Towards diversity-aware, fair and unbiased data management

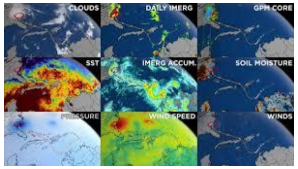
Evaggelia Pitoura

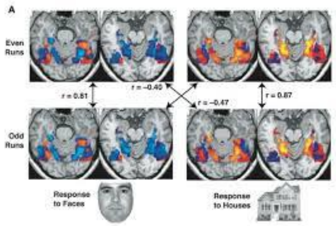
Computer Science and Engineering Department,

University of Ioannina,

Greece

Data is the new oil





Promises of data

Unprecedented amount of data coupled with cheap, widely-available, powerful computing and storage resources

Enormous opportunities

- Accelerate scientific discovery, personalized medicine, smart weather forecasting
- Improve life, personal assistants, recommendations
- Automate tasks
- Transform society, open government
- and more

But:

Should we trust data?

- At a personal level, we use reviews, recommendations, news feed, search results, … to guide our decisions
- And then data are used by organizations for school admission, job employment, insurance rates, and more

Case study: justice

ProPublica

Machine Bias

There's software used across the country to predict future criminals. And it's biased against blacks.

- COMPAS (Correctional Offender Management Profiling for Alternative Sanctions): Commercial tool, uses a risk assessment algorithm to predict some categories of future crime
- Used in courts in the US for bail and sentencing decisions

ProPublica found that

- the false positive rate for African American defendants (people labeled "high-risk" who did not re-offend) nearly twice as high as for White defendants
- Opposite for false negative rate

Prediction Fails Differently for Black Defendants

	WHITE	AFRICAN AMERICAN
Labeled Higher Risk, But Didn't Re-Offend	23.5%	44.9%
Labeled Lower Risk, Yet Did Re-Offend	47.7%	28.0%

Overall, Northpointe's assessment tool correctly predicts recidivism 61 percent of the time. But blacks are almost twice as likely as whites to be labeled a higher risk but not actually re-offend. It makes the opposite mistake among whites: They are much more likely than blacks to be labeled lower risk but go on to commit other crimes. (Source: ProPublica analysis of data from Broward County, Fla.)

The Wisconsin Supreme Court defended the use of COMPAS to inform criminal sentencing decisions

https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Case Study: image search

What images do people choose to represent careers?

In search results:

- evidence for *stereotype exaggeration*
- systematic *underrepresentation of women*

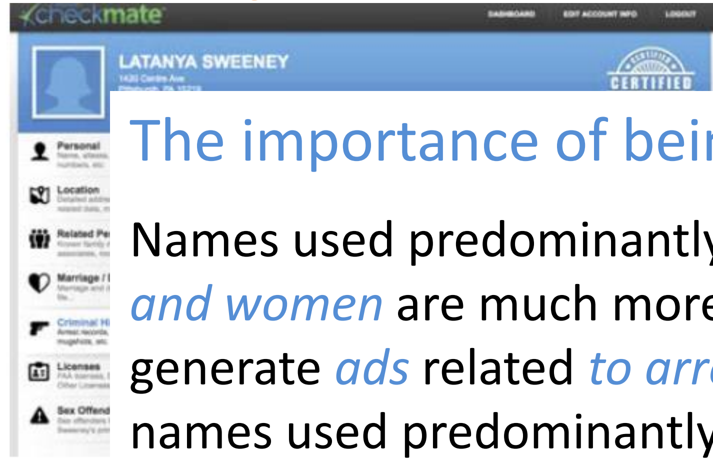
- People rate search results *higher* when they are *consistent* with stereotypes for a career
- Shifting the representation of gender in image search results can *shift people's perceptions* about real-world distributions. (after search slight increase in their believes)

Tradeoff between high user satisfaction results and broader societal goals for equality of representation

- Similar biases in word embeddings

Matthew Kay, Cynthia Matuszek, and Sean A Munson. Unequal representation and gender stereotypes in image search results for occupations. CHI 2015
Tolga Bolukbasi, Kai-Wei Chang, James Y. Zou, Venkatesh Saligrama, Adam Tauman Kalai: Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings. NIPS 2016: 4349-4357

Case Study: ads



The importance of being Latanya

Names used predominantly by *black men and women* are much more likely to generate *ads* related *to arrest records*, than names used predominantly by white men and women.

Adfisher tool to automate the creation of *demographic* and *behavioral* profiles

- setting gender = female results in less ads for high-paying jobs (google ads)

Facebook ad platform is facing charges that it has enabled gender-based discrimination against millions of women in a class action suit

Sweeney L. Discrimination in Online Ad Delivery. Communications of the ACM, Vol. 56 No. 5, Pages 44-54.
https://fairlyaccountable.org/adfisher/

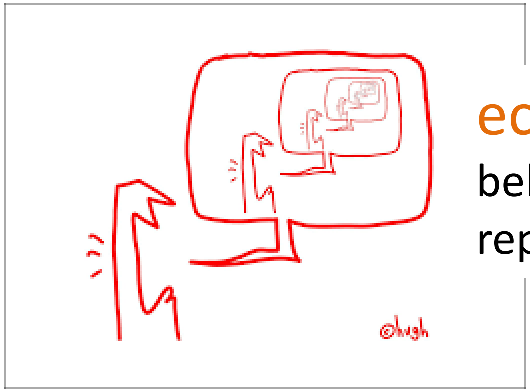
https://techcrunch.com/2018/09/18/facebook-named-in-suit-alleging-job-ads-on-its-platform-unlawfully-discriminated-against-women/?guccounter=1&guce_referrer_us=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_cs=_uzU_r3IoV8zR14P1UckHA

Case study: filter bubbles

personalized searches and recommendations

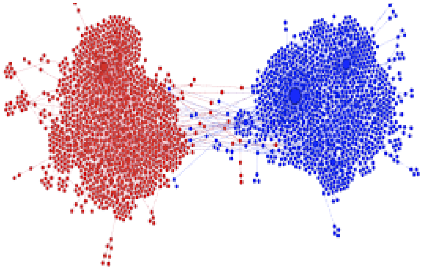
filter bubble a state of intellectual isolation where users become separated from information that disagrees with their viewpoints,

Social media has become the main source of news online with more than 2.4 billion internet users, nearly **64.5%** receive breaking news from social media instead of traditional media



echo chambers: a situation in which information, ideas, or beliefs are amplified or reinforced by communication and repetition inside a defined system

polarity



https://www.forbes.com/sites/nicolemartin1/2018/11/30/how-social-media-has-changed-how-we-consume-news/#18ae4c093c3c

Potential Harms from Automated Decision-Making

Individual Harms		Collective / Societal Harms
Illegal	**Unfair**	
Loss of Opportunity		
Employment Discrimination E.g. Filtering job candidates by race or genetic/health information	E.g. Filtering candidates by work proximity leads to excluding minorities	**Differential Access to Job Opportunities**
Insurance & Social Benefit Discrimination E.g. Higher termination rate for benefit eligibility by religious group	E.g. Increasing auto insurance prices for night-shift workers	**Differential Access to Insurance & Benefits**
Housing Discrimination E.g. Landlord relies on search results suggesting criminal history by race	E.g. Matching algorithm less likely to provide suitable housing for minorities	**Differential Access to Housing**
Education Discrimination E.g. Denial of opportunity for a student in a certain ability category	E.g. Presenting only ads on for-profit colleges to low-income individuals	**Differential Access to Education**
Economic Loss		
Credit Discrimination E.g. Denying credit to all residents in specified neighborhoods ("redlining")	E.g. Not presenting certain credit offers to members of certain groups	**Differential Access to Credit**
Differential Pricing of Goods and Services E.g. Raising online prices based on membership in a protected class	E.g. Presenting product discounts based on "ethnic affinity"	**Differential Access to Goods and Services**
	Narrowing of Choice E.g. Presenting ads based solely on past "clicks"	**Narrowing of Choice for Groups**
Social Detriment		
	Network Bubbles E.g. Varied exposure to opportunity or evaluation based on "who you know"	**Filter Bubbles** E.g. Algorithms that promote only familiar news and information
	Dignitary Harms E.g. Emotional distress due to bias or a decision based on incorrect data	**Stereotype Reinforcement** E.g. Assumption that computed decisions are inherently unbiased
	Constraints of Bias E.g. Constrained conceptions of career prospects based on search results	**Confirmation Bias** E.g. All-male image search results for "CEO," all-female results for "teacher"
Loss of Liberty		
	Constraints of Suspicion E.g. Emotional, dignitary, and social impacts of increased surveillance	**Increased Surveillance** E.g. Use of "predictive policing" to police minority neighborhoods more
Individual Incarceration E.g. Use of "recidivism scores" to determine prison sentence length (legal status uncertain)		**Disproportionate Incarceration** E.g. Incarceration of groups at higher rates based on historic policing data

FUTURE OF PRIVACY FORUM

Fairness, Diversity

Responsible data management

Touches open questions of
- Ethics, and
- Law

Many aspects, in this talk:
- Fairness
- Diversity

Fairness Accountability Transparency

Talk outline

- Fairness
  - What is it?
  - Mitigation tasks
- Diversity
  - What is it?
  - Mitigation tasks
- Discussion of other aspects

What is the cause?

Data

- Correctness and completeness Garbage in, garbage out (GIGO)
  - Poorly selected
  - Incomplete
  - Incorrect
  - Outdated
  - Selected with bias

- Data as a social mirror: perpetuating and promoting historical biases

- Sample size disparity
  - learn on majority (Errors concentrated in the minority class)

What is the cause?

Processing

- Algorithms as black boxes
- Output models that are hard to understand
- Unrealistic assumptions
- Algorithms that do not compensate for input data problems
- Output presentation that is faulty (biased, unfair)
- Personalization and recommendation services that narrow instead of expand user options
- Decision making systems that assume correlation implies causation
- BIAS REINFORCEMENT CYCLE

FAIRNESS

Fairness: definition

Fairness ≡ lack of **discrimination** (treat someone differently)

Protected attribute – output should not depend on the values of these attributes, differences should be explained by other attributes (features)

Two general approaches

- Individual fairness
- Group fairness

Individual fairness

General principle: Similar people should be treated similarly

What does “similar” people mean?

Let *V* be a set of individuals

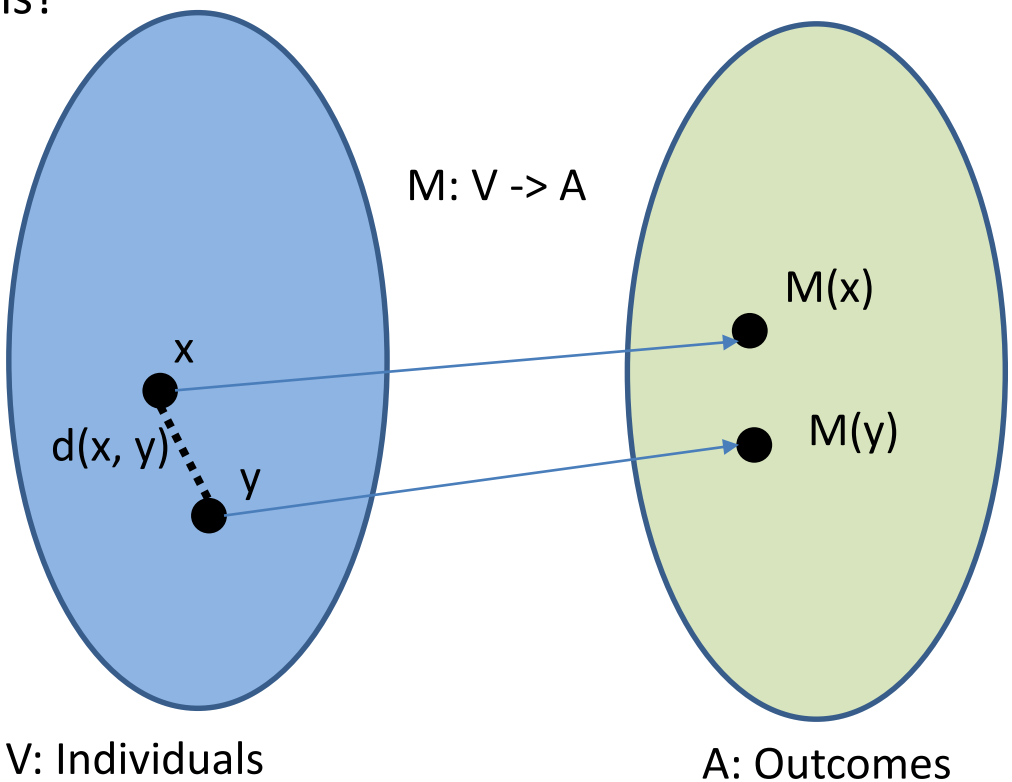
A ***task-specific* distance metric *d*:** V x V -> R

- Expresses *ground truth* (or, best available approximation)
- Public
- Open to discussion and refinement
  - Externally imposed, e.g., by a regulatory body, or externally proposed, e.g., by a civil rights organization

Cynthia Dwork, Moritz Hardt, Toniann Pitassi, Omer Reingold, Richard S. Zemel: *Fairness through awareness*. ITCS 2012: 214-226

Individual fairness

What does "treated similarly" means?



Focus on classification
Classifier *M* maps individuals to outcomes

A: set of classifier outcomes

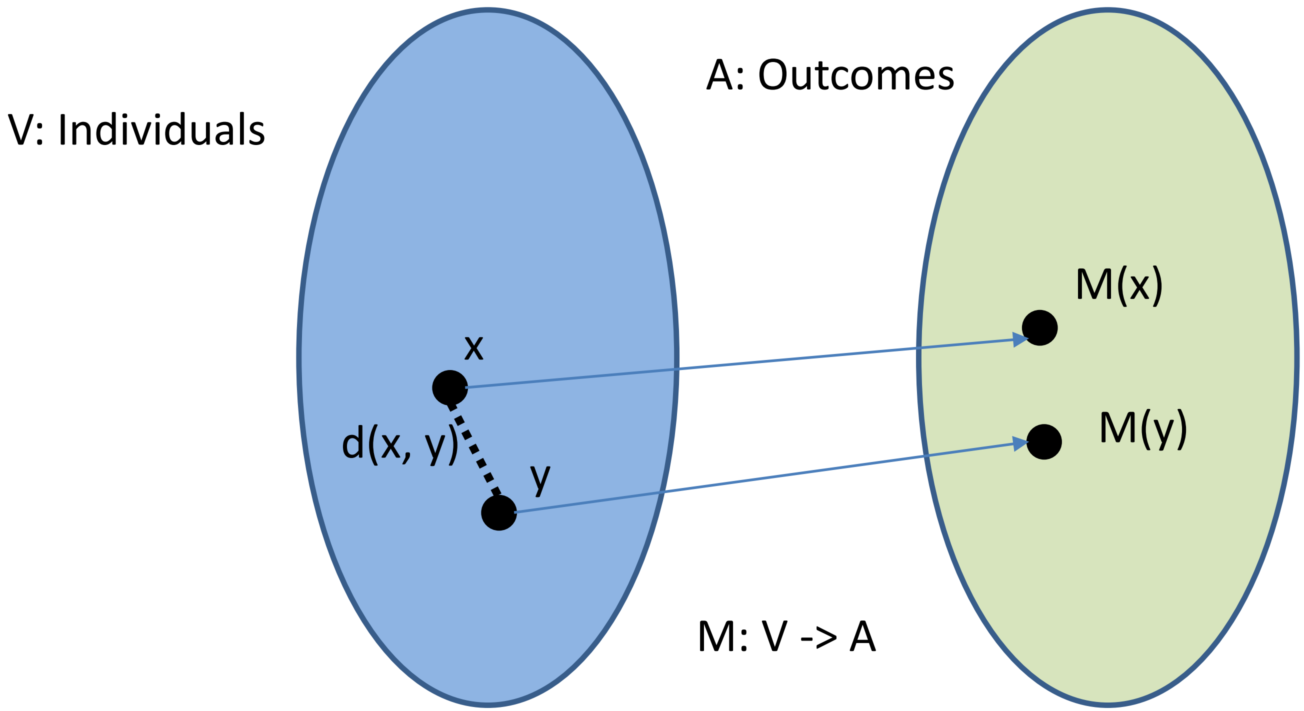
Randomized *mapping* *M: V -> Δ(A)* from individuals to probability distributions over outcomes

- To classify $x \in V$, choose an outcome $a \in A$ according to distribution $M(x)$

Individual fairness

What does "treated similarly" means?

Lipschitz Mapping: a mapping M: V -> Δ(A) satisfies the (*D*, *d*)-Lipschitz property, if for every x, y ∈ V, it holds $D(M(x), M(y)) \leq d(x, y)$



D is a distance between probability distributions
Close individuals map to close distributions

Along this line, other formulations: $d_{IN}(x, y) \leq \epsilon \Rightarrow d_{OUT}(f(x), f(y)) \leq \epsilon'$

Individual fairness

But how are individuals going to be represented as input?

Construct space	Observed space	Decision space
Intelligence	SAT scores	College acceptance
Success in high school	GPA scores	
Propensity to commit crime	Family history	Recidivism
Risk-averseness	Age	

Distances between construct space, observed space and decision space

In many real world situation, there is structural bias in the mapping from CS to OS: unequal treatment of groups

- Researchers have shown that the SAT verbal questions *function differently* for the African-American subgroup

Sorelle A. Friedler, Carlos Scheidegger, Suresh Venkatasubramanian: **On the (im)possibility of fairness.** CoRR abs/1609.07236 (2016)

Group fairness

Protected, or sensitive attribute S

Dataset D divided into groups based on the values of the protected attribute

If S binary and 1 is the "privileged" value, two groups:

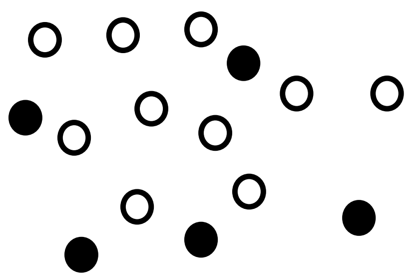
- Privileged group, $S = 1$
- Protected (minority) group, $S \neq 1$

For classification: binary outcome Y, predicted binary outcome $\hat{Y}$

yes the favorable outcome

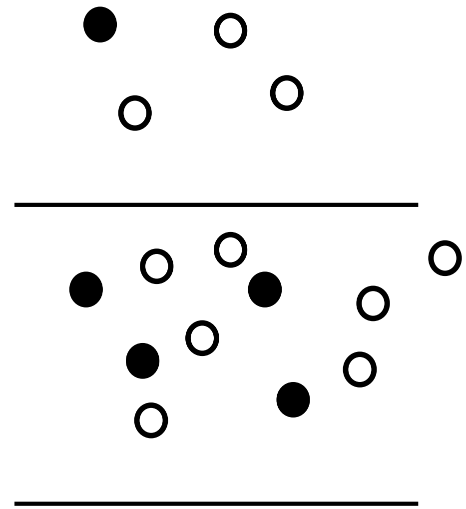
Group fairness

Dataset D



Outcome

Y = Yes



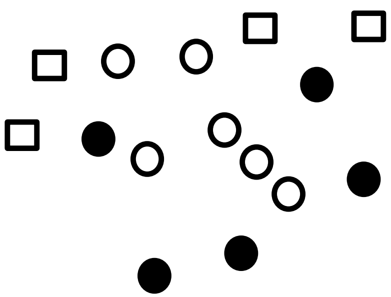
Y = No

Color is the protected attribute
Black group 5 members
White group 10 members
Protected group is the black group

Is this fair?

Group fairness: redundant encoding

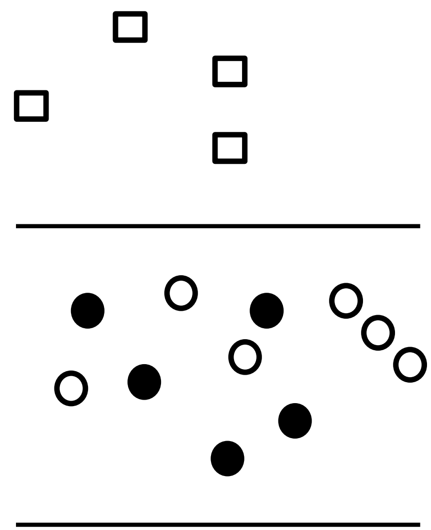
Dataset D



black and white group
black is the protected group
One additional (non protected attribute) shape

Outcome

Y = Yes



Y = No

Selection on shape, not on the protected attribute. Is this fair?

- Redundant encoding, (or, proxies) shape is correlated with color
- Blindness (hiding the value of the protected attribute) does not work

Disparate treatment and impact

Disparate treatment

Illegal practice of treating an entity differently based on a protected characteristic such as race, gender, age, religion, sexual orientation, or national origin.

Disparate impact

Outcome depends on group membership even if people are treated the same way

Disparate impact doctrine solidified in the US after [Griggs v. Duke Power Co. 1971] where a high school diploma was required for unskilled work, excluding black applicants (non job related training)

Redundant encoding

- Discrimination Based on Redundant Encoding

  Redlining:

  well-known form of discrimination based on redundant encoding.

  the practice of arbitrarily denying or limiting financial services to specific *neighborhoods*, generally because its residents are people of color or are poor.“

Illegal in the US

Group fairness

Three basic types of group fairness, based on

- Base rates
- Group-conditioned accuracy
- Calibration

Group fairness: base rates

Compare

$$P[\hat{Y} = yes \mid S = 1]$$

Probability of favorable outcome for privilege group

with

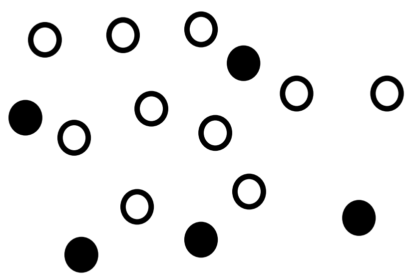
$$P[\hat{Y} = yes \mid S \neq 1]$$

Probability of favorable outcome for minority group

Conditional probabilities evaluated over D

Group fairness: base rates

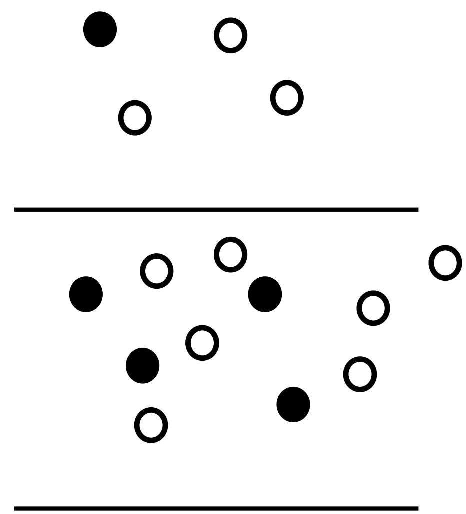
Dataset D



black and white group
black is the protected group

Predicted Outcome

$\hat{Y}$ = Yes



$\hat{Y}$ = No

$$P[\hat{Y} = yes \mid S = "white"] = 3/10$$

$$P[\hat{Y} = yes \mid S = "black"] = 1/5$$

Group fairness: base rates

Compare

$$P[\hat{Y} = yes \mid S = 1]$$

Probability of favorable outcome for privilege group

with

$$P[\hat{Y} = yes \mid S \neq 1]$$

Probability of favorable outcome for minority group

Conditional probabilities evaluated over D

$$\frac{P[\hat{Y} = yes \mid S \neq 1]}{P[\hat{Y} = yes \mid S = 1]}$$

$$1 - (P[\hat{Y} = yes \mid S = 1] - P[\hat{Y} = yes \mid S \neq 1])$$

Group fairness: base rates

Compare

$$P[\hat{Y} = yes \,|S = 1]$$

Probability of favorable outcome for privilege group

with

$$P[\hat{Y} = yes \,|S \neq 1]$$

Probability of favorable outcome for minority group

Conditional probabilities evaluated over D

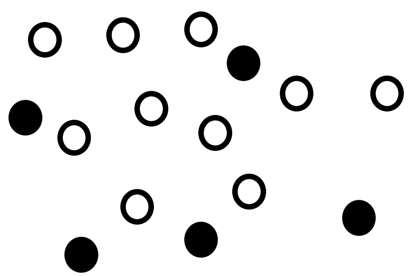
$$\frac{P[\hat{Y} = yes \,|S \neq 1]}{P[\hat{Y} = yes \,|S = 1]}$$

If equal (i.e., ratio 1) demographic parity (statistical parity)
Preserves the input ratio: *demographics of the individuals receiving any outcome same as demographics of the underlying population*

$$1 - (P[\hat{Y} = yes \,|S = 1] - P[\hat{Y} = yes \,|S \neq 1])$$

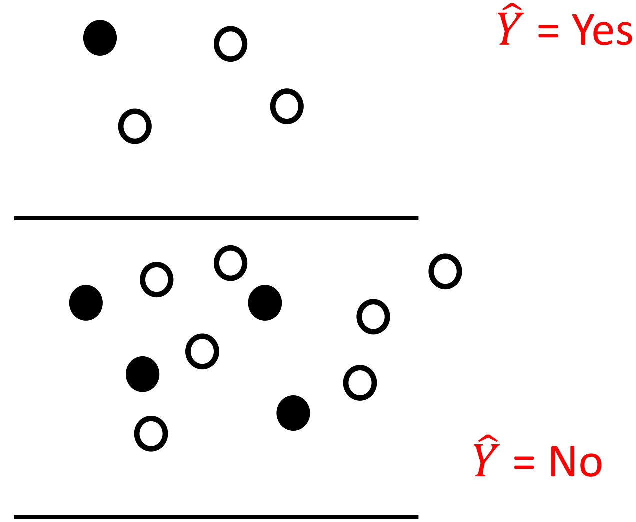
Group fairness: base rates

Dataset D



black and white group
black is the protected group

Predicted Outcome



Does not preserve demographic parity

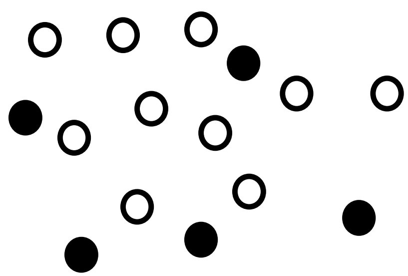
$$P[\hat{Y} = yes \,|S = "white"] = \; 3/10$$

$$P[\hat{Y} = yes \,|S = "black"] = \; 1/5$$

Ratio = 2/3

Group fairness: base rates

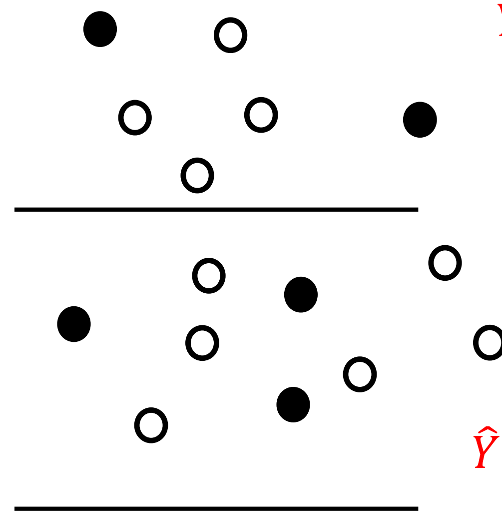
Dataset D



black and white group
black is the protected group

Predicted Outcome

$\hat{Y}$ = Yes



$\hat{Y}$ = No

Preserves demographic parity

$$P[\hat{Y} = yes \mid S = "white"] = 4/10$$

$$P[\hat{Y} = yes \mid S = "black"] = 2/5$$

Group fairness: base rates

Compare

$$P[\hat{Y} = yes \mid S = 1]$$

Probability of favorable outcome for privilege group

with

$$P[\hat{Y} = yes \mid S \neq 1]$$

Probability of favorable outcome for minority group

Conditional probabilities evaluated over D

$$\frac{P[\hat{Y} = yes \mid S \neq 1]}{P[\hat{Y} = yes \mid S = 1]} \leq \tau = 0.8$$

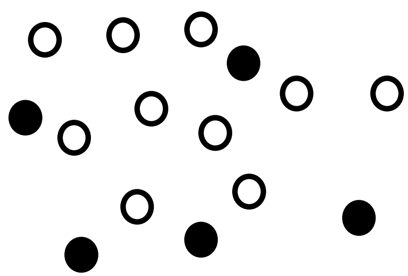
Disparate impact (unintended discrimination)

Based on a generalization of the 80 percent rule advocated by the US Equal Employment Opportunity Commission

Michael Feldman, Sorelle A. Friedler, John Moeller, Carlos Scheidegger, Suresh Venkatasubramanian: **Certifying and Removing Disparate Impact.** KDD 2015

Group fairness: base rates

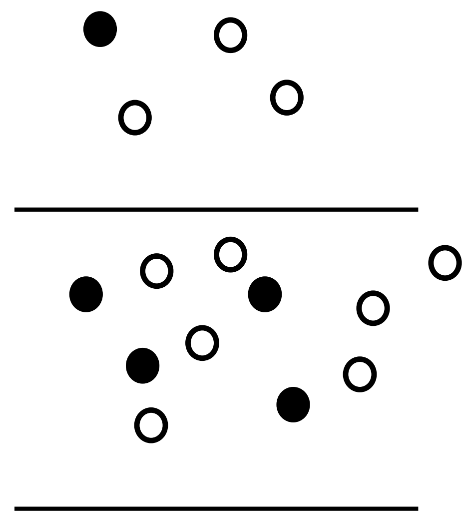
Dataset D



black and white group
black is the protected group

Predicted Outcome

$\hat{Y}$ = Yes



$\hat{Y}$ = No

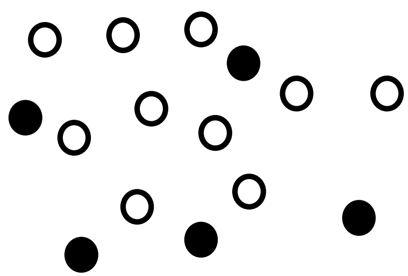
τ = 2/3
Disparate impact

$$P[\hat{Y} = yes \mid S = "white"] = \; 3/10$$

$$P[\hat{Y} = yes \mid S = "black"] = \; 1/5$$

Group fairness: base rates

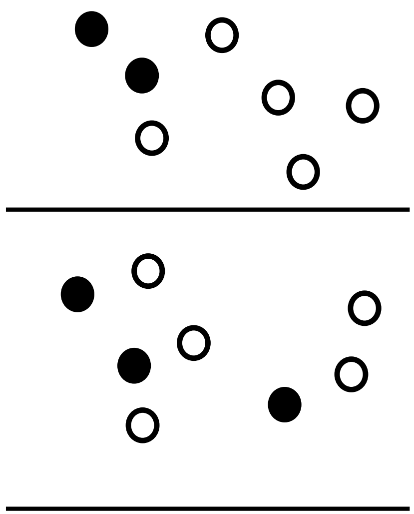
Dataset D



black and white group
black is the protected group

Predicted Outcome

$\hat{Y}$ = Yes



$\hat{Y}$ = No

$$P[\hat{Y} = yes \,|S = "white"] = \; 5/10$$

$$P[\hat{Y} = yes \,|S = "black"] = \; 2/5$$

$\tau = 4/5$
No disparate impact
(no demographic parity)

Group fairness

- Self-fulfilling prophecy

Deliberately *choosing the “wrong" members of the protected group* in order *to build a bad “track record"* for the group
A less malicious vendor simply selects *random members of S* rather than qualified members

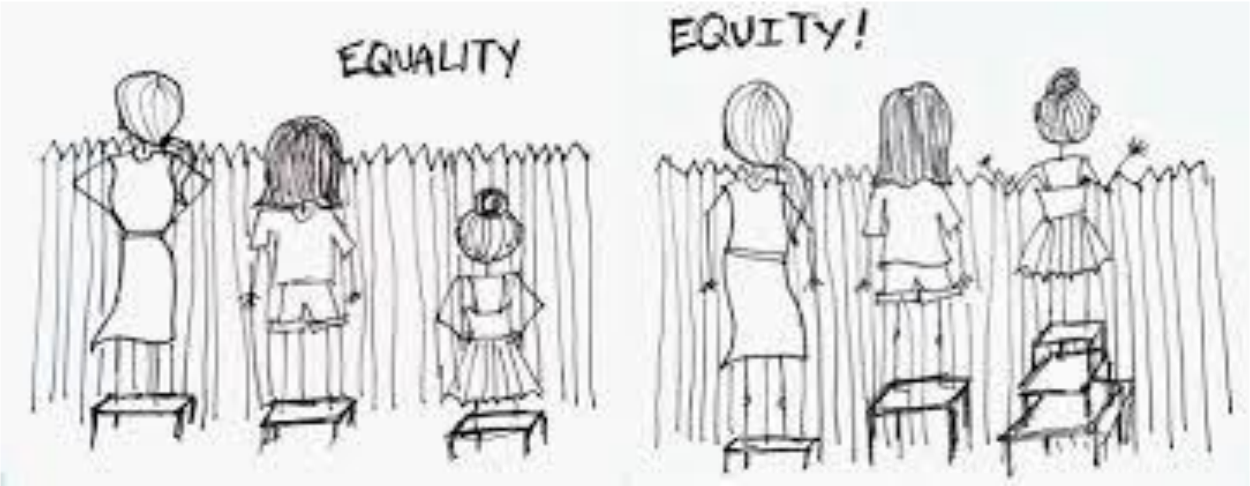
- Reverse tokenism

Deny access to a qualified member of the privileged group
Goal is to create convincing refutations

Discussion

Individual fairness

Statistical parity



Group fairness

Three basic types of group fairness, based on

- Base rates
- Group-conditioned accuracy
- Calibration

Group fairness: accuracy

Considers the performance of the classifier: whether the errors for each group are similar

Look at Y (actual) and $\hat{Y}$ (predicted) for each group

For example:

$P[\hat{Y} = yes \mid Y = yes, S = 1]$ *True positive rate* for privilege class

$P[\hat{Y} = yes \mid Y = yes, S \neq 1]$ *True positive rate* for minority class

Look at 1 – TPR, etc

Many variations with different names, for example:

equalized odds ensures that no error type disproportionately affects any particular group

Group fairness: calibration

Probabilistic classifiers: output the probability that an individual belongs to the positive class

Probability estimates should be *well calibrated*: if the algorithm identifies a set of people as having a probability *p* of constituting positive instances, then approximately a *p* fraction of this set should indeed be positive instances

A predictor that outputs a probability *p* is said to be well calibrated if

$$P[Y = yes \mid \hat{Y} = p] = p$$

We ask the classifier to be well calibrated for both groups, for all *p* values

$$P[Y = yes \mid \hat{Y} = p, S = 1] = P[Y = yes \mid \hat{Y} = p, S \neq 1]$$

It has been shown that a classifier cannot achieve both calibration and Equalized Odds

Jon M. Kleinberg, Sendhil Mullainathan, Manish Raghavan: ***Inherent Trade-Offs in the Fair Determination of Risk Scores.*** *ITCS 2017: 43:1-43:23*

Talk outline

- Fairness
  - What is it?
  - Mitigation tasks
- Diversity
  - What is it?
  - Mitigation tasks
- Discussion of other aspects

Fairness: mitigation

1. Discover/test for unfairness
2. Enforce fairness

Fairness: mitigation

Most work on classification problems, but

- Recommendation algorithms
- Data Integration
  - source data selection
  - entity resolution
  - data cleaning
- Query exploration
- Ranking
- Search
- Crowdsourcing
- Summarization
- Data visualization

Fairness: Testing

- Discrimination discovery: Given a large database of historical decision records, find discriminatory situations and practices

- Test datasets for example, for correlation among protected and other attributes (e.g., Pearson coefficient. mutual information tests)

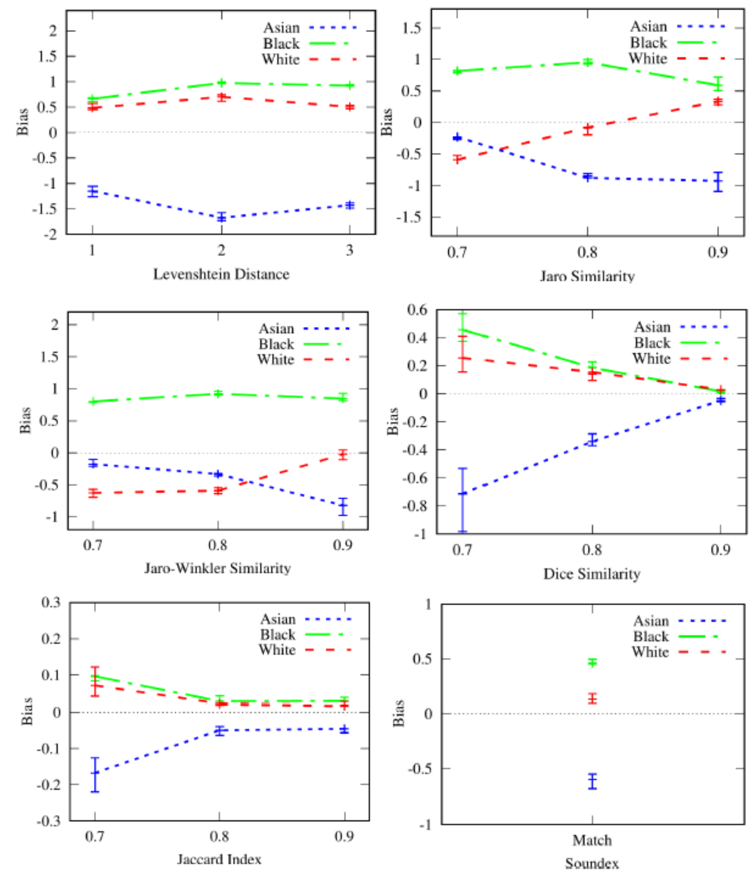
- Test the behavior of specific algorithms

Large body of work and many tools

Fairness: Discovery

Beyond classification: entity resolution

Tested 6 known string matching distances



Preliminary work on name matching

Bias

Mismatch rate for individuals in a specific ethnic group

Compare this to average mismatch rate over all groups

Positive bias: perform better than average

Negative bias: perform worse than average

Not explicit, but redundant encoding (length)

Alexandros Karakasidis, Evaggelia Pitoura: Identifying Bias in Name Matching Tasks. EDBT 2019: 626-629

Ensuring Fairness

Three different approaches

- Pre-processing
  Modify the input data
- Algorithm modification
  Modify the algorithm
- Post-processing
  Modify the output data

Trade-off:
Ensure fairness
Preserve utility

Fairness-aware algorithms: preprocessing

- Reweighting: generate weights for the training examples in each (group, label) combination differently to ensure fairness before classification (Kamiran & Calders, 2012)
- Representation learning:
  - learn a *probabilistic transformation* that edits the attributes and labels in the data with group fairness, individual distortion, and data fidelity constraints and objectives (Calmon et al., 2017)
  - finds a *latent represent*ation that encodes the data well but obfuscates information about protected attributes (Zemel et al., 2013)
- Disparate impact remover: does not modify the labels but edits each attribute so that the marginal distributions based on the subsets of that attribute with a given sensitive value are all equal (Feldman et al., 2015)
- Antidote data: *Add more data* to the input of the recommender to improve fairness with minimum accuracy loss (Rastegarpanah et al., 2019)

Also preferential sampling

Fairness-aware algorithms: modify the algorithm

Design an algorithm that is fair but also preserves as much as possible the quality of the output (utility)

- Depends on the algorithm (classification, how to modify e.g., decision trees)
- For optimization problems:
    - Add regularization constraints to enforce fairness

Recent work on linear ranking functions that use a weight vector to compute a ranking score for items

- A given query f, with a corresponding weight vector, may not satisfy a required fairness constraints.
- Propose a scoring function f' with *a similar weight vector* as f that does *satisfy the constraints*, if one exists.

Abolfazl Asudehy, H. V. Jagadishy, Julia Stoyanovichz, Gautam Das, **Designing Fair Ranking Schemes,** SIGMOD 19

Fairness-aware algorithms: post-processing

Modify the output to enforce fairness and minimize loss of utility

Fair top-k ranking

Given *n* items, select *k* items that maximize the ordering utility and also satisfy a form of statistical parity based fairness

Basic idea:

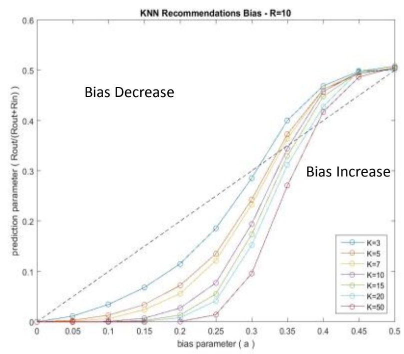
- Given the overall best k-items (the items with the best ordering utility)
- replace the worst items from the majority group in the top-k with the best items from the minority group not in the top-k

Meike Zehlike, Francesco Bonchi, Carlos Castillo, Sara Hajian, Mohamed Megahed, Ricardo A. Baeza-Yates: FA*IR: A Fair Top-k Ranking Algorithm. CIKM 2017: 1569-1578

Fairness/bias in recommendations

Two-sided fairness: Content and user fairness, or bias
User population in groups (men, women)
Recommended items also in groups (based on movie genre)

- Does a recommender increase/decrease bias (preference) of a user group for a specific item group?
  - Women prefer romance but
    does the recommender exaggerates this?



- Group recommendations

Fair and envy-free

V. Tzintzou, E. Pitoura, and P. Tsaparas, Bias Disparity in Recommendation Systems, CoRR abs/1811.01461 (2018)
D. Serbos, S. Qi, N. Mamoulis, E. Pitoura, P. Tsaparas, Fairness in Package-to-Group Recommendations, ACM International Conference on the World Wide Web (WWW), 2017
E. Pitoura, P. Tsaparas, G. Flouris, I. Fundulaki, P. Papadakos, S. Abiteboul, G. Weikum: *On Measuring Bias in Online Information*. SIGMOD Record 46(4): 16-21 (2017)

Talk outline

Fairness
- What is it?
- Mitigation tasks

Diversity
- What is it?
- Mitigation tasks

Discussion of other aspects

DIVERSITY

Diversity

Well studied in Information Retrieval, Search, and Recommendations

Increase *user satisfaction* (initially to address ambiguity and cover all user intents)

- *No useful information is missed*: results that cover all aspects
- *Better user experience*: less boring, more interesting, human desire for discovery, variety, change

Diversity

- Address over-personalization and avoid bias and stereotype reinforcement
- Majority (popularity) based decisions
- Avoid filter bubbles and echo chambers

- *Personal growth*: limited, incomplete knowledge, a self-reinforcing cycle of opinion
- Diversity trumps ability: Diverse perspectives improve collective understanding and collective problem solving (crowdsourcing)

- M Drosou, HV Jagadish, E Pitoura, J Stoyanovich *Diversity in big data: A review* Big data 5 (2), 73-84, 2017
- M Drosou, E Pitoura, *Search result diversification*, SIGMOD record 39 (1), 41-47, 2010

Diversity

Focus on set selection but many applications

Set selection: given a set *P* of *n* items, find a subset *S* $\subseteq$ *P* with the $k \, (\ll n)$ *most diverse* items in *P*

Data Diversity

Besides set selection, many other

- Top-*k* Ranking
- Measures/ranking (centrality measures in graphs (DivRank))
- Diverse results: graph patterns, keyword search, location based queries, skylines queries
- Selecting workers in crowdsourcing
- Data summarization
- Visualization
- Recommendations

Diversity: definitions

Based on:

- Coverage
- Dissimilarity
- DisC diversity
- Novelty (serendipity)
- Network diversity

Diversity: coverage-based definition

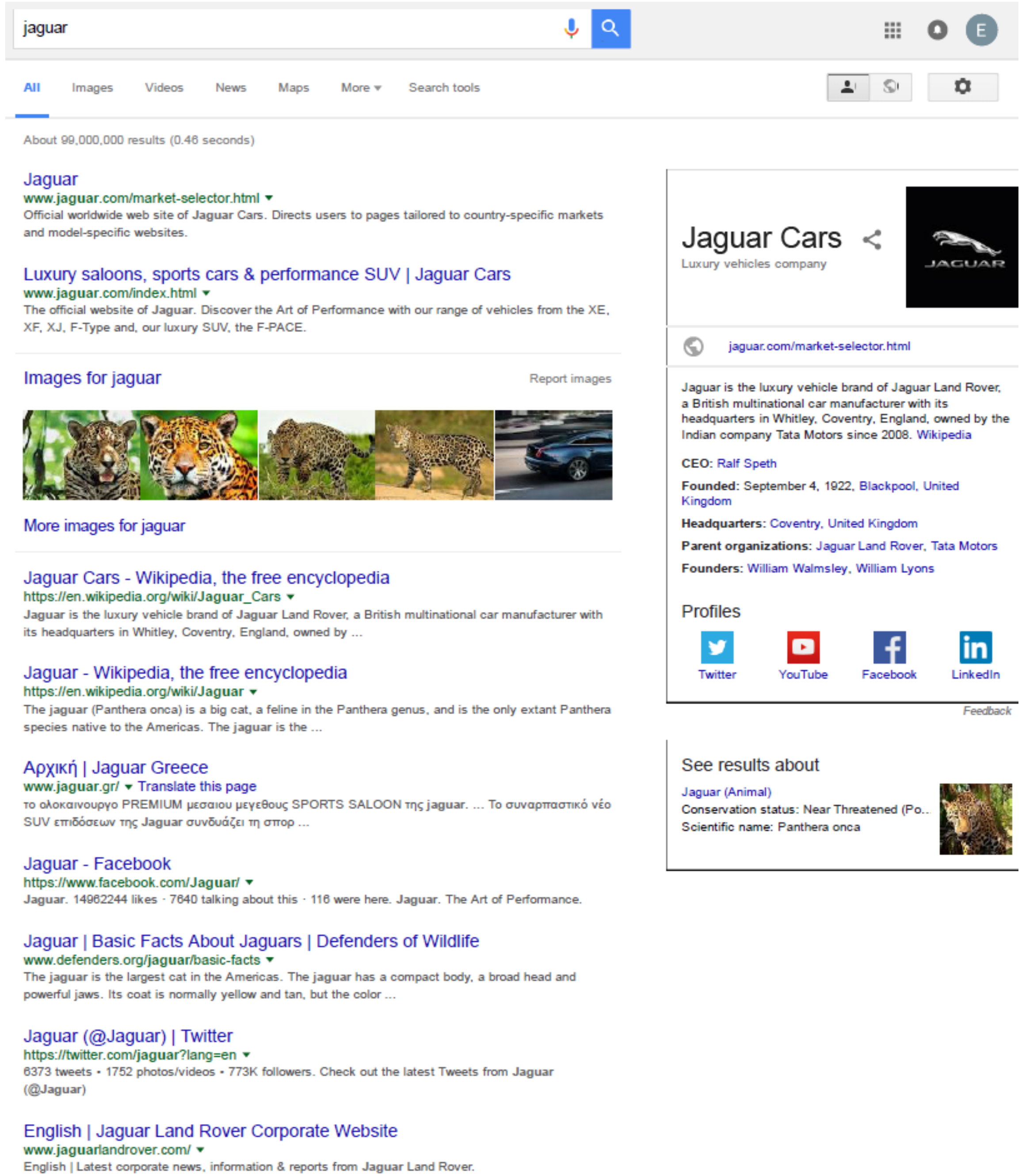
Given a *set of predefined distinct categories* (e.g., concepts, topics, aspects, intents, interpretations, perspectives, opinions, etc)

Select items that cover all (most) of the categories

Rakesh Agrawal, Sreenivas Gollapudi, Alan Halverson, Samuel Ieong: *Diversifying search results*. WSDM 2009

We get the “car” and the “animal” topics but also a “team”, a “guitar”, etc ..

- Assumes “known” topics



Diversity: coverage-based definition

Some similarity with parity based fairness,
if we assume categories = values of the sensitive attribute
(Proportional diversity similar to demographic group parity)

Diversity: distance-based definition

Given multi-dimensional (multi-attribute) items, a *distance measure (metric)* between the items
Select the most different/distant/dissimilar items

- Distance depends on the items and the problem
- Diversity ordering of the attributes

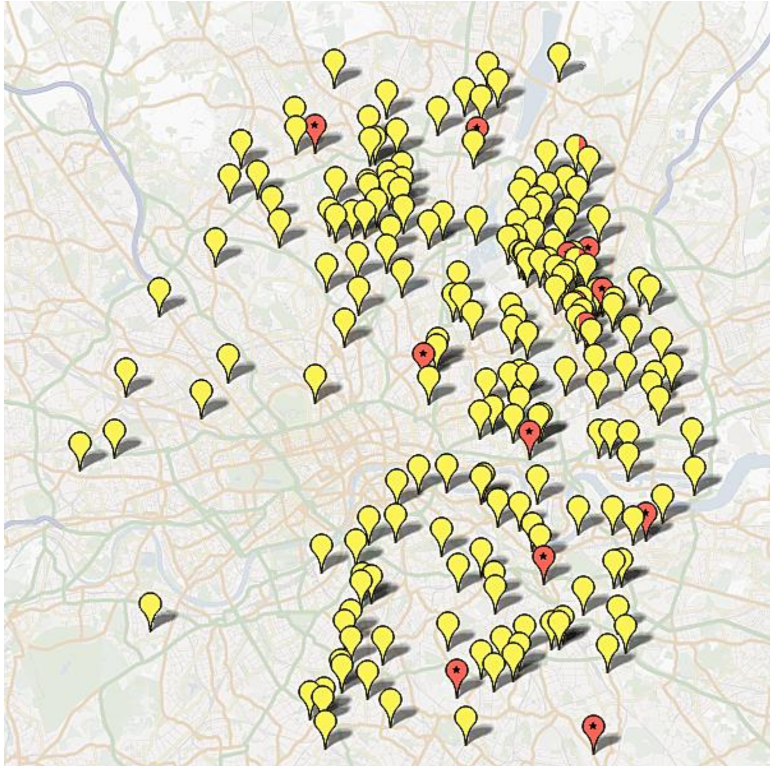
Defining distance/dissimilarity is key

Sreenivas Gollapudi, Aneesh Sharma: *An axiomatic approach for result diversification*. WWW 2009

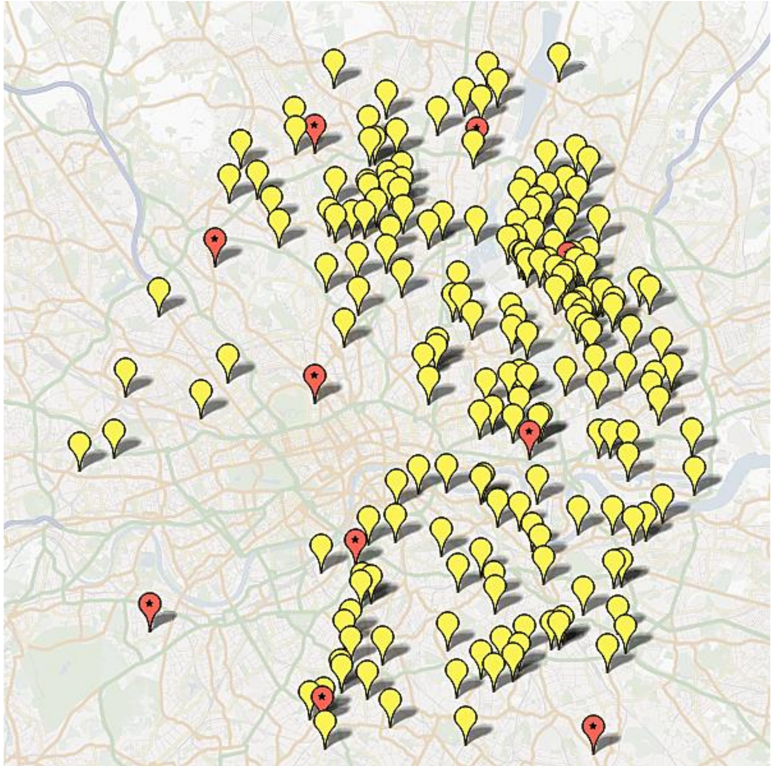
Diversity: distance-based definition

Example: Two-bedroom apartments up to $300K in London

Top based on price *without* (location) diversity



Top based on price *with* (location) diversity

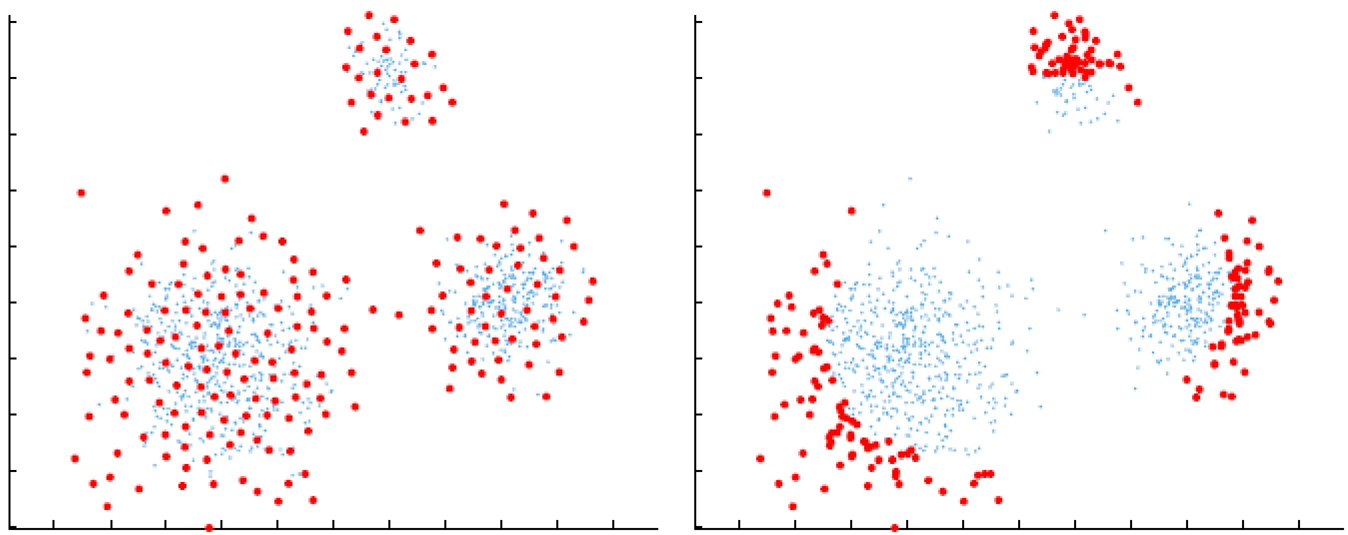


Diversity: distance-based definition

Given a distance measure ***d*** and a function ***f*** measuring the diversity of <u>*set*</u> of *k* items,

$$S^* = \underset{\substack{S \subseteq P \\ |S| = k}}{\operatorname{argmax}} \; f(S, d)$$

$$f_{\mathrm{MIN}}(S,d) = \min_{\substack{p_i, p_j \in S \\ p_i \neq p_j}} d(p_i, p_j) \qquad f_{\mathrm{SUM}}(S,d) = \sum_{\substack{p_i, p_j \in S \\ p_i \neq p_j}} d(p_i, p_j)$$



(a) MaxMin. (b) MaxSum.

Diversity: DisC-diversity

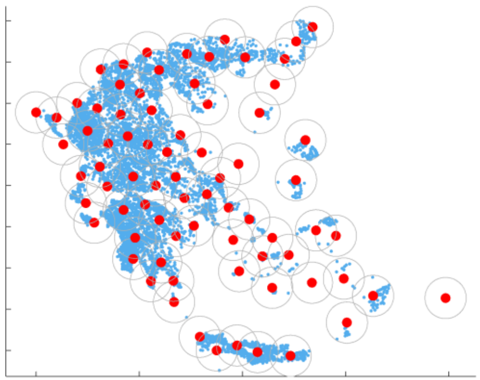
Instead of selecting *k* items, a ***radius r***

Select a representative subset $S \subseteq P$ such that:

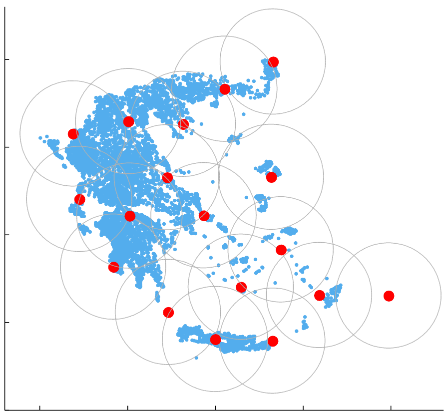
1. For each item *p* in the original set *P*, there is at least one similar item *s* in the selected diverse subset *S*, *d(p, s) <= r* **(coverage)**
2. No two items *s*, *s'* in the diverse subset *S* are similar with each other, *d(s, s') > r* **(dissimilarity)**

Marina Drosou, Evaggelia Pitoura: Multiple Radii DisC Diversity: Result Diversification Based on Dissimilarity and Coverage. ACM Trans. Database Syst. 40(1): 4:1-4:43 (2015)
Marina Drosou, Evaggelia Pitoura: DisC diversity: result diversification based on dissimilarity and coverage. PVLDB 6(1): 13-24 (2012)

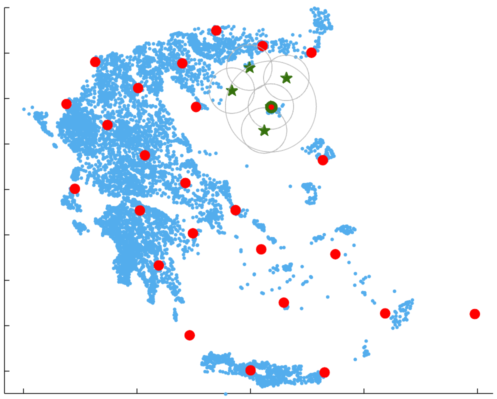
Diversity: DisC-diversity



Zoom-in



Zoom-out



Local zoom

- Small r: more and less dissimilar items (*zoom in*)
- Large r: less and more dissimilar items (*zoom out*)
- Local zooming at specific items

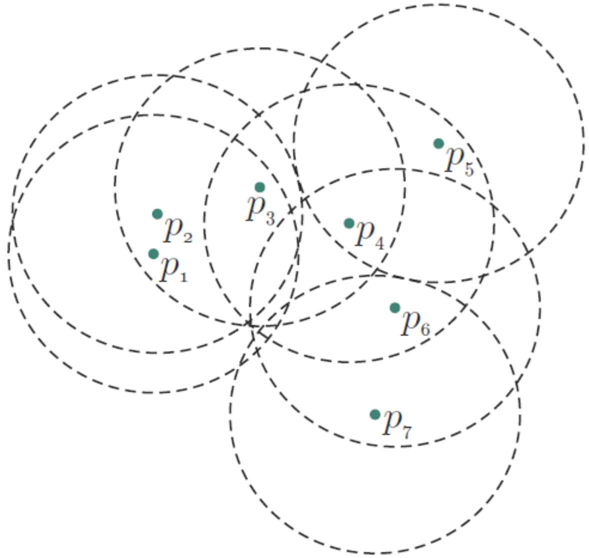
r < smallest distance, |S| = n

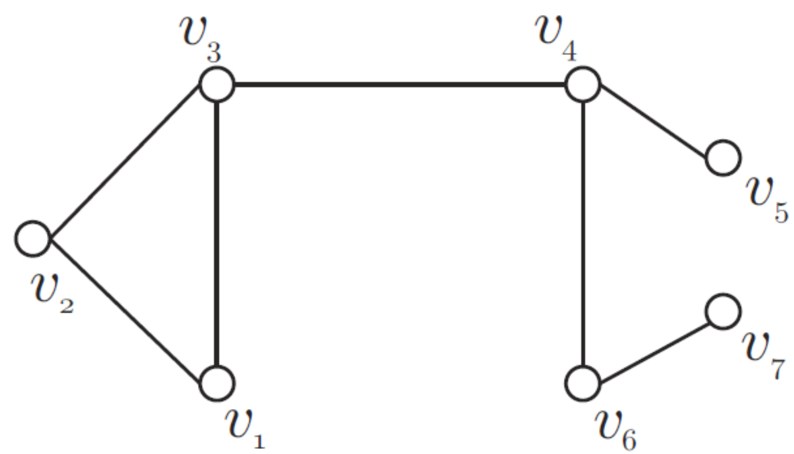
r > largest distance, |S| = 1

Diversity: DisC-diversity

Model the problem as a **graph**

- Items are nodes
- There is an edge between two nodes, if distance ≤ r





Equivalent to finding a minimal

- Independent (no edge about nodes in the set) and
- Dominating (all nodes outside connected with at least one inside)

subset of the corresponding graph (aka maximal independent subset)

Diversity: novelty definition

Given *the history of items* seen in the past, select the items that are the most diverse (coverage, distance) with respect to what a user (or, a community) has seen in the past

- *Marginal relevance*
- *Cascade (evaluation) models:* users are assumed to *scan result lists* from the top down, *eventually stopping* because either their information need is satisfied or their patience is exhausted

Relevant concept: serendipity
represents the “unusualness" or “surprise“
(some notion of semantics – the guitar vs the animal)

Charles L. A. Clarke, Maheedhar Kolla, Gordon V. Cormack, Olga Vechtomova, Azin Ashkan, Stefan Büttcher, Ian MacKinnon: *Novelty and diversity in information retrieval evaluation.* SIGIR 2008
Yuan Cao Zhang, Diarmuid Ó Séaghdha, Daniele Quercia, Tamas Jambor: *Auralist: introducing serendipity into music recommendation.* WSDM 2012

Diversity: Network

Not a single item, but a user in a network
Look at the neighbors of each node

Homophily *"Ομοιος ομοίω αεί πελάζει", (Plato)* **)** *"Birds of a feather flock together"* :
users in a network are similar to their neighbors
Caused by two related social forces

- *Selection:* People seek out similar people to interact with
- *Social influence:* People become similar to those they interact with

Both processes contribute to homophily and lack of diversity, but

- Social influence leads to community-wide homogeneity
- Selection leads to fragmentation of the community

May be reinforced by link recommendation algorithms

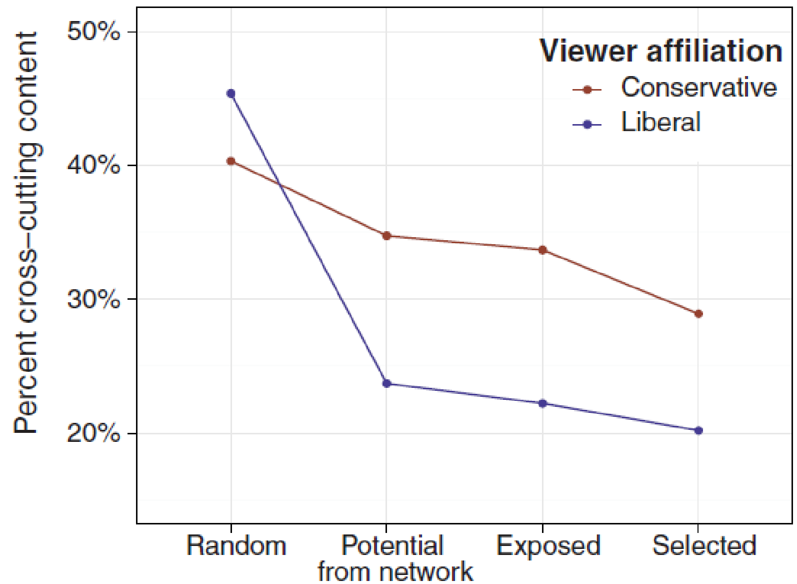
Talk outline

- Fairness
  - What is it?
  - Mitigation tasks
- Diversity
  - What is it?
  - Mitigation tasks
- Discussion of other aspects

Measuring network diversity

Facebook study, three stages of content exposure

- Friends network
- Feeds
- Clicks



Bakshy, Eytan, Solomon Messing, and Lada A. Adamic. *Exposure to Ideologically Diverse News and Opinion on Facebook.* Science 348:1130–1132, 2014

Enforcing diversity

Diversity is just *one of the criteria* in data selection or ranking

Preserve the quality of the output, e.g., relevance in IR or accuracy in recommendations (utility)

MaxSum diversification: maximize the sum (average) relevance (r) and dissimilarity

$$score(S) = (k-1)\sum_{u \in S} r(u) + 2\lambda \sum_{u,v \in S} d(u,v)$$

MaxMin diversification: maximize the minimum relevance (r) and dissimilarity

$$score(S) = \min_{u \in S} w(u) + \lambda \min_{u,v \in S} d(u,v)$$

Diversity-aware algorithms

Greedy set selection

Build the set incrementally, by selecting the item (or, pair of items) with the largest increase of the objective function

- Appropriate re-writing of the maxmin-maxsum dispersion problems in facility location (OR) (approximation bounds)

Constrained optimization problem

Cluster-based algorithms

- Cluster the items and select the cluster centers

Post-processing

Interchange (swap) methods: start with the top-*k* relevant items and replace items that improve the objective function

Grasshoper (random walk on graphs)

Achieving (Network) Diversity

Improve awareness

Blue Feed, Red Feed site -- See Liberal Facebook and Conservative Facebook, Side by Side

http://graphics.wsj.com/blue-feed-red-feed/

Is your news feed a bubble? -- PolitEcho shows you the political biases of your Facebook friends and news feed.

http://politecho.org/

Link recommendation algorithms

Content recommendation algorithms (e.g., feed selection algorithms)

Talk outline

Fairness

- What is it?
- Mitigation tasks

Diversity

- What is it?
- Mitigation tasks

Discussion of other aspects

Fairness and diversity

Diversity of data and opinions

How does diversity of data (or, opinions) presented to individuals or groups affects fairness in decision making?

Does lack of (opinion, data) diversity leads to biased or discriminatory behavior?

Privacy

Privacy legislation cares about *an action* (storage, or use of personal data) independently of the consequences
Discrimination legislation cares about *consequences* (unfair treatment) independently of the mechanism

Some relation between privacy and group fairness

Finding if people having attribute *S* were discriminated is some how similar to inferring attribute *S* from a database in which:

- the attribute S was removed
- a new attribute (the decision), which is based on S, was added

This is similar to trying to reconstruct a column from a privacy-scrubbed dataset

Principles for Algorithmic Transparency and Accountability

1. Awareness
2. Access and redress
3. Accountability
4. Explanation
5. Data Provenance
6. Auditability
7. Validation and Testing

https://www.acm.org/binaries/content/assets/public-policy/2017_usacm_statement_algorithms.pdf

Conclusions

With ~~great power~~ data comes great responsibility

Many interesting problems

testing and detecting (bias, lack of completeness, lack of diversity, ..)

fair (diverse, …) by design

awareness (explanations, visualization, ..)

accountability (through transparency, e.g., provenance)

….

Not just for classical ML but for many ML-supported, or not data management tasks

Thank you!