Data Visualization

Huanrui CHEN

April 28, 2024

My research question is "How does digital surveillance affect public political discussions on the Internet?" The hypothesis I would to discuss is that if the repressive policies and the sense of being monitored will change citizens' behavior, causing them to decrease the frequency of engaging in political discussions online and to use more moderate or indirect language, even if these discussions do not directly violate any laws. Or policies that criminalize dissent and protest do not necessarily affect criticism of the government in anonymous online spaces, and may even fuel discontent.

To address this, I plan to collect data from LIHKG, the largest online political platform in Hong Kong, and use the ChatGPT3.5 Turbo API for data annotation. Subsequently, I will conduct a time-series analysis and sentiment analysis of the data to observe the trends in online political discussions in Hong Kong before and after the enactment of the National Security Law in order to discuss the aforementioned hypothesis and questions. Since both time-series analysis and sentiment analysis are closely related to the timing of posts, and I have not yet collected a complete dataset (spanning seven years) at this stage, I will use simulated data to complete the visualization of the time-series analysis and sentiment analysis.

I manually annotated 41 posts from LIHKG and coded them to check the reliability of the annotations. The annotations include: 1. Sentiment (supportive of the government, neutral, opposing the government) 2. Target (Hong Kong government, mainland Chinese government, unrelated) 3. Degree of aggressiveness in wording (radical, neutral, moderate) 4. Whether non-direct, like ironic or metaphorical expressions were used (Yes, No) 5. Involvement in supporting which of the following topics (Separatism, Subversion, Terrorism, Colluding with Foreign Forces, None).

For the labels and content of the data, please see the attachment CSV document or visit the link: https://github.com/hchen0628/MACS30200/blob/main/Data%20428.xlsx

Dimension	**Data Volume**	**Accuracy (%)**
Sentiment (Support, Neutral, Against Govt)	41	82.93
Target (HK Govt, Mainland Govt, Other)	41	85.37
Degree of aggressiveness in wording (Radical, Neutral, Mild)	41	100.00
Indirect Expression (Yes, No)	41	95.12
Theme Involvement (Separatism, Subversion, Terrorism, Colluding with Foreign Forces, None)	41	92.68

Table 1: Comparison of annotation methods by data volume and accuracy.

Overall, I think ChatGPT 3.5 Turbo performs well in estimating sentiment and degree of aggressiveness in wording, likely because these measurements do not heavily rely on contextual knowledge. However, despite seeming accurate in targeting, indirect expression, and theme involvement, this could be due to a small dataset and the problematic distribution creating bias. Among the forty-one test data currently used, only two instances were considered to violate the national security law in judicial practice, but they were not correctly labeled by ChatGPT 3.5 Turbo as pertaining to one of the four main themes involved in suspected violations of the national security law. That is to say, ChatGPT 3.5 is entirely incorrect in theme labeling. It also defined a post discussing Chinese residents publicly killing their child as supporting terrorism. In contrast, ChatGPT 4.0 was able to correctly label the corresponding post with the appropriate theme, likely due to richer training data, including judicial practice related to the national security law. Therefore, I plan to use ChatGPT 4.0 to label this portion of the data independently. In terms of targets, if the news only mentions a government official's name or just the word "government," it might not accurately determine whether the reference is to the Hong Kong government or the mainland government. This issue could potentially be resolved with more appropriate prompts, such as by default referring to the "government" as Hong Kong government. While ChatGPT 3.5 Turbo performs well in recognizing metaphors, it struggles with detecting sarcasm, a challenge also faced by ChatGPT 4.0. When a post involves derogatory and exaggerated techniques, ChatGPT can identify the sarcasm. However, it struggles with posts that depend on contextual knowledge, like "The whole world has fake democracy, only China has real democracy." My current strategy is to estimate whether a post is sarcastic based on the number of likes and dislikes it receives, because satirical posts can attract a large number of dislikes, but I also worry this method might introduce bias.

Since the data contains multiple labels, I am going to use a 3D image to show the sentiment analysis, including time, percentage of sentiment and type of post. Below is the image generated using the simulated data:

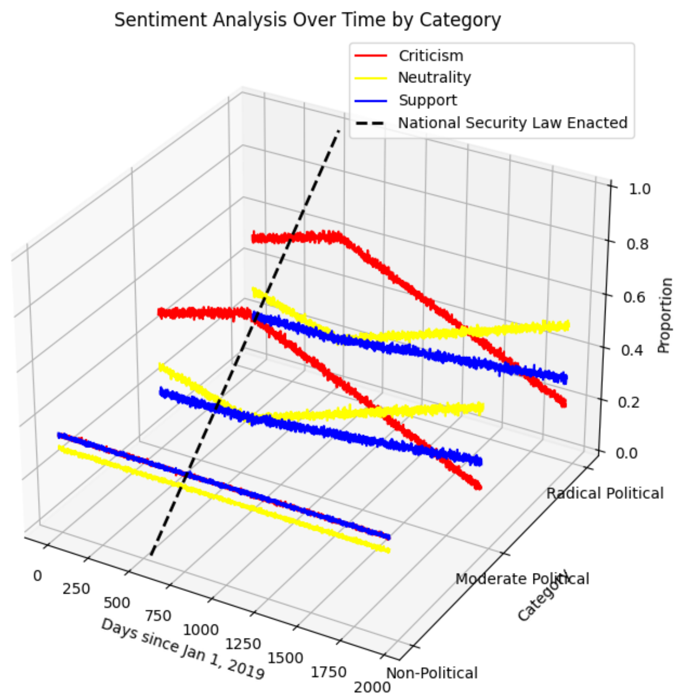

Figure 1: Sentiment Analysis

This image may also be useful for comparing frequencies:

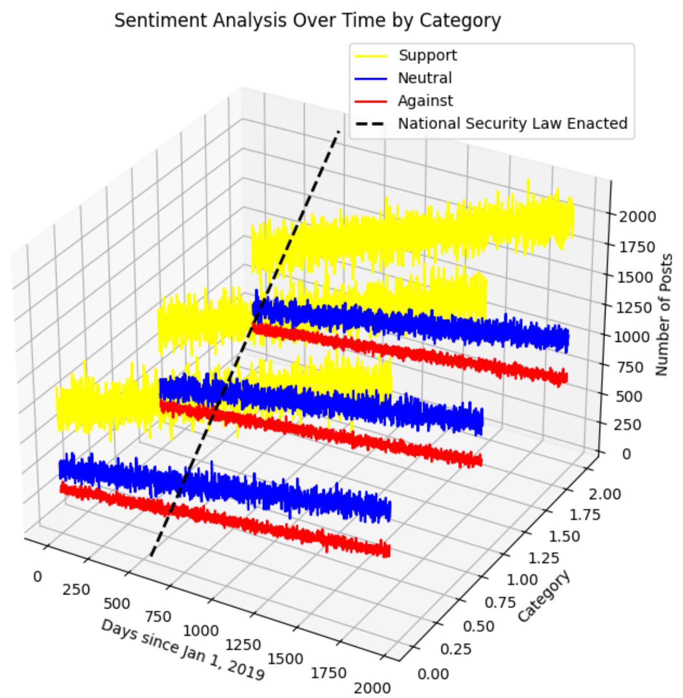

Figure 2: Frequency

Potential image for time series analysis:

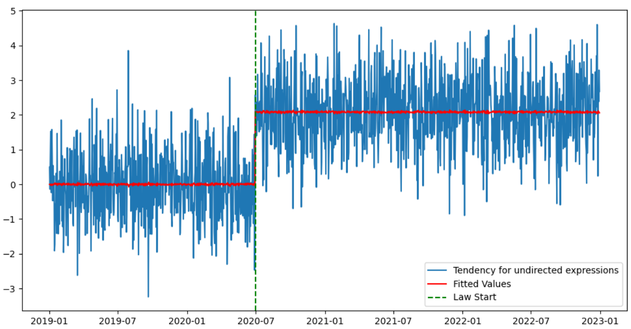

Figure 3: Time series analysis