Text Analysis: Health Misinformation

Mason Womble

Sharon Jepkosgei

Kendall Goodland

“Scientists Warn People to Stop Eating Instant Noodles Due to Cancer and Stroke Risks”

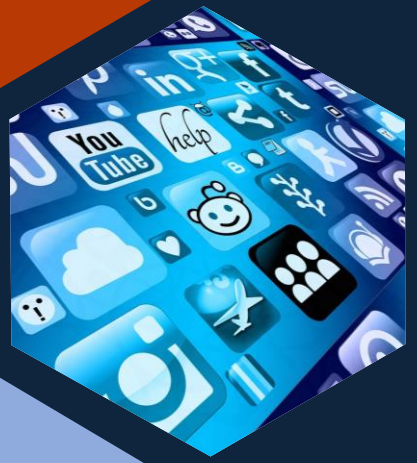
Introduction

- How many times have you seen such unproven statements online?
- Yet, 80% of people online use the internet to search for health information (Zadrozny, 2019)
- Objective – Establish recurring themes, prominent words, and dissemination patterns associated with health misinformation

Research Questions

1.Is health misinformation characterized by specific themes, topics, or words?

2. Which channels and platforms are primarily utilized to disseminate fake health information?



Research Hypothesis

1. Health misinformation exhibits distinct thematic patterns, with certain words and topics recurring more frequently than others.

2. Fake health information is disseminated through specific channels and platforms, contributing to its persistence and amplification.

How we get there

Data

- Datasets with Fake vs. Accurate information
- HealthNewsReview.org - story or release
- CoAID
- YouTube data

Technique

- Text analysis - Natural Language Processing tools

Visualization

- Word Clouds
- Network diagrams

Analysis Plan

Keywords Analysis

- Bag of Words
-Word frequency analysis

Thematic Patterns

N-gram analysis

Language Usage

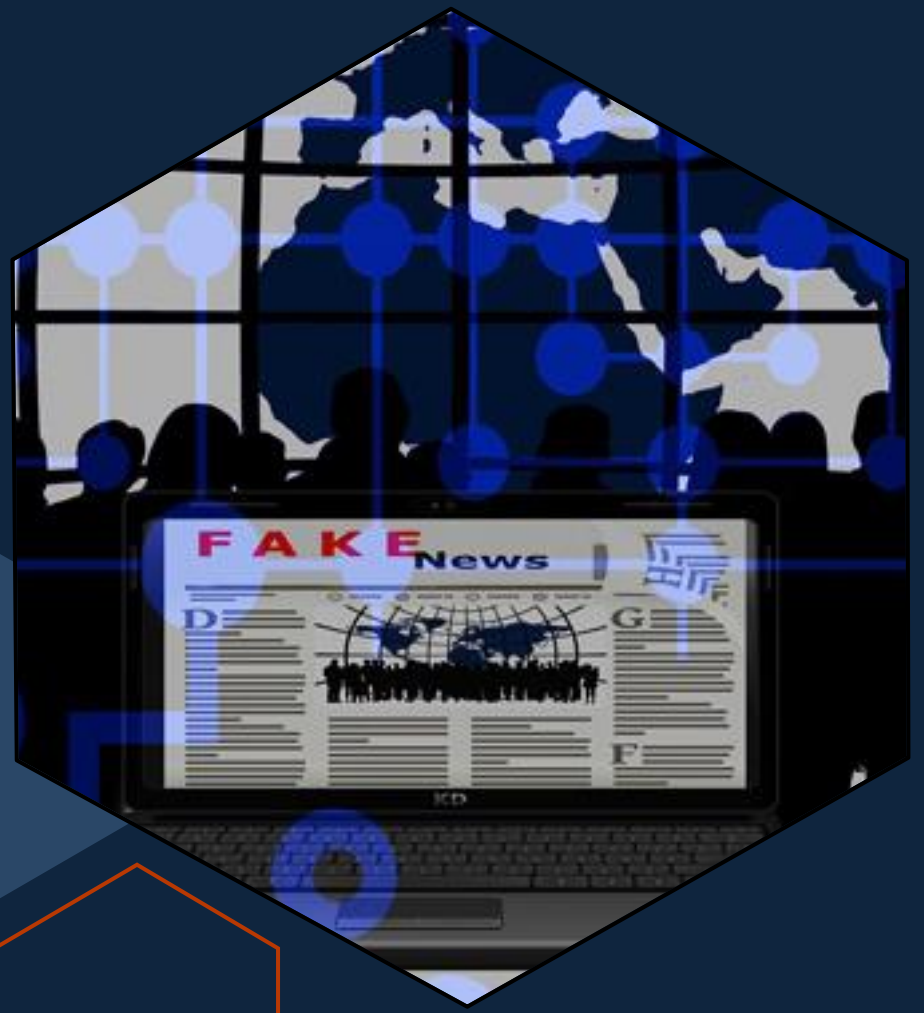
Sentiment analysis:
-NRC Lexicon
-Polarity Score

Topic Modeling

Latent Dirichlet Allocation (LDA)

Summary

Analyzing themes, words, and patterns of information dissemination allows for us to understand variations between fake and credible health news . It may help identify a set of features for future supervised learning.



Works Cited

- Hayawi, K., Shahriar, S., Serhani, M. A., Taleb, I., & Mathew, S. S. (2022). ANTi-Vax: a novel Twitter dataset for COVID-19 vaccine misinformation detection. *Public health*, *203*, 23–30. https://doi.org/10.1016/j.puhe.2021.11.022
- Zadrozny, B. (2019, December 29). *These are the fake health news that went viral in 2019*. NBCNews.com. https://www.nbcnews.com/news/us-news/social-media-hosted-lot-fake-health-news-year-here-s-n1107466




Thank you