

FAKE AND TRUE INFORMATION SPREAD ON INTERNET DATA ANALYSIS

Nadia Sevilla; Mayelyn García

CAP4767 Data Mining. Data Analytics Program. Miami Dade College

Objective

Since information is a service so widespread that shows no signs of slowing down, it becomes increasingly relevant to understand how to distinguish misinformation and to develop critical thinking skills.

This project pretends to offer an insight into the characteristics and structure of disingenuous articles, as well as that of genuine articles.

The graphics present on this dashboard offer a broad biew of the structural difference between the texts included in the dataset.

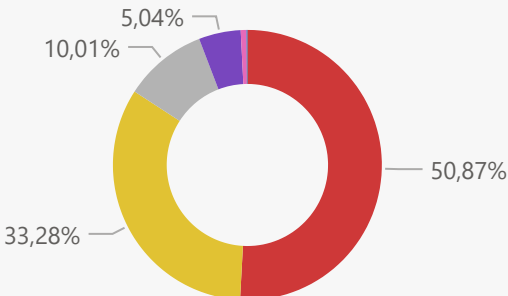
Tools Employed

- Python
- Power BI

Fake News

Main emotions manifested in text

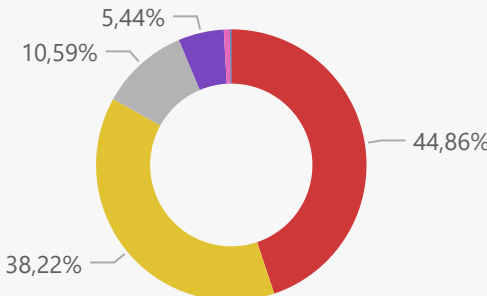
● anger ● joy ● sadness ● fear ● love ● surprise



Real News

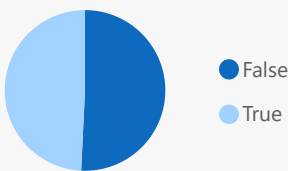
Main emotions manifested in text

● anger ● joy ● sadness ● fear ● love ● surprise



Total texts analyzed
203554

Distribution of fake and real news on this dataset

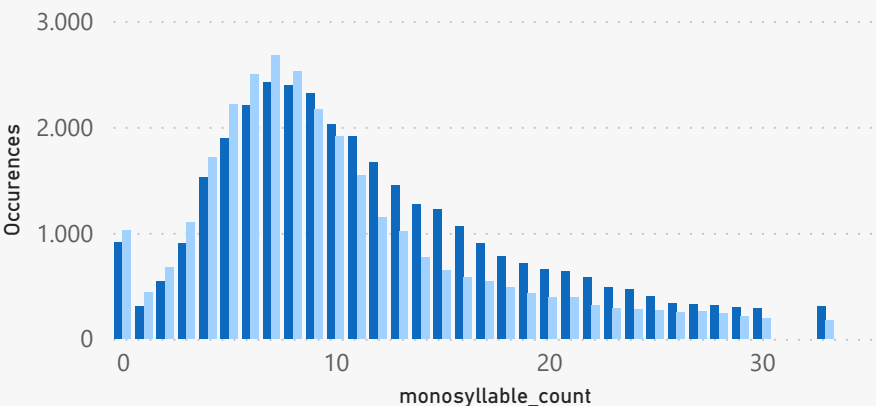


Sources

- Politifact statements.
- WELFake dataset for fake news detection in text data.
- [Gathering-real-news-for-oct-dec-2016](#)
- FakeNewsNet tweets dataset.
- ISOT Fake News dataset.
- An Heuristic-driven Ensemble Framework for COVID-19 Fake News Detection.
- EUvsDisinfo.
- Emotion detection in text using Neural Networks.

Count of monosyllabic words present on articles and statements

● False ● True



Count of polysyllabic words present on articles and statements

● False ● True

