UNIVERSITÉ TOULOUSE CAPITOLE

IRIT
Institut de Recherche en Informatique de Toulouse

Fairness assessment & unfairness mitigation

José Gonzalo Domínguez - jose.dominguez@ut-capitole.fr

1. Measurement

2. Quantitative definition

3. Practical application

4. References

1. Definition

Fairness as **unobservable theoretical construct.**

Measurement model from quantitative social sciences as a framework for understanding fairness in computational systems.

$$q_{tjkl} = \mu_{jkl} + \sum_{i} \tau_{ijklt}$$

$$y_{ijkl} = \mu_{jkl} + \left(\sum_{k^* <= k} \sum_{t=1}^{T_{ijk^*l^*}} w_{ijk^*l^*t} \tau_{ijk^*l^*t} \right),$$

[2] Abigail Z. Jacobs and Hanna Wallach. Measurement and fairness. In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, FAccT '21, 375–385. New York, NY, USA, 2021. Association for Computing Machinery. URL: https://doi.org/10.1145/3442188.3445901, doi:10.1145/3442188.3445901.

2. Individual and group fairness

Individual fairness, which requires that similar people be treated similarly, and **group fairness**, which requires that different groups of people, such as groups defined in terms of different demographic factors, be treated similarly.

```
selection_rates = MetricFrame(
    metrics=selection_rate, y_true=y_true, y_pred=y_true, sensitive_features=sensitive
)
```

[2] Abigail Z. Jacobs and Hanna Wallach. Measurement and fairness. In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, FAccT '21, 375–385. New York, NY, USA, 2021. Association for Computing Machinery. URL: https://doi.org/10.1145/3442188.3445901, doi:10.1145/3442188.3445901.

2. Quantitative definition

→ Statistical or demographic parity

The prediction should be independent of the protected attribute. It is invoked to address disparity under the US Equal Employment Opportunity Comission's "four-fifths rule".

→ Bounded group loss

It formalizes the requirement that the predictor's loss remain below some acceptable level for each protected group.

Alekh Agarwal, Miroslav Dudík, and Zhiwei Steven Wu. Fair regression: quantitative definitions and reduction-based algorithms. In ICML, volume 97 of Proceedings of Machine Learning Research, 120–129. PMLR, 2019. URL: http://proceedings.mlr.press/v97/agarwal19d.html.

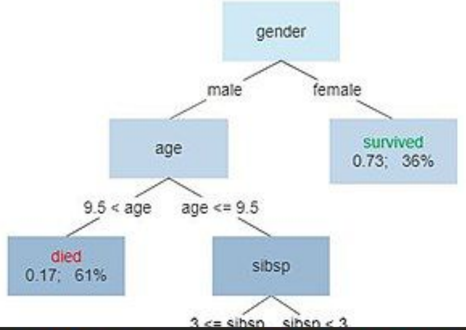
2. Parity constraints

Fairlearn

Group fairness is typically formalized by a set of constraints on the behaviour of the predicator called parity constraints.

Demographic parity in Binary classification

$$E[h(X) \mid A = a] = E[h(X)] \; \forall a$$




```
classifier = DecisionTreeClassifier(min_samples_leaf=10, max_depth=4)
mitigator = ExponentiatedGradient(classifier, constraint)
```

Alekh Agarwal, Miroslav Dudík, and Zhiwei Steven Wu. Fair regression: quantitative definitions and reduction-based algorithms. In ICML, volume 97 of Proceedings of Machine Learning Research, 120–129. PMLR, 2019. URL: http://proceedings.mlr.press/v97/agarwal19d.html.

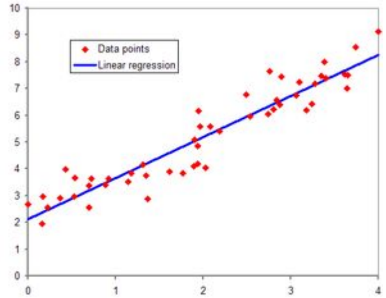
2. Parity constraints

Bounded group loss in Regression.

$$\min_{f \in \mathcal{F}} \mathbb{E}\left[\ell(Y, f(X))\right]$$

$$\text{such that } \forall a \in \mathcal{A}: \quad \mathbb{E}\left[\ell(Y, f(X)) \,\middle|\, A = a\right] \leq \zeta_a .$$

CDF *f* of predictions vector, A sensitive feature, *C* bound.



Alekh Agarwal, Miroslav Dudík, and Zhiwei Steven Wu. Fair regression: quantitative definitions and reduction-based algorithms. In ICML, volume 97 of Proceedings of Machine Learning Research, 120–129. PMLR, 2019. URL: http://proceedings.mlr.press/v97/agarwal19d.html.

3. Practical Application

2. Toward job recommendation

This paper presents a job recommendation algorithm designed and validated in the context of the French Public Employment Service.

It focuses on e-recruitment, i.e. the design and exploitation of recommender systems selecting job ads best suited for job seekers.

- Sensitivity of the data.
- E- recruitment involves rival goods.

Datasets:

- Recsys chanllenge 2017
- CareerBuilder public dataset.

Guillaume Bied, Solal Nathan, Elia Perennes, Morgane Hoffmann, Philippe Caillou, et al.. Toward Job Recommendation for All. IJCAI 2023 - The 32nd International Joint Conference on Artificial Intelligence, Aug 2023, Macau, China. pp.5906-5914, ⟨10.24963/ijcai.2023/655⟩. ⟨hal-04245528⟩

2. MUSE

This paper presents an e-R system called MUlti-head Sparse E-recruitment learned from proprietary data of French public employment service Pôle emploi.

MUSE. 0

It models the main 3 facets relevant to job recommendations such as geographical and skills.

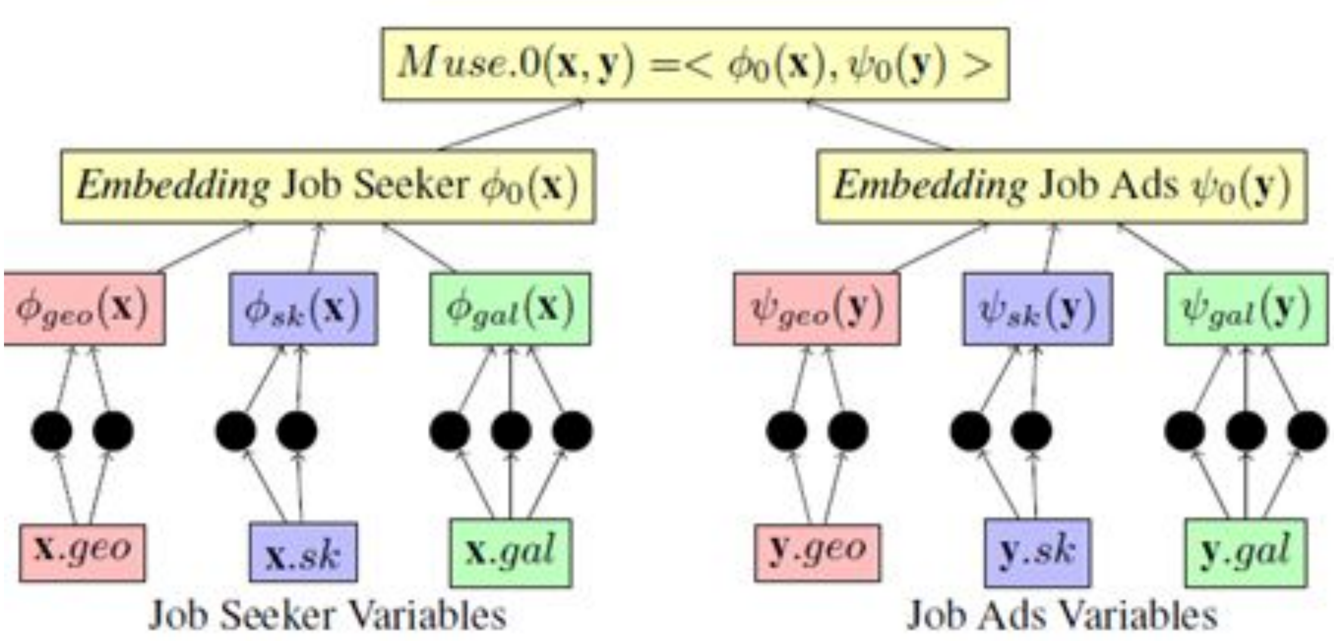
$$\boxed{s(\mathbf{x}, \mathbf{y}) = \langle \phi_0(\mathbf{x}), \psi_0(\mathbf{y}) \rangle}$$

Loss function

$$\mathcal{L}(\phi_0, \psi_0) = \sum_{(x,y,y')} [\langle \phi_0(\mathbf{x}), (\psi_0(\mathbf{y}) - \psi_0(\mathbf{y}')) \rangle + \eta]_+$$

Guillaume Bied, Solal Nathan, Elia Perennes, Morgane Hoffmann, Philippe Caillou, et al.. Toward Job Recommendation for All. IJCAI 2023 - The 32nd International Joint Conference on Artificial Intelligence, Aug 2023, Macau, China. pp.5906-5914, ⟨10.24963/ijcai.2023/655⟩. ⟨hal-04245528⟩

2. MUSE



Guillaume Bied, Solal Nathan, Elia Perennes, Morgane Hoffmann, Philippe Caillou, et al.. Toward Job Recommendation for All. IJCAI 2023 - The 32nd International Joint Conference on Artificial Intelligence, Aug 2023, Macau, China. pp.5906-5914, ⟨10.24963/ijcai.2023/655⟩. ⟨hal-04245528⟩

References

- Abigail Z. Jacobs and Hanna Wallach. Measurement and fairness. In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, FAccT '21, 375–385. New York, NY, USA, 2021. Association for Computing Machinery. URL: https://doi.org/10.1145/3442188.3445901, doi:10.1145/3442188.3445901.
- Alekh Agarwal, Miroslav Dudík, and Zhiwei Steven Wu. Fair regression: quantitative definitions and reduction-based algorithms. In ICML, volume 97 of Proceedings of Machine Learning Research, 120–129. PMLR, 2019. URL: http://proceedings.mlr.press/v97/agarwal19d.html.
- Guillaume Bied, Solal Nathan, Elia Perennes, Morgane Hoffmann, Philippe Caillou, et al.. Toward Job Recommendation for All. IJCAI 2023 - The 32nd International Joint Conference on Artificial Intelligence, Aug 2023, Macau, China. pp.5906-5914, ⟨10.24963/ijcai.2023/655⟩. ⟨hal-04245528⟩.


Thank
You