2023 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)

Emotion Mapping: Sentiment Analysis using Emoji in Twitter Data from Japan in the COVID-19 Era

Ahmed Almohanadi
Tokyo Metropolitan University
Hino, Tokyo, Japan
ahmed-almahanadi@ed.tmu.ac.jp

Shohei Yokoyama
Tokyo Metropolitan University
Hino, Tokyo, Japan
National Institute of Informatics
The University of Tokyo
shohei@tmu.ac.jp

***Abstract*—This research explores the intricate relationship between emojis in Twitter interactions and sentiment interpretation in Japan. Emojis, as part of the digital language lexicon, have become essential in social media communication, especially on platforms like Twitter. However, integrating emojis into sentiment analysis has yet to be heavily explored, presenting an opportunity to better understand localized digital sentiment expression. This paper analyzes a data set within the COVID-19 Era of Twitter messages from Japan, focusing on emojis and the sentiments they convey. An established model was utilized to process Twitter text data and estimate their sentiment analysis, reflecting a more accurate portrayal of users' emotional states. We find a compelling link between emoji usage and the sentiment expressed. Emojis were shown to provide critical sentiment indicators often absent or ambiguous in the text alone. Certain emojis were identified as having a consistent correlation with particular sentiments.**

***Index Terms*—Emoji, Sentiment, Big Data, Twitter, Data Analysis**

I. Introduction

The COVID-19 pandemic caused havoc worldwide due to its impact and destruction of the usual social order. Governments imposed containment measures, including lockdowns, restrictions on movement, and quarantines. The digital network became the primary source of information and expression. In particular, social media use, primarily through shared messages, became prevalent as people expressed their sentiments online and noticed the rise of a new set of emojis. This paper aims to employ sentiment analysis techniques and emojis in Twitter data from Japan during the COVID-19 era to map and understand the emotional responses of the Japanese people. The analysis provides valuable insights into linking sentiment, emojis, and the changes through the pandemic. It sheds light on the changes and impacts of the pandemic on the behavior of social media and texting overall.

Emojis are digital expressions that portray simple but eye-catching artwork comprising facial expressions. According to Kaur et al. [7], emojis can be used to portray emotional reactions among social media users who understand the emoji similarly. The images facilitate non-verbal communication cues on social media that enhance user exchange and interaction. Since their inception around 2005, emojis have become popular tools for expressing joy, fear, and celebration, among others (Bai et al., [3] and Li et al. [9] indicate that emojis are considered a particular form of digital communication as they have a unique social attraction. They have been developed out of an extra-linguistic origin, which is not considered part of the typical vocabulary of any ordinary semantic. Therefore, emojis occupy a unique place in digital communication as they exemplify shared culture at a more granular level. For instance, the emoji that symbolizes a smiley face is commonly used to show a feeling of friendliness or joy. Das [4] notes that the general usage of emojis varies depending on the discussion that people are engaged in at a particular time. The prominence of emojis applies across many parts of the world, from the West to the East. According to Dyer and Kolic [5], the state of California in the United States has been found to have a high concentration and multiplicity of emoji usage, especially along the coastal area. The trend is connected to the high population in the state's central coastal cities, including San Diego, Los Angeles, Palo Alto, San Francisco, and others (Kejriwal et al., [8]). The high usage is also correlated to the linguistic and cultural diversity of the area, which promotes greater exchange and interaction.

II. Related Work

Empirical studies show that social media was a significant source of COVID-19-related data worldwide. Plakhotnik et al. [18] refers to a cross-sectional research study drawing upon college students in Germany to assess their reliance on social media as the source of COVID-19-related information. They found that nearly 38 percent of the learners turned to social media frequently or occasionally as they hunted for information about the pandemic and allied matters (Jiang et al., [6]). Neely et al. [11] conducted a similar research study where they surveyed 1,003 adults based in the United States. Their findings revealed that slightly over three-quarters of respondents depended on social media to gather information about the COVID-19 pandemic. In addition, about 64 percent

of the respondents considered the information accurate, meaning that people should have taken the additional step of verifying the information with a healthcare professional. Neely et al. [11] also note that 59 percent visited social media at least once per week to scan through COVID-19-related information. The findings reveal people's trust in social media as the primary source of information, notwithstanding the problem of fake news and misinformation. Furthermore, research has established that people get to express their mental state on social media. Marzouki et al. [10] analyzed several Facebook posts and found that people use the platform to express negative sentiments, such as depression, fear, anxiety, and stress. The majority of the participants mainly used to portray these negative sentiments. According to Tran and Matsui [16], people increasingly used terms associated with negative feelings and irritation, mainly on Twitter. The findings also revealed that the participants manifested signs of depression as opposed to those who did not use social media. Additionally, several other studies have shown that social media users tend to react adversely toward the spread of COVID-19. Salvi et al. [12] analyzed data from Twitter in Japan on social sensations toward COVID-19 for three months in 2020, from February to April 2020. They established that most of the people reported a feeling of fear regarding the infection. Kaur et al. [7] analyzed data from Twitter in 2020 in February, May, and June (Shi et al., [13]). They established that most of the tweets portrayed negative sentiments. Dyer and Kolic [5] also analyzed Twitter data and found proof of psycho-physical distress. In particular, they found that users were more and more fixated on thinking of death, and there was reduced emotional expression. The research shows that social media, including Twitter, can cause mental illness due to heightened anxiety, stress, and infectious panic (Tran and Matsui, [15]). The situation was worsened by misinformation on COVID-19 and fake news that tended to overstate the perceived threat. Findings on people's reactions through social media indicate that there can be a comparable long-term trend across different regions. Tran and Matsui [16] compared social media reactions across six countries: South Korea, Indonesia, India, Germany, and Thailand. They found a similar pattern based on social media reactions. The similarity in the findings was further supported by the consideration that the countries have a high Twitter usage. The observed trend revealed a dramatic increase in Twitter usage during the COVID-19 pandemic (Suntwal et al., [14]). However, a significant decline was observed afterward, even as each country encountered fresh waves at various times, leading to a surge in reactions on social media. However, the general observation was a general decline in social media reactions in 2022 across all countries (Kaur et al., [7]). Therefore, the general trend was that social media reaction toward the COVID-19 pandemic was exceptionally high at the onset and slowly faded away with every passing of time.

III. Method

The data for this research study emanates from social media, particularly Twitter, in Japan. The platform is the preferred source because of its popularity in the country, promotes a high level of public interaction, and can easily attract participation from unfamiliar people. According to Tran and Matsui [16], data indicated that Japan has the second-highest number of Twitter users after the United States, with more than 50 million users. The platform also offers a more accessible option for collecting and analyzing historical data. The data collected for the research comprised using twitter analysis tool[1] for the period between the start of the pandemic and two years of data. The data was collected in a quarterly manner. In both English and Japanese and must contain one of the recent Emojis that is related to the coronavirus pandemic and has been observed to increase in usage [🧼, 🦠, 🩹, 💉, 😷, 🤢, 🤒, 🤧] [1]. The data has been processed through a cleaning library to remove Twitter's special characters and enhance the accuracy of processing the tweets at a later stage. Then, the data will apply a state-of-the-art integrated platform for natural language processing trained by Twitter data designed by Cardiff University [17].

IV. Results

The continuing rise of emojis, along with the number of tweets, indicates that the effect of the pandemic was and still has affected the usage of Emojis in Tweet linguistics and daily texting. The data provides an interesting sentiment indicator and presents very peculiar trends. The analysis also shows a decrease in negative values and a rise in neutral and positive values over the years. An isolated analysis of each Emoji has been conducted with few findings, but it could lead to future research.

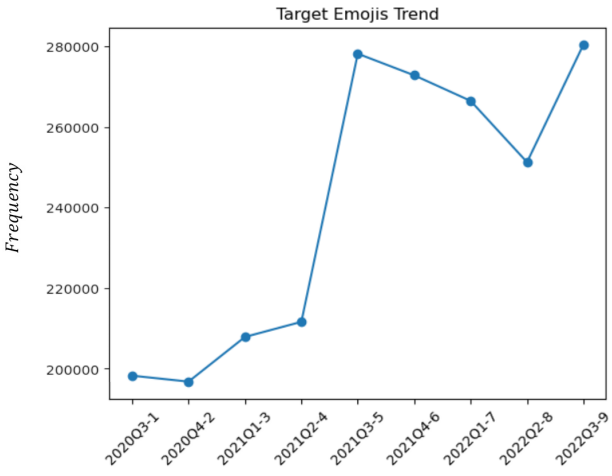

Fig. 1. Target Emojis trend over two years

[1]Twitter data was collected through twitter analysis tool provided by NTT DATA Japan's "Nazuki No Oto"

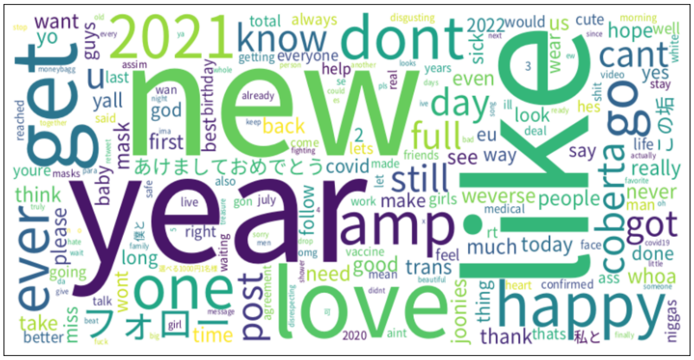

Fig. 2. Word Cloud of associated words with Emojis

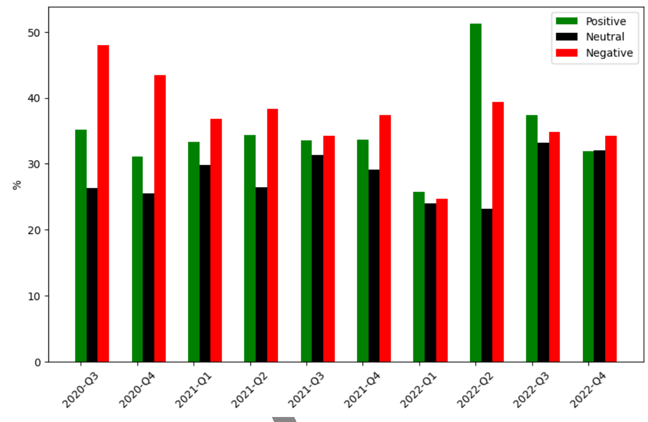

Fig. 3. Sentiminal Values of the collected sample

The following emojis, which represent a sneezing face (🤧), have average positive feedback; the following emojis with mask and nauseated face have negative feedback over the two years, which is understandable (🤢 ,😷), last the following emojis (💉) have neutral value. The following emojis deficient count value which is unreliable for the study and requires a different approach to study them (🧼 , 🦠 , 🩹) The emojis that were singled out [(🤧)] for having a positive correlation with positive sentiment value, having 42% average positive feedback, an indicator that shows a reduction in negativity toward the sneezing action, which was a source of fear when the pandemic began. The syringe has a high neutral value but shallow and equal positive and negative values, which hints at the unsettled feelings toward the vaccination. The most common words that are related to the studied Emojis shown in Fig 2 present an interesting graph, having the words "love" , "happy" and "hope" in association with corona-related Emoji, a lexical that represents an entirely positive sentiment, further investigation into these words and associated emojis reveals vastly unstable trends over the years, the average value is equal. However, when digging deeper into the news and society's condition during those unstable trends, society's feedback can be represented if the correlated Emoji is attached to a specific topic.

V. Discussion

The findings show that emojis can be used to analyze public social media sentiment that is associated with a particular topic , and it is able to gauge the sentimental mood of the society and notice the shift in value and time. This suggests the possibility of monitoring public emotion for any future events or incidents. However, more work needs to be done in describing the different variations and types in meaning of the emoji. This makes it highly challenging to correctly apply, analyze, and get an accurate sentimental analysis of social media users. particularly toward COVID-19 and in general. During the pandemic, Emojis on social media still provided a critical basis for understanding and predicting people's emotions. Even though there is a decline in the use of emojis in expressing sentiments related to the COVID-19 pandemic, there is a need for further research to establish their usage in facilitating communication and interaction among people. In the future, different aspects may be approached within tweet metadata, for example, linking tweet Geodata with Emojis or digging deeper into the lexical content of the tweet itself.

References

[1] Broni, K. (2020) Spread of the coronavirus emoji, Emojipedia. Available at: https://blog.emojipedia.org/spread-of-the-coronavirus-emoji/

[2] Arora, A., Chakraborty, P., Bhatia, M. P. S., and Mittal, P. (2020). Role of emotion in excessive use of twitter during COVID-19 imposed lockdown in India. Journal of Technology in Behavioral Science, 6, 370-377.

[3] Bai, Q., Dan, Q., Mu, Z., and Yang, M. (2019). A systematic review of emoji: Current research and future perspectives. Frontiers in Psychology, 10.

[4] Das, A. (2020). How has the coronavirus (COVID-19) pandemic affected global emoji usage? Journal of Human Behavior in the Social Environment, 31, 425 – 434

[5] Dyer, J., & Kolic, B. (2020). Public risk perception and emotion on Twitter during the Covid-19 pandemic. Applied Network Science, 5(1).

[6] Jiang, J., Thomason, J., Barbieri, F., & Ferrara, E. (2022). Geolocated social media posts are happier: Understanding the characteristics of check-in posts on twitter.

[7] Kaur, S., Kaul, P., & Zadeh, P. M. (2020). Monitoring the dynamics of emotions during COVID-19 using twitter data. Procedia Computer Science, 177, 423–430.

[8] Kejriwal, M., Wang, Q., Li, H., & Wang, L. (2021). An empirical study of emoji usage on Twitter in linguistic and national contexts. Online Social Networks and Media, 24,

[9] Li, M., Chng, E., Chong, A. Y. L., & See, S. (2019). An empirical analysis of emoji usage on Twitter. Industrial Management & Data Systems, 119(8), 1748–1763.

[10] Marzouki, Y., Aldossari, F. S., & Veltri, G. A. (2021). Understanding the buffering effect of social media use on anxiety during the COVID-19 pandemic lockdown. Humanities and Social Sciences Communications, 8(1).

[11] Neely, S., Eldredge, C., & Sanders, R. (2021). Health information seeking behaviors on social media during the COVID-19 pandemic among American social networking site users: Survey Study. Journal of Medical Internet Research, 23(6), e29802.

[12] Salvi, C., Iannello, P., Cancer, A., McClay, M., Rago, S., Dunsmoor, J. E., & Antonietti, A. (2021). Going viral: How fear, socio-cognitive polarization and problem-solving influence fake news detection and proliferation dur-

ing COVID-19 pandemic. Frontiers in Communication, 5. https://doi.org/10.3389/fcomm.2020.562588

[13] Shi, W., Zeng, F., Zhang, A., Tong, C., Shen, X., Liu, Z., & Shi, Z. (2022). Online public opinion during the first epidemic wave of COVID-19 in China based on Weibo data. Humanities and Social Sciences Communications, 9(1). https://doi.org/10.1057/s41599-022-01181-w

[14] Suntwal, A., Brown, S. J., and Brandimarte, L. (2021). Pictographs, Ideograms, and Emojis (PIE): A Framework for Empirical Research Using Non-verbal Cues. https://doi.org/10.24251/hicss.2021.771

[15] V. Tran and T. Matsui, "Tweet Analysis for Enhancement of COVID-19 Epidemic Simulation: A Case Study in Japan," in IEEE Transactions on Public Health, vol. 5, no. 2, pp. 135-145, 2023.

[16] V. Tran and T. Matsui, "COVID-19 case prediction using emotion trends via Twitter emoji analysis: A case study in Japan," in IEEE Transactions on Public Health, vol. 7, no. 3, pp. 215-225, 2023.

[17] Camacho-Collados, J. et al. (2022) TWEETNLP: Cutting-edge natural language processing for social media, arXiv.org. https://arxiv.org/abs/2206.14774.

[18] V. Tran and T. Matsui, "Tweet Analysis for Enhancement of COVID-19 Epidemic Simulation: A Case Study in Japan," in IEEE Transactions on Public Health, vol. 10, no. 2, pp. 135-145, 2022.

[19] Plakhotnik, M.S. et al. (2021) The perceived impact of COVID-19 on student well-being and the mediating role of the university support: Evidence from France, Germany, Russia, and the UK, Frontiers.