



JOHNS HOPKINS
BLOOMBERG
SCHOOL of PUBLIC HEALTH

Department of Biostatistics

BIostatISTICS SEMINAR

Musings on Statistical Models vs. Machine Learning in Health Research

Frank E. Harrell Jr. PhD

Professor of Biostatistics Founding Chair

Vanderbilt University School of Medicine

Department of Biostatistics

Abstract: Health researchers and practicing clinicians are with increasing frequency hearing about machine learning (ML) and artificial intelligence applications. They, along with many statisticians, are unsure of when to use traditional statistical models (SM) as opposed to ML to solve analytical problems related to diagnosis, prognosis, treatment selection, and health outcomes. And many advocates of ML do not know enough about SM to be able to appropriately compare performance of SM and ML. ML experts are particularly prone to not grasp the impact of the choice of measures of predictive performance. In this talk I attempt to define what makes ML distinct from SM, and to define the characteristics of applications for which ML is likely to offer advantages over SM, and vice-versa. The talk will also touch on the vast difference between prediction and classification and how this leads to many misunderstandings in the ML world. Other topics to be covered include the minimum sample size needed for ML, and problems ML algorithms have with absolute predictive accuracy (calibration).

Johns Hopkins Bloomberg School of Public Health, Department of Biostatistics

Monday, November 26, 2018, 12:15:1-15

Room W2008 (Refreshments 12:00pm)

Note: Taking photos during the seminar is prohibited

For disability access information or listening devices, please contact the Office of Support Services at 410-955-1197 or on the web www.jhsph.edu/SupportServices. EO/AA

Department of Biostatistics, 615 N. Wolfe Street, Suite E3527,
Baltimore, MD 21205