IJISRR

Research Article

FORECAST AND ANALYSIS SYSTEM DESIGN OF HOT PUBLIC OPINION BASED ON WEIBO

Juchao Zeng, *Zhan Wen, Jie Chen, Yanying Li, Dehao Ren

Chengdu University of Information Technology, Chengdu, Sichuan, China, 610225.

ABSTRACT

Under the new situation of the vigorous development of the media, all kinds of social organizations post the public opinion information on various social hot issues by organization authentication, and publishing articles or short videos. Weibo provides institutional authentication to guarantee information sources, and people also use Weibo as an important place for their personal opinions. In this paper, we proposed a scheme to forecast and analyze hot public opinion based on Weibo. We collected data from Weibo first, then dealt with these data for analysis. Multiple Linear Regression model were used to forecast public opinion. It is very useful for many companies and organizations to alert related hot events timely and closely, seize the best time to deal with the crisis, minimize the negative impact, and maintain their own stability.

Keywords: Weibo, hot public opinion, forecast, Multiple Linear Regression.

INTRODUCTION

A report by the China Internet Network Information Center (CNNIC)on February 2021 showed that until December 2020, China has 989 million Internet users, and it became the largest digital society in the world. Driven by the huge amount of traffic on the Internet, netizens form public opinion[1] by making opinions and comments. Just one small thing in people's lives may also be facilitated by the Internet, in particular, under the booster of the micro-blog platform, a deep-seated discussion has developed under the public opinion. As a result, in the digital society with a huge group of netizens, netizens are another identity of the common people, public opinion has also become a major concern for the relevant agencies. The micro-blog platform reflects the public opinion of the vast majority of netizens, effectively forecasts and guides the network public opinion[2], which is conducive to solving the problems that the people care about, and is also of great value in resolving the potential public opinion risks.

GENERAL IDEA OF SYSTEM DESIGN

According to the law of public sentiment evolution, Lan Yuexin's[4] research found that the public sentiment of hot and unexpected events on the internet generally includes the stages of the occurrence of events, netizens' heated discussions, the emergence and spread of public sentiment on Weibo, the intervention of relevant organizations, and the disappearance of events, etc. , we can conclude that the prediction of the trend of public opinion should be combined with the attention and emotional orientation of netizens in the event; the micro-blog platform is a way for the majority of netizens to express their views and comments, and the data volume has sufficient reference significance, and the various labels and characteristics of the data themselves can have a lot of uncontrollable influence on public opinion itself. We keep most of the labels of the data, and through data cleansing and sorting, in this paper, a multiple regression model is established in time series data[3] to improve the accuracy of public opinion prediction through training set and test set.

---

*Corresponding Author: Zhan Wen,
Chengdu University of Information Technology, Chengdu, Sichuan, China, 610225.

SYSTEM SCHEME OVERVIEW

Data acquisition

By writing a crawler program to crawl historical events on Weibo, and taking reasonable comment data segments as the training set in a proper time period, and using the crawler program to crawl recent events as the test set, continue to pay attention to the public sentiment of the events of the test set as the reference forecast results of the system; the content of the data collection includes the publishing organization, the release time, the content of the micro-blog, the comment data (including the related information of the reviewer) , the comment quantity, retweets, likes and so on, collates data tags and categorizes and saves data.

Data text analysis

The collected data were processed with the past stop words and the re-word, and then the emotional polarity of the text was analyzed with the emotion dictionary and normalized, we use sentiment analysis to find out how people tend to react to hot issues; we use data visualization to show data that reflects public sentiment, from the time evolution of visual display of the Netizen's attention to hot events peak and turning changes, and further assist public opinion development analysis.

Statistical Data Processing

The data collected from the public and media participation changes[5], changes in regional migration and changes in the depth of topics, and other dimensions to make quantitative statistics, in the authoritative literature at the same time, the definition of objective reflection of changes in public opinion linear independent variables, make simple linear prediction, provide the basis for multiple regression prediction model.

Construction of multiple regression prediction model

Luo Taiye's[6] logistic regression model predicts public opinion only in the changes of basic data, such as the number of comments and

retweets Liu Ding *et al.,* 's[7] optimization of the model of LSTM has a remarkable effect on the time series, but lacks very important affective tendency parameter and a very important link in the prediction link. After a certain amount of data is obtained, a multiple regression prediction model which may be helpful to public opinion prediction is established by combining the data of different dimensions of netizens' attention degree and the normalization parameters of emotional tendency First of all, we need to make sure that we have enough data in the data set, and we need to select the typical hot issues in society to make sure that we have enough data. Second, we need to compare and screen the different kinds of data labels that reflect the interest of the screen name, the influence of

different factors on the development of public opinion is detailed, and finally, the parameters are obtained by comparing or distributing the percentage of influence.

Web development showcase

Development of web interface, through the input of hot events to show the development of the event and public opinion forecast. Multiple regression model is used to predict the trend of hot events and reflect the change of public opinion information, so as to provide auxiliary information for decision-making.
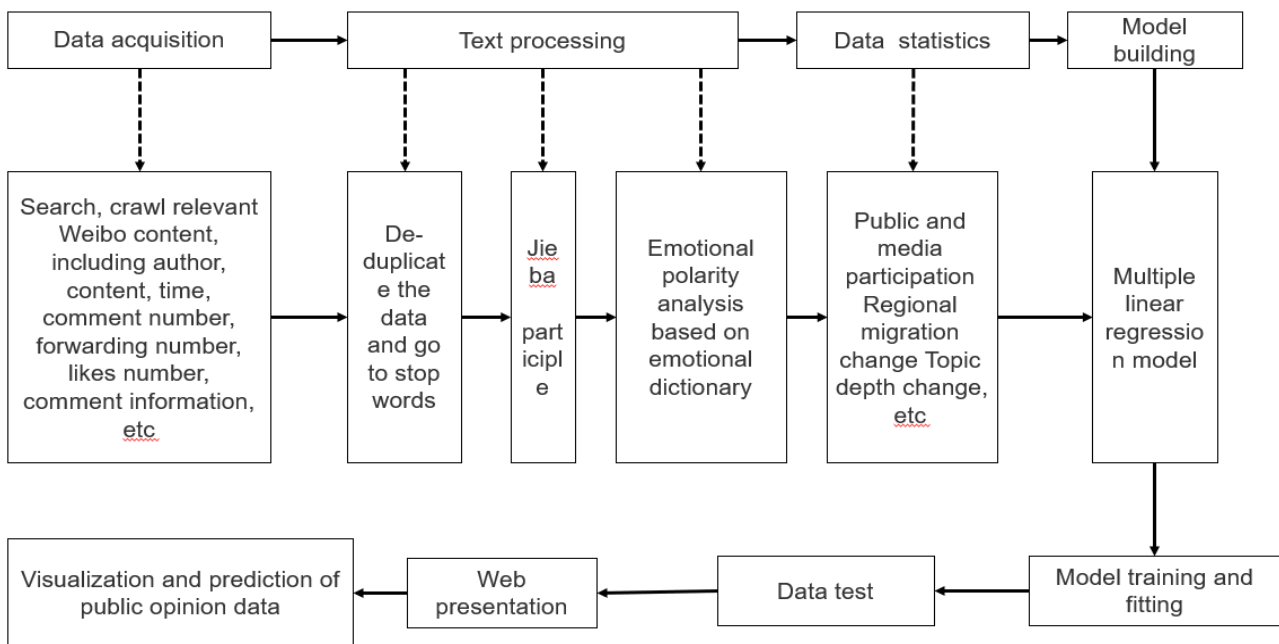

Fig. 1. Flow chart of the design scheme.

CONCLUSION

This paper aims at the premise that Weibo hotspot events may have a certain negative impact on relevant institutions, and crawlers crawl the data of Weibo hotspot events, through sentiment analysis and refining the attention of netizens to establish a multi-regression model for public opinion prediction, and finally, when entering relevant data in the web interface, to achieve data visualization and public opinion prediction of hot events, to help the relevant agencies to their own hot public opinion to make better decisions.

Acknowledgments

REFERENCES

1. Uthirapathy, S.E., Sandanam, D. (2020). Real-time opinion prediction method for emergency public events in social media networks using opinion hit matrix. Revue d'Intelligence Artificielle, Vol. 34, No. 4, pp. 507-514.
2. Edmond Awad, Michael Anderson, Susan Leigh Anderson, Beishui Liao, An approach for combining ethical principles with public opinion toguide public policy, Artificial Intelligence, Volume 287,2020,103349,ISSN 0004-3702.
3. Jingyang Wang, Xiaolei Li, Jiazheng Li, Qiuhong Sun, Haiyao Wang,NGCU: A New RNN Model for Time-Series Data Prediction, Big Data Research, Volume 27,2022,100296,ISSN 2214-5796.
4. Lan Yuexin, Wang Fang, Dong Xilin, etc. . A study on the model of network public sentiment heat of public crisis events[J]. Information Science, 2016,34(2) : 32-36.(in Chinese)
5. Huang jiaoping. Early prediction analysis of negative hot news based on Weibo [D] , 2013.(in Chinese)
6. Luo Taiye. Research on the prediction of popular sentiment of micro-blog based on Logistic model. Forum on statistics and information, 2017,32(10) : 91-95.(in Chinese)
7. Liu Dingyi, Shenyang, Zhan Tianming, *et al.,* . A network consensus prediction method based on Weibo hotspot analysis and LSTM model [J] . Proceedings of the Jiangsu University: Natural Science, 2021,42(5) : 546-553.(in Chinese

****