NewsVallum: Semantics-Aware Text and Image Processing for Fake News Detection system*

Giuliano Armano[1], Sebastiano Battiato[2], Davide Bennato[2], Ludovico Boratto[3], Salvatore M. Carta[1], Tommaso Di Noia[4], Eugenio Di Sciascio[4], Alessandro Ortis[2], Diego Reforgiato Recupero[1]

[1]Università degli studi di Cagliari
{salvatore,diego.reforgiato}@unica.it
[2]Università degli studi di Catania
battiato@dmi.unict.it, dbennato@unict.it, ortis@dmi.unict.it
[3]EURECAT, Centre Tecnológic de Catalunya
ludovico.boratto@acm.org
[4]Politecnico di Bari
{tommaso.dinoia,eugenio.disciascio}@poliba.it

Abstract. As a consequence of the social revolution we faced on the Web, news and information we daily enjoy may come from different and diverse sources which are not necessarily the traditional ones such as newspapers, either in their paper or online version, television, radio, etc. Everyone on the Web is allowed to produce and share news which can soon become viral if they follow the new media channels represented by social networks. This freedom in producing and sharing news comes with a counter-effect: the proliferation of fake news. Unfortunately, they can be very effective and may influence people and, more generally, the public opinion. We propose a combined approach of natural language and image processing that takes into account the semantics encoded within both text and images coming with news together with contextual information that may help in the classification of a news as fake or not.

1 Introduction

Over the past few years, a number of high-profile conspiracy theories and false stories have originated and spread on the Web. After the Boston Marathon bombings in 2013, a large number of tweets started to claim that the bombings were a "false flag" perpetrated by the United States government. Also, the GamerGate controversy started as a blogpost by a jaded ex-boyfriend that turned into a pseudo-political campaign of targeted online harassment. More recently, the Pizzagate conspiracy – a debunked theory connecting a restaurant and members of the US Democratic Party to a child sex ring – led to a shooting in a Washington DC restaurant. These stories were all propagated, in no small part, via the use of "alternative" news sites like Infowars and "fringe" Web communities like 4chan. There are many reasons for the rise in alternative narratives,

* Discussion Paper

ranging from libelous (e.g., to harm the image of a particular person or group), political (e.g,. to influence voters), profit (e.g., to make money from advertising), or trolling. The barrier of entry for such alternative news sources has been greatly reduced by the Web and large social networks which offer the possibility to access to a growing amount of public content that is published worldwide every day [17]. Due to the negligible cost of distributing information over social media, fringe sites can quickly gain traction with large audiences. On the other hand, the growing impact of social media platforms allows any individual to public contents that are then spread globally. The entity of the diffusion of social media content is not related to its quality, but rather it is related to the content virality. However, the news shared on social platforms are not verified by users. As consequence, fake contents that stimulates the interest of people (i.e., viral content) can be spread at large scale in very short time. Social media further changed the way we communicate online. Indeed, social posts contain very short text, often accompanied with a picture aimed to grab attention of users. Indeed, a social media post is more likely to get engagement if a photo accompanies it. Fake news presents itself as a content whose viral spread is possible thanks to the system of values it refers to. People tend to share content that expresses their own vision of the world – political, social, economic – usually without further checking. For this reason, we can try to imagine a process of checking the false content by referring to a series of metrics about the text and the content structure that could be used to identify the digital characteristics of a fake news.

2 Fake News Detection

The problem of detecting fake news is currently widely studied in the literature [7] and there are sources that are known to spread fake news (e.g., both Wikipedia and FakeNewsWatch contain lists of fake-news websites). There are some experiments that try to identify fake news from a socio-structural point of view, such as the Hoaxy project of the Indiana University (`https://hoaxy.iuni.iu.edu/`), which uses as an identification tool of false news the circulation of content among users through social media [26]. Conspiratorial narratives share several elements: the origin within particular social spaces (e.g., VKontakte, Reddit, 4Chan), the reference to an ideologically structured community (e.g., the no vax and alt-right groups), the use of deep fears (e.g., fears towards children). The conversation network topology is another element that in some cases reveals the content of the conversation itself, as shown in the Mapping Twitter Topic Network project.

Although academic research dedicated high effort to fake news detection algorithms in the recent years, there is lack of solutions able to provide an on-demand, real time, news classification. Indeed, such a system requires a wide range of competences ranging from sociology, to combined image and Natural Language Processing analysis, to big data management, Semantic Web, Machine Learning, etc. [14,11,16,25].
The amount of fake news is reaching crisis proportions and it is getting worse.

Among the Vs that characterize Big Data (volume, variety, velocity), we have also the challenge of data veracity. Misinformation dynamics is where the big data concept of data veracity and the problem of fake news connect. Fake news is intentional misinformation and it is also dynamic.
Leveraging artificial intelligence is definitely a must to combat fake news. Artificial intelligence is able to learn behaviors based upon continually improving pattern recognition, so training a system to identify fake news based upon what sort of articles people have flagged as misinformation in the past is well within the reach of today's technology. Moreover, it is straightforward and very effective to apply machine learning tools of Big Data frameworks to come up with classification algorithms with the aim to spot fake news. In the context of fake news, tampered images represent a crucial role in reaching virality and allowing fake news to be shared at large scale. Although methods of multimedia forensics are able to detect tampered images [2,3,12,5], modern applications of AI can generate fake images depicting places, rooms, animals and even existing people. So far, both academic and industrial research laboratories built systems that can recognize faces and common objects in images with increasing accuracy. Now, similar methods are able to create realistic images. In this context, research efforts and development of recognition tools and systems for fake visual contents shared on the Web are needed [4,18]. Also, the exploitation of Natural Language Processing and Semantic Web technologies (lexical and semantic resources) and Cognitive Computation frameworks and tools (e.g., IBM Watson) is definitely a must. Those technologies allow going deeper toward the full comprehension of a given text [15]. They have already been successfully adopted in several artificial intelligence and natural language processing tasks and their combined usage has shown to outperform methods where they were used individually [16,22,24,13]. Their employment will provide huge benefits for the fake news recognition as well.
Moreover, for the purposes of reaching the goal of fake news detection it is important to elaborate a scale of fake news complexity - from hoaxes to conspiracy theories - so as to isolate the digital elements that can help identifying the type of fake news.

2.1 The NewsVallum Approach

The proposed approach, we named **NewsVallum**, is expected to have high impact and to improve the state of the art on a wide range of research fields, ranging from applied artificial intelligence to the continuously disrupting field of the interaction between sociology and technology. A holistic approach is adopted, using the leading edge artificial intelligent techniques focused on image and text analysis. A joint approach to handle these two aspects that are foremost for understanding the characteristics of fake news is of paramount importance. In fact, people enjoy and share contents if they are in accordance with their vision of the world. As this subjective evaluation can be guided by both text and images, contents may be used to influence people towards a specific interpretation (this communicative process is called "framing").

As artificial intelligence is sometimes heralded as the new industrial revolution, a key motto of NewsVallum is to consider deep learning and reinforcement learning as the steam engine of this revolution, whereas data can be represented as its coal. A multimedia forensics approach is also used to make it possible to concentrate on news manipulation strategies not only at a technological level (i.e., using photo-editing programs), but also at a context level (i.e., photos of places not corresponding to the text). In so doing, an analytical interpretation that goes beyond the level of technological manipulation is made, also concerning the contextualization of any given content. Another important characteristic of NewsVallum relies on the massive exploitation of sociological orientation and crowdsourcing. Here, the social dimension of fake news is used as theoretical reference to understand the consequences of the social processes concerning the main technological aspects related to the diffusion and circulation of fake contents. Moreover, the social dimension occurs also when the interpretation of contents is combined with multimedia forensics analysis strategies. This tends to compensate the lack of computational social science that typically characterizes research focused on fake news. In fact, being related to digital platforms, current research works have too often a strong technological bias and tend to forget the social dynamics of underlying the process. Crowdsourcing creates a collaboration between human analysis and computational approach, thus allowing to naturally focus on the human limits and on the technological constraints of fake contents. The NewsVallum approach targets two main goals. First of all, it is aimed at evaluating and assessing the reliability of information that is spread, published and shared through web platforms on a daily and also hourly basis. Indeed, so far the problem of automatic fake news classification has been tackled considering either text or images. In particular, the automatic classification of a piece of text as news has been tackled in [27] by applying a neural network and advanced text processing techniques. Authors in [20] highlighted the importance that images have in framing news stories and [19] studied how to characterize and identify fake images on Twitter during hurricane Sandy. To our knowledge, no proposals have been made able to classify a content as reliable or fake, which take into account the contribution of both text and images. Furthermore, NewsVallum aims to perform a real-time analysis of news, so that the procedures that will be devised and implemented could be used as web brower add-on. Although very challenging to obtain, providing the ability of issuing results in real time will be a crucial feature of the system, also due to the rapid spread that fake news have in social media. This first target will be the foundation for a yet more ambitious second target: the realization of a technology able to give a real time answer to common citizens needing an answer on the veracity of news found in social media and, in general, on the web. Through the exploitation of Linked Data technologies combined with personalized information access [1,6,**?**] the user is allowed to identify good and fake news as well as their sources. Recommendation systems fed by semantics-aware data have been proved to be very effective in the identification of items that can be of interest for (groups of) users [9,10] also in mobile scenarios [23] by leveraging similarities and relatedness between

entities in a knowledge graph [8]. Other than a proactive usage of data coming from Linked Data datasets, exploratory processes of Knowledge graphs [21] by the user may be helpful in finding the right information and, in case, identify fake semantic associations.

2.2 Fake news detection and its sociological impact

NewsVallum targets both text and images together and this has also important sociological aspects and consequences for the following reasons.

The issue of fake news presents itself as the last stage of typical propaganda processes. The novelty lies in the introduction of the computational variable within the disinformation dynamics. According to the sociological theory of media construction of reality, media have a very important role in shaping our perception of the world. As long as mass media and social media keep spreading fake contents, the perception of the world could be modified accordingly, with enormous risks for democracy, security, and civil coexistence. The digital dimension of these dynamics makes these processes very complex, giving small chances to individuals in the task of finding truthful news and information. The properties of digital content such as persistence, replicability, scalability, searchability, greatly enhances the negative consequences that fake news can have on the society. This is the reason why, an innovative approach to the problem of disinformation is needed, able to combine advanced digital technologies (in particular, artificial intelligence, integrated analysis approaches, and multimedia forensics) with the goal of implementing a general framework able to give support in the task of understanding some of the social processes that underlie the current digital society. Only in this way it will be possible to develop a critical thinking positively biased by the underlying digital society. This result can be obtained by devising and implementing sophisticated technologies and approaches for fighting against a phenomenon that may have dangerous effects (on the medium and long run) on the whole society. The important and evolving requirements from the sociological point of view and the consequent technological challenges will be faced exploiting the recent revolutionary advances of AI domain. Indeed, there is evidence that modern artificial intelligence techniques (e.g. deep learning) are good at learning behaviors based on data, avoiding the influence of subjective (i.e., biased) opinions. These systems, which continuously improve the knowledge of patterns in the processed data, can be developed and trained for identifying fake news based on contents that people have flagged as misinformation in the past. Hence, this goal is well within the reach of today's technology.

Important goals that NewsVallum aims to reach for the technologies and algorithms developed in the approach are: i) simplicity, ii) scalability, as well as iii) versatility and reusability. The first goal can be attained considering that the employment of deep learning removes the need for feature engineering, replacing complex engineering-heavy pipelines with simple, end-to-end trainable models. Scalability will be addressed taking advantage of the Moore's law, as the proposed approaches are highly amenable to parallelization on GPUs. Versatility

and reusability can be obtained as the deep learning models to be developed can be trained on additional data without restarting from scratch. This last feature is very important for the domain of fake news, in which continuous online learning is a preferred option.

Fake news detection is currently a topic of central interest at national and international level. At national level, one of the biggest Italian newspaper claims that 2017 was the year of fake news spread and even in 2018 several Italian newspapers discuss how fake news are fueling misinformation in politics, economics, and sport events. There is a specialized national police team entrusted with facing the problem, although it usually lacks of the powerful tools to deal with the problem. In the US, the spread of fake news has been seen as part of the rise of post-truth politics, in which the debate is framed by appeals to emotion disconnected from the details of policy. A study conducted by Stanford Graduate School of Education revealed difficulties of middle, high school and college students experienced in differentiating between advertisements and news articles where the information originated. The same study showed that 44% of all adults get their news from Facebook, and further investigations showed that nearly 40% of content by far-right Facebook pages and 19% of extreme left-leaning pages were false or misleading

3 Conclusion

The growing interest in fake news is motivated by the fact that people are typically not suited to distinguish between good information and fake news, in particular when the source of information is the Internet (and especially social media). In this context, the global nature of such an information-sharing environment allows fake news having a big impact in many fields, including politics, business, and health, at a worldwide scale. For these reasons, we expect that NewsVallum will find several practical applications that concern the verification of online contents. While providing services to detect, authenticate and check the reliability of online contents, the activities and research achievements related to NewsVallum would pursue and accelerate the technological and scientific progress in the field. Moreover, deep learning is currently considered a main research topic in the field of AI, as it creates the conditions for spreading the AI technology in various kinds of applications. NewsVallum is consistent with this trend, as it will give opportunities to young researchers for getting trained in this new and promising research topic. It is worth pointing out that the problem related with fake news, while affecting the information ecosystem, may also have negative impact on the democracy itself. In fact, fake contents may be used to generate hatred towards a social category, ruining the reputation of a political entity, and/or steering political elections with campaigns aimed at spreading false contents. Fake news is also an economic problem. Different economic stakeholders - no matter whether they are individuals, business companies or associations - perform decision-making by gathering information on different topics from digital sources, such as search engines or contents conveyed by social media. This

is the reason why fake news, other than being an issue for the world of information, is also a problem for the world of business. In fact, in a society in which the speed of collecting and evaluating information is a strategic variable, fake news could lead to a wrong understanding of a particular socio-political scenario, with negative economic consequences. The expected impact of NewsVallum is to provide a set of tools to be used while assessing multimedia contents with the goal of unmasking fake news, so that political and/or economic decisions can be based on a proper evaluation of the data at hand.

References

1. Vito Walter Anelli, Tommaso Di Noia, Pasquale Lops, and Eugenio Di Sciascio. Feature factorization for top-n recommendation: From item rating to features relevance. In *Proceedings of the 1st Workshop on Intelligent Recommender Systems by Knowledge Transfer & Learning co-located with RecSys 2017*, pages 16–21, 2017.
2. Sebastiano Battiato, Giovanni Maria Farinella, Enrico Messina, and Giovanni Puglisi. Robust image alignment for tampering detection. *IEEE Transactions on Information Forensics and Security*, 7(4):1105–1117, 2012.
3. Sebastiano Battiato, Giovanni Maria Farinella, Giovanni Puglisi, and Daniele Ravì. Aligning codebooks for near duplicate image detection. *Multimedia Tools and Applications*, 72(2):1483–1506, 2014.
4. Sebastiano Battiato, Oliver Giudice, and Antonino Paratore. Multimedia forensics: discovering the history of multimedia contents. In *Proceedings of the 17th International Conference on Computer Systems and Technologies 2016*, pages 5–16, 2016.
5. Sebastiano Battiato and Giuseppe Messina. Digital forgery estimation into dct domain: A critical analysis. In *Proceedings of the First ACM Workshop on Multimedia in Forensics*, MiFor '09, pages 37–42, New York, NY, USA, 2009. ACM.
6. Vito Bellini, Vito Walter Anelli, Tommaso Di Noia, and Eugenio Di Sciascio. Auto-encoding user ratings via knowledge graphs in recommendation scenarios. In *Proceedings of the 2Nd Workshop on Deep Learning for Recommender Systems*, DLRS 2017, pages 60–66, 2017.
7. Davide Bennato. The shift from public science communication to public relations. the vaxxed case. *JCOM*, 16(02):C02_en–2, 2017.
8. T. Di Noia, V.C. Ostuni, J. Rosati, P. Tomeo, E. Di Sciascio, R. Mirizzi, and C. Bartolini. Building a relatedness graph from linked open data: A case study in the it domain. *Expert Systems with Applications*, 44:354–366, 2016.
9. T. Di Noia, V.C. Ostuni, P. Tomeo, and E. Di Sciascio. Sprank: Semantic path-based ranking for top-n recommendations using linked open data. *ACM Transactions on Intelligent Systems and Technology*, 8(1), 2016.
10. Tommaso Di Noia, Vito Claudio Ostuni, Jessica Rosati, Paolo Tomeo, and Eugenio Di Sciascio. An analysis of users' propensity toward diversity in recommendations. In *Proceedings of the 8th ACM Conference on Recommender Systems*, RecSys '14, pages 285–288, 2014.
11. Amna Dridi and Diego Reforgiato Recupero. Leveraging semantics for sentiment polarity detection in social media. *International Journal of Machine Learning and Cybernetics*, 2017.
12. Fausto Galvan, Giovanni Puglisi, Arcangelo Ranieri Bruna, and Sebastiano Battiato. First quantization matrix estimation from double compressed jpeg images. *IEEE Transactions on Information Forensics and Security*, 9(8):1299–1310, 2014.

13. Aldo Gangemi, Valentina Presutti, and Diego Reforgiato Recupero. Frame-based detection of opinion holders and topics: A model and a tool. *IEEE Comp. Int. Mag.*, 9(1):20–30, 2014.
14. Aldo Gangemi, Valentina Presutti, Diego Reforgiato Recupero, Andrea Giovanni Nuzzolese, Francesco Draicchio, and Misael Mongiovì. Semantic web machine reading with FRED. *Semantic Web*, 8(6):873–893, 2017.
15. Aldo Gangemi, Valentina Presutti, Diego Reforgiato Recupero, Andrea Giovanni Nuzzolese, Francesco Draicchio, and Misael Mongiovì. Semantic web machine reading with fred. *Semantic Web*, 8(6):873–893, 2017.
16. Aldo Gangemi, Diego Reforgiato Recupero, Misael Mongiovì, Andrea Giovanni Nuzzolese, and Valentina Presutti. Identifying motifs for evaluating open knowledge extraction on the web. *Knowl.-Based Syst.*, 108:33–41, 2016.
17. Fabio Giglietto, Luca Rossi, and Davide Bennato. The open laboratory: Limits and possibilities of using facebook, twitter, and youtube as a research data source. *Journal of Technology in Human Services*, 30(3-4):145–159, 2012.
18. Oliver Giudice, Antonino Paratore, Marco Moltisanti, and Sebastiano Battiato. A classification engine for image ballistics of social data. In *International Conference on Image Analysis and Processing*, pages 625–636. Springer, 2017.
19. Aditi Gupta, Hemank Lamba, Ponnurangam Kumaraguru, and Anupam Joshi. Faking sandy: characterizing and identifying fake images on twitter during hurricane sandy. In *Proceedings of the 22nd international conference on World Wide Web*, pages 729–736. ACM, 2013.
20. Paul Messaris and Linus Abraham. The role of images in framing news stories. *Framing public life: Perspectives on media and our understanding of the social world*, pages 215–226, 2001.
21. Roberto Mirizzi, Azzurra Ragone, Tommaso Di Noia, and Eugenio Di Sciascio. Semantic wonder cloud: Exploratory search in dbpedia. In *Proceedings of the 10th International Conference on Current Trends in Web Engineering*, ICWE'10, pages 138–149, 2010.
22. Misael Mongiovì, Diego Reforgiato Recupero, Aldo Gangemi, Valentina Presutti, and Sergio Consoli. Merging open knowledge extracted from text with MERGILO. *Knowl.-Based Syst.*, 108:155–167, 2016.
23. Vito Claudio Ostuni, Giosia Gentile, Tommaso Di Noia, Roberto Mirizzi, Davide Romito, and Eugenio Di Sciascio. Mobile movie recommendations with linked data. In *Availability, Reliability, and Security in Information Systems and HCI - IFIP WG 8.4, 8.9, TC 5 International Cross-Domain Conference, CD-ARES 2013*, pages 400–415, 2013.
24. Diego Reforgiato Recupero, Valentina Presutti, Sergio Consoli, Aldo Gangemi, and Andrea Giovanni Nuzzolese. Sentilo: Frame-based sentiment analysis. *Cognitive Computation*, 7(2):211–225, 2015.
25. Diego Reforgiato Recupero, Sergio Consoli, Aldo Gangemi, Andrea Giovanni Nuzzolese, and Daria Spampinato. A semantic web based core engine to efficiently perform sentiment analysis. In *The Semantic Web: ESWC 2014 Satellite Events*, pages 245–248. Springer, 2014.
26. Chengcheng Shao, Giovanni Luca Ciampaglia, Alessandro Flammini, and Filippo Menczer. Hoaxy: A platform for tracking online misinformation. In *Proceedings of the 25th international conference companion on world wide web*, pages 745–750, 2016.
27. Marin Vuković, Krešimir Pripužić, and Hrvoje Belani. An intelligent automatic hoax detection system. In *International Conference on Knowledge-Based and Intelligent Information and Engineering Systems*, pages 318–325. Springer, 2009.