ECUE21.2 Science des données

Bonnes pratiques

Chloé-Agathe Azencott

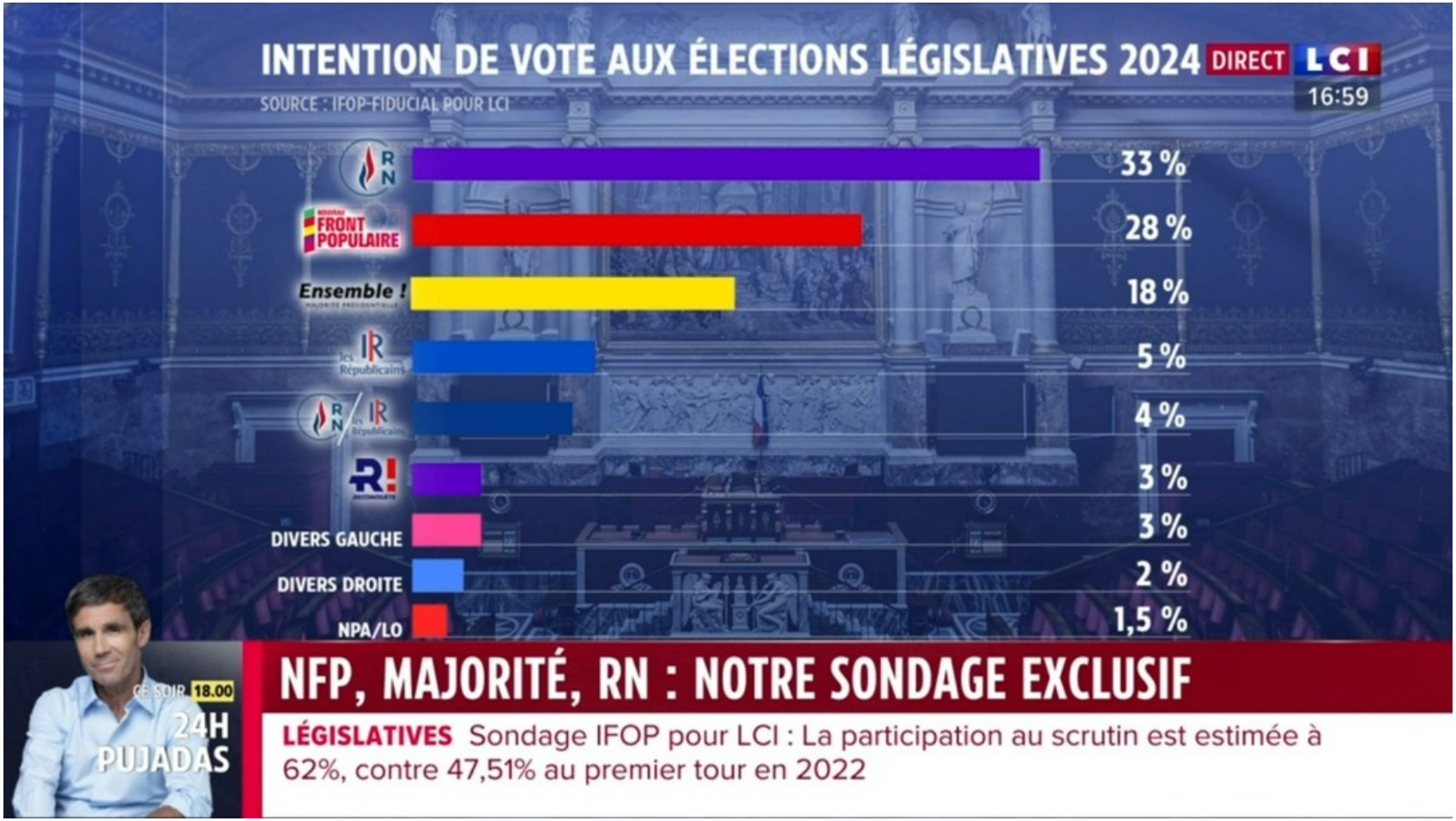
Center for Computational Biology (CBIO)
Mines Paris PSL – Institut Curie – INSERM U900
PSL Research University & PR[AI]RIE, Paris, France

Juin 2024

http://cazencott.info chloe-agathe.azencott@minesparis.psl.eu @cazencott@lipn.info

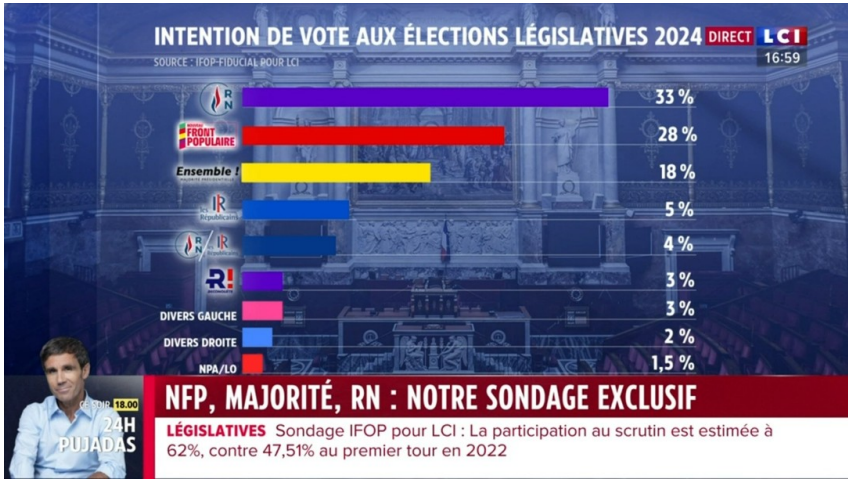
1. Visualisation de données

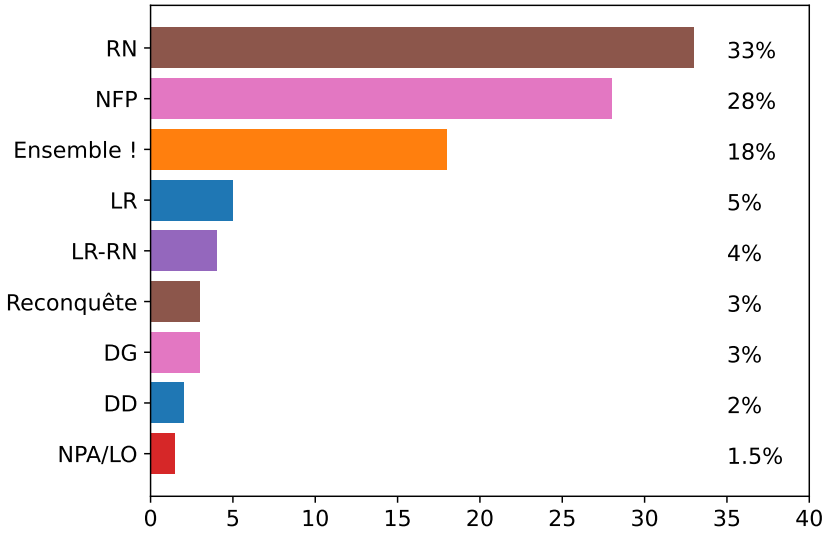
1. Exemple 1



TF1, 17 juin 2024, 16h56

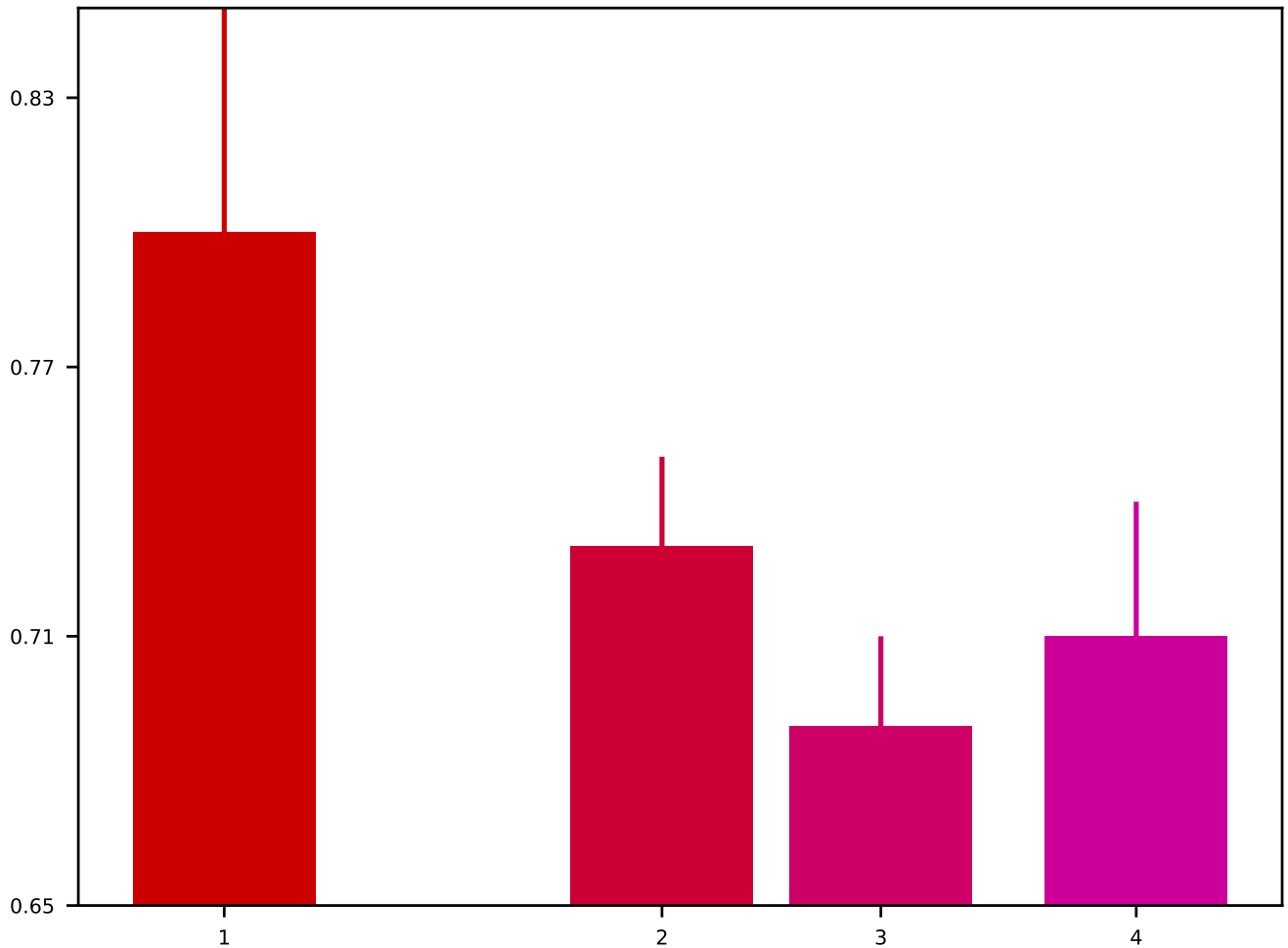
1. Exemple 1


INTENTION DE VOTE AUX ÉLECTIONS LÉGISLATIVES 2024
DIRECT
LCI
16:59
SOURCE : IFOP-FIDUCIAL POUR LCI
33 %
28 %
18 %
5 %
4 %
3 %
3 %
2 %
1,5 %
DIVERS GAUCHE
DIVERS DROITE
NPA/LO
NFP, MAJORITÉ, RN : NOTRE SONDAGE EXCLUSIF
LÉGISLATIVES Sondage IFOP pour LCI : La participation au scrutin est estimée à 62%, contre 47,51% au premier tour en 2022
24H PUJADAS


RN
NFP
Ensemble !
LR
LR-RN
Reconquête
DG
DD
NPA/LO
33%
28%
18%
5%
4%
3%
3%
2%
1.5%
0
5
10
15
20
25
30
35
40

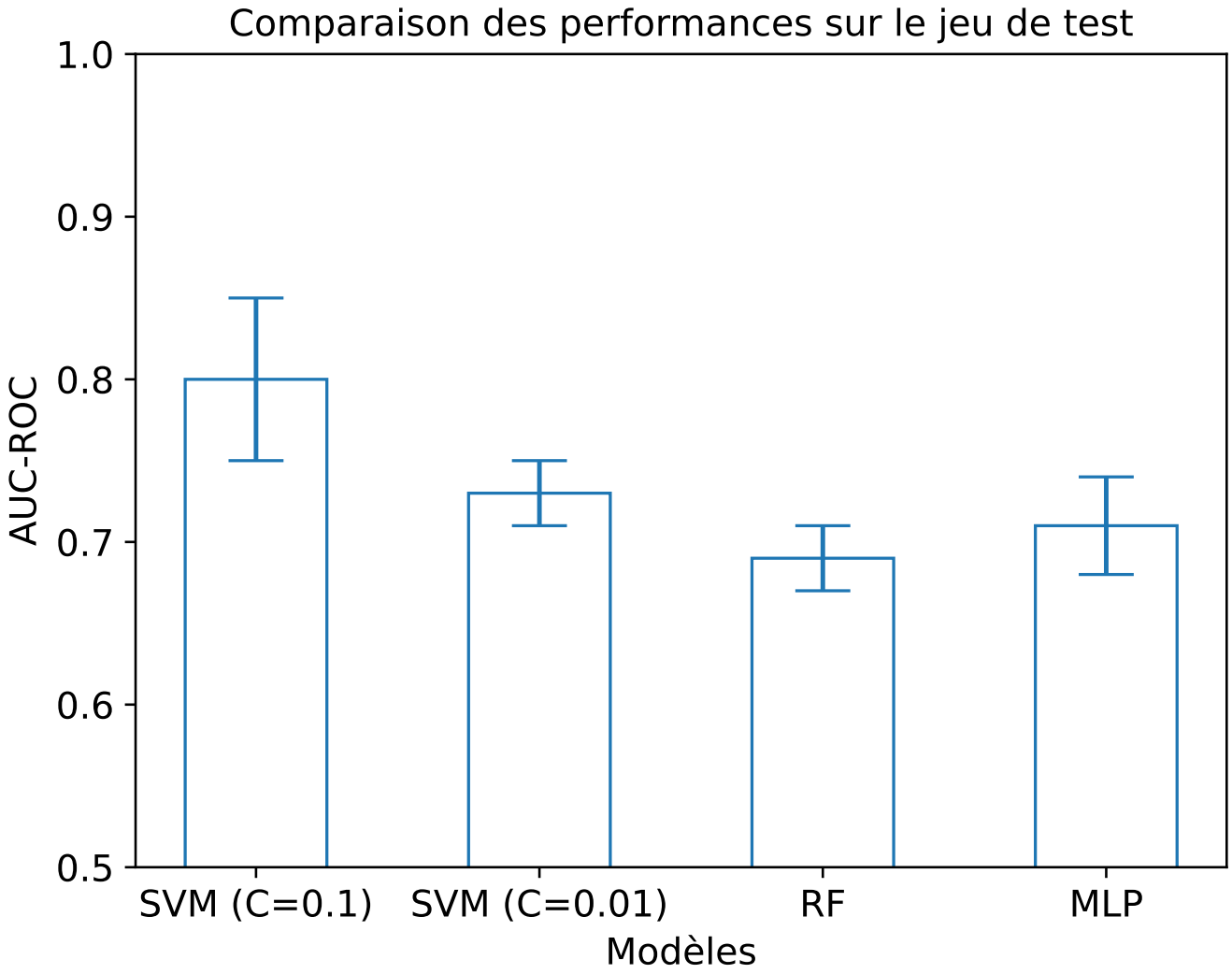
2. Exemple 2

Performance de 4 modèles sur un problème d'apprentissage

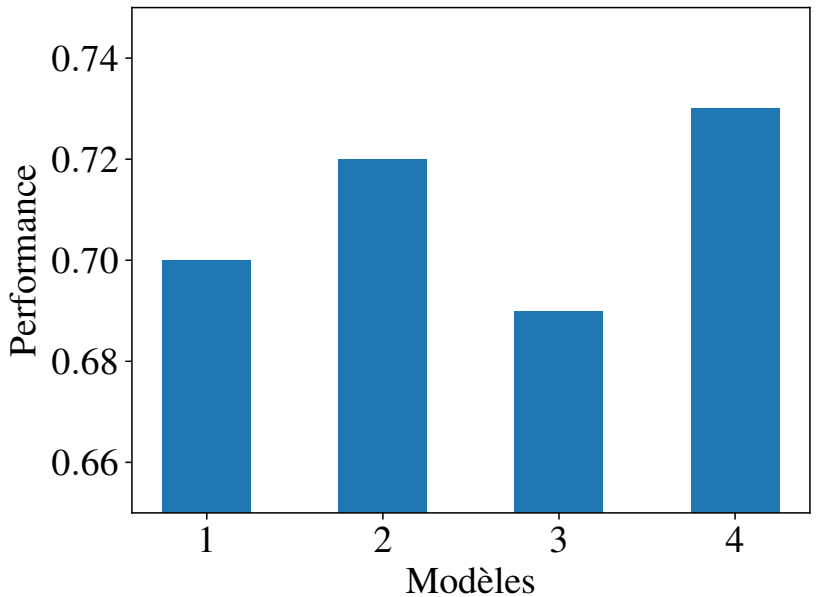


2. Exemple 2

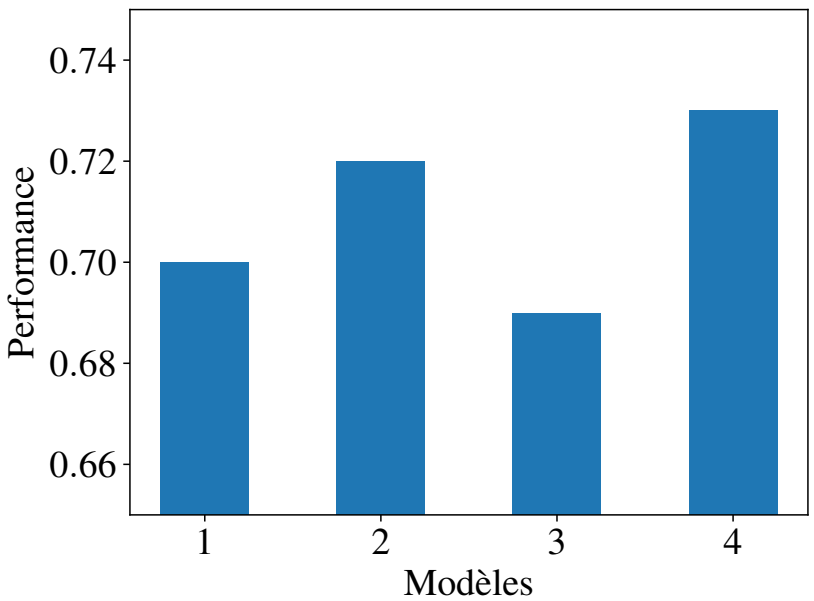
Performance de 4 modèles sur un problème d'apprentissage

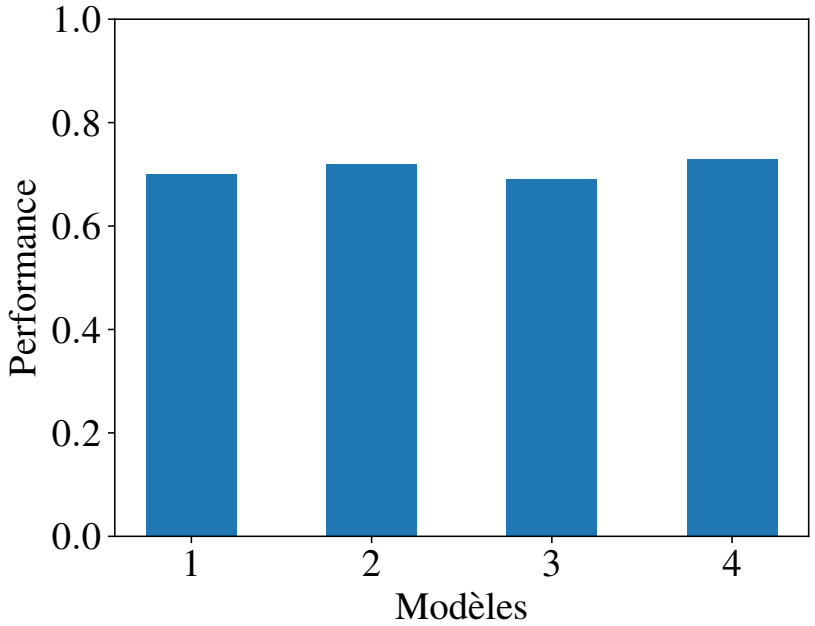


3. Choix des axes (1)

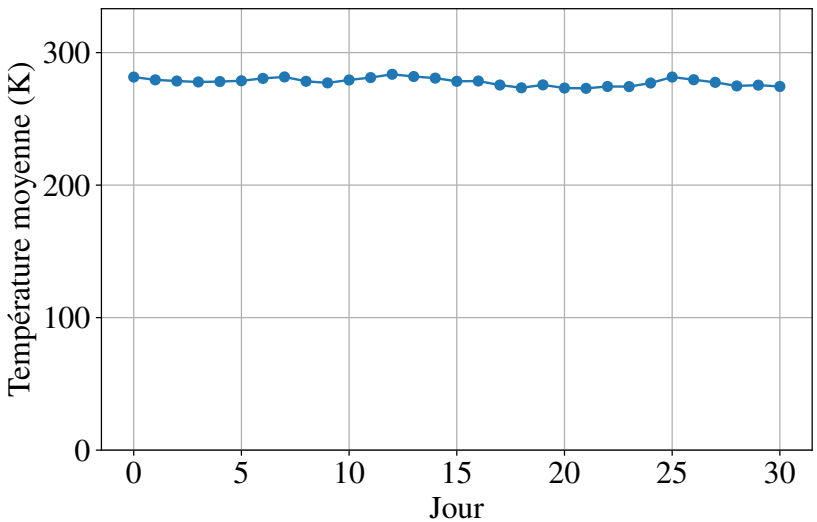


3. Choix des axes (1)

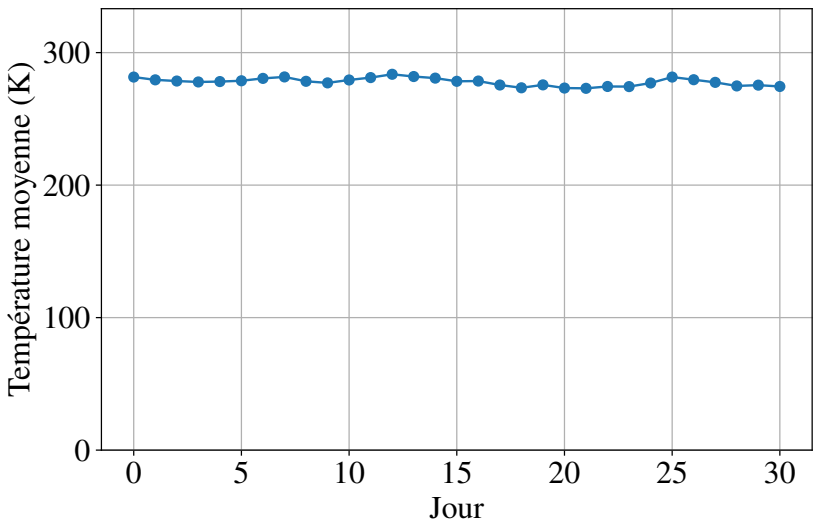


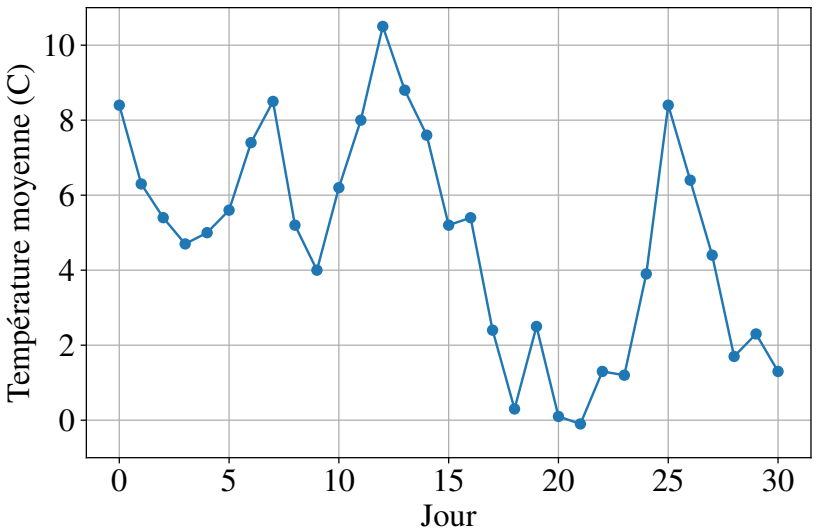


3. Choix des axes (2)

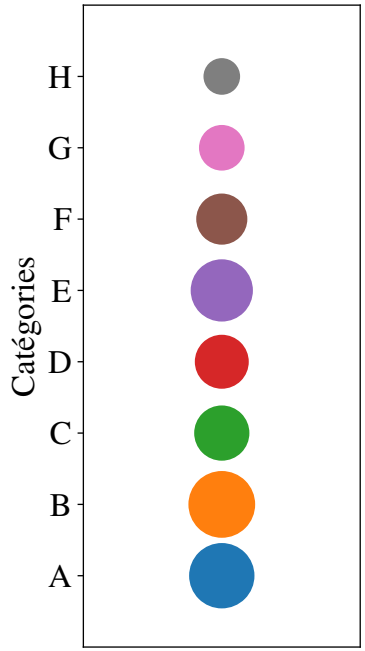


3. Choix des axes (2)

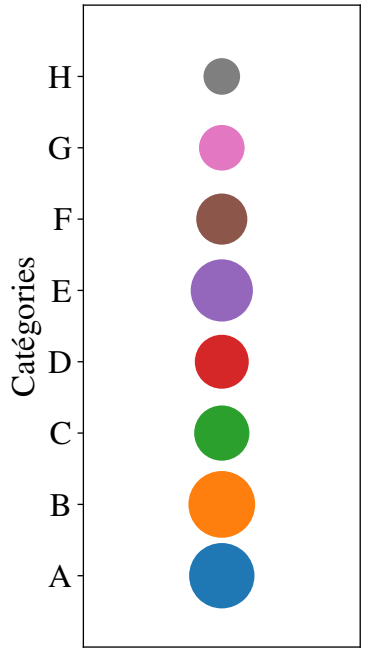


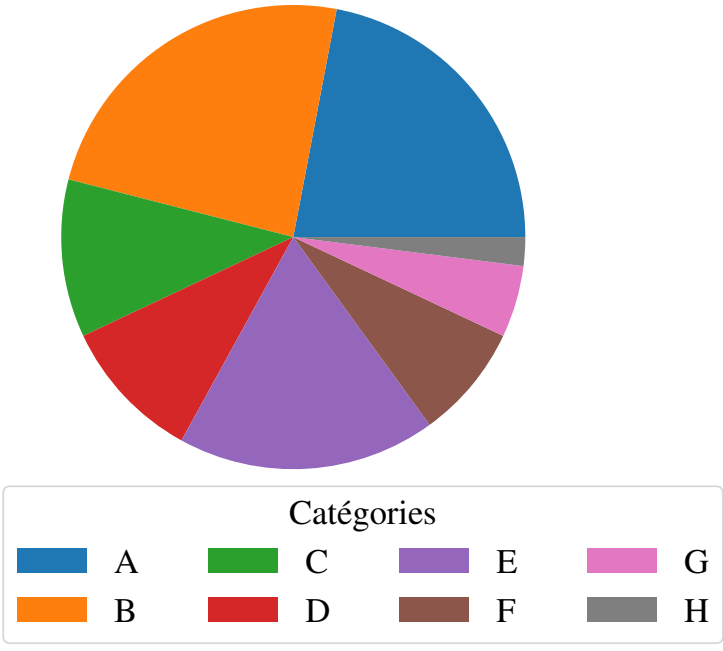


4. Proportional ink

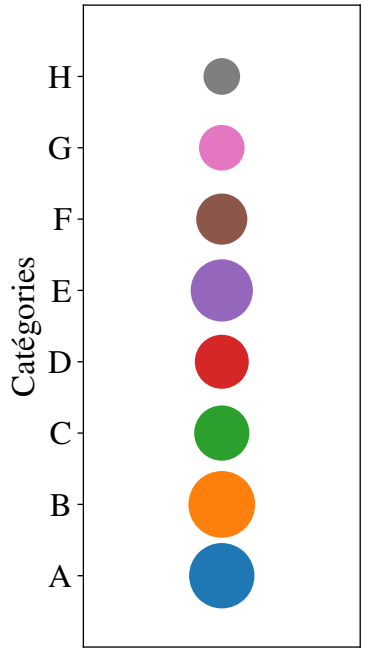


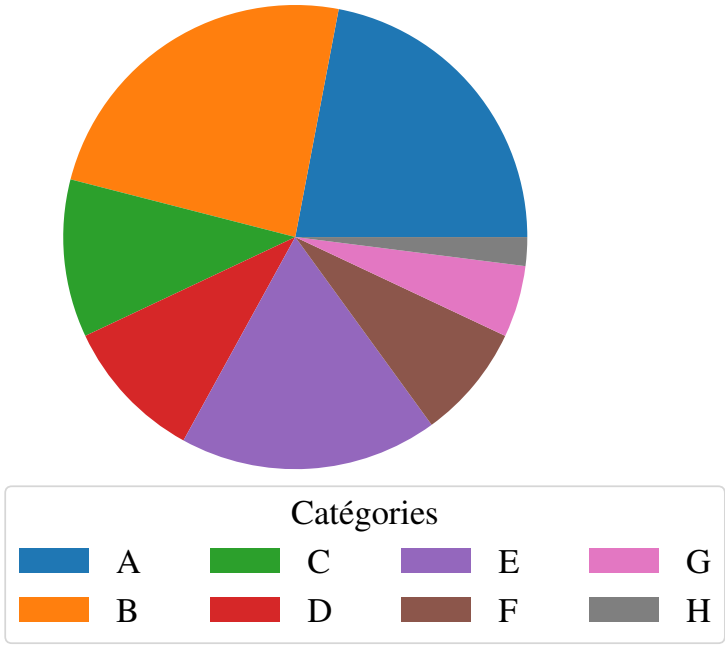
4. Proportional ink

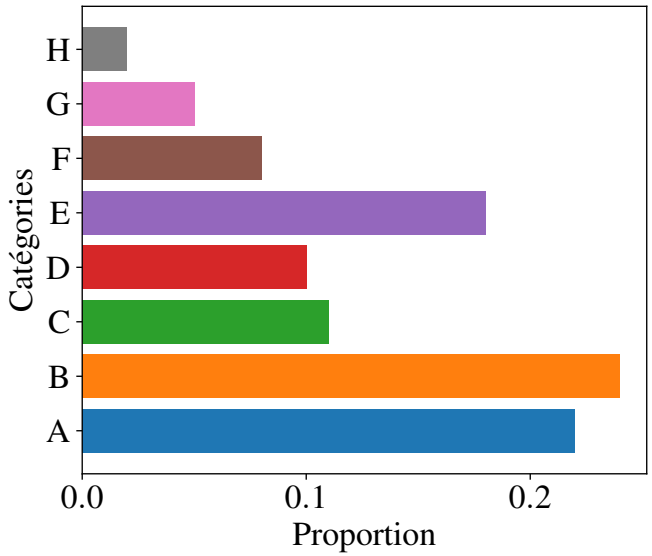




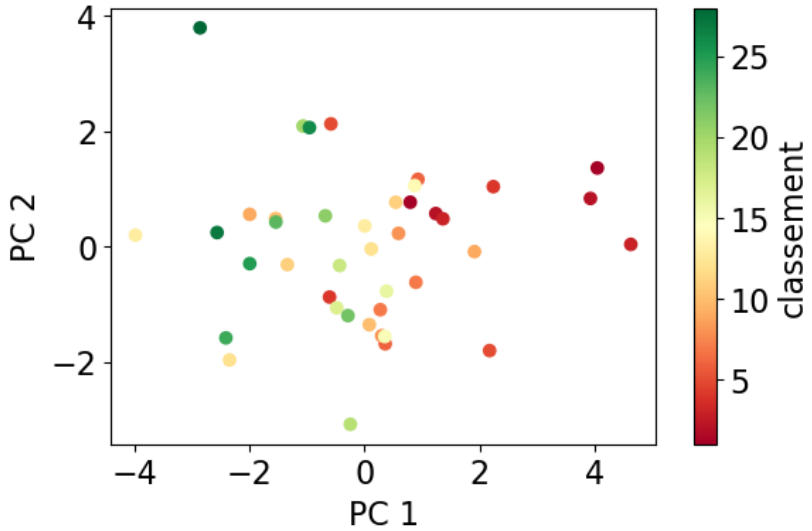
4. Proportional ink





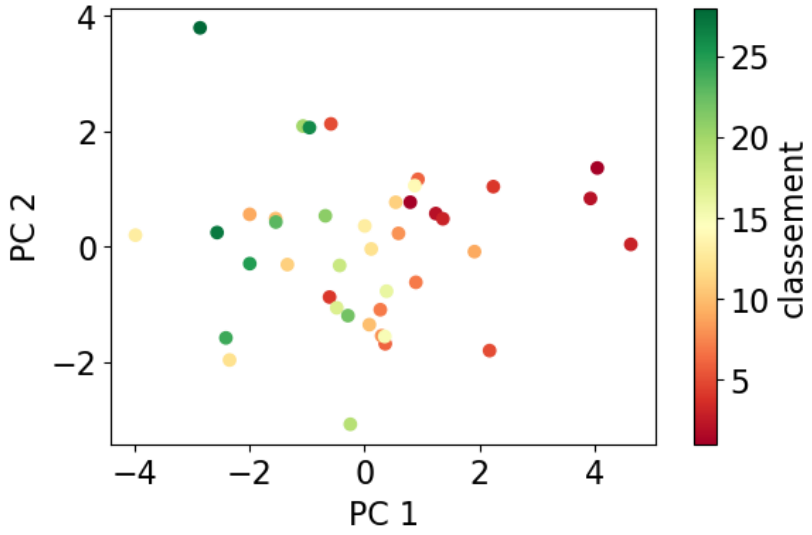


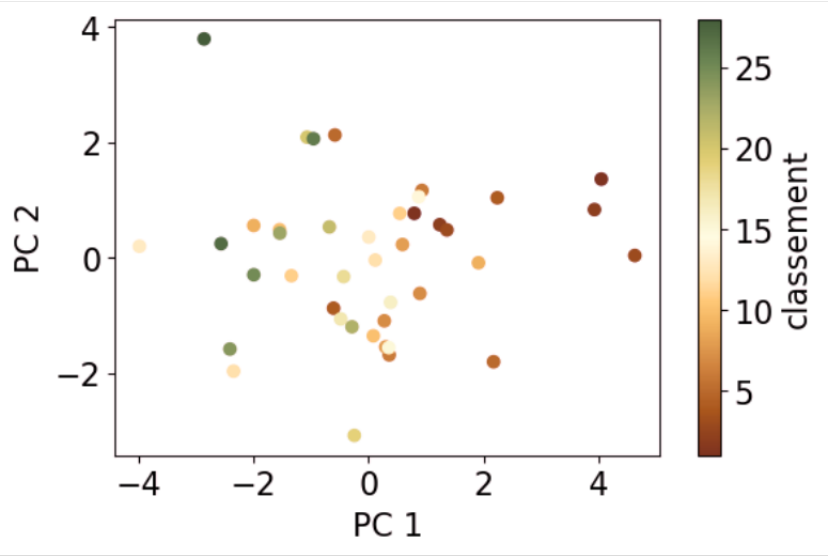
5. Dyschromatopie




```
plt.scatter(...cmap='RdYlGn')
```

5. Dyschromatopie



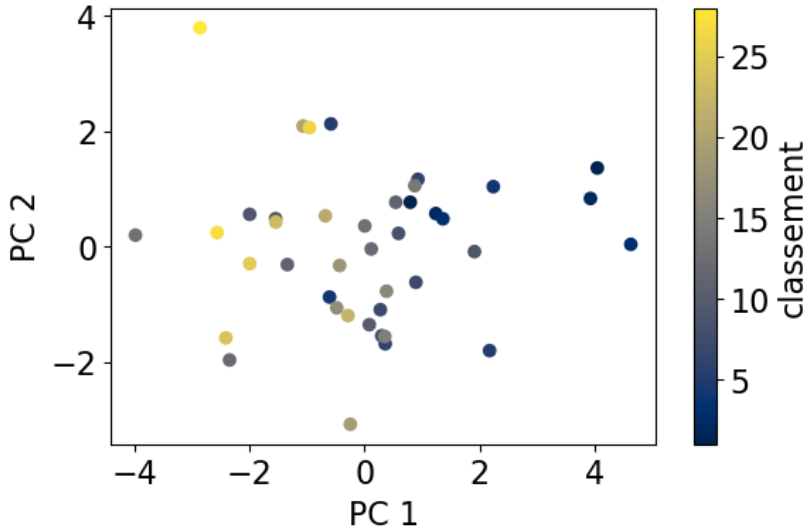



```
plt.scatter(...cmap='RdYlGn')
```

Simulation de deutéranopie par CoBliS

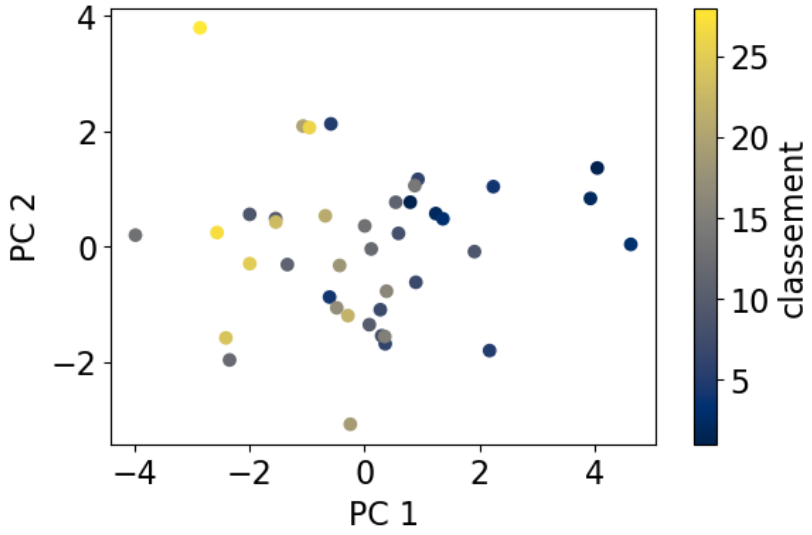
[lien vers CoBliS]

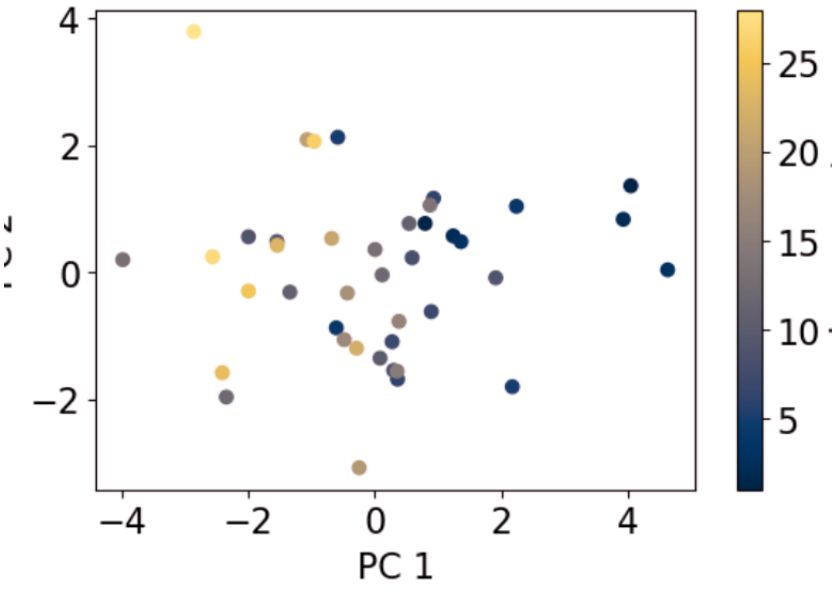
5. Dyschromatopie




```
plt.scatter(...cmap='cividis)
```

5. Dyschromatopie





`plt.scatter(...cmap='cividis)`

Simulation de deutéranopie par CoBliS

[lien vers CoBliS]

2. Équité des algorithmes

1. Recrutement automatisé (Amazon)

source: Reuters [lien]

- Système fortement biaisé en faveur des CV déposés par des hommes
- Pourtant le sexe ne faisait pas partie des variables utilisées

2. Détection de criminels par leur photo

- Article sur arxiv : *Automated Inference on Criminality using Face Images*, Xiaolin Wu & Xi Zhang (2017)
- Motivation : "Unlike a human examiner/judge, a computer vision classifier has absolutely no subjective baggage, having no emotions, no biases whatsoever due to past experience, race, religion, political doctrine, gender, age, etc."

2. Détection de criminels par leur photo

– Article sur arxiv : *Automated Inference on Criminality using Face Images*, Xiaolin Wu & Xi Zhang (2017)

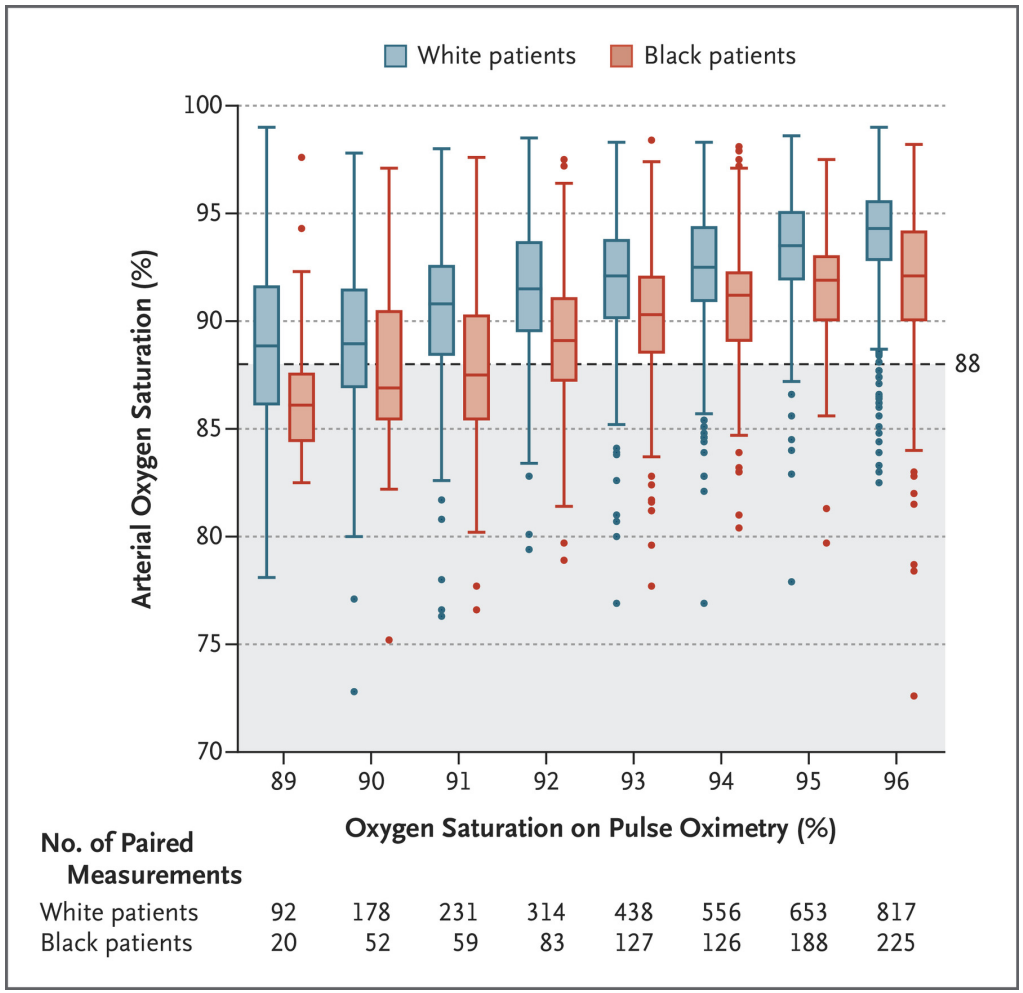
(a) Three samples in criminal ID photo set S_c.

(b) Three samples in non-criminal ID photo set S_n

Figure 1. Sample ID photos in our data set.

3. It's not just AI : oxymètres de pouls

– *Racial Bias in Pulse Oximetry Measurement*, Sjoding et al., New England Journal of Medicine, 2020; 383:2477-2478 [lien]



Accuracy of Pulse Oximetry in Measuring Arterial Oxygen Saturation, According to Race.

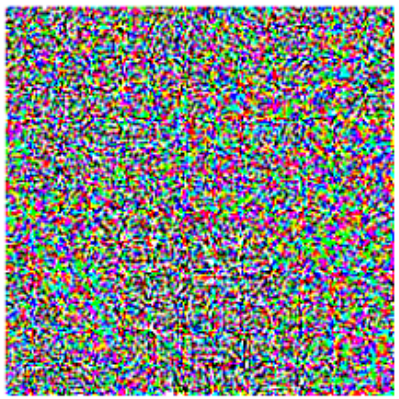
3. Fiabilité

Attaque (bruit gaussien)

$\boldsymbol{x}$

"panda"
57.7% confidence

$+ .007 \times$



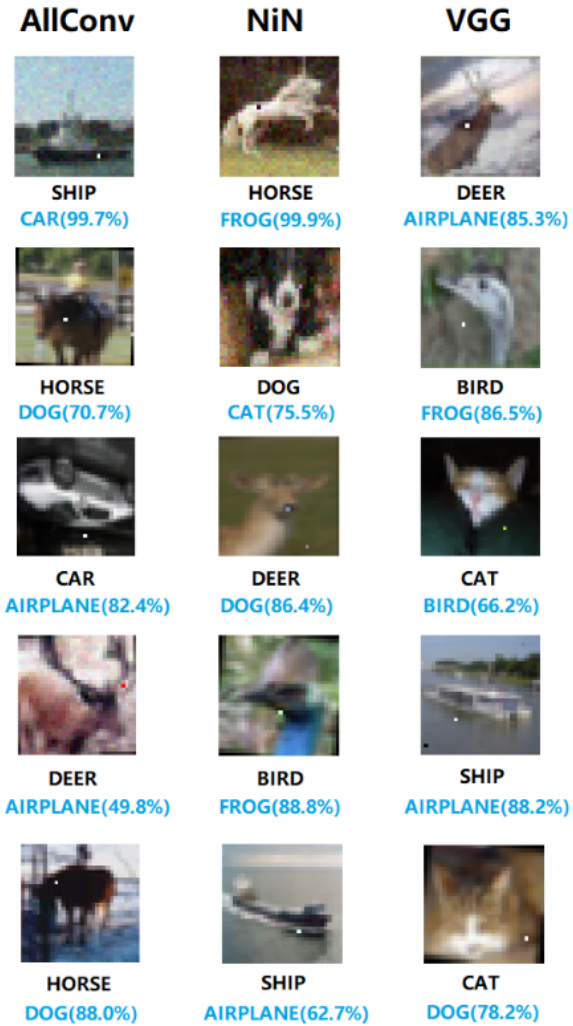
$\text{sign}(\nabla_{\boldsymbol{x}} J(\boldsymbol{\theta}, \boldsymbol{x}, y))$

$=$

$\boldsymbol{x} + \epsilon\text{sign}(\nabla_{\boldsymbol{x}} J(\boldsymbol{\theta}, \boldsymbol{x}, y))$
"gibbon"
99.3 % confidence

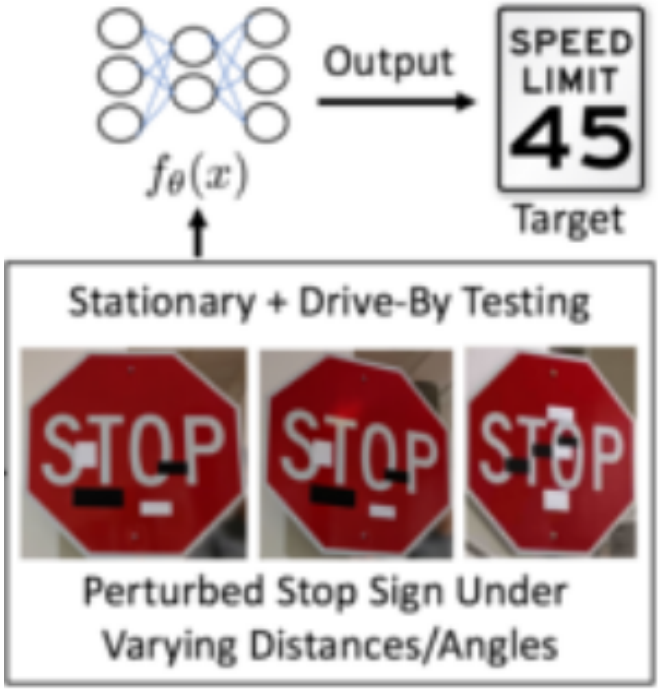
Goodfellow, Shlens & Szegedy (ICLR 2015)

Attaque (1-pixel)



Su, Vargas & Kouichi (IEEE Transactions on Evolutionary Computation 2019)

Attaque (monde réel)



Eykholt et al. (CBPR 2018)

4. Confidentialité

5. Enjeux écologiques