

STANFORD UNIVERSITY
DEPARTMENT OF STATISTICS
DEPARTMENTAL SEMINAR

4:15 p.m., Tuesday, October 19, 1999
Sequoia Hall Rm. 200
(Cookies at 3:45 in 1st Floor Lounge)

*Jerome H. Friedman**
Stanford University

GRADIENT BOOSTING and MULTIPLE ADDITIVE REGRESSION TREES

Abstract

Multivariate function estimation is viewed from the perspective of numerical optimization in function space, rather than parameter space. A gradient-descent "boosting" paradigm is established for rapid fitting of additive basis function expansions, based on any differentiable loss criterion. Focus is on the particular case where each of the basis functions is a (CART) regression tree. Specialized tools are presented for interpreting and visualizing such multiple additive regression tree models. The goal is a class of fast procedures for regression and classification that are competitive in accuracy, while being interpretable, and highly robust especially when mining less than clean data. Connections with the boosting methods of Freund and Schapire, and Friedman, Hastie and Tibshirani are discussed.

* with help from Trevor Hastie and Rob Tibshirani.