

Effectiveness of China's measures
during the COVID-19 epidemic from the
perspective of public opinion
research

Abstract

In this paper, through web crawler technology, I get the epidemic data, Baidu hot search, Sina Weibo hot search comments, video comments, CCTV News data, online shopping data, stock data, etc. Using Python's data processing library and visualization library to visualize the data. In this paper, I use the method of emotion analysis based on dictionary to analyze the crawled public opinion data. After constructing emotion dictionary, negative word dictionary, degree adverb dictionary and stopping word dictionary, I use Hidden Markov Model (HMM) to segment words, TF-IDF algorithm and TextRank algorithm to extract feature words, and use self-defined way to extract text emotion according to the dictionary propensity calculation. Combined with the time development and government measures, this paper analyzes the impact of the epidemic situation and the measures of the Chinese government from multiple perspectives.

Keywords: Internet Public Sentiment Web Crawler Feature Extraction
Sentiment Analysis Visualization Analysis COVID-19

Catalog

Abstract.....	1
1 Introduction.....	3
2 Related theory and technology.....	4
2.1 Web crawler.....	4
2.2 Data analysis.....	4
2.2.1 Visual analysis.....	4
2.2.2 Text sentiment analysis.....	4
2.3 Web application development of Python.....	4
3 Design and Implementation.....	5
3.1 Overall design.....	5
3.2 Data acquisition.....	5
3.3 Public opinion analysis.....	5
3.3.1 Baidu public o pinion.....	5
3.3.2 Sina Weibo public opinion.....	5
3.3.3 News data analysis.....	6
3.3.4 Summary.....	7
3.4 Analysis of returning to school.....	7
3.4.1 University distribution.....	7
3.4.2 Online course software.....	8
3.4.3 Analysis.....	10
3.5 Economic data analysis.....	10
3.5.1 Relationship between GDP and epidemic situation.....	11
3.5.2 Stock trend analysis.....	12
4 Summary and Prospect.....	13

1 Introduction

In the outbreak of COVID-19, the changes in the epidemic situation always attracts our attention. The Chinese government took measures to deal with it. With the development of time, the epidemic was under control in China. The resumption of work and rehabilitation began in succession. Meanwhile, the outbreak of the epidemic was abroad, and the multinational government should be adversely affected. The Chinese government continued to strictly control and prevent the epidemic in China once again.

The Chinese government's control of the epidemic twice is inseparable from the cooperation of policies and measures with the people. Public opinion effectively reflects the implementation of policies and measures. In the current era of Internet popularization, the analysis of network public opinion data can best reflect the real views and emotions of the masses, and the real effect of epidemic prevention measures can be seen. In order to summarize the experience and carry out the work of fighting against the epidemic situation better, I use web crawler technology, data visualization technology and text sentiment analysis technology to analyse public opinion data.

2 Related theory and technology

2.1 Web crawler

Web crawler is a kind of program or script, which can automatically grab the Internet information according to certain rules. It is widely used in the field of Internet search engine. It can automatically collect and obtain the accessible web content, and then provide it to the search engine for further processing, so that users can retrieve the information they need more quickly.

There are six components of web crawler: main control module, web page download module, web page analysis module, the first mock exam module of URL, data cleaning module and data display module.

2.2 Data analysis

2.2.1 Visual analysis

Visual analysis can explain COVID-19's development and related influence, and provide relevant clues and guidelines for epidemic prevention and control strategy.

As an object-oriented interpretive programming language, Python has many visual libraries, such as bokeh, plotly, Matplotlib, Seaborn, pandas, wordcloud and pyechart. In this paper, we mainly use pandas, wordcloud and pyechart.

2.2.2 Text sentiment analysis

Text sentiment analysis is a sub domain of natural language processing. Emotional information in text data is extracted by data mining through computer language. The current research directions include subjective identification, emotional tendency calculation, emotional information classification, etc. the research levels are divided into word level, sentence level and text level according to the granularity of the text to be processed.

This paper uses the emotion analysis method based on dictionary, the steps are as follows: dictionary construction, text preprocessing and affective tendency calculation.

2.3 Web application development of Python

Python has a variety of web development frameworks, this paper uses Python Flask for graduation design.

Flask is a Python based web development framework, with three-tier architecture design, including control layer, business layer, data layer, suitable for the development of lightweight web applications. The structure of flask is extensible, and it is convenient to add other functions.

There are three layer architecture design of Python Flask: control layer, business layer and data layer.

3 Design and Implementation

3.1 Overall design

This paper is divided into six parts: epidemic map, epidemic development, public opinion analysis, study resumption analysis, online shopping analysis and economic analysis. The data of the six parts are analyzed visually. Combined with the development of time and the implementation of measures, the data changes are compared and analyzed to test the effectiveness of policy measures to deal with the epidemic situation.

3.2 Data acquisition

Using web crawler technology, crawling data from official website or API interface, and processing data to obtain target data. The method of data acquisition is legal, and the data source is official and scientific.

3.3 Public opinion analysis

3.3.1 Baidu public opinion

As a Chinese Internet search portal, Baidu covers more than 95% of domestic netizens. To a certain extent, Baidu hot search has become an important reference to reflect the most real and objective concerns of our people.

Crawl Baidu national hot search and Baidu area hot search, according to the hot search content to analyze the issues that people are most concerned about, and look at the changes of people's concerns according to the development of time.

Based on the analysis of Baidu hot search data and regional hot search data from January 2020 to April 2020, we can see that the hot search data are basically the same all over the country, and people in Hubei Province pay more attention to the epidemic.

When the epidemic is serious, some people are pessimistic, and on the whole, the masses generally hold a positive attitude towards the epidemic. With the development of time and the implementation of measures, the attitude of the masses on the epidemic is becoming more and more optimistic and positive.

3.3.2 Sina Weibo public opinion

Sina Weibo is similar to twitter, which is widely used by Chinese people. It has a wide influence. In the process of fighting against epidemic situation, Sina Weibo resolutely implements the important directive spirit of President Xi on epidemic prevention and control work. It sticks to the correct political direction, public opinion orientation and value orientation in epidemic prevention propaganda work, and over 2,6000 government official micro-blog participated in the epidemic information release. More than 550000 related Weibo have been published, with more than 11.4

billion Weibo read.

Compared with news, blogs, forums and other types of users, Weibo users can get the latest information more easily. Hot search on Weibo can see current hot topics, and using Weibo has a strong sense of interaction and participation. Can express and respond in a timely manner, which fully reflects the user's right to speak.

Therefore, the analysis of Weibo public opinion will help to timely judge and guide the factors that form group problems and effectively control social public opinion.

Using web crawler technology, crawl Weibo under the specified hot topic of Weibo within a specified period of time, analyze the emotion of Weibo under this topic, and understand people's mentality to a topic.

This paper analyzed many popular searches on Weibo, such as "The construction of Wuhan Xiaotangshan Model Hospital", "COVID-19" and "Health release", as well as the comment areas of the popular Weibo. We can see that during this epidemic, people paid high attention to the epidemic. With the continuous release of policies and measures, all parties actively cooperate, the mentality of people from the initial fear and hope, to hope to overcome the epidemic.

Thus it can be seen that the state media have implemented the spirit of President Xi's instructions, released authoritative information in a timely manner, disseminated scientific methods, and responded to hot issues, and played an important role in guiding public opinion. It can be seen that the policy measures are more effective at the same time.

3.3.3 News data analysis

During the epidemic period, news is our main source of information. Among the news media, CCTV has a wide audience and is authoritative as an official media. Therefore, the news data analysis refers to CCTV news articles.

Use web crawler technology to crawl all CCTV news releases from December 2019 to April 2020 for data analysis. Use TextRank Algorithm to extract keywords from the news release, record the global word segmentation and count the word frequency. Select the keywords that appear more frequently to make a cloud chart to reflect the main news content of the month.

According to the data analysis, it can be seen that:

In December 2019, Wuhan Health Commission reported that there were patients with pneumonia of unknown cause. According to the news analysis of that month, we can see that it did not cause high attention. In January 2020, at the beginning of January, Wuhan continuously reported patients with unexplained pneumonia, which attracted the attention of all parties. On January 15, the National Center for Disease Control and prevention launched level I emergency response, and the news will release the latest news of pneumonia every day. In February 2020, with the continuous development of the epidemic situation in China, the first level emergency response has been launched one after another. Due to the shortage of medical materials, Hubei Province is in a difficult situation, and medical teams have been sent to support them. The resumption of work and school needs to be postponed. In March

2020, with the policy intervention and the people's cooperation, the epidemic situation reached its peak and began to decline gradually. The resumption of work and production was carried out in an orderly manner according to the epidemic situation in various regions. In April 2020, the epidemic situation in foreign countries worsened. China sent medical teams to carry medical materials to help countries and donated medical materials and equipment to help countries. In China, the resumption of work and production was carried out in an orderly manner.

3.3.4 Summary

After the COVID-19 epidemic appeared in China, people had panic in the face of this unknown infectious disease. The Chinese government promptly took a series of measures, such as isolation, city closure and testing. The public responded positively and had the confidence to overcome the epidemic. With the long-term cooperation of the government and the people, the epidemic situation was gradually under control, and the people were full of hope for overcoming the epidemic situation.

3.4 Analysis of returning to school

3.4.1 University distribution

The college entrance examination website uses web crawler to crawl the college information, and the crawling content is saved as a college_data.csv. Visual analysis of the data was performed using pyecharts, as shown in figures 3.1 to 3.3.

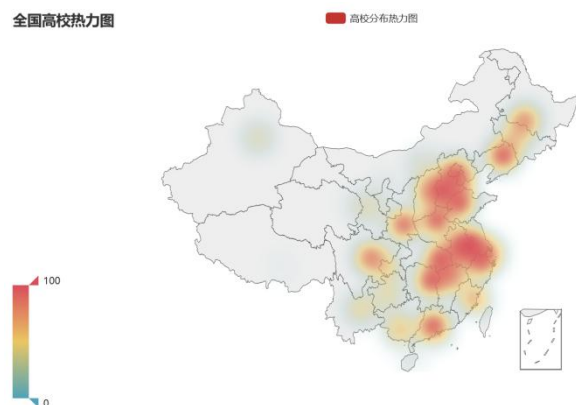


Figure 3.1 thermal diagram of National Universities

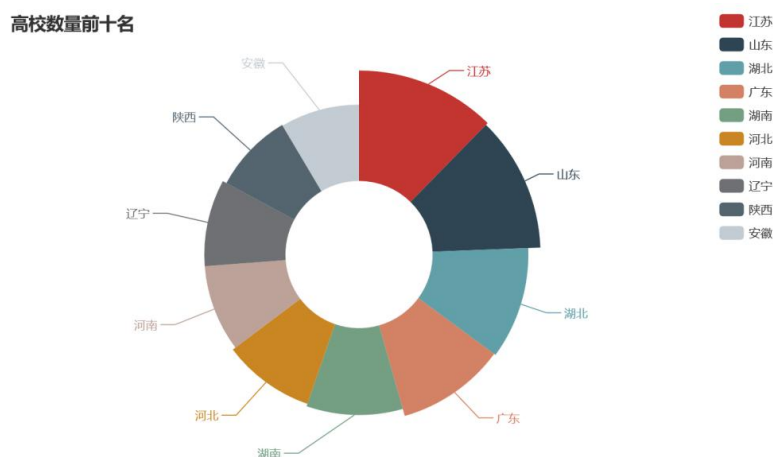


Figure 3.2 top ten universities

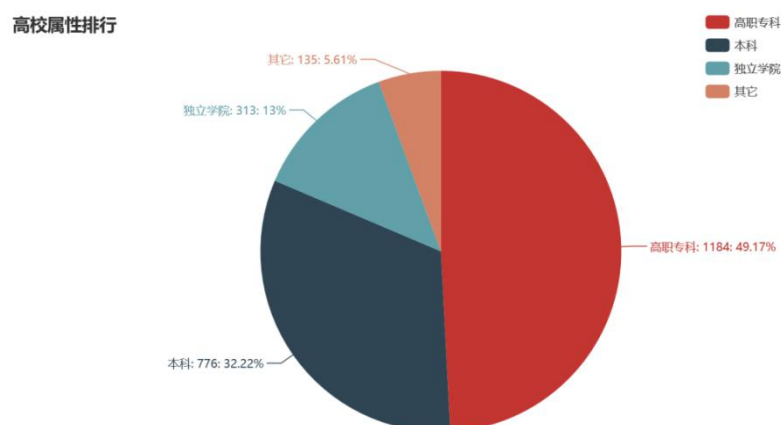


Figure 3.3 ranking of University attributes

From the distribution of the number of colleges and universities in China, the number of colleges and universities in the eastern developed areas is more. In terms of the quality distribution of colleges and universities in China, the top three universities in 211 are Beijing, Jiangsu Province and Shanghai, and the top three universities in 985 are Beijing, Shanghai Province and Shandong Province. Among them, the number of 211 universities in Hubei Province ranks the fifth, and the number of 985 universities ranks the 11th.

3.4.2 Online course software

During the epidemic period, the online teaching method of "suspension of classes without suspension" and "cloud classroom" has become the mainstream in special period. There are many apps for online teaching, and the quality of software varies. Among the first batch of registered educational applications released by the

Ministry of education, nailing was selected. Since then, nailing has become the online course platform choice of most schools. Using web crawler technology to crawl Apple's app store's nail rating and comments, and analyze the status quo of the user experience of the nail.

According to the 500 pieces of data in the first 10 pages, the number of people corresponding to one star, two stars, three stars, four stars and five stars are 320, 14, 10, 21 and 135 respectively. According to the content of the comment area, the proportion of comments with keywords is counted, as shown in Figure 3.4.

```
--在爬取的前10页500条评分数据中
--钉钉1-5评分分值出现的次数:
(array([1, 2, 3, 4, 5]), array([320, 14, 10, 21, 135], dtype=int64))
--统计含主要关键词的评论出现的比例
----五星出现的比例为: 7.60%
----一星出现的比例为: 8.80%
----好评出现的比例为: 9.40%
----差评出现的比例为: 2.80%
----老师出现的比例为: 2.80%
----学生出现的比例为: 6.60%
----分期出现的比例为: 5.80%
----下架出现的比例为: 5.40%
----垃圾出现的比例为: 5.40%
----加油出现的比例为: 1.40%

Process finished with exit code 0
```

Figure 3.4 proportion of scoring keywords

According to the crawling comments, the word segmentation technology and the stop word dictionary are used to screen out the valuable words. After counting the word frequency, the word cloud is drawn, as shown in Figure 3.5.



Figure 3.5 word cloud

According to the analysis, nail software quality is higher, user experience is

better, some student users are dissatisfied with the compulsory download of the school, they vent their emotions on the software score, resulting in low software score. Furthermore, it can be seen that some students have resistance to online classes, and the normalization process of online teaching needs to be further explored.

3.4.3 Analysis

The types of colleges and Universities Participating in the statistics include undergraduate, independent college, higher vocational college, etc. The research shows that local students are more likely to go to local school. The key universities represented by 211 and 985 have a wide range of students, with a large number of foreign students. Kindergartens, primary schools, junior middle schools, senior high schools and technical secondary schools have many local residents. Therefore, the student structure can refer to the situation of local non-key universities.

In view of the differences in student structure, local epidemic situation, physical quality and age, and enrollment status, the regions meeting the three conditions given by the Ministry of education can consider resuming school. The third year of junior high school and the third year of senior high school will return to school first, and the students and teachers with important scientific research tasks will return to school first, and other grades will be determined. The time of resumption of 985 and 211 universities will be determined. It should be discussed carefully in light of the specific situation.

President Xi stressed at the Standing Committee meeting of the Political Bureau of the CPC Central Committee in May 14th that we should strengthen the anti-rebound work in key areas and key places, and highlighted the epidemic prevention work in Hubei Province, Beijing, Heilongjiang Province and Jilin Province. The number of 985 and 211 colleges and universities in Beijing ranks first in China, which coincides with the two sessions. Beijing should be more careful in determining the time for returning to school than other regions. Wuhan, Hubei Province, as the source of domestic epidemic, should continue to strictly prevent and control. Up to the time of the meeting, the time for the resumption of universities in Heilongjiang Province has not yet been determined, and the grade of Jilin Province's reunification has been changed to online teaching. Since June 6th, universities in Beijing have qualified colleges and universities to arrange graduating grades to return to school in succession. The time for the resumption of universities in Hubei Province has yet to be determined, and the above decision is in line with the spirit of President Xi's speech.

3.5 Economic data analysis

The epidemic has had a great impact on the economy. This year's government work report did not put forward a specific target for economic growth. Premier Li Keqiang pointed out in the report that the reason for this is that the global epidemic situation and economic and trade forms are highly uncertain, and economic development is affected by some unpredictable factors. Therefore, it is necessary to analyze the economic data affected by the epidemic situation.

3.5.1 Relationship between GDP and epidemic situation

Gross Domestic Product (GDP) is often used to measure the economic situation of countries. This paper intends to explore the relationship between GDP and the epidemic situation. Countries with GDP over 0.1 trillion in 2019 are selected to make GDP map according to the amount, and epidemic map is drawn according to the epidemic data on March 25, 2020, as shown in figures 3.6 and 3.7 .

2019年GDP达到0.1万亿的国家 (数据单位: 万亿)

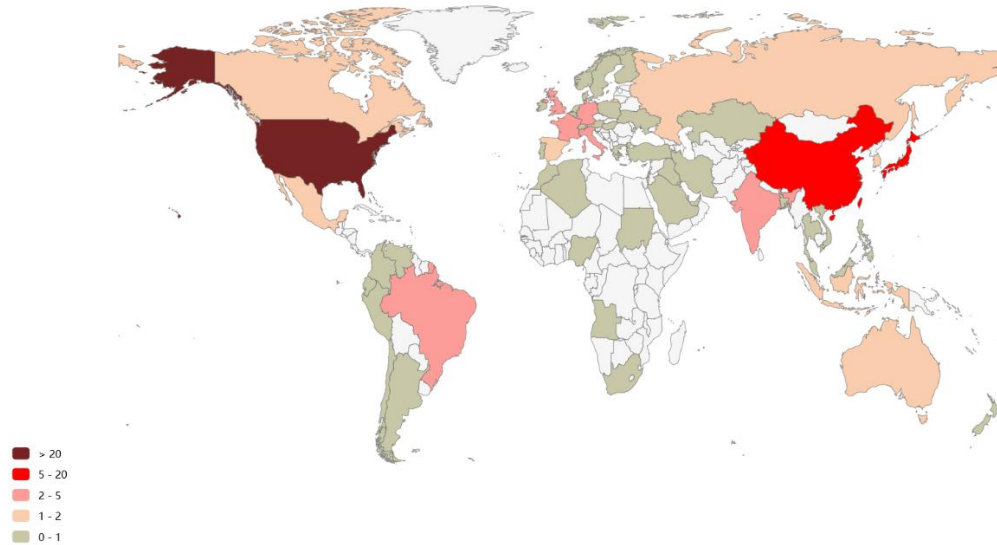


Figure 3.6 GDP map

2019年GDP在0.1万亿以上国家的2020年3月25日疫情确诊情况

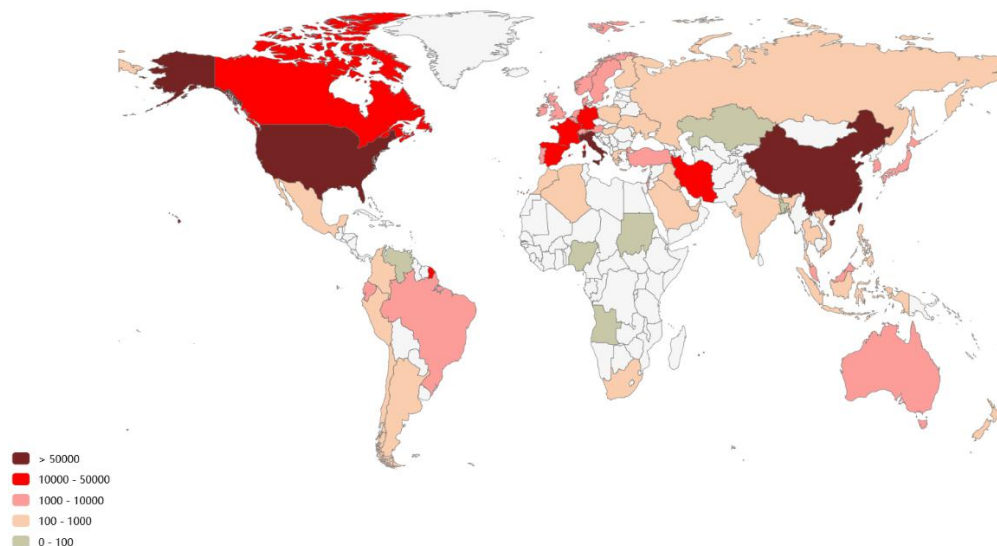


Figure 3.7 Epidemic map

It can be seen that there is a positive correlation between the GDP and the diagnosis of the epidemic situation. Considering the factors such as the total population, population density, medical level, traffic level and policy measures,

countries with higher GDP generally have larger population and population density, generally better traffic conditions, and greater risk of infection among the masses. As the medical level is generally high, they have the ability to treat more people. Therefore, the number of confirmed cases is high.

3.5.2 Stock trend analysis

The outbreak of COVID-19 in China has affected the financial market, and has seriously affected investor confidence. The stock market is under great pressure. In the overall economic downturn, the pharmaceutical industry is optimistic because of the epidemic situation. This paper makes a visual analysis on the representative medical stocks. Moving average and daily trading volume are important indicators to judge the stock market. This paper uses these two indicators to visually analyze the stock market data.

Moving Average (MA) is commonly used to observe the trend of securities price changes, and to count the securities prices within a certain period of time. After averaging the prices, the average value can be connected to form a MA line. MA5 is the five day moving average, MA10 is the ten day moving average, and MA20 is the twenty day moving average, which represent the weighted average prices of stocks on five days, ten days and twenty days respectively.

Stock trading volume refers to the specific number of transactions in a certain period of time, and the daily trading volume reflects the trading situation of the stock on that day.

To facilitate the analysis of the impact of the epidemic on the stock market, python flask is used to create a small stock trend query website. The data is updated in real time according to the date.

According to the MA and stock trading volume of the last four months:

TEDA shares is the largest manufacturer of mask filters in China. After the outbreak of the disease, the demand for respirators increased dramatically. Huahai Pharmaceutical Co., Ltd. imitated the new crown specific drug "redcivir" and put it into clinical trial, and its stock rose sharply. On February 11, Bray Pharmaceutical Co., Ltd. announced that it had successfully imitated redcivir, and its stock rose sharply. On February 14, the Shanghai Stock Exchange announced that it would focus on monitoring the concept stocks related to epidemic prevention and control. After that, Borui pharmaceutical entered a shock correction, but its cumulative increase was still more than 40%.

According to the above analysis, the biomedical industry was originally a weak cycle industry with long-term and stable growth. During the epidemic period, due to the demand relationship, it was chased by capital, and was also valued by the public. Some financial products began to launch "robust medical portfolio series", and investment in biomedical industry became a new trend of investment and financing.

During the epidemic period, from the rise of epidemic related stocks, especially domestic vaccine related stocks, we can see that the Chinese people are full of confidence in overcoming the epidemic situation.

4 Summary and Prospect

The COVID-19 epidemic will inevitably have a great impact on the economy and society. There are inherent contradictions between epidemic prevention and control and economic development. The Chinese government has always insisted on taking the people as the center and immediately implemented strict isolation and closure measures at the initial stage of the outbreak. When the spread of the epidemic has gradually been controlled, it has continued to strictly control the epidemic situation by adopting measures of classification and precise implementation. At the same time, various departments and governments at all levels have introduced a series of policies and measures to promote the economic order to return to normal, providing a set of Chinese experience and scheme for the countries affected by the epidemic.

The main work of this paper is summarized as follows:

(1) In the outbreak of COVID-19, in this graduation project, according to the epidemic situation data and the public opinion data and economic data related to the epidemic situation, according to the research background and significance, related theories and technology, feasibility analysis, demand analysis, design and implementation steps, we used the empirical research method, quantitative analysis method, case study method and interdisciplinary analysis method. A variety of methods, closely linked to current events, combined with time development, spatial transformation, policy measures, from the visual analysis, we can see that the epidemic situation is gradually controlled, public opinion is gradually turning to positive, and the economic situation is gradually becoming stable. Compared with the previous epidemic related papers, this paper's analysis is broader, the time span is larger, and more truly reflects the impact of the epidemic.

(2) In the process of data collection, this paper uses web crawler technology to crawl and collect various epidemic related data, such as epidemic data, Baidu hot search, Sina Weibo hot search and comments, video comment on-screen, CCTV News data, online shopping data, stock data, etc.

(3) In this paper, the crawled data is visually analyzed, and some classic visual libraries of Python are used to analyze and process the data, which makes the results more intuitive.

(4) This paper analyzes the text sentiment of the crawled public opinion data, and adopts the sentiment analysis method based on the dictionary. After constructing the sentiment dictionary, negative word dictionary, degree adverb dictionary and stop word dictionary, the Hidden Markov Model (HMM) is used for word segmentation, TF-IDF Algorithm and TextRank Algorithm are used for feature word extraction, and the emotional tendency of text is calculated according to the dictionary in a user-defined way.

(5) This paper uses Python Flask, combined with web crawler technology Flask Sqlalchemy database function to design and develop a small stock query website, which is convenient for stock data query according to time and stock name.

After being controlled in China, COVID-19 broke out again in Beijing, Jilin Province and Xinjiang Province, and was controlled again. Other countries can

control the epidemic situation only after adopting China's epidemic prevention experience. Up to now, China's measures are the only effective way to control the epidemic. In some countries, such as the United States and India, leaders pay little attention to the life and health of the people, resulting in the increasing number of COVID-19 patients and deaths. Although the appearance of COVID-19 is a natural disaster, the consequences of COVID-19 are actually man-made disasters.