

The 6th International Workshop on Mining Actionable Insights from Social Networks (MAISoN 2021)

Special Edition on Healthcare Social Analytics

Ebrahim Bagheri¹, Diana Inkpen², Christopher C. Yang³, Fattane Zarrinkalam⁴

¹Ryerson University

²University of Ottawa

³Drexel University

⁴Thomson Reuters Labs

bagheri@ryerson.ca, diana.inkpen@uottawa.ca, chris.yang@drexel.edu, fattane.zarrinkalam@thomsonreuters.com

Abstract

The sixth edition of the workshop on Mining Actionable Insights from Social Networks (MAISoN 2021) took place virtually on June 7th, 2021, co-located with 15th International AAAI Conference on Web and Social Media (ICWSM 2021). This year, we organized a special edition with focus on healthcare social analytics. The aim of this edition was to bring together researchers from different disciplines to discuss research that goes beyond descriptive analysis of social media data and instead investigate different techniques that use social media data for building diagnostic, predictive and prescriptive analysis models for health applications. This topic attracted a lot of interest from the community especially because of all the considerations surrounding the impact of social media during the COVID-19 pandemic which has impacted on people's physical, mental and social health issues.

Introduction

With the emergence and growing popularity of social media such as blogging systems, wikis, social bookmarking, social networks and microblogging services, many users are extensively engaged in at least some of these applications to express their feelings and views about a wide variety of social topics as they happen in real time by commenting, tagging, joining, sharing, liking, and publishing posts. According to Statista, there were an estimated 2.65 billion people using social media in 2018, a number projected to increase to almost 3.1 billion in 2021¹. This has resulted in an ocean of data which presents an interesting opportunity for performing data mining and knowledge discovery in many domains including healthcare. The recent highly impressive advances in machine learning and natural language processing present exciting opportunities for developing automatic methods for the collection, extraction, representation, analysis, and validation of social media data for health applications. These methods should be able to simultaneously address the unique challenges of processing social media data and timely discover meaningful patterns identifying emerging health threats.

Copyright © 2021, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

¹<https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/>

Traditional research on healthcare social analytics mainly focuses on descriptive methods such as tracking health trends on social media and tracking infectious disease spread. The main distinguishing focus of this workshop was the use of social media data for building diagnostic, predictive and prescriptive analysis models for health research and applications such as analysing how social media can impact on people's physical, mental and social health issue, and predicting users' health status and recommending solutions to prevent the risk of committing unfortunate actions such as suicide.

The topics of interest of this workshop included but were not limited to the following:

- Social media mining for automatic health monitoring and surveillance.
- Predicting user's health status on social media.
- User behavior analysis and susceptibility prediction with regard to health-related data on social media.
- Predictive models for early detection of trends in health-related issues on Social Media.
- Early detection of disease outbreaks
- Explainable AI for healthcare social media analytics.
- Ethics, bias, and fairness in analysing social media for healthcare applications.
- Analysing health-related misinformation on social media.
- Prescriptive countermeasure methods against formation and circulation of health-related misinformation.
- New datasets and evaluation methodologies to help healthcare social analytics.

Keynotes

The workshop had three invited keynote speakers who shared their perspectives on social media mining for health research and applications.

The first keynote speaker, Dr. Stevie Chancellor, is an Assistant Professor in the Department of Computer Science & Engineering at the University of Minnesota. She presented her talk on "Human-Centered Machine Learning

for Dangerous Mental Health Behaviors Online” to highlight the urgent need to innovate data-driven systems in order to handle the volume and risk of physically dangerous health behaviors discussed on social media and their propagation to others in the community. She pointed out that traditional approaches to prediction have mixed success, because technical solutions oversimplify complex behavior and the unique interactions of communities with both individuals and platforms. Her talk focused on recent research on human-centered machine learning as a lens to make these predictions more ethical and compassionate, as well as technically rigorous. She also talked about her work in Machine Learning for dangerous mental illness behaviors in online communities, like opioid abuse, suicidal ideation, and promoting eating disorders. Finally, she discussed some alarming gaps in data science pipelines for generating labels for training data.

Our second keynote speaker, Dr. Abeed Sarker, is an Assistant Professor at the Department of Biomedical Informatics, School of Medicine, Emory University. In his talk entitled “Obtaining Insights about non-medical use of Prescription Medications from Social Media via Natural Language Processing”, he presented his research, funded by the National Institute on Drug Abuse (NIDA) towards developing the computational infrastructure required for effectively utilize social media data for factors contributing to Prescription medication abuse. Their infrastructure involves manual annotation, supervised classification, information extraction, and combining them with social media meta-data such as geolocation. He presented their various data characterization methods and their validations against traditional sources of data such as the CDC Wonder database and the National Survey on Drug Use and Health.

The final keynote speaker of our workshop, Dr. Jia Xue, is an Assistant Professor at the Factor-Inwentash Faculty of Social Work cross-appointed with the Faculty of Information. In her talk, she discussed the use of Twitter data in sexual assault and family violence research, limitations, and future directions in using social media data in this field. The key point of her talk entitled “Twitter as a tool for understanding Sexual Assault and Family Violence” was to raise awareness of family violence as a significant public health concern due to the COVID-10 pandemic. She also emphasized to the campaigns of “signal for help” and “safe word” launched to facilitate resources of help for family violence survivors. Further, she pointed out that, in the context of telemedicine, useful, validated, and real-time online resources are pivotal to support individuals and families experiencing family violence as well as to strengthen the communication between survivors and non-profit agencies serving family violence victims.

Panel

Our workshop also hosted a panel discussion on “Privacy and Ethical Considerations of Mining Health Data from Social Media” with 5 panelists, namely Dr. Khaled El Emam from University of Ottawa, Dr. Farzaneh Etminani from Halmstad University, Dr. Tristan Henderson from University

of St. Andrews, Dr. Kirsten Ostherr from Rice University, and Dr. Reihaneh Rabbany from McGill University.

The panelists shared their perspectives mainly on two themes:

The first theme was the *ethical issues involved in deriving data from media interfaces that manage the health risks of populations and the financial risks of corporate stakeholders providing services to those populations*. For this topic, they discussed how technologies of quantification and metaclinical risk media contribute to a shift in medicine from individual, reactive care to population-based, preventive care, and how individual users’ health data can be inadvertently used against them as a result. They also discussed how mining data from health apps and health-relevant social media may contribute to redefining “the human” in quantitative terms that elide fundamental aspects of the experience of health and illness.

As the second theme, the panelists talked about “eXplainable AI(XAI) and its importance in healthcare”. Some of panelists argued that AI is not mature enough in clinical practice, but it is expected to be one of the central areas contributing to the next generation of healthcare. Thus, one of the challenges that has hampered pervasive AI in healthcare system is inscrutability. They emphasized that, for a healthcare system to become information-driven, the conclusions and advice derived from data must be explainable and interpretability is considered to be a keystone in this transition and it is crucial for future of personalised medicine specifically.

Contributions

All paper submissions were reviewed by at least two members of the program committee, which consisted of experts from academia and industry alike. In the reviewing process, we ensured that we maintain high quality standards, while at the same time making sure that less-mature yet interesting work was also given a chance to be present and discussed at the workshop. Two full papers and one abstract were accepted based on the quality of the rigor of analysis, results and presentation, and we provide a brief description for each contribution below:

Symptom Extraction from the Narratives of Personal Experiences with COVID-19 on Reddit. The authors (Murray et al. 2021) have quantified the change in discourse throughout a collective timeline of experiences being COVID-19 positive. They have collected discourse in the form of Reddit posts from COVID19Positive (4,610 posts from March 14th, 2020 to May 12th, 2020), and filtered the data to capture stories from people who tested positive for COVID-19. They have also exploited the temporal structure exhibited in these diarised posts to obtain a collective timeline of experiences. Using topic modelling and sentiment analysis, they have quantified the change in discussion of COVID-19 throughout individuals’ experiences for the first 14 days since symptom onset. Their results provided a perspective on the patient experience of COVID-19 that complements other medical data streams and can potentially reveal when mental health issues might appear.

Lonely Road: Speculative Challenges for a Social Media Robot Aimed to Reduce Driver Loneliness. The aim of this conceptual paper (Valle et al. 2021) was to navigate through some “memories” of one possible future, toward stimulating ideation and discussion within the increasingly vital area of safety in smart cities. The authors have explored one idea for how continuous care could be provided to improve drivers’ mental states; in particular, the idea of a “robot” that could positively affect a driver’s health through interactions supported by social media mining on Facebook. They have used a speculative design approach to present some potential challenges and solutions in regard to a robot’s interaction strategy, user modeling, and ethics. For example, to address how to generate appropriate robot activities and mitigate the risk of damage to the driver, they have proposed a hybrid neuro-symbolic recognition strategy leveraging stereotypical and self-disclosed information is described.

Obtaining Insights about Nonmedical use of Prescription Medications from Social Media via Natural Language Processing. In this extended abstract, the author (Sarker 2021) presented his ongoing efforts to leverage social media data, via natural language processing and machine learning methods, for obtaining insights about the nonmedical use of prescription medications.

Workshop Organization

This workshop was organized by: Ebrahim Bagheri, an Associate Professor in the Department of Electrical, Computer and Biomedical Engineering at Ryerson University, Diana Inkpen, a Professor in the School of Electrical Engineering and Computer Science at the University of Ottawa, Christopher C. Yang, a Professor in the College of Computing and Informatics at Drexel University, and Fattane Zarrinkalam, a Research Scientist at Thomson Reuters Labs.

The papers of the workshop were published in the Workshop Proceedings of ICWSM 2021 and the following experts formed our technical program committee:

Steering Committee

- Marcelo G. Armentano (ISISTAN (CONICET-UNICEN))
- Frank W. Takes (University of Amsterdam)

Program Committee

- Anastasia Giachanou (Universitat Politècnica de València)
- Farzaneh Etminani (Halmstad University)
- Oana Inel (Delft University of Technology)
- Amir Karami (University of South Carolina)
- Ugur Kursuncu (University of South Carolina)
- Michael Merrill (University of Washington)
- Joao Palotti (Qatar Computing Research Institute)
- Ashish Sharma (University of Washington)
- Damiano Spina (RMIT University)
- Elad Yom-Tov (Microsoft)
- Galen Weld (University of Washington)

- Arkaitz Zubiaga (Queen Mary University of London)
- Guido Zuccon (University of Queensland)

References

- Murray, C., Mitchell, L., Tuke, S. and Mackay, M. 2021. Symptom Extraction from the Narratives of Personal Experiences with COVID-19 on Reddit. *MAISoN Workshop Proceedings of the 15th International AAAI Conference on Web and Social Media (ICWSM)*.
- Valle, F., Galozy, A., Ashfaq, A., Etminani, K., Vinel, A. and Cooney, M. 2021. Lonely road: speculative challenges for a social media robot aimed to reduce driver loneliness. *MAISoN Workshop Proceedings of the 15th International AAAI Conference on Web and Social Media (ICWSM)*.
- Sarker, A. 2021. Obtaining insights about nonmedical use of prescription medications from social media via natural language processing. *MAISoN Workshop Proceedings of the 15th International AAAI Conference on Web and Social Media (ICWSM)*.