|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Paper | 1 |  | Main message of the paper |  |
| Selective labels | Non-representative sample, selectively labelled train data  *Contraction* as solution | They don’t use SQF data. They just study a problem the SQF data probably has. |  |  |
| Residual Unfairness[[1]](#footnote-1) | Non-representative sample, the difference between train and target pop | Model: Logistic regression  Target: criminal possession of weapon  Train pop: SQF  Target pop: whole NYC pop  PA: race |  |  |
| Bias in Bias out | Bias reversal and bias inheritance | Synthetic data from SQF  Task 1  Model: log reg  Target: possession of contraband (given searched)  Task 2  Model: log reg  Target: being searched  Task 3  Model: log reg  Target: searched and contraband | Bias in does actually not necessarily mean bias out but it depends on a) the data (is it representative?) and b) your task (what is the target, what do you condition on)? |  |
| Precinct or Prejudice |  | Model:  Target: possession of weapon  Train data: SQF limited on the stops due to CPW |  |  |
| Data Transparency |  | Model: six different models (log reg, RF, …)  Target: Arrest  PA: Race, Sex |  |  |
|  |  | Target: Possesion of an illegal weapon  (“task is to predict whether a pedestrian should be stopped on the suspicion of having an illegal weapon or not (positive class”) |  |  |
|  |  |  |  |  |

1. What are the main theoretical concepts and statement the paper makes and how are they illustrated with the SQF dataset?
2. What questions with respect to the dataset are they trying to answer?/ How are they using the dataset? What is the target? What is the feature? What is the model?

In der Präsentation:

Dataset “naiv” verwenden um fairness notions klar zu machen

In der Arbeit:

Zeigen, welche challenges der Datensatz mit sich bringt, wie der Datensatz bis jetzt genutzt wurde, den Datensatz in den größeren Kontext bringen und anhanddessen argumentieren welche fairness Metriken sich hier wie auswirken (und warum sie deshalb passender oder weniger passend sind)

1. Widersprechen sich diese zwei paper? [↑](#footnote-ref-1)