

Legal Regulation of Data Mining Technology

Caishi Ye, Hao Cui, Tianbao Wang

School of Law, Anhui University of Finance and Economics, Anhui, China

Abstract

In recent years, with the development of big data, cloud computing and the Internet of Things, human society has entered the era of big data. Data, as a new resource, has been widely concerned. In China, as a big data country, the relevant data industry grows rapidly and has a huge scale, which causes new social problems and puts forward higher requirements for our legal regulation. Taking supporting data mining and utilization as the basic position, this study provides legal path for standardizing the market operation of data mining technology, provides legal remedy for the endless data property rights disputes, builds legal bottom line for personal privacy protection, escort the high-quality development of big data industry, and provides legal guarantee for national information security.

Keywords

Data Mining Technology; Legal Regulation; Information Security.

1. Introduction

Since entering the information age, the degree of information is constantly improving, the computer hardware is iterated and upgraded, and the amount of data available for us to collect and use shows exponential growth. In the face of such a huge amount of data, we need a tool to maximize its value to develop it, thus -- data mining technology comes into being. Data mining technology as a technical means, on the one hand, for people from the vast sea of data to capture effective information provides convenient; On the other hand, as a new thing, it also causes a series of legal problems, and puts forward higher requirements for the network security standard system. However, the current legal regulation on data mining has not made specific provisions and changes on this technology, and the relevant security management measures have fallen behind the development of new industries and technologies. In order to realize the benign and sustainable development of data mining technology in the information age, we are bound to start from various aspects, promote the legalization and application of data mining technology, and rely on its own advantages, to build a firm network security protection system.

2. Application Situation and Prospect of Data Mining Technology

2.1. Concept of Data Mining Technology

Data Mining refers to the process of using algorithms to retrieve valuable information hidden in massive data. The operation of data mining requires the support of technologies in many fields, such as machine learning, artificial intelligence, database, visualization technology, etc. By analyzing the collected data, the hidden internal relationships among the data are mined and new meaningful connections are extracted from them, so as to provide suggestions for decision-makers, help them reduce risks and make correct decisions. In order to maximize the value of data. The process of data mining mainly includes three parts, the first is to prepare the data, the second is to carry out extensive and deep mining of the prepared data, and the last is to carry out the expression of the results and suggestions.

2.2. The Application of Data Mining Technology in the Industry

In the era of Internet economy, who can obtain the latest market information, grasp the social group's consumption trend and psychology, who will win the market opportunity. In order to obtain these valuable information, we need to resort to data mining technology to achieve, so the technology has become a hot technology in the Internet industry. Data analysts through the complicated data screening, extraction, induction, analysis, summed up the real valuable information for data users, the establishment of a new database. After scientific analysis and prediction, these data are accurately and efficiently used by decision-makers to target consumers' consumer psychology and trends, and play a forward-looking role in occupying emerging markets in advance.

Data mining technology has a wide range of application scenarios and application subjects. Personally, we use a variety of retrieval tools to retrieve effective data from massive data to help us achieve the target task, which is the use of data mining technology in the smallest scenario; For the market players, the collection and analysis of the user's transaction records, trading habits, consumption level, geographical location, payment method and other data, can be the corresponding portrait of their customers, and then provide personalized and differentiated services for them, cultivate customer loyalty, avoid customer fraud, gain competitive advantages, in the fierce market competition in an invincible position; For government agencies, they collect and store massive public data in the process of carrying out public affairs, which provides a reference for the formulation of government policies. The lawful and reasonable use of data mining technology can make the government's decision-making more in line with the public interest.

2.3. The Prospect of Data Mining Technology

Since the birth of data mining technology, has become the "darling" of all walks of life, the market competitor of IT industry, government agencies also attach great importance to the field of data mining, at the same time, Chinese colleges and universities also set up related disciplines for academic research. At the same time, data mining technology in the world has been paid attention to by various countries. A study by IDC of 62 companies in Europe and North America that had adopted smart business technology found that they had an average return of 400% in the third year, with 25% of them achieving 600%. Nowadays, peace and development are the theme of The Times. The competition between big countries is often manifested as technology war and information war. Data mining technology is an important embodiment of this ability. Under the fierce international competition, it is an inevitable trend for China to increase the investment in relevant research and development, improve the data processing ability and enhance the information security guarantee. Therefore, it is expected that data mining will become a mature industry.

3. Social Problems and Legal Risks Caused by Data Mining Technology

While creating huge social wealth and promoting social development, data mining technology also causes a wide range of social problems. From data collection to data mining, to the establishment of databases and the management and use of data, there are conflicts of rights and obligations in almost every stage of the operation of data mining technology.

Although the Network Security Law of the People's Republic of China expressly stipulates that the collection and use of personal information by network operators shall follow the principles of legality, legitimacy and necessity, make public the collection and use rules, express the purpose, method and scope of the collection and use of information, and obtain the consent of the collector. However, driven by huge economic interests, many information enterprises ignore the provisions of the law and use illegal, scientific, secret and indirect ways to collect

user information. When a user needs to use an APP, he must check the "license agreement" about obtaining personal information, otherwise he cannot use the software, and most users will not carefully read the license agreement. In this case, it appears that the user can choose whether or not to provide personal information, but in fact they are forced to collect personal information. On the other hand, information companies often do not distinguish what information they need to collect, but adhere to the "more is better" philosophy in collecting user information. Even if the public does not provide certain information when using a certain software and does not prevent them from providing related services, information companies will collect it, and when there is a dispute, they will say that there is no violation of the law. And I've got the user's authorization. However, when information companies obtain users' private information unrelated to services, they have violated users' right to decide on personal information, or at least reduced users' freedom of exercise. It is obvious that these so-called privacy authorization statements have essentially become the disclaimer of information enterprises rather than the proof of users' rights protection.

Another common phenomenon among the public is "big data kill familiarity". When a user frequently uses an application, the information platform can create a portrait of the user according to the user's transaction records, trading habits, consumption level and other parameters, provide products in line with consumer preferences, and at the same time conduct differential pricing for new and old consumers to obtain higher profits. Such behavior makes us be swayed by the results of the so-called data analysis, as if the best understanding of us is not ourselves, but big data. Moreover, information platforms tend to sell the public information they collect to other market players, leading to the leakage of consumers' personal information, and the same situation happens on other platforms. When a user talks about the need to buy a certain product in a chat app, if he enters another shopping app, he may find that the recommended term contains the product he wants to buy. This is a very scary thing to think about. Generally speaking, the chat content is the private content we don't want to be known, and this information is at risk of being stolen, making us naked in front of big data, with no privacy at all. Therefore, "Big data kill ripe" not only damages the economic interests of consumers, but also infringes on citizens' right to personal information and privacy. However, such operations as information platforms are often hidden and indirect, which lead consumers' consumption behaviors imperceptibly, making it difficult for people to detect. Even if consumers find their own interests are harmed, it is difficult to collect strong evidence to protect their rights.

It cannot be denied that the application of data mining technology to provide customers with personalized services has indeed improved the user experience to a certain extent and met the needs of users efficiently and accurately. However, we cannot ignore the limitations of the technology itself. At present, the data mining technology is not mature, there are problems such as low quality of data, deviation of data interpretation and so on. For example, data analysis may conclude that there is a causal logical relationship between two things from a pile of data, but in fact the two variables are not correlated, only have the same or similar change trend in time or space. In addition, although the data analyst will clean the data before data mining, there is no guarantee that there are still fake, outdated or fundamentally wrong data in the mass data, which greatly reduces the quality of data analysis. In addition, the legal problems caused by data mining technology such as data leakage, data crawler, data trading and data unfair competition need to be solved.

4. The Legal Regulation Path of Data Mining Technology

All the above problems will aggravate the public's negative feelings of aversion, resistance and doubt towards data mining technology and even technological development. In addition to the

improvement of technology and the enhancement of information enterprises' social responsibility, we need more perfect legislation and a more sound legal system to ensure the solution of these social problems and provide legal protection for citizens' right to personal information and privacy. Escort the sound development of data mining technology related industries, provide legal support for national data security, so as to achieve the interactive integration of cutting-edge technology and legal regulation.

4.1. The Principle of Necessity is Concretized in the Legal Regulation of Data Mining Technology

Whether the use of relevant software is subject to consent to the privacy licensing agreement is an issue that we need to focus on in the legislative stage of data collection. It is necessary for lawmakers to think about the legality and rationality of relevant software developers to obtain personal information unrelated to software use, and it is necessary to refine the scope of users' personal information that software developers in each field can obtain, and make specific regulations in the form of legal provisions, rather than giving the public a broad, abstract and difficult to exercise the right of protection.

4.2. Data Should Be Anonymized before Data Mining

After the information enterprise collects the legitimate data, it involves the pre-protection measures for the user's personal information and privacy before data processing. Generally speaking, the anonymous processing of alternative data will not have a great impact on the accuracy of the data and the results of data analysis, nor will it produce a great technical cost, but it can effectively prevent the disclosure of citizens' personal information and infringement of privacy rights, reduce and prevent the occurrence of disputes. The legislature may consider prohibiting any real-name data processing that may damage the legitimate rights and interests of users. This measure will effectively prevent data mining technology from infringing on the rights and interests of personal information and privacy.

4.3. Strict Legal Liability of Data Mining Subjects

Through the collection and analysis of information infringement cases caused by data mining technology in the past ten years, we can conclude that other information enterprises whose information property rights are infringed and the public whose personal information rights and privacy rights are infringed bear most of the burden of proof in safeguarding their rights, while illegal data mining is often concealed and indirect in operation. Moreover, it is easy to erase the traces of illegal mining by technical means, resulting in the victim's difficulty in proving, unable to prove, and no way to safeguard his rights. In this case, the law should be more inclined to protect the well-intentioned and vulnerable information enterprises and individuals, rather than blindly emphasize the fault liability system. In such an urgent situation, lawmakers can consider applying strict liability system to the subject of data mining in related legal proceedings caused by data mining technology, and tilt the burden of proof to the relatively advantageous party. In this way, the data mining subject will consider the legal cost of illegal mining, which is helpful to curb the information infringement behavior from the source.

4.4. Detail the Content of Personal Information Rights

China's Civil Code and Personal Information Protection Law currently do not stipulate that citizens have the right to personal information, but express citizens' personal information as a right (footnote: See Article 111 of the Civil Code of the People's Republic of China and Article 1 and 2 of the Personal Information Protection Law of the People's Republic of China). The mainstream view of the academic circle believes that the content of personal information rights and interests includes information decision, information confidentiality, information inquiry, information deletion, information correction, information blockade, the control and control of

personal information and the exclusion of infringement by others according to law. Although the law clearly stipulates citizens' personal information rights and interests, the specific scope of the above rights and interests is not clear or overlapping. Relevant legal interpretations should also be introduced to help the public understand the content of the protection of personal information rights and interests, and enhance the public's awareness of personal information protection and willingness to seek legal relief when infringed.

However, it is not enough to only allow citizens' personal information protection to stay at the level of rights and interests. In the current situation of data mining technology and citizen information protection in such a fierce conflict, it is an effective measure to establish citizen's information as a legal right.

On this basis, lawmakers can also learn from the rights content of European countries on citizens' right to personal information, and appropriately increase the content of the right to personal information in light of our national conditions. In 2014, the Court of Justice of the European Union issued a ruling in the case of Google Inc and Google Spain v. Spanish Data Protection Authority, ordering Google to remove links to outdated and irrelevant information about the parties involved. The case established the "right to be forgotten" across the European Union. The "right to be forgotten" refers to the right of information subjects to request operators to remove links to their personal information if it is inadequate, inaccurate, irrelevant or beyond the purpose of information processing. However, judging from the judgment of Chinese court on Ren's right to be forgotten in 2015, China does not recognize and protect the "right to be forgotten". Admittedly, the legalization of the right to be forgotten may not be beneficial to the supervision of public opinion, may also cause the instability of practical information resources, and increase the difficulty of the management of information enterprises. However, in the case of a case, if refusing the "right to be forgotten" will indeed lead to an unfair trial, we can still consider transplanting the "right to be forgotten" or adopting the idea of protecting personal information behind the right.

5. Conclusion

With the development of the Internet economy, digital China and China ruled by law have entered a new journey. We need to face the modernization and the world and build a system of legal disciplines with Chinese characteristics. The application of data mining technology should also adhere to the people-oriented value orientation, let science and technology in the track of the rule of law steady and far, answer the questions of China, The Times, the people, the world, to realize the development of technology for the people, science and technology for the benefit of the world. The ultimate purpose of legal regulation path research of data mining technology is to promote the good law and good governance in the digital field and create a better life for the people.

Acknowledgments

This paper belongs to the research result of "Legal Regulation Issues of Data Mining Technology" (No.: 202210378299), a national innovation and entrepreneurship training project of Anhui University of Finance and Economics in 2022.

References

- [1] Tang Qitao. Security Management Measures for Network Data Mining in the Information Age. China Security Science Journal. Vol.32 (2022), No.4. Page 1.
- [2] Tu Yongqian, Wang Xiaotian, "Legislative Improvement of Personal Credit Information Protection under the Background of Big Data". Internet World.2017.5.

- [3] Zhao Zejun, Jiang Lihua. Research on Data Mining Technology and Its Legal Issues. Journal of Zaozhuang University. Vol.36 (2019), Issue 1.
- [4] Qi Aimin, "Conflict of Rights and Legal Regulation in Data Mining", Journal of Guangxi Political Science and Law Administration Cadre Institute. Vol 33 (2018), Issue 4. Pages 7 – 8.
- [5] Ding Yuxiang, Situation and Judicial Development of the Right to Be Forgotten in China: Starting from the First Case of the Right to Be Forgotten in China, Legal Research, 4th issue, 2018. Pp. 35-37.