

Research on intelligent information system of user intelligent behavior data based on computer big data

Siyu Zhang*

Shanghai University Of Electric Power, Shanghai, China

*Corresponding author e-mail: siyuzhang1248@163.com

Abstract. This paper collects and analyzes users' online behaviors through computer big data technology, uses computer big data intelligent analysis system to analyze users' online shopping behavior, and uses information-based data analysis system to detect consumers' online shopping needs under the e-commerce platform. The main technique used in this paper is the computer browser log mining method. In the user's click stream data, the function keys of Tmall and Taobao webpages are used as data information for classification and collection. This paper uses the Bisecting K-means clustering algorithm to mine the state of interest. Finally, the feature maps of interests and behaviors are summarized. By processing four typical types of e-commerce user demand status, including background management type, continuous search type, product browsing type and information search data, and clustering based on page type, an effective method for dynamically changing demand judgment is obtained. The state of online shoppers is analyzed through data processing, which also proves the effectiveness of the computer intelligent information system.

Keywords: Computer, online shopping big data, artificial intelligence, shopping behaviour, Intelligent Information System.

1. Introduction

In recent years, with the rapid development of Internet technology, the scale of e-commerce consumers has continued to expand. By means of computer big data technology, we can deeply mine, analyze and analyze the user's purchasing behavior, and collect the corresponding user behavior data. Computers process and analyze data, and then make decisions to provide reliable technical support for user selection and business operations [1]. In the physical shopping environment, experienced computer big data information collection technology can observe customer behavior at any time, and consumers' needs are easy to find [2]. At present, most e-commerce websites analyze consumer behavior data through advertising, push and other marketing methods.

2. Research framework

The computer data analysis system proposed in this paper can analyze the user's online browsing law based on the analysis of human dynamics and consumer purchase theory, and then obtain the corresponding data analysis system. The research framework of this paper is shown in Figure 1 (the picture is cited in Customer Purchase Data Prediction in E-Commerce: A Conceptual Framework and Research Agenda). Users are driven by interest to generate browsing behavior. These user data are collected by computer big data systems, and interest will decline after a period of time; if the user is affected by information input during the browsing process, the user's interest may increase or decrease [3]. The output of the computer user's state of interest may be: attention, understanding of information, attitude toward commodity preference or disgust, etc. unknown interest state. These interest status information data are collected by the computer big data system, and the computer system selects various situations that the user shows in the actual shopping. The output of these interest states can influence the decision-making of its browsing behavior. During this process, stimuli keep going back and forth, users' interests keep fluctuating, and browsing behavior keeps changing. All data are collected by the computer big data system and entered into the analysis and processing link.

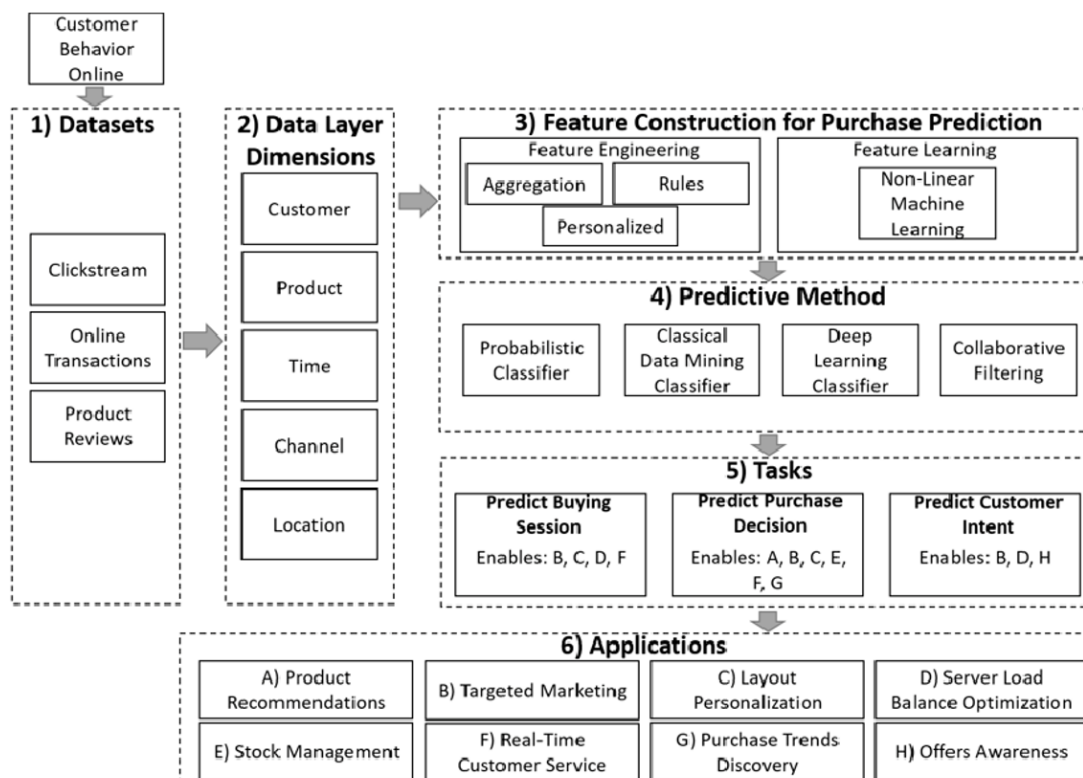


Figure 1. The research framework of this paper

2.1 Data collection

User data sources include user data in e-commerce platforms, user data in social networks, user data in mobile devices, etc. User data is created in the form of "streams", because there is interaction between the three data sources, and their data The content is often overlapping, so it is classified according to transaction, interaction and observation data, and then collected in the process of user consumption or other behaviours through tools such as Needle base.

2.2 Data reprocessing

Data reprocessing includes data preparation, data transformation and data extraction. Data reprocessing determines the quality of mining results. To some extent, data reprocessing often determines the success or failure of data mining. Since there are noise data in the original data, Redundant data and missing values, etc. During the data preparation process, the data is parsed, cleaned, and reconstructed, and missing values are filled to improve the quality of the data to be mined. Then, the unstructured and semi-structured data prepared through the data are processed [4]-[5].

3. Modelling of real-time interest in user purchase behaviour

In online shopping, initially all users see the same content and structure of the website, but once the user's interest changes are identified, the system will adjust the content and structure of the webpage to form a personalized recommendation [6]. Aiming at the difference of users' interests in each web page, this paper firstly classifies pages based on keywords based on click stream data, and then conducts Bisecting K-means clustering modelling of web pages from implicit interest variables, and then analyses the individual users browsing different web pages. Identify the real-time potential interest state [7]. This paper mainly considers feature construction from five aspects, and combines these features to form a training set, a validation set and a test set, as shown in Figure 2 (the picture is quoted from Goal-oriented modelling and verification of feature-oriented product lines).

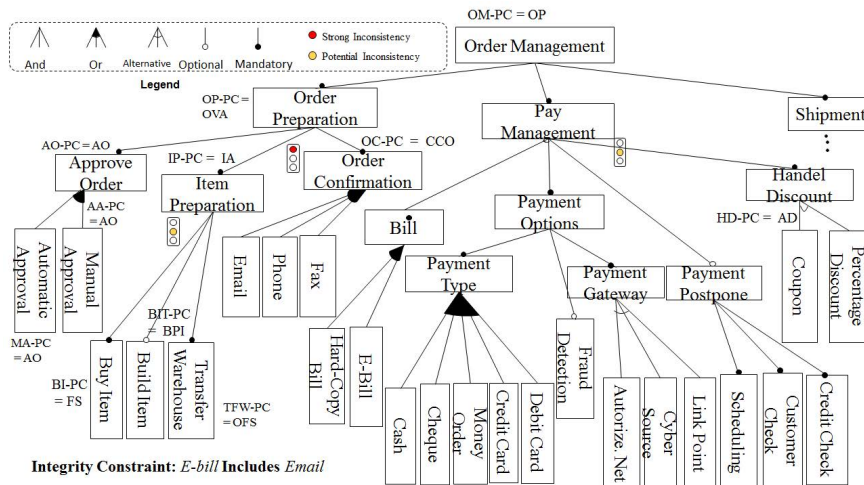


Figure 2. Model diagram for building features

Assuming that the user set is U , the commodity set is I , and the interaction relationship between users and commodities is $R \subseteq U \times I \times O \times T$, where $O = \{0, 1, 2, 3, 4\}$. 1, 2, 3, 4 Represents browsing, favorites, add-ons, and purchases [8]. T is the time when the operation behaviour occurs, if $r \in R, r = \{u, i, o, t\}$ means that user u operates o on commodity i at time t . T_d is the time set for predicting the operation behaviour on the day of the day, and T_{d-j} is the time set for predicting the operation behaviour on the day j days before the day before.

3.1 User commodity pair characteristics.

Pay attention to the user's recent behaviour on the product, which can represent the user's intention to buy or not to buy the product [9]. The extracted features are as follows: the user's various behaviour counts n days before the prediction date, denoted as $|M_n^o|, M_n^o = y | y \in U_{j=1}^n T_{d-j}, (u, i, o, y) \in R$, where $o \in O, u \in U, i \in I$; the total number of all user behaviours n days away from the prediction date, denoted as $|M_n|, M_n = \bigcup_{o=1}^4 M_n^o$; due to the user's shopping behaviour is cyclical, so Take one week as an inspection period, n can be $\{1, 4, 7\}$; the time difference between the time of the user's last actions on the product and the forecast date is recorded as $T_d - T_o$, where T_o is the user's last action on the product.

3.2 User Features.

User characteristics are attributes that describe a user. The extracted features are as follows: the count of all behaviours of user u from the n days before the prediction date, denoted as $|C_n^o|, C_n^o = \{(x, y) | y \in U_{y=1}^n T_{d-j}, (u, x, o, y) \in R\}$, where $o \in O, u \in U$; the total number of all behaviours of the user n days before the prediction date, denoted as $|C_n|, C_n = \bigcup_{o=1}^4 C_n^o$; the average time interval of the user's purchase behaviour; the user clicks to buy Conversion rate $|R_4|/|R|$ (the number of purchases by a user accounts for the number of clicks on a product, where clicks are the collective term for the interaction between users and products). The total number of purchase behaviours of user u is R_4 [10]. $R_4 \subseteq \{(u, i, 4, t) | (u, i, 4, t) \in R\}$.

4. Experimental Design

This paper adopts the method of computer big data system analysis and the method of laboratory field experiment. 10 college students are randomly selected to purchase online products, record user

behavior data, construct click stream data, and identify and manage users' dynamic interests [11]. Then the corresponding user data is analyzed, and the specific process for the experiment is shown in Figure 3 (Scientometrics Review of Urban Logistics Literature: Research Trends, Advanced Theory and Practice) [12]. The experimental process is divided into three stages: preliminary pre-experiment, to obtain the experimenter's daily e-commerce consumption habits through a questionnaire; formal experiment, requiring the experimenter to purchase goods on the e-commerce website, and the extension plug-in embedded in the Google browser will be automatically obtained during the experiment. The experimenter's browsing behavior data; in the final stage, the experimenter's preliminary questionnaire is compared with the behavioral experiment results, and in-depth interviews are conducted, and buyers who successfully pay will receive a subsidy of 30 yuan. 20 yuan subsidy does not buy. The dataset includes data from 14 sessions, 1772 pages of access logs and mouse clicks, in which 7 purchases occurred. These data are automatically collected by the system and analyzed by the big data system.

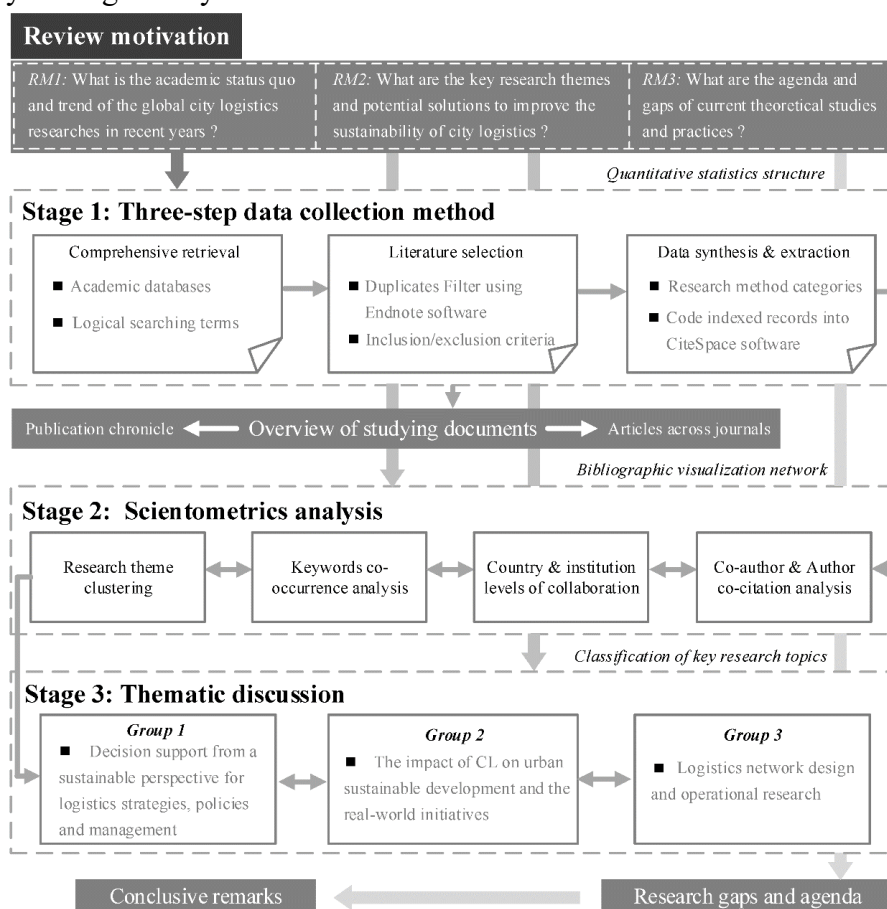


Figure 3. Three-stage experimental design framework

Using a finer data granularity to classify pages based on function key elements, in the processed data set, the page set U contains $B_n = 14$ -page categories, and the descriptive statistics are shown in Table 1.

Table 1. Page category statistics description table

Abbreviation	class name	Frequency	Frequency (%)
H	Home	138	7.79
A	account	96	5.42
S	Pay to buy	7	0.4
D	Add to Cart & Favorites	30	1.69
F	shopping cart	52	2.93
G	commodity	170	9.59

R	Evaluation	11	0.62
B	brand or flagship store	142	8.01
P	price	17	0.96
Y	Popularity	5	0.28
V	sales	4	0.23
T	Product attributes	588	33.18
C	content	438	24.72
O	other	74	4.18

As shown in Table 1, among the 14 types of pages, users browse the most pages of "product attributes" (T) when shopping online, accounting for 33.18%. Product Catalog (C) listing, followed by Product (G) page. Among the pages that can reflect the user's preference for products, the "brand" (B) pages are the most, while the evaluation, price, popularity, and sales are relatively few. Based on the above considerations, the behaviour of browsing the shopping cart and adding items to the shopping cart is attributed to the behaviour of using the shopping cart, and it is believed that the purchase intention interest occurs when the user browses the web for a long time and uses the shopping cart frequently.

5. Conclusion

This paper collects and analyzes the user's network behavior through computer big data technology, and uses the information-based data analysis system to detect the online shopping needs of consumers under the e-commerce platform by mining the user demand status under the e-commerce platform. The background management type, continuous search type, product browsing type and information search data of the computer big data system are processed. Through the data processing system, the system can draw corresponding conclusions, break the existing single e-commerce marketing model, and further detect consumption. Psychological and consumption motivation, and real-time shopping guide according to the user's dynamic demand status, to help users find their own needs.

References

- [1] Yan Kun, Shen Subin. A User Behaviour Prediction Method Based on Smart Home. *Computer Technology and Development*, 30(1) (2020) 6-12.
- [2] Zhao Likun, Wang Yuke. A recommendation algorithm based on artificial intelligence for social network user behaviour data cycle. *Science Technology and Engineering*, 20(28) (2020) 61-64.
- [3] Zhao Yang, Hu Shiyao, Yang Shuqiang, et al. Data-driven power consumption behaviour analysis of users in the electricity sales market environment. *Power Demand Side Management*, 22(4) (2020) 67-74.
- [4] Yang Jie, Li Peilin, Luo Chengchen, et al. Power grid user behaviour analysis based on data mining. *Journal of Yunnan University (Natural Science Edition)*, 42(2) (2020) 44-49.
- [5] Cai Junpeng, Chen Caibo, Chen Dewang, Shen Yong. Research on electric vehicle charging guidance strategy. *Modern Electronic Technology*, 44(1) (2021) 77-81.
- [6] Ge Ning, Chen Xu, Feng Wei. Exploration of New Architecture of Intelligent Mobile Communication. *Chinese Science Foundation*, 34(2) (2020) 48-52.
- [7] Tang Xiaoyue, Zhou Kang, Wang Kai. A spatial context-aware mention target recommendation method. *Journal of Software*, 31(4) (2020) 23-33.
- [8] Li Jian, Zhang Feng, Chen Chong, et al. Research and application of AI-based intelligent voice interactive weather service. *Computer Technology and Development*, 30(12) (2020) 65-69.
- [9] Huang Hehe, Zeng Yuanyuan, Zhang Yi, et al. Prediction method of abnormal crowd aggregation based on group behaviour analysis. *Computer Engineering*, 46(3) (2020) 88-97.
- [10] Li Zeng, Liu Yu, Li Chengcheng. Research on news recommendation algorithm based on user behaviour. *Computer Engineering and Science*, 42(3) (2020) 6.

- [11] Yu Bo, Wang Zhihai, Sun Yadong, et al. Unstructured document sensitive data identification and abnormal behaviour analysis. *Journal of Intelligent Systems*, 16(5) (2021) 99-104.
- [12] Huang Weihua, Fan Xin. Precision Marketing Push System Based on Intelligent Analysis of User Feature Information. *Modern Electronic Technology*, 44(6) (2021) 4-9.