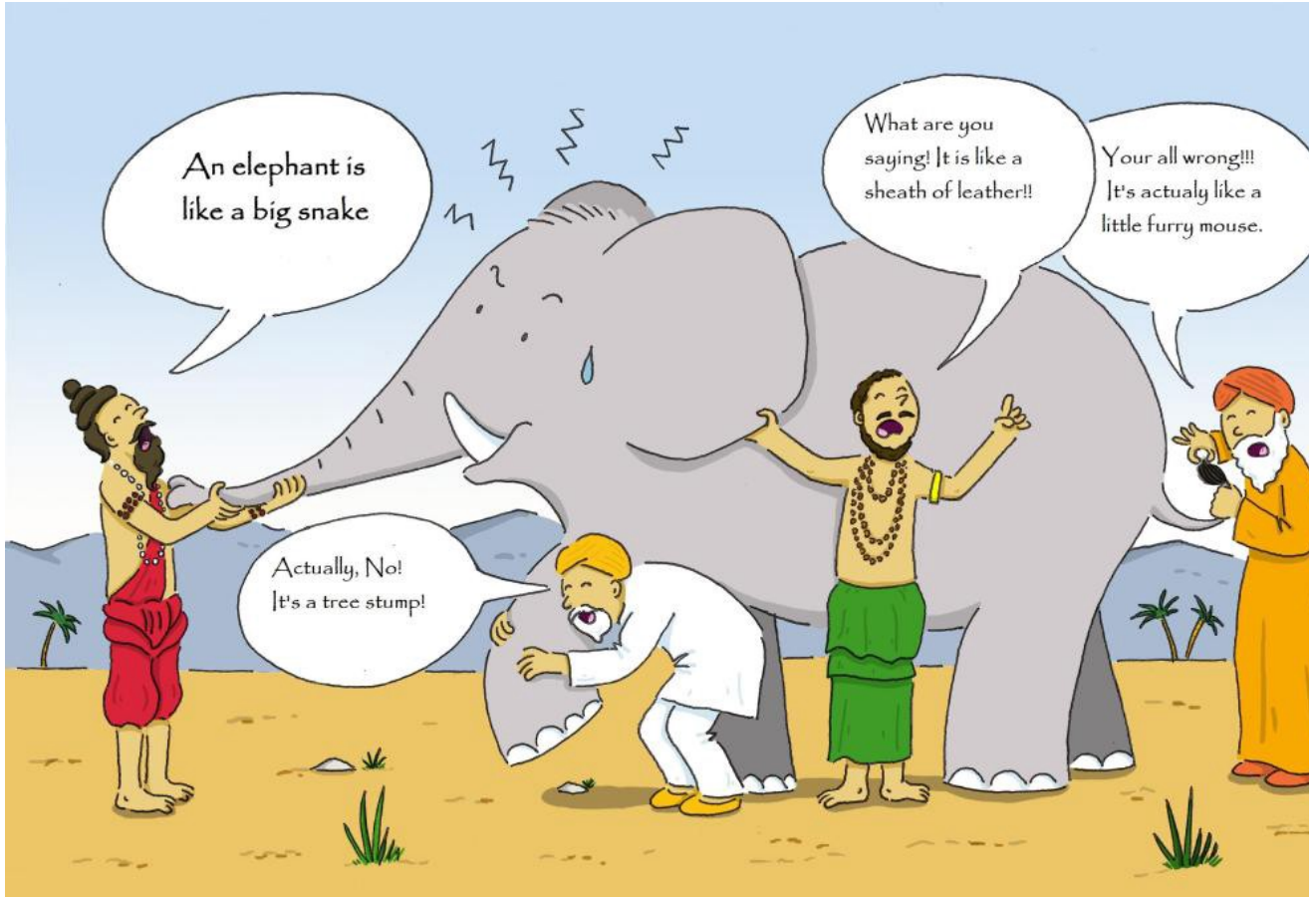
Boosting

Josh Feldman

**You have permission to use these slides

Intuition

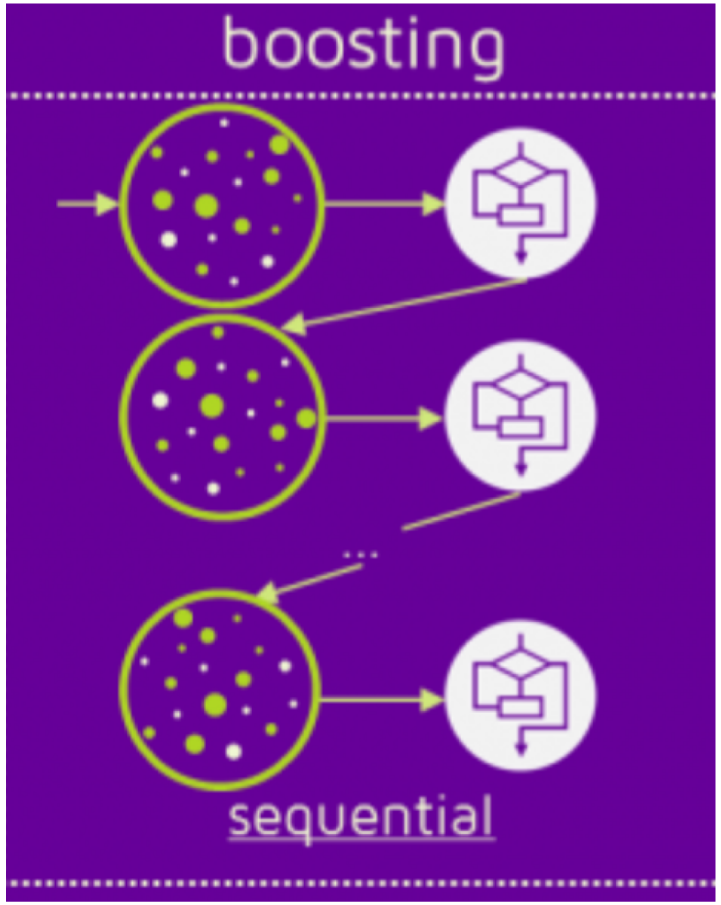
- Often a group of experts is smarter than a single genius, even if the each individual expert isn't as smart as the genius
- In boosting we train our experts sequentially, with each one learning from the mistakes of the previous attempts
- We actually don't want each individual expert to be too "smart", because otherwise our experts will just memorize our examples and not learn any general truths



More Formal Definition

- Suppose we have a set of weak H weak classifiers $\{T_h\}$
- We train each $T_1, T_2, \ldots, T_H$ one at a time, by taking the mistakes of $T_1 \ldots T_{i-1}$ into consideration when training T_n
- We can weight the weak classifiers to make a strong classifier

$$T = \sum_{h \in H} \lambda_h T_h$$



How do we find models T_h and weights λ_h?

There are many boosting algorithms, but we will focus the adaptive boosting classifier or the **adaboost classifier.**

<u>AdaBoost Algorithm</u>

1. Initialize weights W for all of your samples to 1
2. Repeat the following steps for a specified number of iterations:
   a. Fit T_h to your data, weighting your samples by W
   b. Calculate the accuracy a_h of model T_h
   c. Increase the weights in W for samples that were classified incorrectly
3. Let $T = \sum_{h \in H} \lambda_h T_h$ where $\lambda_h = \frac{1}{2} \ln\left(\frac{1-a_h}{a_h}\right)$