subfield of machine learning, which can find meaningful patterns of social interactions from large, un-annotated datasets. I work on developing novel nonparametric topic models to discover and analyze patterns of language usage, emotions, sentiments, and media effects from online social network.

I am developing Elice[1], an online computer science education platform that supports real-time evaluation and feedback on the grouped learning environment. I analyze the student's data collected from Elice, including the set of intermediate codes of solving the exercise, to understand the learning process of students. In the flipped classroom setting in university, I use the machine learning algorithms, to automate the process of (i) generating the relevant feedback for students, (ii) identifying the help-needed students for teaching assistants, and (iii) summarizing how students well understood the lecture for professors.

EDUCATION

KAIST, Ph.D., Computer Science, 2012 − Present

KAIST, M.S., Computer Science, 2012

KAIST, B.S., Computer Science, 2010

PROFESSIONAL EXPERIENCES

Research Assistant, Qatar Computing Research Institute, Doha, Qatar, 2013 − 2014

Research Intern, Microsoft Research Asia, Beijing, PRC, 2012 − 2013

SWE Intern, Google Korea LLC, Seoul, Republic of Korea, 2009

Intern, LG Electronics Inc., Seoul, Republic of Korea, 2009

PUBLICATIONS

[1] Jungkook Park, Yeong Hoon Park, Suin Kim, and Alice Oh. Eliph: Effective visualization of code history for peer assessment in programming education. In *The 20th ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW 2017)*, 2017.

[2] Juhong Park, Suin Kim, and Alice Oh. Analysis of the effect of competition on player immersion and engagement in a mobile game. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, 2017.

[3] Suin Kim, Sungjoon Park, Scott A. Hale, Sooyoung Kim, Jeongmin Byun, and Alice Oh. Understanding editing behaviors in multilingual wikipedia. In *PLOS ONE*, 2016.

[4] Suin Kim, Jae Won Kim, Jungkook Park, and Alice Oh. Elivate: A real-time assistant for student and lecturer on the online cs education platform. In *L@S: Third Annual ACM Conference on Learning at Scale*, 2016.

[5] Suin Kim, Jae Won Kim, Jungkook Park, and Alice Oh. Elice: An online cs education platform to understand how students learn programming. In *L@S: Third Annual ACM Conference on Learning at Scale*, 2016.

---

[1] https://www.elice.io

[6] Sungjoon Park, Suin Kim, Scott A. Hale, Sooyoung Kim, Jeongmin Byun, and Alice Oh. Multilingual wikipedia: Editors of primary language contribute to more complex articles. In *Wiki-ICWSM'15 workshop*, 2015.

[7] Jae Won Kim, Dongwoo Kim, Brian Keegan, Joon Hee Kim, Suin Kim, and Alice Oh. Social media dynamics of global co-presence during the 2014 fifa world cup. In *the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI 2015)*, 2015.

[8] Oul Klara Han, Suin Kim, and Camille Roth. Who are birds of the same feather? epistemic communities in the eu twittersphere. In *The 2nd GESIS Computational Social Science Winter Symposium 2015*, 2015.

[9] Yeooul Kim, Suin Kim, and Alice Oh. A computational analysis of agenda setting. In *The 23rd International World Wide Web Conference (WWW 2014) (2-page poster)*, 2014.

[10] Suin Kim, Ingmar Weber, Li Wei, and Alice Oh. Sociolinguistic analysis of twitter in multilingual societies. In *Hypertext 2014, the 25th ACM Conference on Hypertext and Social Media (HT'14). (Short paper)*, 2014.

[11] Suin Kim, Jianwen Zhang, Zheng Chen, Alice Oh, and Shixia Liu. A hierarchical aspect-sentiment model for online reviews. In *The Twenty-Seventh AAAI Conference on Artificial Intelligence (AAAI-13)*, 2013.

[12] Suin Kim, JinYeong Bak, and Alice Haeyun Oh. Do you feel what i feel? social aspects of emotions in twitter conversations. In *The 6th International AAAI Conference on Weblogs and Social Media (ICWSM 2012). (4-page poster)*, 2012.

[13] Suin Kim, JinYeong Bak, and Alice Oh. Discovering emotion influence patterns in online social network conversations by suin kim, jinyeong bak, and alice oh, with ching-man au yeung as coordinator. *ACM SIGWEB Newsletter*, (Autumn), 2012.

[14] Joon Hee Kim, Dongwoo Kim, Suin Kim, and Alice Oh. Modeling topic hierarchies with the recursive chinese restaurant process. In *The 21st ACM International Conference on Information and Knowledge Management (CIKM 2012)*, 2012.

[15] Dongwoo Kim, Suin Kim, and Alice Oh. Dirichlet process with mixed random measures: A nonparametric topic model for labeled data. *The 29th International Conference on Machine Learning (ICML 2012)*, 2012.

[16] Jin Yeong Bak, Suin Kim, and Alice Oh. Self-disclosure and relationship strength in twitter conversations. In *The 50th Annual Meeting of the Association for Computational Linguistics (ACL 2012). (Short paper)*, 2012.

[17] Jong-bum Woo, Da-jung Kim, Suin Kim, Jaesung Jo, and Youn-kyung Lim. Interactivity sketcher: crafting and experiencing interactivity qualities. In *CHI'11 Extended Abstracts on Human Factors in Computing Systems*, 2011.

[18] Eunkyoung Jee, Suin Kim, Sungdeok Cha, and Insup Lee. Automated test coverage measurement for reactor protection system software implemented in function block diagram. In *Proceedings of the 29th International Conference on Computer Safety, Reliability and Security (SAFECOMP), LNCS 6351*. Springer, 2010.

PATENTS

Lim, Youn-kyung and Woo, Jong-bum and Kim, Suin. 2013. KR Patent 10-1262018, filed March 21, 2011, and issued May 02, 2013.

TALKS AND ACTIVITIES

2016.11. Organizer, Data Science Python Course[2]. CCEI, KAIST, Daejeon

2016.05. Artificial Intelligence at a Glance. CCEI, KAIST, Daejeon

2015.12. Organizer, Data Science Seminar[3] (Lectured by Prof. Wray Buntine, Monash University), KAIST, Daejeon

---

[2]https://academy.elice.io
[3]https://sites.google.com/site/datasciencekaist/

2015.04. Student Presentation: Language Complexity in Multilingual Wikipedia, CHI Post Workshop at KAIST, Daejeon

2014.11. Flipped-class Python lecture for middle school students

2013.05. Guest Lecture at KAIST CS570, Hidden Markov Models

2012.05. Guest Lecture at KAIST CS612/GCT784, Do you feel what I feel? Social Aspects of Emotions in Twitter Conversations

ACADEMIC SERVICES

Invited Reviewer, CHI 2017

Programming Committee, ICWSM-16: Tenth International AAAI Conference on Web and Social Media

Programming Committee, SocialNLP 2016: The 4th International Workshop on Natural Language Processing for Social Media

Reviewer, ACL 2015

Reviewer, WWW 2015

Subreviewer, ICWSM-2015

Subreviewer, ICWSM-2014

Member, Association for Computing Machinery

Member, Association for the Advancement of Artificial Intelligence

Member, Korean Institute of Information Scientists and Engineers

REFERENCES

Alice Oh, Associate Professor, Computer Science, KAIST, alice.oh@kaist.edu

Ingmar Weber, Senior Scientist, Qatar Computing Research Institute, iweber@qcri.qa

Jianwen Zhang, Associate Researcher, Microsoft Research Asia, jiazhan@microsoft.com

RESEARCH STATEMENT

With the Web pervading every part of our lives, we leave an incredible amount of traces behind, and those traces of activities, including posts, comments, clicks, likes, accumulate into zettabytes of information each year. Can such massive data open doors into understanding more about the social nature of humans? They can, but only with the right combination of advanced computational methods, appropriate mathematical models, and deep understanding of the social science literature. Such combination is the basic definition of computational social science, where my research interests lie. Particularly, I aim to contribute to the computational methods for social science, centered around developing new machine learning algorithms for discovering meaningful patterns of social interactions and the underlying process of human social behavior.

Research Accomplishments in Computational Social Science. Computational social science can transcend the limits of traditional social science along the dimensions of size, timeliness, geographic and demographic variety of data. Furthermore, it can propose solutions to research questions that were impractical or even impossible to answer with traditional methods of data collection and analysis. In my research on emotions, I collected and automatically analyzed how people affect each other's emotions in 153,054 conversations, significantly overshadowing the amount of data used in social science research by orders of magnitude [12][4]. With online customer reviews, I developed a machine learning algorithm that can organize the sentences in reviews into sub-topics and corresponding opinions from 30,876 reviews about 1,085 products, and this algorithm can be applied to any dataset of reviews without any human annotation [11]. Most recently, I looked into discovering the role of monolingual and bilingual speakers in multilingual societies [10]. This research included 536,840 people located in Qatar, Switzerland, Quebec, Singapore, and San Francisco, and the fully automated process of collecting and analyzing data means it can be used to answer the question of comparing various types of mulitlingualism in many multilingual societies around the world. This last research project is a transition into a more macro-level study of social behaviors, compared to most of previous research in computational social science with a focus on individuals, and it shows a totally new era of approaching a social science research question with an unprecedented amount and variety of naturally-occurring social behavior data.

Research Goal in Machine Learning. My research goal is to develop new algorithms in machine learning for modeling macro-scale social behavior. I am particularly interested in multilingualism at the level of societies, which has a host of far-reaching implications in this modern era of multilingualism in virtually every society. Large-scale studies on multilingualism using computational methods are just beginning, as multilingual speakers outnumber monolinguals in the world[5], and they try to overcome the limitations of traditional sociolinguistics research at a much smaller scale using controlled study or questionnaire-based methods. I will use what I have learned through past research in topic modeling, including papers published on AAAI [11] and ICML [15]. One major contribution of my research will be discovering the sociality of language use with an analysis of large-scale, naturally occurring observations. Another will be demonstrating the efficacy of machine learning algorithms for problems in social science. A final contribution will be advancing knowledge in sociolinguistics with the power of computation by identifying important questions that can only be answered through massive datasets and advanced computation.

---

[4]Citations are listed in publications section.
[5]A Global Perspective on Bilingualism and Bilingual Education, G. Richard Tucker. 1999.