Detection of Fake News based on readability

Jesús Reyes and Leon Palafox

[1] Instituto Politécnico Nacional
[2] Universidad Panamericana

Abstract. Social media has become one of the principal sources of news consumption, one of their properties is the speed in which the content is created and spread, but the content is not always verified by the users before they share them. This have made easy to generate intentionally false content, there exists different approaches to tackle the detection of fake news mostly of which use English texts for the analysis. We take these works as a basis to analyze and propose some attributes useful in the detection of fake news in Spanish, there exists differences in the way in which the fake news content is generated between language because of the cultural differences and the target audience. We make use of properties of the texts as readability metrics to analyze the difference between content generated by professional journals and not recognized webs, and as a social media approach we use Twitter to analyze how the content is spread between users of this social network.

Keywords: Fake news · Readability · Spanish.

1 Introduction

The detection of fake news is a task which take as basis the veracity of the information to categorize them in true or fake. Nowadays the second most used source of information in accordance with the Mexican news website 'Animal Político' [1] is the social network Facebook (35%). In a survey made by the same website they found that the 22% of the respondents share news content without knowing if the information is true or not. In 2008 the top five websites which publish mostly fake content reached 1.8 million followers.[2]

In this document we propose the combination of some metrics in order to determine if a text published in some news website is potentially fake according to the narrow definition of fake news. From these metrics we are expecting to get a better understanding of the news in Spanish in order to develop a system to detect automatically text in news with potentially fake content.

We take as a definition of fake new we consider the narrow definition of fake news described in [3], in which they claim that a deceptive new is more harmful and less distinguishable than incautiously false news as the former pretends to be truth to better mislead the public, this definition emphasizes both news authenticity and intentions.

Definition 1. *Fake news is intentionally and verifiable false news published by a news outlet. This narrow definition addresses the public's perception of fake news.*

2 Related work

In [4] the authors claim that with the increase of the social media as a principal provider of news at least part of the traditional quality control procedures have disappeared. In their approach referred as stance detection they face the challenge of detect the stance of a claim with regard of another piece of content, they describe this relation as a clickbait detection, i.e., a headline not related with the actual content of the article.

In [5] Buntain and Golbeck tackle the task of classify automatically popular tweets stories as true or fake news, some features are proposed to describe the content of the tweets, these features are grouped in four types: structural, user, content and temporal. The structural features capture proportions of retweets or media shares, the user features refer to properties as interactions, account age, friend/follower counts, the content features include information about the text as polarity and temporary features capture trends of previous features over time.

Readability features has been used in [6], [7] and [8], the readability features uses formulas to count language variables, in a piece of writing in order to provide an index of probably difficulty for readers, some of these metrics are Flech's Reading Ease Score (RES), Dale-Chall, SMOG grading and the Fry graph, in spite of the differences in the way of compute these metrics, all these methods consider that some verbal properties of texts are measurable, taking this consideration as basis we can compute a value which represents the intrinsic complexity of texts, i.e., this value represent the potential difficulty of understand the content of a text, some factors used in the formulas to compute the readability value are the length of phrases, number of syllables, number of words used.

According to Allcot and Gentzkow [9], who analyzed the effect of false stories in the U.S. election of 2016, the fake news are generated from different sources in the web, some of the sites use similar names to well known sources, the fake news are generated with different porpoises for example the revenue they can get by the number of clicks on the news and explicitly show support to some ideology as show support to some of the contenders for the presidency, in the case of Mexico some names and site descriptions claim that they are communicating information that the big media corporations do not want you to known.

There exists many ways in which the authors of fake news spread their content, the more common way is make use use of social networks to sway social opinion. In order to automate the way in which the content will be spread one of the tools used are social bots, these bots are automated accounts that automatically produces content and interacts with humans, trying to mimic and alter their behaviour [10]. In [11] Cai et al. determine some features for their bot detection system, these features are described in three layers: input, which consider timestamps and word embedding, representation, which include behaviour and temporal content, and fusing which represent information of the user through incorporation of both behaviour and content.

3 Dataset

Unlike the works mentioned in the previous section, in this work we does not have news texts labeled as fake or real, we are exploring certain features in order to determine a probability value of potential fake content in the text. In [2] there are mentioned some websites that recurring had been publishing fake news about the 2018 election process in Mexico. From the website ”Nación Unida” [12] a well known fake news site were retrieved 1487 texts, from ”El Heraldo de México” [13] which is a professional journal, 100 samples were retrieved, the objective with these samples is represent both manners in which the news are written.

In order to determine a base space in which we expect to see distributed both ”Nación Unida” [14] and ”El Heraldo de México” samples, we have used the News API to retrieve news from different sources but similar topics.

Source	**Number of samples**
Nación Unida	148
El Heraldo de México	100
CNN	16
Proceso	13
BBC	11
Sopitas	5
Excelsior	4
Aristegui Noticias	4
El país	3
Merca 2.0	3
El Comercio	2
Forbes	2
Animal Político	2
Cultura Colectiva	2

Table 1. Sample distribution of the different sources.

The texts from ”Nación Unida” and ”El Heraldo de México” were pulled from their respective websites using Beautiful Soup [15] because of that some noise were detected were in these texts especially from "Nación Unida", the texts from the rest of the sources were obtained manually from the websites because they are used as a base and it is important to determine were are they distributed in the space according to the proposed metrics without noise.

4 Analysis and results

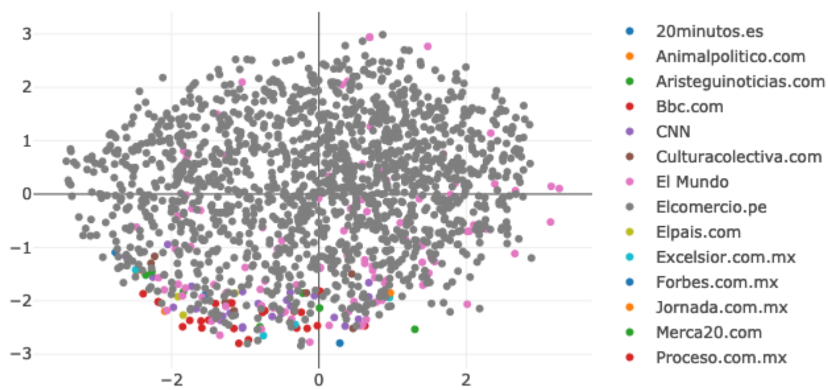
There were used different metrics to determine the a readability score of texts in Spanish, Flesch-Szigriszt index, Fernandez-Huerta, Gutierréz de Polini, Muñoz-Muñoz and Crawford's. In the table below are the max and minimum values for each readability score for the texts from the different sources.

Source	**Crawford**		**Fernandez-Huerta**		**Flesch-Szigriszt**		**Gutierréz de Polini**		**Muñoz-Muñoz**	
	Max	**Min**	**Max**	**Min**	**Max**	**Min**	**Max**	**Min**	**Max**	**Min**
Nación Unida	8.33	0.83	205.60	204.84	73.66	–12.28	33.83	-11.91	178.59	17.94
El Heraldo de México	7.42	-0.50	205.48	204.77	68.50	-0.84	35.02	-14.24	113.2	25.79
CNN	7.01	4.80	205.13	204.91	25.15	-0.68	0.42	-10.38	117.94	73.12
Proceso	6.87	5.21	205.09	204.72	10.05	-14.50	-0.11	-12.56	114.41	69.80
BBC	7.49	4.23	205.11	204.90	20.52	-1.50	-0.59	-6.01	95.52	77.94
Sopitas	6.10	4.89	205.05	204.92	20.15	6.17	4.96	-5.25	108.63	70.99
Excelsior	6.70	5.41	205.09	205.00	18.60	-1.66	-0.06	-6.71	115.00	73.08
Aristegui Noticias	6.59	5.26	204.95	204.86	11.24	2.75	2.95	-5.02	102.24	55.38
El país	7.60	5.15	205.00	204.91	11.63	-5.68	-0.09	-7.01	103.02	93.41
Merca 2.0	7.10	5.99	205.00	204.97	17.82	7.49	-1.21	-3.37	111.83	84.87
El Comercio	6.21	5.83	205.16	205.08	18.96	7.23	1.51	-2.16	99.97	87.38
Forbes	6.33	6.13	204.96	204.88	7.22	-4.22	0.52	-0.62	98.01	68.01
Animal Político	7.05	6.29	204.96	204.85	6.64	-9.39	-2.62	-6.33	114.86	107.55
Cultura Colectiva	6.32	5.94	205.04	205.01	9.93	5.74	2.98	-1.44	109.99	92.20

Table 2. Maximum and minimum index values for the different sources in the dataset.

As can be observed in Table 2 for each source the readability values can get a wider range of values that is among each of the sample spaces there exists samples which are easy to read according to the metrics proposed. In Fig. 1 can be observed the distribution in a 2-D space of the samples using the different metrics, that there exist a region where are distributed the samples obtained from more formal sources. In the frontier of the cloud generated with the samples obtained from "Nación Unida" can be observed the samples gatered from the considered more formal sources as "Proceso", "Aristegui Noticias" and "Forbes" while they considered causal reads as "Cultura Colectiva" and "Sopitas" can be found more closer and inside of the cloud.

Fig. 1. 2-D representation of the samples using T-SNE.



5 Conclusion

As can be observed from the maximum and minimum values that does not exist a well defined distribution of the metrics for the different sources which can help to determine if a text published was made with high standars or is weak in content, but from the visualization in 2-D representing each observation as a combination of different metrics we can observe that there exists difference between the content published by more formal journals in contras to the blog. As future work is needed analyze the phenomenon with more samples from the different sources, because in Fig. 1 can be observed that although most of the samples from "El Heraldo de México" are distributed in the frontier of "Nación Unida" there are samples which are distributed in other zones of the cloud.

References

1. Animal político. https://www.animalpolitico.com/2018/04/mexicanos-noticias-falsas-parametria/. Accessed: 2019-06-10.
2. Verificado. https://verificado.mx/verificado-generadores-divulgadores-noticias-falsas/. Accessed: 2019-06-10.
3. Xinyi Zhou and Reza Zafarani. Fake news: A survey of research, detection methods, and opportunities. *arXiv preprint arXiv:1812.00315*, 2018.
4. Peter Bourgonje, Julian Moreno Schneider, and Georg Rehm. From clickbait to fake news detection: an approach based on detecting the stance of headlines to articles. In *Proceedings of the 2017 EMNLP Workshop: Natural Language Processing meets Journalism*, pages 84–89, 2017.
5. Cody Buntain and Jennifer Golbeck. Automatically identifying fake news in popular twitter threads. In *2017 IEEE International Conference on Smart Cloud (SmartCloud)*, pages 208–215. IEEE, 2017.
6. Verónica Pérez-Rosas, Bennett Kleinberg, Alexandra Lefevre, and Rada Mihalcea. Automatic detection of fake news. *arXiv preprint arXiv:1708.07104*, 2017.
7. Benjamin D Horne and Sibel Adali. This just in: fake news packs a lot in title, uses simpler, repetitive content in text body, more similar to satire than real news. In *Eleventh International AAAI Conference on Web and Social Media*, 2017.
8. Fan Yang, Arjun Mukherjee, and Eduard Dragut. Satirical news detection and analysis using attention mechanism and linguistic features. *arXiv preprint arXiv:1709.01189*, 2017.
9. Hunt Allcott and Matthew Gentzkow. Social media and fake news in the 2016 election. *Journal of economic perspectives*, 31(2):211–36, 2017.
10. Emilio Ferrara, Onur Varol, Clayton Davis, Filippo Menczer, and Alessandro Flammini. The rise of social bots. *Communications of the ACM*, 59(7):96–104, 2016.
11. Chiyu Cai, Linjing Li, and Daniel Zeng. Detecting social bots by jointly modeling deep behavior and content information. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, pages 1995–1998. ACM, 2017.
12. Nación unida. https://nacionunida.com/. Accessed: 2019-06-19.
13. El heraldo de méxico. https://heraldodemexico.com.mx/. Accessed: 2019-06-19.
14. News api. https://newsapi.org/. Accessed: 2019-06-19.
15. Beautiful soup. https://www.crummy.com/software/BeautifulSoup/bs4/doc/.