Analysis of News Headlines by Artificial Intelligence Systems

A Master’s Degree Project Proposal

Chinmay Pai

Interactive Media and Game Development

Worcester Polytechnic Institute

Worcester, MA

Committee

Advisor - Brian Moriarty

Reader –

Reader -

2018

# Abstract

This document intends to propose the analysis of news headlines using the sentimental analysis modules of artificial intelligence systems such as Watson, Google Cloud and Amazon Comprehend. This is part of the creation of the design of a serious game which aims to utilize this research to demonstrate the spread and power of fake news. The game will utilize research carried out in the field of the spread of fake news, as well as inform players as to its spread and means to control and arrest its spread.

# Introduction

The spread of fake news has been a highly topical subject today, which has had a significant impact on the events in society today all over the world. This outreach of propaganda, whether motivated by state actors or other groups, has resulted in highly significant world events in the West over the last three years, such as the vote for Great Britain to leave the European Union, the election and popularity of United States President Donald Trump, and the rise of highly autocratic politics. However, it is hardly a new phenomenon unique to the last few years, especially outside the West in more developing countries with powerful groups outside establishment sources and a perception of being far more reliable than the ‘compromised’ mainstream.

While a good portion of the epidemic has been attributed, especially in the United States, to actively motivated work from state actors as a form of information warfare in order to influence government policy, it is undeniable that a significant portion of propaganda spread is due to people agreeing with the implications, if not necessarily the exact false fact, that is being spread. Having seen several friends and acquaintances believe falsehoods due to the news playing into culturally established narratives, I am not sure that the

I am interested in expressing research into this field into video game form, and therefore would like to create a digital game, potentially in HTML5 in order to better integrate the Node elements of importing the real-life headlines from a variety of real-life news outlets with differing levels of bias. I believe and hope that players experiencing the game will be able to see how fake news spreads across all media and how it may be inhibited.

# Prior Research

Most of the work in the field of the analysis of fake news has been either technical – understanding the spread of fake news through new mass media such as Twitter – or sociological – understanding the nature of fake news and how it influences the people it targets to achieve its ideological aims. Sobieraj and Berry have spoken about how the spread of outrage across all forms of media in the United States, such as television, talk radio, newspaper editorial columns, and political blogs, has been a significant measure of the success of these media in the form of viewership and clicks, which seems to correlate with the generally accepted idea that ‘outrage sells’, and have also measured so-called incivility from both sides of the political spectrum, concluding that while both the left and the right use similar tactics, the right have been shown to use it in greater quantities. Langin has found that fake news spread is not, contrary to popular belief, primarily spread by bots, but by actual humans, whether for malicious or ignorant reasons.

There have also been several ideas regarding the methods by which fake news might be contained. These include the idea of “guardians”, proposed by Vo and Lee, users who can recommend verified facts to users in response to popular misconceptions/fake news about popular figures. Websites such as Politifact and Snopes have taken up this role with mixed success, with allegations of bias from both sides of the political spectrum. Baum also speaks about the possibility of weeding out fake news using algorithmic methods, which utilize bots to either correct or remove fake news, which would require collaboration between sociological academia and computer science experts, as well as careful balancing to ensure that governments do not subvert these systems to tyrannical ends.

# Research Design

The game under discussion, tentatively titled “The Foghorn”, has been designed as a puzzle game where the player controls the head of a media organization whose aim is to rise from a small-town newspaper to a media empire spanning the domains of radio, television, and new media such as the internet using the power of fake news and information warfare to spread their ideology. The

*Explain the basic idea behind the game you want to build, and why you need to conduct experimental research to in order to build it.*

*Then explain the data your research will collect, how it will be collected.*

# Background Preparation

In preparation for the start of the thesis, I have carried out preliminary research in order to display early data results for several artificial intelligence systems which can be used to create guesses for the results that are obtained when analyzing headlines from several sources, such as Google Cloud, IBM Watson and Amazon Comprehend. The headlines were drawn from several news sources such as CNN, Fox News, The New York Times, The Washington Post, Mother Jones and Breitbart. The raw results are available in the appendix provided.

In most cases, the results of the analysis were surprising – even media sources considered highly partisan typically showed high levels of neutrality in the analysis of their headlines as per Amazon Comprehend, Google Cloud, and IBM Watson. Therefore, simple text analysis of headlines and news articles, while informative, is not sufficient to distinguish news sources in terms of “outrage value”, so to speak. Therefore, other methods to observe their reporting is

An observation of the websites of the news organizations, pictures of which are archived in the appendix provided, shows other potential factors which can be seen on their websites. The primary factors observed are positioning and frequency. Positioning certain headlines at the part of the screen which will immediately attract attention affects the

# Evaluation

Explain how your gathered data will be analyzed and evaluated, and (if useful results are produced) how it will be integrated into your proposed game.

# Timetable

# Works Cited

Sarah Sobieraj & Jeffrey M. Berry (2011) From Incivility to Outrage: Political Discourse in Blogs, Talk Radio, and Cable News, Political Communication, 28:1, 19-41, DOI: 10.1080/10584609.2010.542360

Baum, M., Greenhill, K., Metzger, M., Pennycook, G., Rothschild, D., Schudson, M., Sloman, S., et al. (n.d.). The science of fake news. Science, 359(6380), 1094–1096. doi:10.1126/science.aao2998

N. Vo, and K. Lee. The Rise of Guardians: Fact-checking URL Recommendation to Combat Fake News. SIGIR, July 2018

Rose-Stockwell, Tobias (2017) This is How Your Fear and Outrage Are Being Sold For Profit, Medium - https://medium.com/@tobiasrose/the-enemy-in-our-feeds-e86511488de

Langin, K. (n.d.). Fake news spreads faster than true news on Twitter—thanks to people, not bots. Science. doi:10.1126/science.aat5350

Robin, Alexander (2017) How I Scraped 7000+ Articles from a Newspaper Website Using Node, FreeCodeCamp - <https://medium.freecodecamp.org/how-i-scraped-7000-articles-from-a-newspaper-website-using-node-1309133a5070>

# Resources for Future Research

Jang, S., & Kim, J. (n.d.). Third person effects of fake news: Fake news regulation and media literacy interventions. Computers in Human Behavior, 80, 295–302. doi:10.1016/j.chb.2017.11.034

Mihailidis, P., Viotty, S., & Payne, J. (n.d.). Spreadable Spectacle in Digital Culture: Civic Expression, Fake News, and the Role of Media Literacies in “Post-Fact” Society. American Behavioral Scientist, 61(4), 441–454. doi:10.1177/0002764217701217

Vargo, C., Guo, L., & Amazeen, M. (n.d.). The agenda-setting power of fake news: A big data analysis of the online media landscape from 2014 to 2016. New Media & Society, 20(5), 2028–2049. doi:10.1177/1461444817712086

Brigida, M., & Pratt, W. (n.d.). Fake news. North American Journal of Economics and Finance, 42, 564–573. doi:10.1016/j.najef.2017.08.012

M. Rajdev, and K. Lee. Fake and Spam Messages: Detecting Misinformation during Natural Disasters on Social Media (short paper). WI. December 2015.