

Predicting Fake and Real News

Research Questions:

- A) How has fake news changed?
B) Can we predict whether something is fake news or not using NLP?

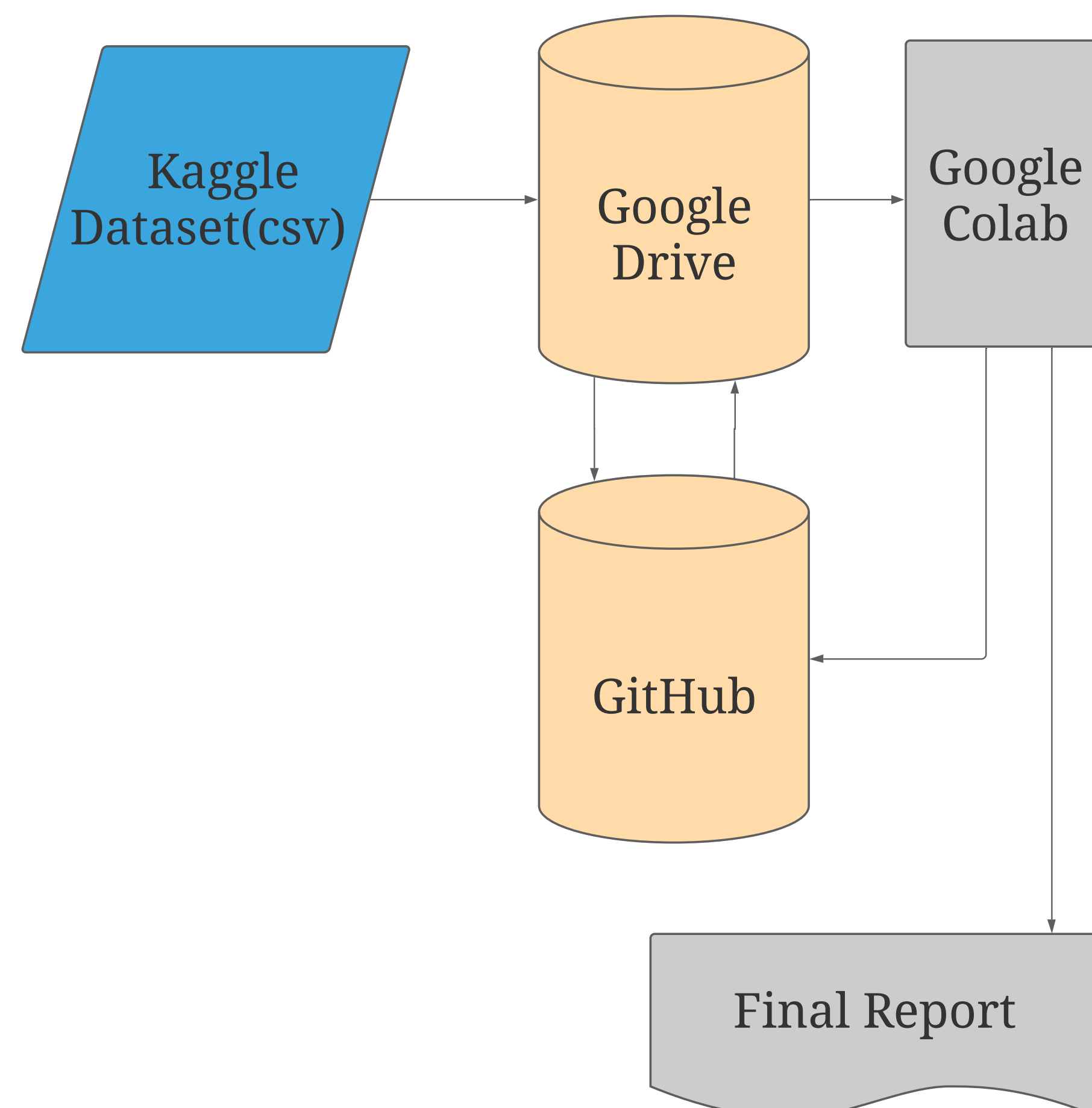
Introduction:

Fake news is not a recent or unique issue, although it has been exacerbated by the ease and speed in which information could be spread. Since the 1890s, fake news has plagued the profession of journalism. Dubbed "yellow journalism", newspaper publishers would run sensationistic news articles. However, fake news was eventually replaced by objective journalism, until recently due to the rise of the internet. (Source:UCSB)

In terms of impacts fake news has on society, University of Michigan lists the following as possibilities:

- Anti-intellectualism
- Antiscience
- Widespread Mistrust

Workflow:



Real News WordCloud:



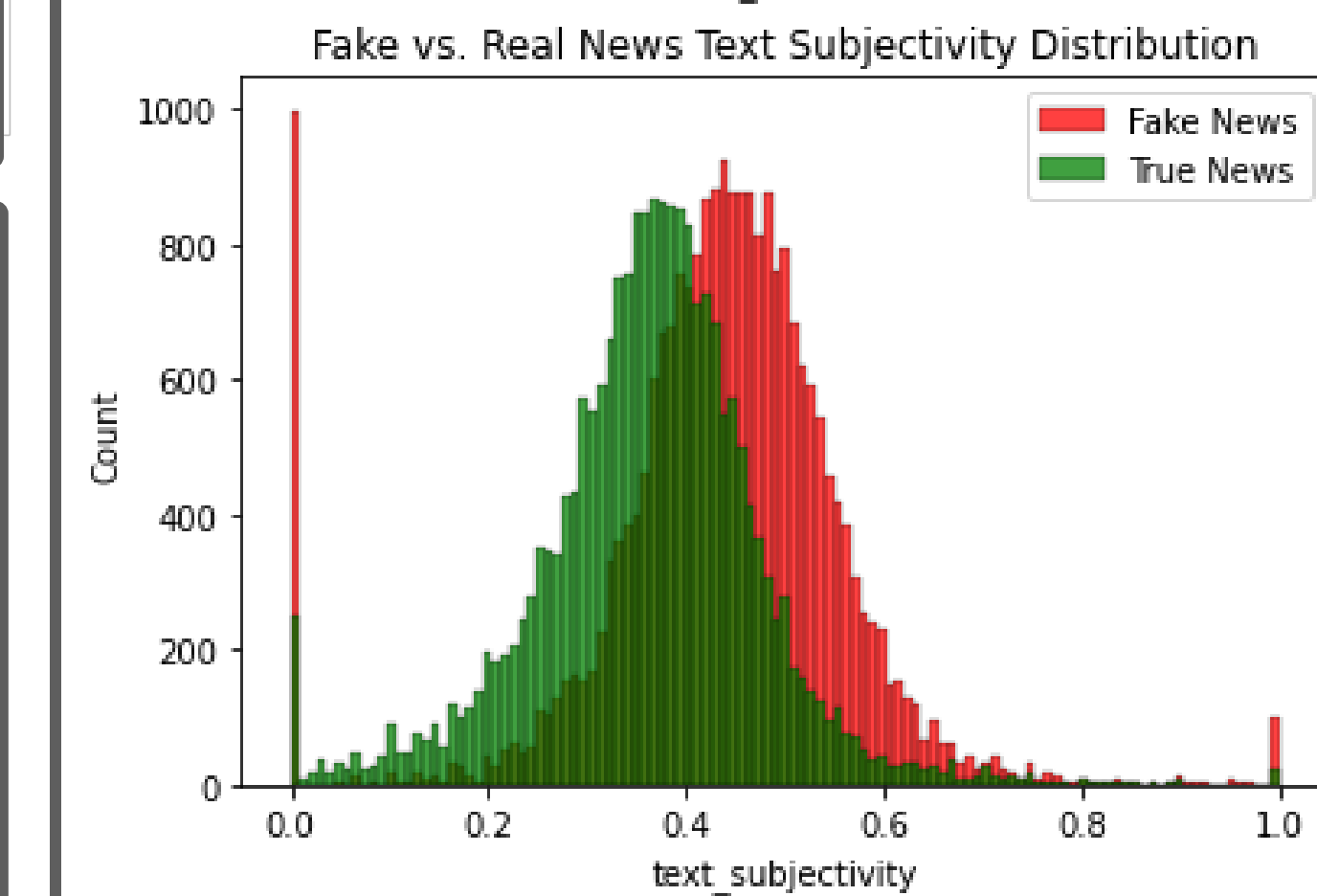
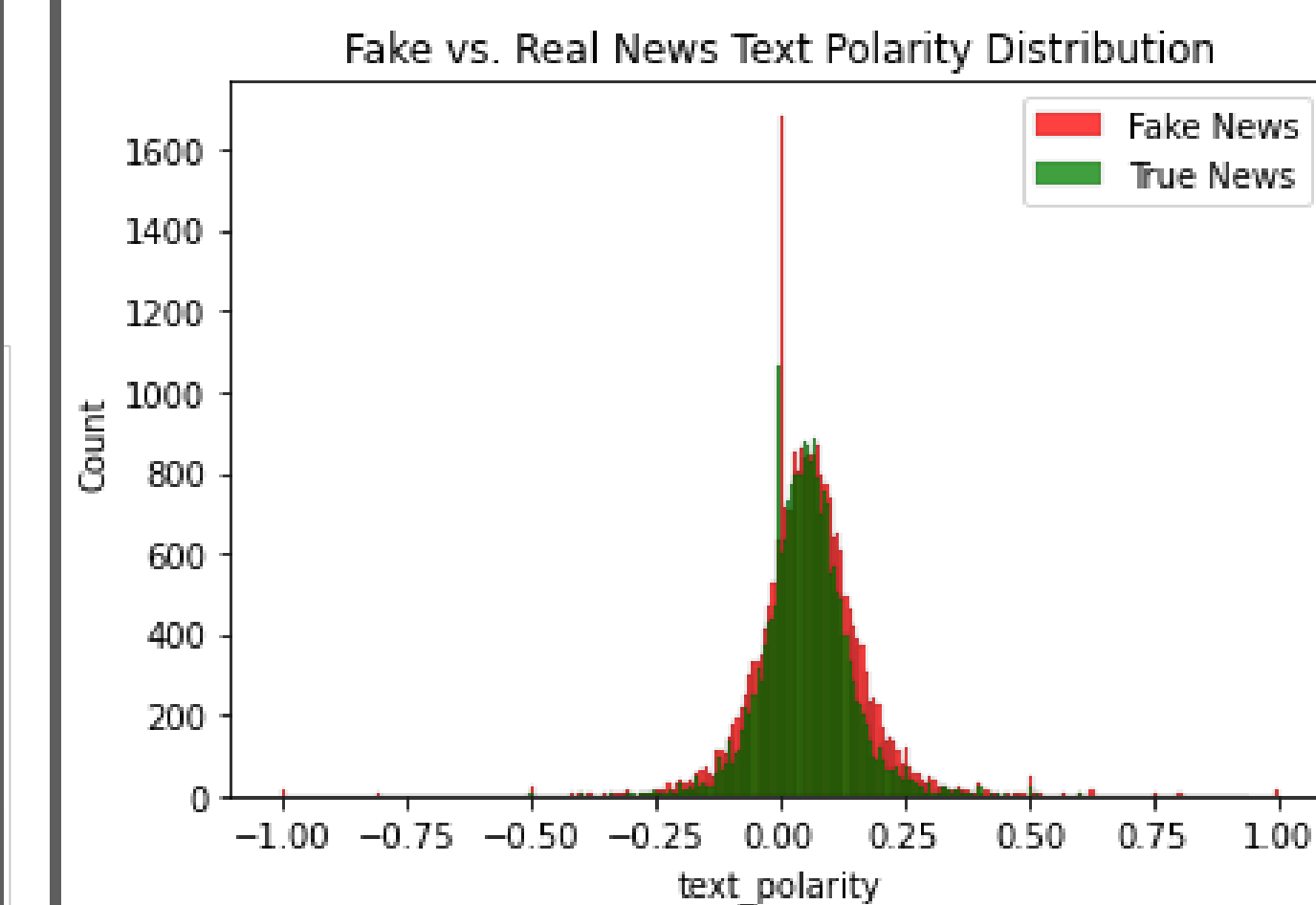
Fake News WordCloud:

**EDA:**

- From wordcloud, can see differences in words used
- Fake news dataset primarily about Donald Trump
- Sources matter!
 - Reuters makes up a large part of the real news data
 - 21st Century Wire, a right wing conservative web blog makes the a large part of fake news data
- Difference in topics discussed:
 - Fake news tends to be nationalistic (america) and more focused on traditional issues (christmas, sheriff)
 - Real news tends to be more about international issues (nato, brussels, london) and more on progressive ideals (transgender)
- Interestingly, it seems real news is more interested in conservatives than fake news is.

Sentiment Analysis

- Tracked polarity and subjectivity scores
- Distribution of polarity scores were not markedly different in true and fake news
- Fake news contains more subjective vocabulary than true news



Predicting Fake News vs. Real News:

- Explain model fitting process:
 - Logistic Regression
 - SVM
 - etc.
- Analysis of performance of models
- ROC Chart/ Binary Confusion Matrix

Conclusion:

- Summarize findings of EDA
- Summarize changes between fake and real news
- Summarize findings from predictive task
- Analyze suitability of tools used for task

Works Cited

- Insert citations: articles, documentation, etc.