The Battlefront of Combating Misinformation and Coping with Media Bias

Yi R. Fung  
University of Illinois Urbana-Champaign  
USA  
yifung2@illinois.edu

Kung-Hsiang Huang  
University of Illinois Urbana-Champaign  
USA  
khhuang3@illinois.edu

Preslav Nakov  
Mohamed bin Zayed University of Artificial Intelligence  
UAE  
preslav.nakov@mbzuai.ac.ae

Heng Ji  
University of Illinois Urbana-Champaign  
USA  
hengji@illinois.edu

ABSTRACT
Misinformation is a pressing issue in modern society. It arouses a mixture of anger, distrust, confusion, and anxiety that cause damage on our daily life judgments and public policy decisions. While recent studies have explored various fake news detection and media bias detection techniques in attempts to tackle the problem, there remain many ongoing challenges yet to be addressed, as can be witnessed from the plethora of untrue and harmful content present during the COVID-19 pandemic, which gave rise to the first social-media infodemic, and the international crises of late. In this tutorial, we provide researchers and practitioners with a systematic overview of the frontier in fighting misinformation. Specifically, we dive into the important research questions of how to (i) develop a robust fake news detection system that not only fact-checks information pieces provable by background knowledge, but also reason about the consistency and the reliability of subtle details about emerging events; (ii) uncover the bias and the agenda of news sources to better characterize misinformation; as well as (iii) correct false information and mitigate news biases, while allowing diverse opinions to be expressed. Participants will learn about recent trends, representative deep neural network language and multimedia models, ready-to-use resources, remaining challenges, future research directions, and exciting opportunities to help make the world a better place, with safer and more harmonic information sharing.

ACM Reference Format:

1 PRESENTER BIOGRAPHIES

Yi R. Fung is a third-year CS Ph.D. candidate at the University of Illinois Urbana-Champaign (UIUC), with research interests in NLP, multimedia reasoning, and computation for social good. Her recent research includes fake news detection and cross-cultural understanding. Yi is a recipient of the NAACL ’21 Best Demo Paper and a UIUC Lauslen and Andrew fellowship. She gave an invited talk at the Harvard Medical School Bioinformatics Seminar.

Kung-Hsiang Huang is a second-year Ph.D. student at the Computer Science Department of UIUC. His research focuses on fact-checking and fake news detection. Prior to joining UIUC, he obtained his B.Eng. in Computer Science from the Hong Kong University of Science and Technology, and his M.Sc. in Computer Science from USC. He is also a co-founder of an AI startup, Rosetta.ai.

Preslav Nakov is a Professor at Mohamed bin Zayed University of Artificial Intelligence. He is President of ACL SIGLEX, Secretary of ACL SIGSLAV, advisory board member of EACL, and editorial board member of Computational Linguistics, TACL, CS&L, IEEE TAC, NLE, AI Communications, and Frontiers in AI. His research on fake news has been featured by over 100 news outlets, including Forbes, Boston Globe, Aljazeera, MIT Technology Review, The Register, WIRED, and Engadget, among others.

Heng Ji is a Professor at the Computer Science Department of the University of Illinois Urbana-Champaign, and an Amazon Scholar. Her research interests focus on NLP, especially Multimedia Multilingual Information Extraction, Knowledge Base Population and Knowledge-driven Generation. She was selected as “Young Scientist” and a member of the Global Future Council on the Future of Computing by the World Economic Forum. She received “AI’s 10 to Watch” Award, NSF CAREER award, Google Research Award, IBM Watson Faculty Award, Bosch Research Award, Amazon AWS Award, ACL2020 and NAACL2021 Best Demo Paper Award.

2 TUTORIAL OUTLINE

2.1 Background and Motivation [20min]
We begin with a selection of real-world examples of fake news and their harmful impact on society, followed by a pedagogical exercise of how humans tend to approach the problem of fake news detection, characterization, and correction. We point out the conceptual distinctions between different types of fake news.
2.2 Fake News Detection [60min]

We discuss detection based on stylistic [6, 19], fact-checking [13, 18, 22], semantic consistency [2, 14], and propagation patterns [1, 26], and their advantages and limitations [21, 27]. We then cover approaches [11, 16] that leverage a combination of these elements for greater representation power and robustness. Importantly, we discuss work that explores the diachronic bias of fake news detection and portability across datasets in different time and language settings [12, 17]. Moreover, we review generating fake news that better aligns with the key topic and the facts [11, 25], e.g., by applying conditioned fake news generation to construct silver-standard data annotations for finer-grained fake news detection [11].

2.3 Fake News Characterization [30min]

We next address some fundamental questions of characterization based on underlying source biases, reporting agenda, propaganda techniques, and target audience [5]. We introduce approaches to detect political and socio-cultural biases in news articles [3, 9, 10]. Next, we discuss the recent EMU benchmark that requires models to answer open-ended questions capturing the intent and the implications of a media edit [7]. We cover methodologies for identifying specific propaganda techniques, e.g., loaded language, appeal to fear, smears, glorifying generalities, whataboutism, etc. [8]. We also discuss recent exploration in predicting the intended target of harmful memes, e.g., person, organization, community, or society [20, 23].

2.4 Corrective Actions [30min]

Finally, we cover research on explaining why a given piece of news is fake through the leverage of reader comments [24], strategies for placing corrective explanations, based on user studies [4], and mitigating media bias through neutral article generation [15].

ACKNOWLEDGEMENTS

This work is supported in part by U.S. DARPA SemaFor Program No. HR001120C0123, AIDA Program No. FA8750-18-2-0014, and KAIROS Program No. FA8750-19-2-1004. The views and the conclusions contained herein are those of the authors and should not be interpreted as representing the official policies of DARPA or the U.S. Government. The U.S. Government is authorized to reproduce and to distribute reprints for governmental purposes notwithstanding any copyright annotation therein.

REFERENCES