USC Viterbi

School of Engineering
*Ming Hsieh Department
of Electrical Engineering*

Ming Hsieh Institute Seminar Series

Ming Hsieh Department of Electrical Engineering

Center for Systems and Control
CommNetS

A Marketplace for Data: An Algorithmic Solution

Munther Dahleh
Massachusetts Institute of Technology

Monday, November 12, 2018
2:00 – 3:00 PM
EEB 132

Abstract: In this work, we aim to create a data marketplace; a robust real-time matching mechanism to efficiently buy and sell training data for Machine Learning tasks. While the monetization of data and pre-trained models is an essential focus of industry today, there does not exist a market mechanism to price training data and match buyers to vendors while still addressing the associated (computational and other) complexity. The challenge in creating such a market stems from the very nature of data as an asset: (i) it is freely replicable; (ii) its value is inherently combinatorial due to correlation with signal in other data; (iii) prediction tasks and the value of accuracy vary widely; (iv) usefulness of training data is difficult to verify a priori without first applying it to a prediction task. As our main contributions we: (i) propose a mathematical model for a two-sided data market and formally define the key associated challenges; (ii) construct algorithms for such a market to function and rigorously prove how they meet the challenges defined. We highlight two technical contributions: (i) a new notion of "fairness" required for cooperative games with freely replicable goods; (ii) a truthful, zero regret mechanism for auctioning a particular class of combinatorial goods based on utilizing Myerson's payment function and the Multiplicative Weights algorithm. These might be of independent interest.

This is joint work with Anish Agarwal, Tuhin Sarkar, and Devavrat Shah.

Biography: Munther A. Dahleh received his PhD degree from Rice University, Houston, TX, in 1987 in Electrical and Computer Engineering. Since then, he has been with the Department of Electrical Engineering and Computer Science (EECS), MIT, Cambridge, MA, where he is now the William A. Coolidge Professor of EECS. He is also a faculty affiliate of the Sloan School of Management. He is the founding director of the newly formed MIT Institute for Data, Systems, and Society (IDSS). Previously, he held the positions of Associate Department Head of EECS, Acting Director of the Engineering Systems Division, and Acting Director of the Laboratory for Information and Decision Systems. He was a visiting Professor at the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA, for the Spring of 1993. He has consulted for various national research laboratories and companies. Dr. Dahleh is interested in Networked Systems with applications to Social and Economic Networks, financial networks, Transportation Networks, Neural Networks, and the Power Grid. Specifically, he focuses on the development of foundational theory necessary to understand, monitor, and control systemic risk in interconnected systems. His work draws from various fields including game theory, optimal control, distributed optimization, information theory, and distributed learning. His collaborations include faculty from all five schools at MIT. Dr. Dahleh is the co-author (with Ignacio Diaz-Bobillo) of the book Control of Uncertain Systems: A Linear Programming Approach, published by Prentice-Hall, and the co-author (with Nicola Elia) of the book Computational Methods for Controller Design, published by Springer. He is four-time recipient of the George Axelby outstanding paper award for best paper in IEEE Transactions on Automatic Control. He is also the recipient of the Donald P. Eckman award from the American Control Council in 1993 for the best control engineer under 35. He is a fellow of IEEE and IFAC. He has given many keynote lectures at major conferences.