# The Value and Pricing of Information

**Haifeng Xu**

Assistant Professor in Computer Science

University of Chicago

Previously: How to quantify value of information

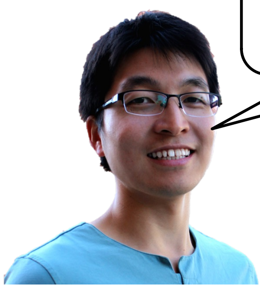
Next: How to price information based on its economic value

# How to Sell My Information Optimally?

Anyone attending Haifeng's talk gets $\$2t$ – if correctly guess his coin toss

AAAI
Association for the Advancement of Artificial Intelligence

I tossed, and learnt the side



- Suppose $t \sim U[0, 100]$; realized value known to James but not me
- Value of my information = $t$
- Post optimal price $p^* = \arg\max_p p \times \boxed{\frac{100-p}{100}} = 50?$
- Sub-optimal!

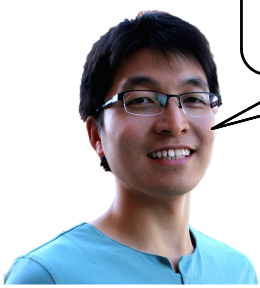
prob of purchase

# How to Sell My Information Optimally?

Anyone attending Haifeng's talk gets $\$2t$ – if correctly guess his coin toss

AAAI
Association for the Advancement of Artificial Intelligence

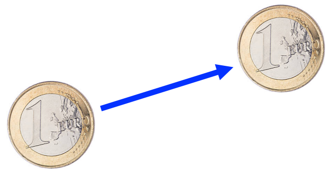
I tossed, and learnt the side



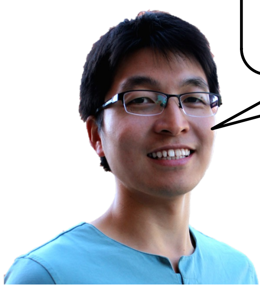
**Question**: What goes wrong?

- Information can be sold in complicated ways
- Here, can add noise to my answer

# How to Sell My Information Optimally?



I tossed, and learnt the side

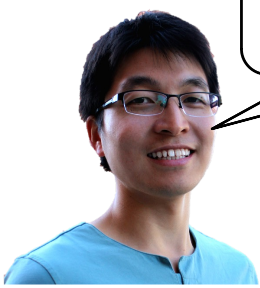


Full information revelation

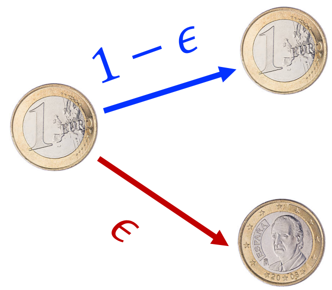
**Question**: What goes wrong?

- Information can be sold in complicated ways
- Here, can add noise to my answer

# How to Sell My Information Optimally?



I tossed, and learnt the side




Noisy information revelation

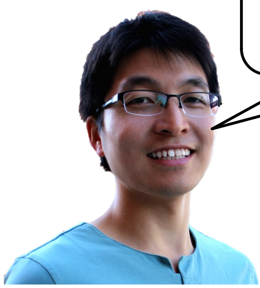
(may adapt noise level based on $t$)

**Question**: What goes wrong?

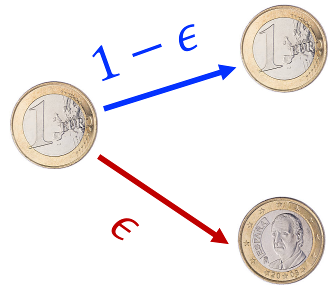
- Information can be sold in complicated ways
- Here, can add noise to my answer

This provides much power for price discrimination – can use different noise level for different $t$

# How to Sell My Information Optimally?



I tossed, and
learnt the side




Noisy information revelation

Fine…but why I should care about this problem?

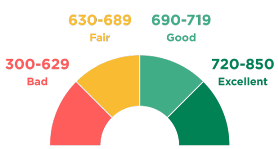
# Applications of Information Pricing


Car/house inspections

House buyers


Financial advices

Investors




Credit report

Loan companies

Consumer data

Small business owners

. . . . .

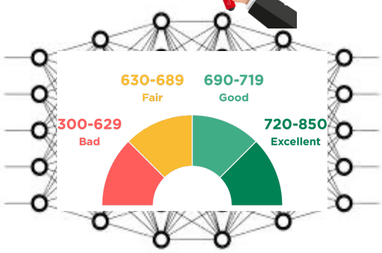
# Applications of Information Pricing

Become more relevant with ML technology


Car/house inspections — House buyers


Financial advices — Investors

Predict default rate ←




Credit report — Loan companies

Predict conversion rate ←

Consumer data — Small business owners

. . . . .

# Applications of Information Pricing

## Pricing for AutoML models

For Vertex AI AutoML models, you pay for three main activities:

- Training the model
- Deploying the model to an endpoint
- Using the model to make predictions


VIDEO

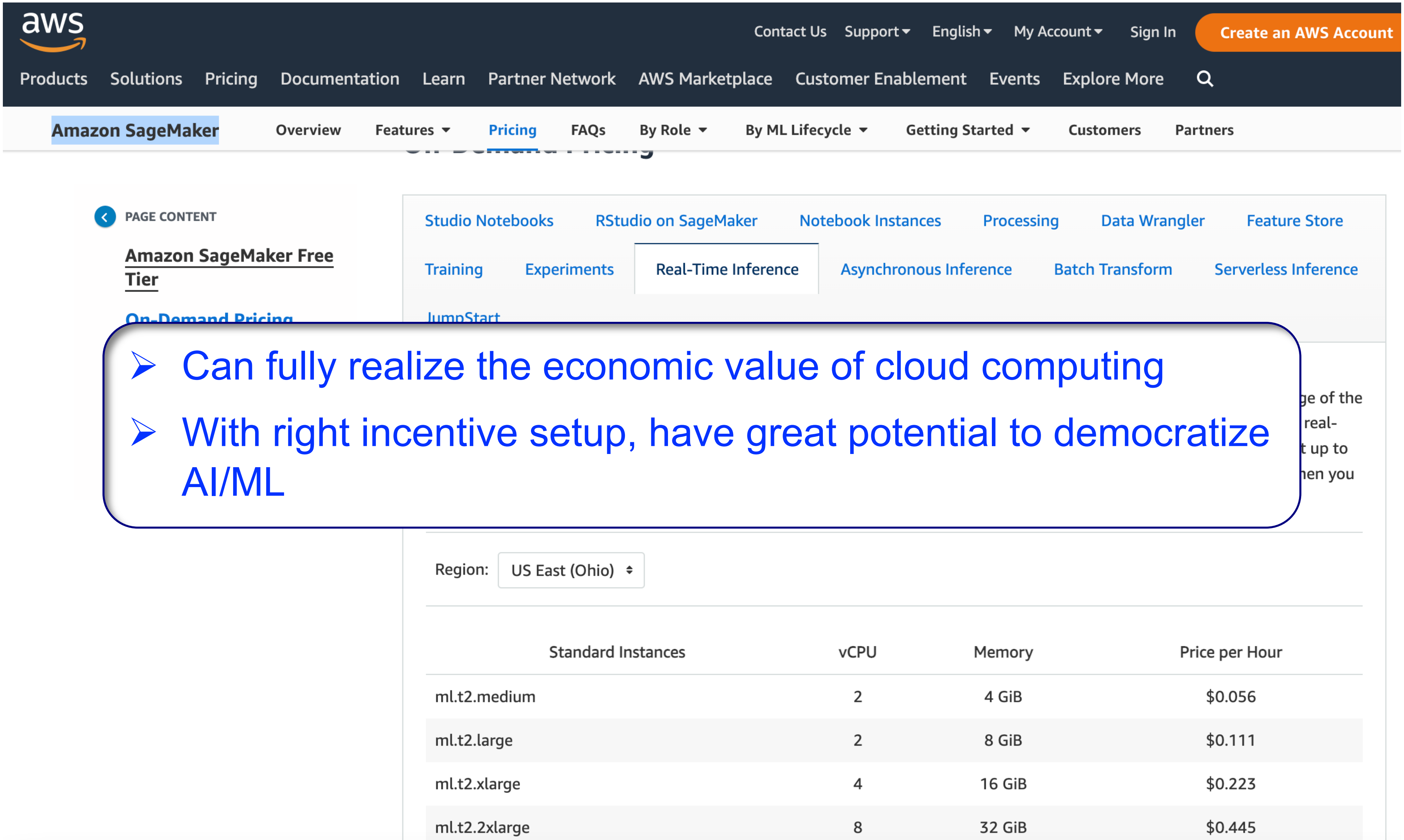
Accelerate ML experimentation and deployment with Vertex AI

Select a model type below for pricing information.

Image data | Video data | Tabular data | Text data

| Operation | Price per node hour (classification) | |
|---|---|---|
| Training | $3.465 | $3.465 |
| Training (Edge on-device model) | $18.00 | $18.00 |
| Deployment and online prediction | $1.375 | $2.002 |
| Batch prediction | $2.222 | $2.222 |

# Applications of Information Pricing




| Standard Instances | vCPU | Memory | Price per Hour |
|---|---|---|---|
| ml.t2.medium | 2 | 4 GiB | $0.056 |
| ml.t2.large | 2 | 8 GiB | $0.111 |
| ml.t2.xlarge | 4 | 16 GiB | $0.223 |
| ml.t2.2xlarge | 8 | 32 GiB | $0.445 |

- Can fully realize the economic value of cloud computing
- With right incentive setup, have great potential to democratize AI/ML

# Plans

- Vignette 1: closed-form optimal mechanism for structured setups
- Vignette 2: algorithmic solution for general setups
- Vignette 3: from distilled data (i.e. information) to raw data

By no means to be comprehensive;
Mainly to introduce the research flavors

# A Model of Information Pricing

- One seller, one buyer
- Buyer is a decision maker who faces a binary choice: an active action 1 and a passive action 0
  - Active action: come to talk, approve loan, invest stock X, etc.
- Payoff of passive action $\equiv 0$
- Payoff of active action $= v(\omega, t) = v_1(\omega)[t + \rho(\omega)]$
  - $\omega$ is a *state of nature*, $t$ is buyer type
  - Assume $v(\omega, t)$ is linear in $t \in [t_1, t_2]$

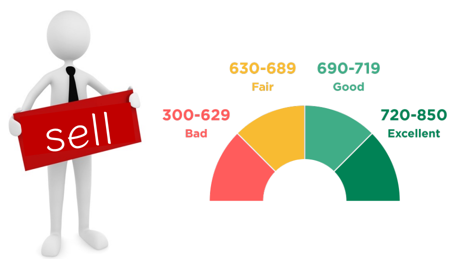
Results generalize to convex $v(\omega, t)$

# A Model of Information Pricing

- One seller, one buyer
- Buyer is a decision maker who faces a binary choice: an active action 1 and a passive action 0
  - Active action: come to talk, approve loan, invest stock X, etc.
- Payoff of passive action $\equiv 0$
- Payoff of active action $= v(\omega, t) = v_1(\omega)[t + \rho(\omega)]$
  - $\omega$ is a *state of nature*, $t$ is buyer type
  - Assume $v(\omega, t)$ is linear in $t \in [t_1, t_2]$
- Information structure:
  - Seller observes $\omega$, and buyer knows $t$

**Mechanism design question**: How can seller optimally sell her information about $\omega$ to the buyer?

# An Example

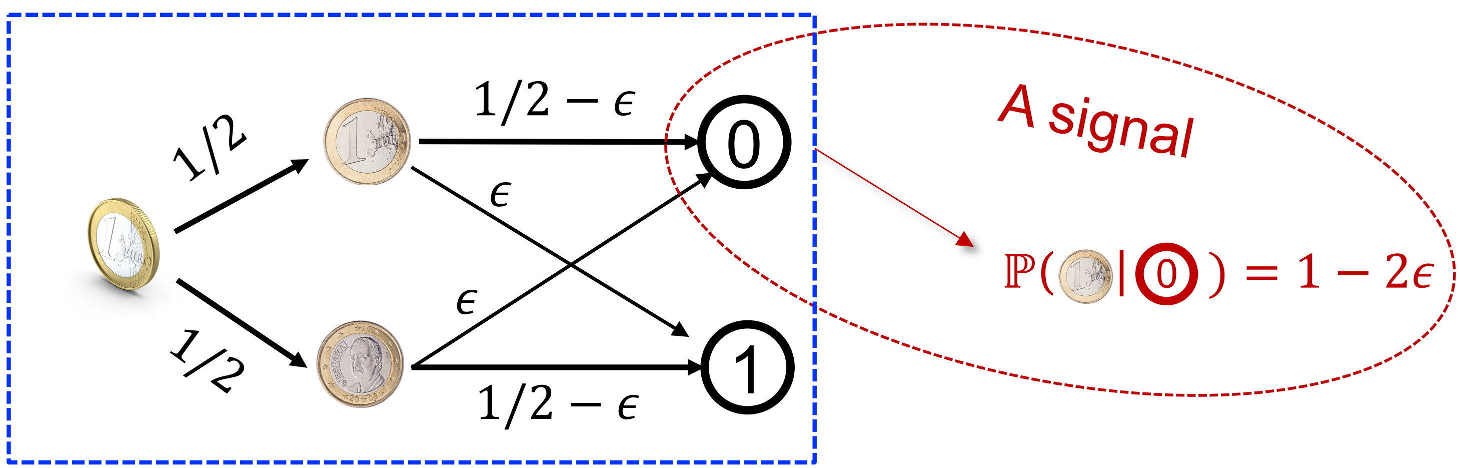




- Buyer is a loan company; action is to approve a loan or not
  - If not approving (action 0), payoff is 0
  - If approving (action 1), payoff is

$$v(\omega, t) = (1 - \underbrace{\omega}_{\substack{\omega \in [0,1] \\ \text{default probability}}})\times \underbrace{t}_{\text{Revenue}} - \underbrace{2}_{\text{operation cost}}$$

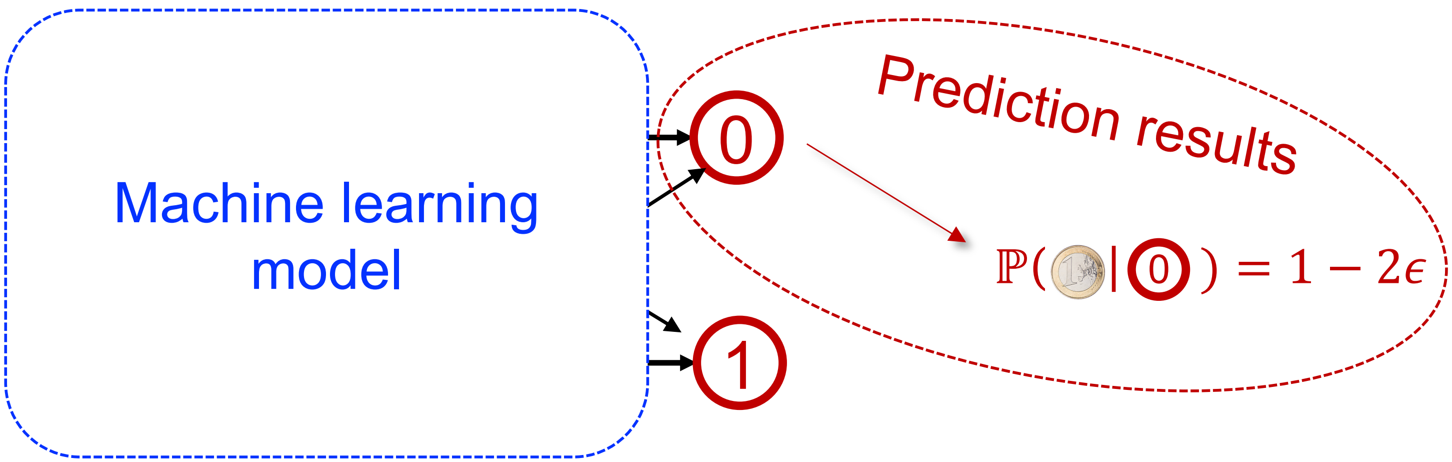
# Selling Signals or Signal Generation Process?




Signal generation process

econ/stat terminology: experiment or signaling scheme

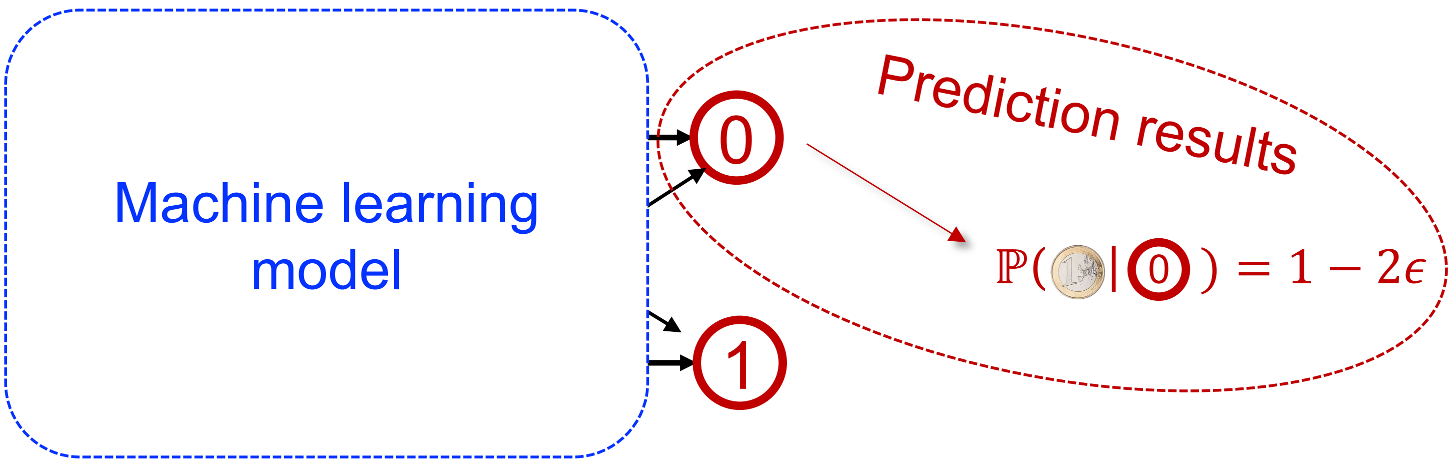
# Selling Signals or Signal Generation Process?




For computer scientists: ML model itself or its prediction results

# Selling Signals or Signal Generation Process?

Next, for convenience, sell "information experiments"




In a *perfect Bayesian world*, these two turns out to be equivalent – price for the model is viewed as expected price over predictions

# Design Space

- Standard revelation principle implies optimal mechanism can w.l.o.g be a menu $\{\pi_t, p_t\}_{t \in T}$
  - $\pi_t: \Omega \to S$ is an experiment (which generates signals) for type $t$
  - $p_t \in \mathbb{R}$ is $t$'s payment
  - Each type is incentivized to report type truthfully

**Concrete design question**: design IC $\{\pi_t, p_t\}_{t \in T}$ to maximize seller's revenue

# How Does It Differ from Selling Goods?

Key differences:

- Each experiment is like an item
  - In this sense, we are selling infinitely many goods
  - In fact, we are even "designing the goods"
- Participation constraint is different
  - Without any information, type $t$'s utility is $\max\{\,\bar{v}(t), 0\,\}$

$$\bar{v}(t) = \int_{\omega\in\Omega} v(\omega, t)\, g(\omega) d\omega$$

Ex-ante expected utility of action 1

# Threshold experiments turn out to suffice

Recall $v(\omega, t) = v_1(\omega)[t + \rho(\omega)]$

**Def.** $\pi_t$ is a threshold experiment if $\pi_t$ simply reveals $\rho(\omega) \geq \theta(t)$ or not for some buyer-type-dependent threshold $\theta(t)$

- Threshold is on $\rho(\omega)$

# Virtual Value Functions

- Recall virtual value function in [Myerson'81]: $\phi(t) = t - \frac{1-F(t)}{f(t)}$

**Def.** Lower virtual value function: $\underline{\phi}(t) = t - \frac{1-F(t)}{f(t)}$

# Virtual Value Functions

- Recall virtual value function in [Myerson'81]: $\phi(t) = t - \frac{1-F(t)}{f(t)}$

**Def.** Lower virtual value function: $\underline{\phi}(t) = t - \frac{1-F(t)}{f(t)}$

Upper virtual value function: $\bar{\phi}(t) = t + \frac{F(t)}{f(t)}$

Mixed virtual value function: $\phi_c(t) = c\underline{\phi}(t) + (1-c)\bar{\phi}(t)$

Note: "upper" or "lower" is due to

$$\underline{\phi}(t) \le t \le \bar{\phi}(t)$$

# The Optimal Mechanism

Depend on two problem-related constants:

$$
\begin{aligned}
V_L &= \max\{v(t_1), 0\} + \int_{t_1}^{t_2} \int_{\omega:\alpha(\omega) \geq -\underline{\phi}(x)} g(\omega)\alpha(\omega)\,\mathrm{d}\omega\mathrm{d}x, \\
V_H &= \max\{v(t_1), 0\} + \int_{t_1}^{t_2} \int_{\omega:\alpha(\omega) \geq -\overline{\phi}(x)} g(\omega)\alpha(\omega)\,\mathrm{d}\omega\mathrm{d}x,
\end{aligned}
$$

Note: $V_L < V_H$

# The Optimal Mechanism

**Theorem ([LSX'21]).**

**1.** If $\bar{v}(t_2) \leq \mathrm{V}_L$, the mechanism with threshold experiments $\theta^*(t) = -\underline{\phi}(t)$ and following payment function represents an optimal mechanism:

$$p^*(t) = \int_{\omega \in \Omega} \pi^*(\omega, t) g(\omega) v(\omega, t) \mathrm{d}\omega - \int_{t_1}^{t} \int_{\omega \in \Omega} \pi^*(\omega, x) g(\omega) v_1(\omega) \mathrm{d}\omega \, \mathrm{d}x$$

Shuze Liu, Weiran Shen and Haifeng Xu, *Optimal Pricing of Information*, Proc. 22th ACM Conference on Economics and Computation (EC 2021)

# The Optimal Mechanism

**Theorem ([LSX'21]).**

**1.** If $\bar{v}(t_2) \leq \mathrm{V}_L$, the mechanism with threshold experiments $\theta^*(t) = -\underline{\phi}(t)$ and following payment function represents an optimal mechanism:

$$p^*(t) = \int_{\omega\in\Omega} \pi^*(\omega,t)g(\omega)v(\omega,t)\mathrm{d}\omega - \int_{t_1}^{t}\int_{\omega\in\Omega} \pi^*(\omega,x)g(\omega)v_1(\omega)\mathrm{d}\omega\,\mathrm{d}x$$

**2.** If $\bar{v}(t_2) \geq \mathrm{V}_H$, the mechanism with threshold experiments $\theta^*(t) = -\bar{\phi}(t)$ and following payment function represents an optimal mechanism:

$$p^*(t) = \int_{\omega\in\Omega} \pi^*(\omega,t)g(\omega)v(\omega,t)\mathrm{d}\omega + \int_{t}^{t_2}\int_{\omega\in\Omega} \pi^*(\omega,x)g(\omega)v_1(\omega)\mathrm{d}\omega\;\mathrm{d}x - \bar{v}(t_2)$$

Shuze Liu, Weiran Shen and Haifeng Xu, *Optimal Pricing of Information*, Proc. 22th ACM Conference on Economics and Computation (EC 2021)

# The Optimal Mechanism

**Theorem ([LSX'21]).**

**3.** If $V_L \le \bar{v}(t_2) \le V_H$, the mechanism with threshold experiments $\theta^*(t) = -\phi_c(t)$ and following payment function represents an optimal mechanism:

$$p^*(t) = \int_{\omega \in \Omega} \pi^*(\omega, t) g(\omega) v(\omega, t) \mathrm{d}\omega - \int_{t_1}^{t} \int_{\omega \in \Omega} \pi^*(\omega, x) g(\omega) v_1(\omega) \mathrm{d}\omega \, \mathrm{d}x$$

where constant $c$ is chosen such that

$$\int_{t_1}^{t_2} \int_{\omega : \rho(\omega) \ge \phi_c^+(x)} g(\omega) v_1(\omega) \mathrm{d}\omega \, \mathrm{d}x = \bar{v}(t_2)$$

Shuze Liu, Weiran Shen and Haifeng Xu, *Optimal Pricing of Information*, Proc. 22th ACM Conference on Economics and Computation (EC 2021)

# Remarks

- Threshold mechanisms are common in real life
  - House/car inspections, stock recommendations: information seller only need to reveal it “passed” or “deserves a buy” or not

- Optimal mechanism has personalized thresholds and payments, tailored to accommodate different level of risk each buyer type can take
  - Different from optimal pricing of physical goods

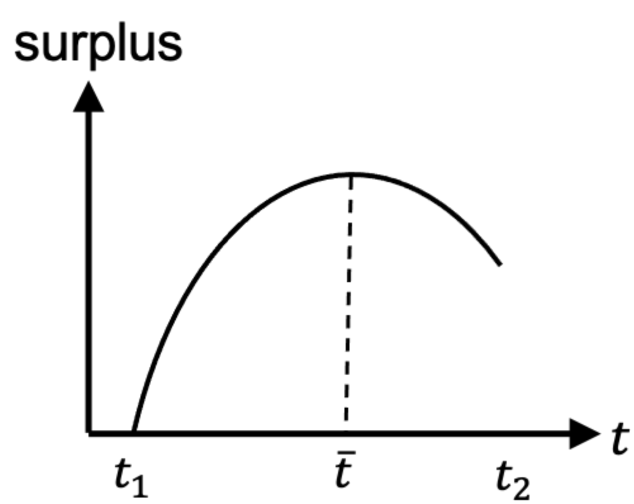
# Remarks

What if seller is restricted to sell the same information to every buyer (e.g., due to regulation)? How will revenue change?

- This is the optimal price (Myerson reserve) in previous example
- Revenue can be arbitrarily worse
- $1/e$ -approximation of optimal revenue if the *value of full information* as a function of $t$ has *monotone hazard rate*

# Additional Properties of Optimal Mechanism

**Proposition 1 ([LSX'21]).** *Buyer surplus* is increasing for $t \in [t_1, \bar{t}]$ and decreasing for $t \in [\bar{t}, t_2]$ where $\bar{t}$ satisfies $\bar{v}(\bar{t}) = 0$.




Recall

$$\bar{v}(t) = \int_{\omega \in \Omega} v(\omega, t)\, g(\omega) d\omega$$

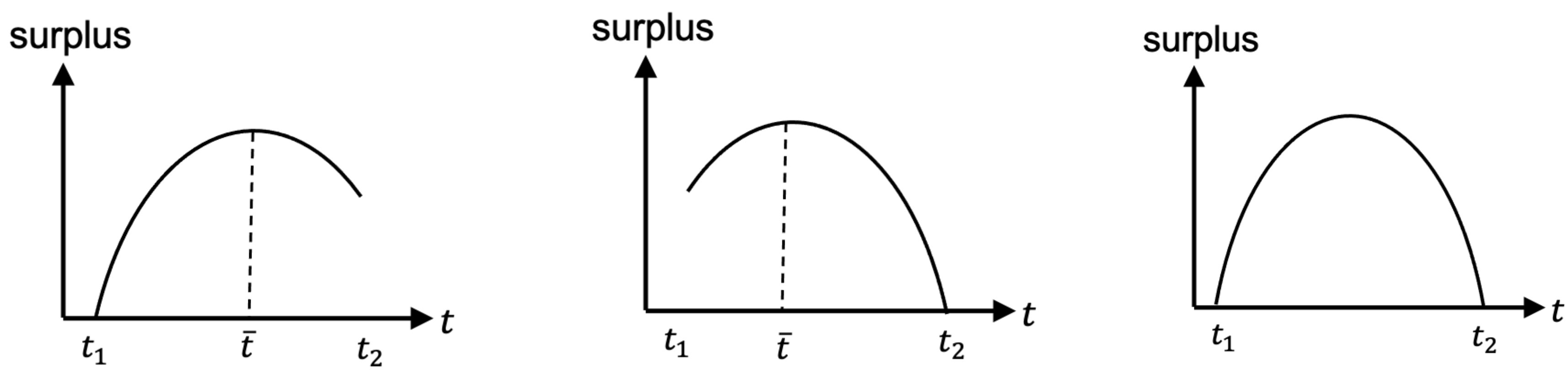
Ex-ante expected utility of action 1

Shuze Liu, Weiran Shen and Haifeng Xu, *Optimal Pricing of Information*, Proc. 22th ACM Conference on Economics and Computation (EC 2021)

# Additional Properties of Optimal Mechanism

**Prop. 2 ([LSX'21]).** Following properties hold in optimal mechanism.

1. In Case 1, surplus of $t_1$ is 0; In Case 2, surplus of $t_2$ is 0; In Case 3, surplus of both $t_1$ and $t_2$ is 0




**Case 1** **Case 2** **Case 3**

Shuze Liu, Weiran Shen and Haifeng Xu, *Optimal Pricing of Information*, Proc. 22th ACM Conference on Economics and Computation (EC 2021)

# Additional Properties of Optimal Mechanism

**Prop. 2 ([LSX'21]).** Following properties hold in optimal mechanism.

1. In Case 1, surplus of $t_1$ is 0; In Case 2, surplus of $t_2$ is 0; In Case 3, surplus of both $t_1$ and $t_2$ is 0
2. *Buyer payment* is increasing in Case 1, decreasing in Case 2, and increase first then decrease in Case 3

➢ These properties all differ from optimal mechanism for selling an item.

Shuze Liu, Weiran Shen and Haifeng Xu, *Optimal Pricing of Information*, Proc. 22th ACM Conference on Economics and Computation (EC 2021)

# A Case I Example

Previous credit score example

- $v(\omega, t) = (1 - \omega)t - 2$, $\omega \in [0,1]$, $t \in [2,3]$ both uniformly at random
- Easy to verify this is Case 1

**Optimal Mechanism**

1. For any buyer type $t \leq 2.5$, optimal mechanism charges $0$ and then reveals no information

# A Case I Example

Previous credit score example

- $v(\omega, t) = (1-\omega)t - 2$, $\omega \in [0,1]$, $t \in [2,3]$ both uniformly at random
- Easy to verify this is Case 1

Optimal Mechanism

1. For any buyer type $t \leq 2.5$, optimal mechanism charges $0$ and then reveals no information
2. For any buyer type $t > 2.5$, optimal mechanism charges $-\frac{1}{4} + \frac{4t-9}{(2t-3)^2}$ and then reveals $\omega \leq \frac{2t-5}{2t-3}$ or not

Note: optimal mechanism reveals no information to some buyer types

# Plans

- Vignette 1: closed-form optimal mechanism for structured setups
- Vignette 2: algorithmic solution for general setups
- Vignette 3: from distilled data (i.e. information) to raw data

# A Generalized Model of Selling Information

- Buyer takes one of $n$ action $a \in [n] = \{1, \cdots, n\}$
- Buyer has an arbitrary utility function $u(a, \omega; t)$

**Mechanism design question**: How can seller optimally sell her information about $\omega$ to the buyer?

- First studied by [Babaioff/Kleinberg/Paes Leme, EC'12], but mechanism is very complex and has extremely large payment

# Existence of Simple "Direct" Mechanisms

**Theorem (Revelation Principle, BBS'18, CXZ'20).** Any information selling mechanism is "equivalent" to a direct and truthful mechanism:

1. Ask buyer to report type $t$
2. Charge buyer $x_t$ and then directly make obedient action recommendation to buyer via a randomized scheme $\pi_t: Q \to [n]$

Moreover, the mechanism is incentive compatible (IC) – it is the buyer's best interest to truthfully report $t$

Dirk Bergemann Alessandro Bonatti Alex Smolin, *The Design and Price of Information*, American Economic Review' 18
Yiling Chen, Haifeng Xu, Shuran Zheng, *Selling Information through Consulting*, SODA 2020

# Existence of Simple "Direct" Mechanisms

**Theorem (Revelation Principle, BBS'18, CXZ'20).** Any information selling mechanism is "equivalent" to a direct and truthful mechanism:

1. Ask buyer to report type $t$
2. Charge buyer $x_t$ and then directly make obedient action recommendation to buyer via a randomized scheme $\pi_t: Q \to [n]$

Moreover, the mechanism is incentive compatible (IC) – it is the buyer's best interest to truthfully report $t$

Dirk Bergemann Alessandro Bonatti Alex Smolin, *The Design and Price of Information*, American Economic Review' 18
Yiling Chen, Haifeng Xu, Shuran Zheng, *Selling Information through Consulting*, SODA 2020

# This Optimal Mechanism is like Consulting!

Consulting Mechanism w/ Bounded Payment [CXZ '20]

1. Elicit buyer type $t$
2. Charge buyer $x_t \leq B$ (bounded payment)
3. Observe realized state $\omega$ and recommend (possibly randomly chosen) action $a$ to the buyer

**Theorem (CXZ'20)**. The optimal payment-limited consulting mechanism can be computed by a convex program.

Less interpretable than previous one, but at least simple to implement



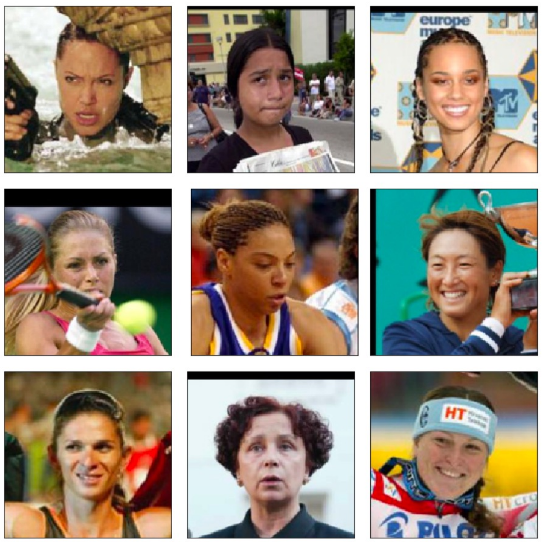
Yiling Chen, Haifeng Xu, Shuran Zheng, *Selling Information through Consulting*, SODA 2020

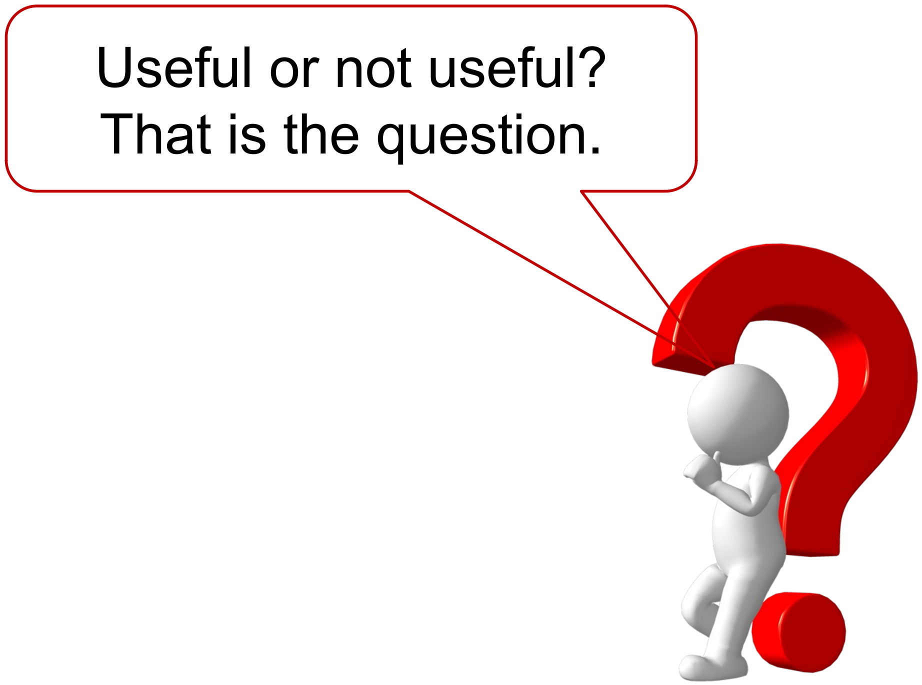
# Plans

- Vignette 1: closed-form optimal mechanism for structured setups
- Vignette 2: algorithmic solution for general setups
- Vignette 3: from distilled data (i.e. information) to raw data

# How to Sell Raw Data to a Machine Learner?

- Unlike prediction outcomes, usefulness of raw data is uncertain





# How to Sell Raw Data to a Machine Learner?

- Unlike prediction outcomes, usefulness of raw data is uncertain

Maybe we can use statistical methods to estimate data value?

- Not easily doable on market
- Statistical methods need to test on data, but if the learner already tried all your data, why she buys?
- Possible rescues: use a trustworthy third party, multi-party secure computation,…

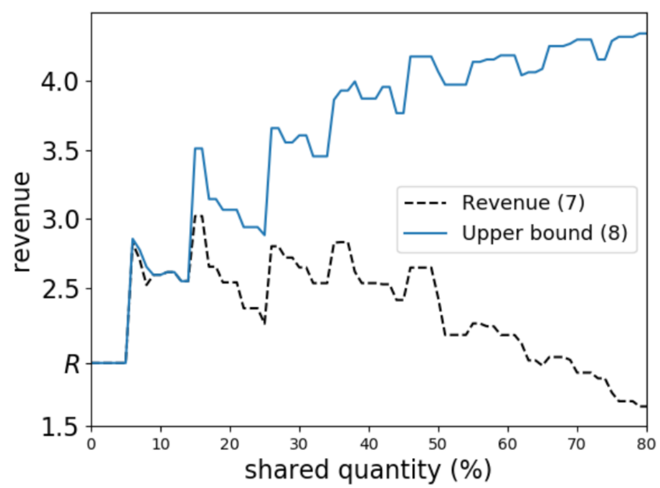
# How to Sell Raw Data to a Machine Learner?

➢ The rescue through better mechanism design

The "free-trial" mechanism [CLX, ICML'22]

1. Reveal a small portion of sample data to update buyer's belief about data usefulness
2. Sell remaining data

Key challenge: needs to figure out right amount of data to reveal






Junjie Chen, Minming Li and Haifeng Xu, *Selling Data To a Machine Learner: Pricing via Costly Signaling*, ICML 2022.

# Summary

- Raw and distilled data (i.e., information) both have economic values
- The pricing of data depends on its economic value
- There are progresses on pricing mechanisms for data/information
- But long way to go….

# Open Directions

- What if signals have error (e.g., predictions of ML algorithms)?
- What if the world is non-Bayesian? Difference between pricing signals vs pricing signal generation processes?
- What is the most practical/efficient/feasible way to sell data? Directly sell raw data, or sell ML model, or sell inferences? Or personalized?
- How to be robust to numerous uncertainty in data and ML models?
- …

## Pricing for AutoML

For Vertex AI AutoML models, you pay for three

- Training the model
- Deploying the model to an endpoint
- Using the model to make predictions

Select a model type below for pricing information.

Image data | Video data | Tabular data | Text data

| Operation | Price per node hour (classification) | Price per node hour (object d |
|---|---|---|
| Training | $3.465 | $3.465 |
| Training (Edge on-device model) | $18.00 | $18.00 |
| Deployment and online prediction | $1.375 | $2.002 |
| Batch prediction | $2.222 | $2.222 |

# References


1. M. Babaioff , R. Kleinberg and R. Paes Leme, *Optimal Mechanisms for Selling Information,* EC 2012
2. Alexander Frankel and Emir Kamenica, *Quantifying Information and Uncertainty*, American Economic Review 2019
3. Yiling Chen, Haifeng Xu, Shuran Zheng, *Selling Information through Consulting*, SODA 2020
4. Shuze Liu, Weiran Shen and Haifeng Xu, *Optimal Pricing of Information*, EC 2021
5. Dirk Bergemann Alessandro Bonatti Alex Smolin, *The Design and Price of Information*, American Economic Review' 18
6. Junjie Chen, Minming Li and Haifeng Xu, *Selling Data To a Machine Learner: Pricing via Costly Signaling*, ICML 2022.
7. Kimon Drakopoulos and Ali Makhdoumi, Providing Data Samples for Free, Management Science 2022

Haifeng: how to value and price distilled data

NEXT

Shuran: how to collect truthful data from strategic agents