Network Vision Analysis

Step by step described by Eduardo Leite, Lara Dal Molin, Madeline Brennan.

This document operationalises the slide deck methods Network Vision Analysis by Janna Joceli Omena.

1. Use memespector to extracts labels and related data from Google vision, Microsoft Aurora, and Clarifai API's.
2. Use Table2Net to make network out of Lula & Bolsonaro and Zika virus labels, for each API (Google vision, Microsoft Aurora, Clarifai)
    a. type of network- normal (monopartide)
    b. Nodes: column defines the nodes (label descriptions) → Semicolon separated
    c. Links: row number
    d. No addl. Settings
    e. Weight the edges
3. Use gephi to build visualizations of these networks
    a. Run ForceAtlas2 (layout)
    b. Run modularity algorithm (statistics)
    c. Change color, highlight 10 densest nodes
        i. Appearance→ nodes→ partition→ modularity class→ colors (select generate 10 colors)
4. Use rawgraphs→ Circlepack
    a. Hierarchy: modularity class, label
    b. Size: occurrences count
    c. color : modularity class (CSV unique)
    d. Label: label (CSV unique)
        i. Select "auto-hide labels"
5. Quali analysis of the complete circlepacks
    a. Analysis of circle packs derived from each API individually, and then comparison between the different API's.
6. Quanti analysis of total tags per dataset, and total unique tags per dataset
7. Match clusters between datasets according to similar language or , compare semantic spaces
8. Run ego networks

For analysis, we performed a qualitative analysis of complete Circlepacks individually, then compared these results between API's. We then matched clusters between datasets according to similar themes and language, and then compared semantic spaces. We also performed a quantitative analysis of unique tags and total tags each API generated per dataset, as well as average confidence scores per dataset.

Findings:

- Google uses fewer labels to identify the images when compared to Microsoft and Clarifai, while Clarifai uses the most of labels.
- With that being said, Google's computer vision API is more often correct. This demonstrates that the number of labels does not represent accuracy of the API, and contrastingly Clarifai, which uses the most labels, is often incorrect.
- Google is the most precise and specific API and that shows in the number of clusters from the Circle Pack visualization. More clusters mean more specific connections between labels.
- By the kinds of words the APIs use, we could identify that Microfost uses a more generic language, which reflects on the little amount of clusters its network has. Since the words are more generic, they are often used to identify lots of different images.
- Google has a more precise and concise vocabulary, going beyond the generic words while trying to describe the image, using its own labels to identify their context.
- Clarifai uses the biggest amount of words and doesn't stick to being objective, often using adjectives and abstract concepts that go beyond what's in the image, but often being incorrect about them.
- That difference of vocabulary translates in the hard time we had by trying to search similar clusters because they often didn't actually match in their totality.
- When analyzing the same kind of images, we could observe that Microsoft stays in a more safe and descriptive way, with more conservative language, while Google goes beyond and identifies them with more specific terms and Clarifai goes beyond that in trying to identify themes that do not actually appear on the image but could be inferred from the image's context in the real world, but again, Clarifai is often incorrect.
- Our findings also match the ego networks between nodes of the APIs. While Google has a more precise network with strong nodes related to the same labels, Microsoft is more generic and Clarifai goes beyond all and matches those labels with all kinds of other labels.