Active Label Acquisition with Personalized Incentives in Assortment Optimization

Mo Liu

IEOR, University of California, Berkeley

2023 INFORMS Annual Meeting

Joint work with Prof. Junyu Cao and Prof. Zuo-Jun Max Shen

Agenda

- ➢ **Active Label Acquisition in Customer Survey**
- Regret of the Prediction Model
- Value of Information
- Upper bound for the value of information
- Guarantees for Assortment Optimization
- Numerical Experiments

Personalized decisions based on customer features

For each customer:

Feature → Prediction model → Preferences → Maximize the revenue → OM decisions

Personalized decisions based on customer features

For each customer:

Feature → Prediction model → Preferences → Maximize the revenue → OM decisions

↑

Personalized information of customers

Personalized decisions based on customer features

For each customer:

Feature → **Prediction model** → Preferences → **Maximize the revenue** → OM decisions

↑

Preferences between different products,
e.g. utility of each product

Personalized decisions based on customer features

For each customer:

Feature → Prediction model → Preferences → Maximize the revenue → OM decisions

↑

Assortment problem, product selection…

Personalized decisions based on customer features

For each customer:

Feature → Prediction model → Preferences → Maximize the revenue → OM decisions

How to build a prediction model?

Personalized decisions based on customer features

For each customer:

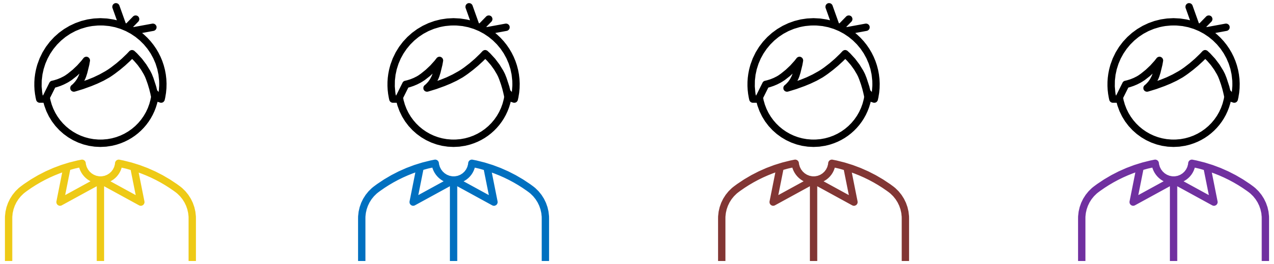
Feature → $\boxed{\text{Prediction model}}$ → Preferences → $\boxed{\text{Maximize the revenue}}$ → OM decisions

↑

Training set: (feature, preferences)
Preferences: label of the customer

How to obtain the true preferences of customers

- Survey customers:
  - Provide a comprehensive survey to customers
  - The response from one customer can reveal the true utility vector (with noise)



  - *"Without costly incentives, most consumers rarely provide this valuable feedback"*

---- by Maytal Saar-Tsechansky et al. (2009)

Incentives in active label acquisition

- Active label acquisition with personalized incentives:

Customer t arrives with type ξ_t →

Incentives in active label acquisition

- Active label acquisition with personalized incentives:

Customer t arrives with type ξ_t → Offer some **incentive** c_t to customer t for completing the survey →

Incentives in active label acquisition

- Active label acquisition with personalized incentives:

Customer t arrives with type ξ_t → Offer some **incentive** c_t to customer t for completing the survey → Customer t decides whether to accept the survey to get the incentive

- Probability of accepting the survey $p(c)$ depends on our offered incentives
  - More incentives we offer → Larger probability of taking the survey

- Can we provide same incentives to all customers?

Benefit of personalized incentives

Provide more incentives to representative customers

Compared to the fixed incentive policy, personalized incentives can:
- ✓ Reduce the size of the training set
- ✓ Reduce the label cost (cumulative incentives)

How to decide personalized incentives?

Personalized decisions based on customer features

Feature → **Prediction model** → Preferences → **Maximize the revenue** → OM decisions

Select customers based on the prediction errors for preferences

Smaller prediction error ? Higher revenue

Reasons:

If the prediction error is small enough to determine the true optimal decisions, then a smaller prediction error will lead to the same decision and obtain the same revenue

Personalized decisions based on customer features

Feature → $\boxed{\text{Prediction model}}$ → Preferences → $\boxed{\text{Maximize the revenue}}$ → OM decisions

Select customers based on the ~~prediction errors for preferences~~

↓

$\boxed{\text{Risk of OM decisions}}$

Personalized decisions based on customer features

Feature → Prediction model → Preferences → Maximize the revenue → OM decisions

~~Select customers~~ based on the ~~prediction errors for preferences~~

Personalized Incentives

Risk of OM decisions

? Behaviors of human

? Risk is a nonlinear of decisions

Agenda

- Active Label Acquisition in Customer Survey
- **Regret of the Prediction Model**
- Value of Information
- Upper bound for the value of information
- Guarantees for Assortment Optimization
- Numerical Experiments

Regret of prediction

In the predict-then-optimize problems:

- Regret of the prediction:
  - Highest possible revenue – the actual revenue of our decisions based on current prediction

❑ Assortment optimization problem:

➢ Revenue of the best assortment – actual revenue of our assortment

Regret of the prediction model

- Type of customer: $\xi \in \{1, \dots, m\}$
- Utility vector: $y \in \mathbb{R}^d$
- Decision vector: $w \in \{0,1\}^d$
- Revenue function: $g(w, \mathbb{E}[y|\xi])$

  – $w^*(y)$: Best decision given the prediction. $w^*(y) = \arg\max_w g(w, y)$

Regret of prediction $\hat{y}$:

$$\ell(\hat{y}, \mathbb{E}[y|\xi]) := \underbrace{\boxed{g(w^*(\mathbb{E}[y|\xi]), \mathbb{E}[y|\xi])}}_{\text{Highest revenue}} - \underbrace{\boxed{g(w^*(\hat{y}), \mathbb{E}[y|\xi])}}_{\text{Actual revenue}}$$

- Given a predictor h, the expected regret of the predictor:

$$\text{Regret}(h) = \mathbb{E}[\ell(h(\xi), \mathbb{E}[y|\xi])]$$

Agenda

- Active Label Acquisition in Customer Survey
- Regret of the Prediction Model
- ➢ **Value of Information**
- Upper bound for the value of information
- Guarantees for Assortment Optimization
- Numerical Experiments

Tradeoff during the survey process

Comprehensive cost at time T:

$$\mathcal{C}(\boldsymbol{c}_T, h_T) \coloneqq \underbrace{\sum_{t=1}^{T} c_t \mathbb{I}\{\text{Customer } t \text{ accept the offer for survey} \mid c_t\}}_{\text{Label cost}} + \beta \cdot \underbrace{\mathrm{Regret}(h_T)}_{\text{Risk of the prediction model } h_T}$$

Market size

Objective: Minimize the expectation of the comprehensive cost

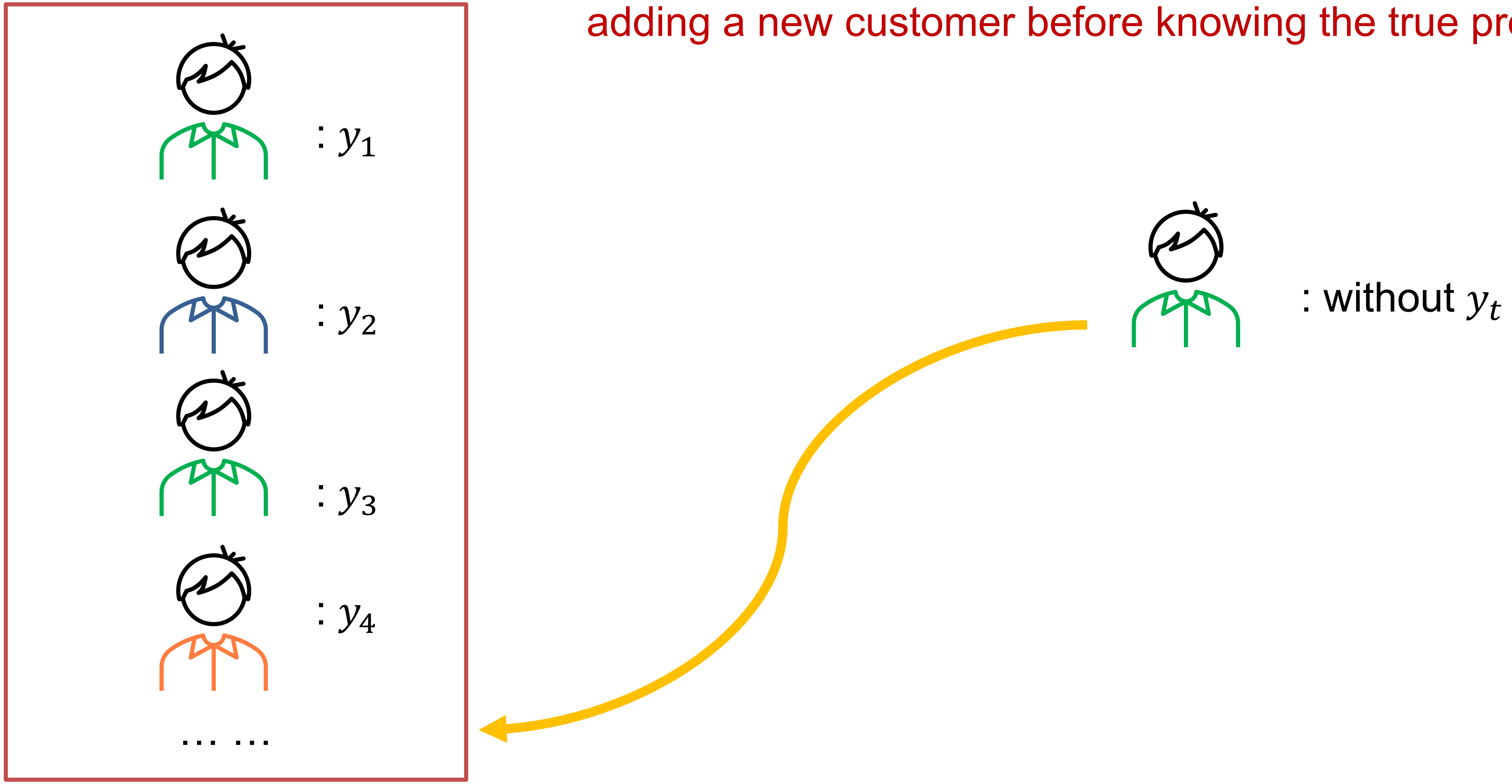
Tradeoff of incentive c_t :

- Too small: Little probability of taking the survey → Lack of data → $\mathrm{Regret}(h_T)$ will be large
- Too large: Waste of label cost（incentive）

Value of information

- Training set:

Value of information: The amount of risk reduction of adding a new customer before knowing the true preference



Value of information

Value of information $V(\xi_t;\, \mathcal{S}_{t-1})$

New predictor h_t when including (ξ_t, y_t)

$$V(\xi_t;\mathcal{S}_{t-1}) := \boxed{\beta\cdot \mathrm{Regret}(h_{t-1})} - \boxed{\beta\cdot\mathbb{E}_{y_t}\left[\mathrm{Regret}\left(\boxed{f\left(\mathcal{S}_{t-1}\bigcup\{(\xi_t,y_t)\}\right)}\right)\middle|\,\xi_t\right]}$$

Risk of predictor h_{t-1}

Expected risk of h_t when knowing (ξ_t, y_t)

It quantifies the expected risk reduction of including the customer t in the training set before knowing y_t

Agenda

- Active Label Acquisition in Customer Survey
- Regret of the Prediction Model
- Value of Information
- ➢ **Upper bound for the value of information**
- Guarantees for Assortment Optimization
- Numerical Experiments

Assortment Problem

- Customers have the no-purchase option 0.
- Suppose y_i follows Gumbel distribution with variance σ
- By MNL choice model, the purchase probability for product i is:

$$\frac{e^{\overline{y_i}/\sigma}}{1+\sum_j e^{\overline{y_j}/\sigma}}$$

- Suppose the price of product i is p_i
- Maximize the revenue of the assortment:

$$\max_{w\in\mathbb{B}^d, \boldsymbol{u}\in\mathbb{R}^d} \frac{\sum_{i\in[d]} u_i p_i w_i}{1+\boldsymbol{u}^T w}$$

$$\begin{aligned} s.t. \quad & w^T\mathbf{1} = z, \\ & u_i = e^{\overline{y_i}^{i}/\sigma}, \qquad \forall i \in [d] \end{aligned}$$

Incentives: upper bound for the value of information

Distance to degeneracy:

$$\nu_S(\hat{y}) := \inf_{w^*(y) \neq w^*(\hat{y})} \{\|\hat{y} - y\|\}$$

- It is defined as the distance between the prediction $\hat{y}$ and the closest vector y that leads to a different decision

Suppose the prediction error for $\hat{y}$ is $\rho(\xi)$, then an upper bound for the value of information is:

$$\varkappa\beta\sqrt{2\min\{z, d-z\}}\boxed{\mu(\xi)} \cdot \boxed{\rho(\xi)} \cdot \boxed{\mathbb{I}\{\nu_S(\hat{y}) \leq \rho(\xi)\}}$$

Test set distribution

Prediction error
(Training set distribution)

Whether the true optimal decision is determined

Insights from the upper bound of value of information

Upper bound

$$\varkappa\beta\sqrt{2\min\{z,d-z\}}\boxed{\mu(\xi)}\cdot\boxed{\rho(\xi)}\cdot\boxed{\mathbb{I}\{\nu_S(\hat{y})\le\rho(\xi)\}}$$

1. If one feature has a higher probability in the test set
   - Its value of information gets larger

2. If one feature has a larger proportion in the training set
   - The prediction error for this feature gets smaller
   - The value of information gets smaller

3. If the prediction error for one sample is smaller than $\nu_S(\hat{y})$:
   - The optimal decision for this sample has been determined
   - Regret for this type of customer is zero
   - We will stop surveying this type of customers

Personalized incentives based on the upper bound of value of information

Range of offered incentive: $c_t \in \{0\} \cup [c_{min}, c_{max}]$

Given a type of customer ξ:

If the upper bound of value of information $\leq c_{min}$:

- Provide zero incentive
- Ignore the feedback of this customer

Otherwise, we offer some incentives between $[c_{min}, c_{max}]$

Agenda

- Active Label Acquisition in Customer Survey
- Regret of the Prediction Model
- Value of Information
- Upper bound for the value of information
- ➢ **Guarantees for Assortment Optimization**
- Numerical Experiments

Near-degeneracy function

Near-degeneracy function Ψ :

$$\Psi(\rho) := \mathbb{P}(\nu_S(\mathbb{E}[y|\xi]) \leq \rho)$$

Ψ describes the difficulty in distinguishing the optimal decision from the sub-optimal decision at a certain prediction error level ρ

Guarantees for the problem

Theorem:

After T iterations, the cumulative label cost is at most $\min\left\{\tilde{O}\left(\sum_{t=1}^{T}\Psi(t^{-\frac{1}{2}})\right), \tilde{O}\left(T^{\frac{1}{2}}\right)\right\}$

- Low-noise condition: For some large $\rho_0 > 0$ and $\kappa > 0$, the near-degeneracy function satisfies:

$$\Psi(\rho) \leq \left(\frac{\rho}{\rho_0}\right)^{\kappa}$$

Low-noise condition is closely related to Hu et al. 2022 and Tsybakov's noise condition

Under low-noise conditions:

✓ The cumulative label cost is at most $\tilde{O}\left(T^{1-\kappa/2}\right)$

Comparison with supervised learning

Under the low-noise condition with $\kappa > 2$:

- Personalized incentives policy requires use finite samples to achieve zero risk
- Fixed incentive policy requires infinite samples to achieve zero risk

Theorem:

Under various conditions of c_{min}, regarding the comprehensive cost:

✓ Our personalized incentive policy $\leq$ fixed incentive policy

Agenda

- Active Label Acquisition in Customer Survey
- Regret of the Prediction Model
- Value of Information
- Upper bound for the value of information
- Guarantees for Assortment Optimization

➢ **Numerical Experiments**

Numerical Experiments: Assortment optimization with contextual information

Use synthetic data:

- Number of product: 10
- Number of types of customers: 5

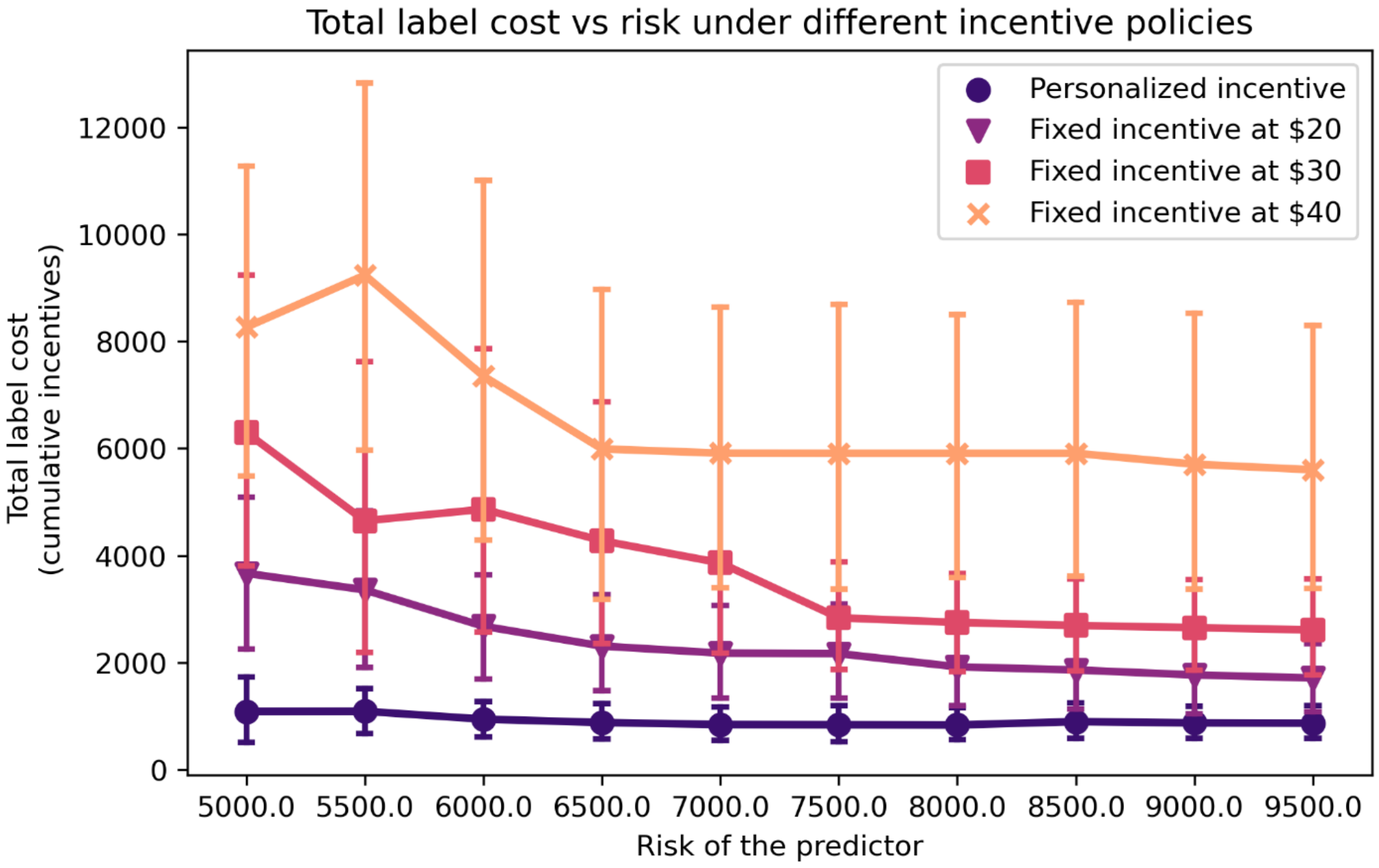
- Dimension of features within each type: 8
- Each type has its own feature and prediction model
- Within each type, we assume the true prediction model is linear

- The offered incentive is either 0 or some value between $[\$20, \$40]$

- $p(c)$ is a linear function between $(\$20, 0.3)$ and $(\$40, 0.9)$

Results: Assortment optimization

- Observation: To achieve the same level of risk, the personalized incentive policy requires much less label cost



Results: Assortment optimization

- When ensuring the excess risk is less than 5000:

	Personalized incentive	Fixed incentive at $20	Fixed incentive at $30	Fixed incentive at $40
Required label cost	1088	3668 (-70%)	6295 (-79%)	8262 (-87%)
Required number of surveyed customers (Size of training set)	30	184 (-84%)	210 (-86%)	206 (-85%)

Thank you

Mo Liu

Personalized incentive given the value of information

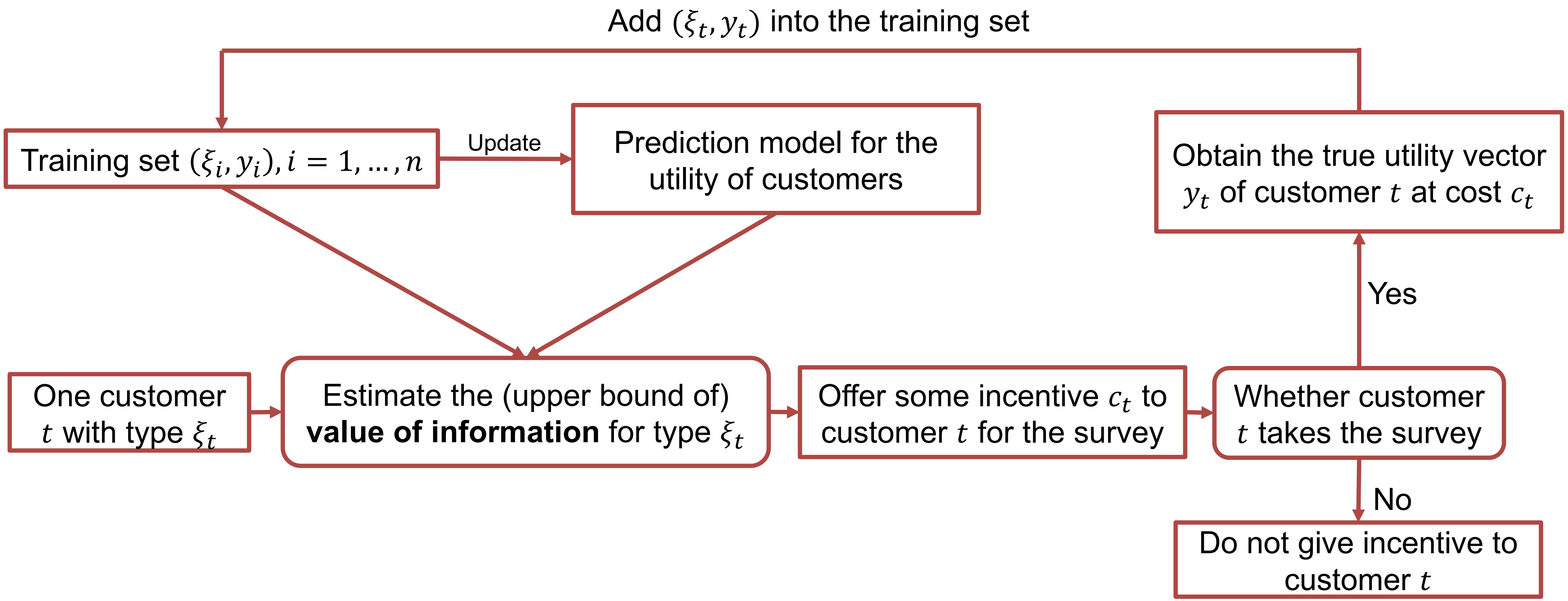
Range of offered incentive: $c_t \in \{0\} \cup [c_{min}, c_{max}]$

- If the offered incentive is zero, we ignore the feedback of this customer
- If the offered incentive is nonzero, the optimal incentive is c^*

$$c^*(V(\xi_t;\, \mathcal{S}_{t-1}), p) \coloneqq \arg \min_{c \in [c_{min}, c_{max}]} \{p(c_t)[c_t - V(\xi_t;\, \mathcal{S}_{t-1})]\}$$

Assumption of $p(c)$: Increasing function with $p(c_{min}) > 0$

Active label acquisition using value of information



If the upper bound is less than c_{min}, we do not offer any incentive

Results: Assortment optimization

- Observation: Using the same amount of label cost, the personalized incentive policy achieves much smaller risk.