Advances in Artificial Intelligence, Big Data and Algorithms
G. Grigoras and P. Lorenz (Eds.)

doi:10.3233/FAIA230881

User Behavior Analysis Based on Big Data and Artificial Intelligence

Tao WANG[a], Yinghui MA[b], Ze WANG[a], Xiaojun WU[a], Long LI[a1]
[a] *Sichuan Zhongdian Qiming Star Information Technology Co., LTD., Chengdu, Sichuan 610000*
[b] *State Grid Qinghai Electric Power Company Information and Communication Company, Xining, Qinghai 810000*

Abstract: Thanks to the support of big data technology, the Internet is gradually expanding towards the Internet of Things, and further upgraded to a better experience of artificial intelligence, which is possible based on big data and artificial intelligence. Taking "Big data and artificial intelligence can be used for analysis of library users" as an example, Explore the domestic and foreign authoritative database input user behavior analysis and other keywords to find relevant literature, According to the comprehensive comparison of literature analysis method, 55 core technical documents were selected for close reading, Through the development process of technology, Clarify the core technologies and methods of user behavior personality analysis and credibility analysis and do the key expansion research, According to the literature research results combined with the needs of the actual development of the library, Exploring a set of big data analysis model of user behavior based on traditional data mining and classification learning, The model can effectively combine the user personality analysis and abnormal behavior detection to solve the above two problems.

Keywords: big data; smart library; user behavior; data modeling

1. Introduction

With the development of technology, big data and artificial intelligence are applied to all walks of life, Users' behaviour analytics based on big data and artificial intelligence has gained the attention of the community. When analyzing users, you can use big data and artificial intelligence to deeply analyze their behavior, promote the application of data mining results to commercial marketing and advertising, realize precise bidding management, and control commercial costs[1]. Based on this, It makes sense to have a discussion on applications that use big data and artificial intelligence to analyze user behavior.

2. The Relationship between Big Data and Artificial Intelligence

Big data mainly refers to massive data, big data mining is the processing of massive amounts of data through a series of means, such as selecting, exchange, analysis and

[1] Corresponding author: Long LI, Sichuan Zhongdian Qiming Star Information Technology Co., LTD.; e-mail: 8805635@qq.com

integration, promote the maximization of mass data, multi-head and intensive application of massive data in all aspects of society, and create new value[2]. And then bring "big development", "big profit", "big technology" and "big knowledge". In artificial intelligence, "deep learning", "adversarial learning", "enhanced learning" and their corresponding "adversarial neural network" and "convolutional neural network" are all closely related to big data mining[3]. In particular, "deep learning" in artificial intelligence can make it possible to collect big data and provide enough sample data for user behavior analysis.

3. The evolution of Data Analysis Technology

3.1 Mathematical Statistics and Statistical Analysis

In early 1990s, When publishing papers on mathematical statistical analysis of user behavior, research can be conducted through borrowing records, questionnaires and other methods. From the beginning of the 21st century, Internet technology began to develop rapidly. Based on the principle of data statistics developed by the website statistics and custom software industry gradually began to commercial use[4]. If eXTReMeTracking provides URL real-time tracking service and user website browsing various statistical information, WebSiteTrafficreport sends user traffic statistics by email, and MiniTab software statistically analyzes users' habits with a user survey[5]. Although the pre-defined curing procedure provides data on only some of the statistical parameters, the scope and accuracy of the predicted user behavior is limited. But for a digital library entered the web 2.0 era, it's important. The analysis and research of deep data mining began to be increasingly extended to the field of library services. For example, Li Panchi analyzes the user borrowing information according to the clustering algorithm, Wei Yuhui et al. used association rules to mine and analyze library circulation data, Wen et al. mined OPAC data using SQL and data warehouse technology[6].

3.2 Data Mining

Cluster analysis forms its own clusters by grouping the elements in a collection according to a certain degree of similarity. The difference distance between the internal elements of the class is small (i. e., more similar between the elements), analyze the statistical characteristics of each category, and find the grouping results of interest in the data mining[7]. A more rigorous mathematical description of the clustering is as shown below: the studied sample set is K, and class M is defined as a non-empty set of K, namely M K, and M =K. Different M classes of Mi that meet the following two conditions (i=1,2,3,4,) are clusters:

$$M_1 \cup M_2 \cup \cdots \cup M_n = K$$
$$M_i \cap M_j = \phi(i \neq j)$$

From condition 1, each one of the samples must belong to one of the categories in the cluster, while condition 2 indicates that each sample belongs to no more than one category. Aggregation is a vital part of data mining, and also an important technology of machine learning in the field of artificial intelligence and pattern recognition. It is a kind

of unsupervised learning. Clustering itself is not a particular method of algorithm, but rather a general task. There are many different clustering algorithms that can be used to complete this task. The main clustering algorithms are segmentation-based, level-based, intensity-based, mesh-based, model-based.

4. Big Data Analysis Model of User Behavior

When analyzing the personalized behavior of users, the massive data generated by users of big data libraries provides effective support for such analysis.

4.1 Personalized Analysis of User Behavior

The user behavior of the library can be analyzed through the human-computer interaction logs, web browsing records, and digital resource downloads of the library. By collecting the behavioral data mentioned above and performing supervised learning, analyzing and predicting the user portrait[8]. Specifically, training data samples can be constructed based on the user's gender information for different response times to each module in the library - coarse-grained training data and fine-grained training data that can be used to validate the binary classification SVM classifier; Training data characterized by the fine-grained data of users' browsing interest and reading interest[9]. Associating the adjacent matrix R with the user and counting the user distribution probability can be used as the prior probability of the Bayesian classifier, Calculate the similarity of users and categories to select user neighbors and category neighbors, Can predict the price-oriented terminal level; Image features and text features posted on postings, message boards and chat platforms can be used when building the data.Vector machines and gradient improvement classifiers can be supported using integration, and users' posting habits can be predicted and analyzed using integration; accumulcumulative release function F through the time users are used to release information every day, different thresholds are set according to the time-series data analysis, divide the active levels; The above user gender, terminal level, platform publishing habits, interests, activity, subject background and other basic elements of library user portrait are gradually clear[10]. Finally, collaborative filtering can form a complete picture of user behavior. Push user personalized service products, Such as recommending users' favorite books, lectures, activities and other information, Complete the big data analysis of user behavior.As shown in Figure 1

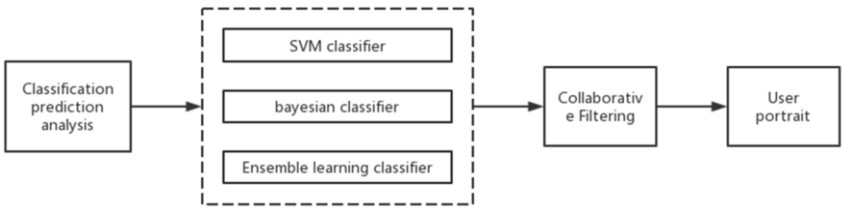

Figure 1 Analysis of user personality information

4.2 Analysis of User Behavior Credibility Data

Smart libraries are established by relying on the Internet, so the emergence of network security issues is inevitable, while the frequent application of interactive devices, lagging software updates, and leakage and tampering of user data are more common., leading to the account theft and rapid spread of various network viruses, reaction, violence, yellow and other abnormal behaviors; how to timely detect and identify unreliable user behavior has become another focus of this paper. At present, the network security manufacturers used for the terminal anti-virus software, firewall, encryption software technology is basically based on the existing virus signature database and scanning engine for data comparison to check the security[11]. Here we deposit these existing virus feature databases and user abnormal behavior data features together as the prior knowledge base, While this can filter out most of the known abnormal behaviors but not for the various abnormal behaviors generated in real time, The existing knowledge base cracked virus updates far less than the latest variant of transmission; In view of the above problems, Domestic and foreign scholars have tried a lot of new technologies, And has made good progress. Such as Rieck et al proposed using machine learning classification algorithm, automatic analysis of malicious behavior, not only can detect the existing characteristics, but also can form a new category add a prior knowledge base, Xin et al cluster analysis according to the traditional data mining methods, feature extraction for virus samples, for a large number of virus samples to realize the effective classification constantly enrich the prior knowledge base. Burguera et al proposed based on the android platform cloud detection mode and Rajab proposed browser built-in virus detection system analysis method, the method is in the era of big data library big data analysis center can real-time monitoring library user operation behavior, should be virus prevention detection work should be brought into the daily management and timely feedback detection results. On the one hand, it can effectively overcome the detection lag caused by the limited computing and storage resources of the user terminal; on the other hand, with the effect of traditional data mining and supervised learning and analysis technology, it can not only push the user personalized product and service information, but also timely issue the network security warning, truly reflecting the service characteristics of intelligent library and providing a good user experience. The following is an introduction to the big data analysis model of library user behavior.

First, all packets generated in real time are collected at the gateway of the big data center. After preliminary data preprocessing, If it is determined to be normal, continue the next data set of training the classification learning technique; If abnormal, conduct data mining of its behavior log, Here is mainly dominated by cluster analysis, Combined with multidimensional association rules and time series analysis, If the new abnormal behavior is confirmed, add to previous library knowledge, warn the user, and if it cannot be verified, a second test is started; Likewise, the ability to detect virus variants through machine learning using classification prediction and personality analysis of users. New feature categories can be added to the previous knowledge base to generate information about the user's personality by analyzing and predicting the rest of the normal behavior Collaborative filtering allows for the sketching of the user's profile. Eventually, the obtained personalized information is sent to the user's terminal to obtain a complete user behavior analysis process. The specific flow chart is As shown in Figure 2：

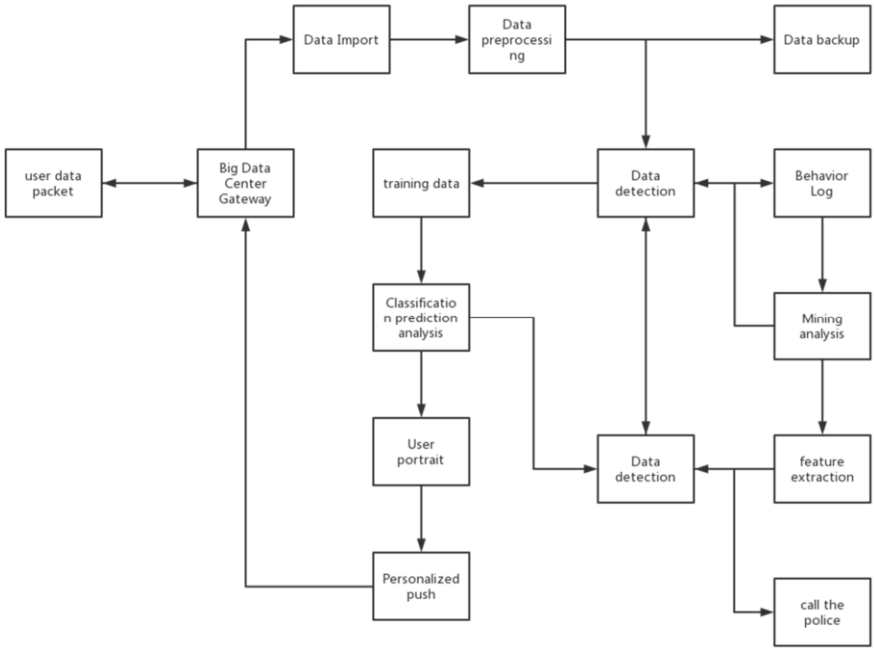

Figure 2 The Big Data analysis model of user behavior

5. Conclusion

To sum up, the use of big data and artificial intelligence for data analysis of user behavior is the inevitable trend of development, and is increasingly important for the development of the industry. Therefore, in the process of using big data and artificial intelligence for data analysis of user behavior, existing big data mining algorithms and artificial intelligence techniques can be utilized when optimizing and improving data mining model libraries. Current library is big data under the development of artificial intelligence technology gradually transformation into wisdom library, with the help of new intelligent interactive equipment and analysis technology for many dream library service concept, this paper, from the perspective of user behavior analysis, relevant literature techniques are studied, and a big data analysis model is proposed that can perform both network security monitoring and analysis of users' personal behavior., in addition to let the library can more timely and accurate to push user personalized service, also maintain the user's security and privacy and library network security, to minimize the impact of harmful behavior.

References

[1] Weldemariam, Komminist, Mohania, et al. Modeling user behavior data in systems of engagement[J]. Future Generations Computer Systems Fgcs, 2017.

[2] Wang Y , Wang B , Huang Y . Comprehensive Analysis and Mining Big Data on Smart E-commerce User Behavior[J]. Journal of Physics Conference Series, 2020, 1616:012016.
[3] Jinping, Xu. Research on Personalized Service of University Library Based on User Behavior[C]// International Workshop on Advances in Social Sciences. 0.
[4] Ebeid E , Heick R , Jacobsen R H . Presenting user behavior from main meter data[C]// IEEE International Conference on Smart Grid Communications. IEEE, 2016.
[5] Zhao Hui, Application demonstration of prison personnel location management and artificial intelligence behavior analysis system based on Internet of Things and big data technology. Zhejiang Province, Zhejiang Junpu Technology Co.
[6] Shi Hecheng, Yu Liqiang. Exploration of big data and artificial intelligence application in the field of supervision--Analysis of false trade financing behavior as an example[J]. Qinghai Finance,2021(02):42-46.
[7] Pan Huiping. User behavior analysis based on big data and artificial intelligence[J]. Modern Computer,2021(02):47-50.
[8] Yang L. The change of cognition, thinking and behavior by big data and artificial intelligence[C]//Shanghai Law Society. Collected Works of Shanghai Law Studies (2019, Vol. 13 General, Vol. 13) - Proceedings of the Internet Justice Research Group of the Shanghai Law Society. [publisher unknown],2019:67-70.
[9] Luo XD,Wen XH. An analysis model of student behavior based on artificial intelligence and big data[J]. Education Observation,2019,8(17):11-13+32
[10] Zheng Erli,Wang Yingji. Data ideology in the era of artificial intelligence--thinking about the impact of big data on values and behavioral activities based on big data[J]. Journalism and Communication Review,2019,72(01):72-79.
[11] Chen FG,Du YY. Behavioral intervention of the insured based on big data and artificial intelligence[J]. Times Finance,2016(29):230.