Explain the Phenomenon of "Ripening" of Big Data Based on Technical and Social Aspects

Rui Gong(✉)

Department of Business, Shanghai University of Engineering Science, Shanghai 201620, China
grsmile1998@163.com

Abstract. Nowadays, life has been occupied by digitalization, which brings many conveniences as well as negative effects. Big data technology can make information explode and spread widely, and can promote the leapfrog development of the country. But at the same time, big data technology is also a double-edged sword, and the "maturity" of big data is one of its shortcomings. In order to effectively curb the phenomenon of big data ripening, ensure the maximum interests of users, and provide theoretical support for dealing with the phenomenon of big data "ripening". First, define the definition of big data "killing ripe"; Secondly, analyze the reasons for the phenomenon of big data "killing ripe" from the social level; Analyze the principle of big data "killing ripe" from the technical level. Finally, countermeasures are put forward to curb the phenomenon of big data maturity. The phenomenon of big data "killing" is the result of the joint action of consumers, platforms and laws. It is suggested to start with enterprises, consumers, relevant departments and national laws to curb the phenomenon of big data "killing".

Keywords: big data "ripening" · Social dimension · Technical level · Coping strategies

1 Introduction

Nowadays, the era has changed into the era of artificial intelligence. With the continuous progress of the information industry, our lives have been occupied by digitalization. The use of big data by businesses and legal institutions has not only brought us a lot of convenience, but also brought negative effects. Big data "killing ripe" is one of them [1]. The phenomenon of big data "killing" is real and growing. For example, Ms. Hu in Zhejiang ordered hotels on Ctrip app, and finally found that the order price she paid was twice the actual listing price of the hotel; When shopping on a shopping platform, Ms. Han in Beijing found that the price she saw was twenty-five yuan higher than what her friends saw; Professor Sun of a well-known university led his team to conduct a research on "taking a taxi with mobile taxi software" and came up with a research report. Through the research report, it is found that the taxi prices of mobile phones of different brands are inconsistent. The survey verified that the phenomenon of big data "killing" really exists. Big data "killing" is essentially big data analysis, which is the product of the rapid development of the Internet in recent years. Many Internet platforms continued to spend

© The Author(s) 2022
G. Ali et al. (Eds.): ISEMSS 2022, ASSEHR 687, pp. 3418–3425, 2022.
https://doi.org/10.2991/978-2-494069-31-2_401

money in order to expand the market in the early stage. After expanding the market, the Internet platform has a huge user base. After several years of accumulation of user data, the platform can analyze the characteristics of each user, so that the Internet platform can provide personalized services for different users, which is good in nature and meet the diverse needs of different users. However, the phenomenon of big data "killing" is that the platform uses good technology in the wrong place to maximize profits for businesses. Therefore, in order to effectively curb the phenomenon of big data "killing" and avoid the damage to the interests of users, it is of great practical significance to analyze the causes of big data "killing" and put forward countermeasures based on this.

2 Definition of Big Data

Literally, big data "kill" refers to the same goods or services. Customers who often use the platform, i.e. regular customers, see that the price is much higher than that of new customers. In real life, not only regular customers but also those who are not price sensitive will encounter "killing". Therefore, big data "killing" is more regarded as "killing people who are not price sensitive" [2]. However, scholars mainly define the "maturity" of big data from three aspects: objective characteristics, subjective purpose of operators and technical path. Scholars who emphasize objective characteristics believe that the so-called "killing off" of big data is a phenomenon of price discrimination [3]. Starting from the subjective purpose of operators, some scholars define "killing" big data as an infringement, which damages the interests of consumers [4]. Scholars who start from the technical path believe that big data "killing off" means that businesses collect consumers' personal information, then draw a picture of consumers according to their consumption habits or preferences, and then implement differentiated pricing through algorithms, so as to finally achieve the purpose of accurate marketing and maximize benefits [5].

To sum up, big data has the following characteristics: Fistly, differentiated pricing. Different consumers will pay different prices for the same commodity or service on the same platform under the same transaction scenario and the same consumption time. Second, the realization of the process needs to rely on technical means. Businesses must rely on big data technology and algorithm technology when carrying out big data "killing" behavior. Third, the form of "ripening" is hidden and difficult to find. Most consumers will only compare the prices of the same service or commodity on different platforms, rather than compare prices with different consumers. Therefore, it is difficult for consumers to be found after being "cooked".

3 Causes of Big Data "Killing"

This chapter will analyze the reasons for the phenomenon of big data "killing" from the social level. In this paper, the social level is divided into consumer plate, relevant department plate, platform enterprise plate and relevant legal plate.

3.1 Consumer Segment

Fistly,the user stickiness is high. Generally, when consumers have a good impression on a certain service or commodity of a certain platform, they will often browse the platform and consume the service or product for a long time, and eventually become loyal users of the platform [6]. Users have a high degree of dependence and stickiness on the platform. When the user stickiness increases, it will improve the tolerance of the "killing of mature" phenomenon of the platform, resulting in frequent "killing of mature" phenomena that can be relied on for survival. Second, consumers do not understand the relevant concepts of big data "ripening". Existing studies have shown that [7] more than 40% of consumers have never heard of the concept of big data "ripening", and only 4.22% of consumers know the concept of data "ripening". Third, consumers' awareness of privacy protection is relatively weak. Previous studies have shown that [7], when checking the privacy agreement, nearly 90% of users will not carefully read the privacy terms. Fourth, consumers' awareness of personal rights and interests protection is weak. Studies have shown that [7], less than 5% of consumers actively take measures to protect their rights after they know that they have been "killed". Even if they have the awareness of protecting their rights, most people do not know how to protect their rights and how to protect their rights is wrong.

3.2 Relevant Departments

To curb the phenomenon of big data "killing" requires the supervision of relevant departments and the coordination and cooperation of multiple departments. Including industry and commerce, industry and information technology, Internet information office and other departments. However, at present, there is still no relatively complete regulatory system among multiple departments, which provides a relatively loose environment for businesses to "kill" big data.

3.3 Platform Enterprise Segment

Big data "killing off" means that businesses, due to their greed [8], implement differentiated pricing for consumers through big data technology and algorithm technology in order to seek private interests.

3.4 Relevant Legal Sections

At present, there is no legal characterization of big data "killing" in China, and the relevant laws and regulations are imperfect. For example: First, the legitimacy of data acquisition is not defined. Second, the ways and means of using consumer data on Internet platforms are not standardized. Third, there is no provision on the relevant punishment after the big data "killing" behavior occurs.

4 Big Data "Kill" Principle

This chapter will analyze the principle of "killing the ripe" of big data from the technical level. There are usually four steps to realize the "killing" behavior of big data. They are: defining labels, data collection and analysis, building user portraits through algorithms, and pricing differentiation through algorithms [9].

First, define the label. Draw different labels for users at the product level. When defining labels for users, we can roughly divide them into four categories, as shown in Fig. 1. The basic tag refers to the user's basic information, including age, occupation, gender, education, etc. The behavior tag refers to the user's personal behavior when using the platform, including the number of likes, purchase records, etc. The preference tag refers to the user's personal preferences, such as whether the user likes food or travel. The prediction tag is a personal prediction of the user's recent behavior, such as whether he plans to travel in the near future and whether he plans to buy a house in the near future.

Second, data collection and analysis. After the label is defined, the data is collected, processed and analyzed. The whole architecture includes data acquisition layer, data processing layer and data service layer. There are two ways for the platform to collect data [10]. One is to collect personal information legally. When logging in and registering, users will generally upload their personal information and sign privacy contracts as required. The platform can legally collect users' personal information. These data information will become the data of the platform. Second, the platform collects users' personal information by itself through algorithms. When users use a certain software, they will generate data such as browsing times. For some shopping software, there will also be data such as the number of orders, evaluation times, return times, and complaint times. These data are combined with platform data and eventually become platform big data. The platform collects this kind of data by itself, and then forms a "data warehouse".

Third, build user portrait through algorithm. After collecting and analyzing user data, the user portrait is constructed by algorithm. According to the overall life cycle, users can be divided into new users, old users, potential users, etc. Then, for different user groups, it subdivides the purchase habits and shopping preferences of users. In practice, users can be divided into many types. Although the division methods are different, the principles are basically the same, that is, users can be divided into different types according to a variety of judgment criteria. As shown in Fig. 2, it is assumed that the platform division divides users into three categories: stable after-sales customers, price insensitive customers and high viscosity customers. If the number of times you place an order for similar goods is less than the average value of users, the total price of the order ranks among the top 30% of all users, and the frequency of coupon use is less than the average value, it indicates that you are a price insensitive customer. Judge whether you are a stable after-sales customer through three aspects: the number of returns and exchanges, the number of complaints, and customer chat and sending statements. Determine whether you are a high viscosity customer by the length of APP use and the depth of APP pages. Thus, the user portrait belonging to you is constructed by algorithm.

Fourth, differential pricing is implemented through algorithm. For different user personas, the algorithm is used to price each user differently. In the process of realizing differentiation, the price or service is adjusted by algorithm to ensure that the transaction

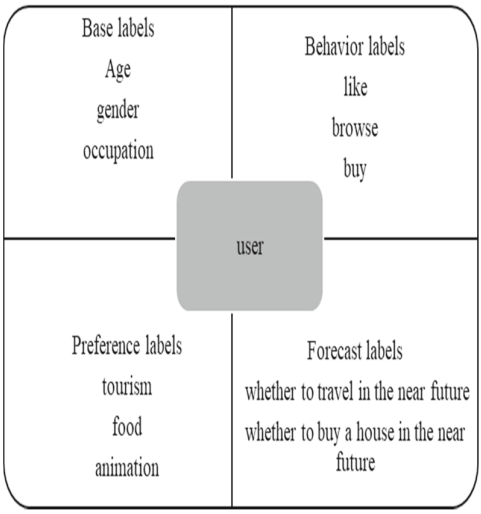

Fig. 1. Label type

is facilitated with the maximum probability and the maximum profit is obtained. Common forms of differentiated pricing include issuing coupons, directly adjusting prices or adding additional services.

5 Coping Strategies

This chapter will put forward countermeasures to curb the phenomenon of "killing" of big data from the four levels of platform enterprises, consumers, regulatory authorities and laws.

5.1 Platform Enterprise Level

Businesses should establish a correct concept and abandon the bad competition with other enterprises. The sales platform should consciously abide by the principles of moral constraints and data transparency. Enterprises should achieve relevant industry norms and industry self-discipline. Businesses should consider social responsibility rather than pursue the maximization of interests, and adhere to the concept of "people-oriented". Protect the rights and interests of consumers. For example, it can make users not accept the merchant's regulations on the collection of users' personal information and the merchant's right to use users' personal information.

5.2 Consumer Level

Consumers can avoid over reliance on a platform. If you use an app for too long, consumers can uninstall the app and reinstall it in a few days. This behavior will trigger an early warning of customer loss on the platform, and the platform will try its best to

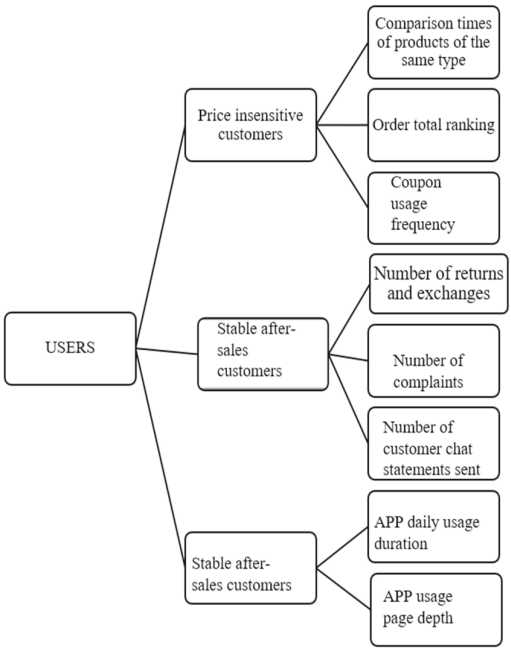

Fig. 2. User portrait

save you and give you new treatment. In online shopping, you can compare prices on different platforms, with friends around you, or with third-party websites. In addition, consumers should improve their awareness of privacy protection and rights protection. Big data technology enables your behavior to be "watched" by the platform, including not only your basic personal information, consumption habits, frequency of platform use, but also the geographical location of users, as well as monitoring your life information through monitoring functions. Therefore, consumers should carefully read the privacy terms when registering the software. When knowing that they have encountered big data "killing" or found that their privacy has been violated, consumers should have the courage to expose big data "killing" or privacy violations to relevant departments.

5.3 Relevant Department Level

Relevant departments should work together to establish a complete supervision system. The big data technology can be used to supervise whether the enterprise implements "killing mature" behavior, and then the obtained analysis results can be fed back to users. At the same time, since big data "killing" depends on big data technology, relevant departments can set up relevant teams to use big data "anti killing" technology to contain this phenomenon.

5.4 Legal Level

The state should improve laws and regulations. Clearly define the "maturity" of big data. Define the legitimacy of consumer data obtained by the platform. Regulate the relevant penalties after big data "killing" behaviors, such as formulating huge fines, increasing illegal costs, etc.

6 Conclusions

This paper first defines and summarizes the characteristics of big data. It is found that big data "killing" mainly has three characteristics: differential pricing for different users, the realization process of big data "killing" depends on technical means, and the form of big data "killing" is relatively hidden and not easy to be discovered by users. Secondly, it analyzes the reasons for the "killing of big data" from the consumer sector, relevant departments, platforms and relevant legal sectors. Finally, analyze the principle of big data "killing ripe" from the technical level. In order to effectively curb the phenomenon of big data "killing", this paper puts forward suggestions and countermeasures from the platform enterprise level, consumer level, relevant department level and national legal level.

The phenomenon of big data "ripening" can be seen everywhere. On the surface, the phenomenon of big data "ripening" is price discrimination, but from a deeper level, it is the lack of personal data protection in China and the weak awareness of consumer rights protection. Therefore, this paper believes that the "killing of big data" is the result of the interaction of consumers, relevant departments, platform enterprises and other parties. It is the common task of consumers, relevant departments and national legal departments to eliminate the phenomenon of big data "killing".

References

1. W. Zhou. regulation of algorithmic collusion. Law, 2020, (01): 40-59.
2. X. F. Qiu, J. Zhang. Analysis on big data maturity and regulation of e-commerce platform. Journal of Hubei open vocational college, 2022,35 (05): 114-116.
3. J. H. Zhu. Legal regulation dilemma of "big data killing" and its solution. Journal of Chongqing University of Posts and Telecommunications (SOCIAL SCIENCE EDITION), 2021,33 (01): 64-72.
4. K. L. Zou, J. M. Liu. Dilemma and solution of legal regulation on big data "killing" only from the perspective of consumer protection law. Price theory and practice, 2018, (08): 47-50.
5. Y. Jiang. Algorithm training and algorithm training: legal regulation of algorithm in the era of artificial intelligence. Hebei law, 2018,36 (12): 142-153.
6. Q. Yuan. Thinking and Analysis on the phenomenon of big data "killing" from the perspective of Matthew effect. Western radio and television, 2022,43 (05): 29-31.
7. Y. N. Xu,Q. Q.Wang, R. Dong et.al. Analysis on the current situation, problems and Countermeasures of "big data slaying". Reform and opening up, 2019, (01): 15-20.
8. F. Yuan, D.J. Wang. Research on the "killing" behavior and improvement strategy of meituan takeout big data. China storage and transportation, 2021, (12): 160-161.
9. Z.G. Wang, H.Y. Li, L.L. Sun. Development and governance of digital Finance: from information asymmetry to data asymmetry. Financial theory and practice, 2019, (12): 25-30.

10. Y. C. Hu, Y.F. Feng. Research on the protection of consumers' fair trading rights in big data ripening. Journal of Shaanxi Normal University (PHILOSOPHY AND SOCIAL SCIENCES EDITION), 2022,51 (01): 161-176.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.