NESS | New England Statistical Society

NEW ENGLAND STATISTICS SYMPOSIUM *SINCE 1987* & NEW ENGLAND STATISTICAL SOCIETY *SINCE 2017*

The NESS Education Committee presents the first
NESS Colloquium Special Series on Ethics in Data Science

Amit Dhurandhar, Research Staff Member
AI Foundations, IBM TJ Watson Research Center

Towards AI Governance with Trustworthy AI

Abstract: The significant advances in AI have led to its usage in many important decision-making scenarios, such as credit scoring, criminal justice, and job recruiting. Accompanying these trends is the need to ensure trust in AI systems. This talk will provide an overview of some of the key components of Trusted AI: fairness, explainability, robustness, transparency, and governance. We will provide real examples that illustrate the importance of these components and describe approaches that IBM Research is taking to address these problems, such as the release of several open-source toolkits.

Speaker Bio: Amit has always been interested in understanding things, even before working on explainable AI his research focused on understanding different AI methods in terms of their statistical behavior. He has worked on projects spanning multiple industries such as Semi-conductor manufacturing, Oil and Gas, Procurement, Retail, Utilities, Airline, Health Care. His current research includes proposing various methods for enhancing trust in systems by developing methods that try to explain or understand their behaviors. His recent work was featured in Forbes and PC magazine with corresponding technical contribution in leading AI research venues such as NeurIPS. His work has helped uncover interesting insights in fields such as Olfaction with papers in reputed journals such as Science and Nature Communications with extensive media coverage (Quartz, New Yorker, Atlantic, Biological Scene, Science News). His research also has received the AAAI deployed application award as well as being selected as Best of ICDM twice. For his work on explainability he was invited to attend a Schloss Dagstuhl seminar in 2019 and 2021. He also Co-led the creation of the AI Explainability 360 open-source toolkit. Besides research impact, his work has also gone into IBM product and his project work has received Outstanding Accomplishment as well as a nomination for Corporate award. He has been an Area Chair and PC member for top AI conferences as well as has served on National Science Foundation (NSF) panels for the small business innovative research (SBIR) program. He also serves on the invention disclosure committee (IDT) in IBM Research.

Date: Thursday, April 8, 2021
Time: 12:00 pm EST, 1 hr. duration
WebEx address for attendees:

https://uconn-cmr.webex.com/uconn-cmr/onstage/g.php?MTID=e4fe84c77fdf1a08ce0886821799288b3

Audio conference only
US Toll: +1-415-655-0002

Access code: 120 076 5184