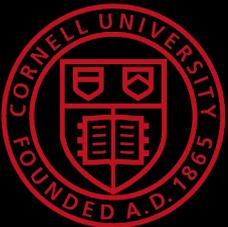


Designing AI Systems with Steerable Long-Term Dynamics

Ashudeep Singh, Adith Swaminathan, Tobias Schnabel, Kianté Brantley,
Aman Agarwal, Marco Morik, Sarah Dean, Zhichong Fang, Yuta Saito,
Thorsten Joachims



Department of Computer Science, Cornell University
Department of Information Science, Cornell University

Funded in part through NSF Awards IIS-2008139, IIS-1615706, IIS-1901168.

Intelligent Online Systems

Ranking function π that ranks items for context x .

svm at DuckDuckGo

duckduckgo.com/?q=svm&it=h&sva=about

svm

Web Images Videos News Maps Meanings Stock Settings

All Regions Safe Search: Moderate Any Time

Support-vector machine

In machine learning, support-vector machines are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis. Wikipedia

Support-vector machine - Wikipedia

W: https://en.wikipedia.org/wiki/Support_vector_machine

In machine learning, support-vector machines (SVMs, also support-vector networks) are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis.

SVM - Gas Gift Card & Promotion Management

W: <https://www.svmorders.com>

SVM is a leading provider of gift card marketing, management & customization services. Find over 250 brands promotional gift card solutions for gas, retail & services.

Silvercorp Metals Inc. (SVM) - Yahoo Finance

<https://finance.yahoo.com/quote/SVM/>

Find the latest Silvercorp Metals Inc. (SVM) stock quote, history, news and other vital information to help you with your stock trading and investing.

Introduction to Support Vector Machine(SVM) | Dimensionless ...

<https://dimensionless.in/introduction-to-svm/>

A Support Vector Machine(SVM) is a yet another supervised machine learning algorithm. It

metal table legs | Etsy

etsy.com/search?q=metal%20table%20legs

metal dining table legs metal coffee table legs metal hairpin table legs table legs

All categories Home & Living Craft Supplies & Tools Art & Collectibles Paper & Party Supplies

Special offers

FREE shipping

On sale

Ready to ship in

1 business day

1 - 3 business days

Price (\$)

Any price

Under \$100

\$100 to \$500

\$500 to \$1,000

Over \$1,000

Custom

Low to High

Style

metal table legs (3,758 Results)

Heavy Duty Metal Table Legs... Modern Urban Metals ***** (4,165)

2 Pack U Shape Legs (1.5" W)... Custom Custom Rattica ***** (1,960)

Heavy Duty Metal Table Legs... Modern Urban Metals ***** (4,165)

2 Pack U Shape Legs (2" Wide - 1/4" Thick)... Custom Custom Rattica ***** (1,960)

Netflix

netflix.com/browse

NETFLIX

Durchsuchen

TOAST IN LONDON

SPECIAL COLLECTIONS

HAPPY! with

CRIMINAL MINDS

Hinterland

REPUBLIC DOYLE

NETFLIX ORIGINALS

THE GREAT BRITISH BAKING SHOW

THE LAST KING OF SCOTLAND

The Mindy Project

HuffPost - Breaking News, U.S., & World

huffpost.com/?guccounter=1&guce_referrer=aHR0cHMhL...

TOP STORIES

Pelosi Announces House Will Vote On Impeachment Process After Weeks Of GOP Protests

Katie Hill Speaks On Resignation: 'I'm Hurt, I'm Angry.'

Court Strikes Down North Carolina Congressional Map

Democratic Former Sen. Kay Hagan Dies At 86

Trump Publicly Revels In Death And Spectacle: 'Boom Boom Boom!'

Pallbearer Who Snubbed Mitch McConnell At Elijah Cummings' Memorial Breaks Silence

U.S. LANDS CHIEF WROTE ANTI-ENVIRONMENTAL SCREENS

Trump's Public Lands Chief Wrote For A Cult Extremist's Magazine

By Chris D'Angelo and Alexander C. Kaufman

UKRAINE CALL WITNESS NO-SHOWS IMPEACHMENT HEARING

Former White House Aide Won't Show For Scheduled Impeachment

Maximizing Utility to Users

Probability Ranking Principle [Robertson, 1977]:

- Sort documents by probability of relevance
→ Optimal ranking y^*
- For most common measures U of ranking quality

$$U(y^* | x) = \max_y [U(y | x)]$$

Query x		
Rank	Item	P(relevant)
1	A	60.99
2	B	58.98
3	C	53.97
4	D	51.00
5	E	49.99
6	F	46.98
7	G	42.97
...

Dynamics of Utility Maximization

Conventional Rankings:

- Unfair allocation of opportunity
- Suboptimal social welfare
- Amplification of existing biases
- Reduced supplier pool
- Polarization

Utility maximization for users

≠

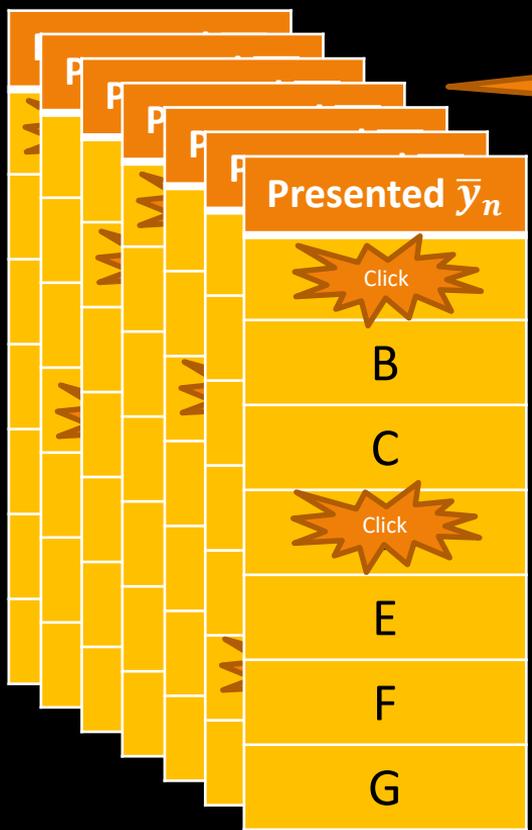
Long-term sustainability of platform

Top News Stories		
Rank	Item	P(read)
1	Times 1	50.99
2	Times 2	50.98
3	Times 3	50.97
...
100	Review 1	49.99
101	Review 2	49.98
102	Review 3	49.97
...

Sustainable Platforms

- 1. Unbiased Estimation of Relevance
2. Fair Treatment of all Platform Participants
3. Steerable Control of Platform Dynamics

Learning-to-Rank from Clicks



Query Distribution
 $x_i \sim P(X)$
Deployed Ranker
 $\bar{y}_i = \pi_0(x_i)$

Learning
Algorithm

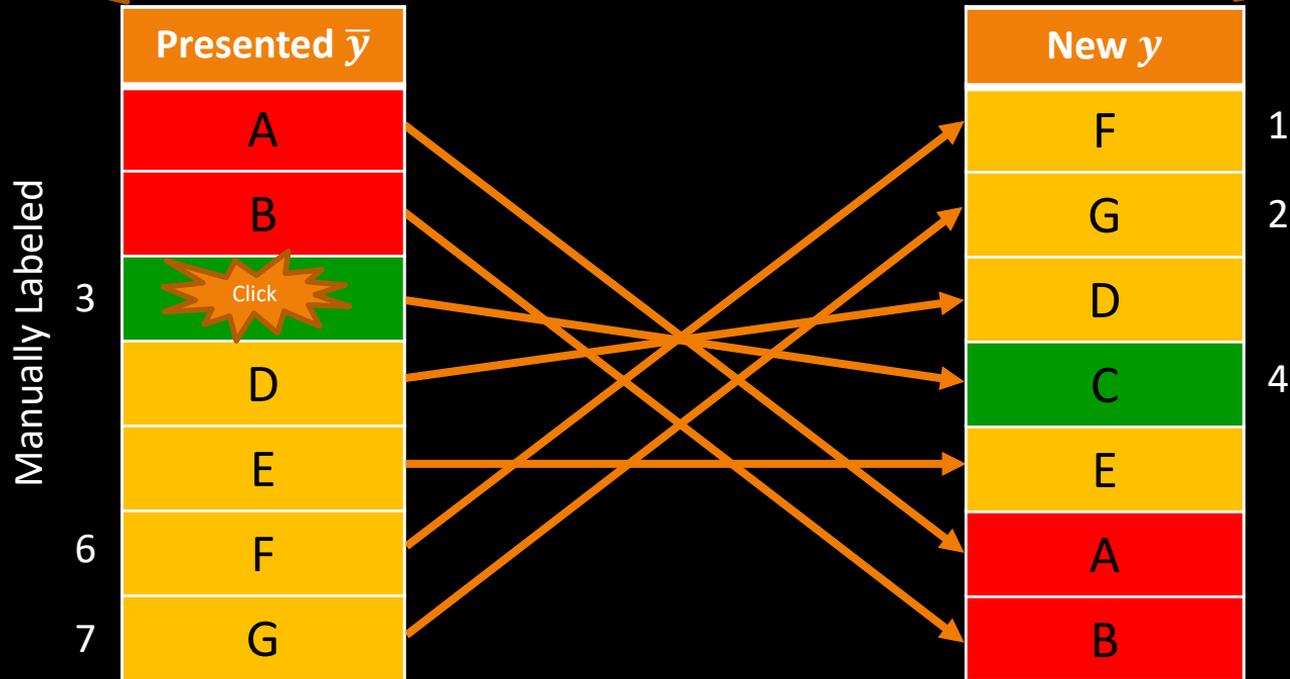
New Ranker
 $\pi(x)$

Should perform
better than
 $\pi_0(x)$

Evaluating Rankings

Deployed Ranker
 $\bar{y} = \pi_0("SVM")$

New Ranker to Evaluate
 $y = \pi("SVM")$



Evaluation with Missing Judgments

- Loss: $\Delta(x, y|rel)$

- Relevance labels $rel_d \in \{0,1\}$
- This talk: rank of relevant documents

$$\Delta(x, y|rel) = \sum_d rank(d|y) \cdot rel_d$$

- Assume:

- Click implies observed and relevant:

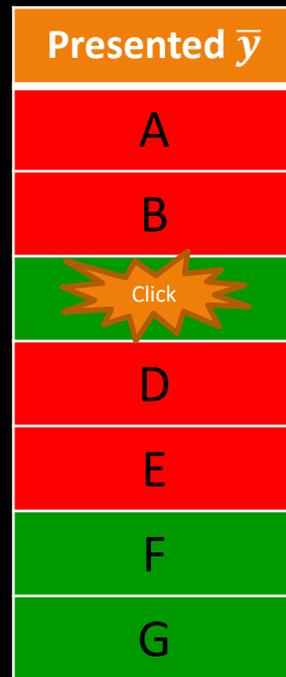
$$(c_d = 1) \leftrightarrow (o_d = 1) \wedge (rel_d = 1)$$

- Problem:

- No click can mean not relevant OR not observed

$$(c_d = 0) \leftrightarrow (o_d = 0) \vee (rel_d = 0)$$

- → Understand observation mechanism



Inverse Propensity Score Estimator

- Observation Propensities $Q(o_d = 1|x, \bar{y}, rel)$
 - Random variable $o_d \in \{0,1\}$ indicates whether relevance label rel_d for is observed

- Inverse Propensity Score (IPS) Estimator:

$$\hat{\Delta}(x, y|rel, o) = \sum_{d:c_d=1} \frac{rank(d|y)}{Q(o_d = 1|x, \bar{y}, rel)}$$

New Ranking

- Unbiasedness: $E_o \left[\hat{\Delta}(x, y | rel, o) \right] = \Delta(x, y|rel)$

Presented \bar{y}	Q
A	1.0
B	0.8
C	0.5
D	0.2
E	0.2
F	0.2
G	0.1

ERM for Partial-Information LTR

- Unbiased Empirical Risk:

$$\hat{V}_{IPS}(\pi) = \frac{1}{N} \sum_{(x,a,c) \in S} \sum_{d:c_d=1} \frac{\text{rank}(d|\pi(x))}{Q(o_d = 1|x, \bar{y}, rel)}$$

Consistent
Estimator of
True
Performance

- ERM Learning:

$$\hat{\pi} = \underset{\pi}{\operatorname{argmin}} \left[\hat{V}_{IPS}(\pi) \right]$$

Consistent
ERM
Learning

- Questions:

- How do we optimize this empirical risk in a practical learning algorithm?
- How do we define and estimate the propensity model $Q(o_d = 1|x, \bar{y}, rel)$?

Propensity-Weighted SVM Rank

- Data: $D = (x_j, d_j, D_j, q_j)^n$



Optimizes convex upper bound on unbiased IPS risk estimate!

- Training QP:

$$w^* = \operatorname{argmin}_{w, \xi \geq 0} \frac{1}{2} w \cdot w + \frac{C}{n} \sum_j \frac{1}{q_j} \sum_i \xi_j^i$$
$$\forall \bar{d}^i \in D_1: w \cdot [\phi(x_1, d_1) - \phi(x_1, \bar{d}^i)] \geq 1 - \xi_1^i$$
$$\vdots$$
$$\forall \bar{d}^i \in D_n: w \cdot [\phi(x_n, d_n) - \phi(x_n, \bar{d}^i)] \geq 1 - \xi_n^i$$

- Loss Bound: $\forall w: \operatorname{rank}(d, \operatorname{sort}(w \cdot \phi(x, d))) \leq \sum_i \xi^i + 1$
- Analogous method with Deep Nets [Agarwal et al., 2019b]

Position-Based Propensity Model

- Model:

$$P\left(c_d = 1 | rel_d, rank(d|\bar{y})\right) = q_{rank(d|\bar{y})} \cdot [rel_d = 1]$$

- Assumptions

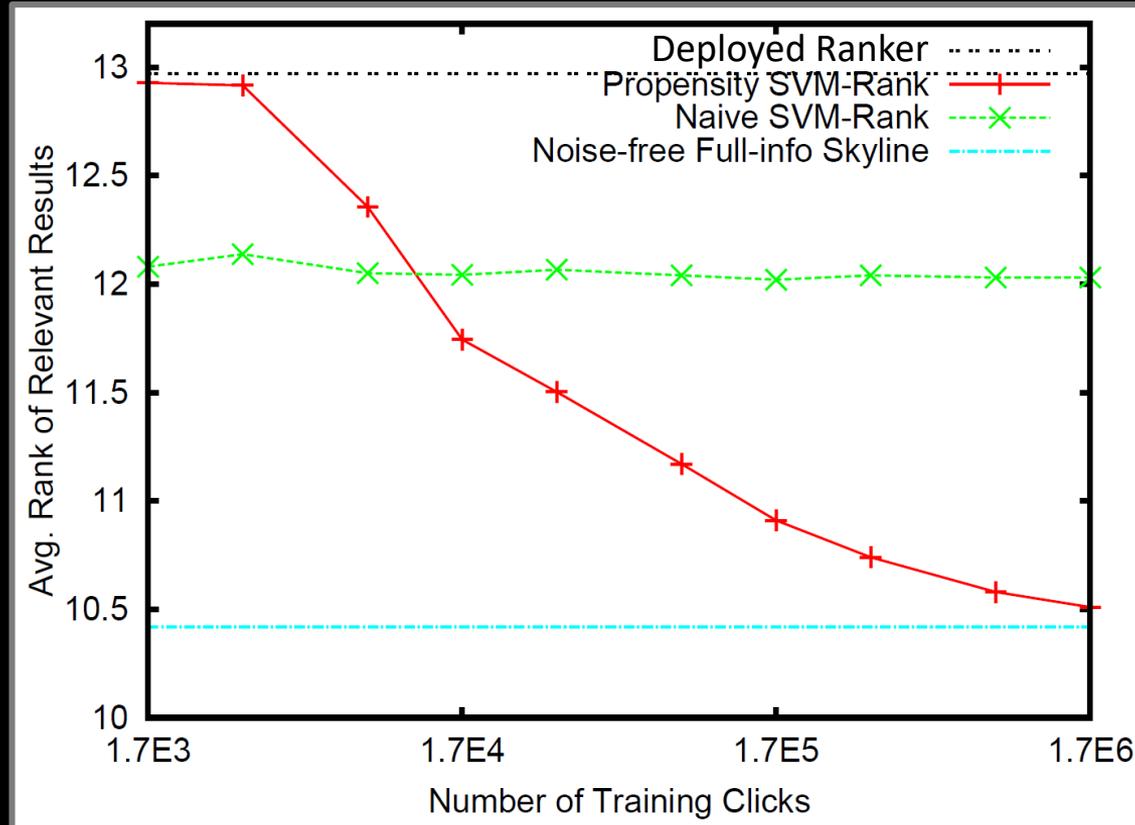
- Examination only depends on rank
- Click reveals relevance if rank is examined

- Estimation

- Estimate q_1, \dots, q_k via small intervention experiments
- See [Joachims et al., 2017] [Agarwal et al., 2019a] [Fang et al., 2019] [Chandar & Carterette, 2018]

Presented \bar{y}	q
A	q_1
B	q_2
C	q_3
D	q_4
E	q_5
F	q_6
G	q_7

Ranking Accuracy vs. Training Data



Sustainable Platforms

1. Unbiased Estimation of Relevance

- Selection bias correction through IPS [Joachims et al. 2017]
Unbiased learning of deep ranking policies [Agarwal et al. 2019]

2. Fair Treatment of all Platform Participants

3. Steerable Control of Platform Dynamics

Dynamics of Utility Maximization

Conventional Rankings:

- Unfair allocation of opportunity
- Suboptimal social welfare
- Amplification of existing biases
- Reduced supplier pool
- Polarization

Utility maximization for users

≠

Long-term sustainability of platform

Query: Software Engineer		
Rank	Item	P(interview)
1	Adam	50.99
2	Bob	50.98
3	Charlie	50.97
...
100	Alice	49.99
101	Barbara	49.98
102	Claire	49.97
...

Exposure
high

Exposure
low

Position-Based Exposure Model

Definition:

Exposure e_j is the probability a users observes item i at position j of ranking y .

$$expo(i|x, y) = e_j$$

Definition:

Exposure of group G of items

$$expo(G|x, y) = \sum_{j \in G} e_j$$

Note: Same as propensity model used earlier.

Rank	Exposure P(observe)
1	e_1
2	e_2
3	e_3
...	...
100	e_{100}
101	e_{101}
102	e_{102}
...	...

Merit-Based Fairness Constraints

$$exposure = f(relevance)$$

- Disparate Exposure:
 - Expected exposure proportional to the expected relevance of the group
- Disparate Impact:
 - Expected revenue (e.g. clicks) proportional to the expected relevance of the group
- Group parity:
 - Expected exposure equal for all groups

Disparate Exposure Constraint

Group Exposure and Merit

$$\text{expo}(G|x, \pi) = \sum_{i \in G} \text{expo}(i|x, y) \quad \text{rel}(G|x) = \sum_{i \in G} \text{rel}(i|x)$$

Group Fairness Constraint

$$\frac{\text{expo}(G_0|x, y)}{\text{rel}(G_0|x)} = \frac{\text{expo}(G_1|x, y)}{\text{rel}(G_1|x)}$$

→ Make exposure proportional to relevance

Computing the Best Fair Ranking

Goal: Maximize ranking quality while fair to items.

$$y = \operatorname{argmax}_y [DCG(y|x)]$$
$$s.t. \quad \frac{\operatorname{expo}(G_0|x, y)}{\operatorname{rel}(G_0|x)} = \frac{\operatorname{expo}(G_1|x, y)}{\operatorname{rel}(G_1|x)}$$

→ Computationally hard and typically infeasible!

Probabilistic Ranking Policies $\pi(y|x)$

Exposure and Quality for $\pi(y|x)$

$$expo(i|x, \pi) = \sum_j \mathbb{P}_{i,j} e_j$$

$$DCG(\pi|x) = \sum_i \sum_j e_j \mathbb{P}_{i,j} rel_i$$

$\mathbb{P}_{i,j}$ = Prob that item i is ranked at position j

e_j = exposure at position j

π

y_1	y_2	y_3	y_4
A	B	A	B
B	A	C	C
C	C	B	A
D	D	D	G
E	E	E	F
F	F	F	E
G	G	G	D

0.52 0.23 0.20 0.05

Marginal Rank Distribution \mathbb{P}

π

	y_1	y_2	y_3	y_4
A	A	B	A	B
B	B	A	C	C
C	C	D	B	A
D	D	C	D	G
E	E	E	E	F
F	F	F	F	E
G	G	G	G	D

0.52 0.23 0.20 0.05



\mathbb{P}

	1	2	3	4	5	6	7
A	0.72	0.23	0.05	0	0	0	0
B	0.28	0.52	0.20	0	0	0	0
C	...						
D				$\mathbb{P}_{i,j}$			
E							
F							
G							

Computing the Best Fair Policy

- Optimal \mathbb{P}^* is solution of linear program

$$\mathbb{P}^* = \operatorname{argmax}_{\mathbb{P}}$$

s. t.

$$[rel^T \mathbb{P} e]$$

$$1^T \mathbb{P} = 1$$

$$\mathbb{P} 1 = 1$$

$$0 \leq \mathbb{P} \leq 1$$

$$rel_2 g_1^T \mathbb{P} e = rel_1 g_2^T \mathbb{P} e$$

DCG

P is doubly
stochastic

Fairness

Computing π^* from \mathbb{P}^*

Birkhoff-von Neumann decomposition

$$\mathbb{P}^* = \theta_1 P_1 + \dots + \theta_k P_k$$

where $P_1 \dots P_k$ are permutation matrices and $\theta_i \geq 0$ with $\sum_i \theta_i = 1$.

$$\rightarrow \text{Ranking policy } \pi^*(y|x) = \begin{cases} \theta_i & \text{if } (y = P_i) \\ 0 & \text{else} \end{cases}$$

Summary of Method

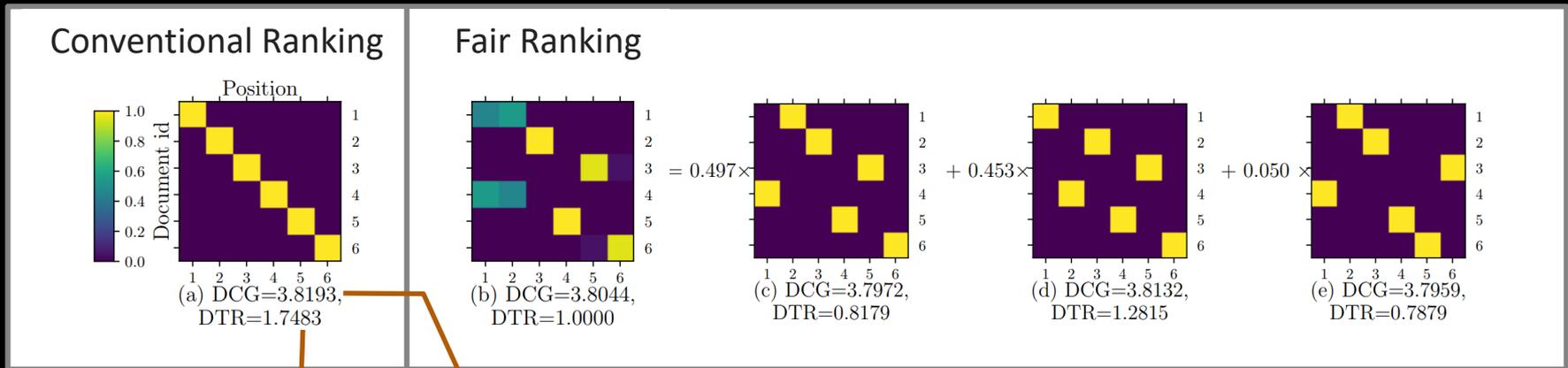
1. Estimate relevances r for query x
2. Define (merit-based) fairness constraint
3. Solve linear program for marginal rank matrix

$$\begin{aligned} \mathbb{P}^* = \operatorname{argmax}_{\mathbb{P}} \quad & [rel^T \mathbb{P} q] \\ \text{s.t.} \quad & \mathbf{1}^T \mathbb{P} = \mathbf{1} \\ & \mathbb{P} \mathbf{1} = \mathbf{1} \\ & 0 \leq \mathbb{P} \leq 1 \\ & \mathbb{P} \text{ is fair} \end{aligned}$$

4. Compute ranking policy π^* from \mathbb{P}^* via Birkhoff-von Neumann
5. Sample ranking y from π^*

Example

- Six items, two groups
- Relevances: $\text{rel}(G_1) = \{82\%, 81\%, 80\%\}$, $\text{rel}(G_2) = \{79\%, 78\%, 77\%\}$



Relative
Unfairness

Quality

Sustainable Platforms

1. Unbiased Estimation of Relevance

- Selection bias correction through IPS [Joachims et al. 2017]
Unbiased learning of deep ranking policies [Agarwal et al. 2019]

2. Fair Treatment of all Platform Participants

- Item fairness through fairness of exposure [Singh & Joachims, 2018]
Fair ranking through Nash-fair division [Saito & Joachims 2022]
Fair policy learning [Singh & Joachims, 2019] [Yadav et al. 2021]

→ 3. Steerable Control of Platform Dynamics

Beyond Microeconomics

Macroeconomic Control of AI Platform
Long-term Sustainability of the Platform

Macro-Metrics: user satisfaction, supplier pool, polarization, etc.
Macro-Interventions: exposure allocation, diversification, novelty, etc.



Microeconomic Optimization of AI Platform
Short-term Utility Maximization of Participants

Micro-Metrics: engagement through clicks, purchases, likes, etc.
Micro-Interventions: ranking, artwork, push-notifications, etc.



Towards Steerable Dynamics

Macroeconomic Control of AI Platforms

Long-term Sustainability of the Platform

Macro-Metrics: user satisfaction, supplier pool size, polarization, discrimination, ...

Macro-Interventions: exposure allocation, diversification, novelty, external regulations, ...

Macro-Interventions

Micro/Macro Abstraction and Interface

Optimal micro-interventions consistent with macro-interventions

Microeconomic Optimization of AI Platforms

Short-term Utility Maximization of Participants

Micro-Metrics: engagement through clicks, purchases, likes, streams, ...

Micro-Interventions: ranking, artwork, push-notifications, upsell, ...

Translating Macro to Micro

Macroeconomic Control of AI Platforms

Weekly/Monthly Metrics

User: Show user TJ at least δ_{TJ} new artists; do not send more than 3 push messages; ...
Item: Show new artist A to at least δ_A users; give items from supplier B at least δ_B exposure; ...

Macro-Interventions

Micro/Macro Abstraction and Interface

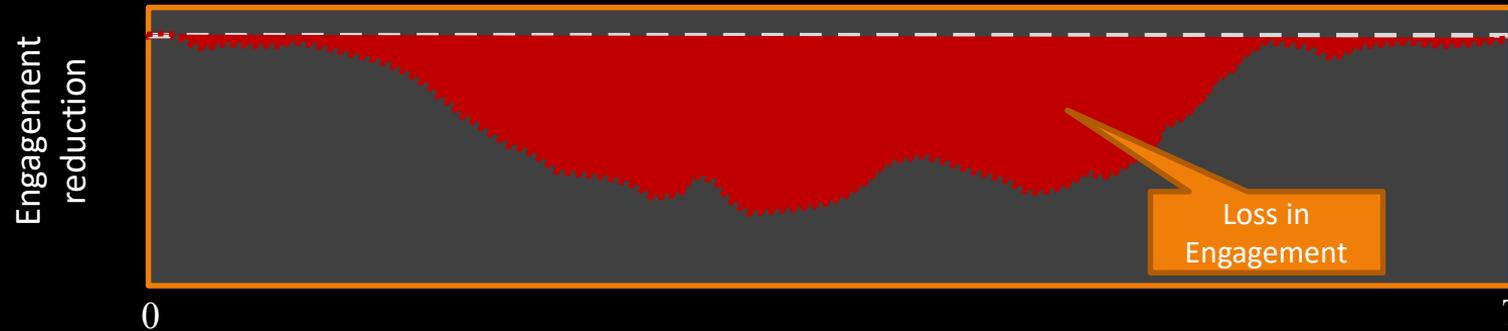
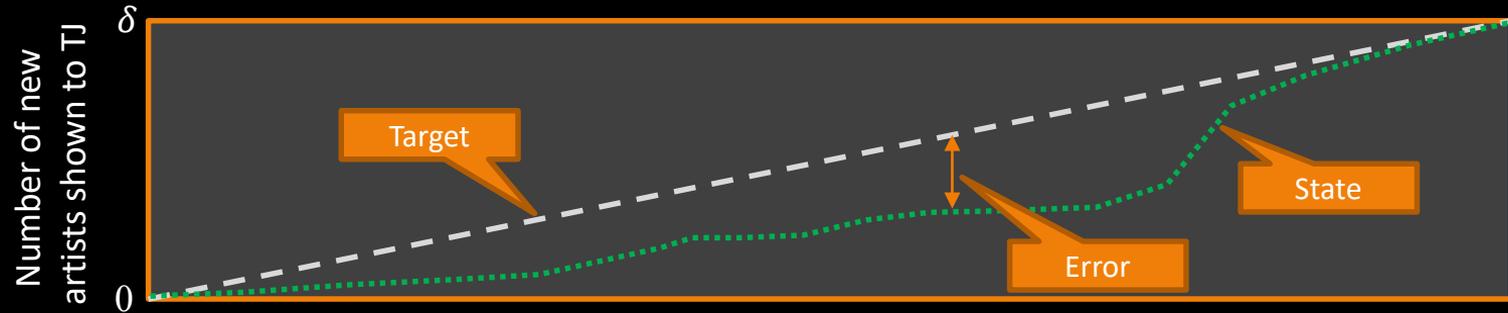
Optimal micro-interventions consistent with macro-interventions

Microeconomic Optimization of AI Platforms

Session Metrics

Micro-Metrics: engagement through clicks, purchases, likes, streams, ...
Micro-Interventions: ranking, artwork, push-notifications, upsell, ...

Reactive Controller



Mo

Tu

We

Th

Fr

Sa

Su

T

P-Controller

- **Group G:**

All artists i that are novel to TJ

- **Control Error:**

$$err(G|t) = \delta \frac{t}{T} - \sum_{i=1}^t expo(G|x_i, y_i)$$

- **Policy:**

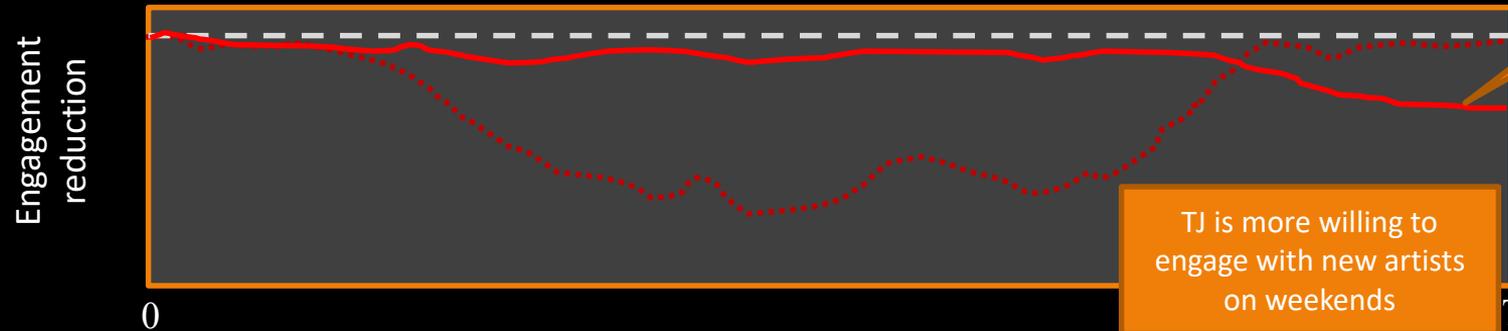
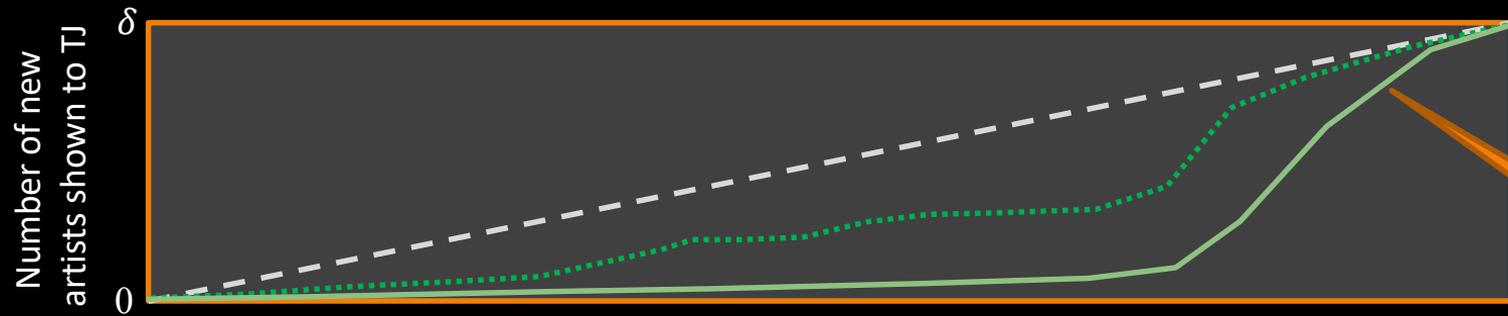
$$\pi(x) \stackrel{\text{def}}{=} \underset{i}{\text{argsort}}[rel(i|x) + \lambda \cdot 1[i \in G] \cdot err(G|t)]$$

Relevance of
item i

Boost
members of G
by $\lambda \cdot error$

Control error
for group
target

Planning Controller



TJ is more willing to engage with new artists on weekends

Planning Controller

Model Predictive Controller

- **Group G:**

All artists i that are novel to TJ

- **Model:**

Sample $S = ((x_1, rel_1), \dots, (x_N, rel_N)) \sim P(S_{t...T})$ as model of which future queries to expect.

- **Policy:**

$$\max_{\mathbb{P}_0, \mathbb{P}_1 \dots \mathbb{P}_N} rel_t^T \mathbb{P}_0 e + \frac{t-T}{N} \sum_{k=1}^N rel_k^T \mathbb{P}_k e$$

DCG of current ranking \mathbb{P}_0

Expected Future DCG

s. t. $\forall \mathbb{P}_i: \mathbb{P}_i$ is doubly stochastic

$$\sum_{i=1}^{t-1} expo(G|x_i, y_i) + G_t^T \mathbb{P}_0 e + \frac{t-T}{N} \sum_{k=1}^N G_k^T \mathbb{P}_k e \geq \delta$$

Past Exposure

Current Exposure

Expected Future Exposure

Target Exposure

Extensions

- Multiple constraints
- Soft constraints
- Computational efficiency

Towards Steerable Dynamics

Macroeconomic Control of AI Platforms

Long-term Sustainability of the Platform

Macro-Metrics: user satisfaction, supplier pool size, polarization, discrimination, ...

Macro-Interventions: exposure allocation, diversification, novelty, external regulations, ...

- Causal Modeling
- Connections to Social Sciences
- Regulatory Policy

Macro-Interventions

Micro/Macro Abstraction and Interface

Optimal micro-interventions consistent with macro-interventions

Control Theory

Microeconomic Optimization of AI Platforms

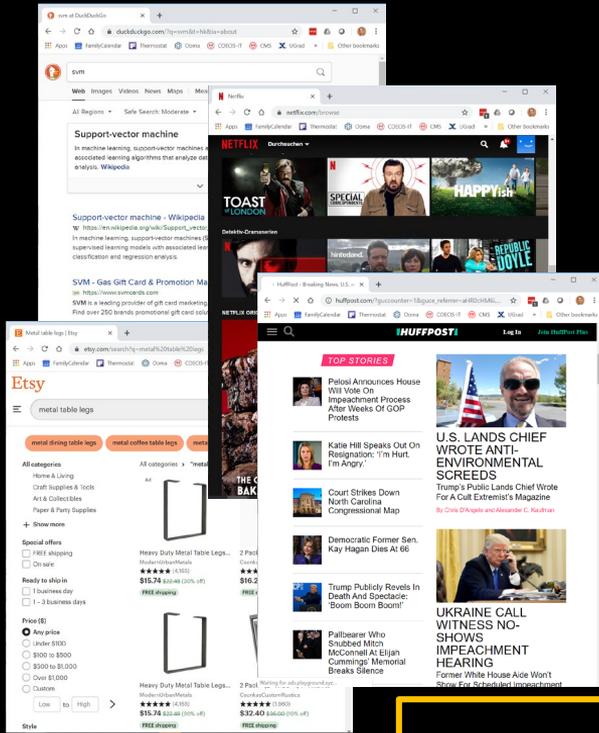
Short-term Utility Maximization of Participants

Micro-Metrics: engagement through clicks, purchases, likes, streams, ...

Micro-Interventions: ranking, artwork, push-notifications, upsell, ...

Research for Sustainable AI Platforms

- Unbiased estimation
- Fairness
- Steerable long-term dynamics
- Transparency
- Privacy



<http://www.joachims.org>



GATES HALL

