



Introduction to the special issue on social data analytics in medicine and healthcare

Alejandro Rodríguez-González¹ · Athena Vakali² · Miguel A. Mayer³ · Takashi Okumura⁴ · Ernestina Menasalvas-Ruiz¹ · Myra Spiliopoulou⁵

Published online: 26 October 2019
© Springer Nature Switzerland AG 2019

1 Introduction

The growing availability and accessibility of diverse and relevant health-related data resources, and the rapid proliferation of technological developments in data analytics are contributing to make the most of extracting the power of these datasets. This allows us to improve diagnosis and decision making, shorten the development of new drugs from discovery to marketing approval, facilitate early outbreak detection, improve healthcare professionals training and reduce costs to name but a few examples.

Extracting the knowledge to make this a reality is still a daunting task: on the one hand, data sources are not integrated; they contain private information and are not structured. On the other hand, we still lack context- and privacy-aware algorithms to extract the knowledge after a proper curation and enrichment of the datasets.

In recent years technology has made it possible not only to get data from many healthcare settings (hospitals, primary care settings, laboratories, etc.), it also allows information to be obtained from the society itself (sensors, Internet of Things (IoT) devices, social networks, etc.). For instance, social media platforms are a new source of data coming from all the population groups.

For this reason, the organization of the current special issue responds to the necessity in collecting the last efforts that have been made in these areas of research. In 2018, a Special Track on Social and Medical Data Analytics was organized as part of the 31st IEEE CBMS International Symposium on Computer-Based Medical Systems (CBMS 2018) in Karlstad, Sweden.

After the special track, we invited some of the authors that presented their papers to submit an extended version of their work to this special issue in International Journal of Data Science and Analytics. The reviewers performed accurate reviews considering, in those papers that were originally sent to the special track, the extensions made. From the total number of papers received, a total of five papers were selected to be included in this special issue.

We are pleased to introduce this collection of papers. We want to thank to all the reviewers who helped with the review of the received papers. In the following section, a brief description of the selected papers for this special issue will be shown.

✉ Alejandro Rodríguez-González
alejandro.rg@upm.es
Athena Vakali
avakali@csd.auth.gr
Miguel A. Mayer
miguelangel.mayer@upf.edu
Takashi Okumura
taka@wide.ad.jp
Ernestina Menasalvas-Ruiz
ernestina.menasalvas@upm.es
Myra Spiliopoulou
myra@ovgu.de

¹ Universidad Politécnica de Madrid, Madrid, Spain
² Aristotle University of Thessaloniki, Thessaloníki, Greece
³ Universitat Pompeu Fabra, Barcelona, Spain
⁴ Kitami Institute of Technology, Kitami, Japan
⁵ Otto-von-Guericke-Universität Magdeburg, Magdeburg, Germany

2 The five papers selected for this special issue

In the first selected paper [1], the authors focused their study on the use of mobile technologies to solve the deficiency associated to the use of interviews or self-report questionnaires with patients in reference to evaluate symptoms of neuropsychiatric disorders such as tinnitus. The authors presented the results obtained by using TrackYourTinnitus

mobile crowdsensing platform. Concretely, they focused on showing the existing discrepancies between prospective and retrospective assessments, showing that “the prospective variation of tinnitus loudness does not differ between the users who retrospectively rate tinnitus loudness as “varying” and the ones who retrospectively rate it as “non-varying.””.

The second accepted paper [2] presented a study with the goal of defining an auxiliary diagnostic method for obstructive sleep apnea that can support the decision to support polysomnography. The paper was based on the use of a specific set of risk and diagnostic factors. The data was retrieved after performing polysomnography between January and May 2015 to a total of 194 patients that fulfilled the inclusion criteria, being 63% male. The study was focused on the development of two bayesian classifiers to build the models (Naïve Bayes and Tree Augmented Naïve Bayes) through the use of 38 variables identified by literature review. The results confirm the necessity of a pre-evaluation prior to polysomnography.

In the third accepted paper [3], the authors focused their efforts on analyzing the the sentiments expressed in online medical forums that discuss about Lyme disease. As part of their goals, the authors aimed to identify the categories that can characterize Lyme disease discourse and investigate the strategies for the automatic classification of the medical forum posts into those categories.

The fourth accepted paper [4] focused on the association between sleep and exercise patterns. While recent studies analyzed these patterns in the data captured by digital devices, the authors focused their work on conducting a large-scale data study of exercise and sleep with an analysis of 8 months for 20K users, combining search query logs, location information and aggregated social media data. The authors analyzed the factors correlated with better sleep

and more effective exercise, confirming these relationships through casual inference analysis.

Finally, the fifth accepted paper extended [5] their preliminary work, focused on hospital clustering analysis to study hospital monthly admission behavior for different diseases. In this work, they included other aspects of disease data and the combination of different views of disease data. The approach tackled clustering complex networks with the goal of determine the effect of disease networks on characterizing the underlying clustering structure of disease-specific hospital networks.

References

1. Pryss, R., Probst, T., Schlee, W., Schobel, J., Langguth, B., Neff, P., Spiliopoulou, M., Reichert, M.: Prospective crowdsensing versus retrospective ratings of tinnitus variability and tinnitus-stress associations based on the TrackYourTinnitus mobile platform. *Int J Data Sci Anal* (2018). <https://doi.org/10.1007/s41060-018-0111-4>
2. Ferreira-Santos, D., Rodrigues, P.P.: A clinical risk matrix for obstructive sleep apnea using Bayesian network approaches. *Int J Data Sci Anal* (2018). <https://doi.org/10.1007/s41060-018-0118-x>
3. Alnashwan, R., Sorensen, H., O’Riordan, A., Hoare, C.: Accurate classification of socially generated medical discourse. *Int J Data Sci Anal* (2018). <https://doi.org/10.1007/s41060-018-0128-8>
4. Farajtabar, M., Kiciman, E., Nathan, G., White, R.W.: Modeling behaviors and lifestyle with online and social data for predicting and analyzing sleep and exercise quality. *Int J Data Sci Anal* (2018). <https://doi.org/10.1007/s41060-018-0136-8>
5. Albarakati, N., Obradovic, Z.: Multi-domain and multi-view networks model for clustering hospital admissions from the emergency department. *Int J Data Sci Anal* (2018). <https://doi.org/10.1007/s41060-018-0147-5>

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.