

Ethics of Algorithmic Decision Making



Liwen Peng

Mentor: Sarah Teichman



What People May Be Concerned about...

- Privacy
- Freedom
- Independence
- Fairness



Case Study about
Fairness :
Algorithmic
Decision Making
in Criminal
Justice System



Case Study about Fairness : Algorithmic Decision Making in Criminal Justice System

- Software/Algorithm: COMPAS provided by Northpointe
- Concern : Software is biased against black people.
 - MORE likely to falsely flag black defendants as future criminals;
 - MORE likely to mislabel white defendants as low risk
- However, Northpointe argues that COMPAS is fair.
- Who is right?



Definition of Fairness

- Equal accuracy
- Equal False positive rates and false negative rates
- Equal False discovery rates and false omission rates
- Equal positive decision rates

	Y=1	Y=0	$P(Y=1 D)$	$P(Y=0 D)$
D=1	True positive	False positive	$P(Y=1 D=1)$: Positive predictive value	$P(Y=0 D=1)$: False discovery rate
D=0	False negative	True negative	$P(Y=1 D=0)$: False omission rate	$P(Y=0 D=0)$: Negative predictive value
$P(D=1 Y)$	$P(D=1 Y=1)$: True positive rate	$P(D=1 Y=0)$: False positive rate		
$P(D=0 Y)$	$P(D=0 Y=1)$: False negative rate	$P(D=0 Y=0)$: True negative rate		$P(D=Y)$: Accuracy

D: decision
Y: true outcome



Evaluate COMPAS Decision

- Data from:
<https://farid.berkeley.edu/downloads/publications/scienceadvances17/>
- Use BROWARD_CLEAN as the data set
- Use logistic regression and decision tree as comparisons
- Use {age, sex, number juvenile misdemeanors, number juvenile felonies, number prior crimes, and crime degree} as covariates



Evaluate COMPAS Decision

	Classify method	Race	Overall Accuracy	False Positive Rate	False Negative Rate	False Discovery Rate	False Omission Rate	Positive Decision Rate
1	Logistic	white	0.6646091	0.1789137	0.6184971	0.4590164	0.2939560	0.2510288
2	Logistic	black	0.6492637	0.3404255	0.3611860	0.3506849	0.3507853	0.4886212
3	COMPAS	white	0.6666667	0.2460064	0.4913295	0.4666667	0.2647975	0.3395062
4	COMPAS	black	0.6706827	0.3989362	0.2587601	0.3529412	0.2981366	0.5689424
5	Tree	white	0.6728395	0.1853035	0.5838150	0.4461538	0.2837079	0.2674897
6	Tree	black	0.6706827	0.3005319	0.3584906	0.3219373	0.3358586	0.4698795



Evaluate COMPAS Decision

- Result
 - Accuracies are nearly equal for both groups
 - False positive rate is higher for black people (about twice as that of white defendants)
 - False negative rate is higher for white people (about twice as that of black defendants)
- COMPAS decisions show fairness if we look at accuracy



Fairness Conclusion

- Hard to tell which definition is the most appropriate.
- These definitions cannot be satisfied at the same time.



REASONS OF BIAS

- A result of the difference in conviction history, which itself is linked to inequities in the criminal justice system.
- On a national scale, Black people are more likely to have prior convictions on their record than white people are.



How to Reduce Bias

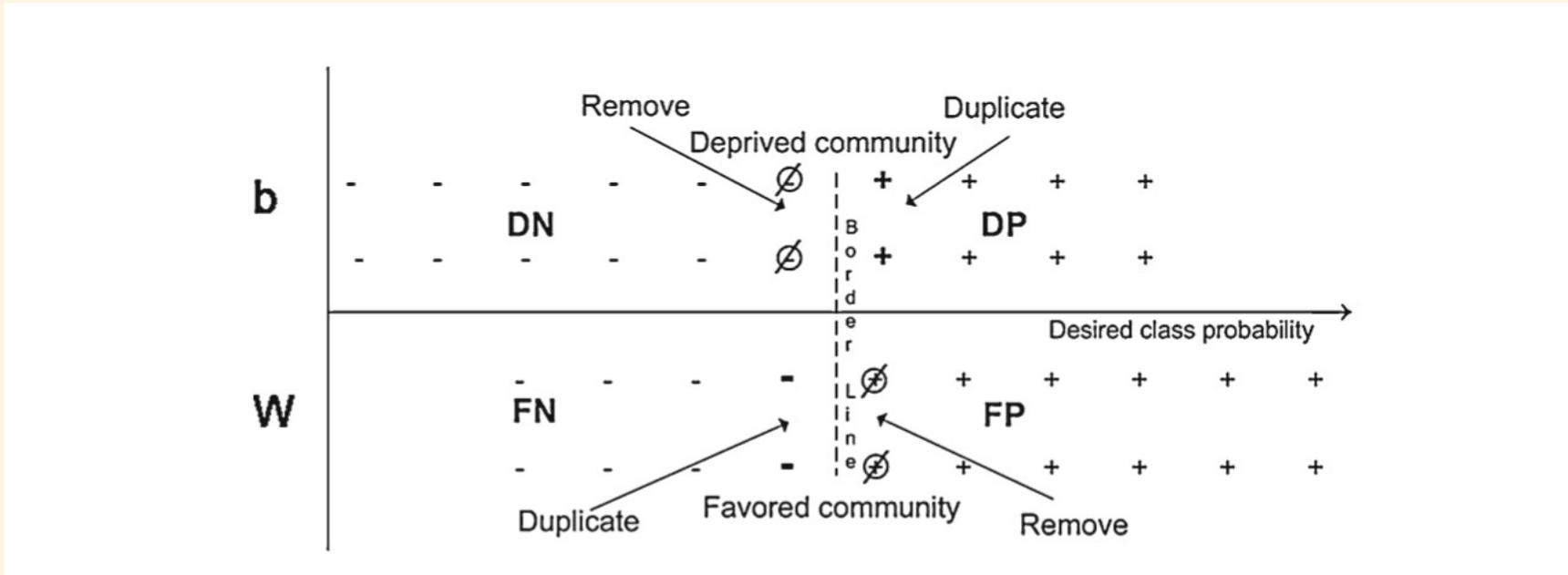
- Pre-processing: change training data
- In-processing: change algorithm
- Post-processing: change prediction results



Try
sampling
from Pre-
processing
Techniques



Basic Idea



DP: deprived community with positive outcome;
 DN: deprived community with negative outcome;
 FP: favored community with positive outcome;
 FN: favored community with negative outcome;



Try Sampling

- The same data frame as before
- Use logistic regression
- Logistic model training on original training data and use this model to predict training set outcomes as rank score for preferential sampling.



Try Sampling

	Method	Race	Overall Accuracy	False Positive Rate	False Negative Rate	False Discovery Rate	False Omission Rate	Positive Decision Rate
1	Original	white	0.6646091	0.1789137	0.6184971	0.4590164	0.2939560	0.2510288
2	Original	black	0.6492637	0.3404255	0.3611860	0.3506849	0.3507853	0.4886212
3	Uniform Sampling	white	0.6687243	0.1916933	0.5838150	0.4545455	0.2853107	0.2716049
4	Uniform Sampling	black	0.6506024	0.3537234	0.3450135	0.3537234	0.3450135	0.5033467
5	Preferential Sampling	white	0.6666667	0.2044728	0.5664740	0.4604317	0.2824207	0.2860082
6	Preferential Sampling	black	0.6465863	0.3776596	0.3288410	0.3631714	0.3426966	0.5234270



Try sampling

-
- No big effect
 - Try other methods to reduce bias

Any Questions?



Thanks



Special thanks to my mentor Sarah



Sources

- <https://farid.berkeley.edu/downloads/publications/scienceadvances17/>
- <https://www.annualreviews.org/doi/pdf/10.1146/annurev-statistics-042720-125902>
- <https://link.springer.com/article/10.1007/s10115-011-0463-8>
- <https://biztimes.com/business-ethics-2/>