

MICHIGAN STATE UNIVERSITY
Department of Statistics and Probability

COLLOQUIUM

Rebecca Nugent
Carnegie Mellon University

Before Teaching Data Science, Let's First Understand How People Do It

Tuesday, March 12, 2019
10:20 AM - 11:10 AM
Refreshments 10:00 AM
C405 Wells Hall

Abstract

In the last few years, the number of (primarily graduate) programs in Data Science has grown to the hundreds. Most of these programs were built on a foundation of already existing courses in several computing-oriented departments; less effort, understandably under the constraint of development efficiency, has been spent on understanding the integration of all of the necessary skills or how people from diverse backgrounds and disciplines approach or think about data science. The Department of Statistics & Data Science at Carnegie Mellon is inside the Dietrich College of Humanities and Social Sciences. In addition, our undergraduate program teaches about a third of the campus population every semester (Statistics, Math, Computer Science, Business, etc), so our sequences are taken by hundreds of students with incredibly diverse future degrees ranging from English Rhetoric to Chemistry to Statistics & Machine Learning. In the interest of characterizing how students with very diverse backgrounds approach or even think about Data Science, we have been developing ISLE (Interactive Statistics Learning Environment), an interactive platform that removes the computing cognitive load and lets students explore Statistics & Data Science concepts in both structured and unstructured ways. The platform also supports student-driven inquiry and case studies. We track every click, word used, and decision made (e.g., which graphs are designed/explored before settling on a final histogram) throughout the entire data analysis pipeline from loading the data to the final written report. Models of the students' online behavior and decisions also include performance metrics as well as what areas they're choosing to study. The platform is flexible enough to allow adaptation, providing different modes of data analysis instruction, active learning opportunities, and exercises for different subsets of the population – allowing us to research the science of data science while we teach it.

To request an interpreter or other accommodations for people with disabilities, please call the Department of Statistics and Probability at 517-355-9589.