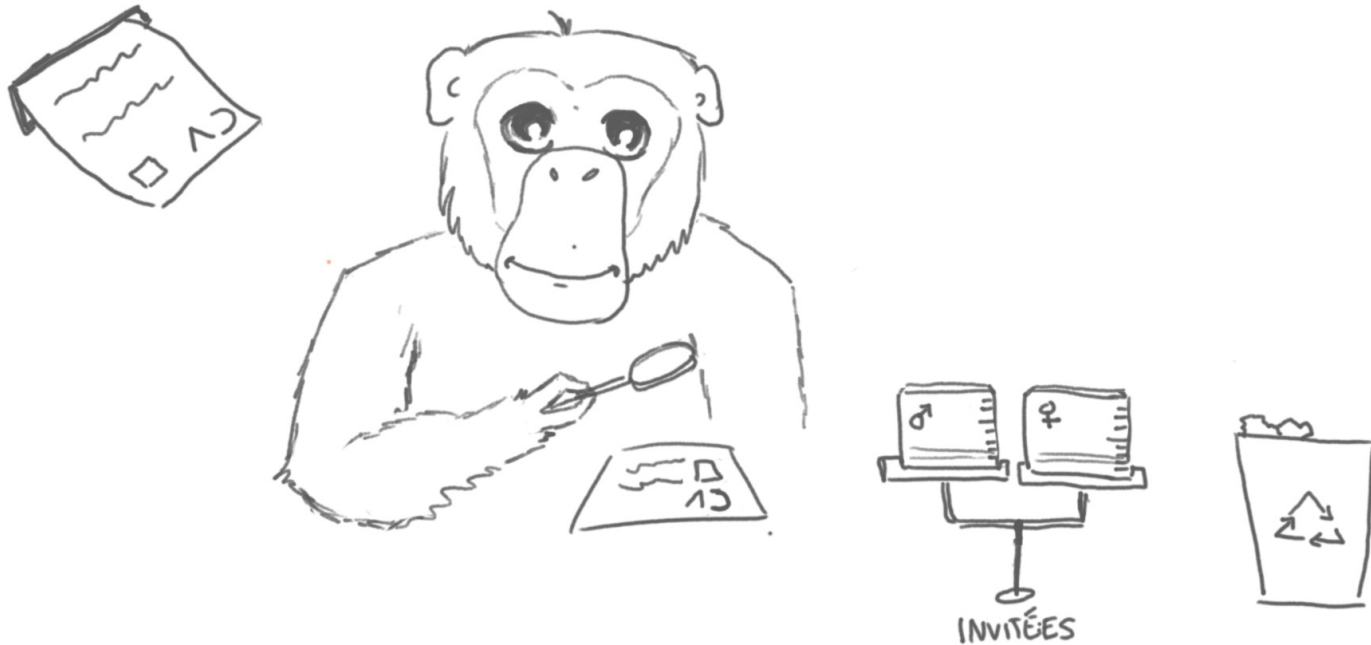


Découvrir les biais d'un algorithme

Journée d'étude « Quantification des discriminations en France »

20.11.2020 ▪ Linda Fernsel

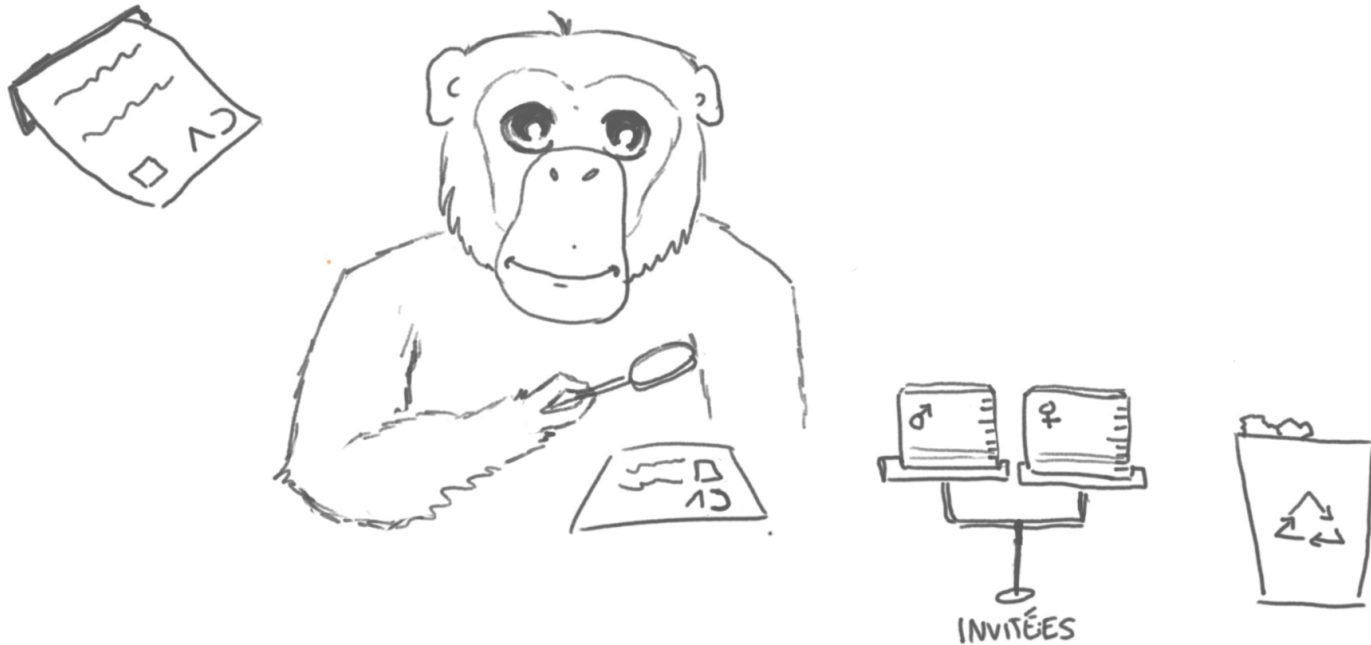
Procès de recrutement



Apprendre à classifier



Procès de recrutement

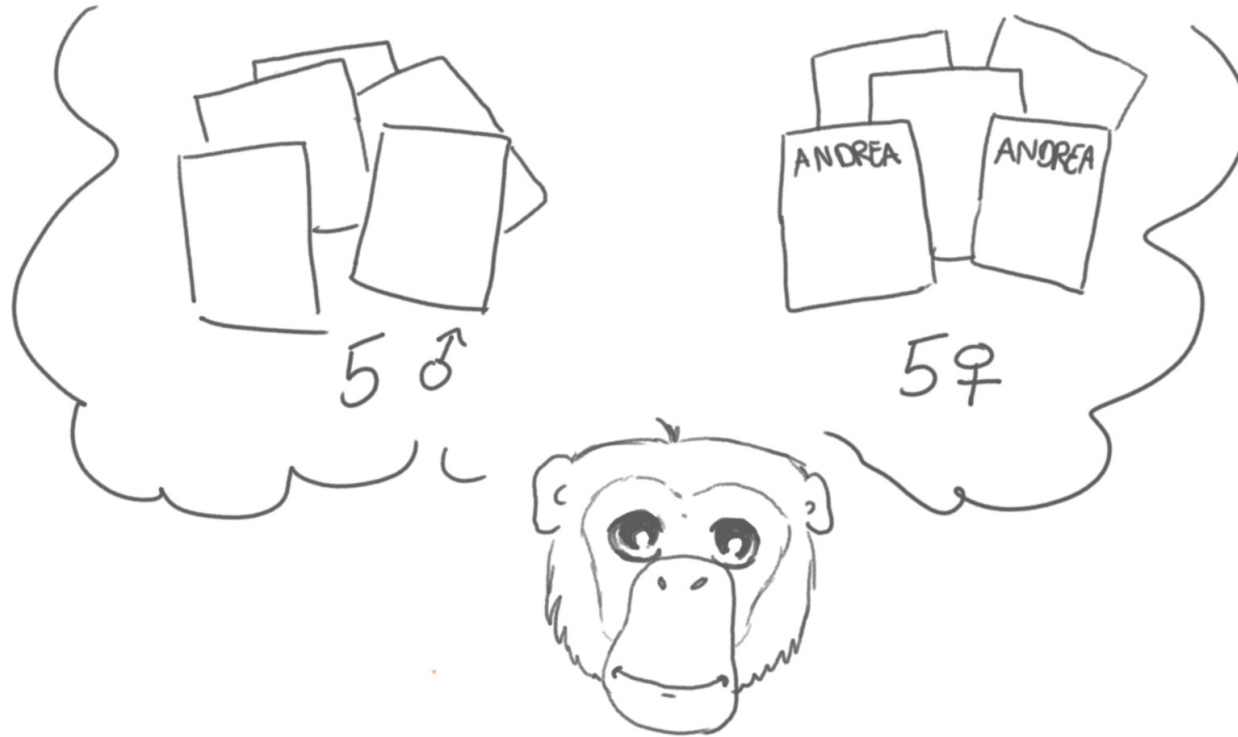


Pourquoi est-ce que le singe invite plus d'hommes que de femmes ?

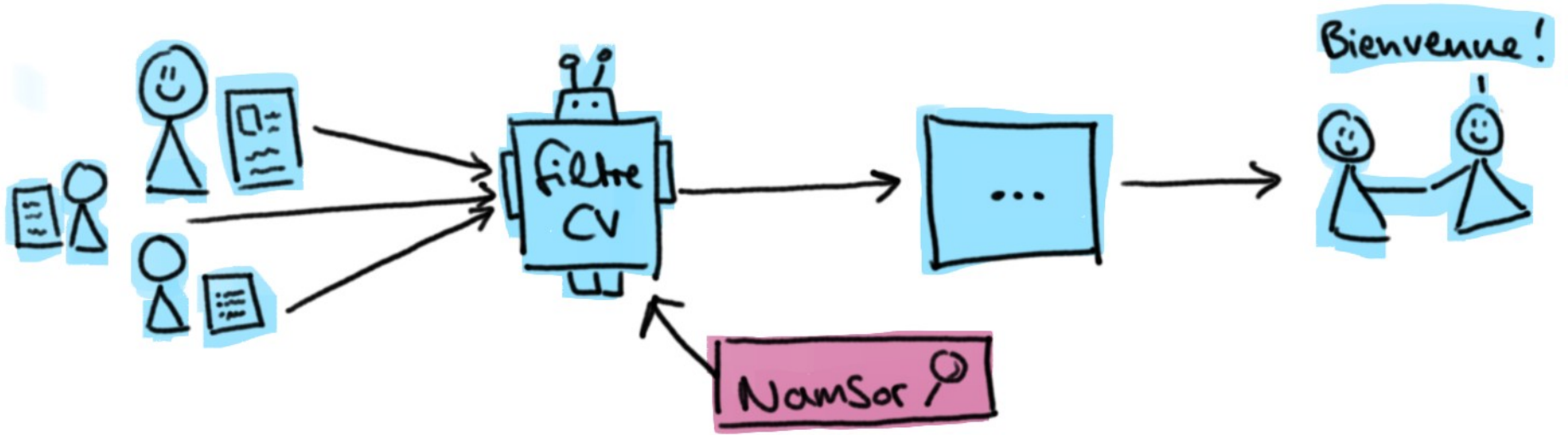
Application de l'appris



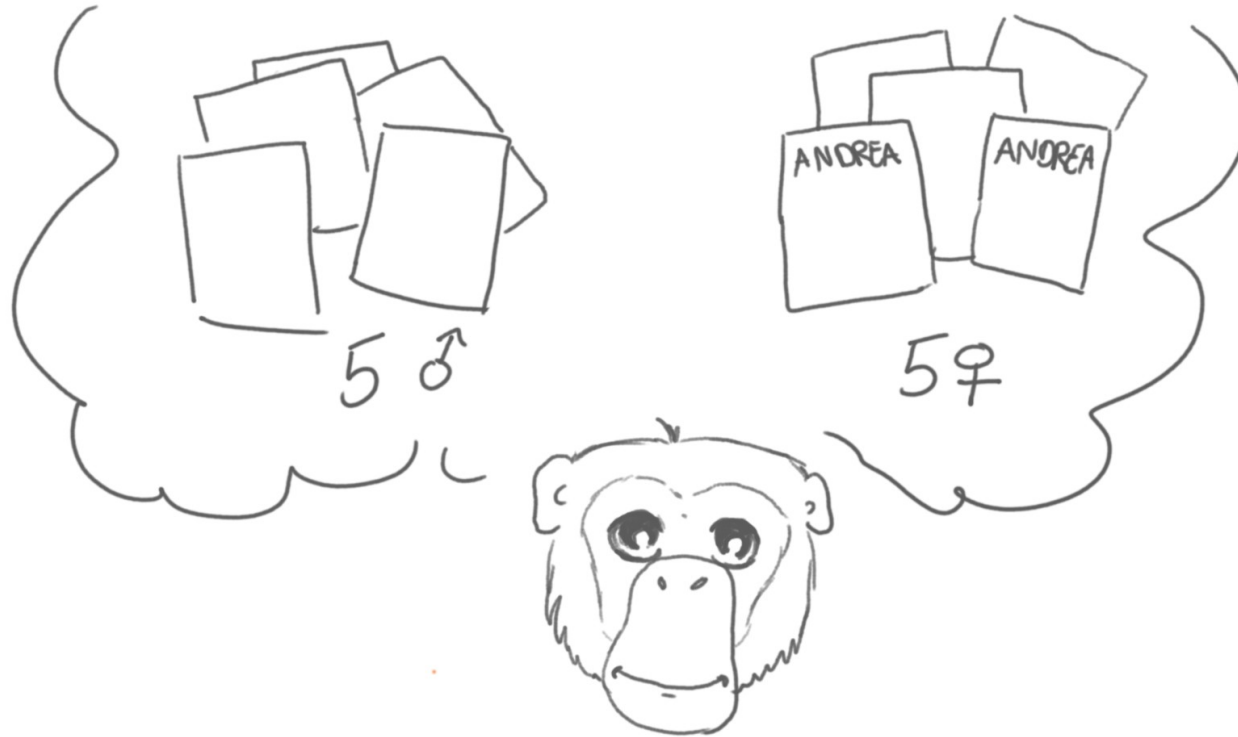
L'erreur



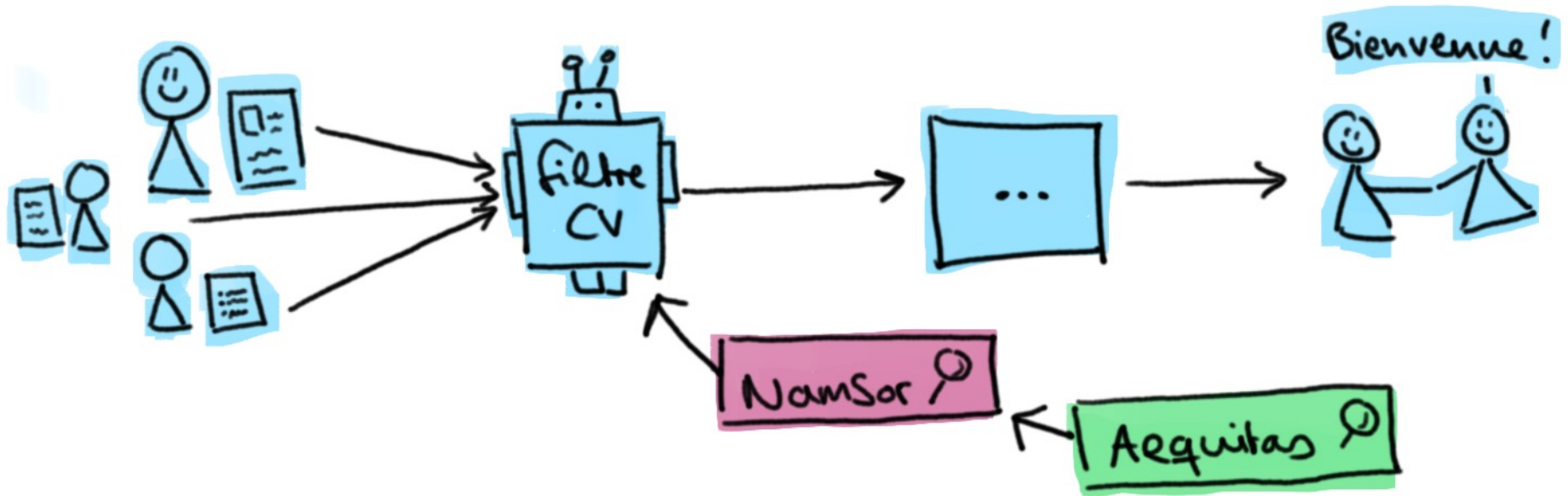
NamSor dans le processus de recrutement



L'erreur



Verifier la qualité de résultats



« Aequitas » nécessite :

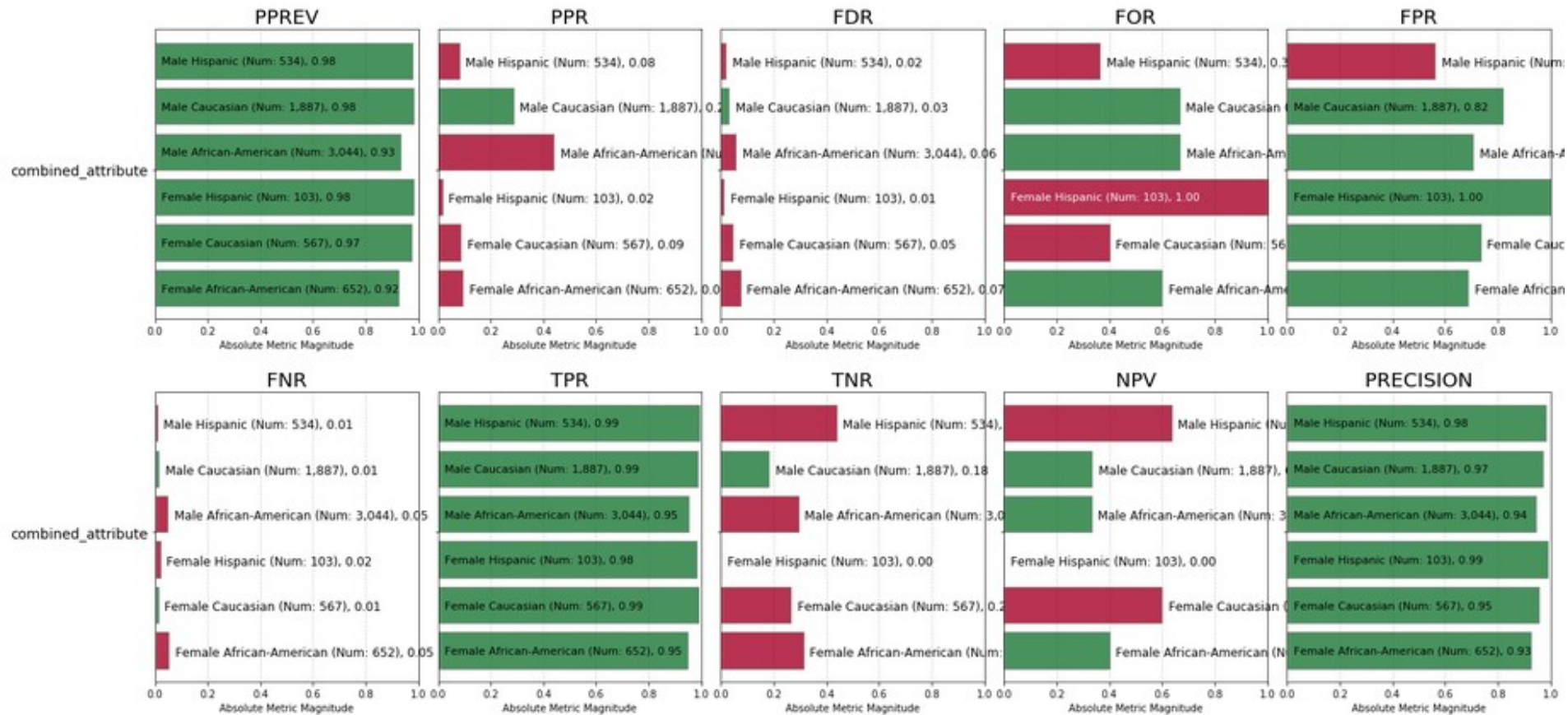
ID	Classes	Classe prédite	Prob. que résultat soit incorrect	
1	Other Male	Hispanic	0.024501	0.0
3	African-American Male	African-American	0.142035	1.0
4	African-American Male	Asian	0.388947	0.0

1 si le résultat est correct, 0 si non



Classe de référence : « Caucasian Male »

« Aequitas » calcule et visualise



Linda Fernsel



fernsele@htw-berlin.de
lindafernsele.de

Lire plus

- <http://www.datasciencepublicpolicy.org/projects/aequitas/>
- Saleiro et al. *Aequitas: A Bias and Fairness Audit Toolkit* (2019) <https://arxiv.org/pdf/1811.05577.pdf>
- Fernsel. *Algorithmic Fairness from Theory to Practical Application. A Fairness Audit of NamSor's Gender and Ethnicity Classification Algorithms.* (2020)
<https://github.com/namsor/namsor-compass/blob/master/algorithmic-fairness-fernse-linda-2020.pdf>