Fake news classification

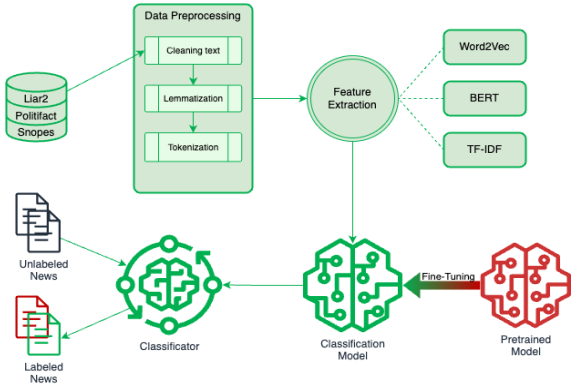
Authors:
Gia Minh Le
Nikita Ovsyannikov
Sebastian Sierocki

BACKGROUND:

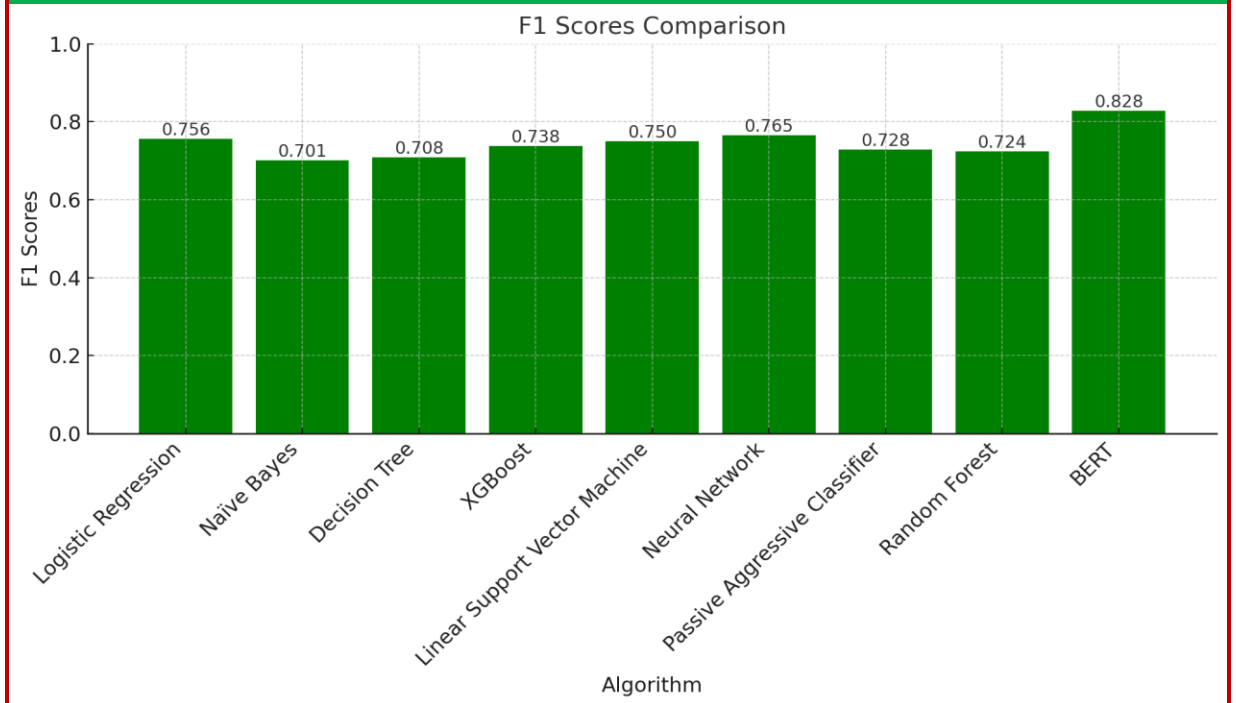
- Fake news has emerged as a critical issue in the digital age, spreading misinformation at an unprecedented scale. Detecting fake news is essential to safeguarding public trust and combating its harmful effects
- In this project, we utilize multiple datasets containing 67812 articles to develop a robust classification approach

The word cloud below highlights frequent terms associated with fake news:

METHODS:

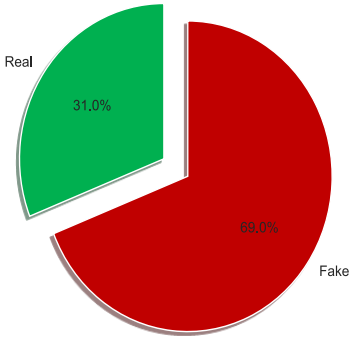


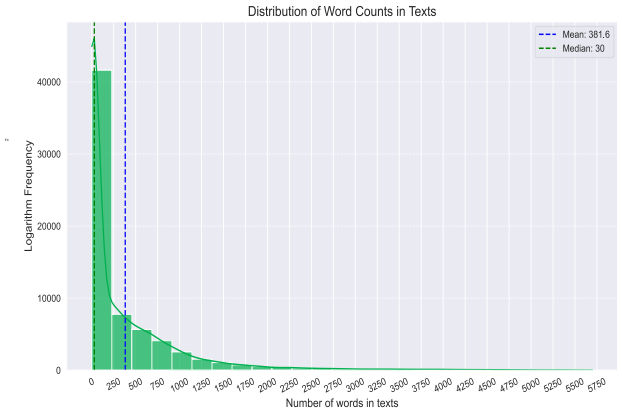
After training and testing with the dataset, **BERT** outperformed other models with the highest F1 score, showcasing its ability to capture complex patterns in text, especially for classification tasks.



Future work: the overall scores are moderate due to the inherent challenges of processing and classifying nuanced news data.

ORIGINAL DATASET VISUALIZATION:





DATA REFRENCES:

POLITIFACT Snopes

LIAR