Sentiment Analysis of Public Opinion on the Russia-Ukraine Conflict

Yiwen Lu, Siheng Xiong, Zhaowei Li

Outline

- Introduction
    - Overview
    - Methodology
- Dataset
- LDA analysis
    - Theory
    - Result Analysis
- Visualization - Demo

Introduction

Overview

Starting from Feb 24th, 2022, Russia began an open-military conflict of Ukraine. From then on, related rumors spread online and make public confused about what is actually going on in the frontline. In this project we tend to figure out what people on social media care about, and their opinions about the conflict.

Methodology

- We plan to prepare a new dataset relating to this issue from twitter and label the dataset automatically with different sentimental analysis tools.
- Then analyze the relationship between the results and data characteristics like date, location, keyword and sentiment.
- Finally we will apply Latent Dirichlet Allocation(LDA) model to different groups and develop a UI to visualize the result.

https://www.bbc.com/news/world-europe-56720589

Data Preprocess

Data Source

A data source was obtained from kaggle, which uses the twitter API to fetch tweets that contain specific keywords like "Ukraine boarder", "Ukraine troops".

Data Preprocess

- Use a pretrained RoBERTa model trained on English tweets to perform sentimental analysis and yields 1.6% pos tweets, 45.48% neutral tweets and 52.92% negative tweets.
- Link the data with exact dates(events), countries and sentiment analysis results.
- For each tweet, we need to convert the attitude:
  'POS', 'NEU', 'NEG' into discrete variables, such as '1' ,'0' ,'-1'

- For each country, we average the scores by countries ,and the 'mean' represents the overall attitude of this country to this conflict. e.g. -0.91: very pessimistic

LDA Analysis

LDA brief overview

**Topic modeling: Latent Dirichlet Allocation**
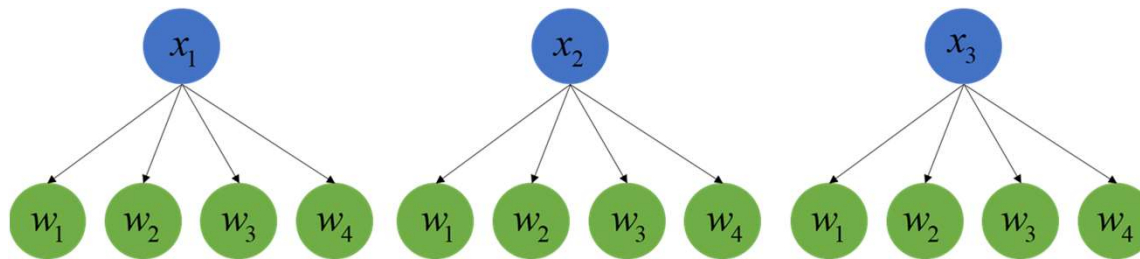
**Generative procedure: D comments, N words, V-word vocabulary, K topics**
- **For each topic** $1, \ldots, K$
  - **Draw a multinomial over words** $\varphi \sim \mathrm{Dir}(\beta)$
- **For each comment** $1, \ldots, D$
  - **Draw a multinomial over topics** $\theta \sim \mathrm{Dir}(\alpha)$
  - **For each word** w_{N_d}
    - **Draw a topic** $Z_{N_d} \sim \mathrm{Mult}(\theta_D)$ **with** $Z_{N_d} \in [1 \ldots K]$
    - **Draw a word** $W_{N_d} \sim \mathrm{Mult}(\varphi)$

Result Analysis

- **Keywords for different sentiments**

positive   neutral   negative
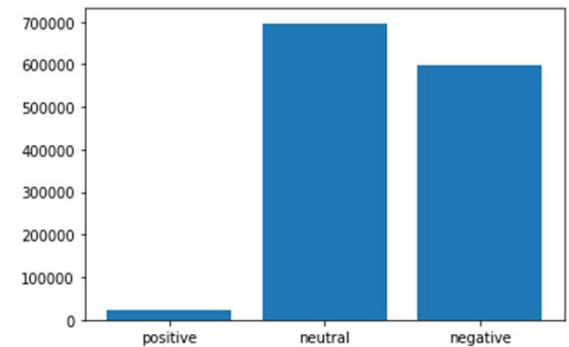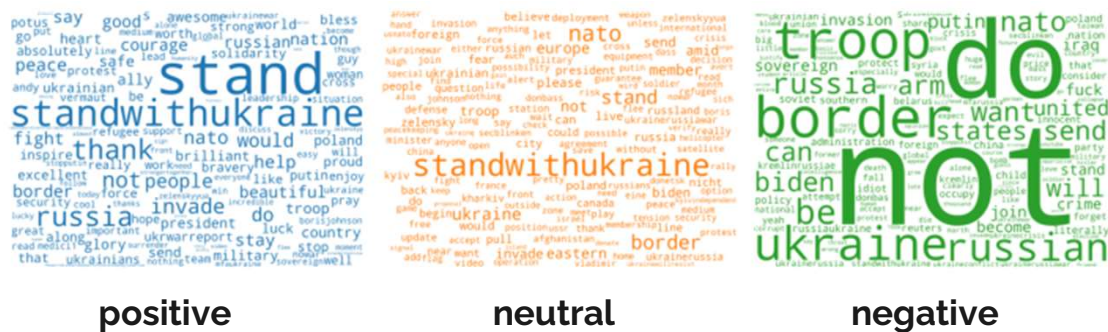
- **Keywords related to certain topics**

Biden   **NATO**   **Putin**   **Zelensky**   **Poland**
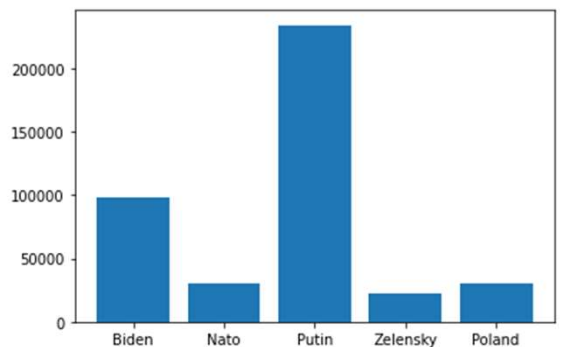
Result Analysis

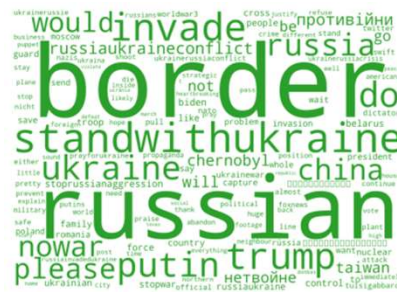- **Keywords for different dates**
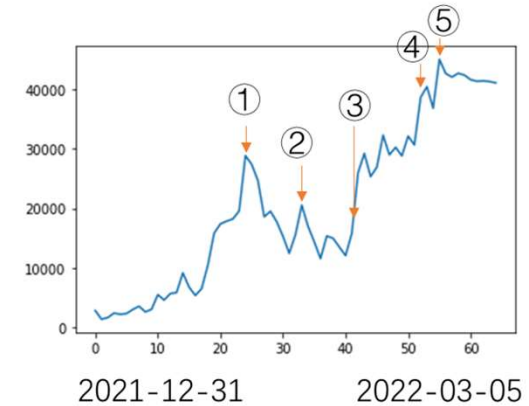

2022-01-24

2022-02-02

| ① | 2022-01-24 | NATO put troops on standby |
②	2022-02-02	Putin accused West of trying to draw Russia into war
③	2022-02-10	Russia and Belarus began 10 days of military maneuvers
④	2022-02-21	Russia recognized the independence of Donetsk People's Republic and Luhansk People's Republic
⑤	2022-02-24	Russia started the war

2022-02-10

2022-02-21

2022-02-24

Result Analysis

- **Keywords for different user locations**

Ukraine

Russia

USA

Germany

India
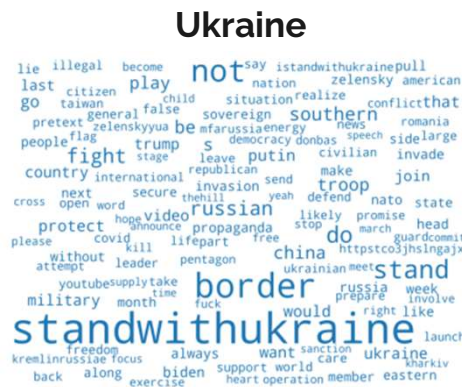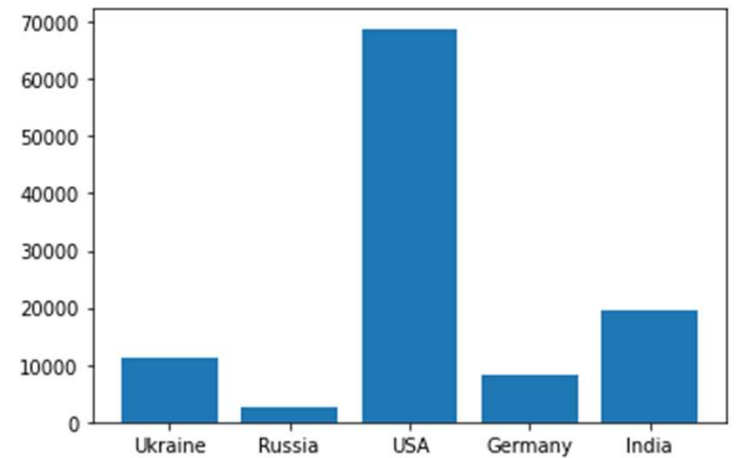
Visualization Demo

Visualization

The Russian and Belarusian Armies began Maneuvers in Belarus

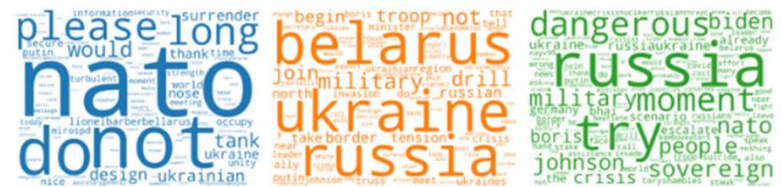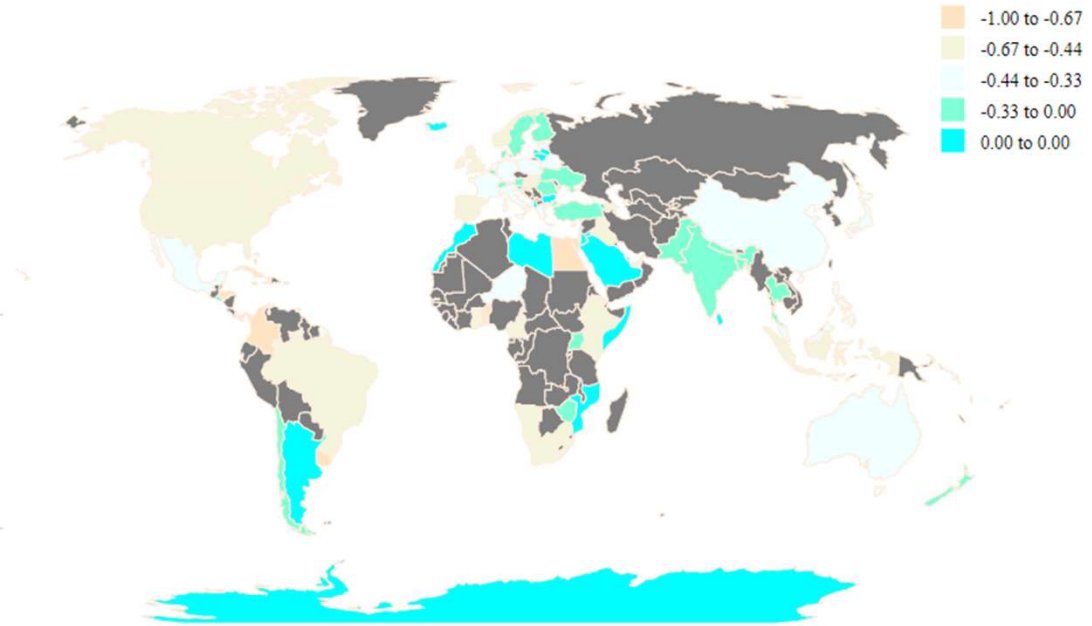**World** » **RUSSIA** | February 10, 2022, Thursday // 09:47

sputniknews.com

Legend:
- -1.00 to -0.67
- -0.67 to -0.44
- -0.44 to -0.33
- -0.33 to 0.00
- 0.00 to 0.00
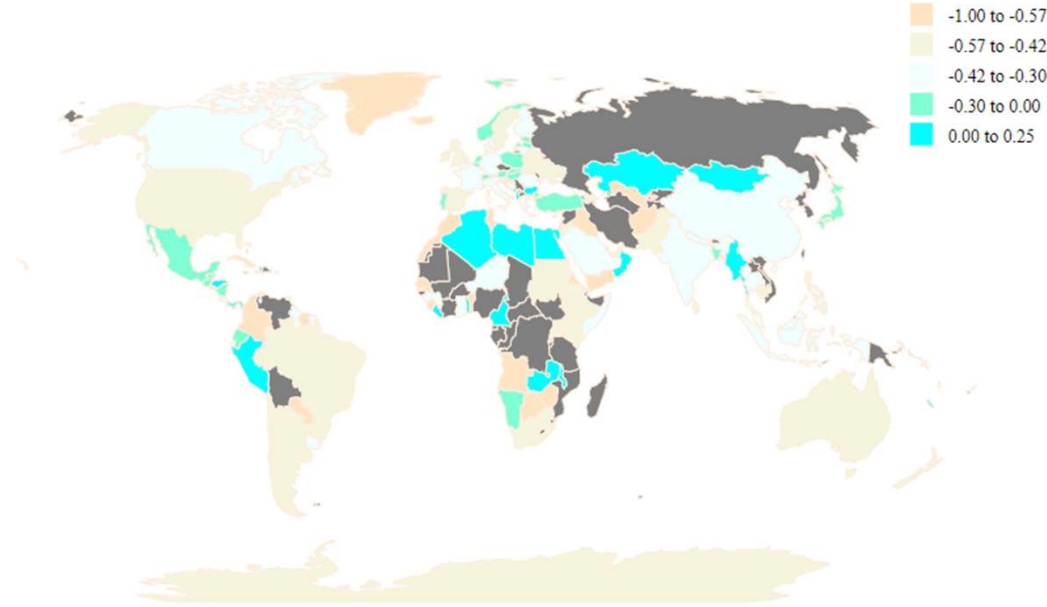
Visualization

Visualization

Public Opinion towards Russia-Ukraine War across the World

Select Date to view: 2022-03-03

Discussion & Conclusion

Conclusion

LDA does help for public opinion extraction.

- We successfully saw group opinions and its transition in events by date, countries, and senntimental expressions.

Discussion

As we finish the whole pipelines, there are still rooms for improvement and consideration.

- When the rumor flies, it could be hard for the lda model to catch keywords. - Inherent weakness
- Our model can only reflect public opinion on Twitter - Data Bias

Public Opinion towards Russia-Ukraine War across the World

Select Date to view: 2022-02-28

-1.00 to -0.67
-0.67 to -0.46
-0.46 to -0.33
-0.33 to -0.17
-0.17 to 0.00

Thanks! Q&A