

Data Science Seminar

Hosted by the Department of Mathematical Sciences

- Date: Tuesday, September 21, 2021
- Time: 12:00pm - 1:00pm
- Room: Via Zoom
- Speaker: Dr. Haoda Fu (Eli Lilly and Company)
- Title: Our Recent Development on Cost Constrained Machine Learning Models

Abstract

Suppose we can only pay $\$100$ to diagnose a disease subtype for selecting the best treatments. We can either measure 10 cheap biomarkers or 2 expensive ones. How can we pick the optimal combinations to achieve the highest diagnostic accuracy? This is a nontrivial problem. In a special case where each variable costs the same, the total cost constraint will be reduced to an L_0 penalty which is the best subset selection problem. Until recently, there is no good solution even for this special case. Traditional algorithms can only solve up to ~ 35 variables for best subset selections. Thanks to algorithm breakthroughs in the field of optimization research, we have modified and extended a recently developed algorithm to handle our cost constraint problems with thousands of variables. In this talk, we will introduce the background of this problem, methods development, and theoretical results. We will also show an impressive example of dynamic programming. It will tell a story on how algorithms can make a difference in computing. We hope that through this presentation, the audience can have a feel of modern statistics, which combines computer science, statistics, and algorithms.

Biography of the speaker: Dr. Fu is a Research Fellow and an Enterprise Lead for Machine Learning, Artificial Intelligence, and Digital Connected Care from Eli Lilly and Company. He is a Fellow of ASA (American Statistical Association). He is also an adjunct professor of biostatistics at the University of North Carolina Chapel Hill and Indiana University School of Medicine. Dr. Fu received his Ph.D. in statistics from the University of Wisconsin - Madison in 2007 and joined Lilly after that. Since he joined Lilly, he has been very active in statistics methodology research. His publications vary in Bayesian adaptive design, survival analysis, recurrent event modeling, personalized medicine, indirect and mixed treatment comparison, joint modeling, Bayesian decision making, and rare events analysis. In recent years, his research focuses on machine learning and artificial intelligence. His research has been published in various top journals including JASA, JRSS, Biometrika, Biometrics, ACM, IEEE, JAMA, Annals of Internal Medicine, etc. He has been teaching topics of machine learning, AI in large industry conferences, and FDA workshops. He was on the board of directors for statistics organizations and was the program chair and committee chair of ICOSA, ENAR, and the ASA Biopharm section.

This talk is endorsed by the [Data Science Transdisciplinary Area of Excellence at Binghamton University](#).

From:

<http://www2.math.binghamton.edu/> - **Department of Mathematics and Statistics,
Binghamton University**

Permanent link:

<http://www2.math.binghamton.edu/p/seminars/datasci/092121>



Last update: **2021/09/24 17:58**