An intelligent optimization method of E-commerce product marketing

Fang Cui[1] · Haihua Hu[2] · Ying Xie[3]

Abstract

In order to improve the marketing effect of e-commerce products, based on machine learning algorithms, this paper constructs an e-commerce product marketing model based on machine learning and SVM. Moreover, this paper studies the classic reinforcement learning algorithm Q-learning and proposes an improved Q-learning algorithm. In addition, this paper uses the method of mean standardization to reduce the noise impact of the reward signal caused by the unfixed time interval between decision points, further constructs a standardization factor for the deviation caused by the asynchronous update of the time interval in the iterative process of the Q value function and updates the standardization factor according to the update method of the value function, and proposes the Interval-Q algorithm. At the same time, in view of the fact that traditional reinforcement learning algorithms cannot effectively deal with the observable part of customer status in direct marketing scenarios, based on the deep reinforcement learning DQN model, this paper combines with the idea of hybrid model to propose a DQN model based on dual networks. Finally, this paper uses public data sets for model training and simulation environment construction and then evaluates the algorithm proposed in this paper from different angles and analyses model performance based on examples. From the research results, it can be seen that the precision marketing model constructed in this paper has a good effect and can be applied to practice.

Keywords Machine learning · SVM · E-commerce · Product marketing

1 Introduction

In simple terms, precision marketing is to use the results of big data analysis and the Internet and big data mining technology to segment the market and customers and have targeted and in-depth communication with segmented customers to understand their needs, so as to push them the most needed products and services at the most suitable time. In this way, it cannot only improve marketing efficiency but also greatly reduce the company's operating costs, and it will help companies achieve low-cost and high-efficiency marketing goals. Under normal circumstances, a company's precision marketing is divided into

the following steps: the first is to collect data and establish a database; the second is to segment the market and target customers; the third is to provide targeted products and services based on the results of the segmentation; the fourth is to implement different marketing programs for different target groups and provide personalized marketing services; the fifth is to continuously improve products and services based on marketing results to meet user needs [1].

In the context of the Internet, precision marketing uses cloud computing, data mining and other technologies based on big data to segment markets and target customers and provide users with personalized products and services. Moreover, it deeply explores the needs of users, and achieves accurate products and services on the basis of meeting user needs, and achieves the purpose of precision marketing. The core of precision marketing is consumer-centric, and everything is consumer-oriented. First, segment the market and conduct in-depth communication with target customers to understand their needs. Then, it formulates marketing strategies based on the target customers' behavioral characteristics and consumption habits. Whether it is a real customer or a potential customer, precision

✉ Haihua Hu
hhh@xauat.edu.cn

[1] International Business School, Shaanxi Normal University, Xi'an 710119, China

[2] School of Management, Xi'an University of Architecture and Technology, Xi'an 710055, China

[3] School of Economic and Management, Northwest University, Xi'an 710127, China

marketing must formulate different marketing plans based on their behavioral characteristics and habit preferences to achieve personalized recommendations. Finally, considering the long-term interests of enterprises, the purpose of precision marketing is to improve marketing efficiency. For real customers, it is very necessary to cultivate their loyalty to products and services, so precision marketing on the other hand can enhance customer loyalty and form a stable customer base [2].

In the context of the Internet era, the application of big data is becoming more and more important for an enterprise. The application of big data to marketing is now a new marketing method that many companies are trying. In the context of increasing data on the Internet, precision marketing has data support and technical support. The company organizes and analyses the collected customer information and formulates different marketing plans for different groups of customers to achieve the goal of precision marketing. The reason why big data can be used in precision marketing is not only that the use of big data allows companies to have a definite target direction for marketing, but also has a great role in achieving precision marketing and improving overall marketing efficiency.

There are four main significances for precision marketing through big data. First, the company uses the data accumulated in the past to analyse the needs of customers at different levels, thereby launching a variety of suitable products and services and promoting the efficiency of corporate marketing. Second, through big data technology, companies organize and analyse collected data and information, and the resulting analysis results are fed back to all marketing links of the company, thereby improving the quality of products and services. Third, corporate data are divided into structured data and unstructured data. Structured data are better for statistics. For unstructured data, special technical means are needed, and big data technology can process various web pages, pictures, sounds and other unstructured data to help companies solve data processing problems. Fourth, through the analysis results of big data, companies can understand their product marketing effects, predict in advance which products consumers are interested in, and what products and services are suitable for people at different levels, so as to prepare for marketing in advance [3].

2 Related work

The literature [4] proposed that precision marketing is very important for corporate marketing and proposed the concept of Internet precision marketing. In the context of the Internet, many traditional communication methods of enterprises have been broken, and consumers can directly communicate with enterprises. Based on this communication mode, enterprises can achieve precise marketing based on the content of communication. The literature [5] proposed the marketing concept of segmented marketing. Segmented marketing is specifically to segment consumers, label them differently in accordance with established standards, and then provide precise and personalized services. The literature [6] introduced the two concepts of big data and precision marketing and made a case analysis on how big data is applied to precision marketing, and studied the relationship between big data and precision marketing through case analysis. The literature [7] studied the conditions for big data technology to achieve precision marketing and believed that in the Internet era, precision marketing should be combined with mobile Internet technology to accurately locate targets. The literature [8] studied the role of big data in precision marketing of media and believed that social media still has a large potential market. Using big data technology to explore these potential markets can better help the media's precision marketing. The literature [9] combined big data with management, and studies data mining of enterprise production and operation information through big data technology and then grasped consumer behavior preferences. Moreover, it used the results of the analysis as an important reference for the application of big data to precision marketing. The literature [10] studied the factors that affect the results of precision marketing and believed that the differentiation of target customers is an important factor in determining whether precision marketing can be achieved. Only by categorizing target customers can the big data precision marketing work efficiently and orderly. The literature [11] analyzed the security of big data precision marketing and believed that under the background of the Internet, big data marketing should pay attention to protecting corporate secrets and data security. The literature [12] studied data mining methods, and believed that data cannot be separated from the technical support of cloud computing, and believed that cloud computing technology is the most effective among many big data technology mining methods. The literature [13] discussed and analyzed the application scenarios and related fields of big data. The literature [14] analyzed the validity and safety of data and put forward corresponding suggestions on how to ensure the safety and validity of data. The literature [15] studied the impact of big data technology on the development of enterprises from the perspective of business model, and proposed a new business model combined with big data technology. The literature [16] discussed the relationship between big data and marketing from the perspective of corporate marketing, specifically analyzed the application of big data in this field, and put forward suggestions on how to use big data to achieve precision and efficiency in

marketing. The literature [17] discussed the importance of big data, connects big data with the commercial value of enterprises, and believed that using big data to improve the commercial value of enterprises has high requirements for the application of big data. The literature [18] studied the specific application of big data and introduced the application of big data technology in audit work. This innovative work model has brought many conveniences to audit work. The literature [18] pointed out that in the era of big data, rural small and medium financial institutions should seize development opportunities, use big data technology to mine user information, explore the development of value-added business services and create new marketing management paths for business operations. The literature [19] studied the advantages of big data in insurance marketing, including improving corporate risk management and control capabilities, improving the marketing efficiency of various marketing channels and providing conditions for precision marketing. The literature [20] studied the application of precision marketing in e-commerce. The literature [21] analyzed how precision marketing is implemented in the Internet, including the establishment of databases and the application of related technologies. The literature [22] studied the advantages of precision marketing from personalized services and believed that personalized precision marketing is an effective means to improve customer service levels and help companies to accurately locate. Moreover, it believed that personalized precision marketing is a key measure for e-commerce-related companies to carry out precision marketing, and the cooperation between personalized precision marketing and e-commerce platforms is a new direction for future marketing. The literature [23] proposed that the essence of precision marketing is data marketing whether it is with the help of Internet platforms or e-commerce platforms. To successfully realize precision marketing, we need to consider from the following perspectives, namely, market, customers, marketing channels and marketing strategy, market segmentation, target customer locking, marketing channel innovation and marketing strategy formulation. These are inseparable from massive data support, and these data are obtained through the user's buying behavior and browsing behavior. Moreover, analyzing and predicting the collected data can provide theoretical support for precision marketing. The literature [24] focused on the relationship between big data and precision marketing, how to apply big data mining technology to process and analyse large amounts of data, so as to achieve precision marketing. The literature [25] summarized the marketing concept of precision marketing, and believed that precision marketing should pay attention to the three factors of "time", "degree" and "efficiency". "Time" is timing and rhythm, "degree" is precision and frequency, and "effect" is effect.

3 Elements of reinforcement learning

Because the theoretical framework of reinforcement learning is MDP, MDP is introduced first, and then the basic components of reinforcement learning: strategies, rewards, and value functions are elaborated in order to discuss the reinforcement learning methods later.

(1) Markov decision process

MDP is generally used to formalize the description of reinforcement learning problems. Further, if both the state space and the behavior space are finite, then the finite Markov Decision Process (finite MDP) is used to describe. Unless otherwise specified, the MDP mentioned in this paper refers to finite MDP. Under normal circumstances, we can use a four-tuple $\langle S, A, p, r \rangle$ to represent finite MDP, among them: s represents a finite state space, defined as a finite set $\{S_1, S_2, \ldots, S_N\}$. N is the size of the state space. At the same time, A represents a finite behavior space, which is defined as a finite set $\{A_1, A_2, \ldots, A_M\}$. M is the size of the behavior space. P represents the state transition function, and r represents the instant reward function [26].

Consider that the state of the environment is S_t at any discrete time point t in the MDP. At this moment, if the agent takes action A_t, it will get a limited instant reward R_t, and the environment state will also change accordingly and move to the next state S_{t+1}. Among them, the state S_{t+1} at the next moment is determined according to the state transition function P, and the limited reward R_t is obtained according to the reward function r.

In addition, according to the different ways of state transition, it can be divided into two types: deterministic state transition and random state transition. Among them, in the deterministic state transition, the next state to which the transition is made is deterministic. The reward function is only related to the current state and the behavior to be taken, and it can be regarded as an evaluation of instant reward. In the random state transition, the next state transferred to is uncertain, but a random variable and the reward function depend on the next state in addition to the current state and the behavior taken. Therefore, it can be regarded as an evaluation of long-term effects.

(2) Strategy and return

The so-called strategy refers to the mapping from state s to behavior a, which is usually represented by the symbol π. Similarly, strategies also have deterministic strategies and random strategies. Among them, the deterministic strategy is defined as $\pi : S \rightarrow A$, and its output is an action sequence. The randomness strategy is defined as $\pi : S \times A \in [0, 1]$, and its output is the probability of the action taken in this state [27]:

$$\pi(s,a) = p[A_t = a | S_t = s] \tag{1}$$

The task of reinforcement learning is to learn a strategy π from the continuous interaction with the environment to maximize the cumulative reward (cumulative reward) generated, that is, maximize the return. We assume that the reward sequence received after the moment of death is $\{R_t, R_{t+1}, R_{t+2}, \ldots\}$, then the calculation of the reward using the discount cumulative reward method can be expressed as:

$$G_t = \sum_{k=0}^{\infty} \gamma^k R_{t+k} \tag{2}$$

In the above formula, G_t is the return and $\gamma < 1$ is a constant, which is called the discount factor. In particular, in the process of interaction between the system and the environment, if a task can be naturally divided into segments with an end time slice, the task is called an episode task. If the task cannot be decomposed into several fragments, and the entire task needs to be continued continuously, the task is called a continuous task.

(3) Value function

Because the goal of reinforcement learning is to maximize the cumulative reward, it is natural to use the above-mentioned return calculation method to measure the value of a state. However, since the state sequence generated in a certain state can have many situations, the return G_t is a random variable, not a definite value, and it cannot be optimized as an objective function. However, its expectation is a certain value, so it can be used as the definition of the state value function: when the agent adopts strategy π, its return obeys a distribution, and the expected return value of state s is defined as the value function of state s. The formula is defined as:

$$v_\pi(s) = E_\pi[G_t | S_t = s] = E_\pi\left[\sum_{k=0}^{\infty} \gamma^k R_{t+k} | S_t = s\right] \tag{3}$$

In some cases, it is more useful to record the value of the state behavior than just the value of the state. Therefore, in the state s, the expected value of the return generated by the agent after taking action a is defined as the state behavior value function, and the formula is defined as:

$$q_\pi(s,a) = E_\pi[G_t | S_t = s, A_t = a]$$
$$= E_\pi\left[\sum_{k=0}^{\infty} \gamma^k R_{t+k} | S_t = s, A_t = a\right] \tag{4}$$

The expression form of the above state value function is very inconvenient in actual calculation and programming. Therefore, through further derivation, the Bellman equation of the state value function can be obtained as follows:

$$v_\pi(s) = E_\pi[G_t | S_t = s]$$
$$= E_\pi\left[\sum_{k=0}^{\infty} \gamma^k R_{t+k} | S_t = s\right]$$
$$= E_\pi\left[R_t + \gamma \sum_{k=0}^{\infty} \gamma^k R_{t+k+1} | S_t = s\right]$$
$$= \sum_a \pi(a|s) \sum_{s'} p(s'|s,a) \tag{5}$$
$$\left[r(s,a,s') + \gamma E_\pi\left[\sum_{k=0}^{\infty} \gamma^k R_{t+k+1} | S_{t+1} = s'\right]\right]$$
$$= \sum_a \pi(a|s) \sum_{s'} p(s'|s,a)[r(s,a,s') + \gamma v_\pi(s')]$$
$$\forall s \in S$$

In the above formula:

$$p(s'|s,a) = pr\{S_{t+1} = s' | S_t = s, A_t = a\} \tag{6}$$

Indicates the probability that the environment will transition to the next state s' after the agent executes the behavior in the state s.

$$r(s,a,s') = E\{R_t | S_t = s, A_t = a, S_{t+1} = s'\} \tag{7}$$

represents the expected reward generated when the environment moves to the next state s' after the agent takes action a in the current state s.

In the same way, the Bellman equation of the state behavior value function can also be obtained as follows:

$$q_\pi(s,a) = \sum_{s'} p(s'|s,a)\left[r(s,a,s') + \gamma \sum_{a'} \pi(a'|s')q_\pi(s',a')\right]$$
$$\forall s \in S, \ \forall a \in A$$
$$\tag{8}$$

The purpose of calculating the value function is to learn the optimal strategy from the data when constructing the learning algorithm. Because each strategy corresponds to a state value function, the optimal strategy corresponds to the optimal state value function. Therefore, the optimal state value function $v_*(s)$ is defined as the value function with the largest value among all strategies, that is:

$$v_*(s) = \max_\pi v_\pi(s) \tag{9}$$

Similarly, the optimal state behavior value function $q_*(s,a)$ is defined as the state behavior value function with the largest value among all strategies, that is:

$$q_*(s,a) = \max_\pi q_\pi(s,a) \tag{10}$$

The Bellman optimal equations of the optimal state value function $v_*(s)$ and the optimal state behavior value function $q_*(s,a)$ can be obtained as shown in the following formula:

$$v_*(s) = \max_a \sum_{s'} p(s'|s,a)[r(s,a,s') + \gamma v_*(s')] \qquad (11)$$

$$q_*(s,a) = \sum_{s'} p(s'|s,a)\left[r(s,a,s') + \gamma \max_{s'} q_*(s',a')\right] \qquad (12)$$

Finally, if the optimal state behavior value function is obtained, the optimal strategy can be determined by directly maximizing $q_*(s,a)$, as shown in the following formula:

$$\pi_*(a|s) = \begin{cases} 1, & \text{if } a = \arg\max_{a \in A} q_*(s,a) \\ 0. & \text{otherwise} \end{cases} \qquad (13)$$

It can be seen that the value function is actually a prediction of future rewards (returns). Then, for a state s, if the instant reward it receives at this time is relatively low, it does not mean that its state value is low. Because if a higher reward is produced in the state after the state s, it will still achieve a high state value. This is the reason why reinforcement learning can consider the long-term profit maximization in the learning process, and it is also the reason why this paper chooses to use reinforcement learning technology to apply to marketing problems to obtain long-term profit maximization.

4 Reinforcement learning algorithm based on value function

The strategy iteration method mainly includes two steps of strategy evaluation and strategy improvement, and these two steps are alternated in sequence during the learning process. Specifically, when evaluating strategies, the algorithm needs to perform multiple iterations in each round. Moreover, in each iteration, it is necessary to scan each state in the state space according to the improved strategy in the previous round, and then use the Bellman equation to update the state value. After constant iteration, the value function finally converges to a fixed point, and then enters this round of strategy improvement steps. When improving the strategy, the algorithm uses the value function obtained from the previous round of strategy evaluation to generate a new strategy in a greedy manner. Then, the above two steps are continuously looped until it converges to the optimal strategy. In particular, when evaluating the strategy, the calculation of the state value $v_\pi(s)$ needs to use the value function $v_\pi(s')$ of its subsequent state, so the idea of bootstrapping is used.

In the following, the strategy iteration process of the state behavior value function is taken as an example to introduce in detail:

$$\pi_0 \xrightarrow{E} q_{\pi_0} \xrightarrow{I} \pi_1 \xrightarrow{E} q_{\pi_1} \xrightarrow{I} \pi_2 \xrightarrow{E} q_{\pi_2} \xrightarrow{I} \pi_3 \xrightarrow{E} \cdots \xrightarrow{I} \pi_* \xrightarrow{E} q_{\pi_*}$$
$$(14)$$

is all random. In the strategy evaluation, for each state behavior value, the Bellman equation is used to update the value function until q_{k+1} is stable, and the current round of iteration ends, as shown in the following equation. Among them, k represents the number of iterations in this round.

$$q_{k+1}(s,a) = \sum_{s'} p(s'|s,a)\left[r(s,a,s') + \gamma \sum_{a'} \pi(a'|s')q_k(s',a')\right] \qquad (15)$$

When the strategy is improved, for each state, the greedy strategy is used to improve the current strategy, as shown below.

$$\pi(s) = \arg\max_a q(s,a) \qquad (16)$$

Before converging to the optimal strategy, the strategy of each round is better than the strategy of the previous round. When strategy π is stable, the dynamic programming process ends, and the optimal value function $q_*(s,a)$ and optimal strategy π_* can be obtained.

(2) Value iteration

The value iteration algorithm is a simplification of the strategy iteration algorithm process, and it also includes two steps of strategy evaluation and strategy improvement. Different from strategy iteration, when making strategy improvement, it does not need to wait until the value function completely converges before making the improvement. Instead, it performs a strategy improvement every time the state behavior space is scanned (updated). In this way, the convergence speed of the value function is accelerated. For each state-behavior pair, the update formula for value iteration is as follows. Among them, l is the number of iterations.

$$q_{l+1}(s,a) = \sum_{s'} p(s'|s,a)\left[r(s,a,s') + \gamma \max_{a'} q_l(s',a')\right] \qquad (17)$$

Until q_{l+1} stabilizes, the algorithm ends. At this point, because the value function has converged, the optimal strategy can be obtained by directly applying the greedy method to the value function.

The Monte Carlo method uses the idea of averaging experience to replace the idea of random variables, so there is no need to provide environmental models. Among them, experience refers to doing many experiments according to this strategy, resulting in many episodes. Each plot is an experiment, and the average refers to the average value. According to the different solution methods, it can be

divided into the first-visit Monte Carlo method and the every-visit Monte Carlo method. In the process of interacting with the environment, when a plot is over, MC will use the data in the plot to update the value function and improve the strategy. In addition, the update of the value function uses an incremental (Incremental) method of calculating the mean. Taking the state value function as an example, the update formula is:

$$V_{k+1}(S_t) = V_k(S_t) + \alpha(G_t - V_k(S_t)) \tag{18}$$

Among them, $V_{k+1}(S_t)$ represents the estimated value of $V_{k+1}(S_t)$ at the $k+1$ th iteration, α represents the learning rate, and G_t represents the expected reward from the beginning of the state s to the end of the plot.

The so-called time difference refers to the deviation of the observation of the same event or variable at two consecutive times. In this paper, unless otherwise specified, the time difference algorithm mentioned refers to a one-step update algorithm, that is, the value function is directly learned according to the next time step, and then the value function of the next state is estimated. We assume that in state S_t, the agent takes action A_t, and the state transfers to S_{t+1} when reward R_t is obtained. Then, the time difference δ_t at time t is expressed as:

$$\delta_t = R_t + \gamma V(S_{t+1}) - V(S_t) \tag{19}$$

The update formula of the value function can be expressed as:

$$V(S_t) \leftarrow V(S_t) + \alpha(R_t + \gamma V(S_{t+1}) - V(S_t)) \tag{20}$$

α is the step length parameter, which controls the learning rate. $R_t + \gamma V(S_{t+1})$ is called TD target, which corresponds to G_t in formula (18). The difference between the two is that the TD target uses the bootstrapping method to estimate the current value function.

From a mathematical point of view, function approximation methods include parametric function approximation methods and non-parametric function approximation methods. Among them, parameterized function approximation can be divided into linear function approximation and nonlinear function approximation.

(1) Parametric linear function approximation

Parametric function approximation can be defined as a mapping from the parameter space to the approximate value function space. Therefore, the state value function and the state behavior value function can be approximated by a set of parameters $\theta \in \mathbb{R}^n$, respectively. In this paper, the approximate state value function and the approximate state behavior value function are denoted as: $\hat{v}(s; \theta)$ and $\hat{q}(s, a; \theta)$, as shown below:

$$\hat{v}(s; \theta) \approx v_\pi(s) \tag{21}$$

$$\hat{q}(s, a; \theta) \approx q_\pi(s, a) \tag{22}$$

It can be found that in the approximate state value function $\hat{v}(s; \theta)$, for each state, it no longer stores its state value, but only stores a set of parameters. The value function obtained from the known state is extended to those untouched states. For a large-scale state space, this is equivalent to compressing its state space. Because of $\theta \in \mathbb{R}^n$, the space storage overhead is $O(n)$. When the state scale is large and discrete, n is much smaller than the cost $|S|$ of storing values for each state.

From the table-based value function update process, it can be seen that both the MC method and the TD method are updated toward a certain target value. This target value is G_t in the Monte Carlo method. In the time difference method, it is:

$$R_t + \gamma Q(S_{t+1}, A_{t+1}) \tag{23}$$

Therefore, if the update process of the tabular value function is extended to the value function approximation process, then the process of approximating the value function $\hat{v}(s; \theta)$ can actually be regarded as a supervised learning process. The data and label pair is $\langle S_t, v_\pi(S_t) \rangle$. Among them, $v_\pi(S_t)$ is equivalent to G_t in the Monte Carlo method and $R_t + \gamma Q(S_{t+1}, A_{t+1})$ in the time difference method. Taking the approximation of the state value function as an example, the objective function to be trained can be expressed as the following formula:

$$\arg\min_\theta(v_\pi(s) - \hat{v}(s; \theta))^2 \tag{24}$$

That is, by finding the parameter vector θ, the mean square error between the approximate value function $\hat{v}(s; \theta)$ and the true value function $v_\pi(s)$ is minimized. In addition, depending on the update method, the update method of the value function can be divided into an incremental update method and a batch update method.

The most commonly used incremental update method is the gradient descent method. Therefore, the update formula of the parameters can be obtained from the above formula:

$$\theta \leftarrow \theta + \alpha[v_\pi(s) - \hat{v}(S_t; \theta)]\nabla_\theta \hat{v}(S_t; \theta) \tag{25}$$

In the incremental learning method, each update of the model requires interaction with the environment, and the data are discarded once it is used, which leads to inefficient use of samples. The so-called batch update method is to extract a data set from a given period of experience:

$$D = \{\langle S_1, v_1^\pi \rangle, \langle S_2, v_2^\pi \rangle, \ldots, \langle S_T, v_T^\pi \rangle\} \tag{26}$$

The best fitting function $\hat{v}(s; \theta)$ is found.

$$LS(\theta) = \sum_{t=1}^T (v_t^\pi - \hat{v_t^\pi}(S_t, \theta))^2 \tag{27}$$

This method is equivalent to Experience Replay. It revisits the experience over a period of time and updates the parameters, so the sample utilization efficiency is high.

Taking the state value function as an example, the linear function approximation method can be expressed as:

$$\hat{v}(s;\theta) = \sum_{i=1}^{n} \phi_i(s)\theta_i = \theta^T \phi(S) \tag{28}$$

Among them, $\theta = (\theta_1, \cdots, \theta_n) \in \mathbb{R}$.

$$\phi(s) = (\phi_1(s), \ldots, \phi_n(s))^T \tag{29}$$

The above formula is the eigenvector of state s, and $\phi_i(s)$ is called eigenfunction or basis function. Common basis functions are: polynomial basis functions, Fourier basis functions and radial basis functions. The linear function approximation method is not only simple and easy to understand, but also theoretically converges to the global optimum. The disadvantage is that the characterization ability is weak, and the type of basis function and the number of parameters need to be specified in advance, which limits the approximation ability of the function.

(2)   Parameterized nonlinear function approximation

Neural network can be seen as a method of parameterizing basis functions. Since the parameters of the basis function are learned from the data, its representation ability is greatly improved. Each layer of the neural network can be regarded as a new basis function based on the previous layer, that is, a feature. In this sense, the latter layer is the abstraction of the previous layer, which can achieve the effect of dimensionality reduction for high-dimensional input. Therefore, compared with the linear approximator, the advantage of neural network is that it has strong characterization ability and approximation ability, and can approximate the objective function with arbitrary precision. However, it also faces many problems: the samples of deep learning are independent, and the target distribution is fixed, while the state before and after reinforcement learning is related, and the target distribution is constantly changing, which often leads to unstable convergence, etc., problem. Therefore, a training method suitable for non-static, non-independent uniformly distributed data is needed to obtain the approximate function.

5 Model construction

The three main steps in precision marketing are the selection of target audiences, the selection of communication strategies and the measurement of marketing results. Using big data theory to conduct information mining on consumption data can provide data support for target audience selection, communication strategy selection and marketing result measurement, so as to conduct accurate mining of potential customers, accurate selection of marketing methods and accurate evaluation of marketing effects. The precision marketing process under big data is shown in Fig. 1.

According to the behavior of potential consumers, all potential customers are divided into multiple groups, Customers in each group have several obvious consumption and life characteristics. There are many kinds of marketing promotion methods, which can be selected according to the characteristics of the potential customer group, so as to maximize the marketing effect and minimize the marketing cost. In the process of selecting marketing methods, factors such as potential customer groups, marketing methods, marketing effects, and marketing costs need to be considered, as shown in Fig. 2.

Accurate customer management is the result of quantitative and precise marketing activities under the Internet and big data technology. Through the digital quantitative calculation of each element of marketing activities, it realizes the optimization control of the marketing process, the accurate management of precise customers, and finally realizes the optimization of marketing results. In the context of big data, the online accurate customer management plan is shown in Fig. 3.

In the era of big data, the browsing, clicking, leaving messages, comments, consumption and other information of hundreds of millions of consumers in the physical and online markets are silently recorded, which constitutes a
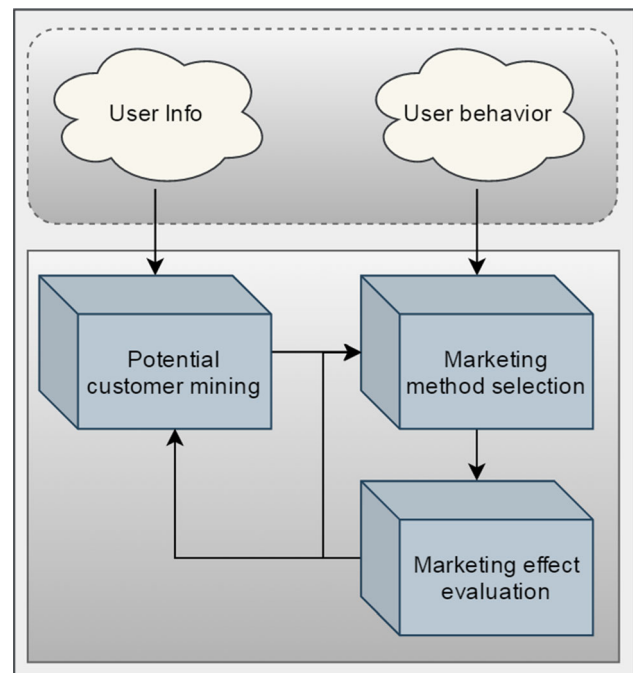

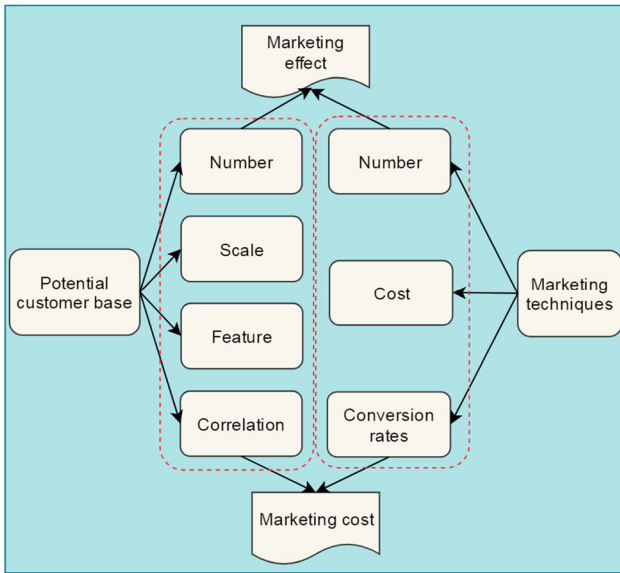**Fig. 1** The precision marketing process under big data

Fig. 2 Considerations for choosing marketing methods

digital activity map of consumers. The digital activity map records basic information, such as consumer living habits, social data and consumption information such as consumption records and consumption characteristics, as shown in Fig. 4. It records consumer life, work and consumption in a comprehensive and three-dimensional way. Through the analysis of the digital activity map, the potential consumer demand of consumers can be accurately obtained.

We need to segment hundreds of millions of consumers based on the huge amount of fragmented information and discover potential customers of Baby Online.For different
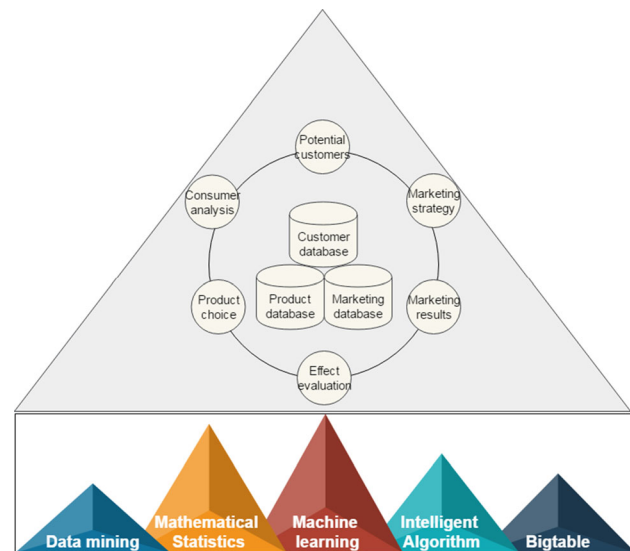

Fig. 3 Customer relationship management scheme in the context of big data

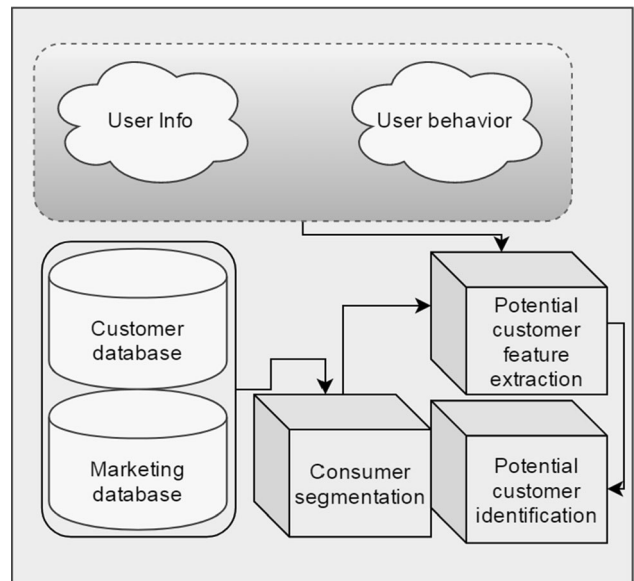**Fig. 4** Consumer Information Structure
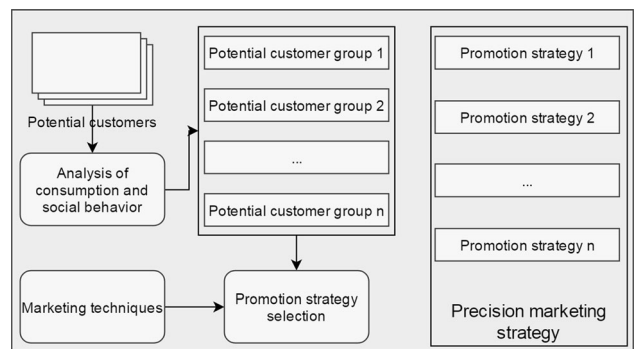

Fig. 5 Consumer segmentation process

Fig. 6 Marketing strategy selection process

Table 1 Statistical table of structural domain names

| Chinese domain	English domain	Chinese domain	English domain
Area code	ORG_CODE	Required quantity	QTY_DEMOND
Sales area code	SALE_REG_CODE	The actual amount	QTY_ORDER
Retail Account Code	CUST_CODE	Record time	INSERT_TIME
Address	CUST_ADDRESS	Store ID	SHOP_ID
Phone	CUST_PHONE	Execution ID	REC_ID
Store type	CUST_TYPE_CODE	User name	USER_NAME
Sales area code	SALE_REG_CODE	User ID	USER_ID
Years	YEAR_MONTH	Store country bureau code	NATIONAL_CODE
Which week	YEAR_WEEK	Store Name	SHOP_NAME
Order number	ORDER_ID	Store owner name	SHOP_USER
Order date	ORDER_DATE	Shop address	SHOP_ADDRESS
Retail Account Code	CUST_CODE	Is it a new store	ISNEW
Order amount	AMT_ORDER_SUM	Shop phone	SHOP_PHONE
Sales area code	SALE_REG_CODE	Longitude	LONGITUDE
Years	YEAR_MONTH	Latitude	LATITUDE
Which week	YEAR_WEEK	Total purchases of retail households	ORDER_QTY
Order number	ORDER_ID	Number of purchases by retail households	ORDER_NUM
Cigarette Specification Code	BRAND_CODE		

Table 2 Statistical diagram of the final cluster centers

	Clustering			
	1			1
ORDER_QTY	3.5499	ORDER_QTY		3.5499
ORDER_NUM	46	ORDER_NUM		46

Table 3 Statistical table of the number of cases in each cluster

Clustering	
1	10,666
2	7227
3	8205
Effective	25,839
Missing	0

Fig. 7 Final cluster center

(a) ORDER_QTY

(b) ORDER_NUM
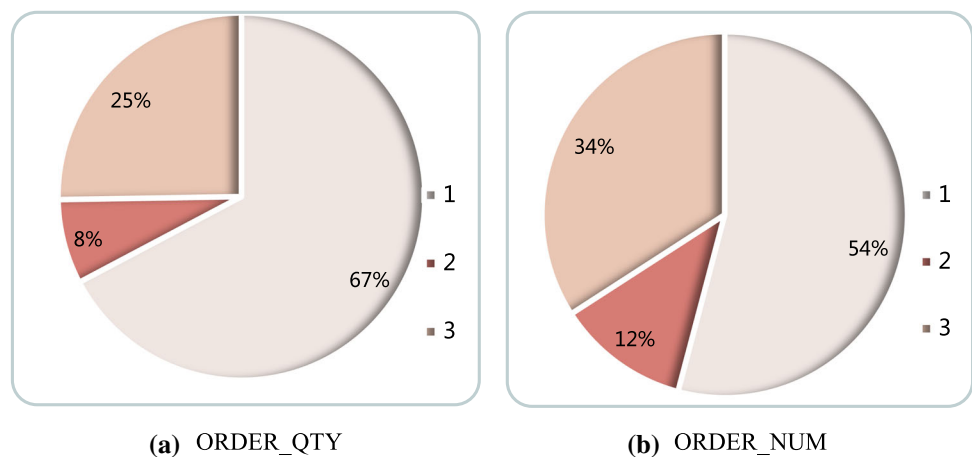
products, its potential customers are different and have different characteristics. These characteristics can be discovered from historical customer and marketing records. Based on these characteristics, combined with big data
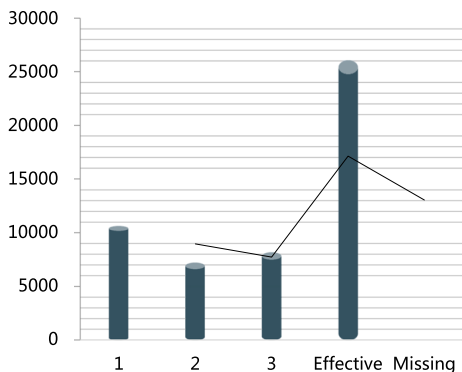
Fig. 8 Statistical diagram of the number of cases in each cluster
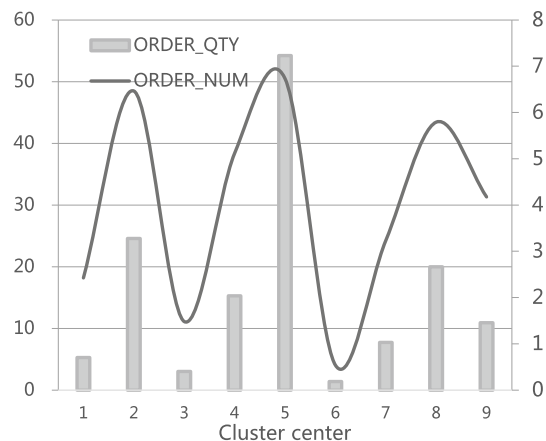
Table 4 Cluster analysis table

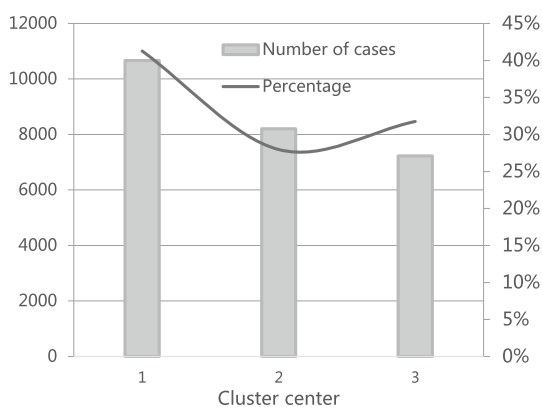| Cluster center	Number of cases	Percentage	Value type
1	10,665.6	41.28	High value
2	8205.24	27.97	MIDDLE value
3	7226.55	31.75	Low value

Fig. 9 Cluster analysis diagram

information, consumers can be segmented, and then potential maternal and child consumption customers of Baby Online can be determined. The process is shown in Fig. 5.

Through the analysis of the consumption and social characteristics of potential customer groups, Baobao can choose suitable and effective marketing strategies online.

Fig. 10 Statistical diagram of the final cluster centers

Table 6 Value clustering results of retail households

Cluster center	Number of cases	Percentage	Value type
1	2358	9.13	Low
2	3980	15.40	In
3	2726	10.55	Low
4	3246	12.56	In
5	1783	6.90	High
6	2533	9.81	Low
7	3018	11.68	Low
8	3238	12.53	In
9	3215	12.44	Low

Fig. 11 Statistical diagram of the value clustering results of retail households

Table 5 Statistical table of the final cluster centers

	1	2	3	4	5	6	7	8	9
ORDER_QTY	0.705	3.275	0.401	2.036	7.231	0.182	1.028	2.667	1.454
ORDER_NUM	18	48	11	38	51	4	24	43	31

Table 7 Value classification table of regional retail households

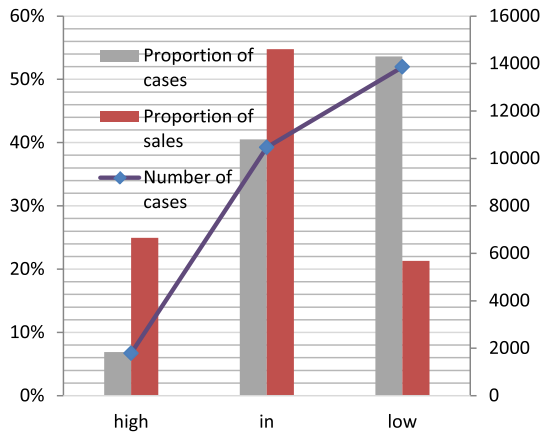| Value type	Number of cases	Proportion of cases (%)	Proportion of sales (%)
High	1783	6.90	24.96
In	10,465	40.50	54.75
Low	13,850	53.61	21.29

Fig. 12 Value classification diagram of regional retail households

Using appropriate channels for promotion and publicity at the right time can achieve a multiplier effect. On the premise of ensuring the marketing effect, it effectively reduces the marketing cost of Baby Online. The process is shown in Fig. 6.

6 E-commerce precision marketing case analysis

Company A mainly sells products through e-commerce model. As a top-ranking company in the industry, after years of development, it has achieved tremendous development and has made significant contributions to the

country's tax profits. For company A's precision marketing, the relevant structured domain names are roughly as shown in Table 1.

Drawing lessons from the idea of the RFM model, this paper selects the total purchase amount (ORDER_QTY), and the number of purchases (ORDER_NUM) corresponding to retail households as two index parameters, as the analysis dimension, to perform cluster analysis. Before clustering, relevant data have been cleaned up, and outliers have been processed. Therefore, for each dimension, the result set of clustering is determined by the data itself, which allows the function of the clustering algorithm to be fully utilized to ensure that customer clustering in each dimension achieves better results. This paper uses SPSS statistical tools to cluster the data in 2014, and tries to cluster three categories according to the dimensions in the above table, and divide the retail households into three value segments: high, medium and low. The results obtained are shown in Table 2 and Fig. 7.

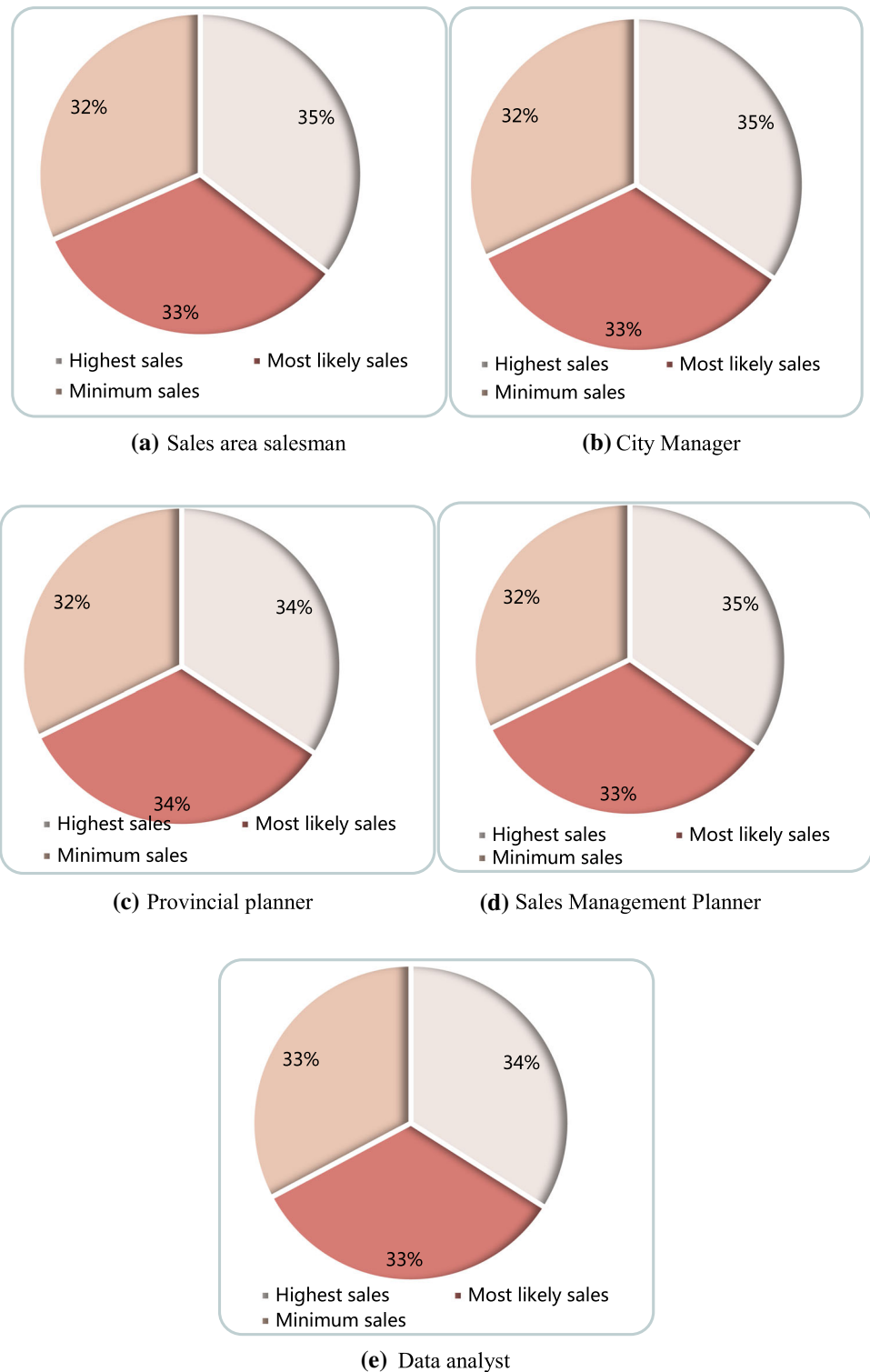The number of cases in each cluster is shown in Table 3 and Fig. 8.

Through the analysis of the clustering results, the following rules are found in Table 4 and Fig. 9.

According to the above classification results, there is an unreasonable distribution of the number of cases of the three value types, and the number of high-value customer cases is too high, which does not meet the 28 principle, and the result of the solid classification is not accepted. After adjusting the number of clusters and combining multiple clustering analyses, the final clustering results are shown in Table 5 and Fig. 10.

Table 8 A brand sales forecast in the study area (unit: ten thousand boxes)

Expert group	Predictive classification	Regional sales	Expert group	Predictive classification	Regional sales
Sales area salesman	Highest sales	14.14	Sales management planner	Highest sales	14.645
	Most likely sales	13.13		Most likely sales	13.938
	Minimum sales	12.625		Minimum sales	13.635
City Manager	Highest sales	14.645	Data analyst	Highest sales	14.645
	Most likely sales	14.14		Most likely sales	14.342
	Minimum sales	13.635		Minimum sales	14.14
Provincial planner	Highest sales	14.443			
	Most likely sales	14.14			
	Minimum sales	13.635			

Fig. 13 Forecast of brand sales in the study region

(a) Sales area salesman

(b) City Manager

(c) Provincial planner

(d) Sales Management Planner

(e) Data analyst

According to empirical judgment, the final clustering results of the order quantity (ORDER_QTY) and the order number (ORDER_NUM) are shown in Table 6 and Fig. 11.

The value of regional retail households is classified, and the results are shown in Table 7 and Fig. 12.

As can be seen from the above figure, the sales of Brand A of Company A are mainly concentrated in high-value

and medium-value retail households. The number of high-value and medium-value retail households accounted for about 46% of the total retail households, but the sales of high-value and medium-value retail households accounted for nearly 80%, so they are the key retail households of Company A.

From the above analysis results, we can identify the key retail households of Company A. Next, we will study the value changes of these retail households.

For the sales forecast in the study area, we divide it into three levels, namely, the highest sales volume, the most likely sales volume, and the lowest sales volume, which are scored by selected group members. At the same time, the sales weights of the three gears are different, the highest sales weight is 0.2, the most likely sales weight is 0.5, and the lowest sales weight is 0.3. Finally, the weighted average method is used to obtain the sales forecast. The whole process adopts a recurring method. After 3 rounds of investigation, the average value of the 3 results is collected, and the final result is obtained, as shown in Table 8 below, and the corresponding statistical diagram is shown in Fig. 13.

When forecasting sales, it is required that the team members must not discuss with each other, do not have horizontal contact, and can only communicate with the sales management planner. The final results and opinions have been repeatedly consulted, summarized and revised. Moreover, the final forecast data can provide a basis for the decision-making of superior leaders.

7 Conclusion

Aiming at the problem that traditional supervised learning and unsupervised learning methods can only maximize the instantaneous revenue of a single event when dealing with direct marketing problems, but cannot guarantee that the long-term revenue can be maximized in sequential decision-making problems, this paper chooses to use the value function-based reinforcement learning method for research. Then, in order to better adapt to actual needs, based on three aspects: the unfixed time interval of marketing decision points in the direct marketing scenario, the large data load and the low learning efficiency, and the partial observability of customer status, this paper improves the reinforcement learning algorithm based on the value function, and uses a simulation environment to verify the effectiveness of the proposed method.

Among them, this paper mainly introduces the development and application of big data, and the concept and theory of precision marketing; analyses the bottlenecks and problems encountered by company A in the development of precision marketing; introduces the design of the precision marketing system of A company; uses retail household data and the KMEANS algorithm to subdivide the value of retail households; uses the Boston matrix to classify the market; combines with the ARMA model, Holt-Winter model and Delphi expert survey method to comprehensively predict sales. In summary, this paper provides a reference for company A's brand marketing to improve the efficiency and quality of the company's precision marketing, and perfect its existing precision marketing system.

Compliance with ethical standards

Conflict of interest The authors declare that they have no competing interests.

References

1. Kozlenkova IV, Samaha SA, Palmatier RW (2014) Resource-based theory in marketing. J Acad Mark Sci 42(1):1–21
2. Eckhardt GM, Houston MB, Jiang B et al (2019) Marketing in the sharing economy. J Mark 83(5):5–27
3. Tiago MTPMB, Veríssimo JMC (2014) Digital marketing and social media: why bother? Bus Horiz 57(6):703–708
4. Moorman C, Day GS (2016) Organizing for marketing excellence. J Mark 80(6):6–35
5. Harmeling CM, Moffett JW, Arnold MJ et al (2017) Toward a theory of customer engagement marketing. J Acad Mark Sci 45(3):312–335
6. Kannan PK (2017) Digital marketing: a framework, review and research agenda. Int J Res Mark 34(1):22–45
7. Wedel M, Kannan PK (2016) Marketing analytics for data-rich environments. J Mark 80(6):97–121
8. Leeflang PSH, Verhoef PC, Dahlström P et al (2014) Challenges and solutions for marketing in a digital era. Eur Manag J 32(1):1–12
9. Germann F, Ebbes P, Grewal R (2015) The chief marketing officer matters! J Mark 79(3):1–22
10. McDonagh P, Prothero A (2014) Sustainability marketing research: past, present and future. J Mark Manag 30(11–12):1186–1219
11. Hillebrand B, Driessen PH, Koll O (2015) Stakeholder marketing: theoretical foundations and required capabilities. J Acad Mark Sci 43(4):411–428
12. Barrales-Molina V, Martínez-López FJ, Gázquez-Abad JC (2014) Dynamic marketing capabilities: toward an integrative framework. Int J Manag Rev 16(4):397–416
13. Samaha SA, Beck JT, Palmatier RW (2014) The role of culture in international relationship marketing. J Mark 78(5):78–98
14. Voorhees CM, Brady MK, Calantone R et al (2016) Discriminant validity testing in marketing: an analysis, causes for concern, and proposed remedies. J Acad Mark Sci 44(1):119–134

15. Krishna A, Schwarz N (2014) Sensory marketing, embodiment, and grounded cognition: a review and introduction. J Consum Psychol 24(2):159–168

16. Gneezy A (2017) Field experimentation in marketing research. J Mark Res 54(1):140–143

17. Ashley C, Tuten T (2015) Creative strategies in social media marketing: an exploratory study of branded social content and consumer engagement. Psychol Mark 32(1):15–27

18. Key TM, Czaplewski AJ (2017) Upstream social marketing strategy: an integrated marketing communications approach. Bus Horiz 60(3):325–333

19. Berger J, Humphreys A, Ludwig S et al (2020) Uniting the tribes: using text for marketing insight. J Mark 84(1):1–25

20. Carins JE, Rundle-Thiele SR (2014) Eating for the better: a social marketing review (2000–2012). Public Health Nutr 17(7):1628–1639

21. Kozlenkova IV, Hult GTM, Lund DJ et al (2015) The role of marketing channels in supply chain management. J Retail 91(4):586–609

22. Heimbach I, Kostyra DS, Hinz O (2015) Marketing automation. Bus Inf Syst Eng 57(2):129–133

23. Chang YT, Yu H, Lu HP (2015) Persuasive messages, popularity cohesion, and message diffusion in social media marketing. J Bus Res 68(4):777–782

24. Luxton S, Reid M, Mavondo F (2015) Integrated marketing communication capability and brand performance. J Advert 44(1):37–46

25. Buhalis D, Foerste M (2015) SoCoMo marketing for travel and tourism: empowering co-creation of value. J Destin Mark Manag 4(3):151–161

26. Huang J, Duan Z, Kwok J et al (2019) Vaping versus JUULing: how the extraordinary growth and marketing of JUUL transformed the US retail e-cigarette market. Tobacco Control 28(2):146–151

27. Appel G, Grewal L, Hadi R et al (2020) The future of social media in marketing. J Acad Mark Sci 48(1):79–95