EPSTEIN INSTITUTE SEMINAR ▪ ISE 651

Interpreting Black-Box Supervised Learning Models Via Accumulated Local Effects

ABSTRACT - For many supervised learning applications, understanding and visualizing the effects of the predictor variables on the predicted response is of paramount importance. A shortcoming of black-box supervised learning models (e.g., complex trees, neural networks, boosted trees, random forests, nearest neighbors, local kernel-weighted methods, support vector regression, etc.) in this regard is their lack of interpretability or transparency. Partial dependence (PD) plots, which are the most popular general approach for visualizing the effects of the predictors with black box supervised learning models, can produce erroneous results if the predictors are strongly correlated, because they require extrapolation of the response at predictor values that are far outside the multivariate envelope of the training data. Functional ANOVA for correlated inputs can avoid this extrapolation but involves prohibitive computational expense and subjective choice of additive surrogate model to fit to the supervised learning model. We present a new visualization approach that we term accumulated local effects (ALE) plots, which have a number of advantages over existing methods. First, ALE plots do not require unreliable extrapolation with correlated predictors. Second, they are orders of magnitude less computationally expensive than PD plots, and many orders of magnitude less expensive than functional ANOVA. Third, they yield convenient variable importance/sensitivity measures that possess a number of desirable properties for quantifying the impact of each predictor.

Dr. Daniel Apley
Professor
Dept. of Industrial Engineering & Management Sciences,
Northwestern University

SPEAKER BIO – Dan Apley is a Professor of Industrial Engineering & Management Sciences at Northwestern University. His research and teaching interests are at the interface of engineering modeling, statistical analysis, and predictive analytics, with particular emphasis on improving operations of complex manufacturing and other enterprise systems. He received the NSF CAREER award, the IIE Transactions Best Paper Award, the Technometrics Wilcoxon Prize, and the Journal of Quality Technology Lloyd S. Nelson award. He has served as Editor-in-Chief of Technometrics and of the Journal of Quality Technology. He is a Fellow of the American Statistical Association and served as Chair of the Quality Statistics & Reliability Section of INFORMS and Director of the Manufacturing and Design Engineering Program at Northwestern.

USC Viterbi
School of Engineering
Daniel J. Epstein Department of
Industrial and Systems Engineering

TUESDAY, APRIL 12, 2022
3:30 PM – 4:50 PM
ZOOM/ONLINE *PLEASE EMAIL OWH@USC.EDU FOR PASSWORD*