

# A Study on the Emotional Analysis of Abandoned Surrogacy Events Based on Text Mining

Guanlan Liang<sup>1</sup>, Xunbing Shen<sup>1,\*</sup>

<sup>1</sup>Jiangxi University of Chinese Medicine, Nanchang, Jiangxi, China

**Abstract.** In late January 2021, news that actress Zheng Shuang had surrogacy abroad and had wanted to give up her children sparked a public outcry. This paper takes Zheng Shuang's comments on the topic of surrogacy and her abandonment as the research object. Firstly, the web crawler technology is used to grab and mine the comment text, and then the ROSTCM software is used to analyze the text data to explore the comment topics of Weibo network users after the abandonment event and the analysis of their emotional tendencies to the event.

## 1 Introduction

According to the 47th Statistical Report on Internet Development in China released by China Internet Network Information Center (CNNIC) in February 2021, as of December 2020, the number of Internet users in China has reached 989 million, with the Internet penetration rate reaching 70.4%. More and more users choose to access news or hot events through the Internet in the first place. On January 18, 2021, news surfaced that actress Zheng Shuang had given birth to two children by surrogacy abroad and then wanted to give them up to be raised abroad. In a very short period of time, Weibo trending topics such as "Zheng Shuang abandoned her surrogate mother" and "China bans surrogacy in any form" topped the charts, garnered more than 5 billion views.

As the main body of this incident is a public figure, and involves surrogacy laws and regulations and special social ethical issues, it has aroused heated discussion among netizens. Therefore, this paper analyzes netizens' Weibo comments on this event, explores the topics of Weibo users' comments, and further understands the changes of netizens' emotional positioning.

## 2 Literature review

At present, there are two main methods of text sentiment analysis, one is based on sentiment dictionary, and the other is based on machine learning.

With the help of machine learning, Pang et al. used Naive Bayes (NB), Maximum Entropy (ME) and Support Vector Machine (SVM) to conduct sentiment analysis[1]. Yang et al. divided comments on Twitter into three categories: positive, neutral and negative[2]. Lin et al. uses a Bi LSTM and Attention model to reduce the dependence of sentiment analysis methods on humans[3]. Li Lei et al. also used this model and compared it with the analysis

effect based on deep learning[4]. Luo et al. propose to merging naive Bayes and causal rule for small sample data sentiment analysis scenarios[5]. S. Kumano and K. Nomura proposed a multitask RS removal method, which was an extension of the method in the field of emotional computing[6]. K. Chawla et al. proposed techniques to incorporate affect lexica, which captured fine-grained information about a word's psycholinguistic and emotion orientation, into the training for Word2Vec SkipGram, Word2Vec CBOW, and GloVe using a Joint Learning approach[7]. There is also a wealth of research based on sentiment dictionaries. Liu et al. used K-means clustering algorithm to construct sentiment dictionary and further carried out sentiment analysis on online product reviews[8]. Mohamad and Moattar proposed a new approach that combines neural networks with emotion dictionaries[9]. Felipe et al. added the variable of time evolution[10].

Due to the particularity and diversity of Chinese grammar, the research on the emotional tendency of text appears relatively late in China, and the research mainly relies on the Web platform. In the early stage, Zhou Shengchen et al. summarized the research on Chinese microblog sentiment analysis and put forward the need to focus on the development of relevant dictionaries or corpora for the filtering and emotion mining of network language[11]. At present, through the literature review of the past ten years, Zhong Jiawa et al. found that there are mainly three text sentiment analysis methods, which are based on sentiment dictionary and rules, based on traditional machine learning and based on deep learning, and all of them have advantages and disadvantages. At the same time, the method based on multi-strategy mixing gradually becomes an important direction for improvement[12]. For example, Song Yan took Weibo topics as an empirical case to study and verify the fine-grained viewpoint mining process and system functional modules, which further extended the traditional fine-

\* Corresponding author: 20060923@jxutcm.edu.cn

grained viewpoint mining model[13]. Fan Hao and Li Pengfei proposed a model based on word vector and two-way GRU cyclic neural network, which improved the accuracy of sentiment classification in networked short texts[14]. Du Yixian et al. used HowNet sentiment dictionary and sentiment vocabulary ontology to conduct sentiment analysis of public opinion in the COVID-19 epidemic[15]. Zhou Ziyu and Liu Fang'ai proposed a novel neural network, named Filter Gate Network based on Multi-head attention (FGNMH) to solve the problem that the attention mechanism in ASC brings noise and captures contextual words that are irrelevant to the current aspect[16].

To sum up, the research at home and abroad mainly focuses on the theoretical and technical exploration of data mining technology, and there are few empirical research results based on typical cases. At the same time, through sorting out the research process of the emotional tendency of text, it can be found that there is no most suitable method for the identification of the emotional tendency of Chinese microblog comment text. The current mainstream method is to use the emotion dictionary combined with Chinese grammar rules to score. Therefore, this paper uses text data mining technology to capture Weibo comments related to Zheng Shuang's surrogacy abandonment, and then uses ROSTCM to conduct emotional orientation analysis, so as to perceive the impact of this event on the public.

### 3 Sentiment analysis of microblog comments

The basic research idea of the article is to first use a web crawler to collect user comment information on topics related to "Zheng Shuang surrogacy and abandonment". After pre-processing such as cleaning the collected data, word frequency is counted and word cloud map is drawn to facilitate intuitive analysis of hot and concerned words in the later stage. Then perform sentiment analysis on the topic comment information, judge the positive and negative sentiment ratio of the topic, and conduct a more specific analysis based on this.

#### 3.1 Text Mining and Pre-processing

Text mining refers to the discovery of hidden patterns in large text collections. Web text mining technology includes text capture, text collection, text preprocessing, text analysis and other processes. This article uses Gooseeker web crawler software to capture microblog comments on related microblog topics. The software can capture effective related data such as blog content, comments, user ID, and save the captured data as Excel, which is conducive to later data analysis. Fig. 1 shows the process of mining the comment text of Zheng Shuang's surrogacy and abandonment.

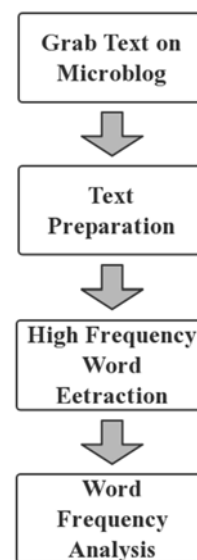


Fig. 1. Flow chart of microblog comment text mining.

#### 3.2 Word Frequency Statistics

This paper searched the keyword "Zheng Shuang's Surrogacy Abandoned", and obtained a total of 814 comments on blog posts from January 18 to 22, 2021.

The ROST Content Mining software is used for word segmentation of text. In the process of word segmentation, we need to make custom word segmentation for the obtained text. The words involved in this study, such as "Zheng Shuang", "Zhang Heng" and "surrogacy", need to be added to the "custom word segmentation filter word list", and then carry out word frequency statistical analysis. The first 27 high-frequency words are sorted out as shown in Table 1 below.

Table 1. Some words and word frequency statistical table.

Words	Num.	Words	Num	Words	Num.
Ban	104	Sentencing	31	Recording	23
Moral	103	Speak up	30	Law	23
Children	96	Artist	29	Cheating	22
Bottom line	81	Woman	29	Parents	22
China Television Artists Association	79	China	28	Man	22
Beyond	77	Entertainment	27	Oppose	21
Female	67	Country	26	Strongly	21
Surrogacy Agency	59	Star	25	Prohibit	17
Fans	31	Disgusting	23	Public Opinion	15

According to the word frequency statistics, in the topic of "Zheng Shuang's surrogacy and abandonment", Weibo users are more concerned about whether the actress will be banned, the moral issues of surrogacy and

abandonment, the future of the children, and the rights of women and men. The two words with the highest frequency are "Ban" and "Moral", which shows the attitude of Internet users. In addition, from the words "Disgusting", "Oppose" and "Prohibit", we can initially perceive that most netizens maintain a negative attitude towards this matter, accompanied by anger and boredom.

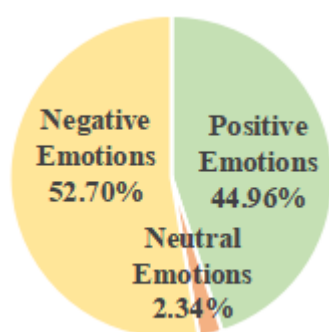
### 3.3 Text Sentiment Analysis

ROSTCM has a custom word list, emotion calculation, positive emotion, negative emotion, neutral emotion classification calculation and other functions.

This software was used to conduct emotional analysis on the text data of "Zheng Shuang's surrogacy and abandonment", which was divided into three categories of positive emotions, negative emotions and neutral emotions. Statistical results of emotional distribution were obtained (Table 2), and the results were plotted into a pie chart (Fig. 2).

**Table 2.** A table of sentiment analysis on Weibo texts of surrogacy and abandonment incidents.

	Affective Tendency Value	Number	Percentage
Positive Emotions	$(5, +\infty)$	366	44.96%
Neutral Emotions	$[5, 5]$	19	2.34%
Negative Emotions	$(-\infty, 5)$	429	52.70%
<i>Segmental Statistical Results of Positive Emotions</i>			
Low	$(5, 15]$	220	60.11%
Medium	$(15, 25]$	98	26.78%
High	$(25, +\infty)$	48	13.11%
<i>Segmental Statistical Results of Negative Emotions</i>			
Low	$[-15, 5)$	383	89.28%
Medium	$[-25, -15)$	32	7.46%
High	$(-\infty, -25)$	14	3.26%



**Fig. 2.** A figure of Sentiment analysis on surrogate pregnancy abandonment event.

According to the rules of text emotion calculation, 814 microblog texts were calculated and counted, and then their polarity was judged. In this event, there were 429 microblogs with negative emotions, accounting for 52.70% of the total, which is more than the sum of the microblogs with positive emotions (366) and intermediate emotions (19). It can be seen that netizens' emotional tendency towards this event is more negative, and this event also has a certain negative impact on the society. By observing the

data of negative emotions and combining with the analysis of Weibo users' historical information, it is found that most Weibo users with negative comments are female. On the one hand, they are emotional, on the other hand, they are familiar with laws and secular common sense. Therefore, it is only natural for people to express their condemnation of abandoning their children and their disdain for surrogacy. However, it should also be noted that there are a lot of comments with positive emotions. In addition to fans' support for their idols, most of them are comments like "strongly support the punishment of both parties in surrogacy transaction". Especially after the official media made their comments in the later stage, positive comments on microblogs dominated. It shows the netizens' approval of seriously dealing with the behavior of surrogacy and hope that the stars can play a positive guidance.

### 4 Summary

Starting from the event of "Zheng Shuang's surrogacy abandoning", this paper, based on web crawler technology, conducts word frequency statistics and judgment of emotional polarity intensity on relevant real-time microblog comment texts.

According to the statistics of high-frequency words, netizens have launched heated discussions on this topic, including Zheng Shuang was banned from broadcasting and television, whether surrogacy should be punished, the bottom line of morality, surrogacy agencies and so on. It can be seen that this surrogacy abandonment event is a social event involving a series of issues such as class, ethics and gender. Most netizens stand in opposition to surrogacy, condemn the abandonment of care, denounce the standpoint of surrogacy agencies. A small number of netizens believed that it was not all the woman's fault, and that the man was also to blame, and felt distressed and understanding. In addition, from the China Television Artists Association said Zheng Shuang beyond the moral bottom line and other official media can be seen in their opposition to surrogacy abandoned attitude.

Through the analysis of emotional tendency and the extraction of polarity intensity, it can be seen that the majority of netizens' attitudes are negative, which is far more than positive emotions. This is understandable, after all, surrogacy is banned in China. Abandoning is against social morality and public order good customs. Some positive emotions, including fans' affection for their idols and netizens' approval of the official ban, are not always in favor of surrogacy. Most netizen masses are not willing, do not accept, also do not allow this kind of surrogacy to abandon the behavior.

This incident, on the one hand, exposed the huge risks and ethical crisis of surrogacy, on the other hand, it also pushed the dialectical needs of surrogacy to the stage again. There is an objective market demand for surrogacy, such as infertility and sexual minorities. If surrogacy were to be banned altogether, it would likely drift further underground and into darker corners, even abroad. But both official and public opinion are firmly opposed to commercial surrogacy. In a word, the management system of surrogacy needs to be further improved.

## Acknowledgments

This study was partially supported by the grants from the Planned Project of Social Sciences in Jiangxi Province (No. 18JY24) and the project of "1050 Young top-notch talent" of Jiangxi University of Traditional Chinese Medicine.

## References

1. Pang B, Lee L, Vaithyanathan S, et al. (2002) Thumbs up: Sentiment Classification using Machine Learning Techniques. *Empirical Methods in Natural Language Processing*, 79-86.
2. Yang M, Tu W, Wang J, et al. (2017) Attention Based LSTM for Target Dependent Sentiment Classification. *National Conference on Artificial Intelligence*, 5013-5014.
3. Lin Xingmin, Ho Chunheng, Xia Luting, Zhao Ruoyi. (2021) Sentiment analysis of low-carbon travel APP user comments based on deep learning. *Sustainable Energy Technologies and Assessments*, 44.
4. Li Lei, Wu Xuhui, Liu Ji. (2021) Sentiment Analysis Model of Bi-LSTM with Key Opinion Target Recognition and Deeper Selfattention. *Journal of Chinese Computer Systems*, 42(03): 504-509.
5. Luo Yongjian, Yang Xiaohua, Ouyang Chunping, et al. (2021) Merging Naive Bayes and Causal Rules for Text Sentiment Analysis. *Journal of Physics: Conference Series*, 1757(1).
6. S. Kumano, K. Nomura. (2019) Multitask Item Response Models for Response Bias Removal from Affective Ratings. *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*, Cambridge, UK, 1-7.
7. K. Chawla, S. Khosla, N. Chhaya, K. Jaidka. (2019) Pre-trained Affective Word Representations. *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*, Cambridge, UK, 1-7.
8. Liu F, Wei F, Yu K, et al. (2017) Sentiment classification of reviews on automobile website by combining word2vec and dependency parsing. *Proceedings of the International Conference on Smart Computing and Communication*, 206-221.
9. Mohamad Beigi Omid, Moattar Mohammad H.. (2020) Automatic construction of domain-specific sentiment lexicon for unsupervised domain adaptation and sentiment classification. *Knowledge-Based Systems*, 213.
10. Felipe Bravo-Marquez, Arun Khanchandani, Bernhard Pfahringer. (2021) Incremental Word Vectors for Time-Evolving Sentiment Lexicon Induction. *Cognitive Computation* (prepublish).
11. Zhou Shengchen, Qu Wenting, Shi Yingzi, et al. (2013) Summary of Chinese Microblog Sentiment Analysis. *Computer Applications and Software*, 30(03): 161-164+181.
12. Zhong Jiawa, Liu Wei, Wang Sili, et al. (2021) text sentiment analysis methods and application were reviewed. The data analysis and knowledge discovery, 1-15.
13. Song Yan. (2020) Research on Multi-level Fine-grained Attribute Mining of Social Media Text Information. *Information Science*, 38(11): 98-103.
14. Fan Hao, Li Pengfei. (2021) Short Text Sentiment Analysis Based on FastText Word Vector and Bidirectional GRU Cyclic Neural Network -- A Case Study of Microblog Comment Text. *Information Science*, 39(04): 15-22.
15. Du Yixian, Xu Jiapeng, Zhong Linying, et al. (2021) Multi-dimensional Feature Analysis and Visualization of Online Public Opinion Situation and Sentiment: A Case Study of COVID-19 Epidemics. *Journal of Geo-information Science*, 23(02): 318-330.
16. Zhou Ziyu, Liu Fang'ai. (2021) Filter gate network based on multi-head attention for aspect-level sentiment classification. *Neurocomputing*, 441: 214-225.