PI_ExpAI – Twitter bots detection project

César Michaud, Robin Müller, Florian Polier

Goals

- Apply bots detection techniques to analyse the public debate during the 2022 French presidential election
- Estimate the energetic impacts of bots
- Determine which political candidates use bots
- Build a suitable ML model and gather a custom dataset

State of the art

Annotated datasets

- TwiBot-20: A Comprehensive Twitter Bot Detection Benchmark
- Midterm-2018: Scalable and generalizable social bot detection through data selection
- Cresci-17: The Paradigm-Shift of Social Spambots: Evidence, Theories, and Tools for the Arms Race

Models - two main approaches

- Scalable and generalizable social bot detection through data selection
- Detection of Novel Social Bots by Ensembles of Specialized Classifiers / Botometer

Datasets

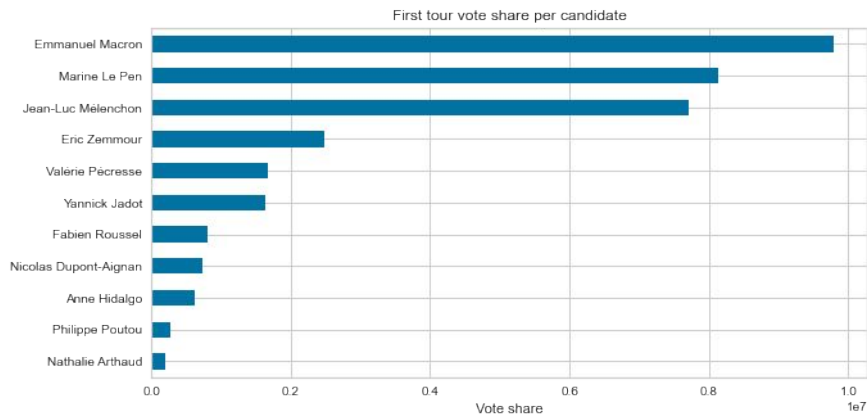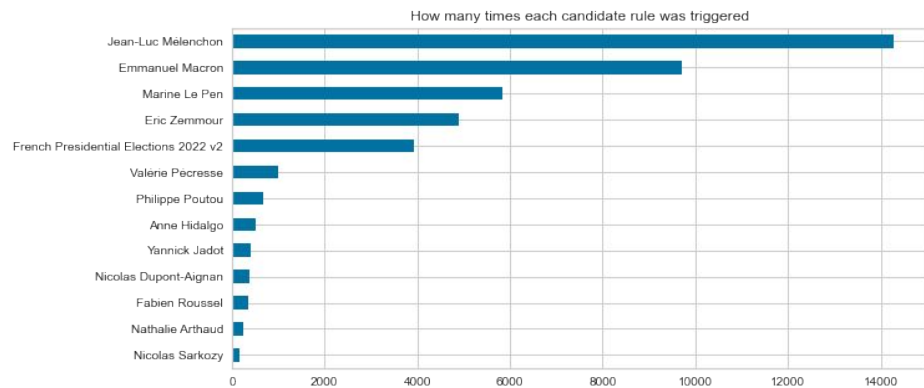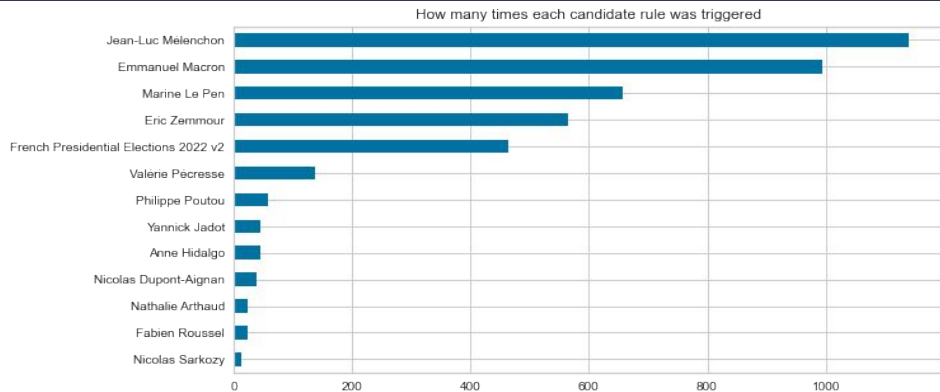|	cresci-17	midterm-2018	TwiBot-20
# of users	3474	8092	5237
# of bots	7543	42446	6589
Property	✓	✓	✓
Semantic	✗	✓	✓
Neighbor	✗	✗	✓
Annotation methodology	Expert	Automated	Crowdsourcing

Models

- Model trained on a single dataset
  - 0.82 - 0.99 accuracy on the same dataset
  - Poor performance on the other datasets
- Model trained on multiple datasets
  - Better performance (accuracy > 0.8)
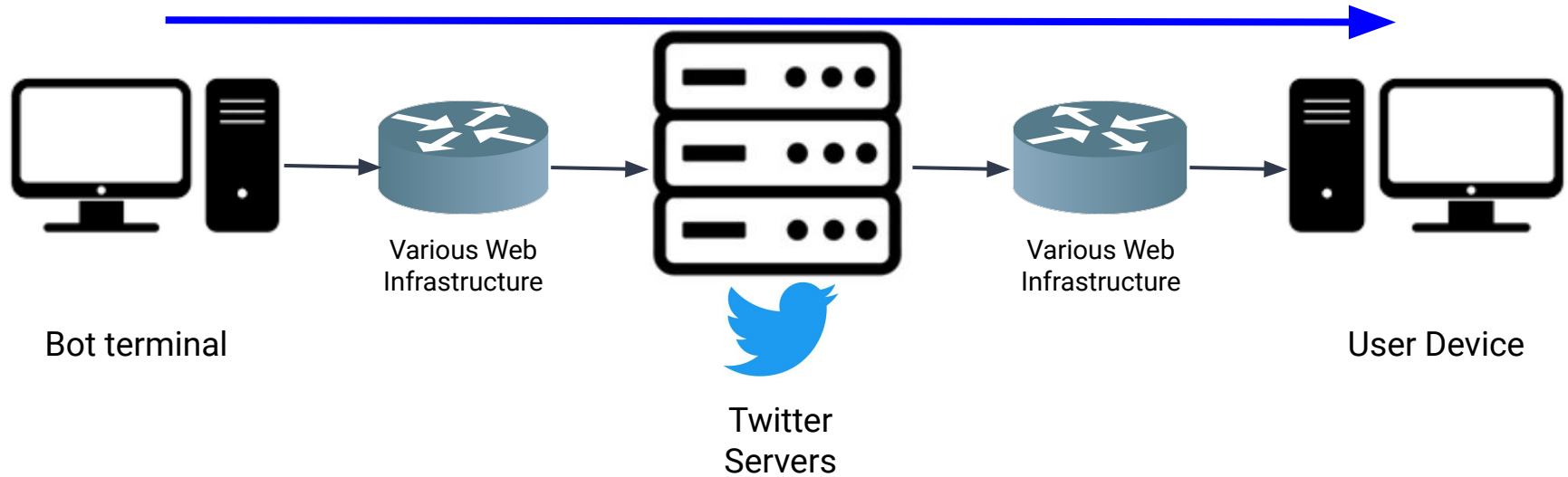- Unresolved issue with the model
  - False positives

Analysis on our French dataset

Principal Component Plot
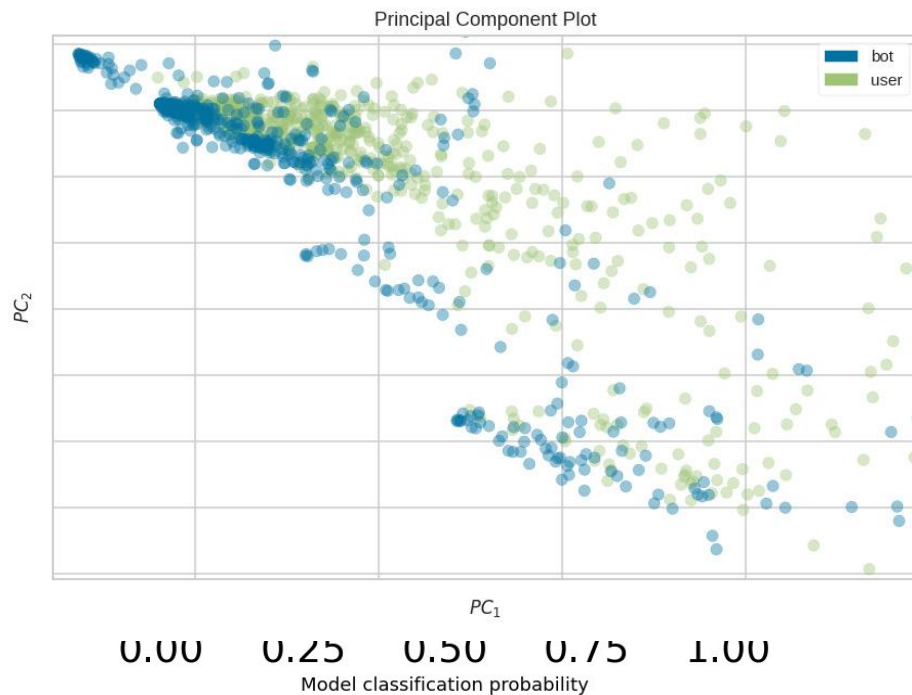
Analysis on our French dataset – 2

Energy cost

Energy cost

Bot terminal

Various Web Infrastructure

Various Web Infrastructure

Twitter Servers
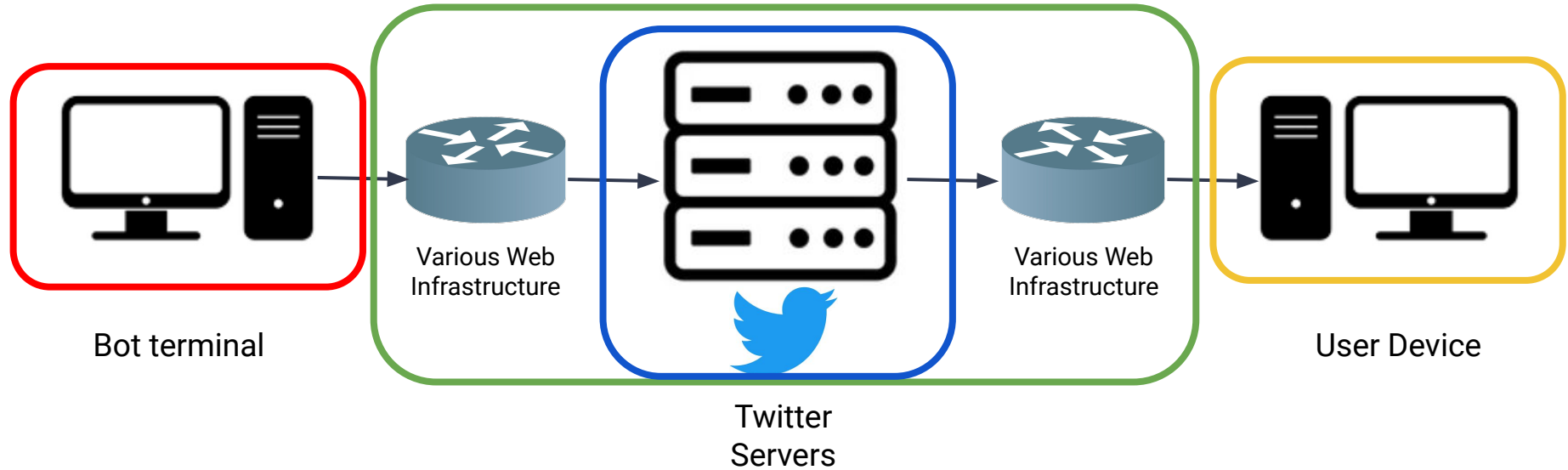
User Device

- Cost of bot terminals   : 1.75 TWh
- Cost of web infrastructure   : 4.05 - 6.75 TWh
- Cost of user devices   : 58.9 GWh
- Cost of tweets posted   : 1.4 - 2.2 GWh

Different bots obtained

Retweet bots :

- Rarely post (if ever)
- Sometimes focus on a single subject, sometimes on
- Would like to present visual example but hard to represent in a single image

Examples :

- DuffaudR
- kagan_sumner
- Guitou_8888
- colinjarrett587
- BronyRetweets
- AGrothendick

Different bots obtained

Clout bots :

- Exclusively share, retweet or post about one company (usually a news website)
- Open about the use of bot

Examples :

- StarvisionNews
- P_Mundo
- RedditLFI

Starvision News @StarvisionNews · 21s

Bahrain and Egypt sign 10 agreements and memoranda of understanding

starvisionnews.com
Bahrain and Egypt sign 10 agreements and memoranda of understandi...
Bahrain and Egypt have signed 10 agreements, memoranda of understanding and joint executive programs in the fields of economy, ...

Different bots obtained

Spam bots :

- Actually posts content
- Often repetitive
- Goal unclear as link between posts is sometimes difficult to establish

Examples :

- CianGaia (changed content since analysis)
- whataweekhuh

Different bots obtained

Spam bots :

- Actually posts content
- Often repetitive
- Goal unclear as link between posts is sometimes difficult to establish

Examples :

- CianGaia (changed content since analysis)
- whataweekhuh

Conclusion

Take away messages

- Objectives achieved

- Energy consumed by bots is negligible on a global scale

- Different types of bots, but most only spread content and do not create it

- Next time, a more suitable problematic

Thank you for your attention !

Questions ?