

Researchers find flaws in using source reputation for training automatic misinformation detection algorithms

May 16 2023, by Carol Peters



Credit: Pixabay/CC0 Public Domain

Researchers at Rutgers University have found a major flaw in the way that algorithms designed to detect "fake news" evaluate the credibility of online news stories.

Most of these algorithms rely on a credibility score for the "source" of the article, rather than assessing the credibility of each individual article, the researchers said.

"It is not the case that all [news articles](#) published by sources labeled 'credible' (e.g., The New York Times) are accurate, nor is it the case that every article published by sources labeled 'non-credible' publications are ['fake news'](#)," said Vivek K. Singh, an associate professor at the Rutgers School of Communication and Information and co-author of the study "Misinformation Detection Algorithms and Fairness Across Political Ideologies: The Impact of Article Level Labeling," published on OSFHome.

"Our analysis shows that labeling articles for [misinformation](#) based on the source is as bad an idea as just flipping a coin and assigning true/false labels to [news stories](#)," added Lauren Feldman, an associate professor of journalism and media studies at the School of Communication and Information, who is another co-author of the paper.

The researchers found using source-level labels for credibility isn't a [reliable method](#), with article-level labels matching 51% of the time. This labeling process has important implications for tasks such as the creation of robust fake news detectors and for audits on fairness across the political spectrum.

To address this problem, the study offers a [new dataset](#) of journalistic quality individually labeled articles and an approach for misinformation detection and fairness audits. The findings of this study highlight the need for more nuanced and reliable methods of detecting misinformation in online news and provide valuable resources for future research in this area.

Researchers assessed the credibility and political leaning of 1,000 news articles and used these article-level labels to build misinformation detection algorithms. Then, they evaluated how the labeling methodology (source level versus article level) impacts the performance of misinformation detection algorithms.

Their aim was to explore the impact of article-level labeling on the process and determine whether the bias that exists when applying machine-learning approach at the source level still exists when applying the same machine-learning approach to individual articles, and in addition, to learn if bias is reduced when dealing with individually labeled articles.

The authors presented their paper at the 15th [Association for Computing Machinery](#) Web Science Conference 2023, held from April 30-May 1 in Austin, Texas.

A joint effort between journalism, [information science](#) and computer science professionals, the authors, in addition to Singh and Feldman, include Jinkyung Park, a Ph.D. alumna of the School of Communication and Information; Rahul Dev Ellezhuthil, a computer science master's degree student; School of Communication and Information doctoral student Joseph Isaac; and Christoph Mergerson, a Ph.D. alumnus of the School of Communication and Information and an assistant professor of race and media at the University of Maryland.

The authors said algorithms used to detect misinformation in online articles function the way they do "mainly because there is a dearth of fine-grained labels defined at the news article level. We acknowledge that labeling each news article may not be feasible given the massive volume of news articles that are published and disseminated on the web. At the same time, there are reasons to question the validity of datasets labeled at the source level."

"Validating online news and preventing the spread of misinformation is critical for ensuring trustworthy online environments and protecting democracy," the authors wrote, adding that their work "aims to increase public confidence in misinformation detection practices and subsequent corrections by ensuring the validity and fairness of results," and their

dataset and the conceptual results "aim to pave the way for more reliable and fair misinformation detection algorithms."

More information: Jinkyung Park et al, Misinformation Detection Algorithms and Fairness across Political Ideologies: The Impact of Article Level Labeling, [DOI: 10.17605/OSF.IO/QWNSE](https://doi.org/10.17605/OSF.IO/QWNSE)

Provided by Rutgers University

Citation: Researchers find flaws in using source reputation for training automatic misinformation detection algorithms (2023, May 16) retrieved 11 December 2023 from <https://techxplore.com/news/2023-05-flaws-source-reputation-automatic-misinformation.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.