### Learning Socially Optimal Information Systems from Egoistic Users

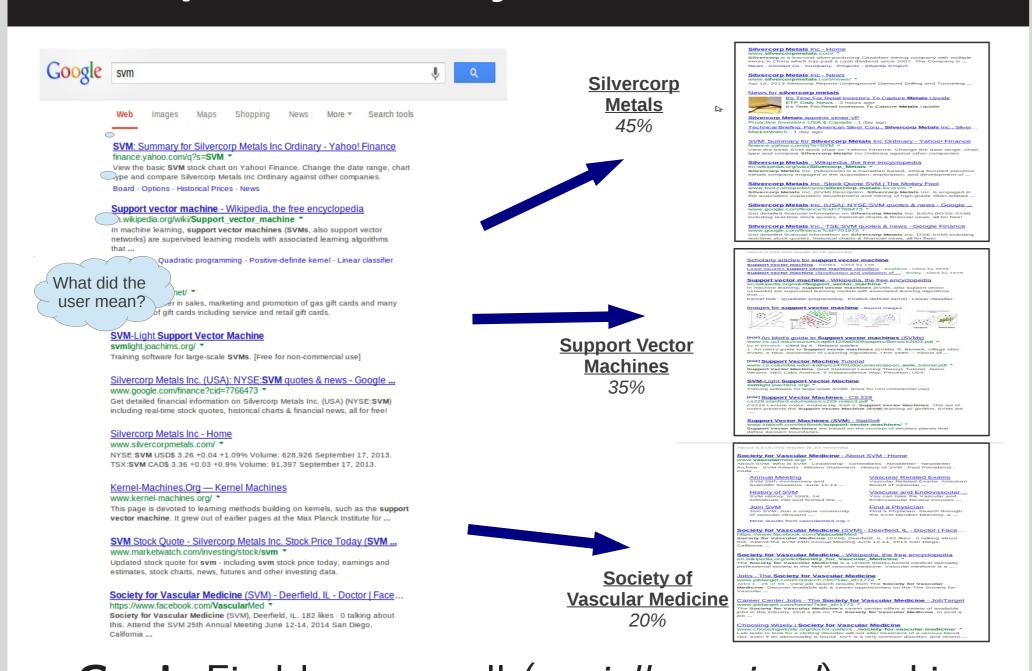
Karthik Raman, Thorsten Joachims ({karthik,tj}@cs.cornell.edu)



#### Problem Overview

Learn robust systems that collectively satisfy a population of diverse users from user feedback.

#### **Example: Diversity in Search**



**Goal:** Find best overall (socially optimal) ranking.

#### User Feedback

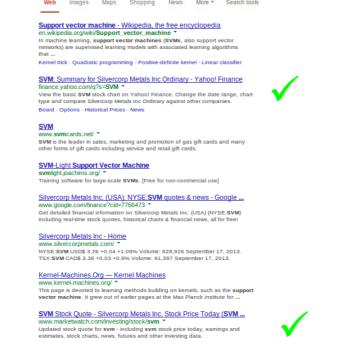
#### **Egoistic** user feedback

- User's choice not social.
- Conflicting choices.

Weak, noisy & biased feedback.

- Cannot regard as cardinal labels.
- Treat as preferences.

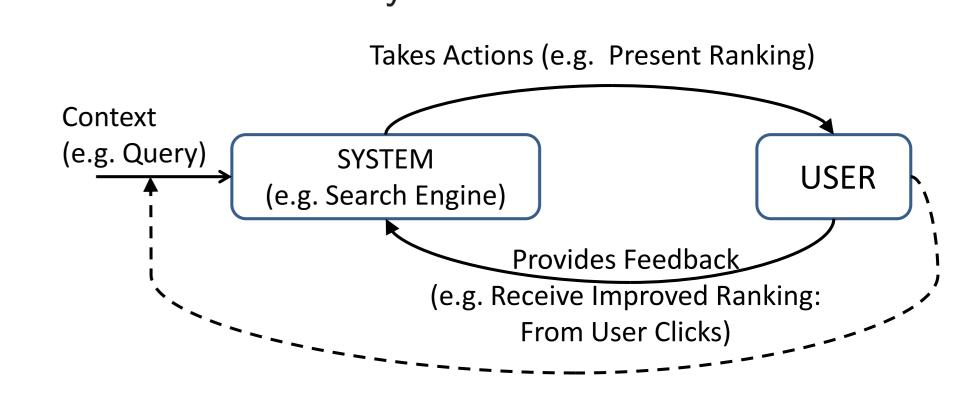
# Web Images Maps Shopping News Monin Search bools Support vector machine - Wildpedia. The fine encyclopedia en whigelist on yells (Segment vector machine) encyclopedia grow in the control of the contro



## Web Images Maps Shopping News More \* Search tools SYM: Summary for Silvercorp Metals Inc Ordinary - Yahoo! Finance finance, yahoo.com/q?s=SVM \* View the basic SVM stock chart on Yahoo! Finance. Change the date range, chart type and compare silvercorp Metals Inc Ordinary against other companies. Board - Options - Historical Prices - News Support vector machine - Wikipedia, the free encyclopedia envikipedia org/wiki/Support, vector machine \* In machine learning, support vector machine \* In machine learning nodels with associated learning algorithms that ... Kemel trick - Quadratic programming - Positive-definite kemel - Linear classifier SVM www.svmcards.net/ \* SVM Light Support Vector Machine svmlight\_loachins.org/ \* Training software for large-scale SVMs. [Free for non-commercial use] Silvercorp Metals Inc. (USA): NYSE-SVM quotes & news - Google ... www.google.com/mance/exicle/Tr66473 \* Get detailed francial information on silvercorp Metals inc. (USA) (NYSE-SVM) including real-time stock quotes, historical charts & financial news, all for free! Silvercorp Metals Inc. - Home www.silvercorpmetals.com/ \* NYSE-SWM LOSS 3.26 + 0.03 + 0.99% Volume: 93,397 September 17, 2013. TSX:SVM CAOS 3.36 + 0.03 + 0.99% Volume: 93,397 September 17, 2013. TSX:SVM CAOS 3.36 + 0.03 + 0.99% Volume: 93,397 September 17, 2013. TSX:SVM Stock Quote - Silvercorp Metals Inc. Stock Price Today (SVM ... www.marketwatch.com/firvesting/stock/svm \* Updated stock quote for svm - including svm stock price today, earnings and estimates, stock charts, news, futures and other investing data. Kernel-Machines.Org — Kernel Machines www.kernel-machines.org \*

#### Learning from User Preferences

- ► Builds on **Coactive Learning** [SJ12,RJSS13].
- ► Given context **x**, predict object **y**.
- ► Goal: Optimize social utility  $U(x, y) = E[U_i(x, y)]$ .
  - $\triangleright$  **U**<sub>i</sub> is **personal** utility of user type **i** (w.p.  $\mathbf{p}_i$ ).
- ▶ User preferences: Feedback tends to improves personal utility:  $U_i(x, \bar{y}) \ge_{\alpha, \delta} U_i(x, y)$ .
  - ▶ Not social utility.



#### Modeling Utility: Submodularity

- Model personal utility of users as submodular in individual elements.
- ▶ **Diminishing returns:** Marginal benefit diminishes.
- ► Example: Coverage Function
- ▶ Given ranking/set  $y=(d_{i_1},\ldots,d_{i_n})$  and position-discount factors  $\gamma_1 \geq \gamma_2 \geq \ldots \geq \gamma_n \geq 0$ , aggregate features using submodular function F:

$$\phi_{\mathsf{F}}^{\mathsf{j}}(\mathsf{x},\mathsf{y}) = \mathsf{F}(\gamma_1 \phi^{\mathsf{j}}(\mathsf{x},\mathsf{d}_{\mathsf{i}_1}), \gamma_2 \phi^{\mathsf{j}}(\mathsf{x},\mathsf{d}_{\mathsf{i}_2}), \dots, \gamma_{\mathsf{n}} \phi^{\mathsf{j}}(\mathsf{x},\mathsf{d}_{\mathsf{i}_{\mathsf{n}}}))$$

- $\triangleright \phi^{j}(x, d_{i})$  is  $j^{th}$  feature of  $d_{i}$ .
- Model personal utility as linear in submodular aggregate:  $U_i(x, y) = w_{*,i}^T \phi_F(x, y)$ 
  - ▶ Submodular aggregation leads to diversity.
- ▶ Computing ranking ≈ Submodular maximization
- ► Use simple, efficient greedy algorithm.
- ► Approximation guarantee of  $\frac{1}{2}$  (under partition matroid constraint).
- Example of Diversity:

Doc	Words		
$d_1$	ma:3 le:3		
$d_2$	ma:5 le:2		
$d_3$	ma:2 le:5		
d <sub>4</sub>	me:3 si:5		
<b>d</b> <sub>5</sub>	me:6 si:2		
de	me:3 si:1		

Word	Weight	
machine	5	
learning	7	
metal	4	
silver	6	

Posn	Doc	ma	le	me	si
1	$d_3$	2	5	0	0
2	d <sub>4</sub>	0	0	3	5
3	$d_2$		2	0	0
MAX of Col		5	5	3	5

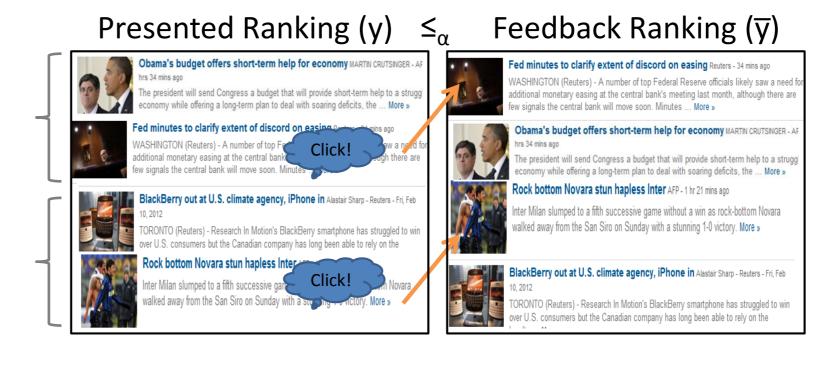
Doc	Marginal Benefit			
	lter1	lter2	Iter3	
$d_1$	3*5 + 3*7	(3-2)*5	(3-2)*5	
$d_2$	5*5 + 2*7	(5-2)*5	(5-2)*5	
$d_3$	2*5 + 5*7	-	_	
d <sub>4</sub>	3*4 + 5*6	3*4 + 5*6	_	
$d_5$	6*4 + 2*6	6*4 + 2*6	(6-3)*4	
$d_6$	3*4 + 1*6	3*4 + 1*6	0	

#### Social Perceptron for Ranking

- 1. Initialize weight vector  $\mathbf{w_1} \leftarrow \mathbf{0}$ .
- 2. Given context  $\mathbf{x_t}$  present user with  $\mathbf{y_t} \leftarrow \operatorname{argmax_v} \mathbf{w_t}^{\top} \phi(\mathbf{x_t}, \mathbf{y})$ .
- 3. Observe user clicks  $\mathcal{D}$ .
- 4. Construct preference feedback:  $\bar{y}_t \leftarrow PairedFeedback(y_t, \mathcal{D})$ .
- 5. Update weight vector:

$$\bar{\mathbf{w}}_{t+1} \leftarrow \mathbf{w}_t + \phi(\mathbf{x}_t, \bar{\mathbf{y}}_t) - \phi(\mathbf{x}_t, \mathbf{y}_t)$$

- 6. Clip to be non-negative:
  - $\mathbf{w}_{t+1}^{\mathsf{J}} \leftarrow \max(\bar{\mathbf{w}}_{t+1}^{\mathsf{J}}, \mathbf{0})$
- 7. Repeat from step 2.



PairedFeedback: Form pairs and swap if only lower element is clicked.

Referred to as the **SoPer-R** algorithm.

Also provide an algorithm for learning diverse sets called the **SoPer-S** algorithm.

See paper for more details

#### Theoretical Analysis

#### $\alpha_i$ , $\delta_i$ -Informative Feedback:

Characterize feedback  $\bar{\mathbf{y}}$  in terms of  $\alpha_{\mathbf{i}}, \delta_{\mathbf{i}}, \boldsymbol{\xi}$  as:

$$E_{\bar{y}}[U_{i}(x,\bar{y})] \geq (1+\delta_{i})U_{i}(x,y) \\
+\alpha_{i}(U_{i}(x,y^{*,i})-U_{i}(x,y))-\bar{\xi}$$

- ightharpoonup where  $\mathbf{y}^{*,i}$  is optimal for user  $\mathbf{i}$
- and y is the presented object.Note that this is a
- characterization (not an assumption).
- Does not assume anything about social utility.
- Used to prove regret bounds.

**Regret:** Define the regret after **T** iterations as:

$$\frac{1}{T} \sum_{t=1}^{T} (U(x_t, y_t^*) - E[U(x_t, y_t)]).$$

Note: In terms of social utility and social optimal.

#### Regret Bound

If 
$$\delta_i \geq \left(\Gamma_F \cdot \frac{1-p_i}{p_i}\right)$$
, average regret of the SoPer-R is:

$$\leq \frac{1}{\eta \mathsf{T}} \sum_{\mathsf{t}=0}^{\mathsf{T}-1} \mathsf{E}_{\mathsf{i}}[\mathsf{p}_{\mathsf{i}}\bar{\xi}_{\mathsf{t}}] + \frac{\mathsf{R}\|\mathsf{w}_{*}\|}{2\eta} \\ + \frac{\sqrt{15}\mathsf{R}\|\mathsf{w}_{*}\|}{\eta\sqrt{2\mathsf{T}}} \\ \text{with } \eta = \mathsf{min}_{\mathsf{i}} \ \mathsf{p}_{\mathsf{i}}\alpha_{\mathsf{i}}.$$

#### **Understanding the bound:**

- ▶ Does not depend on number of dimensions only radius of ball R.
- ightharpoonup Decays gracefully with weak feedback:  $lpha_i$ s.
- Need not converge to optimal (due to NP-hardness of submodular maximization).
- Bound is loose as solution improves.

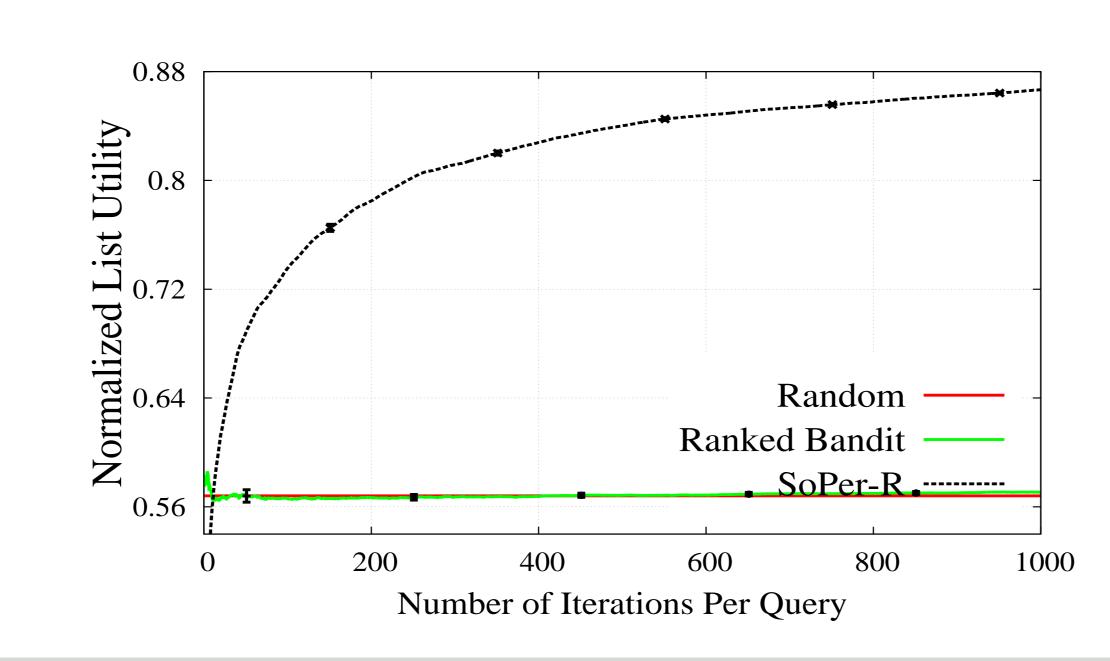
Similar bound for **SoPer-S** algorithm as well.

#### **Experimental Results**

- ➤ Offline experiments on standard **TREC 6-8 Interactive** search diversification dataset.
  - ▶ Queries have 7-56 user types with binary relevance labels.
- Simulated user behavior: Scan rankings top to bottom. Click on first document relevant to them (with some error).
- ► Utility: Normalized DCG-Coverage function upto rank 5.

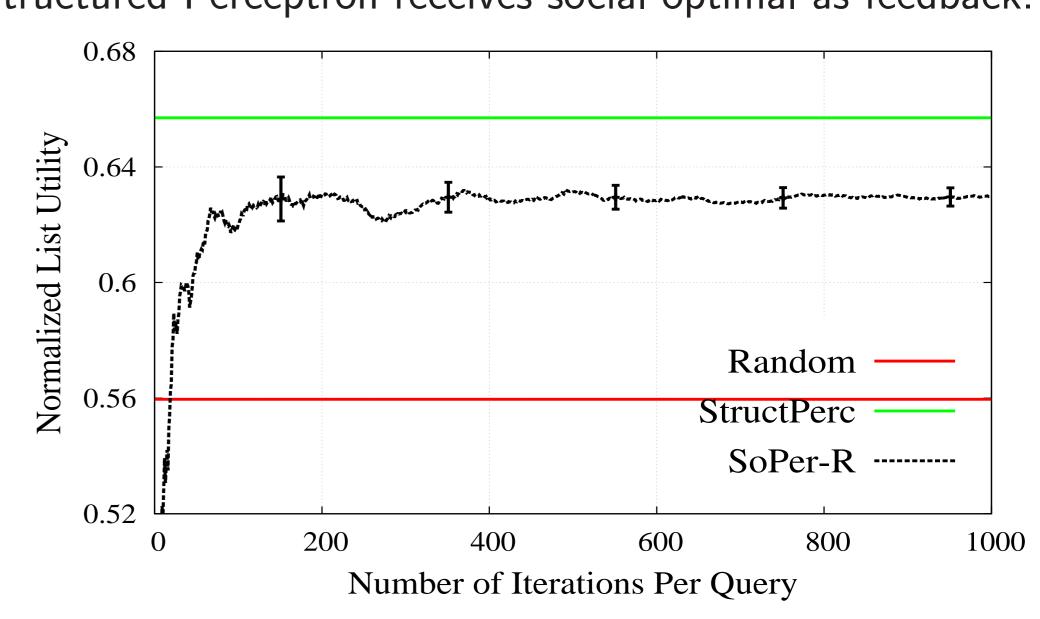
#### Single Query Diversification:

- ► Learning to diversify for single query.
- ► Compare with RankedBandit (Array of coupled MABs).



#### Cross-Query Diversification:

- Learning to diversify given **any** query.
- Structured Perceptron receives social-optimal as feedback.



- ► First method to diversify across queries from preferences.
- ► **Robust** to model mis-specification.

TrueSocialF	SoPer-R (Varying Submodular Function)			Rand
	MAX	SQRT	LIN	
MAX	$.630 \pm .007$	$.620\pm.006$	$.618\pm.006$	$.557\pm.006$
SQRT	$.656\pm.007$	$.654\pm.007$	$.684\pm.006$	$.610\pm.007$
LIN	$.500\pm.006$	$.504\pm.006$	$.566\pm.007$	$\textbf{.474} \pm \textbf{.007}$

► **Robust** to feedback noise (.631 vs .630).