Surveys

- Classes on Tuesday (Nov 26)?
- One (long) lecture for project presentation or two separate lectures?

CS6501: Topics in Learning and Game Theory (Fall 2019)

Learning From Strategically Revealed Samples

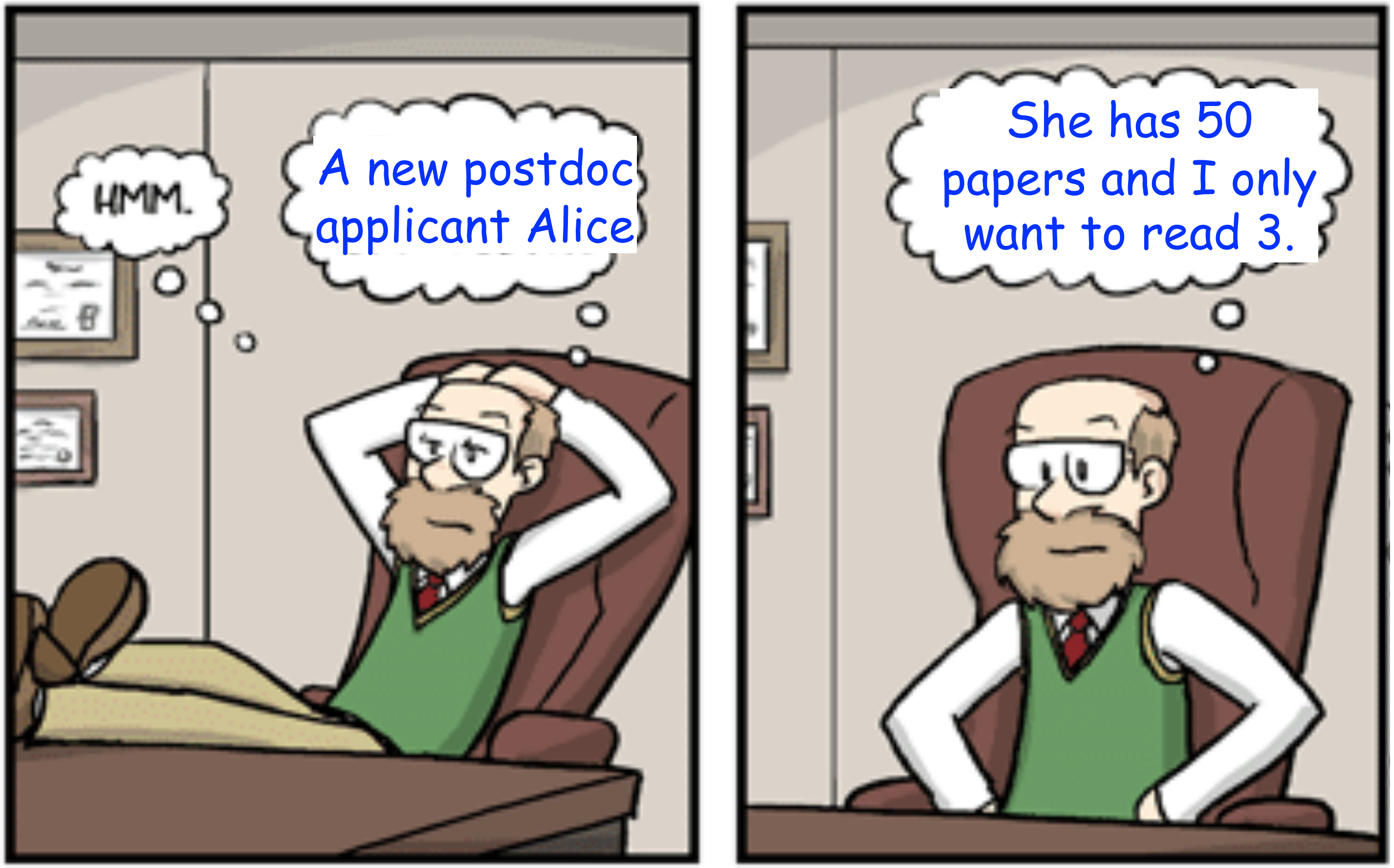
Instructor: Haifeng Xu

Part of slides by Hanrui Zhang

Outline

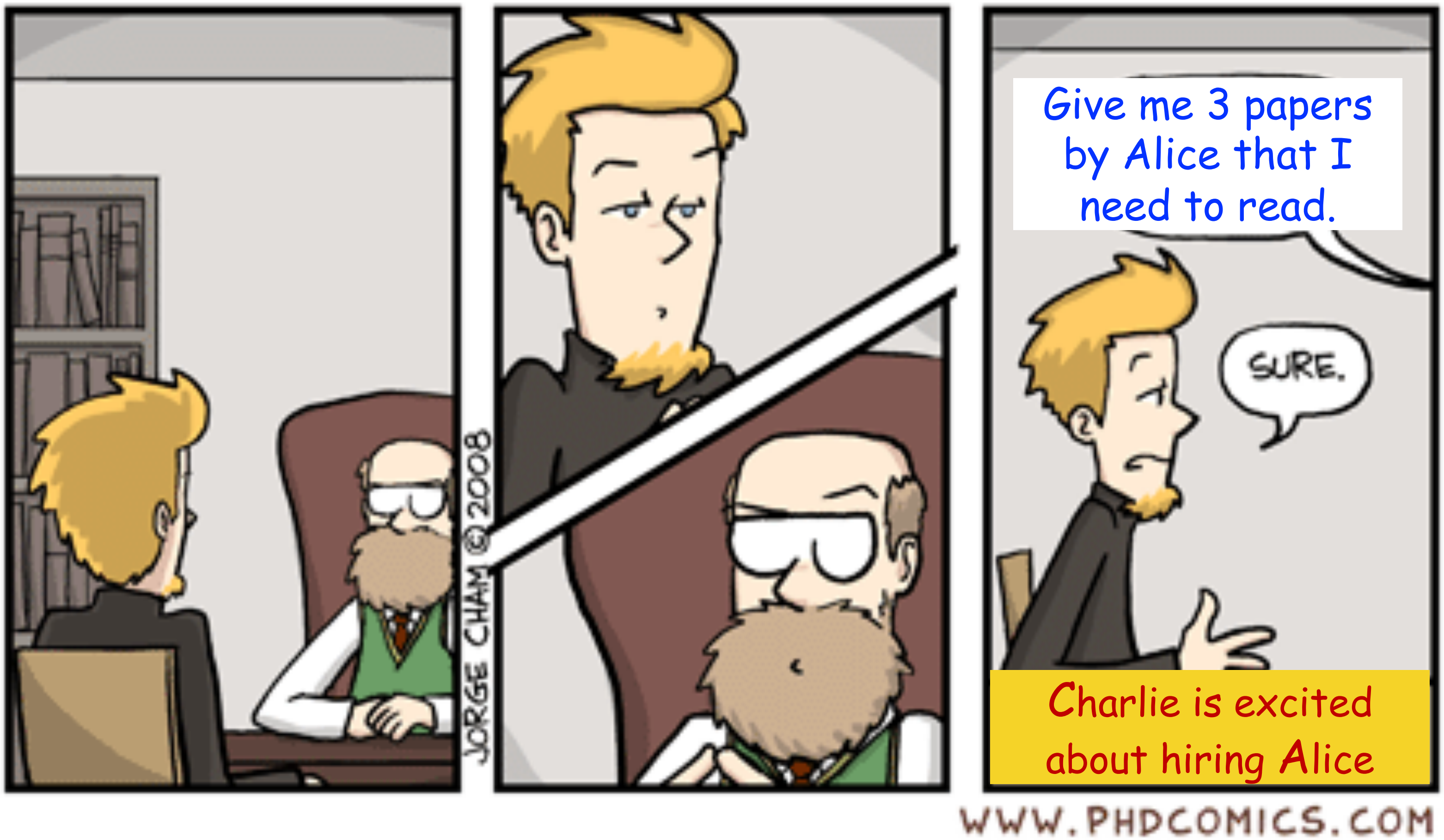
- Introduction and An Example
- Formal Model and Results
- Learning from Strategic Samples: Other Works

Academia in the Era of Tons Publications



The Trouble of Bob, a Professor of Rocket Science

Academia in the Era of Tons Publications



Current postdoc Charlie is happy . . .

Academia in the Era of Tons Publications

I got to pick best 3 papers to persuade Bob, so that he will hire Alice.

Charlie shall pick best 3 papers by Alice – I need to calibrate for that

They know what each other is thinking…

Abstracting the Problem

- Setup: (binary-)classify distributions with label $l \in \{g, b\}$
  - Opposed to classic problem of classifying samples drawn from distributions
- Goal: accept good ones $(l = g)$ and reject bad ones $(l = b)$
- Previous example: a postdoc candidate = a distribution (over papers)

Principal Reacts by Committing to a Policy

- Principal (Bob) commits to and announces a policy to agent Charlie
  - He decides whether to accept l (hire Alice) based on agent's report

Agent's Problem

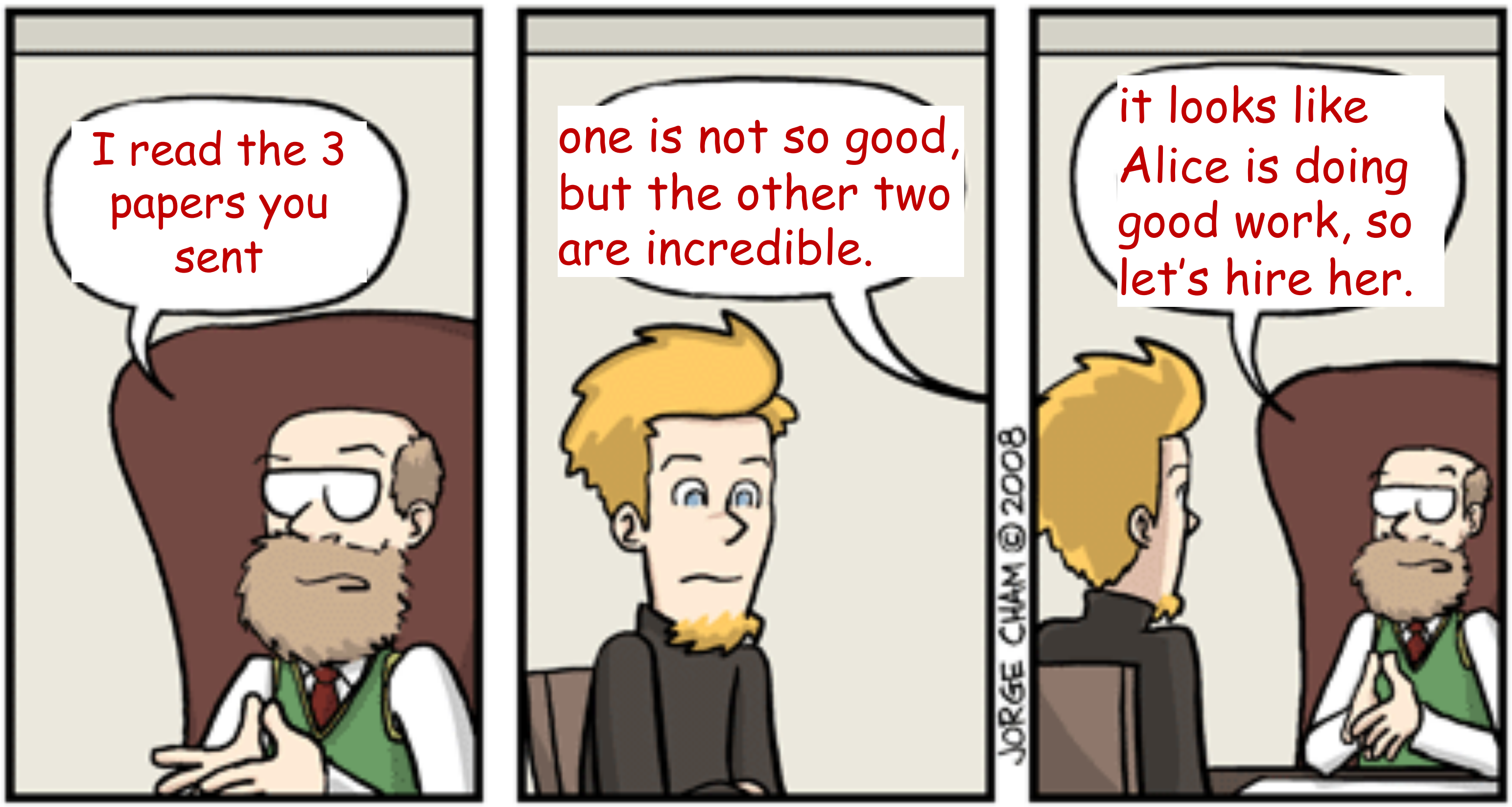
- Has access to $n(= 50)$ samples (papers) from distribution l (Alice)
  - Assume samples are i.i.d.
- Can choose $m(= 3)$ samples as his report

Agent's Problem

- Has access to $n(= 50)$ samples (papers) from distribution l (Alice)
  - Assume samples are i.i.d.
- Can choose $m(= 3)$ samples as his report
- Agent (Charlie) sends his report to Bob principal (Bob), aiming to persuade Bob to accept distribution l (Alice)

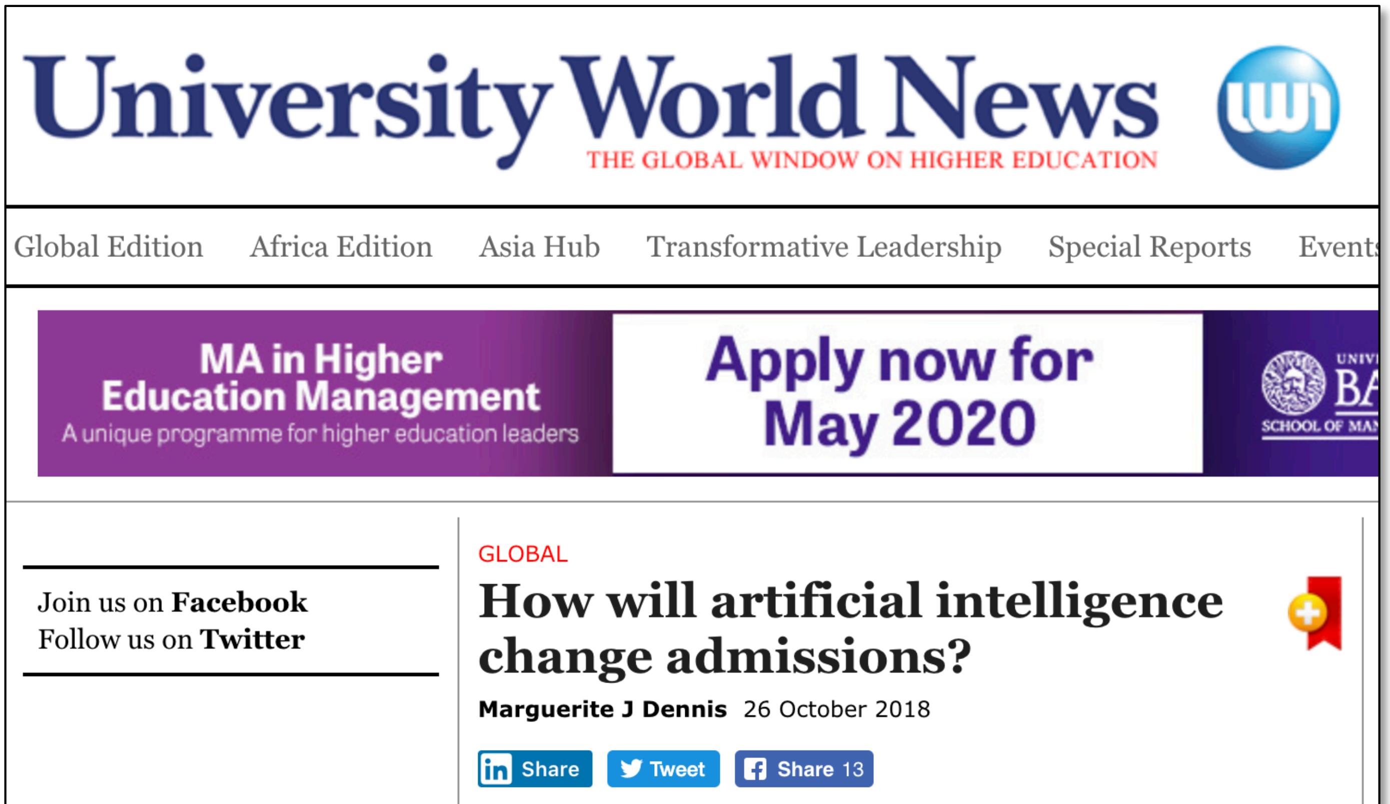
Principal Executes Based on His Policy

- Bob observes Charlie's report, and makes a decision according to the policy he announced



Strategic Classifications are Everywhere

- University admissions
  - Students academic records are selectively revealed


University World News
THE GLOBAL WINDOW ON HIGHER EDUCATION

Global Edition | Africa Edition | Asia Hub | Transformative Leadership | Special Reports | Events

Join us on **Facebook**
Follow us on **Twitter**

GLOBAL

How will artificial intelligence change admissions?

Marguerite J Dennis 26 October 2018

Share | Tweet | Share 13

Strategic Classifications are Everywhere

- University admissions
  - Students academic records are selectively revealed
- Classify loan lending decisions
  - Borrowers will selectively report their features

Strategic Classifications are Everywhere

- University admissions
  - Students academic records are selectively revealed
- Classify loan lending decisions
  - Borrowers will selectively report their features
- Decide which restaurants to go based on Yelp rating
  - Platform may selectively showing you ratings
- Hiring job candidates in various scenarios

Strategic Classifications are Everywhere

- University admissions
  - Students academic records are selectively revealed
- Classify loan lending decisions
  - Borrowers will selectively report their features
- Decide which restaurants to go based on Yelp rating
  - Platform may selectively showing you ratings
- Hiring job candidates in various scenarios
- Note: this problem deserves study even you do classification manually instead of using an automated classifier
  - E.g., deciding where to hold the next Olympics based on photographs of different city locations

Outline

- Introduction and An Example
- Formal Model and Results
- Learning from Strategic Samples: Other Works

The Model: Basic Setup

- A distribution $l \in \{g, b\}$ arrives, which can be good ($l = g$) or bad ($l = b$)
- An agent has access to n i.i.d. samples from l, from which he chooses a subset of exactly m samples as his report
  - Agent's goal: persuade a principal to accept l
- Principal observes agent's report, and decides whether to accept
  - Principal's goal: accept when $l = g$ and reject when $l = b$
  - Want to minimize her probability of mistakes

The Model: the Timeline

a distribution $l \in \{g, b\}$ arrives

l generates n iid samples $D = \{d_1, \cdots, d_n\}$

Agent receives D and report $R = \{r_1, \dots, r_m\} \subseteq D$

Objective: maximize prob of accepting l

Objective: accept g and reject b

Principal commits to a policy $\Pi(R) \in [0, 1]$ that maps report R to probability of accepting R

probability $p = \Pi(R)$ of accepting l given report R

Simpler Case: Agent is NOT Strategic

- This is the same as distinguishing two distributions from samples
  - You have m samples from distribution either g or b
  - Want to tell which one it is, with high probability (you almost can never be 100% certain)

Fact: Let $\epsilon = \max_S [g(S) - b(S)]$ be total variation (TV) distance between g, b. Then $\Omega(1/\epsilon^2)$ samples to distinguish g, b with constant success probability.

Note: $g(S) = \Pr_{x \sim g}(x \in S)$ is accumulated probability for $x \in S$

Simpler Case: Agent is NOT Strategic

- This is the same as distinguishing two distributions from samples
  - You have m samples from distribution either g or b
  - Want to tell which one it is, with high probability (you almost can never be 100% certain)

Fact: Let $\epsilon = \max_S [g(S) - b(S)]$ be total variation (TV) distance between g, b. Then $\Omega(1/\epsilon^2)$ samples to distinguish g, b with constant success probability

Formally,

$$||g - b||_{TV} = \int_{x:g(x)>b(x)} [g(x) - b(x)]dx$$

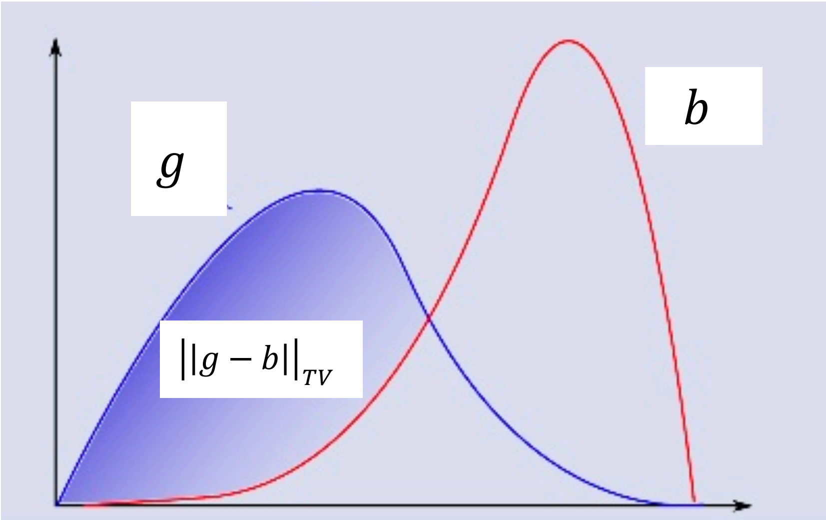


Illustration of TV distance

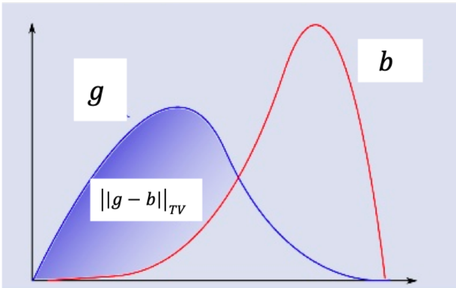
Simpler Case: Agent is NOT Strategic

- This is the same as distinguishing two distributions from samples
  - You have m samples from distribution either g or b
  - Want to tell which one it is, with high probability (you almost can never be 100% certain)

Fact: Let $\epsilon = \max_S [g(S) - b(S)]$ be total variation (TV) distance between g, b. Then $\Omega(1/\epsilon^2)$ samples to distinguish g, b with constant success probability

Proof

- First, compute $S^* = \arg\max_S [g(S) - b(S)]$
- Idea: try to estimate value of $l(S^*)$ where $l \in \{g, b\}$
  - Why? This statistics has largest gap among g, b
- How to estimate $l(S^*)$ from samples?
  - Calculate fraction of samples in S^*
- $\Omega(1/\epsilon^2)$ samples suffices to distinguish random variable $g(S^*)$ from $b(S^*)$



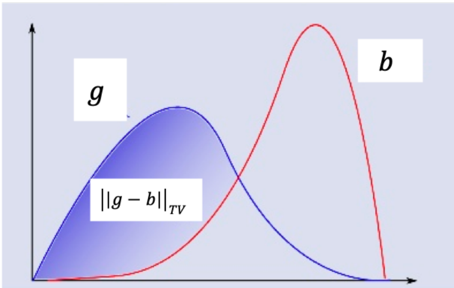
Simpler Case: Agent is NOT Strategic

- This is the same as distinguishing two distributions from samples
  - You have m samples from distribution either g or b
  - Want to tell which one it is, with high probability (you almost can never be 100% certain)

Fact: Let $\epsilon = \max_S [g(S) - b(S)]$ be total variation (TV) distance between g, b. Then $\Omega(1/\epsilon^2)$ samples to distinguish g, b with constant success probability

Remarks

- When agent is not strategic, performance depends on TV distance in the form of $\Omega\left(\frac{1}{\epsilon^2}\right)$



Strategic Agent: An Example

"Tough" World

- A good candidate writes a good paper w.p. 0.05
- A bad candidate writes a good paper w.p. 0.005
- All candidates have $n = 50$ papers, and the professor wants to read only $m = 1$ good candidate

Q: What is a reasonable principal policy?

Strategic Agent: An Example

"Tough" World

- A good candidate writes a good paper w.p. 0.05
- A bad candidate writes a good paper w.p. 0.005
- All candidates have $n = 50$ papers, and the professor wants to read only $m = 1$ good candidate

Q: What is a reasonable principal policy?

- Accept iff the reported paper is good
  - Good candidate is accepted with prob $p_g = 1 - (1 - 0.05)^{50} \approx 0.92$
  - A bad candidate is accepted with prob $p_b = 1 - (1 - 0.005)^{50} \approx 0.22$
- What happens if agent not strategic? → almost cannot distinguish
- Strategic selection actually helps principal!

Strategic Agent: An Example

"Easy" World

- A good candidate writes a good paper w.p. ~~0.05~~ 0.95
- A bad candidate writes a good paper w.p. ~~0.005~~ 0.05
- All candidates have $n = 50$ papers, and the professor wants to read only $m = 1$ good candidate

Strategic Agent: An Example

"Easy" World

- A good candidate writes a good paper w.p. ~~0.05~~ 0.95
- A bad candidate writes a good paper w.p. ~~0.005~~ 0.05
- All candidates have $n = 50$ papers, and the professor wants to read only $m = 1$ good candidate

Policy: Accept iff the reported paper is good

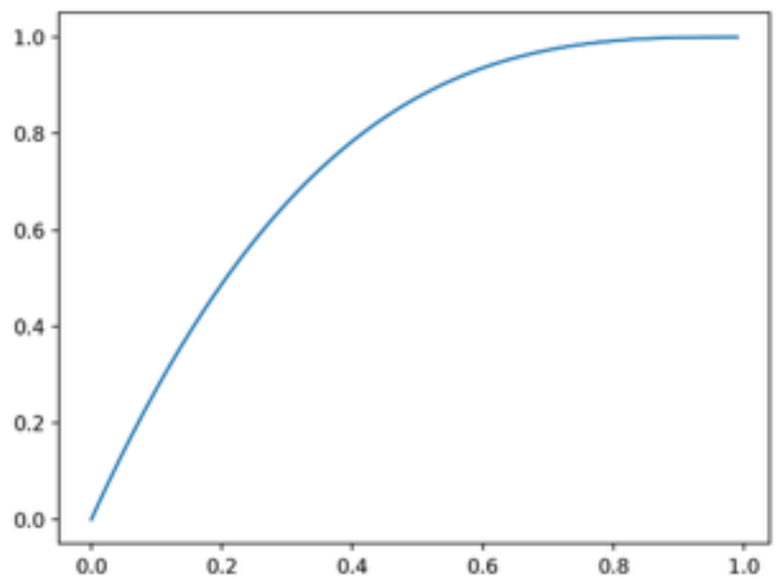
- Good candidate is accepted with prob $p_g = 1 - (1 - 0.95)^{50} \approx 1$
- A bad candidate is accepted with prob $p_b = 1 - (1 - 0.05)^{50} \approx 0.92$
- What happens if agent not strategic? → can distinguish easily
- Here, strategic selection hurts principal!

General Results: One Sample

Theorem: Any pareto optimal deterministic policy satisfies:

1. It orders sample space based on likelihood ratio $g(x)/b(x)$
2. Limiting acceptance probability satisfy: $p_g + (1 - p_b)^r = 1$ where $r = \max_x g(x)/b(x)$ is maximum likelihood ratio

➢ That is, principle tries to use the "most distinguishable" sample



Pareto frontier when $r = 3$

Note: can define

error rate = $\min \frac{p_b}{p_g}$

General Results: One Sample

Theorem: Any pareto optimal deterministic policy satisfies:

1. It orders sample space based on likelihood ratio $g(x)/b(x)$
2. Limiting acceptance probability satisfy: $p_g + (1 - p_b)^r = 1$ where $r = \max_x g(x)/b(x)$ is maximum likelihood ratio

- That is, principle tries to use the "most distinguishable" sample
- In strategic environment, likelihood ratio $g(x)/b(x)$ matters
  - Opposed to TV distance in non-strategic setting

Multiple Samples:

Theorem: There is a deterministic policy:

1. Which orders the sample space
2. Whose limiting error rate is at most $\exp(-m(1 - r^{-0.5})^2/2)$

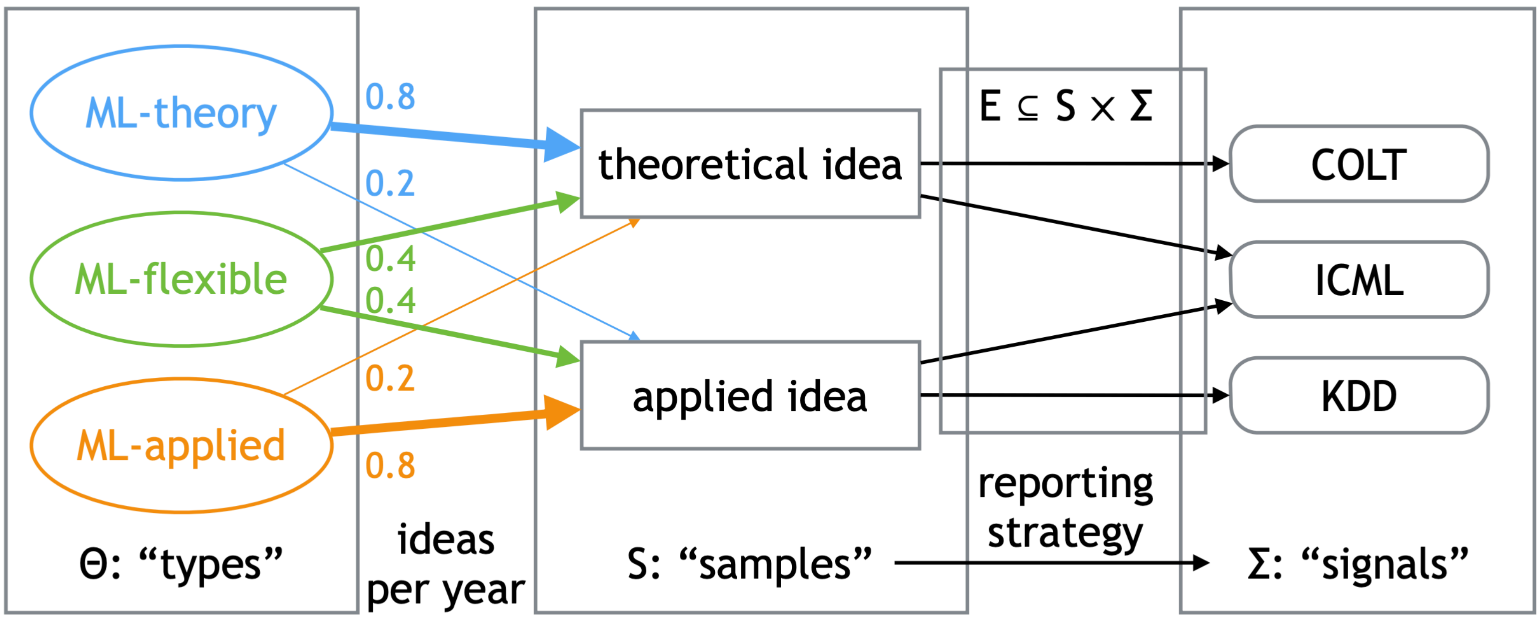
- This is not exactly optimal
- But, error rate decrease exponentially in m
- What is optimal like?
  - It is open, we don't know

Outline

- Introduction and An Example
- Formal Model and Results
- Learning from Strategic Samples: Other Works

Many Other Work in this Space

➢ Learning from samples that are strategically transformed

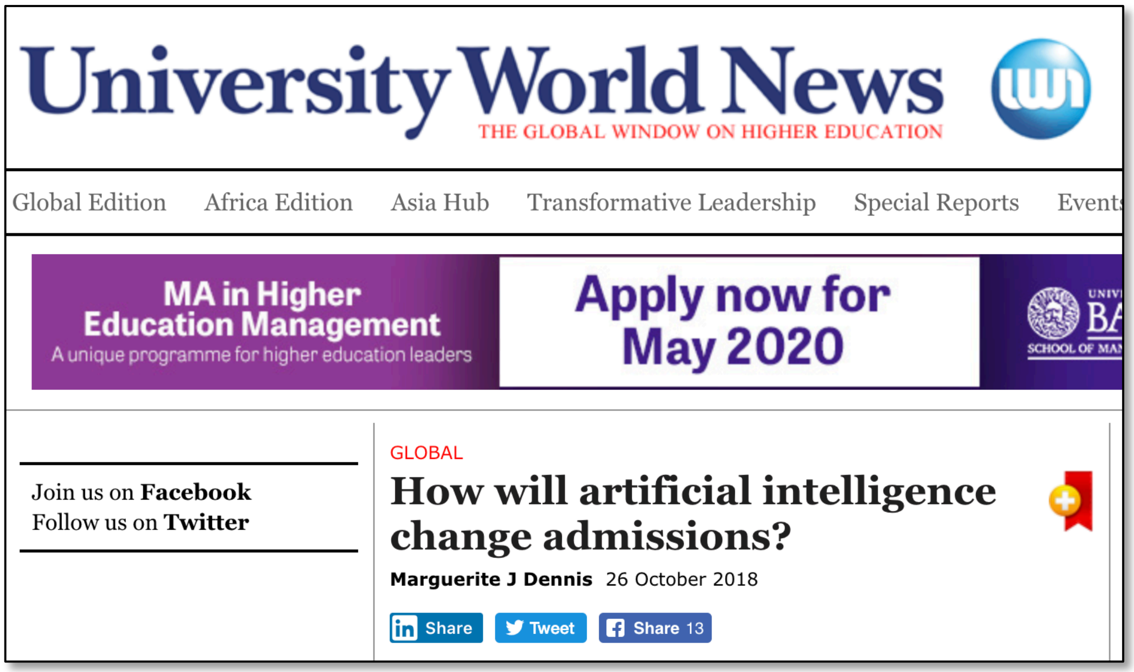


Many Other Work in this Space

- Strategic behaviors are costly

When Strategic Behaviors are Costly

- How to induce the correct strategic behaviors



Paper: How Do Classifiers Induce Agents To Invest Effort Strategically by Kleinberg and Raghavan

Thank You

Haifeng Xu

University of Virginia

hx4ad@virginia.edu