A News Verification Browser for the Detection of Clickbait, Satire, and Falsified News

Victoria Rubin
Western University, Faculty of Information and Media Studies, vrubin@uwo.ca

Chris Brogly
Western University, Faculty of Information and Media Studies, cbrogly@ivey.uwo.ca

Nadia Conroy
Western University, Faculty of Information and Media Studies, nadia.k.conroy@gmail.com

Yimin Chen
Western University, Faculty of Information and Media Studies

Sarah E. Cornwell
Western University, Faculty of Information and Media Studies

See next page for additional authors

Follow this and additional works at: https://ir.lib.uwo.ca/fimspub
Part of the Library and Information Science Commons, and the Science and Technology Studies Commons

Citation of this paper:
Rubin, Victoria; Brogly, Chris; Conroy, Nadia; Chen, Yimin; Cornwell, Sarah E.; and Asubiaro, Toluwase V., "A News Verification Browser for the Detection of Clickbait, Satire, and Falsified News" (2019). FIMS Publications. 259.
https://ir.lib.uwo.ca/fimspub/259
Authors
Victoria Rubin, Chris Brogly, Nadia Conroy, Yimin Chen, Sarah E. Cornwell, and Toluwase V. Asubiaro

This article is available at Scholarship@Western: https://ir.lib.uwo.ca/fimspub/259
A News Verification Browser for the Detection of Clickbait, Satire, and Falsified News

Victoria L. Rubin¹, Chris Brogly¹, Nadia Conroy¹,³, Yimin Chen¹, Sarah E. Cornwell¹, and Toluwe V. Asubiaro¹,²

¹Language and Information Technology Research Lab (LiT.RL), Faculty of Information and Media Studies, The University of Western Ontario, London, Ontario, Canada. ²E. Latunde Odeku Medical Library, College of Medicine, University of Ibadan, Ibadan, Nigeria ³Media Sonar Technologies, London, Ontario, Canada.

DOI: 10.21105/joss.01208

Software
- Review
- Repository
- Archive

Submitted: 12 December 2018
Published: 09 March 2019

License
Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License (CC-BY).

Description-in-Brief

The LiT.RL News Verification Browser is a research tool for news readers, journalists, editors or information professionals. The tool analyzes the language used in digital news web pages to determine if they are clickbait, satirical news, or falsified news, and visualizes the results by highlighting content in color-coded categories. Although the clickbait, satire, and falsification detectors perform to certain accuracy levels on test data, during real-world internet use accuracy may vary. The browser is not a replacement for digital literacy and is not always correct. All processing is completed on the local machine - results are not sent to or from a remote server. Results may be saved locally to a standard SQLite database for further analysis.

Summary

Widespread adoption of internet technologies has changed the way that news is created and consumed. The current online news environment is one that incentivizes speed and spectacle in reporting at the cost of fact-checking and verification, encouraging the proliferation of misinformation and disinformation. The LiT.RL News Verification (NV) Browser is a system that offers a first step counter-measure by automatically detecting and highlighting clickbait (to 94% accuracy on a test set of 5670 texts), satirical fake news (to 84% accuracy on a test set of 95 texts), and fabricated news (to 71% accuracy on a test set of 28 texts). The browser was built to study the effectiveness of these deception detectors when applied to online news websites, where the accuracy of these detectors may vary considerably given the variety of styles, formats and content of online news. The system is presented to users as a set of assistive technologies built into a stand-alone browser tuned to identify different varieties of fakes (Rubin, 2017; Rubin, Chen, & Conroy, 2015). Our algorithms look for patterns of subtle lexico-syntactic features in text. Images, audio and video formats are unsupported. The core of the functionality is in the natural language processing (NLP) of textual data and automated classification of results with machine learning using support vector machines.


Each is implemented as a separate overlay option through three tabs at the bottom right of the browser. Currently in its proof of concept stage, the system runs over a batch of website content, such as news feed on a website, and visualizes the results by highlighting content in red, orange or green, by analogy with the traffic stop-light.

Limitations

This browser is meant to augment our human discernment, rather than replace it, by highlighting potentially false information which may require further scrutiny. Digital literacy is key for everyone to effectively evaluate potential misinformation online, and the LiT.RL Browser is NOT a replacement for that (Chen, Conroy, & Rubin, 2015b). During real-world internet use, the accuracy of the detectors in this tool may vary. News readers’ critical thinking remains key to navigating the increasingly fraught online information landscape. The News Verification Browser is a research tool - not a replacement for a day-to-day web browser. It should not be used where security is critical.

Availability

The source code is openly available on GitHub (Rubin et al., 2018) under the GPLv3 license for anyone in the research and development community to use or improve on. The public can download the browser for experimentation on their own computers with “no strings attached”.

Acknowledgements


References


