Data Augmentation for Fairness under Spatio-Temporal Distribution Shift and Class Imbalance

S. Siamak Ghodsi
L3S-Hannover & Dept. of Math and CS, Free University Berlin

INTRODUCTION

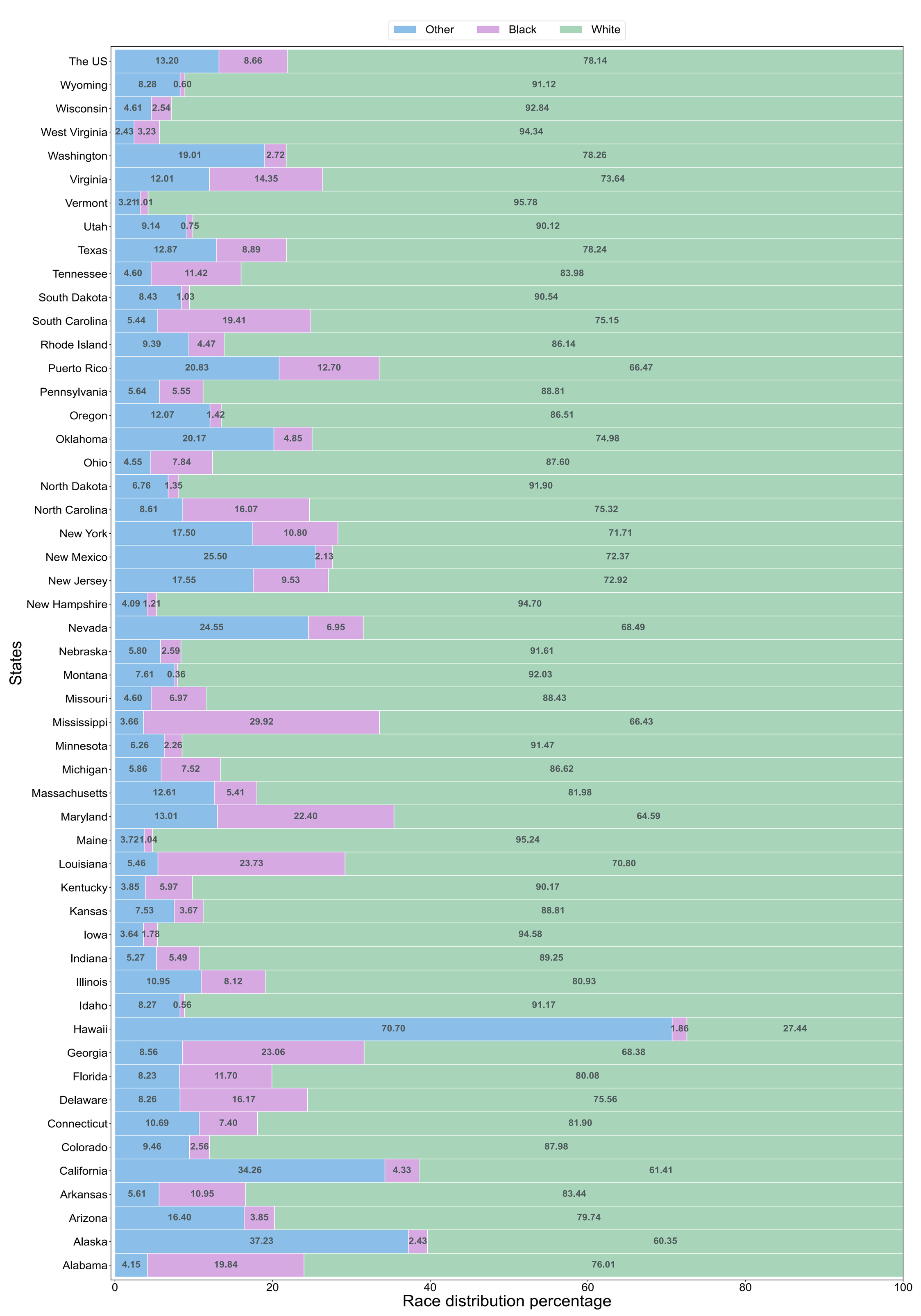
Distribution shift refers to difference in training and deployment sets. It could happen due to spatial (e.g. geographical) and temporal differences in data. Demographic dissimilarities could easily result in biased models. In the presence of class imbalance bias, training under distribution shift can lead to highly discriminatory decisions.

MOTIVATION

The recently released Census dataset for the United States provides a good example of distribution shift and imbalance. It has spatial and temporal shifts among each state's data to the other states and has a wide range of imbalance ratio (IR) difference between states.

CASE STUDY

Spatial distribution shifts and Imbalance bias: The racial distribution of a the 2019 datasets shows a lot of differences.



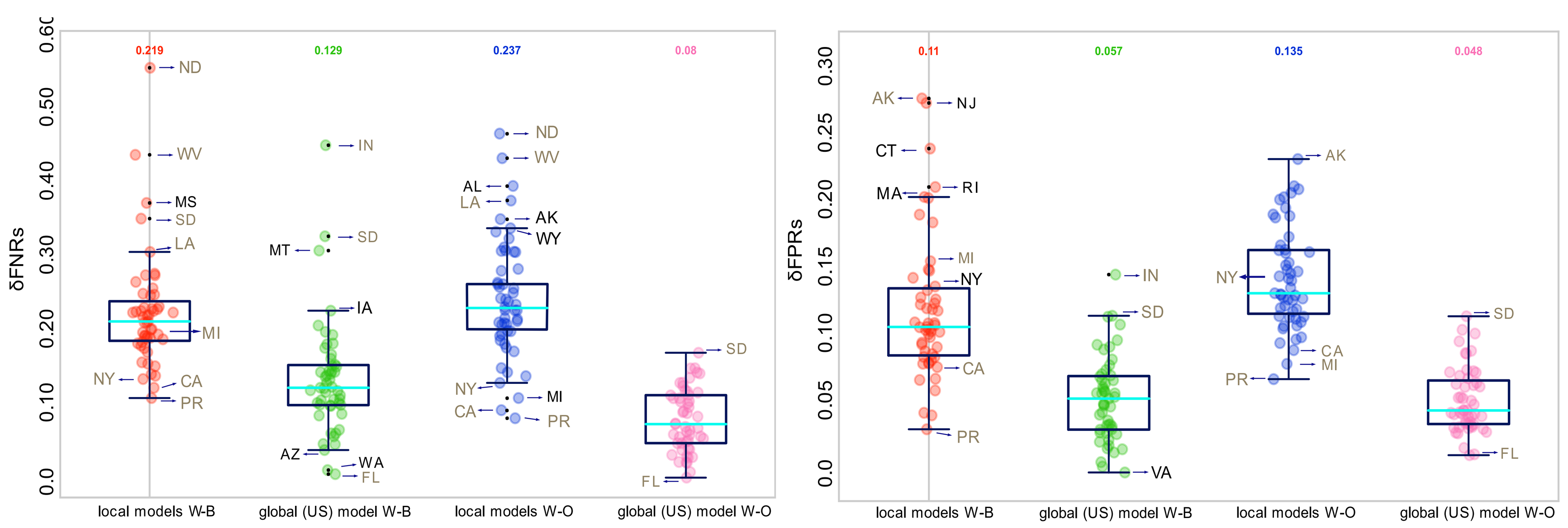
RESEARCH QUESTIONS

1) **Local vs Global model:** *How would local models learned from particular states compare to a global model trained upon data from the whole US, w.r.t both predictive and fairness-related performance?*

$$\delta FNR = \left| P(\hat{Y}=0|Y=1,g=w) - P(\hat{Y}=0|Y=1,g=b/o) \right|$$

$$\delta FPR = \left| P(\hat{Y}=1|Y=0,g=w) - P(\hat{Y}=1|Y=0,g=b/o) \right|$$

$$Eq.Op = |\delta FPR| + |\delta FNR|$$

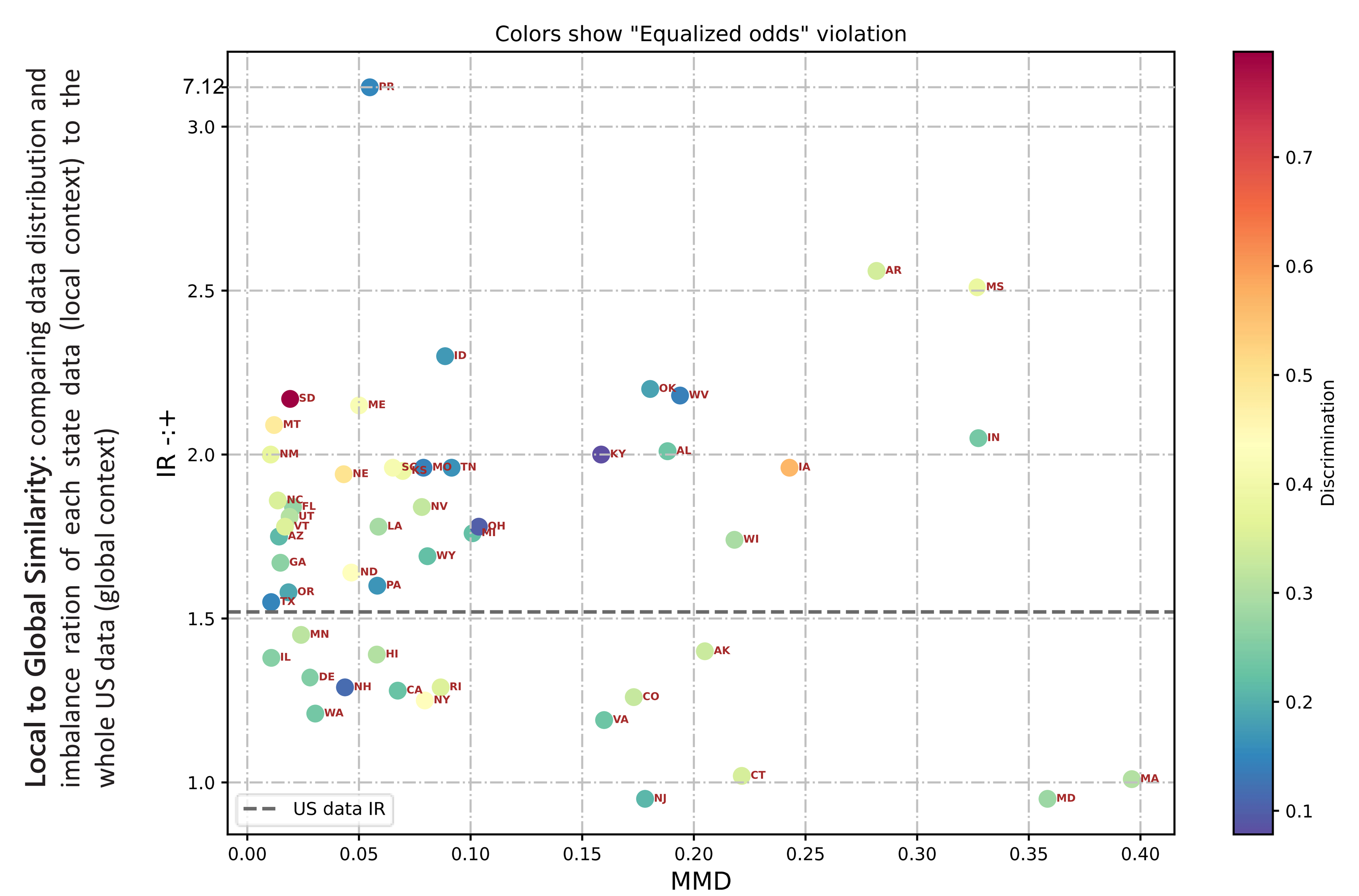


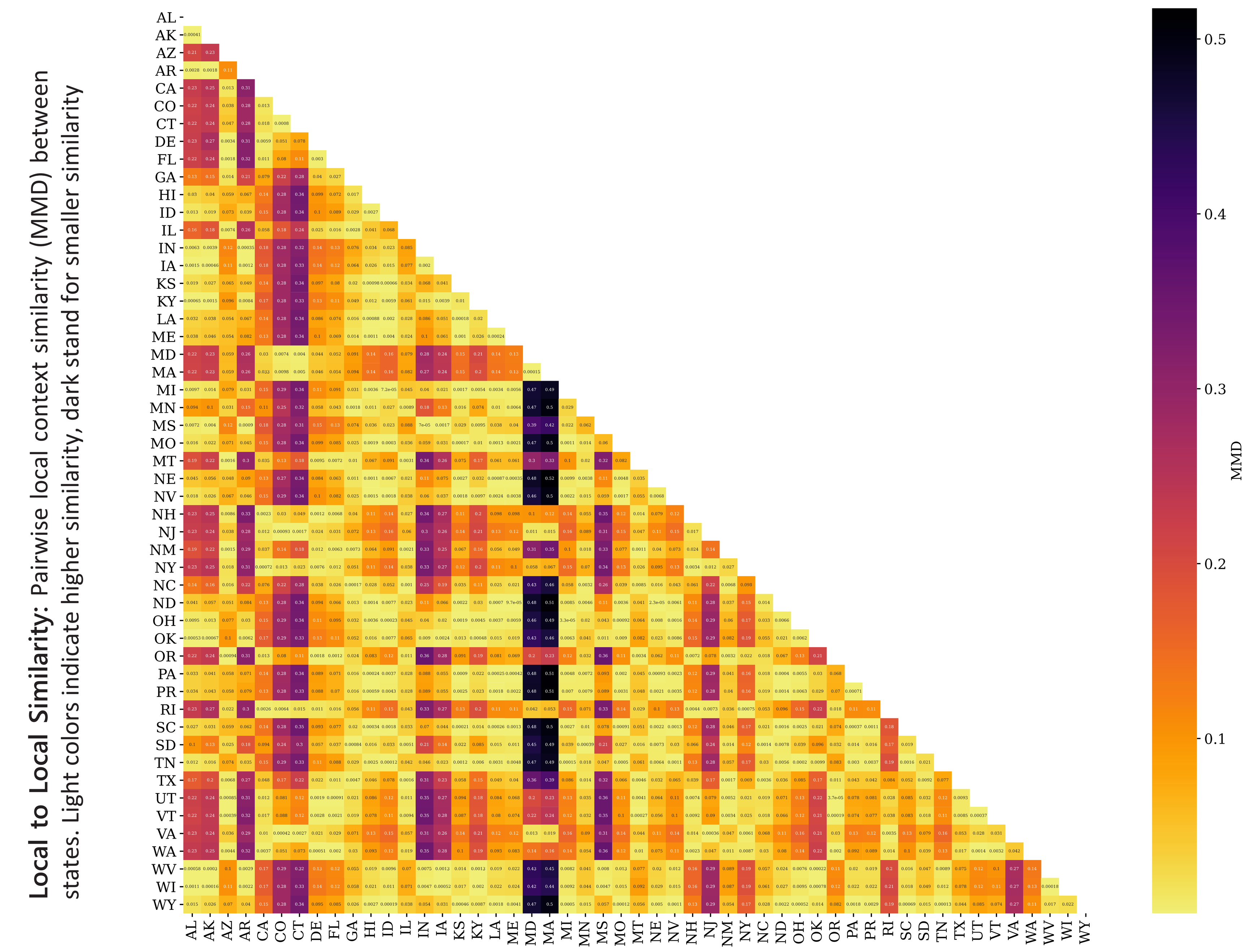
2) **Using context similarity to understand spatial differences:** *How can we detect context similarity, i.e., similar states, which can be used to predict how a model will perform to a different context/state, w.r.t both predictive and fairness-related performance?*

$$MMD^2(P,Q) = \|\mu_P - \mu_Q\|^2_{\mathcal{H}}$$

$$MMD^2(P,Q) = \left\| \frac{1}{n}\sum_{i=1}\phi(x_i) - \frac{1}{m}\sum_{i=1}\phi(v_i) \right\|^2_{\mathcal{H}}$$

$$MMD^2(X,V) = \frac{1}{m(m-1)}\sum_i\sum_{j\neq i}k(\mathbf{x_i},\mathbf{x_j}) - 2\frac{1}{m.m}\sum_i\sum_j k(\mathbf{x_i},\mathbf{v_j}) + \frac{1}{m(m-1)}\sum_i\sum_{j\neq i}k(\mathbf{v_i},\mathbf{v_j})$$





CHALLENGES AND FUTURE DIRECTIONS

Challenges: As seen in boxplots, a global model performs less discriminatively compared to local models, but still it doesn't perform fairly/similarly on all the states. So, a problem is high variance in deployment discrimination-score on different states (**unreliability of global model**). Another problem is estimating an application range for local models (**clustering similar context that perform similarly**). But how is similarity defined? Spatial neighbors (**geopolitical similarity**) can be similar, **semantically similar contexts** (based on a similarity score e.g. **MMD**) can be also similar.

Idea: building an augmented fair local-model using the similarity notion that outperforms each single local model and is comparable or even better (less discriminative) than the global model.

Question: Model augmentation using similar context vs Synthetic data augmentation. Which one would perform better?

CONTRIBUTIONS

1. Ghodsi, Siamak, Harith Alani, and Eirini Ntoutsi. "Context matters for fairness - a case study on the effect of spatial distribution shifts." arXiv preprint arXiv:2206.11436 (2022). **Submitted**
2. Ghodsi, Siamak, Harith Alani, and Eirini Ntoutsi. "A context-aware fair learning model using local similarity for augmentation." **Under progress**
3. Ghodsi, Siamak, Vasilis Iosifidis, Arjun Roy, and Eirini Ntoutsi. "Fair-SMOTEBoost: Sub-group correction for parity-based cumulative fairness-aware boosting." **Under progress**

Find more about me:

- Webpage: https://siamakghodsi.github.io/
- Email: ghodsi@{l3s.de, zedat.fu-berlin.de}
- Twitter: @SiamakGhodsi

NoBIAS

This work has received funding from the European Union's Horizon 2020 research and innovation programme under Marie Sklodowska-Curie Actions (grant agreement number 860630) for the project "NoBIAS - Artificial Intelligence without Bias".