

Weekly update - 2023-03-28 to 2023-04-04

Paula Jeniffer dos Santos Viriato¹[0000-0003-0900-1686]

Institute of Computing, University of Campinas, Campinas - SP, Brazil
p234831@dac.unicamp.br

1 Dataset Proposals for Customer Segmentation

During the week, I analyzed two datasets in Portuguese. One specialized in posts, and the other specialized in Folha de São Paulo news. News and posts can have the same textual format. Below are examples of the contents (one line from each dataset):

Facebook Posts of Amazon Tourism

- **status_id:** 1549033165419320_2139793586343272;
- **status_message:** "Everyone enjoys a hot roll. And those who deny it, lied. We open at 6:00 pm! Batista Campos - Rua d...";
- **link_name:** "Timeline Photos";
- **status_type:** "photo";
- **status_published:** 2018-05-30 15:22:24;
- **num_reactions:** 6;
- **num_comments:** 1;
- **num_shares:** 1;
- **num_likes:** 6.

News of the Brazilian Newspaper

- **Title:** "Rio police arrest security guard for drug lord in Rocinha favela";
- **Content:** "The police of Rio arrested this Monday afternoon (9) a man appointed as personal security of...";
- **Date:** 2017-09-10;
- **Category:** everyday;
- **Subcategory:** null;

Customer Segmentation algorithms will be applied to both datasets, as shown in Figure 1, and I will compare the results. Mainly on the dataset *Facebook Posts of Amazon Tourism*, an analysis of the psychographic profile will be carried out. The main question at this point is: how can one dataset be related to the other?

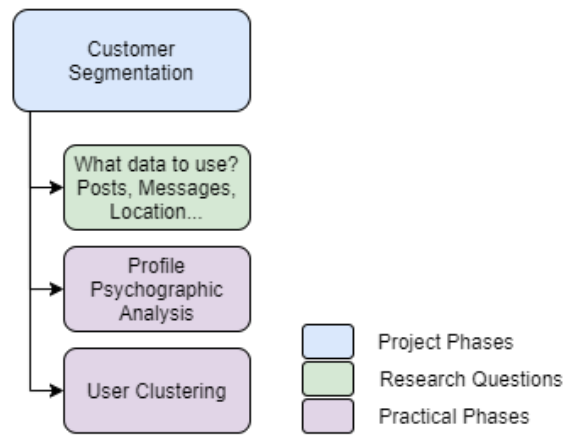


Fig. 1: Customer Segmentation

2 Next steps

This week it was intended to extend searches on *invisible machines*, but it was not possible. In the week that begins today, the proposal is as follows:

- Apply clustering algorithms on the two datasets previously presented;
- Apply a psychographic profile analysis on the *Facebook Posts of Amazon Tourism* dataset;
- Search for a textually descriptive marketing dataset;
- Research on *Invisible Machines*;
- Read the paper *Predicting the Need for Xai from High-Granularity Interaction Data*;
- Project schedule generation;
- Explore courses within the scope of the project:
 - Neuromarketing and mental triggers;
 - Inbound marketing;
 - Digital marketing.
- MC750 Course Planning for Digital Marketing Assessment.