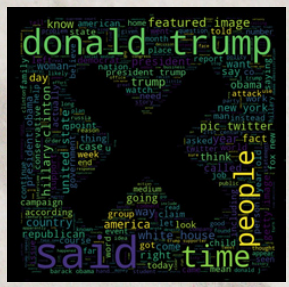
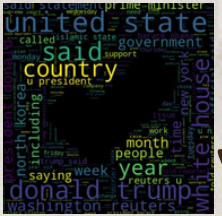
???TRUE News???
¿¿¿FAKE News¿¿¿

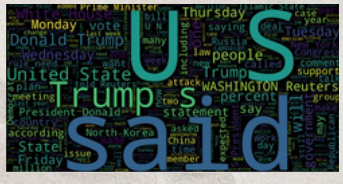
• Fake News Detection •

Data Engineering

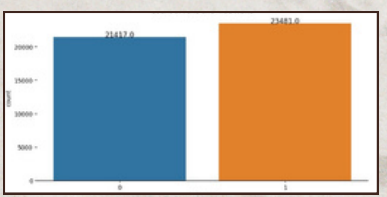
data visualization

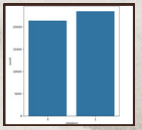






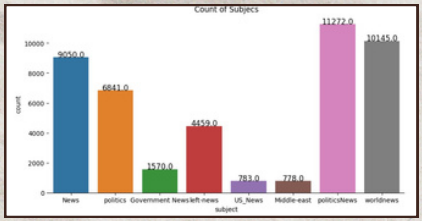
Data visualization in our project highlights ***trends*** and ***patterns*** in news articles, making it easier to interpret the distribution of *fake vs. real news* and model performance.





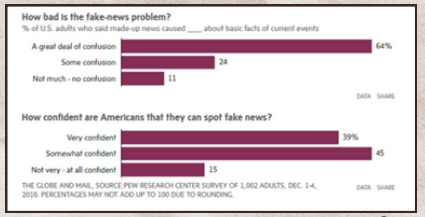
PS: The rise of digital media has fueled the ***spread of misinformation***, making it harder to distinguish between real and fake content. To tackle this, we propose a machine learning system that uses ***NLP techniques*** to accurately classify news articles, providing ***real-time insights*** to combat misinformation.

data cleaning • data-preprocessing



Model Report

	precision	recall	f1-score	support
0	0.99	0.99	0.99	7047
1	0.99	0.99	0.99	7562
accuracy			0.99	14609
macro avg	0.99	0.99	0.99	14609
weighted avg	0.99	0.99	0.99	14609



Formulae

$$\text{TF-IDF}(t, d, D) = \text{TF}(t, d) \times \text{IDF}(t, D)$$

$$\text{TF}(t, d) = \frac{\text{Number of times term } t \text{ appears in document } d}{\text{Total number of terms in document } d}$$

$$\text{IDF}(t, D) = \log\left(\frac{\text{Total number of documents in corpus } D}{\text{Number of documents containing term } t}\right)$$

TF-IDF measures the importance of a word in a document relative to a collection of documents. It enhances text analysis by highlighting distinctive terms, crucial in NLP tasks like fake news detection.

source

kaggle

We used two Kaggle datasets: ***'Fake.csv'*** and ***'Real.csv'*** for training and testing, with text features like headlines and content for feature extraction and classification.

TF-IDF Algorithm

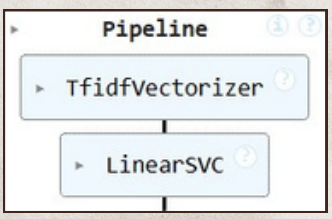
99% accurate • detection •

variables

- News Source
- Language
- Article Category
- Author
- Label
- Publication Date

Features

Univariate Selection
Correlation Analysis
Mutual Information
Tree-Based Models
Lasso (L1)
Info Gain



References

https://youtu.be/f7ckMaEJvh4?si=QiCgLxnDqGQu-c72
https://youtu.be/2o-_bMKGi_o?si=g8_DnHgUAPg2gK5E
https://github.com/kapilsinghnegi/Fake-News-Detection
https://www.projectpro.io/article/fake-news-detection-project/854
https://sjcit.ac.in/wp-content/uploads/2022/11/1SJ18CS098-soniya-cj-1.pdf

Team • Members

37 - Manasi Mali
43 - Riya Kondawar
45-Suhani Choudhary

conclusion

Our project effectively leverages AI and NLP to combat misinformation by accurately detecting and classifying fake news, contributing to a more informed digital landscape