

Benefits of Loss-Minimization View

- immediate generalization to other loss functions and learning problems
 - e.g. squared error for regression
 - e.g. logistic regression
(by only changing one line of AdaBoost)
 - sensible approach for converting output of boosting into conditional probability estimates
 - helpful connection to regularization
 - basis for proving AdaBoost is statistically “consistent”
 - i.e., under right assumptions, converges to best possible classifier
- [Bartlett & Traskin]