



GDRR Program Opening Workshop August 5-9, 2019

SPEAKER TITLES/ABSTRACT

Bo Li

University of Illinois

“Gradient Boosting Trees for Spatial Data Prediction”

In recent years, statistical machine learning approaches have been extremely popular largely due to its superior performance in prediction. Of all the commonly used machine learning tools, the gradient boosting tree is usually the favored vehicle for many practitioners. On the popular data analytics competition platform Kaggle, gradient boosting is the winning algorithm for almost every structured data. Besides its superior prediction performance, the gradient boosting trees also enjoy the interpretability of a non-parametric additive model and its fitting algorithm can be paralleled. In this project, we extend this powerful machine learning technique to the realm of spatial data analysis. The proposed approach does not require any parametric assumption on spatial correlations and enjoy all the advantages of gradient boosting. We illustrate the potential of the data with application on prediction of HIV new diagnose rates for all counties of the United States.