Scalable statistical learning from dependent, high-dimensional, and structured data

Network, spatial, and temporal data are dependent, high-dimensional, and structured data, presenting challenges as well as opportunities for scalable statistical learning. My talk will highlight some of the challenges and some of the opportunities arising from structured data in high-dimensional settings, using large-scale data on air pollution and professional networks as examples. I will conclude with a classic and modern challenge educators face in the age of COVID-19: The problem of monitoring student progress based on structured data collected in online educational assessments. To monitor student progress, I will introduce visual learning progression maps that help assess how students progress towards learning targets and help detect disadvantaged students who need more, and different support than other students.