

STANFORD UNIVERSITY
DEPARTMENT OF STATISTICS
DEPARTMENT SEMINAR

4:15 p.m., Tuesday, January 29, 2002
Sequoia Hall Room 200
(Cookies at 3:45 in 1st Floor Lounge)

Hugh Chipman
University of Waterloo

Treed Generalized Linear Models

Tree models can be an effective and interpretable tool for supervised learning problems (i.e., regression and classification). A recent variation on trees is the "treed model", which includes a more sophisticated model in each terminal node of the tree, such as a linear regression. This talk considers generalized linear models as a broader class of terminal node models. Specific examples include binary and Poisson regression. A Bayesian approach to this problem offers several advantages, including regularization through careful specification of prior distributions, a stochastic search in the tree space, and the potential to improve predictions by model averaging. Data mining applications in areas such as marketing, insurance, and drug discovery will be discussed.