Paper 341-2010

Tuning Bootstrap Forest and Gradient Boosting Fits

John Sall, SAS

ABSTRACT

It can make a big difference how you choose the tuning parameters for tree-based fits, and it turns out that advice that is commonly given is not confirmed in the experiments we did. The Bootstrap Forest fitter has two major tuning parameters: the number of trees, and the number of terms to randomly select at each split. Gradient Boosting has three major tuning parameters: the number of splits per tree, the learning rate, and the number of layers. Choosing bad tuning parameters can lead to bad fits, even negative holdback R-Squares.

No paper was submitted for publication in the *Proceedings*. Check http://support.sas.com/rnd/papers/ or contact the author.

CONTACT INFORMATION

John Sall
SAS
john.sall@sas.com