

# Historical Review of the Key Technologies of Brand Influence Monitoring of Large Enterprises

Guang Chen<sup>1\*</sup>, Chengcheng Fu<sup>2</sup>, Rui Liu<sup>3</sup>, Ruixin Chen<sup>4</sup>

chengguang@sgeri.sgcc.com.cn<sup>1</sup>, fuchengcheng@sgcc.com.cn<sup>2</sup>, liurui@sgeri.sgcc.com.cn<sup>3</sup>;  
chenruixin@sgeri.sgcc.com.cn<sup>4</sup>

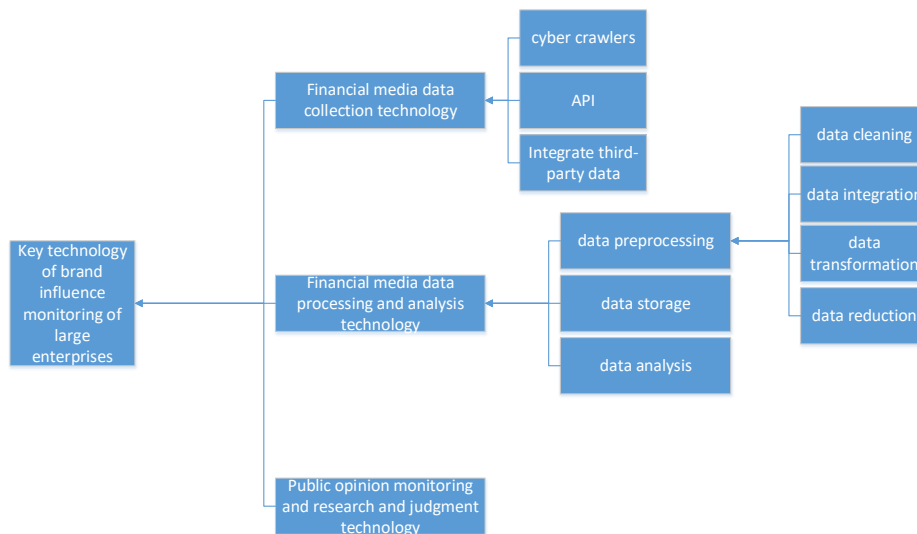
State Grid Energy Research Institute Co. Ltd., Beijing 102209, China

**Abstract.** The key technologies of brand influence monitoring of large enterprises are mainly composed of financial media data collection, data processing and analysis, public opinion monitoring and research and judgment technologies. Among them, the data collection methods of the financial media shall mainly include three kinds: network crawler, API and third-party data collection; data processing and analysis technologies are divided into different technologies such as data pretreatment, data storage and data analysis; public opinion monitoring and analysis technologies mainly include public opinion analysis, data cleaning, automatic information collection, etc. With the help of the above technologies, large enterprises can evaluate and monitor their brand influence, provide effective tools for enterprises to enhance the brand value, expand the international influence of the brand, optimize the enterprise development environment, and provide important support for shaping a good corporate image.

**Keywords:** brand value; brand influence; monitoring

## 1 Introduction

The key technologies of brand influence monitoring of large enterprises are mainly composed of financial media data collection, data processing and analysis, public opinion monitoring and research and judgment technologies (See Figure 1 for more details). Among them, the data collection methods of the financial media shall mainly include three kinds: network crawler, API and third-party data collection; data processing and analysis technologies are divided into different technologies such as data pretreatment, data storage and data analysis; public opinion monitoring and analysis technologies mainly include public opinion analysis, data cleaning, automatic information collection, etc. With the help of the above technologies, large enterprises can evaluate and monitor their brand influence, provide effective tools for enterprises to enhance the brand value, expand the international influence of the brand, optimize the enterprise development environment, and provide important support for shaping a good corporate image[1].



**Figure1:** Overall technical framework diagram

## 2 Financial media data collection technology

If the data sources of financial media are classified, it can be divided into intra-group data and non-group data; if divided into data format, including structured data and semi-structured data. Large enterprises need to make reasonable planning for the data to be collected according to the construction objectives, and select the data collection method according to the specific types of data. For data within the group, deformed data can be imported through interface or file, and data docking can be realized through database and other devices. Due to the large amount of network data outside the group, in the process of big data collection, the network crawler or public API can be reasonably used, so as to successfully collect the relevant text, audio and other information in the website[2].

The data of online social network mainly comes from major online social networking sites, including Sina Weibo, Tencent Weibo, Renren, etc. There are three main collection methods as follows:

**Web crawler way.** By simulating the user's access on the browser, capture the accessed HTML page, and obtain the required user and relationship data for the page analysis. Due to the different access control modes of different websites and the different structure of HTML pages, it is necessary to crawl the crawler of different websites. At the same time, the website also limits the frequency of access, so the frequency of access should also be controlled[3].

**API way.** All major social platforms provide open platforms for the H party development institutions, such as the open platform of Sina Weibo, which adopts OAuth2. [ ] 0 authorization agreement, and provides the SDK and access API of different languages, through authorization and access, W can obtain the required user and relationship information. However, the open platform has restrictions on the number and frequency of API visits for different levels of developers, and the amount of data obtained is relatively small.

Integrate third-party data. There are many shared data published by scientific research institutions on the Internet, and there are also special data sharing websites, such as Data Hall, Climbing Alliance, etc. The platform can collect these shared data and submit them to the platform for preservation[4]. The data acquisition module uses the collector for different online social networking sites and different collection methods, and configures and monitors the acquisition tasks through the performance layer. The acquisition module manages the collection task uniformly, and the collector collects the data according to the collection task. If the collection is interrupted, it can be collected again. The module also needs to process the data obtained by the collector, clean out the redundant data and dirty data, and persist the data after a unified storage format. The key problems to solve in the data acquisition module are the control of the grasping task and the control of the grasping frequency, W and how to conduct a unified and efficient storage of the data.

### **3 Financial media data processing and analysis technology**

#### **3.1 data preprocessing**

When the data collection work is completed, the data structure is very complex due to the different sources and formats of the data, coupled with the large numbers. Therefore, before the practical application of big data, it is necessary to carry out pre-processing work, build one or multiple sets of data selection standards according to the use requirements of the final data, and process the data according to the corresponding standards, and then send them to the readers. In the process of big data pre-processing, the main steps include data extraction, data cleaning, etc. The big data obtained in the process of data collection may have the characteristics of structural diversification. After the data extraction process, it may help people to transform the originally complex data into a single or easy to handle structure type, so as to achieve the purpose of rapid data processing[5]. In the use of big data, not all data is of use value to people, some data information is what readers do not care about, and some data has obvious errors. Therefore, in the process of data preprocessing, the staff needs to carefully filter the data to drive out the useless or wrong information, so as to achieve the effective improvement of the data preprocessing effect. Among them, there are four conventional methods for data preprocessing:

**Data cleaning:** Processing routines usually include filling in missing data values, smoothing noisy data, identifying or removing outliers, and resolving inconsistent problems.

**Data integration:** in the future, data from multiple data sources will be merged together to form a consistent data storage, such as integrating data from different databases into a data warehouse for storage. Later, data cleaning is sometimes required to eliminate possible data redundancy[6].

**Data transformation:** it is mainly to convert the data into a form suitable for mining, such as to scale the attribute data, so that it falls into a relatively small specific interval. This is particularly important for those mining algorithms based on distance. Including smoothing processing, aggregation processing, data generalization processing, normalization, and attribute construction.

**Data reduction:** on the premise of not affecting the mining results, the data is compressed by numerical aggregation and deleting redundant features to improve the quality of the mining mode and reduce the time complexity.

### **3.2 Data storage**

After the data preprocessing work is completed, the staff needs to store the collected data, so as to realize the unified management of the data. Generally speaking, people often use Hadoop body architecture to solve large-scale data storage problems. The architecture was researched and developed by Apache Foundation and belongs to the infrastructure of distributed system. In the process of use, even if users do not know enough of the distributed underlying details, they can also realize the deep development of distributed programs, through the application of the power of the cluster, and finally realize the efficient storage of data. Hadoop You can also build a file library in a distributed system, and the software is called HDFS. HDFS has a high fault tolerance rate in the process of use, which can be used in the low price hardware device, in order to achieve the effective improvement of data transmission efficiency, suitable for application in the big data environment. Using this architecture to build the big data platform can effectively reduce the cycle cost. With the establishment of the big data storage and management platform, it is also necessary to fully consider the construction of the security system of the synchronous platform. Staff can, through the destruction and encryption of data, promote the integrity of big data in the storage process, and effectively improve the security of the entire storage platform.

### **3.3 Data analysis**

In the process of big data analysis, it can be divided into two parts. Among them, the bottom layer is the basic analysis layer, and above this is the business analysis layer. The important role of the basic analysis layer is to provide the analysis foundation and service for the business analysis layer. In the process of work, the business analysis layer is mainly to accurately establish the analysis objects according to the different needs of users, and integrate the basic analysis services in the future, so as to provide users with more professional analysis results. In the actual analysis work, the staff can establish a keyword cloud library for each label according to the established multi-dimensional semantics, and conduct a deep analysis of each news report according to the specific algorithm used in the past, and automatically do the classification work, and send it to a fixed label. In the process of automatic classification, the classification results may be problematic at the beginning, and it takes a long time of learning and training to gradually adapt. In addition, the relevant staff should also do a good job of classification and error correction, timely update the classification algorithm and keyword cloud library, and finally achieve the reasonable optimization of automatic classification results.

Solve the data integration problem with different data sources and formats. Due to the large number of financial media platforms, in order to improve the efficiency of ELT, all data are converted from EXF file to Flat Text file format after entering the ELT system. At the same time, in order to improve the data transmission speed, the Push and Pull methods are adopted to reduce the impact on the source system. Push means the new data in the source system to the ETL program through FTP or copy; Pull refers to the direct access to the source database through DRDA or ODBC database protocol.

The Support Vector Machine (SVM) is a widely used classifier in machine learning. With the increase of the number of examples used for training, the training complexity increases exponentially, which seriously affects the classification efficiency of SVM. Therefore, the instance sample needs to be selected to reduce the sample number. Example sample selection is an important data preprocessing technology in machine learning. Its purpose is to pick out the

appropriate subset from the original instance sample data set, and use this subset to train classifier models with similar or higher accuracy than those trained using the original sample data set.

At present, the problem of fine-grained classification is the research focus in the object and scene recognition tasks. At the same time, the research results of the sound, image and text information of media resources are few, which is also the future research direction.

In addition, in the process of network monitoring, the emotion analysis of financial media is an essential part. The text is divided into two or more types of positive or derogatory meaning based on the meaning and emotional information expressed in the text. Generally speaking, the purpose of emotion analysis is to find out the attitude of the speaker / author on certain topics or on the two poles of a text. By adopting the method of weighted calculation, the index is applied to emotion analysis, using the emotion classification algorithm based on the classification of subjective text, according to the classification process of a view belongs to a kind of probability size, can further on the basis of the classification of classification, in order to better response.

#### **4 Public opinion monitoring and research and judgment technology**

The research of online public opinion is the research of social science, which is a marginal subject including political science, sociology, communication science and informatics. In the study of online public opinion, political science mainly studies the development of online public opinion and democracy, the connection with public policy, the relationship between online public opinion and election and the feasibility of being directly democratic; sociology mainly studies the social background, social stratification, process, change, trend and reason, the influence of online public opinion on personality, the relationship, law and court decision, and predicting social movement through online public opinion; the strategy and communication route of online public opinion direction. Because the amount of information is very huge, only rely on artificial method is difficult to cope with huge amounts of information collection and processing, need modern information technology to strengthen the study of related information technology, the network public opinion analysis, form a set of automated network public opinion analysis system, timely response to network public opinion, by passive prevention into active comb, to control and guide[7].

For all kinds of network public opinion, the government and the news media should be able to make timely feedback to prevent small problems, and even nip in the bud. Its premise is that modern information technology must be used to analyze the network public opinion, so as to control and guide it. However, the Internet has a large amount of information, so it cannot rely on manual methods to collect and process the characteristics of massive information online. It is necessary to establish a network public opinion analysis system for the purpose of "coping with network public opinions, preventing passive blocking, and becoming active combing and guidance". The technical core of such systems is mainly:

Public opinion analysis engine. This is the core function of the public opinion analysis system, which mainly includes tendency analysis, hot topics (sensitive topics) identification, statistical report, trend analysis, alarm system, emergency analysis and other contents.

Data cleaning function. Pre-process the collected information, such as format conversion, data cleaning, and data statistics.

Automatic information collection function. The existing information collection technology is mainly to automatically obtain the page information from the Internet through the link relationship between the network pages, and as the link continues to expand to the whole network.

It should be pointed out that the core technology of the public opinion analysis system lies in the public opinion analysis engine, and the most important technologies involved include text classification, clustering, view tendency identification, theme detection and tracking, automatic summary and other computer text information content identification technology.

## 5 The Conclusion

Relying on the media data collection, data processing and analysis, public opinion monitoring and analysis technology, large enterprises to evaluate their own brand influence and monitoring, which can enhance brand value for the enterprise, expand the brand international influence, optimize the environment of enterprise development, to provide effective tools, to provide important support for shaping a good corporate image.

**Acknowledgments.** This work is supported by the science and technology project of State Grid Corporation of China (The key technologies and applications of brand influence evaluation and monitoring fits into the layout of "One body and four wings", grant no.1400-202257240A-1-1-ZN).

## References

- [1] KAVAK B ,İBRAHİMOĞLU N .The Influence of Self Concept and Self Monitoring on Brand Loyalty : An Investigation on University Students[J].Hacettepe Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi,2006,(2):109-125.
- [2] Sun C ,Xu S .Discover The Influence of Brand Image for M·A·C Cosmetics[C]//Wuhan Zhicheng Times Cultural Development Co., Ltd..Proceedings of 4th International Conference on Innovations in Economic Management and Social Science (IEMSS 2022).Hubbard School of Journalism and Mass Communication,University of Minnesota;Tisch School of The Arts,New York University;,2022:10.DOI:10.26914/c.cnkihy.2022.079843.
- [3] Du Xiaoping. Research on big data analysis technology of financial media monitoring and supervision [J]. Radio and TV Technology, 2023,50 (06): 124-127.DOI: 10.16171 / j.cnki.rtbe.20230006023.
- [4] Wu Minyu. Analysis of the influence and development trend of big data technology on financial media [J]. Science and Technology Information, 2023,21 (08): 44-47. DOI: 10.16661 / j.cnki.1672-3791.2211-5042-7820.
- [5] Chen. On online public opinion monitoring and disposal [J]. Industrial Innovation Research, 2020, (11): 71-72.
- [6] Zhang Feifei. Research on the mechanism of research and role of online public opinion [J].Theory and contemporary, 2016, (03): 20-21.

[7] Sun Yuyuan, Wang Xuan, Lu Yuliang. Summary of deep learning model safety studies [J].Information countermeasures technology, 2023,2 (Z1): 93-112.