

Design and Implementation of a Microblog Public Opinion Visualization System Based on Flask and ECharts

Yun Liu and Dong Xiao

Information Center, Hainan Vocational College of Political Science and Law, Haikou 571100, China

Abstract. Conducting public opinion analysis on Weibo data is of significant importance to decision-making processes in government, enterprises, and personal contexts. This paper presents the design and implementation of a real-time microblog public opinion visualization system based on Flask and ECharts. The system aims at effectively monitoring Weibo data in real time, performing sentiment analysis, and trend prediction. The system adopts a B/S architecture, employing distributed web crawling techniques to efficiently gather Weibo data. For sentiment analysis, the SnowNLP library is utilized. On the frontend, the ECharts charting library dynamically displays the sentiment data, offering users an intuitive interface and interactive experience. In designing the system, considerations have been given to data processing efficiency, real-time capabilities, scalability, and data security to cater to the diverse needs of different users. By integrating these technologies, the system provides comprehensive insights into the prevailing sentiments, trending topics, and patterns within the vast volume of Weibo content.

Keywords: Weibo public opinion Analysis, Data Visualization, Sentiment Analysis.

1. Introduction

With the rapid development of Internet technology and the widespread application of social media, Weibo has become an important channel for the public to express opinions and obtain information. The immediacy and widespread nature of Weibo make information dissemination extremely fast, and public opinion dynamics change rapidly, which poses new challenges to public opinion monitoring and analysis. The analysis of Weibo public opinion can help governments, enterprises and individuals understand the focus of public attention, predict possible social trends, and make more informed decisions. Especially in public emergencies, the rapid response and widespread dissemination of public opinions on Weibo are particularly critical for emergency management of events and guidance of public opinion. Therefore, researching and developing a system that can monitor, accurately analyze and efficiently visualize Weibo public opinion in real time has important practical significance and far-reaching social value.

As a key technology for understanding the dynamics of social public opinion, public opinion visualization has attracted widespread attention from researchers around the world. Abroad, the research on public opinion visualization has become relatively mature. Researchers use various data mining and infographic technologies to conduct in-depth analysis of social media data. For example, American scholars build complex information dissemination models to conduct real-time monitoring and visual display of public opinion data on platforms such as Twitter, helping governments and enterprises respond to public concerns in a timely manner (Neppalli et al., 2017). The European research team focuses on the integration and analysis of cross-platform public opinion data, providing a more comprehensive public opinion analysis perspective by integrating data from different sources (Moody, 2016). Domestically, scholars have made significant progress in public opinion monitoring and analysis, especially in the mining and visualization of Weibo public opinions. For example, some research focuses on sentiment analysis and topic discovery of Weibo public opinions, by applying natural language processing technology to extract public emotions and focus of attention, and display the analysis results in the form of charts (Liu Yashu et al., 2019). In addition, domestic researchers have also explored public opinion visualization methods that combine geographical information systems (GIS) and big data analysis technology to display the evolution of public opinion in multiple

dimensions such as maps and timelines, providing a new perspective for the analysis of the spatiotemporal characteristics of public opinion (Zhang Jiaomeng et al., 2022). Although certain research results have been achieved in public opinion visualization at home and abroad, there are still some challenges and shortcomings. For example, issues such as how to effectively process and analyze large-scale unstructured public opinion data, how to improve the accuracy and real-time nature of public opinion analysis, and how to design more intuitive and easier-to-understand visual charts all require further research and exploration. In addition, with the diversification of social media platforms and the complexity of information dissemination methods, how to comprehensively utilize multi-source data for public opinion analysis is also an important direction of current research.

In view of the importance of Weibo public opinion analysis and the shortcomings of existing research, this article aims to design and implement a Weibo public opinion visualization system based on Flask and ECharts. This system efficiently collects Weibo data through distributed crawler technology, uses advanced data processing algorithms for data cleaning and analysis, and achieves intuitive visual display of data through the ECharts chart library. The main contributions of this article include: proposing a public opinion analysis framework that combines back-end data processing and front-end visualization; realizing real-time monitoring and dynamic display of Weibo public opinion data; providing a user-friendly operation interface so that non-professional users can Easily conduct public opinion analysis and decision support. Through this study, we hope to provide new perspectives and tools for the analysis and application of Weibo public opinion, thereby providing scientific basis for social management and public decision-making.

2. System requirements analysis

Weibo public opinion has the characteristics of large data volume, fast update speed, rich content and obvious emotional tendency. User demands for the Weibo public opinion analysis system mainly focus on the following aspects: first, the system can process and analyze large-scale Weibo data, including text content, user information, timestamps, etc.; secondly, the system can provide real-time public opinion Monitoring and early warning functions to quickly respond to hot social events; in addition, the system can conduct in-depth analysis of Weibo public opinion, such as sentiment analysis, topic mining and trend prediction; finally, the system has a friendly user interface and visual display functions, making Non-professional users can also intuitively understand the public opinion analysis results.

Based on the analysis of user needs, this system should have the following functions: first, realize the automatic collection of Weibo data, including blog posts, comments, reposts, user information, etc.; secondly, have a data preprocessing function that can clean and remove the collected data. and format conversion; thirdly, the system should include a public opinion analysis module to implement functions such as sentiment analysis, topic mining, and trend prediction; in addition, the system should provide real-time monitoring and early warning mechanisms to promptly discover and notify users of social hot spots of concern; finally, the system It should have the function of data visualization and display the results of public opinion analysis intuitively through charts and other forms.

In terms of performance, first of all, the system should have high-efficiency data processing capabilities and be able to quickly process and analyze large-scale Weibo data sets; secondly, the system has high real-time requirements and should be able to monitor and analyze Weibo data streams in real time. analysis; in addition, the system should have good scalability and be able to adapt to the growing amount of data and emerging analysis needs; finally, the system should ensure the security and stability of data and protect user data from unauthorized access and loss.

In summary, the design and implementation of this system will be optimized around user needs and system performance to ensure that the system can effectively serve the monitoring, analysis and visual display of Weibo public opinion.

3. Introduction to related technologies

3.1. Flask framework

Flask is a lightweight web application framework written in Python. Its core features are simplicity and flexibility, which allows developers to quickly build small projects and can also serve as the infrastructure for large applications (Grinberg, 2018). Flask mainly relies on the Werkzeug WSGI toolbox and Jinja2 template engine, which provide Flask with a stable WSGI environment and powerful template rendering functions. The Flask framework adopts a modular design, and developers can choose to install various extensions according to their needs, such as Flask-SQLAlchemy for database operations, Flask-Login for user authentication, etc. In addition, Flask supports RESTful request distribution, making building RESTful APIs simple and fast.

3.2. ECharts chart library

ECharts is an open source data visualization library based on JavaScript, which provides developers with rich chart types and highly customizable visualization options. ECharts supports multiple data formats such as JSON and arrays, and can be flexibly combined with various data sources. ECharts has powerful interactive capabilities, including data zooming, roaming, dragging, etc., as well as rich chart styles and themes, making data display more intuitive and vivid. ECharts also provides detailed API documentation and rich configuration items, allowing developers to customize charts according to needs to meet the visualization needs in different scenarios.

3.3. Other related technologies

Ajax (Asynchronous JavaScript and XML) is a technology that exchanges data with the server and updates parts of a web page without reloading the entire page. Through Ajax, asynchronous transmission and partial updating of data can be achieved, improving the response speed and user experience of Web applications. Ajax is usually used in combination with technologies such as XML, HTML, CSS and JavaScript to achieve dynamic content loading and user interaction.

MySQL is a relational database management system based on SQL (Structured Query Language) language and provides powerful data storage, query and processing functions. MySQL is widely used in applications of all sizes, from small personal projects to large enterprise-level applications. Its features include high performance, reliability, ease of use, and good security. MySQL supports multiple operating systems and programming languages, and can be seamlessly integrated with Web frameworks such as Flask to provide stable data storage solutions for Web applications.

The above technologies provide a solid technical foundation for the Weibo public opinion visualization system proposed in this study, enabling the system to efficiently process data, provide real-time public opinion monitoring and analysis, and display public opinion analysis results through intuitive charts.

4. System architecture design

4.1. Overall architecture of the system

This Weibo public opinion visualization system adopts a layered B/S architecture model to support efficient data processing and intuitive user interaction. The overall architecture of the system consists of three main layers: front-end display layer, back-end logic layer and data storage layer. The front-end display layer is responsible for direct interaction with users and provides data visualization interface and user operation functions. In order to achieve cross-platform compatibility and responsive user experience, the front end is built using technologies such as HTML5, CSS3 and JavaScript, and uses the ECharts chart library to achieve dynamic visual display of data. As the data processing center of the system, the back-end logic layer uses the Flask framework to build RESTful services, handle user requests from the front-end, and call the corresponding data processing modules.

The back-end logic layer is also responsible for interacting with the data storage layer, performing data addition, deletion, modification and query operations, and feeding the processing results back to the front-end. The data storage layer uses the MySQL database and is responsible for storing original Weibo data, user information, and analyzed and processed public opinion data. The database design takes into account the requirements for efficient storage and fast query in a big data environment, ensuring that the system can process and store large amounts of Weibo data.

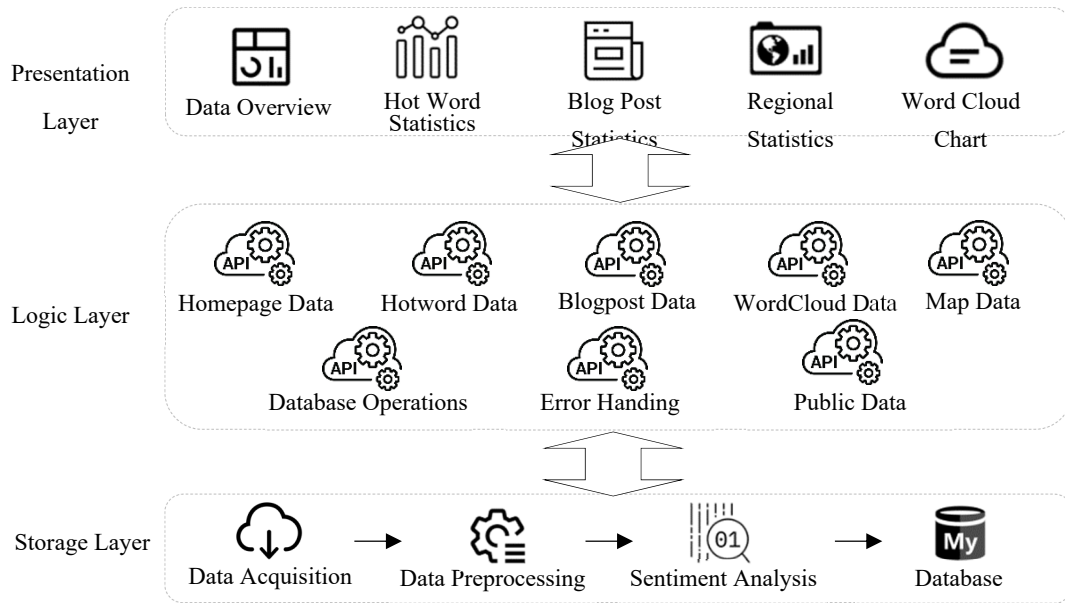


Figure 1. System Functional Architecture Diagram

4.2. Front-end design

Front-end design focuses on providing an intuitive and friendly user interface and interactive experience. The user interface adopts a responsive design to ensure that the system has good display effects on different devices. The main interface of the system includes modules such as data overview, article publishing trends, popular blogger rankings, hot word statistics, geographical distribution, word cloud diagrams, etc., and dynamically displays the visualization results of public opinion data through the ECharts chart library. The user interaction design focuses on simplifying the operation process. Through clear navigation and intuitive operation buttons, users can easily perform data query, chart switching, report generation and other operations. In addition, the system also provides customized setting options, allowing users to adjust public opinion keywords, chart parameters and interface layout according to personal needs.

4.3. Backend design

Back-end design is the core of system data processing, and its design goal is to ensure the accuracy and efficiency of data processing. The back-end design of this system uses the Flask framework, which is an ideal choice for this system because of its lightweight and highly scalable nature.

The data processing process is a key part of the back-end design, which includes data collection, pre-processing, analysis and storage. The data collection phase uses distributed crawler technology to obtain original data from the Weibo platform. The pre-processing phase cleans, deduplicates and formats the data. The analysis phase uses natural language processing technology to conduct sentiment analysis and topic mining, and finally stores the processing results. into the MySQL database.

The back-end logic layer has designed multiple API interfaces to handle various requests sent by the front-end. These interfaces include database operation interface , home page data interface , hot word data interface , blog post data interface , word cloud graph data interface , map data interface , public

data interface , and error handling interface . The database operation interface uses the pymysql library to encapsulate the connection, query, update and other related operations that interact with the MySQL database. The homepage data interface is used to obtain homepage data, including counting the number of likes of blogs and comments, city distribution, time distribution, etc., and preparing to generate chart data for the homepage. The hot word data interface is used to count the number and frequency of occurrences of specific hot words in comment data, and to obtain data such as the list of comments related to hot words. The blog data interface classifies blog data and returns relevant statistical results. The word cloud graph data interface is responsible for generating word cloud images by reading the comment text content and using the WordCloud library to create word clouds, and then converting the generated word cloud images. The map data interface obtains the total number of likes in different regions in the comment data by querying the database, and returns data for map visualization. The public data interface provides functions for obtaining public data such as comments, articles, blogs, etc. These data are usually used for data processing and analysis in other modules. The error information processing interface is used to render an error page and display error information when an error occurs. These interfaces return data in JSON format for visual display by the front end. The visualization implementation framework is shown in Figure 2.

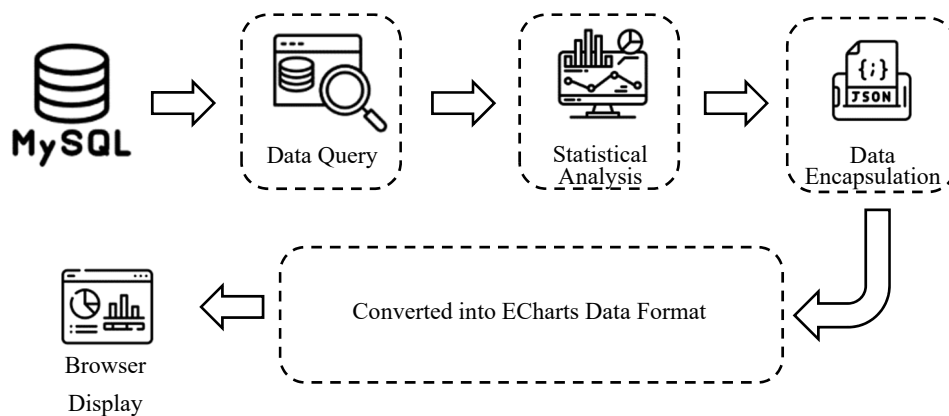


Figure 2. Visualization Implementation Framework

In terms of database design, the system uses the relational database MySQL and designs a reasonable database model to store Weibo data, user information and analysis results. The database's table structure is optimized to support efficient data retrieval and complex query operations. At the same time, the system also implements database security mechanisms, including data backup, recovery, and access control, ensuring data security and integrity.

5. System function implementation

5.1. Implementation of data collection and preprocessing module

Data collection is the basis of the Weibo public opinion visualization system, which directly determines the accuracy and effectiveness of subsequent analysis. This system uses distributed crawler technology and combines Weibo's open API interface to achieve efficient collection of Weibo data. The crawler module is designed to be multi-threaded and can handle a large number of requests at the same time to ensure the real-time and integrity of data collection. The collected Weibo data includes text content, user information, interactive information such as likes and forwarding, timestamps, etc. These data first need to be preprocessed before they can be used for subsequent analysis. The preprocessing module mainly includes steps such as data cleaning, deduplication and format conversion. Data cleaning mainly removes irrelevant information, such as hyperlinks, irrelevant topics, etc. Deduplication is to eliminate duplicate data and ensure the uniqueness of the data. Format conversion is to convert unstructured Weibo text into structured data for easy storage and analysis .

Sentiment analysis is the key to this system. By analyzing Weibo text, we can determine its emotional tendency and mark relevant emotional tags. This system uses Python's SnowNLP library to implement this function. SnowNLP uses a pre-built Chinese emotion dictionary. Each entry in the dictionary corresponds to an emotion value. The value of positive emotion words is positive, the value of negative emotion words is negative, and the value of neutral words is zero. By performing word segmentation and part-of-speech tagging on the input text, each word is given an emotional score based on the emotional dictionary. If the word is in the dictionary, its sentiment value is returned directly; otherwise, it is inferred based on the context of the word and the relationship between known sentiment words. The overall sentiment score of the text is obtained by aggregating the sentiment values of all sentiment words in the text. The formula is:

$$S = \sum_{i=1}^n score_i \times weight_i \quad (1)$$

In formula (1), S is the overall sentiment score of the text, n is the number of sentiment words in the text, $score_i$ is the sentiment value of the i -th sentiment word, and $weight_i$ is the weight of the i -th sentiment word.

This system uses the sentiments method of the SnowNLP library to calculate sentiment analysis results for the incoming text. The sentiments attribute returns a score between 0 and 1. The closer the score is to 1, the more positive the sentiment is, and the closer the score is to 0, the more negative the sentiment is. This system marks Weibo texts that are greater than or equal to 0.6 as positive, those that are less than or equal to 0.4 as negative, and others as neutral. The relevant code is as follows:

```
for blog in bloglist:
    item = list(blog)
    value = ''
    if SnowNLP(item[4]).sentiments >= 0.6:
        value = 'Pos'
    elif SnowNLP(item[4]).sentiments <= 0.4:
        value = 'Neg'
    else:
        value = 'Neu'
    item.append(value)
```

The emotional statistical effect of Weibo articles after sentiment analysis is shown in Figure 3.

ID	作者	点赞数	转发数	评论数	内容	发布时间	情感
497996044121338	中医妇儿林波	477	33	63	#中医# 今天劝退了一个来看咳嗽的宝妈。孩子支原体肺炎咳了半个月，听人介绍过来，据推荐者说她的孩子三天就好了。我说个人体质不一样，饮食忌口配合程度，西医药的伤害程度都不一样，不能保证你这个小孩三天治好咳嗽。何况有些不咳了也要再吃一段时间中药，不咳了只是没有咳嗽症状，肺里的内环境还没有那么快恢复。巴拉巴拉说了半天，一定要我说大概多久能好，我说前台退号，本人技术有限，保证不了每个就诊的都能治愈。花钱买的是中医服务，所有的医疗行为都不对疗效负责。如果花钱就一定要保证治好，肿瘤医院的医生已经被砍成十八段了。	2023-12-17	正面
4976335074625482	峰哥一切为了健康	63	33	31	#中医# 治疗支原体肺炎，古方确实有效都是从古代名医那里传下来的。①鼻咽吹。鼻叶、菊花、石膏、杏仁。这四味药加点冰糖，用开水泡喝。可以清肺散气，散热解表。适用于风热感冒引起的咳嗽、发烧、头痛等症。②竹叶石膏汤。竹叶、石膏、甘草。这三味药加点白糖，用开水泡喝。可以清肺泻火。解表清热。适用于里热感冒引起的咳嗽、口臭、痰稠等症。③桑杏汤。桑叶、杏仁、桔梗、石膏。这四味药加点白糖，用开水泡喝。可以宣肺清热，化痰止咳。适用于风热感冒引起的咳嗽、胸闷、呼吸不畅等症。④麻杏石甘汤。麻黄、杏仁、石膏、甘草。这四味药加点白糖，用开水泡喝。可以宣肺解表，清肺泻火。适用于风热感冒引起的咳嗽、发烧、头痛等症。⑤玉屏风散：由黄芪、白术、防风、荆芥、薄荷、升麻、葛根组成，具有益气固表、解表止汗的功效。适用于感冒、表虚自汗、气虚表弱等症。以上5个方子，都是从古代名医那里传下来的，很有效果。如果大家有条件的，可以试一试，但是要注意，不要自己随便买药吃，一定要在医生的指导下用药。关注我每天学一些中医知识，做自己的神医。感谢祖国天下无疾，愿所有人都精神饱满，健康快乐每一天！#国医的精诚力量##感冒发烧咳嗽##养生#口碑好做祖国长城卫士	2023-12-07	正面
4993862888327481	中医田应旭	86	2	2	#中医##鼻窦炎# 发点增长病种，10岁女孩，支原体肺炎后引发鼻窦炎，每天发作额头疼，一诊14剂，每周额头痛轻痛2-3次，发作时间变短，服药后大量排涕，守方吃吃停药即可。这种没啥难度，治过太多，5年前刚出道的时候鼻窦炎头痛如刀劈斧砍的都躺下来过2武汉武汉顺时岐舍中医门诊部	2024-01-25	负面
4988730196628383	植物人史军	66	6	32	#为什么会两个月内频繁发烧#支原体，甲流，乙流，中招了都会发烧，要注意查一下病原体，好在现在的试纸用起来很方便，最好还是不要中招，不仅仅是发烧的事儿，中招之后，咳嗽要持续很长时间我是元旦前那周中的乙流，一周时间好了，但是咳嗽持续到今天才真正缓解	2024-01-10	正面
4988752189463596	圈内皇叔	238	81	82	#为什么会两个月内频繁发烧#甲流乙流，支原体，支原体肺炎，还有普通感冒，都会有发烧的症状。发烧时除了吃药，物理降温也很重要，还要注意补充水喝电解质。快要过年了，大家注意身体啊！	2024-01-10	负面
4988754202988395	圈内皇叔	263	86	86	#为什么会两个月内频繁发烧# 这个冬季，支原体、甲流、乙流像瘟疫似的，你方刚要我登场，甚至还有和我一起登场，交叉感染的情况。不少中招的人出现频繁的发烧情况，大家注意保暖，勤加锻炼，身体一旦出现不适，48小时内就医，按时吃药，还要多喝热水和补充电解质。还有一个月就要过年了，希望大家都能健健康康，平平安安！	2024-01-10	负面

Figure 3. Sentiment Analysis Results of Microblog Posts

5.2. Implementation of data storage module

The data storage module is responsible for storing preprocessed Weibo data and analysis results. This system uses the MySQL database and is designed to include Weibo data tables, user information tables, comment information tables, sentiment analysis tables, word frequency tables, etc. The Weibo data table stores the basic information of Weibo, such as Weibo ID, content, likes, number of retweets, number of comments, publishing time, user information, etc. The user information table records the user's basic information and behavioral characteristics. The comment information table mainly stores Weibo ID, publication time, number of likes, location, commenter's name, commentator's gender, commentator's address, etc. The sentiment analysis table and word frequency table store the emotional tendencies and word frequency calculation results of Weibo respectively .

Table 1. Microblog Post Table

Field	Type	Null	Key	Default
id	bigint	YES		NULL
likeNum	bigint	YES		NULL
commentsLen	text	YES		NULL
reposts_count	bigint	YES		NULL
content	text	YES		NULL
created_at	varchar	YES		NULL
detailUrl	text	YES		NULL
authorAvatar	text	YES		NULL
authorName	varchar	YES		NULL
authorDetail	text	YES		NULL
isVip	bigint	YES		NULL

Table 2. Comment Table

Field	Type	Null	Key	Default
articleId	bigint	YES		NULL
created_at	varchar	YES		NULL
likes_count	bigint	YES		NULL
region	varchar	YES		NULL
content	text	YES		NULL
authorName	varchar	YES		NULL
authorGender	text	YES		NULL
authorAddress	text	YES		NULL
authorAvatar	text	YES		NULL

5.3. Implementation of visual display module

The visual display module is the part of the system that directly interacts with users. It presents complex data analysis results to users in the form of intuitive charts. This system uses ECharts chart library and china.js map library to implement various types of charts, such as bar charts, line charts, pie charts, and area charts. The display content includes article publishing trend charts, popular blogger ranking charts, hot word statistics charts , regional distribution charts , word cloud charts , etc., allowing users to intuitively understand the distribution, trends and correlations of public opinion data. The design of the visualization module focuses on the user's operating experience and the interactivity of charts. Users can explore data and discover patterns through simple operations such as clicking, dragging, and zooming. Charts are also designed with aesthetics and information clarity in mind to ensure users can quickly understand the meaning of the data. In addition, the visualization module also provides customized chart configuration options, allowing users to adjust the chart style and parameters according to needs. The display effect is shown in Figure 4.

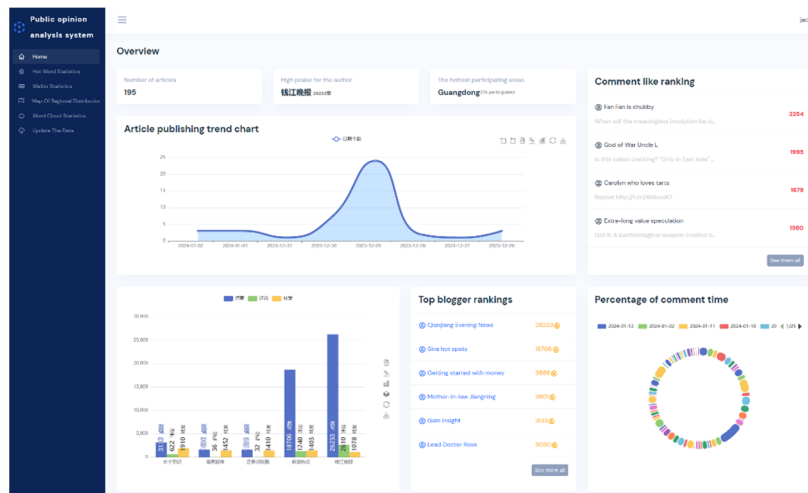


Figure 4. Screenshot of Part of the System

6. Conclusion and outlook

The Weibo public opinion visualization system based on Flask and ECharts efficiently collects data through distributed crawler technology and uses SnowNLP for sentiment analysis, realizing functions such as real-time monitoring, sentiment analysis, topic mining, and trend prediction. The system adopts B/S architecture, has a user-friendly interface and modular design, and provides a RESTful API interface to ensure the scalability and maintainability of the system. Although the system has achieved remarkable results in processing Weibo data and providing intuitive analysis, it also has limitations, such as limited ability to collect other social media data, the accuracy of sentiment analysis needs to be improved, and there are performance bottlenecks in large-scale data processing. . Future work will be dedicated to expanding the scope of data collection, optimizing the sentiment analysis model , and optimizing the system architecture to handle larger amounts of data. In addition, the research will also explore more refined public opinion analysis models, introduce real-time data stream processing technology, and develop applications for specific fields to further enhance the system's intelligence level and decision-making support capabilities.

Acknowledgments

This work was funded by the Hainan Provincial Natural Science Foundation of China (No. 721RC751)

References

- [1] Peng Shuguang, Wang Mengmei, Zhao Qibo, Shen Liubao, Peng Yujie, & Lu Xueyan. Epidemic information visualization system for ECharts[J]. Fujian Computer,2022, 38(4), 81-83.
- [2] Chen Junsheng, & Peng Lifan. Design and implementation of big data visualization system based on Python+Echarts[J]. Journal of Anhui Electronic Information Vocational and Technical College, 2019,18(4), 6-9.
- [3] Li Jing, Huang Jie, Yuan Hui, Zhu Guowei, Zhang Xianfei, & Wang Xinyin. Design and implementation of network threat visual analysis system in big data environment[J]. Journal of South Central University for Nationalities (Natural Science Edition), 2022,41(1), 79-86.
- [4] Wang Keqin, Gao Zhijiao, Qiao Yanan, Li Jing, & Tong Shurong. User demand identification and evolutionary trend mining in online reviews[J]. Mechanical Science and Technology,2023, 42(7), 1070-1080.
- [5] Chen Xingshu, Chang Tianyou, Wang Haizhou, Zhao Zhilong, & Zhang Jie. Spatiotemporal analysis of the evolution of public opinion on the "New Coronavirus Epidemic" based on Weibo data[J]. Journal of Sichuan University (Natural Science Edition), 2020, 57(2), 409.
- [6] Fan Luqiao, Gao Jie, & Duan Banxiang. Domestic popular tourist attractions data visualization system based on Python+Flask+ECharts[J]. Modern Electronic Technology, 2023, 46(9), 126-130.
- [7] Huang Shijing, Wu Chuanhui, Yuan Qinjian, & Xia Jingran. Research on spatiotemporal evolution differences of public opinion on public health emergencies based on sentiment analysis[J]. Information Science, 2022, 40(6), 149-159.