Understanding Misinformation: A Case Study of COVID-19 Social Media Posts

Group 2: Teshinee Kukamjad, Jolina Hor, Sam Reade

Motivation

March 2020 - COVID-19 declared a pandemic

- What happened in between? → **constant spread of misinformation**

Referenced past works like Murugesan et. al (2022) and Farhoudinia et. al (2024) for project, however, there is still **research gap in finding most important features in misinformation detection**

Research Question

Which social media features are most important for distinguishing real vs misinformation posts?

Dataset

Real	Fake	Total
5,600	5,100	10,700

Fighting an Infodemic: COVID-19 Fake News Dataset (Patwa et al., 2021)

- 10,700 English-language social media posts and articles collected through web scraping and the Twitter API.
- **Annotation**: Human annotators manually verified posts to ensure accuracy
- **Dimensions:** 10,700 x 2

posts	label
Our daily update is published. States reported 734k tests 39k new cases and 532 deaths. Current hospitalizations fell below 30k for the first time since June 22.	real
Alfalfa is the only cure for COVID-19.	fake
President Trump Asked What He Would Do If He Were To Catch The Coronavirus https://t.co/3MEWhusRZI #donaldtrump #coronavirus	fake

Data Cleaning

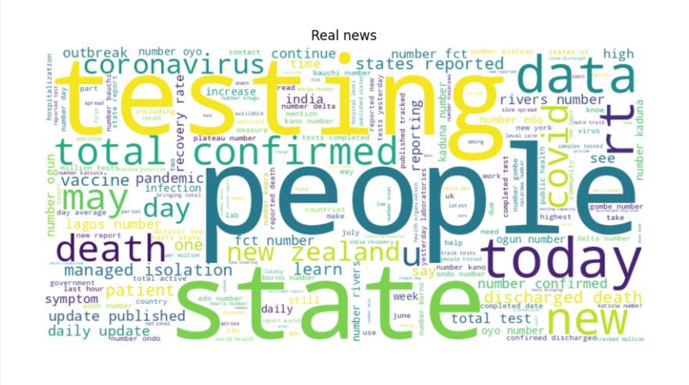
- Converted text to lowercase
- Replaced:
    a. URLs (https://...) ➔ 'link'
    b. Mentions (@user) ➔ 'mention'
    c. Numbers (e.g., 24k, 3.5M, 12,345) ➔ 'number'
- Applied **TweetTokenizer** to split each posts into list of tokens
- Removed Standard English stopwords from NLTK library
- Removed top 20 common words

Attribute	Fake	Real
Unique words	19728	22916
Avg words per post	21.65	31.97
Avg chars per post	143.26	218.37

(Patwa et al., 2021)

Word Clouds

Real News

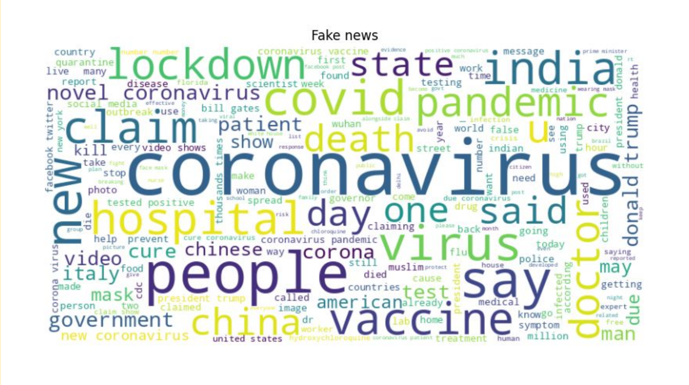


Dominant words: People, Testing, State

Common words: Data , Today, Confirmed

Emphasizes verified facts and real time updates

Fake News



Dominant words: Coronavirus

Common words: People, Vaccine, Hospital, Pandemic, Death

Amplifies fear and uncertainty

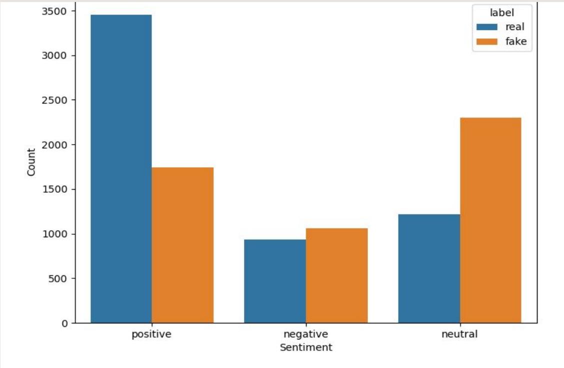
Methods

For each posts, we extract the following features:

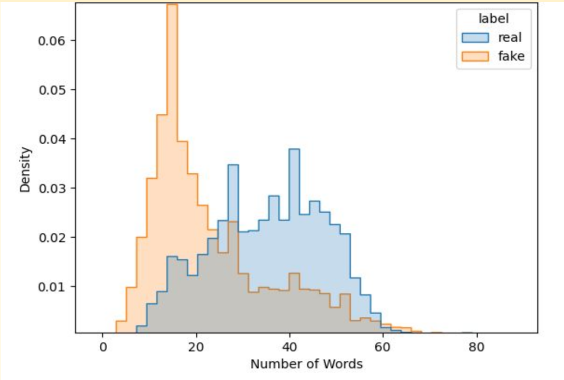
- Length
- Hashtag count
- Link presence
- Sentiment polarity
- Important words (identified using TF-IDF weighting)

which are used to train a Random Forest model to classify posts as real or fake

Sentiment Label



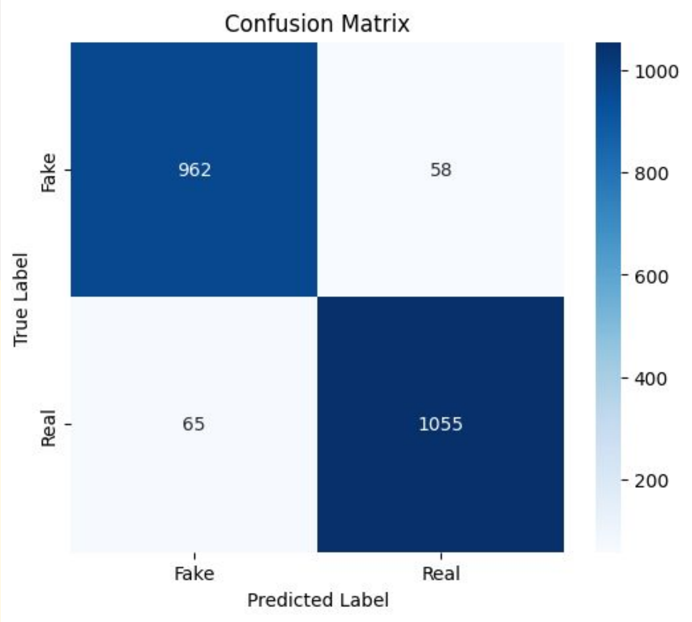
Post Length Distribution



Performance

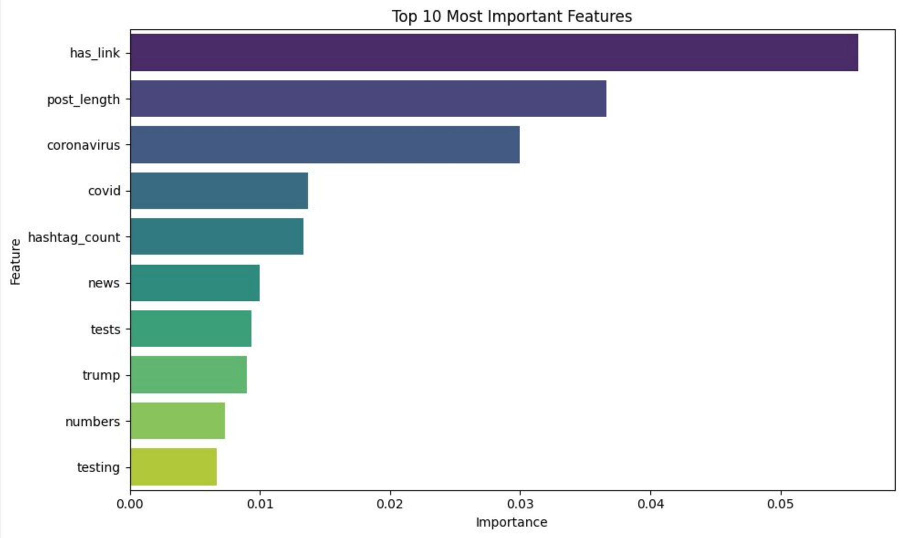
Classification Report				
	Precision	Recall	F1-Score	Support
0	0.94	0.94	0.94	1020
1	0.95	0.94	0.94	1120
Accuracy			**0.94**	2140
Macro Avg	0.94	0.94	0.94	2140
Weighted Avg	0.94	0.94	0.94	2140

where 0 indicates fake news and 1 indicates real news

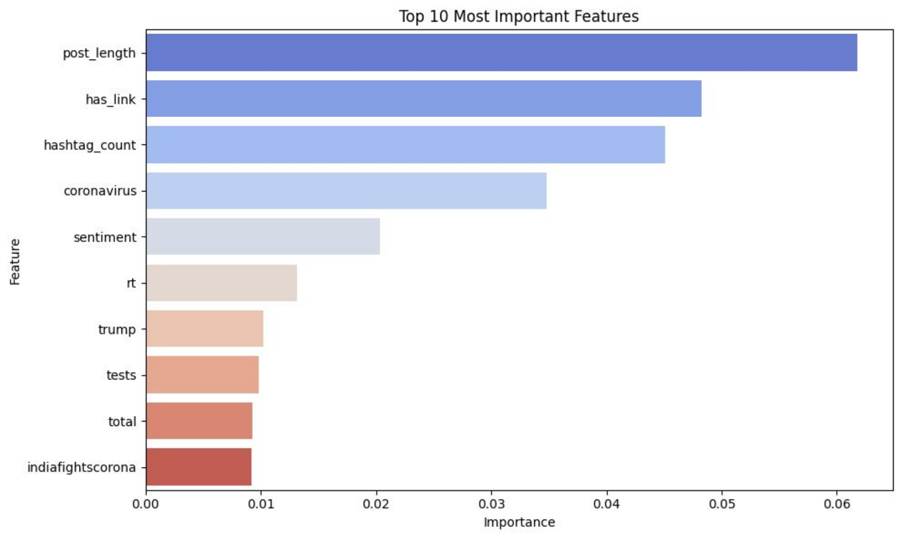


Global Feature Importance

Permutation



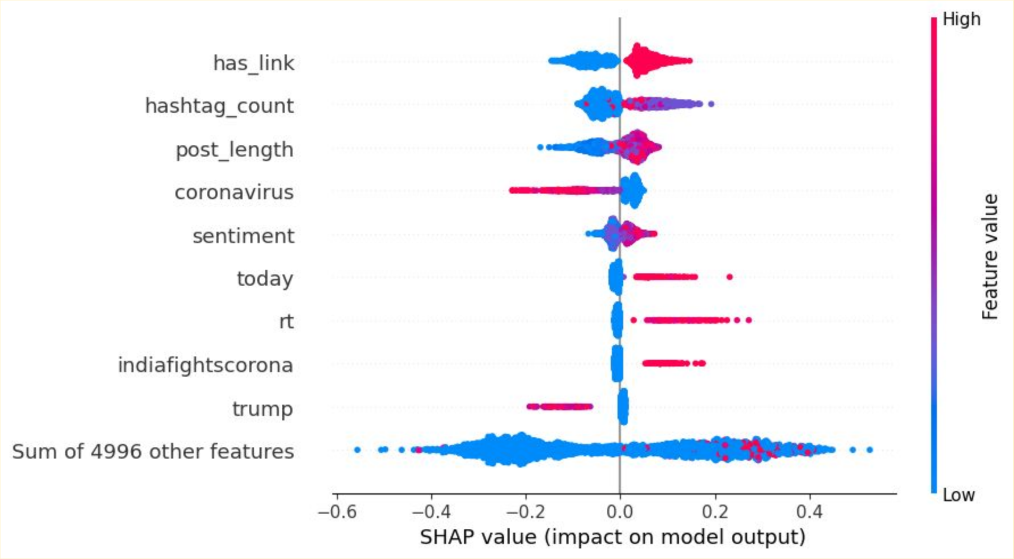
Impurity-Based



SHAP *(SHapley Additive exPlanations)*

Feature Importance (Real Posts):

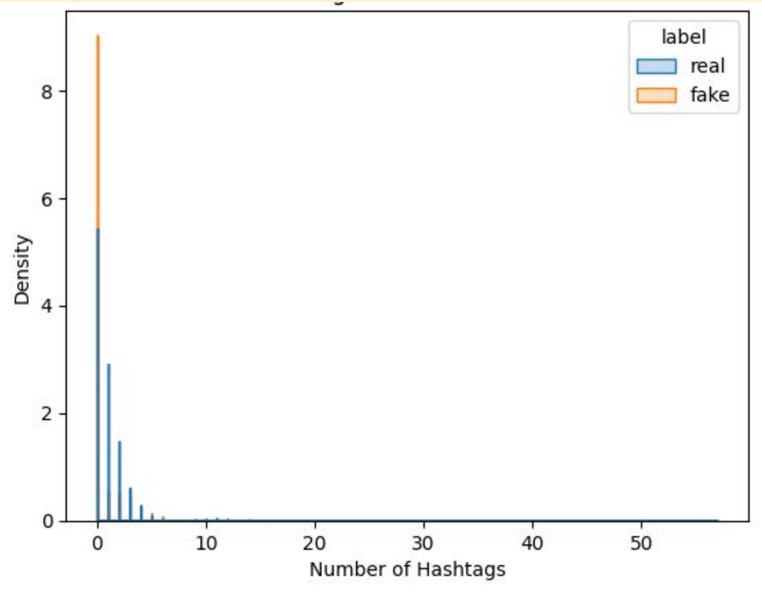
- compares what model predicts with and without a feature
- assigns each feature a contribution value (SHAP)
- values push observations toward one classification (label)



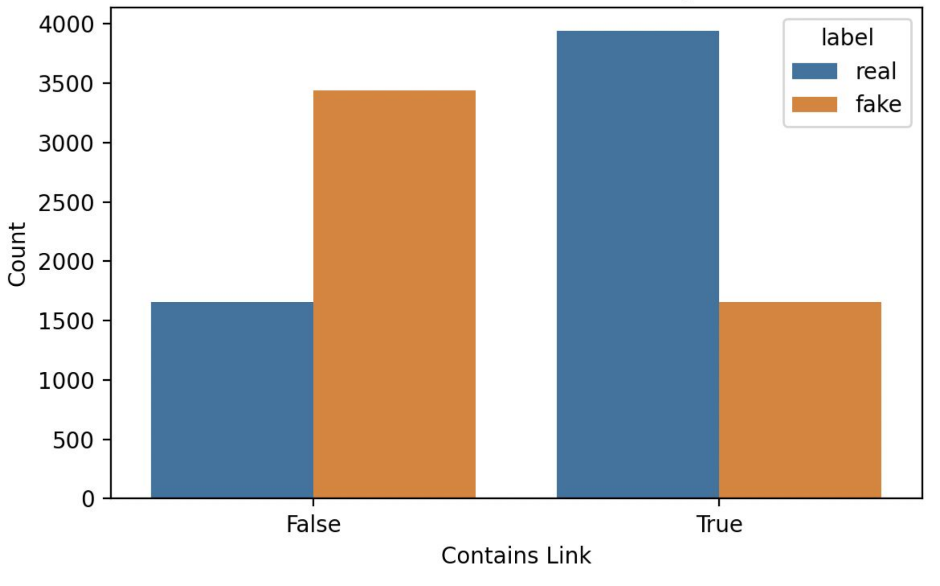
Feature	Value	Impact
Links on Posts	High	+
Hashtags	High	+
Word Length	High	+
Sentiment	Low	–
“trump”	High	–
Other Features	Low	–

Visualizations (1)

Hashtag Count

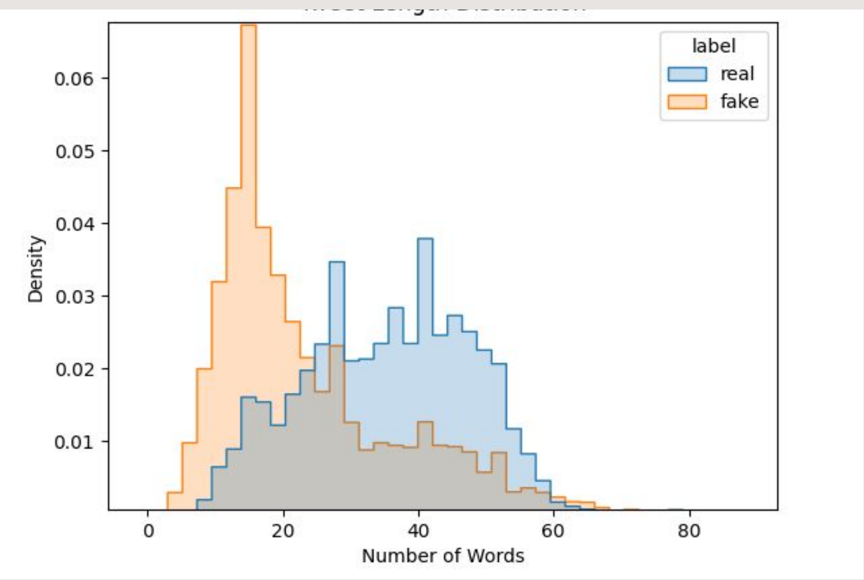


Link Distribution

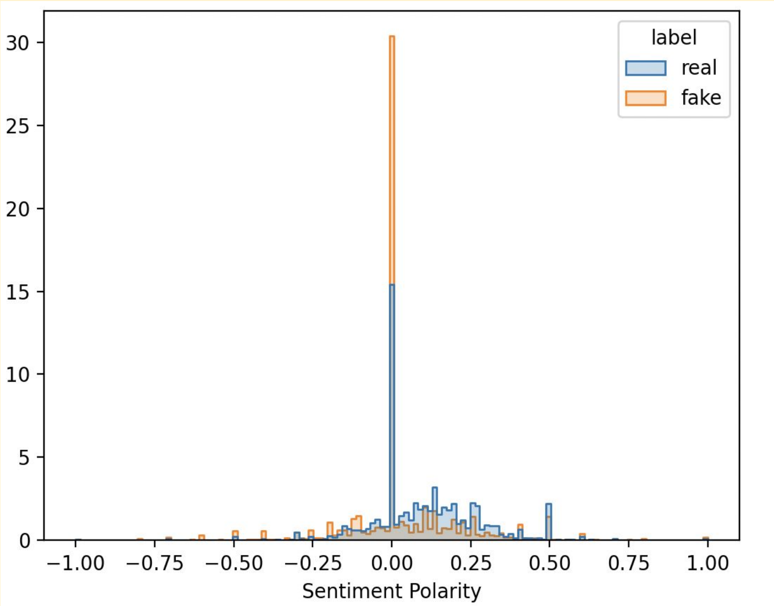


Visualizations (2)

Post Length Distribution



Sentiment Distribution



Visualization Analysis

Real News:

- **Longer** post content
- Usually **include external sources** that are more informative, referring to evidence for credibility
- **Greater variation in sentiment, especially positively**, possibly in relation to informative and uplifting news

Fake News:

- **Shorter** content to quickly grab attention
- Typically contain f**ewer or no external links**, indicating lack of reliable sources
- **Neutral/flat language** to avoid detection or attempt objectivity
- **Slightly more prevalence in negative sentiment**, potentially used to provoke readers

References

Cáceres, M. M. F., Sosa, J. P., Lawrence, J. A., Sestacovschi, C., Tidd-Johnson, A., Rasool, M. H. U., Gadamidi, V. K., Ozair, S., Pandav, K., Cuevas-Lou, C., Parrish, M., Rodriguez, I., & Perez Fernandez, J. (2022). The impact of misinformation on the COVID-19 pandemic. *AIMS Public Health, 9*(2), 262–277. https://doi.org/10.3934/publichealth.2022018

Farhoudinia, Bahareh, Selcen Ozturkcan, and Nihat Kasap. "Emotions Unveiled: Detecting COVID-19 Fake News on Social Media." *Humanities and Social Sciences Communications*, vol. 11, 2024, article no. 640. Palgrave Macmillan, https://doi.org/10.1057/s41599-024-03083-5. SSRN, https://ssrn.com/abstract=4833527.

Murugesan, S., and K. Pachamuthu. "Fake News Detection in the Medical Field Using Machine Learning Techniques." International Journal of Safety and Security Engineering, vol. 12, no. 6, 2022, pp. 723–727. https://doi.org/10.18280/ijsse.120608.

References

Nelson, T., Kagan, N., Critchlow, C., Hillard, A., & Hsu, A. (2020). The danger of misinformation in the COVID-19 crisis. *Missouri Medicine, 117*(6), 510–512. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7721433/

Patwa, P., Sharma, S., Pykl, S., Guptha, V., Kumari, G., Akhtar, M. S., Ekbal, A., Das, A., & Chakraborty, T. (2021). Fighting an infodemic: COVID-19 fake news dataset. In T. Chakraborty, K. Shu, H. R. Bernard, H. Liu, & M. S. Akhtar (Eds.), *Combating Online Hostile Posts in Regional Languages during Emergency Situations* (pp. 21–29). Springer. https://doi.org/10.1007/978-3-030-73696-5_3