COVID-19 discussions in social media: cross-country comparison

Name: Kiki Ferawati

Laboratory's name: Social Computing

Supervisor's name: Professor Eiji Aramaki

Abstract

The COVID-19 pandemic is a worldwide phenomenon that disrupted the world. As a significant source of information during the pandemic, social media provided insights and interesting discourses about events that occurred during this timeframe. Various topics related to COVID-19 were discussed on Twitter (now X), one of the most popular platforms for social media analysis. Different countries adopted different measures in handling the pandemic, which is reflected in social media posts from these countries. This dissertation aims to provide a cross-country comparison of responses to preventive measures for COVID-19, specifically vaccination and mask-wearing policies.

In the first part of the study, we analyzed data related to COVID-19 vaccinations in Japan and Indonesia, two Pacific nations with large Twitter-using populations. We observed differences in the ratio of side effects reported in official reports versus those reported on Twitter. We also discussed the potential of using Twitter as a surveillance tool for monitoring COVID-19 vaccine side effects, noting that fever appears to be over-represented.

As one of the most widely adopted policies, mask-wearing generates discussions in countries with different societal values, such as the United States and Japan. This study employed a synchronized data annotation approach to address the multilingual challenge in English and Japanese, resulting in annotation guidelines covering both languages for mask-related cases. The results were extended by building a classifier to annotate the entire dataset, providing a comprehensive understanding of the stance on masking and mask-wearing across all states and prefectures in the observed areas.

The correlation analysis between social media results and demographic data in each country assesses whether the opinions found on social media are consistent with other existing studies about public responses to COVID-19. This study demonstrates that social media data can capture real-world phenomena, such as the relationship between education level and income to the stance on mask-wearing policy, and highlights that this topic is an interesting field for exploration in natural language processing research.