Learning Human Preferences

From Clicks to Conversations

Suryanarayana Sankagiri

IIT Madras

18th November 2025

My Academic Journey

Segmentation of Hindustani concert music

NCC ’16, ISMIR ’16

B.Tech.

2012-2016

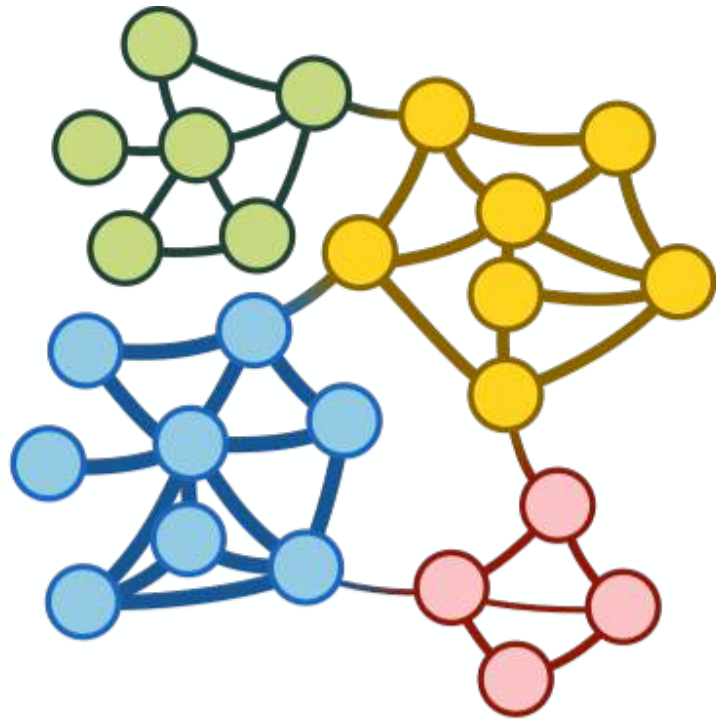
My Academic Journey

Segmentation of Hindustani concert music

NCC ’16, ISMIR ’16

B.Tech.

2012-2016

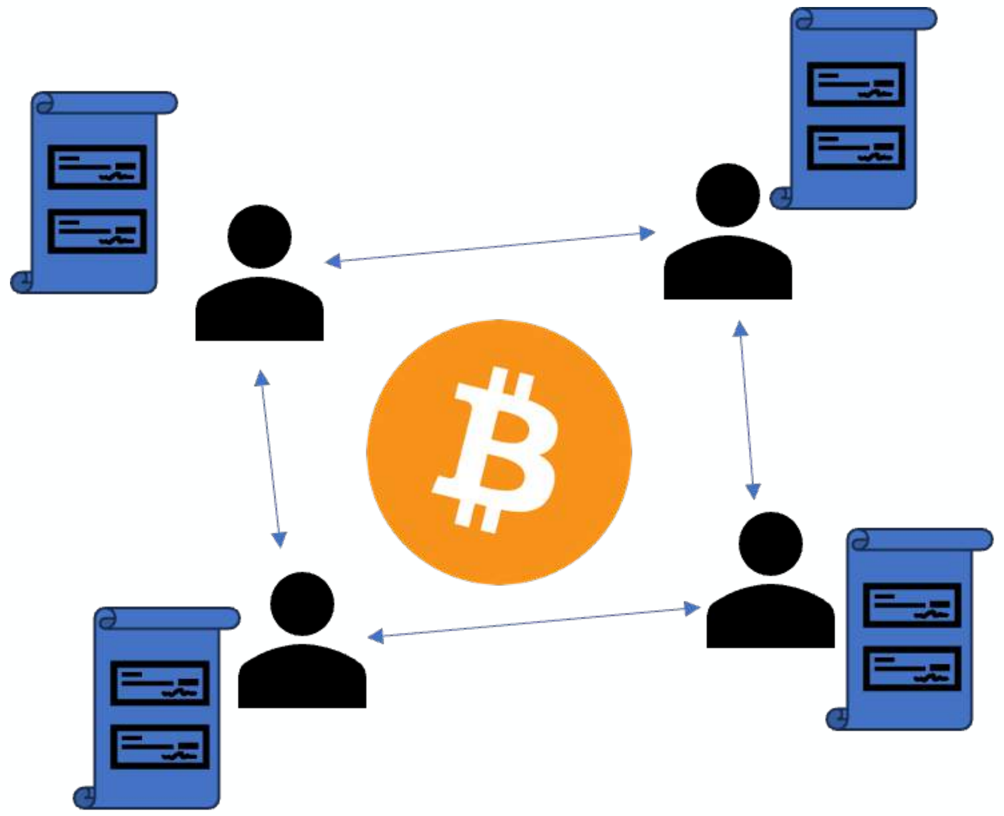


Community detection in random graphs

IEEE TIT ’19

M.S.

2016-2018



Blockchain protocols for security and efficiency

FC’ 21, INFORMS SS ’23, IEEE ToN ’25

Ph.D

2018-2022

UNIVERSITY OF ILLINOIS URBANA-CHAMPAIGN

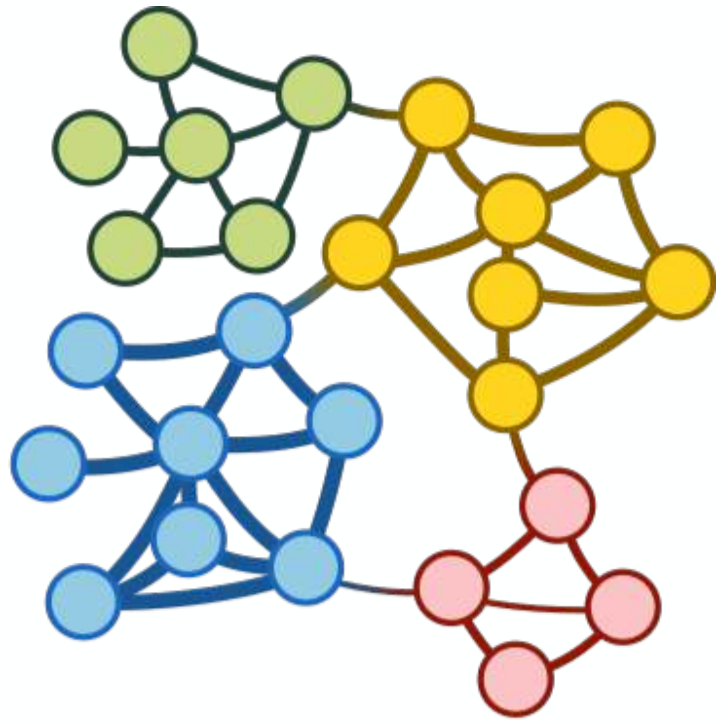
My Academic Journey

Segmentation of Hindustani concert music

NCC ’16, ISMIR ’16

B.Tech.

2012-2016

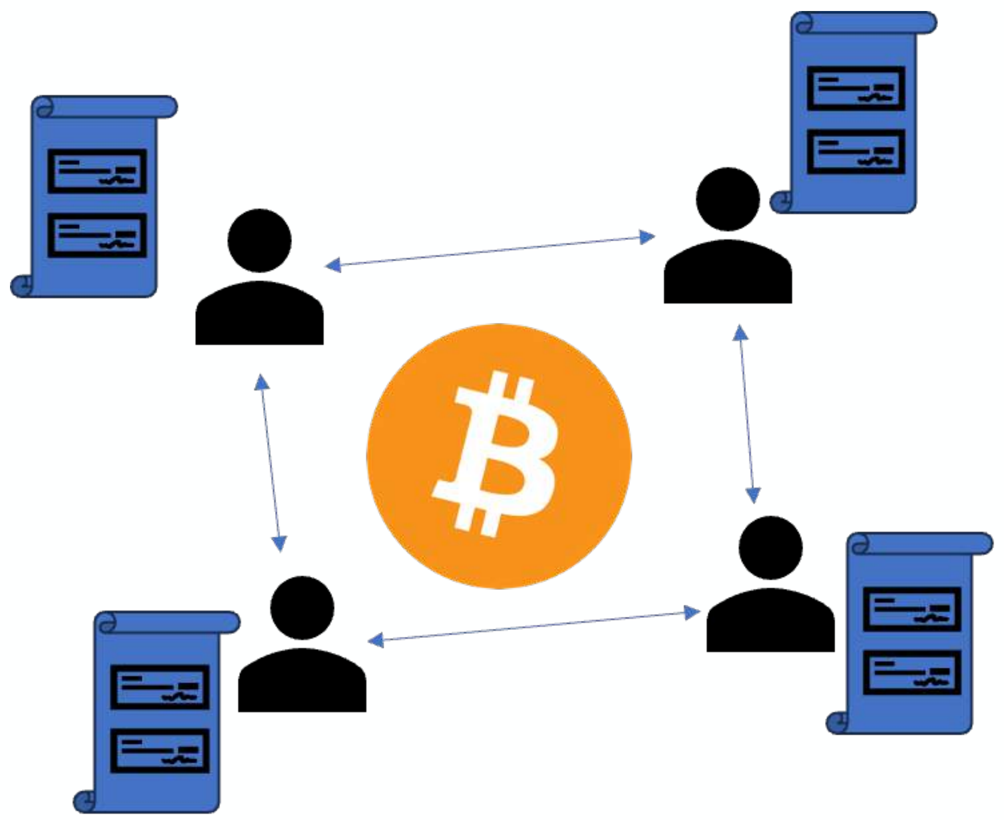


Community detection in random graphs

IEEE TIT ’19

M.S.

2016-2018

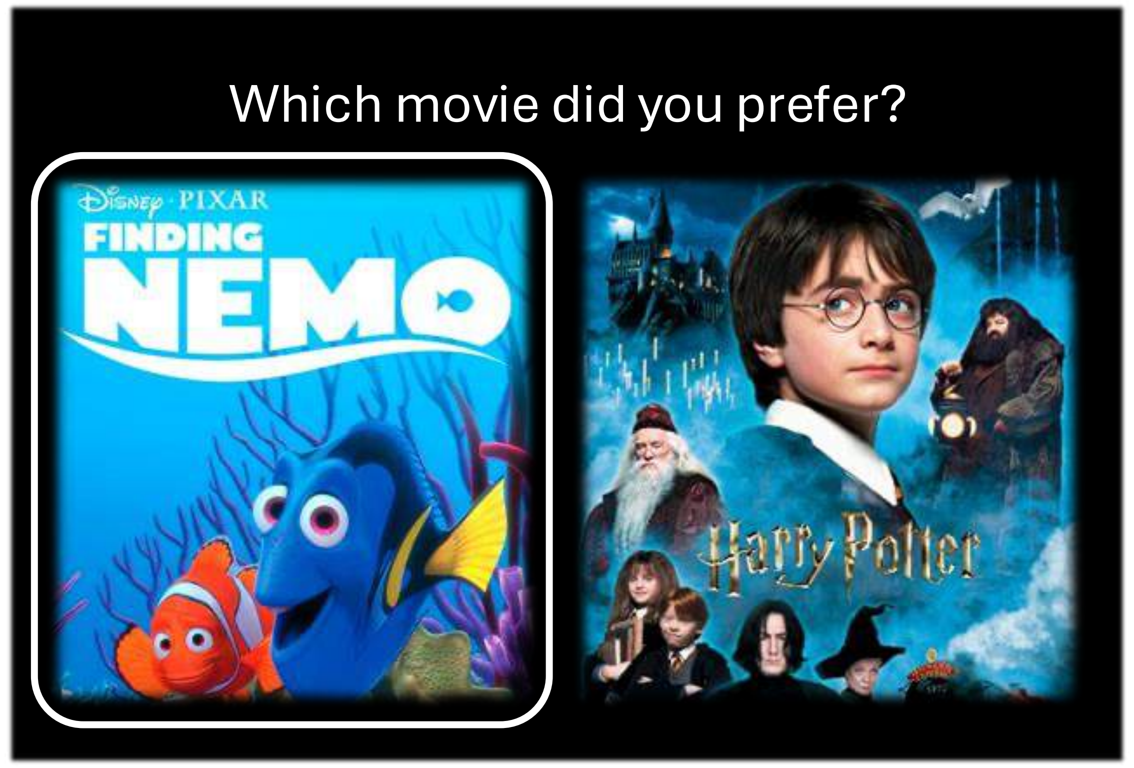


Blockchain protocols for security and efficiency

FC’ 21, INFORMS SS ’23, IEEE ToN ’25

Ph.D

2018-2022



Recommender systems that learn from comparisons

ICML ’25, UAI ’25

Postdoc

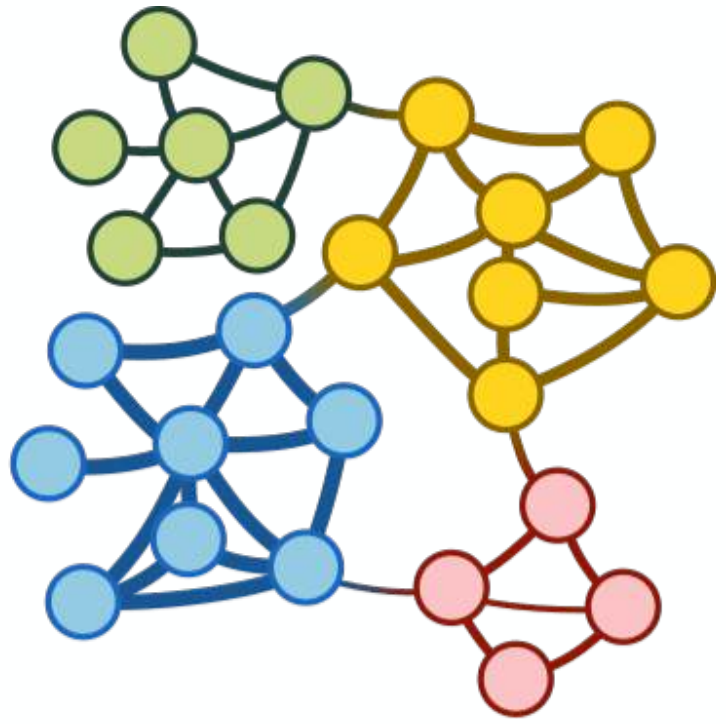
2022-2025

UNIVERSITY OF ILLINOIS URBANA-CHAMPAIGN

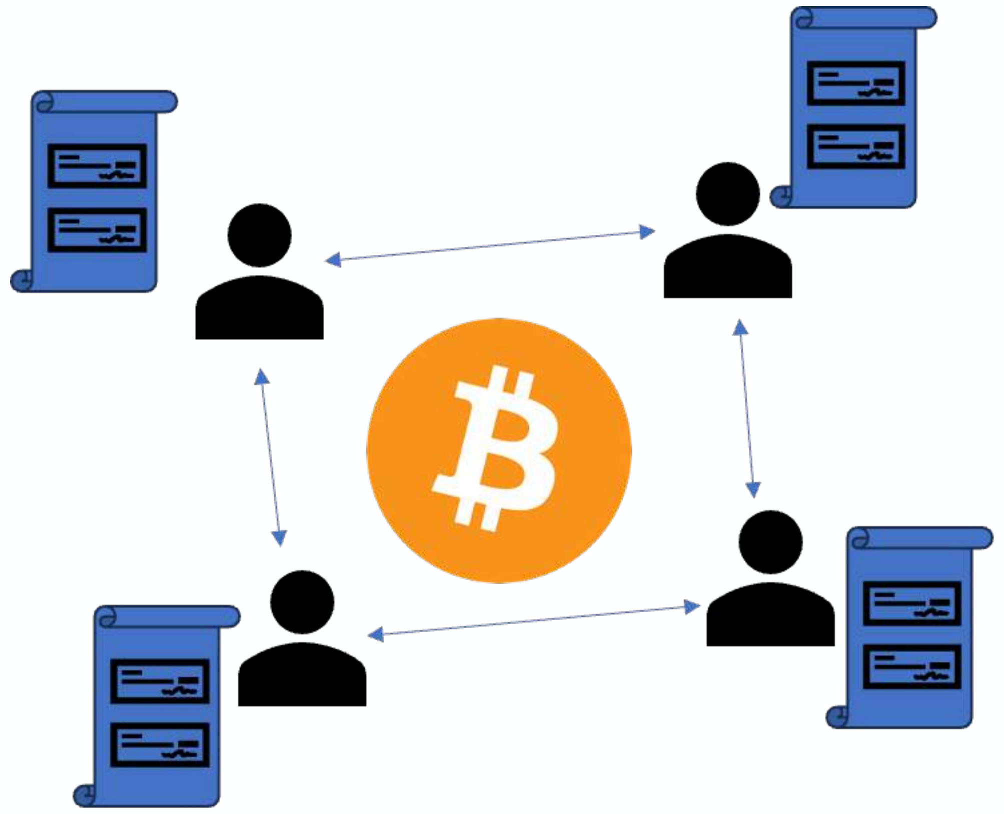
EPFL

Theme: Theoretical Systems Researcher

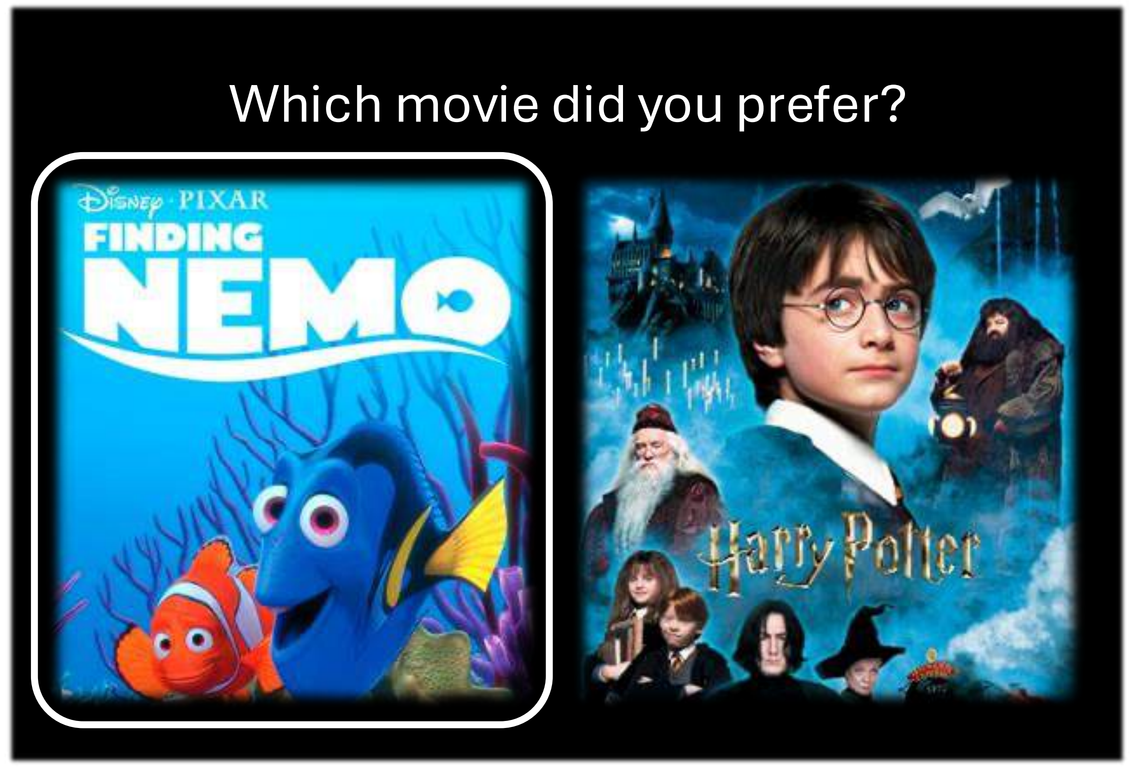
Segmentation of Hindustani concert music



Community detection in random graphs



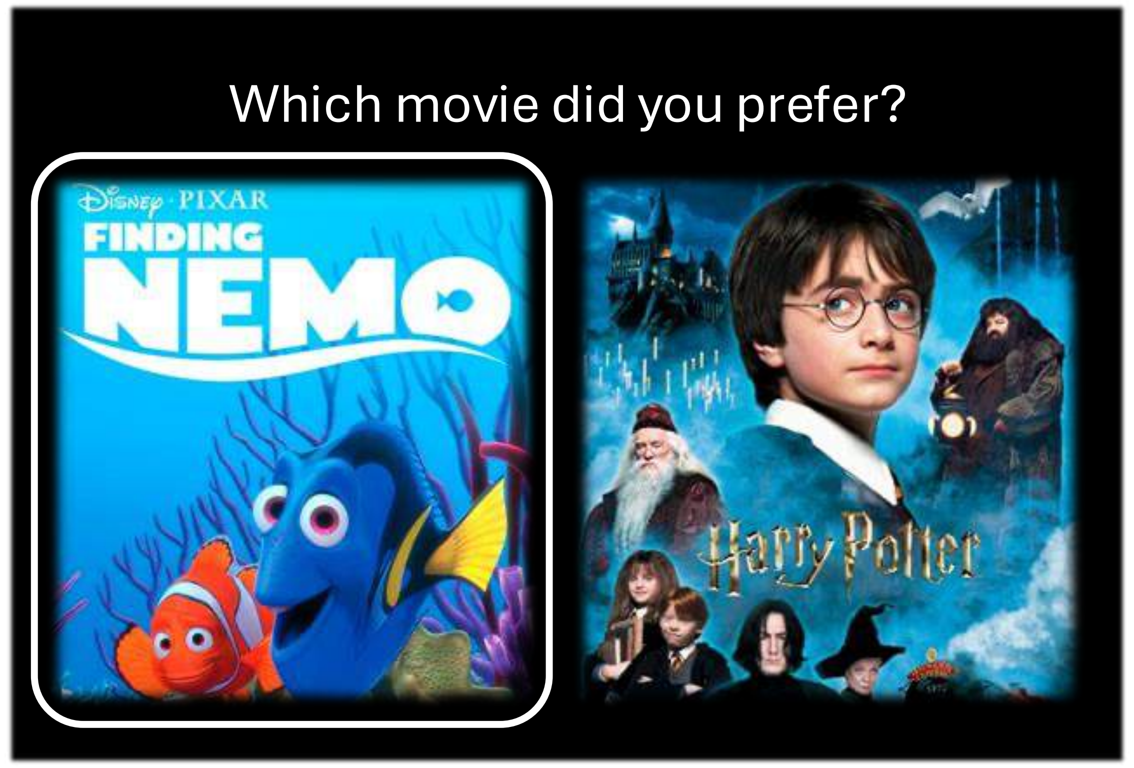
Blockchain protocols for security and efficiency



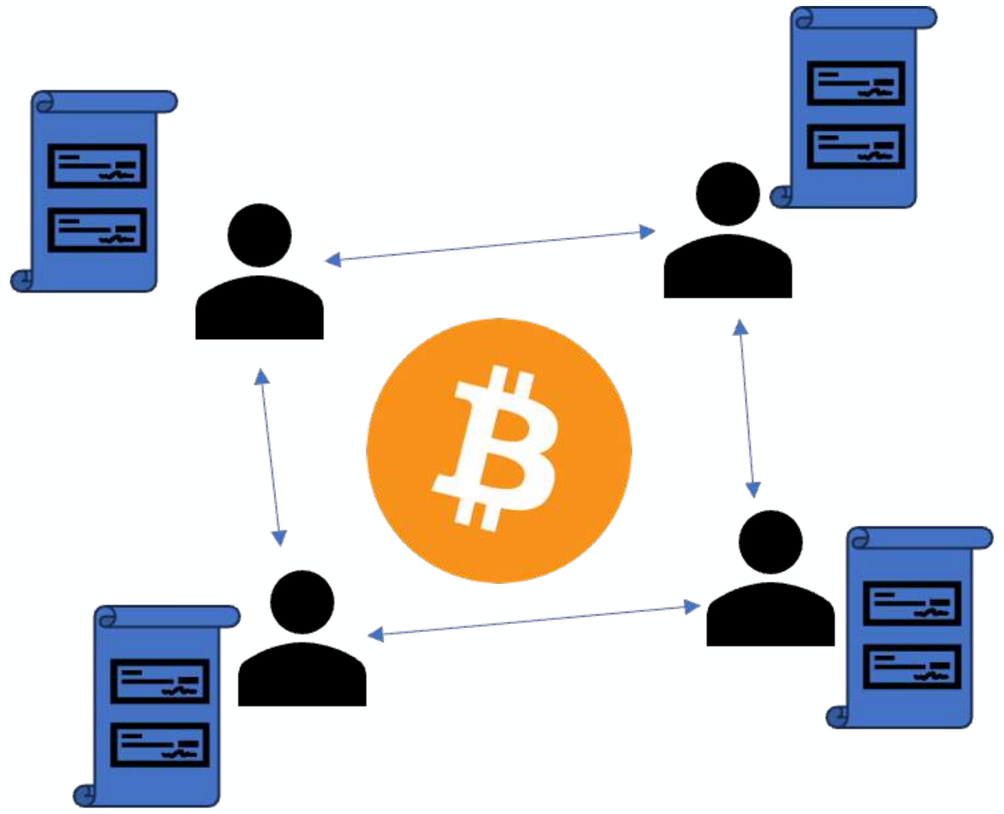
Recommender systems that learn from comparisons

Theme: Theoretical Systems Researcher

Community detection in random graphs

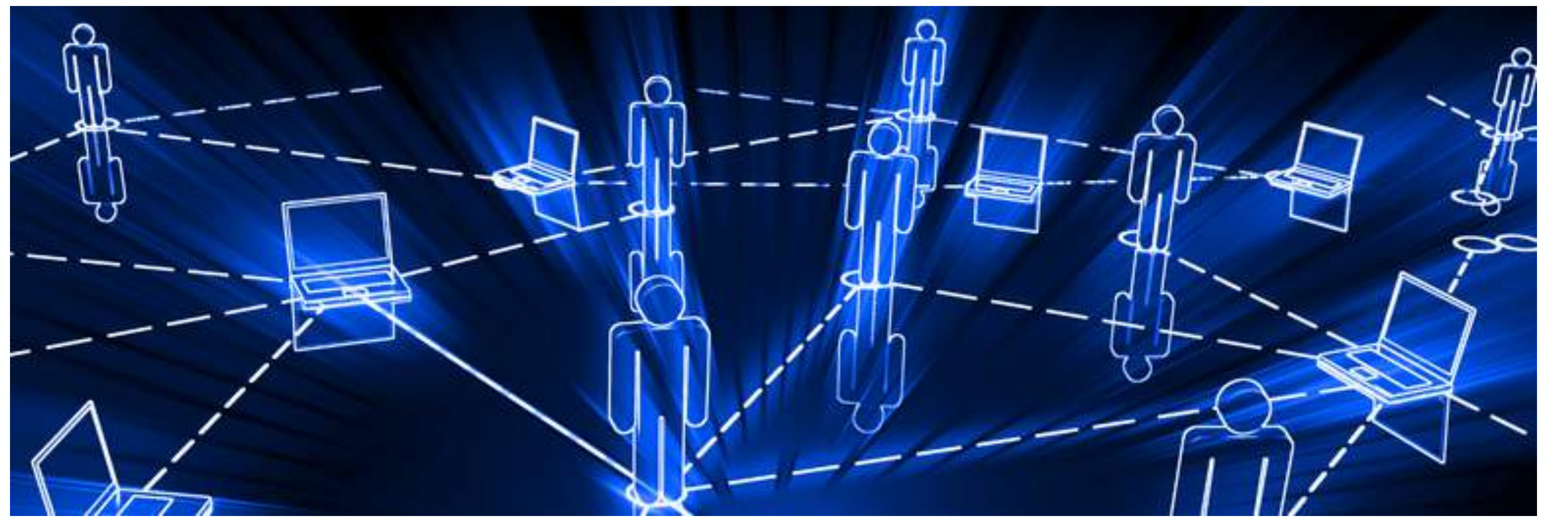


Recommender systems that learn from comparisons



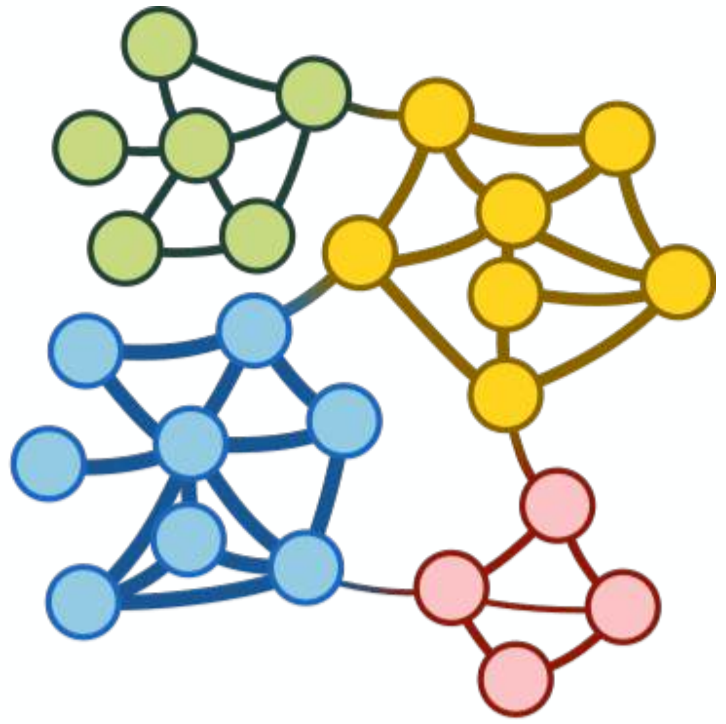
Blockchain protocols for security and efficiency

Individual utilities influence actions

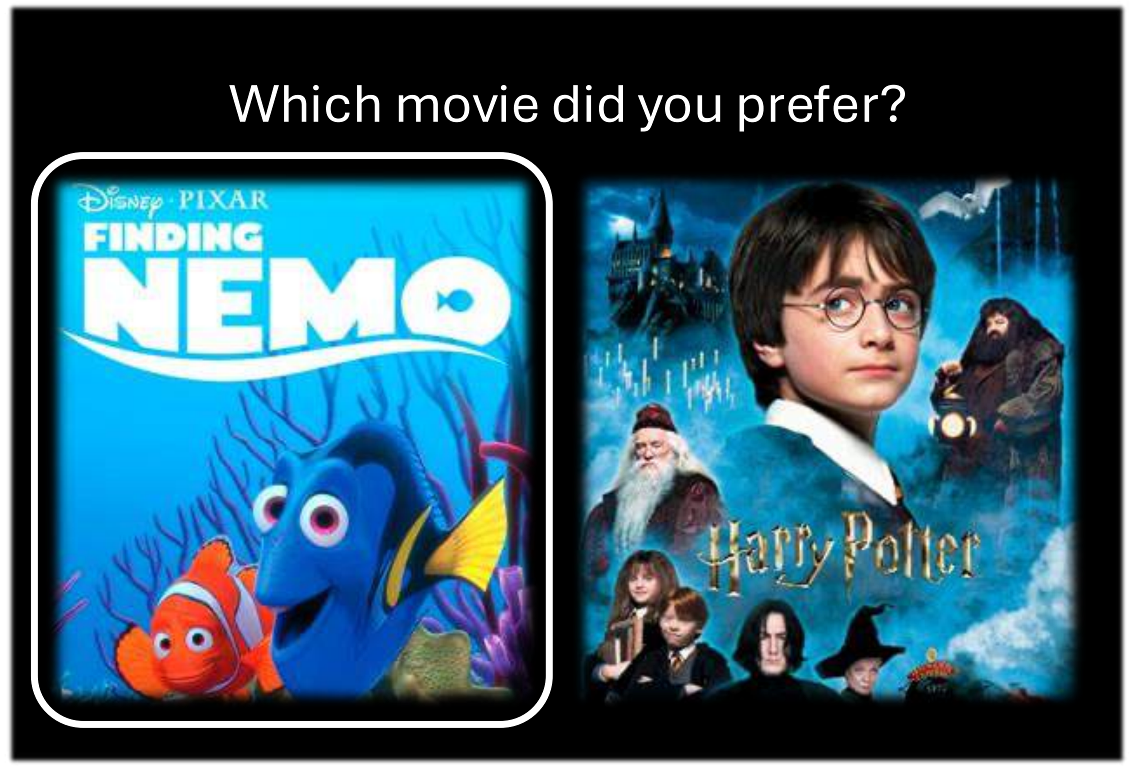


Human-in-the-loop systems

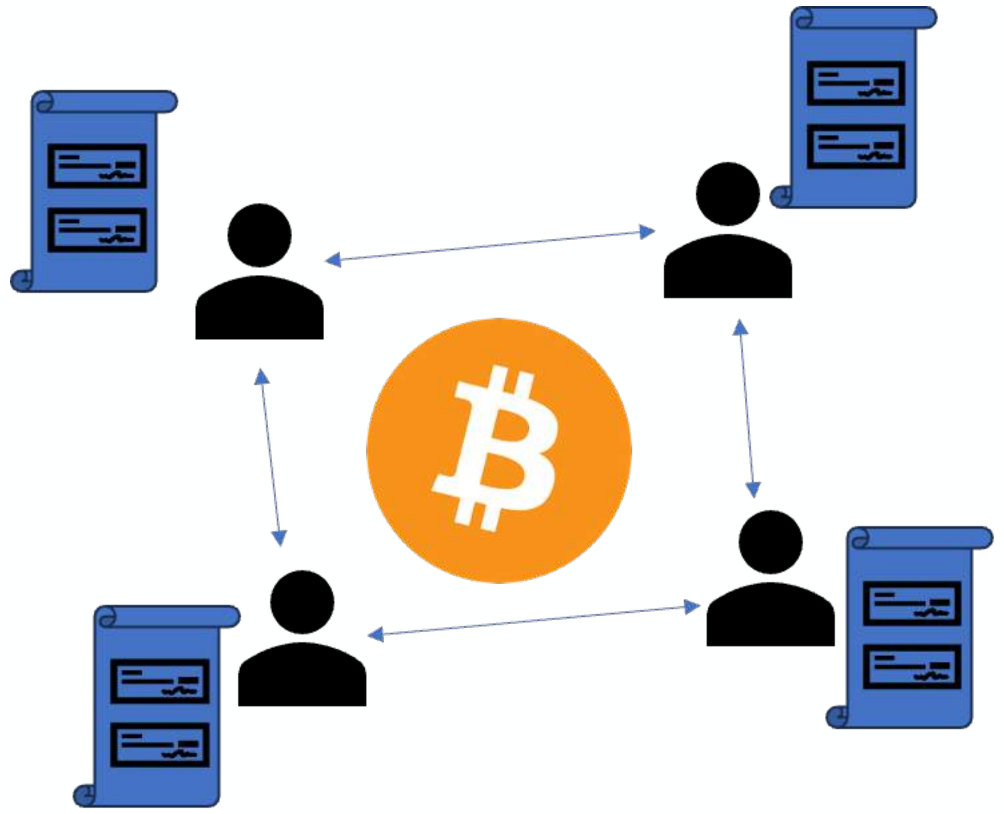
Theme: Theoretical Systems Researcher



Community detection in random graphs

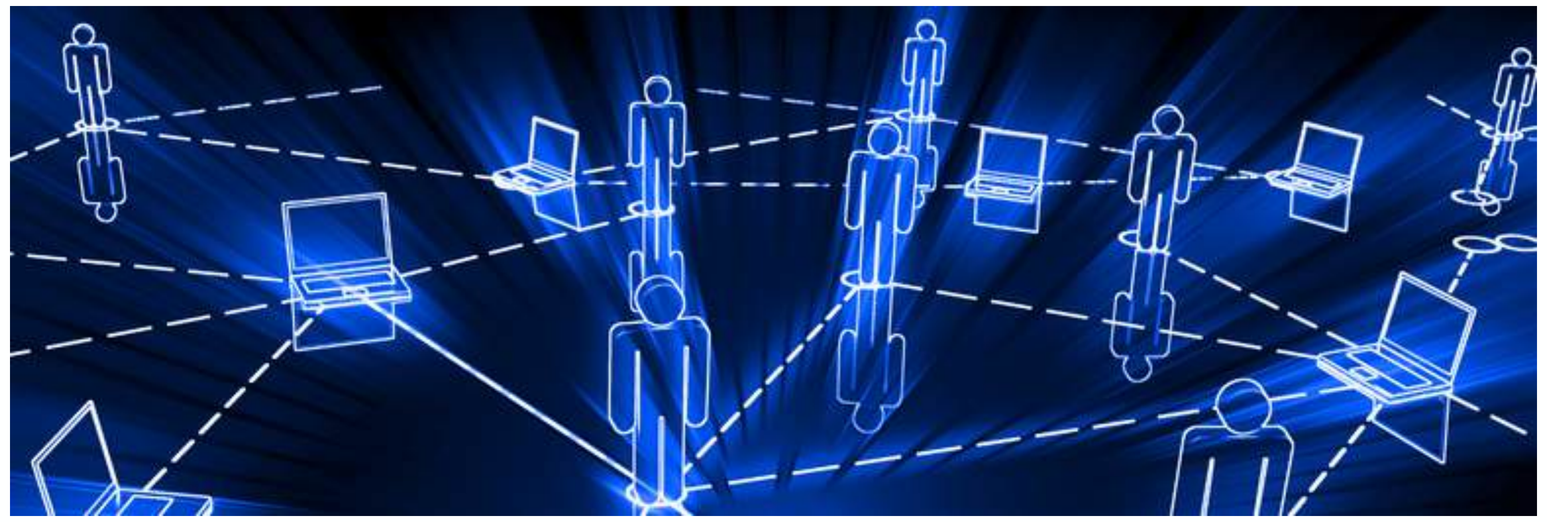


Recommender systems that learn from comparisons

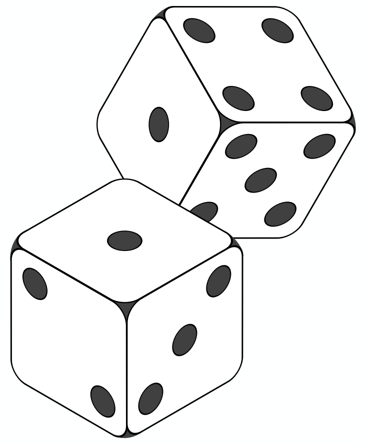


Blockchain protocols for security and efficiency

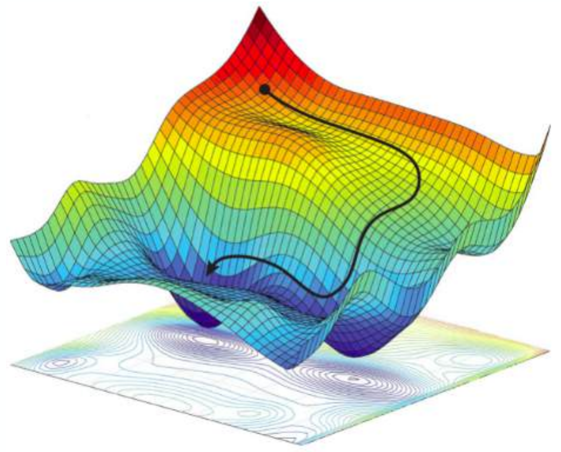
Individual utilities influence actions



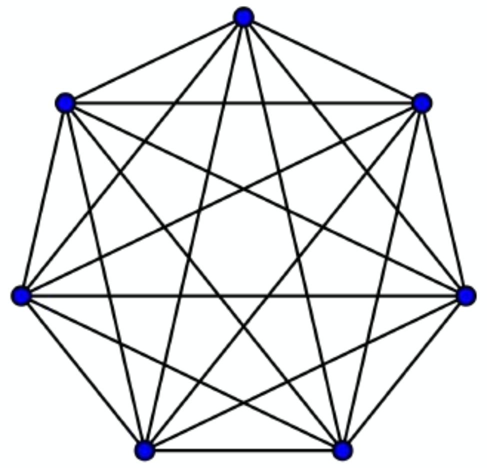
Human-in-the-loop systems



Probability



Optimisation



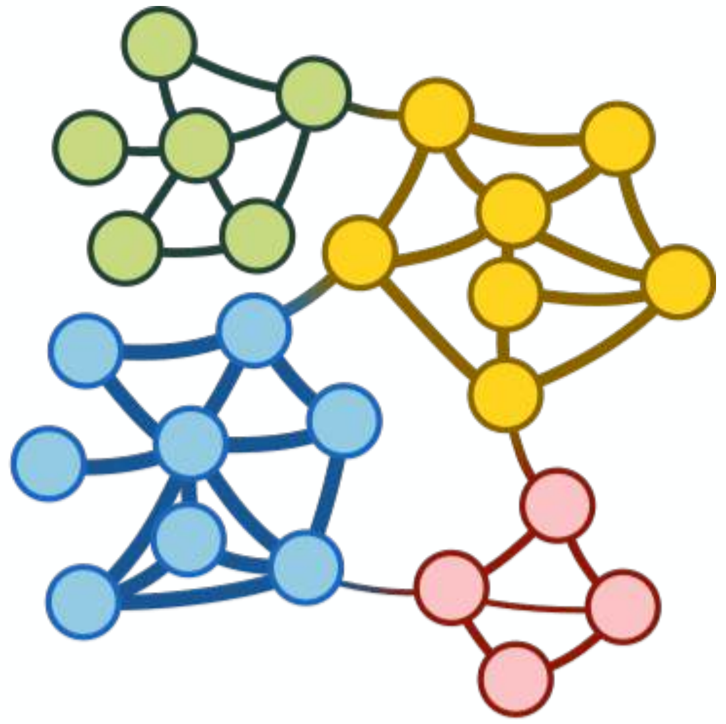
Networks

Modelling

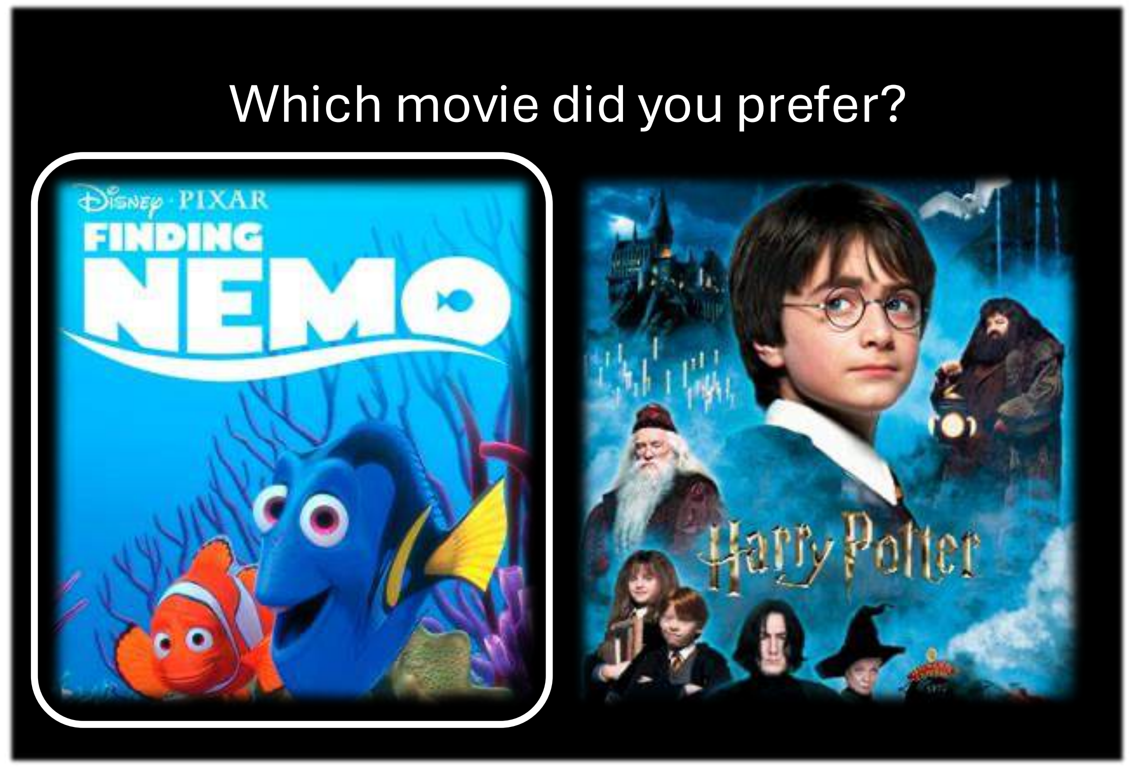
Analysis

Insights

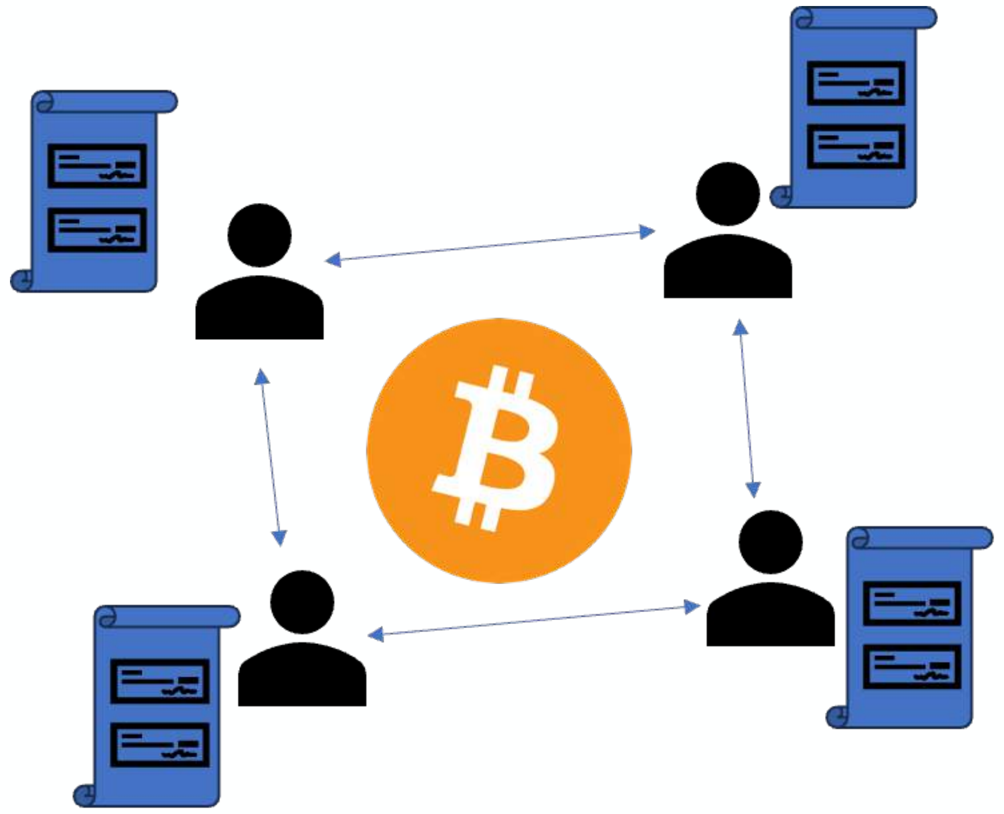
Theme: Theoretical Systems Researcher



Community detection in random graphs

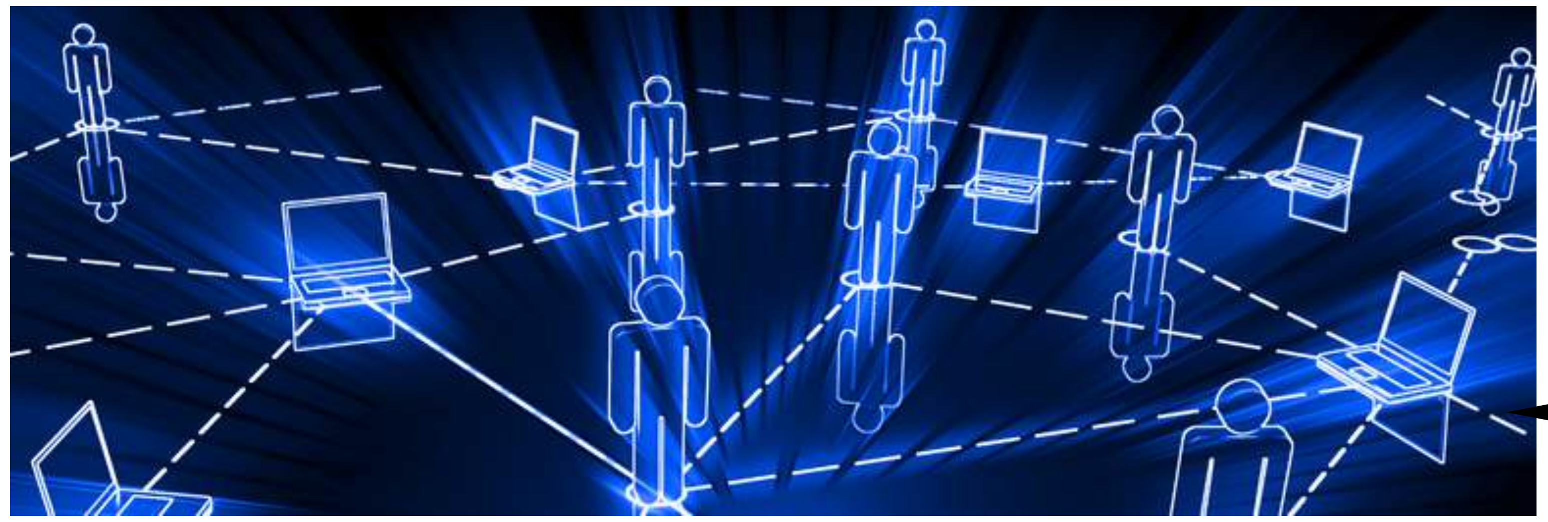


Recommender systems that learn from comparisons



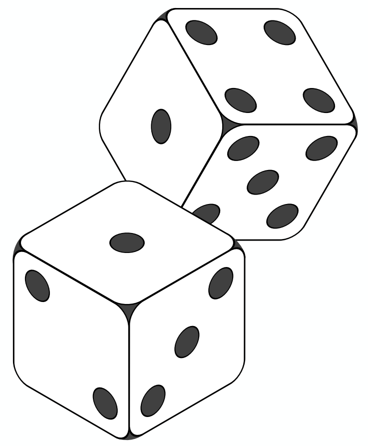
Blockchain protocols for security and efficiency

Learning → **Individual utilities influence actions** ← **Control**

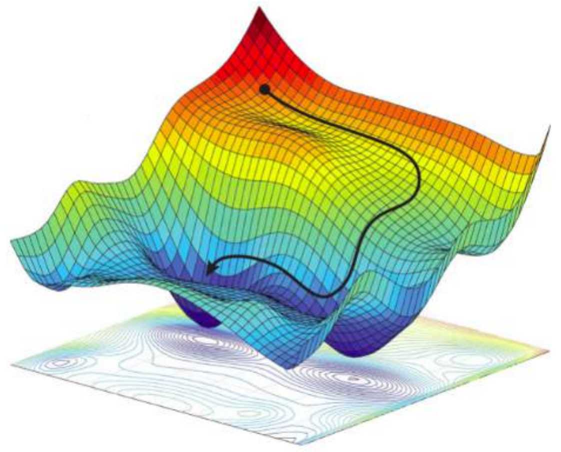


Human-in-the-loop systems

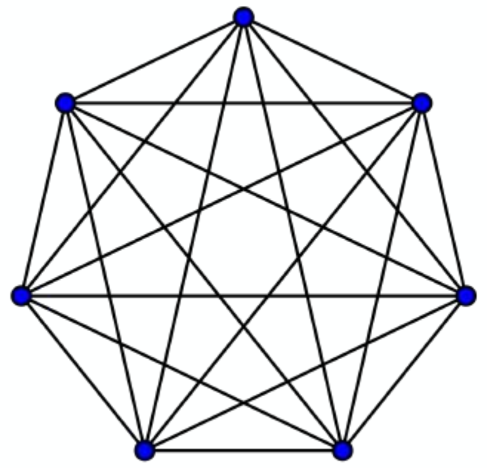
Next-gen systems need both learning and control



Probability



Optimisation



Networks

Modelling

Analysis

Insights

Towards Recommender Systems That Learn From Comparisons

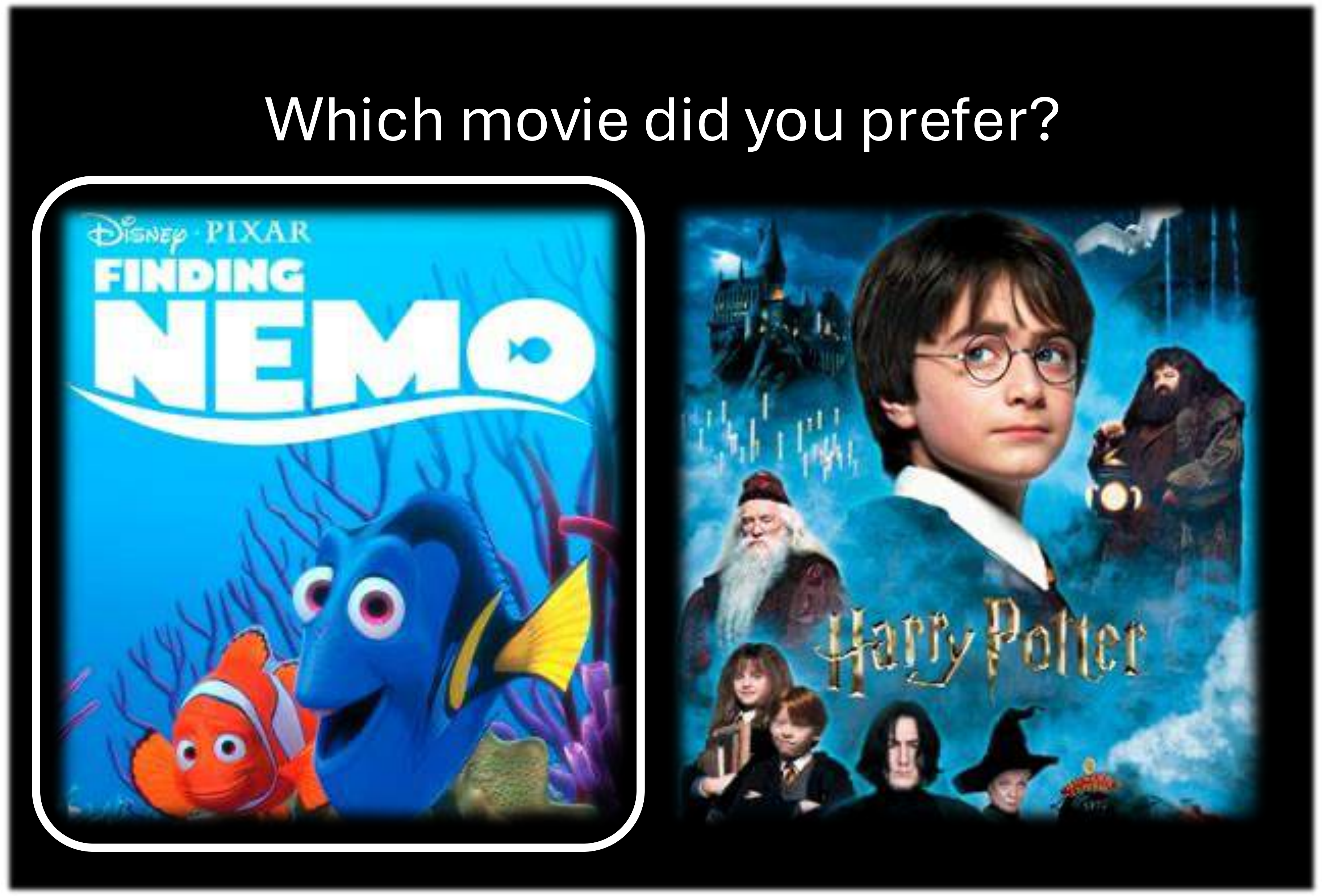
Goal: recommend good items ⇒ learn user utilities from interactions

Ratings reflect user utilities

Towards Recommender Systems That Learn From Comparisons

Ratings reflect user utilities

Why seek comparisons?

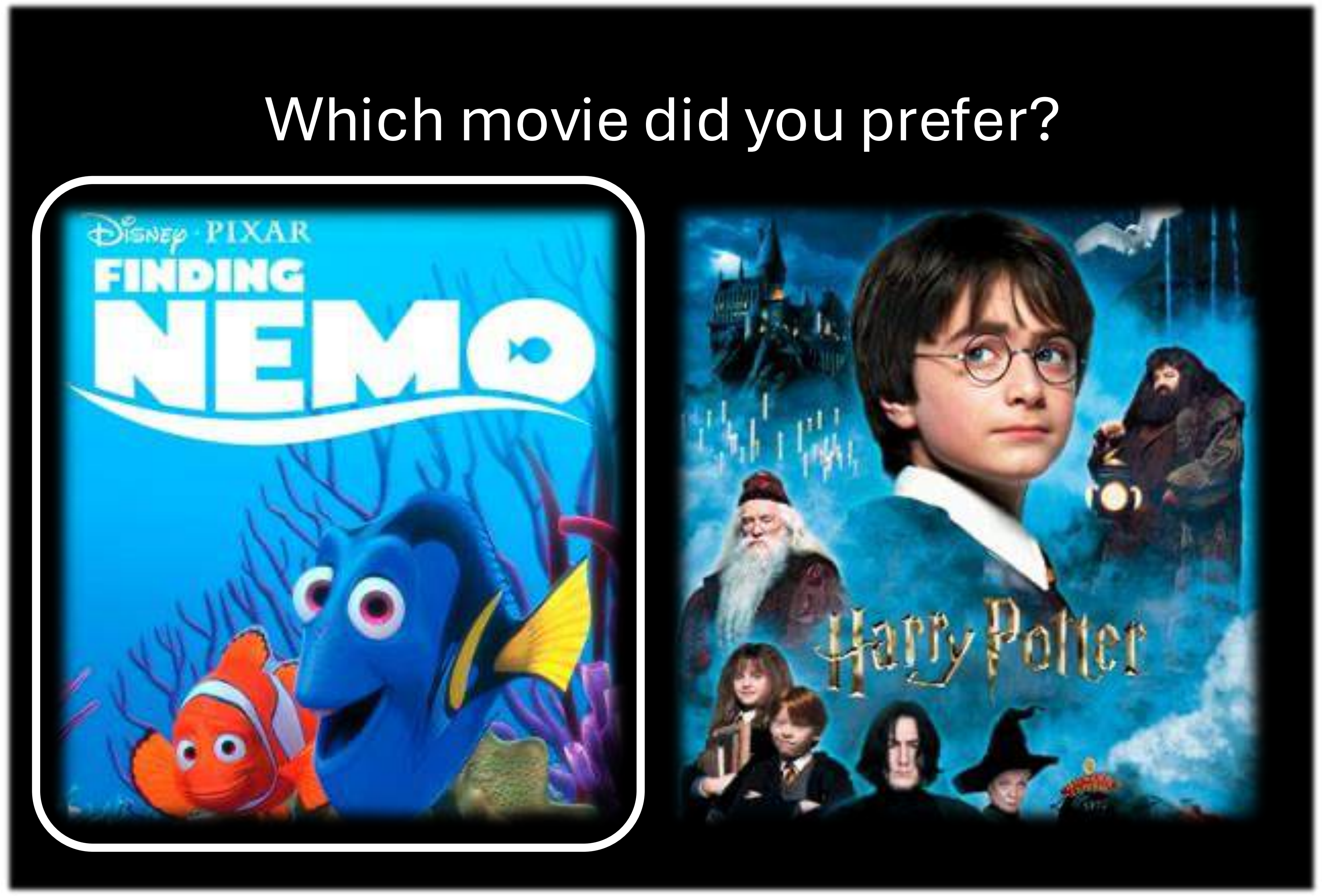


Towards Recommender Systems That Learn From Comparisons

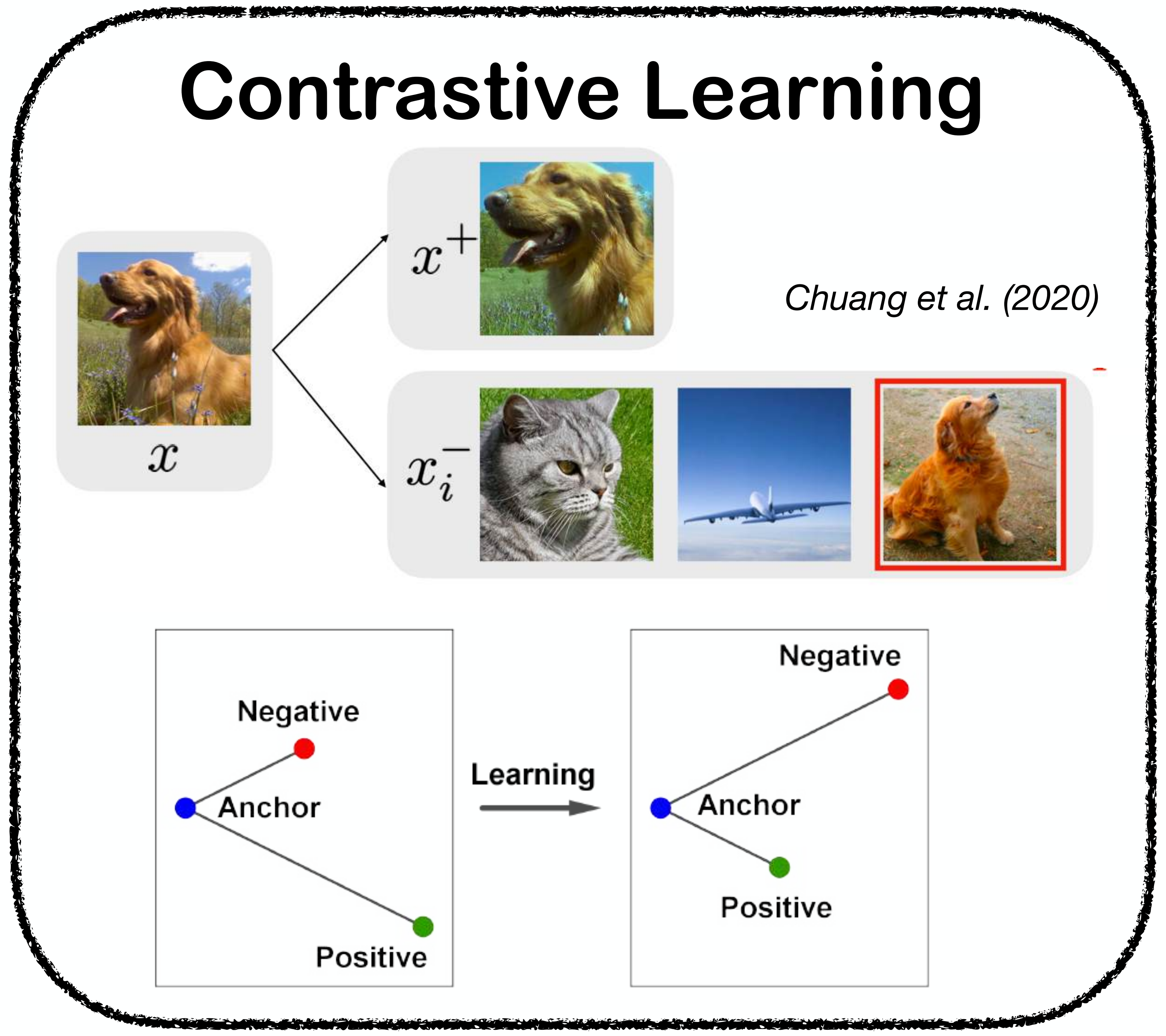
Ratings reflect user utilities

Ratings are lossy due to discretisation

Comparisons convey fine-grained ranking information

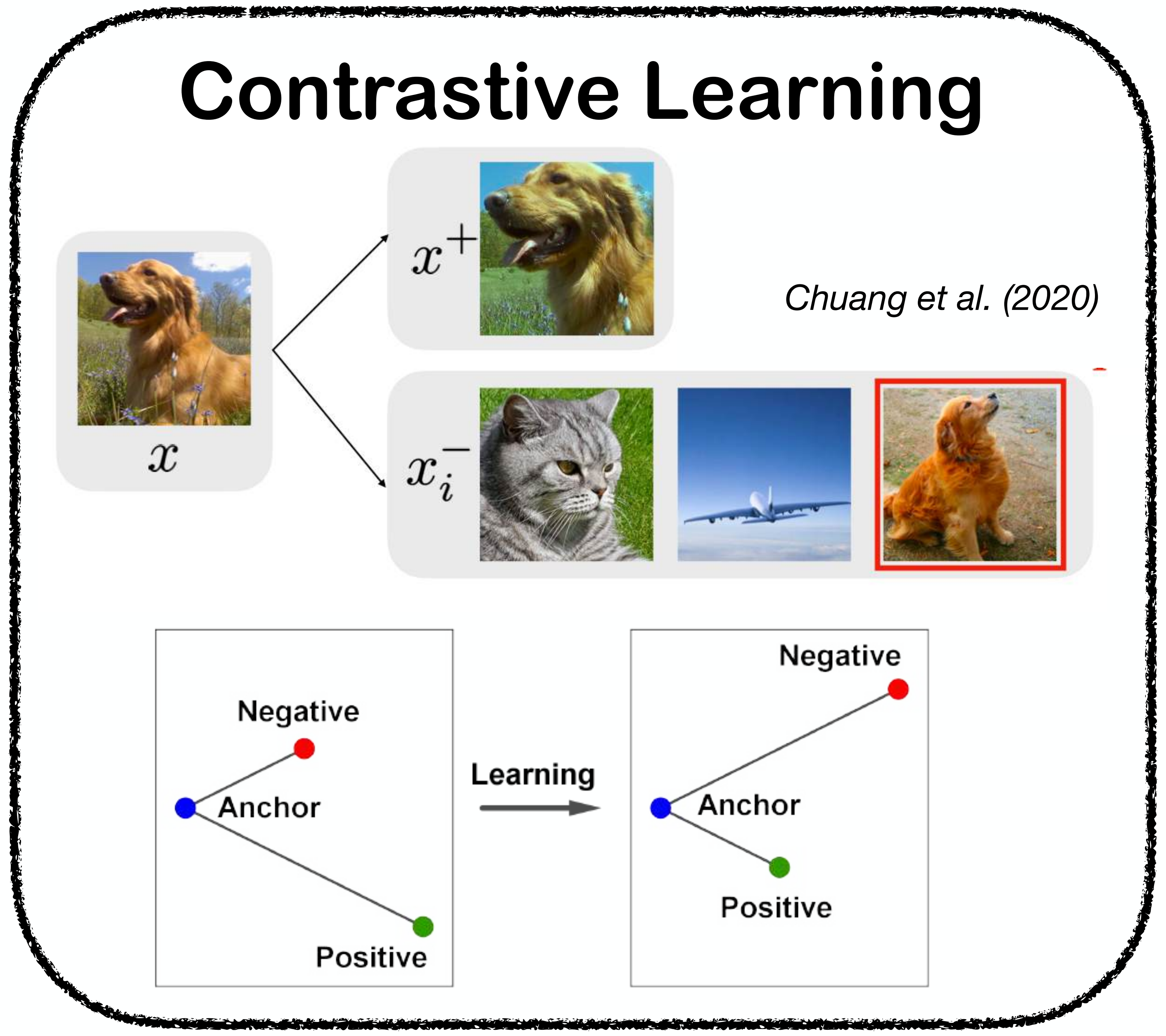


Learning From Comparisons: Examples

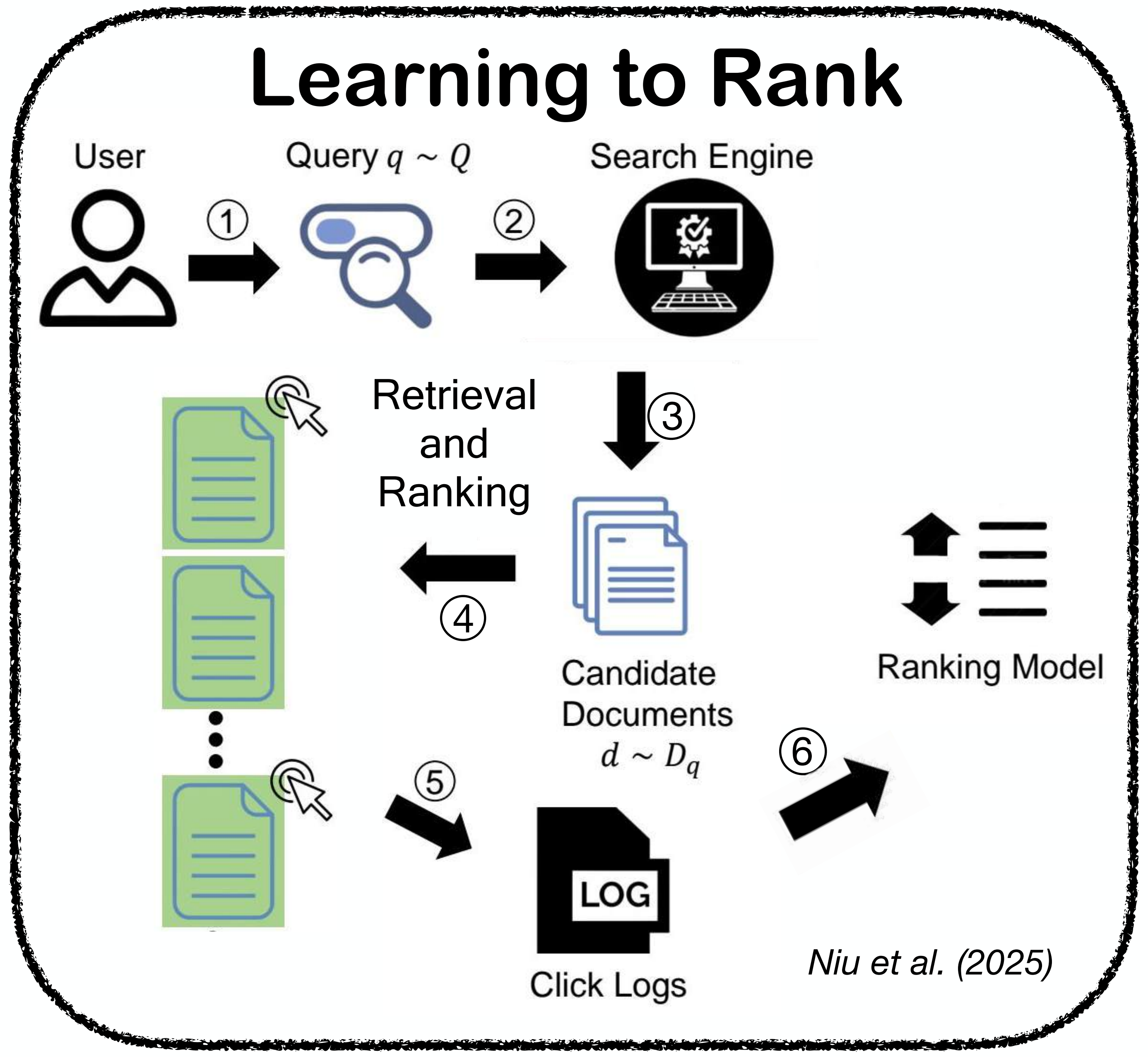


artificial comparisons

Learning From Comparisons: Examples

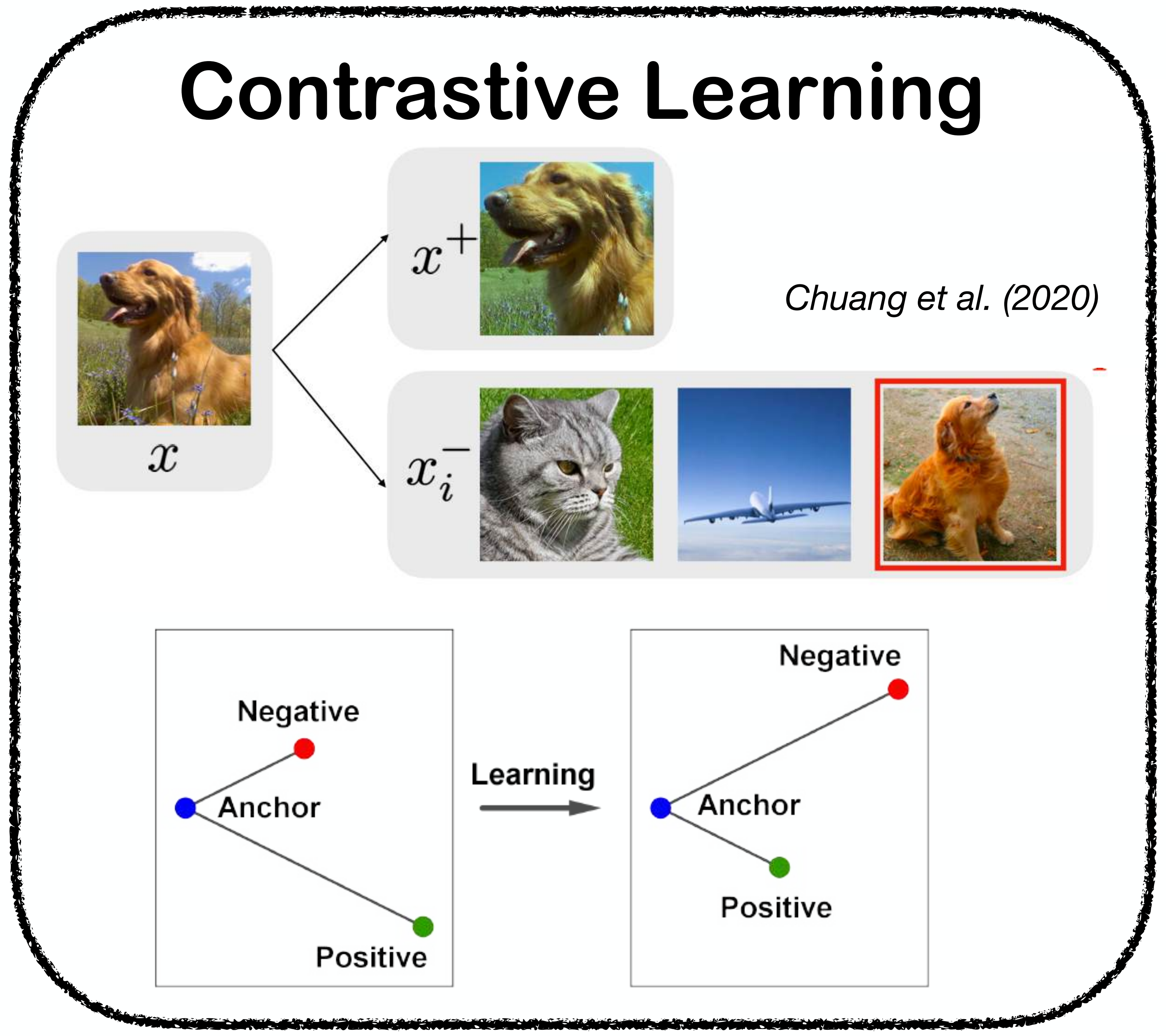


artificial comparisons

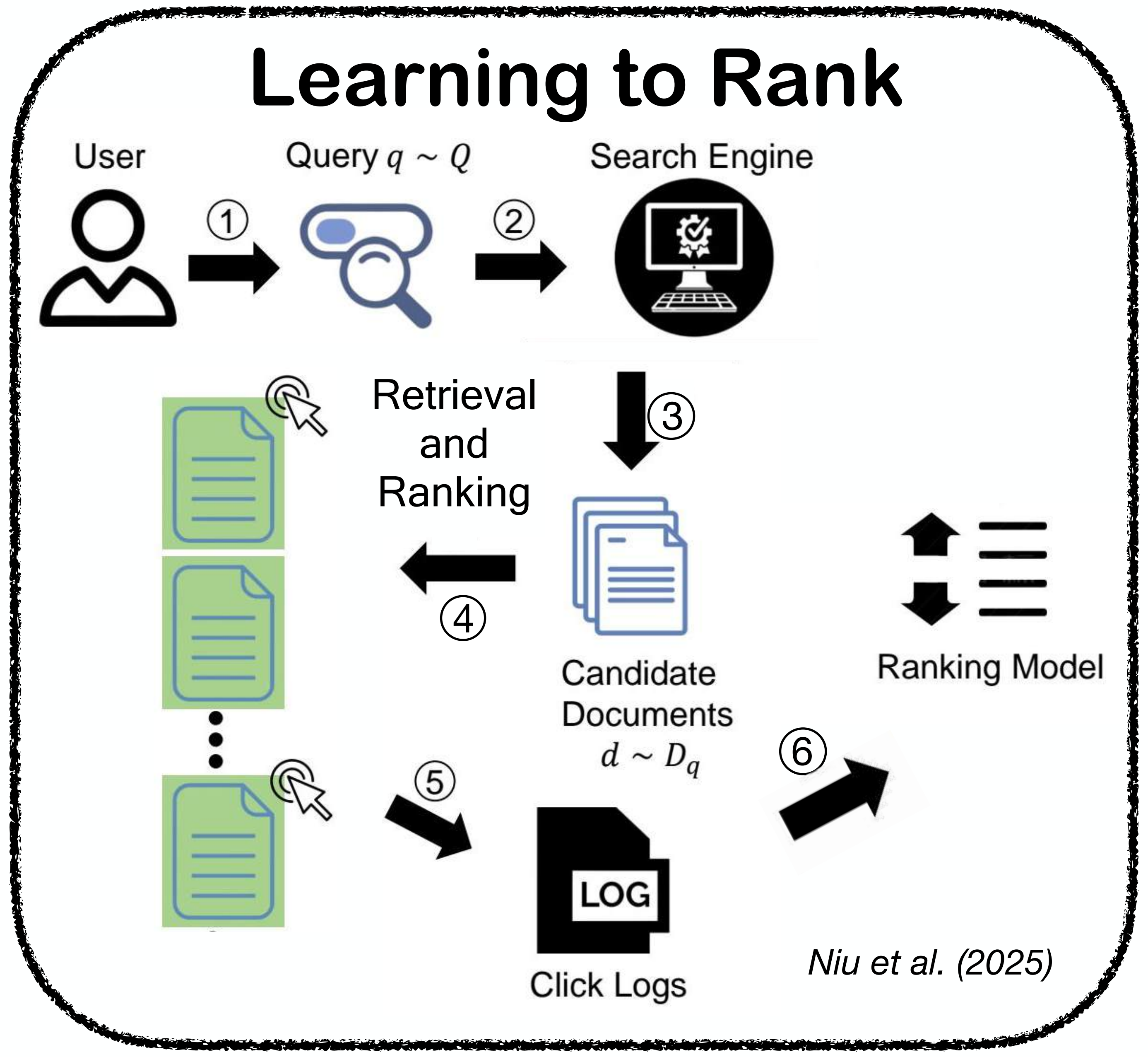


implicit human comparisons

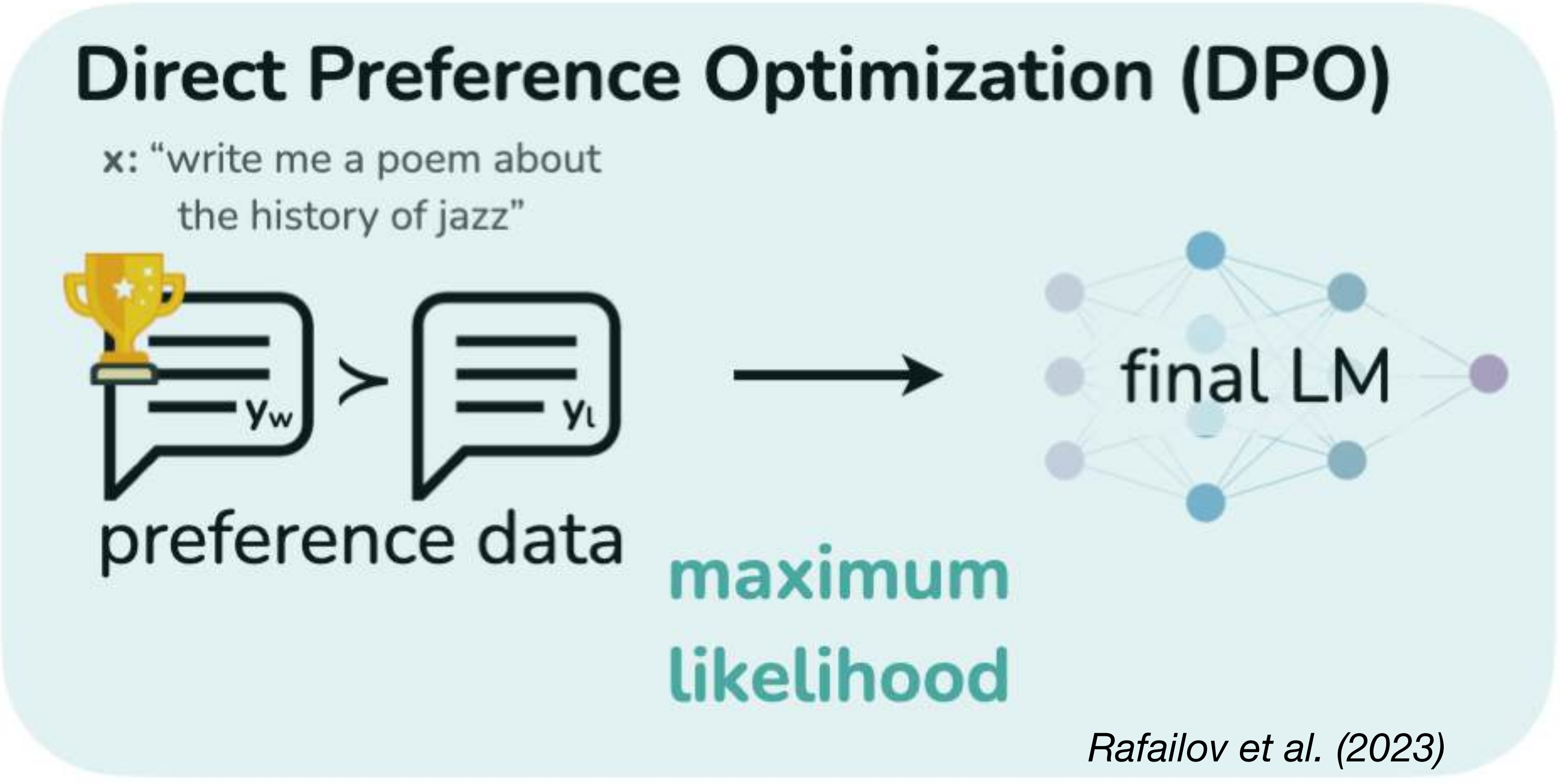
Learning From Comparisons: Examples



artificial comparisons

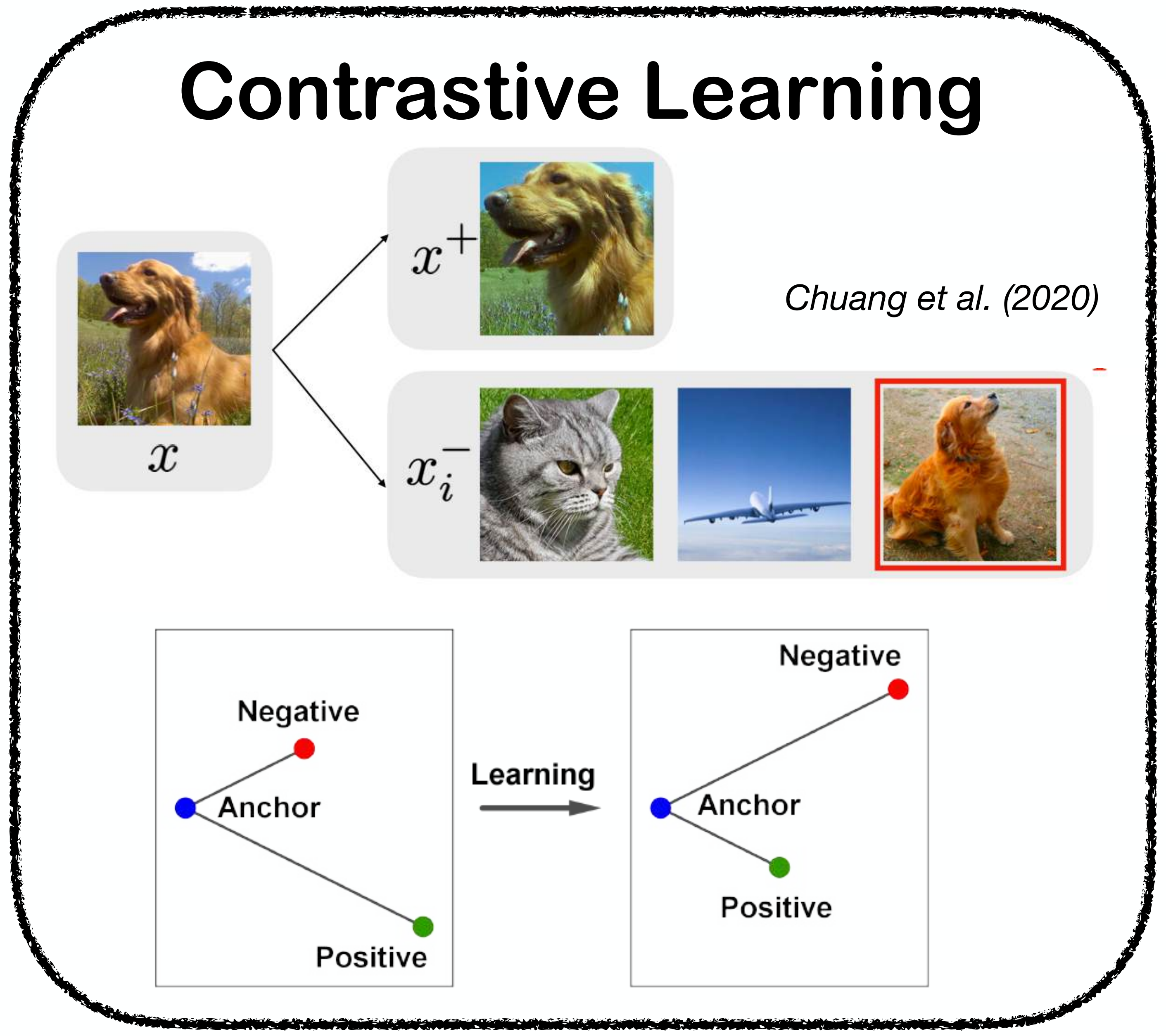


implicit human comparisons

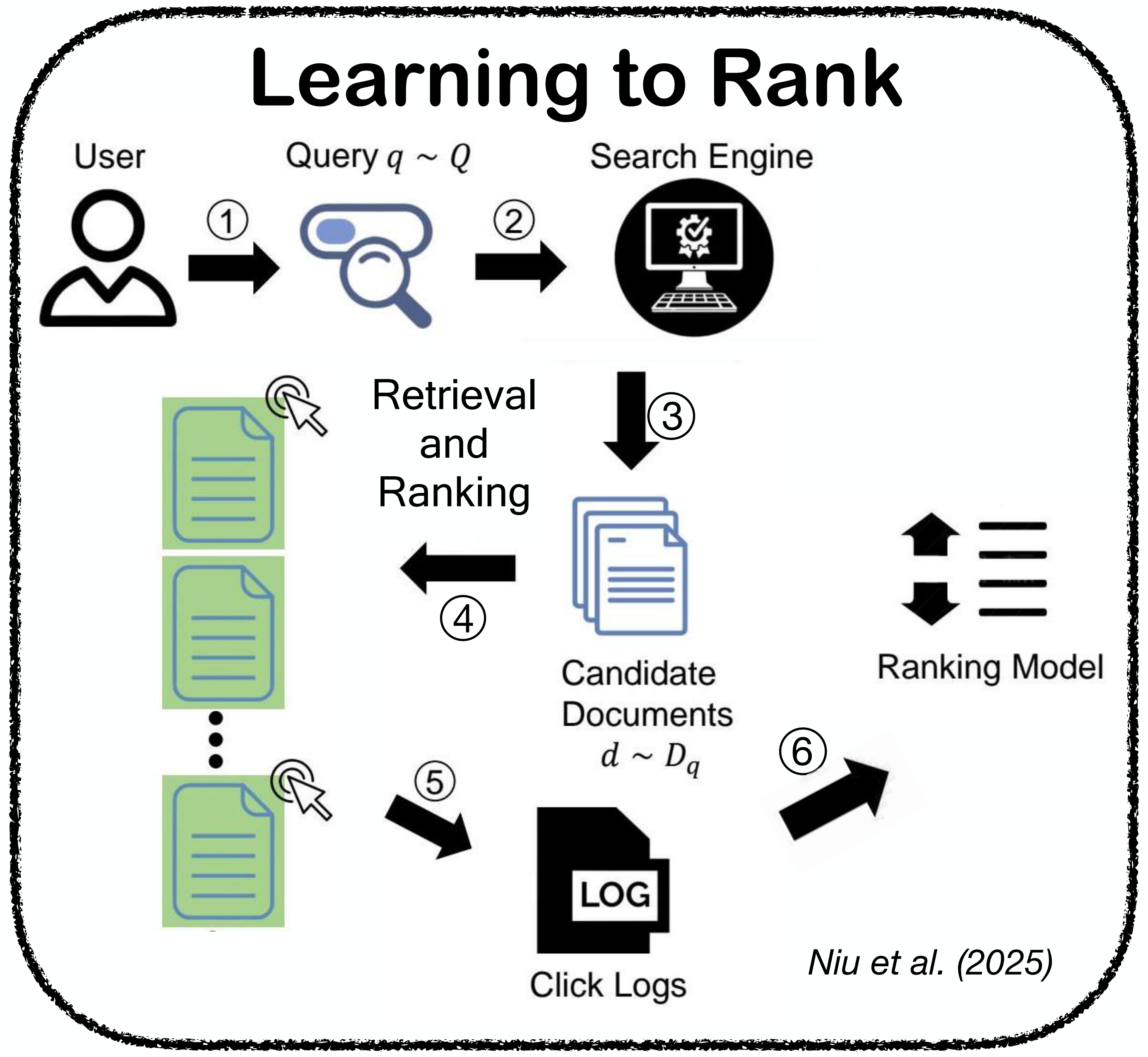


explicit human comparisons

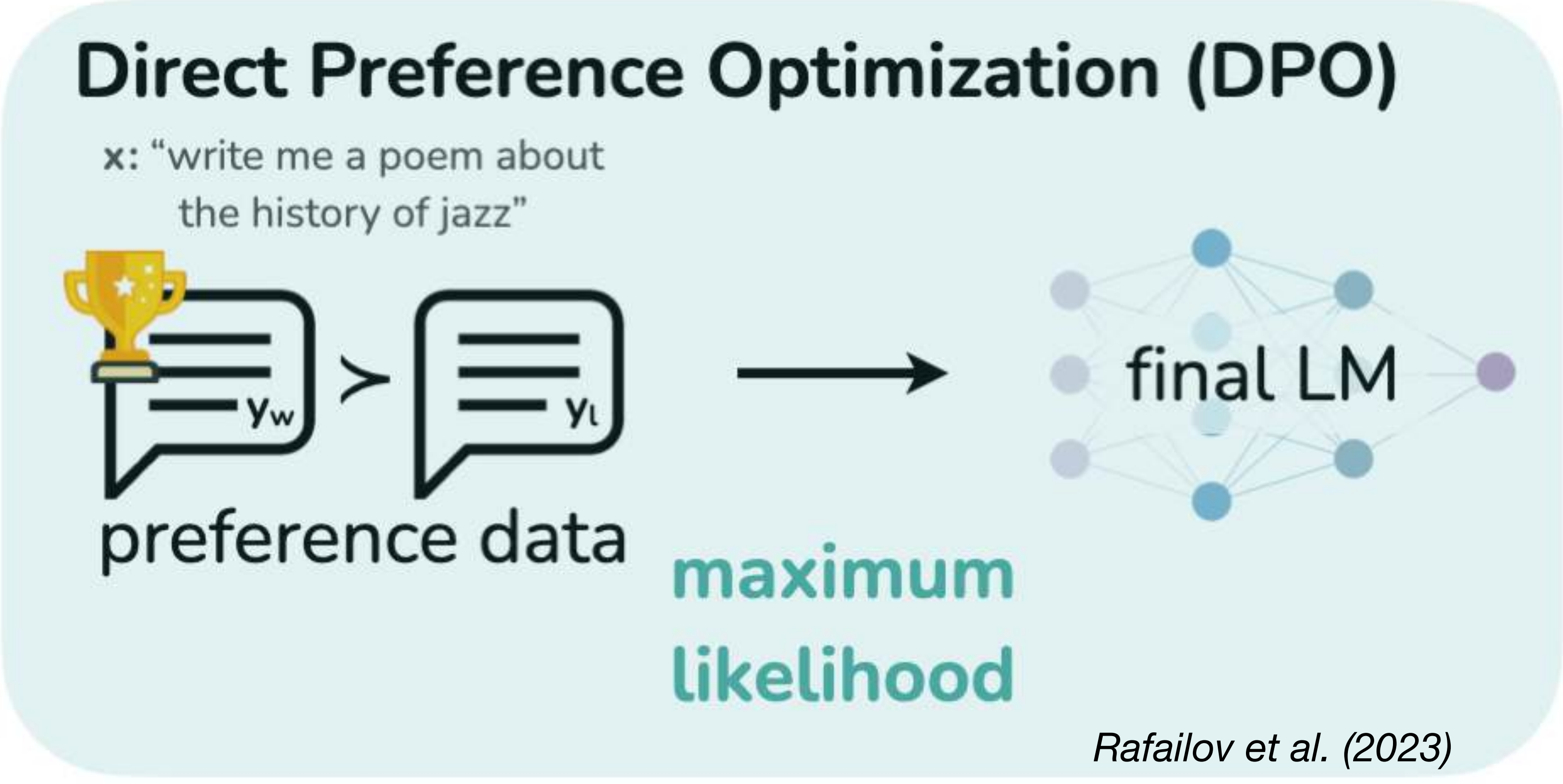
Learning From Comparisons: Examples



artificial comparisons



implicit human comparisons

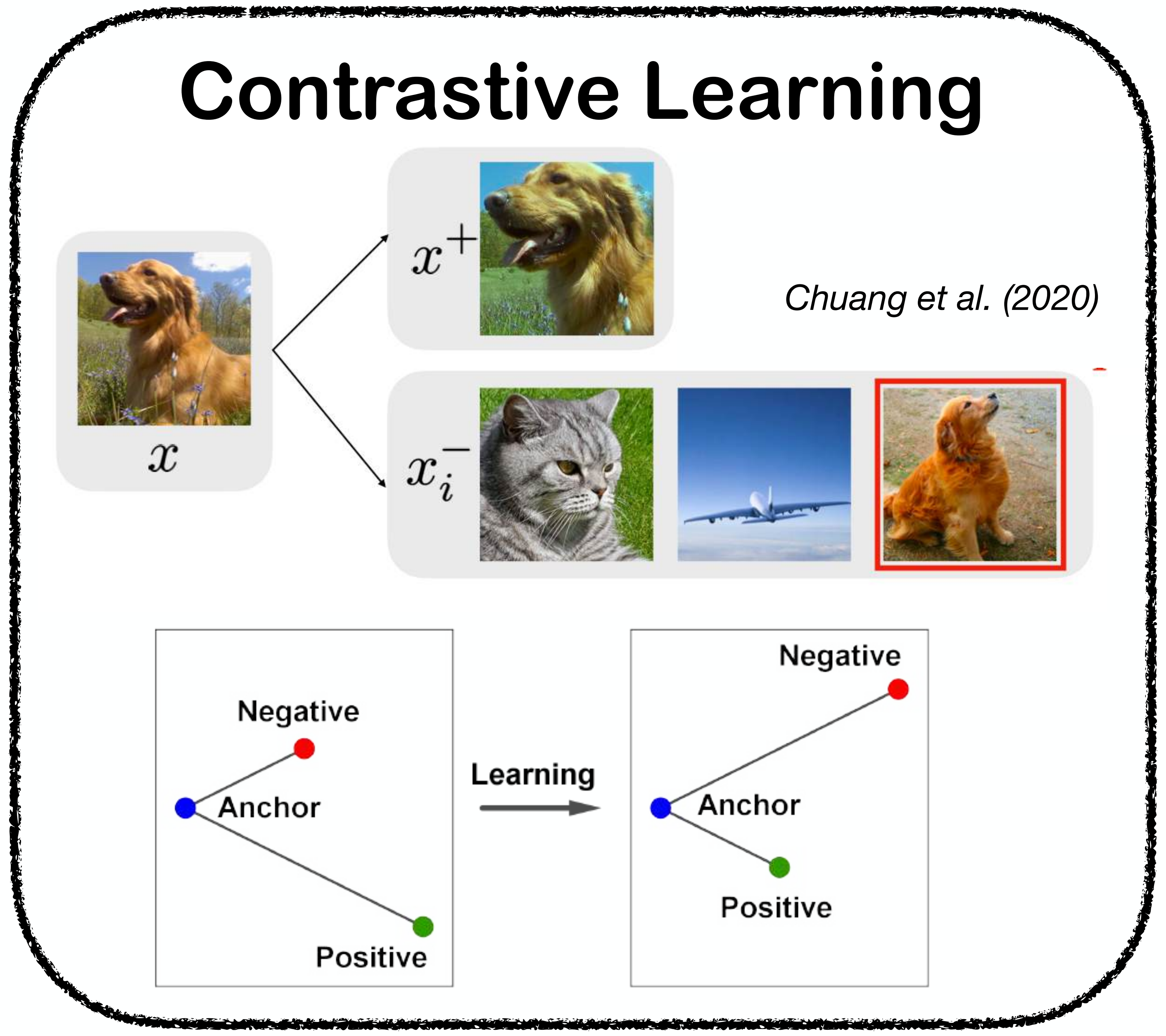


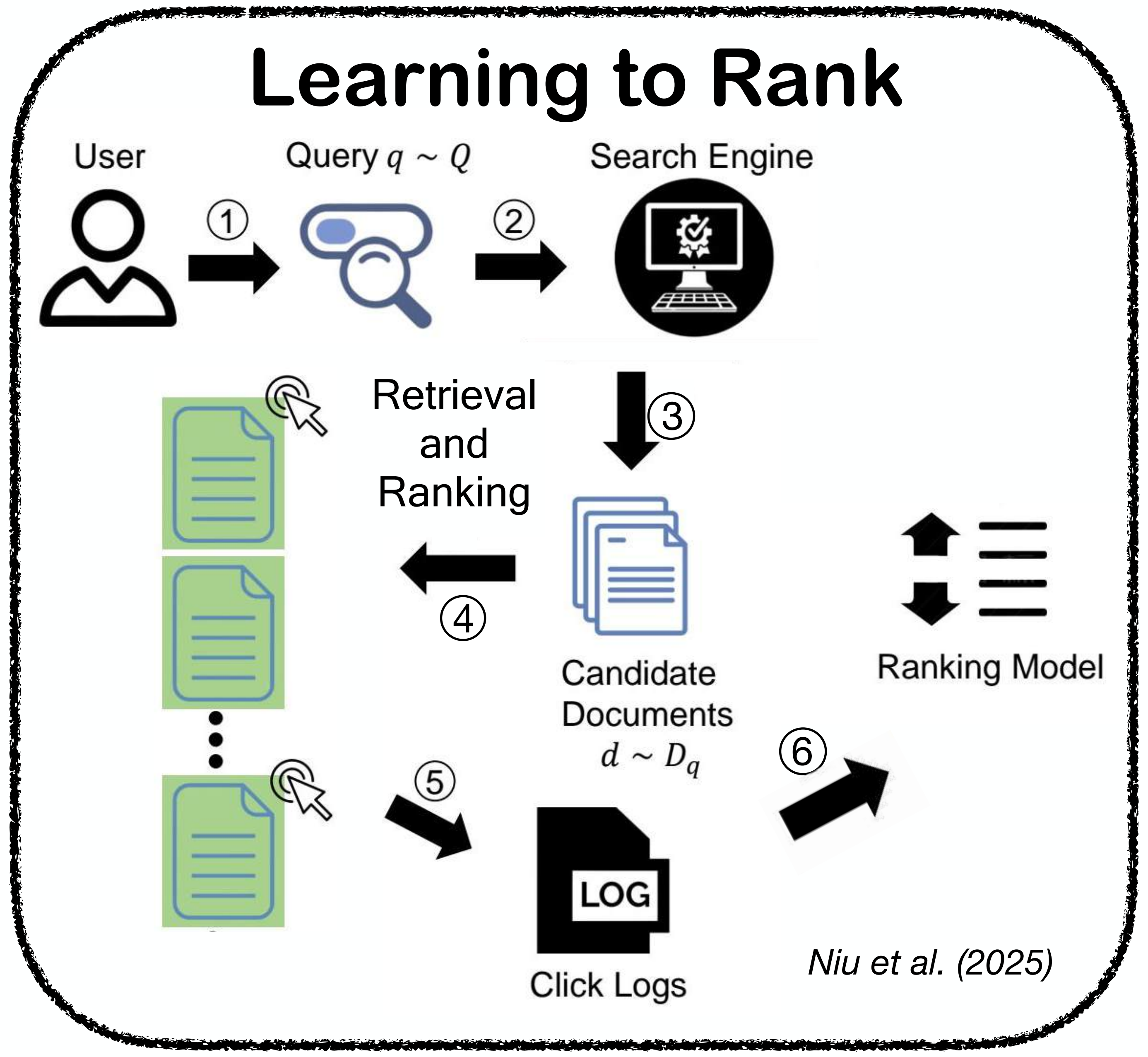
recommender system design

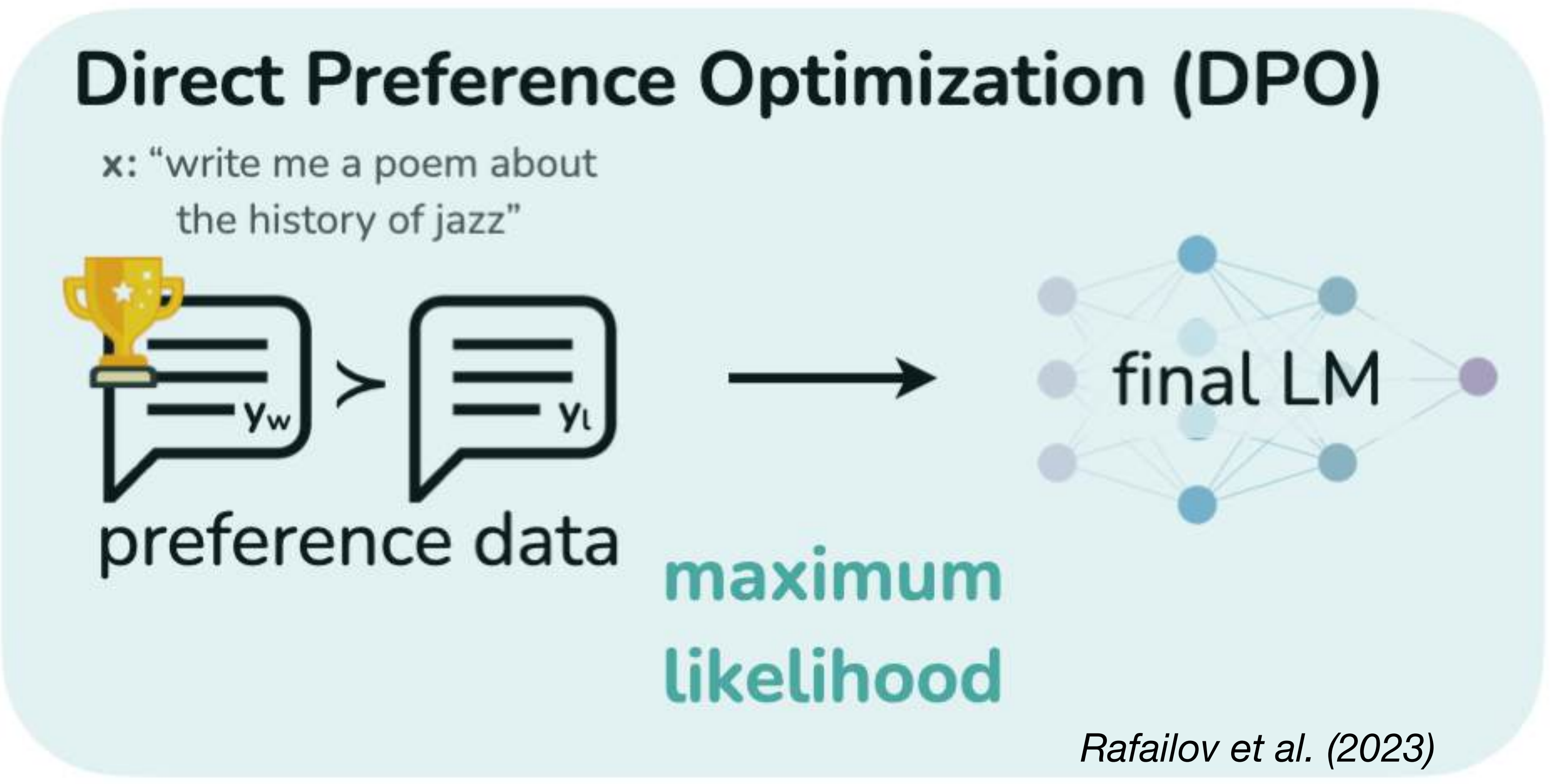
explicit human comparisons

More expensive, more valuable →

Learning From Comparisons: Examples







recommender system design

artificial comparisons | **implicit human comparisons** | **explicit human comparisons**

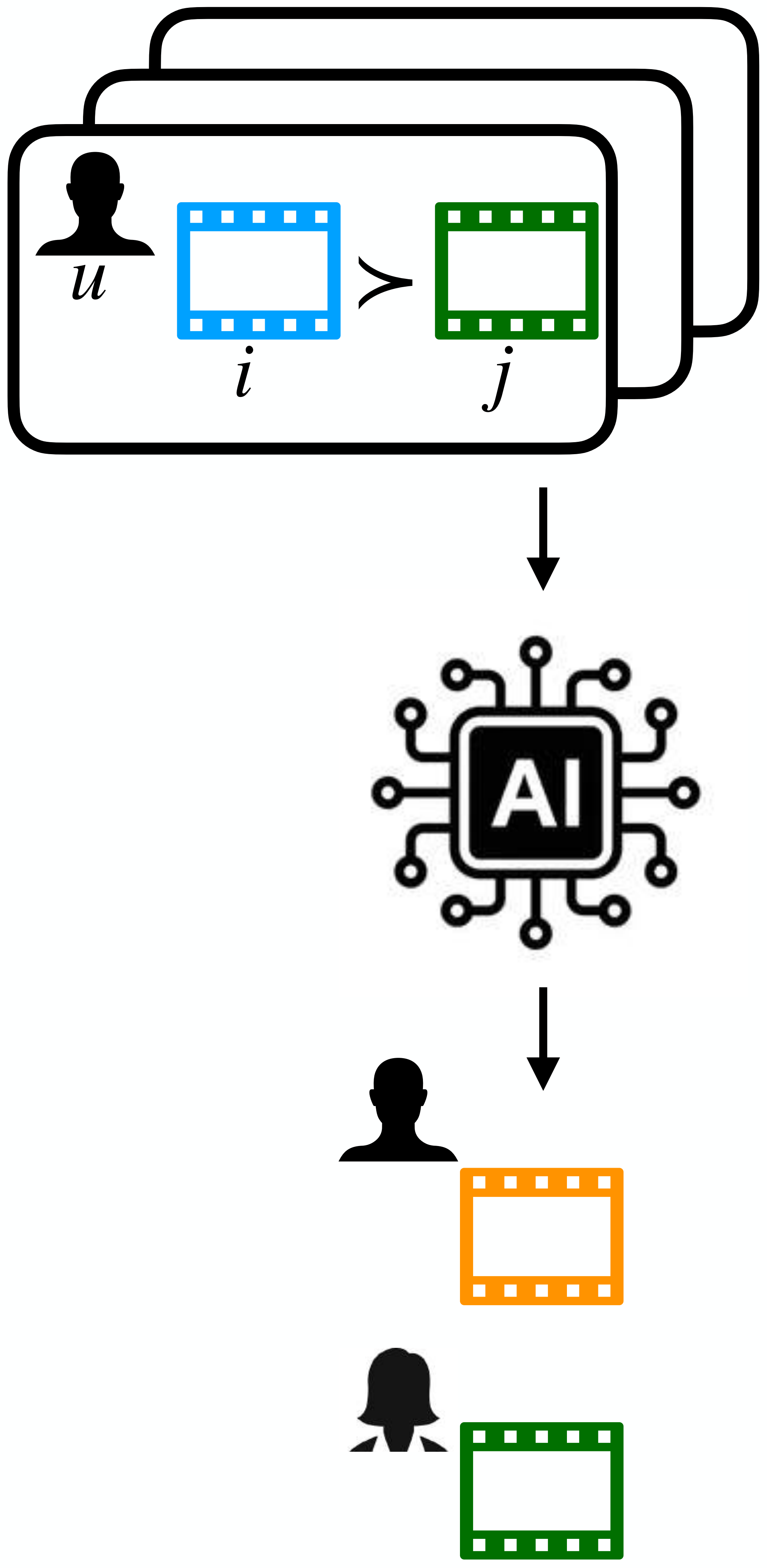
More expensive, more valuable →

- **How much data** is needed to learn personalised utilities from comparisons?
- **What comparisons** should be sought in order to learn utilities quickly?

Learning From Comparisons Is Efficient

Task: Learn personalised tastes of each user, given offline comparison data

How many comparisons are needed? What algorithm should we use?

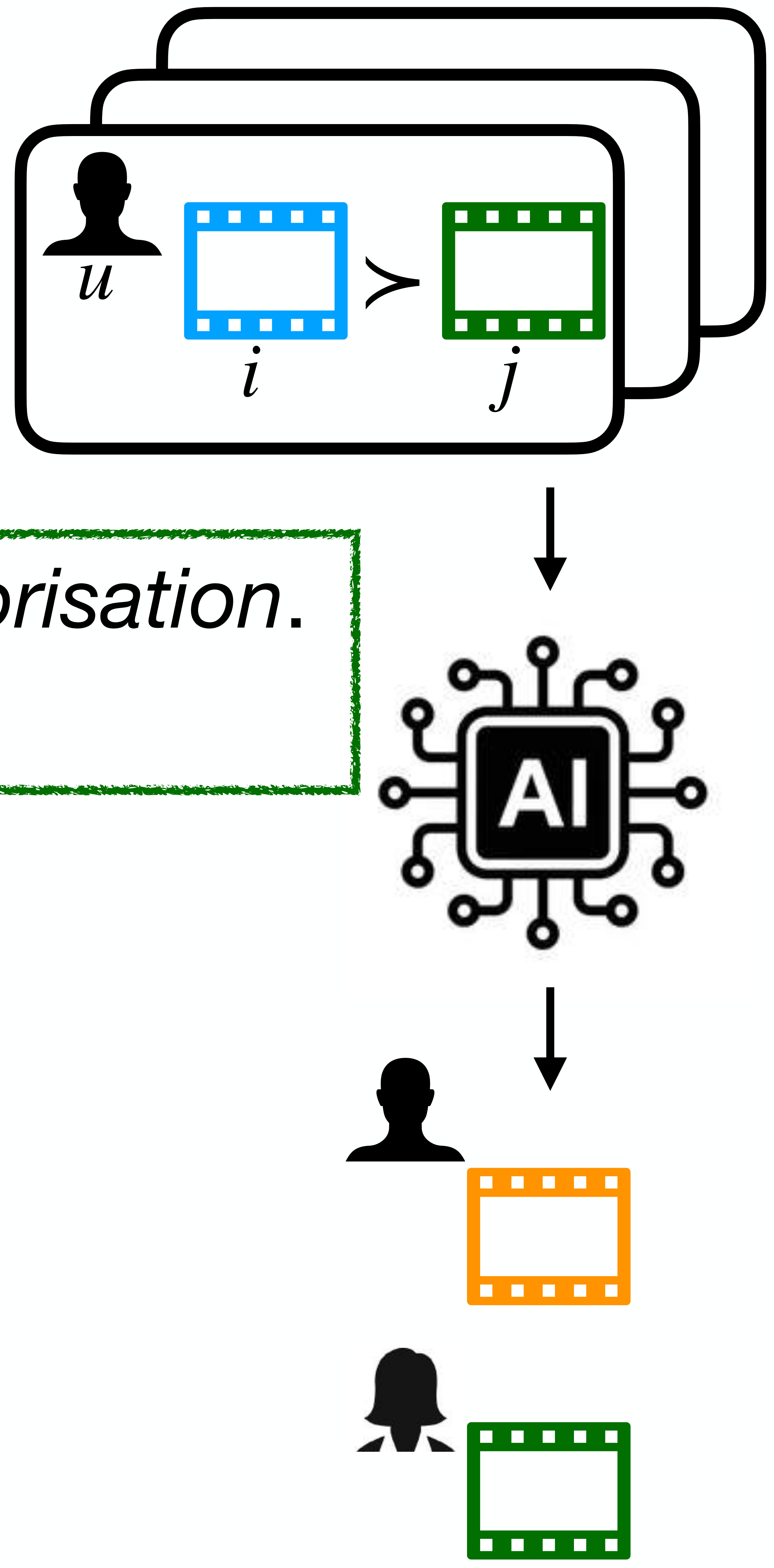


Learning From Comparisons Is Efficient

Task: Learn personalised tastes of each user, given offline comparison data

How many comparisons are needed? What algorithm should we use?

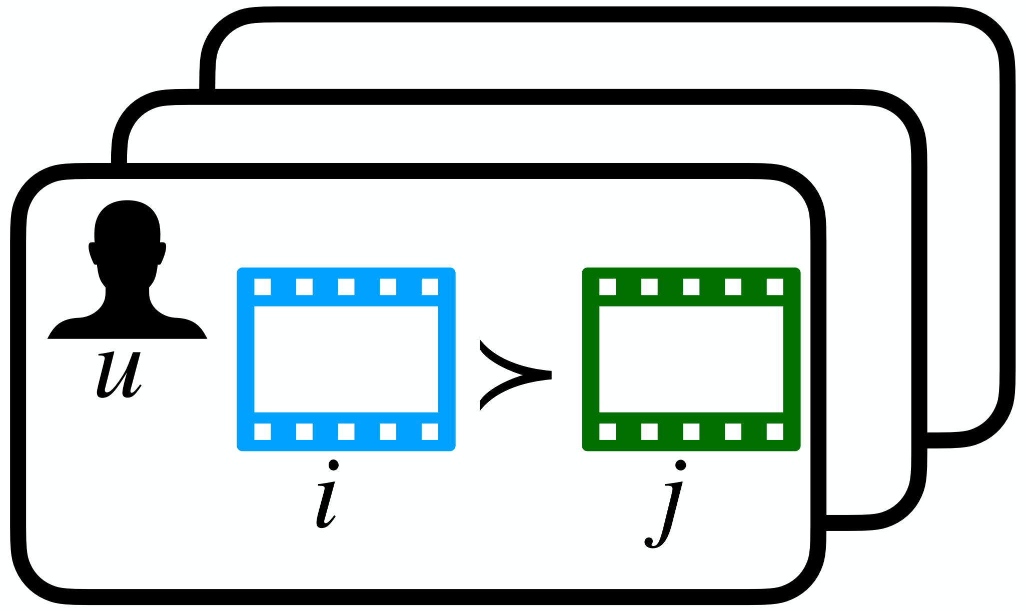
Recommendations with Sparse Comparison Data: Nonconvex Matrix Factorisation.
S. Sankagiri, J. Etesami, and M. Grossglauser (*ICML 2025*)



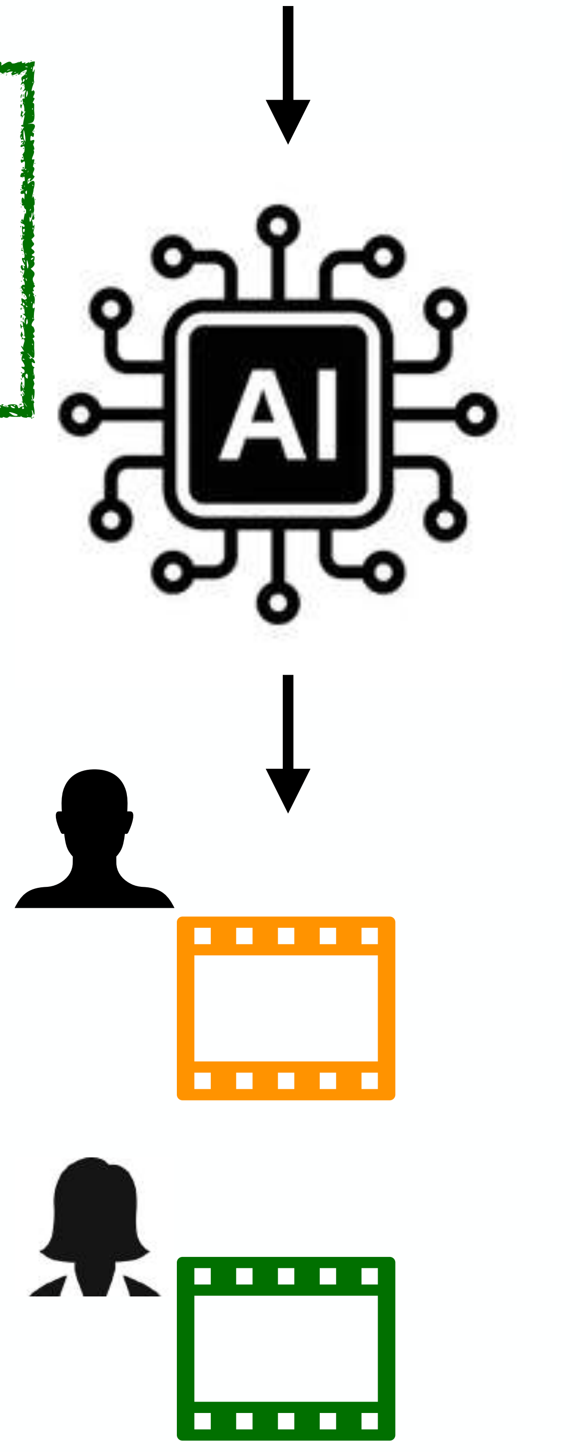
Learning From Comparisons Is Efficient

Task: Learn personalised tastes of each user, given offline comparison data

How many comparisons are needed? What algorithm should we use?



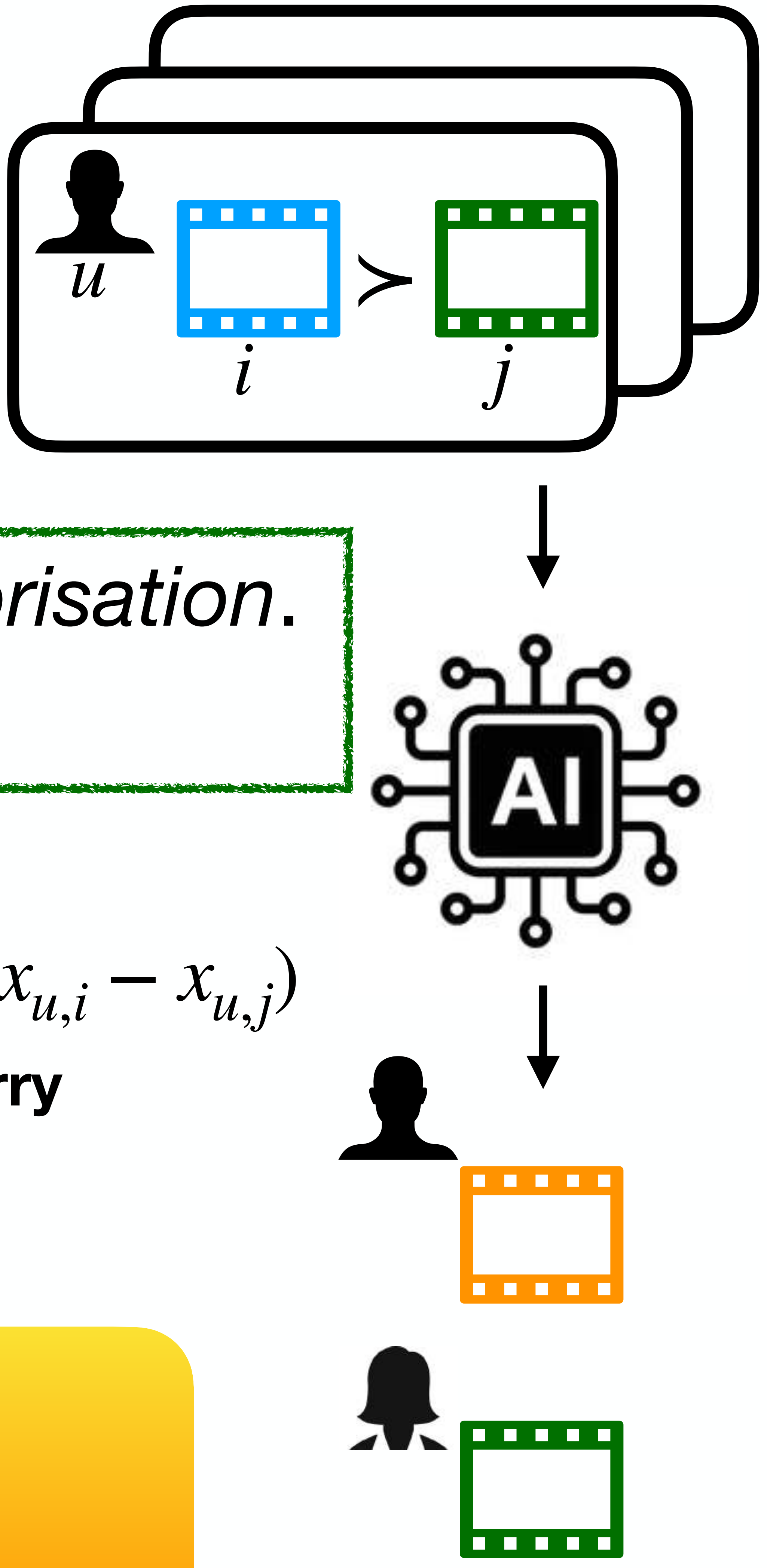
Recommendations with Sparse Comparison Data: Nonconvex Matrix Factorisation.
S. Sankagiri, J. Etesami, and M. Grossglauser (*ICML 2025*)



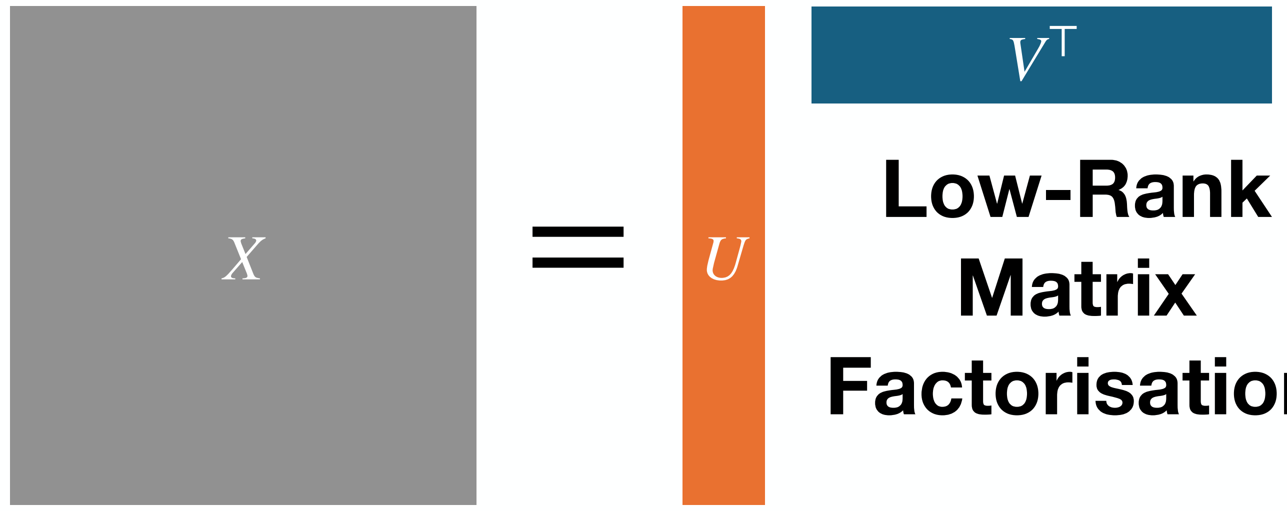
Learning From Comparisons Is Efficient

Task: Learn personalised tastes of each user, given offline comparison data

How many comparisons are needed? What algorithm should we use?



Recommendations with Sparse Comparison Data: Nonconvex Matrix Factorisation.
S. Sankagiri, J. Etesami, and M. Grossglauser (*ICML 2025*)

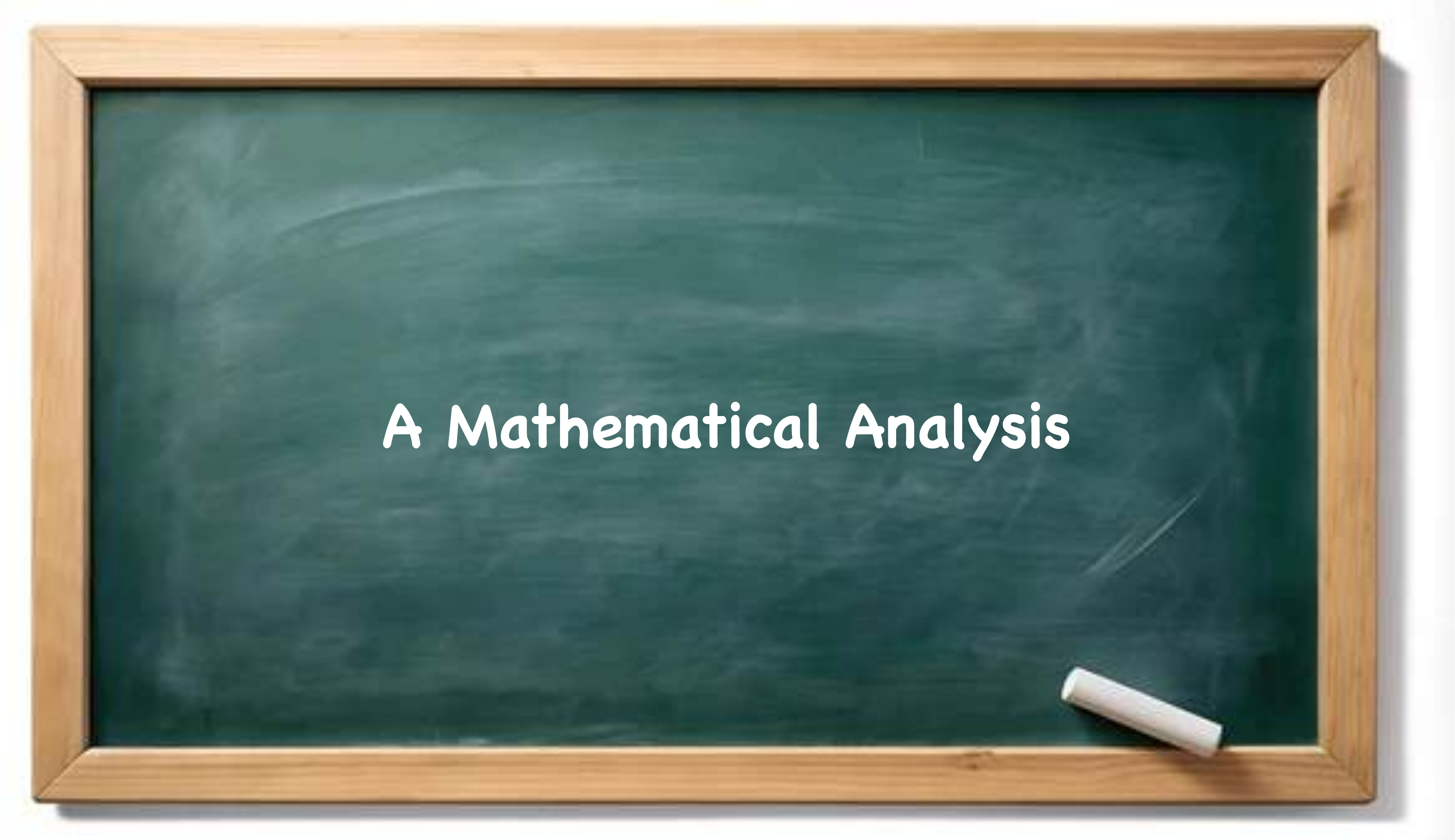


Low-Rank Matrix Factorisation → Learning Model ← **Bradley-Terry Model**

$$\mathbb{P}(i \succ j \mid u) = \sigma(x_{u,i} - x_{u,j})$$

- Personalisation possible with only **few comparisons per user**
- Nonconvex objective can be provably minimised via **gradient descent**

Reasoning About Human Choice



Simple Math Captures Human Choice

Random Utility Model (RUM)

Item i has latent utility x_i

From choice set C,

user picks

$$i^* = \arg\max_{i \in C} \; x_i + \varepsilon_i$$

noisy utility maximisation

Simple Math Captures Human Choice

Random Utility Model (RUM)

Item i has latent utility x_i

From choice set C,

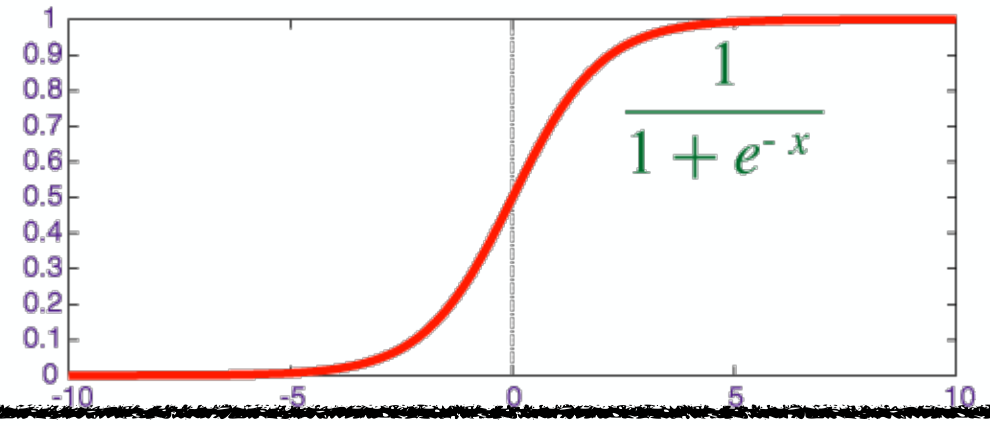
user picks

$$i^* = \arg\max_{i \in C} \; x_i + \varepsilon_i$$

noisy utility maximisation

Bradley-Terry-Luce (BTL) Model

- RUM: $\varepsilon \sim$ i.i.d. Gumbel
- $$\mathbb{P}(i \mid C) = \frac{e^{x_i}}{\Sigma_{j \in C}\, e^{x_j}}$$
  - softmax over utilities
- $\mathbb{P}(i \succ j) = \sigma(x_i - x_j)$



Simple Math Captures Human Choice

Random Utility Model (RUM)

Item i has latent utility x_i

From choice set C,

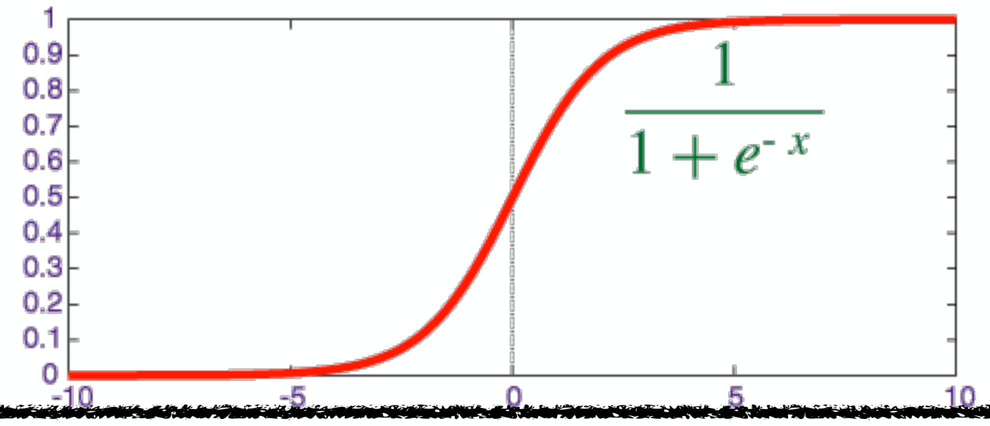
user picks

$$i^* = \arg\max_{i \in C} \; x_i + \varepsilon_i$$

noisy utility maximisation

Bradley-Terry-Luce (BTL) Model

- RUM: $\varepsilon \sim$ i.i.d. Gumbel
- $\mathbb{P}(i \mid C) = \dfrac{e^{x_i}}{\sum_{j \in C} e^{x_j}}$
  - softmax over utilities
- $\mathbb{P}(i \succ j) = \sigma(x_i - x_j)$



Pick out the older person

Simple Math Captures Human Choice

Random Utility Model (RUM)

Item i has latent utility x_i

From choice set C,

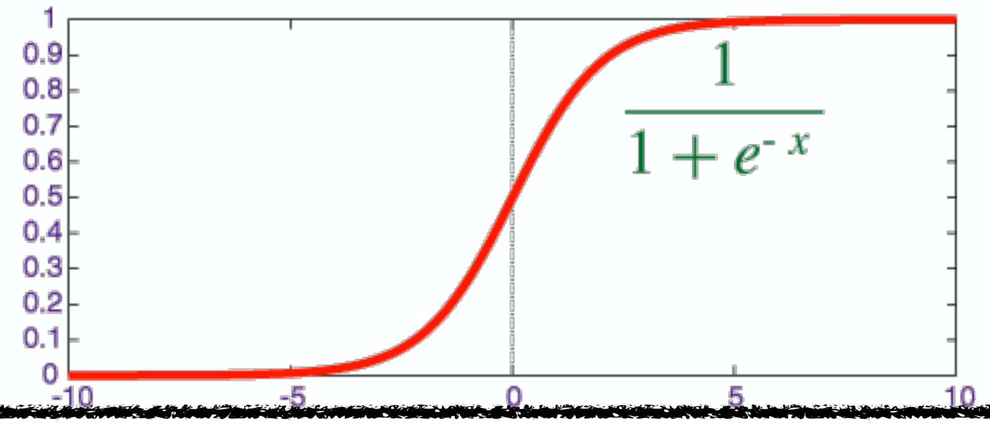
user picks

$$i^* = \arg\max_{i \in C} \; x_i + \varepsilon_i$$

noisy utility maximisation

Bradley-Terry-Luce (BTL) Model

- RUM: $\varepsilon \sim$ i.i.d. Gumbel
- $\mathbb{P}(i \mid C) = \dfrac{e^{x_i}}{\Sigma_{j \in C} \, e^{x_j}}$
  - softmax over utilities
- $\mathbb{P}(i \succ j) = \sigma(x_i - x_j)$



Pick out the older person

Simple Math Captures Human Choice

Random Utility Model (RUM)

Item i has latent utility x_i

From choice set C,

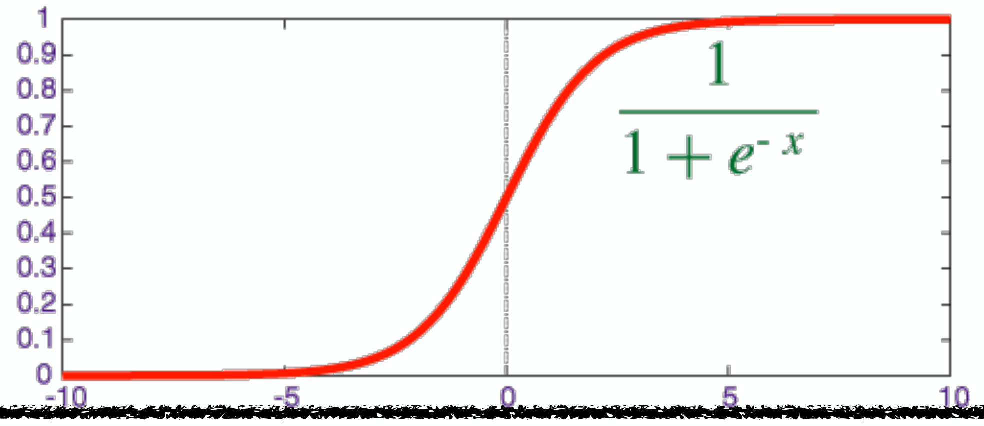
user picks

$$i^* = \arg\max_{i \in C} \; x_i + \varepsilon_i$$

noisy utility maximisation

Bradley-Terry-Luce (BTL) Model

- ✳ RUM: $\varepsilon \sim$ i.i.d. Gumbel
- ✳ $\mathbb{P}(i \mid C) = \dfrac{e^{x_i}}{\Sigma_{j \in C} \; e^{x_j}}$
  - ◉ softmax over utilities
- ✳ $\mathbb{P}(i \succ j) = \sigma(x_i - x_j)$



Pick out the older person

Useful Features

- learning from comparison data ≡ estimating utilities
- strong guarantees
- easily interpretable

not always accurate

a good first-order model

Towards Personalisation

BTL identifies global order

- Recommendation systems need user-dependent rankings ⇒ personalised utilities
- $\mathbb{P}(i \succ j \mid u) = \sigma(x_{u,i} - x_{u,j})$

Towards Personalisation

BTL identifies global order

- Recommendation systems need user-dependent rankings ⇒ personalised utilities

- $\mathbb{P}(i \succ j \mid u) = \sigma(x_{u,i} - x_{u,j})$

Example from Transportation

- $x_{u,i}$: user u's utility for commute mode i
- $x_{u,i} = \beta^\top y_{u,i} \leftarrow$ measured (time, cost, etc)
- Predicted usage of new metro in SF

Towards Personalisation

BTL identifies global order

- Recommendation systems need user-dependent rankings $\Rightarrow$ personalised utilities

- $\mathbb{P}(i \succ j \mid u) = \sigma(x_{u,i} - x_{u,j})$

- For recommendations: need to learn utilities from scratch $X = \{x_{u,i} : u \in \{n_1 \text{ users}\}, i \in \{n_2 \text{ items}\}\}$

- Without further assumptions: $O(n_2 \log n_2)$ comparisons per user

- Each user compares all items! Infeasible

Example from Transportation

- $x_{u,i}$: user u's utility for commute mode i
- $x_{u,i} = \beta^\top y_{u,i} \leftarrow$ measured (time, cost, etc)
- Predicted usage of new metro in SF

Towards Personalisation

BTL identifies global order

- Recommendation systems need user-dependent rankings $\Rightarrow$ personalised utilities
- $\mathbb{P}(i \succ j \mid u) = \sigma(x_{u,i} - x_{u,j})$

- For recommendations:
  need to learn utilities from scratch
  $X = \{x_{u,i} : u \in \{n_1 \text{ users}\}, i \in \{n_2 \text{ items}\}\}$
- Without further assumptions:
  $O(n_2 \log n_2)$ comparisons per user
- Each user compares all items! Infeasible

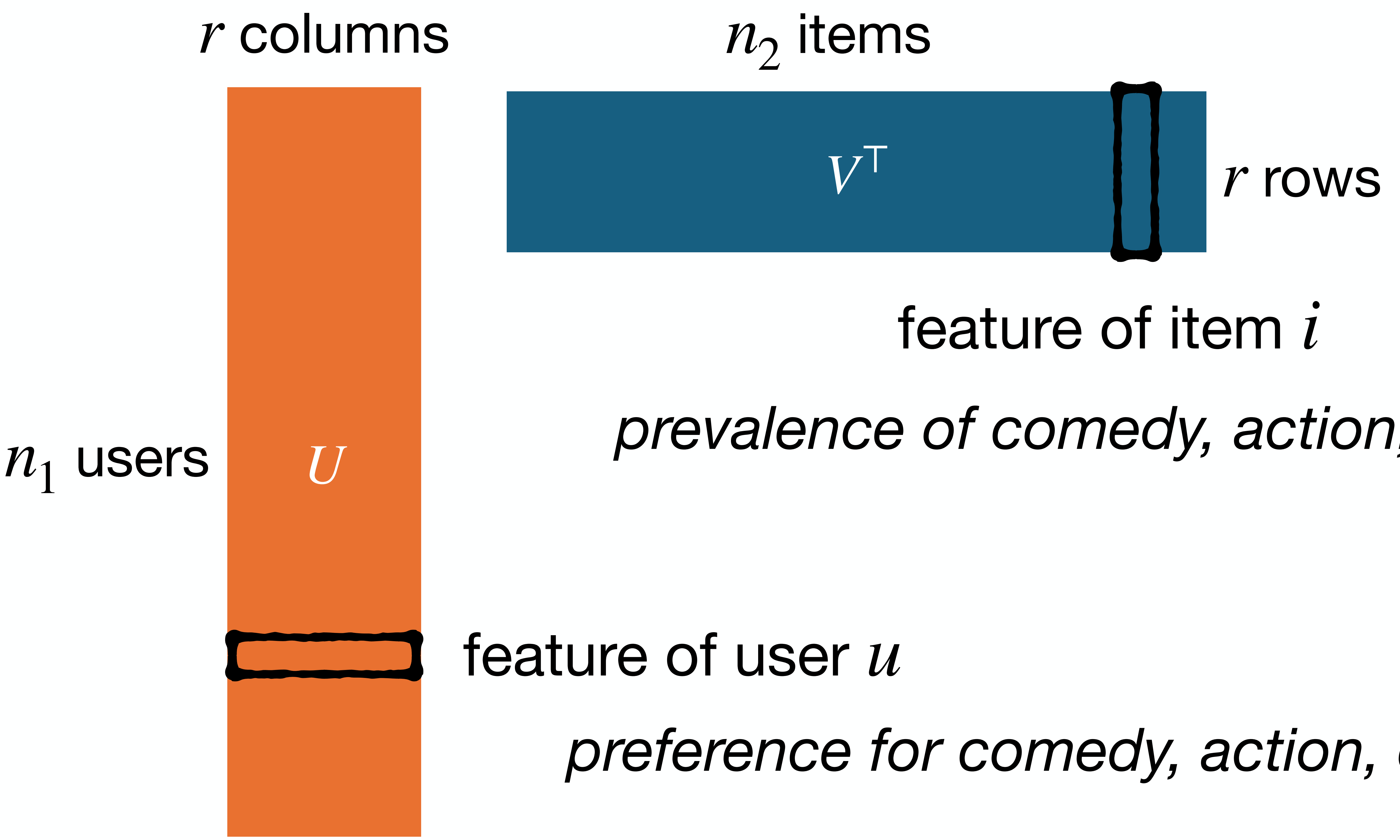
Example from Transportation

- $x_{u,i}$: user u's utility for commute mode i
- $x_{u,i} = \beta^\top y_{u,i} \leftarrow$ measured (time, cost, etc)
- Predicted usage of new metro in SF

Additional Structure Needed

- Assume $x_{u,i} = U_u^\top V_i$
- U_u, V_i: r dimensional *feature vectors*
- Learning $X \leftrightarrow$ learning (U, V)

Interpreting the Parameters



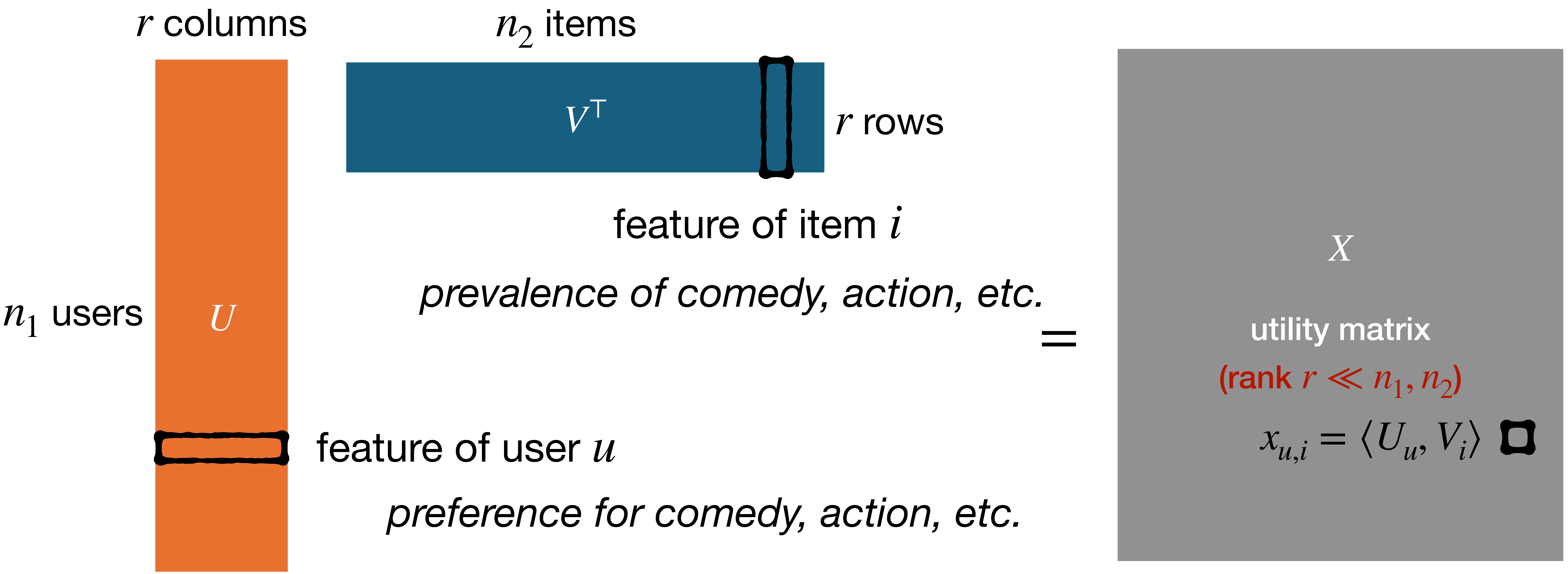
feature of item i

prevalence of comedy, action, etc.

feature of user u

preference for comedy, action, etc.

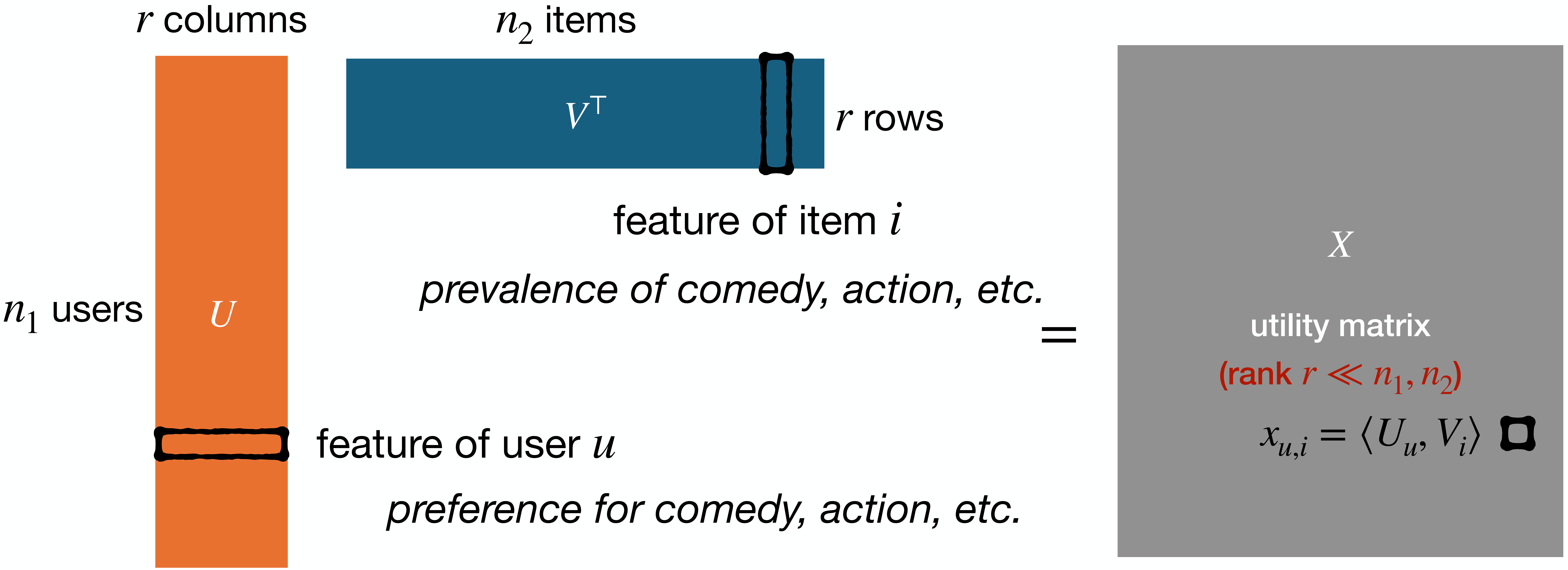
Interpreting the Parameters



<u>Parameter Reduction:</u>

$n_1 n_2 \to nr$
$(n = n_1 + n_2)$

Interpreting the Parameters



Parameter Reduction:

$n_1 n_2 \to nr$

$(n = n_1 + n_2)$

Learning personalised utilities $\leftrightarrow$ learning X $\leftrightarrow$ learning (U, V)

Goal: estimate $Z = (U, V)$ from a dataset of pairwise comparisons

Fundamental Learning Problem

Dataset: $\Omega = \left\{ \text{user } u; \text{items } (i, j); \text{outcome } y = \mathbf{1}\{i \succ j\} \right\}; \ |\Omega| = m$

Fundamental Learning Problem

Dataset: $\Omega = \Big\{\text{user } u; \text{items } (i,j); \text{outcome } y = \mathbf{1}\{i \succ j\}\Big\};\ |\Omega| = m$

Ground-truth Z^* of size $n \times r$

(u, i, j) chosen uniformly at random

$\mathbb{P}(i \succ j \,|\, u) = (1 + \exp(x_{u,i} - x_{u,j}))^{-1}$

Fundamental Learning Problem

Dataset: $\Omega = \left\{\text{user } u; \text{items } (i,j); \text{outcome } y = \mathbf{1}\{i \succ j\}\right\};\ |\Omega| = m$

Ground-truth Z^* of size $n \times r$

(u, i, j) chosen uniformly at random

$\mathbb{P}(i \succ j \,|\, u) = (1 + \exp(x_{u,i} - x_{u,j}))^{-1}$

Maximum Likelihood Estimation:

$$\min_X \sum_{k=1}^{m} \log\left(1 + \exp(x_{u_k,i_k} - x_{u_k,j_k})\right) - y_k\left[x_{u_k,i_k} - x_{u_k,j_k}\right]$$

Fundamental Learning Problem

Dataset: $\Omega = \Big\{\text{user } u; \text{items } (i, j); \text{outcome } y = \mathbf{1}\{i \succ j\}\Big\}; \ |\Omega| = m$

Ground-truth Z^* of size $n \times r$

(u, i, j) chosen uniformly at random

$\mathbb{P}(i \succ j \mid u) = (1 + \exp(x_{u,i} - x_{u,j}))^{-1}$

Maximum Likelihood Estimation:

$$\min_{Z=(U,V)} \sum_{k=1}^{m} \log(1 + \exp(\langle U_{u_k}, V_{i_k} - V_{j_k} \rangle)) - y_k \langle U_{u_k}, V_{i_k} - V_{j_k} \rangle$$

Fundamental Learning Problem

Dataset: $\Omega = \left\{\text{user } u; \text{items } (i,j); \text{outcome } y = \mathbf{1}\{i \succ j\}\right\}; \ |\Omega| = m$

Ground-truth Z^* of size $n \times r$

(u, i, j) chosen uniformly at random

$\mathbb{P}(i \succ j \,|\, u) = (1 + \exp(x_{u,i} - x_{u,j}))^{-1}$

Maximum Likelihood Estimation:

$$\min_{Z=(U,V)} \sum_{k=1}^{m} \log(1 + \exp(\langle U_{u_k}, V_{i_k} - V_{j_k}\rangle)) - y_k\langle U_{u_k}, V_{i_k} - V_{j_k}\rangle$$

Key Research Questions:

(Statistical) How many samples m are required to recover ground-truth Z^*?

Fundamental Learning Problem

Dataset: $\Omega = \{\text{user } u; \text{items } (i,j); \text{outcome } y = \mathbf{1}\{i \succ j\}\}; \ |\Omega| = m$

Ground-truth Z^* of size $n \times r$

(u, i, j) chosen uniformly at random

$\mathbb{P}(i \succ j \,|\, u) = (1 + \exp(x_{u,i} - x_{u,j}))^{-1}$

Maximum Likelihood Estimation:

$$\min_{Z=(U,V)} \sum_{k=1}^{m} \log(1 + \exp(\langle U_{u_k}, V_{i_k} - V_{j_k}\rangle)) - y_k \langle U_{u_k}, V_{i_k} - V_{j_k}\rangle$$

Key Research Questions:

(Statistical) How many samples m are required to recover ground-truth Z^*?

(Algorithmic) How do we solve the above nonconvex optimisation problem?

Theoretical Guarantees

Theorem: (Sankagiri et al, ICML 2025)

Given: $m = \mathcal{O}(nr^2 \log(n))$ *noiseless* comparison samples.

Algorithm: gradient descent with *warm start, small step size*

Guarantee: exponential convergence rate w.h.p:

$$\|\hat{Z}_t - Z^*\| \leq (1-\alpha)^t \, \|\hat{Z}_0 - Z^*\|$$

$$\min_{Z=(U,V)} \sum_{k=1}^{m} \log(1 + \exp(\langle U_{u_k}, V_{i_k} - V_{j_k} \rangle)) - y_k \langle U_{u_k}, V_{i_k} - V_{j_k} \rangle$$

Theoretical Guarantees

Theorem: (Sankagiri et al, ICML 2025)

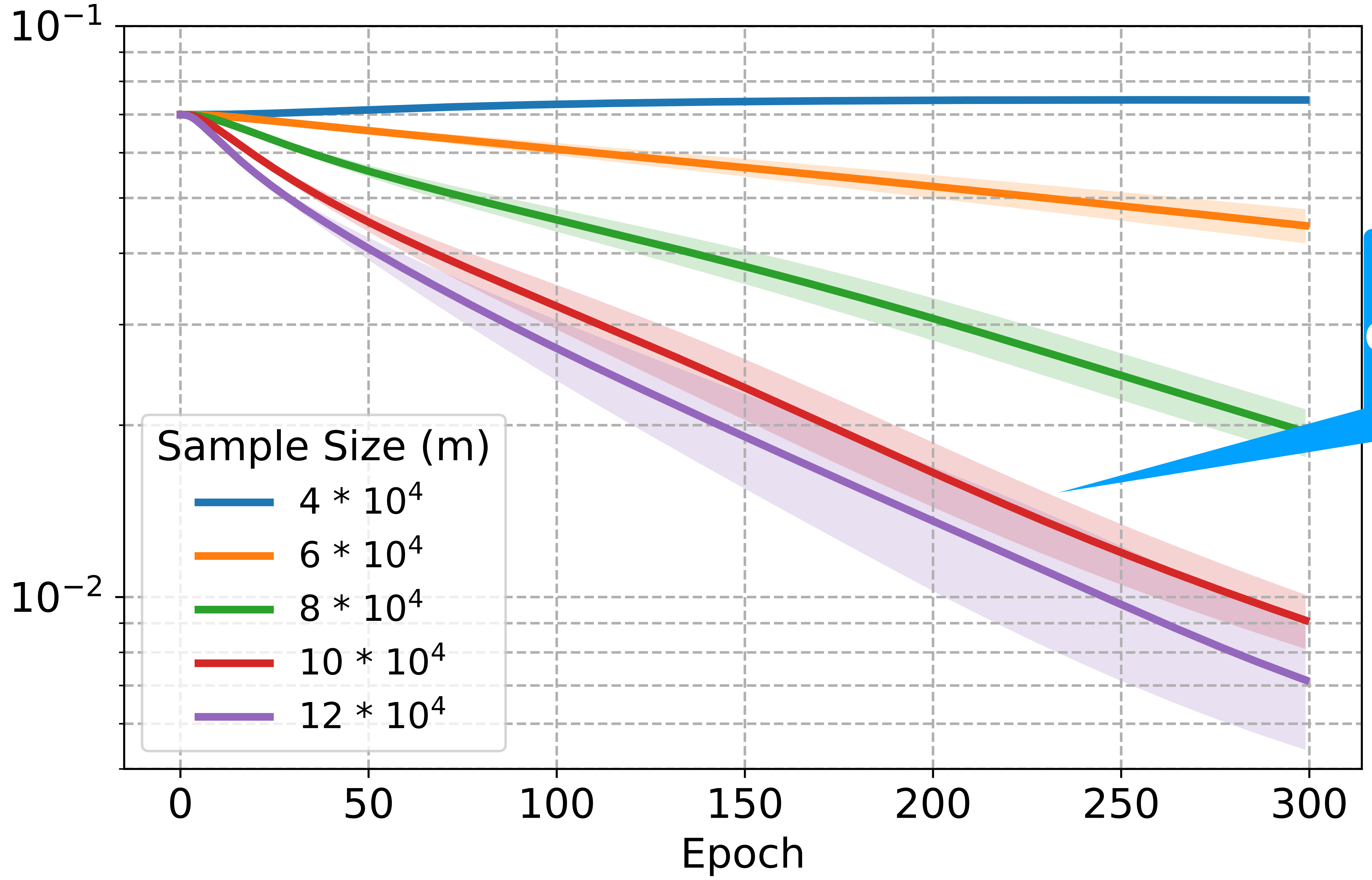
Given: $m = \mathcal{O}(nr^2 \log(n))$ *noiseless* comparison samples.

Algorithm: gradient descent with *warm start, small step size*

Guarantee: exponential convergence rate w.h.p:

$$\|\hat{Z}_t - Z^*\| \leq (1-\alpha)^t \, \|\hat{Z}_0 - Z^*\|$$

Learning a utility matrix of rank 3
2000 users and 3000 items



50 comparisons per user

Error metric (RMSE): $\dfrac{\|X_t - X^*\|_F}{\sqrt{n_1 n_2}}$

Theoretical Guarantees

Theorem: (Sankagiri et al, ICML 2025)

<u>Given:</u> $m = \mathcal{O}(nr^2 \log(n))$

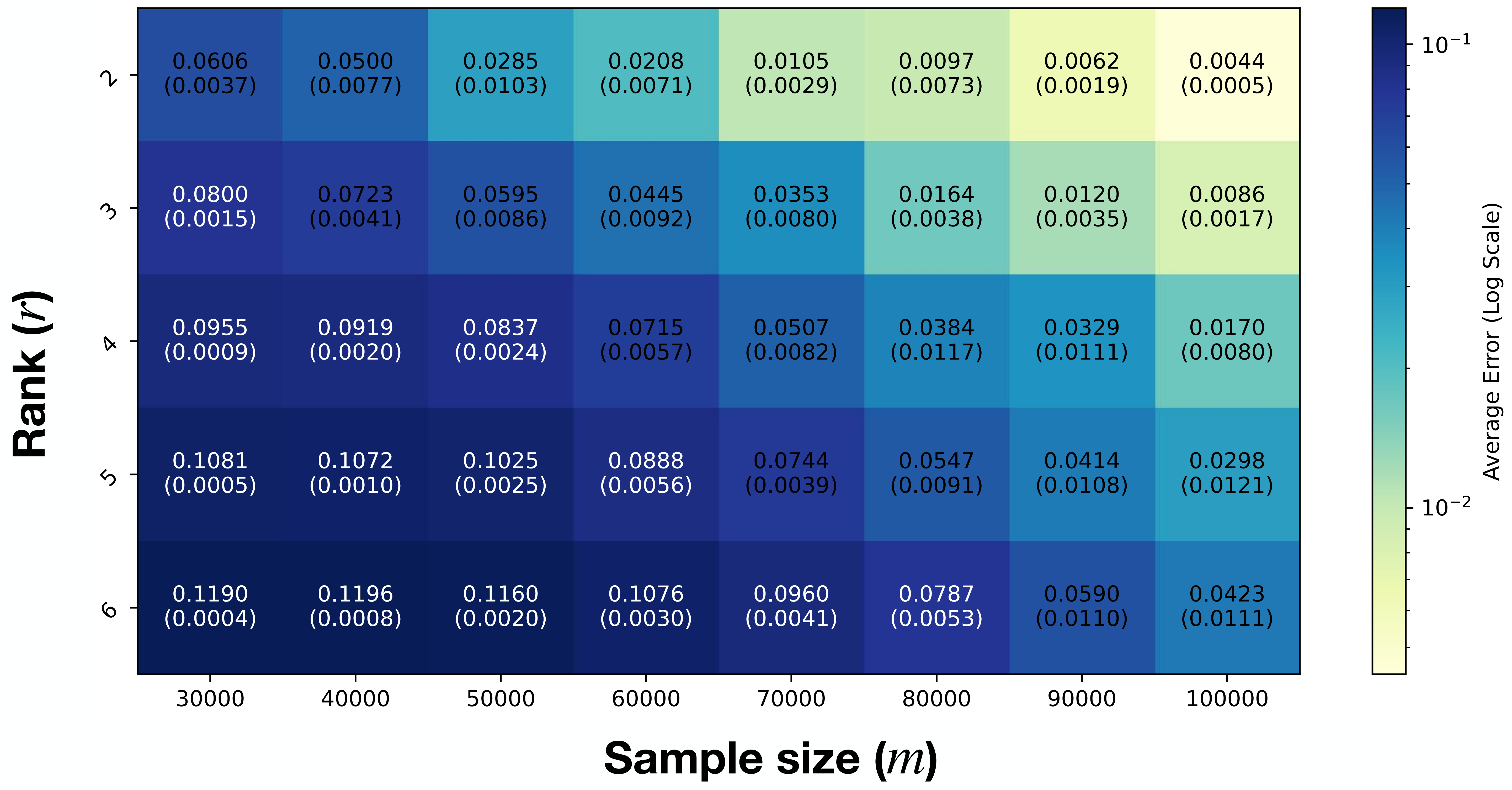
noiseless comparison samples.

<u>Algorithm:</u> gradient descent with *warm start, small step size*

<u>Guarantee:</u> exponential convergence rate w.h.p:

$$\|\hat{Z}_t - Z^*\| \leq (1-\alpha)^t \, \|\hat{Z}_0 - Z^*\|$$

Learning a utility matrix of rank 3
2000 users and 3000 items



Error metric (RMSE): $\dfrac{\|X_t - X^*\|_F}{\sqrt{n_1 n_2}}$

Comparison With Prior Work

Theoretical Guarantees

- Learning from ratings well-studied:

$$y_{u,i} = x_{u,i} + \varepsilon_{u,i}, (u, i) \in \Omega$$

- Our results $\sim$ SoTA for this problem

first such results for comparisons

Comparison With Prior Work

Theoretical Guarantees

- Learning from ratings well-studied:

$$y_{u,i} = x_{u,i} + \varepsilon_{u,i}, (u, i) \in \Omega$$

- Our results $\sim$ SoTA for this problem

first such results for comparisons

- Comparisons as informative as ratings

Comparison With Prior Work

Theoretical Guarantees

- Learning from ratings well-studied:

$$y_{u,i} = x_{u,i} + \varepsilon_{u,i}, (u, i) \in \Omega$$

- Our results ~ SoTA for this problem

first such results for comparisons

- Comparisons as informative as ratings

Empirical Observations

- Comparison-based learning problem studied before, without guarantees

- Data: infer comparisons from ratings:

$$y_{u,i} = 5, y_{u,j} = 3 \Rightarrow \quad u : i > j$$

- Prediction quality matches ratings!

Comparison With Prior Work

Theoretical Guarantees

- Learning from ratings well-studied:

$$y_{u,i} = x_{u,i} + \varepsilon_{u,i}, (u, i) \in \Omega$$

- Our results $\sim$ SoTA for this problem

first such results for comparisons

- Comparisons as informative as ratings

Empirical Observations

- Comparison-based learning problem studied before, <u>without guarantees</u>
- <u>Data:</u> infer comparisons from ratings:

$$y_{u,i} = 5, y_{u,j} = 3 \Rightarrow \quad u : i \succ j$$

- Prediction quality matches ratings!

No dataset with user comparisons

Comparison With Prior Work

Theoretical Guarantees

- Learning from ratings well-studied:

$$y_{u,i} = x_{u,i} + \varepsilon_{u,i}, (u, i) \in \Omega$$

- Our results $\sim$ SoTA for this problem

first such results for comparisons

- Comparisons as informative as ratings

Empirical Observations

- Comparison-based learning problem studied before, without guarantees

- Data: infer comparisons from ratings:

$$y_{u,i} = 5, y_{u,j} = 3 \Rightarrow u : i > j$$

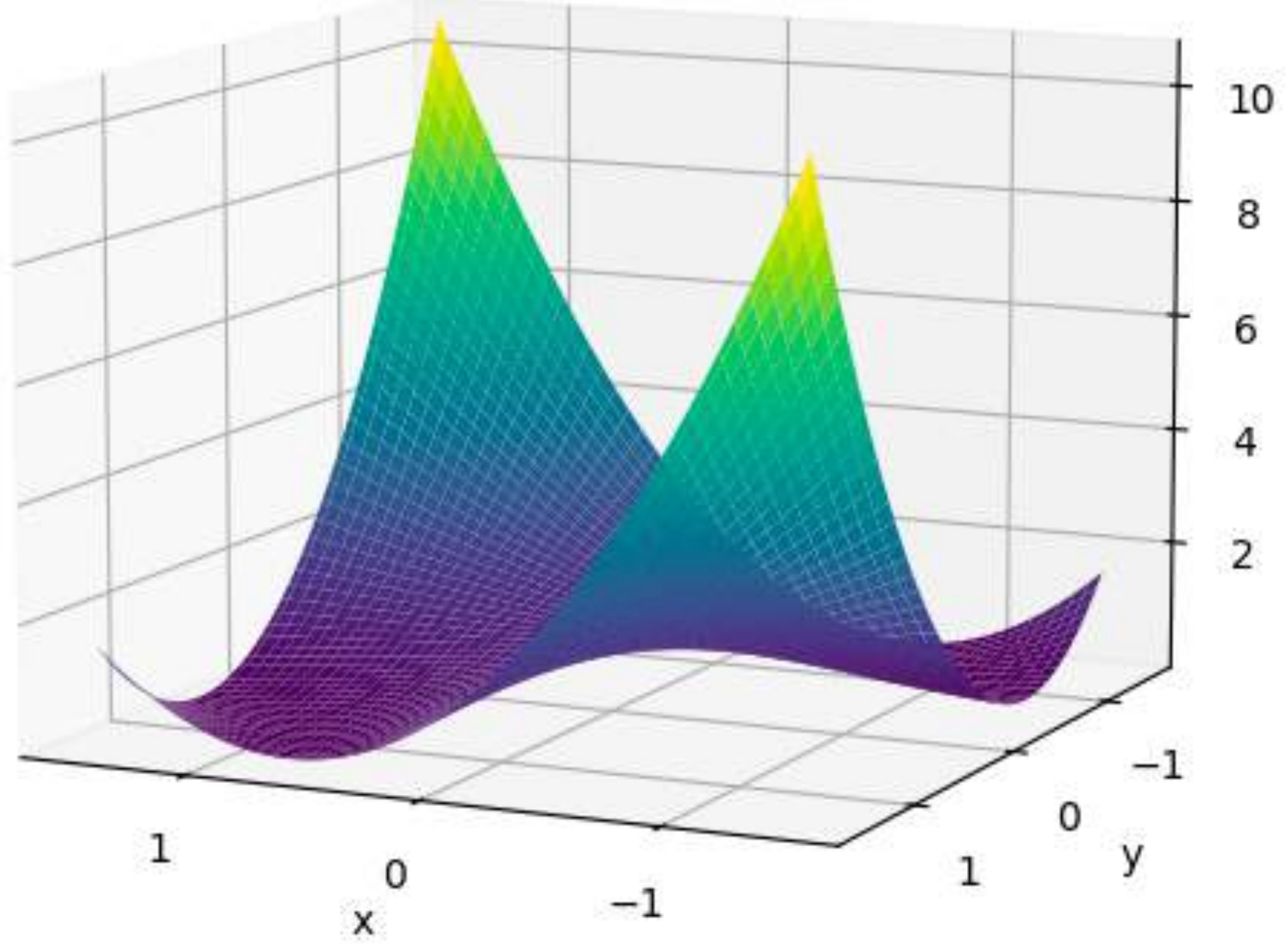
- Prediction quality matches ratings!

No dataset with user comparisons

Open Question: ratings or comparisons —what's better?

Gradient Descent Works Despite Nonconvexity

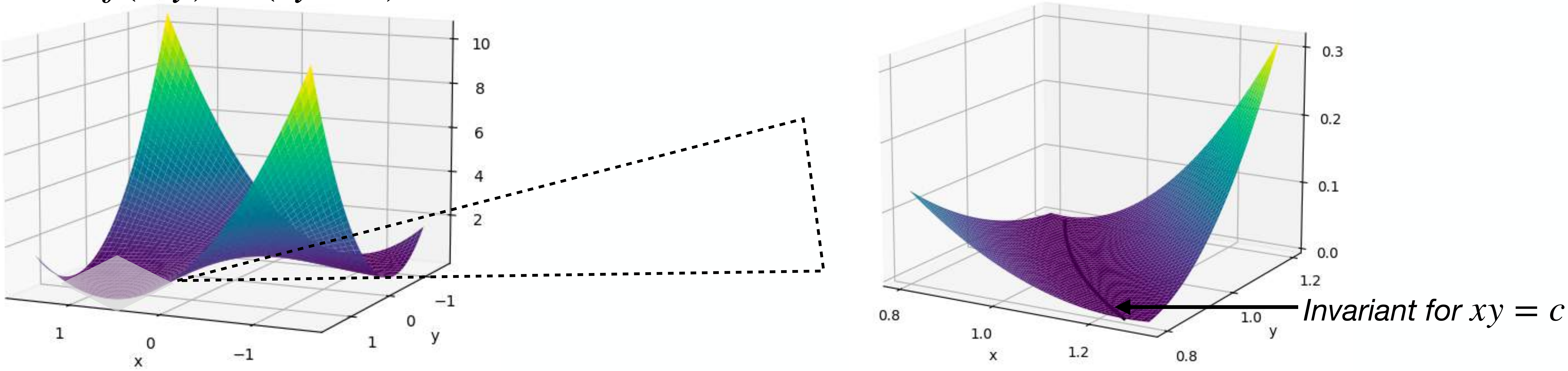
$$f(x, y) = (xy - 1)^2$$



Globally non-convex

Gradient Descent Works Despite Nonconvexity

$f(x, y) = (xy - 1)^2$

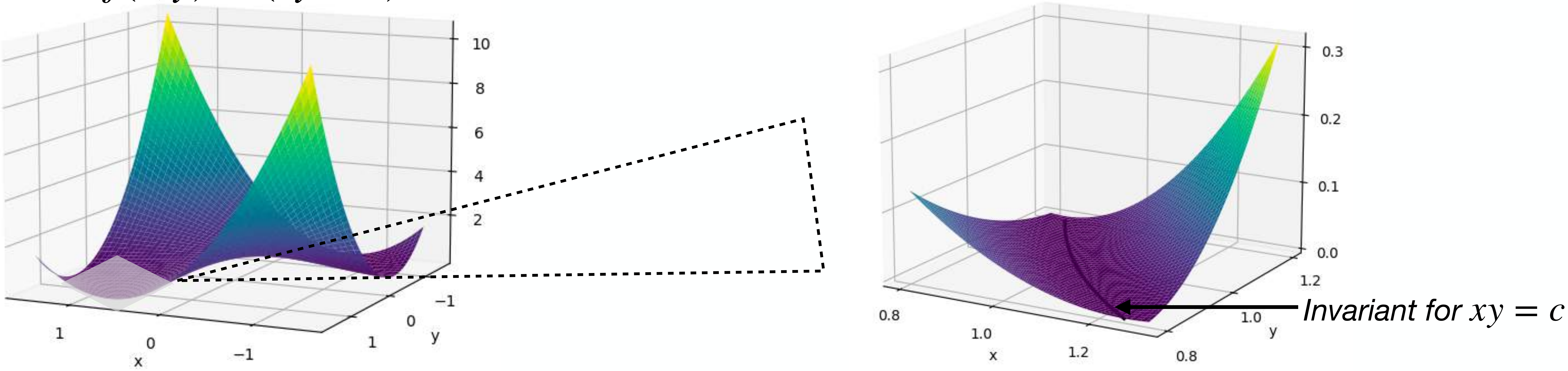


Globally non-convex

Locally strongly convex

Gradient Descent Works Despite Nonconvexity

$f(x, y) = (xy - 1)^2$



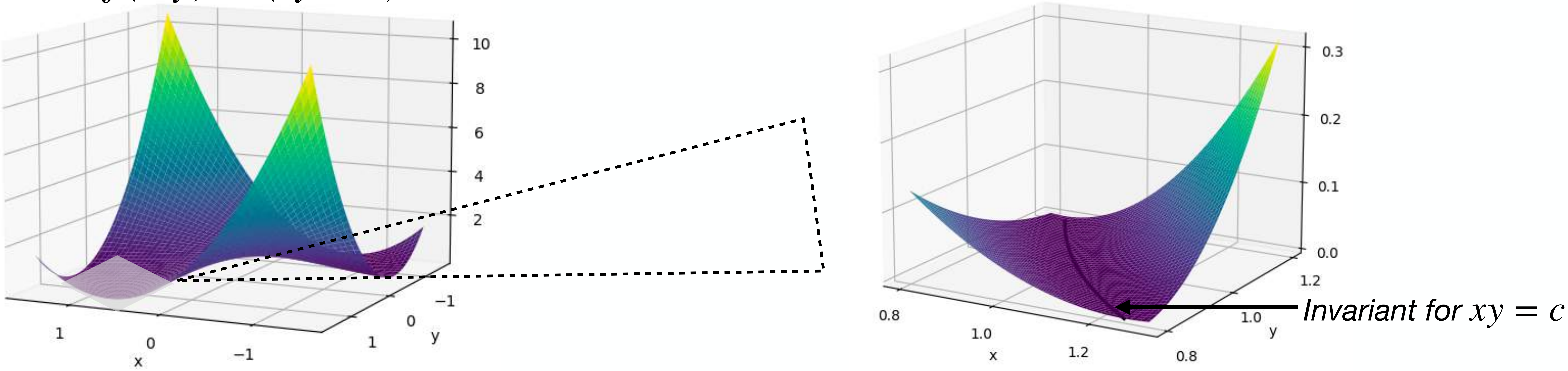
Globally non-convex

Locally strongly convex

- Overcome rotational invariance: $\Delta(Z) = Z - Z^* R(Z),\ R(Z) = \min_{R^T R = I} \|Z - Z^* R\|_F$

Gradient Descent Works Despite Nonconvexity

$f(x, y) = (xy - 1)^2$



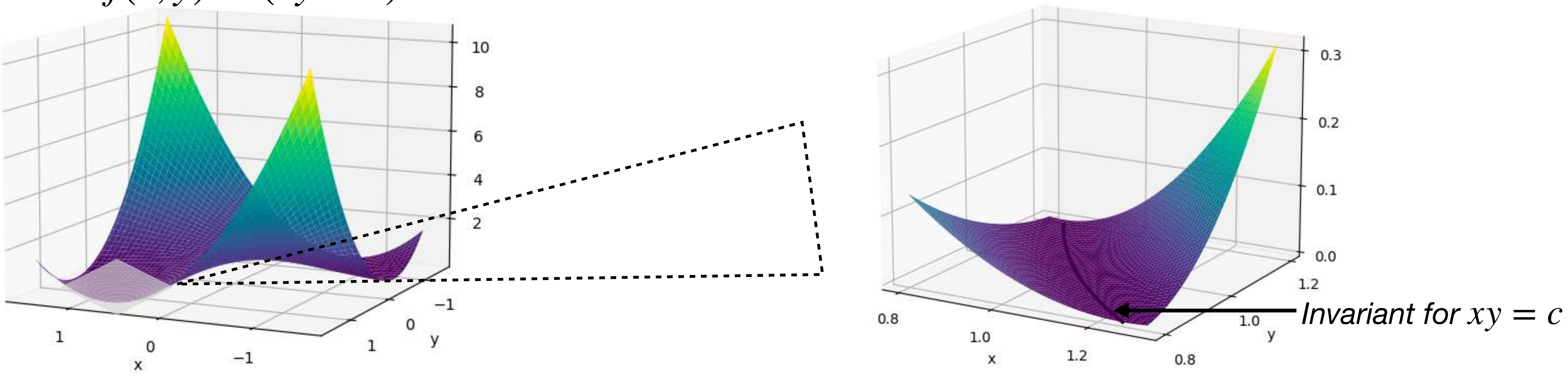
Globally non-convex

Locally strongly convex

- Overcome rotational invariance: $\Delta(Z) = Z - Z^*R(Z),\ R(Z) = \min_{R^TR=I} \|Z - Z^*R\|_F$
- Show
    1. **strong convexity** $\langle \nabla \mathscr{L}(Z), \Delta(Z) \rangle \geq c_1 \|\Delta(Z)\|_F^2$

Gradient Descent Works Despite Nonconvexity

$f(x, y) = (xy - 1)^2$



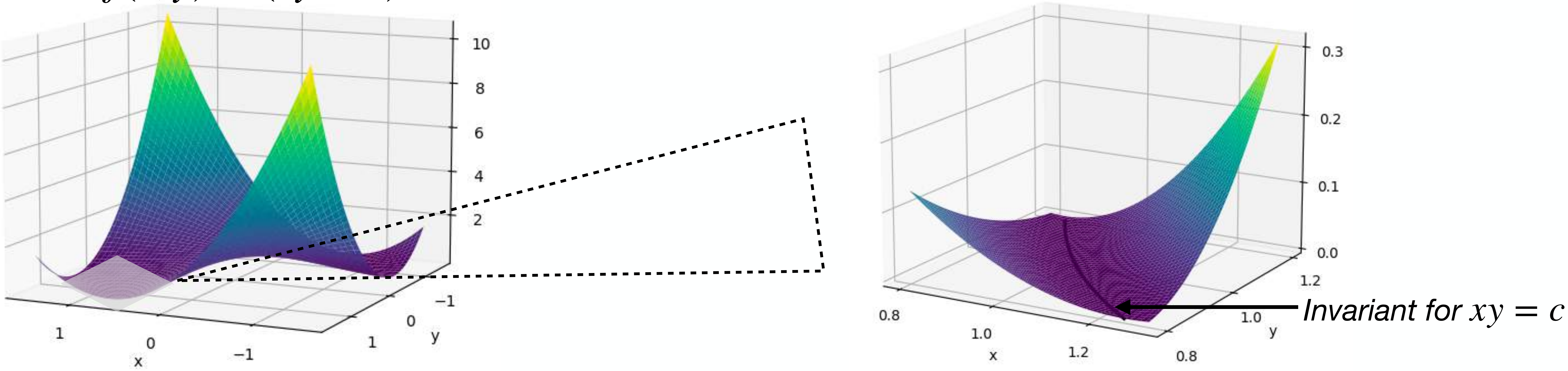
Globally non-convex

Locally strongly convex

- Overcome rotational invariance: $\Delta(Z) = Z - Z^* R(Z),\ R(Z) = \min_{R^T R = I} \|Z - Z^* R\|_F$
- Show
  1. **strong convexity** $\langle \nabla \mathscr{L}(Z), \Delta(Z) \rangle \geq c_1 \|\Delta(Z)\|_F^2$
  2. **smoothness** $\|\nabla \mathscr{L}(Z)\|^2 \leq c_2 \|\Delta(Z)\|_F^2$

Gradient Descent Works Despite Nonconvexity

$f(x, y) = (xy - 1)^2$



Globally non-convex

Locally strongly convex

- Overcome rotational invariance: $\Delta(Z) = Z - Z^*R(Z),\ R(Z) = \min_{R^TR=I} \|Z - Z^*R\|_F$
- Show
  1. **strong convexity** $\langle \nabla \mathscr{L}(Z), \Delta(Z) \rangle \geq c_1 \|\Delta(Z)\|_F^2$
  2. **smoothness** $\|\nabla \mathscr{L}(Z)\|^2 \leq c_2 \|\Delta(Z)\|_F^2$

  $\Rightarrow$ gradient descent gives exponential convergence

Establishing Concentration Bound

Goal: With high probability, for all Z: $\langle \nabla \mathscr{L}(Z), \Delta(Z) \rangle \gtrsim \|\Delta(Z)\|_F^2$

Establishing Concentration Bound

Goal: With high probability, <u>for all Z</u>: $\langle \nabla \mathscr{L}(Z), \Delta(Z) \rangle \gtrsim \|\Delta(Z)\|_F^2$

1. **Taylor Approximation:** Show $\langle \nabla \mathscr{L}(Z), \Delta(Z) \rangle \geq p^\top \bar{A} p$

- $p = \text{vec}\left(\Delta(Z)\right)$ is a *location-dependent* vector

- *i.i.d.*
  $\bar{A} = \frac{1}{m} \sum_{k=1}^{m} A_k(u_k, i_k, j_k)$ is a *data-dependent* matrix; $\mathbb{E}[A] \approx I$

Establishing Concentration Bound

Goal: With high probability, <u>for all Z</u>: $\langle \nabla \mathscr{L}(Z), \Delta(Z) \rangle \gtrsim \|\Delta(Z)\|_F^2$

1. **Taylor Approximation:** Show $\langle \nabla \mathscr{L}(Z), \Delta(Z) \rangle \geq p^\top \bar{A} p$

- $p = \text{vec}\left(\Delta(Z)\right)$ is a *location-dependent* vector
- $\bar{A} = \frac{1}{m} \sum_{k=1}^{m} A_k(u_k, i_k, j_k)$ is a *data-dependent* matrix; $\mathbb{E}[A] \approx I$

i.i.d.

2. **Concentration:** $\mathbb{P}(\|\bar{A} - \mathbb{E}[A]\|_{op} \geq t) \leq 2nr \exp\left(-\frac{mt^2}{nr^2}\right)$

Matrix Bernstein inequality

Establishing Concentration Bound

Goal: With high probability, <u>for all Z</u>: $\langle \nabla \mathscr{L}(Z), \Delta(Z) \rangle \gtrsim \|\Delta(Z)\|_F^2$

1. **Taylor Approximation:** Show $\langle \nabla \mathscr{L}(Z), \Delta(Z) \rangle \geq p^\top \bar{A} p$
   - $p = \text{vec}\left(\Delta(Z)\right)$ is a *location-dependent* vector
   - $\bar{A} = \frac{1}{m} \sum_{k=1}^{m} A_k(u_k, i_k, j_k)$ (*i.i.d.*) is a *data-dependent* matrix; $\mathbb{E}[A] \approx I$
2. **Concentration:** $\mathbb{P}(\|\bar{A} - \mathbb{E}[A]\|_{op} \geq t) \leq 2nr \exp\left(-\frac{mt^2}{nr^2}\right)$ — Matrix Bernstein inequality
3. **Conclusion:** $p^\top \bar{A} p \simeq p^\top \mathbb{E}[A] p \simeq \|p\|_2^2 = \|\Delta(Z)\|_F^2$

Establishing Concentration Bound

Goal: With high probability, <u>for all Z</u>: $\langle \nabla \mathscr{L}(Z), \Delta(Z) \rangle \gtrsim \|\Delta(Z)\|_F^2$

1. **Taylor Approximation:** Show $\langle \nabla \mathscr{L}(Z), \Delta(Z) \rangle \geq p^\top \bar{A} p$
   - $p = \text{vec}\left(\Delta(Z)\right)$ is a *location-dependent* vector
   - *i.i.d.* $\bar{A} = \frac{1}{m} \sum_{k=1}^{m} A_k(u_k, i_k, j_k)$ is a *data-dependent* matrix; $\mathbb{E}[A] \approx I$

Open: How general is this new proof technique?

2. **Concentration:** $\mathbb{P}(\|\bar{A} - \mathbb{E}[A]\|_{op} \geq t) \leq 2nr \exp\left(-\frac{mt^2}{nr^2}\right)$

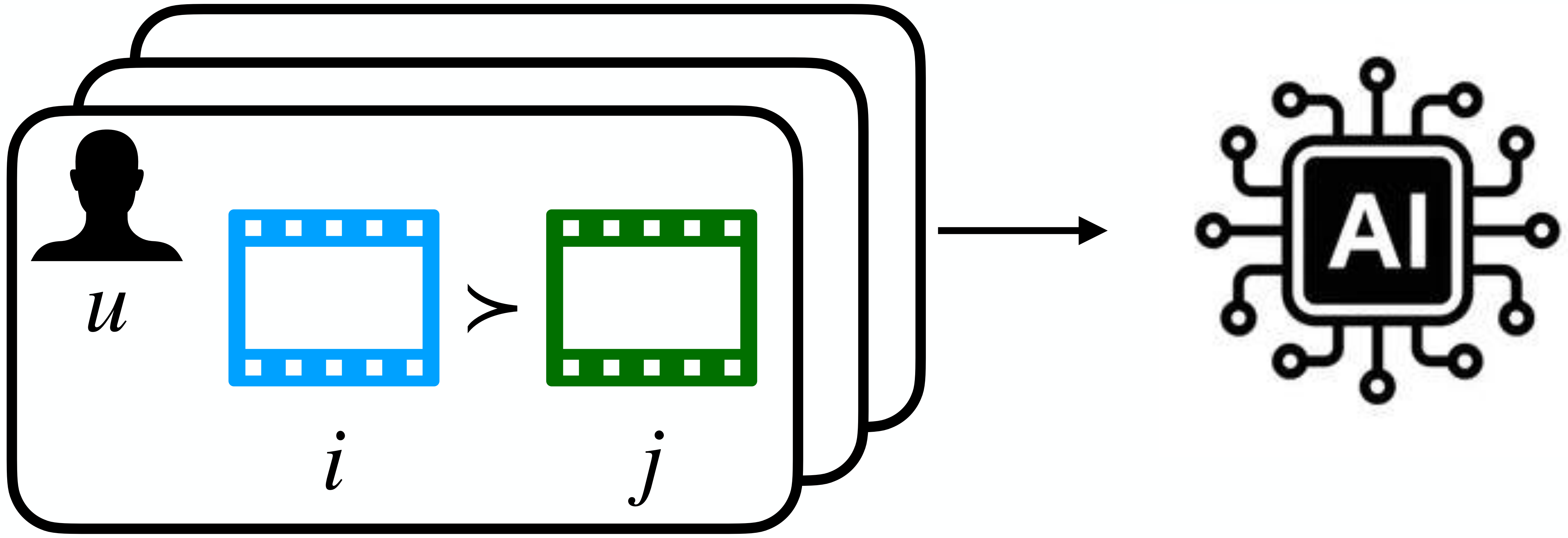
Matrix Bernstein inequality

3. **Conclusion:** $p^\top \bar{A} p \simeq p^\top \mathbb{E}[A] p \simeq \|p\|_2^2 = \|\Delta(Z)\|_F^2$

Journey So Far

Learning From Offline Comparisons is Efficient

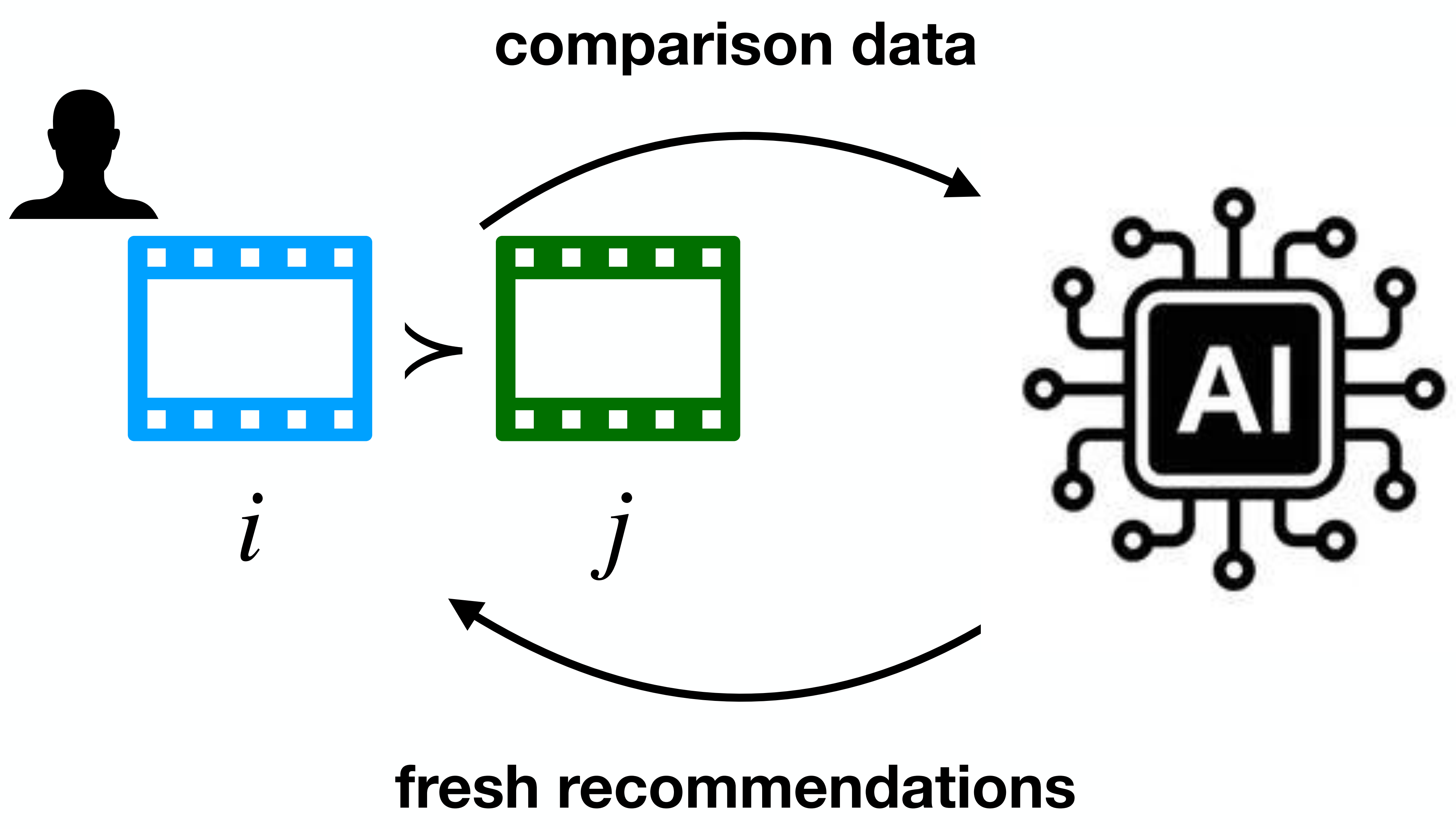
Fixed dataset (offline learning)



Journey So Far

Learning From Offline Comparisons is Efficient

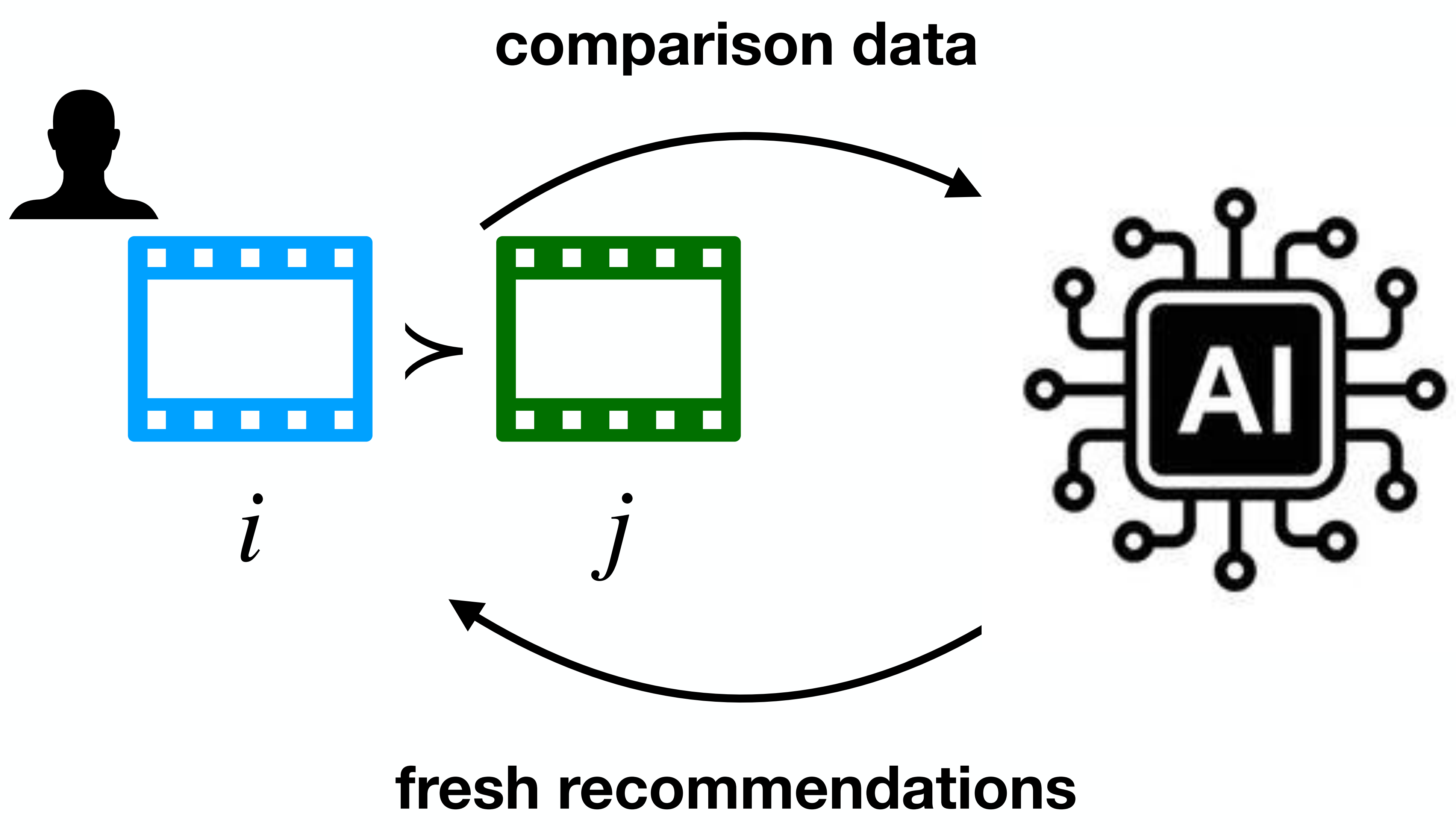
Interactive setting (online learning)



The Next Big Question

What Comparisons Should Be Sought?

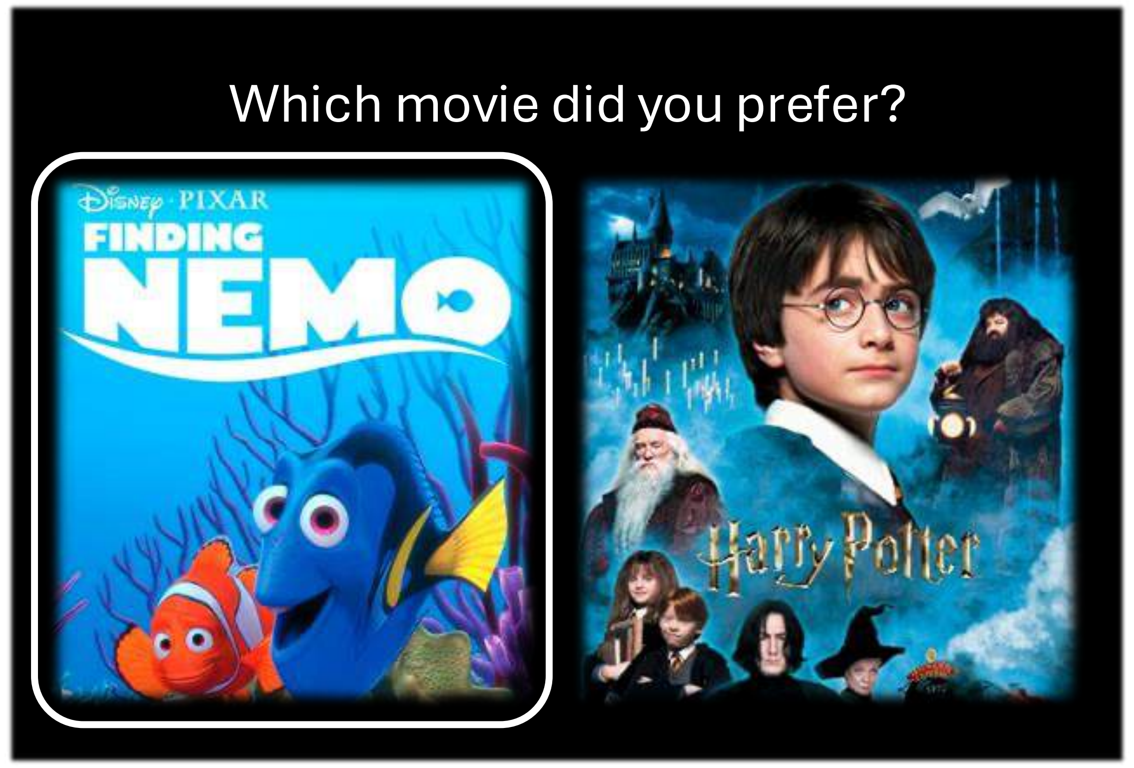
Interactive setting (online learning)



The Next Big Question

What Comparisons Should Be Sought?

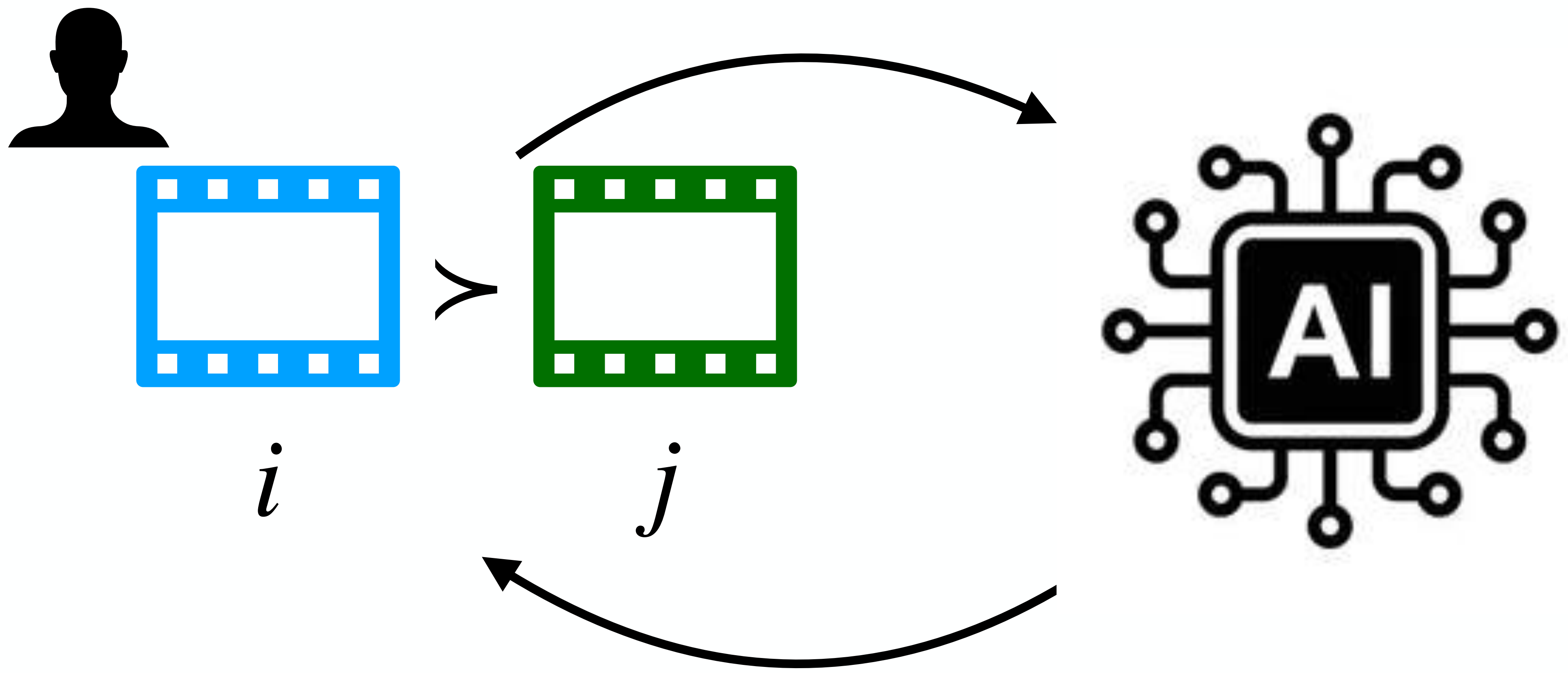
Interactive setting (online learning)



Watched just now | **Watched last week**

What to recommend? | *What to compare?*

comparison data



fresh recommendations

Comparing with consumption history effectively reuses past experiences

Comparing with History is Efficient

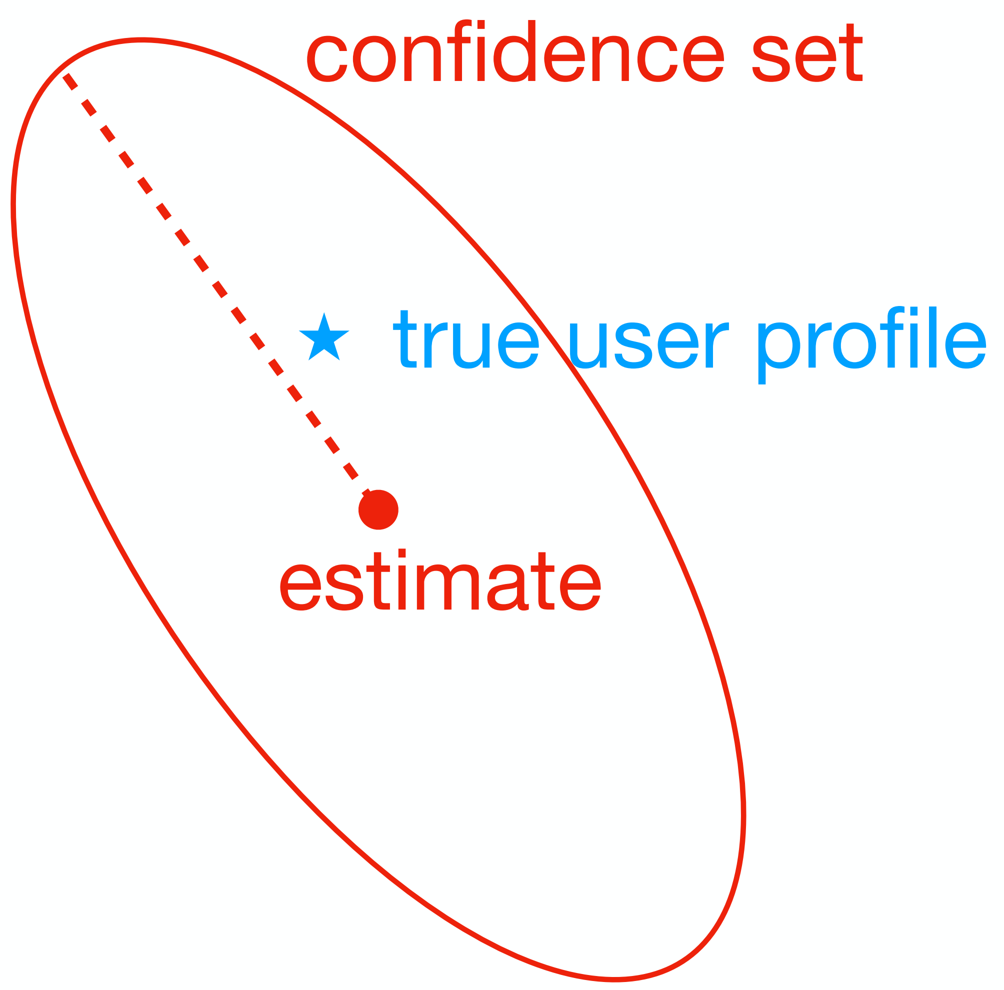
Recycling History: Efficient Recommendations From Contextual Duelling Bandits.

S. Sankagiri, P. Fatemi, J. Etesami, and M. Grossglauser (*Under Review*)

Comparing with History is Efficient

Recycling History: Efficient Recommendations From Contextual Duelling Bandits.

S. Sankagiri, P. Fatemi, J. Etesami, and M. Grossglauser (*Under Review*)

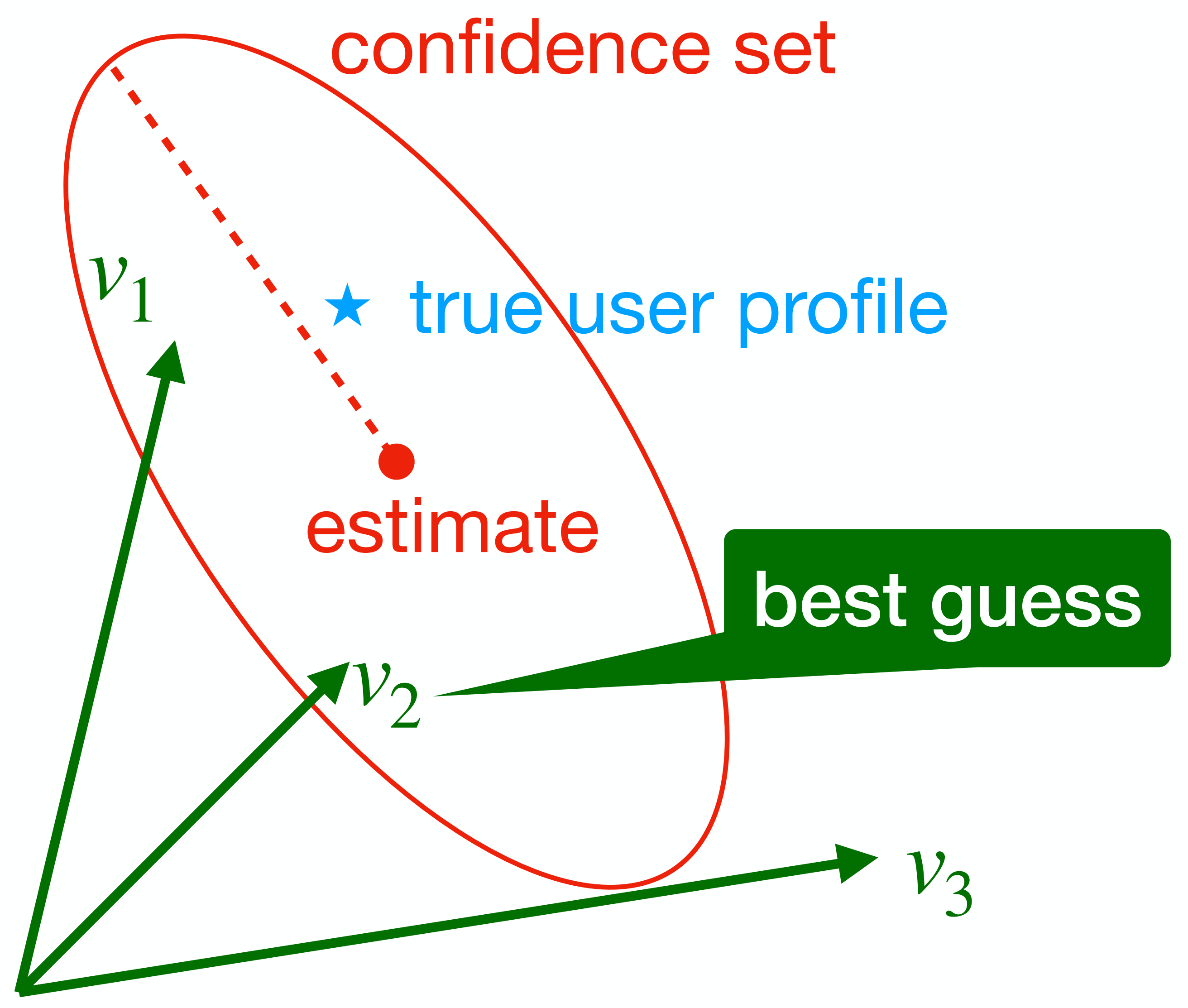


Comparing with History is Efficient

Recycling History: Efficient Recommendations From Contextual Duelling Bandits.

S. Sankagiri, P. Fatemi, J. Etesami, and M. Grossglauser (*Under Review*)

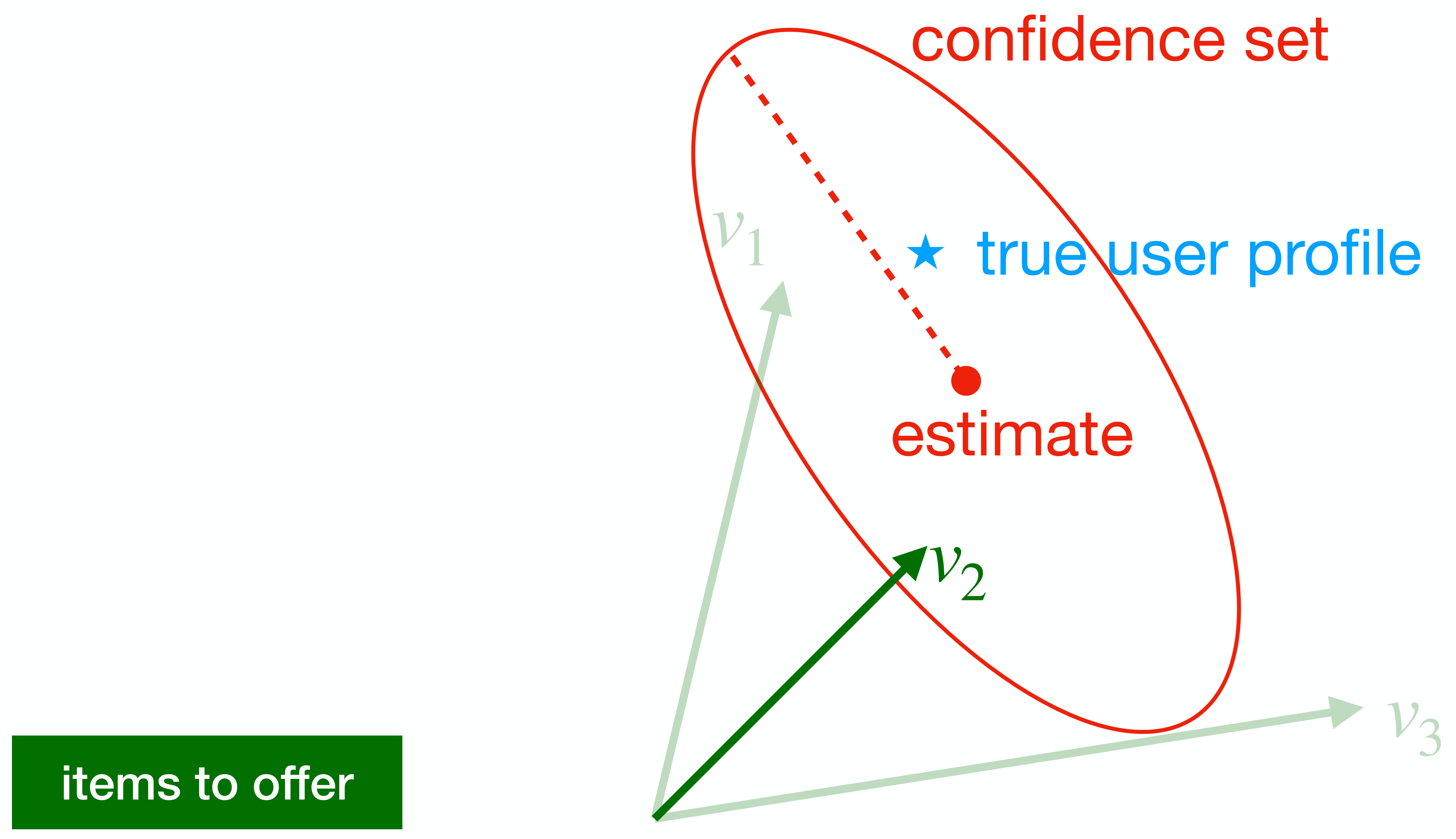
items to offer



Comparing with History is Efficient

Recycling History: Efficient Recommendations From Contextual Duelling Bandits.

S. Sankagiri, P. Fatemi, J. Etesami, and M. Grossglauser (*Under Review*)

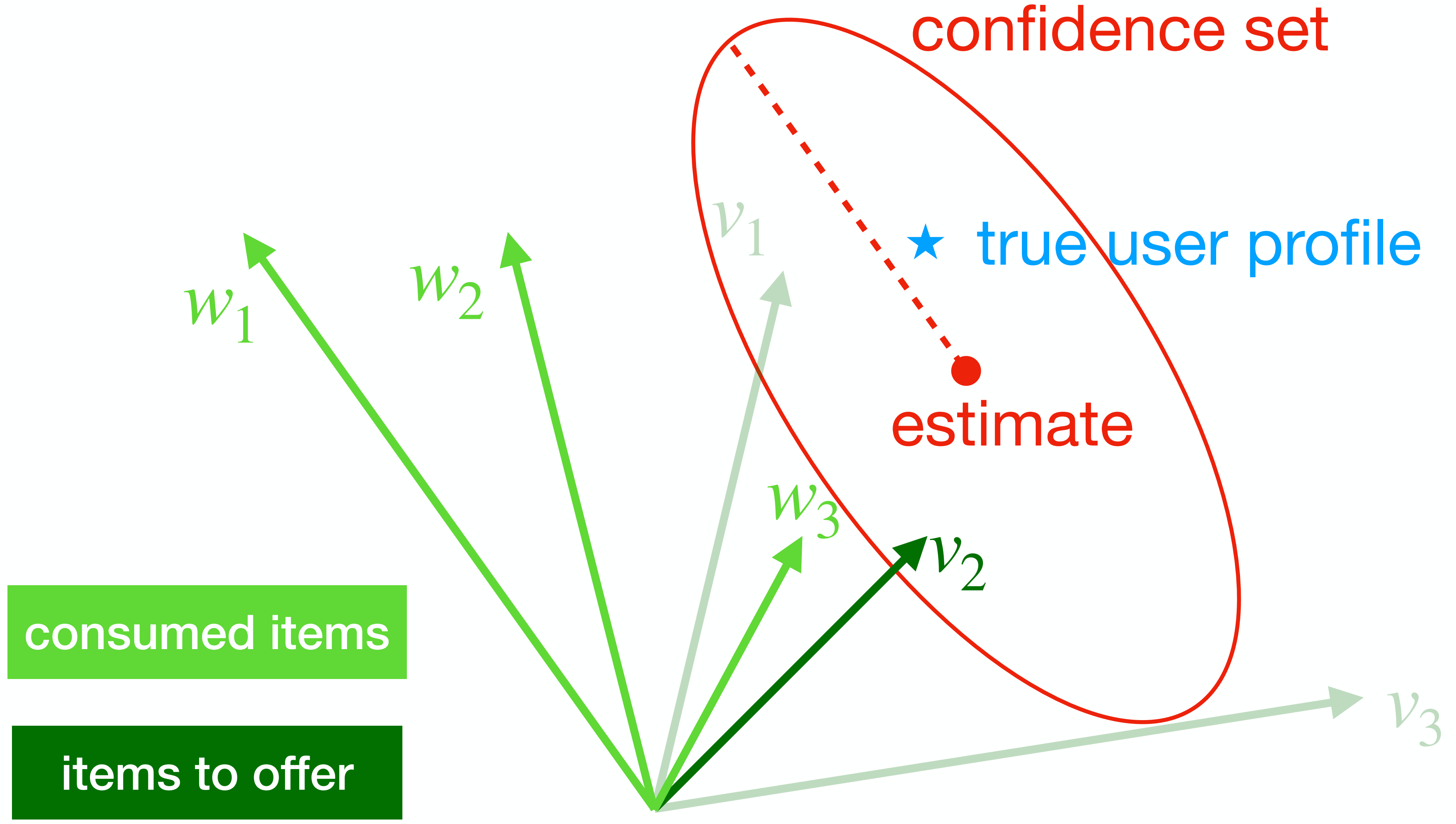


- Recommend the best guess item

Comparing with History is Efficient

Recycling History: Efficient Recommendations From Contextual Duelling Bandits.

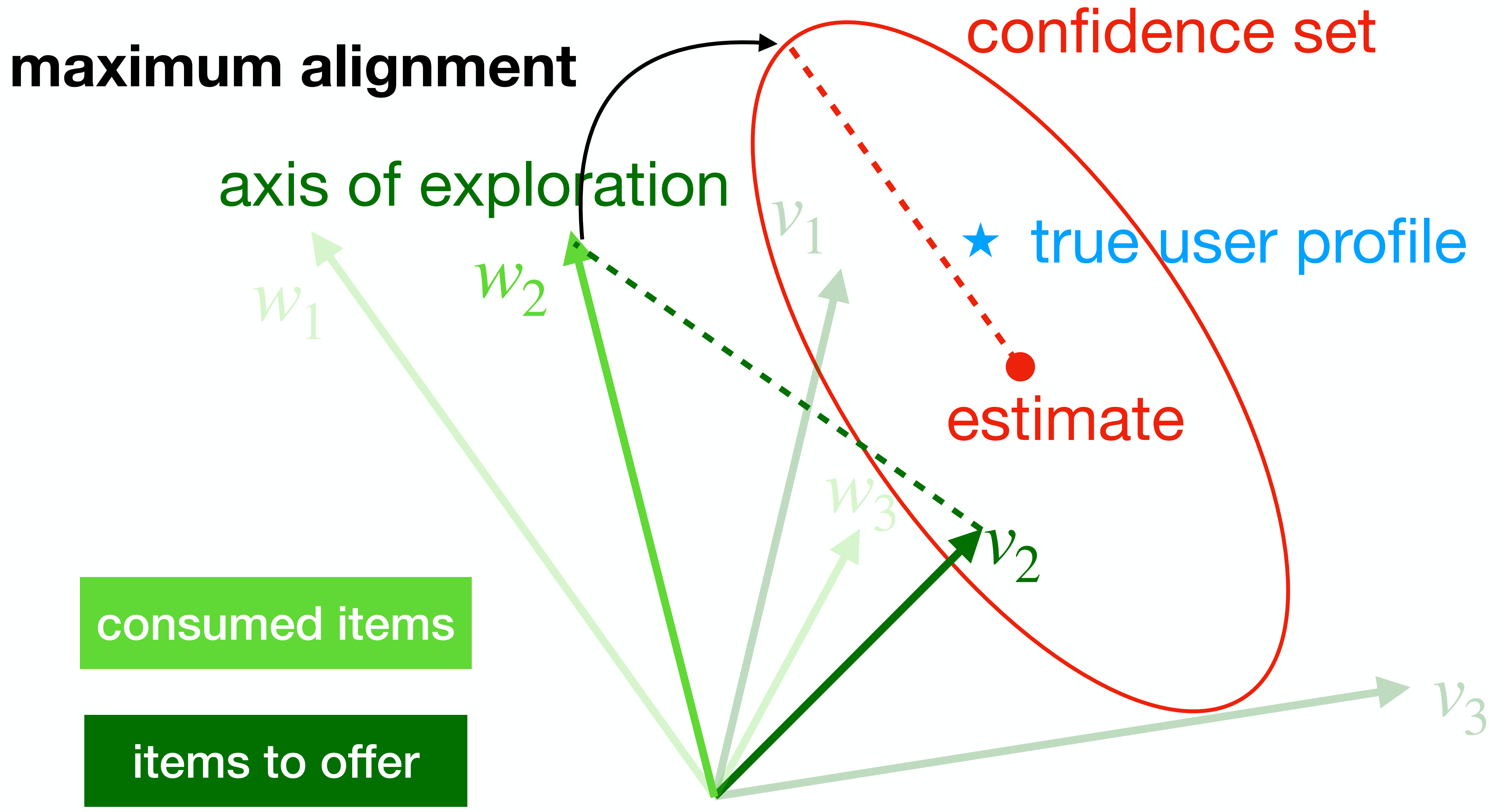
S. Sankagiri, P. Fatemi, J. Etesami, and M. Grossglauser (*Under Review*)



- Recommend the best guess item

Comparing with History is Efficient

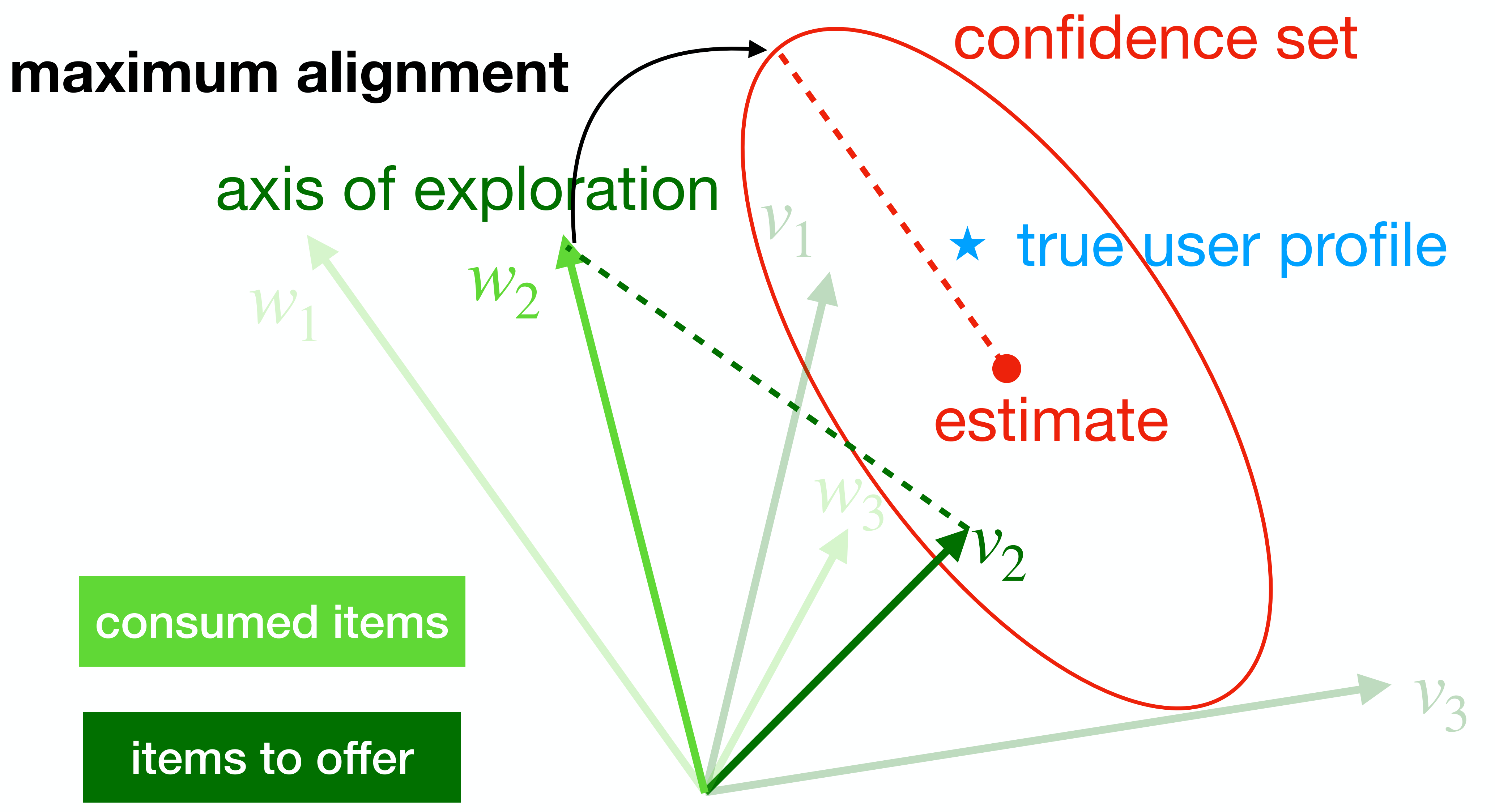
Recycling History: Efficient Recommendations From Contextual Duelling Bandits.
S. Sankagiri, P. Fatemi, J. Etesami, and M. Grossglauser (*Under Review*)



- Recommend the best guess item
- Compare with an item to probe along axis of uncertainty

Comparing with History is Efficient

Recycling History: Efficient Recommendations From Contextual Duelling Bandits.
S. Sankagiri, P. Fatemi, J. Etesami, and M. Grossglauser (*Under Review*)



- Recommend the best guess item
- Compare with an item to probe along axis of uncertainty
- **Mitigates exploration-exploitation tradeoff!**

Experimental Results

Experimental Results

Real Data Doesn't Obey BTL

Pick the older person

 $\succ$ 

Fading memory effects

Context effects

Incompatibility effects

Real Data Doesn't Obey BTL

Pick the older person

Fading memory effects

Context effects

Incompatibility effects

Real Data Doesn't Obey BTL

Pick the older person

Fading memory effects

Context effects

Incompatibility effects

Real Data Doesn’t Obey BTL

Pick the older person

<u>Observation</u>

Consistent if back-to-back

Inconsistencies if spaced out

Fading memory effects

Context effects

Incompatibility effects

Real Data Doesn't Obey BTL

Pick the older person

<u>Observation</u>

Consistent if back-to-back

Inconsistencies if spaced out

Fading memory effects

Pick the option most similar to the target

target

Context effects

Incompatibility effects

Real Data Doesn't Obey BTL

Pick the older person

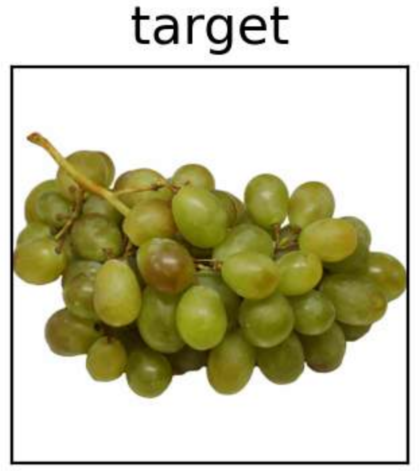
<u>Observation</u>

Consistent if back-to-back

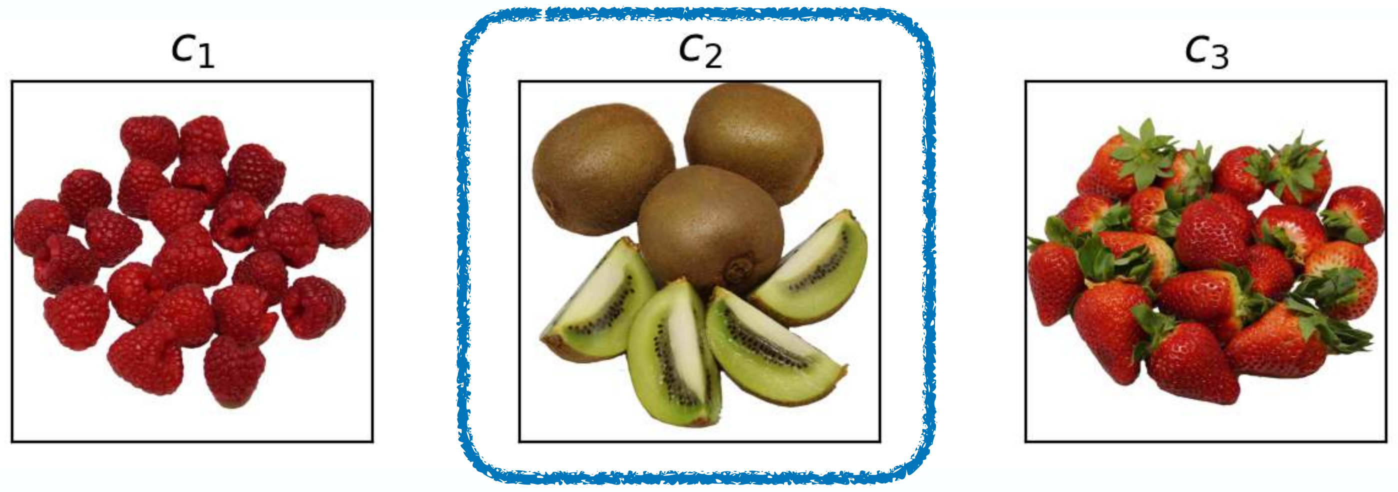
Inconsistencies if spaced out

Fading memory effects

Pick the option most similar to the target



Case 1



Context effects

Incompatibility effects

Real Data Doesn't Obey BTL

Pick the older person

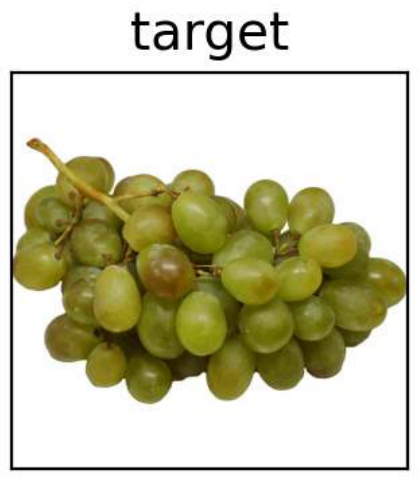
<u>Observation</u>

Consistent if back-to-back

Inconsistencies if spaced out

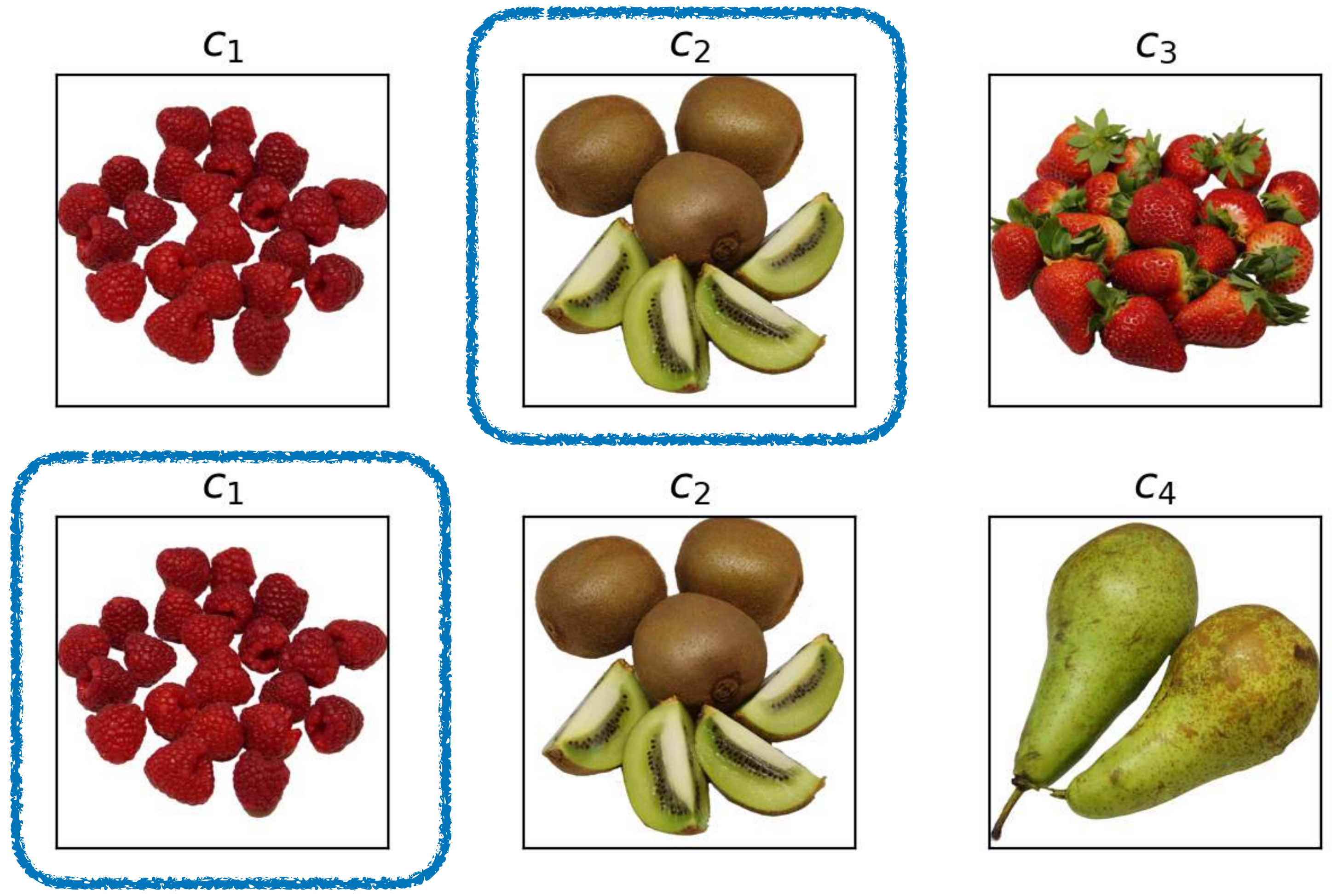
Fading memory effects

Pick the option most similar to the target



Case 1

Case 2



Context effects

Incompatibility effects

Real Data Doesn't Obey BTL

Pick the older person

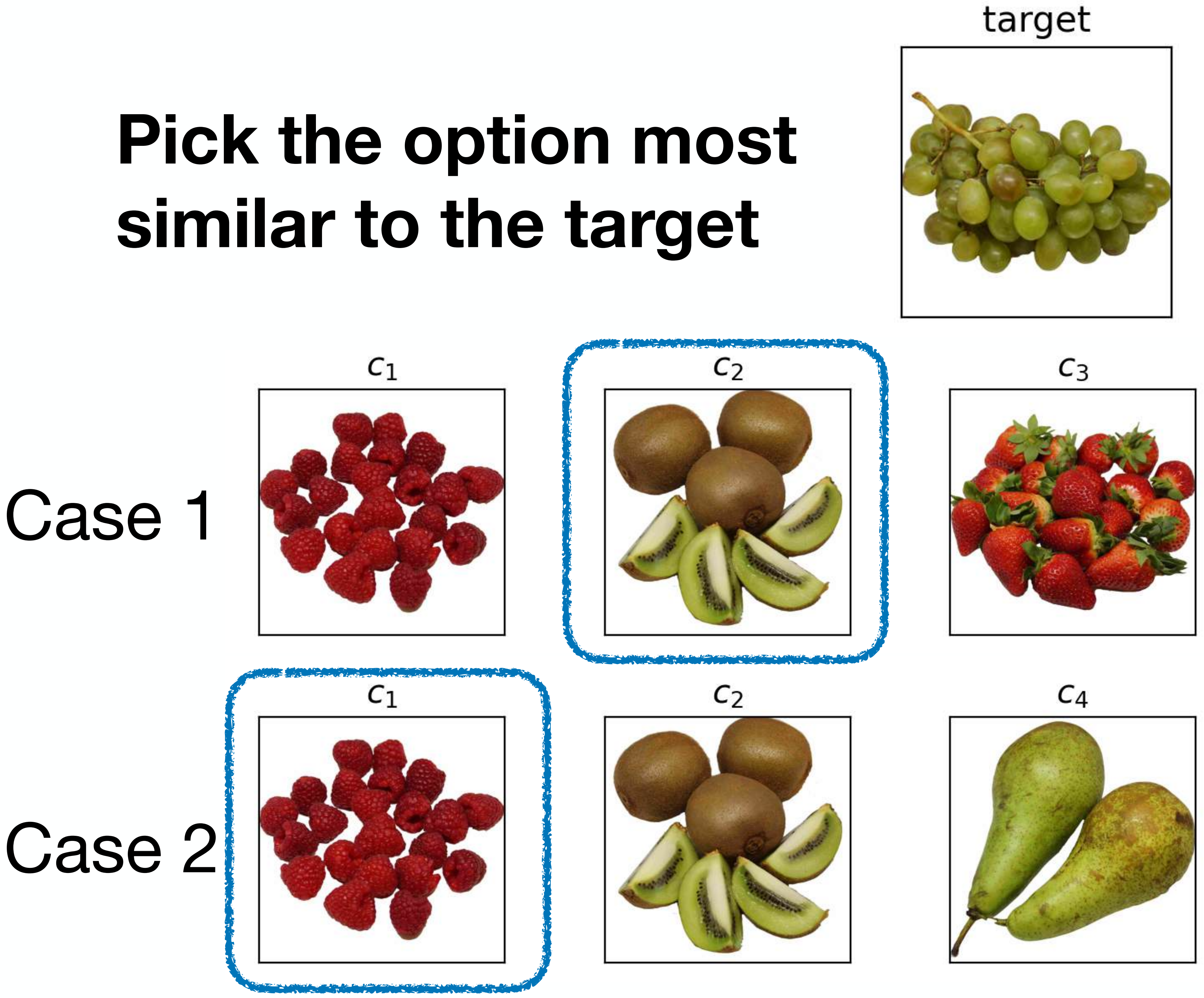
Observation

Consistent if back-to-back

Inconsistencies if spaced out

Fading memory effects

Pick the option most similar to the target



Observation

Statistically significant deviation from static similarity model

Context effects

Incompatibility effects

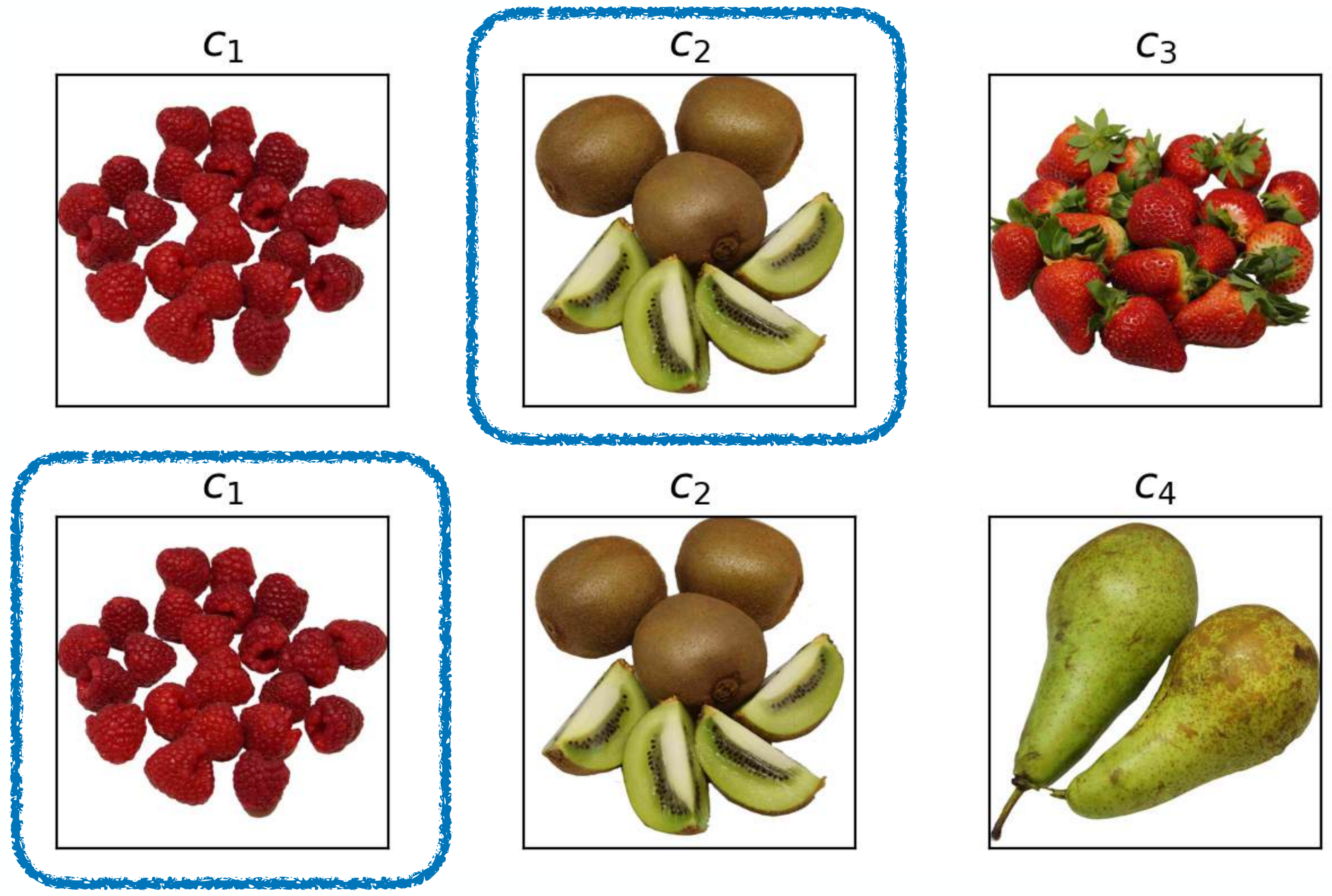
Real Data Doesn't Obey BTL

Pick the older person

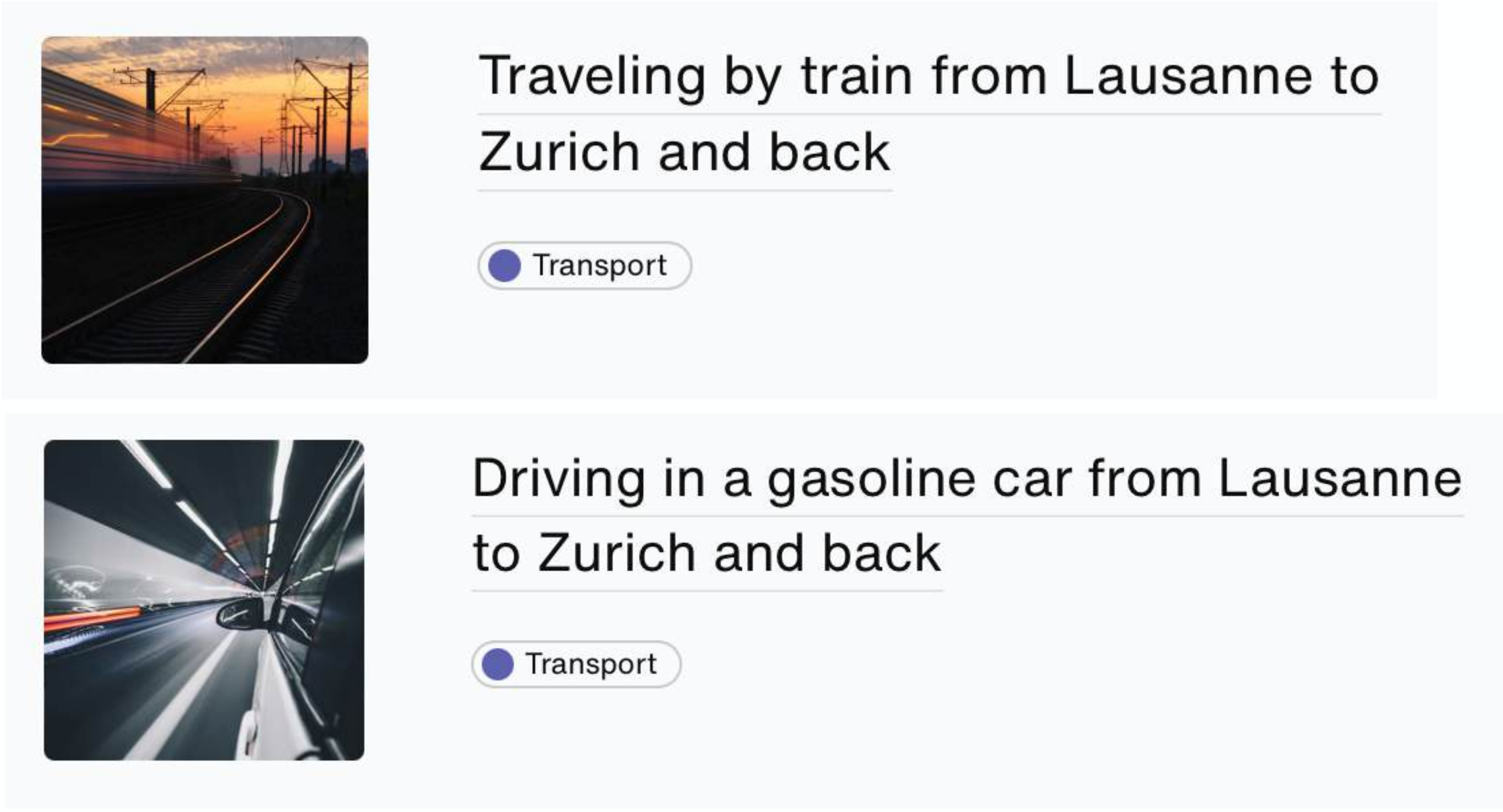
Pick the option most similar to the target

Case 1

Case 2



Pick action with larger CO_2



courtesy: climpact.ch

<u>Observation</u>

Consistent if back-to-back

Inconsistencies if spaced out

Fading memory effects

<u>Observation</u>

Statistically significant deviation from static similarity model

Context effects

Incompatibility effects

Real Data Doesn't Obey BTL

Pick the older person

Observation

Consistent if back-to-back

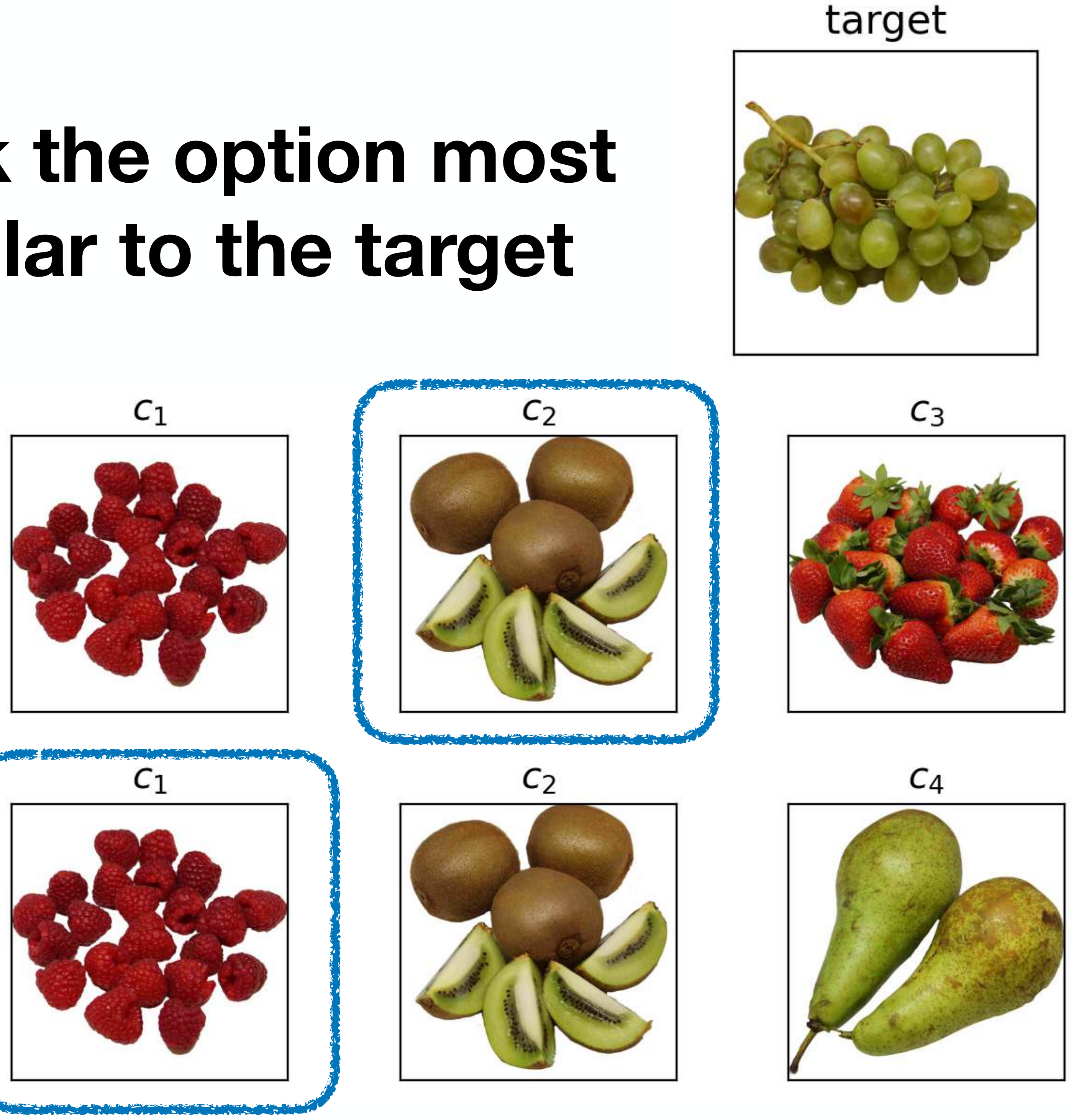
Inconsistencies if spaced out

Fading memory effects

Pick the option most similar to the target

Case 1

Case 2



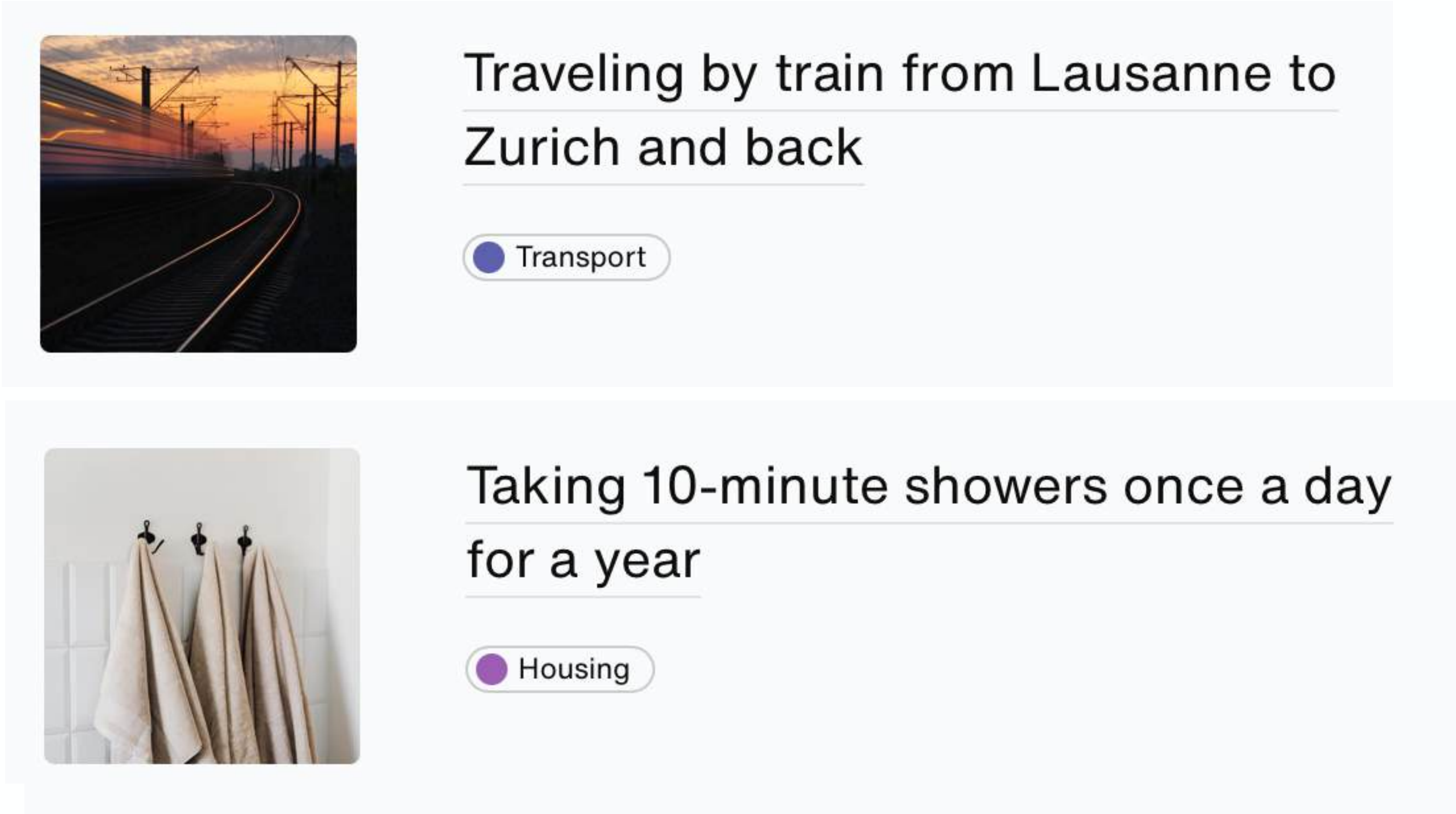
Observation

Statistically significant deviation

from static similarity model

Context effects

Pick action with larger CO_2



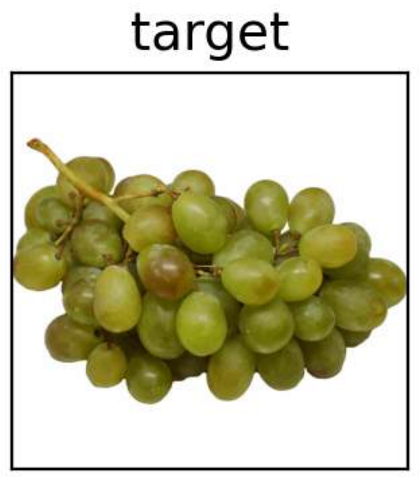
courtesy: climpact.ch

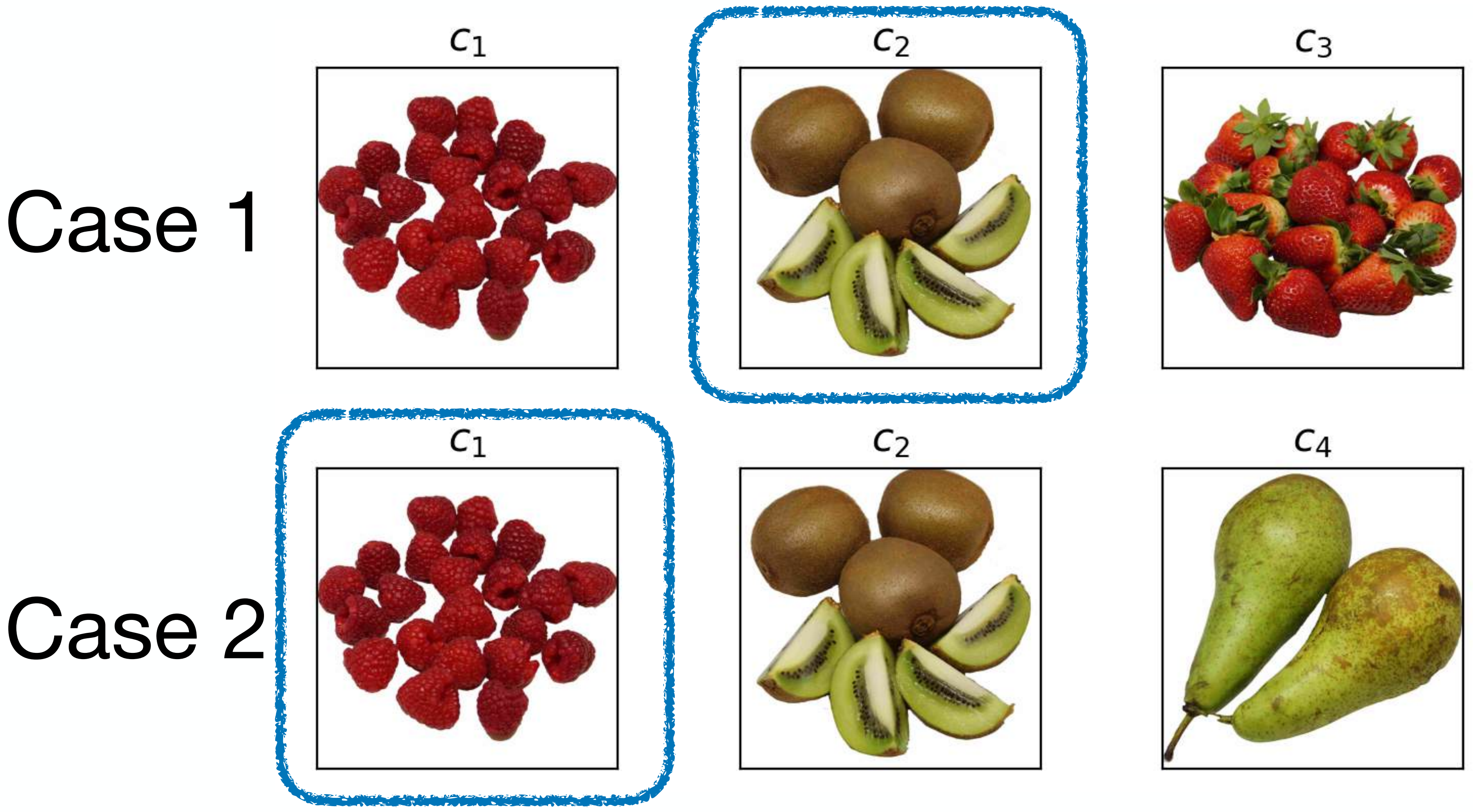
Incompatibility effects

Real Data Doesn't Obey BTL

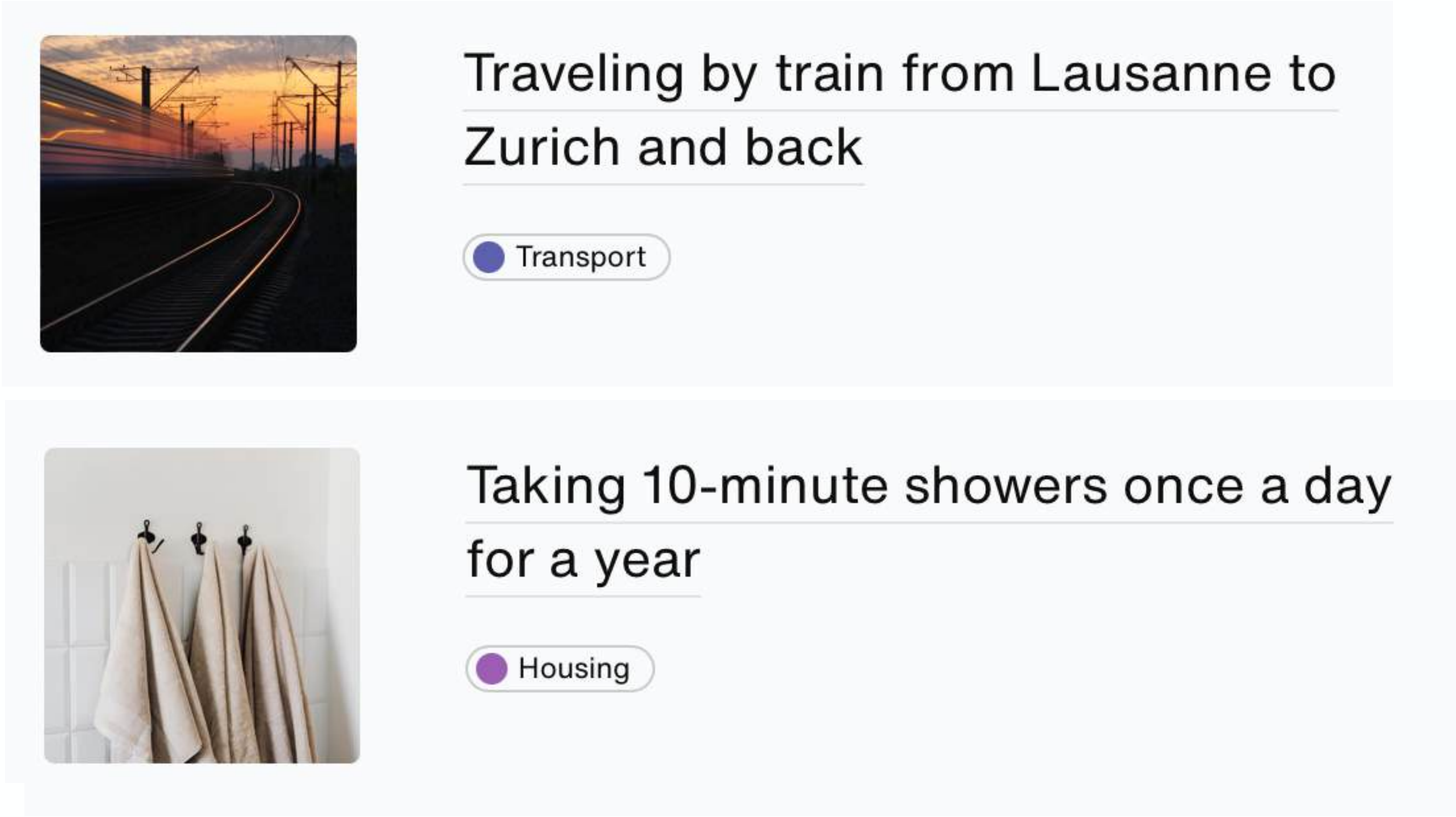
Pick the older person

Pick the option most similar to the target





Pick action with larger CO_2



courtesy: climpact.ch

<u>Observation</u>

Consistent if back-to-back

Inconsistencies if spaced out

Fading memory effects

<u>Observation</u>

Statistically significant deviation

from static similarity model

Context effects

<u>Observation</u>

Variance in user response larger

For apples-to-oranges comparisons

Incompatibility effects

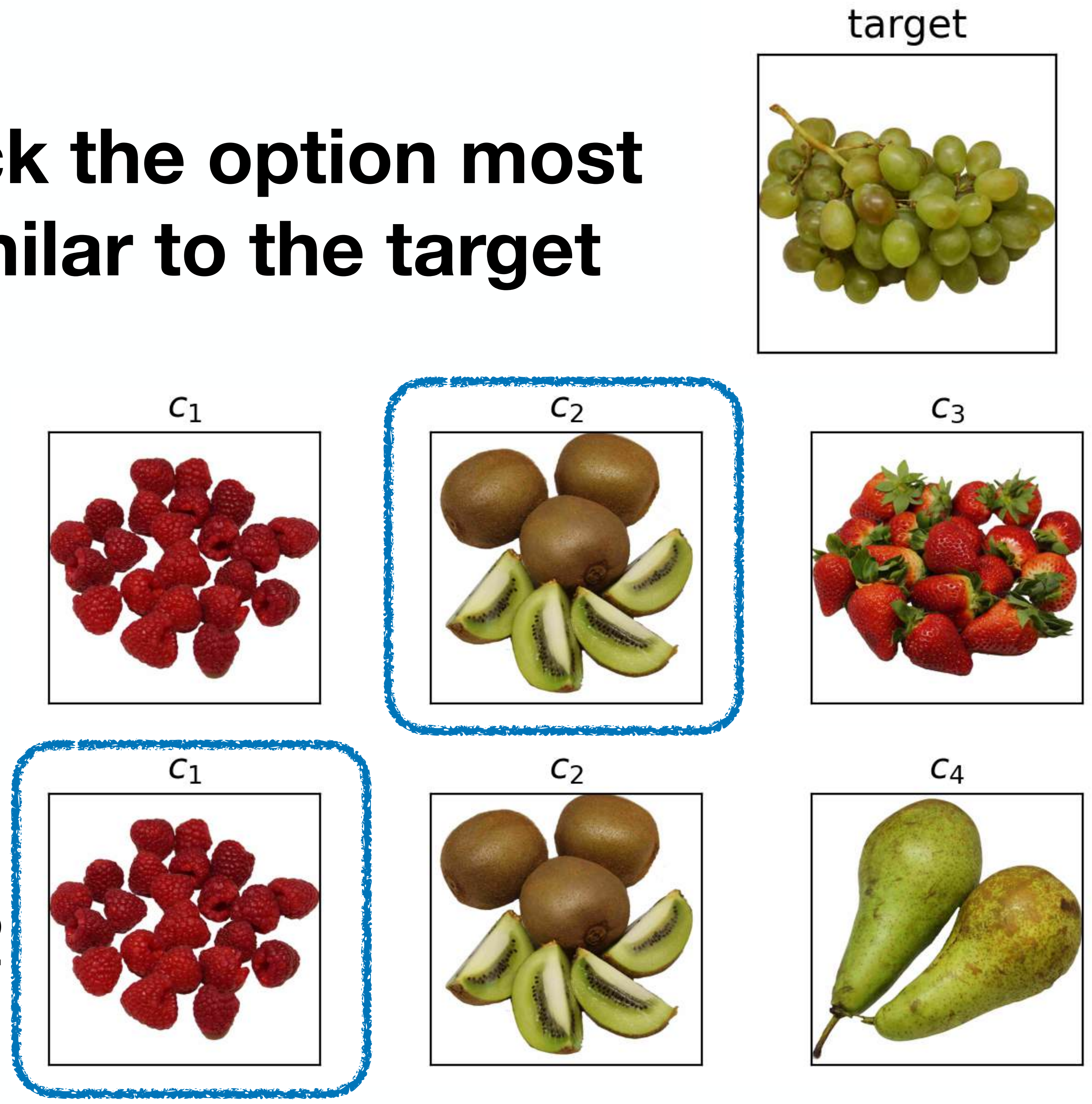
Real Data Doesn't Obey BTL

Pick the older person

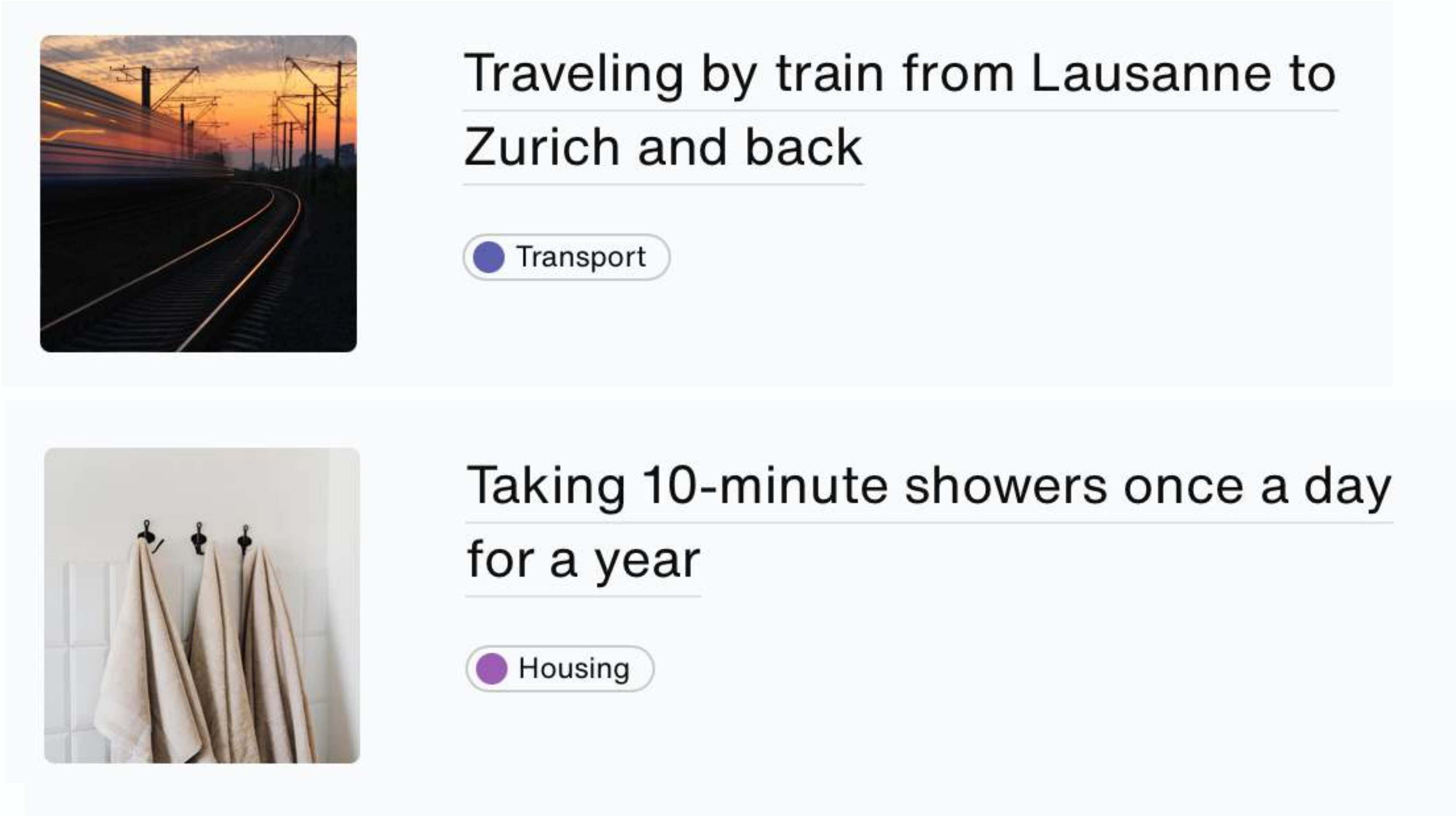
Pick the option most similar to the target

Case 1

Case 2



Pick action with larger CO2



courtesy: climpact.ch

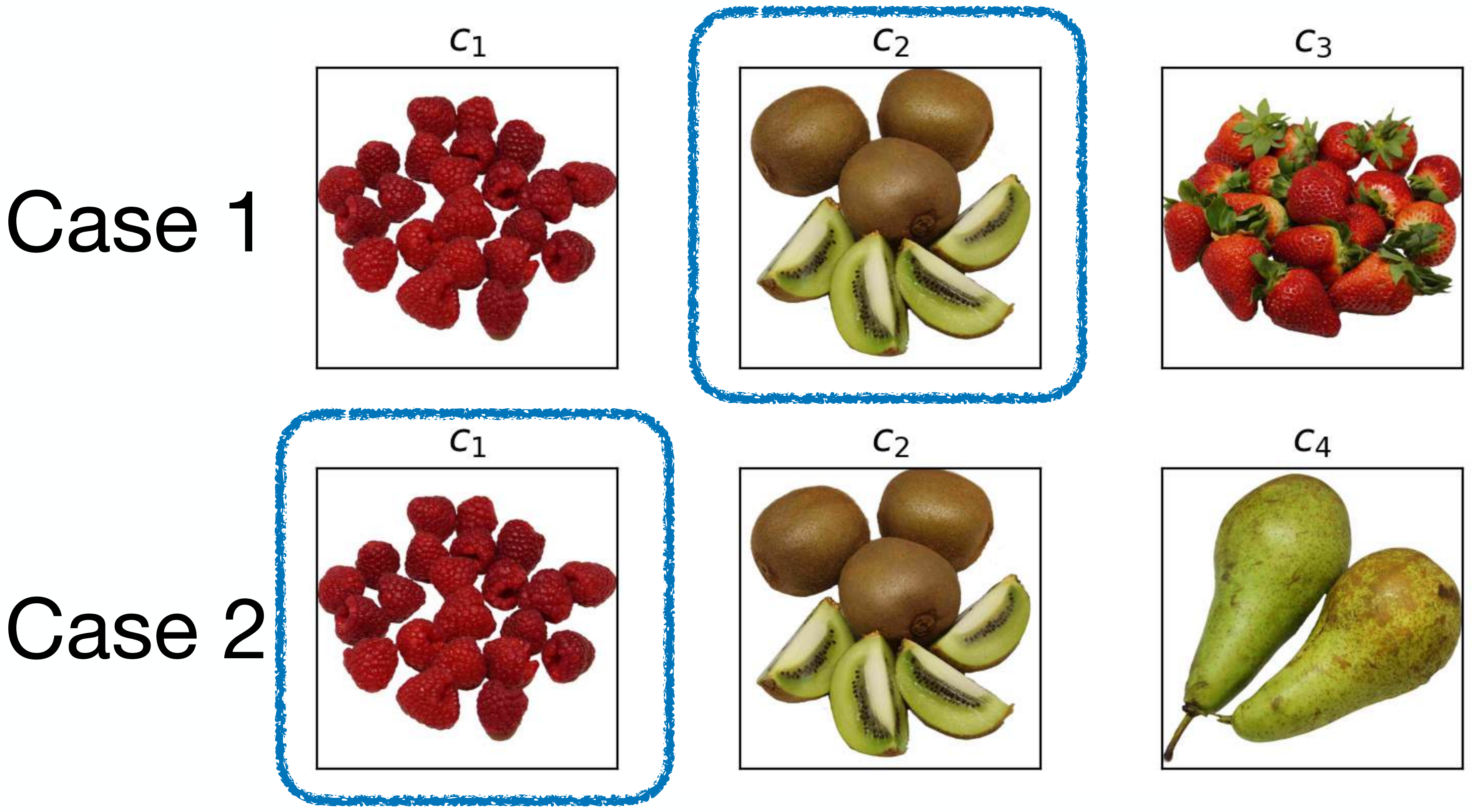
Research Question 1: What's the right <u>model</u> to capture these effects?

Fading memory effects | **Context effects** | **Incompatibility effects**

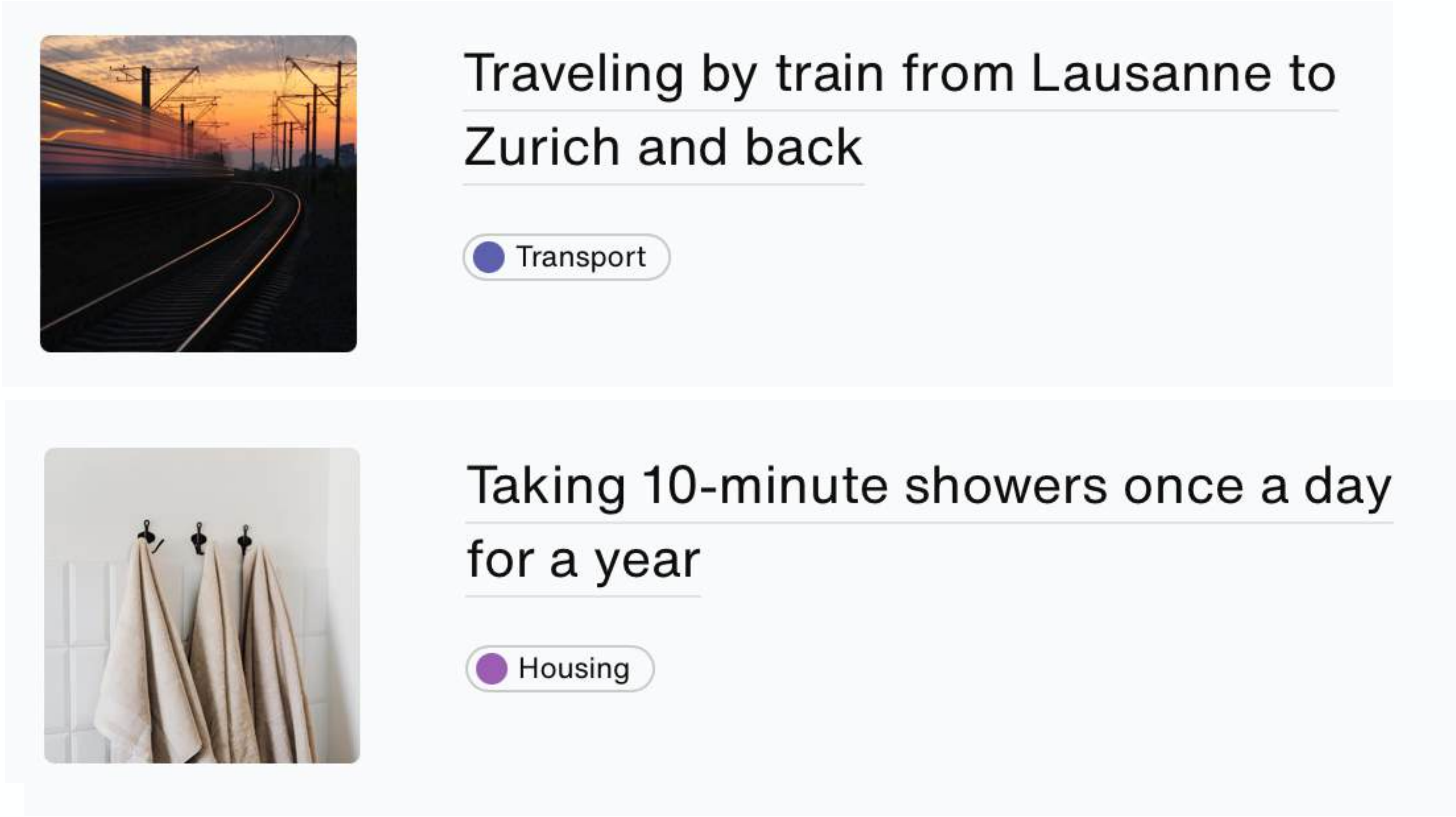
Real Data Doesn't Obey BTL

Pick the older person

Pick the option most similar to the target



Pick action with larger CO_2



courtesy: climpact.ch

Research Question 1: What's the right <u>model</u> to capture these effects?

Research Question 2: Do new models imply different <u>algorithms</u> for inference?

Fading memory effects | **Context effects** | **Incompatibility effects**

My Research Contributions

- *Recommendations with Sparse Comparison Data: Nonconvex Matrix Factorisation*.
  S. Sankagiri, J. Etesami, and M. Grossglauser (*ICML 2025*)

- *Recycling History: Efficient Recommendations From Contextual Duelling Bandits.*
  S. Sankagiri, P. Fatemi, J. Etesami, and M. Grossglauser (*Under Review*)

- *Measuring IIA Violations in Similarity Choices with Bayesian Models*.
  H. Correa, **S. Sankagiri**, D. Figueiredo, and M. Grossglauser (*UAI 2025*)

- *Ranking Items from Discrete Ratings: The Cost of Unknown User Thresholds.*
  O. Villemaud, **S. Sankagiri**, and M. Grossglauser (*Under Review*)

M. Grossglauser
EPFL, Switzerland

D. Figueiredo
UFRJ, Brazil

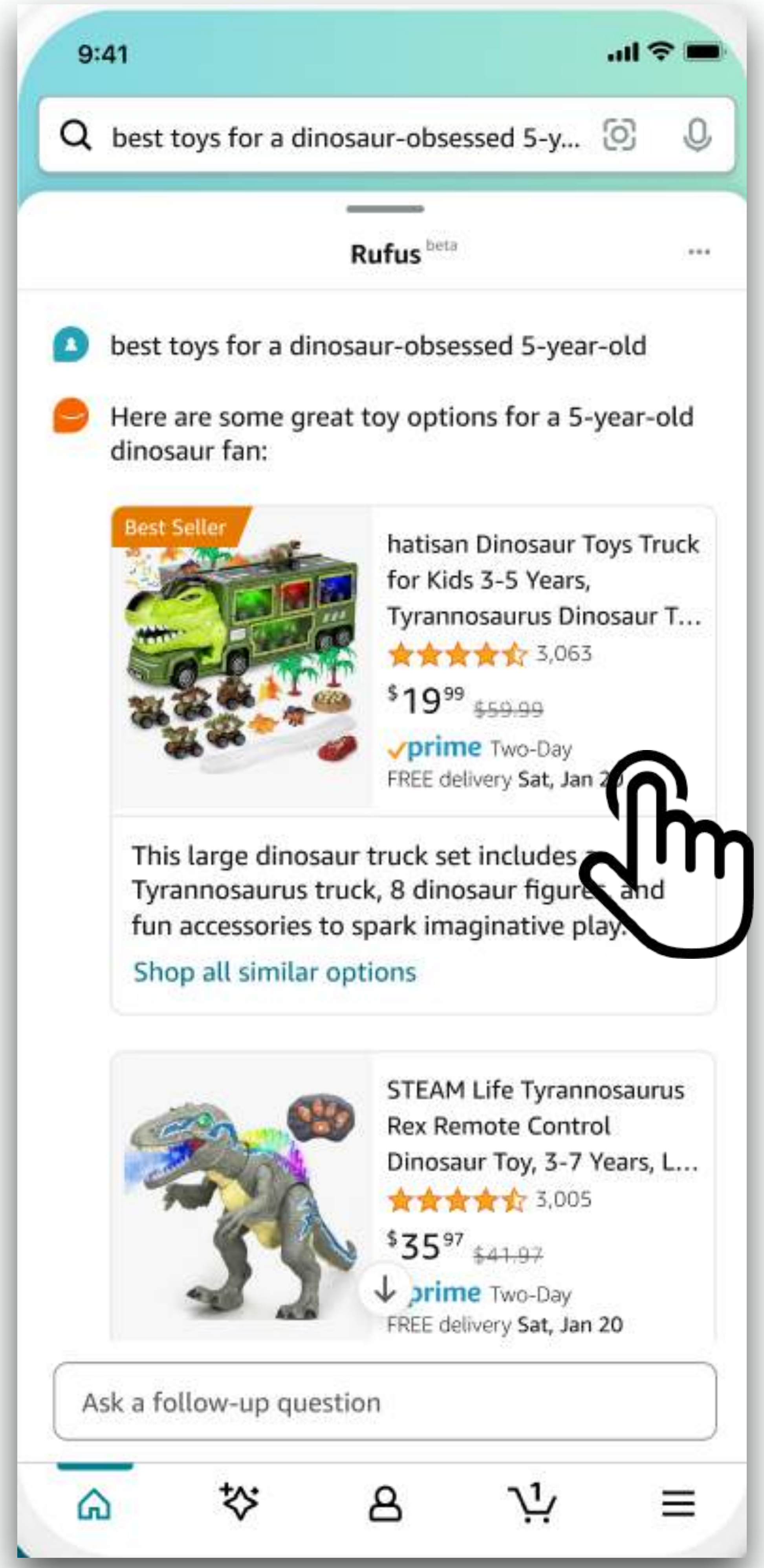
J. Etesami
TUM, Germany

O. Villemaud
EPFL, Switzerland

H. Correa
UFRJ, Brazil

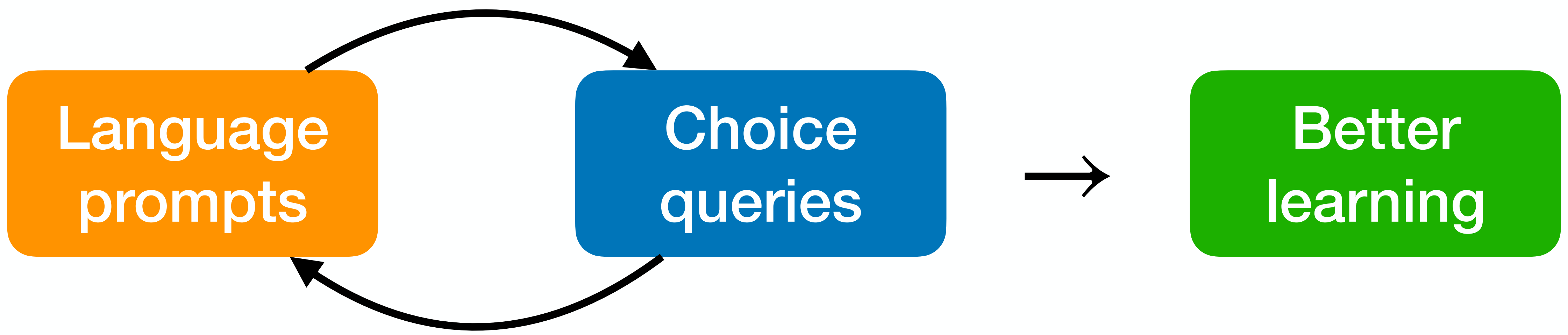
P. Fatemi
TUM, Germany

Vision: From Clicks to Conversations

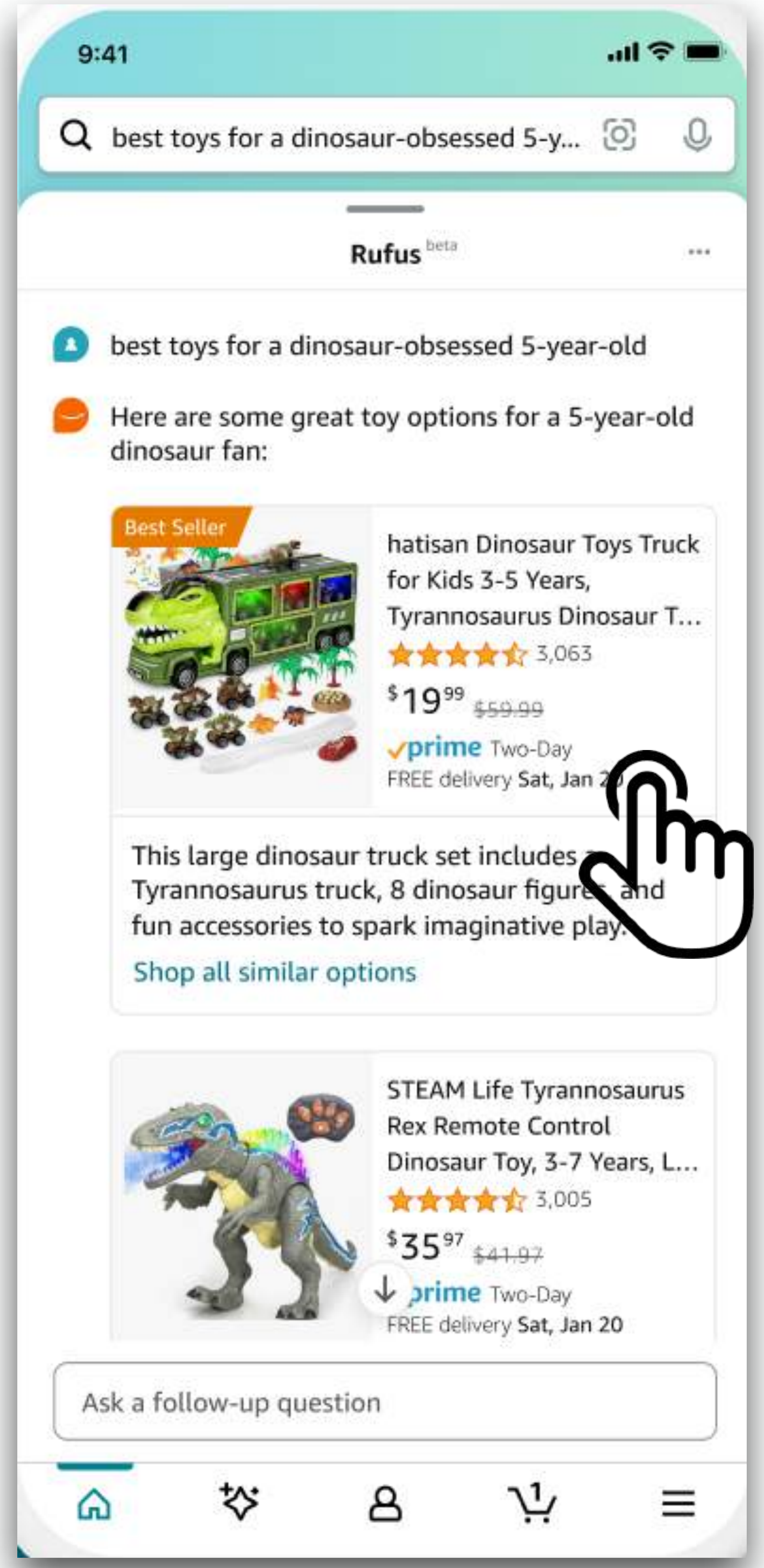


AI chatbots understand natural language

Users can reveal preferences through choices, not just words

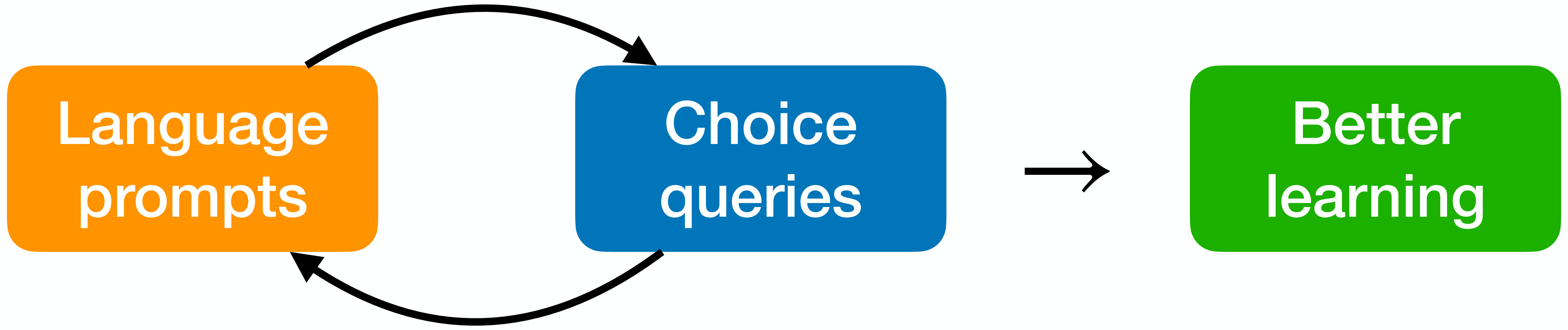


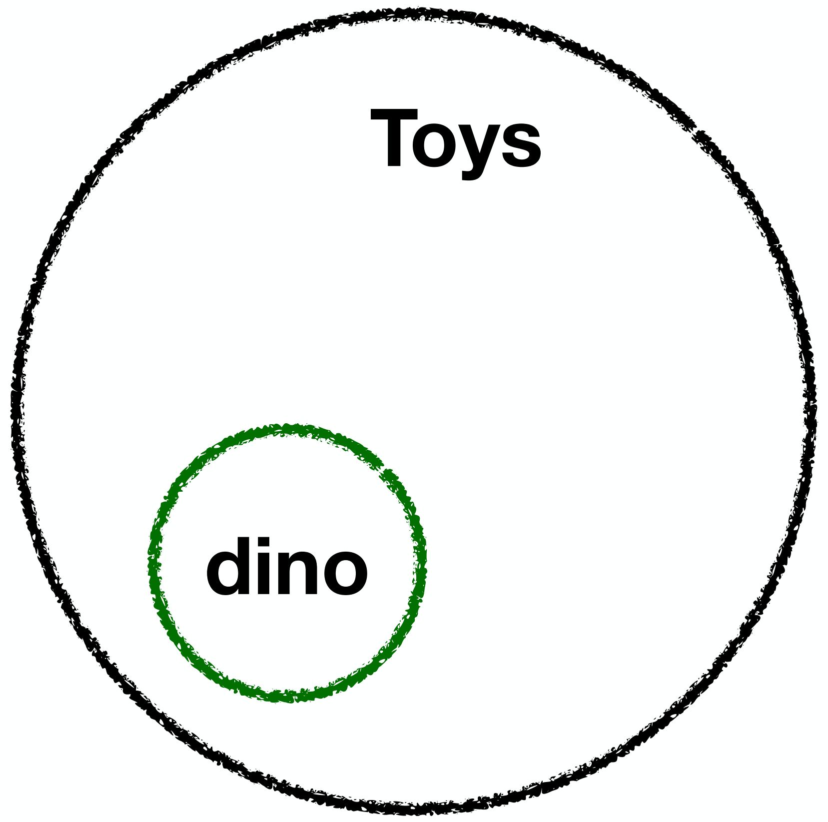
Vision: From Clicks to Conversations



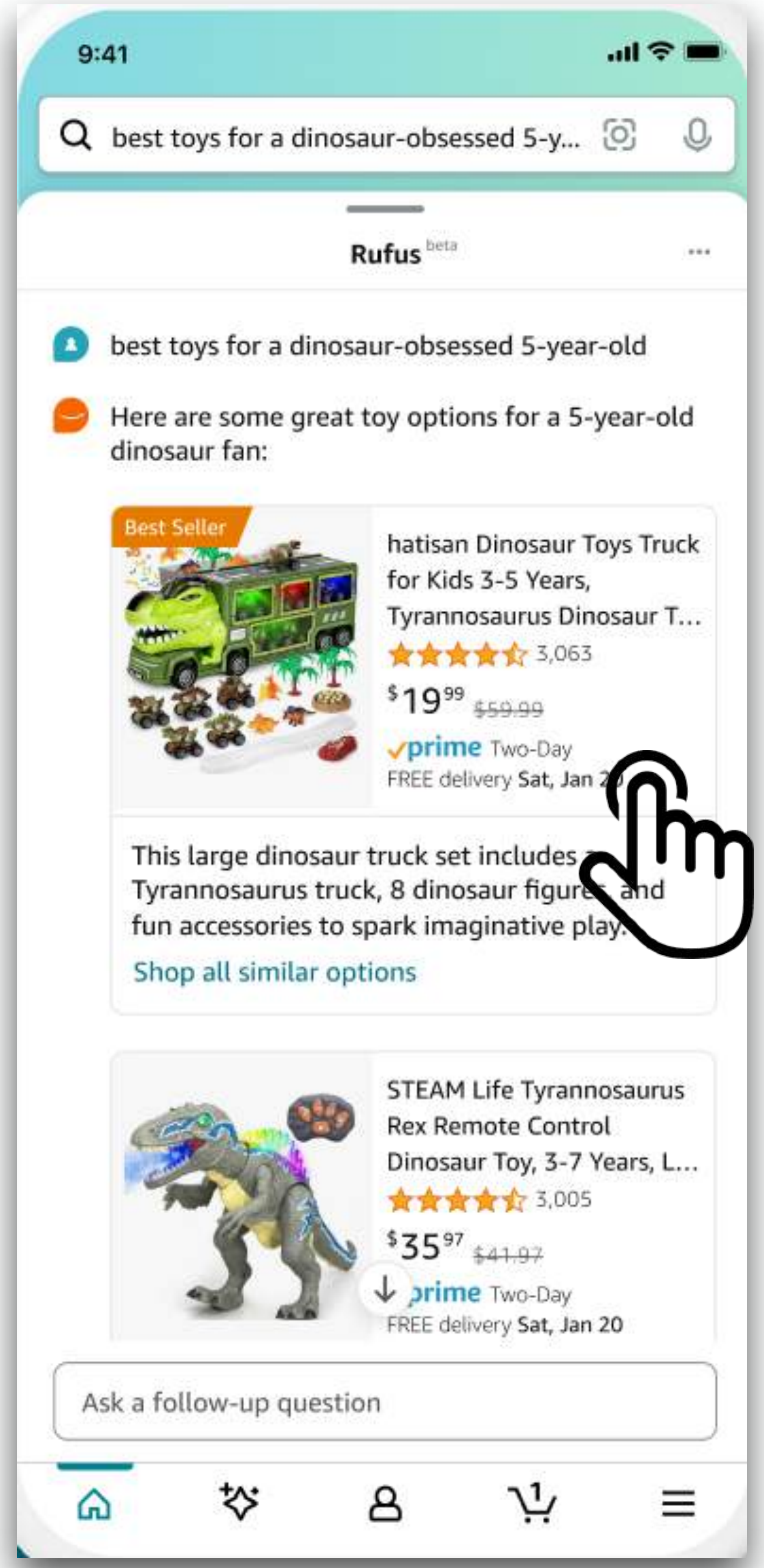
AI chatbots understand natural language

Users can reveal preferences through choices, not just words



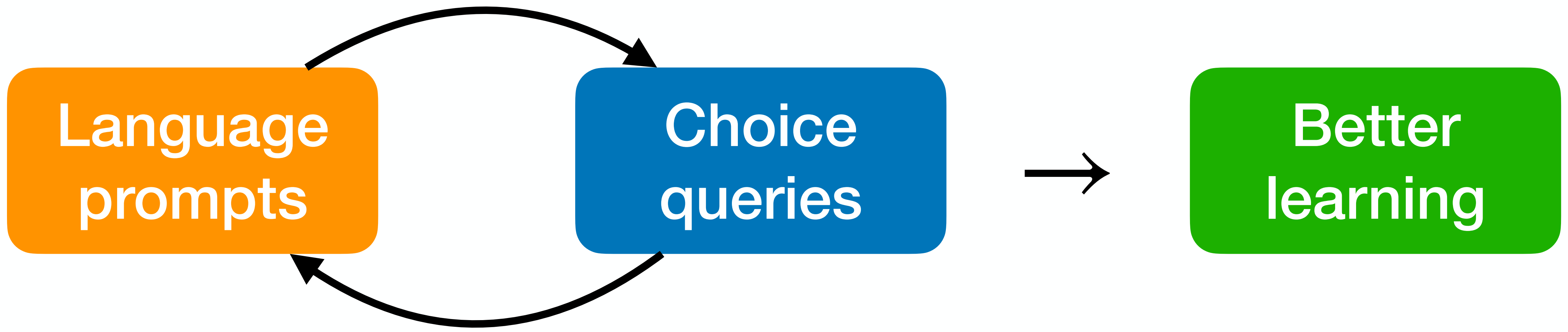


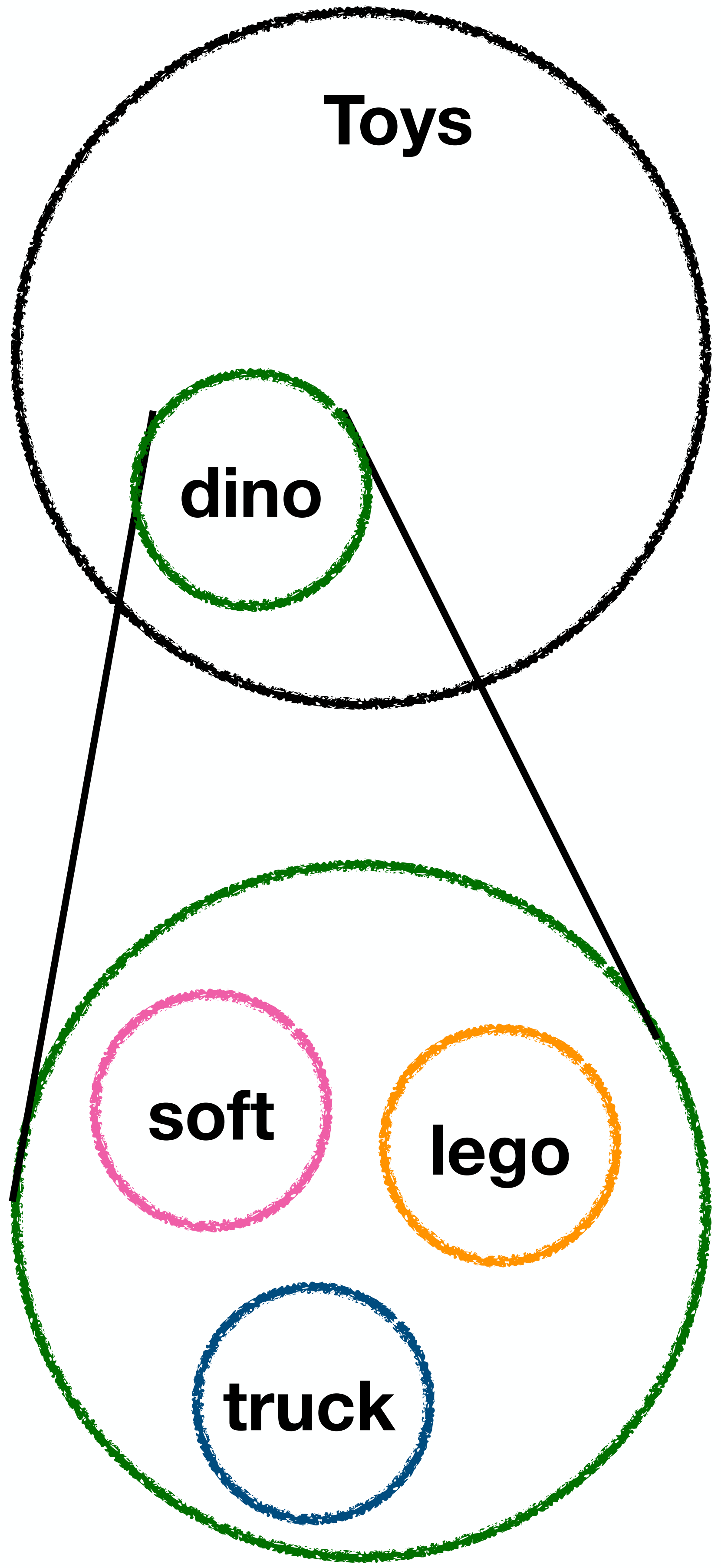
Vision: From Clicks to Conversations



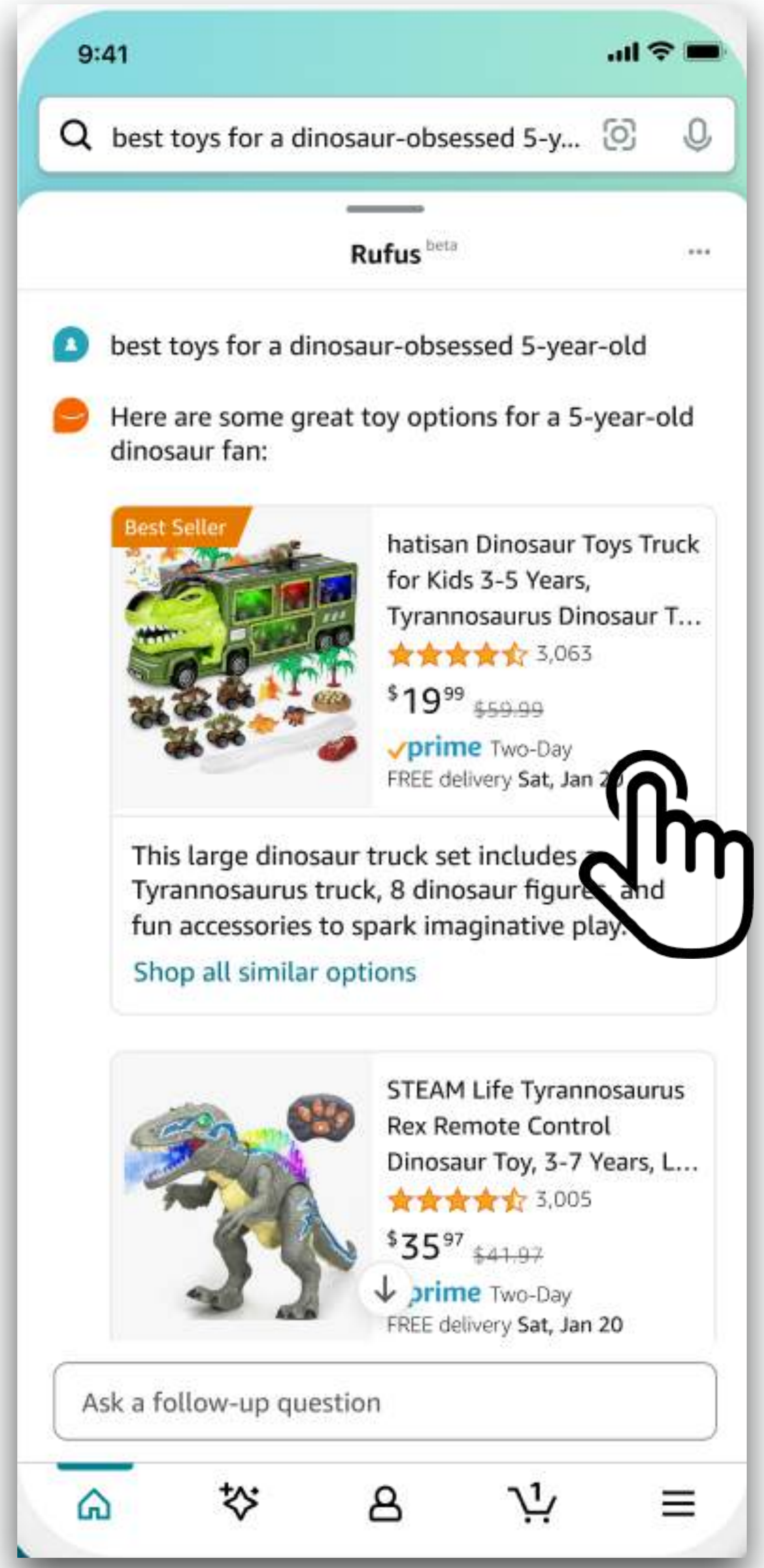
AI chatbots understand natural language

Users can reveal preferences through choices, not just words



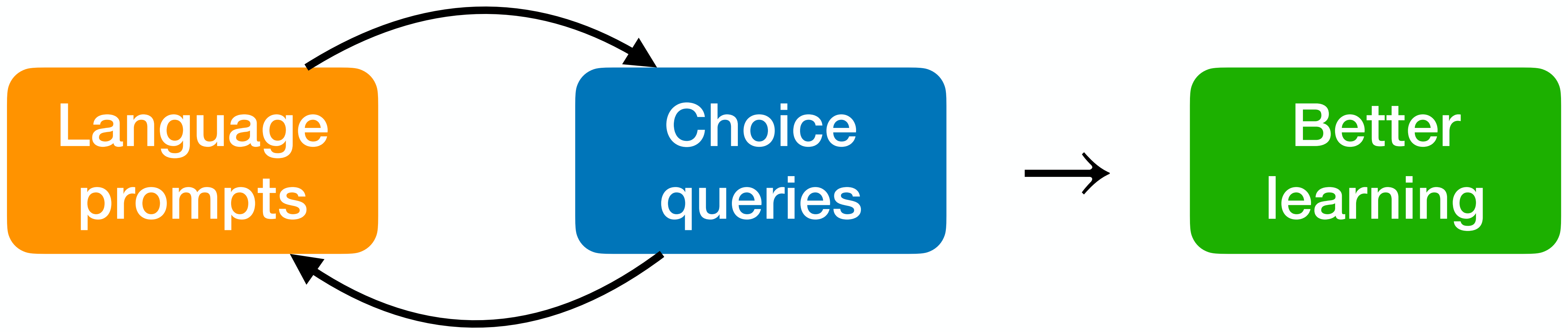


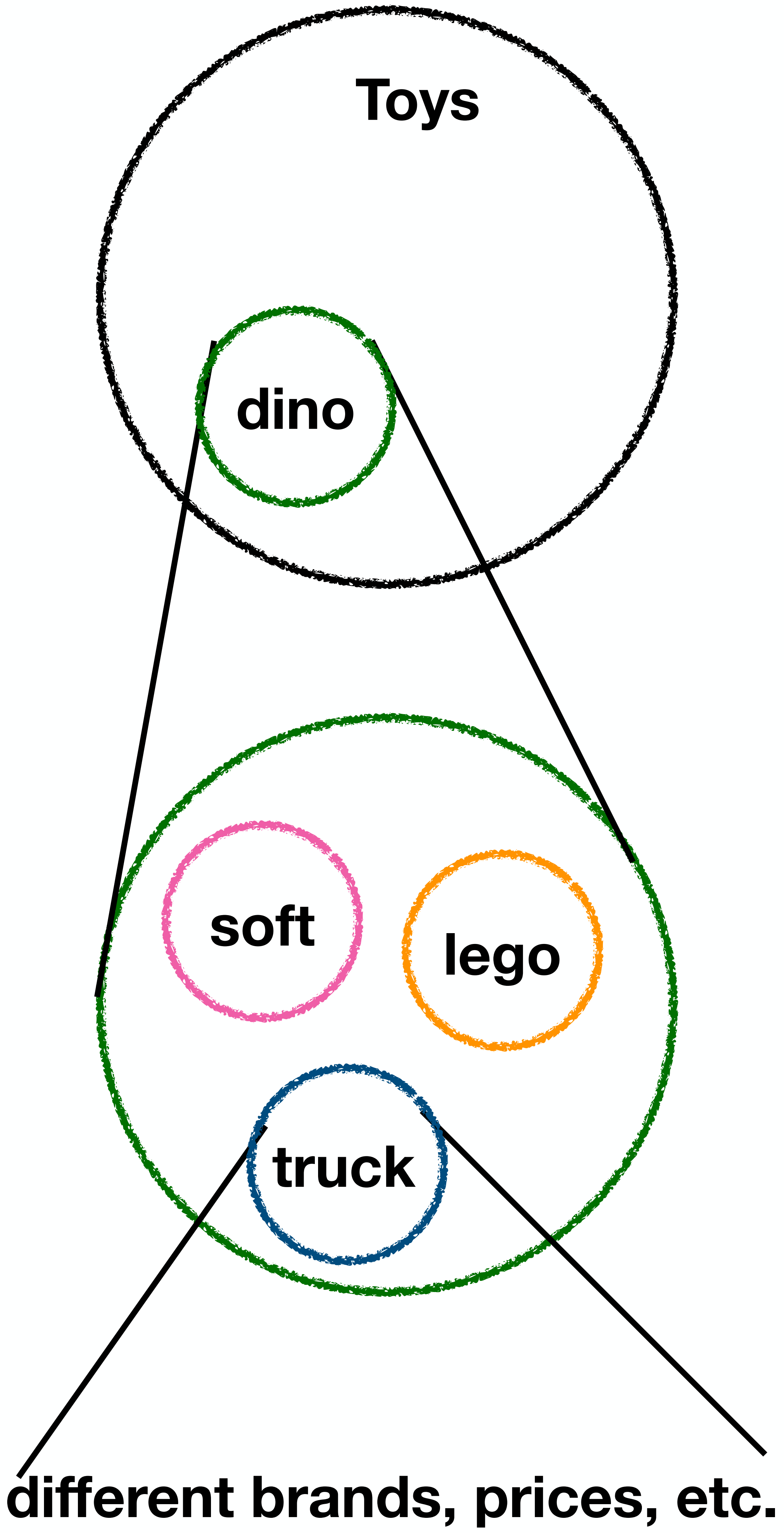
Vision: From Clicks to Conversations



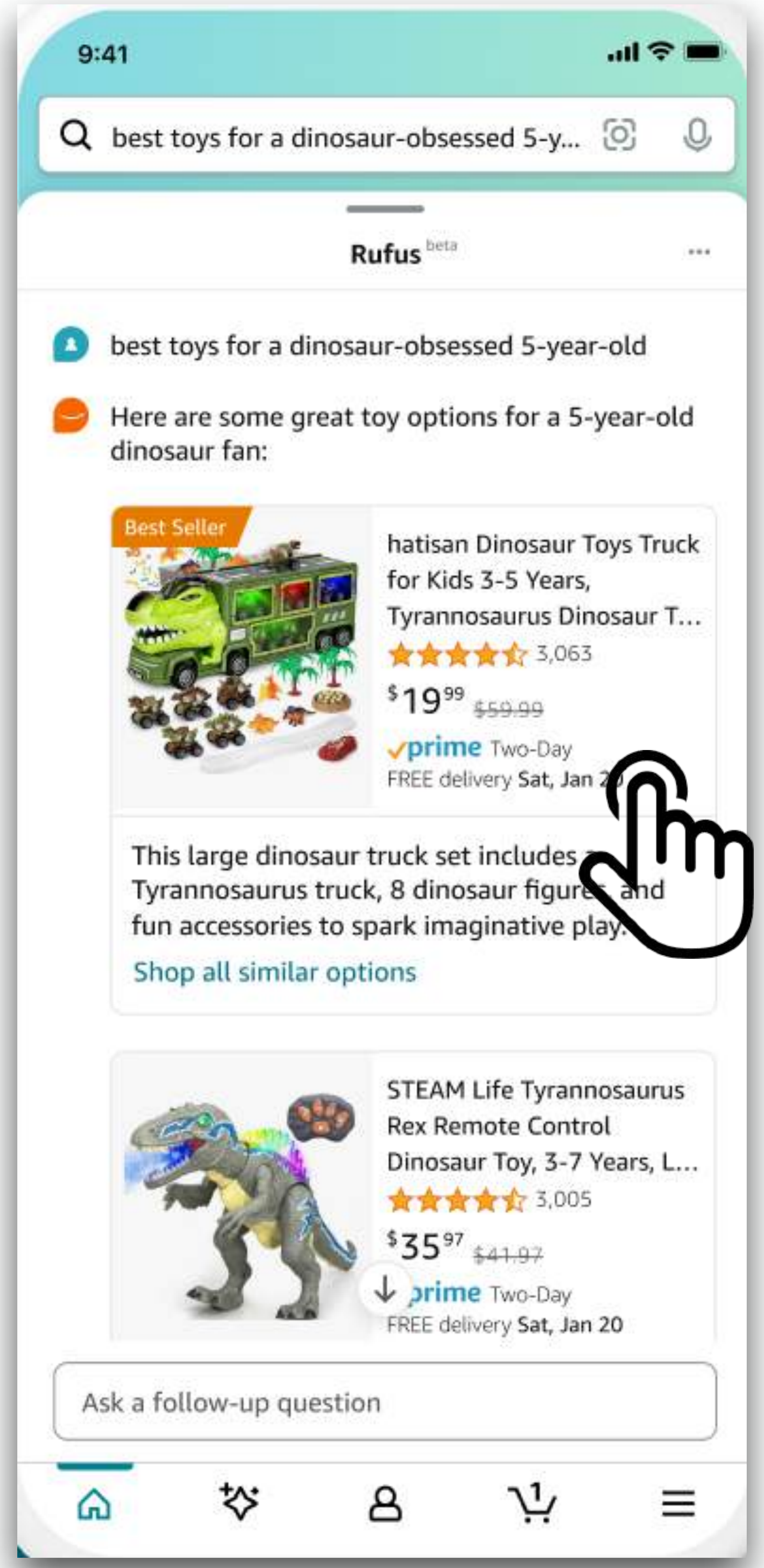
AI chatbots understand natural language

Users can reveal preferences through choices, not just words



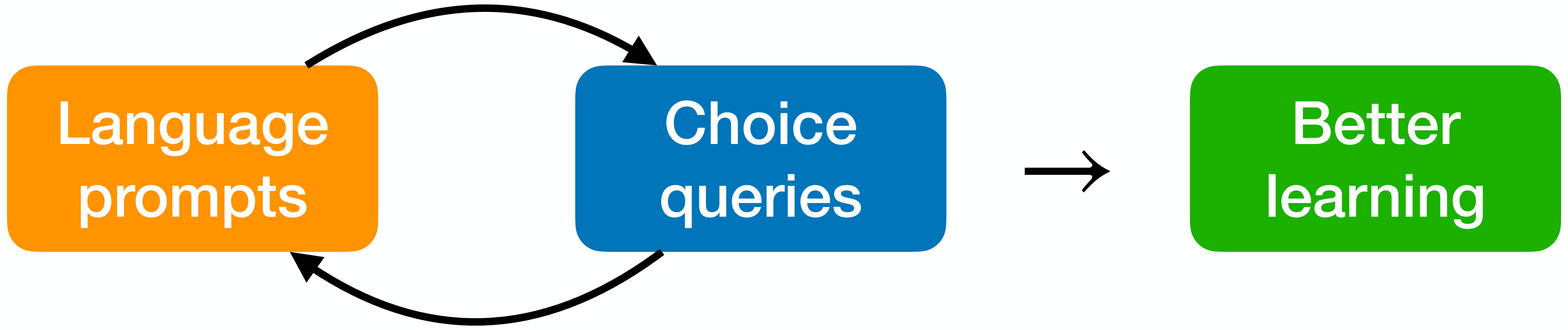


Vision: From Clicks to Conversations

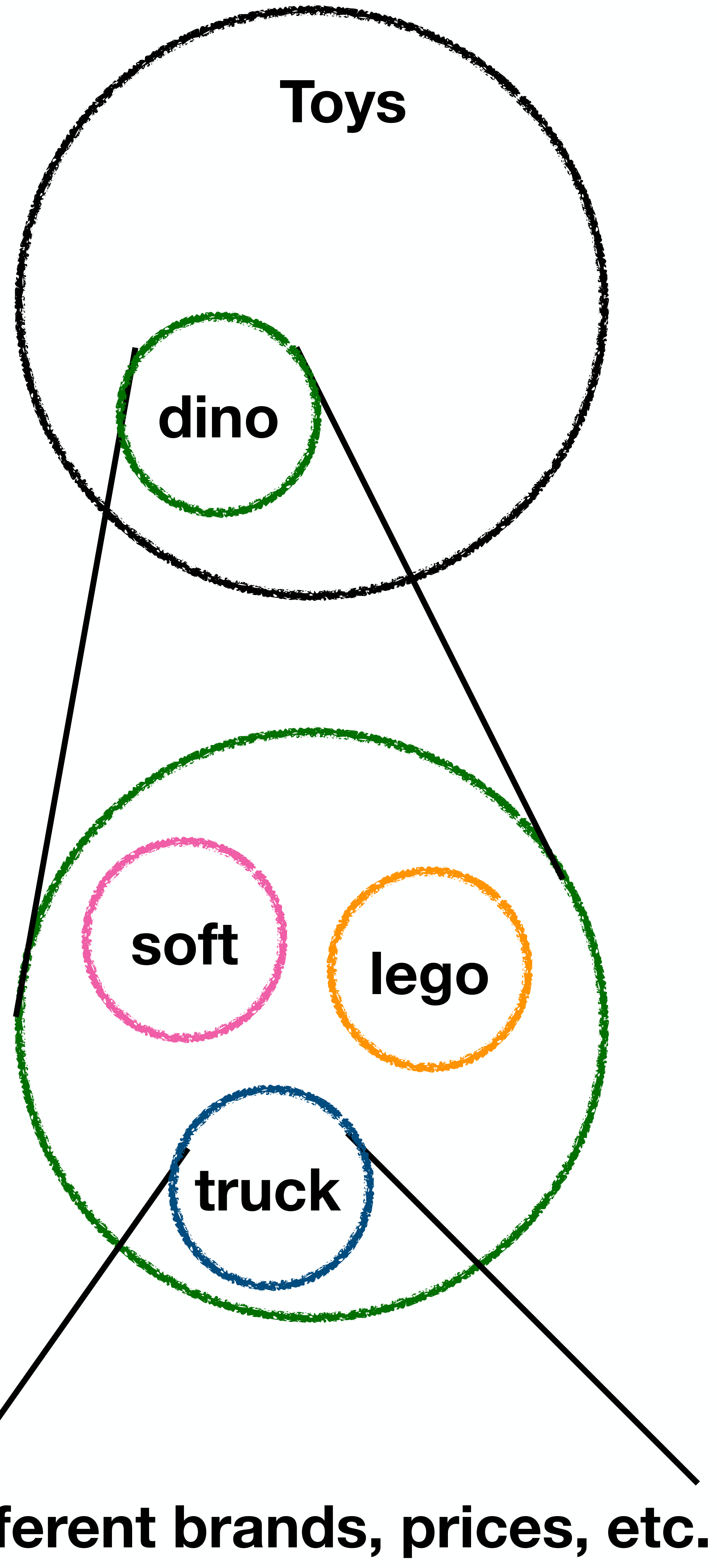


AI chatbots understand natural language

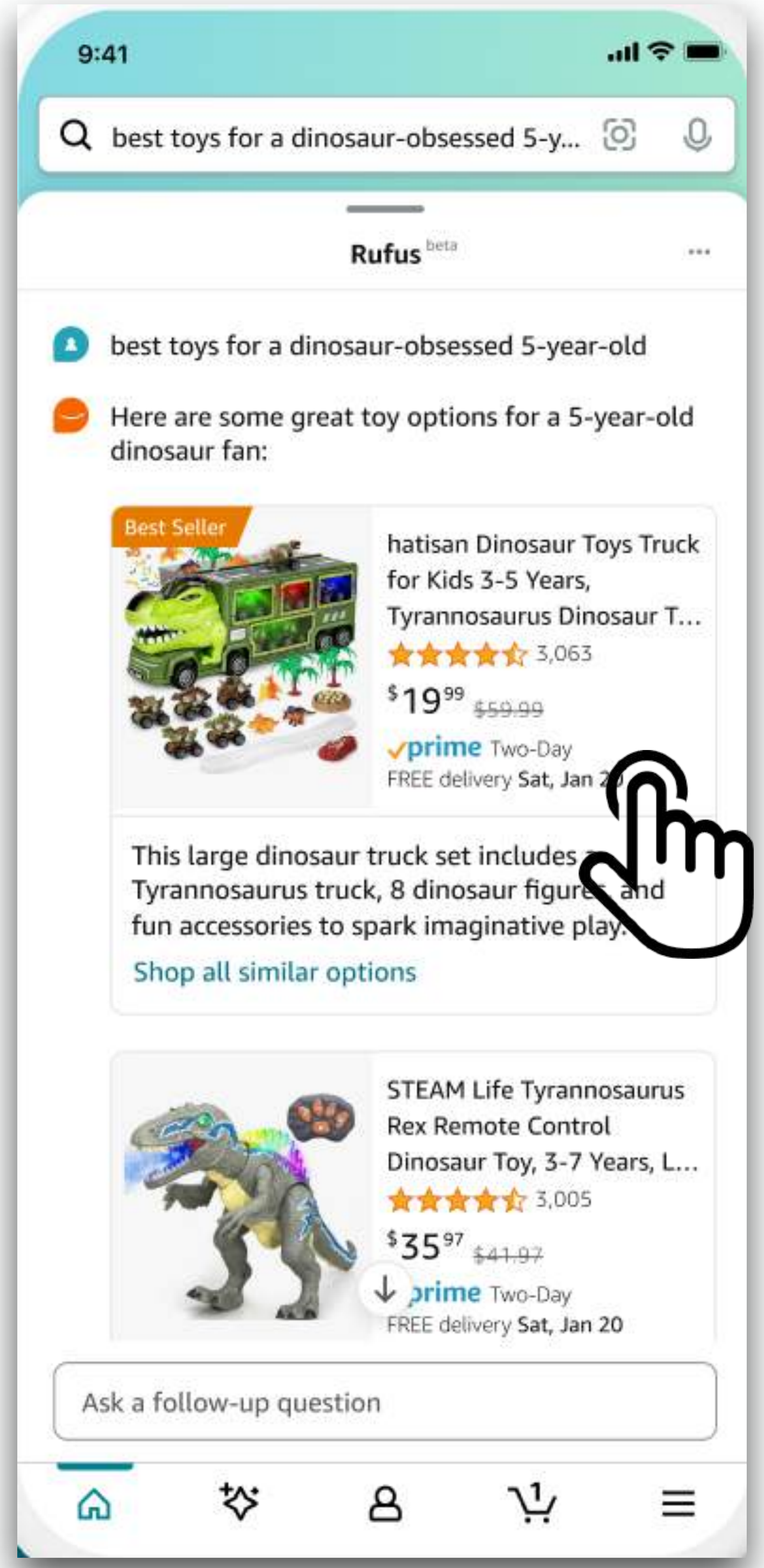
Users can reveal preferences through choices, not just words



- What queries to ask?
- When to ask choice queries?



Vision: From Clicks to Conversations

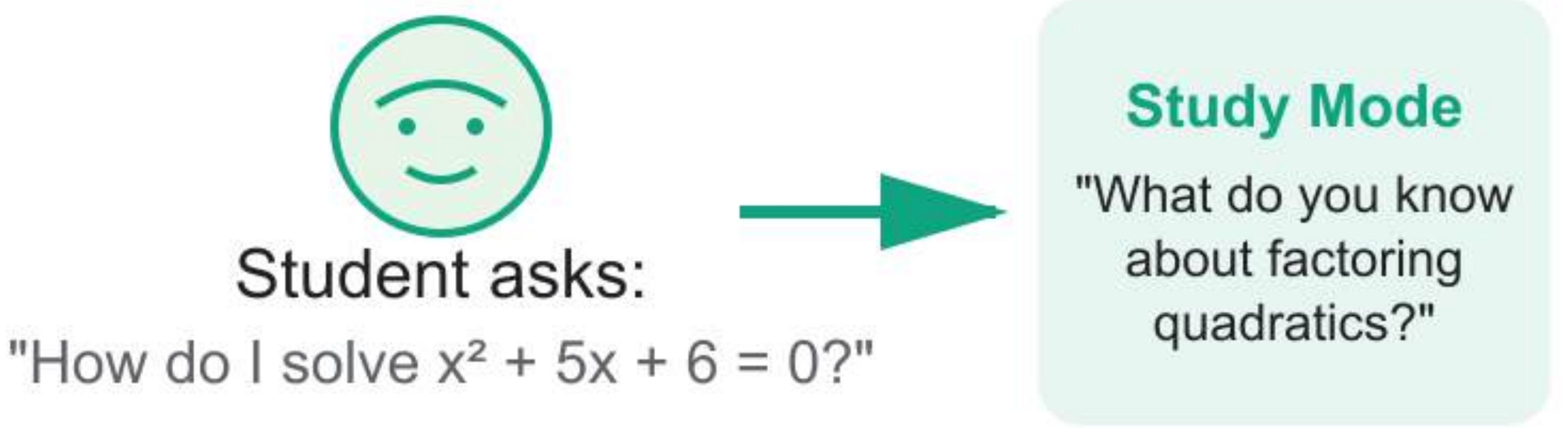


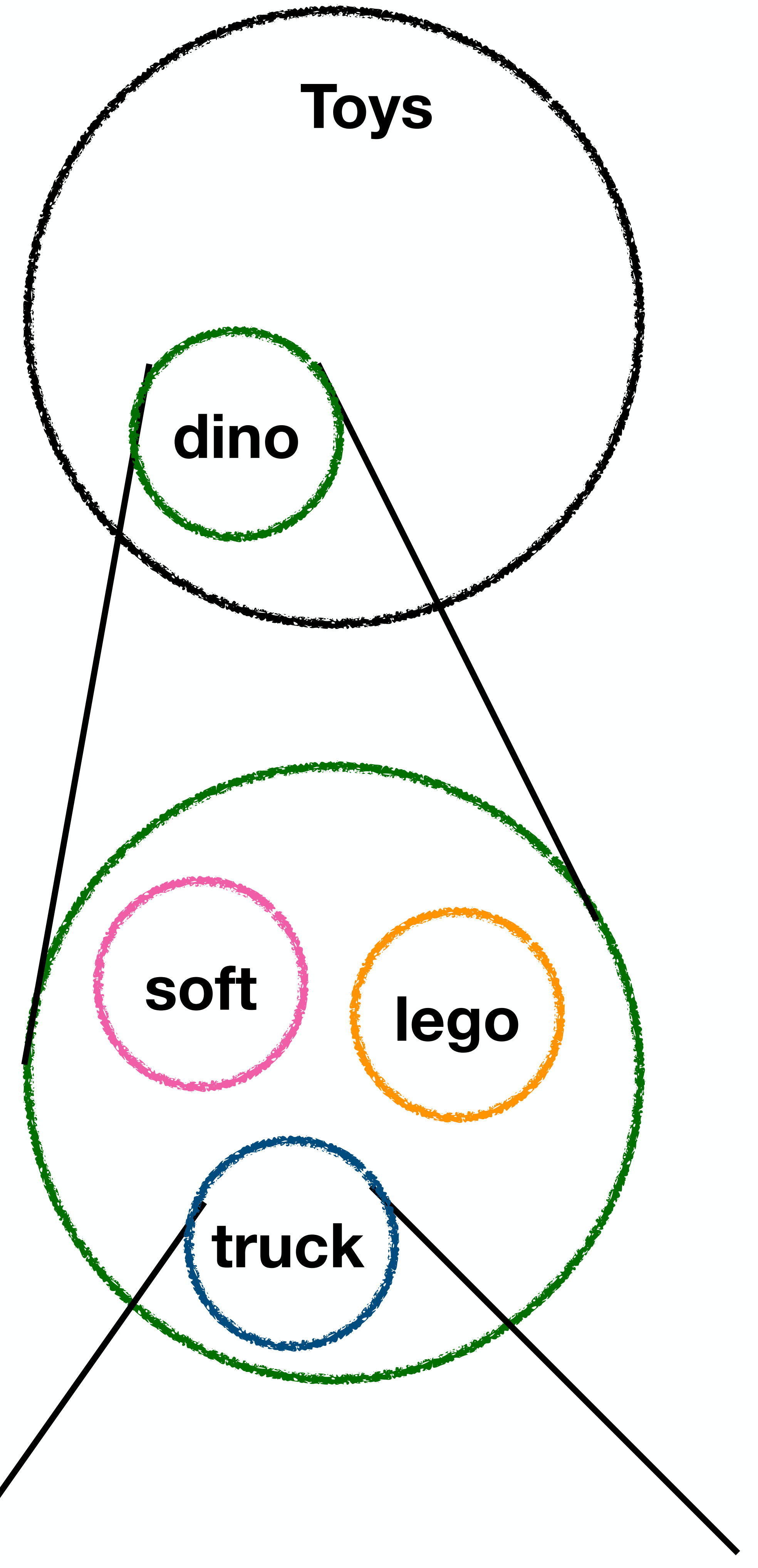
AI chatbots understand natural language

Users can reveal preferences through choices, not just words

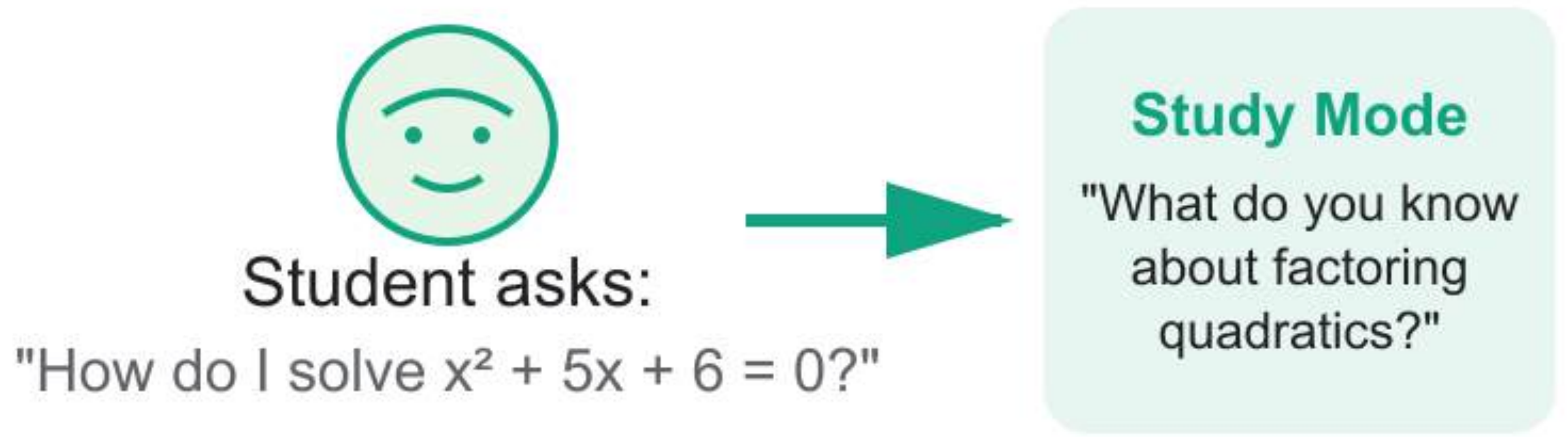
Language prompts ⇄ Choice queries → Better learning

- What queries to ask?
- When to ask choice queries? **Known!**

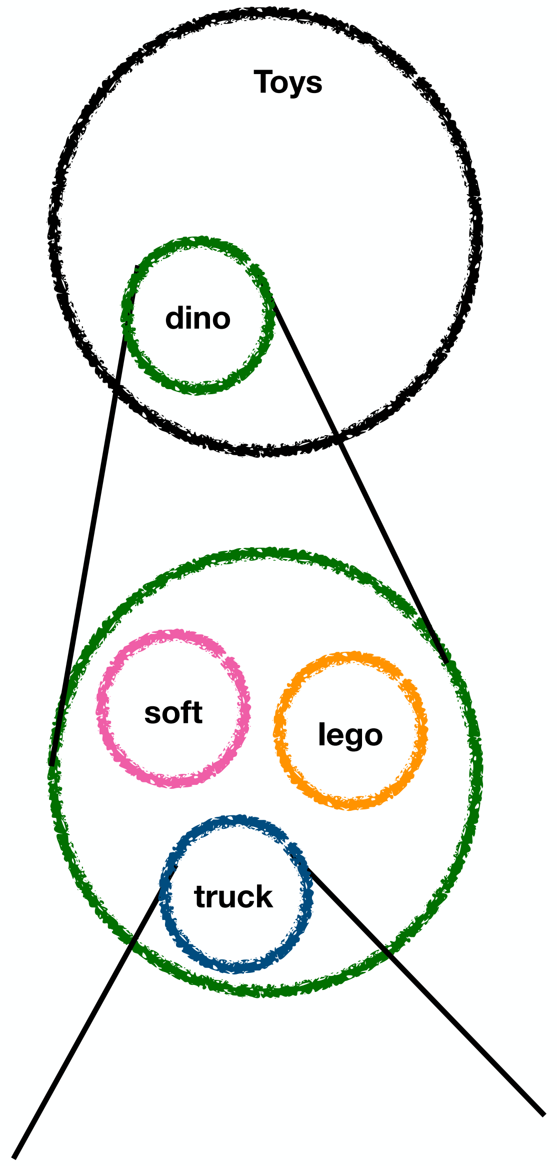




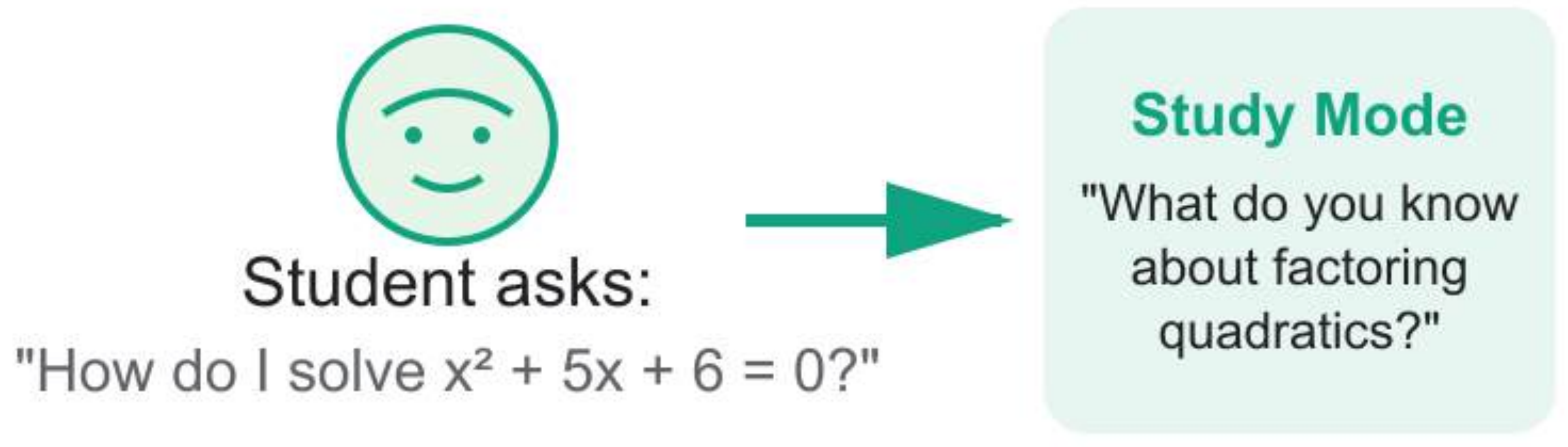
Long-Term Preference Learning & Growth



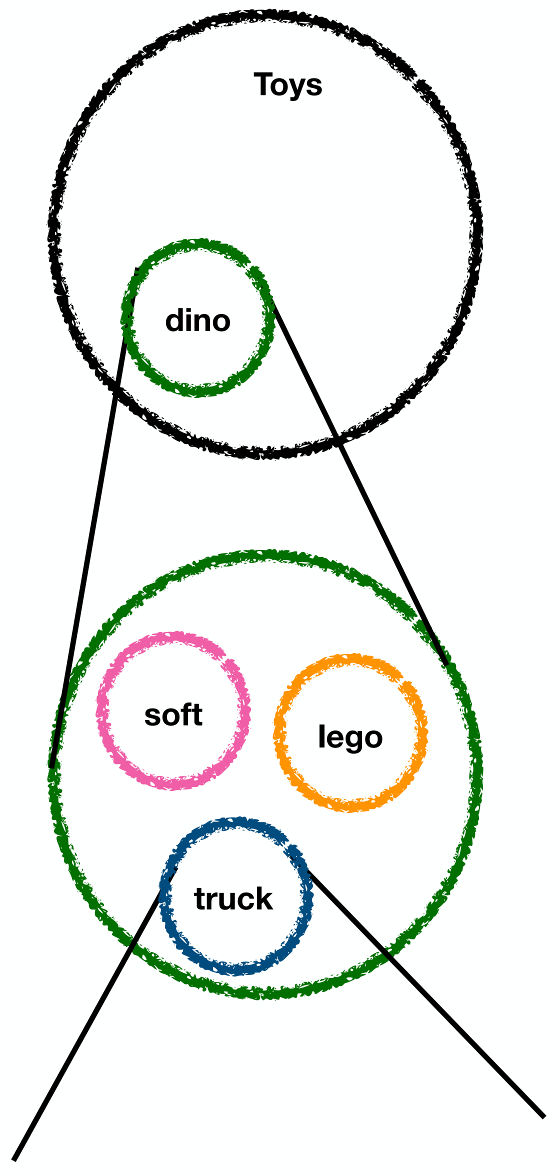
Interactive learning of preferences:
iterative state estimation



Long-Term Preference Learning & Growth

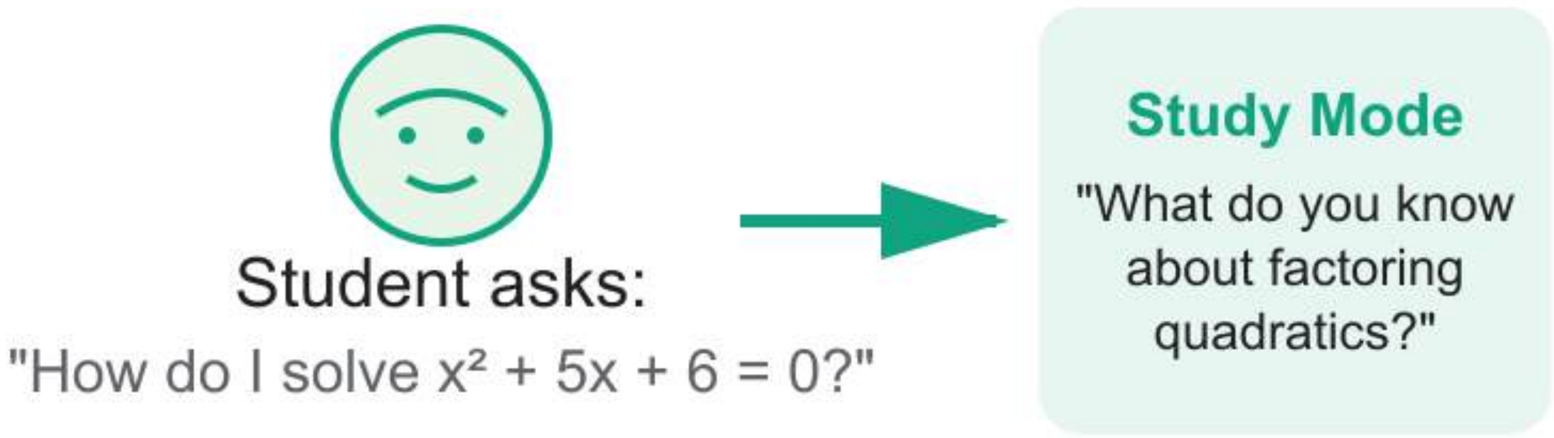


Interactive learning of preferences:
iterative state estimation



- Options shown can shape preferences
- E.g., career counselling, PhD advising

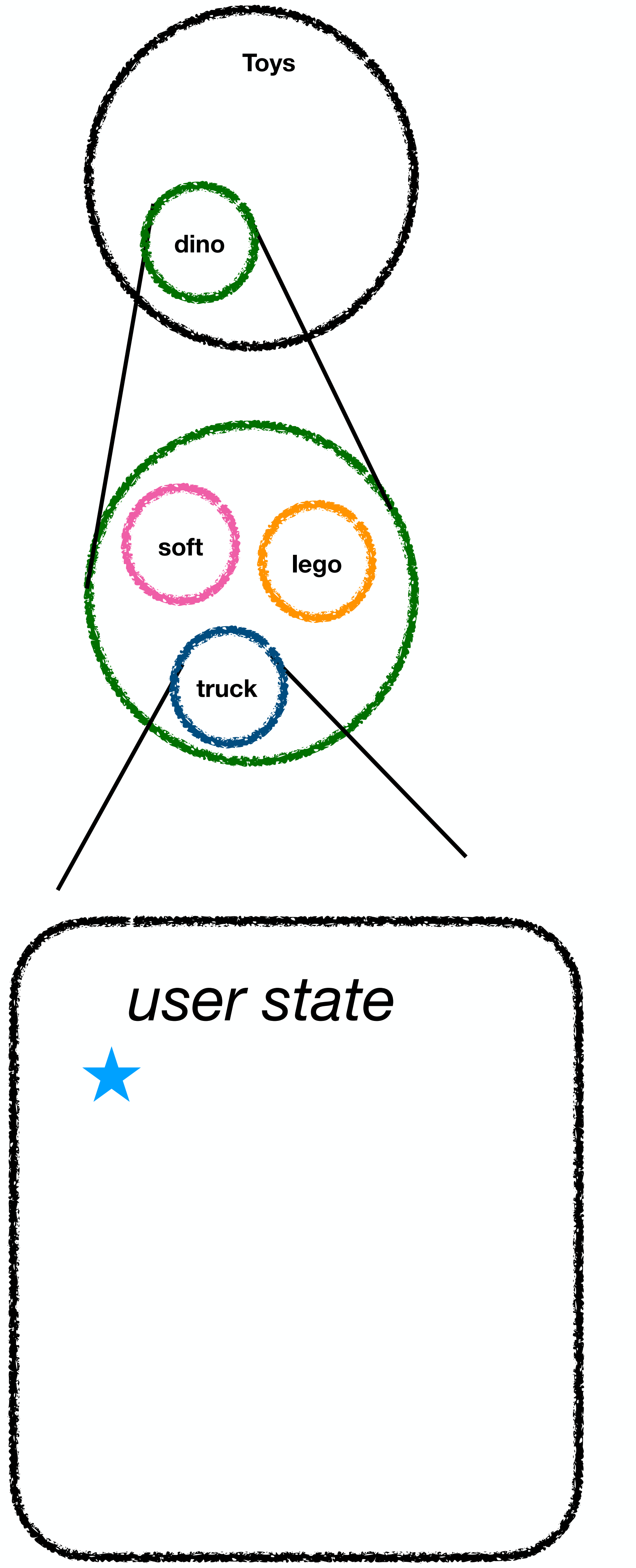
Long-Term Preference Learning & Growth



Interactive learning of preferences:
iterative state estimation

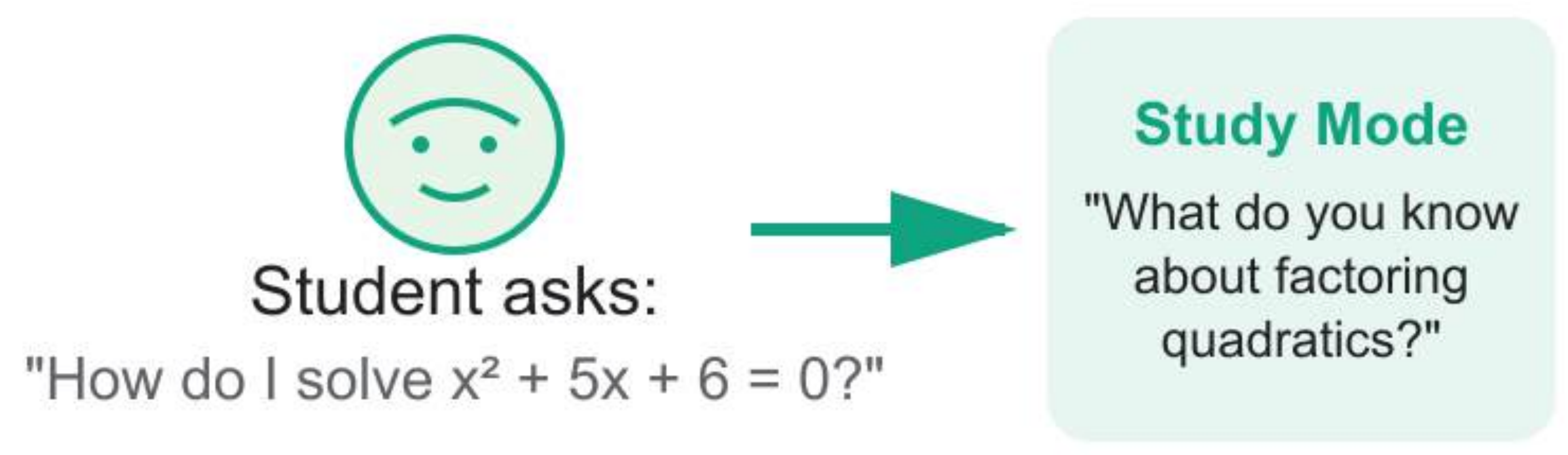
- Options shown can shape preferences
- E.g., career counselling, PhD advising

How can an AI chatbot help a user grow?
A control problem

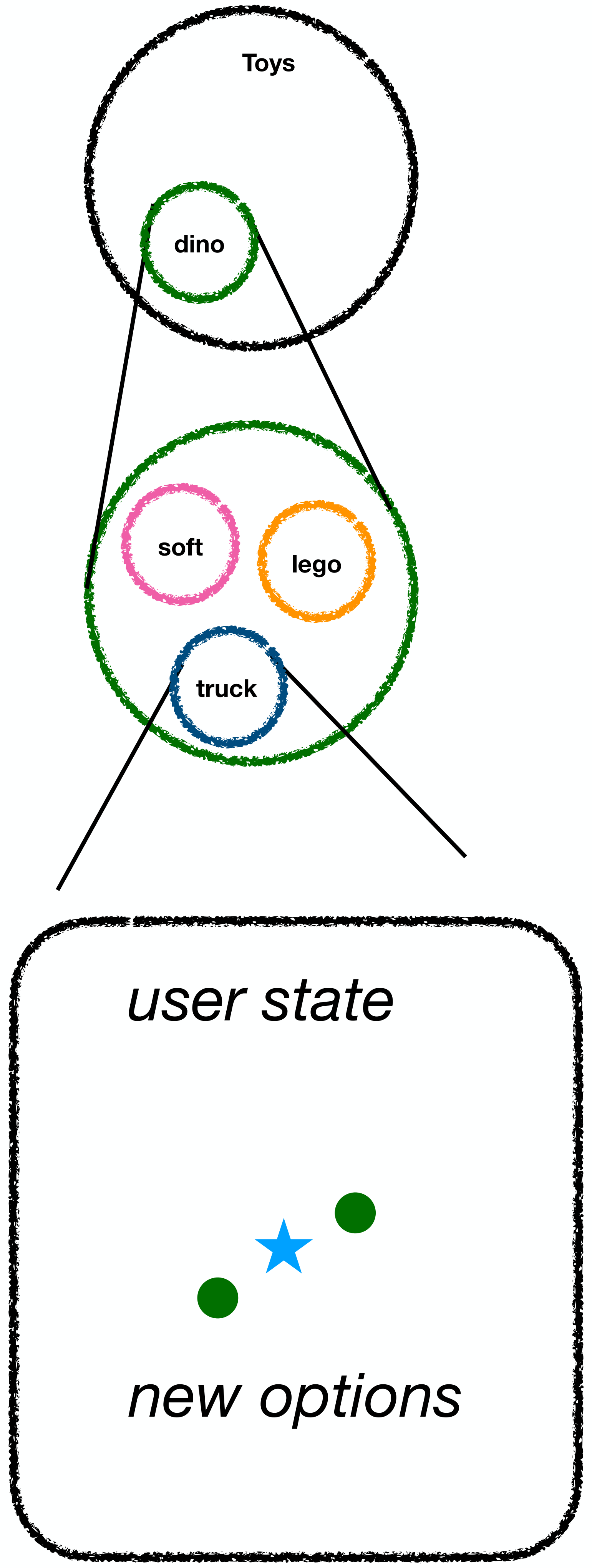


preference state space

Long-Term Preference Learning & Growth



Interactive learning of preferences:
iterative state estimation



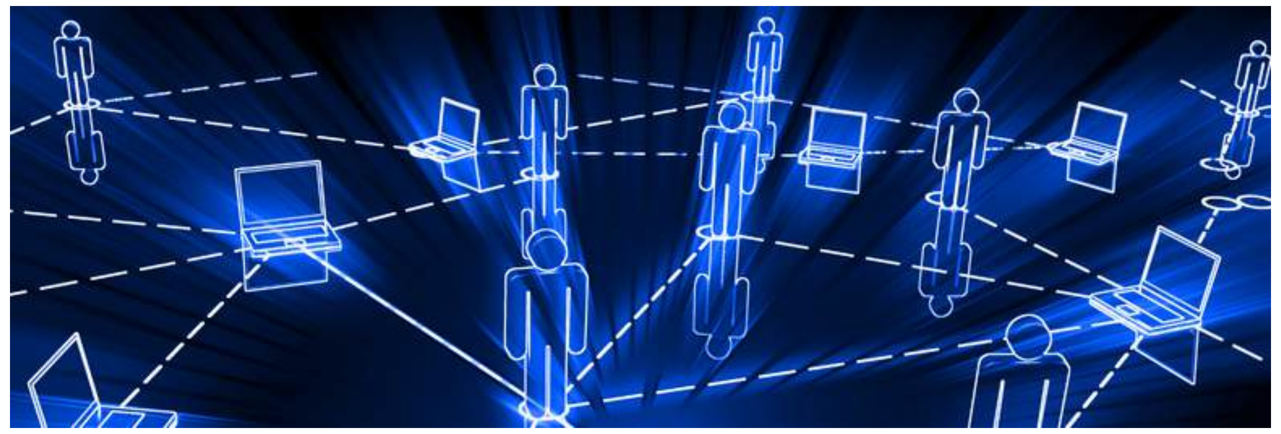
- Options shown can shape preferences
- E.g., career counselling, PhD advising

How can an AI chatbot help a user grow?
A control problem

preference state space

Concluding Remarks

Broad Theme



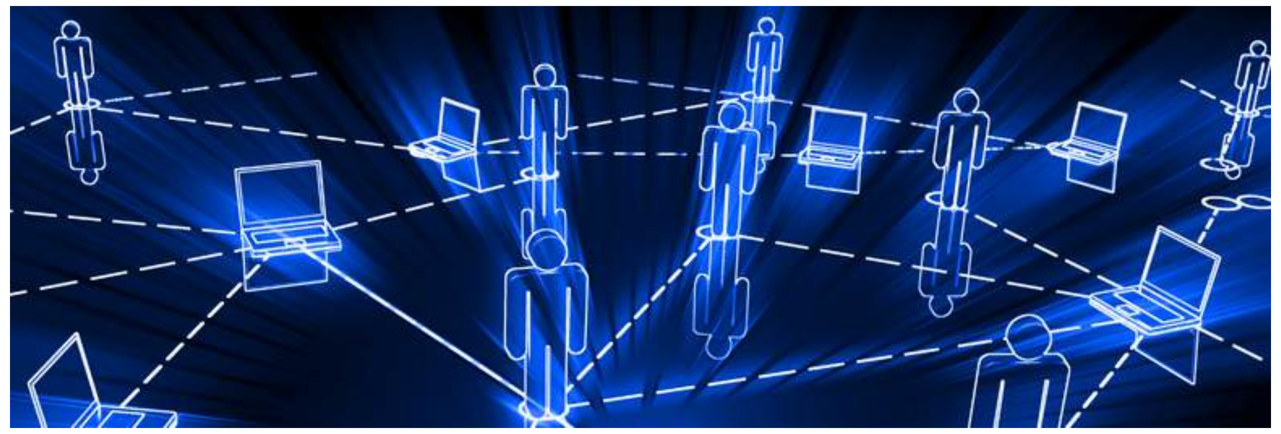
Theory of human-in-the-loop systems

Current Focus

Recommender systems that
learn from comparisons

Concluding Remarks

Broad Theme

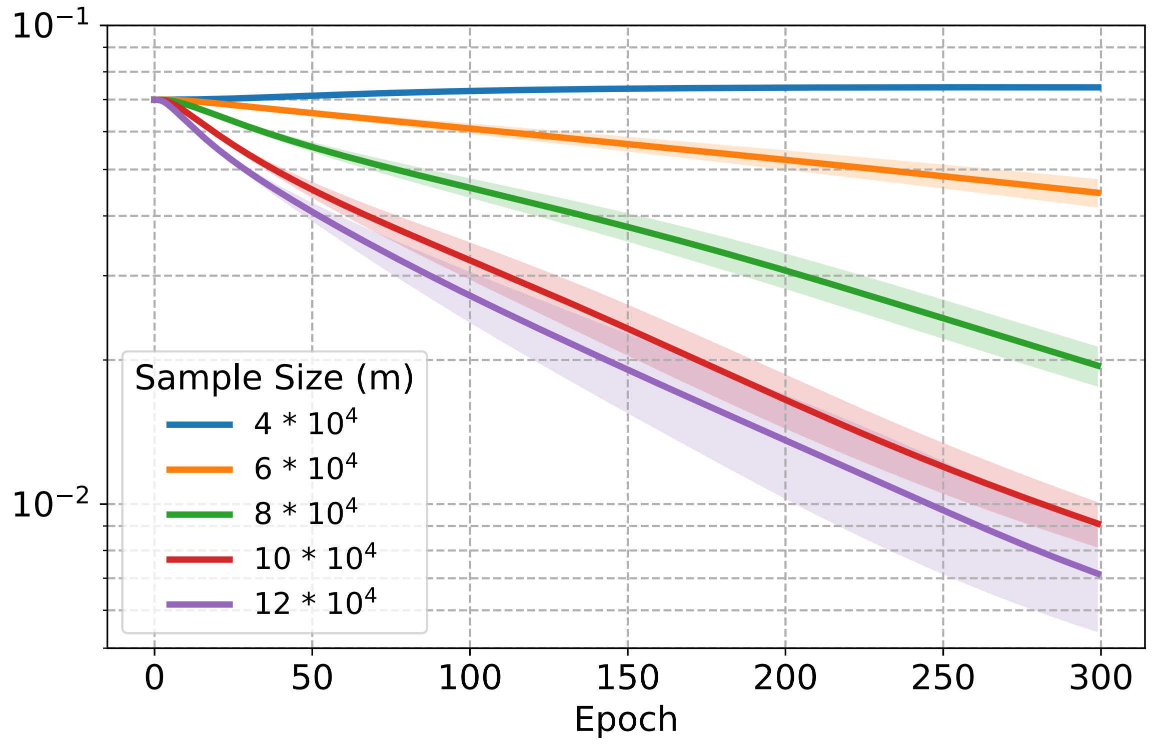


Theory of human-in-the-loop systems

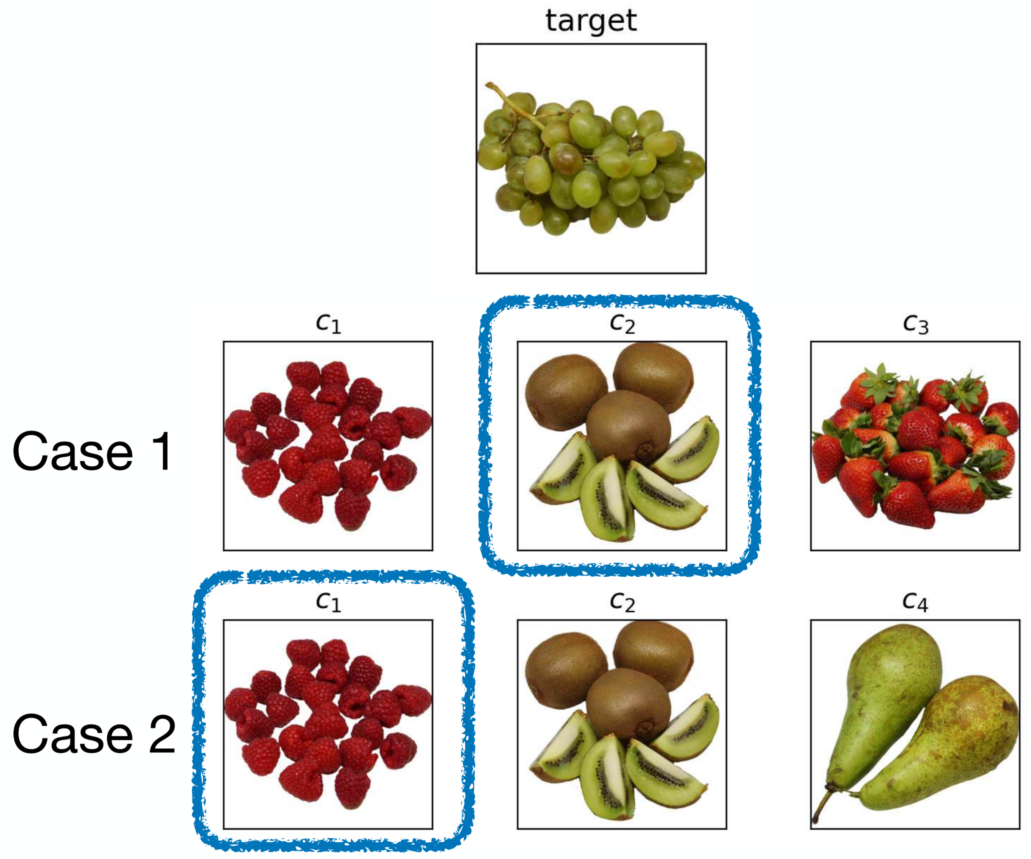
Current Focus

Recommender systems that learn from comparisons

Theoretical Guarantees

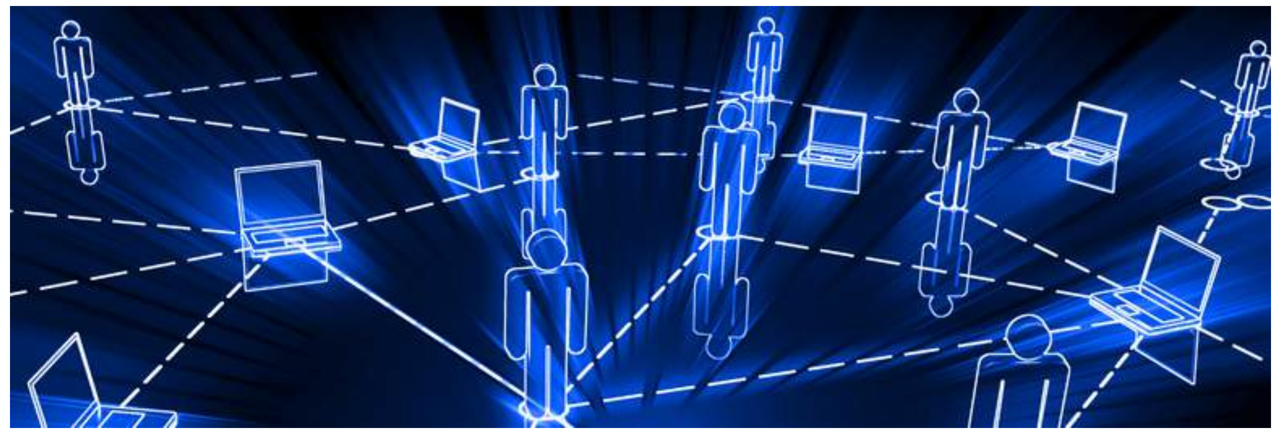


Empirical Observations



Concluding Remarks

Broad Theme

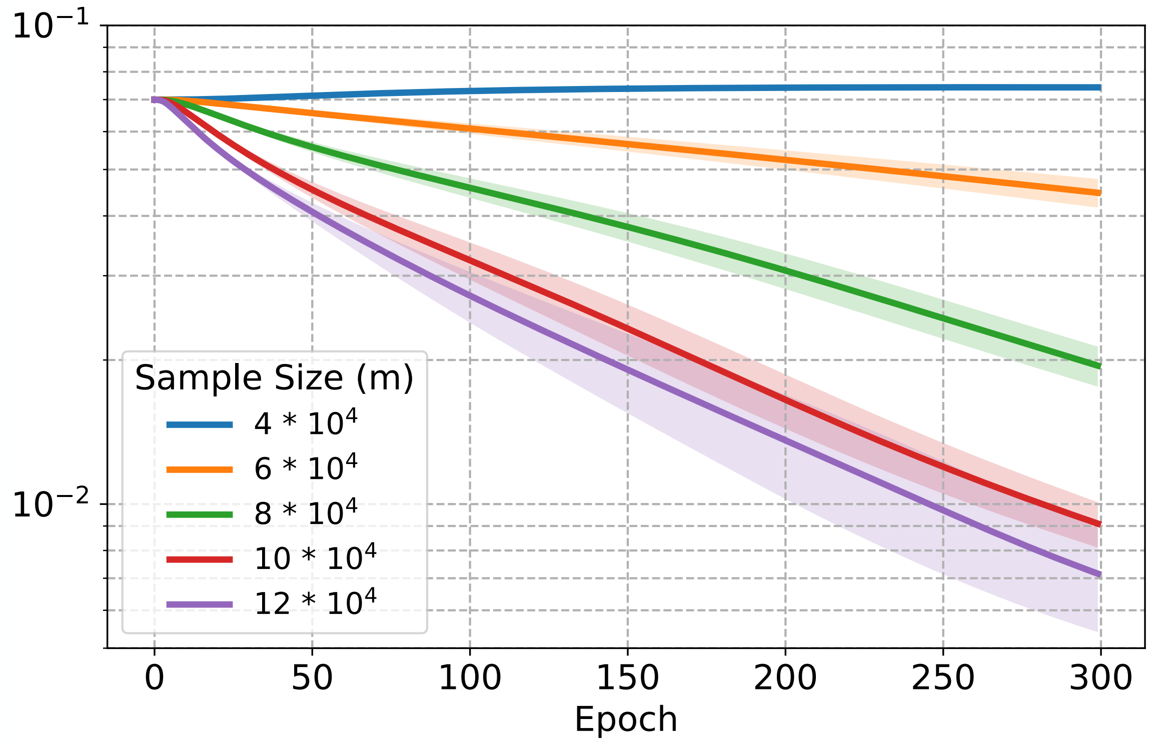


Theory of human-in-the-loop systems

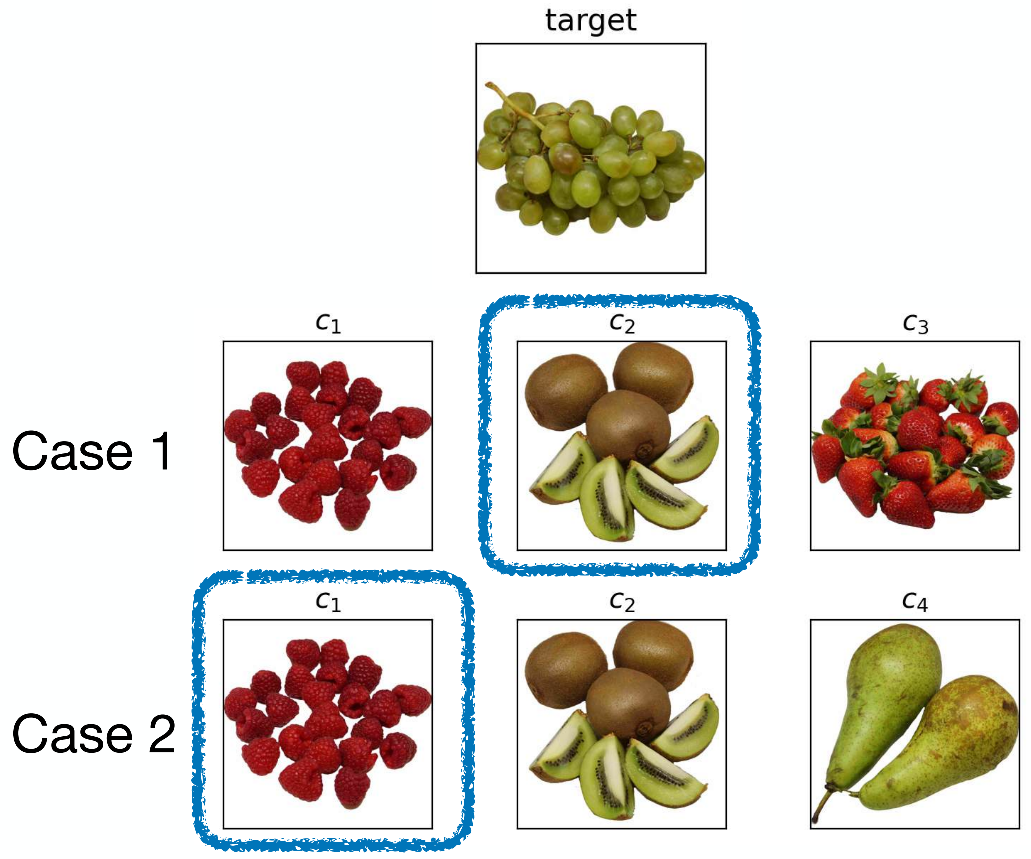
Current Focus

Recommender systems that learn from comparisons

Theoretical Guarantees

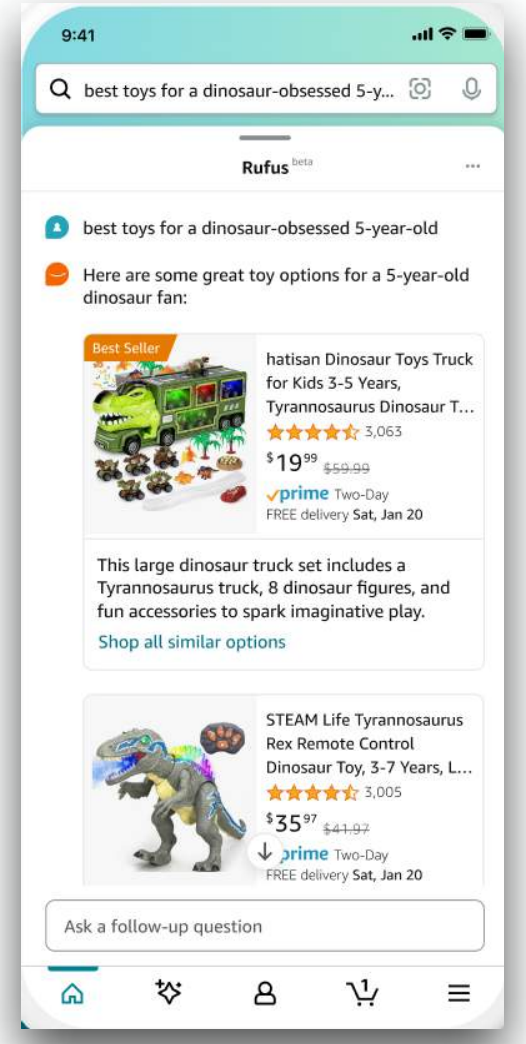


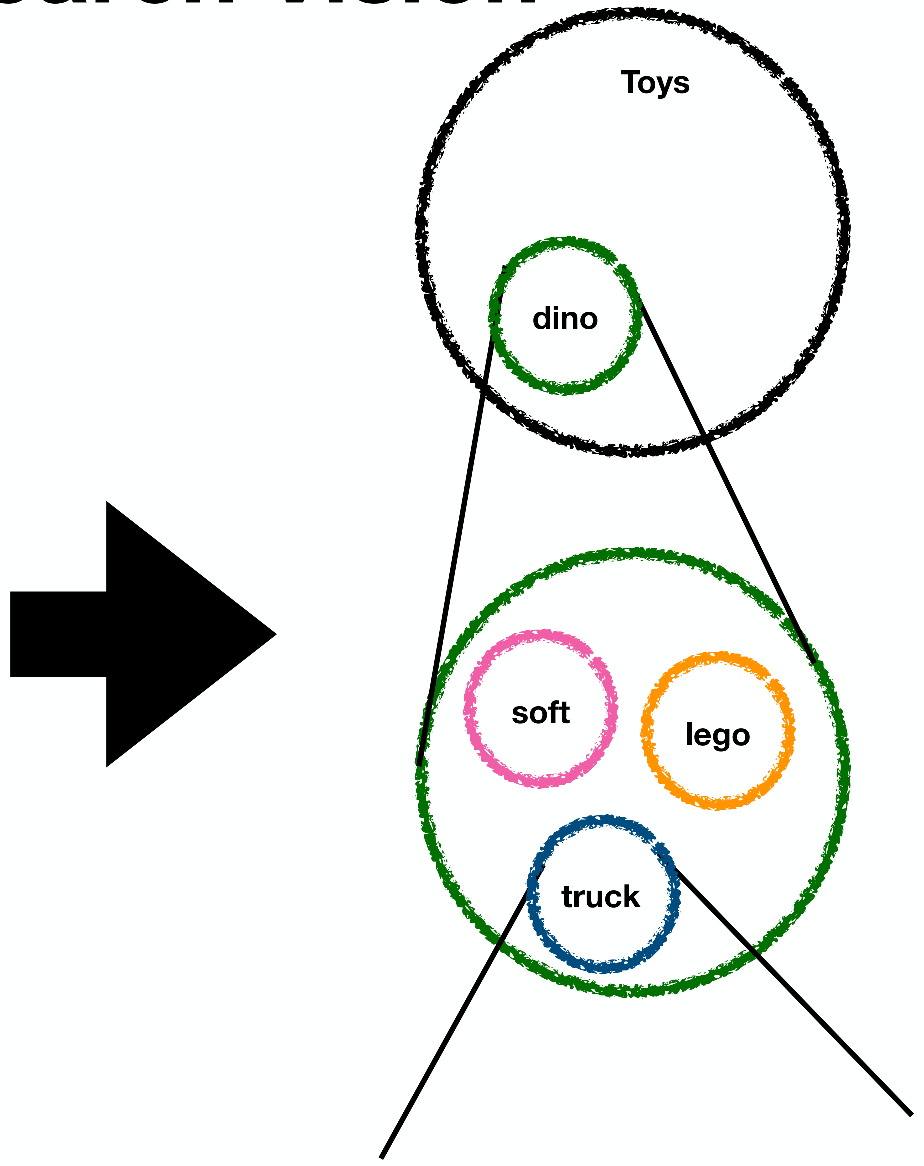
Empirical Observations



Thank you!

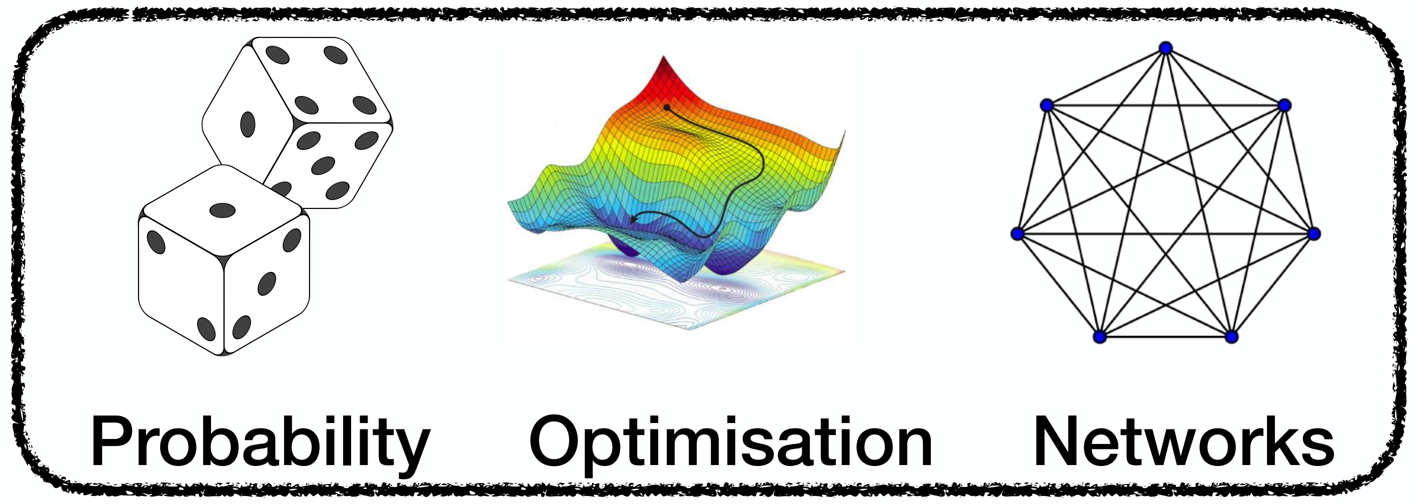
Research Vision





interactive recommender systems

Methodology



with industry, AI practitioners