Liar Liar!
Pants on Fire!

Fake News Detector powered by Machine Learning

Faiz Hasmani

Shruti Bangera

Yazhini Gopal

Nida Iqbal

Syed Muddassir Raza

Nora Younes

Reason for Topic Choice:

In a time when we can spread and receive information quickly, thanks to the technology and 4G networks , media literacy is very important.

Detecting fake news has become more important than ever for the sake of a peaceful society. A prediction tool would be a terrific way to help people make more educated decisions on news.

Only 26% of Americans are confident in their ability to detect fake news

THAT MEANS THAT

242,868,000

AMERICANS

WOULD NOT BE ABLE TO TELL FACT FROM FAKE

Source: Statista

A large set of True and Fake news would be used to create a model which would predict whether an input is True or False based on that model.

- The primary Dataset is from **University of Victoria and Kaggle Dataset.**
- **Top 20 Fake/Real** words using Naive Bayes classifier.

Questions We Hope to Answer:

- ➔ Is the article True of False?
- ➔ What is the accuracy of the prediction model?
- ➔ Are Fake news articles generally more negative in nature?

Description of the Data Exploration phase of the project.

As a group, we went through several research articles and data sources looking for options for our dataset. We knew we wanted:

- Source of the article
- Topic such as World News, Sports, Politics, etc.
- The Dataset to already determine if the article was Real/Fake

Description of the Analysis phase of the project.

Once deciding on a dataset, as a group, we needed to clean each dataset and join the Real and Fake CSV files into one. From here we used 1's and 0's to identify if the article was Real of Fake. We then analyzed the data set to find:

- Frequency of words in Real articles and Fake articles

Technologies and Tools used -

Technologies :

- Python
- SQLAlchemy
- Natural Language Processing
- Machine Learning
- HTML
- CSS
- Flask

Tools :

- Tableau
- Jupyter Notebook
- PostgreSQL

Recommendations for Future Use

This project was designed to work on political news and not others such as pop culture.

What the team would have done differently?

Quotes for illustration purposes only