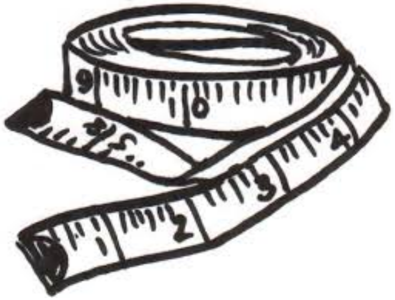


Weights and Measures: Prediction in the Era of Big Data

Lev Reyzin
UIC, Math Department

In a world of Big Data…

- Google/Yahoo! approaching 1B daily visitors.
- 3B yearly doctor’s visits in the US (records digitized).
- CERN currently storing > 100 petabytes of data.
- The NSA collects…. everything.

In a world of Big Data…

- Google/Yahoo! approaching 1B daily visitors.
  What ads to show?
- 3B yearly doctor's visits in the US (records digitized).
  How to treat them? (w/o performing too many tests.)
- CERN currently storing > 100 petabytes of data.
  Which of the events contain an interesting signal?
- The NSA collects…. everything.
  Who is a terrorist?

Some Opportunities and Challenges

Can handle rich / interesting classes of functions.

Should be really fast (linear time, faster?)

Want to leverage all the data.

It may not fit on one machine.

Can finally predict accurately.

Must predict quickly! (Or otherwise limited.)

Weights and Measures

- Weighing – keep track of "performance" without needing to remember the history.

- Measuring – given the weights, how to properly "measure" them to determine the correct outcome?

Optimal and Efficient Contextual Bandits

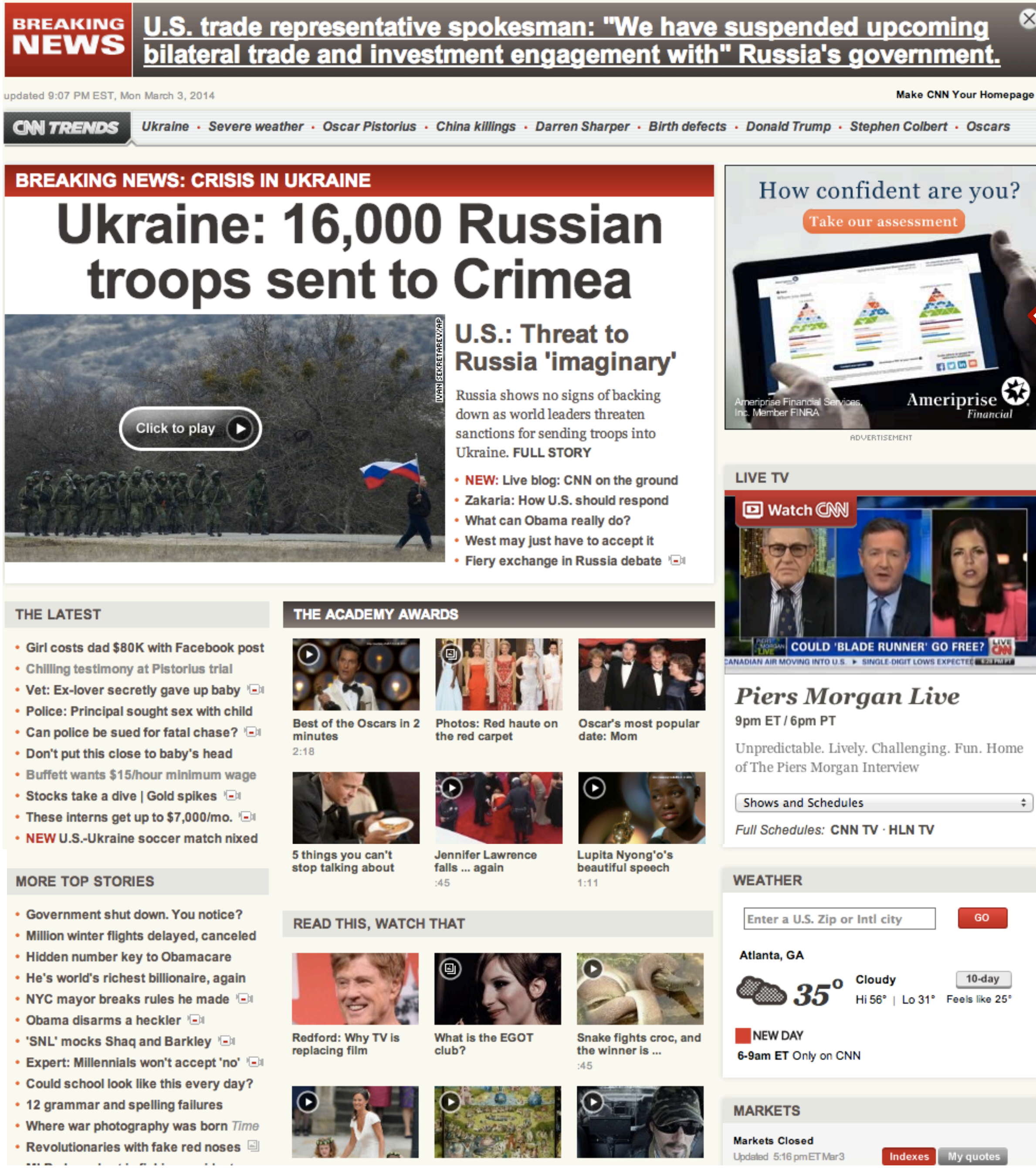
“the world is a bandit problem”

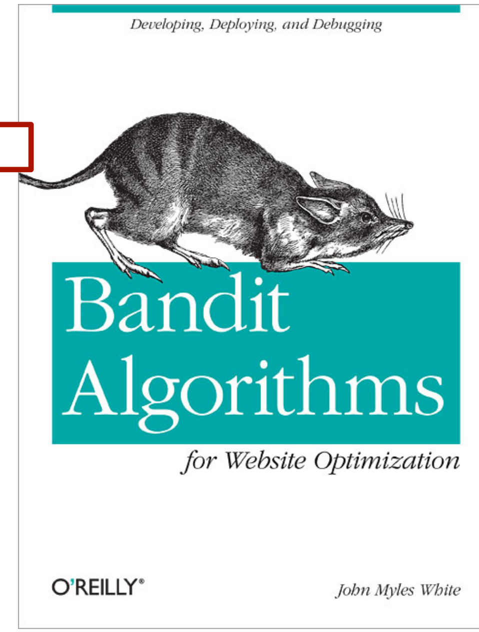
with

Alina Beygelzimer, Miro Dudik, Daniel Hsu, Satyen Kale, Nikos Karampatziakis, John Langford, Lihong Li, Rob Schapire, Tong Zhang

What is the Bandit Setting?

- The name bandit refers to slot machines in a casino.
- You choose actions (e.g. which machine to play), one after another. These actions come with rewards.
- Goal is to minimize your regret – informally, how well you did compared to how well you could have done.
  - Bound expected regret or bound regret w.h.p.


BREAKING NEWS
U.S. trade representative spokesman: "We have suspended upcoming bilateral trade and investment engagement with" Russia's government.
updated 9:07 PM EST, Mon March 3, 2014
Make CNN Your Homepage
CNN TRENDS
Ukraine · Severe weather · Oscar Pistorius · China killings · Darren Sharper · Birth defects · Donald Trump · Stephen Colbert · Oscars
BREAKING NEWS: CRISIS IN UKRAINE
Ukraine: 16,000 Russian troops sent to Crimea
Click to play
U.S.: Threat to Russia 'imaginary'
Russia shows no signs of backing down as world leaders threaten sanctions for sending troops into Ukraine. FULL STORY
NEW: Live blog: CNN on the ground
Zakaria: How U.S. should respond
What can Obama really do?
West may just have to accept it
Fiery exchange in Russia debate
How confident are you?
Take our assessment
Ameriprise Financial
Ameriprise Financial Services, Inc. Member FINRA
ADVERTISEMENT
LIVE TV
Watch CNN
COULD 'BLADE RUNNER' GO FREE?
Piers Morgan Live
9pm ET / 6pm PT
Unpredictable. Lively. Challenging. Fun. Home of The Piers Morgan Interview
Shows and Schedules
Full Schedules: CNN TV · HLN TV
THE LATEST
Girl costs dad $80K with Facebook post
Chilling testimony at Pistorius trial
Vet: Ex-lover secretly gave up baby
Police: Principal sought sex with child
Can police be sued for fatal chase?
Don't put this close to baby's head
Buffett wants $15/hour minimum wage
Stocks take a dive | Gold spikes
These interns get up to $7,000/mo.
NEW U.S.-Ukraine soccer match nixed
THE ACADEMY AWARDS
Best of the Oscars in 2 minutes 2:18
Photos: Red haute on the red carpet
Oscar's most popular date: Mom
5 things you can't stop talking about
Jennifer Lawrence falls ... again :45
Lupita Nyong'o's beautiful speech 1:11
MORE TOP STORIES
Government shut down. You notice?
Million winter flights delayed, canceled
Hidden number key to Obamacare
He's world's richest billionaire, again
NYC mayor breaks rules he made
Obama disarms a heckler
'SNL' mocks Shaq and Barkley
Expert: Millennials won't accept 'no'
Could school look like this every day?
12 grammar and spelling failures
Where war photography was born Time
Revolutionaries with fake red noses
READ THIS, WATCH THAT
Redford: Why TV is replacing film
What is the EGOT club?
Snake fights croc, and the winner is ... :45
WEATHER
Enter a U.S. Zip or Intl city
GO
Atlanta, GA
35°
Cloudy
Hi 56° | Lo 31°
10-day
Feels like 25°
NEW DAY
6-9am ET Only on CNN
MARKETS
Markets Closed
Updated 5:16 pm ET Mar3
Indexes
My quotes


Developing, Deploying, and Debugging
Bandit Algorithms
for Website Optimization
O'REILLY
John Myles White

9/21/2010

Regretting the dead

Tags: Exploration,Machine Learning,Online — jl@ 9:31 pm

Nikos pointed out this new york times article about poor clinical design killing people. For those of us who study learning from exploration information this is a reminder that low regret algorithms are particularly important, as regret in clinical trials is measured by patient deaths.

The New York Times

Research

TARGET CANCER

New Drugs Stir Debate on Rules of Clinical Trials

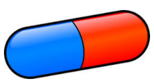
Monica Almeida/The New York Times, left

Two Cousins, Two Paths Thomas McLaughlin, left, was given a promising experimental drug to treat his lethal skin cancer in a medical trial; Brandon Ryan had to go without it.

Contextual Bandits [Auer et al. '02]

context:

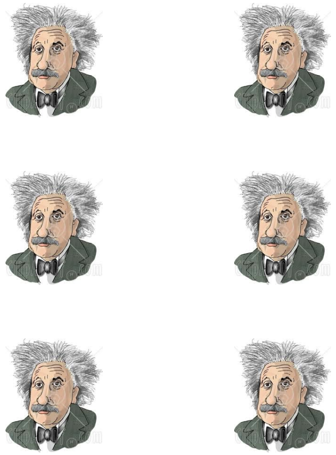
1 

2 

3 

⋮

k 

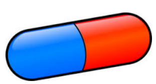


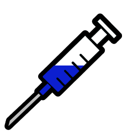
N experts/function
think of N >> K

Contextual Bandits [Auer et al. ’02]

context: x_1

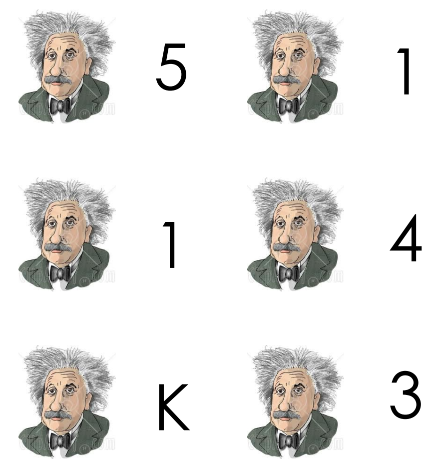
1 

2 

3 

⋮

k 

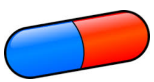


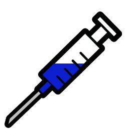
N experts/function
think of N >> K

Contextual Bandits [Auer et al. '02]

context: x_1

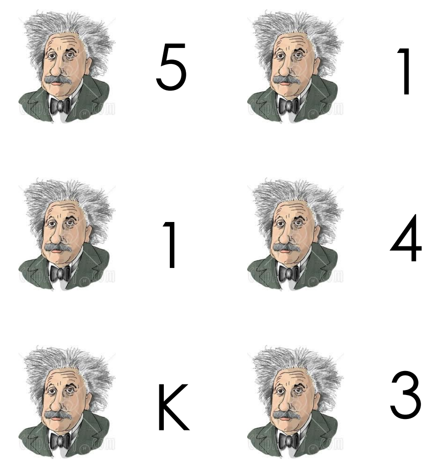
1  yes

2 

3 

⋮

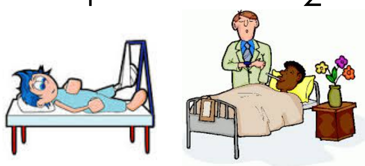
k 



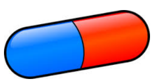
N experts/function
think of N >> K

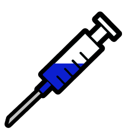
Contextual Bandits [Auer et al. '02]

context: x_1 x_2



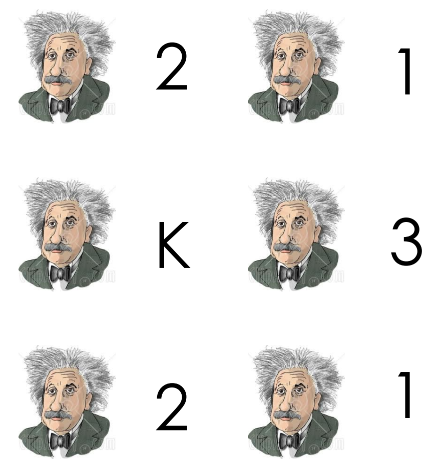
1  yes

2 

3  yes

⋮

k 



N experts/function
think of N >> K

Contextual Bandits [Auer et al. '02]

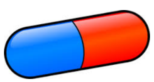
context: x_1 x_2 x_3 x_4 … x_T

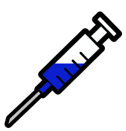





…

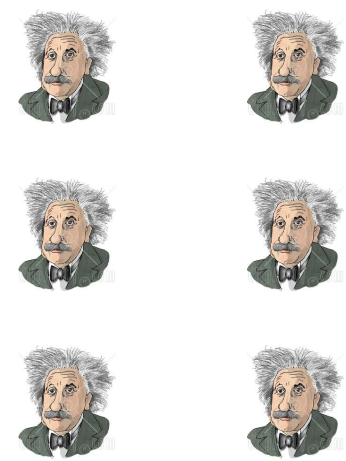
1  yes … yes

2 

3  yes no

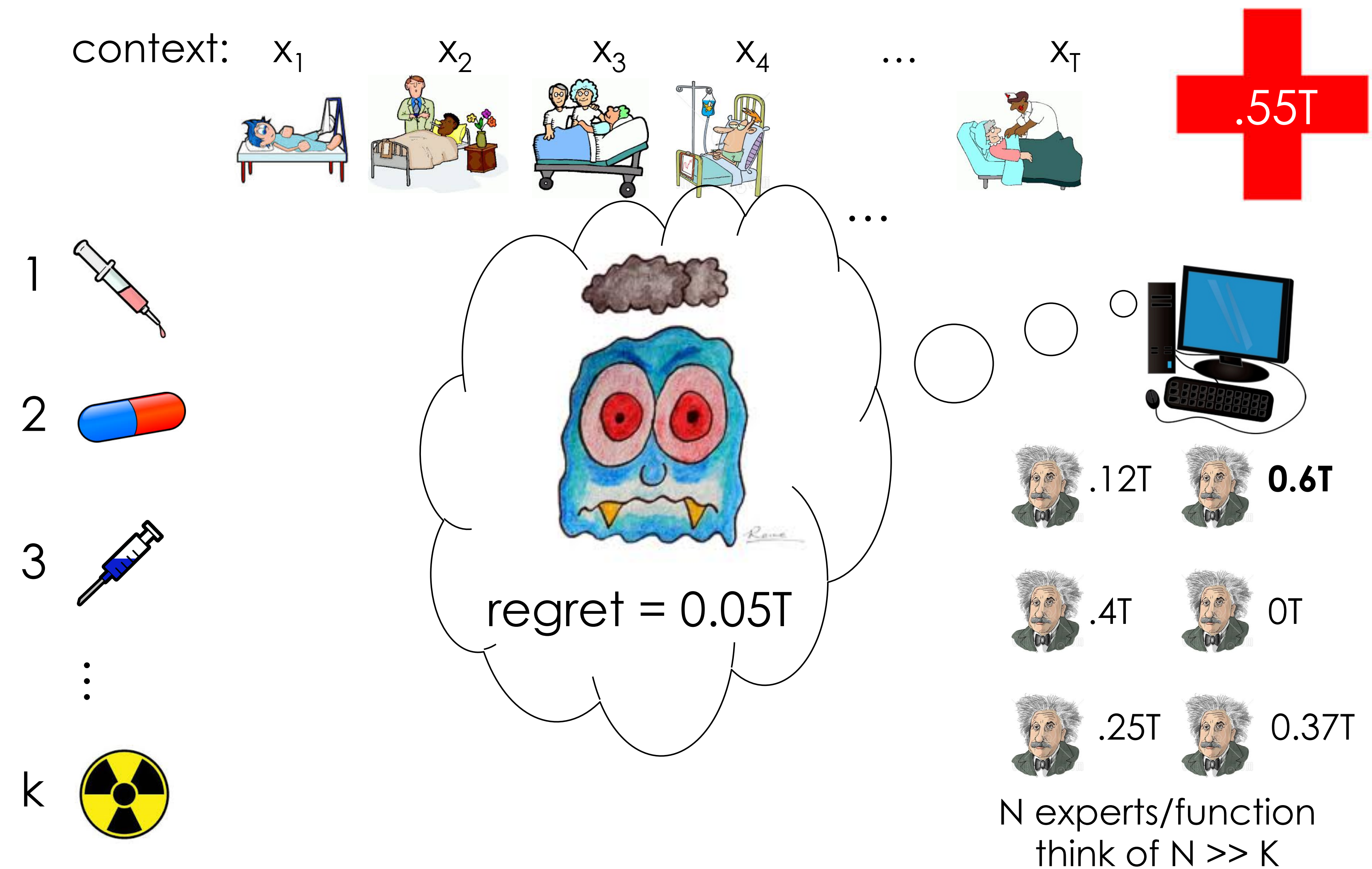
⋮

k  no



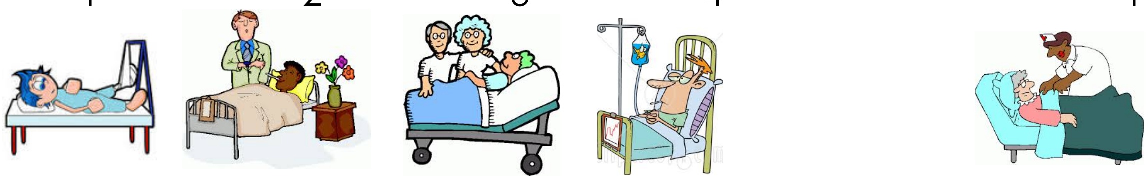
N experts/function
think of N >> K

Contextual Bandits [Auer et al. '02]



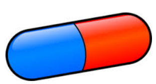
Contextual Bandits [Auer et al. '02]

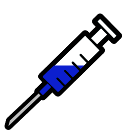
context: x_1 x_2 x_3 x_4 … x_T



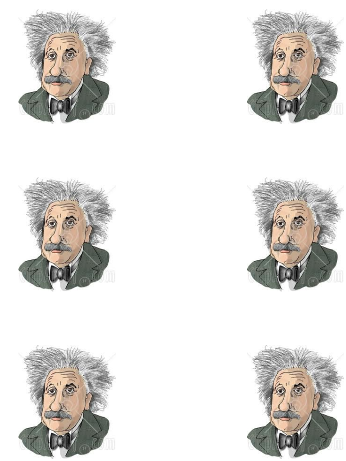
…

1 

2  ← the contexts & rewards can come from a distribution (stochastic) or be adversarial

3 

The experts can be present (contextual) or not. →



⋮

k 

- harder than supervised (usual) learning:

  In the bandit setting, we do not know the rewards of actions not taken.

- many applications:

  Medicine, ad auctions, finance, …

- exploration/exploitation dilemma:

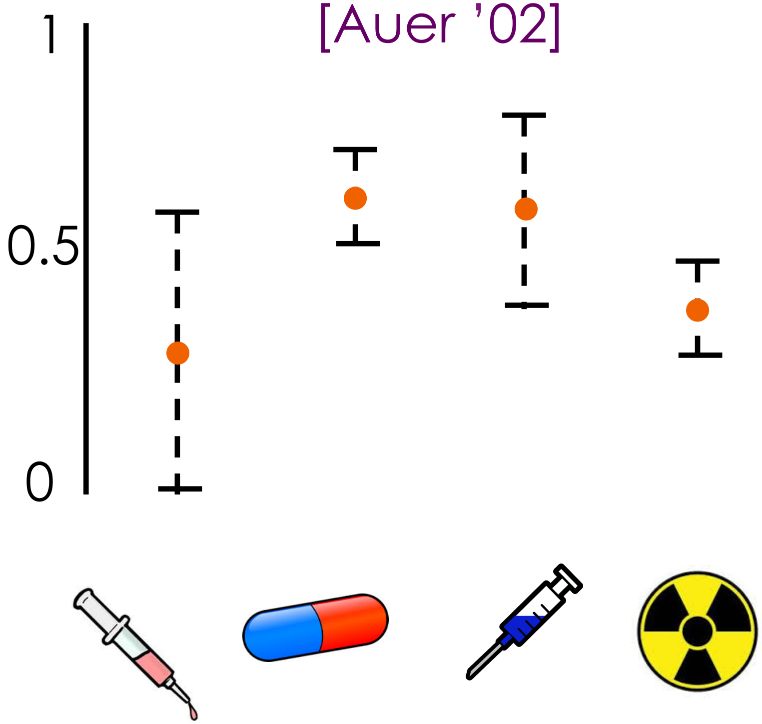
  exploit policies you've learned to be good?

  explore policies you're not sure about?

- $\Omega(TK \ln N)^{1/2}$ is a known lower bound on regret [Auer et al. '02] even in the stochastic setting.
- Algorithms achieving $\tilde{O}(KT \text{ polylog } N)^{1/2}$ regret are said to be optimal.
- greedily first exploring (acting randomly) then exploiting (following best policy) cannot be optimal. Optimal algorithms must be adaptive.

Upper Confidence Bounds

[Auer '02]

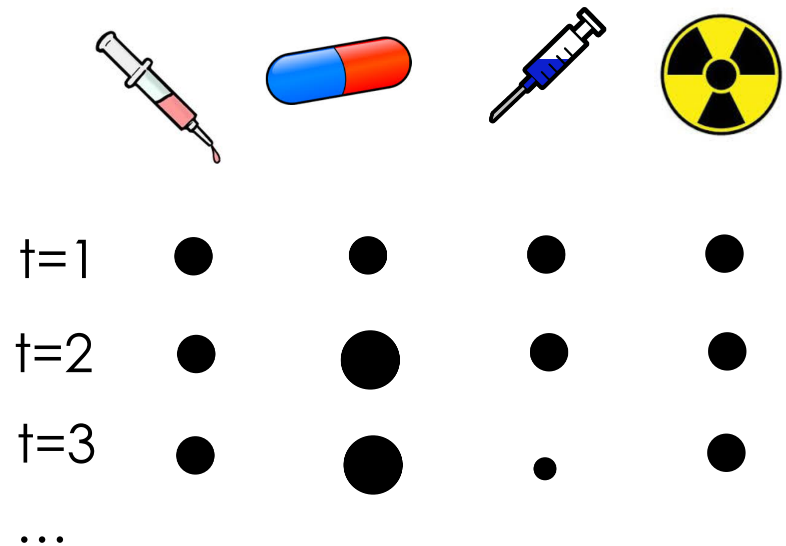


At every time step:

1) choose action with highest UCB
2) update confidence bound of the arm pulled.

EXP3 Exponential Weights

[Auer et al '02]



At every time step:

1) sample from distribution defined by weights (mixed w/ uniform)
2) "exponential" weight updates

UCB vs EXP3: A Comparison

UCB

[Auer '02]

- **Pros**

  Optimal in stochastic setting.

  Succeeds w.h.p.

- **Cons**

  Fails in adversarial setting.

  Not optimal in the contextual setting.

EXP3 & Friends

[Auer et al. '02]

- **Pros**

  Optimal for adversarial and stochastic settings.

  Adaptable to the contextual setting

- **Cons**

  Succeeds only in expectation.

Algorithm	Optimal?	High Prob?	Context?
Exp4 [ACFS '02]	Yes	No	Yes
epoch-geedy [LZ '07]	No	Yes	Yes
Exp3.P[ACFS '02] UCB [Auer '00]	Yes	Yes	No

Algorithm	Optimal?	High Prob?	Context?
Exp4 [ACFS ’02]	Yes	No	Yes
epoch-geedy [LZ ’07]	No	Yes	Yes
Exp3.P[ACFS ’02] UCB [Auer ’00]	Yes	Yes	No
Exp4.P [BLL**R**S ’11]	Yes	Yes	Yes

EXP4.P
[Beygelzimer-Langford-Li-R-Schapire '11]

Main Theorem [BLLRS '11]: For any $\delta > 0$, w.p. at least $1-\delta$, EXP4P has regret at most **$O(KT \ln (N/\delta))^{1/2}$** in adversarial contextual bandit setting.

Combines advantages of Exponential Weights and UCB.

- Optimal for both the stochastic and adversarial settings
- Works for contextual case (also non-contextual case)
- A high probability result

Ideas Behind Exp4.P

(all appeared in previous algorithms)

- exponential weights

  keep a weight on each expert that drops exponentially in the expert's (estimated) performance

- upper confidence bounds

  adds an upper confidence bound on each expert's estimated reward

- ensuring exploration

  make sure each action is taken with some minimum probability

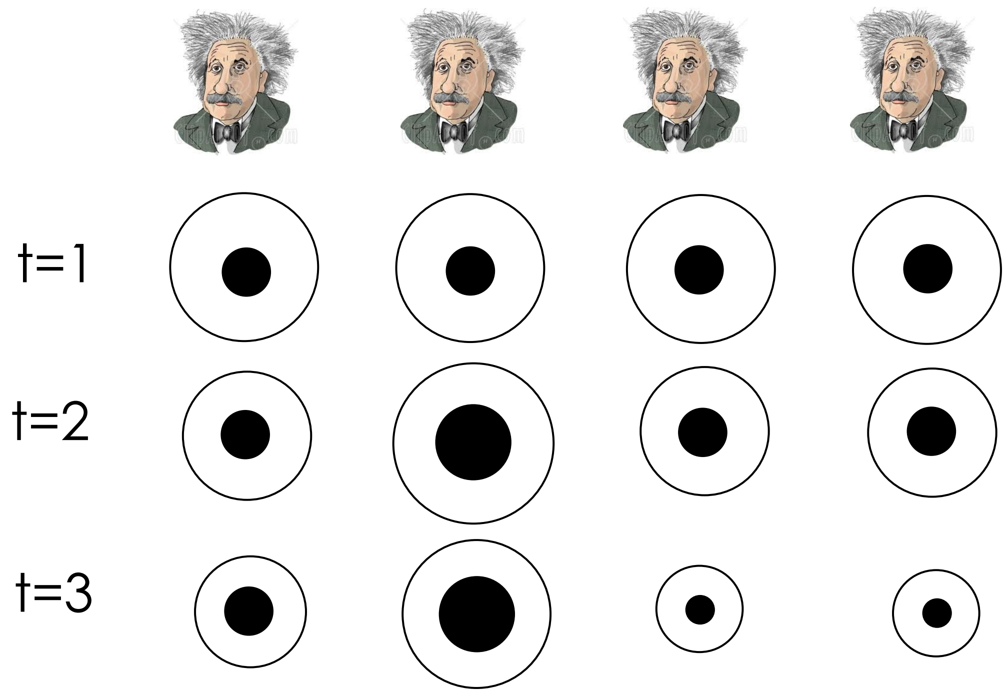
- importance sampling

  give rare events more importance to keep estimates unbiased

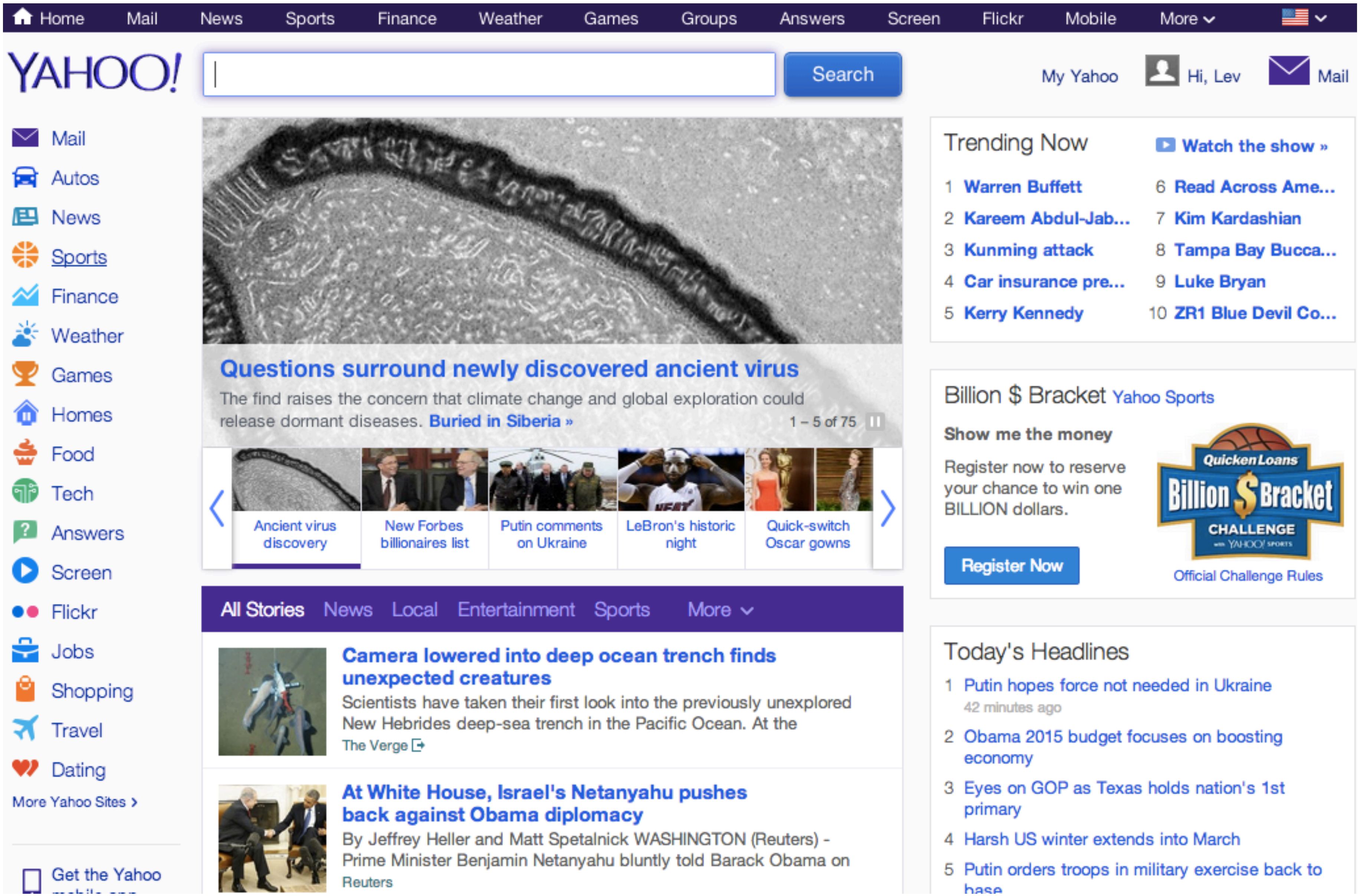
Key Insights

key insights
(on top of UCB/ EXP)

1) exponential weights and confidence bounds "stack"
2) generalized Bernstein's inequality for martingales



Algorithm	Optimal?	High Prob?	Context?	Efficient?
Exp4 [ACFS '02]	Yes	No	Yes	No
epoch-geedy [LZ '07]	No	Yes	Yes	Yes
Exp3.P [ACFS '02] UCB [A '00]	Yes	Yes	No	Yes
Exp4.P [BLL**R**S '10]	Yes	Yes	Yes	**No**


Home
Mail
News
Sports
Finance
Weather
Games
Groups
Answers
Screen
Flickr
Mobile
More
YAHOO!
Search
My Yahoo
Hi, Lev
Mail
Mail
Autos
News
Sports
Finance
Weather
Games
Homes
Food
Tech
Answers
Screen
Flickr
Jobs
Shopping
Travel
Dating
More Yahoo Sites ›
Get the Yahoo
Questions surround newly discovered ancient virus
The find raises the concern that climate change and global exploration could release dormant diseases. Buried in Siberia »
1 – 5 of 75
Ancient virus discovery
New Forbes billionaires list
Putin comments on Ukraine
LeBron's historic night
Quick-switch Oscar gowns
All Stories
News
Local
Entertainment
Sports
More
Camera lowered into deep ocean trench finds unexpected creatures
Scientists have taken their first look into the previously unexplored New Hebrides deep-sea trench in the Pacific Ocean. At the
The Verge
At White House, Israel's Netanyahu pushes back against Obama diplomacy
By Jeffrey Heller and Matt Spetalnick WASHINGTON (Reuters) - Prime Minister Benjamin Netanyahu bluntly told Barack Obama on
Reuters
Trending Now
Watch the show »
1 Warren Buffett
2 Kareem Abdul-Jab...
3 Kunming attack
4 Car insurance pre...
5 Kerry Kennedy
6 Read Across Ame...
7 Kim Kardashian
8 Tampa Bay Bucca...
9 Luke Bryan
10 ZR1 Blue Devil Co...
Billion $ Bracket Yahoo Sports
Show me the money
Register now to reserve your chance to win one BILLION dollars.
Register Now
Quicken Loans
Billion $ Bracket
CHALLENGE
Official Challenge Rules
Today's Headlines
1 Putin hopes force not needed in Ukraine
42 minutes ago
2 Obama 2015 budget focuses on boosting economy
3 Eyes on GOP as Texas holds nation's 1st primary
4 Harsh US winter extends into March
5 Putin orders troops in military exercise back to

Evaluation

- We chose a policy class for which we could efficiently keep track of the weights.
  - Created 5 clusters, using user features.
  - Policies mapped clusters to article choices.
  - Ran on personalized news article recommendations for Yahoo! front page.
- We used a learning bucket on which we ran the algorithms and a deployment bucket on which we ran the greedy (best) learned policy.

Experiments

Reported estimated (normalized) click-through rates on front page news. Over 41M user visits. 253 total articles. 21 candidate articles per visit.

	EXP4P	EXP4	ε-greedy
Learning eCTR	1.0525	1.0988	1.3829
Deployment eCTR	**1.6512**	1.5309	1.4290

Why does this work in practice?
[McMahan '11]

A General Efficient Algorithm? [DHKKL**R**Z '11]

EXP4.P's regret grows only logarithmically with N.

this suggests

We could compete with a huge set of policies!
(e.g. **N=K**^{100} becomes **10 log**$^{1/2}$ **K** in the regret)

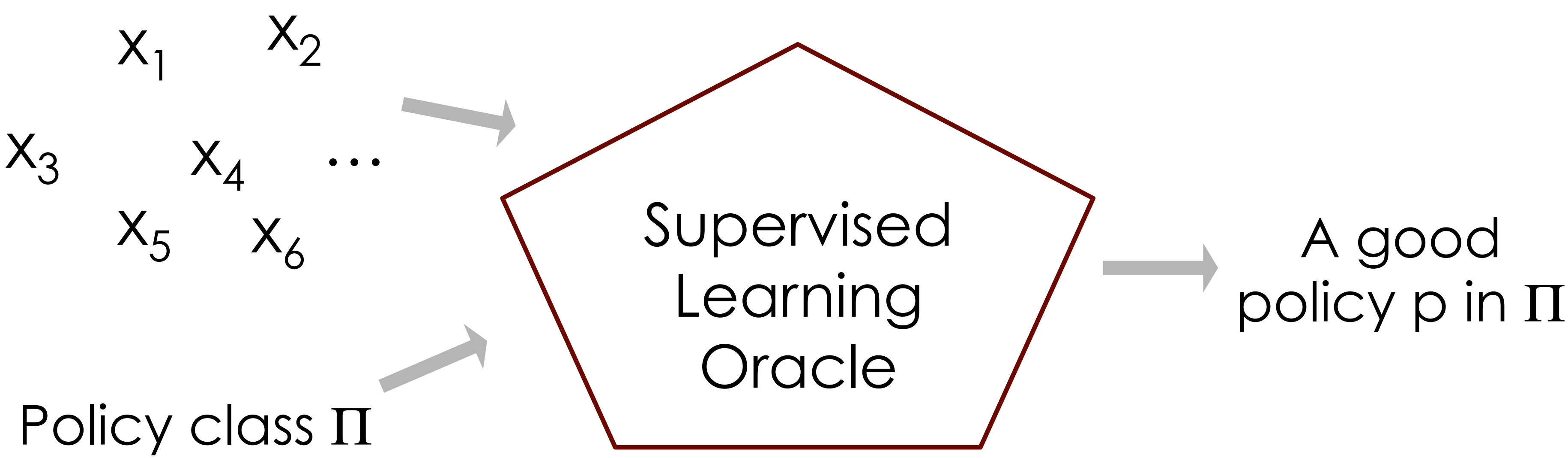
however

Exp4.P explicitly "keeps track" of all policies. Reading in all recommendations, for large N, would take too long.

Reduce to Supervised Learning!

(Idea from [Langford-Zhang '07])

- "Competing" with an exponentially large set of policies is commonplace in supervised learning.
- Recommendations of the policies/functions don't need to be explicitly read when the policy class has structure!

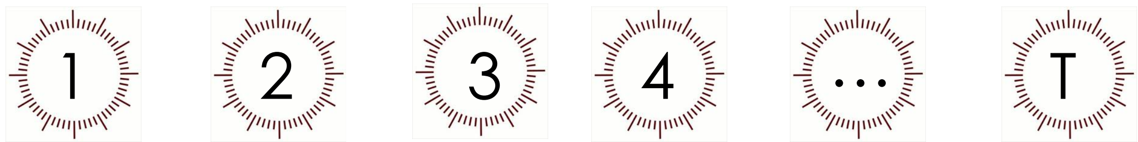


Reduce to Supervised Learning!

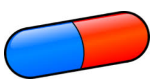
(Idea from [Langford-Zhang '07])

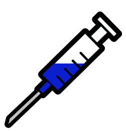
- "Competing" with an exponentially large set of policies is commonplace in supervised learning.
- Recommendations of the policies/functions don't need to be explicitly read when the policy class has structure!

context: x_1 x_2 x_3



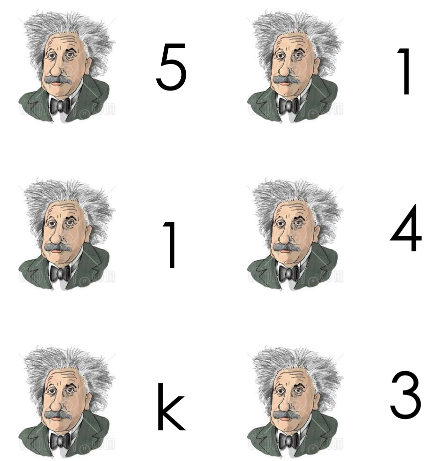
1  yes

2  no

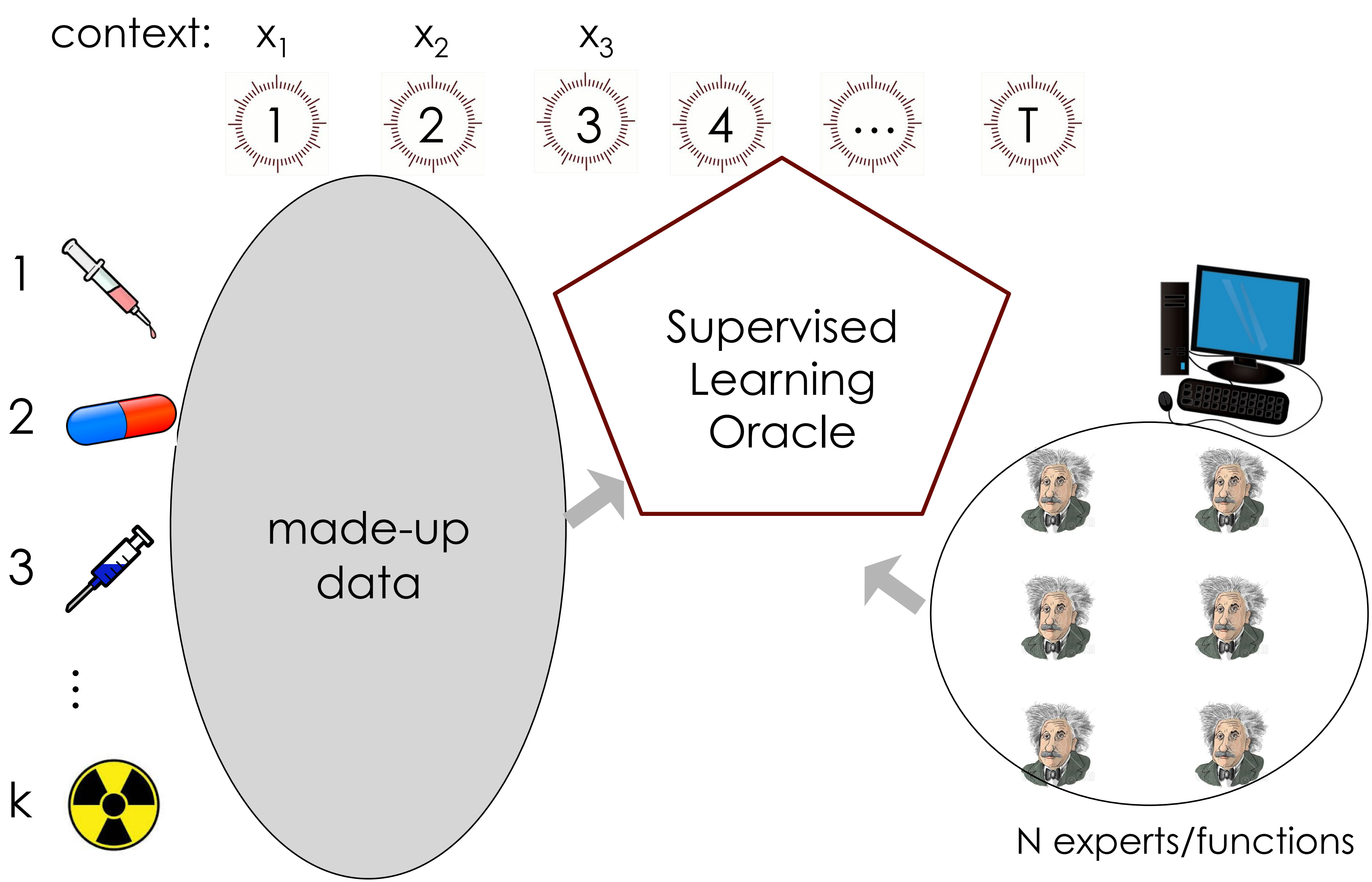
3 

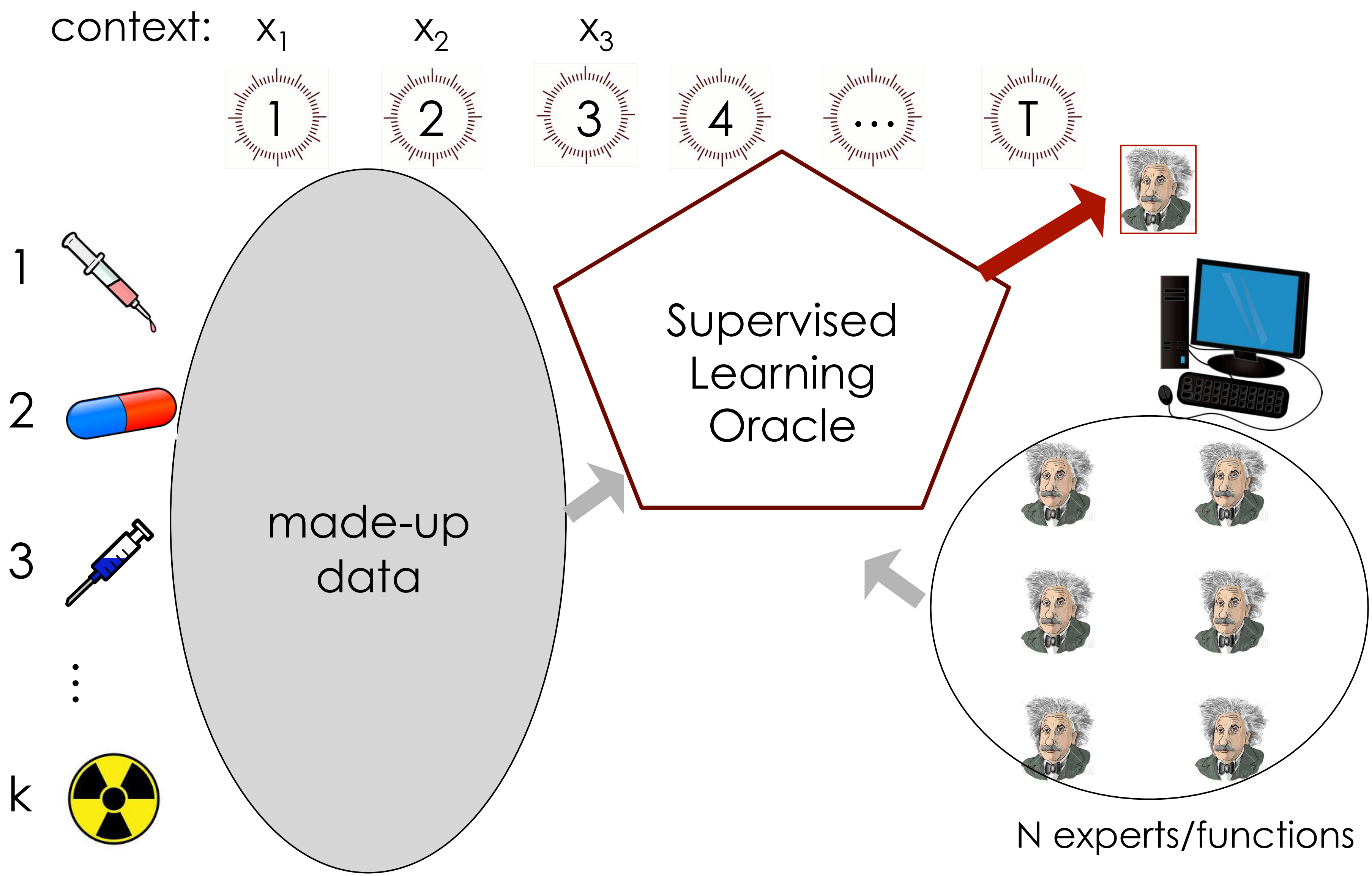
⋮

k 



N experts/functions


context: x_1 x_2 x_3
1
2
3
4
...
T
1
2
3
k
made-up data
Supervised Learning Oracle
N experts/functions


context: x_1 x_2 x_3
1
2
3
4
...
T
1
2
3
⋮
k
made-up data
Supervised Learning Oracle
N experts/functions

Thm: [Dudik-Hsu-Kale-Karampatziakis-Langford-R-Zhang '11]:
For any $\delta > 0$, w.p. at least $1-\delta$, given access to a learning oracle, R-UCB has regret $O((KT \ln (NT/\delta))^{1/2}$ in the stochastic contextual bandit setting and runs in time poly(K,T, ln N).

Thm: [Dudik-Hsu-Kale-Karampatziakis-Langford-R-Zhang '11]: For any $\delta > 0$, w.p. at least $1-\delta$, given access to a learning oracle, **R-UCB** has regret $O((KT \ln (NT/\delta))^{1/2})$ in the stochastic contextual bandit setting and runs in time **poly(K,T, ln N).**

if arms chosen among only "good" policies w/ variance < 2K, we win

can prove this exists via a minimax theorem

↓

this can be softened to occasionally allow choosing of bad policies

via "randomized" upper confidence bounds

↓

creates a problem of how to choose arms as to satisfy the constraints

expressed as convex optimization problem

↓

solvable by ellipsoid algorithm

can implement a separation oracle with the supervised learning oracle

Thm: [Dudik-Hsu-Ka … d-**R**-Zhang '11]:
For any δ>0 … learning
oracle, R-… hastic
conte…).

if arm … we win

thi … ies

creat … raints

can implement … d learning oracle

Big theoretical breakthrough!

But not practical.

(and needs stochastic assumption)

new on arXiv (2/14): Taming the Monster: A Fast and Simple Algorithm for Contextual Bandits
[Agarwal-et al. '14]

A research goal of mine: make it work in adversarial model.

Applying to Public Health

UIC SCHOOL OF PUBLIC HEALTH
UNIVERSITY OF ILLINOIS AT CHICAGO

Search
GO

About SPH | Academics | Admissions | Departments | Research

Research
- Announcements
- Seed Funding Competition
- Annual Research and Practice Awards Day
- New Awards
- Getting To Know UIC SPH Researchers
- Faculty Research Interests
- SPH Proposal Development
- Community Connections
- Current Sponsored Activities
- Centers Institutes & Programs
- Traineeships & Fellowships
- Contact Office of Research Services

Research / Centers Institutes & Programs /

Community Outreach Intervention Projects

ABOUT COIP

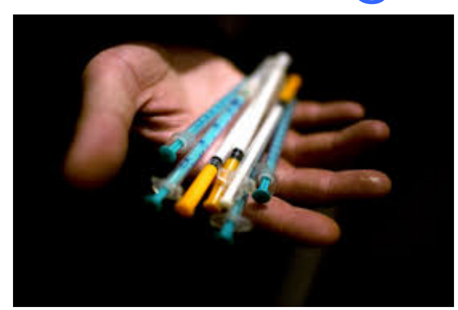
The Community Outreach Intervention Projects ("COIP"), School of Public Health, University of Illinois at Chicago was founded in 1986 to address HIV/AIDS, particularly among people who use drugs. COIP operates from storefront sites in Austin, Humboldt Park, West Englewood, South Chicago, and Uptown. Other neighborhoods are served by COIP's motorhome and mobile van units. COIP's interventions are known for their use of the Indigenous Leader Outreach Model, which employs former drug users to deliver services and assist in conducting research.

COIP's services include street outreach, counseling and testing for HIV, syphilis and other infectious diseases associated with substance use, case management for people who are HIV positive, syringe exchange, drug abuse and risk reduction counseling, support groups, educational activities, and a program that enhances linkages to care for HIV positive women exiting jail. COIP also makes many referrals to other providers such as drug treatment programs. Through a collaboration with UIC's Community Clinic Network, all COIP's storefront sites provide free medical, mental health and pharmacy care for people living with HIV, and one site offers free dental care.

COIP also conducts research to better understand HIV/AIDS in Chicago communities. Recent studies have examined or

Consortium for Modeling and Analysis of Treatments and Interventions

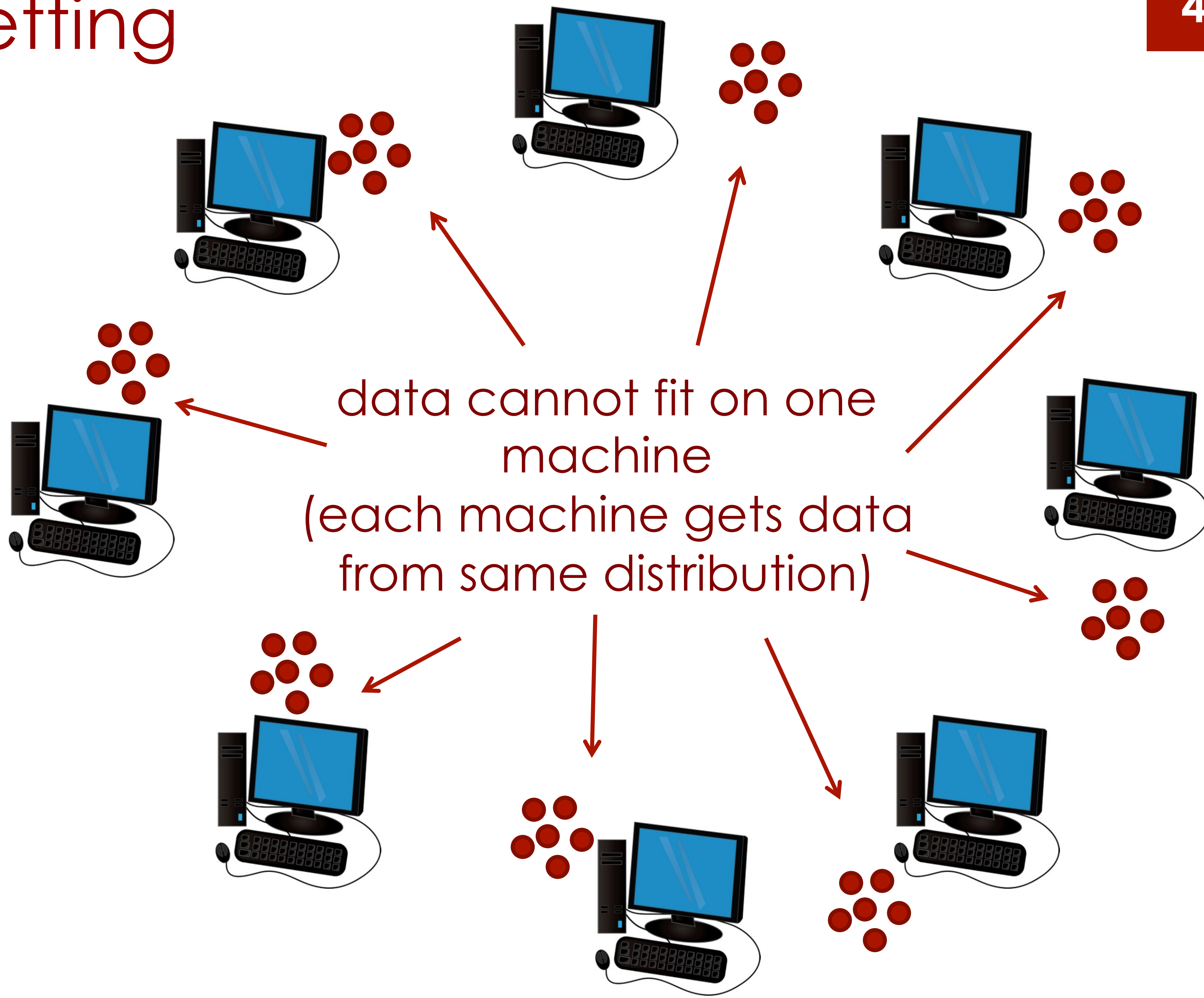
c-mati.org



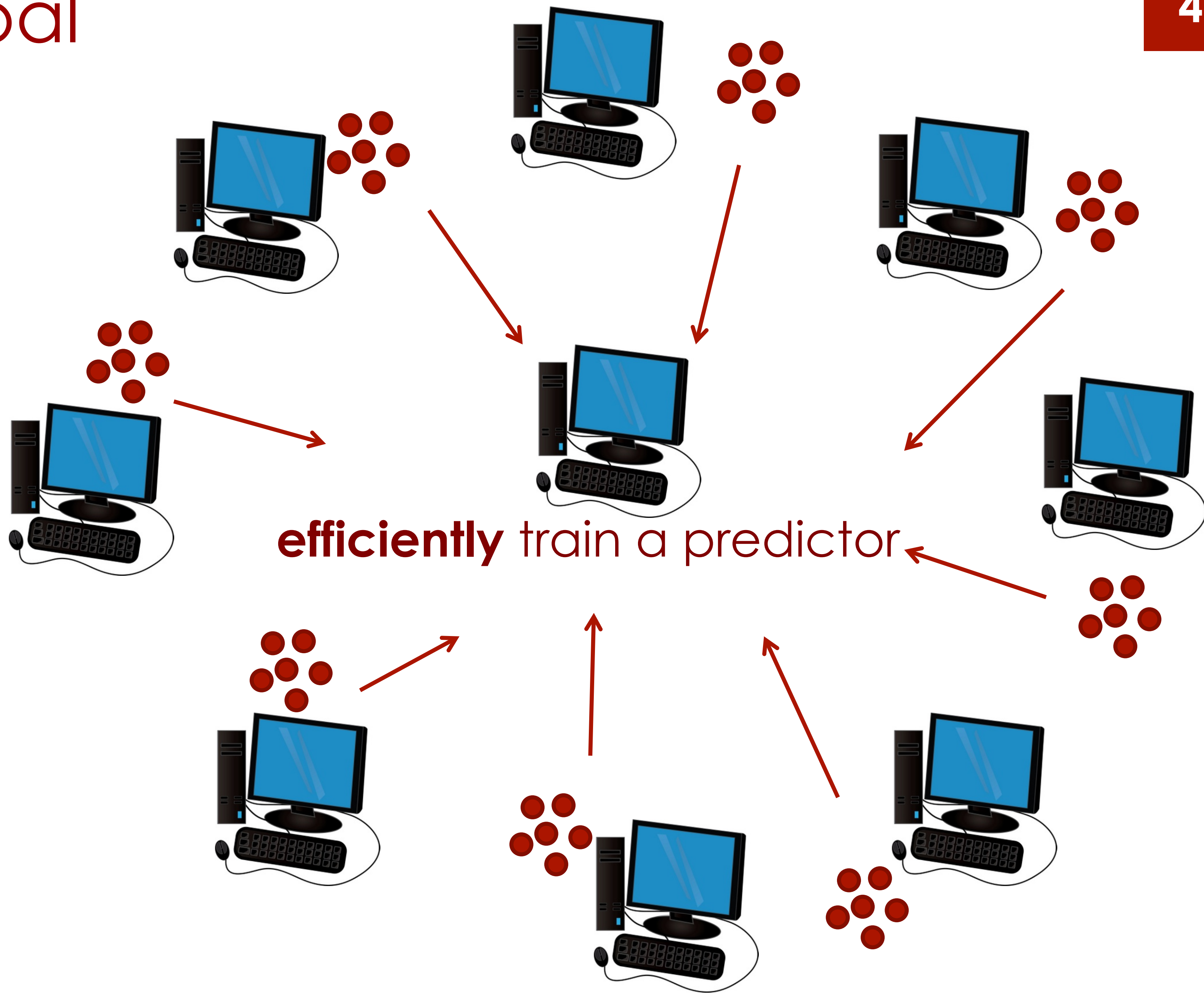
Distributed Learning
"the unreasonable effectiveness of data"

with
Jeff Cooper

Setting


data cannot fit on one machine
(each machine gets data from same distribution)

Goal

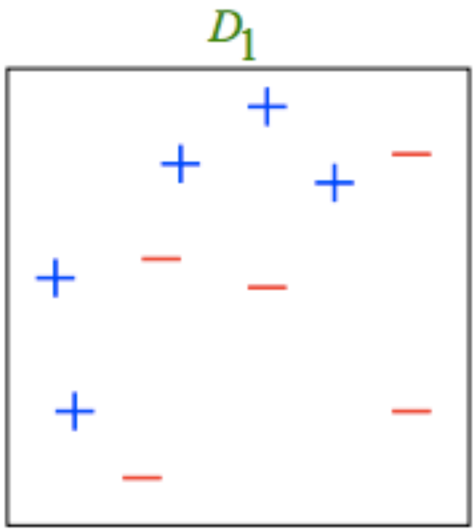


Distributing AdaBoost

- What is boosting?
  - AdaBoost [Freund-Schapire '97] combines "weak" predictors into a strong predictor.
  - Weak predictors are easy to design → lets us make strong predictors by "magic"!
- Why distribute boosting?
  - Probably the best "off the shelf" algorithm.
  - Lends itself to being "distributed" but no adequate solution had been found.
  - Is a weigh + measure approach!

AdaBoost in Pictures (Slides from Schapire)

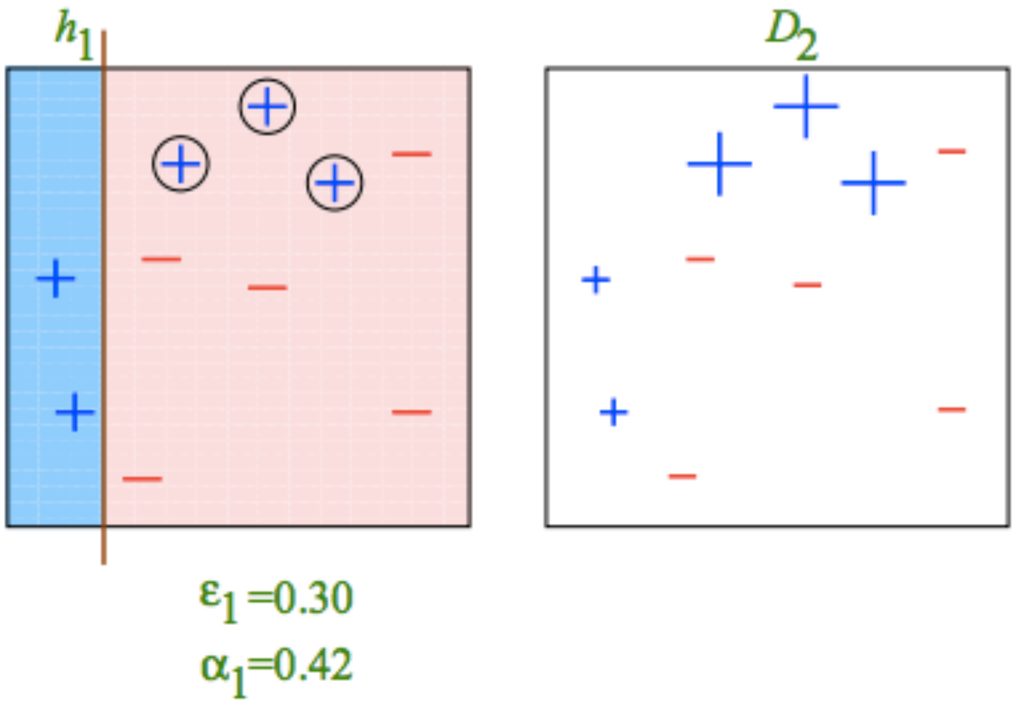
Toy Example



weak classifiers = vertical or horizontal half-planes

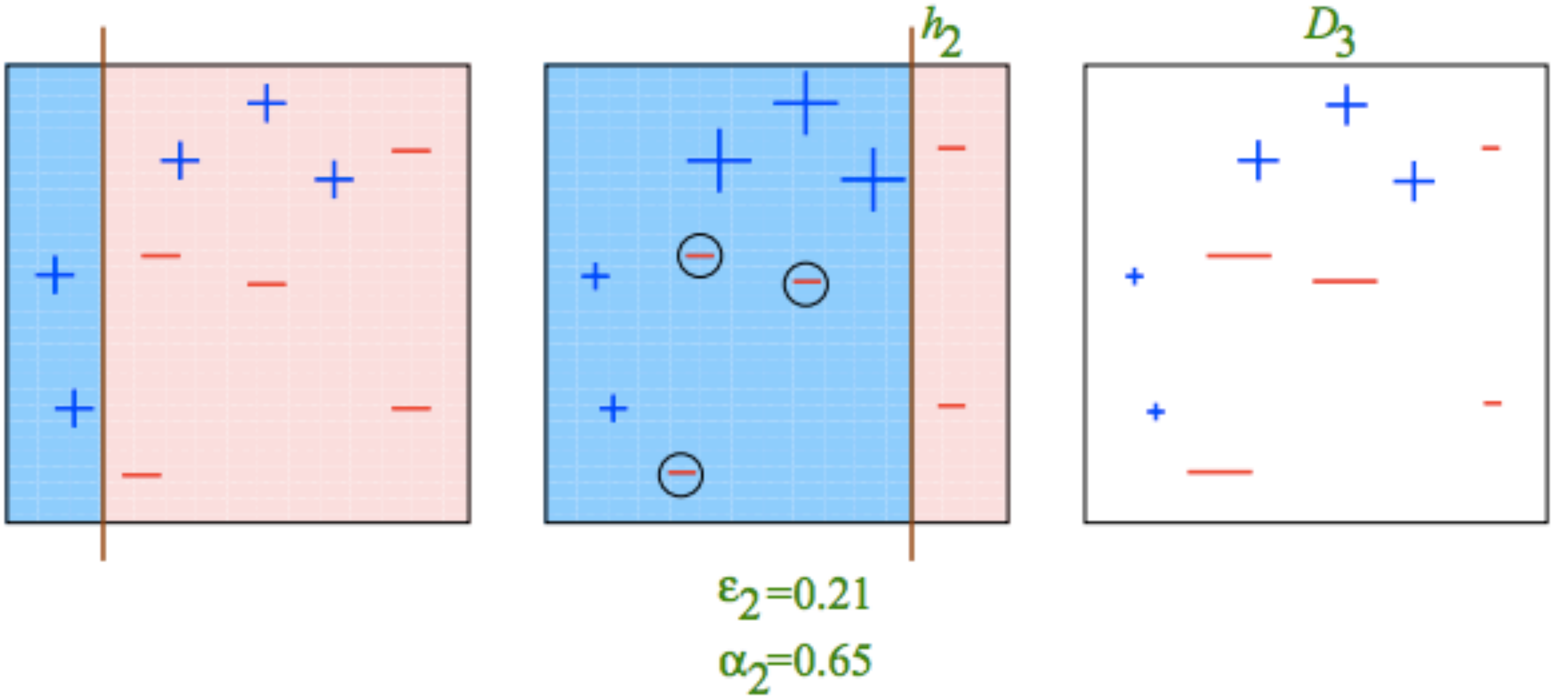
AdaBoost in Pictures (Slides from Schapire)

Round 1



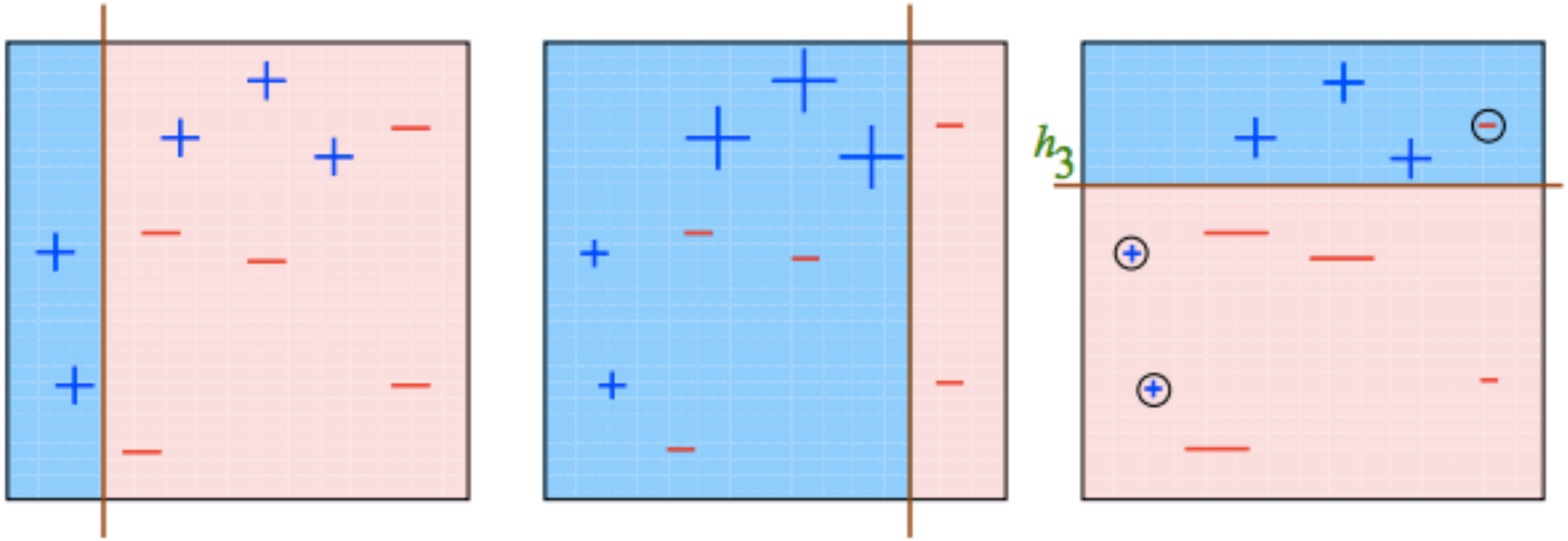
AdaBoost in Pictures (Slides from Schapire)

Round 2



AdaBoost in Pictures (Slides from Schapire)

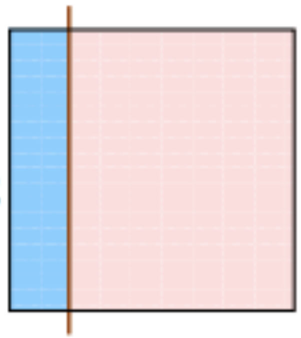
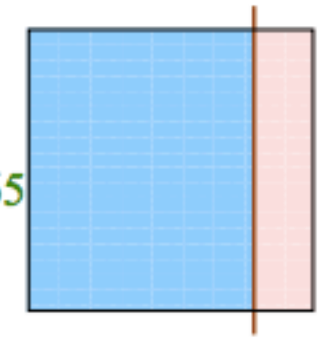
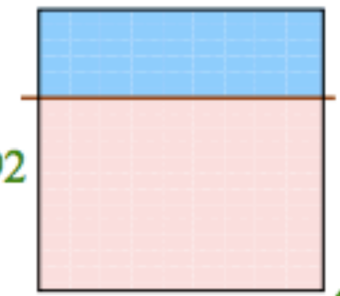
Round 3

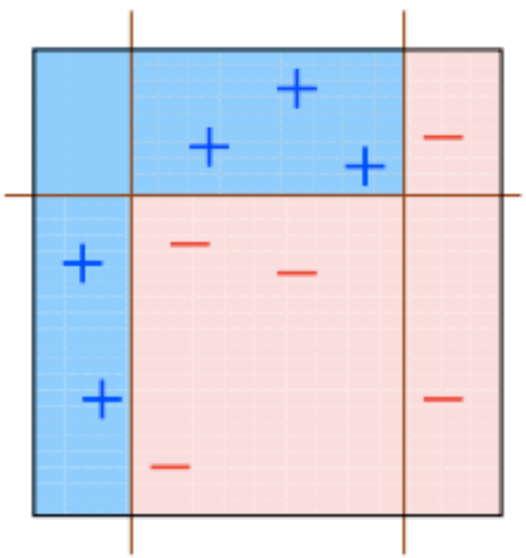


$\varepsilon_3 = 0.14$

$\alpha_3 = 0.92$

Final Classifier

$H_{\text{final}} = \text{sign}\Big(0.42$  $+ 0.65$  $+ 0.92$  $\Big)$

$=$ 

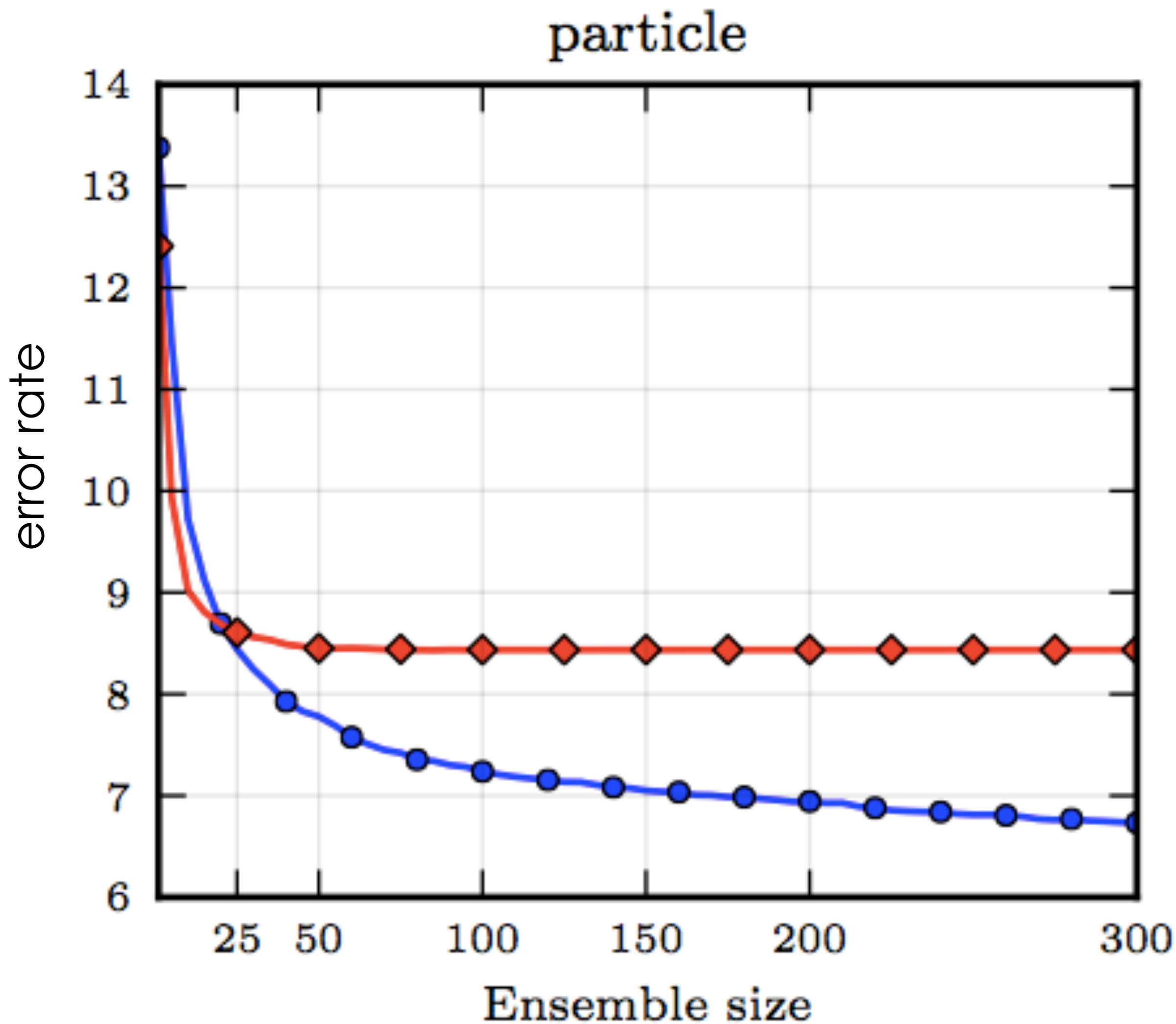
Past Work on Distributed Ensembles

- Bagging [Breiman '96], iVoting [Breiman '99], etc.
  - Upside: easy to distribute for large data [Basilico et al. '11]
  - Downside: not true "boosting" algorithms and do not reach the error rates of AdaBoost
- Boosting in Distributed PAC Model [Balcan et al. '12]
  - A harder model where communication complexity was studied.
- Filterboost [Bradley-Schapire '01], etc.
- **"The Distributed Boosting Alg."** [Lazarevic-Obradovic '01]
  - Upside: Best-yet practical distributed boosting algorithm
  - Downside: unable to reach AdaBoost error rates, uses a lot of communication.

The Distributed Boosting Algorithm [Lazarevic-Obradovic '01]

- Data is split among K machines. The machines "boost" in parallel.
- On each round of boosting:
  1. Each machine sends its weak learner to all other machines.
  2. Each machine computes and sends the local error rate of the "majority learner" to all other machines.
  3. All machines update their local weights based on this info.
- At the end, every machine has full predictor.

DistBoost vs AdaBoost



DistBoost vs **AdaBoost** (if all data fit on one machine)

The Distributed Boosting Algorithm [Lazarevic-Obradovic '01]

- Drawbacks:
  1. Uses lots of communication.
  2. Unable to recover full accuracy of boosting.

     Why? Each site overspecializes to its own data.

- Two different fixes [Cooper-R '14]:
  1. `PreWeak`: Smartly restrict weak learner selection.
  2. `AdaSampling`: Each machine selects its most informative data to send to a central processor.

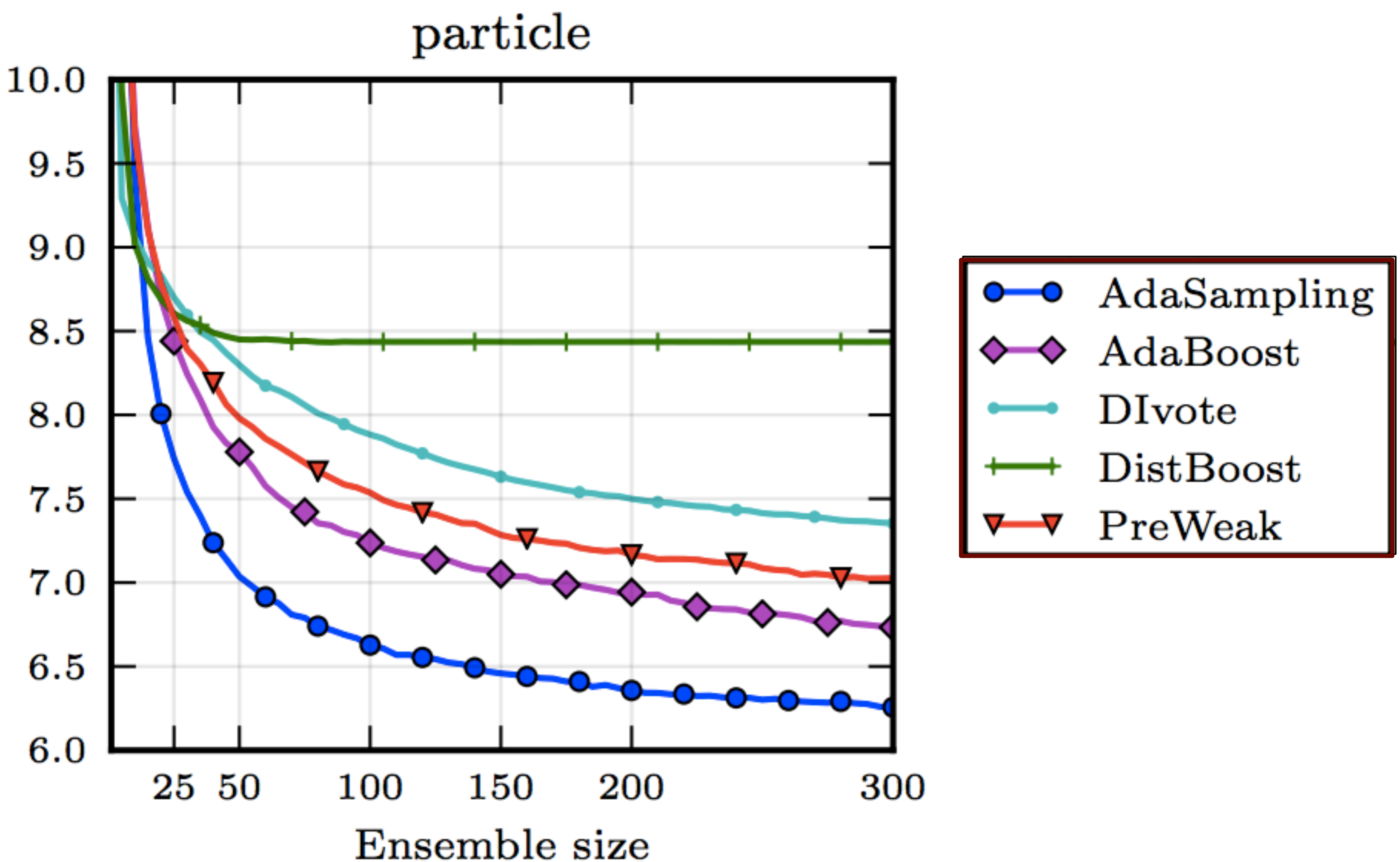
Adaptive Sampling
[Cooper-R '14]

- Main Idea: each machine uses AdaBoost to figure out which examples are "most informative"
  - Not always "hardest examples" – these could be noise.
  - Touches on margins theory

- These examples are sent to main processor, which boosts just on them.

- Rationale based on a game-theoretic view of boosting.

Empirical Results



A Big-Data Example
(Over 1M clicks/non-clicks)

Error rates on Yahoo! click data.

	Stumps	Depth 3
AdaBoost	36.48	35.02
DistBoost	37.23	35.54
DIvote	40.35	37.36
PreWeak	**36.29**	35.58
AdaSampling	37.14	**35.36**
AdaBoost(1/10)	36.76	35.68

Feature-Efficient Prediction

with
Yi Huang, Brian Powers

Feature-Efficient Prediction Examples

- Medical testing

  Want to predict what patients are sick with, but tests might be expensive or dangerous.

- Displaying internet results

  Want to give users the best results you can, but if you don't give results within 300 milliseconds, users will leave.

Model

- Goal is to do supervised learning, using a limited number of features in test-time.
  - Given a budget on total cost: on each example, the learner must look at no more features than allowed by the budget.
  - Each feature has an associated cost.
  - Budget only limited in test data, not training.

- Predictors that do this are feature-efficient.

Lots of work on this problem

- Sequential analysis: when to stop sequential clinical trials. [Wald '47] and [Chernoff '72]
- PAC learning with incomplete features. [Ben-David-Dichterman '93] and [Greiner et al. '02]
- Robust prediction with missing features. [Globerson-Roweis '06]
- Learning linear functions by few features [Cesa-Bianchi et al. '10]
- Incorporating feature costs in CART impurity [Xu et al. '12]
- MDPs for feature selection [He et al. '13]

A “Weigh + Measure” Idea [R ’11]

- An ensemble is usually a weighted vote of many simple rules.
- The simple rules are usually feature-efficient.
- Take a vote of only a few of the rules.

AdaBoostRS [R '11]

Training: train AdaBoost (or any ensemble).

Prediction:

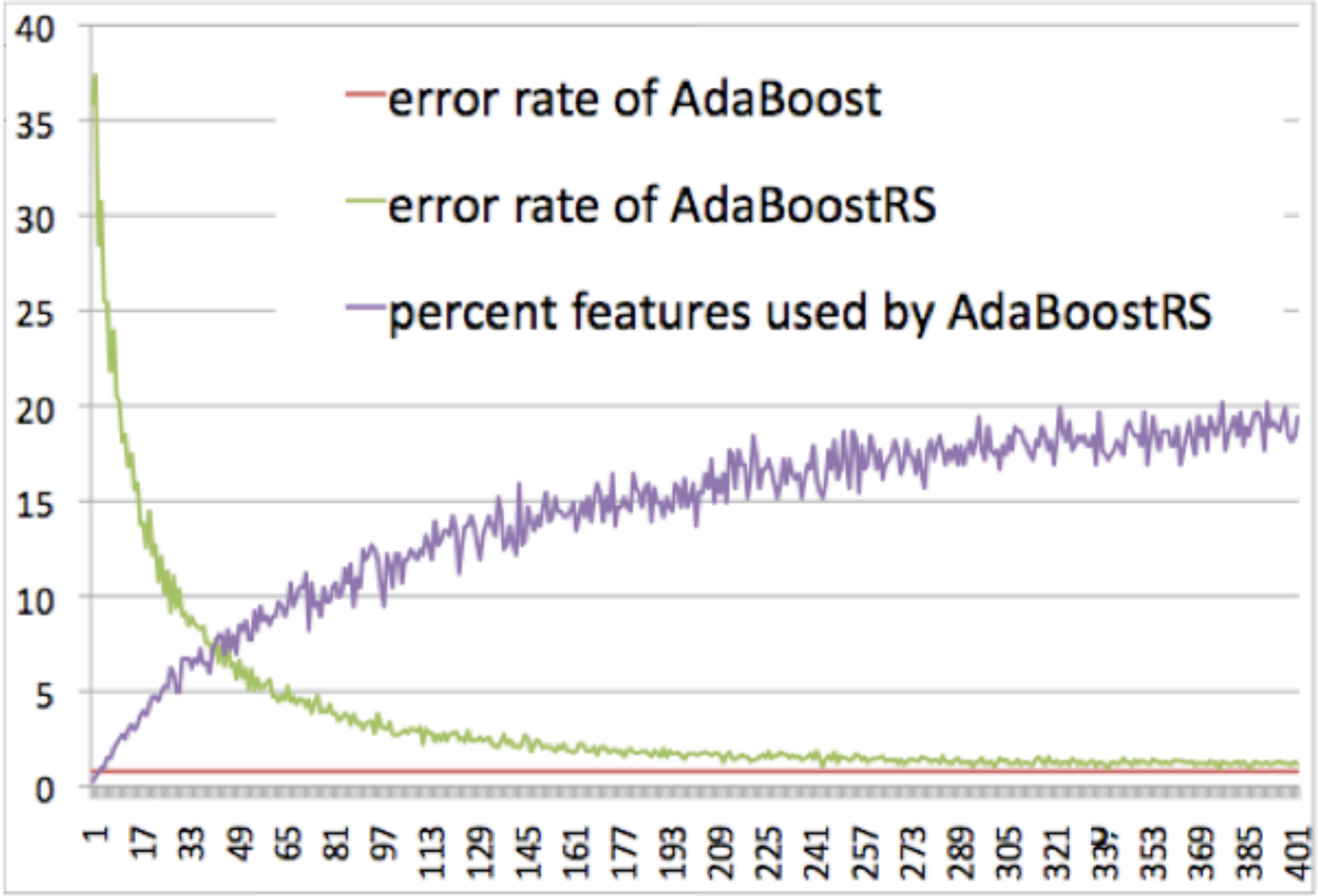
1. Sample the weak learners depending on their voting weights and feature costs.
2. Take a importance-weighted vote of the sampled weak learners.

Intuition:

If ensemble has strong preference, sampling will converge fast. If ensemble is split, who cares?

(Thm resembles margin bound [Schapire et al. '98])

Experiments with AdaBoostRS



On ocr17 dataset. x-axis is number of samples taken.

Room for Improvement

Can we improve by moving the optimization into training?

Turns out: yes, by a lot! [Huang-Powers-R '14]

- Naïve idea: train AdaBoost until budget runs out
- Improvement: choose weak learners more wisely

`AdaBoost (S )` where: $S \subset X \times \{-1, +1\}$

given: $(x_1, y_1), \ldots, (x_m, y_m) \in S$
initialize $D_1(i) = \frac{1}{m}$
for $t = 1, \ldots, T$ **do**
    train base learner using distribution D_t.
    get $h_t \in \mathcal{H} : X \to \{-1, +1\}$.
    choose $\alpha_t = \frac{1}{2} \ln \frac{1+\gamma_t}{1-\gamma_t}$, where $\gamma_t = \sum_i D_t(i) y_i h_t(x_i)$.
    update $D_{t+1}(i) = D_t(i) \exp(\alpha_t y_i h_t(x_i)) / Z_t$,
end for
output the final classifier $H(x) = \mathrm{sign}\left(\sum_{t=1}^{T} \alpha_t h_t(x) \right)$

`AdaBoostBT(S,B,C)` where: $S \subset X \times \{-1,+1\}$, $B > 0$, $C : [n] \to \mathbb{R}^+$

given: $(x_1, y_1), \ldots, (x_m, y_m) \in S$
initialize $D_1(i) = \frac{1}{m}$, $B_1 = B$
for $t = 1, \ldots, T$ **do**
  train base learner using distribution D_t.
  get $h_t \in \mathcal{H} : X \to \{-1,+1\}$.
  if the total cost of the unpaid features of h_t exceeds B_t **then**
    set $T = t - 1$ and **end for**
  else set B_{t+1} as B_t minus the total cost of the unpaid features of h_t, marking them as paid
  choose $\alpha_t = \frac{1}{2}\ln\frac{1+\gamma_t}{1-\gamma_t}$, where $\gamma_t = \sum_i D_t(i) y_i h_t(x_i)$.
  update $D_{t+1}(i) = D_t(i)\exp(\alpha_t y_i h_t(x_i))/Z_t$,
end for
output the final classifier $H(x) = \mathrm{sign}\left(\sum_{t=1}^{T} \alpha_t h_t(x)\right)$

How to choose weak learner h_t?

Training error of AdaBoost is bounded by [Freund & Schapire '97]

$$\hat{\Pr}[H(x) \neq y] \leq \prod_{t=1}^{T} \sqrt{1-\gamma_t^2}$$

With budgets, we need to consider two effects:

- high edges make individual terms smaller
- low costs allow for more terms in the product

Two Optimizations
[Huang-Powers-R '14]

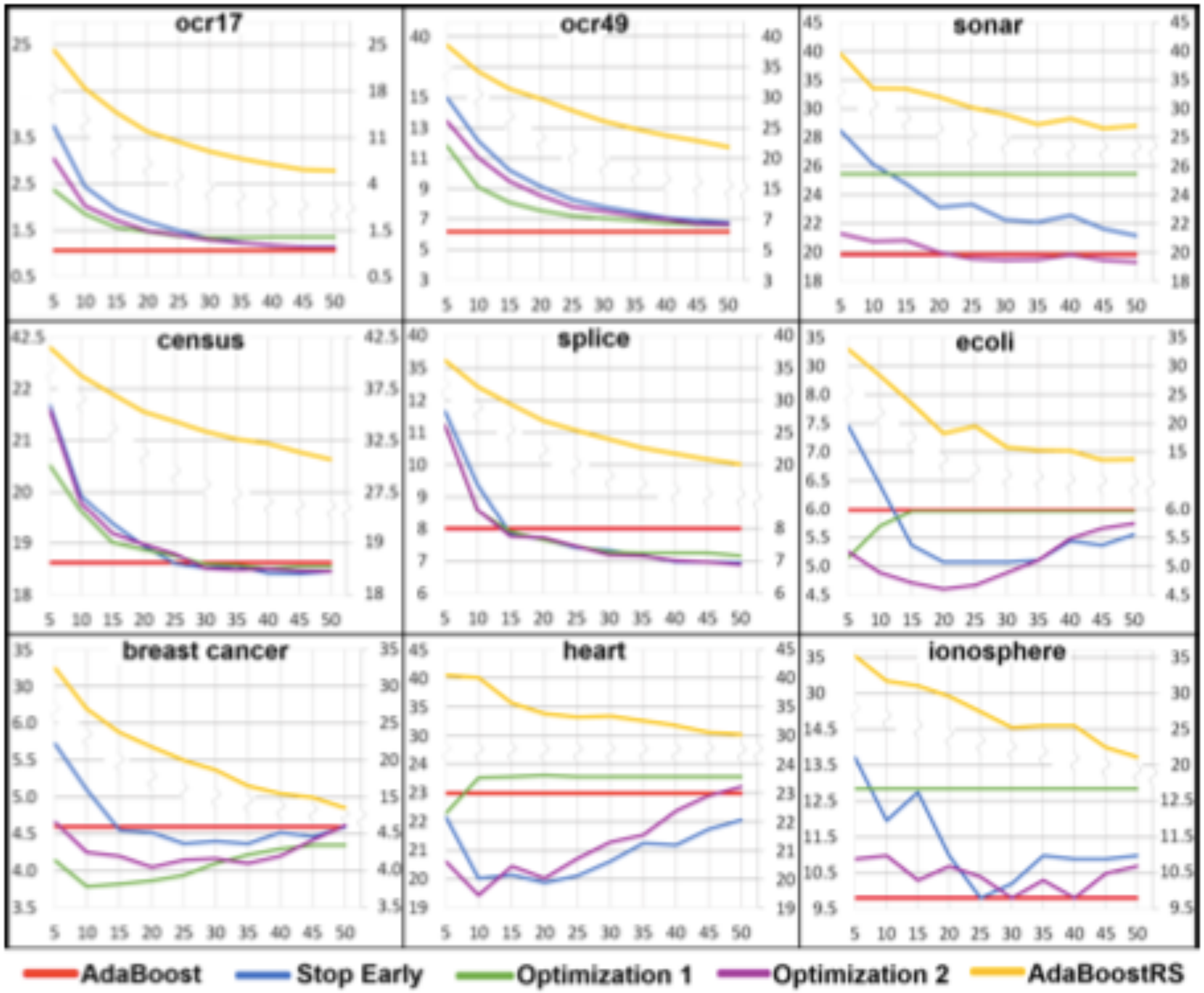
First idea: assume all future rounds will behave like current. Leads to optimization

1) $$h_t = \mathrm{argmin}_{h \in \mathcal{H}} \left(\left(1 - \gamma_t(h)^2\right)^{\frac{1}{c(h)}} \right)$$

Second idea: smoothed version of first.

2) $$h_t = \mathrm{argmin}_{h \in \mathcal{H}} \left(\left(1 - \gamma_t(h)^2\right)^{\frac{1}{(B - B_t) + c(h)}} \right)$$

Experiments with costs ~ U(0,2)



Discussion

Summary

- Gave first optimal and (theoretically) efficient algorithm for contextual bandits by reducing to supervised learning.

- Gave a distributed ensemble algorithm with empirical performance matching boosting.

- Showed how to make any ensemble algorithm "feature efficient".

Observations

- Progress occurred in stages – eg. contextual algorithm was first made optimal but inefficient, then theoretically efficient, then practical (by others).
- All these problems originate from practical concerns.
- Each of these areas heavily used theory, but included practical elements, verified on large data. E.g. distributed boosting's taking care of noisy examples.

Open Problems

- A practical, optimal contextual bandit algorithm that works in adversarial setting.
- Distributed and feature-efficient contextual bandits. (Reduction to supervised learning doesn't preserve these properties.)
- Better models for trading off error / prediction time.
- Many others…

The End

Thank you! Any Questions?