Designing AI Systems with Steerable Long-Term Dynamics

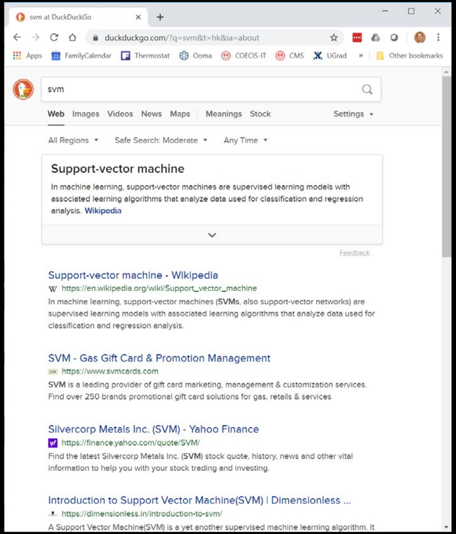
Ashudeep Singh, Adith Swaminathan, Tobias Schnabel, Kiante Brantley, Aman Agarwal, Marco Morik, Sarah Dean, Zhichong Fang, Yuta Saito,
Thorsten Joachims

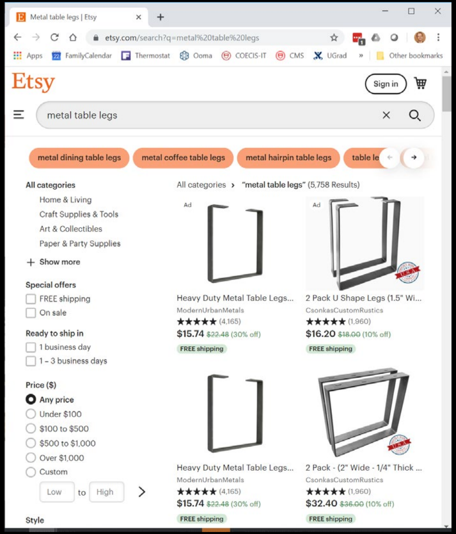
Department of Computer Science, Cornell University
Department of Information Science, Cornell University

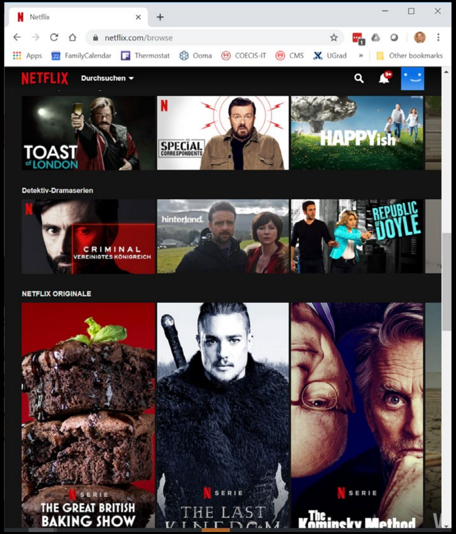
Funded in part through NSF Awards IIS-2008139, IIS-1615706, IIS-1901168.

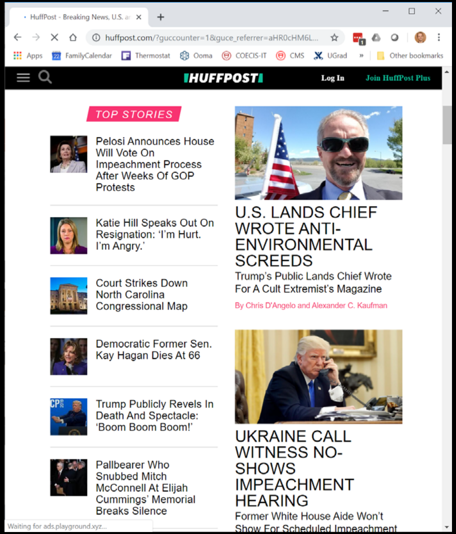
Intelligent Online Systems

Ranking function π that ranks items for context x.

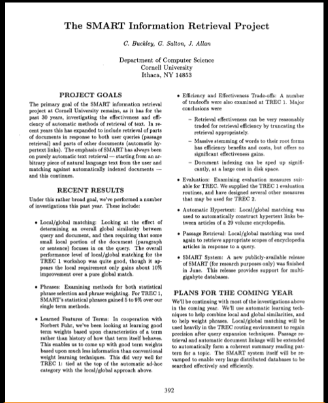






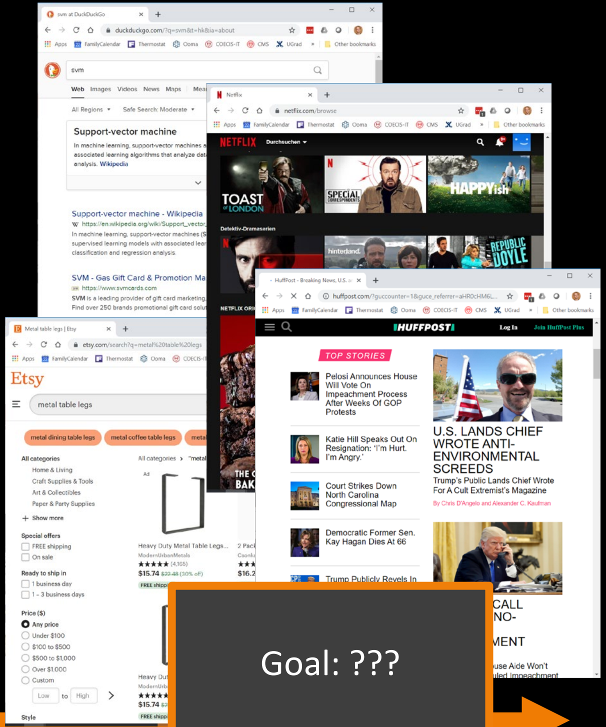


What is the ideal ranking?



Goal: Maximize utility of rankings to the users.

©FastCompany



Goal: ???

1970

1994

2022

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

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

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

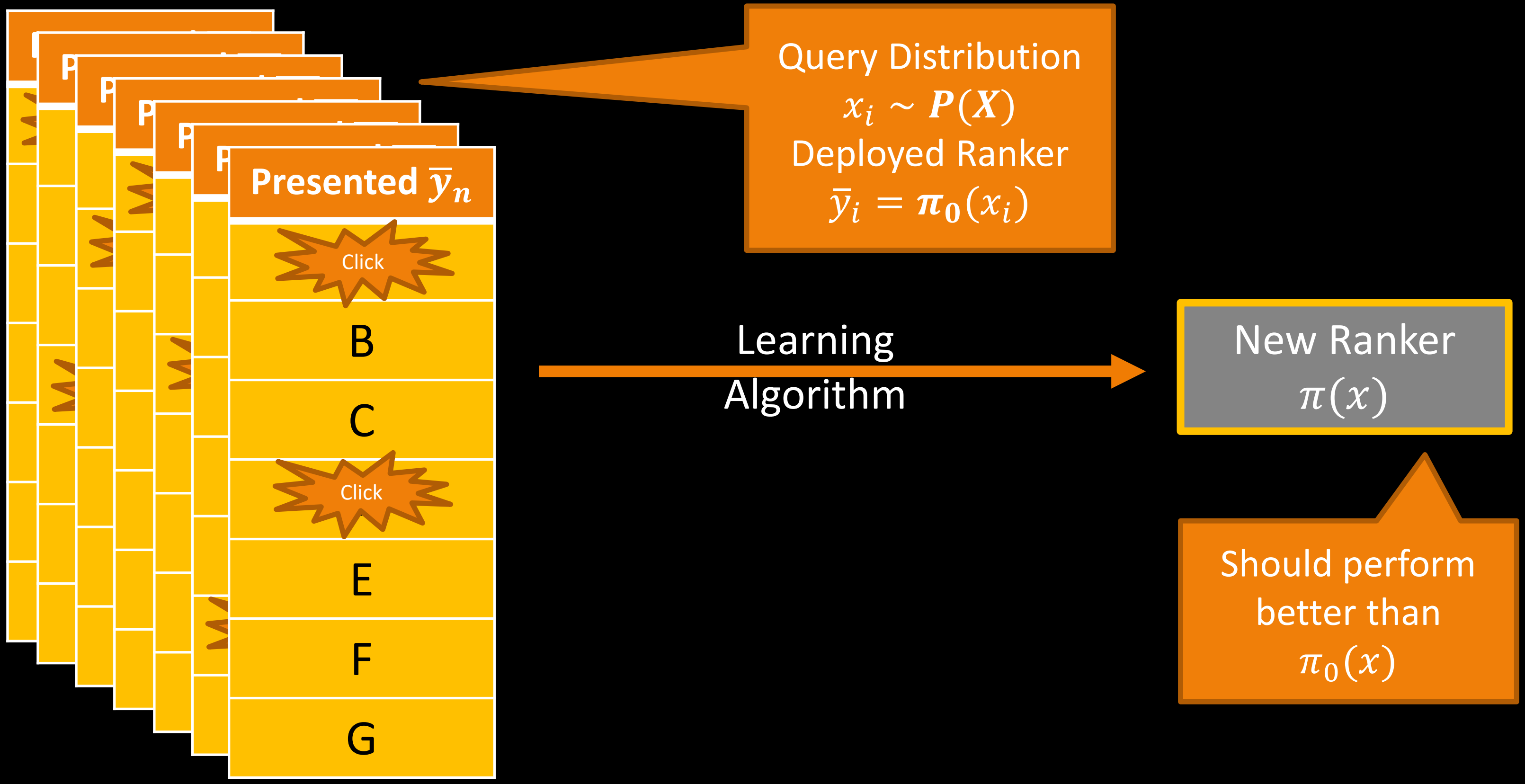
Top News Stories

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
…	…	…

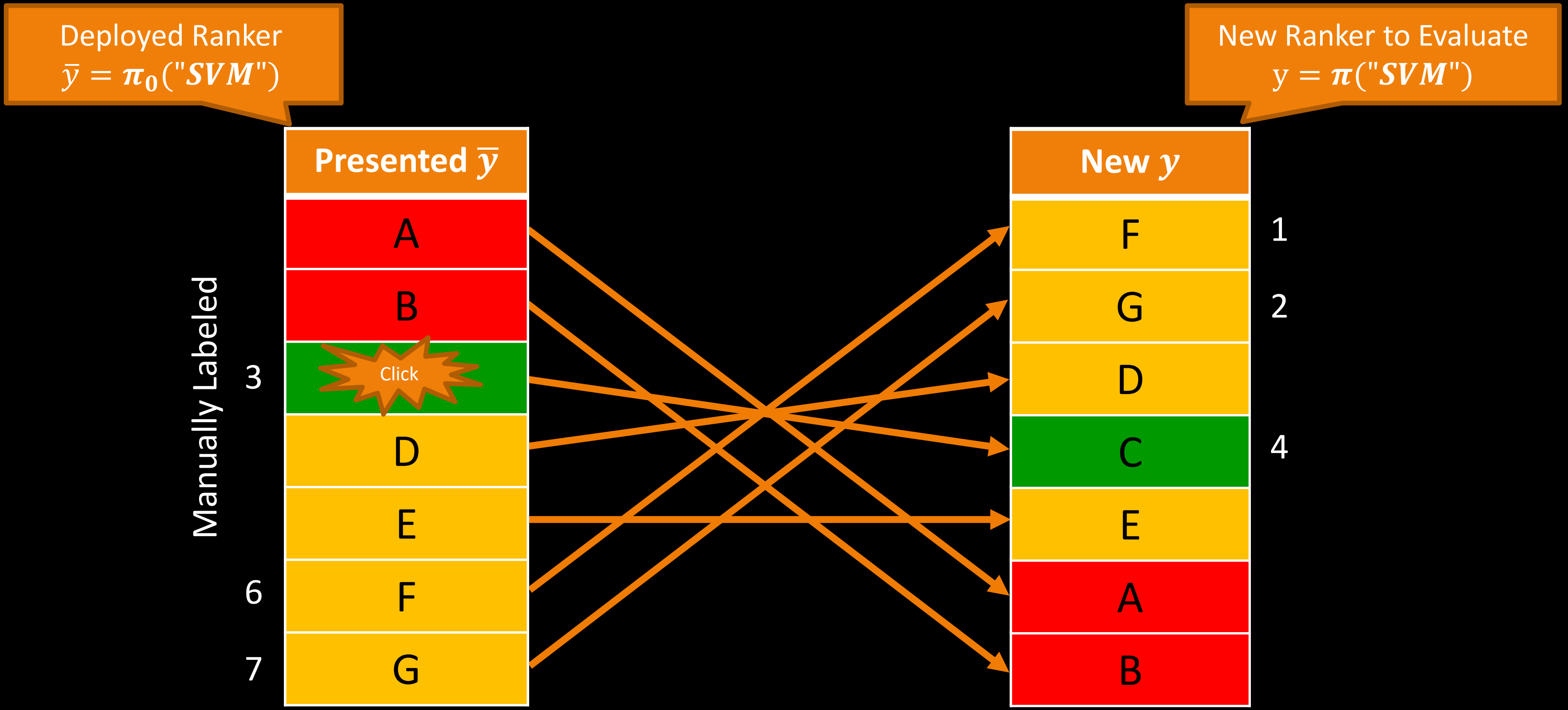
Sustainable Platforms

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

Learning-to-Rank from Clicks



Evaluating Rankings



Evaluation with Missing Judgments

- Loss: $\Delta(x, y|rel)$
  - Relevance labels $\mathrm{rel}_d \in \{0,1\}$
  - This talk: rank of relevant documents

$$\Delta(x, y|rel) = \sum_{\mathrm{d}} rank(d|y) \cdot rel_d$$

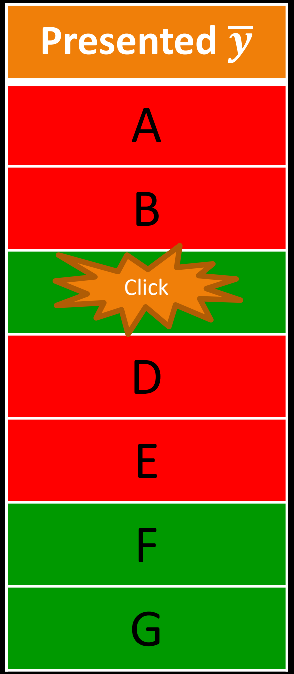
- Assume:
  - Click implies observed and relevant:

$$(c_d = 1) \leftrightarrow (o_d = 1) \wedge \left(\mathrm{rel}_d = 1\right)$$

- Problem:
  - No click can mean not relevant OR not observed

$$(c_d = 0) \leftrightarrow (o_d = 0) \vee (\mathrm{rel}_d = 0)$$

- → Understand observation mechanism



[Wang et al., 2016] [Joachims et al., 2017]

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

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

[Horvitz & Thompson, 1952] [Rubin, 1983] [Zadrozny et al., 2003] [Langford, Li, 2009] [Wang et al., 2016] [Joachims et al., 2017]

ERM for Partial-Information LTR

- Unbiased Empirical Risk:

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

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 $Q(o_d = 1|x, \bar{y}, rel)$?

[Joachims et al., 2017]

Propensity-Weighted SVM Rank

- Data: $\mathrm{D} = \left(x_j, d_j, D_j, q_j\right)^n$

Query Clicked Others Propensity

Optimizes convex upper bound on unbiased IPS risk estimate!

- Training QP:

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

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

[Herbrich at al., 1999] [Joachims et al., 2002] [Joachims et al., 2017]

Position-Based Propensity Model

- Model:

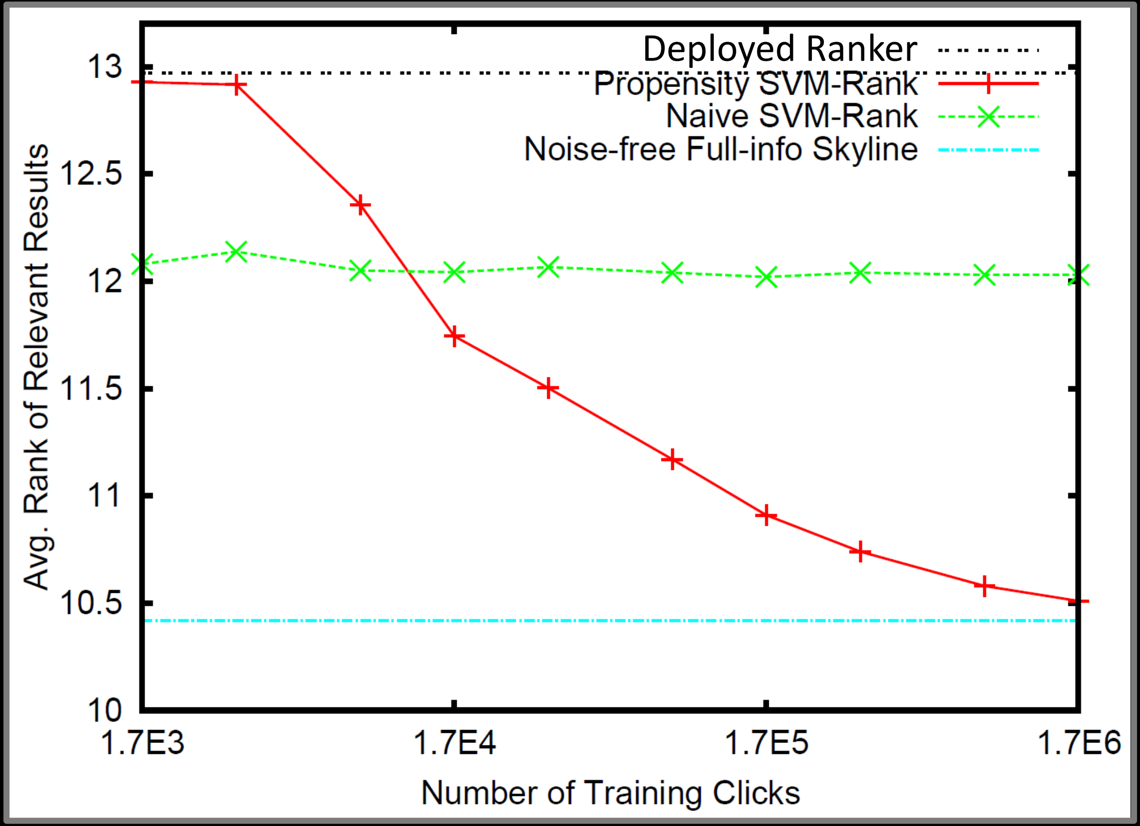
$$P\left(c_d = 1 | rel_d , rank(d|\bar{y})\right) = q_{rank(d|\bar{y})} \cdot [rel_d = 1]$$

- 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

[Richardson et al., 2007] [Chuklin et al., 2015] [Wang et al., 2016]

Ranking Accuracy vs. Training Data



[Joachims et al., 2017]

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

$\neq$

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}
…	…

[Craswell et al. 2008] [Singh & Joachims 2018]

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

[Biega et al. 2018] [Singh & Joachims 2018]

Disparate Exposure Constraint

Group Exposure and Merit

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

Group Fairness Constraint

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

→ Make exposure proportional to relevance

[Singh & Joachims 2018]

Computing the Best Fair Ranking

Goal: Maximize ranking quality while fair to items.

$$y = \quad \operatorname{argmax}_y \quad [DCG(y|x)]$$

$$s.t. \quad \frac{expo(G_0|x,y)}{rel(G_0|x)} = \frac{expo(G_1|x,y)}{rel(G_1|x)}$$

→ Computationally hard and typically infeasible!

[Singh & Joachims 2018]

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

[Singh & Joachims 2018]

Marginal Rank Distribution $\mathbb{P}$

π

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

[Singh & Joachims 2018]

Computing the Best Fair Policy

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

$$\mathbb{P}^* = \operatorname{argmax}_{\mathbb{P}} \quad \left[rel^T \mathbb{P} e\right]$$

$$s.t. \quad \mathbb{1}^T \mathbb{P} = \mathbb{1}$$

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

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

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

DCG

P is doubly stochastic

Fairness

[Singh & Joachims 2018]

Computing π^* from $\mathbb{P}^*$

Birkhoff-von Neumann decomposition

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

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

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

[Singh & Joachims 2018]

Summary of Method

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

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

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

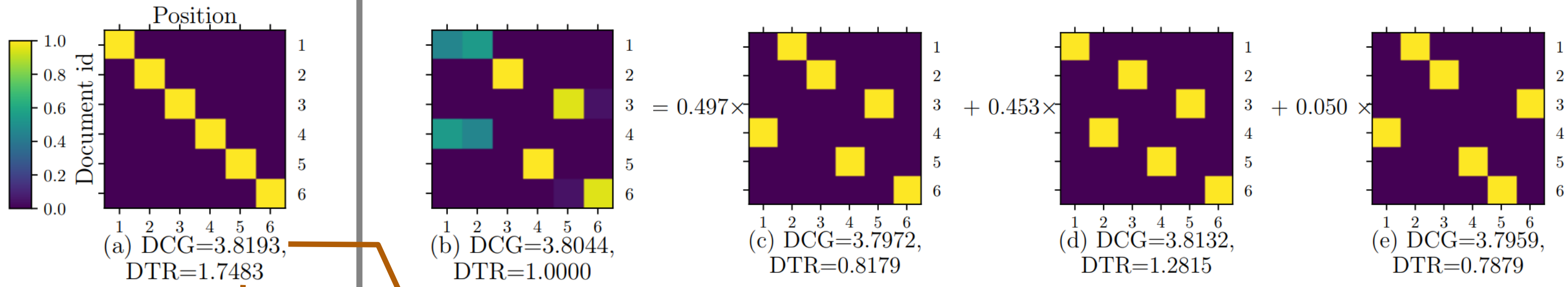
[Singh & Joachims 2018]

Example

- Six items, two groups
- Relevances: $\mathrm{rel}(\mathrm{G}_1) = \{82\%, 81\%, 80\%\}, \mathrm{rel}(G_2) = \{79\%, 78\%, 77\%\}$

Conventional Ranking

Fair Ranking



Relative Unfairness

Quality

[Singh & Joachims 2018]

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

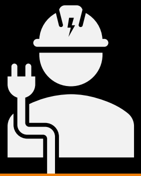
Beyond Microeconomics

Macroeconomic Control of AI Platforr

Long-term Sustainability of the Platform

Macro-Metrics: user satisfaction, supplier pool, polarization, c

Macro-Interventions: exposure allocation, diversification, novelty, e



Microeconomic Optimization of AI Platf

Short-term Utility Maximization of Participant

Micro-Metrics: engagement through clicks, purchases, like

Micro-Interventions: ranking, artwork, push-notification

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

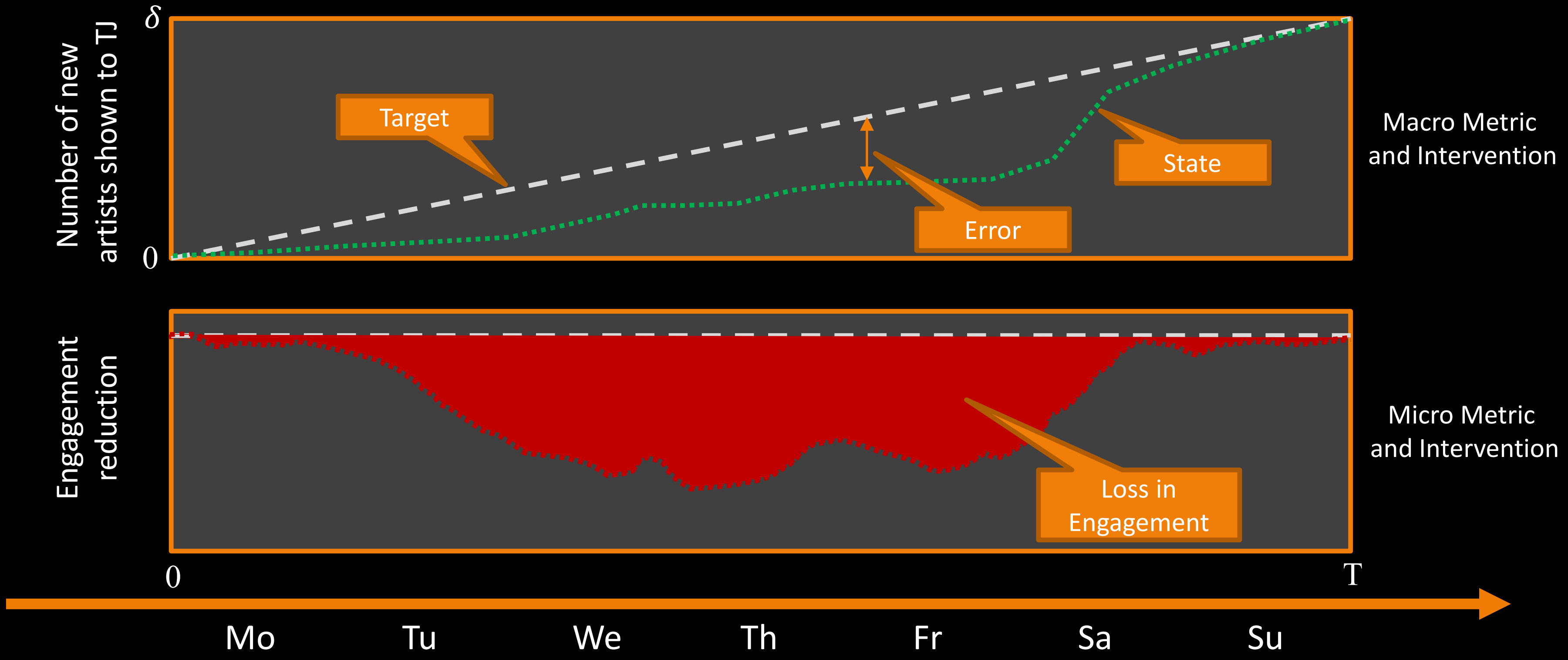
Microeconomic Optimization of AI Platforms

Session Metrics

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

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

Reactive Controller



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:**

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

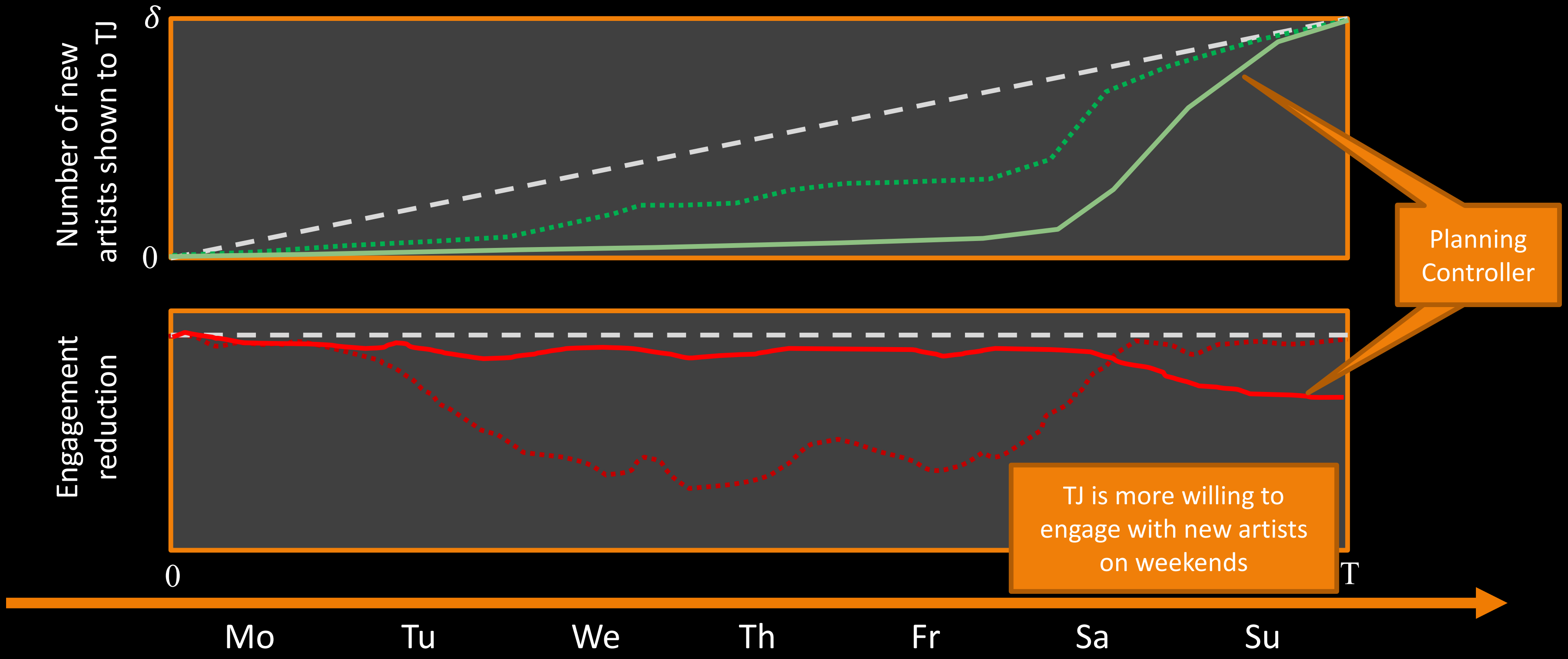
Relevance of item i

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

Control error for group target

[Morik et al, 2020]

Planning Controller



Model Predictive Controller

- **Group G:**

  All artists i that are novel to TJ
- **Model:**

  Sample $S = \left((x_1, rel_1), \ldots, (x_N, rel_N)\right) \sim P(S_{t \ldots T})$ as model of which future queries to expect.
- **Policy:**

DCG of current ranking $\mathbb{P}_0$

Expected Future DCG

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

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

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

Target Exposure

Past Exposure

Current Exposure

Expected Future Exposure

Extensions

- Multiple constraints
- Soft constraints
- Computational efficiency

Towards Steerable Dynamics

- Causal Modeling
- Connections to Social Sciences
- Regulatory Policy

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

Control Theory

Optimal micro-interventions consistent with macro-interventions

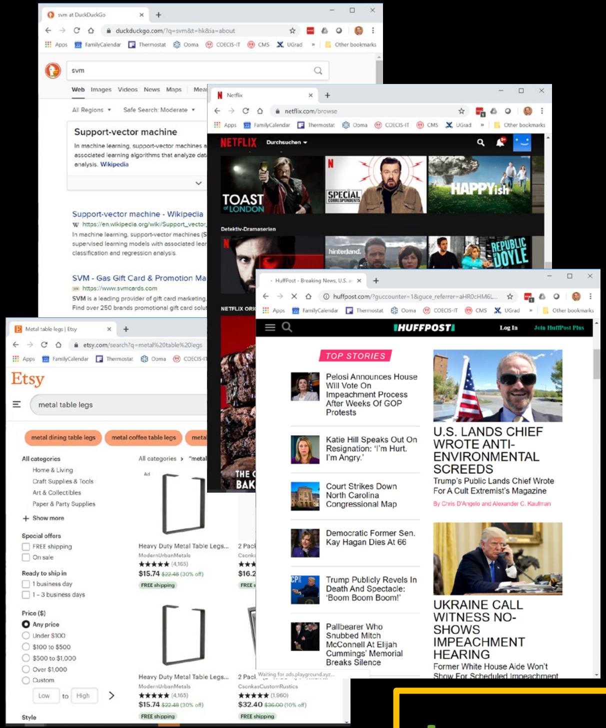
Microeconomic Optimization of AI Platforms

Short-term Utility Maximization of Participants

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

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

Research for Sustainable AI Platforms



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

http://www.joachims.org


GATES HALL