Research on Efficient Work Innovation Effect Evaluation Based on Real-Time Online Public Opinion Data Classification and Detection Algorithm

Yu-liang WANG

Shenzhen Polytechnic

Shenzhen, Guangdong, 518055, China

Abstract:Various heterogeneous network resources are collected through web crawlers, and after preprocessing such as data cleaning, they are written into the data center. And implement data mining technologies such as classification, clustering, association rules and anomaly detection. If an enterprise wants to maintain a good development ability and maintain a correct development trend, it must pay attention to and do a good job in ideological and political work, and take ideological and political work as the guiding light of its own operation and development. Improved the stability and accuracy of clustering results. The news topics are clustered to form topics, and the topic tracking technology is used to track specific news topics to achieve personalized services.

Keywords: Efficient Work, Innovation Effect Evaluation, Real-Time Online，Public Opinion Data Classification

1. INTRODUCTION

Zhaizhen Coal Mine is a large-scale mine with an annual production capacity of 200×104t, which was put into operation in 1993 [1]. The coal under construction accounts for more than 75% of the total mine reserves. The surface corresponding to the seventh mining area is Zhaizhen Town Government, hospitals, schools [2], commercial streets, etc. The building complex has more than 4 million tons of recoverable coal reserves. In 2003, it was demonstrated by Shandong University of Science and Technology and adopted the trend of broad strips [3]. Skills are the skills and methods used to complete specific tasks and achieve output, and are the actual operation and operation process to realize ideas and complete output [4]. A comprehensive manifestation of operational ability, tool control ability and recognition of high-quality standards.

With the rapid development of network technology, the Internet has become [5] a part of people's life and work, especially with the rise of 5G technology, it is foreseeable that the Internet will have a more and more profound impact on my country's political and economic development. It is an important product [6] derived from the Internet. It can quickly attract the attention of netizens. Netizens express their views through forums, Weibo, Tieba [7] and BBS, which in turn generate pressure from public opinion, which has attracted great attention from management departments at all levels [8].

Compared with newspapers and other media, the Internet has the advantages of product digitization, distribution network [9], and convenient distribution. The remarks, suggestions or emotions expressed by news media or individuals on the Internet are a special form of social public opinion on the Internet [10]. In recent years, with the increasing importance of Internet public opinion, it has gradually become a research hotspot [11]. It is a clear and definite original record of past social activities, with a distinct original record, which is the same as the essential characteristics of archives [12], and determines the archive properties of network public opinion data information. Due to the characteristics of virtuality, concealment, divergence, permeability and randomness of the network itself [13], more and more people are willing to express their real thoughts through such channels [14].

The Internet has gradually become the main place for the generation and dissemination of public opinion, and network public opinion plays an increasingly [15] important role in social life. The Qicai District is a joint mining area, and a total of 3 downhills are arranged, namely the Qicai track down the mountain and the [16] Qicai transportation Go down the mountain and Qicai Huifeng go down the mountain. The Qicai filling system is reconstructed on the basis of the original mining area, that is, the existing system is fully utilized, and the return wind of the [17] Qicai is used to transport the gangue down the mountain. For new development trends and new development strategies, it is also necessary to adapt to the development of the times, move towards transformation, and realize the development of modernization [18]. Among them, ideological and political work plays an important role, which is not only the driving force for transformation, but also the key to the success of reform [19].

At present, China's online public opinion is in a very active state. On the one hand, the development of the Internet has brought great convenience and people can easily obtain a large amount of information; social problem. At the same time, there are also problems such as difficult information identification [20], numerous online rumors, and difficulty in semantic analysis. Therefore, in order to achieve the purpose of creating a good cyberspace and grasping the correct orientation of public opinion [21], it is of great significance to strengthen the analysis of online public opinion. Such information can easily lead to a series of social problems. Therefore, relevant departments must [22] strengthen the analysis and monitoring of network public opinion, and provide positive guidance to network public opinion when necessary [23].

Manual collection refers to manually browsing online platforms, such as portal websites, blogs, forums, post bars, etc., to track sensitive issues [24], obtain valuable data and information through manual monitoring, and observe the trend of social public opinion after aggregation and analysis. In

addition to the content information, the data information on the network includes many useful [25] structural signs. Although many international organizations have developed a large number of agreements to unify the format of these data information to a certain extent, most of these agreements focus on the representation of data. At this stage, my country is in a period of rapid development the impact of external culture has caused different degrees of ideological impact on the majority of the workforce. Under the situation that the business environment and social environment are changing, the enterprise must pay attention to the change of the thinking form of the majority of employees.

2. THE PROPOSED METHODOLOGY

2.1 The Real-Time Online Public Opinion Data Classification

However, such a single-label classification system cannot adapt to the characteristics of network public opinion, and it also buries the information points that people are interested in in a single field. The public opinion conveyed by the crowd in the media is accompanied by emotional and often ambiguous characteristics, which makes the analysis of public opinion more complicated. The engine is the key. The engine is mainly realized by data mining technology, including classification, clustering, association rules and anomaly detection, and finally can realize hot spot identification. In order to ensure the representativeness of public opinion data, the target website that collects public opinion data needs to have daily activities. It has many users and covers a wide range of topics.

Generally speaking, Weibo data covers a wide range, information is updated quickly, and data output is large. Manual collection and search engine collection are the most commonly used Internet public opinion collection technologies, but the disadvantage of these two collection methods is that they can extract the text of key information. The content is relatively short, and it is easy to cause low accuracy and recommended recall rate when doing related operations due to sparse data. The task of the document classification system is to automatically determine the category associated with the document according to the content of the document under a given classification system. From a mathematical point of view, document classification is a mapping process, which maps uncategorized documents into existing categories. Different from traditional text data, Weibo data contains some special characters, including: before and after the text The common # character identifies the central content of the microblog data: there may also be a label composed of symbols in the microblog data, that is, when a user wants other users to see the microblog messages or comments he/she sends.

To achieve innovation, we first need to analyze the current problems: first, the ideological and political work of public institutions is neglected. As far as the current development is concerned, the influence of diversification and economic benefits dominate the development, while other aspects of development are ignored.

2.2 The Public Opinion Data Classification and Detection Algorithm

The public opinion data can be divided into text, picture, audio and video in terms of data types. Since audio and video need to process a large amount of information, and the amount of data carried is less than that of text and pictures, pictures are relatively the network public opinion information

that the text can express is less. The collected data is not standardized and needs to be further processed, such as aggregation, sampling, feature creation, feature subset selection and variable transformation. The preprocessed data will be written to the data center for invocation by the public opinion analysis engine.

The meaning of the association rules between item sets is. Put it into the queue to be crawled, and the IP address obtained by parsing can automatically crawl the information of the webpage and save it, and put the crawled URL into the crawled queue. , get new links from crawled URLs. Preprocess the text. Text preprocessing is to preliminarily process the raw web pages collected by the public opinion collector, and then perform topic discovery and digital modeling on the preliminarily processed web page text. The most important part of text preprocessing is to perform feature selection and Chinese language. Participle. The classification idea of this method is very simple. According to the arithmetic average, a center vector representing the class is generated for each type of text set, and then when a new text arrives, the new text vector is determined, and the distance between the vector and each type of center vector is similar. Spend.

Due to the imbalance in the number of samples, when traditional classification methods are used to process Weibo texts, more majority classes will eventually be identified, but minority classes are often ignored. For example, the identification of Weibo public opinion events can be regarded as a binary classification problem. For the acquired web page information, only the body content is often required. For example, when analyzing a shopping website, only the information of the item needs to be retained, that is, the structured information.

2.3 The Efficient Work Innovation Effect Evaluation

Two commonly used extraction methods are template and web page library structure information extraction. The goal of anomaly detection is to find objects that are different from most of the data, also known as outlier detection, deviation detection. Anomaly detection methods mainly include model-based technology, proximity-based technology and density-based technology.

On the Internet, users often discuss an event, and events and topics discussed by a large number of users are called hot topics. According to the content of the data center, such as the keywords of the webpage, the number of replies, the number of forwarding, etc., the heat value of each topic can be calculated by using data mining technology. In single-label text classification, there is no overlap between categories, and a text can only have one label.

However, in real life, things may be ambiguous, and some texts can belong to multiple categories at the same time. For example, a news text related to farmers' land subsidies can belong to "people's livelihood" at the same time. In the development of ideological and political work, it is necessary to Working methods for effective innovation. The object of ideological and political work is the employees of the vast enterprises. It should be people-oriented, go deep into the actual life of employees, grasp the ideological context of employees, and diagnose and treat some contradictions and problems existing in the workforce. Effective solutions. We must adhere to the people-oriented principle, take the actual needs of employees as the focus of management, all

management is based on the actual needs of employees, respect employees, create a fair environment, and realize their value.

In single-label text classification, there is no overlap between categories, and a text can only have one label. However, in real life, things may be ambiguous, and some texts can belong to multiple categories at the same time. There are two common multi-label text classification, one is to convert the problem into a single-label classification problem, the other is The single-label classification algorithm is transformed to adapt to multi-label classification. For example, the ML-KNN algorithm is to extend the traditional K-nearest neighbor algorithm into a lazy learning multi-label K-nearest neighbor algorithm. The development of internal ideological and political work in the enterprise also requires a working mechanism. Innovate and adapt to improve with the times.

3. CONCLUSIONS

Starting from the significance and background of network public opinion monitoring, this paper introduces the domestic and foreign research on network public opinion monitoring system and classification technology in detail. At the forefront of the times, closely follow the pace of the times, combine the specific development needs of the enterprise and the changing situation, correct attitude and understanding, and use methods such as the classification, clustering, association rules and anomaly detection to achieve hot spot identification, topic tracking and Sentiment analysis and other functions. There are massive and heterogeneous data resources on the network.

4. REFERENCES

[1]Ma Kunyun. Research on online product quality risk assessment and early warning based on user reviews [D]. Nanjing University of Science and Technology, 2018.

[2] Ye Jian, Zhao Hui. Public opinion analysis model based on large-scale barrage data monitoring and sentiment classification [J]. Journal of East China Normal University (Natural Science Edition), 2019, 000(003):86-100.

[3] Li Shengfei, Zhao Jie, Gong Zheng, et al. Research on identification of emergency-related institutions based on public opinion data [J]. Standard Science, 2021(5):5.

[4] Guo Peng. Research on network packet classification algorithm based on deep reinforcement learning [J]. Journal of Jiamusi University: Natural Science Edition, 2020, 38(2):3.

[5] Wu Hao. Research on course recommendation based on data-supported Apriori classification algorithm [J]. China Information Technology Education, 2019(20):4.

[6] Jin Yong, Li Li. Accurate and efficient network public opinion detection and early warning method based on multi-level geography: CN109388749A[P]. 2019.

[7] Zhou Tao, Yang Cuicui, Lv Meijing. Research on Internet Public Opinion Monitoring and Early Warning in Colleges and Universities - Funded by the Central University of Finance and Economics for Party Building and Ideological and Political Work Theory Research Projects [J]. Friends of English in Middle Schools: Research on Teaching Methods in Foreign Languages, 2018(11): 233-234.

[8] Peng Jianxin. Data quality detection based on intelligent classification algorithm [J]. Journal of Jinan University: Natural Science Edition, 2018, 32(5):5.

[9] Zhang Min. Research on the characteristics and laws of public opinion dissemination based on ARIMA model—

taking the early data of Weibo platform as an example [J]. Advances in Applied Mathematics, 2022, 11(5):11.

[10] Xia Yixue, Yuan Ye, Zhang Wencai, et al. Research on abnormal data monitoring and application of network public opinion for big data [J]. Modern Intelligence, 2018, 38(6):6.

[11] Liu Xiaoming. Research on the Effectiveness of Internet Public Opinion Handling of College Ideological and Political Counselors under the Background of New Media [J]. Modern Education Forum, 2020, 3(10):39-40.

[12] Kuang Chuwen. Research and design of campus network public opinion monitoring system based on data mining [J]. Digital Technology and Application, 2021, 39(5):4.

[13] Yang Hongtian. Research on Weibo topic detection based on feature enhancement and convolutional neural network [D]. Nanjing University, 2019.

[14] Zou Dingling, Duan Fengfeng. Research on hot public opinion monitoring and coping strategies based on implicit data analysis [J]. Journal of News Research, 2018(3):3.

[15] Zhen Jia. Research on the extraction and classification of hot words of public opinion based on campus network traffic [D]. Hubei University, 2018.

[16] Yuan Chuanqi. An intelligent analysis system of network public opinion based on network data mining algorithm:, CN108107973A[P]. 2018.

[17] Guo Lejiang, Xiao Lei, He Song, et al. Research on network public opinion analysis based on big data and artificial intelligence [J]. Yangtze River Information and Communication, 2021(3):6.

[18] Zhang Tian. Research on Ideological and Political Education Work Based on Network Public Opinion Monitoring [J]. Reference for Middle School Political Teaching, 2021(45):1.

[19] Liu Yifei. Research on key technologies of network public opinion information identification and analysis [D]. University of Electronic Science and Technology of China, 2020.

[20] Chen Xiaobai. Research on the Improvement of Shenzhen Public Security Bureau's Ability to Dispose of Police-related Public Opinions Based on 4R Crisis Management Theory [D]. Shenzhen University, 2019.

[21] Zhu Jiafang. Research on Distributed Disaster Public Opinion Information Acquisition and Analysis Methods Considering Time and Space Theme Characteristics [D]. Zhejiang University, 2019.

[22] Fan Shenghua. Research on the Supervision Mechanism of Internet Public Opinion in the Era of Big Data [D]. Central China Normal University, 2019.

[23] Wang Huiquan, Xie Dongfang, Liu Lu. Research on Internet Public Opinion Monitoring Based on Big Data Thinking and Methods [J]. China Radio and Television Journal, 2019(2):6.

[24] Lv Xiangnan. Innovation of ideological and political education in colleges and universities under the background of network public opinion [J]. Journal of Corporate Culture, 2018, 000(006):230.

[25] Liu Yu. Research on Internet Public Opinion Topics Detection and Tracking Technology [D]. Dalian Maritime University, 2018.