DISSERTATION DEFENSE

Ashkan Kazemi

Human-Centered Natural Language Processing for Countering Misinformation

Thursday, November 16, 2023
11:00am – 1:00pm
3725 Beyster
Hybrid – Zoom
Passcode: OhP5L8Zw

ABSTRACT: As curbing the spread of online misinformation has proven to be challenging, we look to artificial intelligence and natural language technology for helping individuals and society counter and limit it. Despite current advances, state-of-the-art natural language processing (NLP) and AI still struggle to automatically identify and understand misinformation. Humans exposed to harmful content may experience lasting negative consequences in real life, and it is often difficult to change one's mind once they form wrong beliefs. Addressing these interwoven technical and social challenges requires research and understanding into the core mechanisms that drive the phenomena of misinformation.

This thesis introduces human-centered NLP tasks and methods that can help prioritize human welfare in countering misinformation. We present findings on the differences in how people of different backgrounds perceive misinformation, and how misinformation unfolds in different conditions such as end-to-end encrypted social media in India. We build on this understanding to create models and datasets for identifying misinformation at scale that put humans in the decision making seat, through claim matching, matching claims with fact-check reports, and query rewriting that scale the efforts of fact-checkers. Our work highlights the global impact of misinformation, and contributes to advancing the equitability of available language technologies through models and datasets in a variety of high and low resources and languages.

We also make fundamental contributions to data, algorithms, and models through: multilingual and low-resource embeddings and retrieval for better claim matching, reinforcement learning for reformulating queries for better misinformation discovery, unsupervised and graph-based focused content extraction through introducing the Biased TextRank algorithm, and explanation generation through extractive (Biased TextRank) and abstractive (GPT-2) summarization.

Through this thesis, we aim to promote individual and social wellbeing by creating language technologies built on a deeper understanding of misinformation, and provide tools to help journalists as well as internet users to identify and navigate around it.

CHAIR: Prof. Rada Mihalcea and Asst. Research Scientist Veronica Perez-Rosas