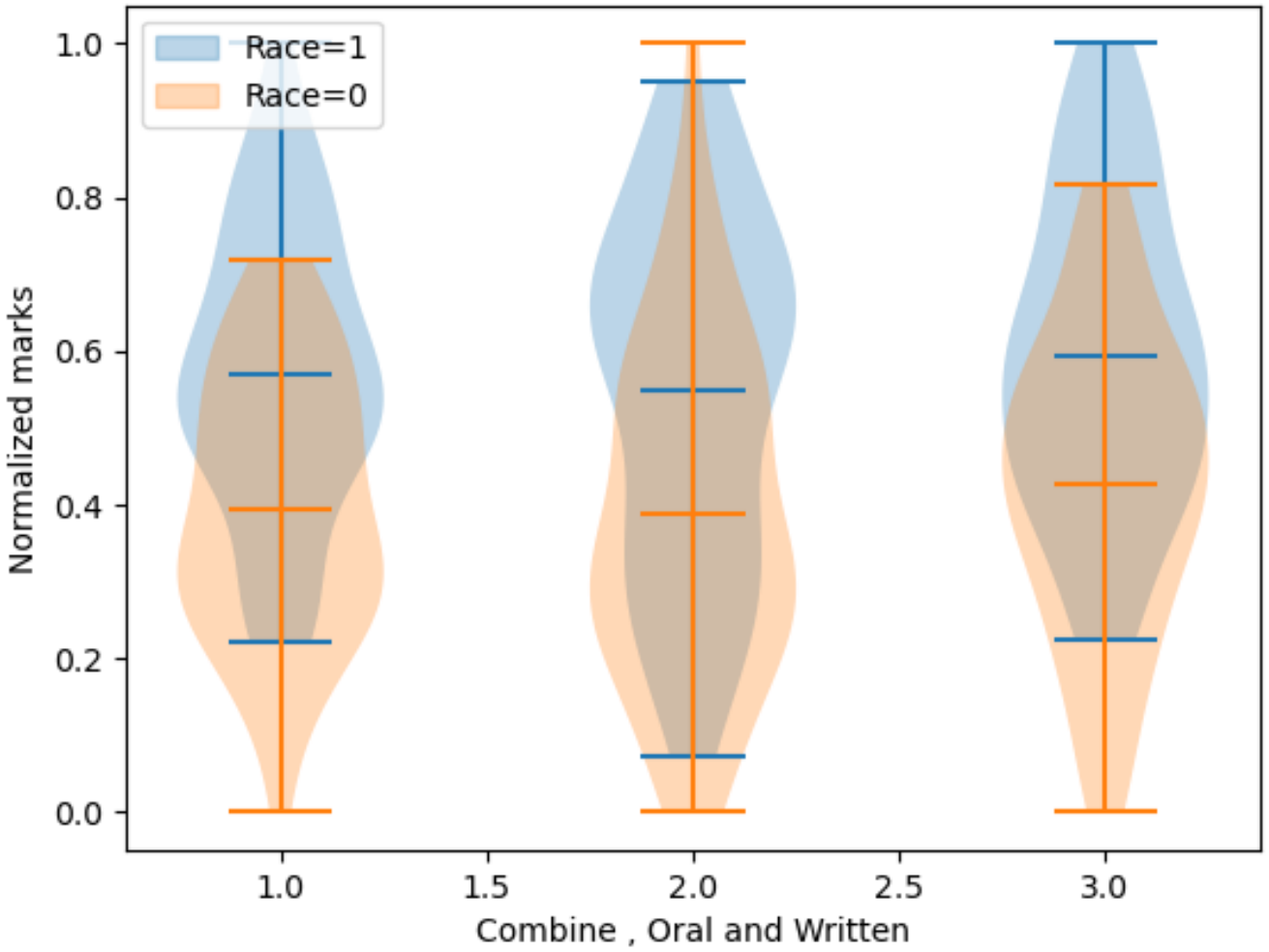
TRANSPARENCY AND FAIRNESS IN AI AND BIG DATA ALGORITHMS

Ricci Use Case

The Ricci Dataset

- Exam to firefighters to have a promotion (Captain or Lieutenant)
- Features :
  - *Oral, Written and Combine score (Combine score is 60% Written and 40% Oral)*
  - *Race (White, Hispanic or Black)*
  - *Position for the promotion (Captain or Lieutenant)*

The Ricci Dataset



Race is the sensitive attribute

68 white participant (Race=1) with 41 promotion

50 black and hispanic participant (Race=0) with 15 promotion

Fairness Metrics

- Mean difference between privileged and unpriviliged group show the fraction of privileged people which have more positive outocomes, We want him around 0
- For the original test set : 0.1
- Disparity impact show the rate of positive outcomes over the unprivileged group divide by the rate of positive outcomes over the privileged group. We want it around 1
- For the original test set : 0.78

Reweight + Classifier

- Reweighing transforms the dataset to have more equity between the privileged and unpivileged groups
- Pre processing technique : we need a classifier after reweighting the dataset
- With Logistic regression:
  - *Mean difference* = 0.025
  - *Disparate impact =1.06*
- With Random Forest:
  - *Mean difference =1e-16*
  - *Disparate impact =1*

Grid Search Reduction

- In-processing technique
- Returning the deterministic classifier with the lowest empirical error subject to fair classification constraints among the candidates searched
- Mean difference =0.08
- Disparate imapct =0.83

Results

	Mean difference	Disparate impact
Original	1.086957e-01	0.782609
Reweight + Logistic regression	2.542373e-02	1.058824
Reweight + Random forest	1.110223e-16	1
Grid Search Reduction	8.091787e-02	0.828645