Prediction Method of Consumption Behaviour on Social Network Oriented to User Mental Model

Han Yin(✉)

School of Traffic and Transportation, Xi'an Traffic Engineering Institute, Xi'an 710300, China
xinqiba001@sina.com

Abstract. The current consumption behavior prediction method is mainly to use historical data modeling, through finding the laws of the data, to predict the user's consumption behavior. But the consumption psychology of users in social network will be impacted by the information in social network. The current usage methods ignore the impact of users' psychology on consumption behavior. In order to improve the above defects, the prediction method of social network consumption behavior oriented to user mental model is studied. Psychological characteristics refer to the stable characteristics of psychological activities. After understanding the psychological characteristics of users, the user's social network consumption psychological model is established. Dynamic identification is performed according to the user's personal preferences. According to the fit between user preferences and commodity characteristics, the utility value of commodities is obtained, and the social network consumption behavior is predicted using the differentiation ensemble learning model. The experimental results show that the average prediction accuracy is up to 90.52%, about 15% higher than the original method, and the proposed method has good stability for different conditions.

Keywords: User mental model · Social network · Consumption behavior · Behavior prediction · Subdivision ensemble learning model

1 Introduction

With the way of shopping from offline to online migration of a large scale, the market of online stores is huge, but the same fierce competition. In an era of booming big data and the Internet, the interaction between shopping platforms and social networking platforms has led to a proliferation of valuable data. Relevant businesses and departments make full use of these valuable data and information, target consumers' behaviors in a small way to form targeted operational strategies, and predict the future short-term and long-term economic situation to carry out regulation and maintenance in a higher dimension. Consumer behavior theory is a theory based on a series of factors, such as consumer psychology, desire, priority choice, etc. When consumers have a strong desire to buy goods, the utility of goods is very large; conversely, the utility of goods

W. Fu and L. Yun (Eds.): ADHIP 2022, LNICST 469, pp. 534–545, 2023.
https://doi.org/10.1007/978-3-031-28867-8_39

is small. Businesses can attract a large number of new consumers at low cost if they follow up the data information in real time, on the one hand, to understand consumers, on the other hand, to monitor competitors' trends, and on the other hand, to mobilize commodity prices to launch a variety of preferential activities. After all, in a competitive market, commodity prices matter. But note here that information contains value, but data is cluttered, worthless, even misleading and burdensome if not analyzed in a decent and appropriate way. In addition to understanding how consumers make choices about products and services based on information, the study of consumer behavior can also understand the relevant experience and satisfaction of consumers after making behavioral decisions. It is inseparable from the marketing activities of the enterprise market and is the basis of marketing decision-making [1]. In addition to online consumer buying data reflecting business information, prophase behavior data are also of great value. When businesses accurately grasp consumption habits and preferences, they can develop personalized recommendation and marketing methods to reduce operating costs and increase profits. At present, the prediction of consumer behavior is mainly based on the analysis of consumer online search behavior, comparison and selection information before consumption and historical consumption, and the prediction result of whether the user will buy a certain kind of goods or not is obtained by using the feature engineering of selection model and residual grey model and the method of model fusion, classifier and residual grey model. But also faces many problems, such as cold start, do not consider the content of the product label information.

User's behavior is determined by their psychological state. Studying user's consumption psychology will help to understand user's needs and provide multi-intelligentized services. The traditional research method of consumer psychology based on questionnaire is not only time-consuming and labor-intensive, but also subjective. With the coming of big data era, it is possible to use historical behavioral data to model user psychology. Mental model is the result of the interaction between the user's knowledge system and the current environment, which is similar to the user experience theory. It has qualitative characteristics as well as uncertain characteristics. Psychological model embodies a way of thinking, and becomes a visual angle for people to know things consciously or unconsciously. A mental model is a known experience in which the user knows something new. When users are faced with a product, they tend to understand the new product from their own experience, from the stored knowledge structure and past experience [2]. Using the user's mental model can understand the market of the product from the user's point of view, and get a better user experience. According to the above analysis, the consumer behavior of social network users contains a lot of information that can assist the business to make business decisions. Considering the advantages of user mental model in analyzing user behavior, this paper will study the prediction method of social network consumption behavior oriented to user mental model. The K-means clustering algorithm is used to construct the consumer psychology model of users' social network. The user preference is expressed according to the user emotion calculation in the construction of the user psychological model, so as to improve the accuracy of prediction. Predict the consumption behavior of social network users through commodity utility evaluation, and improve the overall prediction effect.

2 Prediction Method of Consumption Behavior on Social Network Oriented to User Mental Model

2.1 Establishing a Consumer Psychological Model of User Social Network

2.1.1 Predict Consumer Psychology of Users

In order to use the user's mental model, we need to master the user's mental state and the help of cognitive psychology, i. e., the psychology of information processing. Social network users will have different consumption psychology when they spend money. And different consumption psychology may lead to the consumer behavior decision. The psychology of consumers can be divided into quirks, beauty, convenience, reality, difference, habits psychology, economic psychology, different consumer psychology of packaging preferences and needs are different. Social networks provide users with a variety of services, users can enjoy the social network brought about by a variety of services, such as making friends, access to information, games, instant messaging. Every day, users use the Web, especially social networks, to generate huge amounts of data, such as tweets, comments, ratings of products, articles or news likes, check-ins at places of interest, clicking/browsing favorite products or news, favorites and purchases of products of interest, uploading travel pictures to social networks, uploading videos of their own filming to video sites, and so on. User-centered concept has been continuously infiltrated in various industries and fields, and the behavior of users in social networks is the reflection of their true thoughts and thoughts, which plays an important role in analyzing and understanding users' preferences [3]. Therefore, by analyzing the consumption data of user social network, we can build a universal consumer psychological model of user social network, which can help to predict consumer behavior more accurately.

The user behavior data set is processed and the number of occurrences of each behavior is counted. In general, the most frequent click-and-view behavior occurs when users browse social networks. Therefore, click-and-view behavior is the main behavior in data sets. Further, the number of collections and the number of last purchases are similar, and the number of purchases occurred the lowest. This shows that users will be after the collection of goods, the probability of purchase will be greatly increased. To the user clicks to look at the behavior and the frequency of the same item to buy statistics, for the same item, only once to buy the majority of users. The frequency decreases with the increase of the number of purchases. In addition, the analysis of online consumer behavior, data types can be collected complex, data types, including the consumer ID of the integer data, long integer data that time stamps. Data dimensions vary, including the conversion rates of goods obtained through purchase and browsing. The data angle is also very broad, including indicates consumer sex, the age and so on personal information; Represents the product traffic, the sales volume and so on merchant information. Screening This example is intended to determine whether a purchase will occur based on a large amount of previous data. The final data of this study includes consumer ID, commodity category ID, consumer behavior and timestamp of each behavior.

2.1.2 Using k-Means Clustering Algorithm to Build the Consumer Psychology Model of Social Network Users

This paper uses k-means clustering algorithm to cluster the consumption data and consumption psychology of users, and establishes the consumption psychology model of social network users. Consumer psychology is divided into four categories. Consumers in the first category mainly pursue the singularity of the goods, and pursue new packaging, new technology and new technology. Consumers in the second category have higher requirements for the beauty brought by packaging, and attach great importance to the craft, color and shape of the goods, while attaching importance to the economy and uniqueness of packaging. The third category of consumers in the consumption of the main focus on habitual, higher brand loyalty, and pay attention to the practicality of goods. The fourth category of consumers in the consumer's main psychological reality, followed by convenience, in the purchase of convenient, convenient packaging on the basis of more attention to the cost-effective goods.

According to the current types of users' consumption psychology, the cluster is classified by the clustering algorithm, and the user's consumption-related data and consumption preferences are taken as the feature vectors of the clustering algorithm. The initial cluster centers of each fern cluster are randomly selected, and the distance between the cluster centers and the samples to be clustered is calculated according to the following formula:

$$d\left(c_i, x_j\right) = \sqrt{\sum_{j=1}^{n} \left(c_i - x_j\right)^2} \tag{1}$$

Among them, c_i is the cluster center of each consumer psychology cluster, and x_j is the data of consumer behavior to be clustered. The parameters of k-means algorithm are trained by building feature training set manually. After clustering all the consumption data, the distance center of each cluster is updated. The data with the shortest distance from other data in each cluster is calculated as a new cluster, and the clustering process is repeated again until no new cluster center is generated. After the mapping relationship between user's psychology and user's social network consumption behavior is established, the user's consumption psychological model can be obtained.

2.2 Dynamic Identification of Users' Social Network Consumption Preferences

Static personalized recommendations have the ability to identify the user's past preferences, but if the user's preferences change, recommendations based on past preferences do not always provide the user with the greatest utility. As user preferences evolve, users tend to change their behavior over time. The ever- changing preferences of users have a great influence on the accuracy of recommendations. However, user preferences may not necessarily evolve gradually, but can also evolve rapidly between two consecutive periods of time, which means that user preferences may be similar to or completely different from those of the previous period of time in a continuous period of time. Therefore, it has been studied that the user's preference is expressed by a series of behaviors before the purchase. Such as buying, viewing, like points. Generally speaking, user behavior can be divided into goal feedback and assistant feedback. Therefore, considering the preferences of future users can provide more accurate recommendations [4].

In the field of service application, the situation involved is different according to the different application needs, and the more common one is to classify the situation into service requester situation, service provider situation and service situation. In the field of pervasive computing, because pervasive computing is the fusion of information space and physical space, people can get digital services at any time, anywhere and transparently in this fusion space, so situational information plays a vital role in it, and contextual and situational awareness computing has been attached great importance.

The factor of user preference is expressed as the user's preference for product features, including the feature dimension of user preference and the degree of user preference. The reason why users buy goods is that the information reflected by the content of goods accords with the user's preference, and because the user's preference is variable, this paper selects two time intervals of half a month and a month to obtain the historical purchase data of users in these two time periods, and expresses the user's preference through the calculation of user's emotion in the user's psychological model. The specific steps are as follows:

(1) Computation method for user preference of a single user t moment: at any time as t moment, obtain historical purchase data of the user for the previous week up to the t moment, compute the emotional value q of the purchasers' b based on all user behavior data sets of the same category of goods purchased in the history, use the product label as the dimension of user preference, and use the emotional value q as the degree of user preference. If there are k dimensions of user preference, then the individual user t moment preference can be expressed as follows [5]:

$$P_t = \{(b_1 : q_1), (b_2 : q_2), \cdots, (b_k : q_k)\} \tag{2}$$

(2) Computation method for user preference for $t + 1$ moment of a single user: take half a month as time span, half a month after commencement of t moment as $t + 1$ moment, obtain historical purchase data from the t moment to the $t + 1$ moment of the user, take the historical purchase behavior of users of the same type of commodities purchased in the past as the data set, compute the emotional value s of the b of the purchaser, the commodity label as the dimension of user preference, and the emotional value q as the degree of user preference. During this period of time, several uncontrollable factors may change the dimension of user preference or the degree of user preference. Therefore, the updating rules for user preference are as follows: if a new commodity buyer arises, the corresponding emotional value shall be added directly; if several commodities of the same type are purchased, the emotional value of the purchaser $q' = q \times w + q' \times w'$ shall be updated as the new user preference in accordance with the method of weighted proportional accumulation. If the updated user preferences have k' dimensions, the $t + 1$ moment preference of a single user can be expressed in the following formula.∘

$$P_{t+1} = \left\{\left(b_1 : q_1'\right), \left(b_2 : q_2'\right), \cdots, \left(b_{k'} : q_{k'}'\right)\right\} \tag{3}$$

Based on the dynamic identification of consumers' social network consumption preference, this paper predicts consumers' social network consumption behavior by using the subdivision ensemble learning model.

2.3 Predicting Consumption Behavior of Social Network Users

In the actual user social network purchase process, the user's choice depends on his assessment of the utility of the commodity in the transaction process $w_{s,u,i}$, which is a combination of the user's preference and the characteristics of the commodity. Denote the latent factor vectors of users and products as r_u and s_i, respectively, there are:

$$w_{s,u,i} = r_u^T s_i \tag{4}$$

The expression above indicates that the more consistent the user's preferences are with the characteristics of the product, the more the product is in line with the consumption habits of the user, and the higher the utility value of the product $w_{s,u,i}$; on the contrary, the more deviation the characteristics of the product deviate from the user's preferences, and the more inconsistent the user's consumption habits are, the lower the utility value of the product $w_{s,u,i}$ is. For rational users, the probability of choosing a certain item depends on the relative utility of the item in the whole set, and there is a positive correlation between them. Specifically, in the online consumption forecasting problem, users will consider the final opportunity cost and benefit in each transaction process, and choose the goods with utility value not less than $w_{s,u,ioc}$ to purchase. That is, in an arbitrary purchase, the utility of the commodity k purchased by the user shall not be less than that of the best alternative:

$$w_{s,u,k} \geq \max w_{s,u,i} = w_{s,u,ioc} \tag{5}$$

In a user's historical purchase, the final purchase k is logged and therefore always known; the best alternative is not, and the platform does not directly know and record the user's preferences, so any product that appears in the purchase is likely to be the best alternative. To obtain the best alternative to each purchase, the utility value of each purchase is calculated in terms of formula (4) for each application of the product in question. In fact, although the user's preference is not directly reflected in the product sequence of the purchase behavior, it can still be used to analyze the user's preference in the purchase behavior. For example, the number of clicks a user makes on different items in a sequence reflects the user's preferences to some extent. Based on this, this paper proposes a behavior sequence utility function $F(s, i)$, using $F(s, i)$ to analyze the product sequence can be used to estimate and judge the utility of the product to the user. Form of definition: The sequence of clicks made by a user in a historical purchase is recorded as cq in order, the length of cq is N, and the position in which the commodity k constitutes a set $P(s, i)$, the sequence utility of the commodity k in a historical purchase is:

$$F(s, i) = \sum\nolimits_{k \in P(s,i)} e^{-(N-k)/N} \tag{6}$$

The sequential utility of each commodity in all purchasing behavior can be calculated by the above expression. Based on the sequential utility of goods and the dynamic identification of consumer psychology and consumer preference, a subdivision ensemble learning model is used to predict consumer behavior.

The category of a consumer is related not only to his initial state, but also to the state of other consumers in the consumer's social network who are closely related to him. Ordinary consumers will listen to the advice of the buyer, the purchase impulse, become the buyer. In order to facilitate the research, it is assumed that the information that consumers receive is positive, that is, after the general consumers obtain the information, they will have the desire to buy and become potential consumers or buy and become buyers directly. Set rules for the conversion of conduct between various types of consumers as follows [6]:

R1: Interaction between general consumers and purchasers, general consumers may obtain certain product information and then transform into potential consumers;
R2: A potential consumer may become a buyer after interacting with the buyer;
R3: After a period of time, buyers no longer have the desire to disseminate information on the conversion of immunity to consumers;
R4: Immunized consumers will no longer accept information from purchasers, but may continue to purchase products over time:
R5: Buyers will spread consumer information to the general consumer, potential consumers and immune consumers do not have the ability to spread.
R6: General consumer contact with the buyer may lead to direct purchase behavior and thus into the buyer.

The process is as follows:

(1) Amount of information exchanged: The amount of information exchanged when a consumer comes into contact with other types of consumers. At the moment t, agent i and agent j contact, agent i and agent j will change the amount of information in i.
(2) Amount of updated information: the total amount of information on changes in Agent i shall be calculated and updated within the t time limit.
(3) Comparative information content: the conversion of consumer categories shall be carried out according to the information content of agent i and the threshold of information content for each category of consumers.

According to the change of consumers' consumption psychology, historical consumption data and social interaction information are collected, and DSEM is used to predict the consumers' consumption behavior.

DSEM is a nested ensemble learning model, which subdivides the dataset by sample filtering. The model is built from multiple Bagging modules, each of which contains several base classifiers. The training set of each Bagging module in the DSEM is obtained by sample filtering. DSEM can be independently trained to n Bagging modules by sample filtering layer by layer, and then the final learning device can be obtained by set strategy for these n Bagging modules. To implement sample filtering, DSEM sets a strong rule binding policy in each of the Bagging modules, giving the sample that cannot be filtered through the rules to the next Bagging module. Strong rule-combining strategy means that when Bagging model predicts a sample, more than 70% of the base learner must predict the sample into the same class before the model can give the prediction results. For example, in a binary classification problem, the positive sample is labeled as l, the negative sample is labeled as 0, the base learner is $b(x)$, and the strong rule classifier consisting of the base learner is $B(x)$. The strong rule-binding strategy of $b(x)$ is [7]:

$$B(x) = \begin{cases} 1, \sum b(x) \geq r \times E \\ 0, \sum b(x) \leq (1 - r) \times E \\ np, eles \end{cases} \quad (7)$$

The E represents the number of base learners contained in the Bagging module; the r is the strong regular coefficient, $r \in [0.7, 1]$. By setting the combination strategy of strong rules, the model can avoid the random guessing of uncertain samples, so that each layer of Bagging model can only predict the sure samples and filter the uncertain samples to the next layer. Through filtering layer by layer, DSEM can subdivide the original dataset into several subsets, and use different Bagging models to fit each subset to improve the accuracy of prediction.

Each layer of the Bagging module is trained based on SFTraing, giving misclassification and samples that have not passed the rules to the next layer of the model. The training framework based on sample filtering is shown in Fig. 1.

When a test sample enters the DSEM model, it is first predicted by the first Bagging module. If the sample passes through the strong rules of the Bagging module, the classification of the test sample by the Bagging module is used as the prediction result. So far, we have completed the research on the prediction method of consumption behavior of social network oriented to user mental model.

3 Test Study

3.1 Test Preparation and Process Design

This section will test the predictive effect of this method, so as not to fail to achieve the expected research goal in the practical application of this method.

Firstly, the data of social network consumption behavior is preprocessed according to the following contents, using userbehavior.csv data set for experiments, and the actual consumption behavior corresponding to the consumption behavior data is used as a reference for the prediction accuracy.

Data preprocessing mainly includes four operations: duplicated data de-duplicating, missing data processing, unreasonable data processing and data format conversion. Data de-duplication refers to the processing of the user order table. All the fields in the user order table have the same values, so it is necessary to de-reprocess them. There are still some fields in the user order table that have different numbers of items but the values of other fields are the same. This kind of data is de-duplication and the first article of duplicate data is kept. Missing data processing is mainly for the user's age and the

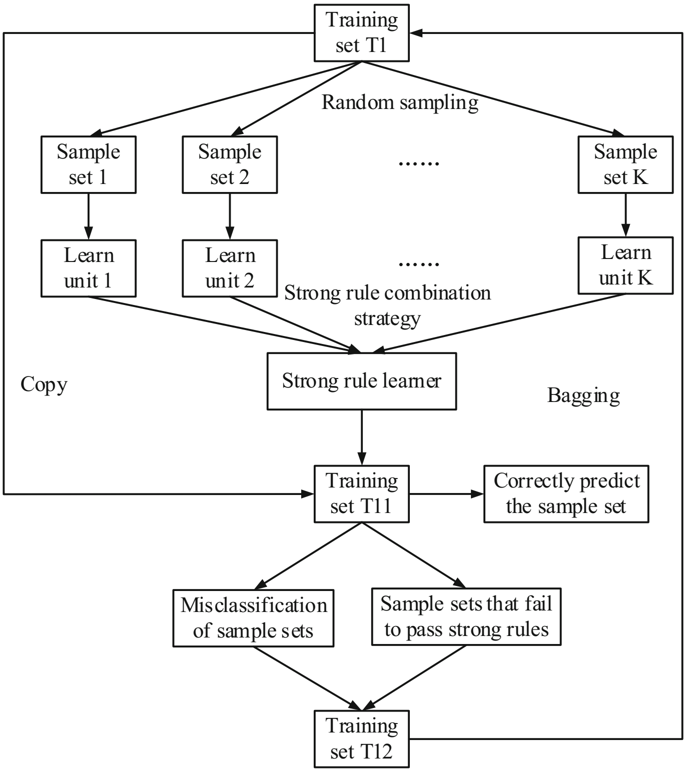

Fig. 1. Training framework based on sample filtering

parameters of the product field, delete the missing data for too much, the missing data for the average filling.

Based on this dataset, several experiments were designed to evaluate the performance of behavioral prediction methods by using some common indexes such as precision, recall, F1 value.

The formula for calculating the accuracy and recall rates is as follows:

$$P = \frac{TP}{TP + FP} \tag{8}$$

$$R = \frac{TP}{TP + FN} \tag{9}$$

Among them, *TP* indicates that the forecast is positive, and the actual forecast is correct. The *FP* indicates that the forecast is positive and the actual forecast is wrong. *TN* indicates that the forecasts are negative and that the forecasts are actually wrong. *FN* indicates that the forecast is negative and that the actual forecast is correct.

The following is the formula for calculating the F1 value:

$$F\_s = \frac{\left(1 + \beta^2\right)R \times P}{\beta^2(R + P)} \tag{10}$$

The F1 value is a tradeoff between accuracy and recall, and it measures the effectiveness of the classification based on the ratio of the recall to the weighted importance of accuracy as determined by the manually set coefficient β. F1 value can reflect the prediction performance of positive samples. Comparing the prediction method proposed in this paper with the behavior prediction method based on support vector machine and the prediction method based on Bayesian model, the performance of consumption behavior prediction method is evaluated by comparing the values of three indexes.

3.2 Test Results

The accuracy and recall of the three methods for predicting the consumption behavior of social network users on the same historical consumption behavior dataset are shown in Table 1.

Table 1. Comparison of accuracy and recall rates of consumer behavior forecasts

Serial number	Method in this paper		SVM based forecasting method		Prediction method based on bayes	
	Accuracy/%	Recall rate/%	Accuracy/%	Recall rate/%	Accuracy/%	Recall rate/%
1	94.28	88.43	82.18	75.66	75.71	73.57
2	90.85	87.95	70.53	76.08	86.14	69.46
3	87.43	84.09	78.75	80.14	63.85	73.76
4	88.17	86.43	84.07	81.55	74.08	64.41
5	91.25	88.47	77.83	77.08	74.91	57.65
6	90.59	85.19	80.62	77.32	60.57	79.34
7	94.27	87.94	72.39	76.83	79.73	63.42
8	87.68	86.26	74.24	79.18	63.64	74.83
9	91.38	87.87	81.64	80.71	84.76	73.49
10	89.26	83.25	73.05	77.46	85.38	72.99

Analysis of the data in Table 1 shows that the accuracy and accuracy of this method in predicting consumer behavior in social networks are significantly higher than those in the other two consumer behaviors. The average prediction precision and accuracy of this method are 90.52% and 86.59% respectively, 77.53% and 78.20% respectively based on SVM, 74.88% and 79.29% respectively based on Bayesian method. From the numerical point of view, the prediction effect of this method is better. Among the different forecasting methods, the difference between the accuracy and the accuracy of the prediction method based on the Bayesian model is small, and the difference between the prediction accuracy and the accuracy of the prediction method based on the support vector machine is obvious. It shows that this method is more stable for different social network users to predict consumer behavior.

The F1 values of the three methods under different coefficients β are compared as shown in Fig. 2.

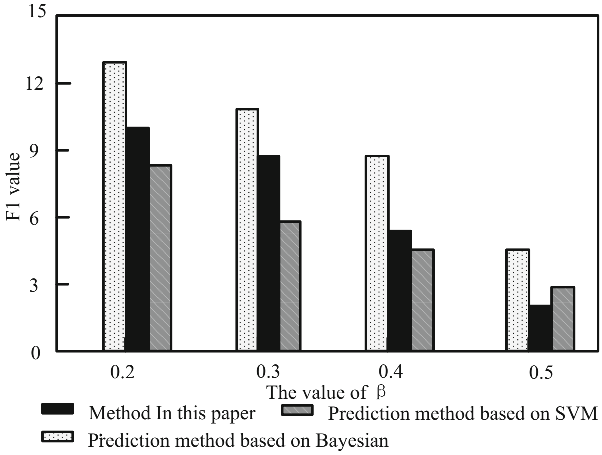

Fig. 2. Comparison of F1 values under different adjustment coefficients

Analysis of the information in Fig. 2 shows that under different coefficients β, the predicted F1 value of the method is higher than that of the other two comparison methods. According to the definition of F1 value, the prediction performance of this method is more comprehensive, and the accuracy and recall are more balanced. To sum up, the consumer behavior prediction method based on user mental model proposed in this paper is effective and can predict consumer behavior accurately, which provides technical support for marketing strategy.

4 Conclusion

In the field of modern consumer behavior research, scholars believe that the ultimate goal of consumer behavior is to make themselves satisfied. With the development of network economy, it can not only affect the offline operation decision of market economy, but also affect the online consumption decision of consumers. Since the vast majority of users do not give explicit feedback (grading and commenting) on the goods they have consumed, the implicit feedback of users plays an important role in grasping users' needs and describing users' psychology. User psychological model can analyze the user's consumption data, can help businesses more accurately analyze the user's psychology and thus accurate sales. This paper studies the consumer behavior prediction method of social network based on user mental model, and uses this method to predict the consumer behavior and consumer propensity in social network. The accuracy of the method is verified by experiments. The method can accurately predict the consumption behavior of users in the social network from their consumption psychology. At present, the research on social networks is mainly comprehensive, and the professional research on social networks is not mature. In future research, professional social networks need to be the focus of research, and the prediction of user consumption behavior needs to be constantly updated to better grasp the needs of users.

References

1. Zhang, F., Ni, R.: Research on online users consumption prediction model based on SOMTE. Math. Practice Theory **50**(15), 49–59 (2020)
2. Zhou, L.: Empirical analysis on the prediction of tourism consumption price index construction for tourism consumption of urban and rural residents—Test based on VAR model. Prices Monthly **06**, 1–5 (2019)
3. Ye, Z., Chen, E.: Study on the prediction of natural gas consumption based on residual grey model. J. Chongqing Inst. Technol. **32**(09), 99–102 (2018)
4. Hu, X., Zhang, H., Dong, J., et al.: Prediction model of user buying behavior based on CNN-LSTM. Comput. Appl. Softw. **37**(06), 59–64 (2020)
5. Han, L., Jiang, G., Wang, H.: Fast mining simulation of online behavioral information characteristics under big data analysis. Comput. Simul. **36**(6), 346–349 (2019)
6. Hu, X., Zhang, H., Dong, J., et al.: Prediction of ensemble learning-based new users' repurchase behavior on e-commerce platform. Modern Electron. Tech. **43**(11), 115–119+124 (2020)
7. Ge, S., Ye, J., He, M.: Prediction model of user purchase behavior based on deep forest. Comput. Sci. **46**(09), 190–194 (2019)