Fairness and Transparency in Ranking

Carlos Castillo / UPF
chato@acm.org

WSSC
Research Group on Web Science and Social Computing

Data and Algorithmic Bias Workshop (DAB) at CIKM'18
Turin, Italy, 2018-10-22

Ranking in IR

Objective: provide maximum relevance to searche**r**

Order by decreasing probability of being relevant

However, we sometimes care about the searche**d** items

Carbonell, J., & Goldstein, J. (1998, August). The use of MMR, diversity-based reranking for reordering documents and producing summaries. In *Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 335-336). ACM.

When searche**d** utility matters

Finding a local business	Business success
Purchasing a product or service	Marketing success
Recruiting a candidate for a job	Career success
Discovering events or groups to join	Social success
Learning about a political candidate	Political success
Dating/mating	Affective/reproductive success

What is discrimination?

X <u>discriminates</u> against someone Y in relation to Z if:

1. Y has property P and Z does not have P
2. X treats Y <u>worse</u> than s/he treats or would treat Z
3. It is <u>because</u> Y has P and Z does not have P that X treats Y worse than Z

Disadvantageous differential treatment

Kasper Lippert-Rasmussen: Born Free and Equal? A Philosophical Inquiry Into the Nature of Discrimination. Oxford University Press, 2013.

Group discrimination

X group-discriminates against Y in relation to Z if:

1. X generically discriminates against Y in relation to Z
2. P is the property of belonging to a socially salient group
3. This makes people with P worse off relative to others
   or X is motivated by animosity towards people with P,
   or by the belief that people with P are inferior
   or should not intermingle with others

Kasper Lippert-Rasmussen: Born Free and Equal? A Philosophical Inquiry Into the Nature of Discrimination. Oxford University Press, 2013.

Statistical discrimination

X <u>statistically discriminates</u> against Y in relation to Z if:

1. X group-discriminates against Y in relation to Z
2. P is <u>statistically relevant</u>
   (or X believes P is statistically relevant)

Kasper Lippert-Rasmussen: Born Free and Equal? A Philosophical Inquiry Into the Nature of Discrimination. Oxford University Press, 2013.

Example (statistical / non-statistical)

a. Not hiring a highly-qualified woman because women have <u>a higher probability</u> of taking parental leave (statistical discrimination)

b. Not hiring a highly-qualified woman because <u>she has said</u> that she intends to have a child and take parental leave (non-statistical discrimination)

Kasper Lippert-Rasmussen: Born Free and Equal? A Philosophical Inquiry Into the Nature of Discrimination. Oxford University Press, 2013.

In statistical machine learning

An algorithm developed through statistical machine learning can statistically discriminate if we:

1. <u>Disregard intentions/animosity</u>
2. Understand <u>statistically relevant</u> as any information derived from training data

Castillo, C. (2018). Algorithmic discrimination. Assessing the impact of machine intelligence on human behaviour: an interdisciplinary endeavour (1st HUMAINT Workshop).

Fairness in ranking is ...

1. A **sufficient presence** of elements of the protected group
   - Absence of statistical (group) discrimination
   - Prevent allocative harms to a group
2. A **consistent treatment** of elements of both groups
   - Absence of individual discrimination
3. A **proper representation** of disadvantaged groups
   - Prevent representational harms to a group

Representational harms

Representational harms occur when systems reinforce the subordination of some groups along the lines of identity (Kate Crawford)

- Sexualized search results
  Circa 2013, "black women" but in general "(race) women"
- Stereotyped search suggestions
  Google now blacklist many "(nationality) are ..." completions
- Automatic image tagging errors

Noble, S. U. (2018). Algorithms of Oppression: How search engines reinforce racism. NYU Press.
Crawford, K. (2017). The Trouble with Bias. Keynote at NIPS.

Possible sources of unfairness

Biases in training data

Expert or editorially provided rankings

(e.g., all protected items ranked lower than nonprotected)

Biases in user behavior

Clicks and user feedback

(e.g., if women preferred ads for jobs that pay less)

Biases in document construction

(e.g., completion of different CV sections by men/women)

Why fair rankings might be needed?

1. Biases in training data harming searcher utility
2. Legal mandates and voluntary commitments to

   equal representation, or

   positive actions
3. Ensuring technology embodies certain values

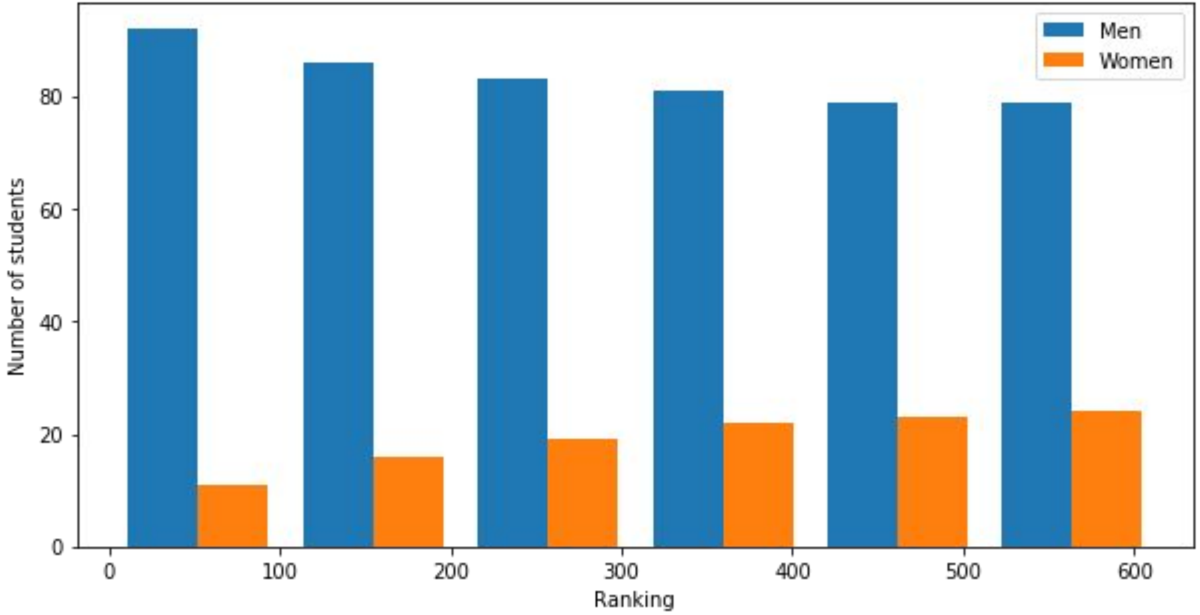
Example: job search

	Position 1	2	3	4	5	6	7	8	9	10	top 10 male	top 10 female	top 40 male	top 40 female
Economist	f	m	m	m	m	m	m	m	m	m	90%	10%	73%	27%
Market analyst	f	m	f	f	f	f	f	m	f	f	20%	80%	43%	57%
Copywriter	m	m	m	m	m	m	f	m	m	m	90%	10%	73%	27%

Top-10 results for 3 professions in XING (a recruitment site, similar to LinkedIn, that is a market leader in Germany and Austria)

Zehlike, M., Bonchi, F., Castillo, C., Hajian, S., Megahed, M., & Baeza-Yates, R. (2017). FA*IR: A fair top-k ranking algorithm. In Proc. of the ACM on Conference on Information and Knowledge Management (pp. 1569-1578). ACM.

Example: university admissions



Ranking of men and women admitted to an engineering school in Chile in 2013.

Zehlike, M., & Castillo, C. (2018). Reducing Disparate Exposure in Ranking: A Learning To Rank Approach. arXiv preprint arXiv:1805.08716.

Diversity

Introduced (20+ years ago!) to:

1. Increasing variety by maximizing marginal relevance
2. Accounting for uncertain intent ("hedging bets")

Making sure that people searching for a luxury car would not get only results about *Panthera onca*

Carbonell, J., & Goldstein, J. (1998). The use of MMR, diversity-based reranking for reordering documents and producing summaries. In *Proc. SIGIR, the 21st annual International Conference on Research and Development in Information Retrieval* (pp. 335-336). ACM.

Diversity ≠ Fairness

Concerned with searche**r** utility

Symmetric

Concerned with searche**d** utility

Asymmetric

Focus on a protected group:
a socially salient, disadvantaged group

Measuring Fairness in Rankings

Methods for measuring fairness

Exposure-based

Singh and Joachims 2018

Probability-based

Yang and Stoyanovich 2017, Zehlike et al. 2017

Methods for measuring fairness

Exposure-based

Singh and Joachims 2018

Probability-based

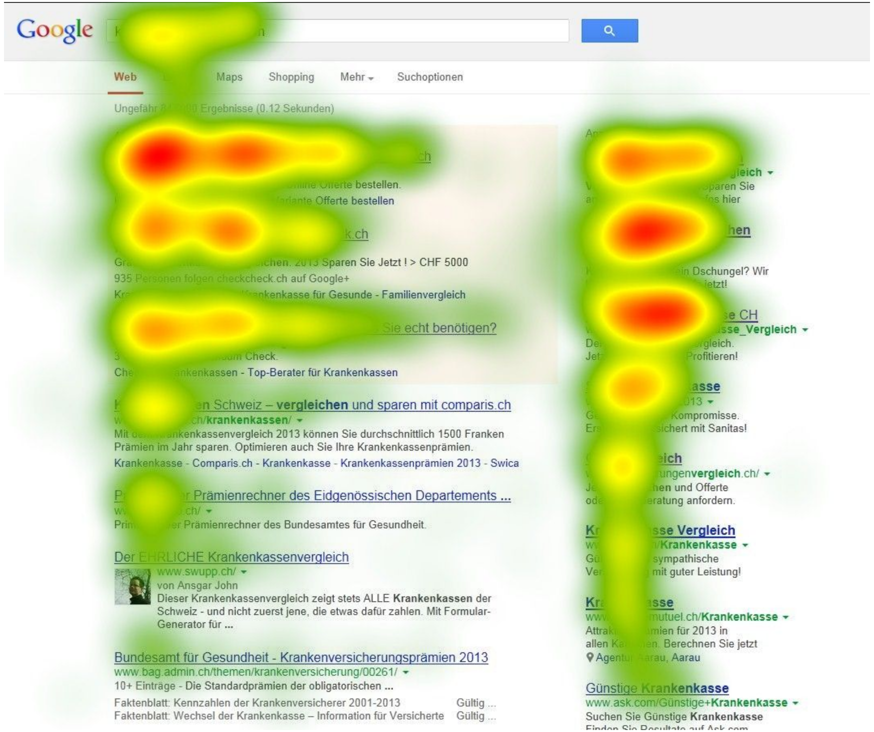
Yang and Stoyanovich 2017, Zehlike et al. 2017

Disparate exposure

Each position in a ranking has a certain probability of being examined v_i

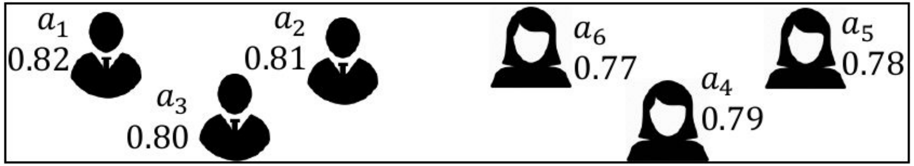
A ranking is fair if

$$\underset{i \in G_0}{E(v_i)} \simeq \underset{i \in G_1}{E(v_i)}$$



Singh, A., & Joachims, T. (2018). Fairness of Exposure in Rankings. In Proc. of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (pp. 2219-2228). ACM.

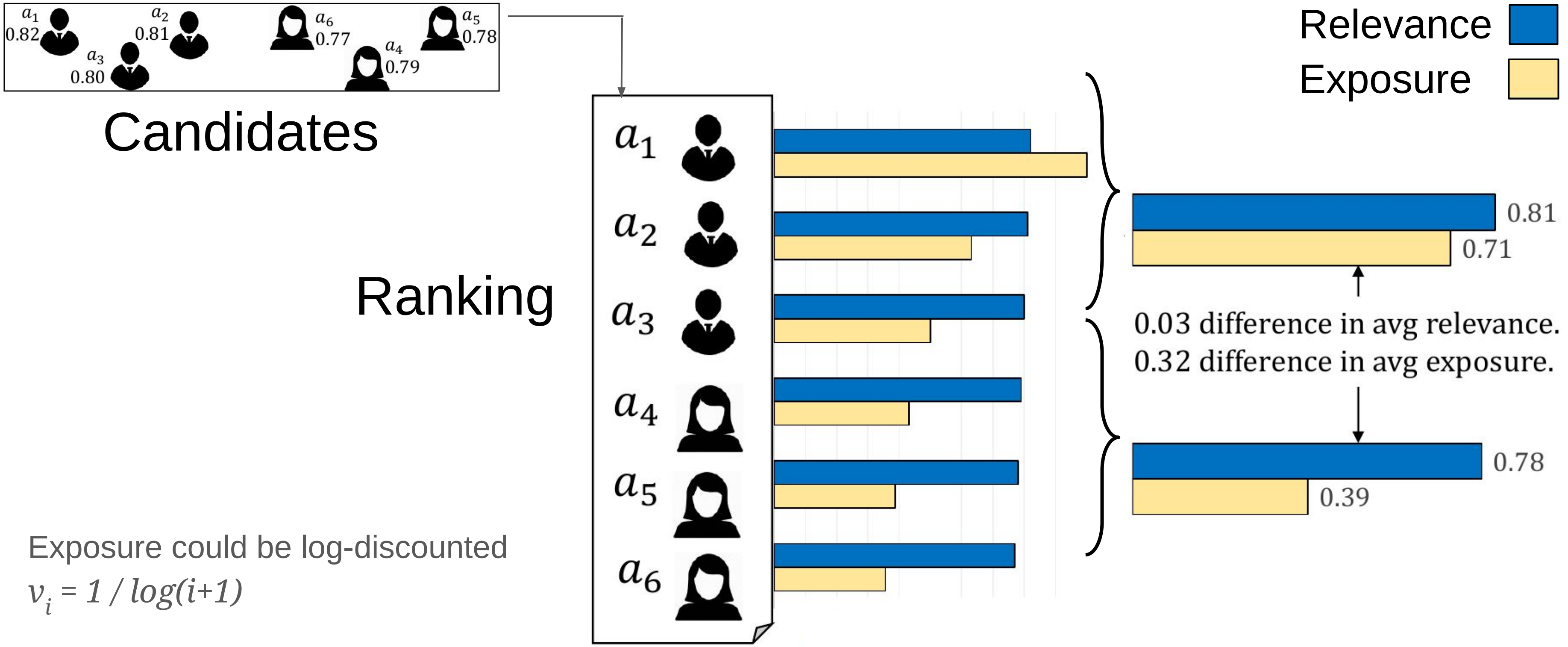
Disparate exposure: example



Candidates
(and their relevance)

Singh, A., & Joachims, T. (2018). Fairness of Exposure in Rankings. In Proc. of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (pp. 2219-2228). ACM.

Disparate exposure: example



Exposure could be log-discounted

$v_i = 1 / log(i+1)$

Singh, A., & Joachims, T. (2018). Fairness of Exposure in Rankings. In Proc. of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (pp. 2219-2228). ACM.

Disparate exposure

Utility-normalized exposure disparity ("Disparate Treatment Ratio"):

$$\mathrm{DTR}(G_0, G_1|\mathbf{P}, q) = \frac{\mathrm{Exposure}(G_0|\mathbf{P})/\mathrm{U}(G_0|q)}{\mathrm{Exposure}(G_1|\mathbf{P})/\mathrm{U}(G_1|q)}$$

$$\mathrm{Exposure}(G_k|\mathbf{P}) = \frac{1}{|G_k|} \sum_{d_i \in G_k} \sum_{j=1}^{N} \mathbf{P}_{i,j}\mathbf{v}_j$$

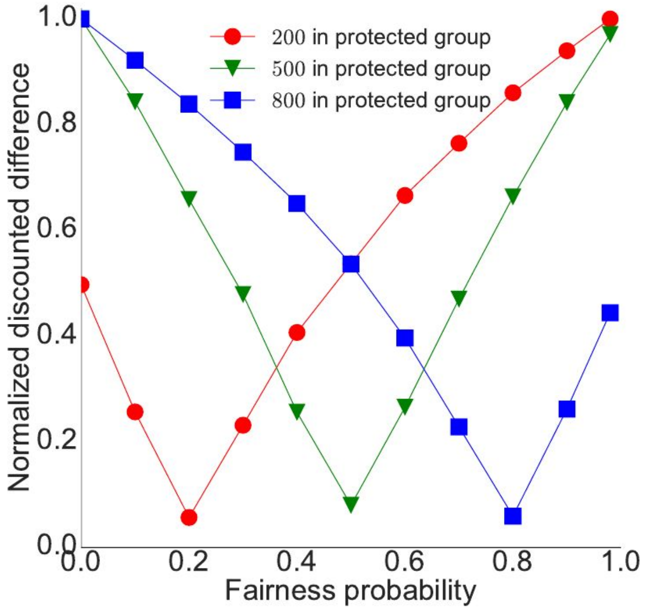
Expected click-through rate disparity ("Disparate Impact Ratio"):

$$\mathrm{DIR}(G_0, G_1|\mathbf{P}, q) = \frac{\mathrm{CTR}(G_0|\mathbf{P})/\mathrm{U}(G_0|q)}{\mathrm{CTR}(G_1|\mathbf{P})/\mathrm{U}(G_1|q)}$$

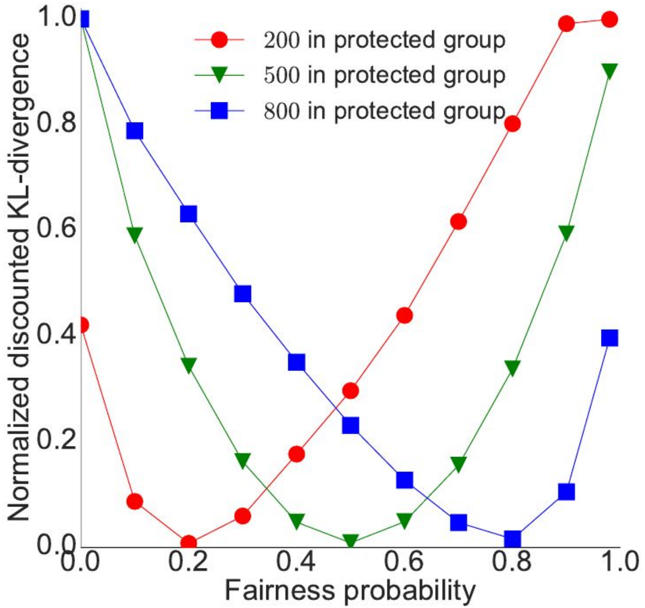
$$\mathrm{CTR}(G_k|\mathbf{P}) = \frac{1}{|G_k|} \sum_{i \in G_k} \sum_{j=1}^{N} \mathbf{P}_{i,j}\mathbf{u}_i\mathbf{v}_j.$$

Singh, A., & Joachims, T. (2018). Fairness of Exposure in Rankings. In Proc. of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (pp. 2219-2228). ACM.

Alternative: ad-hoc functions



$$\mathrm{rND}(\tau) = \frac{1}{Z}\sum_{i=10,20,\ldots}^{N} \frac{1}{log_2 i}\left|\frac{|S^+_{1\ldots i}|}{i} - \frac{|S^+|}{N}\right|$$



$$D_{KL}(P||Q) = \sum_i P(i) log \frac{P(i)}{Q(i)}$$

Yang, K., & Stoyanovich, J. (2017). Measuring fairness in ranked outputs. In Proc. of the 29th International Conference on Scientific and Statistical Database Management (p. 22). ACM.

Methods for measuring fairness

Exposure-based

Singh and Joachims 2018

Probability-based

Yang and Stoyanovich 2017, Zehlike et al. 2017

Ranking as randomized merging

1. Rank protected and unprotected separately

2. For each position:
- Pick protected with probability *p*
- Pick nonprotected with probability *1-p*

Continue until exhausting both lists

rank	gender
1	M
2	M
3	M
4	M
5	M
6	F
7	F
8	F
9	F
10	F

$p=0$

rank	gender
1	M
2	M
3	F
4	M
5	M
6	F
7	M
8	F
9	F
10	F

$p=0.3$

rank	gender
1	M
2	F
3	M
4	F
5	M
6	F
7	M
8	F
9	M
10	F

$p=0.5$

Yang, K., & Stoyanovich, J. (2017). Measuring fairness in ranked outputs. In Proc. of the 29th International Conference on Scientific and Statistical Database Management (p. 22). ACM.

Fair representation condition

Given parameters p, α and a set of size k

Let $F(x;p,k)$ be the cumulative distribution function of a binomial distribution with parameters p, k

A ranking of k elements having x protected elements has the fair representation condition with probability p and significance α if $F(x;p,k) > \alpha$

Zehlike, M., Bonchi, F., Castillo, C., Hajian, S., Megahed, M., & Baeza-Yates, R. (2017). FA*IR: A fair top-k ranking algorithm. In Proc. of the ACM on Conference on Information and Knowledge Management (pp. 1569-1578). ACM.

Example: fair representation condition

Suppose p=0.5, k=10, α=0.10

$F(1, 0.5, 10) = 0.01 < 0.10 \Rightarrow$ if 1 protected element, fail

$F(2, 0.5, 10) = 0.05 < 0.10 \Rightarrow$ if 2 protected elements, fail

$F(3; 0.5, 10) = 0.17 > 0.10 \Rightarrow$ if 3 protected elements, pass

$F(4; 0.5, 10) = 0.37 > 0.10 \Rightarrow$ if 4 protected elements, pass

...

Ranked group fairness (unadjusted)

Given parameters p, α and a list of size k

The list has the ranked group fairness condition if

for every $k \leq n$

the prefix of size k of the list
has the (p, α)-fair representation condition

Ranked group fairness condition

Given parameters p, α and a list of size n

Let $F(x;p,n)$ be the cumulative distribution function of a binomial distribution with parameters p, n

A ranking of n elements having x protected elements has the fair representation condition with probability p and significance α if $F(x;p,n) > \alpha$

Zehlike, M., Bonchi, F., Castillo, C., Hajian, S., Megahed, M., & Baeza-Yates, R. (2017). FA*IR: A fair top-k ranking algorithm. In Proc. of the ACM on Conference on Information and Knowledge Management (pp. 1569-1578). ACM.

Examples: ranked group fairness

Can be expressed with a vector

p \ k	1	2	3	4	5	6	7	8	9	10	11	12
0.1	0	0	0	0	0	0	0	0	0	0	0	0
0.2	0	0	0	0	0	0	0	0	0	0	1	1
0.3	0	0	0	0	0	0	1	1	1	1	1	2
0.4	0	0	0	0	1	1	1	1	2	2	2	3
0.5	0	0	0	1	1	1	2	2	3	3	3	4
0.6	0	0	1	1	2	2	3	3	4	4	5	5
0.7	0	1	1	2	2	3	3	4	5	5	6	6

Problem: multiple hypothesis testing

Zehlike, M., Bonchi, F., Castillo, C., Hajian, S., Megahed, M., & Baeza-Yates, R. (2017). FA*IR: A fair top-k ranking algorithm. In Proc. of the ACM on Conference on Information and Knowledge Management (pp. 1569-1578). ACM.

Ranked group fairness (adjusted)

Given parameters p, α and a list of size k

The list has the ranked group fairness condition if

for every $k \leq n$

the prefix of size k of the list

has the (p, α_c)-fair representation condition

Where $\alpha_c > \alpha$ is adjusted to make the failure probability of a ranking generated by Yang-Stoyanovich equal to α

Zehlike, M., Bonchi, F., Castillo, C., Hajian, S., Megahed, M., & Baeza-Yates, R. (2017). FA*IR: A fair top-k ranking algorithm. In Proc. of the ACM on Conference on Information and Knowledge Management (pp. 1569-1578). ACM.

Probability-based measure

Given a ranking of n elements …

… and a probability p:

The ranked group fairness is the minimum alpha such that the ranking passes the ranked group fairness at p, α

… and a significance α:

The ranked group fairness is the maximum p such that the ranking passes the ranked group fairness at p, α

Zehlike, M., Bonchi, F., Castillo, C., Hajian, S., Megahed, M., & Baeza-Yates, R. (2017). FA*IR: A fair top-k ranking algorithm. In Proc. of the ACM on Conference on Information and Knowledge Management (pp. 1569-1578). ACM.

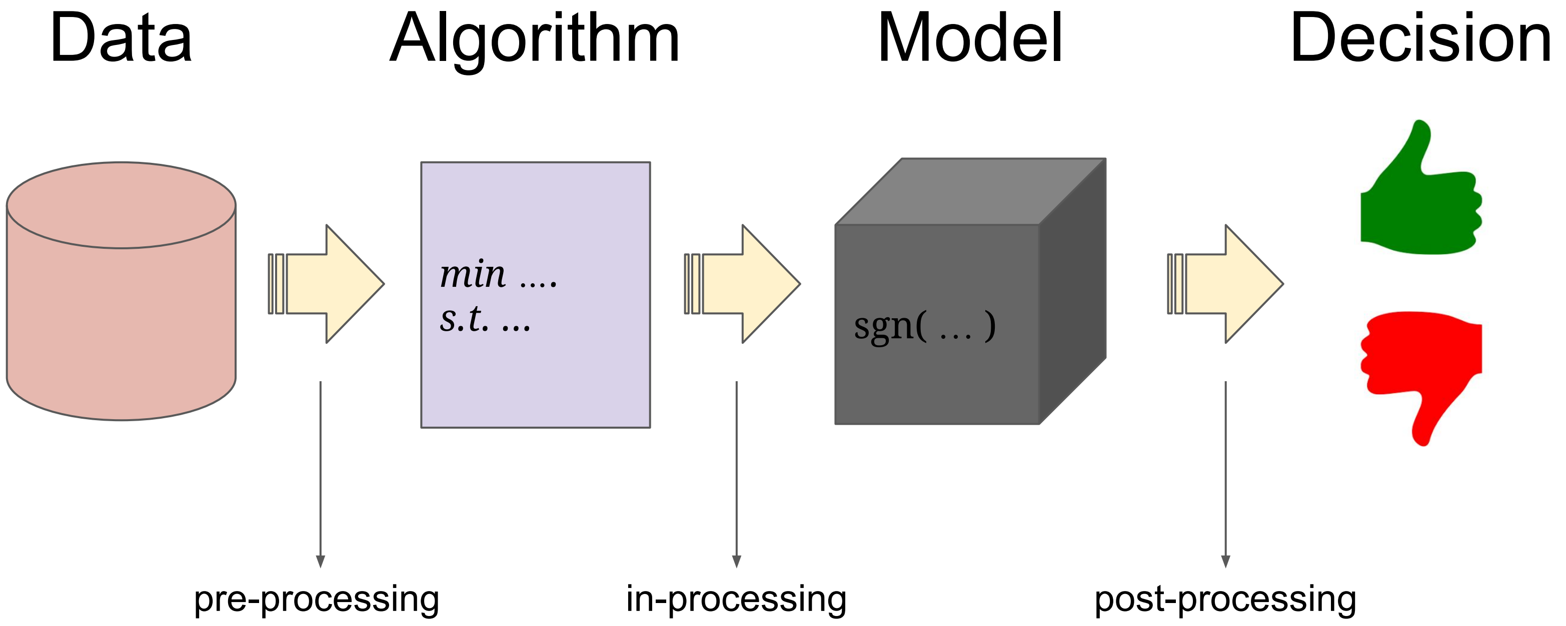
Multinomial FA*IR

Single-table approach to substitute CDF test does not trivially extend to multiple classes (yields a test that is too strict)

One additional dimension is needed ...

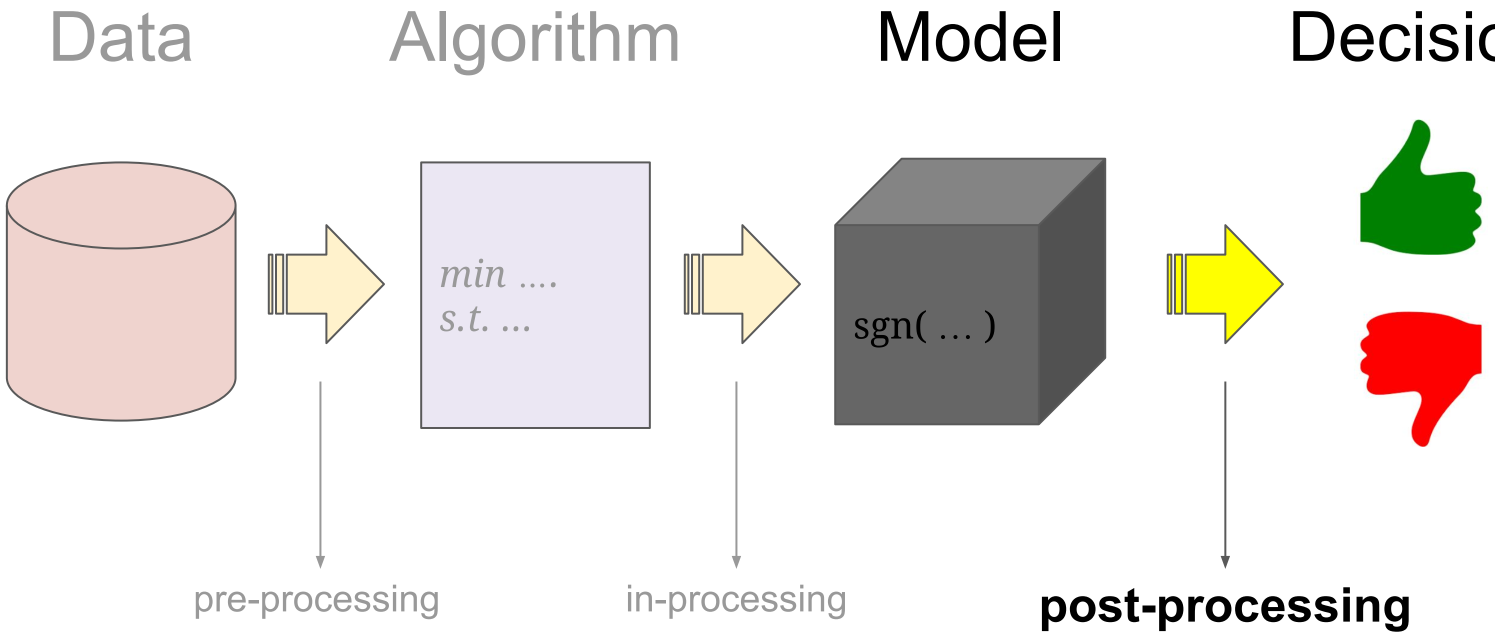
Creating Fair Rankings

Fairness: (pre,post,in)-processing



Hajian, S., Bonchi, F., & Castillo, C. (2016). Algorithmic bias: From discrimination discovery to fairness-aware data mining. In Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining (pp. 2125-2126). ACM.

Post-processing methods



Hajian, S., Bonchi, F., & Castillo, C. (2016). Algorithmic bias: From discrimination discovery to fairness-aware data mining. In Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining (pp. 2125-2126). ACM.

Single protected attribute (e.g., FA*IR)

Rank separately protected P and nonprotected N

Determine the *minimum number* of protected elements required at every ranking position using p, α

For every position

> **If** *enough* protected elements: pick next from best of P, N
> **else**: pick next from P

Zehlike, M., Bonchi, F., Castillo, C., Hajian, S., Megahed, M., & Baeza-Yates, R. (2017). FA*IR: A fair top-k ranking algorithm. In Proc. of the ACM on Conference on Information and Knowledge Management (pp. 1569-1578). ACM.

Multiple protected attribs (Celis et al.)

$$\arg\max_{x \in R_{m,n}} \sum_{i \in [m], j \in [n]} W_{ij} x_{ij} \quad \text{s.t.} \quad L_{k\ell} \leq \sum_{1 \leq j \leq k} \sum_{i \in P_\ell} x_{ij} \leq U_{k\ell} \quad \forall\, \ell \in [p], k \in [n]$$

x_{ij} is whether we place item i in position j

$R_{m,n}$ is the constraint that each item goes in one position only

W_{ij} is the utility of placing in position i the item j (non-decr.)

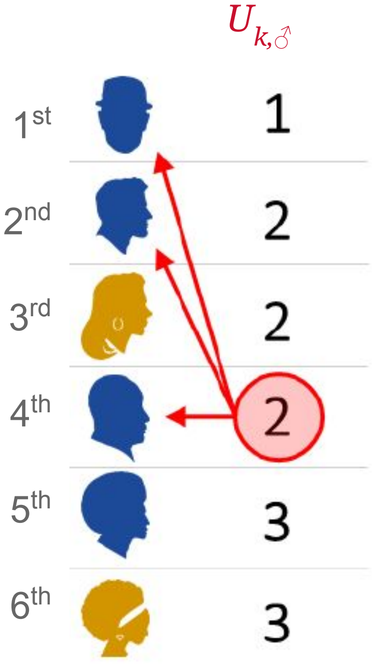
U_{kl} is the **given** max. number of items of class l up to pos k

Celis, L. E., Straszak, D., & Vishnoi, N. K. (2018). Ranking with fairness constraints. In *Proc. of 45th International Colloquium on Automata, Languages, and Programming (pp. 28:1-28:15).*

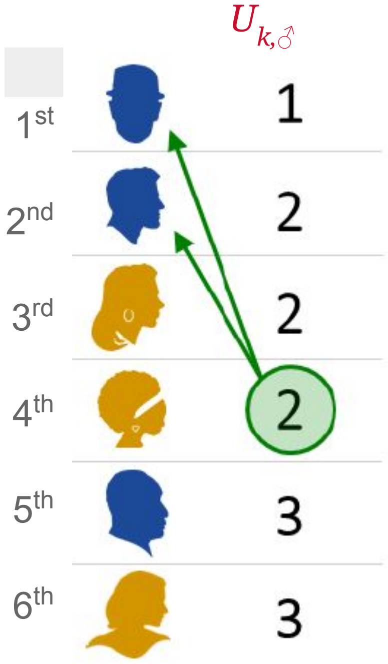
Example (Celis et al.)

W_{ij}								
1st	97	93	89	81	73	72	64	62
2nd	94	90	86	79	71	69	61	60
3rd	90	86	82	75	68	66	59	57
4th	78	74	71	65	58	57	51	49
5th	74	71	68	62	56	55	48	47
6th	71	68	65	59	53	52	46	45

Optimal unconstrained   Optimal constrained



Optimal unconstrained



Optimal constrained

Celis, L. E., Straszak, D., & Vishnoi, N. K. (2018). Ranking with fairness constraints. In *Proc. of 45th International Colloquium on Automata, Languages, and Programming (pp. 28:1-28:15).*

Results in Celis et al.

Let Δ = max. number of constrained attributes of an element

If Δ = 1: solvable in polynomial time using an LP relaxation

If Δ > 1: approximately solvable in polynomial time
using an LP relaxation, violates constraints
by at most a (Δ+2) factor

Singh and Joachims

Probabilistic ranking $\boldsymbol{P}$

$P_{i,j}$ is probability of placing document i in position j

$$\mathrm{U}(\mathbf{P}|q) = \sum_{d_i \in \mathcal{D}} \sum_{j=1}^{N} \mathbf{P}_{i,j}\, u(d_i|q)\, \mathbf{v}_j$$

$$\mathrm{Exposure}(G_k|\mathbf{P}) = \frac{1}{|G_k|} \sum_{d_i \in G_k} \sum_{j=1}^{N} \mathbf{P}_{i,j} \mathbf{v}_j$$

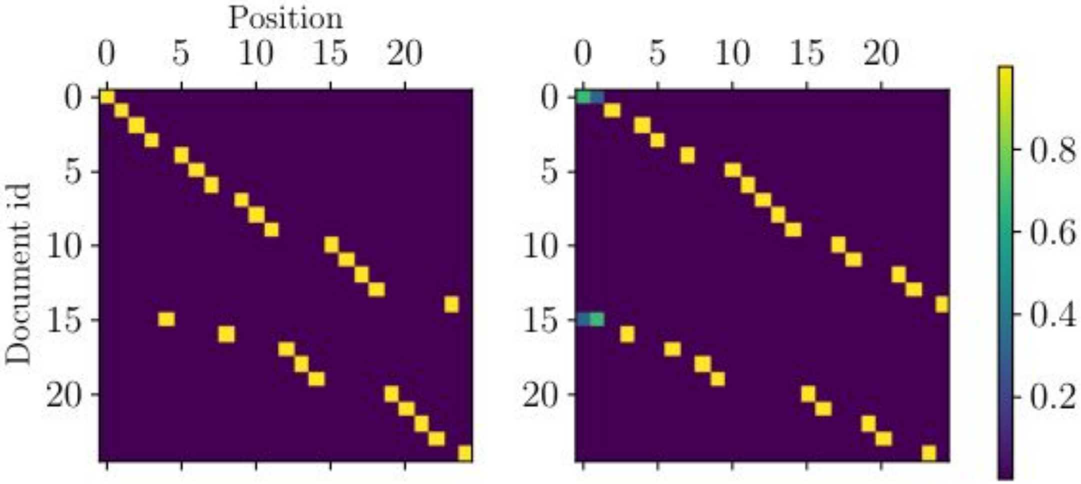
Maximize utility and reduce DTR and DIR

(utility-normalized exposure or predicted click-through rates)

Singh, A., & Joachims, T. (2018). Fairness of Exposure in Rankings. In Proc. of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (pp. 2219-2228). ACM.

Singh and Joachims (cont.)

Experimental results: (a) unconstrained and (b) fair ranking



(a) DCG=5.2027 (b) DCG=5.1360

Singh, A., & Joachims, T. (2018). Fairness of Exposure in Rankings. In Proc. of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (pp. 2219-2228). ACM.

Amortized fairness

Every element should receive attention or exposure (a_i) proportional to its utility (r_i)

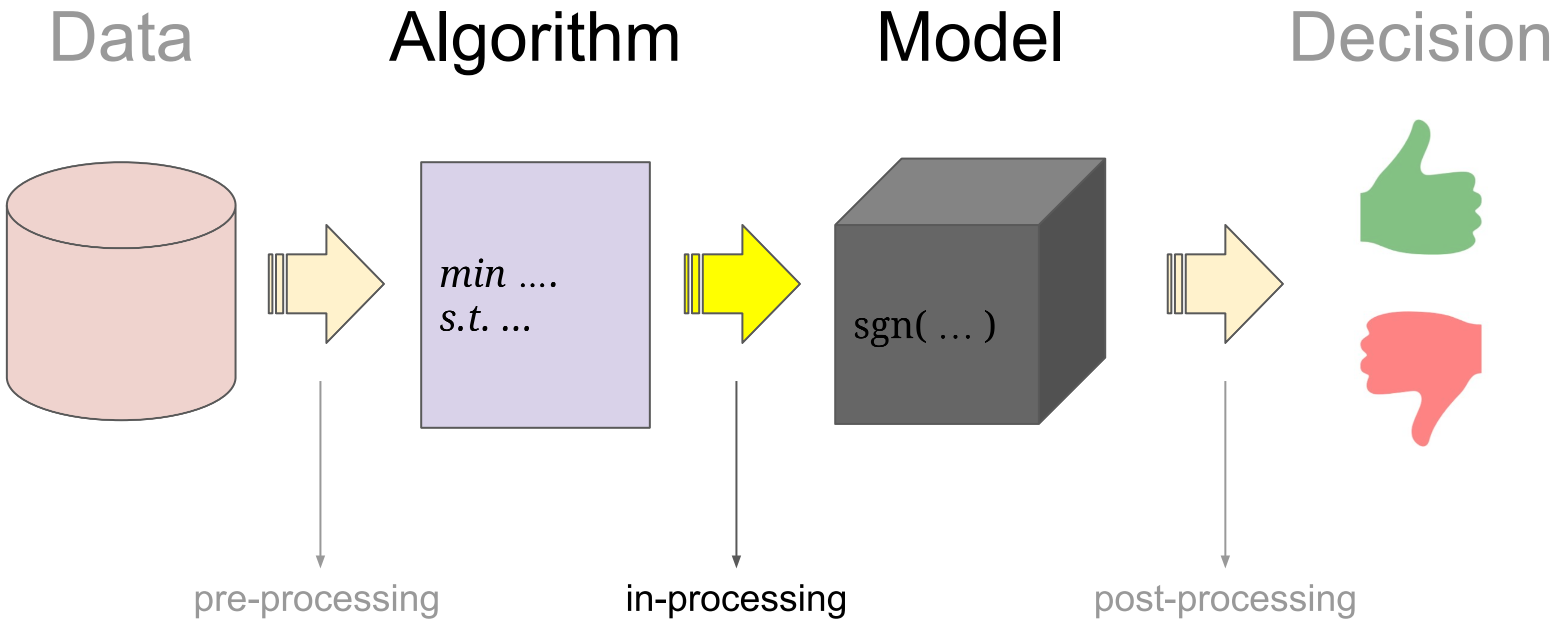
$$\frac{\sum_{l=1}^{m} a_{i1}^{l}}{\sum_{l=1}^{m} r_{i1}^{l}} = \frac{\sum_{l=1}^{m} a_{i2}^{l}}{\sum_{l=1}^{m} r_{i2}^{l}}, \forall u_{i1}, u_{i2}.$$

This should should be achieved across all m queries

At every query, consider past accumulated attention/utility deficits or surpluses, and correct them to the extent possible while honoring quality constraints

Biega, A. J., Gummadi, K. P., & Weikum, G. (2018). Equity of Attention: Amortizing Individual Fairness in Rankings. Proc. of SIGIR.

In-processing methods



Hajian, S., Bonchi, F., & Castillo, C. (2016). Algorithmic bias: From discrimination discovery to fairness-aware data mining. In Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining (pp. 2125-2126). ACM.

DELTR [Zehlike & Castillo 2018]

Optimize a combination of two losses:

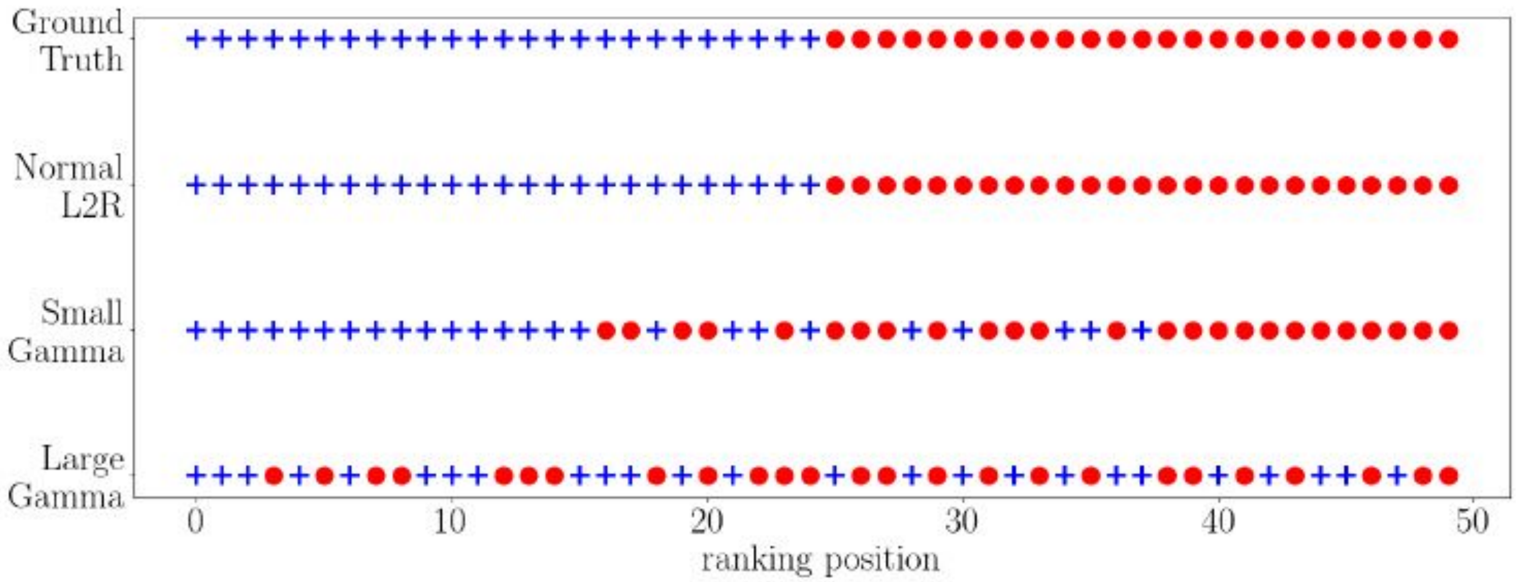
- L = loss due to difference between ranking predictions and training elements
- U = loss due to expected different exposure

$$L_{DELTR}\left(y^{(q)}, \hat{y}^{(q)}\right) = L\left(y^{(q)}, \hat{y}^{(q)}\right) + \gamma U\left(\hat{y}^{(q)}\right)$$

$$U(\hat{y}^{(q)}) = \max\left(0, \text{Exposure}(G_0|P_{\hat{y}^{(q)}}) - \text{Exposure}(G_1|P_{\hat{y}^{(q)}})\right)^2$$

Zehlike, M., and Castillo, C. (2018). Reducing Disparate Exposure in Ranking: A Learning To Rank Approach. Preprint arXiv:1805.08716.

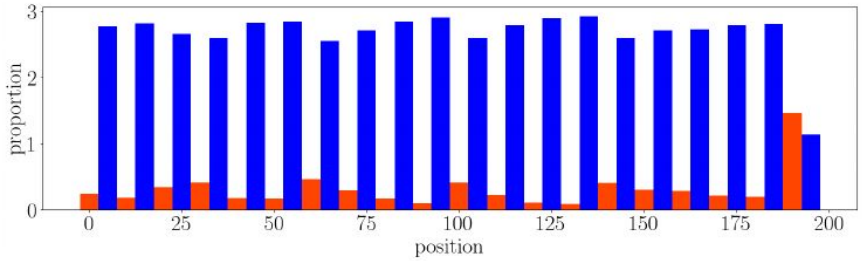
DELTR: Synthetic example



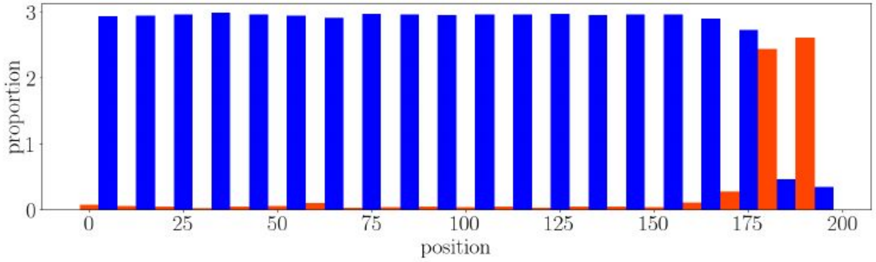
Zehlike, M., and Castillo, C. (2018). Reducing Disparate Exposure in Ranking: A Learning To Rank Approach. Preprint arXiv:1805.08716.

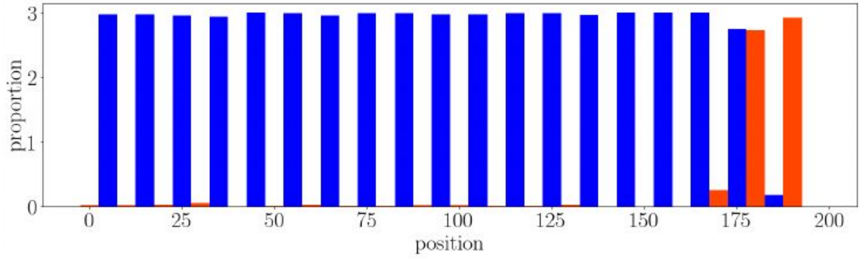
DELTR: W3C Corpus (TREC Expert)

"Color-blind"

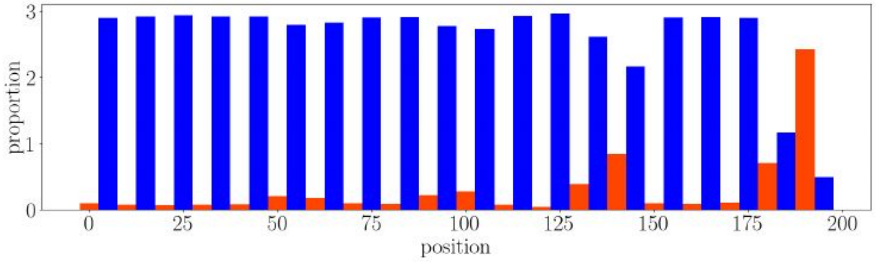


DELTR (small gamma)



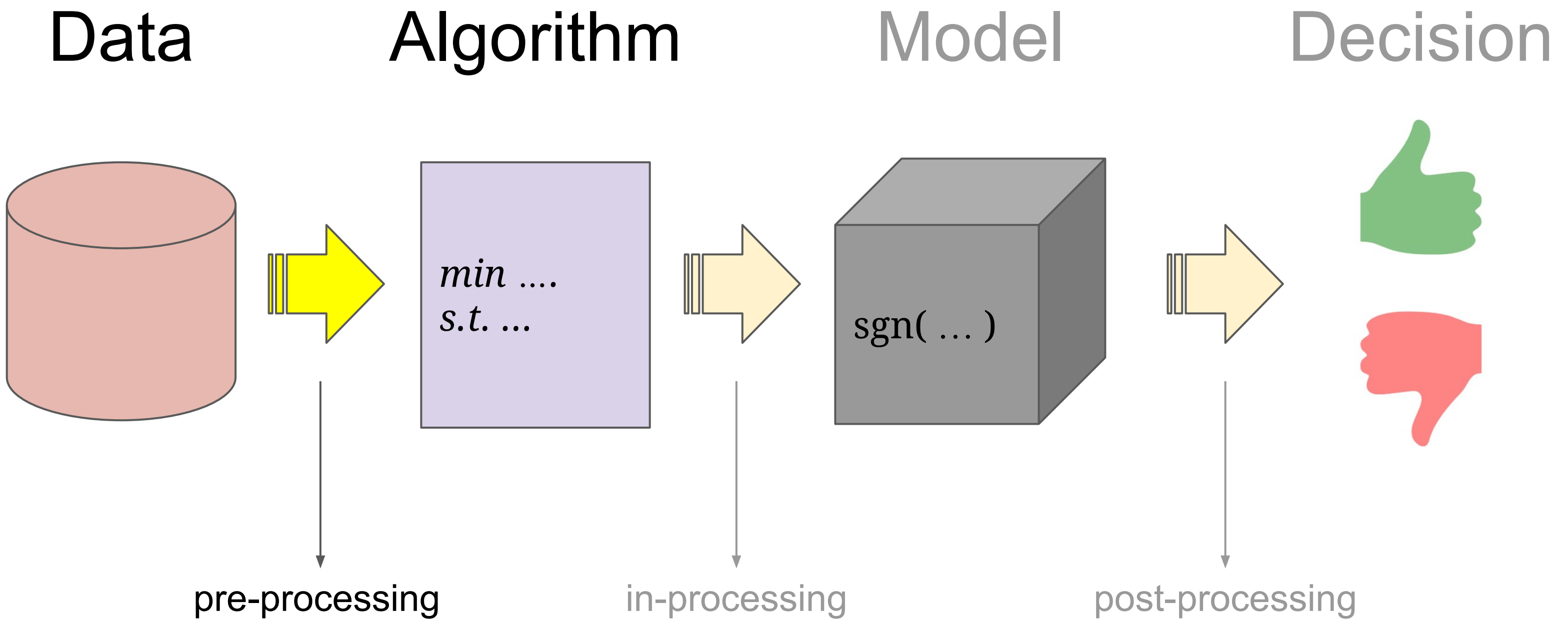


Learning to Rank



DELTR (large gamma)

Pre-processing methods



Hajian, S., Bonchi, F., & Castillo, C. (2016). Algorithmic bias: From discrimination discovery to fairness-aware data mining. In Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining (pp. 2125-2126). ACM.

Pre-processing training data

1. Before training a LTR system
   - Ensure rankings given as input satisfy a fair ranking condition
2. Train the LTR as usual
3. Profit!

Preliminary experiments promising, more remains to be done

Transparency in Rankings

Transparency: why and how?

Why:

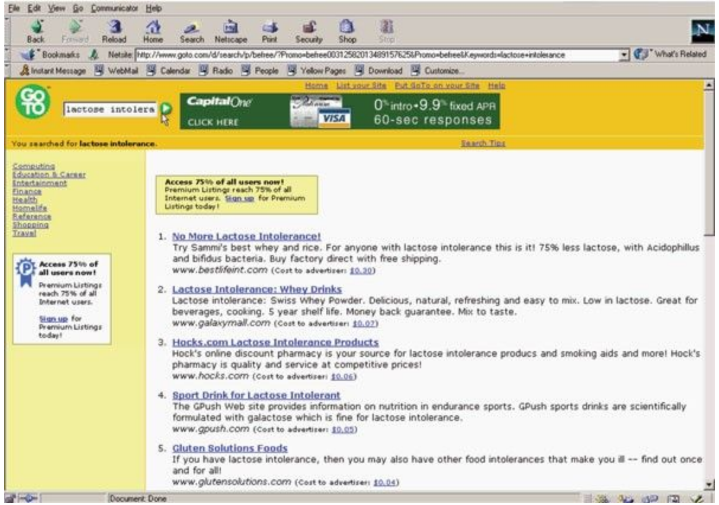
- Being able to test (**safety**)
- Supporting **ethics** compliance
- Ensuring **objectives** are aligned
- Making **trade-offs** visible

How:

- Explanations tend to be **contrastive**: why P and not Q?
- Explanations should empower user to **challenge** rankings

Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. Preprint arXiv:1702.08608.
Miller, T., Howe, P., & Sonenberg, L. (2017). Explainable AI: Beware of inmates running the asylum. In IJCAI-17 Workshop on Explainable AI (XAI) (Vol. 36).

GoTo.com (1997) and ad transparency



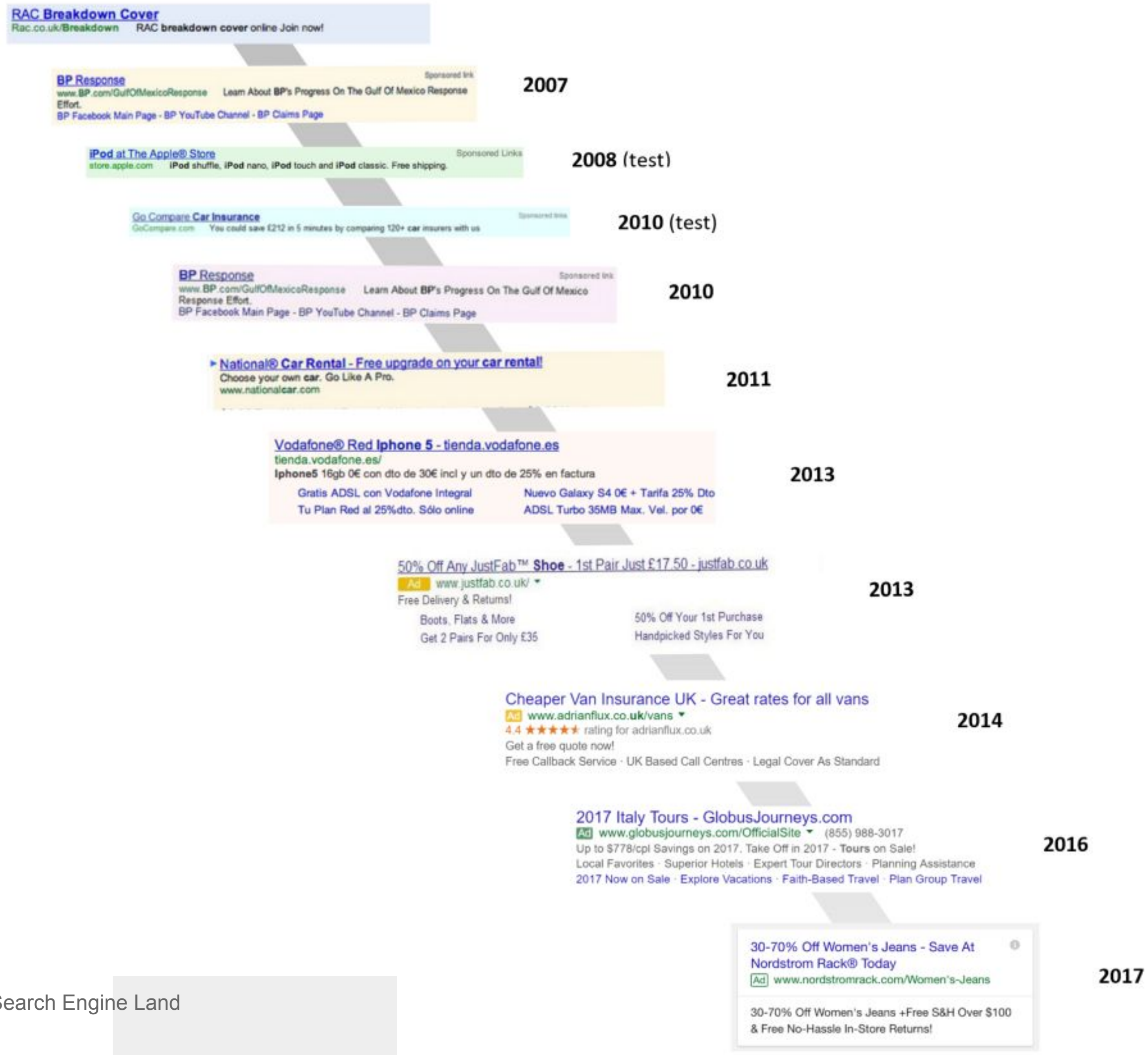
Ad labeling optional

- AOL ("sponsored links")
- Netscape("partner search results")
- CompuServe ("premium pages")
- GoTo ("featured listings")

Letter in 2002 to US' FTC from consumer advocacy organization leads to a warning by FTC in 2002, regulation ca. 2013.

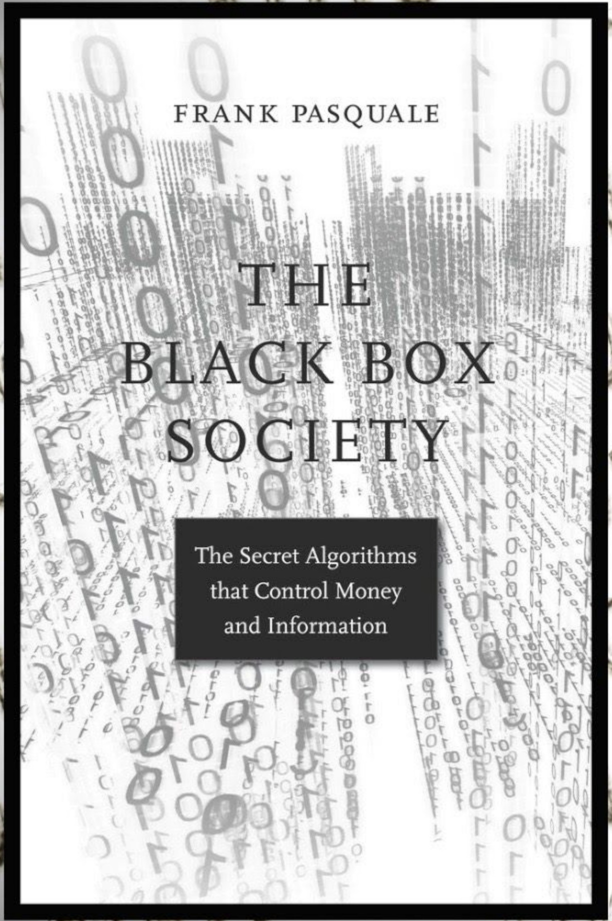
Hansell, S. (2001). Clicks for Sale; Paid Placement Is Catching On in Web Searches. New York Times.
Sinclair, A. (2004). Regulation of Paid Listings in Internet Search Engines: A Proposal for FTC Action. BUJ Sci. & Tech. L., 10, 353.

Advertising transparency is becoming transparent (!)



Marvin, G. (2017): A visual history of Google ad labeling in search results. Search Engine Land


Donald J. Trump
@realDonaldTrump
Following
Google search results for “Trump News” shows only the viewing/reporting of Fake New Media. In other words, they have it RIGGED, for me & others, so that almost all stories & news is BAD. Fake CNN is prominent. Republican/Conservative & Fair Media is shut out. Illegal? 96% of...
4:24 AM - 28 Aug 2018
Donald J. Trump
@realDonaldTrump
Following
....results on “Trump News” are from National Left-Wing Media, very dangerous. Google & others are suppressing voices of Conservatives and hiding information and news that is good. They are controlling what we can & cannot see. This is a very serious situation-will be addressed!
4:34 AM - 28 Aug 2018



Pasquale, F. (2015). The black box society: The secret algorithms that control money and information. Harvard University Press.

Transparency in algorithmic rankings

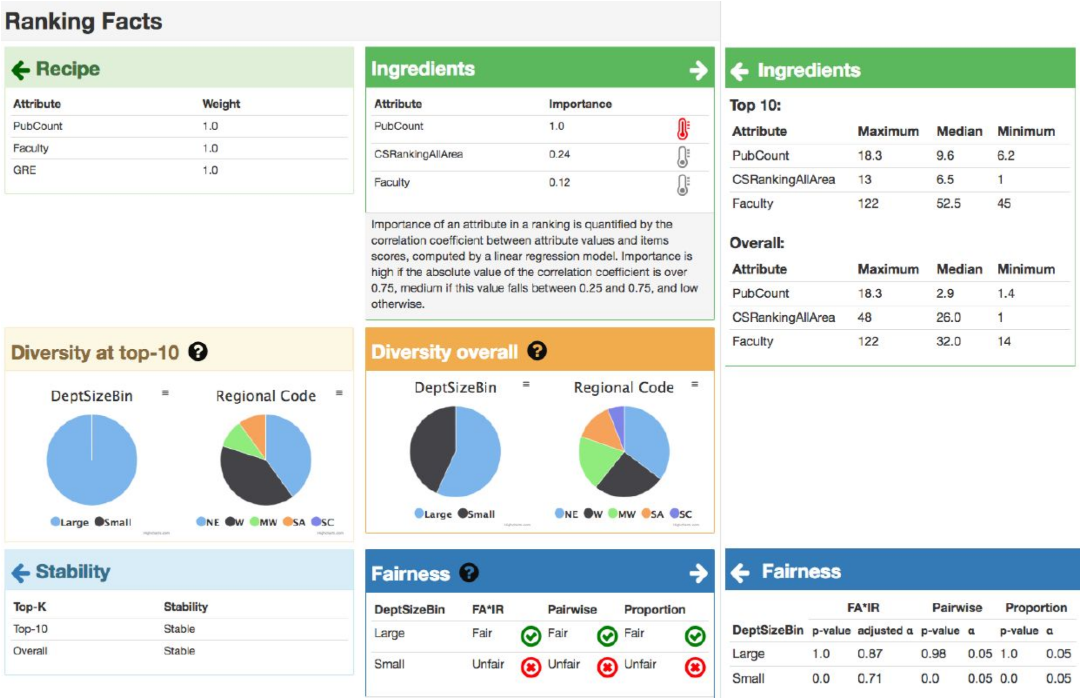
"Broadcast television can be monitored by anyone … **If the nightly television news does not cover a protest, the lack of coverage is evident** … However, **there is no transparency in algorithmic filtering**: how is one to know whether Facebook is showing [news about a protest] to everyone else but him or her, whether there is just no interest in the topic, or whether it is the algorithmic feedback cycle that is depressing the updates in favor of a more algorithm-friendly topic …?"

Tufekci, Z. (2017). Twitter and tear gas: The power and fragility of networked protest. Yale University Press.

Nutritional labels for rankings

Provide transparency about ranking factors, composition of the list, and fairness test

Example ranking labels for a ranking of computer science departments ▸



Yang, K., Stoyanovich, J., Asudeh, A., Howe, B., Jagadish, H. V., & Miklau, G. (2018). A Nutritional Label for Rankings. In Proc. SIGMOD (pp. 1773-1776). ACM.

Perturbation-based method

Suppose the score is a linear function of features, and documents are ranked by decreasing score ▸

	x_0	x_1	x_2	$score = 0.2x_0 + 0.3x_1 + 0.5x_2$
d_0	1	1	1	1
d_1	0.5	0.5	1	0.75
d_2	1	0	0.7	0.55

Feature x_2 has the highest weight but even if it were 0.6 for d_0 (lower than any other), document d_0 still would be at the top

In contrast, changing feature x_1 to 0 would change the ranking, hence x_1 is a better explanation

ter Hoeve, M., Schuth, A., Odijk, D., & de Rijke, M. (2018). Faithfully Explaining Rankings in a News Recommender System. arXiv preprint arXiv:1805.05447.

Transparency can help us researchers

We should avoid (at least) two pitfalls in our work:

- Sneaking positive/affirmative action without a consensus or where it is not welcome
- Certifying an algorithm that is part of an unfair system or is used in conditions of unfairness

⇒ we should be the first to provide transparency!

Barocas, S. (2017). What is the problem to which fair machine learning is the solution? AI Now Experts Workshop on Bias and Inclusion

Conclusions

This is just the beginning

Fairness in ranking is less explored than fairness in ML

There is no single solution and perhaps there will never be

Paraphrasing Solon Barocas: *«What is the problem to which fair ranking is the solution?»*

> The answer is that different solutions address different problems, which is totally fine!

Thank you!