

AdaBoost and Loss Minimization

- many (most?) learning and statistical methods can be viewed as minimizing **loss** (a.k.a. cost or objective) function measuring **fit to data**:
 - e.g. least squares regression $\sum_i (F(x_i) - y_i)^2$
- AdaBoost also minimizes a loss function
- **helpful** to understand because:
 - clarifies goal of algorithm and useful in proving convergence properties
 - decoupling of algorithm from its objective means:
 - faster algorithms possible for same objective
 - same algorithm may generalize for new learning challenges