Comparing User Comments on Twitter and Weibo A case study of Russia-Ukraine Conflict

Wenxuan Li

***Abstract*—The Russia-Ukraine conflict has been a hot issue on many platforms since it began. Additionally, different platforms have various regulatory restrictions and highly diverse user bases at various times. We evaluate the emotional differences between user comments on Twitter and Weibo on the same issue at various periods using the Bidirectional-LSTM-based Emotion Detection Model and Hashtags to detect the emotions of the comment data on the Russian-Ukrainian war from Twitter and Weibo.**

***Index Terms*—Russia-Ukraine conflict, Emotion detection, Bidirectional-LSTM, Twitter, Weibo**

I. Introduction

In this year, Russia formally declared war on Ukraine on February 24. This contemporary conflict sparked a significant online debate, and internet users from all over the world left comments on the occurrence on various social media platforms. As a result, we wish to examine how, through the comments of netizens from many nations, we can track changes in the common people's perceptions of the incident from the start of the conflict to the present. Depending on the social media platform, users have a tendency to rate the event differently. For instance, while Twitter users' remarks are serious, Weibo users' comments are hilarious.

We analyzed the communication function, social function, review mechanism, user group, and environmental maintenance of Twitter and Weibo, two well-known social media platforms with comparable functions in China and internationally. In general, Twitter is more powerful and offers a freer forum for conversation. We will compare the key attitudes of user comments on Twitter and Weibo and base our analysis mostly on Twitter comments related to the crisis between Russia and Ukraine.

TABLE I
Comparison of main features of Twitter and Weibo

General Comparison	Twitter	WeiBo
Communication Function	Two-way Internet communication function	Cannot communicate with other social media
Social Function	"what's happening" call for users to record news events	Events in the entertainment circle attract more attention from users
Review mechanism	Does not censor users	Comments posted by users are monitored and deleted if deemed harmful
User group	Users are mainly distributed abroad	Users are mainly ethnic Chinese
Environmental maintenance	Full service for users' reporting and notifies the results at the fastest speed	Blocking and reporting functions of Weibo are very unfriendly to users

II. Data & Model

The two main categories of our data are Emotion Detection Data and Russo-Ukrainian War Commentary Data from Weibo and Twitter. It is important to note that a Bidirectional-LSTM model is created using the emotion detection data and is then used to finish the emotion detection process. The data that is ultimately used for the major sentiment analysis is the Russo-Ukrainian War Commentary Data from Twitter and Weibo. The established Bidirectional-LSTM emotion detection model will be used to label the data from Twitter, and the text-based hashtag will be used to label the data from Weibo.

A. Emotion Detection Data

Through Kaggle, an emotion detection competition, we were able to acquire data on emotion recognition. The training set, test set, and verification set each comprise one third of the total amount of data. The training set has 16,000 text records, whereas the test and verification sets each contain 2,000 text records. The data contains six different types of labels: happiness, sadness, rage, fear, love, and surprise. The picture below displays the numerical distribution for each label in the three subsets:

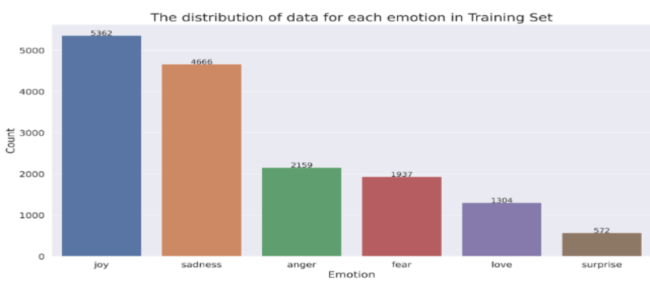

Fig. 1. The distribution of each label in the training set

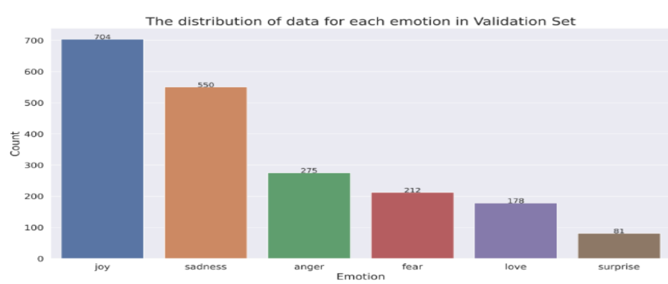

Fig. 2. The distribution of each label in the validation set

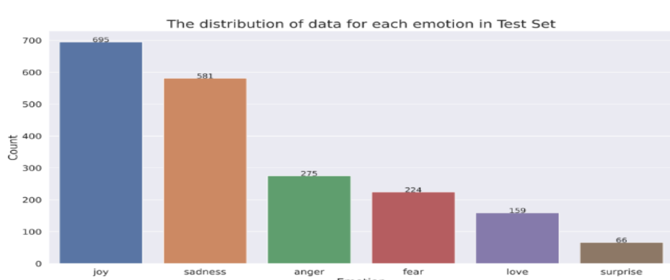

Fig. 3. The distribution of each label in the test set

Additionally, we must examine each text's length because we will subsequently use this data to develop an emotion recognition model. As displayed below:

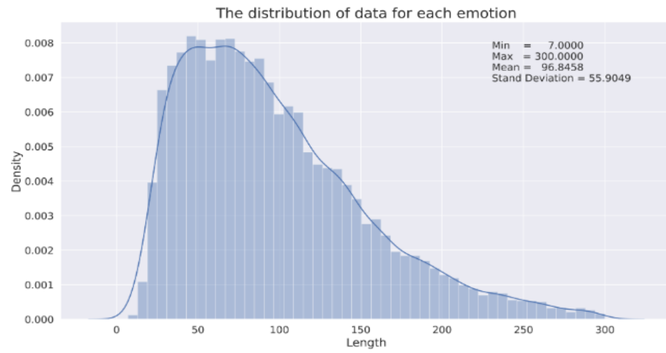

Fig. 4. The distribution of text length in the data

From this we know that the maximum length of the text in the data is 300, which is the data we need for subsequent processing. Then we need to process the text next, we have processed the text as follows:

1) Convert all letters to lowercase
2) Split abbreviations in the text, such as "you're" will be changed to "you are"
3) Remove emails and tags from text
4) Remove special chars and accented chars from text
5) Convert all words in the text to their base form, e.g. "ran" to "run"

In addition, the labels must be encoded. In this study, we encode each emotion as 0, 1, 2, 3, 4, and 5 in turn: happiness, rage, love, sadness, fear, and surprise. Next, in order to accomplish all of these goals, we record 10,000 distinct words in the tokenizer. The length of all sequences is padded to 300, which is the longest text length in the dataset, once the text data is transformed to numeric sequences. So far, we have finished processing the data before creating the model.

The Bidirectional-LSTM model that we developed completes the modeling process by reading text data from both positive and negative directions through two LSTM models, which helps the model comprehend a sentence's meaning more fully. In addition to this model, we create a matrix of size (10000, 100) using the GloVe pre-training model, which is then utilized to transform each word into a 100-dimensional word vector. The hyperparameters, epoch=25 and batch_size=120, are finally chosen by training on the verification set.

	precision	recall	f1-score	support
anger	0.90	0.95	0.93	261
fear	0.90	0.87	0.88	233
joy	0.93	0.96	0.94	675
love	0.89	0.81	0.85	176
sadness	0.98	0.95	0.96	597
surprise	0.70	0.79	0.74	58
accuracy			0.93	2000
macro avg	0.88	0.89	0.88	2000
weighted avg	0.93	0.93	0.93	2000

Fig. 5. The prediction ability of Bidirectional-LSTM Emotion Detection Model

B. Russo-Ukraine War Commentary Data

Kaggle is the source of the comment data on the Twitter battle between Russia and Ukraine. The data spans the months of February 2022 and July 2022, totaling 25,671 data points. To finish classifying this data, we employ the proven Bidirectional-LSTM Emotion Detection Model.

The 34,480 total bits of data that make up the Weibo comment data on the Russia-Ukraine war are taken from MANGODB. We use many hashtags to describe various emotions, and we finish the label by determining which hashtag best captures the sentiment expressed in the text.

III. Analysis

A. Twitter Users' Comment React Distribution

We start by examining the data from labeled Twitter comments. The pie chart in the following graphic represents how we represented the sentiment distribution of the data from Twitter comments. We can see that about 60% of users are angry, with close to 20% of people being sad. We draw the conclusion that foreign Twitter users generally have a poor perception of the turmoil in Russia.

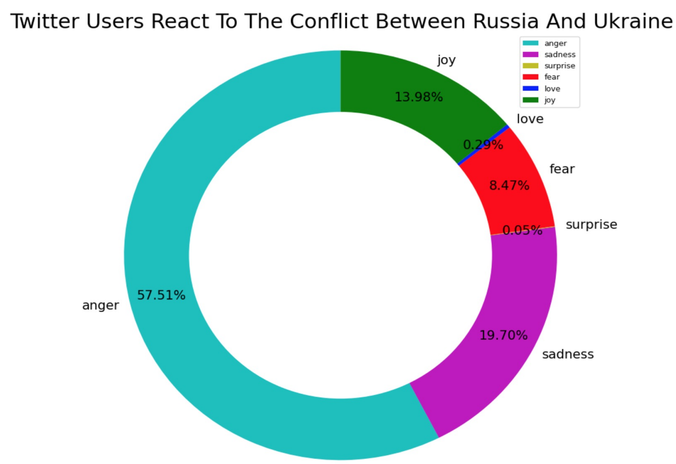

Fig. 6. Twitter Users React To the Conflict Between Russia and Ukraine

B. Twitter Users' Comment React changed over time

Using python's visual analysis, we can clearly see the distribution of Twitter users' reactions to events and their changes over time:

1) The number of angry users has been at its highest level, peaking in May
2) The number of users expressing likes and surprises always has the lowest peak, and the number of users expressing surprises continues to decrease
3) The number of people feeling happy went from being roughly on par with

those feeling fearful in February to more than 20% in July.

4) Macroscopically, little change

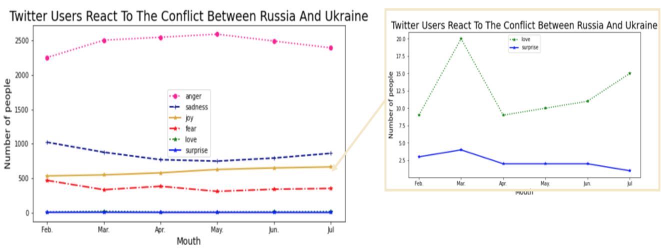

Fig. 7. Twitter Users React To the Conflict Between Russia and Ukraine Over Time

C. Why Twitter Users' Comment React has changed

There were additional Russian military triumphs while Russian forces slaughtered Ukrainian citizens, according to the timeline of the significant events in the Russia-Ukraine conflict, which began in May. It's one of the factors contributing to the rise in angry remarks. One of the factors contributing to the decline in users who feel loved is the train station massacre on April 8 that claimed the lives of 57 innocent victims. No human likes to witness the devastation brought on by conflict. However, the number of happy comments kept increasing, and we discovered that it was put in a location far from the conflict. It might be challenging for people of different social classes to have the same emotional resonance. On the other hand, contrary to what we previously believed, we've come to the conclusion that Twitter users' sentiments won't shift all that much unless something shocking occurs.

Since the outbreak of the incident from February to July, although there have been big and small news, the situation of military confrontation between the two countries has remained in general, and there has not been a big change (such as Ukraine's surrender or the United States' entry into the war, etc.). Therefore, the attitude of Twitter users and the scale of the event also maintained a positive correlation, without much fluctuation.

D. Comparison of comments on Twitter and Twitter

In the previous analysis, we compared review sentiment over time series. Next, we wanted to compare the sentiment distribution of the Russian and Ukrainian conflicts on Weibo and Twitter.

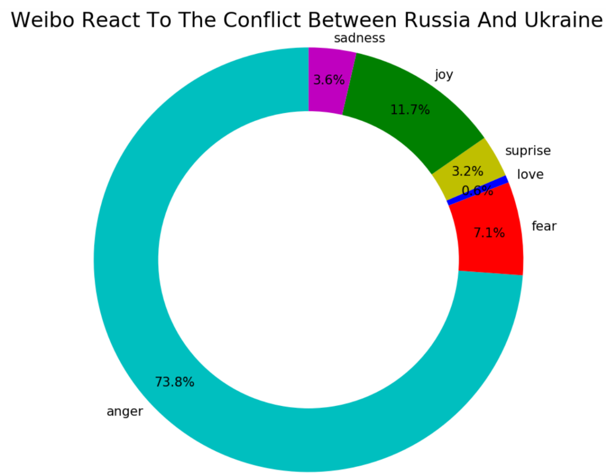

Fig. 8. Weibo Users React To the Conflict Between Russia and Ukraine

The remarks in Figures 6 and 8 share a lot of similarities, as can be seen. The most prevalent emotion is anger. Anger was voiced by roughly 57.5% of users on Twitter and by about 73.8% of users on Weibo. The number of comments that contain the word "love" is relatively low.

The reviews on the two platforms are likewise highly dissimilar at the same time. Twitter received less comments about the conflict between Russia and Ukraine than Weibo. The biggest disparity in comments with the sad label was found on the two sites. It was 3.6% on Weibo compared to 19.7% on Twitter, a difference of more than five times.

That is what we might anticipate. Since war is unpopular, the proportion of people who are angry is at its highest. One the one hand, Internet connectivity in China is not widely available. As a result, they are less aware of the situation than Twitter users are, and as a result, there are more comments that express surprise. On the other side, because those who can remark on Twitter may more easily and directly learn about the tragic repercussions of the situation in Russia and Ukraine, there are many more depressing comments on Twitter than on Weibo.

CONCLUSION

Weibo and Twitter are two well-known social media platforms in China and the rest of the globe, respectively, and they serve essentially the same purposes. Weibo and Twitter are similar in terms of user base, review mechanism, user group, and environmental upkeep, according to our analysis of their communication, social, and environmental functions. To contrast the two social media, we opted to pick a trending issue and analyze the sentiment of user comments made on it. The confrontation between Russia and Ukraine was the subject matter we settled on.

Since the beginning of the crisis between Russia and Ukraine about ten months ago, things have gotten more difficult. There are many users on Weibo and Twitter who express or analyze their own feelings and opinions. Our findings show that from the start of the battle to the present, the majority of people have been upset and furious over this tragedy. We believe that people are consistently opposed to starting wars because doing so would place a great deal of strain on the nation and its people. From Russia's perspective, they are defending their own nation's authority, whereas Ukraine appears to be looking for better development. But the fires of the battlefield were directed at civilians by both sides of the conflict.

To sum up, we believe that Twitter has more powerful functions and a more free communication environment than Weibo.

REFERENCES

[1] G.Srinivasa Raju,M.AjayDilipKumar & S.Suryanarayana Raju.(2019).Quantifying Inference Learning Model to Explore Twitter User Emotions. International Journal of Innovative Technology and Exploring Engineering (IJITEE)(10).

[2] Santhosh Baboo S. & Amirthapriya M..(2021).Comparison of Machine Learning Techniques on Twitter Emotions Classification. SN Computer Science(1). doi:10.1007/S42979-021-00889-X.

[3] CBalabantaray R,Mohammad Mudasir & Sharma Nibha.(2012).Multi-Class Twitter Emotion Classification: A New Approach. International Journal of Applied Information Systems(1). doi:10.5120/ijais12-450651.

[4] Yuwei Chuai and Jichang Zhao*. Anger Can Make Fake News Viral Online. Frontiers in Physics, 10:970174, 2022, doi:10.3389/fphy.2022.970174.

[5] Lange-Ionatamishvili, E., Svetoka, S., & Geers, K. (2015). Strategic communications and social media in the Russia Ukraine conflict. Cyber war in perspective: Russian aggression against Ukraine, 103-111.

[6] Gao, Q., Abel, F., Houben, G. J., & Yu, Y. (2012, July). A comparative study of users' microblogging behavior on Sina Weibo and Twitter. In International conference on user modeling, adaptation, and personalization (pp. 88-101). Springer, Berlin, Heidelberg.

[7] Lin, X., Lachlan, K. A., & Spence, P. R. (2016). Exploring extreme events on social media: A comparison of user reposting/retweeting behaviors on Twitter and Weibo. Computers in human behavior, 65, 576-581.

[8] Chen, S., Zhang, H., Lin, M., & Lv, S. (2011, December). Comparision of microblogging service between Sina Weibo and Twitter. In Proceedings of 2011 International Conference on Computer Science and Network Technology (Vol. 4, pp. 2259-2263). IEEE.