Sonal, Spring'23
Identifying Fake News on Social Media
(1284 words)

1. INTRODUCTION

Fake news has become an increasingly prevalent issue in today's digital age. The term "fake news" refers to fabricated or misleading information that is presented as if it were true, often with the intent to mislead or deceive readers. Social media platforms have become a primary source of fake news due to their wide reach and ease of sharing content. Fake news can have a significant impact since it may impact public opinion, affect political elections, and even injure people or organizations. The spread of false information may also destroy public confidence in established news organizations, making it more challenging for consumers to discriminate between true and false news. Multiple factors, such as the absence of fact-checking, the simplicity of creating and distributing content online, and the use of bots or automated accounts to spread misleading information, contribute to the spread of fake news. Fact-checking, media literacy education, and other measures have been taken to try to stop the spread of false news as knowledge and concern about its effects have grown in recent years. Identifying and stopping the spread of fake news is an ongoing challenge, and requires a multidisciplinary approach that involves journalists, researchers, and technology companies. It can be difficult to distinguish real news from fake news because of the volume of information shared on social media. It's essential to recognize fake news on social media to prevent its adverse effects. In this study, topic modeling using LDA (Latent Dirichlet Allocation) will be used to analyze the text of news to distinguish between fake news and real news will be performed to overcome this challenge.

2. RESEARCH QUESTION

Can topic modeling techniques be used to accurately identify fake news on social media?

3. METHOD

3.1. DATA

The dataset used in this study is the "Fake and Real News Dataset" from Kaggle:
- https://www.kaggle.com/clmentbisaillon/fake-and-real-news-dataset

The dataset consists of two CSV files, one containing the fake news articles and another containing the real news articles. The time-period of the dataset is between January 2016 and December 2017. This dataset consists of 44,918 news articles, including 21,416 real news articles and 23,502 fake news articles across both CSV files. The dataset includes metadata such as the title, text, subject, and date of the article.

3.2. DATA ANALYSIS

For exploratory analysis, a word cloud analysis is conducted to find out the most frequent words used. The LDA model allows better understanding of individual topics and relationships between the topics. A visualization of the topics and high frequency words will be generated.

Data analysis is performed on two datasets, one containing fake news articles and the other containing real news articles. First, the data is preprocessed by converting the text to lowercase, removing punctuation, tokenizing, removing stop words, and stemming. The preprocessed text is then used to create a corpus, which is a collection of documents from both the datasets. A dictionary is created from the corpus and words that occur less than 10 times or more than 50% of the documents are filtered out. The bag-of-words corpus is used to train a topic modeling algorithm called Latent Dirichlet Allocation (LDA). To visualize the top topics, word cloud is generated to find out the most frequent words.

The LDA model is then evaluated using a coherence score, which measures the semantic similarity between the top words in each topic. The coherence score of 0.43523023927525994 is obtained by the LDA model. The top topics and their associated words are printed to the console. The analysis concludes with the creation of a word cloud visualization of the most frequent words in the corpus. Further, an inter-topic distance map is created to visualize the distance between topics in the topic space. Additionally, the efficiency of the model is evaluated using log perplexity, where a score of -7.980899644119352 is obtained.

4. RESULTS

Figure 1 is a word cloud image of all the words from the corpus that was created using the WordCloud module in Python.
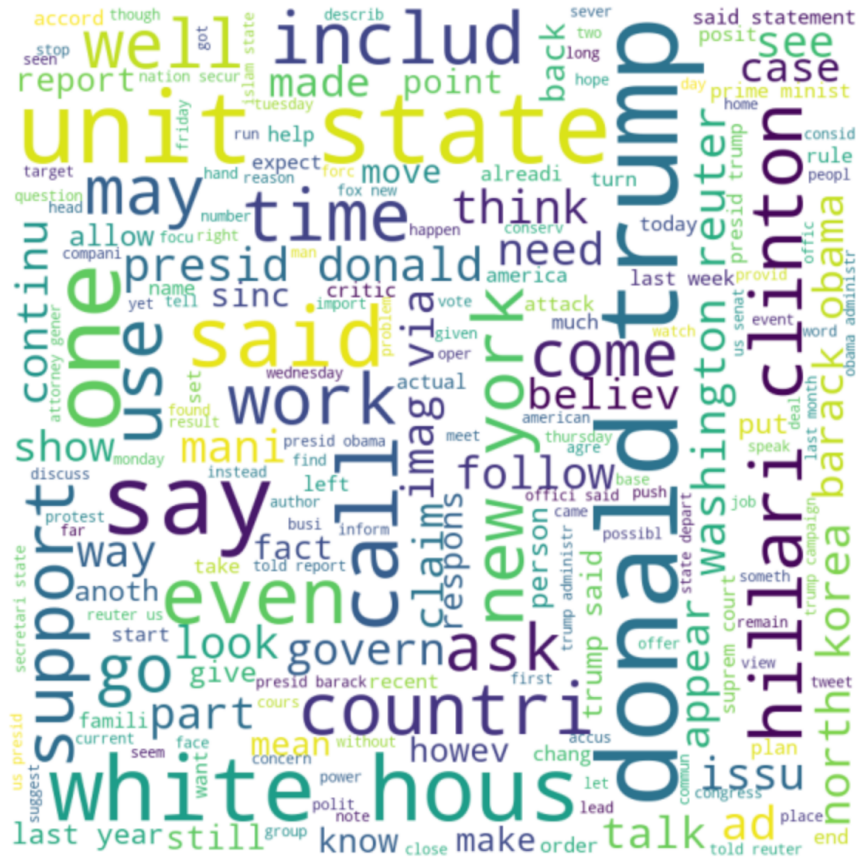

Figure 1: Word Cloud Based on

The analysis conducted on the corpus using the genism module resulted in the generation of 10 topics. Each topic is a combination of keywords, and each keyword contributes a certain weight to the topic. Figure 2 displays the 10 topics generated by the analysis.

To visualize the results of the topic modeling, pyLDAvis was used to create an interactive chart, which is displayed in Figure 3. The chart displays the top 30 terms for each of the 10 topics generated by the LDA model. The size of the circles

represents the frequency of the term within the corpus, and the distance between the circles represents the similarity between the topics.

```
Topics:
Topic: 0
Words: 0.009*"year" + 0.008*"percent" + 0.008*"tax" + 0.008*"compani" + 0.008*"million" + 0.007*"us" + 0.007*"billion" + 0.006*"govern" + 0.006*"plan" + 0.006*"fund"
Topic: 1
Words: 0.010*"like" + 0.008*"one" + 0.008*"peopl" + 0.008*"say" + 0.007*"go" + 0.007*"time" + 0.007*"get" + 0.006*"know" + 0.006*"even" + 0.006*"think"
Topic: 2
Words: 0.025*"polic" + 0.016*"kill" + 0.013*"attack" + 0.010*"arrest" + 0.009*"offic" + 0.007*"two" + 0.007*"peopl" + 0.007*"report" + 0.006*"charg" + 0.006*"one"
Topic: 3
Words: 0.184*"trump" + 0.033*"donald" + 0.021*"white" + 0.016*"hous" + 0.012*"us" + 0.008*"irma" + 0.008*"tweet" + 0.008*"twitter" + 0.007*"campaign" + 0.007*"washington"
Topic: 4
Words: 0.013*"group" + 0.011*"milit" + 0.011*"muslim" + 0.011*"rohingya" + 0.011*"peopl" + 0.008*"right" + 0.008*"protest" + 0.008*"polit" + 0.007*"nation" + 0.007*"parti"
Topic: 5
Words: 0.011*"peopl" + 0.009*"refuge" + 0.008*"citi" + 0.007*"year" + 0.007*"countri" + 0.007*"border" + 0.006*"immigr" + 0.005*"govern" + 0.005*"mexico" + 0.005*"home"
Topic: 6
Words: 0.014*"investig" + 0.011*"report" + 0.011*"russian" + 0.010*"us" + 0.009*"clinton" + 0.009*"russia" + 0.008*"offici" + 0.008*"former" + 0.007*"intellig" + 0.007*"inform"
Topic: 7
Words: 0.026*"parti" + 0.025*"republican" + 0.025*"elect" + 0.021*"vote" + 0.021*"democrat" + 0.017*"senat" + 0.011*"support" + 0.010*"poll" + 0.010*"campaign" + 0.010*"candid"
Topic: 8
Words: 0.013*"state" + 0.011*"us" + 0.010*"reuter" + 0.008*"govern" + 0.008*"minist" + 0.008*"unit" + 0.008*"north" + 0.006*"korea" + 0.006*"militari" + 0.006*"countri"
Topic: 9
Words: 0.031*"court" + 0.020*"law" + 0.017*"state" + 0.015*"rule" + 0.010*"case" + 0.009*"right" + 0.009*"order" + 0.008*"legal" + 0.008*"justic" + 0.008*"us"
```

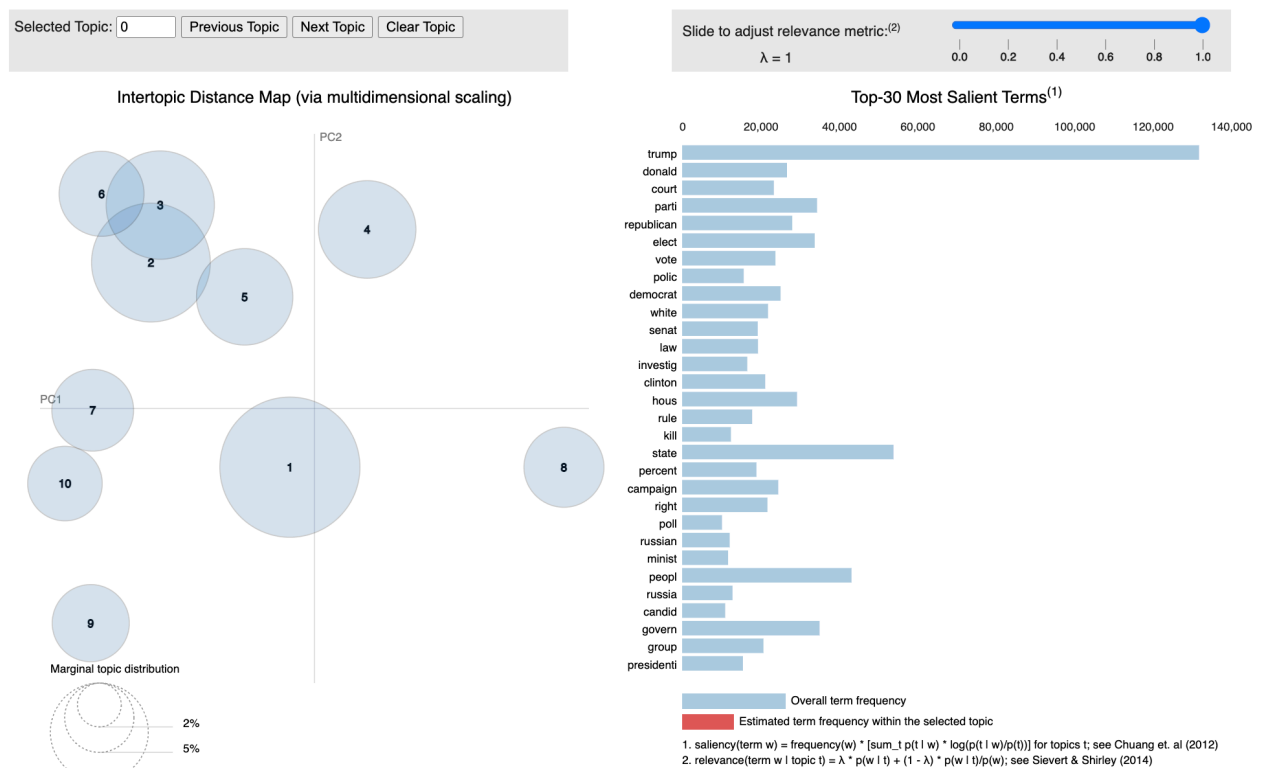**Figure 2:** Top 10 Topics Generated by Topic Modeling using LDA

Figure 3: Visualization of Top 10 Topics Using LDA

The interactive chart provided by pyLDAvis allows for a more detailed exploration of the topics generated by the LDA model. Users can hover over the circles to view the frequency and relevance of the terms within each topic. They can also click on the circles to view the individual terms and their contributions to the topic. Overall, the combination of the word cloud image, the list of generated topics, and the interactive chart provides a comprehensive view of the topics and their relationships within the corpus.

5. CONCLUSION

Answering to the question "Can topic modeling techniques be used to accurately identify fake news on social media?", the inter-topic distance map, coherence score, and log perplexity are all important metrics for evaluating the performance of a topic modeling algorithm.

From the topic model, we can see that the model has identified 10 topics with varying probabilities for each word. Topic 3 seems to be related to Donald Trump and his administration, while topic 7 is related to political parties, elections, and campaigns. Topic 2 is related to crime and law enforcement, and topic 4 seems to be related to human rights issues. Overall, the topics seem to cover a broad range of topics, indicating that the model is able to capture different themes in the news articles.

The analysis evaluates the quality of the LDA model based on coherence score and log perplexity. The coherence score of 0.43523023927525994 indicates that the LDA model is producing topics that are reasonably coherent. Generally, a higher coherence score indicates better topic coherence. However, the interpretation of coherence scores can be subjective, and the optimal score may vary depending on the application. In this case, a score of 0.435 is a decent score, indicating that the topics produced by the LDA model are moderately coherent. While, the log perplexity value of -7.980899644119352 indicates that the LDA model has a good fit on the corpus. Generally, a lower perplexity value is better, and a value close to zero indicates a perfect fit. In this case, the value is negative, which is not unusual for language models. However, the absolute value of the perplexity is relatively small, which suggests that the model is performing well on the given corpus.

However, topic modeling techniques can be used to analyze and understand the content of social media posts, including identifying patterns in the language used in fake news posts. By comparing the topics used in real and fake news posts, it may be

possible to develop a model that can accurately classify social media posts as real or fake news.

Additionally, an inter-topic distance map can be useful for visualizing the relationships between topics and identifying potential overlaps or similarities between topics that may indicate shared themes or topics between real and fake news posts. However, further research and analysis would be needed to determine the effectiveness of topic modeling techniques in accurately identifying fake news on social media.

REFERENCES

[1] Davidson, T., Warmsley, D., Macy, M., & Weber, I. (2017, May). Automated hate speech detection and the problem of offensive language. Proceedings of the International AAAI Conference on Web and Social Media, 11(1), 512-515.

[2] Kapadia, S. (2019, April 14). Topic modeling in python: Latent dirichlet allocation (lda): How to get started with topic modeling using lda in python.

[3] Ghanem, Bilal & Rosso, Paolo & Rangel Pardo, Francisco. (2020). An Emotional Analysis of False Information in Social Media and News Articles. ACM Transactions on Internet Technology. 20. 1-18. 10.1145/3381750.

[4] Xu, Kuai & Wang, Feng & Wang, Haiyan & Yang, Bo. (2020). Detecting fake news over online social media via domain reputations and content understanding. Tsinghua Science and Technology. 25. 20-27. 10.26599/TST.2018.9010139.