An Interview with Dr. Huan Liu, Winner of ACM SIGKDD 2022 Innovation Award

ABSTRACT

In recognition of outstanding technical innovations that have greatly influenced the direction of research and development in the field of knowledge discovery and data mining, Huan Liu, Ph.D. is the recipient of the 2022 ACM SIGKDD Innovation Award. Conferred annually on an individual or group, Dr. Liu was honored for his outstanding contributions to the foundation, principles and applications of social media mining and feature selection for data mining. Dr. Liu is currently professor of computer science and engineering at Arizona State University (ASU). He sat down with SIGDKDD Explorations to discuss his career, his research, and the biggest challenges facing the field.

CONGRATS ON RECEIVING THE ACM SIGKDD INNOVATION AWARD. TELL US ABOUT YOURSELF, DR. LIU. WHERE ARE YOU FROM, WHERE DID YOU STUDY, AND HOW DID YOU BEGIN YOUR PROFESSIONAL CAREER?

I grew up in Shanghai and earned my undergraduate degree in computer science and electrical engineering at Shanghai Jiao Tong University. I then came to the United States to pursue my Ph.D. in computer science at the University of Southern California. My doctoral thesis – titled "Knowledge-Based Grasp Planning for Robot Hands" – was on the use of artificial intelligence to control a robot's hand, which paved the way for my future work on machine learning, data mining, and social media mining.

WHEN DID YOU DECIDE TO GO INTO THE FIELD OF KNOWLEDGE DISCOVERY AND DATA MINING?

My first job after my PhD study was at Australia Telecom Research Labs in Melbourne where my focus shifted from robots to telecommunication networks. At the research labs, we turned to machine learning to address the bottleneck problem in knowledge acquisition. Subsequently, when I started my academic career at National University of Singapore (NUS) in 1994, I started teaching Introduction to AI, Machine Learning. In collaboration with my colleague, I also taught first graduate course on "data mining and data warehousing" at NUS, when there were no textbooks yet on the subject. I was lucky that I found my way into data mining early by following my interests and with the help of my mentors and collaborators.

WHAT DOES THE RECOGNITION FROM SIGKDD MEAN TO YOU?

It means a lot to receive the Innovation Award from ACM SIGKDD. It is a prestigious award. If you look at past winners and their work, they made great impact and pushed the field forward. It is a real honor to be among them. Recognition is not the most important part of the job as a researcher and professor, but it does matter because it inspires me as well as my students to continue exploring and innovating. The Innovation Award is a recognition bestowed by the best experts in our field. It's a true honor for me and my team.

WHAT DOES INNOVATION MEAN TO YOU? HOW DOES IT IMPACT YOUR APPROACH TO YOUR WORK?

As an academic and researcher, one of my goals has been to identify an emerging area and establish a niche, putting myself in a position to pioneer new innovations, while also creating opportunities to nurture next generations of researchers and practitioners in our field. Early in my career, I was fortunate enough to have been at the forefront of feature selection, which led to my appointment in computer science and engineering at Arizona State University. Moving back to the United States and re-establishing myself was a tough challenge because of many differences between the US and Singapore. However, it also provided me an opportunity to establish new research programs that would allow my work to make unique impact. One of my early Ph.D. students at ASU proposed to conduct his research of data mining with social media in early 1990s. At the same time, I was fortunate to be exposed to network

science during a research visit. The confluence of different lines of research led to our new focus on social computing and social media mining. Ever since then, we've followed the data and tried to learn from and collaborate with experts from different disciplines, including social sciences, psychology, and information science, to investigate unprecedented research issues and develop novel algorithms. In the process, my students and I also developed the world first textbook on "Social Media Mining: An Introduction" published by Cambridge University Press in 2014. It is a widely used text all over the world.

YOU HAVE GRADUATED 34 Ph.D. STUDENTS AT ASU, 13 OF WHOM BECAME PROFESSORS. WHAT ADVICE DO YOU HAVE FOR DATA SCIENTISTS AT THE BEGINNING OF THEIR CAREERS?

I encourage my students early in their careers to focus on fundamental and practical problems by asking many questions. For example, feature selection is such a practical problem in data mining and machine learning; it still plays an important role in many real-world applications. Data mining is naturally a practical field, and students are encouraged to ask themselves about their research, "Why is this necessary?" We try to find simple solutions to tackle original problems. In our research, we want to identify fundamental problems with practical impact. We don't just train problem solvers; we train problem finders. When they become critical thinkers and ask interesting questions, students can unsurprisingly put themselves in a situation where their research in data mining will drive innovation. I am fortunate to have worked with many talented and outstanding students.

CAN YOU SHARE SOMETHING YOU ARE CURRENTLY WORKING ON THAT INTERESTS YOU?

Much of my work today is in collaboration with my students. There is still substantial focus on social media, as it remains the most popular medium for communication and expression of human behavior. Recent research I've been involved with includes causal learning with social media data, contrastive learning for labeled data scarcity, combating disinformation on social media, understanding cyberbullying, and applying machine learning and data mining in applications and education.

WHAT IS THE BIGGEST CHALLENGE FACING THE FIELD?

The biggest challenge in my view is attracting good students to conduct long-term research. The private sector pays so well that academia is losing great talent to industry, especially those born in the U.S. A bachelor's degree in computer science in some cases now suffices for those seeking good jobs in data mining and knowledge discovery in the enterprise sector, and as a result, we have few U.S. citizens pursuing doctoral degrees. It is important to strike a healthy balance between short-term development and long-term research for a sustainable and better future.