

A Note of Caution

- tempting (but **incorrect!**) to conclude:
 - AdaBoost is **just** an algorithm for minimizing exponential loss
 - AdaBoost works **only because** of its loss function
 - ∴ more powerful optimization techniques for same loss should work even better
- **incorrect** because:
 - **other** algorithms that minimize exponential loss can give very **poor** generalization performance compared to AdaBoost
- for example...