

Introduction

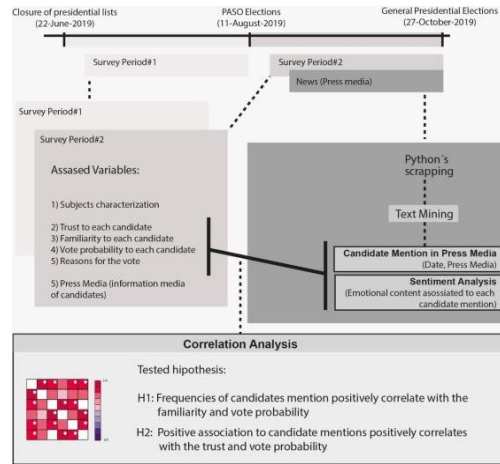
Social media offers the possibility of instant access to political news. Due to the enormous growth of political news on social media, the affective content of headlines could influence the political decision-making process. Every newspaper contains a large number of headlines with affective content about politicians. Computational models have been developed to characterize, identify, predict and influence the perceived subjective valence of politicians from simple texts (e.g. tweets). However, most of these tools are based on English texts, and mostly at sentence level, ignoring the possibility of different polarities for each target in the analyzed text. Therefore, developing sentiment analysis resources for news headlines could be useful to analyse electoral behaviour.

The aim of this study was to assess the subjective perception of individuals for each presidential formula/force from the news headlines of the main newspapers in Argentina during the 2019 elections.

Hypothesis: A HIGHER POSITIVE ASSOCIATION IN NEWS HEADLINES INDUCES AN INCREASE IN THE PERCEIVED TRUST OF CANDIDATES, WHICH AY FAVOUR THEIR ELECTIONS

Methodology and procedures

This work is part of a research project involving Cognitive Experiments and a Social Study conducted during the 2019 Argentinean presidential elections (See Poster A19, on Poster Session A, March 16, 18:30hrs). During the election period, surveys were conducted to characterise familiarity (F), trust (T) and voting probability (VP) for each candidate, while a web scraping bot was designed to scrape news published in the main local media between 21/07/2019 - 21/10/2019 (Fig. 1). The total number of newspaper articles that constituted the base was 22,510 (Table). Using text mining algorithms, the frequency of mention of each presidential formula in total by each media outlet was evaluated. For this purpose, any mention of the presidential candidate (e.g. Alberto Fernández), the vice-presidential candidate (e.g. Cristina Fernández de Kirchner or Cristina Fernández or CFK), the party candidate (e.g. Frente de Todos) or any general reference (e.g. Kirchnerismo) was considered as a mention of the formula. Mentions of other candidates from the same political force (e.g. Axel Kilicoff) were not included in this analysis.



MEDIA PRESS	TOTAL NEWS	HEADLINES
INFOBAE	5956	374
LA NACION	4862	393
CLARIN	4113	412
AMBITO	1739	241
CRONISTA	1513	268
LA IZQUIERDA DIARIO	1275	260
PERFIL	992	70
PAGINA 12	810	82
TIEMPO ARG	712	95
POPULAR	346	39
PRENSA OBRERA	192	19

The analysed candidates correspond to those of the general elections: A. Fernandez (AF), M. Macri (MM), R. Lavagna (RL), N. Del Caño (NDC), J.L. Espert (JLE) and J. Gomez Centirion (JGC). Finally, the relationship between the subjective variables was evaluated by correlations with each other, and with the frequency of mention (See Poster A19, on Poster Session A, March 16, 18:30hrs). From the news dataset, headlines mentioning at least one candidate were extracted and a manual sentiment analysis was carried out to extract from them the subjective perception of whether the headline favours the image of a candidate (positive), disfavours it (negative), or simply describes a fact (neutral). In the end, 2253 headlines were used for this analysis (Table).

Table 1. Total scrapped news and headlines employed in the Social Study.

Fig 1. Social Study conducted during Argentinean Presidential Elections (2019)

Results

Krippendorff's alpha values indicate the agreement of the participants on the headline labels. On the other hand, Krippendorff alpha values on an ordinal scale refer to the degree of agreement with the emotional valence of the headlines. Among the results (Fig.2), MM's degree of agreement stands out with a Krippendorff value of $\alpha=0.619$, AF $\alpha=0.587$, NDC $\alpha=0.464$, JLE $\alpha=0.456$ and RL $\alpha=0.441$. Values greater than 0.4 are acceptable as good agreement. In the case of JGC, the low alpha ($\alpha=0.09$) is explainable by the low n of headlines mentioning it. The results indicate high values of inter-agreement among the participants for most of the candidates.

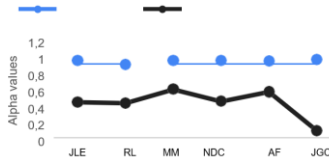
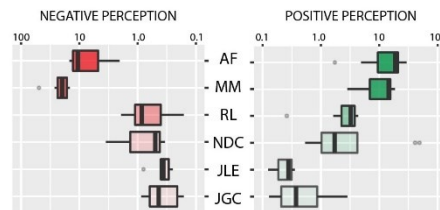


Fig 2. Krippendorff's alpha values for each candidate.

As previously observed for mention analysis, significant higher positive targets were observed again for the first two candidates [AF and MM], though MM shown significant higher negative targets (Fig. 3A). Through a Lineal Regression Analysis between the variables from both datasets, the means of each candidate's Trust significantly better correlates with the Positive mentions [$R^2=0.87$; $p=0.02$], than for Familiarity [$R^2=0.78$; $p=0.04$] (Fig. 3B).

A. Sentiment analysis of news headlines



B. Correlations between subjective variables and positive perception

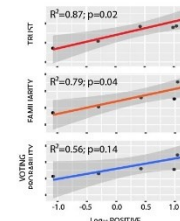


Fig. 3. Sentimental analysis of the news headlines mentioning at least to a candidate. A. Log10 of Relative Positive and Negative values of each news outlet for each candidate/political force. B. Regression and Pearson correlations between the Log10 of positive mentions and the mean of the trust, familiarity, or voting probability for each candidate. R^2 and p values are shown.

Conclusions

The main goal of sentiment analysis is to extract sentiments and opinions expressed in social media and classify them into different polarities. In the present study, a pioneering approach in applying algorithms to a different grammatical structure (news headlines) was carried out. According to the results, the inter-subjective agreements were optimal according to recent research. Therefore, the present tool was able to characterise the participants' perception of subjective agreement. Moreover, a **strong correlation was observed between positive mentions of candidate in the news headlines with Trust and Familiarity** (although with a lesser effect). **The results suggest that news headlines may have a social impact on decision-making at the cognitive-political level.** Although this approach alone is not conclusive and does not prove causality, a strong and significant regression between variables could allow us to support our hypothesis in a more "ecological" scenario.