2022 Spring SMOA

YouTube Political Channel Comment Analysis

Analyzing channel characteristics from different perspectives

using News Tornado and News Wawawa as examples

Po-Jen Hsieh

Research Question

Research Question

- Do different political commentary channels attract different topics of comments?
- Are there any unique characteristics in the comments made by suspected internet trolls?
- Can differences in stance be identified through word embeddings trained on different texts?
- Do word embeddings trained on different texts imply different latent political biases?

Exploratory Data Analysis

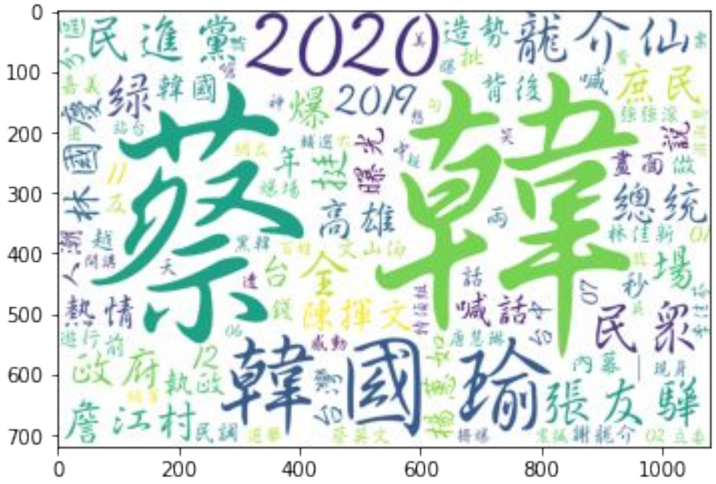
Exploratory Data Analysis

Dataset Overview

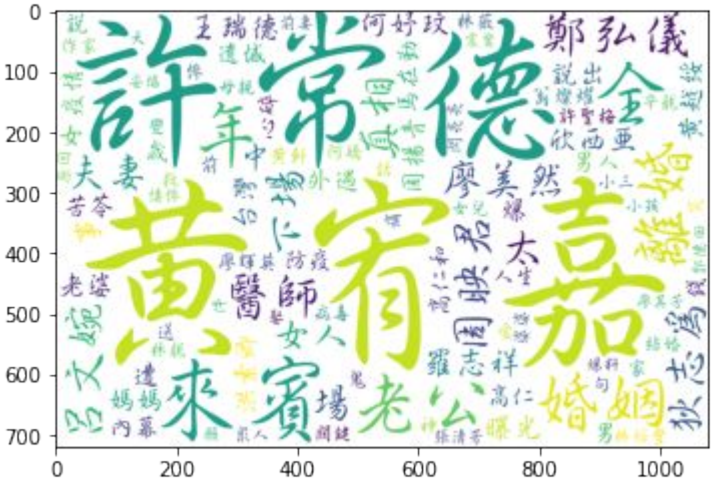
- Data counts: News Tornado 180,578 entries, News Wawawa 167,487 entries.
- Columns: videoId, commentId, authorDisplayName, textOriginal, likeCount, publishedAt.
- Time coverage:
  - News Tornado: from 2019-01-01 00:02:14 to 2020-08-05 06:27:10.
  - News Wawawa: from 2014-06-11 11:12:26 to 2020-07-18 13:43:05.

Exploratory Data Analysis

News Tornado



News Wawawa



Definition

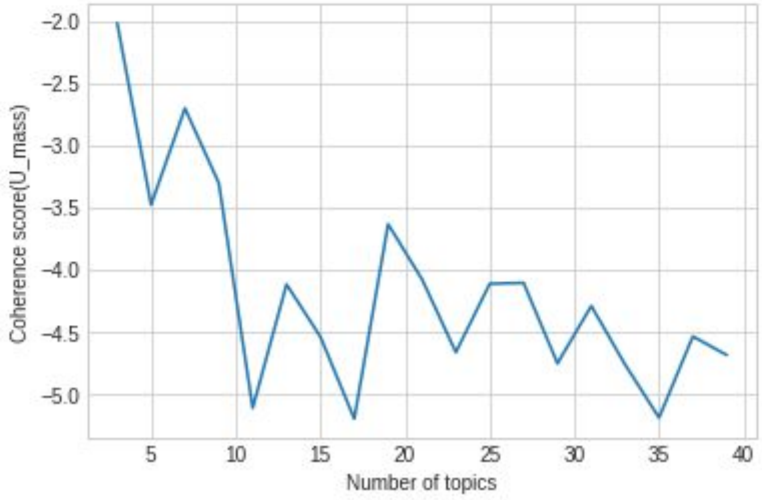
What defines a potential internet troll?

- Literature review: Active during weekdays, high popularity response, multiple IP locations.
- My naive approach: For both News Tornado and News Wawawa, calculate for each unique author:
    - Total number of comments
    - Average likes
    - Weekday commenting ratio
    - Classified as a suspected internet troll if meeting the following criteria:
        - Total number of comments above the 90th percentile
        - Average likes above the 90th percentile
        - Weekday commenting ratio over 85%

Topic Modeling

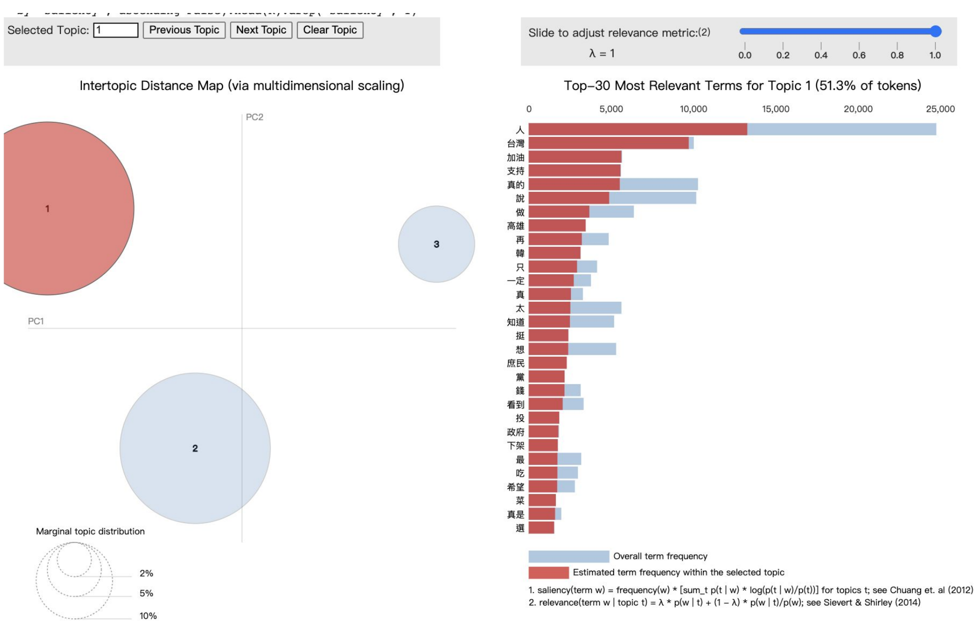
Topic Modeling—LDA

- Text: Includes all comments from News Tornado and News Wawawa
- Unit: By author
- Number of topics: 3



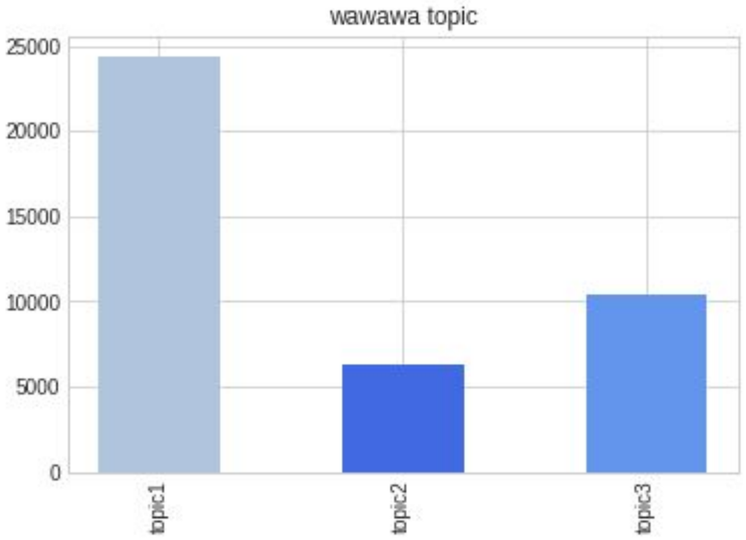
Topic Modeling—LDA

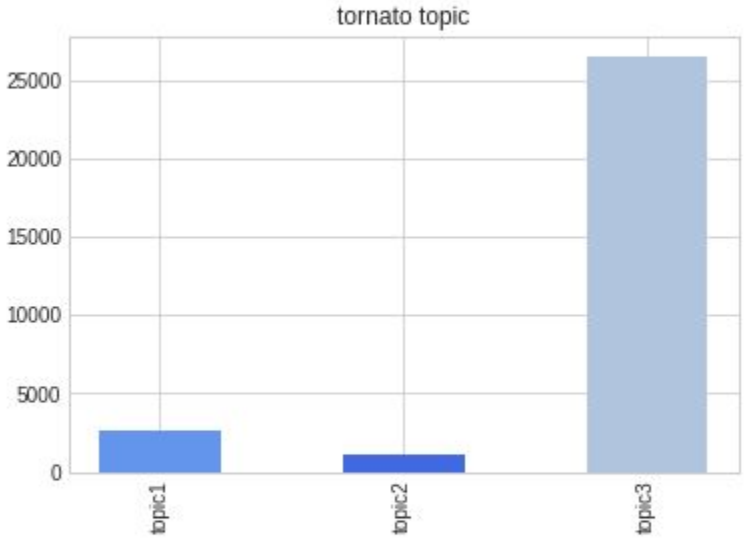
Topic Modeling—LDA



Do different political commentary channels attract different topics of comments?

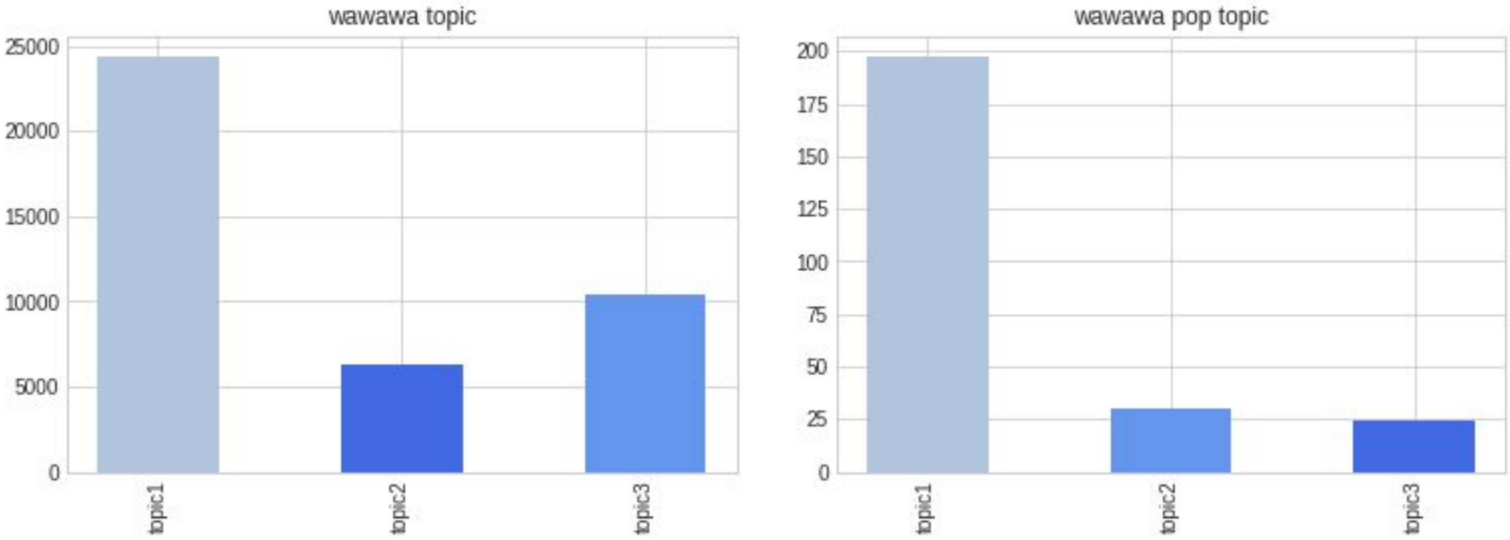
Topic Modeling—LDA



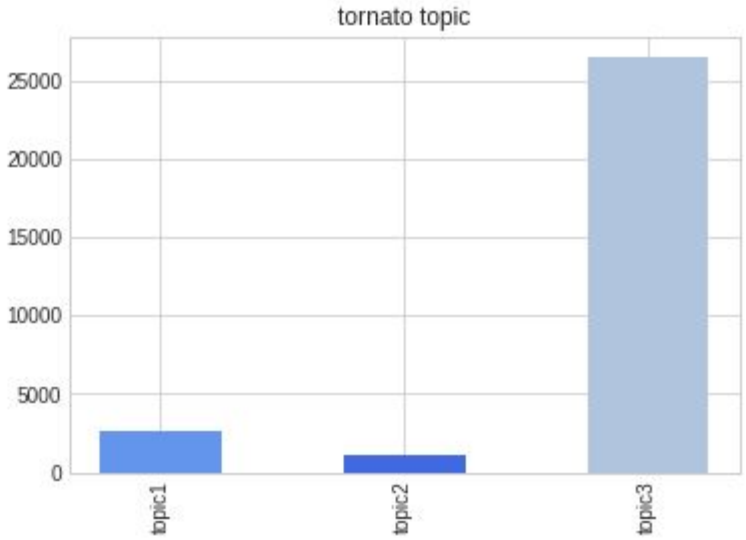


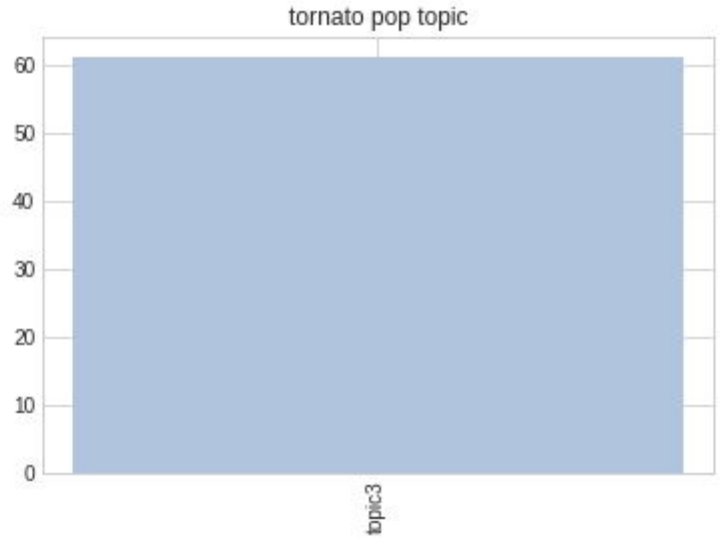
Are there any unique characteristics in the comments made by suspected internet trolls?

Topic Modeling—LDA



Topic Modeling—LDA





Word Embedding

Word Embedding

I trained the following 5 types of word2vec models with the texts below:

- Wikipedia Chinese (Backup from 2022/06/01)
- All comments from News Tornado
- All comments from News Wawawa
- Comments from suspected internet trolls on News Tornado
- Comments from suspected internet trolls on News Wawawa

Can differences in stance be identified through word embeddings trained on different texts?

Word Embedding

- The word2vec trained on Wikipedia data, showing the top 100 words most similar to Tsai Ing-wen and Han Kuo-yu.

Han(Wiki)

Tsai(Wiki)

Word Embedding

- The word2vec trained on all comments of News Wawawa, showing the top 100 words most similar to Tsai Ing-wen and Han Kuo-yu.

Han(Wawawa)

Tsai(Wawawa)

Word Embedding

- The word2vec trained on all comments of News Tornado, showing the top 100 words most similar to Tsai Ing-wen and Han Kuo-yu.

Tsai(Tornato)

Han(Tornato)

Word Embedding

- The word2vec trained on comments of suspected internet trolls on News Tornado, showing the top 100 words most similar to Tsai Ing-wen and Han Kuo-yu.

Tsai(Tornato pop)

Han(Tornato pop)

Do word embeddings trained on different texts imply different latent political biases?

Word Embedding

- USA: ['USA', 'Democratic Party', 'Republican Party', 'Washington D.C.', 'White House', 'Trump']
- China: ['China', 'CCP', 'Communist Party', 'Beijing', 'Zhongnanhai', 'Xi Jinping']
- Concepts: ['betrayal', 'evil', 'dictatorship', 'hypocrisy', 'threat', 'oppression', 'freedom', 'human rights', 'rule of law', 'equality']
- Using a method similar to Garg et al. (2017), compare each concept to see whether it is closer to China or the USA

Word Embedding

	Wikipedia	News Wawawa	News Tornado
betrayal	China	China	China
evil	US	China	US
dictatorship	US	China	China
hypocrisy	US	China	N/A
threat	US	US	US
oppression	China	China	China
freedom	China	China	China
human rights	US	US	China
rule of law	China	China	China
equality	China	China	China

Thanks!