Incentive Mechanisms for Data: The Peer Prediction Approach

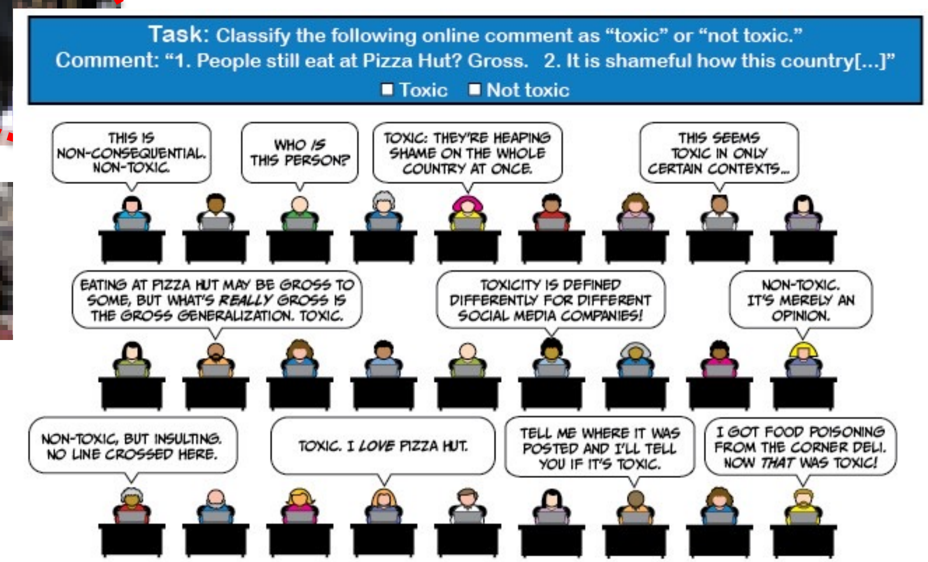
Yang Liu

Assistant Professor, UC Santa Cruz

Amazon Visiting Academics, Human Labeled Data@Amazon Search

Summer School on Game Theory and Social Choice 2022

Human annotations are prone to errors..



Preferred caption:

A young boy wearing a orange helmet riding a scooter.

Noisy caption:

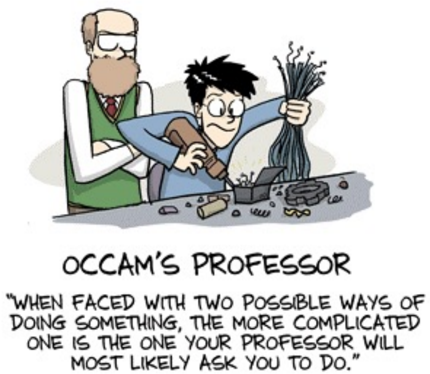
A person scating on the road.

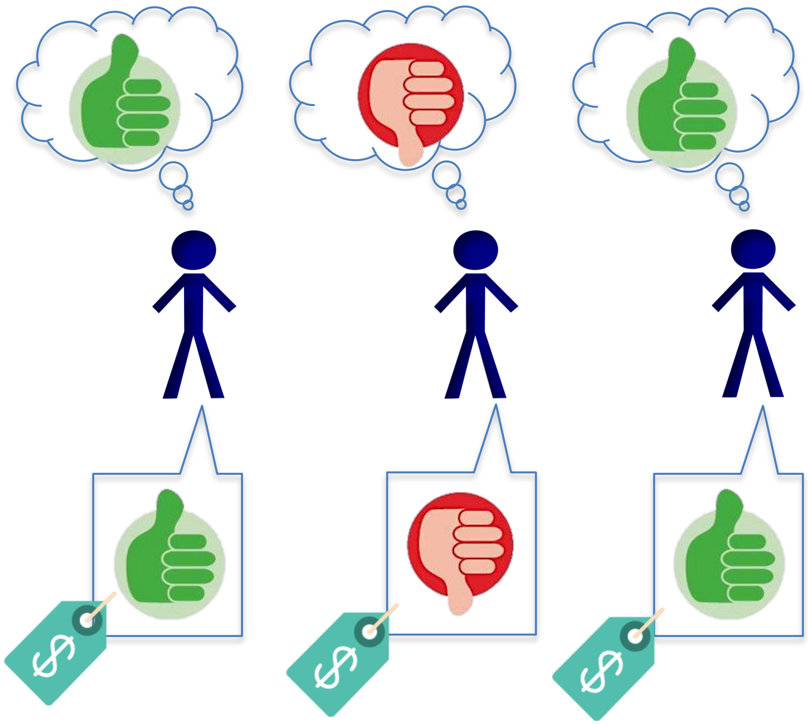
Harms of noisy annotations

- Harms training accuracy
- Invalidate model performance
- False sense of fairness
- Leads to biased treatments

..and more

Incentives

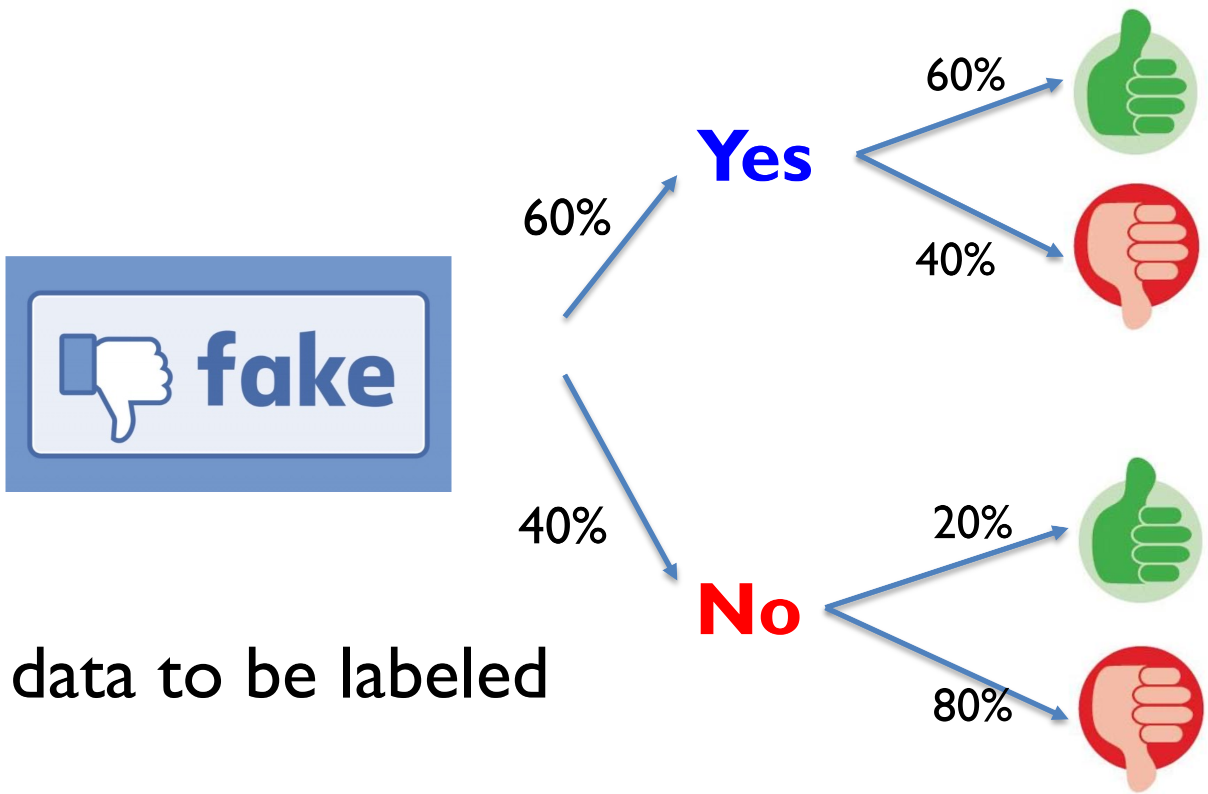





Your Data Is Crucial to a Robotic Age. Shouldn't You Be Paid for It?

Eduardo Porter
ECONOMIC SCENE
MARCH 6, 2018

Running example



N data to be labeled

X	Y	Y_i or $\mathbb{P}(Y \mid Y_i)$
feature	True label	Prediction

Main Question

Can we incentivize high-quality prediction when the ground truth is unavailable?

Main Question

Can we incentivize high-quality prediction when the ground truth is unavailable?

Two meanings:

- **Incentivize truthful reporting**

Main Question

Can we incentivize high-quality prediction when the ground truth is unavailable?

Two meanings:

- **Incentivize truthful reporting**
- **Accurate prediction get higher expected rewards**

Information Elicitation with Ground-truth

Strictly Proper Scoring Rules (SPSR)

- How likely will it rain tomorrow?

Strictly Proper Scoring Rules (SPSR)

- How likely will it rain tomorrow? 

- The principal pays:

Report of agent i

SPSR: $S(q_i, Y)$

Strictly Proper Scoring Rules (SPSR)

- How likely will it rain tomorrow?

- The principal pays:

Report of agent i

$$\text{SPSR: } S(q_i, Y)$$

Ground truth $Y \in \{0,1\}$

Strictly Proper Scoring Rules (SPSR)

- How likely will it rain tomorrow?

- The principal pays:

Report of agent i

SPSR: $S(q_i, Y)$

Ground truth $Y \in \{0,1\}$

- Truthfulness: $S(q_i, Y)$ is SPSR if and only if

$$\forall p_i, q_i \neq p_i, \mathbb{E}_{Y\sim p_i}[S(p_i, Y)] > \mathbb{E}_{Y\sim p_i}[S(q_i, Y)]$$

Strictly Proper Scoring Rules (SPSR)

- How likely will it rain tomorrow?

- The principal pays:

Report of agent i

SPSR: $S(q_i, Y)$

Ground truth $Y \in \{0,1\}$

- Truthfulness: $S(q_i, Y)$ is SPSR if and only if

$$\forall p_i, q_i \neq p_i, \mathbb{E}_{Y \sim p_i}[S(p_i, Y)] > \mathbb{E}_{Y \sim p_i}[S(q_i, Y)]$$

Belief of agent i

Strictly Proper Scoring Rules (SPSR)

- How likely will it rain tomorrow?

- The principal pays:

Report of agent i

SPSR: $S(q_i, Y)$

Ground truth $Y \in \{0,1\}$

- Truthfulness: $S(q_i, Y)$ is SPSR if and only if

$$\forall p_i, q_i \neq p_i, \mathbb{E}_{Y\sim p_i}[S(p_i, Y)] > \mathbb{E}_{Y\sim p_i}[S(q_i, Y)]$$

Belief of agent i

From agent i's perspective

Strictly Proper Scoring Rules (SPSR)

- Example: $S(q_i, Y) = 1 - (q_i - Y)^2$

Strictly Proper Scoring Rules (SPSR)

- Example: $S(q_i, Y) = 1 - (q_i - Y)^2$

- Accurate predictions get higher rewards

Strictly Proper Scoring Rules (SPSR)

- Example: $S(q_i, Y) = 1 - (q_i - Y)^2$

- Accurate predictions get higher rewards
  - p^* - true distribution of Y (fixed)

$$\underbrace{\mathbb{E}_{Y \sim p^*}[S(q_i, Y)]}_{\text{The true expected reward}} = const - \underbrace{||p^* - q_i||}_{\text{A divergence function}}$$

True belief: $p = 0.6$ for an event happening

Report a q in [0,1]

Expected scores

$$
\begin{aligned}
&0.6 \cdot \left(1-(1-q)^2\right)+0.4 \cdot \left(1-(0-q)^2\right) \\
=&1-0.6 \cdot (1-2q+q^2)-0.4 \cdot q^2 \\
=&-q^2+1.2q+0.4
\end{aligned}
$$

$q^* = 0.6$

Truth telling > Deviation

Challenges

Costly or impossible (e.g., significant delay) to verify reports against observable ground truth

- Label collection
- Peer review/grading
- Personal record
- Q: will people land on Mars by 2030?

Peer Prediction: A family of algorithms mechanisms to truthfully elicit private or high quality signals at equilibria

[Prelec 04, Miller et al. 05, Jurca & Faltings 09, Witkowski & Parkes 12, Radanovic & Faltings 13, Dasgupta & Ghosh 13, Shnayder et al. 16]

Peer Prediction

Key Idea: verify the reports against one another

- **Reward** = how well each report correlates with other reports
- **Truthful equilibrium**

Person A	Person B	Payment for A
		\$1.50
		\$0.10
		\$0.30
		\$1.20

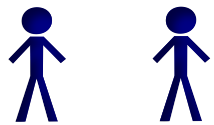
A Scoring rule: $S(q_A, q_B)$

Design Goal

- A Scoring rule: $S(q_A, q_B)$
- Truthfulness at Bayesian Nash Equilibrium

$$\forall q_A \neq p_A, \mathbb{E}_{p_B|p_A}[S(p_A, p_B)] > \mathbb{E}_{p_B|p_A}[S(q_A, p_B)]$$

Peer Prediction

- Two agents *i, -i* 

- Reporting cateogrical signals

$$y_i, y_{-i} \in \{0,1\}$$

- Predicting *peer agent's prediction*

$$\mathrm{SPSR}: S(q_i, y_{-i}), \quad q_i = \Pr(y_{-i}|y_i)$$

Eliciting Informative Feedback: The Peer-Prediction Method, Miller et al. 2005.

Peer Prediction

$$\text{SPSR: } S(q_i, y_{-i}), \quad q_i = \Pr(y_{-i} | y_i)$$

- Inherits guarantees from SPSR at a **Bayesian Nash Equilibrium**
  - If everybody else is truthfully reporting, the best to do is also truthful reporting

- Caveat: Has to know the information structure to update

Eliciting Informative Feedback: The Peer-Prediction Method, Miller et al. 2005.

Bayesian Truth Serum

"What is right is not always popular and what is popular is not always right."

--- Albert Einstein

Solution: Bayesian Truth Serum

- A surprisingly more popular answer is the correct answer
- Each participant also answers the question: *how much they believe that others will agree with themselves (f)*

$$\mathrm{BTS}(x=i,f) = \underbrace{\log\frac{\bar{x}_i}{\bar{f}_i}}_{\text{information score}} - \underbrace{\sum_j \bar{x}_j \log\frac{f_j}{\bar{x}_j}}_{\text{prediction penality}}$$

$$\bar{x}_i = \text{average of } (x_n = i)$$

$$\bar{f} = \text{geometric average of} (f_n)$$

Drazen Prelec, Bayesian Truth Serum, Science 2004, Nature 2017

Multi-task Setting

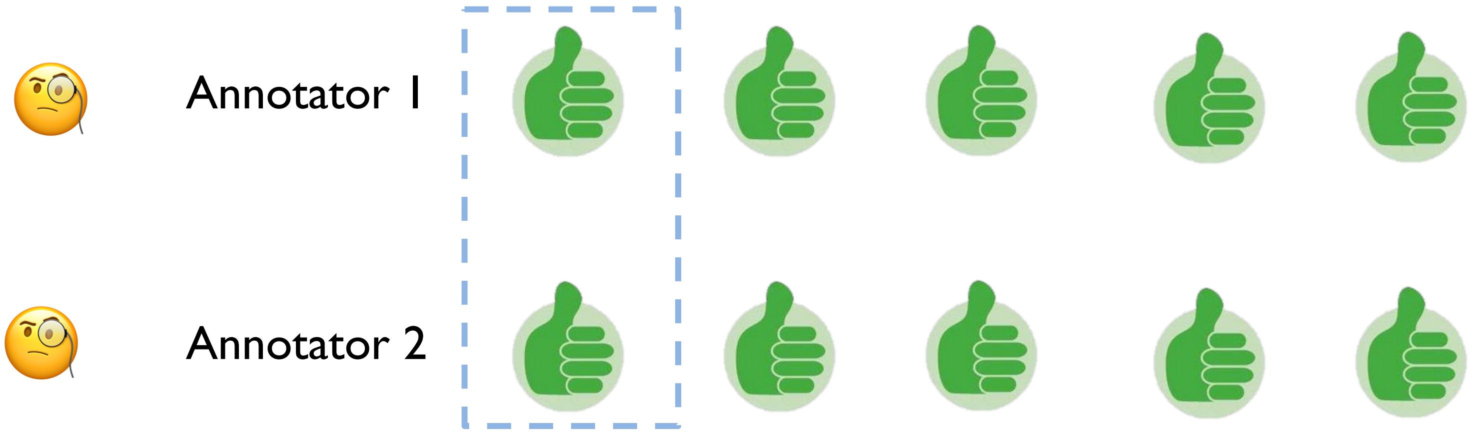
PP and BTS are single task mechanisms

- Each has limitations

Using correlation among tasks

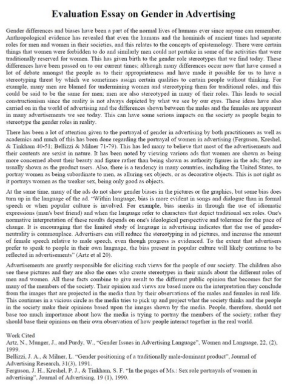
- N – A set of agents (index i)
- M – A set of tasks (index k)
- $Y_k \in \{0,1\}$ – the ground truth of task k
- $Z_{i,k}$ – the report of agent i on task k

What went wrong with agreement



Using agreement metric = 1 😀

Cheap signal



Evaluation Essay on Gender in Advertising

Gender differences and biases have been a part of the normal lives of humans ever since anyone can remember. Anthropological evidence has revealed that even the humans and the hominids of ancient times had separate roles for men and women in their societies, and this relates to the concepts of epistemology. There were certain things that women were forbidden to do and similarly men could not partake in some of the activities that were traditionally reserved for women. This has given birth to the gender role stereotypes that we find today. These differences have been passed on to our current times; although many differences occur now that have caused a lot of debate amongst the people as to their appropriateness and have made it possible for us to have a stereotyping threat by which we sometimes assign certain qualities to certain people without thinking. For example, many men are blamed for undermining women and stereotyping them for traditional roles, and this could be said to be the same for men; men are also stereotyped in many of their roles. This leads to social constructionism since the reality is not always depicted by what we see by our eyes. These ideas have also carried on in the world of advertising and the differences shown between the males and the females are apparent in many advertisements we see today. This can have some serious impacts on the society as people begin to stereotype the gender roles in reality.

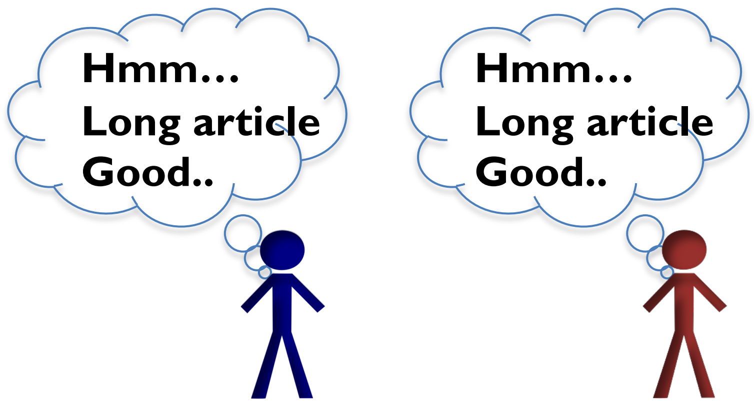
There has been a lot of attention given to the portrayal of gender in advertising by both practitioners as well as academics and much of this has been done regarding the portrayal of women in advertising (Ferguson, Kreshel, & Tinkham 40-51; Bellizzi & Milner 71-79). This has led many to believe that most of the advertisements and their contents are sexist in nature. It has been noted by viewing various ads that women are shown as being more concerned about their beauty and figure rather than being shown as authority figures in the ads; they are usually shown as the product users. Also, there is a tendency in many countries, including the United States, to portray women as being subordinate to men, as alluring sex objects, or as decorative objects. This is not right as it portrays women as the weaker sex, being only good as objects.

At the same time, many of the ads do not show gender biases in the pictures or the graphics, but some bias does turn up in the language of the ad. "Within language, bias is more evident in songs and dialogue than in formal speech or when popular culture is involved. For example, bias sneaks in through the use of idiomatic expressions (man's best friend) and when the language refer to characters that depict traditional sex roles. One's normative interpretation of these results depends on one's ideological perspective and tolerance for the pace of change. It is encouraging that the limited study of language in advertising indicates that the use of gender-neutrality is commonplace. Advertisers can still reduce the stereotyping in ad pictures, and increase the amount of female speech relative to male speech, even though progress is evidenced. To the extent that advertisers prefer to speak to people in their own language, the bias present in popular culture will likely continue to be reflected in advertisements" (Artz et al 20).

Advertisements are greatly responsible for eliciting such views for the people of our society. The children also see these pictures and they are also the ones who create stereotypes in their minds about the different roles of men and women. All these facts combine to give result to the different public opinion that becomes fact for many of the members of the society. Their opinion and views are based more on the interpretation they conclude from the images that are projected in the media than by their observations of the males and females in real life. This continues in a vicious circle as the media tries to pick up and project what the society thinks and the people in the society make their opinions based upon the images shown by the media. People, therefore, should not base too much importance about how the media is trying to portray the members of the society; rather they should base their opinions on their own observation of how people interact together in the real world.

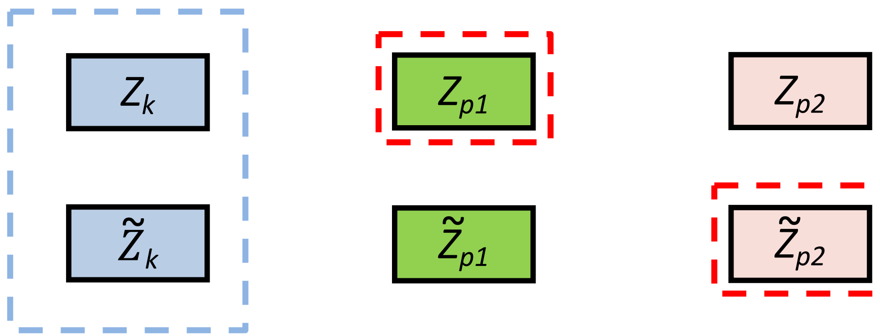
Work Cited

Artz, N., Munger, J., and Purdy, W., "Gender Issues in Advertising Language", Women and Language, 22, (2), 1999.
Bellizzi, J. A., & Milner, L. "Gender positioning of a traditionally male-dominant product", Journal of Advertising Research, 31(3), 1991.
Ferguson, J. H., Kreshel, P. J., & Tinkham, S. F. "In the pages of Ms.: Sex role portrayals of women in advertising", Journal of Advertising, 19 (1), 1990.



- Cheap signal for collusion
- Common mistakes
- Better equilibrium for agents to follow

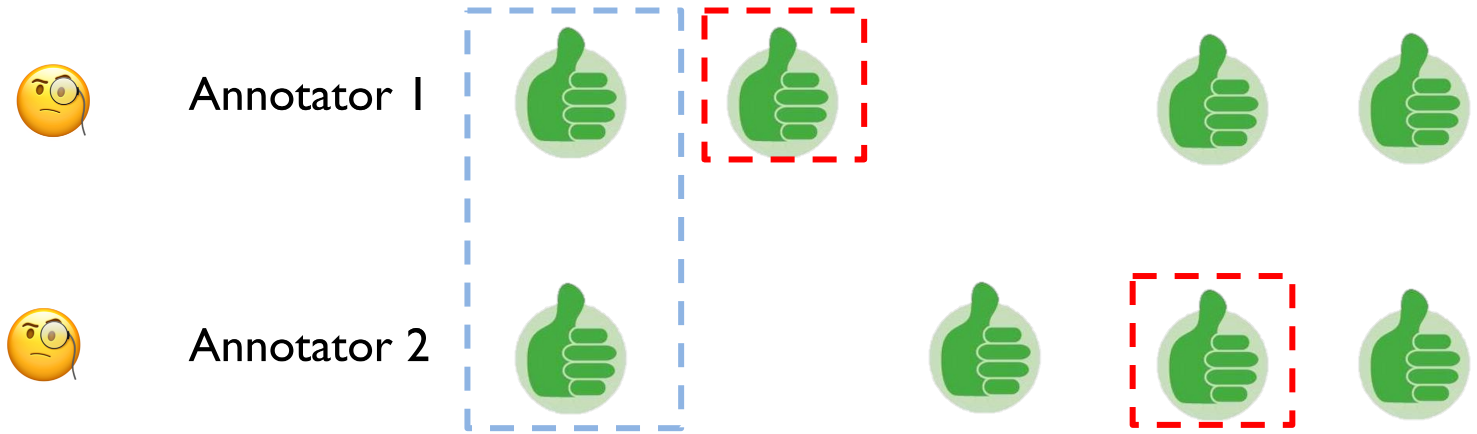
Correlated Agreement



$$\mathrm{Eval}(Z_k, \tilde{Z}_k) = \underbrace{\mathrm{Eval}(Z_k, \tilde{Z}_k)}_{\text{Evaluation on the same task}} - \boxed{\underbrace{\mathrm{Eval}(Z_{p_1}, \tilde{Z}_{p_2})}_{\text{Evaluation on different tasks}}}$$

Informed Truthfulness in Multi-Task Peer Prediction, Shnayder et al. EC 2016.

Correlated Agreement



Correlated agreement = 1 – 1 = 0

Why CA works?

Expectation of CA calibrates:

$$\boxed{\underbrace{\mathbb{P}(Z = \tilde{Z})}_{\text{Joint Distribution}}} - \boxed{\underbrace{\mathbb{P}(Z = \tilde{Z}_{\text{rand}})}_{\text{Punishment Distribution}}}$$

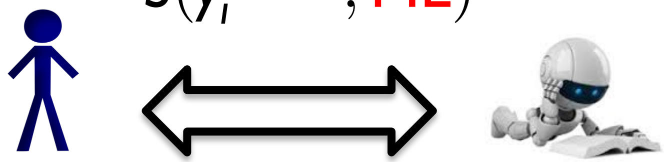
Agreement

Be different!

Theorem: CA induces truthful report at a Bayesian Nash Equilibrium.

ML aided Peer Prediction

$S(y_i^{\text{report}}, \text{ML})$



$f(\mathbf{x}) \rightarrow \mathbf{y}$

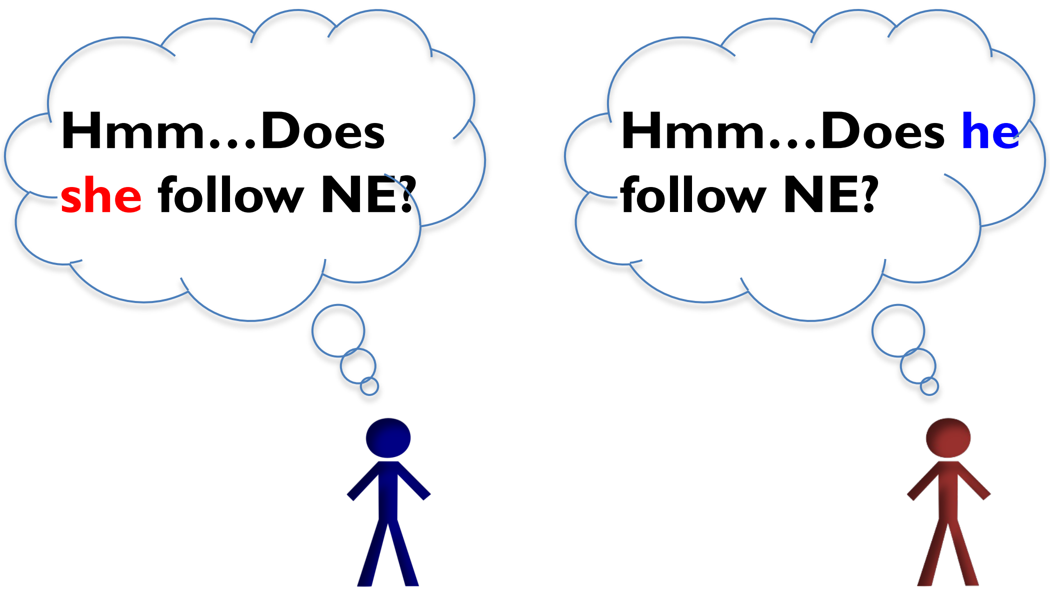
$$\mathbb{E}[S(\tilde{Y}_i, \text{ML}(\mathcal{I}_{-i}))] > \mathbb{E}[S(Z_i \neq \tilde{Y}_i, \text{ML}(\mathcal{I}_{-i}))]$$

- No **ground-truth** verification
- **Predict** via ML
- No requirement of others' report

Machine Learning aided Peer Prediction. Liu and Chen. EC 2017.

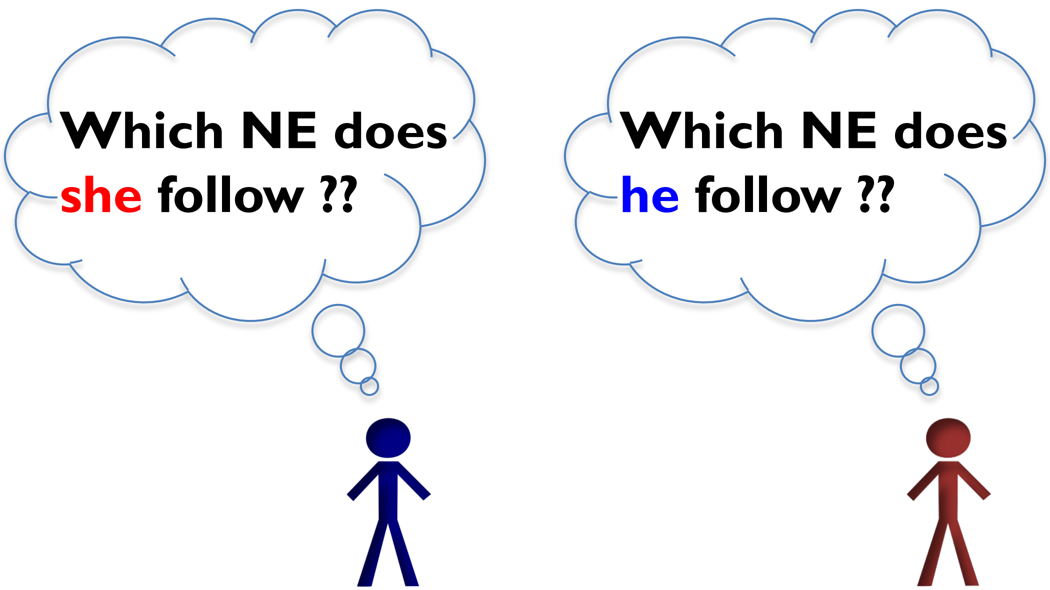
Caveats and What We Want

Equilibria v.s. **Dominant Truthfulness**



Caveats and What We Want

Equilibria v.s. **Dominant Truthfulness**



Surrogate Scoring Rules (SSR)

Previous work

Our work

Strictly Proper Scoring Rules

SSR

SSR Mechanism

Known ground truth

Noisy ground truth with known bias

No ground truth

✅ Truthfulness
✅ Reward accuracy

Surrogate Scoring Rules. Liu et al. EC 2020.

Surrogate Scoring Rules (SSR)

$$R(q_i, Z; e_Z^+, e_Z^-)$$

A noisy ground truth with known error rates:

- $e_Z^+ := \Pr(Z = 0 | Y = 1)$ (False negative rate)
- $e_Z^- := \Pr(Z = 1 | Y = 0)$ (False positive rate)

Surrogate Scoring Rules (SSR)

$$R(q_i, Z; e_z^+, e_z^-)$$

A noisy ground truth with known error rates:

- $e_z^+ := \Pr(Z = 0 | Y = 1)$ (False negative rate)
- $e_z^- := \Pr(Z = 1 | Y = 0)$ (False positive rate)

➢ Unbiasedness property (how SSR are defined):

$$\mathbb{E}_Z[R(q_i, Z; e_z^+, e_z^-)] = \mathbb{E}_Y[S(q_i, Y)], \forall q_i$$

An implementation of SSR

$$R(q_i, Z=1) = \frac{(1 - e_Z^-) \cdot S(q_i, 1) \; - \; e_Z^+ \cdot S(q_i, 0)}{1 - e_Z^- - e_Z^+}$$

$$R(q_i, Z=0) = \frac{-e_Z^+ \cdot S(q_i, 1) \; + \; (1 - e_Z^-) \cdot S(q_i, 0)}{1 - e_Z^- - e_Z^+}$$

An implementation of SSR

$$R(q_i, Z=1) = \frac{(1-e_Z^-)\cdot S(q_i,1) \; - \qquad e_Z^+ \cdot S(q_i,0)}{1-e_Z^- - e_Z^+}$$

$$R(q_i, Z=0) = \frac{-e_Z^+ \cdot S(q_i,1) \; + \; (1-e_Z^-)\cdot S(q_i,0)}{1-e_Z^- - e_Z^+}$$

Degenerate to SPSR when $e_Z^- = e_Z^+ = 0$

An implementation of SSR

$$R(q_i, Z = 1) = \frac{(1 - e_Z^-) \cdot S(q_i, 1) \; - \qquad e_Z^+ \cdot S(q_i, 0)}{1 - e_Z^- - e_Z^+}$$

$$R(q_i, Z = 0) = \frac{-e_Z^+ \cdot S(q_i, 1) \; + \; (1 - e_Z^-) \cdot S(q_i, 0)}{1 - e_Z^- - e_Z^+}$$

- Weights are designed for unbiasedness property:

> Lemma 1. For this implementation,
> $$\mathbb{E}_{Z|Y}[R(q_i, Z; e_Z^+, e_Z^-)] = S(q_i, Y), \forall q_i, Y$$

Roadmap

Previous work

Our work

Strictly Proper Scoring Rules

SSR

SSR Mechanism

Known ground truth

Noisy ground truth with known bias

No ground truth

✅ Truthfulness
✅ Reward accuracy

✅ Truthfulness
✅ Reward accuracy

Multi-task Setting

- N – A set of agents (index i)

Multi-task Setting

- N – A set of agents (index i)
- M – A set of tasks (index k)

Multi-task Setting

- N – A set of agents (index i)
- M – A set of tasks (index k)
- $Y_k \in \{0,1\}$ – the ground truth of task k

Multi-task Setting

- N – A set of agents (index i)
- M – A set of tasks (index k)
- $Y_k \in \{0,1\}$ – the ground truth of task k
- $p_{i,k}$, $q_{i,k}$ – the belief/report of agent i on task k

Multi-task Setting

- N – A set of agents (index i)
- M – A set of tasks (index k)
- $Y_k \in \{0,1\}$ – the ground truth of task k
- $p_{i,k}$, $q_{i,k}$ – the belief/report of agent i on task k
- Each task is assigned to at least 3 agents

SSR Mechanisms

When score a prediction $q_{i,k}$:

$$\text{SSR: } R(q_{i,k}, Z; e_z^+, e_z^-)$$

SSR Mechanisms

When score a prediction $q_{i,k}$:

$$\text{SSR: } R(q_{i,k}, Z; e_z^+, e_z^-)$$

- ➢ Construct Z
- ➢ Estimate e_z^+, e_z^-
- ➢ Apply SSR

SSR Mechanisms

When score a prediction $q_{i,k}$:

SSR: $R(q_{i,k}, Z; e_z^+, e_z^-)$

1

- Construct Z
- Estimate e_z^+, e_z^-
- Apply SSR

SSR Mechanisms

When score a prediction $q_{i,k}$:

$$\text{SSR: } R(q_{i,k}, Z; e_Z^+, e_Z^-)$$

1

2

- Construct Z
- Estimate e_Z^+, e_Z^-
- Apply SSR

SSR Mechanisms

When score a prediction $q_{i,k}$:

SSR: $R(q_{i,k}, Z; e_z^+, e_z^-)$

- Construct Z
- Estimate e_z^+, e_z^-
- Apply SSR

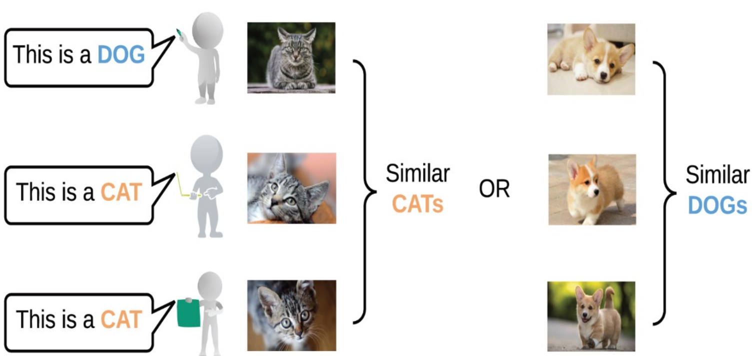
Construct Z

- For a task k, uniformly randomly pick an agent $j \neq i$, draw $Z = 1$ with probability $q_{j,k}$

the report of agent j on task k

HOC

Using **h**igh-**o**rder **c**onsensus to infer the noise transition matrix



Surrogate Scoring Rules, ACM EC 2020. **Liu**, Wang and Chen.

Clusterability as an Alternative to Anchor Points When Learning with Noisy Labels, ICML 2021. Zhu, Song, **Liu.** **Best paper award at IJCAI workshop on weakly supervised representation learning.**

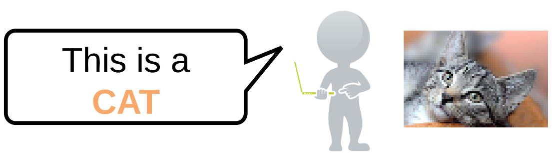
Data → (**without training**) → **Noise rate**

Solver and Implementation: https://github.com/UCSC-REAL/HOC

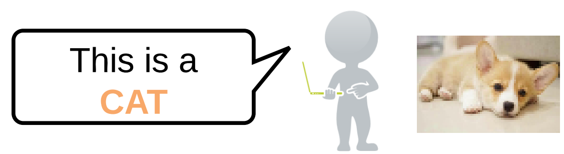
Calculate the probability

- Binary classification: Cat or Dog
- 1st-order (2 patterns)

Pattern "**CAT**"



+



True: CAT. Report: CAT (Correct)

$$p \cdot (1 - e_{\mathrm{cat}})$$

True: DOG. Report: CAT (Wrong)

$$(1 - p) \cdot e_{\mathrm{dog}}$$

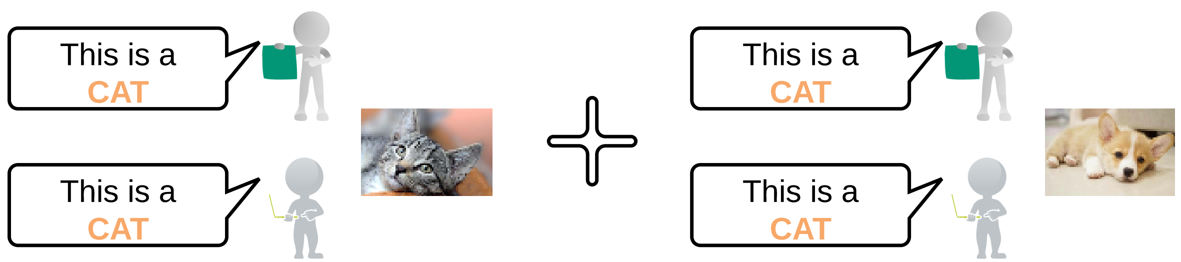
Population of true Cat

Noise rate of class Cat

Calculate the probability (Binary example)

- 2nd-order (4 patterns)

Pattern "(CAT, CAT)"



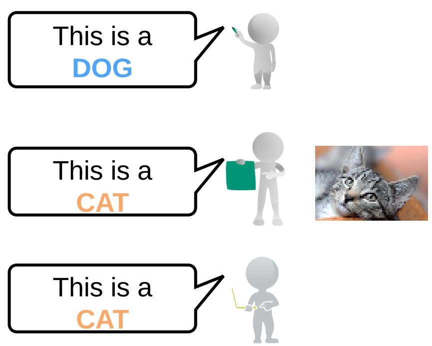
True: CAT. Report: CAT, CAT (Correct) True: DOG. Report: CAT, CAT (Wrong)

$$\underbrace{\text{True: CAT. Report: CAT, CAT (Correct)}}_{p \cdot (1 - e_{\text{cat}})^2} + \underbrace{\text{True: DOG. Report: CAT, CAT (Wrong)}}_{(1-p) \cdot e_{\text{dog}}^2}$$

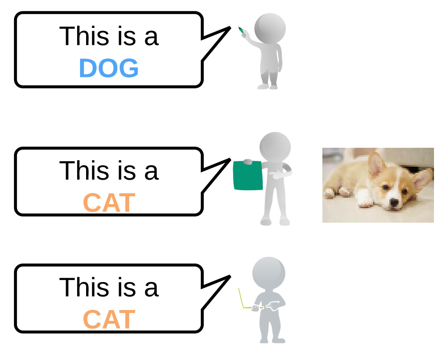
Calculate the probability (Binary example)

- 3rd-order (8 patterns)

Pattern "(**DOG**, **CAT, CAT**)"



+



True: CAT. Report: DOG, CAT, CAT

$$p \cdot e_{\mathrm{cat}} \cdot (1 - e_{\mathrm{cat}})^2$$

True: DOG. Report: DOG, CAT, CAT

$$(1 - p) \cdot (1 - e_{\mathrm{dog}}) \cdot e_{\mathrm{dog}}^2$$

High-Order Consensuses (HOC)

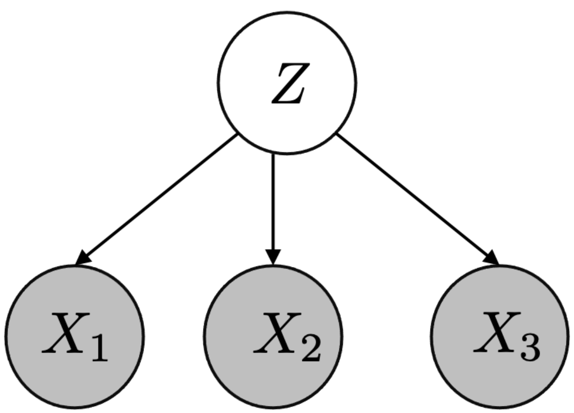
Consensus Equations

- 1st-order (K equations): $\boldsymbol{c}^{[1]} := \boldsymbol{T}^{\top}\boldsymbol{p}$
- 2nd-order (K^2 equations): $\boldsymbol{c}_r^{[2]} := (\boldsymbol{T} \circ \boldsymbol{T}_r)^{\top}\boldsymbol{p},\ r \in [K]$
- 3rd-order (K^3 equations): $\boldsymbol{c}_{r,s}^{[3]} := (\boldsymbol{T} \circ \boldsymbol{T}_r \circ \boldsymbol{T}_s)^{\top}\boldsymbol{p},\ r, s \in [K]$

$\mathbf{T} :=$ Noise transition matrix

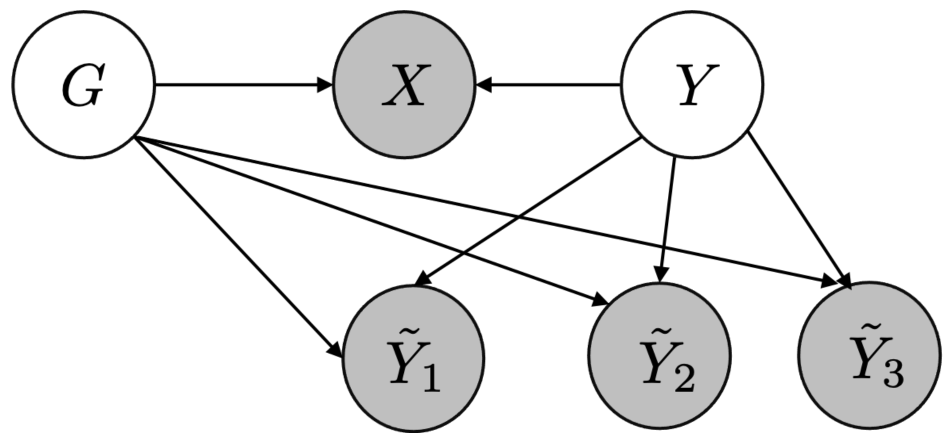
Three noisy labels are necessary and sufficient

Kruscal's Identifiability results: the identifiability of unobserved model ***Z* -> *X*** relies on the informativeness of three observations ***X*.**



Kruskal, J. B. Linear algebra and its applications, 18(2):95–138, 1977.

Our theorems: (1) Three labels are necessary and sufficient at instance level (2) Informative features help too.



Identifiability of Label Noise Transition Matrix, **Liu**, 2022.

Apply SSR

SSR property: $\mathbb{E}_{Z|Y_k}\left[R\left(p_{i,k}, Z; \widehat{e_Z^+}, \widehat{e_Z^-}\right)\right] = S(\mathrm{p}_{i,k}, Y_k)$

SSR mechanisms inherit two properties of SPSR:

- Incentivizing truthful reporting
- Accurate predictions get higher rewards

Theorem 1. Under A1~A4, in SSR mechanisms, truthful reporting is the uniform dominant strategy when $\mathrm{M}, \mathrm{N} \to \infty$

Apply SSR

SSR property: $\mathbb{E}_{Z|Y_k}\left[R\left(p_{i,k}, Z; \widehat{e_Z^+}, \widehat{e_Z^-}\right)\right] = S(\mathrm{p}_{i,k}, Y_k)$

SSR mechanisms inherit two properties of SPSR:

- Incentivizing truthful reporting
- Accurate predictions get higher rewards

Theorem 1. Under A1~A4, in SSR mechanisms, truthful reporting is the uniform dominant strategy when $\mathrm{M}, \mathrm{N} \rightarrow \infty$

$$\epsilon \sim 0\left(\frac{1}{\mathrm{N}} + \frac{1}{\sqrt{M}}\right)$$

Other challenges & extensions

Learning to design optimal mechanism

Workers are effort sensitive $e_i \in \{H, L\}$

- Exerting effort leads to better data

$$p_{i,e_i} = \Pr(s' = s | s, e_i) \quad 1 \geq p_H > p_L \geq 0.5$$

- Exerting effort incurs **unknown** cost

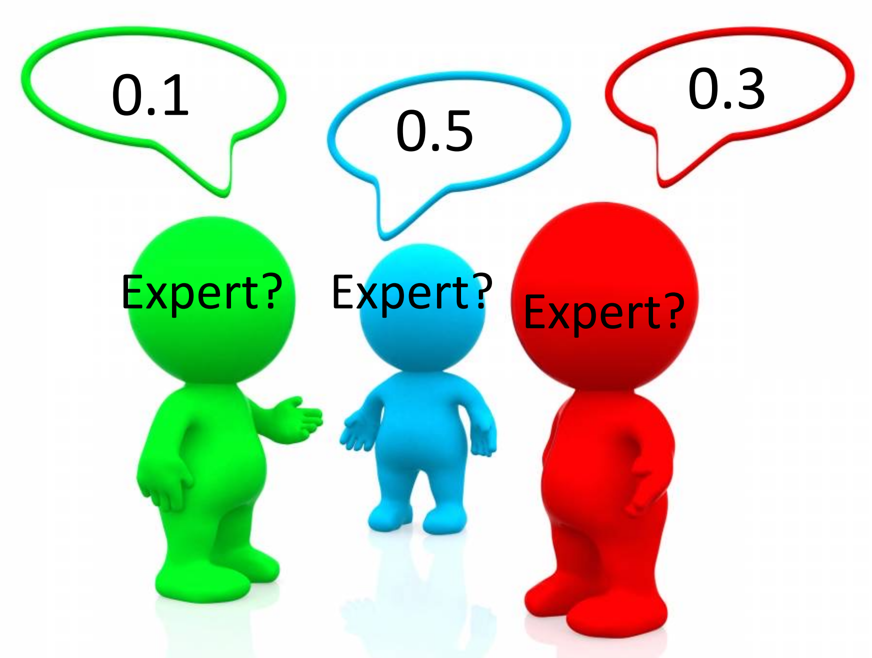
$$c \in F(c),\ c \in [0, c_{\max}]$$

- **Goal**: learnig to design best payment mechanism

Sequential Peer Prediction: Learning to Elicit Effort using Posted Prices. Liu and Chen. AAAI 2017.

Aggregation using peer prediction

Use peer prediction mechanisms to identify experts



Forecast Aggregation via Peer Prediction. Wang et al. HCOMP 2021.

Online Aggregation Using Peer Prediction

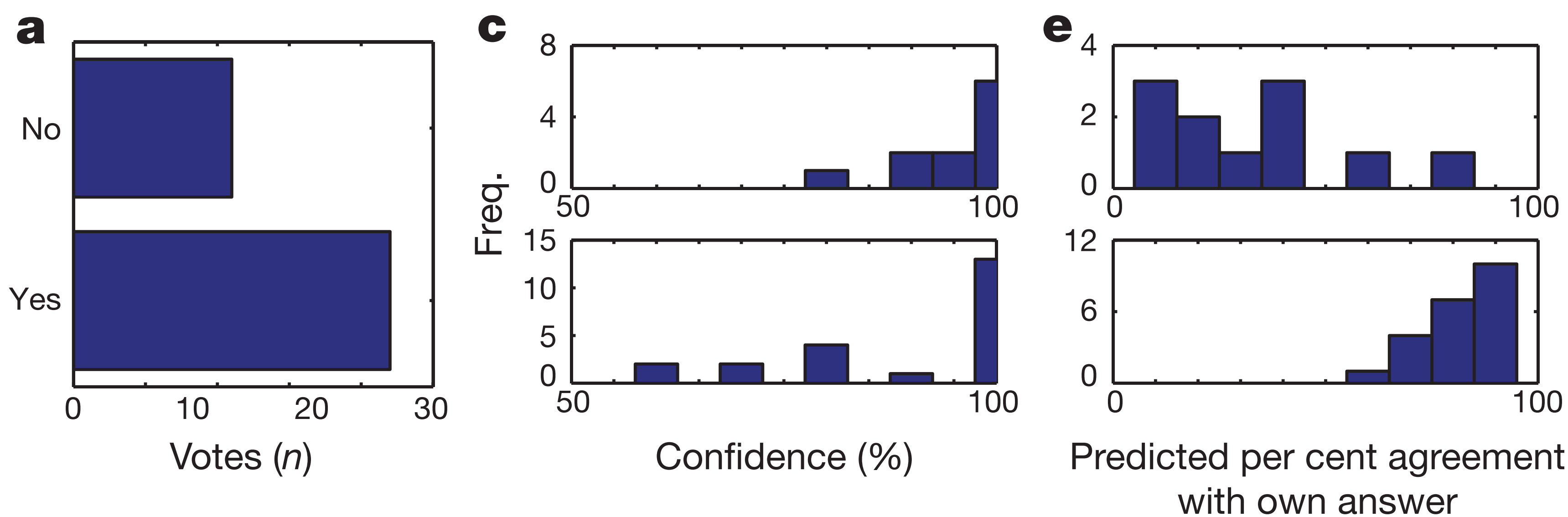
- Maintain a set of weights for each agent $w_i(t)$
- Aggregate: $\sum_{i=1}^{N} \frac{w_i(t)}{\sum_j w_j(t)} p_i(t)$
- At each round t, observe outcome of the event; each agent i incurs a score $(S_t(p_i(t), y_t))$
- Update weight of i using $w_i(t+1) := w_i(t) \cdot (1 + \eta \cdot S_t(p_i(t), y_t))$

How to compute the weights w.o. y_t??

Online Learning Using Only Peer Prediction. Liu and Helmbold. AISTATS 2020.

Aggregation Using BTS

Philadelphia is the capital of Pennsylvania, yes or no?



Answer "No" received: [Vote: ~30%, Peer Prediction: ~20%]

30% < 50%,but

30% > 20% <= surprisingly more popular

We are developed a machine learning version of this method.

Takeaways

- A lot of practical challenges
  - budget constraints
  - human-level information structure
- A lot of practical concerns
  - Interpretability of the mechanisms
- Human-in-the-loop => Machine-in-the-loop

Questions?

yangliu@ucsc.edu

Aggregation Using BTS

An answer is the correct answer only if (surprisingly popular)

Percentage of the answer > Percentage of peer predicted answer

Is this a fake news?

What's your answer? How many others would agree with you?

Expert 1: (NO, 20%) (*I know I have the minority answer*)

Expert 2: (YES,70%) (*Easy, most ppl know*)

Expert 3: (YES, 80%) (*Easy,most ppl know*)

NO: 0.3333 > (20%+30%+20%)/3 = 0.2333 (TRUTH)

YES: 0.6666 < (0.8+0.7+0.8)/3 = 0.7777