



HAL
open science

Secredisdata 2018: Special session on sentiment, emotion, and credibility of information in social data (Turin 2018)

Farah Benamara, Cristina Bosco, Elisabetta Fersini, Gabriella Pasi, Viviana Patti, Marco Viviani

► To cite this version:

Farah Benamara, Cristina Bosco, Elisabetta Fersini, Gabriella Pasi, Viviana Patti, et al.. Secredisdata 2018: Special session on sentiment, emotion, and credibility of information in social data (Turin 2018). Benamara, Farah; Bosco, Cristina; Fersini, Elisabetta; Pasi, Gabriella; Patti, Viviana; Viviani, Marco. IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA 2018), IEEE Computer Society, pp.638-640, 2018, 10.1109/DSAA.2018.00082 . hal-03044244

HAL Id: hal-03044244

<https://hal.science/hal-03044244>

Submitted on 10 Dec 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

SeCredISData 2018: Special Session on Sentiment, Emotion, and Credibility of Information in Social Data

Farah Benamara
IRIT
Paul Sabatier University
Toulouse, France
Email: benamara@irit.fr

Cristina Bosco, Viviana Patti*
Dept. of Computer Science
University of Turin
Turin, Italy
Email: {bosco, patti}@di.unito.it

Elisabetta Fersini, Gabriella Pasi, Marco Viviani
Dept. of Informatics, Systems, and Communication
University of Milano-Bicocca
Milan, Italy
Email: {fersini, pasi, marco.viviani}@disco.unimib.it

Abstract—The Social Web represents nowadays the principal means to support and foster social interactions among people through Web 2.0 technologies. Individuals interact in virtual communities to pursue mutual interests or goals, by exchanging multiple kinds of contents (i.e., textual, acoustic, visual), the so-called User-Generated Content (UGC). In this context, the SeCredISData Special Session is especially devoted at discussing the implications that the analysis of big social data has in tackling open issues related to society from different perspectives. On one side, there is the need to push forward the research on emotion and sentiment, and the investigation of affective cognitive models and their possible integration into intelligent systems. On the other side, it is urgent to address the issue of on-line information credibility assessment, in an era where trusted intermediaries have disappeared and people must rely only on their cognitive capacities to judge information. The Special Session is therefore aimed at promoting the development of models and applications able to tackle these issues.

Keywords—sentiment analysis; credibility; social media; opinion mining; fake news detection; hate speech detection

I. INTRODUCTION

The rise of social media and the availability of big social data, in particular in the form of User-Generated Content (UGC), represent nowadays a challenge in the fields of data analysis and data science. The multiple kinds of contents (i.e., textual, acoustic, visual) that people share within the Social Web can be employed to different purposes and analyzed under different perspectives to tackle several issues related to society.

On the one hand, the focus of the SeCredISData Special Session¹ is given to the study and the application of affective computing and sentiment analysis [1], [2] to social data, which can impact on monitoring, analyzing and counteracting discrimination and hate speech, which are increasingly spreading phenomena in our countries also in combination with the pervasiveness of social media [3]–[5]. A further aspect to be taken into account in the automatic analysis

of sentiment is the use of figurative language [6]–[9], also considering that often, sarcastic messages are the ones that spread more virulently. Furthermore, also the applications of sentiment analysis and emotion detection in social media for the development of education, entertainment, cultural heritage, health, e-government, and games is considered as interesting object of investigation [10].

On the other hand, by considering the process of ‘disintermediation’ that affects social media, characterized by the absence of traditional trusted intermediaries in the process of information generation and diffusion [11], [12], the Special Session will also investigate the problem of assessing the credibility of UGC spreading among and across virtual communities [13]. The diffusion of fake news, hoaxes, rumors, fake reviews, inaccurate health information, can have a negative impact on society with respect to different aspects [14], from influencing political elections, producing harmful effects if connected to the health of patients, to generating hate and discrimination phenomena. For all these reasons, the study and the development of approaches that can help people in automatically assess the level of credibility of information is a fundamental research issue in the last years [13].

II. SCOPE AND TOPICS OF INTEREST

The aim of the SeCredISData Special Session is to cover different aspects related to data analysis applied to big social data, by addressing to a heterogeneous community of researchers who has data science as a common denominator. Big social data analysis is interdisciplinary by its nature, and combines disciplines such as natural language processing, social network analysis, multimedia management, social media analytics, trend discovery, information retrieval, computational linguistics. Therefore, areas of interest to SeCredISData in particular, and to DSSA 2018² in general include, but are not limited to:

- Subjectivity, sentiment, and emotion detection;
- Sentiment-based indexing, search, and retrieval;

*The work of C. Bosco and V. Patti was partially funded by Progetto di Ateneo/CSP 2016 (*Immigrants, Hate and Prejudice in Social Media*, S1618_L2_BOSC_01).

¹<http://www.ir.disco.unimib.it/secredisdata2018/>

²<https://dsaa2018.isi.it>

- Sentiment topic detection and trend discovery;
- Multimodal emotion and sentiment detection;
- Irony and sarcasm detection;
- Affect in natural language and multimodal interaction;
- Emotion models and ontologies of emotions;
- Hate speech detection and ethical issues in affect and opinion detection;
- Summarization of affective data;
- Sentiment analysis and emotion detection in social media applications;
- Opinion spam, group spam, fake news detection;
- Credibility/reliability of health-related information;
- Multimedia content credibility;
- Information/misinformation diffusion;
- Trust and reputation in virtual communities;
- Retrieval of credible information;
- Credibility gold standard datasets generation;
- Fact-checking and crowdsourcing credibility.

III. SUBMISSIONS AND ACCEPTED PAPERS

The SeCredISData Special Session has received a good number of submissions, i.e., eleven, which demonstrates the interest of the scientific community about the problems related to the analysis of the sentiment, the emotion and the credibility of information that diffuses on-line, in social media in particular. Of the eleven submissions, six were accepted to be presented at the conference and published on the Proceedings of DSAA 2018, with an acceptance rate of about 55%.

The accepted articles deal with interesting issues that are fairly evenly distributed among the topics of interest of the Special Session. A first paper, entitled: “Willingness to Share Emotion Information on Social Media: Influence of Personality and Social Context”, presents a user study, carried out on 190 subjects, exploring the willingness of Web users to share emotion information either through self-reporting or through facial expressions. The study tries to find correlations between sharing and feedback mechanisms, as well as sharing and personality traits. In the paper entitled: “Utilizing Information from Tweets for Detection of Sentiment-based Interaction Communities on Twitter”, the authors investigate the task of detecting semantically-meaningful communities on Twitter, by proposing a set of modifications to the traditional way edges are built and weighted between interacting nodes of a Twitter graph. The idea of introducing semantics in standard analytical metrics and tasks is promising. Another paper, entitled: “An Empirical Analysis of the Role of Amplifiers, Downtoners, and Negations in Emotion Classification in Microblogs”, presents the impact of amplifiers, downtoners, and negations on emotion words and on document classification, by considering tweets. One of the interesting aspects of the paper is that modifiers have been studied in sentiment analysis but rarely in emotion analysis. The paper:

“Portability of Aspect Based Sentiment Analysis: Thirty Minutes for a Proof of Concept” introduces an unsupervised method for the detection of finer-grained sentiment aspects of product/services in reviews, by integrating graph-based extraction rules over dependency trees and distributional semantics techniques, and addressing the interesting issue of domain portability. In: “‘Behind the words’: psychological paths underlying the un/supportive stance toward immigrants in social media environments”, the authors aim at studying supportive and unsupportive attitudes in social media, with particular reference at exploring psychological paths underlying social media stances toward immigrants in the state of need, a problem of sad actuality. Finally, the paper entitled: “A Methodological Template to Construct Ground Truth of Authentic and Fake Online Reviews”, discusses an interesting research issue, i.e., the way of building ground truth datasets in the area of on-line credibility assessment, where the problem of collecting labeled data with respect to the credibility of information has not been solved yet in the literature. This would help researchers to better evaluate ongoing research, which constitutes one of the major issues in opinion spam detection [15].

IV. CONCLUSION

The SeCredISData Special Session had the objective to promote the development of methods and approaches for the analysis of big social data that can have positive repercussions on society. In particular, the Special Session was addressed to all those researchers who employ data science as a common denominator to tackle issues related to the analysis of the sentiment, the emotion and the credibility of on-line information, in social media and virtual communities in particular. The success of the Special Session demonstrates the interest of the scientific community in these aspects, and the need to continue research in this area.

ACKNOWLEDGMENT

The SeCredISData Organizers would like to thank the members of the Program Committee for their valuable contribution: Snehasish Banerjee, Valerio Basile, Stefano Cresci, Rossana Damiano, Rino Falcone, Carlos A. Iglesias, Giovanni Livraga, Mihai Lupu, Stefania Marrara, Marcelo Mendoza, Stuart E. Middleton, Véronique Moriceau, Malvina Nissim, Nicole Novielli, Symeon Papadopoulos, Barbara Poblete, Paolo Rosso, Giancarlo Ruffo, Manuela Sanguinetti, Michael Sirivianos.

REFERENCES

- [1] F. Benamara, M. Taboada, and Y. Y. Mathieu, “Evaluative language beyond bags of words: Linguistic insights and computational applications,” *Computational Linguistics*, vol. 43, no. 1, pp. 201–264, 2017.
- [2] C. Bosco, V. Patti, and A. Bolioli, “Developing corpora for sentiment analysis: The case of irony and senti-tut,” *IEEE Intelligent Systems*, vol. 28, no. 2, pp. 55–63, 2013.

- [3] C. Bosco, V. Patti, M. Bogetti, M. Conoscenti, and G. Ruffo, "Tools and resources for detecting hate and prejudice against immigrants in social media," in *Proc. of AISB Annual Convention*, 2017, pp. 79–84.
- [4] M. Sanguinetti, F. Poletto, C. Bosco, V. Patti, and M. Stranisci, "An italian Twitter corpus of hate speech against immigrants," in *Proc. of the 11th Int. Conf. on Language Resources and Evaluation (LREC 2018)*. Miyazaki, Japan: ELRA, May 7-12 2018.
- [5] M. Anzovino, E. Fersini, and P. Rosso, "Automatic identification and classification of misogynistic language on twitter," in *Proc. of the Int. Conf. on Applications of Natural Language to Information Systems (NLDB)*, 2018, pp. 57–64.
- [6] J. Karoui, F. Benamara, V. Moriceau, V. Patti, C. Bosco, and N. Aussenac-Gilles, "Exploring the impact of pragmatic phenomena on irony detection in tweets: A multilingual corpus study," in *Proc. of the 15th Conf. of the European Chapter of the Association for Computational Linguistics*, 2017, pp. 262–272.
- [7] D. Nozza, E. Fersini, and E. Messina, "A multi-view sentiment corpus," in *Proc. of the 15th Conf. of the European Chapter of the Association for Computational Linguistics*, vol. 1, 2017, pp. 273–280.
- [8] E. Fersini, F. A. Pozzi, and E. Messina, "Detecting irony and sarcasm in microblogs: The role of expressive signals and ensemble classifiers," in *Proc. of the IEEE Int. Conf. on Data Science and Advanced Analytics (DSAA)*, 2015, pp. 1–8.
- [9] E. Sulis, D. I. H. Fariás, P. Rosso, V. Patti, and G. Ruffo, "Figurative messages and affect in Twitter: Differences between# irony,# sarcasm and# not," *Knowledge-Based Systems*, vol. 108, pp. 132–143, 2016.
- [10] F. Bertola and V. Patti, "Ontology-based affective models to organize artworks in the social semantic web," *Inf. Process. Manage.*, vol. 52, no. 1, pp. 139–162, 2016.
- [11] B. Carminati, E. Ferrari, and M. Viviani, *A Multi-dimensional and Event-Based Model for Trust Computation in the Social Web*. Springer Berlin Heidelberg, 2012, pp. 323–336.
- [12] ———, "Security and trust in online social networks," *Synth. Lect. on Information Security, Privacy, & Trust*, vol. 4, no. 3, pp. 1–120, 2013.
- [13] M. Viviani and G. Pasi, "Credibility in social media: opinions, news, and health information—a survey," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 7, no. 5, p. e1209, 2017.
- [14] V. Patti, R. Damiano, and C. Bosco, "Ethical implications of analyzing opinions, emotions and interactions in social media," in *Seventh International Conference on Affective Computing and Intelligent Interaction Workshops and Demos, AII Workshops 2017, San Antonio, TX, USA, October 23-26, 2017*. IEEE Computer Society, 2017, pp. 153–158.
- [15] M. Viviani and G. Pasi, "Quantifier guided aggregation for the veracity assessment of online reviews," *International Journal of Intelligent Systems*, vol. 32, no. 5, pp. 481–501, 2017.