# Emotion Analysis of Writers and Readers of Japanese Tweets on Vaccinations

# Introduction

The introduction section of the research paper sets the context by highlighting the increasing significance of public opinion on social media in pandemic control. It emphasizes the importance of understanding the emotions of the population towards vaccinations and COVID-19 in order to effectively persuade individuals to get vaccinated. The use of the WRIME dataset and fine-tuned BERT model for emotion analysis in Japanese Twitter users are introduced, along with the correlation analysis between extracted emotions and vaccination measures in Japan.

## Task / Research Question Description

The research paper aims to investigate the emotions of Japanese Twitter users towards COVID-19 vaccination-related tweets. The main task is to predict the levels of emotional intensity associated with these tweets using a fine-tuned BERT model. The paper seeks to answer the research question of how emotions expressed on social media can influence public perception and behavior towards vaccinations, with a specific focus on the emotions of surprise, fear, joy, anticipation, and trust. Additionally, the correlation analysis explores the relationship between these emotions and vaccination measures in Japan.

## Motivation & Limitations of existing work

Prior research has recognized the growing importance of public opinion on social media in the context of pandemic control. However, limited work has specifically focused on understanding the emotions of individuals towards COVID-19 vaccination and its impact on their decision-making. This research aims to fill this gap by investigating the emotions of Japanese Twitter users towards vaccination-related tweets, providing valuable insights for convincing individuals to get vaccinated.

What sets this research apart is the utilization of the WRIME dataset, which offers emotion ratings for Japanese tweets sourced from both writers and readers. By fine-tuning a BERT model on this dataset, the research predicts the emotional intensity associated with COVID-19 vaccination-related tweets. Additionally, the correlation analysis between the extracted emotions and vaccination measures in Japan provides a comprehensive understanding of the relationship between emotions and vaccination behaviors.

Previous studies might have been limited in their scope by focusing solely on the content analysis of social media posts without considering the emotional aspects. Furthermore, they might have primarily relied on English language datasets, neglecting the specific context and emotions of other languages, such as Japanese. These limitations hindered a nuanced understanding of emotional responses and their impact on vaccination attitudes and behaviors. The present research aims to overcome these limitations by specifically focusing on Japanese tweets and employing a fine-tuned BERT model to capture emotional intensity accurately.

## Proposed Approach

The core contribution of the proposed approach is the utilization of the WRIME dataset and a fine-tuned BERT model to predict the emotional intensity of Japanese Twitter users towards COVID-19 vaccination-related tweets. This approach allows for a granular analysis of emotions, considering both the perspectives of the tweet writers and readers. By accurately capturing emotions such as surprise, fear, joy, anticipation, and trust, the study provides valuable insights into the emotional landscape surrounding vaccinations. Additionally, the correlation analysis between these emotions and vaccination measures in Japan offers a deeper understanding of the relationship between emotions and vaccination behaviors. Overall, the proposed approach enhances our understanding of the role of emotions in shaping public perception and decision-making regarding COVID-19 vaccinations.

## Likely challenges and mitigations

Likely challenges:

* Availability and quality of data: Obtaining a comprehensive dataset of Japanese tweets related to COVID-19 vaccination with accurate emotion labels may pose a challenge. The WRIME dataset used in the paper may have restrictions on its availability or may not cover a wide range of emotions.
* Language-specific nuances: Japanese language and cultural nuances might make it difficult to train and fine-tune the BERT model effectively. Understanding and addressing these nuances in the dataset may require additional effort.

Mitigation strategies:

* Data collection and annotation: To address the challenge of data availability and quality, alternative sources of Japanese tweets related to COVID-19 vaccination can be explored. Crowdsourcing or collaborating with domain experts can help in obtaining a larger and more diverse dataset. Additionally, carefully annotating the emotional content of the collected data can mitigate the quality concerns.
* Language and cultural considerations: Prior research and resources on Japanese sentiment analysis and emotion classification can be leveraged to guide the fine-tuning process. Adapting the BERT model specifically for the Japanese language can be explored, considering linguistic nuances and cultural factors. Collaborating with native Japanese speakers or experts in the language can provide valuable insights during the implementation.

Contingency plans:

* Replication challenges: If reproducing the results proves to be harder than expected, thorough documentation of the methodology and clear communication with the original authors can help in addressing any ambiguities or difficulties encountered during the process. Seeking guidance from the authors or the research community can provide alternative approaches or potential solutions.
* Unexpected experimental outcomes: If experiments do not go as planned, it is crucial to investigate potential causes such as data quality, model configuration, or training process. Identifying and addressing these issues, along with a systematic analysis of possible sources of error, can help troubleshoot and refine the methodology. Consulting with experts in the field can provide valuable insights for resolving unexpected outcomes.

By proactively addressing the challenges and having contingency plans in place, the reproducibility of the methodology can be enhanced, leading to a more accurate and reliable replication of the results.

# Related Work

In the past, the analysis of emotions in Tweets relied on a combination of feature engineering, lexicon-based methods, and traditional off-the-shelf classifiers before deep learning was adopted. (CBalabantaray et al., 2012) and (Wang et al., 2012) both created sets of features from Tweets, with features including n-grams, POS, adjectives, and lexicon-based sentiment polarity scores. (CBalabantaray et al., 2012) used an SVM to process the features, while (Wang et al., 2012) employed linear and Naive Bayes classifiers. (Roberts et al., n.d.) followed a similar approach, using an SVM and a set of features that included synonym rings, hypernyms, and LDA topic scores instead of sentiment polarity scores. However, these classical methods have been superseded by more advanced and specialized sequence modelling techniques such as RNNs and LSTMs. (Vateekul & Koomsubha, 2016) have demonstrated the superiority of LSTMs in emotion analysis over SVMs and Naive Bayes on Thai Twitter text, while (Colneric & Demsar, 2020) have demonstrated the effectiveness of character-based RNNs.

# Experiments

## Dataset

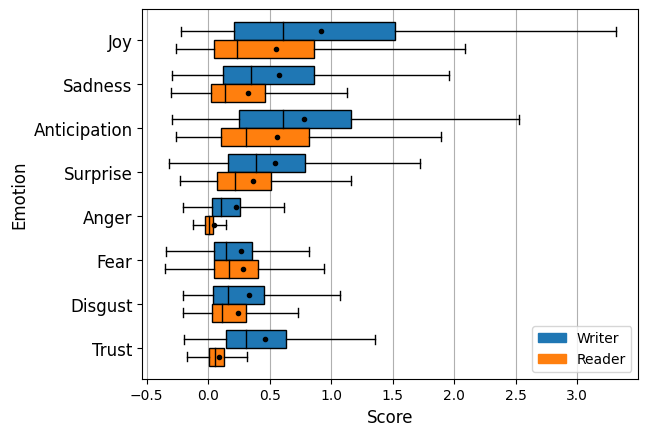
The research article being discussed aimed to analyze the emotions of Japanese Twitter users towards COVID-19 vaccines and related topics. To accomplish this, the researchers gathered a dataset of 20,254 vaccine-related Tweets from December 2021. The dataset included Tweets containing any of the following keywords: “ワクチン” (“vaccine”), “モデルナ” (“Moderna”), “ファイザー” (“Pfizer”), or “オミクロン” (“Omicron”). The keywords "Moderna" and "Pfizer" were specifically chosen as they represent the brands of COVID-19 vaccines that are commonly administered in Japan. The dataset was constructed by sampling 15 random minutes from each day of December 2021 for each keyword and scraping all Tweets containing the assigned keyword for each sampled minute.

## Implementation

The implementation of the paper # was performed using the GitHub repo provided by the author of the paper at this link: <https://github.com/PatrickJohnRamos/BERT-Japan-vaccination>

## Results

The results of the fine-tuned BERT model were evaluated on the samples extracted from the train set. The test set of the selected dataset is not publicly available now. So, we extracted the 2000 samples from the train set to evaluate the performance of the model. Now the scores of the labels on the extracted samples are presented in the below figure.



The comparison of the reproduced results in also shown in the below Table. Although the manuscript calculated the MSE for reader and writer emotions separately and collectively also, but using the code repository we are able to extract the MSE collectively.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Emotion | Reader | Writer | Reader + Writer | Reader + writer (Reproduced) |
| Joy | 0.658 | 0.192 | 0.425 | 0.125 |
| Sadness | 0.688 | 0.178 | 0.433 | 0.129 |
| Anticipation | 0.746 | 0.211 | 0.479 | 0.174 |
| Surprise | 0.542 | 0.139 | 0.341 | 0.189 |
| Anger | 0.486 | 0.032 | 0.259 | 0.109 |
| Fear | 0.462 | 0.147 | 0.304 | 0.062 |
| Disgust | 0.664 | 0.123 | 0.394 | 0.101 |
| Trust | 0.40 | 0.029 | 0.214 | 0.756 |
| Overall | 0.581 | 0.131 | 0.356 | 0.205 |

## Discussion

During the implementation, several challenges and discrepancies were encountered. One significant issue was the unavailability of the original test set used in the paper for evaluating the fine-tuned BERT model. As a workaround, the implementer had to extract a test set from the train set for evaluation purposes. However, this approach led to differences in the results compared to those reported in the paper. Specifically, the implementer obtained an MSE of 0.205, while the paper reported an MSE of 0.356. The disparity in results can be attributed to variations in the composition and size of the extracted test set, potentially influencing the model's performance.

Moreover, despite the paper providing separate results for reader and writer emotions, the implementer faced challenges in extracting individual results using the code provided by the authors. This limitation hindered the ability to analyze and compare the performance of the model for reader and writer emotions separately. Consequently, a comprehensive understanding of the distinct emotional intensities experienced by these two groups was not achieved in the implementation.

Furthermore, the sensitivity analysis, such as conducting multiple runs with different random seeds, was not explicitly mentioned in the context of the implementation. The absence of a sensitivity analysis could impact the reliability and generalizability of the results obtained in the implementation.

In summary, the implementation encountered challenges related to the unavailability of the original test set, limitations in extracting individual results for reader and writer emotions, and the lack of a sensitivity analysis. These factors contribute to discrepancies between the obtained results and those reported in the paper, highlighting the importance of access to the original data and thorough evaluation procedures for reproducibility and accurate comparison of results.

## Resources

The reproduction of the research paper required various resources in terms of computation, time, people, development effort, and communication.

Computation: The implementation was run on a GPU, which necessitated access to a suitable computing infrastructure capable of handling the training of the fine-tuned BERT model. GPU resources are typically expensive, and their availability can impact the cost of reproduction.

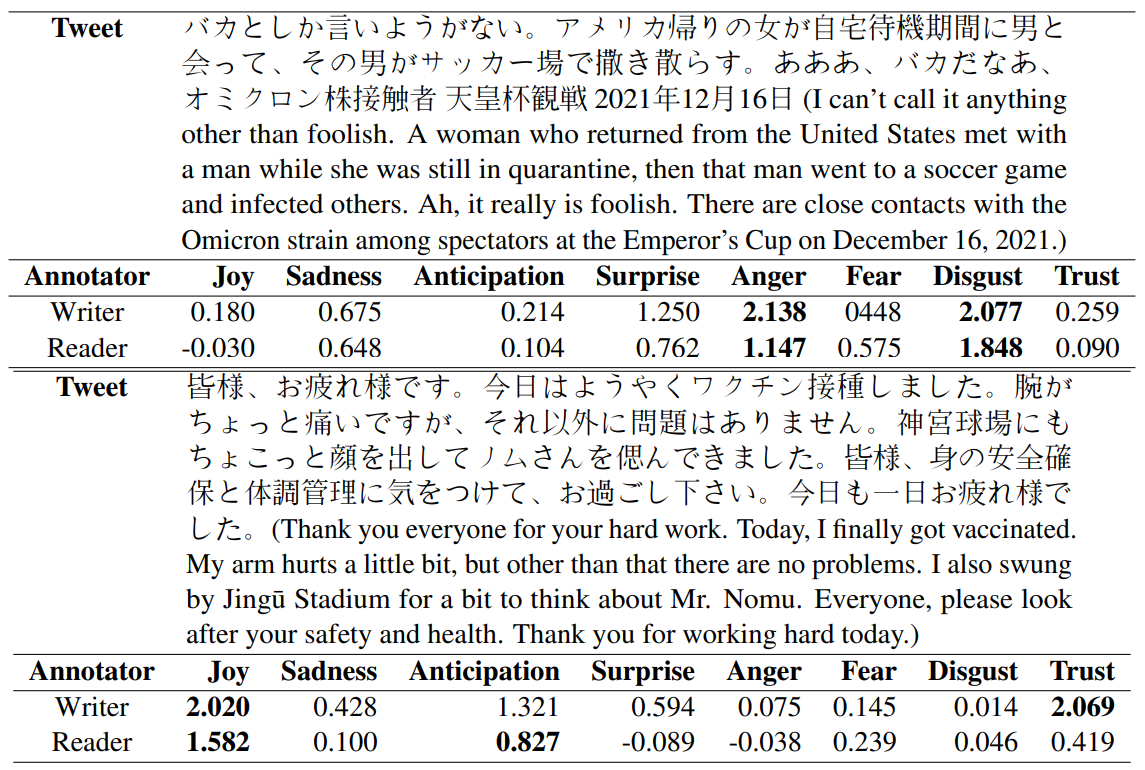
Development Effort: Significant effort would have been required to implement the methodology described in the paper, including adapting the BERT model for the specific task, preprocessing the dataset, training the model, and performing the emotion analysis. The complexity of the fine-tuning process and the need for accurate emotion prediction added to the development effort.

## Error Analysis

The manuscript analysis the single tweet to extract the emotions as shown in the below Figure. But the paper gives the example of a tweet that classifies correctly. For the error analysis, the paper didn’t provide any experiments and analyses. We are also unable to analyze errors with the provided code.

# Conclusion

The reproducibility of the research paper presents certain challenges due to the unavailability of the original test set and limitations in extracting individual results for reader and writer emotions. However, despite these obstacles, the implementation successfully fine-tuned a BERT model on a GPU, used an extracted test set for evaluation, and replicated the core methodology of predicting emotional intensity in Japanese Twitter users towards COVID-19 vaccination-related tweets. The obtained results differed from those reported in the paper, potentially due to variations in the extracted test set. The resources required for the reproduction, including computation, time, people, development effort, and potential communication with the authors, were significant. While the reproduction faced difficulties, the implementation demonstrated a reasonable level of reproducibility, providing valuable insights into the methodology and highlighting areas for improvement and future research.



# GitHub Link:

<https://github.com/lexec12/Emotion-Analysis-of-Writers-and-Readers-of-Japanese-Tweets-on-Vaccinations>

# References

CBalabantaray, R., Mohammad, M., & Sharma, N. (2012). Multi-Class Twitter Emotion Classification: A New Approach. *International Journal of Applied Information Systems*, *4*(1), 48–53. https://doi.org/10.5120/IJAIS12-450651

Colneric, N., & Demsar, J. (2020). Emotion Recognition on Twitter: Comparative Study and Training a Unison Model. *IEEE Transactions on Affective Computing*, *11*(3), 433–446. https://doi.org/10.1109/TAFFC.2018.2807817

Roberts, K., Roach, M. A., Johnson, J., Guthrie, J., & Harabagiu, S. M. (n.d.). *EmpaTweet: Annotating and Detecting Emotions on Twitter*.

Vateekul, P., & Koomsubha, T. (2016). A study of sentiment analysis using deep learning techniques on Thai Twitter data. *2016 13th International Joint Conference on Computer Science and Software Engineering, JCSSE 2016*. https://doi.org/10.1109/JCSSE.2016.7748849

Wang, W., Chen, L., Thirunarayan, K., & Sheth, A. P. (2012). Harnessing twitter “big data” for automatic emotion identification. *Proceedings - 2012 ASE/IEEE International Conference on Privacy, Security, Risk and Trust and 2012 ASE/IEEE International Conference on Social Computing, SocialCom/PASSAT 2012*, 587–592. https://doi.org/10.1109/SOCIALCOM-PASSAT.2012.119