

Ranking Methods in Machine Learning

Example 1: Recommendation Systems

The screenshot shows the Amazon.com homepage for user Shivani Agarwal. The page features a navigation bar with the Amazon logo, user name, and links for Today's Deals, Gifts & Wish Lists, and Gift Cards. Below the navigation bar is a search bar and a 'Shop All Departments' button. The main content area is titled 'Today's Recommendations For You' and displays a carousel of four books. Each book is accompanied by its cover image, title, author, format, price, and a star rating. Navigation arrows are visible on the left and right sides of the carousel. At the bottom of the page, there is a button labeled 'Improve Your Recommendations'.

Amazon.com: Recommended for You

amazon.com Hello, Shivani Agarwal. We have [recommendations](#) for you. (Not Shivani?) FREE 2-Day Shipping: See details

Shivani's Amazon.com | Today's Deals | Gifts & Wish Lists | Gift Cards Your Account | Help

Shop All Departments Search All Departments GO Cart Wish List

Your Amazon.com | Your Browsing History | Recommended For You | Rate These Items | Improve Your Recommendations | Your Profile | Your Communities | Learn More

Shivani, Welcome to Your Amazon.com (If you're not Shivani Agarwal, click here.)

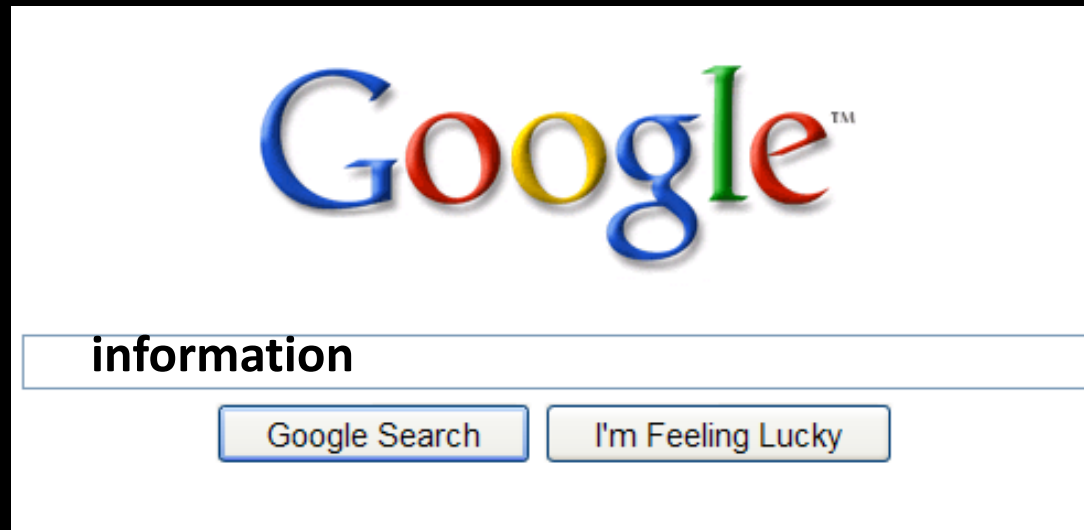
Today's Recommendations For You

Here's a daily sample of items recommended for you. Click here to [see all recommendations](#). Page 1 of 44

Book Title	Author	Format	Price	Rating	Action
Extremal Graph Theory	Béla Bollobás	Paperback	\$20.42	★★★★★	Fix this recommendation
Introduction to Modern Cryptography	Jonathan Katz, Yehuda Lindell	Hardcover	\$68.39	★★★★★ (4)	Fix this recommendation
The Laplacian on a Riemannian...	Steven Rosenberg	Paperback	\$38.70	★★★★★ (3)	Fix this recommendation
Basic Probability Theory (Dover...)	Robert B. Ash	Paperback	\$13.57	★★★★★ (4)	Fix this recommendation

[Improve Your Recommendations](#)

Example 2: Information Retrieval



Example 2: Information Retrieval

information - Google Search - Windows Internet Explorer

http://www.google.com/#hl=en&source=hp&q=information&rlz=1W1FUJB_en&aq=f&aqi=g10&aqi=&soq=&fp=18ec2db39eb50b9d

File Edit View Favorites Tools Help

Google information Search Share Sidewiki Bookmarks Check Translate AutoFill information

Information - Google Search

Web Images Videos Maps News Shopping Gmail more

Google information Search Advanced Search


Web Show options... Results 1 - 10 of about 2,290,000,000 for information [definition]. (0.19 seconds)

Information - Wikipedia, the free encyclopedia
Information as a concept has many meanings, from everyday usage to technical settings. The concept of **information** is closely related to notions of ...
[Etymology](#) - [As sensory input](#) - [As an influence which leads to ...](#)
en.wikipedia.org/wiki/Information - [Cached](#) - [Similar](#)

Information theory - Wikipedia, the free encyclopedia
Information theory is a branch of applied mathematics and electrical engineering involving the quantification of **information**. ...
en.wikipedia.org/wiki/Information_theory - [Cached](#) - [Similar](#)

Information Please
Infoplease.com, a free, authoritative, and respected reference for Internet users, provides a comprehensive encyclopedia, almanac, atlas, dictionary, ...
[Countries](#) - [United States](#) - [This Day In History](#) - [Biography](#)
www.infoplease.com/ - [Cached](#) - [Similar](#)

Local business results for information near Allston, MA - Change location



A **Federal Reserve Bank: General Information**
www.bos.frb.org - [\(617\) 973-3000](#) - [More](#)

B **Dana-Farber Cancer Institute**
www.dana-farber.org - [\(617\) 632-3000](#) - [95 reviews](#)

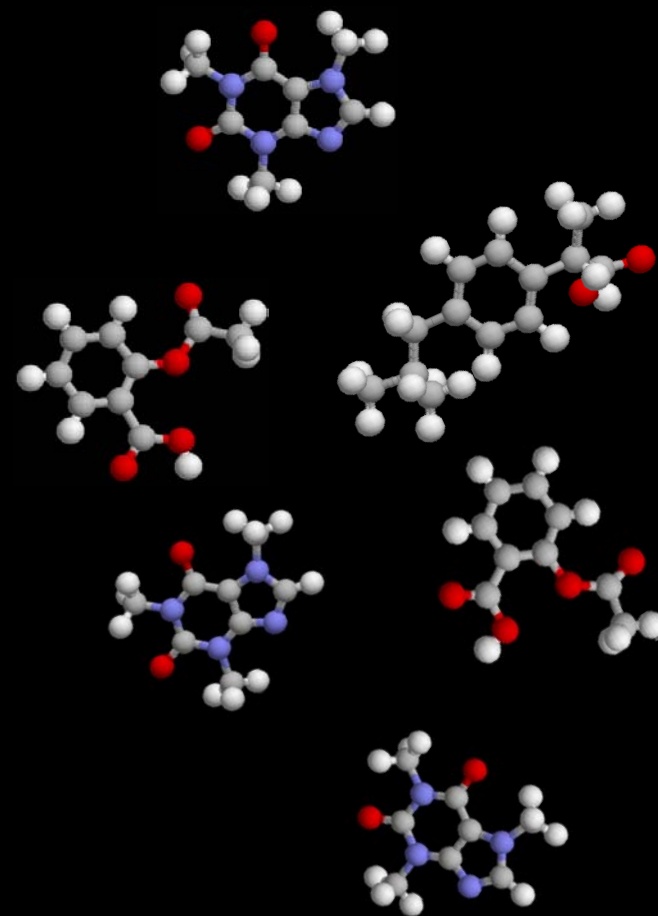
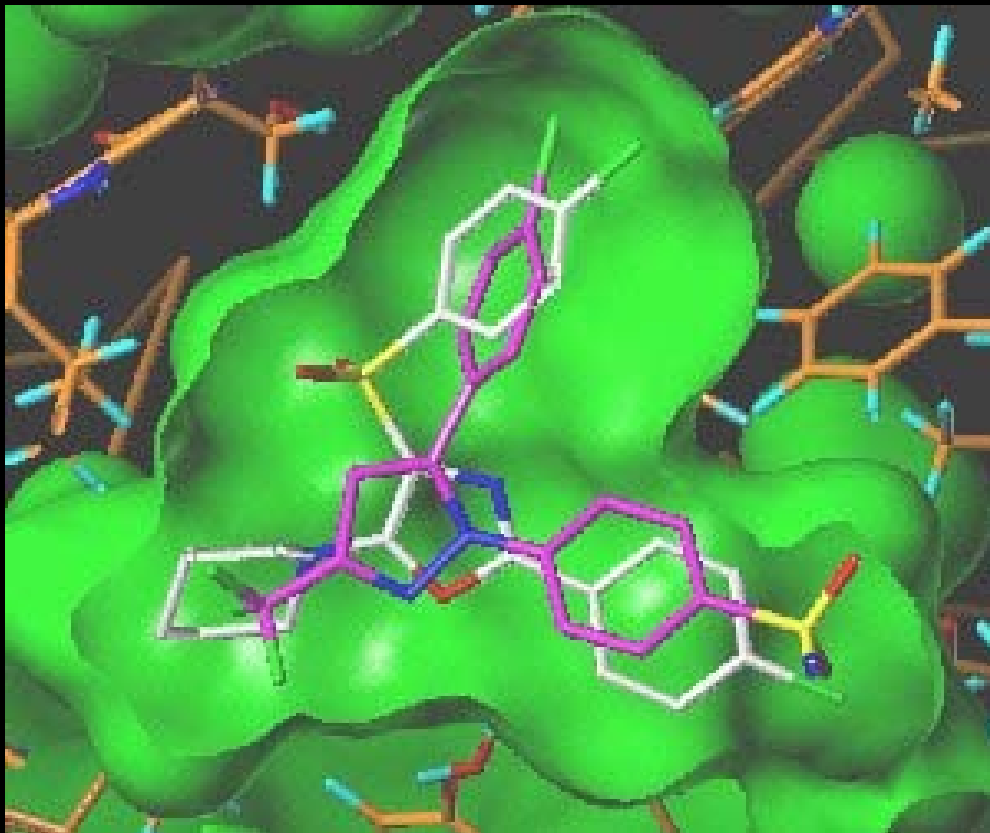
Sponsored Links

Looking For Information?
Find The Info You're Looking For With Google. Make It Your Homepage!
Google.com/Homepage

Information at Amazon
Low Prices on **Information**
Free 2-Day Shipping w/ Amazon Prime
www.Amazon.com/Books

[See your ad here »](#)

Example 3: Drug Discovery



Problem: Millions of structures in a chemical library.
How do we identify the most promising ones?

Example 4: Bioinformatics

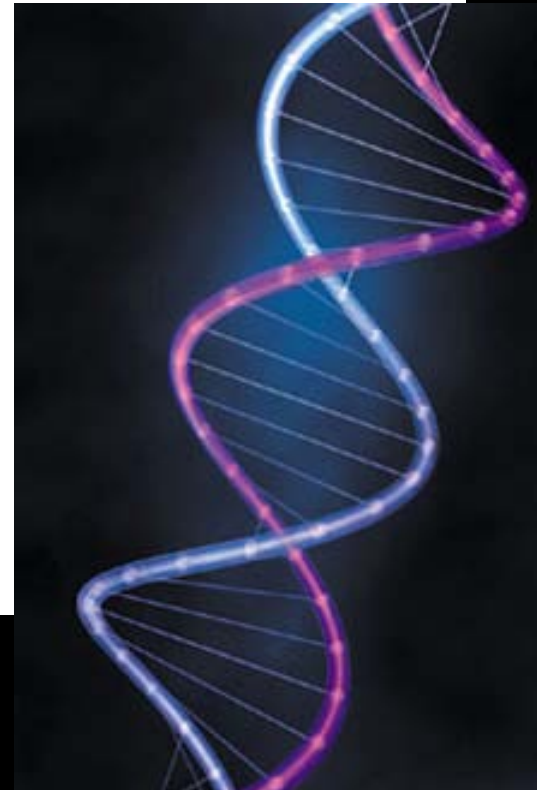


Searching for genetic determinants in the new millennium

N.J. Risch

Human genetics is now at a critical juncture. The molecular methods used successfully to identify the genes underlying rare mendelian syndromes are failing to find the numerous genes causing more common, familial, non-mendelian diseases . . .

Nature **405**:847–856, 2000



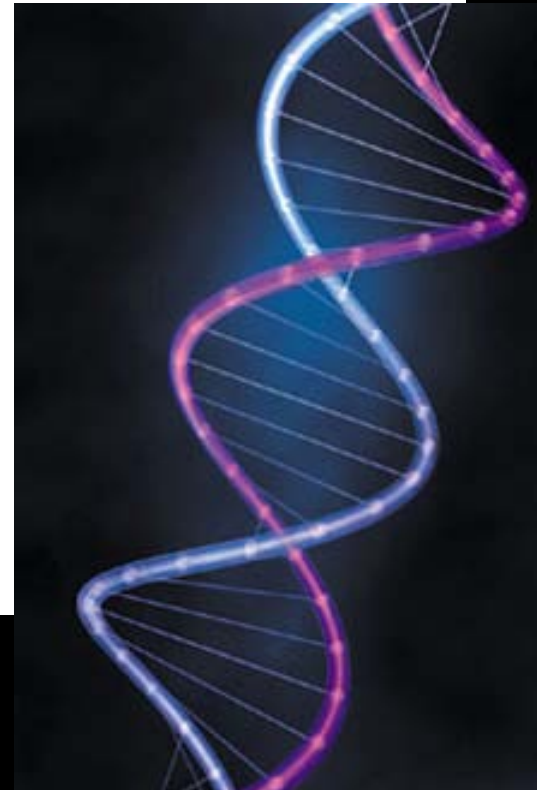
Example 4: Bioinformatics



Searching for genetic determinants in the new millennium

N.J. Risch

With the human genome sequence nearing completion, new opportunities are being presented for unravelling the complex genetic basis of nonmendelian disorders based on large-scale genomewide studies . . .



Nature **405**:847–856, 2000

Types of Ranking Problems

Instance Ranking

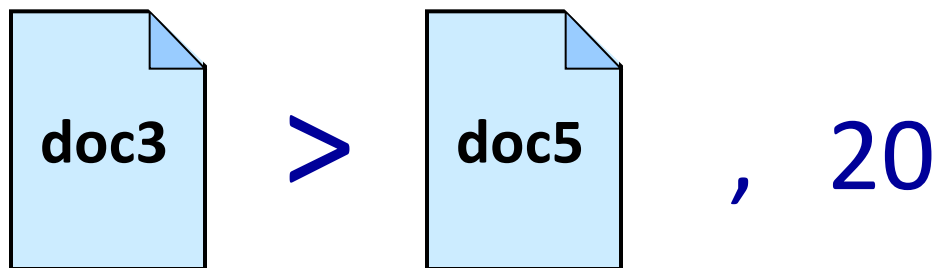
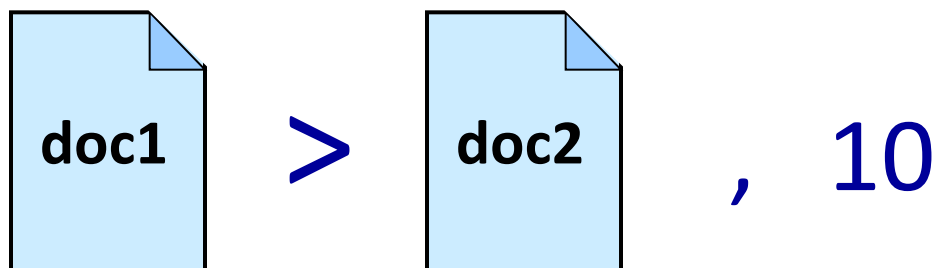
Label Ranking

Subset Ranking

Rank Aggregation

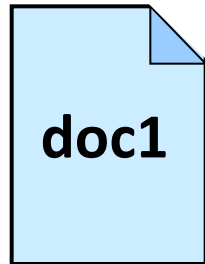
?

Instance Ranking

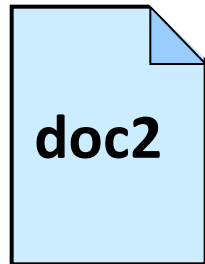


...

Label Ranking



sports > politics
health > money
...

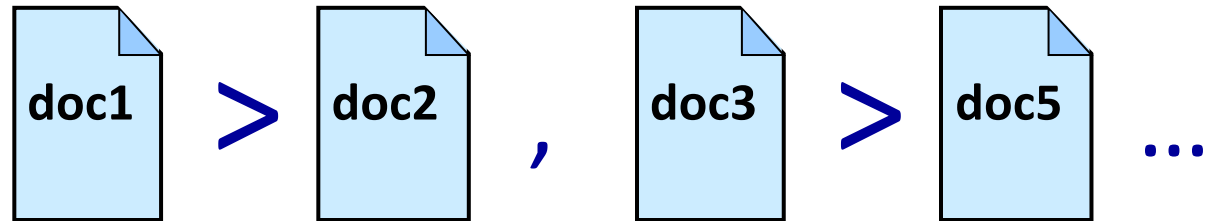


science > sports
money > politics
...

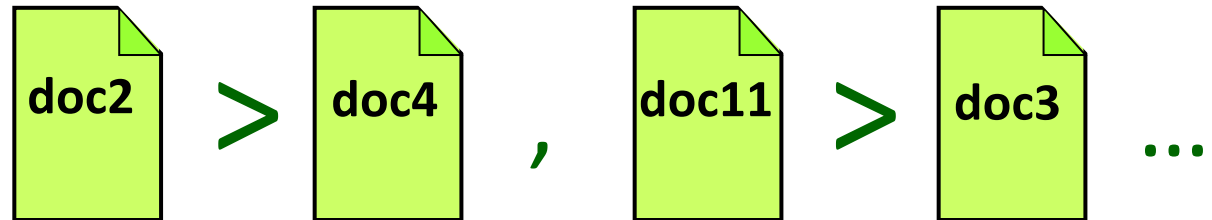
...

Subset Ranking

query 1

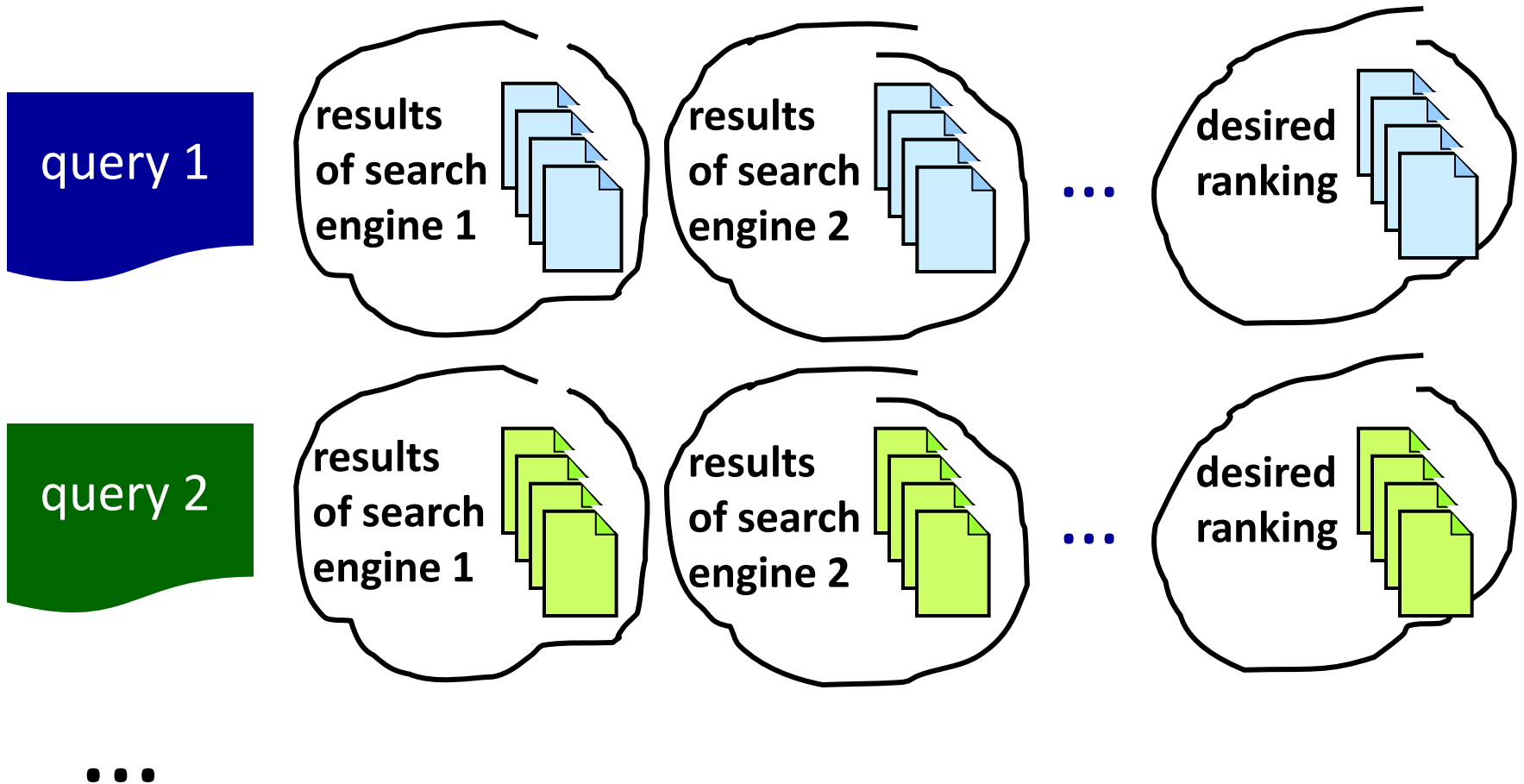


query 2



...

Rank Aggregation



Types of Ranking Problems

Instance Ranking

Label Ranking

Subset Ranking

Rank Aggregation

?

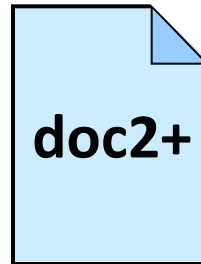
This tutorial

A diagram illustrating the focus of the tutorial. It features a vertical list of ranking problem types: Instance Ranking, Label Ranking, Subset Ranking, Rank Aggregation, and a question mark. Two orange hand-drawn arrows originate from the 'Instance Ranking' and 'Subset Ranking' items and converge on the text 'This tutorial' located at the bottom right of the slide.

Part I
Theory & Algorithms
[for Instance Ranking]

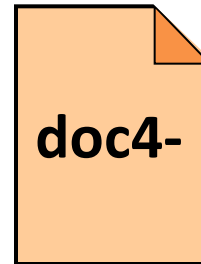
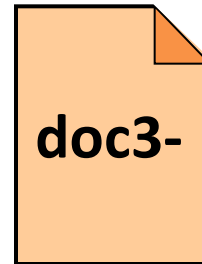
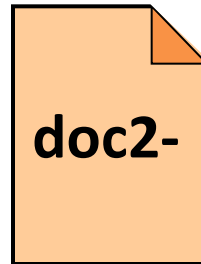
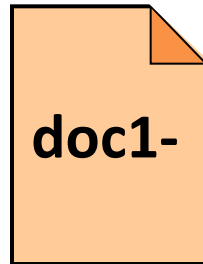
Bipartite Ranking

Relevant (+)



...

Irrelevant (-)



...

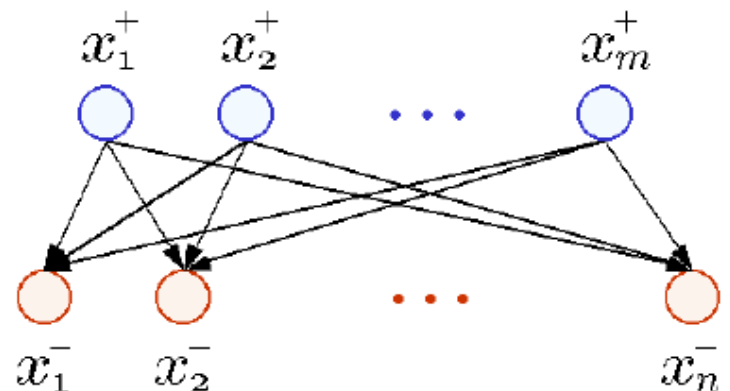
Bipartite Ranking

- ▶ Instance space X
- ▶ **Input:** Training sample $S = (S_+, S_-)$:

$S_+ = (x_1^+, \dots, x_m^+) \in X^m$ (positive examples)

$S_- = (x_1^-, \dots, x_n^-) \in X^n$ (negative examples)

- ▶ **Output:** Ranking function $f : X \rightarrow \mathbb{R}$



Bipartite Ranking

▶ Instance space X

▶ **Input:** Training sample $S = (S_+, S_-)$:

$$S_+ = (x_1^+, \dots, x_m^+) \in X^m \quad (\text{positive examples})$$

$$S_- = (x_1^-, \dots, x_n^-) \in X^n \quad (\text{negative examples})$$

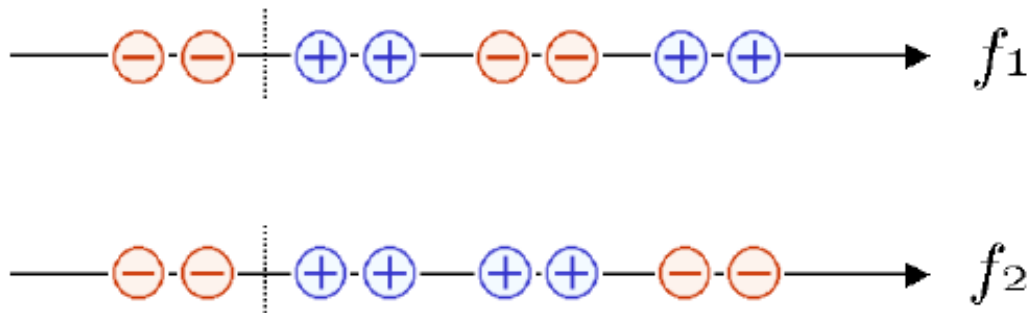
▶ **Output:** Ranking function $f : X \rightarrow \mathbb{R}$

▶ Expected error: $\mathbf{er}(f) = \mathbf{P}_{(x,x') \sim \mathcal{D}_+ \times \mathcal{D}_-} [f(x) < f(x')]$

▶ Empirical error: $\widehat{\mathbf{er}}_S(f) = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n \mathbf{1}(f(x_i^+) < f(x_j^-))$

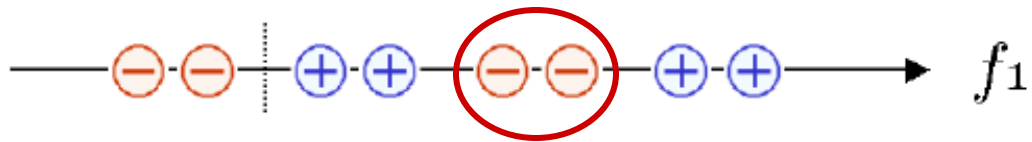
Is Bipartite Ranking Different from Binary Classification?

Example 1

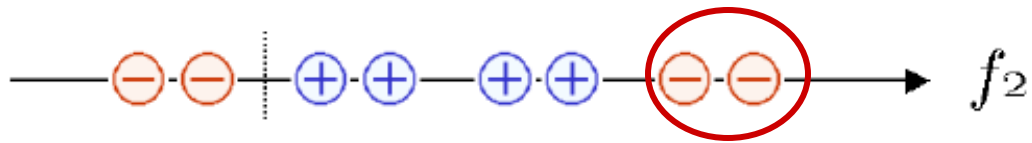


Is Bipartite Ranking Different from Binary Classification?

Example 1



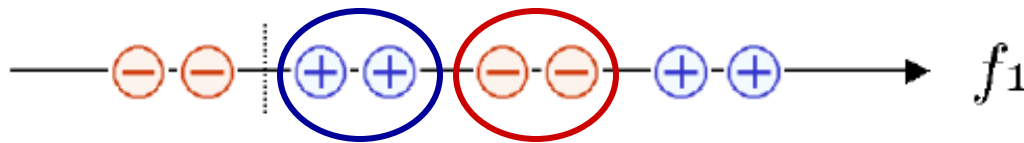
Classification error = $\frac{1}{4}$



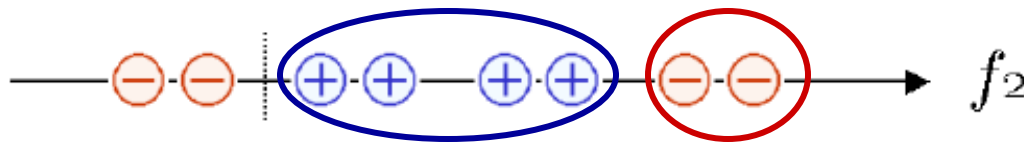
Classification error = $\frac{1}{4}$

Is Bipartite Ranking Different from Binary Classification?

Example 1



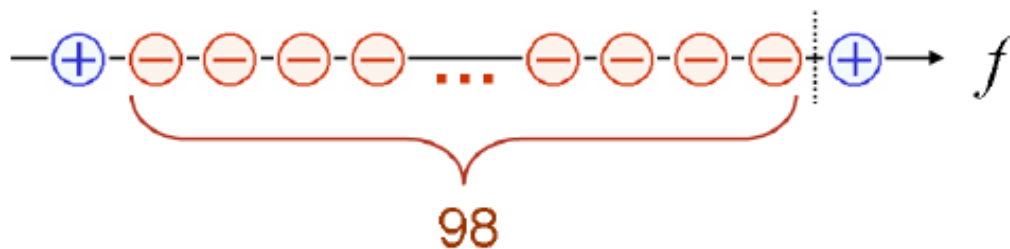
Classification error = $\frac{1}{4}$
Ranking error = $\frac{1}{4}$



Classification error = $\frac{1}{4}$
Ranking error = $\frac{1}{2}$

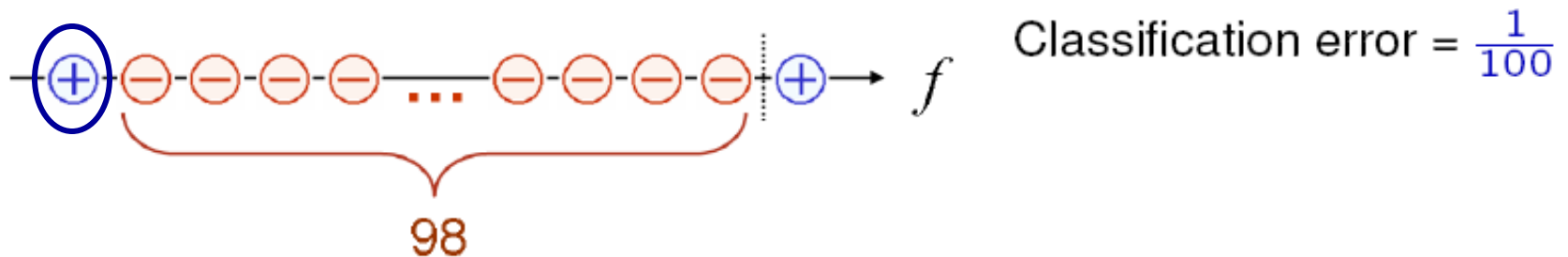
Is Bipartite Ranking Different from Binary Classification?

Example 2



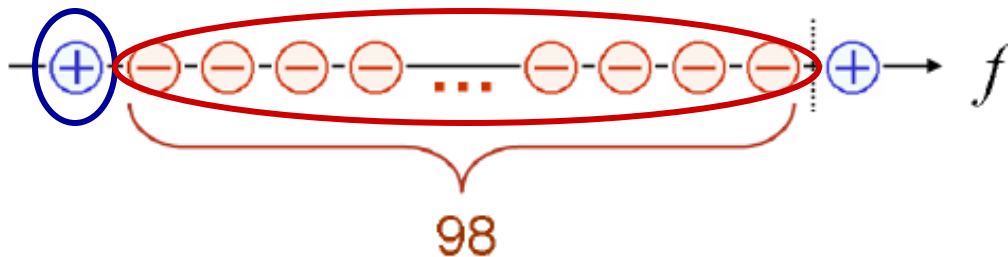
Is Bipartite Ranking Different from Binary Classification?

Example 2



Is Bipartite Ranking Different from Binary Classification?

Example 2



Classification error = $\frac{1}{100}$
Ranking error = $\frac{1}{2}$

Bipartite Ranking: Basic Algorithmic Framework

Minimize a convex upper bound on the empirical ranking error, possibly with some regularization, over some class of ranking functions:

$$\min_{f \in \mathcal{F}} \left[\frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n \ell(f, x_i^+, x_j^-) + \lambda N(f) \right]$$

where

$\ell(f, x_i^+, x_j^-)$: convex upper bound on $\mathbf{1}(f(x_i^+) < f(x_j^-))$

$N(f)$: regularizer

$\lambda > 0$: regularization parameter

\mathcal{F} : class of ranking functions

Bipartite RankSVM Algorithm

$$\min_{f \in \mathcal{F}_K} \left[\frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n \ell_{\text{hinge}}(f, x_i^+, x_j^-) + \frac{\lambda}{2} \|f\|_K^2 \right]$$

$$\ell_{\text{hinge}}(f, x_i^+, x_j^-) = \left(1 - (f(x_i^+) - f(x_j^-)) \right)_+ \quad [u_+ = \max(u, 0)]$$

\mathcal{F}_K = reproducing kernel Hilbert space (RKHS)
with kernel function K

$$N(f) = \frac{\|f\|_K^2}{2}$$

[Herbrich et al, 2000; Joachims, 2002; Rakotomamonjy, 2004]

Bipartite RankSVM Algorithm

Introducing slack variables and taking the Lagrangian dual results in the following convex quadratic program (QP) over mn variables $\{\alpha_{ij} : 1 \leq i \leq m, 1 \leq j \leq n\}$:

$$\min_{\alpha} \left[\frac{1}{2} \sum_{i=1}^m \sum_{j=1}^n \sum_{k=1}^m \sum_{l=1}^n \alpha_{ij} \alpha_{kl} \phi(x_i^+, x_j^-, x_k^+, x_l^-) - \sum_{i=1}^m \sum_{j=1}^n \alpha_{ij} \right]$$

$$\text{subject to } 0 \leq \alpha_{ij} \leq C \quad \forall i, j$$

where

$$\phi(x_i^+, x_j^-, x_k^+, x_l^-) = (K(x_i^+, x_k^+) - K(x_i^+, x_l^-) - K(x_j^-, x_k^+) + K(x_j^-, x_l^-))$$

$$C = \frac{1}{\lambda mn}$$

Can be solved using a standard QP solver, or more efficient methods (e.g. Chapelle & Keerthi, 2010).

Bipartite RankBoost Algorithm

$$\min_{f \in \mathcal{L}(\mathcal{F}_{\text{base}})} \left[\frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n \ell_{\text{exp}}(f, x_i^+, x_j^-) \right]$$

$$\ell_{\text{exp}}(f, x_i^+, x_j^-) = \exp \left(- \left(f(x_i^+) - f(x_j^-) \right) \right)$$

$\mathcal{L}(\mathcal{F}_{\text{base}})$ = linear combinations of functions in some
base class $\mathcal{F}_{\text{base}}$

[Freund et al, 2003]

Bipartite RankBoost Algorithm

Input: $(S_+, S_-) \in X^m \times X^n$.

Initialize: $D_1(x_i^+, x_j^-) = \frac{1}{mn}$ for all $i \in \{1, \dots, m\}, j \in \{1, \dots, n\}$.

For $t = 1, \dots, T$:

- Train weak learner using distribution D_t ; get weak ranker $f_t \in \mathcal{F}_{\text{base}}$.
- Choose $\alpha_t \in \mathbb{R}$.
- Update: $D_{t+1}(x_i^+, x_j^-) = \frac{1}{Z_t} D_t(x_i^+, x_j^-) \exp(-\alpha_t (f_t(x_i^+) - f_t(x_j^-)))$

where $Z_t = \sum_{i=1}^m \sum_{j=1}^n D_t(x_i^+, x_j^-) \exp(-\alpha_t (f_t(x_i^+) - f_t(x_j^-)))$.

Output final ranking: $f(x) = \sum_{t=1}^T \alpha_t f_t(x)$.

k-partite Ranking

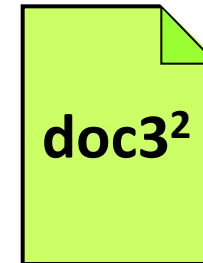
Rating *k*



...

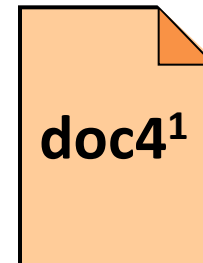
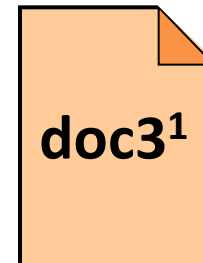
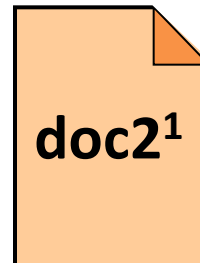
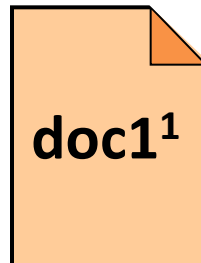
⋮

Rating 2



...

Rating 1



...

k -partite Ranking

▶ Instance space X

▶ **Input:** Training sample $S = (S_1, S_2, \dots, S_k)$:

$$S_k = (x_1^k, \dots, x_{n_k}^k) \in X^{n_k} \quad (\text{examples of rating } k)$$

⋮

$$S_2 = (x_1^2, \dots, x_{n_2}^2) \in X^{n_2} \quad (\text{examples of rating } 2)$$

$$S_1 = (x_1^1, \dots, x_{n_1}^1) \in X^{n_1} \quad (\text{examples of rating } 1)$$

▶ **Output:** Ranking function $f : X \rightarrow \mathbb{R}$

▶ Empirical error:

$$\widehat{\mathbf{er}}_S(f) = \left(\frac{1}{\sum_{1 \leq a < b \leq k} n_a n_b} \right) \sum_{1 \leq a < b \leq k} \sum_{i=1}^{n_b} \sum_{j=1}^{n_a} (b - a) \mathbf{1}(f(x_i^b) < f(x_j^a))$$

k -partite Ranking: Basic Algorithmic Framework

Minimize a convex upper bound on the empirical ranking error, possibly with some regularization, over some class of ranking functions:

$$\min_{f \in \mathcal{F}} \left[\left(\frac{1}{\sum_{1 \leq a < b \leq k} n_a n_b} \right) \sum_{1 \leq a < b \leq k} \sum_{i=1}^{n_b} \sum_{j=1}^{n_a} \ell(f, x_i^b, x_j^a, (b-a)) + \lambda N(f) \right]$$

where

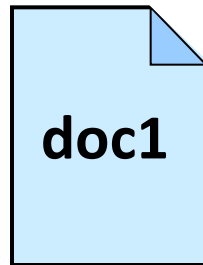
$\ell(f, x_i^b, x_j^a, (b-a))$: convex upper bound on $(b-a) \mathbf{1}(f(x_i^b) < f(x_j^a))$

$N(f)$: regularizer

$\lambda > 0$: regularization parameter

\mathcal{F} : class of ranking functions

Ranking with Real-Valued Labels



y_1



y_2



y_3

...

Ranking with Real-Valued Labels

- ▶ Instance space X
- ▶ Real-valued labels $Y = \mathbb{R}$
- ▶ **Input:** Training sample $S = ((x_1, y_1), \dots, (x_m, y_m)) \in (X \times \mathbb{R})^m$
- ▶ **Output:** Ranking function $f : X \rightarrow \mathbb{R}$

- ▶ Empirical error:

$$\widehat{\mathbf{er}}_S(f) = \frac{1}{\binom{m}{2}} \sum_{1 \leq i < j \leq m} |y_i - y_j| \mathbf{1} \left((y_i - y_j)(f(x_i) - f(x_j)) < 0 \right)$$

Ranking with Real-Valued Labels: Basic Algorithmic Framework

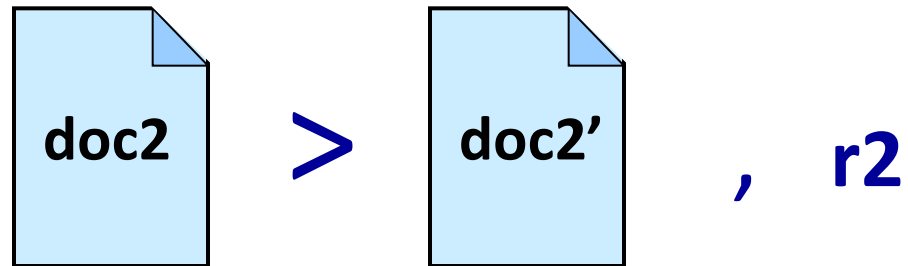
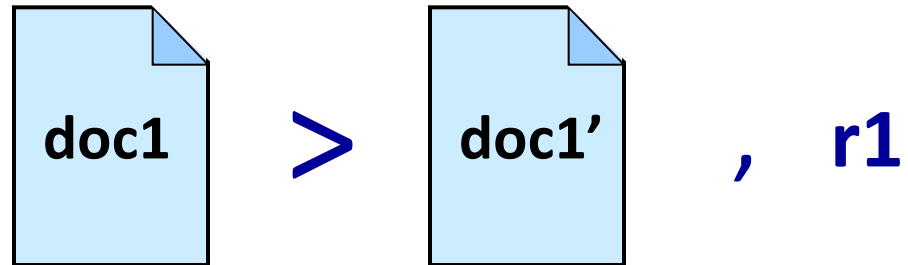
Minimize a convex upper bound on the empirical ranking error, possibly with some regularization, over some class of ranking functions:

$$\min_{f \in \mathcal{F}} \left[\frac{1}{\binom{m}{2}} \sum_{1 \leq i < j \leq m} \ell(f, (x_i, y_i), (x_j, y_j)) + \lambda N(f) \right]$$

where

- $\ell(f, (x_i, y_i), (x_j, y_j))$: convex upper bound on
 $|y_i - y_j| \mathbf{1} \left((y_i - y_j)(f(x_i) - f(x_j)) < 0 \right)$
- $N(f)$: regularizer
- $\lambda > 0$: regularization parameter
- \mathcal{F} : class of ranking functions

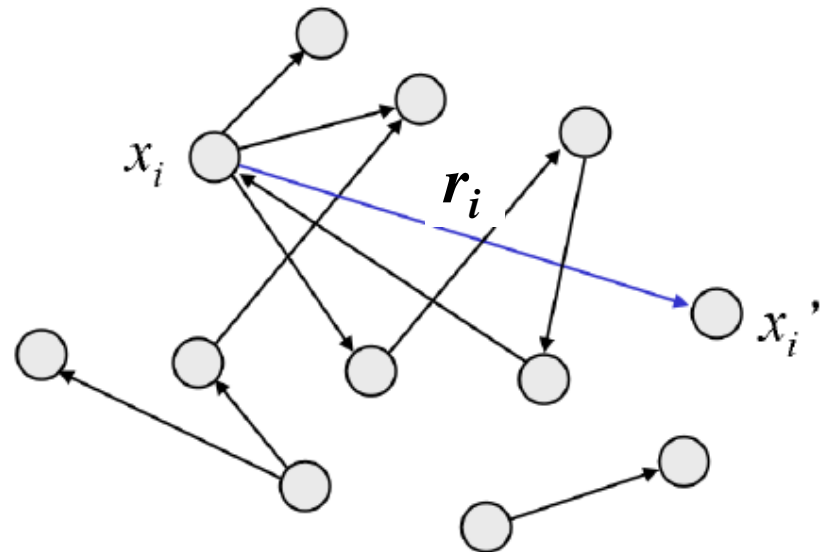
General Instance Ranking



...

General Instance Ranking

- ▶ Instance space X
- ▶ **Input:** Training sample $S = ((x_1, x'_1, r_1), \dots, (x_m, x'_m, r_m)) \in (X^2 \times \mathbb{R}_+)^m$
- ▶ **Output:** Ranking function $f : X \rightarrow \mathbb{R}$



General Instance Ranking

- ▶ Instance space X
- ▶ **Input:** Training sample $S = ((x_1, x'_1, r_1), \dots, (x_m, x'_m, r_m)) \in (X^2 \times \mathbb{R}_+)^m$
- ▶ **Output:** Ranking function $f : X \rightarrow \mathbb{R}$

▶ Empirical error:
$$\widehat{\mathbf{er}}_S(f) = \frac{1}{m} \sum_{i=1}^m r_i \mathbf{1}(f(x_i) < f(x'_i))$$

General Instance Ranking: Basic Algorithmic Framework

Minimize a convex upper bound on the empirical ranking error, possibly with some regularization, over some class of ranking functions:

$$\min_{f \in \mathcal{F}} \left[\frac{1}{m} \sum_{i=1}^m \ell(f, x_i, x'_i, r_i) + \lambda N(f) \right]$$

where

- $\ell(f, x_i, x'_i, r_i)$: convex upper bound on $r_i \mathbf{1}(f(x_i) < f(x'_i))$
- $N(f)$: regularizer
- $\lambda > 0$: regularization parameter
- \mathcal{F} : class of ranking functions

General RankSVM Algorithm

$$\min_{f \in \mathcal{F}_K} \left[\frac{1}{m} \sum_{i=1}^m \ell_{\text{hinge}}(f, x_i, x'_i, r_i) + \frac{\lambda}{2} \|f\|_K^2 \right]$$

$$\ell_{\text{hinge}}(f, x_i, x'_i, r_i) = \left(r_i - (f(x_i) - f(x'_i)) \right)_+ \quad [u_+ = \max(u, 0)]$$

\mathcal{F}_K = reproducing kernel Hilbert space (RKHS)
with kernel function K

$$N(f) = \frac{\|f\|_K^2}{2}$$

[Herbrich et al, 2000; Joachims, 2002]

General RankBoost Algorithm

$$\min_{f \in \mathcal{L}(\mathcal{F}_{\text{base}})} \left[\frac{1}{m} \sum_{i=1}^m \ell_{\text{exp}}(f, x_i, x'_i, r_i) \right]$$

$$\ell_{\text{exp}}(f, x_i, x'_i, r_i) = r_i \exp \left(- \left(f(x_i) - f(x'_i) \right) \right)$$

$\mathcal{L}(\mathcal{F}_{\text{base}})$ = linear combinations of functions in some
base class $\mathcal{F}_{\text{base}}$

[Freund et al, 2003]

Application to Information Retrieval (IR)

information - Google Search - Windows Internet Explorer

http://www.google.com/#hl=en&source=hp&q=information&rlz=1W1FUJB_en&aq=f&aqi=g10&aqi=&soq=&fp=18ec2db39eb50b9d

File Edit View Favorites Tools Help

Google information Search Share Sidewiki Bookmarks Check Translate AutoFill information

Information - Google Search

Web Images Videos Maps News Shopping Gmail more

Google information Search Advanced Search


Web Show options... Results 1 - 10 of about 2,290,000,000 for information [definition]. (0.19 seconds)

Information - Wikipedia, the free encyclopedia
Information as a concept has many meanings, from everyday usage to technical settings. The concept of **information** is closely related to notions of ...
[Etymology](#) - [As sensory input](#) - [As an influence which leads to ...](#)
en.wikipedia.org/wiki/Information - [Cached](#) - [Similar](#)

Information theory - Wikipedia, the free encyclopedia
Information theory is a branch of applied mathematics and electrical engineering involving the quantification of **information**. ...
en.wikipedia.org/wiki/Information_theory - [Cached](#) - [Similar](#)

Information Please
Infoplease.com, a free, authoritative, and respected reference for Internet users, provides a comprehensive encyclopedia, almanac, atlas, dictionary, ...
[Countries](#) - [United States](#) - [This Day In History](#) - [Biography](#)
www.infoplease.com/ - [Cached](#) - [Similar](#)

Local business results for information near Allston, MA - Change location



A **Federal Reserve Bank: General Information**
www.bos.frb.org - [\(617\) 973-3000](#) - [More](#)

B **Dana-Farber Cancer Institute**
www.dana-farber.org - [\(617\) 632-3000](#) - [95 reviews](#)

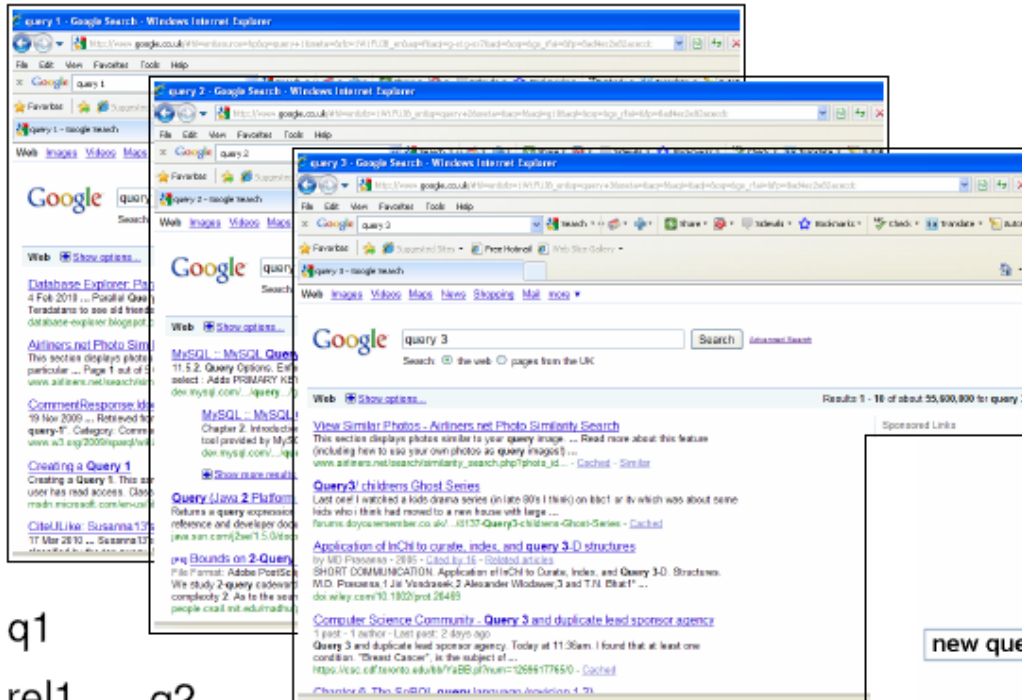
Sponsored Links

Looking For Information?
Find The Info You're Looking For With Google. Make It Your Homepage!
Google.com/Homepage

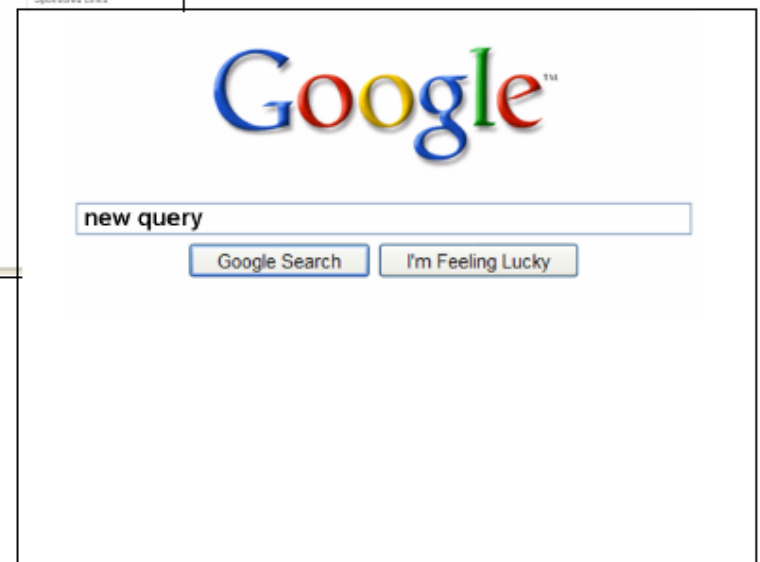
Information at Amazon
Low Prices on **Information**
Free 2-Day Shipping w/ Amazon Prime
www.Amazon.com/Books

[See your ad here >](#)

Learning to Rank in IR

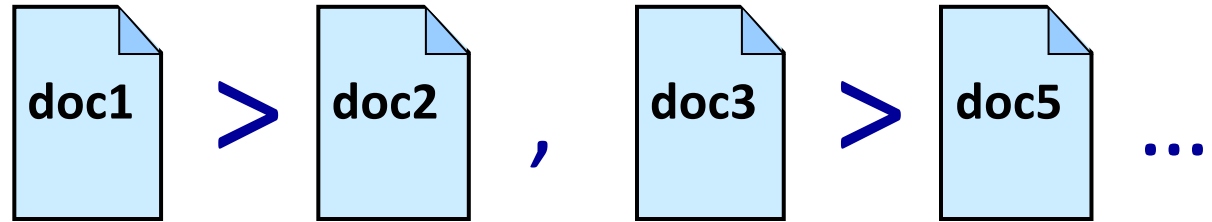


q1
rel1 q2
 rel2 q3
 rel3

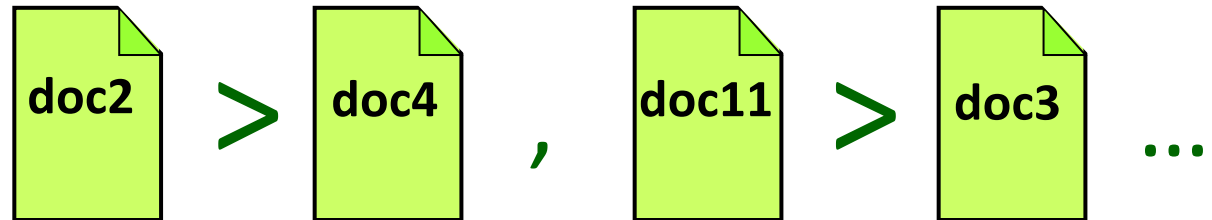


General Subset Ranking

query 1



query 2



...

General Subset Ranking

- ▶ Query space Q
- ▶ Document space D
- ▶ Query-document feature mapping $\phi : Q \times D \rightarrow \mathbb{R}^d$
- ▶ **Input:** Training sample $S = (S^1, \dots, S^m)$:

$$S^i = ((\phi_1^i, \phi_1^{i'}), \dots, (\phi_{n_i}^i, \phi_{n_i}^{i'})) \in (\mathbb{R}^d \times \mathbb{R}^d)^{n_i}$$

where

$$\phi_j^i = \phi(q^i, d_j^i), \quad \phi_j^{i'} = \phi(q^i, d_j^{i'})$$

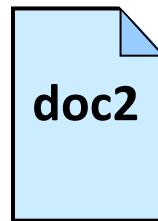
- ▶ **Output:** Ranking function $f : \mathbb{R}^d \rightarrow \mathbb{R}$

Subset Ranking with Real-Valued Relevance Labels

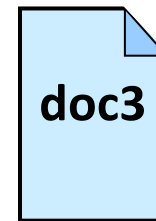
query 1



y_1 ,



y_2 ,



y_3 ...

query 2



y_1 ,



y_2 ,



y_3 ...

...

Subset Ranking with Real-Valued Relevance Labels

- ▶ Query space Q
- ▶ Document space D
- ▶ Query-document feature mapping $\phi : Q \times D \rightarrow \mathbb{R}^d$
- ▶ **Input:** Training sample $S = (S^1, \dots, S^m)$:

$$S^i = ((\phi_1^i, y_1^i), \dots, (\phi_{n_i}^i, y_{n_i}^i)) \in (\mathbb{R}^d \times \mathbb{R})^{n_i}$$

where

$$\phi_j^i = \phi(q^i, d_j^i), \quad y_j^i = \text{relevance of } d_j^i \text{ to } q^i$$

- ▶ **Output:** Ranking function $f : \mathbb{R}^d \rightarrow \mathbb{R}$

RankSVM Applied to IR/Subset Ranking

Standard RankSVM

$$\min_{f \in \mathcal{F}_K} \left[\left(\frac{1}{\sum_{i=1}^m \binom{n_i}{2}} \right) \sum_{i=1}^m \sum_{1 \leq j < k \leq n_i} \ell_{\text{hinge}} \left(f, (\phi_j^i, y_j^i), (\phi_k^i, y_k^i) \right) + \frac{\lambda}{2} \|f\|_K^2 \right]$$

$$\ell_{\text{hinge}} \left(f, (\phi_j^i, y_j^i), (\phi_k^i, y_k^i) \right) = \left(1 - \left(\text{sign}(y_j^i - y_k^i) \cdot (f(\phi_j^i) - f(\phi_k^i)) \right) \right)_+,$$

convex upper bound on

$$\mathbf{1} \left((y_j^i - y_k^i)(f(\phi_j^i) - f(\phi_k^i)) < 0 \right)$$

[Joachims, 2002]

RankSVM Applied to IR/Subset Ranking

RankSVM with Query Normalization & Relevance Weighting

$$\min_{f \in \mathcal{F}_K} \left[\frac{1}{m} \sum_{i=1}^m \left[\frac{1}{\binom{n_i}{2}} \sum_{1 \leq j < k \leq n_i} \ell_{\text{hinge}}^{\text{rel}}(f, (\phi_j^i, y_j^i), (\phi_k^i, y_k^i)) + \frac{\lambda}{2} \|f\|_K^2 \right] \right]$$

$$\ell_{\text{hinge}}^{\text{rel}}(f, (\phi_j^i, y_j^i), (\phi_k^i, y_k^i)) = \left(|y_j^i - y_k^i| - \left(\text{sign}(y_j^i - y_k^i) \cdot (f(\phi_j^i) - f(\phi_k^i)) \right) \right)_+,$$

convex upper bound on

$$|y_j^i - y_k^i| \mathbf{1} \left((y_j^i - y_k^i)(f(\phi_j^i) - f(\phi_k^i)) < 0 \right)$$

[Agarwal & Collins, 2010; also Cao et al, 2006]

Ranking Performance Measures in IR

Mean Average Precision (MAP)

Binary Labels: $y_j \in \{0, 1\}$

$$\text{MAP}_S(f) = \frac{1}{m} \sum_{i=1}^m \left[\frac{1}{|\{j : y_j^i = 1\}|} \sum_{j: y_j^i = 1} \text{prec}_{r_j^i}^i(f) \right]$$

r_j^i = rank of document d_j^i for query q^i

$\text{prec}_r^i(f)$ = fraction of positives in top r documents for query q^i

Ranking Performance Measures in IR

Normalized Discounted Cumulative Gain (NDCG)

General Real-Valued Labels: $y_j \in \mathbb{R}$

$$\text{NDCG}_S(f) = \frac{1}{m} \sum_{i=1}^m \left[\frac{1}{Z_i} \sum_{r=1}^{n_i} \frac{2^{y_{\pi_r^i}} - 1}{\log_2(r+1)} \right]$$

π_r^i = index of document ranked at position r for query q^i

Z_i = normalization constant

$$\text{NDCG}@k_S(f) = \frac{1}{m} \sum_{i=1}^m \left[\frac{1}{Z_i} \sum_{r=1}^k \frac{2^{y_{\pi_r^i}} - 1}{\log_2(r+1)} \right]$$

Ranking Algorithms for Optimizing MAP/NDCG

- ▶ SVMMAP [Yue et al. 2007]
- ▶ SVMNDCG [Chapelle et al. 2007]
- ▶ LambdaRank [Burges et al. 2007]
- ▶ AdaRank [Xu & Li 2007]
- ▶ Regression-based algorithm [Cossock & Zhang 2008]
- ▶ SoftRank [Taylor et al. 2008]
- ▶ SmoothRank [Chapelle & Wu 2010]