

STANFORD UNIVERSITY
DEPARTMENT OF STATISTICS
DEPARTMENTAL SEMINAR

4:15 p.m., Tuesday, October 22, 2002
Sequoia Hall Room 200
(Cookies at 3:45 in 1st Floor Lounge)

Saharon Rosset
Department of Statistics
Stanford University

Boosting as a Regularized Path to a Maximum Margin Classifier

Boosting is a method for incrementally fitting an additive model given a loss function and a basis or dictionary of "weak learners". The idea originated in the classification literature and the success of the resulting "Adaboost" algorithm (Freund and Schapire 95) has been analyzed from two distinct perspectives: the "gradient descent" view (advocated by Friedman, Hastie and Tibshirani (00)) and the "margin maximizing" approach, popular in the machine learning literature.

In this paper we build a unifying framework for both views, illustrating that gradient-based 2-class boosting approximately (and sometimes exactly) converges to a "margin maximizing" linear separator. This property holds both for the exponential loss criterion of Adaboost and for the logistic log-likelihood criterion of LogitBoost.

We follow (Efron et al 02) to show that the path of models traced by an idealized boosting algorithm follows the path of L_1 -constrained optimal solutions to the loss criterion. We prove that the idealized path converges to a "margin maximizing" separator as the constraint (or regularizer) is relaxed. This regularization effect justifies "early stopping" in boosting and illustrates that margin maximization does not necessarily correspond to improved generalization performance. It also gives us a direct analogy between boosting and support vector machines - both find an "optimal separator" in their non-regularized form, but they differ in the distance measure defining optimality and in the effect of regularization.

This is joint work with Ji Zhu and Trevor Hastie