On Boosting Sparse Parities

Lev Reyzin

UIC

@ AAAI 2014
Boosting

weak learning
achieve some error
$\varepsilon < \frac{1}{2}$ on all
distributions

boosting

strong learning
achieve any error
$\varepsilon > 0$ on all
distributions

[Schapire ’90]
[Freund-Schapire ’97]
Boosting

**weak learning**
achieve some error
$\epsilon < \frac{1}{2}$ on all
distributions

[Schapire '90]

[Freund-Schapire '97]

**strong learning**
achieve any error
$\epsilon > 0$ on all
distributions

lots of work here

Boosting algorithms:
AdaBoost, LogitBoost, SmoothBoost, LPBoost, BrownBoost...

Explanations:
Occam bounds, game theory, margins, compression,...
Boosting

**weak learning**
achieve some error $\varepsilon < \frac{1}{2}$ on all distributions

**boosting**

[Schapire ’90]  
[Freund-Schapire ’97]

**strong learning**
achieve any error $\varepsilon > 0$ on all distributions

Our work asks: how can we design better weak learners?

Lots of work here

Boosting algorithms:  
AdaBoost, LogitBoost, SmoothBoost, LPBoost, BrownBoost...

Explanations:  
Occam bounds, game theory, margins, compression,...
What do we want in a weak learner?

- simplicity
- coverage
- richness
- diversity
- error
- optimizability
- evaluability
What do we want in a weak learner?

sparse parity functions
turn out to be a reasonable candidate
(in part due to recent advances).

And experimentally, they are very competitive (see poster)!

Suggests further study is worthwhile.
Main Open Problems

Better formalize a what makes a good weak learner.

Find better weak learners!
I’m not suggesting parities are the answer, just the beginning.