

Fairness in Multi-Stage Selection Processes

Vitalii Emelianov

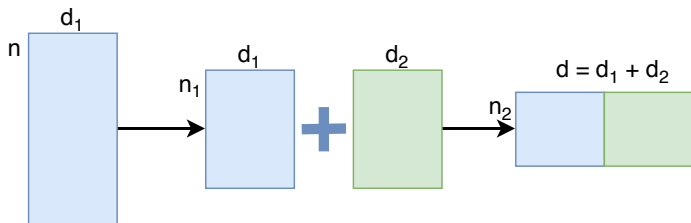
advisors: Patrick Loiseau, Nicolas Gast

May, 2019

2-Stage Selection

Each of n candidates is

- described by d features (x_1, \dots, x_d)
- endowed with a binary label $y \in \{0, 1\}$ which is not known

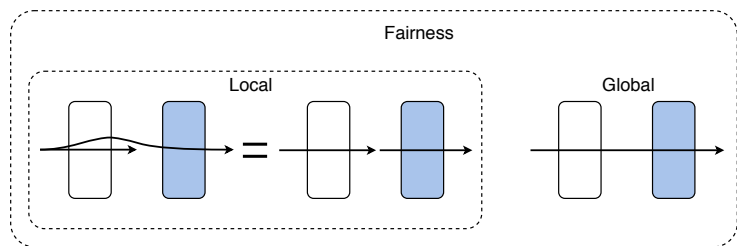


Assumptions:

- 1 knowledge of $P(y, x_1, \dots, x_d)$
- 2 number of candidates $n \rightarrow \infty$

Fair 2-Stage Selection

There is a sensitive attribute $x_s \in \{x_1 \dots x_d\}$ (e.g. race or gender)



Takeaways:

- later x_s is observed, higher the GF/LF performance gap
- if you want to be LF, reveal x_s soon (GF/LF difference is small)
- if you want to be GF, reveal x_s later (GF more fair at intermediate stages, has larger utility)