

Multi-objective Approaches to Fair Machine Learning

Xin Yao

Research Institute of Trustworthy Autonomous Systems (RITAS)

Department of Computer Science and Engineering

Southern University of Science and Technology (SUSTech), Shenzhen, China

and

School of Computer Science, University of Birmingham, UK

As the rapid development of artificial intelligence (AI) and its real-world applications in recent years, AI ethics has become increasingly important. It is no longer a nice feature to consider, but a must for both AI research and applications. First, this talk first tries to recall what classical ethics is about from an historical perspective. It tries to understand how technology ethics and AI ethics grow out of the broad ethics field. Specific features of AI ethics will be discussed. Second, a brief review of current research into AI ethics will be given. Key research topics will be extracted from a large number of reports to give a more concrete picture of most important issues covered in AI ethics. Third, we will examine the fairness issue in AI ethics and demonstrate how an algorithmic approach could help machine learning to be fairer. In other words, the results from machine learning will have less biases. Finally, some open research questions will be touched upon.

Biosketch of the speaker:

Xin Yao is a Chair Professor of Computer Science at the Southern University of Science and Technology, Shenzhen, China, and a part-time Professor of Computer Science at the University of Birmingham, UK. His major research interests include evolutionary computation, ensemble learning and search-based software engineering. More recently, he has been working on AI ethics, especially fairness. He is an IEEE fellow, a former (2014-15) president of IEEE Computational Intelligence Society (CIS) and a former (2003-08) Editor-in-Chief of IEEE Transactions on Evolutionary Computation. His research work won the 2001 IEEE Donald G. Fink Prize Paper Award, 2010, 2016 and 2017 IEEE Transactions on Evolutionary Computation Outstanding Paper Awards, 2010 BT Gordon Radley Award for Best Author of Innovation (Finalist), 2011 IEEE Transactions on Neural Networks Outstanding Paper Award, and many other best paper awards. He received a Royal Society Wolfson Research Merit Award in 2012, the IEEE CIS Evolutionary Computation Pioneer Award in 2013, and the 2020 IEEE Frank Rosenblatt Award.