USER BEHAVIOR MODALISATION IN MOBILE ADVERTISING

Faustine BOUSQUET, El hassan DAOUDI, Samy ZAROUR

Joseph AKA BROU
Marwa ELATRACHE
Caroline MARTIN
Tharshika NAGARATNAM
Omar SECK

Outline

1. INTRODUCTION
2. DATA PRESENTATION
3. DATA EXPLORATION
4. DIFFICLUTIES
5. PERSPECTIVES & CONCLUSION

TABMO
CREATIVE MOBILE DSP

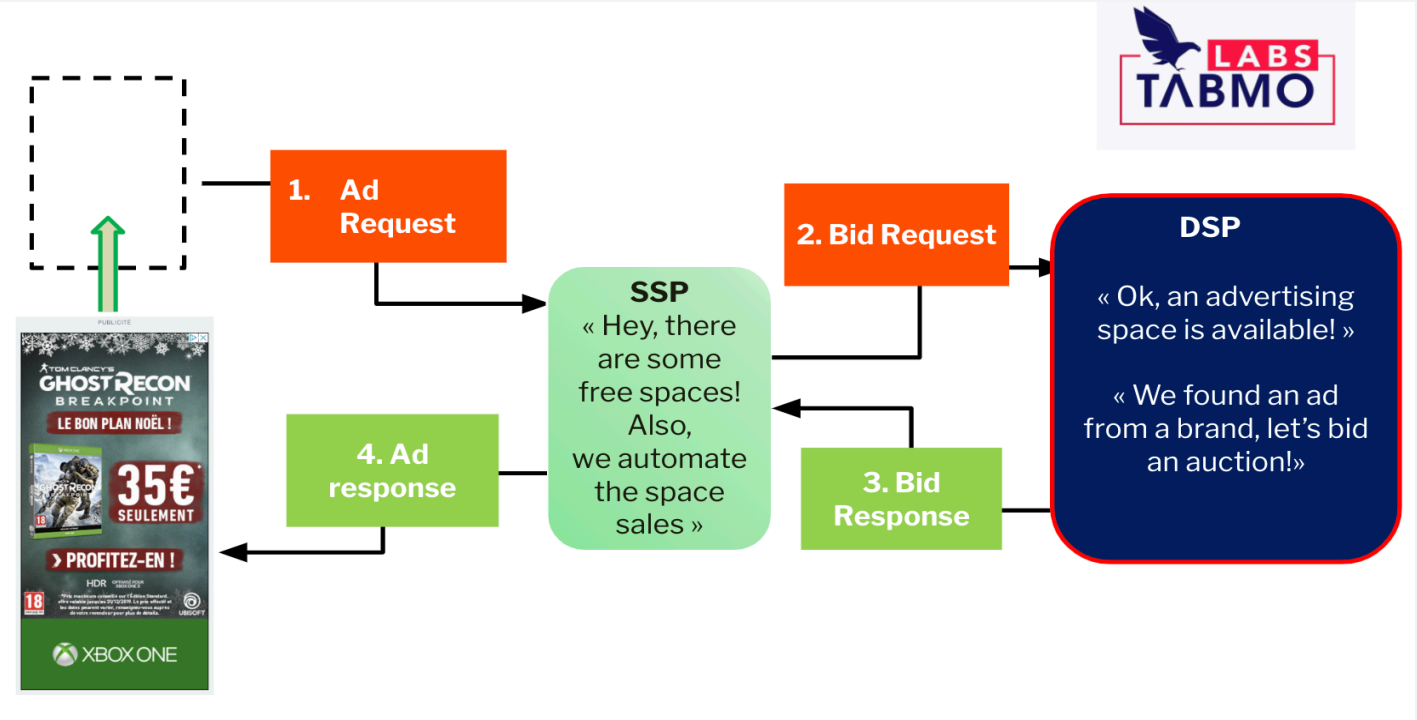
AD YOU LIKE.

Who is TaBmo?

- Adtech company
- Set up in 2013

Main missions

- DSP - Demand Side Platform
- Broadcast ads on mobile

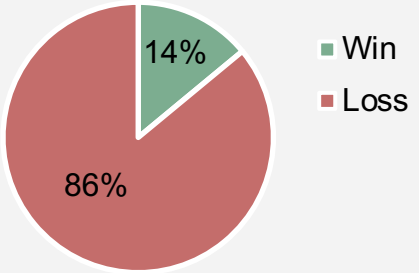

LABS TABMO
1. Ad Request
2. Bid Request
DSP
« Ok, an advertising space is available! »
« We found an ad from a brand, let's bid an auction!»
SSP
« Hey, there are some free spaces! Also, we automate the space sales »
3. Bid Response
4. Ad response
PUBLICITÉ
TOM CLANCY'S GHOST RECON BREAKPOINT
LE BON PLAN NOËL !
35€ SEULEMENT
PROFITEZ-EN !
HDR
18
XBOX ONE

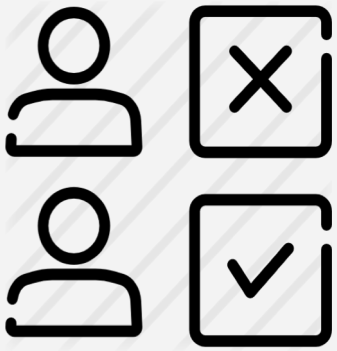
Key values

500 000 Bid Request

500 Bid Response

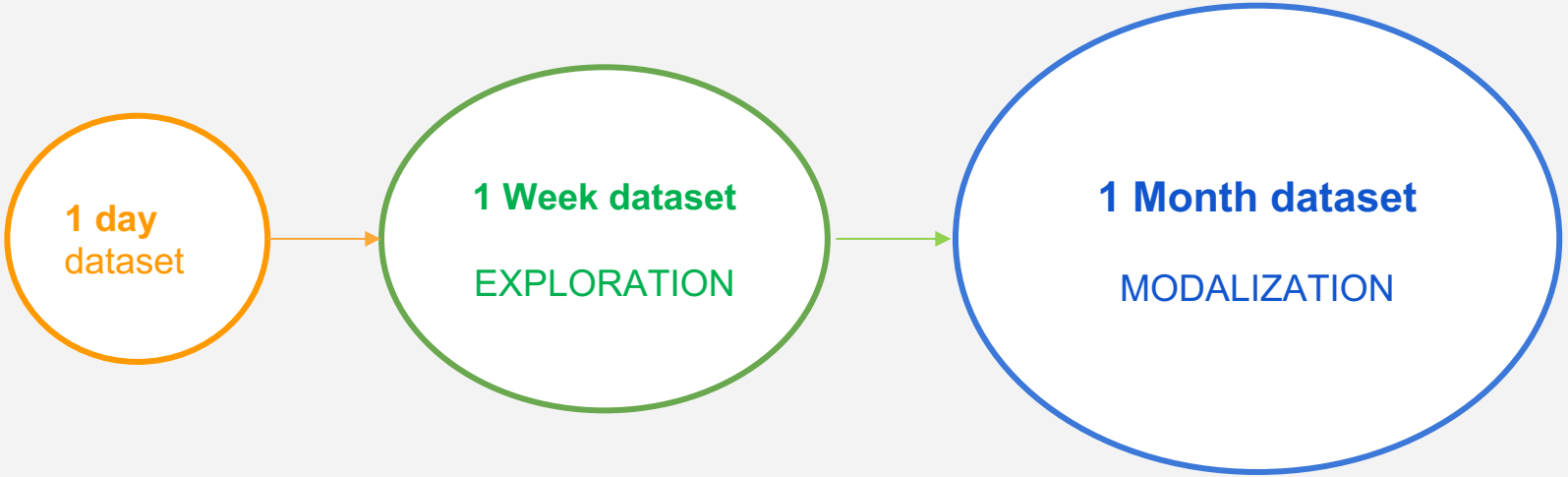
70 Win ⇒ 14%

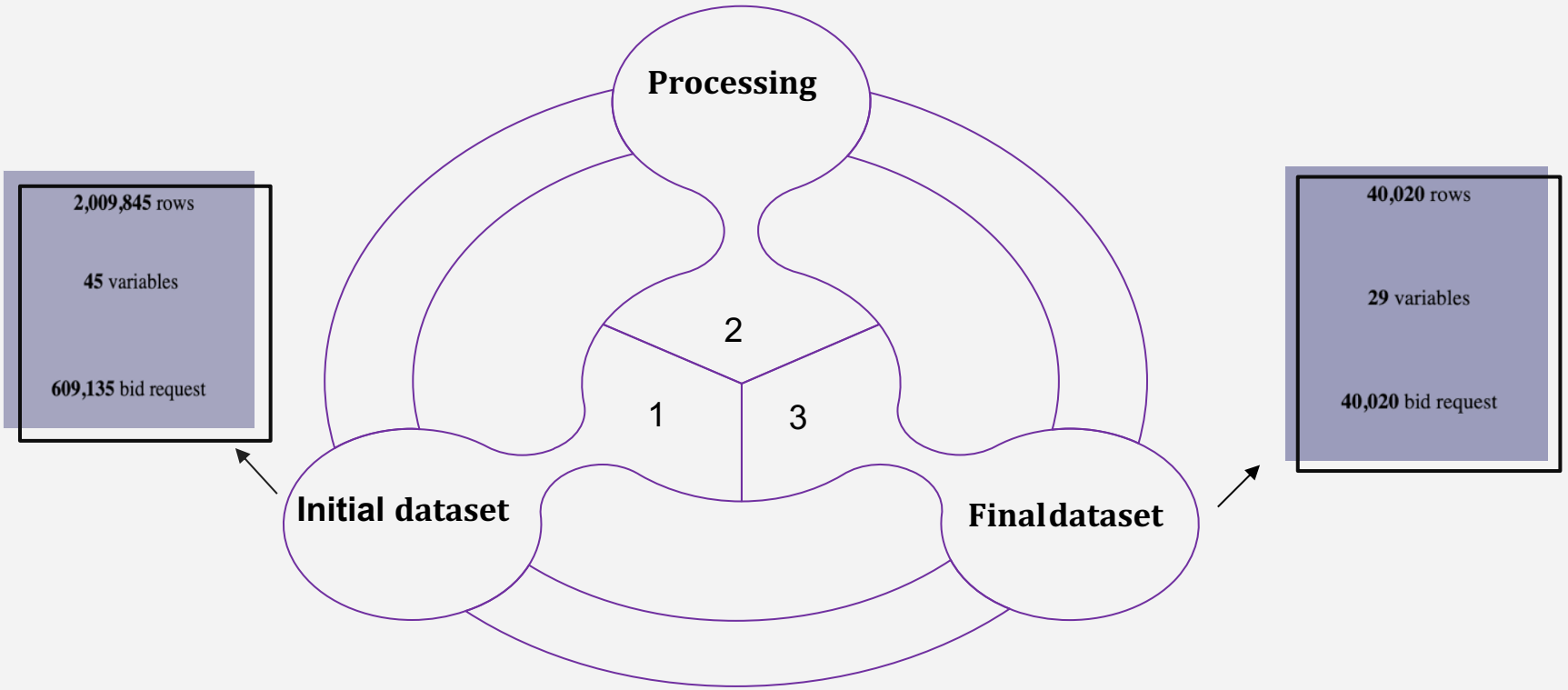




Will a user click on the ad?

Geographic	Device	Time	RTB	Front-end




Processing
2
1
3
Initial dataset
Final dataset
2,009,845 rows
45 variables
609,135 bid request
40,020 rows
29 variables
40,020 bid request

What is a Shiny App?

Why use a Shiny App?

Live Demo

Multiple Correspondance Analysis

What is the goal?

- Identify a group of auctions with similar profil
- Identify the association between categories
- Keep the most contributing variables
- To visualize and synthesize the information

MCA steps

- Retrieve cleaned data
- Compute the MCA
- Choose the number of dimension
- Interpret the visualizations
- Study the groups of variables

Our issues

- Low inertia (less than 1%)
- Complex data
- Large volume of data
- Lack of computing power

Data comprehension

- What do these variables mean?
- What variables are useless and why?

Cleaning

- Incoherence of data
- Data aggregation

Complex feature engineering

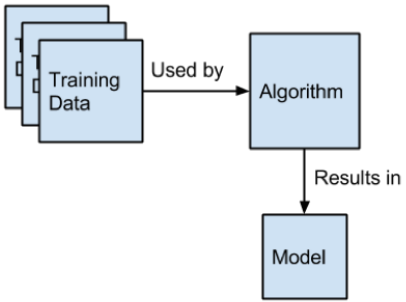
- How can we handle these conflicts?
- What transformation should we apply?

Huge data / Long computing

- High number of modalities for some variables

Predictive model

- SVM
- Logistic regression
- Random forest

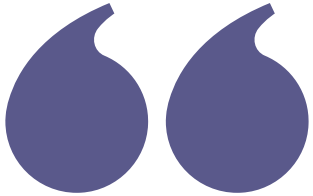


Fraud detection

- Click farms
- Ad stacking
- Pixel stuffing
- other frauds

Computing power

- Huge amount of data
- Time consuming



Thank you for your attention