

AdaBoost

**Freund and Schapire:
Experiments with a new Boosting Algorithm**

**Schapire and Singer:
Improved Boosting Algorithms using Confidence-rated
Predictions**

Ian Fasel

October 23, 2001

Resampling for Classifier Design

- Arcing – “adaptive reweighting and combining”
 - reusing or selecting data in order to improve classification
- Two most popular:
 - Bagging (Breiman, 1994)
 - AdaBoost (Freund and Schapire, 1996)
- Combine the results of multiple “weak” classifiers into a single “strong” classifier.

The horse-track problem

Gambler wants to write a computer program to accurately predict winner of horse races.

- Gathers Historical horse-race data.
 - input vector:
 - * Odds
 - * dry or muddy
 - * jockey
 - output label:
 - * win or lose
- Discovers:
 - easy to find rules of thumb that are "often" correct.
 - hard to find a single rule that is very highly accurate.