Sentiment Response to News Shock II

Columbia University

Zheng Hui(zh2683)

Advisor: Professor John Kender

Abstract

Different cultures should have different emotional responses to major events. People of different countries have their own views and emotions according to their culture, thoughts, habits and customs. Our main goal is to distinguish cultural differences in posts on social media. These posts have different sentiments on the major events of COVID-19 in different countries. On the same theme, while some may find satisfaction, others may express disgust or anger. Some cultures express emotion more directly, but others express it more ambiguously, identifying this emotion as a key topic in the field of natural language processing. During our work last semester, we decided to choose the UK and China as our research objects. They have completely different policies on COVID-19, and major events will have different sentiments on social platforms. Therefore, during the same period, the populations of these countries may have different attitudes. We successfully compared the similarities and differences in people's reactions to major events in social media.

Our study is divided into three main parts. Firstly, we built upon the framework of our previous semester's work but focused on examining the different reactions of China and the United States to the Silicon Valley Bank bankruptcy event. Secondly, we enhanced the performance of our Chinese language model, with its accuracy now consistent with the state-of-the-art. Lastly, we investigated the performance difference of a specific language model that uses zero-shot learning on various domains.

1.Introduction

In our previous study (Hui et al., 2022), we investigated the distinct responses of the United Kingdom and China to the novel Covid-19 shock. We believed that these two countries, with their unique cultural backgrounds, would exhibit diverse reactions. In this paper, we aim to enhance the analysis by expanding the range of data labels beyond the simplistic positive, negative, and neutral categories(Figure 1). To achieve this, we try to incorporate a five-point emotion scale(Figure 2), providing a more nuanced understanding of sentiments. Moreover, we have updated our sentiment analysis model for Chinese Weibo, transitioning from the previous approach to utilizing SKEP which stands for Sentiment Knowledge Enhanced Pre-training(Tian et al., 2020), which represents the latest state-of-the-art technique.

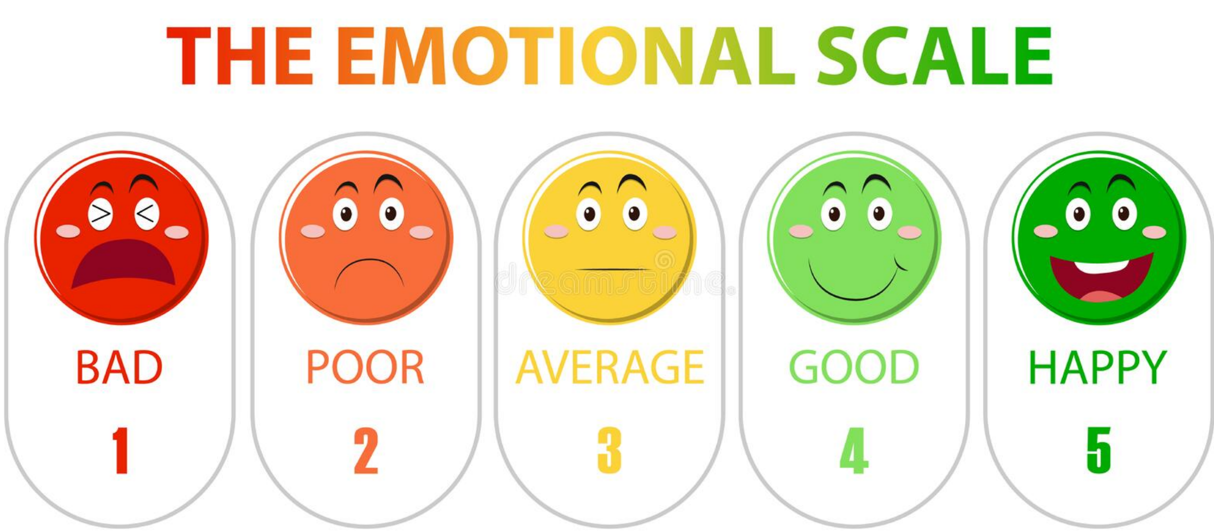


Figure 1&2: Five-point Emotion Scale(Kari Dunn Buron and Mitzi Curtis,2003)

In addition to these methodological advancements, we have identified and examined a noteworthy event, namely the bankruptcy of Silicon Valley Bank(Silicon Valley Bank, with the parent company, SVB Financial Group, filing for Chapter 11 bankruptcy. The bank has suffered losses due to its investment strategy, and depositors are being paid out by the FDIC. The situation may have broader implications for the tech industry, particularly for venture capital-backed companies that rely on Silicon Valley Bank). Through our proposed framework, we conduct a comparative analysis of public opinions in China and the United States, exploring both the similarities and differences. Lastly, we delve into the concept of zero-shot learning(Palatucci et al., 2008), leveraging large language models in various languages to train language-specific models.

2.RELATED WORK

2.1 Sentiment response to news shock

Previous work(Hui et al., 2022) compared the emotional fluctuations of significant events on Chinese Weibo and British Twitter during the COVID-19 pandemic, aiming to understand the similarities and differences between the two cultures. The emotional analysis is based on text analysis of Twitter data from the UK and Weibo data from China. by using Artificial intelligence technology and natural language processing tools are used to process data, analyze emotional trends and topics. By analyzing the main languages of the two countries, one can better understand people's reactions to these events, and explore the differences and similarities between different cultures. These results can help policymakers and the public better understand the impact of public events on different cultures and countries, and improve people's emotional health through appropriate interventions.

2.2 Sentiment Analysis on social media

Sentiment analysis in social media that explored the methods, social media platforms used and its application. Social media contains a large amount of raw data that has been uploaded by users in the form of text, videos, photos and audio. The data can be converted into valuable information by using sentiment analysis. Past studies, such as WITHDRAWN: A systematic study of sentiment analysis for social media data(Drus et al., 2019) and Topic-level sentiment analysis of social media data using deep learning(Pathak et al., 2021), have also investigated this field.

2.3 Large Language Models(LLMs)

A large language model (LLM), also known as a large neural language model, is a type of artificial intelligence model designed to understand and generate human language. These models are trained using large amounts of textual data and can perform a wide range of tasks, including text summarization, sentiment analysis, machine translation, and others. LLMs are distinguished by their large scale, incorporating billions of parameters, which enable these models to learn complex patterns present in linguistic data. LLMs commonly feature deep learning architectures such as Transformers, which help them achieve impressive performance across various natural language processing (NLP) tasks. Over the years, LLMs have gained popularity in NLP-related applications such as OpenAI's GPT-3 and GPT-4 (Brown et al., 2020) as well as Meta's LLaMA model (Touvron et al., 2023).

2.4 Zero-shot Learning

Recently, a lot of work has also focused on small sample learning and zero sample learning. Data labeling usually takes a lot of time and experience, so small sample and zero sample learning is very valuable and worthy of further research. Some in the field of sentiment analysis Zero-shot learning works such as Zero-Shot Learning for Cross-Lingual News Sentiment Classification (Pelicon et al., 2020) show that this is an area of growing interest.

3.MATERIAL & METHODS

3.1 Data Collection

In this study, we sourced data from both English Twitter and Chinese Weibo. To collect Chinese data, we utilized the same web scraping software as in our previous work(Hui et al., 2022), targeting the keywords "硅谷银行" (Silicon Valley Bank in Chinese) and "SVB" on both platforms. Additionally, we developed a new program to search for data using Twitter's API, targeting the same keywords which are “Silicon Valley Bank” and "SVB". Samples of the collected data are illustrated in Figures 3 and 4. We have collected about 45k data from Weibo and 65k data from Twitter to randomly choose a 30mins window from a day and continue for a one month period(two weeks before SVB went bankrupt and two weeks after).

Choosing Twitter and Weibo as representatives of social media platforms to collect opinions is primarily due to the following reasons:

Wide user base: Twitter and Weibo are highly popular social media platforms worldwide with a substantial user count. Through the collection of opinions on these platforms, a more extensive audience can be reached, resulting in diverse opinions and viewpoints from different backgrounds and regions.

Publicity and transparency: These platforms are open to the public, permitting the free exchange and discussion of opinions that is visible to all. This openness promotes transparency and fairness, facilitating a more open and traceable opinion collection procedure.

Real-time and interactive functionality: Twitter and Weibo's distinct characteristic is the rapid dissemination of information, including the ability for users to engage in live interaction with one another. This aspect enables opinion collection to gather a large amount of feedback and recommendations within a brief period, leading to real-time responses and discussions.

372	371	If SVB goes under it wou	0
373	372	Looks like Silicon Valley	-1
374	373	51% of tech companies ir	0
375	374	Ouch Silicon Valley Bank	0
376	375	Silicon Valley Bank launc	1
377	376	Afternoon in California r	0
378	377	Silicon Valley Bank is loo	0
379	378	Muy interesante la char	-1
380	379	I just called a VC to ask h	-1
381	380	I feel bad for Silicon Vall	-1
382	381	Oh nooooo #SVB	[illegible]

Figure 3: Twitter data with artificial labeled

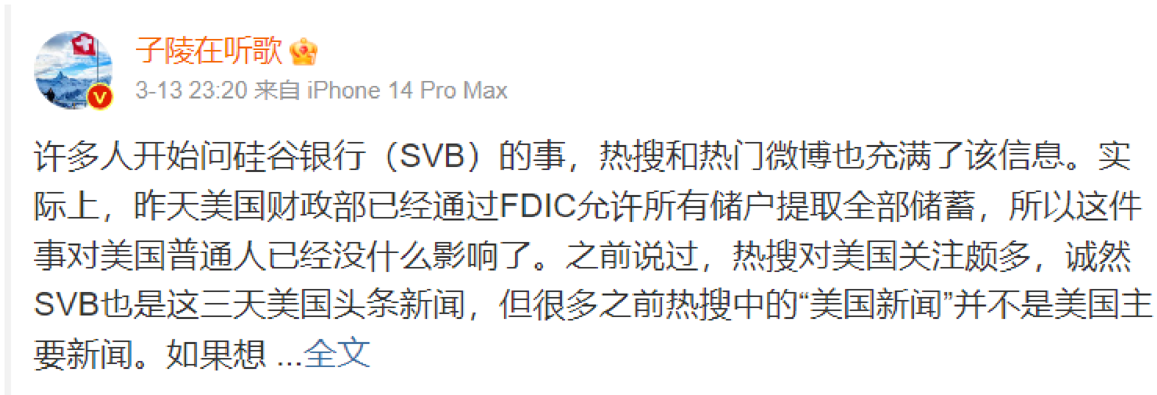

子陵在听歌
3-13 23:20 来自 iPhone 14 Pro Max
许多人开始问硅谷银行（SVB）的事，热搜和热门微博也充满了该信息。实际上，昨天美国财政部已经通过FDIC允许所有储户提取全部储蓄，所以这件事对美国普通人已经没什么影响了。之前说过，热搜对美国关注颇多，诚然SVB也是这三天美国头条新闻，但很多之前热搜中的“美国新闻”并不是美国主要新闻。如果想 ...全文

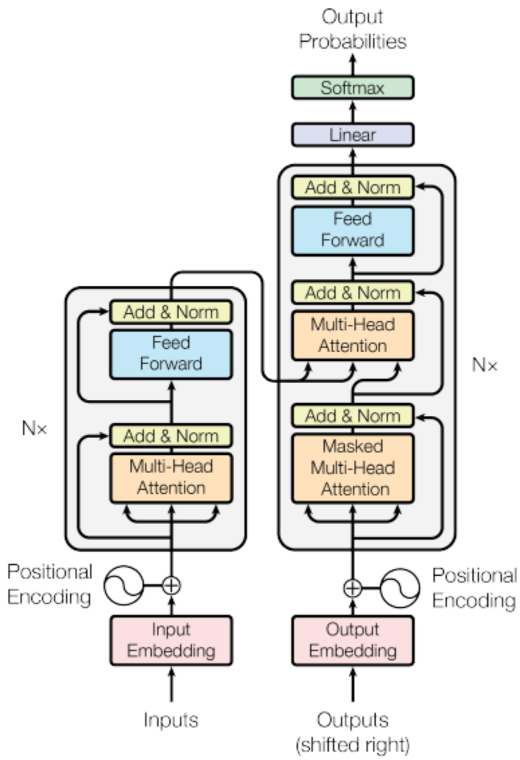
Figure 4: Weibo data sample

3.2 Classification Model

A. Model

In this research, we utilized Twitter Bert(Nguyen et al., 2020) for the Twitter dataset and performed fine-tuning with our dataset, similar to our previous semester. For the Weibo dataset, we adopted a novel pre-trained model named Sentiment Knowledge Enhanced Pre-training(Tian et al., 2020) and conducted fine-tuning to align with the state-of-the-art methods.

To ensure the accuracy and relevance of our sentiment analysis, we updated our methodology by employing SKEP, a cutting-edge pre-training model designed specifically for sentiment analysis tasks. This state-of-the-art model incorporates advanced sentiment knowledge, enabling us to capture the intricacies of emotions expressed in social media posts on the Chinese platform Weibo. Both of the models are based on Transformer mechanism(Vaswani et al., 2017) and build on BERT(Devlin et al 2018) and RoBerta(Liu, 2019).



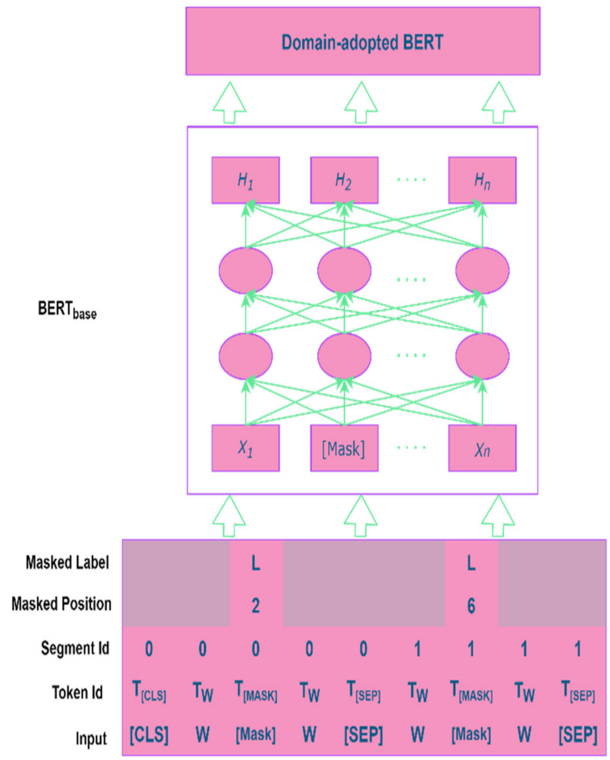


Figure 5: Transformer mechanism(Vaswani et al., 2017) *Figure 6: BERT example*

For Chinese Weibo data, we use SKEP(Tian et al., 2020) to analyze the data. The SKEP model introduces key modifications to RoBERTa, particularly in the aspect-sentiment pair prediction using multi-label classification. Instead of predicting pairs using the representation of the entire sequence, SKEP utilizes the final state of the classification token [CLS]. The prediction process employs a sigmoid activation function, which enables the occurrence of multiple tokens simultaneously. The objective Lap for the aspect-sentiment pair is denoted as follows:

$$\hat{\mathbf{y}}_a = \text{sigmoid}(\widetilde{\mathbf{x}}_1 \mathbf{W}_{ap} + \mathbf{b}_{ap})$$

$$L_{ap} = -\sum_{a=1}^{a=A} \mathbf{y}_a \log \hat{\mathbf{y}}_a$$

Figure 7: how output vector calculated by SKEP(Tian et al., 2020)

B. Fine-ture

Our approach which was similar to what we used in last semester is to freeze all the layers except the last few (typically the last convolutional layers and the fully-connected layers) and retrain them with the new dataset. Since the amount of data is small, the similarity is high. In this case, we only need to modify the last few layers or only the output category of the final softmax layer which can project the classification to three different classes(Positive, Negative, Netural). And what is new for this semester is after a half of the training data feed in to our neural network(TwitterBERT/SKEP) then we freeze all layers and apply child-turing(Xu et al., 2021) then we feed in another half of the training data. By applying this approach, we successfully build a stronger model.

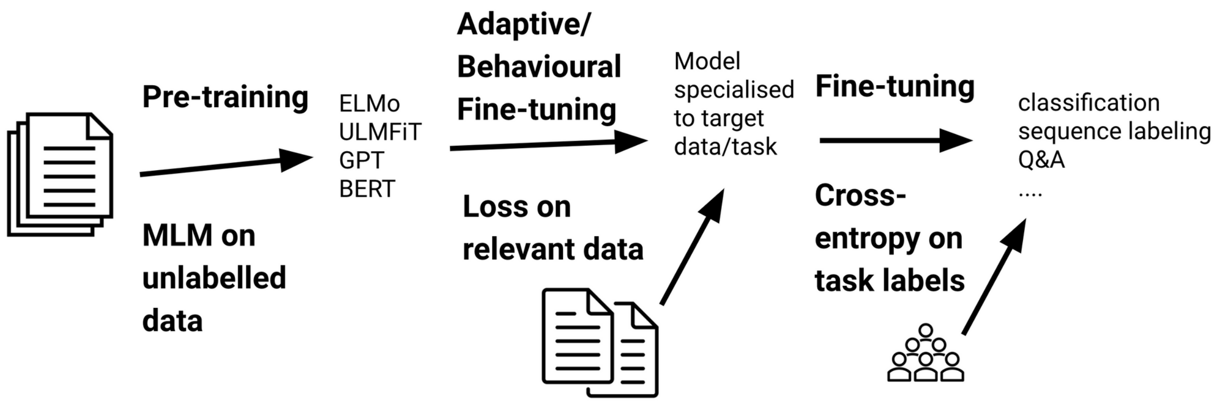


Figure 7: Pre-training model and fine-tuning example, by Jerry Wei, BERT: Why it's been revolutionizing NLP,Towards Data Science

After fine-tuning, our English model achieved an accuracy of 0.89, a 0.01 improvement from last semester, and a macro-F1 score of 0.72. Meanwhile, our Chinese model achieved an accuracy of 0.88, a 0.04 improvement from last semester, and a macro-F1 score of 0.67.

C. Levaging LLMs

We are also leveraging LLMs to better understand how the model performs in different languages other than the fine-tuning dataset. Which it provides is another way to look at how people's emotions shift during shock news.

4.RESULTS

4.1 Twitter Data

We successfully collected over 100,000 Twitter data using its built-in API. However, for the purpose of balanced comparison, we downsized the dataset to a random sample of 65,000 Twitter data. This step aimed to facilitate better comparison between two distinct cultures. The dataset was collected daily from February 23 to March 25, 2023, before and after two weeks following the bankruptcy declaration of Silicon Valley Bank on March 10.

Input: Set of tuples X, size of samples n
Output: Set of sample tuples S

$S \leftarrow \varnothing$
for $i = 1 \rightarrow n$ **do**
  $x \leftarrow selectRandomTuple(X)$
  $S \leftarrow S \cap \{x\}$
end for

Figure 7: Pseudocode for Simple random sampling for downsampling

The US dataset contains a total of 65000 Twitter and by our model the detailed principal component analysis are shown in the figure below.

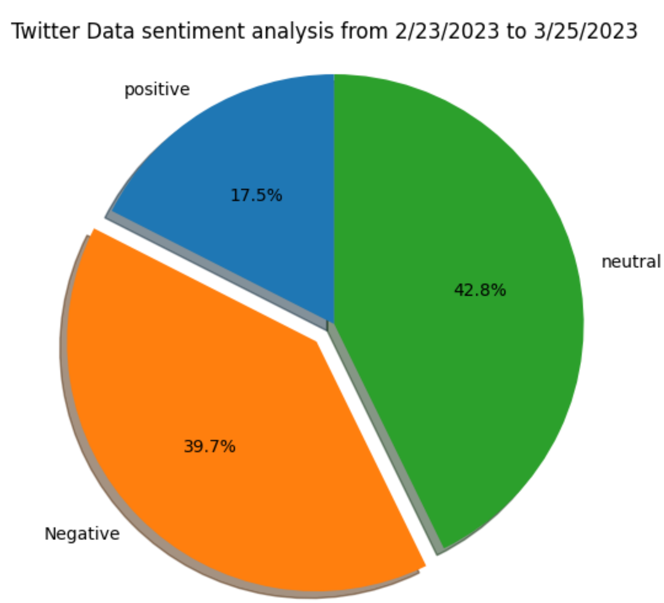


Figure 8: Pie chart showing total sentiment count

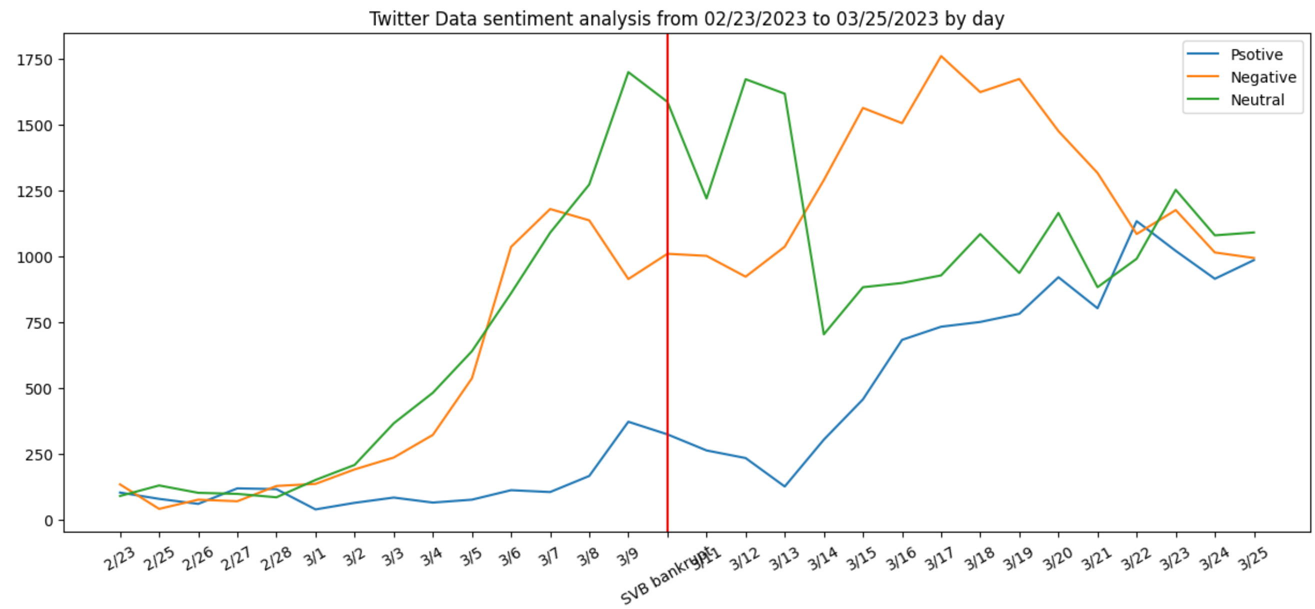


Figure 9: correlation between three different emotions during 2/23/2023 -3/25/2023

Above figure shows sentiments on social media related to keyword Silicon Valley Bank or SVB over a one month period center of event SVB's Bankrupt.

4.2 Weibo Data

We used crawlers to collect 45000 data from Weibo. The dataset was collected daily from February 23 to March 25, 2023, before and after two weeks following the bankruptcy declaration of Silicon Valley Bank on March 10. The key word using for Weibo web crawlers is “硅谷银行” or “SVB”.

The China dataset contains a total of 45000 Twitter and by our model the detailed principal component analysis are shown in the figure below.

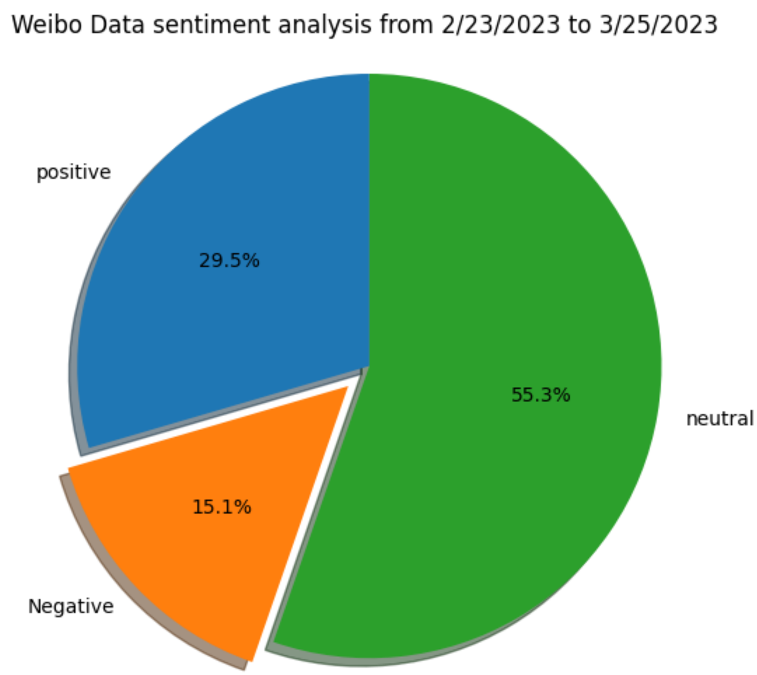


Figure 10: Pie chart showing total sentiment count

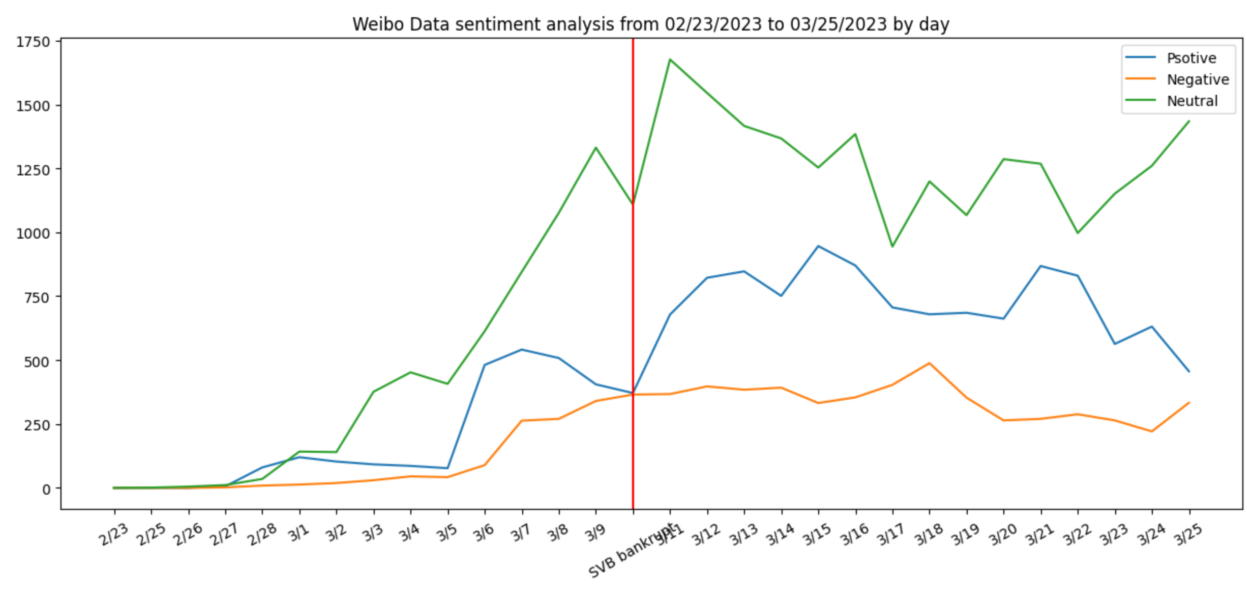


Figure 11: correlation between three different emotions during 2/23/2023 -3/25/2023

4.3 Comparison

According to available data on Weibo, 29.5% of individuals hold a positive view regarding Silicon Valley Bank's failure while 15.1% hold an opposing perspective. The majority of individuals, approximately 55.3%, hold neutral views. Within China , the collapse of Silicon Valley Bank may represent an example of the shortcomings of the Western financial system and, in contrast, serve as evidence of the comparative stability and reliability of the Chinese financial system. A number of Chinese Weibo users utilize the bank's collapse to criticize and highlight the deficit cies of the US financial system while also advocating for China's own more sound and stable financial system. In considering the collapse, many individuals may make comparisons with the effectiveness of China's own financial supervision and argue that China's financial supervision is strict and thus better positioned to prevent similar incidents .

Analysis of data from Twitter indicates that 17.5% of individuals hold a positive view in relation to Silicon Valley Bank's failure while 39.7% hold a negative view. The majority of individuals, approximately 42.8%, hold neutral views regarding the incident. Generally speaking, Silicon Valley Bank holds a relatively low proportion of deposits owned by the public. The institution caters primarily to technology companies and venture capitalists. As such, these groups have expressed concern.

In the United States, the collapse of Silicon Valley Bank may be perceived as an event within financial markets, similar to other bank failures or financial institutions facing distress.

The failure of Silicon Valley Bank could trigger discussions on the financial regulatory system, including the effectiveness of financial supervision, risk management, and institutional reform. This incident could also prompt calls for increased scrutiny and regulation of the fintech industry to prevent similar occurrences from arising in the future.

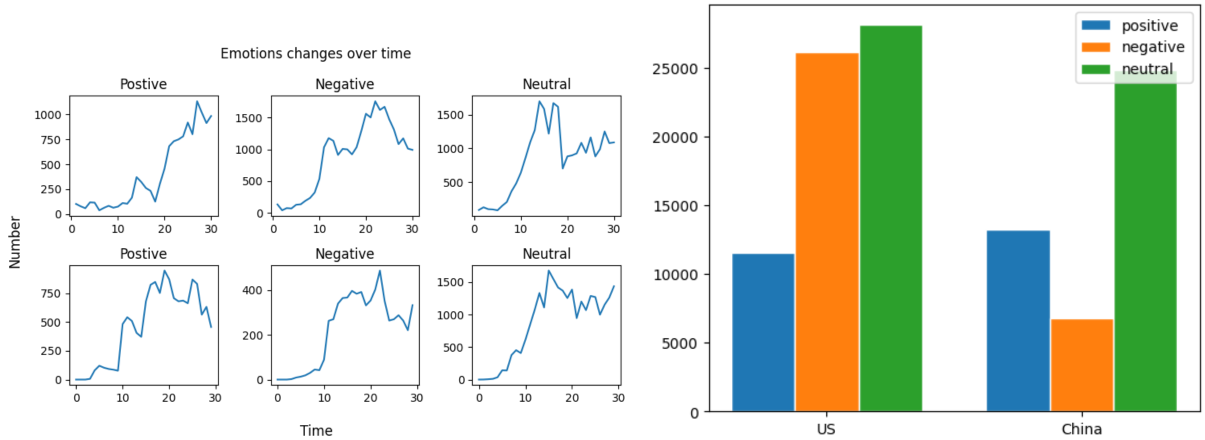


Figure 12: Sentiments for each emotions

There are several possible reasons for the varying perceptions of the collapse of Silicon Valley Bank. Some individuals may perceive the event positively because they view it as a result of financial system reform, which they believe may ultimately promote financial stability and transparency. Conversely, others May Perceive the Collapse Negatively, Fearing the Potential Knock-On Effects on the Wider Economy, or BeCAUSE of Personal Interests Vestem. For Example, Customers and Shareholders of the Bank May Sustain Financial Losses Due to this. View towards the event may include individuals who lack deposits in the bank or general knowledge about the intricacies of banking operations. These individuals may also adopt a wait-and-see approach until more information becomes available.

4.4 Redefine sentiment labels

We attempted to replicate the categorization of sentiment labels in our study, transitioning from a basic positive-negative-neutral classification to a more intricate 1-5 sentiment scale. However, this modification resulted in a decline in model performance. Specifically, our few-shot model exhibited signs of overfitting. This outcome could be attributed to the insufficiency of our model's capacity and robustness or the inherent nature of social media data, which often lack extreme sentiments, thus posing challenges in discerning nuanced emotional divisions.

4.5 Cross Language explore

We also study the cross-lingual robustness of a given language model by using LLM. Specifically, we use the English language model and English Twitter data for fine-tuning, and then evaluate on the Chinese test set. We also perform this evaluation in reverse. Our exploration yields surprising results, among current MLMs and LLMs, which have shown excellent cross-lingual adaptability, including zero-shot scenarios. This area deserves further study.

Model	Chinese	English
Chinese(SKEP)	0.88	0.84
English(TwitterBERT)	0.79	0.89

Table 1: ross-lingual robustness

5.FUTURE WORK

Multimodal learning is a popular and highly sought-after field in machine and deep learning. In our project, we only used the text input of Tweets and Weibo posts, but users frequently include images in their social media posts. Multimodal learning allows for combining natural language processing and computer vision, which can improve sentiment analysis accuracy by eliminating ambiguities and examining the data from multiple angles.

REFERENCE

1. *Zheng Hui, Zihang Xu, John Kender; (n.d.).Sentiment response to news shock.*
2. *Buron, Kari Dunn, and Mitzi Curtis. The incredible 5-point scale: Assisting students with autism spectrum disorders in understanding social interactions and controlling their emotional responses. AAPC Publishing, 2003.*
3. *Tian, Hao, et al. "SKEP: Sentiment Knowledge Enhanced Pre-training for Sentiment Analysis." Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. 2020.*
4. *Palatucci, Mark, et al. "Zero-shot learning with semantic output codes." Advances in neural information processing systems 22 (2009).*
5. *Drus, Zulfadzli, and Haliyana Khalid. "Sentiment analysis in social media and its application: Systematic literature review." Procedia Computer Science 161 (2019): 707-714.*
6. *Pathak, Ajeet Ram, Manjusha Pandey, and Siddharth Rautaray. "Topic-level sentiment analysis of social media data using deep learning." Applied Soft Computing 108 (2021): 107440.*
7. *Brown, Tom, et al. "Language models are few-shot learners." Advances in neural information processing systems 33 (2020): 1877-1901.*
8. *Touvron, Hugo, et al. "Llama: Open and efficient foundation language models." arXiv preprint arXiv:2302.13971 (2023).*
9. *Pelicon, Andraž, et al. "Zero-shot learning for cross-lingual news sentiment classification." Applied Sciences 10.17 (2020): 5993.*
10. *Dat Quoc Nguyen, Thanh Vu, and Anh Tuan Nguyen. 2020. BERTweet: A pre-trained language model for English Tweets. In Proceedings of the 2020 Conference on Empirical*

Methods in Natural Language Processing: System Demonstrations, pages 9–14, Online. Association for Computational Linguistics.

11. *Liu, Yinhan, et al. "Roberta: A robustly optimized bert pretraining approach." arXiv preprint arXiv:1907.11692 (2019).*
12. *Devlin, Jacob, et al. "Bert: Pre-training of deep bidirectional transformers for language understanding." arXiv preprint arXiv:1810.04805 (2018).*
13. *[10]Vaswani, Ashish, et al. "Attention is all you need." Advances in neural information processing systems 30 (2017).*